



Video-based step measurement in sport and daily living.

DUNN, Marcus David.

Available from the Sheffield Hallam University Research Archive (SHURA) at:

<http://shura.shu.ac.uk/20193/>

A Sheffield Hallam University thesis

This thesis is protected by copyright which belongs to the author.

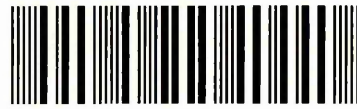
The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Please visit <http://shura.shu.ac.uk/20193/> and <http://shura.shu.ac.uk/information.html> for further details about copyright and re-use permissions.

Learning and IT Services
Collegiate Learning Centre
Collegiate Crescent Campus
Sheffield S10 2BP

102 044 771 0



REFERENCE

ProQuest Number: 10700838

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10700838

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Video-based step measurement in sport and daily living

Marcus David Dunn

A thesis submitted in partial fulfilment of the requirements of
Sheffield Hallam University
for the degree of Doctor of Philosophy

March 2014

Collaborating organisation: International Tennis Federation

Abstract

Current knowledge of tennis player-surface interactions with different court surfaces is limited. The measurement of player step and movement strategy would aid the understanding of tennis player-surface interaction. However, this has not yet been performed: no readily available motion analysis tool is capable of measuring spatio-temporal parameters of gait during match-play tennis. The purpose of this project was to develop, validate and use a motion analysis tool designed to measure player location and foot-surface contacts during match-play tennis.

Single camera video footage, obtained from the 2011 Roland Garros Qualifying Tournament, was manually digitised to characterise step and movement strategy during men's and women's forehand groundstrokes. Player movements were consistent with previous notational analyses; however gender differences were highlighted for step frequency. Initial findings were limited by manual analysis, e.g. manual digitising subjectivity and low sample size: an objective and automated system was required.

A markerless, view-independent, foot-surface contact identification (FSCi) algorithm was developed. The FSCi algorithm identifies foot-surface contacts in image sequences of gait by quantifying the motion of each foot. The algorithm was validated using standard colour image sequences of walking and running obtained from four unique camera perspectives: output data were compared to three-dimensional motion analysis. The FSCi algorithm identified data for 1243 of 1248 foot-surface contacts; root-mean-square error (RMSE) was 52.2 and 103.4 mm for shod walking and running respectively (all camera perspectives). Findings demonstrated that the FSCi algorithm measured basic, spatio-temporal parameters of walking and running, e.g. step length and step time, without interfering with the activity being observed. Furthermore, analyses were independent of camera view.

Video footage obtained from the 2011 ATP World Tour Finals was used to develop a combined player tracking and foot-surface contact identification (PT-FSCi) algorithm. Furthermore, a graphical user interface was developed. The PT-FSCi algorithm was used to analyse twenty match-play tennis rallies: output data were compared to manual digitising. The PT-FSCi algorithm tracked player position and identified data for 832 of 890 foot-surface contacts during match-play tennis. RMSE for player position and foot-surface contacts was 232.9 and 121.9 mm respectively. The calculation of step parameters required manual intervention: this reflected the multi-directional nature of tennis. This represents a limitation to the current algorithm however the segmentation of player movement phases to allow the automatic calculation of step parameters.

The analysis of this data indicated that top ranked tennis players can win rallies using movement strategies previously considered to be defensive. Furthermore, step length data indicated that shorter step lengths formed the majority of step strategy. The largest 25% of steps were observed behind the baseline, aligned with deuce and advantage court sidelines. This reflected lunging and turning manoeuvres at lateral extremes of player movement.

The single camera system that has resulted from this project will enable the International Tennis Federation to characterise player step and movement strategy during match-play tennis. This will allow a more informed approach to player-surface interaction research. Furthermore, the system has potential to be used for different applications, ranging from sport to surveillance.

Acknowledgements

I would like to thank Dr Simon Goodwill, Dr Jon Wheat and Professor Steve Haake for their continued support and advice throughout this project. Their experience has proved both insightful and motivating. My thanks are also extended to everyone within the Centre for Sports Engineering Research at Sheffield Hallam University.

I am very grateful to the International Tennis Federation for their sponsorship of this project. In particular, I would like to thank Dr Stuart Miller, Jamie Capel-Davies and James Spurr for their invaluable assistance and continued support for this project.

Finally, I would like to thank Dr Karl Cooke of British Swimming (formerly the Lawn Tennis Association), for his assistance in obtaining video footage of elite, match-play tennis.

Contents

Abstract	ii
Acknowledgements	iii
Contents	iv
List of Figures	x
List of Tables	xxi
Nomenclature	xxiii
1 Introduction	1
1.1. Motivation for research	1
1.2. Aim and objectives	3
2 Literature review	5
2.1. Introduction	5
2.1.1. History of tennis	5
2.1.2. Science and tennis	7
2.1.3. Tennis player movement and injury	9
2.1.4. Conclusion	12
2.2. Motion analysis techniques	13
2.2.1. Non-intrusive motion analysis technologies in tennis	14
2.2.2. Non-intrusive player motion analysis	15
2.2.3. Sport stadia	20
2.2.4. Conclusion	25
2.3. Gait measurement techniques	26
2.3.1. Holistic approaches	26
2.3.2. Model-based approaches	28
2.3.3. Model-based heel strike detection	31
2.2.4. Background modelling	33
2.2.5. Conclusion	35
2.4. Chapter findings	36
2.4.1. Player tracking	36
2.4.2. Heel strike extraction	37
2.4.3. Conclusion	37
3 Player step and movement characterisation at the 2011 Roland Garros Qualifying Tournament	38
3.1. Introduction	38
3.2. Aim and objectives	38

3.3.	Measuring player step and movement strategy at the 2011 Roland Garros Qualifying Tournament	39
3.3.1.	Single camera calibration	39
3.3.2.	Footage collection in sport stadia	41
3.3.3.	Elevated calibration plane	42
3.3.4.	Player step and movement definitions	43
3.3.5.	Analysis software	44
3.4.	Results	46
3.5.	Discussion	48
3.5.1.	Rally characterisation	48
3.5.2.	Filming and analysis	49
3.6.	Conclusion	50
4	Single camera position reconstruction	52
4.1.	Introduction	52
4.2.	Aim and objectives	53
4.3.	Monocular photogrammetry	53
4.3.1.	Camera model	53
4.3.2.	Camera-plane model	57
4.3.3.	Position reconstruction	59
4.4.	Ground plane position reconstruction assessment	61
4.4.1.	Competitive tennis environments: data collection and modelling	61
4.5.	Results	68
4.6.	Discussion	72
4.7.	Conclusion	75
5	Single camera player position reconstruction using an elevated calibration plane	77
5.1.	Introduction	77
5.2.	Aim and objectives	77
5.3.	Out-of-plane error	78
5.4.	Simulating sport stadia camera views	79
5.4.1.	Player position reconstruction	84
5.5.	Results	86
5.6.	Discussion	93
5.7.	Conclusion	96

6	An automatic technique for identifying foot-surface contacts	98
6.1.	Introduction	98
6.2.	Aim and objectives	98
6.3.	Vision-based foot-surface contact identification	99
6.3.1.	Single camera methods	99
6.3.2.	Logical image processing	100
6.3.3.	Synthetic walking data	103
6.3.4.	Foot-region segmentation	106
6.3.5.	Foot-region inter-frame motion	108
6.4.	Development of an algorithm to measure foot-surface contacts	111
6.4.1.	Colour image collection	112
6.4.2.	Background segmentation	113
6.4.3.	Shadow removal	114
6.4.4.	Inter-frame motion	117
6.4.5.	Foot segmentation	118
6.4.6.	Foot-surface contact threshold	121
6.4.7.	Geometric rules	126
6.5.	Application to walking and running	128
6.6.	Conclusion	133
7	Validation of an automatic technique for identifying foot-surface contacts in walking and running	134
7.1.	Introduction	134
7.2.	Aim and objectives	134
7.3.	Laboratory validation study	135
7.3.1.	Participants and procedures	135
7.3.2.	Experimental setup	135
7.3.3.	Criterion data treatment	138
7.3.4.	FSCi algorithm data treatment	140
7.3.5.	Reconstruction plane elevation	141
7.3.6.	Data analysis	142
7.4.	Results	144
7.4.1.	Reconstruction plane elevation	145
7.4.2.	Foot-surface contacts	147
7.4.3.	Step parameters and reconstruction plane elevation	154

7.5.	Discussion	157
7.5.1.	Analysis time and identification rate	157
7.5.2.	Foot-surface contact time	159
7.5.3.	Foot-surface contact position	162
7.5.4.	Step parameters	165
7.5.5.	Reconstruction plane elevation	166
7.6.	Conclusion	166
8	A semi-automatic technique for player tracking and foot-surface identification at the 2011 ATP World Tour Finals	168
8.1.	Introduction	168
8.2.	Aim and objectives	169
8.3.	Development of a semi-automatic technique to identify player position and foot-surface contacts	169
8.3.1.	Image collection and intrinsic camera calibration	169
8.3.2.	Player segmentation	173
8.3.3.	Court line and shadow processing	174
8.3.4.	Player tracking windows	178
8.3.5.	Assessing player tracking	180
8.3.6.	Foot-surface contact identification	181
8.3.7.	Extrinsic camera calibration and position reconstruction	182
8.3.8.	Graphical user interface	183
8.4.	Application to match play tennis	187
8.5.	Conclusion	192
9	Validation of a semi-automatic technique for player tracking and foot-surface contact identification at the 2011 ATP World Tour Finals	193
9.1.	Introduction	193
9.2.	Aim and objectives	193
9.3.	Match play tennis validation study	193
9.3.1.	Camera calibration parameters	194
9.3.2.	PT-FSCi tennis rally analysis	196
9.3.3.	Manual tennis rally analysis	197
9.3.4.	Data analysis	198
9.4.	Results	199

9.4.1.	Player tracking	200
9.4.2.	Foot-surface contact and step parameters	202
9.4.3.	Player step and movement characterisation	204
9.5.	Discussion	209
9.5.1.	Analysis time and user intervention	209
9.5.2.	Player tracking assessment	211
9.5.3.	Foot-surface contacts and step parameters	212
9.5.4.	Rally characterisation	215
9.6.	Conclusion	217
10	Conclusions	219
10.1.	Introduction	219
10.2.	Summary of research	219
10.2.1.	Single camera filming	219
10.2.2.	Manual player step and movement strategy analysis	219
10.2.3.	Two-dimensional position reconstruction	220
10.2.4.	Two-dimensional player position reconstruction	220
10.2.5.	Foot-surface contact identification	220
10.2.6.	Tennis player tracking and foot-surface contact identification	221
10.2.7.	Tennis rally analysis using the PT-FSCi algorithm	222
10.2.8.	Conclusion	223
10.3.	Current and future development	223
10.3.1.	Analysis procedure	223
10.3.2.	Programming language	223
10.3.3.	Outdoor filming	224
10.3.4.	Gait analysis tool	224
	References	225
	Personal Bibliography	240
	Appendix 1	241
A.1.1.	Permission to film and accreditations obtained via the Fédération Française de Tennis.	241
A.1.2.	Ethics application form: Tennis player step and movement characterisation.	242
A.1.3.	Risk assessment: Tennis player step and movement characterisation.	250

Appendix 2	252
A.2.1. Age, mass and stature of players analysed at the 2011 Roland Garros Qualifying Tournament.	252
A.2.2. Custom MATLAB analysis script for manually digitised tennis rally parameters.	252
Appendix 3	262
A.3.1. Permission to film and accreditations obtained via the Lawn Tennis Association.	262
A.3.2. Ethics application form: Semi-automatic tennis player step and movement characterisation.	263
A.3.3. Risk assessment: Semi-automatic tennis player step and movement characterisation.	272
Appendix 4	274
A.4.1. Scale (1:30) tennis court model.	274
Appendix 5	275
A.5.1. Internal report: Assessment of Calibration Techniques.	275
Appendix 6	289
A.6.1. Ethics application form: Validation of an automatic foot-surface contact identification algorithm.	289
A.6.2. Risk assessment: Validation of an automatic foot-surface contact identification algorithm.	297
A.6.3. Participant information sheet: Validation of an automatic foot-surface contact identification algorithm.	300
Appendix 7	302
A.7.1. Ethics application form: Validation of an automatic foot-surface contact identification algorithm.	302
A.7.2. Risk assessment: Validation of an automatic foot-surface contact identification algorithm.	311
A.7.3. Participant information sheet: Validation of an automatic foot-surface contact identification algorithm.	314
A.7.4. Foot-surface contact identification algorithm profile: most expensive processes.	316

List of Figures

Figure 1.1.	Adrian Quist (1930; A) and Novac Djokovic (2011; B) have won multiple Grand Slam tournaments. The athleticism of ‘modern era’ player movement is both striking and topical in tennis media.	2
Figure 1.2.	Motion analysis tool development stages, described within chapters.	4
Figure 2.1.	<i>Jeu de paume</i> (adapted from Haake <i>et al.</i> , 2007).	6
Figure 2.2.	Wingfield’s lawn tennis (adapted from Haake <i>et al.</i> , 2007).	6
Figure 2.3.	Average speed of 20 fastest servers at Grand Slam events since 2002 (taken from Miller, 2006).	7
Figure 2.4.	Rule changes to the racket from 1978 to 2002 (adapted from Haake <i>et al.</i> , 2007).	8
Figure 2.5.	A: Manually digitised image. B: ‘Attack angle’ (θ) and base of support (adapted from Starbuck <i>et al.</i> , 2013).	11
Figure 2.6.	The application of electromagnetic sensors (A), inertial sensors (B), accelerometers (C) and markers (D) is intrusive and would violate rules of match-play tennis.	14
Figure 2.7.	‘Official Review’ graphic presented to players and the public.	15
Figure 2.8.	Camera locations used for analysis by SAGIT (adapted from Martínez-Gallego <i>et al.</i> , 2013).	17
Figure 2.9.	Screenshots from TennisSense illustrating a rear camera view and fixed camera filming locations (red circles about tennis court; adapted from Conaire <i>et al.</i> , 2009).	18
Figure 2.10	Feature extraction: camera image (A), foreground image (B) and radial map (C: adapted from Connaghan, Moran and O’Connor, 2013).	19
Figure 2.11.	A: The All-England Lawn Tennis and Croquet Club at Wimbledon. B: Incremental elevation of stadia seating (adapted from John, Sheard and Vickery, 2007).	21
Figure 2.12.	Player (A) and ball (B) tracking from broadcast video (adapted from Yan, Christmas and Kittler, 2005).	23

Figure 2.13.	A: Perspective camera filming location. B: Player foreground segmentation and identified player location (green cross) used for player position tracking (adapted from Mauthner <i>et al.</i> , 2008).	24
Figure 2.14.	Manual and automatic estimates for player position: a spike in estimated position (bottom right) reflects a jumping motion (adapted from Mauthner <i>et al.</i> , 2008).	25
Figure 2.15.	A: Reflective tape on anatomical landmarks is floodlit. B: Point-light display (adapted from Kozlowski and Cutting, 1977).	26
Figure 2.16.	Holistic gait recognition features (angular transform) derived from a silhouette (adapted from Boulgouris, Plataniotis and Hatzinakos, 2006).	27
Figure 2.17.	Inter-foot distance (diamonds: top) and tertiary signal (asterisks: bottom) highlighting inter-foot distance maxima (positive) and minima (negative). Distance maxima indicate dual-stance (adapted from Johnson and Bobick, 2001).	28
Figure 2.18.	Anatomical silhouette segmentation for sagittal (A) and oblique-sagittal (B) walking (adapted from Goffredo <i>et al.</i> , 2008). Red and green dashed lines illustrate anatomical segmentation and walking direction respectively.	29
Figure 2.19.	Model-based recognition of running (A) and walking (B), joint angle data are derived using the leading edge (highlighted in white) of lower-limbs (adapted from Yam, Nixon and Carter, 2002).	30
Figure 2.20.	A: Sagittal camera view of walking. B: Corner proximity image for heel-strike extraction (adapted from Bouchrika and Nixon, 2006).	31
Figure 2.21.	A: Accumulator map for extracted foreground pixels. B: Filtered heel strike candidates (adapted from Jung and Nixon, 2013).	32
Figure 2.22.	Evaluation of six background models using 100 sample images. Analysis time (green) and total segmentation error rate (blue) for indoor applications (adapted from Hassanpour, Sedighi and Manashty, 2011).	34

Figure 2.23.	Evaluating of shadow removal by colour space. F^* is an efficiency measure quantifying misclassified foreground and background pixels (adapted from Benedek and Szirányi, 2007).	35
Figure 3.1.	Overview of Roland Garros; singles tennis matches performed on court seven (yellow ring) were filmed.	39
Figure 3.2.	Elevated camera setup at the 2011 Roland Garros Qualifying Tournament.	42
Figure 3.3.	Digitised ground and elevated calibration plane locations (red and yellow filled circles respectively) highlighted in a combined image.	43
Figure 3.4.	Screenshot of analysis system used to manually digitise and quantify player step and movement strategy.	44
Figure 3.5.	Camera image with calibration points (C1 - C4) and global coordinate system plotted.	45
Figure 4.1.	Pinhole camera geometry illustrating the real world point Q projected as q in the image plane (adapted from Bradski and Kaehler, 2008).	54
Figure 4.2.	Sample image (original format; A) and with distortions (radial and tangential) applied to the z (imaginary) axis (B), illustrating the spherical effect of the lens system.	56
Figure 4.3.	Point p_c (camera coordinate) is related to point P_o (global coordinate) by applying the rotation matrix R and translation vector t (adapted from Bradski and Kaehler, 2008).	58
Figure 4.4.	Camera-plane model describing extrinsic camera parameters (position and orientation) at the 2011 Roland Garros Qualifying Tournament.	59
Figure 4.5.	Camera setup at the real (A) and scale model (B) of the 2011 ATP World Tour Final.	62
Figure 4.6.	Checkerboard calibration at the 2011 Roland Garros Qualifiers: checkerboard corner extraction (A), tennis court field-of-view (B) and extrinsic checkerboard parameters (C).	63
Figure 4.7.	Combined camera image of real and scale model tennis courts of the 2011 ATP World Tour Finals. Crosshairs provide a criterion measure for photogrammetric assessment.	65

Figure 4.8.	Real (A) and model (B) camera perspectives of the 2011 Roland Garros Qualifying Tournament. Images illustrate perspective projection and lens distortion (arrows and rings, values in pixels). $[R \ t]$ illustrates the homography between court (XYZ) and camera (xy) coordinate systems.	64
Figure 4.9.	Internal (blue pluses) and external (red pluses) reconstruction points and calibration points (C1 - C4) for ATP (A) and Roland Garros (B) models.	67
Figure 4.10.	Internal and external ($n = 124$) RMSE (mm) for 2D-DLT (red: incremented calibration coordinates) and planar reconstruction (green) in X , Y and R directions for the RG model.	70
Figure 4.11.	Internal and external ($n = 158$) RMSE (mm) for 2D-DLT (red: incremented calibration coordinates) and planar reconstruction (green) in X , Y and R directions for the ATP model.	71
Figure 5.1.	Reconstruction error incurred when the reconstructed point and reconstruction plane are not coplanar (adapted from Holden-Douilly <i>et al.</i> , 2011).	78
Figure 5.2.	Coarsely meshed point cloud data for a human participant in standing (A) and running (B) postures.	82
Figure 5.3.	Simulated ($n = 657$) camera locations (blue squares: arrows indicate optical axis) relative to running point cloud data (black figure at court centre) for RG.	83
Figure 5.4.	Processed camera image of reprojected point cloud data and reprojected COM (red) and ground plane COM (green) locations.	84
Figure 5.5.	A: Sample image features used for ground and elevated plane position reconstruction (red and blue diamonds respectively). B: Representation of ground (red), elevated (blue) and out-of-plane, e.g. ± 200 mm (green), reconstruction planes.	85
Figure 5.6.	Resultant direction RMSE (mm) for player COM position using Ground and Elevated plane definitions.	87
Figure 5.7.	Mean (cross) and standard deviation (error bars) player position reconstruction error for RG using standing (A) and running (B) posture data. Black vertical lines illustrate RG, ATP and maximum stadia elevation angles (solid, dash-dot and dots respectively).	89

Figure 5.8.	Mean (cross) and standard deviation (error bars) player position reconstruction error for ATP using standing (A) and running (B) posture data. Black vertical lines illustrate RG, ATP and maximum stadia elevation angles (solid, dash-dot and dots respectively).	91
Figure 5.9.	Maximum position reconstruction error for all camera azimuth angles at each camera elevation angle for RG (A) and ATP (B). Black vertical lines illustrate RG, ATP and maximum stadia elevation angles (solid, dash-dot and dots respectively).	92
Figure 6.1.	Inter-frame difference image of walking results in a 'double image' (adapted from Zhang, Zhou and Zhu, 2010).	100
Figure 6.2.	Filled circle with basic image properties highlighted.	101
Figure 6.3.	Logical image of tennis racket and ball (A), removal of small objects (B), dilation (C), erosion (D), convex hull (E), skeletonised image (F) and skeletonised image endpoints (yellow rings).	102
Figure 6.4.	Erosion (left) and dilation (right) of a 3×3 logical image (centre), using a 3×3 structuring element.	103
Figure 6.5.	Sample images rendered in Visual 3D of walking from oblique frontal (A), oblique rear (B) and sagittal (C) perspectives.	104
Figure 6.6.	Column-wise concatenation of the image plane region $R(x,y,p)$.	104
Figure 6.7.	Silhouette segmentation based on silhouette height (H; adapted from Goffredo <i>et al.</i> , 2008). Red and green dashed lines illustrate anatomical segmentation and walking direction respectively.	106
Figure 6.8.	Binary images of lower-limbs (A), foot-region masks (B) and foot-region inter-frame differences (C). Connected perimeter and convex-hull skeleton pixels are highlighted in green and blue respectively. Skeleton endpoint and skeleton-perimeter pixels are highlighted in red and blue filled circles respectively.	107
Figure 6.9.	$MaskFDI_B$ areas for foot-region masks one and two (red and blue respectively) during a walking sequence. Areas exceeding mean t_{FSC} (black horizontal line) indicate foot-surface contacts.	109
Figure 6.10.	Collated foot-surface contact data (red crosses) for sagittal (A), oblique frontal (B) and oblique rear (C) camera views.	110

Figure 6.11.	Image collection and analysis steps for foot-surface contact algorithm.	111
Figure 6.12.	Schematic of experimental setup: dashed rectangle represents the motion capture volume filmed by four network cameras, e.g. NCam 1 - 4.	113
Figure 6.13.	Absolute image differencing (A - C), binary image extraction (D) and morphological operations (E - F).	114
Figure 6.14.	Lower-body (A), lower-body convex hull (B), HSV absolute image differencing (C - E) and binary image extraction (F).	115
Figure 6.15.	Three-plane binary image concatenation.	116
Figure 6.16.	Lower-limb binary image with shadow suppressed (A), cropped inter-frame difference image (B) and binary image extraction (C - E).	117
Figure 6.17.	Convex hull of $HS-ADI_B$ (A), skeletonised image of $HS-ADI_B$ convex hull with inferior endpoints highlighted (B) and $HS-ADI_B$ perimeter pixels with skeleton endpoints transformed into the image coordinate system (red circles: C).	119
Figure 6.18.	Initial circular mask (A), initial foot-region (B), initial foot-region (zoomed) with identified coordinates and refined mask highlighted (C), refined foot-region (D), product of refined foot mask and FSC_B (E) and largest connected component (F).	120
Figure 6.19.	Refined foot-regions, i.e. $RefinedFoot_B$, obtained from cameras 1 - 4 (A - D respectively) with refined foot mask footprints highlighted (blue circles).	122
Figure 6.20.	Refined foot-region areas ($AreaRefinedFoot$) for frontal (A) and sagittal (B) perspectives. Solid and dashed lines indicate foot-regions one and two respectively.	122
Figure 6.21.	Frontal camera perspective of stance with refined foot-regions (blue circles) and skeleton-perimeter pixels (red filled circles) highlighted.	123
Figure 6.22.	Sagittal camera perspective of stance with refined foot-regions (blue circles) and skeleton-perimeter pixels (red filled circles) highlighted.	124

Figure 6.23.	M_t for frontal (A) and sagittal (B) camera perspectives. Solid and dashed lines indicate M_t for foot-regions ones and two respectively.	125
Figure 6.24.	Normalised refined foot-region areas (red) and corresponding foot-region inter-frame differences (blue) for frontal (A) and sagittal (B) camera perspectives. Solid and dashed lines indicate foot-regions one and two respectively.	126
Figure 6.25.	Geometric rules applied to a foot-surface contact candidate (green filled circle and red cross).	129
Figure 6.26.	Logical foot-region masks (red and blue circles) and foot-surface contacts (filled green circles and red cross) for four camera perspectives.	128
Figure 6.27.	Stance (highlighted by red rectangle) is evidenced by accumulating foot-surface contact data points (green filled circles) about the stance foot for both walking (top) and running (bottom) gait (temporal direction is right-to-left). Self-occlusion (walking) is highlighted by blue dashed rectangle (images 6 - 8).	130
Figure 6.28.	Collated foot-surface contact data (green filled circles) superimposed on image t_F of walking (A - D) and running (E - H) image sequences for four camera perspectives.	132
Figure 7.1.	Experimental setup: motion analysis cameras, e.g. MAC 1 - 8 (A) and red circles (B), network cameras, e.g. NCam 1 - 4 (A) and green circles (B), filmed the motion capture volume, e.g. C ₁ - C ₄ (A) and blue circles (B), and LED light box, e.g. LED (A) and yellow circle (B).	136
Figure 7.2.	Network camera calibration: checkerboard extraction (A), extrinsic checkerboard parameters (B) and network camera image with motion capture volume superimposed (C).	137
Figure 7.3.	Foot marker motion between heel-strike (blue), foot-surface contact (green) and toe-off (red) for walking: progression is indicated by colour (blue-to-red).	139
Figure 7.4.	Example criterion foot-surface contact data (filled red circles). Foot markers (red pluses and blue dashed lines), resultant step length and step time (coloured arrows) are highlighted for illustration.	140

Figure 7.5.	Foot-surface contact RMSE (single walking trial) using incrementally elevated reconstruction planes (red crosses).	142
Figure 7.6.	Spatial step components, i.e. step length, step width, step distance and step angle (θ), relative to direction of progression (adapted from Huxham <i>et al.</i> , 2006).	143
Figure 7.7.	Proportion of image sequences requiring manual intervention (left) and manual intervention type (right).	145
Figure 7.8.	LOA (indicated by red and black lines) for all foot-surface contacts ($n = 1243$) in the X direction using reference (A) and elevated (B) reconstruction planes. Red, green, blue and magenta data identify network cameras 1 - 4 respectively.	146
Figure 7.9.	LOA (indicated by red and black lines) for all foot-surface contacts ($n = 1243$) in the Y direction for reference (A) and elevated (B) reconstruction planes. Red, green, blue and magenta data identify network cameras 1 - 4 respectively.	146
Figure 7.10.	Proportion of stance identified for walking (A) and running (B).	150
Figure 7.11.	Mean (green circles) and standard deviation (green error bars) foot contact locations relative to criterion locations (blue circles) for barefoot walking measured by cameras 1 - 4 (A - D respectively).	151
Figure 7.12.	Mean (green circles) and standard deviation (green error bars) foot contact locations relative to criterion locations (blue circles) for shod walking measured by cameras 1 - 4 (A - D respectively).	152
Figure 7.13.	Mean (green circles) and standard deviation (green error bars) foot contact locations relative to criterion locations (blue circles) for barefoot running measured by cameras 1 - 4 (A - D respectively).	153
Figure 7.14.	Mean (green circles) and standard deviation (green error bars) foot contact locations relative to criterion locations (blue circles) for shod running measured by cameras 1 - 4 (A - D respectively).	154
Figure 7.15.	Optimal reconstruction plane elevation for task and condition (A) and condition (B) relative to the reference plane for all network cameras.	156

Figure 7.16.	Foot marker motion between heel-strike (blue), foot-surface contact (green) and toe-off (red) for running: progression is indicated by colour (blue-to-red).	160
Figure 8.1.	Image collection and analysis steps for player tracking and foot-surface contact identification.	169
Figure 8.2.	Camera setup at the 2011 ATP World Tour Finals (A) and sample camera image (B).	170
Figure 8.3.	Single camera calibration: checkerboard extraction (A) and extrinsic checkerboard parameters (B).	170
Figure 8.4.	Deinterlaced player images (cropped) during a forehand groundstroke.	171
Figure 8.5.	Extracted fields without (A) and with (B) a row-averaging filter applied.	172
Figure 8.6.	Manually defined region (red rectangle) superimposed on image t_1 (A) and image t_n (B) and resultant background image (C).	172
Figure 8.7.	Extracted binary image (A) and original image (B) with player position (manual and estimated in yellow and red respectively) and tracking window plotted.	174
Figure 8.8.	Player image (A), identified court lines (B), processed court lines and superimposed player outline (C), combined player and court line silhouette (D).	175
Figure 8.9.	Binary images derived from RGB (A) and HS (B) colour spaces. RGB pixel values and shadow perimeter are superimposed on the combined foreground player image for illustration (C).	177
Figure 8.10.	Elliptical grouping for a discontinuous binary image, relative to previous player location (red filled circle).	177
Figure 8.11.	Tracking windows superimposed on combined images of a baseline run with original image interval (A) and increased image interval (B).	179
Figure 8.12.	Imposition of a default tracking window (green rectangle) if an automated tracking window (red rectangle) exceeds image boundaries (black rectangle).	179
Figure 8.13.	Horizontal player coordinate data (black) and smoothing spline (blue). Coordinates exceeding 95% confidence intervals about the smoothing spline are highlighted (red circles).	180

Figure 8.14.	Pixels above player COM (red filled circle) are set to zero (A → B).	182
Figure 8.15.	Manually digitised sideline-baseline and sideline-serviceline intersections.	182
Figure 8.16.	Screenshot of main GUI window. User controls are located to the left and data visualisations to the centre and right of the screen.	184
Figure 8.17.	Analysis controls and information (A) and threshold sliders (B).	185
Figure 8.18.	Foot-surface contact controls (A) and data collection options (B).	186
Figure 8.19.	Module for reviewing player segmentation as well as player position and foot-surface contact data.	187
Figure 8.20.	Sequence of foreground player images with RGB pixel values and shadow perimeter superimposed for illustration.	189
Figure 8.21.	Player tracking residuals and 95% confidence intervals (red dashed rectangle).	190
Figure 8.22.	Player foreground image (A) and zoomed-in region (B).	191
Figure 9.1.	Camera setup at the 2011 ATP World Tour Finals (A) and sample camera image (B).	194
Figure 9.2.	Single camera calibration: checkerboard extraction (A) and extrinsic checkerboard parameters (B).	195
Figure 9.3.	Camera-plane model describing extrinsic camera parameters (position and orientation) at the 2011 ATP World Tour Finals.	196
Figure 9.4.	Foot-surface contact data (analysed by PT-FSCi algorithm) accepted for analysis or flagged for removal.	199
Figure 9.5.	Manually digitised player location (green cross and error bars) relative to normalised foreground player coordinates for all analysed images (n = 6612). Colour bar indicates foreground coordinate frequency.	201
Figure 9.6.	Proportion of foot-surface contacts identified during match-play tennis.	202
Figure 9.7.	Player position (green dotted line) and foot-surface contact (red filled circles) data relative to the camera-plane model (zoomed-in) for a single rally.	204
Figure 9.8.	Player-court occupancy map (colour bar indicates duration in seconds) for match-play tennis rallies (n = 20).	205
Figure 9.9.	Distribution of tennis player speed (resultant direction) during	206

match-play tennis rallies (n = 20).

Figure 9.10.	Distribution of step length during match-play tennis rallies (n = 20).	207
Figure 9.11.	Player-court location map for step lengths ≥ 1.12 m (colour bar indicates frequency) for match-play tennis rallies (n = 20).	208
Figure 9.12.	Segmented player images and incorrectly identified foot-surface contact data (magenta diamonds about lower-limbs).	210

List of Tables

Table 3.1.	Men's and women's tennis rally characteristics on clay (N = 40).	46
Table 3.2.	Men's and women's step and movement characteristics for tennis rallies on clay (N = 40).	47
Table 3.3.	Player displacement (resultant direction) for men's and women's tennis rallies on clay (N = 40).	48
Table 4.1.	RMSE (mm) for <i>X</i> , <i>Y</i> and <i>R</i> directions using 2D-DLT (incremented calibration points) and planar reconstruction for a scale model of the 2011 Roland Garros Qualifying Tournament.	68
Table 4.2.	Table 4.2. RMSE (mm) for <i>X</i> , <i>Y</i> and <i>R</i> directions using 2D-DLT (incremented calibration points) and planar reconstruction for a scale model of the 2011 ATP World Tour Finals.	69
Table 4.3.	Mean calibration plane pixel-scale (mm; <i>X</i> , <i>Y</i> and <i>R</i> directions), calibration point image distortion (pixels; C1 – C4) and residual r^2 (<i>X</i> and <i>Y</i> directions) for ATP and RG models.	72
Table 5.1.	Mean \pm standard deviation (mm) resultant direction (<i>R</i>) player reconstruction error for all camera perspectives (n = 657).	88
Table 7.1.	Network camera calibration residuals and extrinsic parameters.	138
Table 7.2.	Step frequency and resultant velocity for walking and running tasks.	144
Table 7.3.	Accepted and rejected foot-surface contact candidates.	144
Table 7.4.	RMSE (mm) for all foot-surface contact position data (n = 1243) in <i>X</i> , <i>Y</i> and <i>R</i> directions.	147
Table 7.5.	LOA (absolute and ratio), r^2 and RMSE for foot-surface contact (FSC) parameters during barefoot (n = 391) and shod (n = 362) walking.	148
Table 7.6.	LOA (absolute and ratio), r^2 and RMSE for foot-surface contact (FSC) parameters during barefoot (n = 251) and shod (n = 238) running.	149

Table 7.7.	Mean camera pixel-scale (mm) for <i>X</i> and <i>Y</i> directions and schematic of camera locations.	150
Table 7.8.	LOA (absolute and ratio), r^2 and RMSE for walking step parameters.	155
Table 7.9.	LOA (absolute and ratio), r^2 and RMSE for running step parameters.	156
Table 7.10.	Resultant direction RMSE (mm) for foot-surface contact parameters.	163
Table 9.1.	Images, duration and strokes for match-play tennis rallies ($N = 20$).	194
Table 9.2.	Camera calibration residuals and extrinsic camera parameters ($n = 20$).	195
Table 9.3.	Manual digitising standard error of measurement ($N = 5$).	197
Table 9.4.	Foot-surface contact data removed from analyses ($n = 591$).	200
Table 9.5.	LOA (absolute and ratio), r^2 and RMSE for player position ($n = 6612$) and player velocity ($n = 6572$) during match-play tennis.	202
Table 9.6.	LOA (absolute and ratio), r^2 and RMSE for foot-surface contact parameters ($n = 832$) during match-play tennis.	203
Table 9.7.	LOA (absolute and ratio), r^2 and RMSE for step length and step time during match-play tennis ($n = 762$).	203
Table 9.8.	Mean and standard deviation player travel, absolute player velocity, step number, step length and step rate for match-play tennis rallies ($n = 20$).	206
Table 9.9.	Resultant direction RMSE (mm) for foot-surface contacts.	214

Nomenclature

Abbreviations

ANOVA	Analysis of Variance
ATP	Association of Tennis Professionals
BMIT	Time independent background model
CIE	International Commission on Illumination
CMOS	Complementary Metal–Oxide–Semiconductor
CMYK	Cyan, Magenta, Yellow and Key (black)
COF	Coefficient of Friction
COM	Centre of mass
CPU	Central Processing Unit
DLT	Direct Linear Transformation
2D-DLT	Two-dimensional Direct Linear Transformation
ES _B	Between factor effect size
FIFA	Fédération Internationale de Football Association
FSCi	Foot-Surface Contact identification
GMM	Gaussian mixture model
GUI	Graphical User Interface
HSV	Hue, Saturation and Value
IBBM	Improved basic background model
ITF	International Tennis Federation
LED	Light Emitting Diode
LOA	Limits of Agreement
LTABM	Long-term average background model
LUX	SI unit of illuminance (equal to one lumen per square metre)
MABM	Moving average background model
MAC	Motion Analysis Corporation
NaN	Undefined numeric value
NCam	Network Camera
PT-FSCi	Player-Tracking and Foot-Surface Contact identification
RG	Roland Garros
RGABM	Running Gaussian average background model

RGB	Red, Green and Blue
RMSE	Root-Mean Square Error
SEM	Standard Error of Measurement
TEM	Technical Error of Measurement
TIFF	Tagged Image File Format
YCbCr	Luma, Blue difference and Red difference

1 Introduction

The following chapters contain a three-year study into the development, validation and use of a motion analysis tool designed to measure foot-surface contact position in match-play tennis.

1.1. *Motivation for research*

Tennis is a popular racket sport that attracts millions of players and spectators worldwide (Fernandez, Mendez-Villanueva and Pluim, 2006). Competitive tennis events are regulated by the International Tennis Federation (ITF) which is responsible for monitoring tennis to improve safety, performance and participation while preserving the sports' integrity. As such, a fundamental role of the ITF is to determine the rules and specifications of tennis to help regulate the sport (ITF, 2010).

A defining characteristic of tennis is that it can be played on a variety of court surfaces (ITF, 2010). Indeed, Grand Slam tournaments, i.e. Wimbledon, Roland Garros, Australian Open and US Open, are played on grass, clay and acrylic surfaces. Mechanical properties of tennis courts used for competition, i.e. friction, energy restitution, dimension, etc., must meet standards published by the ITF to ensure safety of use and consistency between competitions (ITF, 2010). However different types of tennis injury have been associated with play on different court surfaces (Pluim *et al.*, 2006), highlighting limited knowledge concerning how players interact with court surfaces (Miller, 2006).

Tennis is an evolving sport. The 'wooden racket era' of tennis reflects a period when game style was characterised by style and finesse. Following the introduction of aluminium, oversized rackets in 1975, the 'modern era' of tennis refers to a game now characterised by more powerful strokes, higher rates of ball spin and more athletic court movements (Fernandez, Mendez-Villanueva and Pluim, 2006). In 'modern era' tennis, players travel 8 – 12 m during an average rally, changing direction four times (Fernandez, Mendez-Villanueva and Pluim, 2006). However player movement patterns differ between grass, clay and acrylic tennis surfaces. O'Donoghue and Ingram (2001) identified that rallies played at Roland Garros (clay surface) were longer and consisted

of more baseline play than rallies played at the US and Australian Opens (acrylic surfaces). Similarly, rallies at the US and Australian Opens were longer and consisted of more baseline play than rallies played at Wimbledon (grass surface). As such, mechanical court surface characteristics, e.g. coefficient of friction (COF), have been suggested to influence player movement patterns and subsequently, player injury risk (Girard *et al.*, 2007).

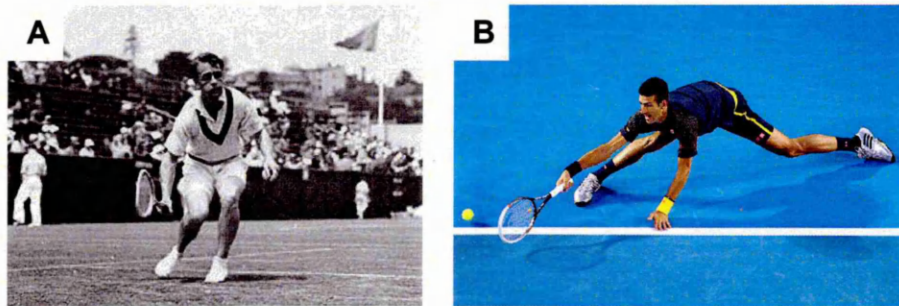


Figure 1.1. Adrian Quist (1930; A) and Novak Djokovic (2011; B) have won multiple Grand Slam tournaments. The athleticism of ‘modern era’ player movement is both striking and topical in tennis media.

Laboratory based research has attempted to characterise the tennis shoe-surface interface during tennis forehand foot plants using different surfaces. However Stiles and Dixon (2006) highlighted problems with simulating sport specific manoeuvres in laboratory conditions. Despite distinctly different mechanical characteristics between carpet, acrylic and artificial turf, the authors found no differences for biomechanical measures, e.g. peak translation COF, peak ground reaction force (vertical and horizontal). Biomechanical parameters were assessed as they were considered to be indicative of the human response to impacts with different surfaces. The lack of change to foot-surface contact characteristics reflected individual movement strategies for task-oriented skills, e.g. inter-participant variability (Stiles and Dixon, 2006). The role of gait strategy mediation has been previously identified in slip research, e.g. in response to known slip hazards (Chambers *et al.*, 2003), and in running research, e.g. to reduce injury risk (Derrick, 2004) and to allow optimal performance (Hardin, Bogert and Hamill, 2004). Stiles and Dixon (2006) suggested their findings were influenced by gait strategy mediation prior to the forehand foot plant they recorded. However, the authors had no method of quantifying this or, more importantly, qualifying whether player movement data were representative of real tennis.

Current knowledge of tennis player-surface interactions with different court surfaces is limited (Miller, 2006). Spatio-temporal gait parameters are commonly assessed for many activities, e.g. slip research (Chambers *et al.*, 2003), running (Derrick, 2004; Hardin, Bogert and Hamill, 2004) and sprinting (Hunter, Marshall and McNair, 2004). In tennis, Hughes and Meyers (2005) developed a normative profile of player movement sequences for notational analyses; the quantification of gait parameters was not considered. Only in squash has the work of Pereria, Wells and Hughes (2001) attempted to link footwork to the outcome of rallies, providing a more detailed representation of games. However, the explicit measurement of spatio-temporal parameters of gait in tennis has not yet been performed.

To characterise tennis player interactions with different court surfaces, spatio-temporal parameters of gait during match-play must be measured. There is limited data on tennis player gait strategy in match-play tennis. The key reason for this is that there is no readily available motion analysis tool that is capable of measuring gait parameters during match-play tennis. The ITF, who sponsor this project, require such a tool to measure spatio-temporal parameters of gait during match-play tennis. This will allow the ITF to inform future research into tennis shoe-surface interaction.

1.2. *Aim and objectives*

Aim

To develop, validate and use a motion analysis tool designed to measure player location and foot-surface contacts during match-play tennis.

Objectives:

1. To review relevant literature on motion analysis techniques; specifically those used to track sports performers and identify gait strategy.
2. To collect pilot footage at match-play tennis events to identify real world filming constraints and characterise gait strategy of tennis players.
3. To develop a bespoke tool designed to measure tennis player position and foot-surface contact location from footage of match-play tennis.
4. To validate the tool using laboratory and match-play tennis data.

5. To use the tool to obtain player position and foot-surface contact location data for match-play tennis.

The following chapters address the project aim and objectives identified above. Figure 1.2 identifies specific development stages of the motion analysis tool, e.g. boxes A - H, which are addressed by ensuing chapters.

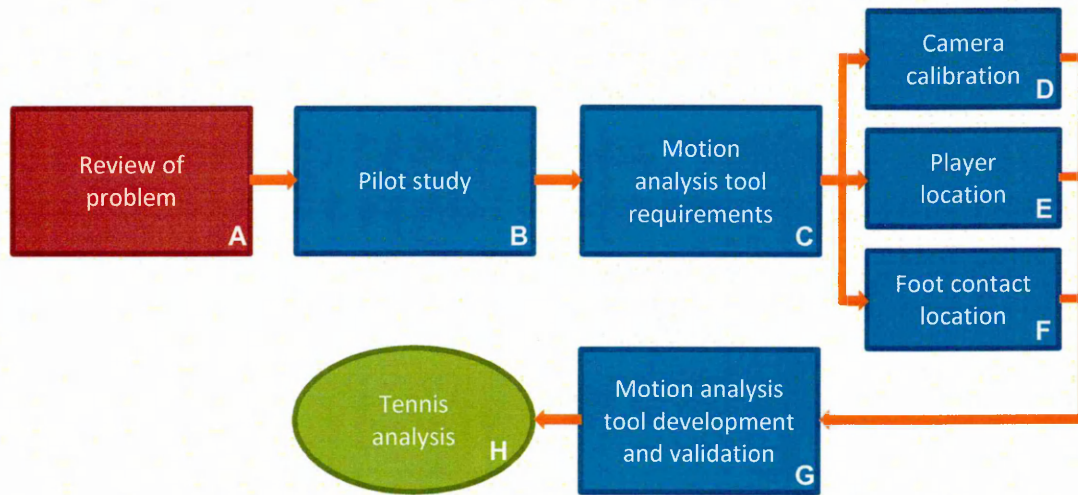


Figure 1.2. Motion analysis tool development stages, described within chapters.

2 Literature review

2.1. Introduction

This chapter provides a review existing literature related to the aim and objectives identified in section 1.2. Specifically, this chapter relates to box A of the development stage flow diagram, presented in Figure 1.2.

2.1.1. History of tennis

The origins of tennis are not clear. Descriptions and illustrations of ball games thought to be early incarnations of modern tennis date back to twelfth century Europe. Accounts of medieval Europe usually derive from the Church, as is much of literary material from the time. Such drawings, writings and descriptions help form current understanding of the historical development of tennis. One fascinating and sinister account of a game was written by Caesarius of Heisterbach between 1219 and 1223 in the '*Dialogus miraculorum*' (Gillmeister, 1988). Caesarius described a band of demons that, in a valley steaming with sulphurous vapours, divided into two teams and began hitting a young Parisian clerk's soul to each other, catching it mid-air with their hands (Gillmeister, 1988).

Although Caesarius' account is folkloric, the religious figures cited in the tale are traceable and importantly, elements within the story echo more accepted accounts of tennis from the period. *Jeu de paume* (meaning '*game of palm*') was a courtyard game played in twelfth century France (Haake *et al.* 2007). The game was played with a ball that was struck by hand to opposing players. The use of hands in *jeu de paume* (Figure 2.1) highlights the word '*tenez*' (meaning '*hold*'), from which the English word '*tennis*' is thought to descend. Similarities exist between Caesarius' tale and *jeu de paume*. Indeed, it is possible that Caesarius was recounting his observations of *jeu de paume* but added the sinister context: it was considered indecorous of clergy members to play ball games at the time. However, both accounts place the origins of tennis in twelfth century France. The game later evolved into *Royal Tennis* or *Real Tennis*, which is accepted as the precursor for modern tennis (ITF Tennis, 2013).

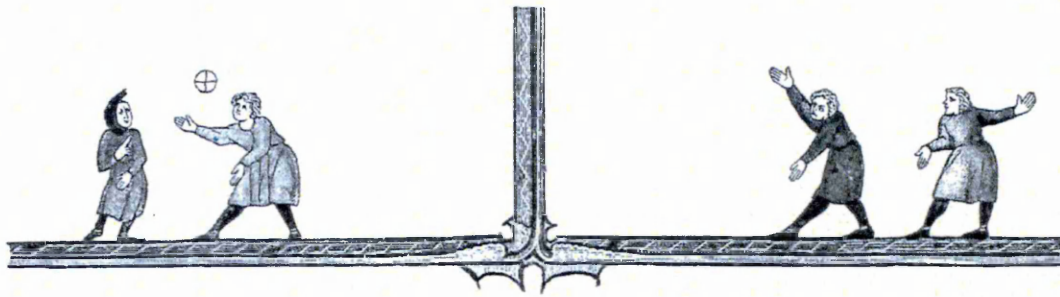


Figure 2.1. *Jeu de paume* (adapted from Haake *et al.*, 2007).

Real tennis, which is still played today, is played on an asymmetric court with a net and is enclosed by walls. A low bouncing, cork ball is used and can be played off of walls and the sloping gallery roofs around the court. Real tennis continued to develop throughout the middle ages but was initially considered elitist due to expensive facilities needed to play the game. A marked change to the format of tennis came when Major Wingfield patented his game ‘*sphairistikè*’ (meaning ‘*ball game*’ in Greek) in 1874 (Haake *et al.*, 2007). Wingfield brought the game of tennis to a much wider audience since ‘*sphairistikè*’ or lawn tennis as it later became known, could be played on practically any size or shape lawn and only required rackets, a net and ball (Figure 2.2).

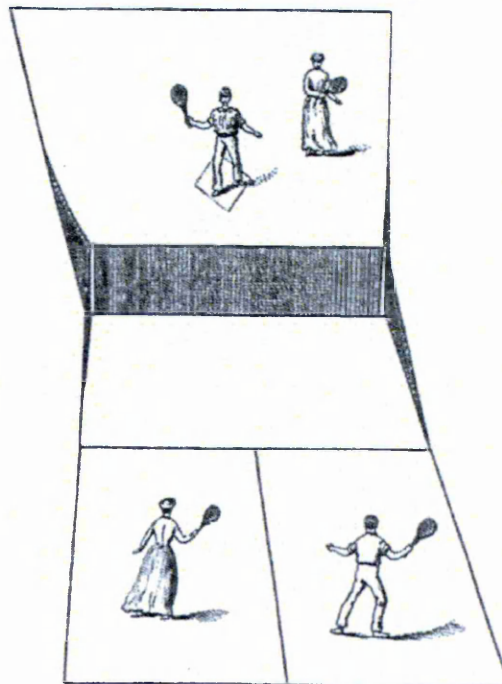


Figure 2.2. Wingfield’s lawn tennis (adapted from Haake *et al.*, 2007).

Lawn tennis was introduced to the members of the All England Croquet Club in 1875. Following the games popularity, the club changed its name to the All England Croquet and Lawn Tennis Club in 1877. The first competitive lawn tennis tournament was held at Wimbledon that same year, where standardised rectangular courts and a scoring system were formally introduced. Following the first Wimbledon Championships, the popularity of lawn tennis accelerated with The Lawn Tennis Association being formed in 1888 and the International Tennis Federation in 1913. Today, tennis is a global sport; the athleticism of tennis player movement helps to keep the sport entertaining and relevant, an element that the ITF wish to preserve (ITF, 2010).

2.1.2. Science and tennis

Science, irrespective of interest or discipline, is characterised by a research-question approach to epistemology (Winter and Fowler, 2009). Following the first Wimbledon tournament in 1877, questions about the influence of the serve on the outcome of the game arose. Spencer Gore (winner of Wimbledon that year) replied “...*Did you know that Mr Jones has figured out that 376 games have been won on serves and only 225 games on returns. Does that seem fair to you?*” (Coe, 2000). Gore’s response reflects a match or notational analysis. Since this first competitive event, the popularity of tennis has increased as has the desire to win; technological innovation in tennis is driven by the financial implications of success.

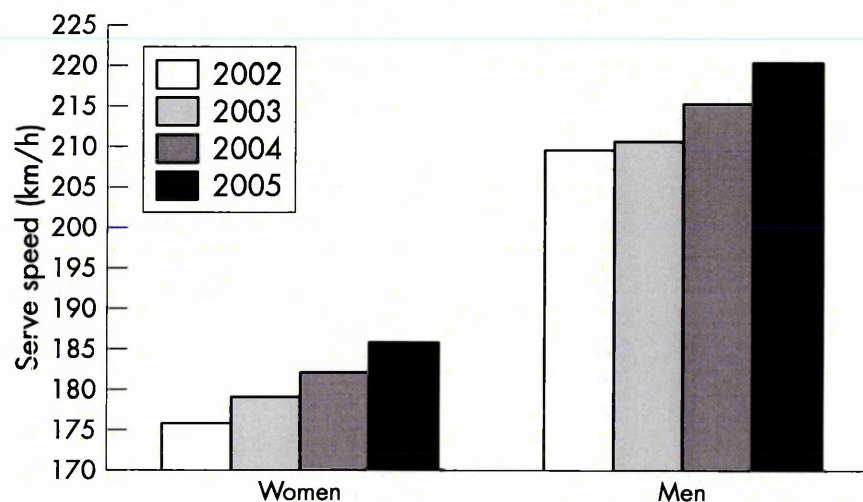


Figure 2.3. Average speed of 20 fastest servers at Grand Slam events since 2002 (taken from Miller, 2006).

Miller (2006) quantified the 20 fastest serve speeds recorded at Grand Slam events between 2002 and 2005 (Figure 2.3). Serve speeds increased year-on-year, highlighting the impact of racket development to serve speed (Miller, 2006). Haake *et al.* (2007) provide a historical account of technological developments to the racket and ball, in relation to tennis performance. Simulations by Haake *et al.* (2007) suggest that serve speeds have increased by 17.5% between the 1870s and 2007, with 25% of that improvement occurring since the 1970s. Furthermore, the time to react to an incoming serve would have fallen by 15% between the 1870s and 2007; again 25% of that reduction occurring since the 1970s. Surprisingly, until 1978 no rules existed concerning the racket or ball (Haake *et al.*, 2007). This highlights that in the past, rule changes by the ITF, e.g. Figure 2.4, have been reactionary in nature.

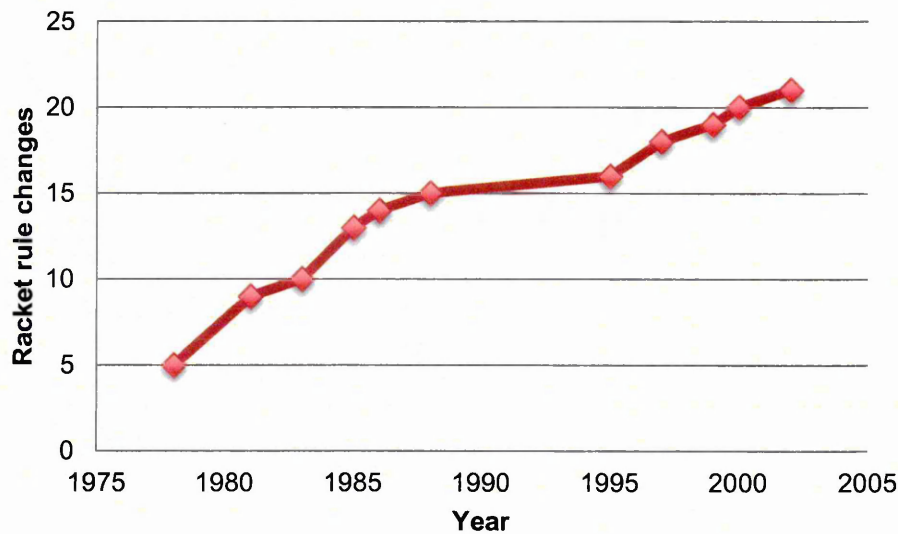


Figure 2.4. Rule changes to the racket from 1978 to 2002 (adapted from Haake *et al.*, 2007).

As a result of continual technological development in tennis, the role that the ITF performs has become a balancing act between preserving the sports' integrity and improving its appeal to the public. It is important that the ITF are research active in order to monitor and regulate tennis effectively. This 'research activity' can be evidenced by the three international congresses, 'Tennis Science & Technology', hosted by the ITF (2000, 2003 and 2007). Rule changes made by the ITF are now informed by research. The ITF's research activity also highlights a proactive role in the monitoring and assessment of new technologies that are introduced to the game of tennis.

2.1.3. *Tennis player movement and injury*

Technological developments have facilitated the increased pace of modern tennis (Haake *et al.*, 2007; Miller, 2006). However, increases to game pace also impact upon players and their on-court movement (Miller, 2006). Playing tennis, as participating in other sports, increases risk of injury due to physical exertion (Hjelm, Wener and Renstrom, 2010). Tennis injuries are commonly reported as overuse injuries or muscle and ligament strains and sprains, reflecting the various demands placed on anatomical structures (Bylak and Hutchinson, 1998). Indeed, tennis has a unique ‘injury profile’ when compared to other sports (Pluim *et al.*, 2006). However, tennis is an evolving sport. As previously noted, the ‘wooden racket era’ of tennis reflects a period when game style was characterised by style and finesse. At that time, injuries were predominantly to the hands and arms; injuries to the feet and back occurred less frequently and with lower severity (Frey, 1969). Following the introduction of aluminium, oversized rackets in 1975, the ‘modern era’ of tennis refers to a game now characterised by more powerful strokes, higher rates of ball spin and more athletic court movements (Fernandez, Mendez-Villanueva and Pluim, 2006). The ‘modern era’ of tennis, therefore, has different physiological requirements of players and as such, frequently injured sites differ to those of the ‘wooden racket era’.

In a recent review of 28 epidemiological tennis injury studies published between 1976 and 2005, Pluim *et al.* (2006) identified that the lower extremities now comprise the most frequently injured sites in tennis (31 – 67%), followed by the upper extremities (20 – 49%) and trunk (3 – 22%). Furthermore, the review highlights a progression from predominantly upper extremity injuries, (four studies), to lower extremity injuries, (23 studies; Pluim *et al.*, 2006). The review also noted that the nature of lower extremity injuries were predominantly acute injuries, in contrast to chronic, upper extremity injuries. This apparent shift in tennis injury profile again reflects the increased pace and intensity of ‘modern era’ tennis, highlighting the importance of understanding player-surface interactions.

Tennis movements typically consist of an initial split step followed by a combination of side steps and strides to reach an incoming ball (Hughes and Meyers, 2005). Within a rally, approximately 80% of strokes are played within 2.5 m of the players’ ready position; 10% of strokes are played between 2.5 – 4.5 m of the ready position and less

than 5% of strokes are played beyond 4.5 m of the ready position (Fernandez, Mendez-Villanueva and Pluim, 2006). During an average rally, players travel 8 – 12 m and change direction four times, constituting 300 – 500 high intensity efforts during a three set match (Fernandez, Mendez-Villanueva and Pluim, 2006).

However, little is known of the impact that tennis play on different court surfaces has on player-surface interactions (Miller, 2006). Notational analyses suggest that player movement patterns differ between grass, clay and acrylic tennis surfaces. O'Donoghue and Ingram (2001) identified that rallies played at Roland Garros (clay surface) were longer and consisted of more baseline play than rallies played at the US and Australian Opens (acrylic surfaces). Similarly, rallies at the US and Australian Opens were longer and consisted of more baseline play than rallies played at Wimbledon (grass surface). As such, court surface characteristics, e.g. COF, have been suggested to influence both player movement patterns and subsequently, player injury risk (Girard *et al.*, 2007).

The ratio of 'available' and 'utilised' COF describes the stability of the shoe-surface interface; if the ratio is greater than one, a slip should not occur (Redfern *et al.*, 2001). However friction models that assume friction is entirely a material property, e.g. Amontons-Coulomb model, are not appropriate during dynamic loading conditions such as human locomotion. Shoe-surface friction is a dynamic quantity, dependent on contact area, pressure, velocity, contact time and numerous other variables (Chang *et al.*, 2001). In 1991, Chapman *et al.* demonstrated that utilised COF on a squash court varied depending on foot contact type, e.g. heel or whole foot, and court surface contamination, e.g. dust or water. Chapman *et al.* (1991) controlled foot contact via stroke type, e.g. side-step or lunging forehand, and demonstrated that limiting friction could be exceeded (resulting in a slip) on dusty surfaces with a whole foot contact but not a heel contact, and on damp surfaces with a heel contact but not a whole foot contact.

Running and sliding are functional movement strategies in tennis (Fernandez, Mendez-Villanueva and Pluim, 2006). The ability to regulate utilised COF highlights the dynamic nature of player interactions with the court surface. In tennis, recent laboratory based research demonstrated that changes to utilised COF to enable sliding was the result of player movement coordination. Damm *et al.* (2013) measured kinetic and

kinematic data during forehand groundstroke manoeuvres performed on clay and acrylic surfaces. Damm *et al.* (2013) reported that for the clay surface utilised COF was higher than for the same manoeuvre performed on acrylic, despite the clay surface providing a lower slip resistance value in pendulum tests. Damm *et al.* (2013) suggested that in the case of clay, players attempted to elicit sliding as a result of prior experience. The regulation of friction demand in relation to mechanical surface properties reflects movement coordination of player-surface interaction. Research by Starbuck *et al.* (2013) supported the findings of Damm *et al.* (2013). Starbuck *et al.* (2013) compared lower-limb kinematics during forehand groundstroke manoeuvres performed by expert and novice players on clay and acrylic tennis courts. Larger sliding distances on the clay court reflected greater attack angles, e.g. B (Figure 2.5), particularly for experienced players (Starbuck *et al.*, 2013). This indicated that experienced players adopted a larger base of support to facilitate sliding.

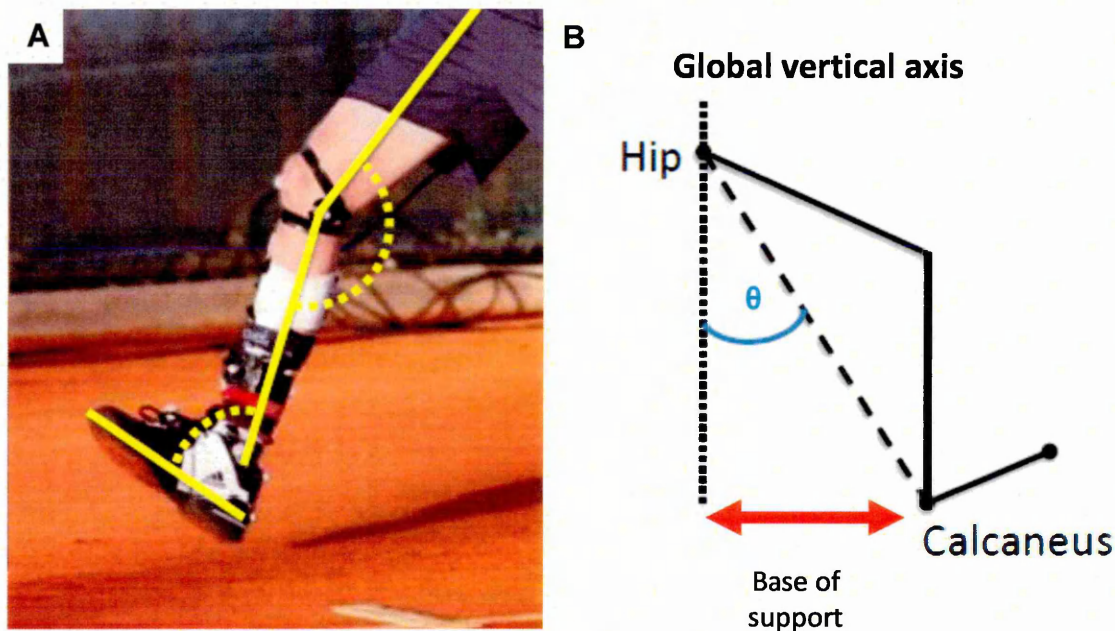


Figure 2.5. A: Manually digitised image. B: 'Attack angle' (θ) and base of support (adapted from Starbuck *et al.*, 2013).

The findings of Starbuck *et al.* (2013) suggest different step strategies between expert and novice forehand groundstroke manoeuvres performed on clay and acrylic tennis courts. However simulating tennis specific manoeuvres, in a laboratory or the field, can be problematic; measured data can be questionable or invalid. For example, Stiles and Dixon (2006) found no differences in peak ground reaction force (vertical and

horizontal) for a laboratory analysis of the tennis forehand foot plant, despite distinctly different mechanical characteristics between carpet, acrylic and artificial turf. Furthermore, when repeating Girard *et al.*'s (2007) field based analysis of plantar pressure for specific manoeuvres on different tennis court surfaces, Eckl, Kornfeind and Baca (2011) reported limited success. In particular, Eckl, Kornfeind and Baca (2011) could not confirm higher loads under the hallux and toe regions observed by Girard *et al.* (2007) for manoeuvres performed on acrylic court surfaces. Eckl, Kornfeind and Baca (2011) suggested that participants in their study did not perform movements as aggressively as participants assessed by Girard *et al.* (2007). Regardless, the validity of existing field- or laboratory-based studies cannot be confirmed because spatio-temporal parameters of gait during match-play tennis have not been quantified.

2.1.4. Conclusion

Tennis is an international sport that attracts millions of players and spectators worldwide. As a result of technological developments in tennis, the ITF has become increasingly research active in order to effectively monitor and regulate the game. Notational analyses have demonstrated that player movement strategy is affected by court surface; however player-surface interactions are not well understood. Laboratory and field based research has revealed important characteristics of tennis manoeuvres; however the validity and repeatability of current research is questionable. A logical first step in the process of characterising player-surface interactions would be to measure tennis player step and movement characteristics during match-play tennis. Player step and movement characteristics could then be used to inform field- or laboratory-based research.

2.2. *Motion analysis techniques*

The aim of current work is to develop a motion analysis tool designed to measure player location and foot-surface contacts during match-play tennis. Prior to an assessment of suitable motion analysis techniques (section 2.4), consideration must be given to the data collection environment. The ITF determine the rules of tennis (ITF Rules of Tennis, 2013). The rules are explicit and state that players, playing conditions or anything within the field-of-vision of players cannot be interfered with during match-play. Motion analysis aims to provide objective measurements of position. Motion analysis techniques can be categorised into 'intrusive' and 'non-intrusive' techniques:

- Intrusive techniques: Motion analysis where the player or environment is altered in some way. This might be a laboratory environment or an application of one or more markers to a performer or their equipment in a realistic setting.
- Non-intrusive techniques: Motion analysis where the player or environment is not altered. Non-intrusive techniques are desirable as they can preserve the validity of measurements; however measurement accuracy can be reduced.

Motion analysis techniques that would interfere with players during match-play, e.g. electromagnetic sensors, inertial sensors, accelerometers or markers (Figure 2.6), would violate the rules of tennis and therefore cannot be used (ITF Rules of Tennis, 2013). A new rule on performance analysis technology (ITF, 2013) may make this possible in the future; however, in order to ensure that the player's actions or the environment are not altered, only non-intrusive motion analysis techniques will be considered.

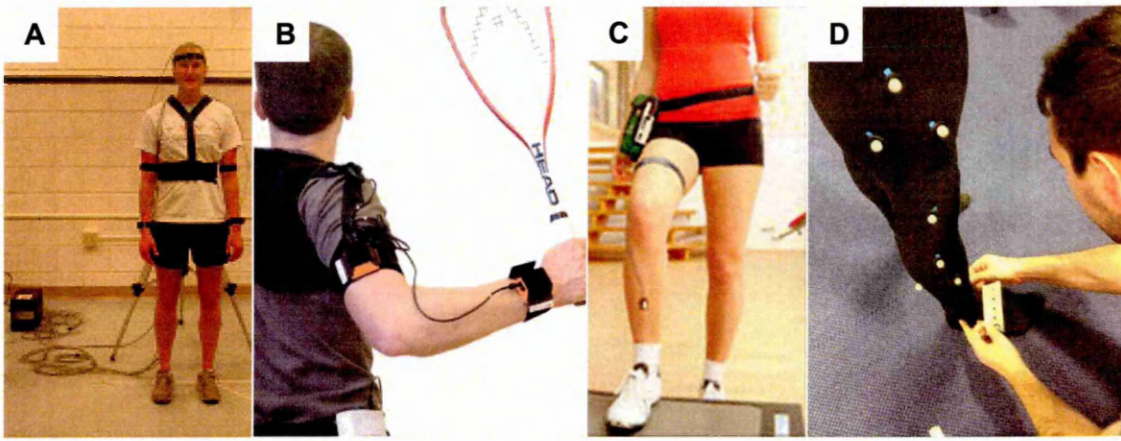


Figure 2.6. The application of electromagnetic sensors (A), inertial sensors (B), accelerometers (C) and markers (D) is intrusive and would violate rules of match-play tennis.

2.2.1. Non-intrusive motion analysis technologies in tennis

In 1980, the Cyclops line calling system became the first non-intrusive motion analysis system to be used during match-play tennis (Pallis, 2004). For serves that landed close to service lines, Cyclops would determine whether the ball landed in or out of the service box. Cyclops achieved this by monitoring five infrared beams aligned with service lines; if one of the beams behind the service line was broken the serve was called out.

Cyclops was superseded in 2006 by the multi-camera line-calling system, Hawk-Eye. Hawk-Eye was patented in 2001 (Sherry and Hawkins, 2001) and improved upon Cyclops by identifying ball-surface location in relation to any court line with high accuracy. The ITF reported maximum uncertainty values of ± 0.9 mm in an initial validation in 2004 (ITF, 2010). Using high-speed photogrammetry (Capel-Davies and Miller, 2007), the ITF have deemed Hawk-Eye accurate enough to officiate in tennis. Hawk-Eye was first used to officiate at match-play tennis in 2005 and is now used at nearly every major tennis tournament (ITF, 2010). Hawk-Eye uses between four and ten fixed cameras to track the three-dimensional trajectory of the tennis ball in relation to the tennis court (Hawk-Eye, 2013). The triangulation of individual camera data, e.g. two-dimensional ball position, is used to calculate ball position in relation to the tennis court (used as a calibration object); this can then be used to determine ball-surface contact area used to officiate line-calls. Analyses also account for ball deformation, rolling and slip during impact. Hawk-Eye also provides a visual representation of data

used in line-calling decisions (Figure 2.7). Visualisations are powerful representation of the complex analyses being performed to determine a simple line-call. Visualisations are presented to players and spectators alike during match-play, providing a novel and engaging element to match-play tennis.



Figure 2.7. ‘Official Review’ graphic presented to players and the public.

Hawk-Eye has subsequently been applied to other sports including cricket, snooker and most recently football (Hawk-Eye, 2013). In snooker, Hawk-Eye is used to enhance television broadcast material (Hawk-Eye, 2013). In cricket, Hawk-Eye is used both to enhance television broadcast material and assist umpiring decisions regarding the leg before wicket rule (Hawk-Eye, 2013). Most recently Hawk-Eye has been licenced by FIFA to install Goal Line Technology systems worldwide (Hawk-Eye, 2013). Hawk-Eye is a powerful, non-intrusive ball-tracking system. However the system has not been reported to measure foot-surface contacts during match-play tennis. Hawk-Eye has reported the ability to track player location during match-play tennis (Hawk-Eye, 2013). However, little information is known about the methods used to track player location. Furthermore, the accuracy of player location data has not been reported. Hawk-Eye does not provide access to their analysis systems to external researchers: therefore additional analysis tools cannot be developed, e.g. foot-surface contact measurement. Finally, Hawk-Eye systems are reliant on multiple cameras that are fixed in typically inaccessible locations. This would limit the use of any system that was developed.

2.2.2. Non-intrusive player motion analysis

Prozone is a multi-camera player tracking system used in football (Di Salvo *et al.*, 2006). Eight cameras are fixed to provide full coverage of the football pitch; cameras

are initially calibrated using a linear four-point calibration transformation and then refined using a proprietary, 50-point algorithm to reduce errors induced by lens distortions (Di Salvo et al, 2006). Prozone identifies and tracks players based on shirt colour and player movement thresholds, e.g. velocity. Furthermore, Prozone is capable of resolving complex player tracking problems during unpredictable team sports such as football (Di Salvo *et al.*, 2006). Di Salvo *et al.* (2006) validated Prozone using players running through a predefined system of shuttle runs; Prozone player velocity estimates were compared to velocities derived from timing gate data. Di Salvo *et al.* (2006) reported high player tracking accuracy; mean velocity error was $0.127 \text{ m}\cdot\text{s}^{-1}$, concluding that the system allowed the real-time tracking of multiple players without requiring special equipment, e.g. transmitters or colour coded clothing. Prozone is predominantly used in football for post-match analyses of player movement strategy; Prozone has not been reported to measure a player's step strategy. Furthermore the restricted access and high-cost of multiple camera systems such as Prozone (Di Salvo *et al.*, 2006) limit the development of bespoke applications.

SAGIT is a computer tracking system that automatically (albeit with operator supervision) tracks player position using video sequences (Perš *et al.*, 2002). The SAGIT system has been applied to different sports including handball (Perš *et al.*, 2002), squash (Vučković *et al.*, 2010) and tennis (Martínez-Gallego *et al.*, 2013). The SAGIT system uses video footage (384×288 pixels) recorded from two standard cameras fitted with wide-angle lenses, located 10 m above a court, e.g. Figure 2.8. Following background subtraction, the SAGIT system segments and tracks players based on colour and shape matching. Player location was converted from pixel to real-world coordinates by first accounting for lens distortion, position was then determined by scaling and translating coordinates between the sensor and court plane (Perš *et al.*, 2002). When compared to manual digitising, Perš *et al.* (2002) reported root-mean square-errors (RMSE) of less than 0.6 m for player location on a handball court. However Vučković *et al.* (2010) demonstrated that, when performing realistic movements, e.g. unknown player locations or movement paths, error in player distance could range between 1 and 21 m (~10%) in one minute. Player position, estimated by the SAGIT system, was therefore dependent on the nature and location of player movement and not desirable for the analysis of match-play tennis.

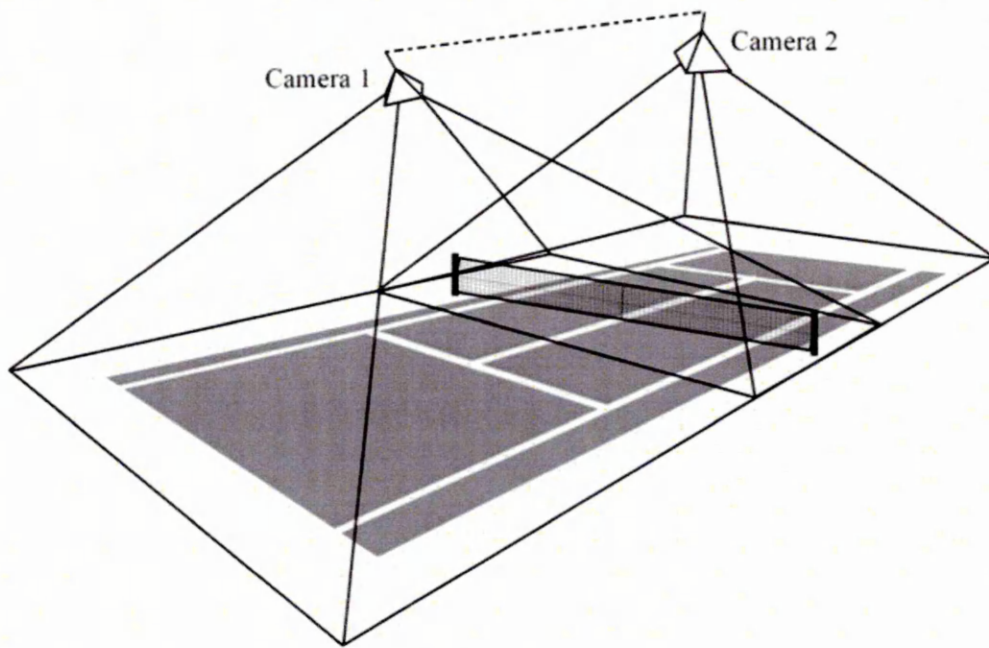


Figure 2.8. Camera locations used for analysis by SAGIT (adapted from Martínez-Gallego *et al.*, 2013).

TennisSense is a multi-camera system for player tracking and extraction of semantic information in tennis (Conaire *et al.*, 2009). Conaire *et al.* (2009) tracked tennis balls and players as well as detecting when balls were struck, providing coaches with a method for reviewing player movement and stroke choice. The system obtained footage from nine networked, pan, tilt and zoom cameras. Four cameras located at each end of the court (eight in total), e.g. Figure 2.9, were used for coaching / revision purposes; an overhead camera was used to perform all analyses. To identify player location, Conaire *et al.* (2009) used a background subtraction approach using images from the overhead camera. A background model was developed using an adapted Stauffer and Grimson (1999) model to account for lighting fluctuations and to suppress shadow; foreground regions were then extracted. Ball trajectories were identified by inter-frame subtraction; motion blur helped identify correct ball candidates. Ball strikes were identified by comparing the direction and crossing time of incoming and outgoing ball trajectories. A trajectory break of half a second was considered a 'ball strike' as this reflected the ball being classified as a foreground player region. To convert pixel coordinates, e.g. ball and player, to real-world coordinates, Conaire *et al.* (2009) corrected for lens distortions and calibrated cameras using the OpenCV camera calibration library.

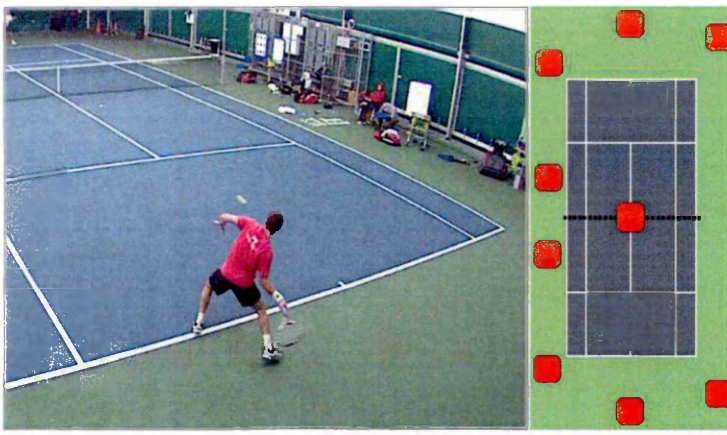


Figure 2.9. Screenshots from TennisSense illustrating a rear camera view and fixed camera filming locations (red circles about tennis court; adapted from Conaire *et al.*, 2009).

However, Conaire *et al.* (2009) applied a box filter to player foreground images. The box filter had a fixed resolution of 29×29 pixels; image resolution was 640×480 pixels. While this reduced processing time, the box filter would have affected the accuracy of player position estimates. Using a UbiSense 3D position tag-tracking system, Conaire *et al.* (2009) reported median player position errors of 1.10 and 1.09 m for each player. Conaire *et al.* (2009) demonstrated that the TennisSense tracked tennis player position and identified ball strike events in tennis matches. However player position estimates were limited and only one of nine cameras were used for analysis.

In addition to player tracking, Connaghan, Moran and O'Connor (2013) used the TennisSense system to perform content annotation using images from rear camera views, e.g. Figure 2.10. Connaghan, Moran and O'Connor (2013) automatically identified events such as services, forehand / backhand strokes and change of ends with an average precision and recall rate of 0.84 and 0.86 respectively. This reduced the need to manually index and annotate key events in tennis matches. Feature extraction, such as serve and forehand / backhand classification, was performed as described by Conaire *et al.* (2010). The foreground image was first divided into 16 pie segments centred about the player centroid, e.g. C (Figure 2.10). The largest distances of foreground pixels in each segment (relative to player centroid) were then recorded throughout entire strokes; this allowed for feature extraction. However features were not invariant of viewing angle; only rear camera views were suitable, e.g. A (Figure 2.10).

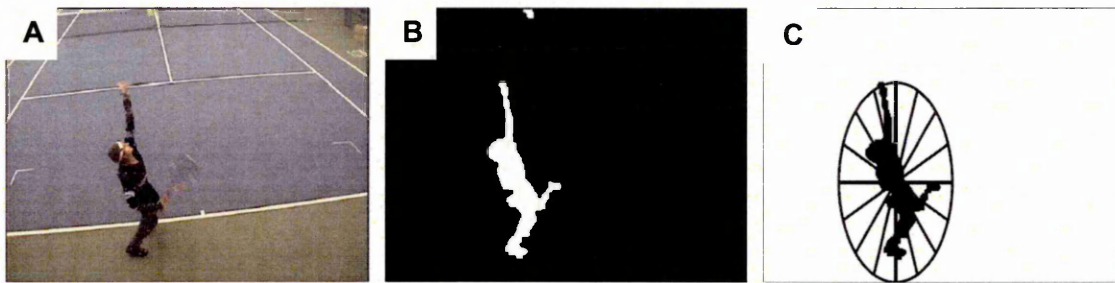


Figure 2.10. Feature extraction: camera image (A), foreground image (B) and radial map (C: adapted from Connaghan, Moran and O'Connor, 2013).

Automatic annotation performed by Connaghan, Moran and O'Connor (2013) identified key tennis events considered useful for tactical analyses and coaching. Whilst event annotation is useful for coaching purposes, player position estimates will be limited by the use of the player tracking method described by Conaire *et al.* (2009). Connaghan, Moran and O'Connor (2013) did not analyse player step strategy; this was not the focus of their work. However, images used for feature extraction, e.g. tennis stroke annotation, might allow the application of gait measurement techniques (described in section 2.3). However Conaire *et al.* (2010) noted that extracted stroke features were not view invariant. View invariance might also limit the success of gait measurement, e.g. multidirectional player movement (Fernandez, Mendez-Villanueva and Pluim, 2006); however image requirements to measure gait in match-play tennis are not yet known.

Tracking player position without interfering with play has been demonstrated by numerous systems, e.g. Di Savlo *et al.* (2006); Martínez-Gallego *et al.* (2013); Conaire *et al.* (2009); Connaghan, Moran and O'Connor (2013). Systems such as SAGIT and TennisSense represent lower cost and more accessible player tracking technologies than Prozone; however limitations exist for each system. Conaire *et al.* (2010) highlighted that camera view was an important consideration when developing a multi-purpose camera system, e.g. player tracking and feature extraction. No camera system for tracking players has measured player gait strategy. Gait parameters cannot be measured using an overhead camera. The feasibility of camera systems should therefore be assessed in relation to the operational environment.

2.2.3. *Sport stadia*

System development must focus on typical and accessible filming locations to increase the flexibility of future match-play tennis analyses. An unobtrusive approach to data collection must be adopted in order to conform to the rules of tennis (ITF Rules of Tennis, 2013). Overhead camera locations are not typically accessible, particularly at match-play tennis events. Furthermore, it is impracticable to physically install a fixed camera system at every tennis venue. Match-play tennis is typically performed in front of a public audience who require seating. Public stadia seating therefore represent the most typical and accessible filming locations.

Basic assumptions about filming conditions can therefore be made. Sport stadia provide seating that maximises spectator capacity without compromising an individual spectator's view. Seating elevation increments typically range between 60 – 150 mm (John, Sheard and Vickery, 2007) to preserve spectator sightline, e.g. B (Figure 2.11). Furthermore, sport stadia for tennis are typically circular, providing 360° viewing angles, e.g. Centre court and No. 1 court; A (Figure 2.11). The rigid transform of a world coordinate $[X, Y, Z]^T$ to a camera coordinate $[x, y, z]^T$, e.g. tennis court and camera coordinate systems, can be expressed as:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos(az) & \sin(az) & 0 \\ -\sin(el) \times \sin(az) & \sin(el) \times \cos(az) & \cos(el) \\ \cos(el) \times \sin(az) & -\cos(el) \times \cos(az) & \sin(el) \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} r \times \cos(el) \times \cos(az) \\ r \times \cos(el) \times \sin(az) \\ r \times \sin(el) \end{bmatrix} \quad [2.1]$$

where az is camera azimuth angle, el is camera elevation angle and r is camera radius.

Rotations about the optical axis, e.g. camera tilt, in equation 2.1 are ignored. However camera azimuth, elevation and radius are physical constraints determined by sport stadia. Equation 2.1 can be used to assess basic filming requirements.

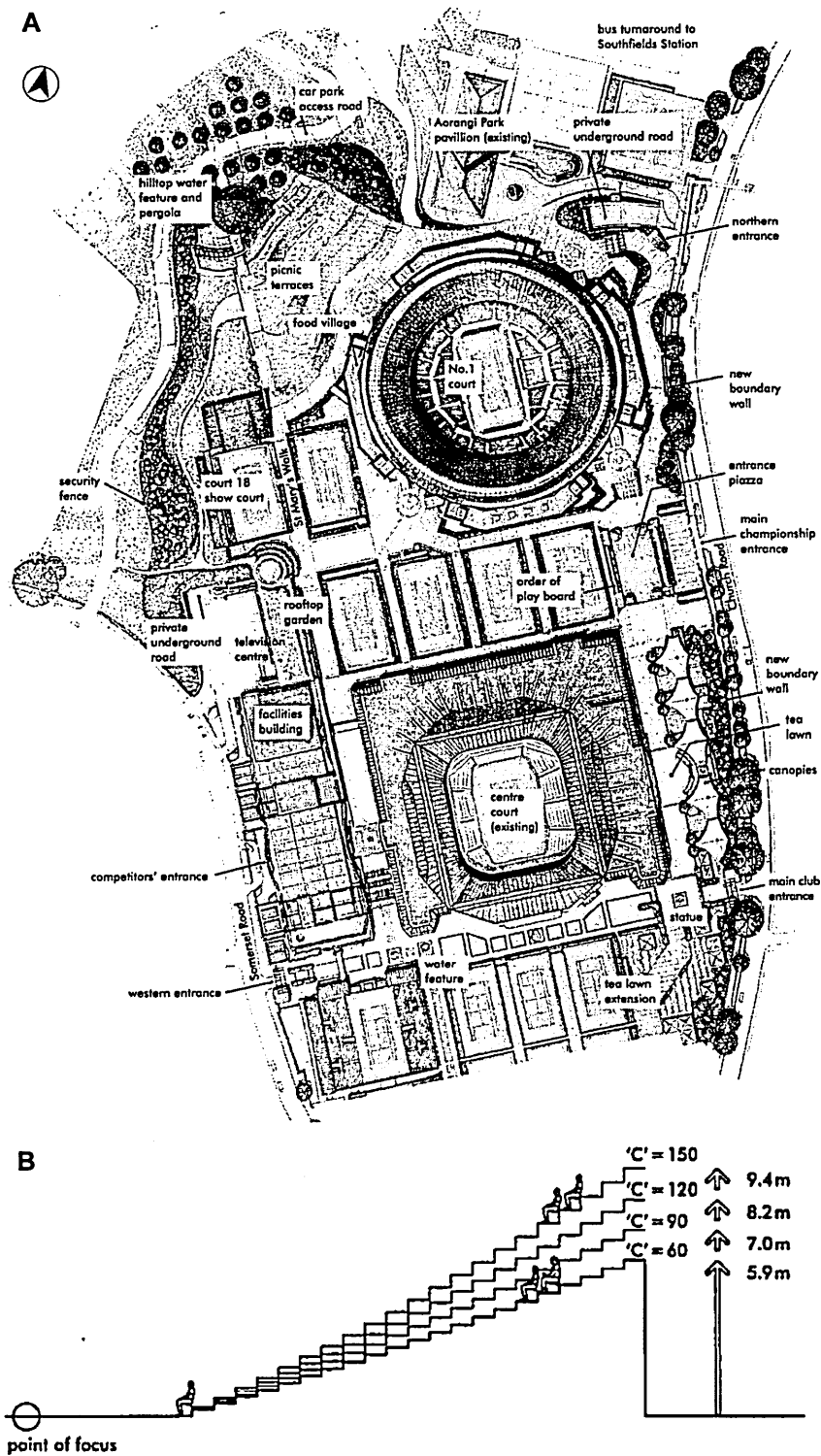


Figure 2.11. A: The All-England Lawn Tennis and Croquet Club at Wimbledon. B: Incremental elevation of stadia seating (adapted from John, Sheard and Vickery, 2007).

Photogrammetric systems require calibration. In the case of stereoscopic calibration, i.e. multi-camera systems, calibration accuracy is dependent on the angle at which the optical axes of cameras intersect. Chen and Davis (2000) demonstrated that the optimal camera intersection angle for stereoscopy was 90° ; however 3D reconstruction uncertainty was acceptable for camera intersection angles of $40 - 140^\circ$. Camera intersection angles beyond this range would result in poor target resolution and introduce error into the measurement system. To perform stereoscopic calibration, time synchronised or genlocked camera images are required. A typical viewing distance, i.e. radius from point of focus to spectator (B: Figure 2.11), is 30 m (John, Sheard and Vickery, 2007). The arc, given by $\theta \times r$, provides the length of camera cabling required to link cameras. To record synchronised camera images from two cameras, between 20.9 and 77.3 m of generator locking (genlock), mains power and camera data cabling would be required for the camera intersection angles of $40 - 140^\circ$. In the case of multi-camera systems using networked cameras, similar cabling constraints would exist.

At competitive sport events, event organisers require that all stadia seating, walkways and fire escapes remain accessible. Camera systems requiring 20 – 80 m of cabling placed around sport stadia would be impracticable to implement and potentially unsafe. Single camera filming is not restricted in the same way as multi-camera systems and represents the most flexible approach for sport stadia filming. Therefore single camera filming is a more feasible approach for developing a motion analysis tool to measure player location and foot-surface contacts during match-play tennis.

Using single camera, television broadcast footage, Yan, Christmas and Kittler (2005) developed an image processing algorithm to track tennis ball and player movements. The authors used background subtraction to identify player location; ball location was identified using inter-frame subtraction. Impact events, e.g. ball-racket and ball-surface contacts; were also extracted to enhance broadcast footage, e.g. Figure 2.12. The approach was suitable for basic spatial analyses of match-play tennis. However the algorithm was developed to optimise its application to broadcast footage of varying quality; the accuracy of measurements was not prioritised or assessed.

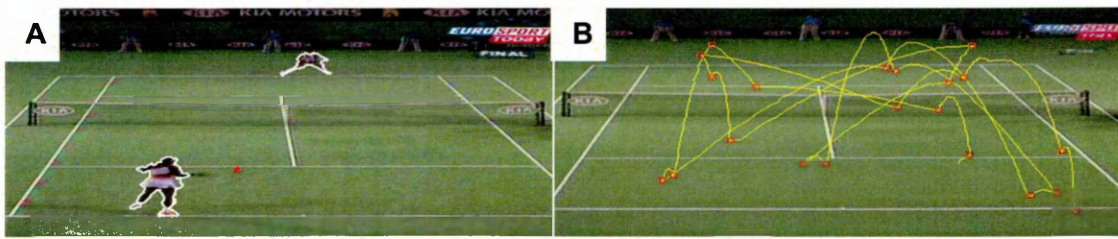


Figure 2.12. Player (A) and ball (B) tracking from broadcast video (adapted from Yan, Christmas and Kittler, 2005).

Jiang *et al.* (2009) presented a novel tennis player tracking algorithm that used television broadcast footage; tracking accuracy was also assessed. Jiang *et al.* (2009) extracted foreground player regions using the Hue and Value channels of HSV colour space images; the authors reported this to be more successful than processing RGB colour space images. Jiang *et al.* (2009) determined player position using the centroid of foreground player regions; the authors did not account for camera perspective. Jiang *et al.* (2009) manually identified ground truth player position data and reported ‘close’ results for automatic tracking. However, player position estimates would not have represented player position on-court, e.g. perspective error. Player position measured by Yan, Christmas and Kittler (2005) would have been similarly limited by perspective error. The accuracy of player position measurements obtained from broadcast footage is unlikely to be suitable for research purposes. Furthermore, the focus of analyses could not be controlled because of the nature of broadcast footage; this would be disadvantageous.

Mauthner *et al.* (2008) presented a single camera algorithm for tracking beach volleyball players during competition. A static, low elevation camera, e.g. Figure 2.13, was used to identify and track players without applying markers. Court line intersections were used to identify the Homography between image and world coordinates, enabling position reconstruction from perspective camera images. Mauthner *et al.* (2008) used an integral histogram tracker (HSV colour space) to identify player's shirts from a background model. To identify player position, Mauthner *et al.* (2008) subsequently extracted upper and lower-limbs using the YCbCr colour space. This was necessary due to the similarity of skin and sand pixel values in the RGB colour space. The extraction of the lower-limbs enabled Mauthner *et al.* (2008) to

estimate the player's ground position, e.g. B (Figure 2.13), and not violate planarity assumptions of two-dimensional motion analysis.

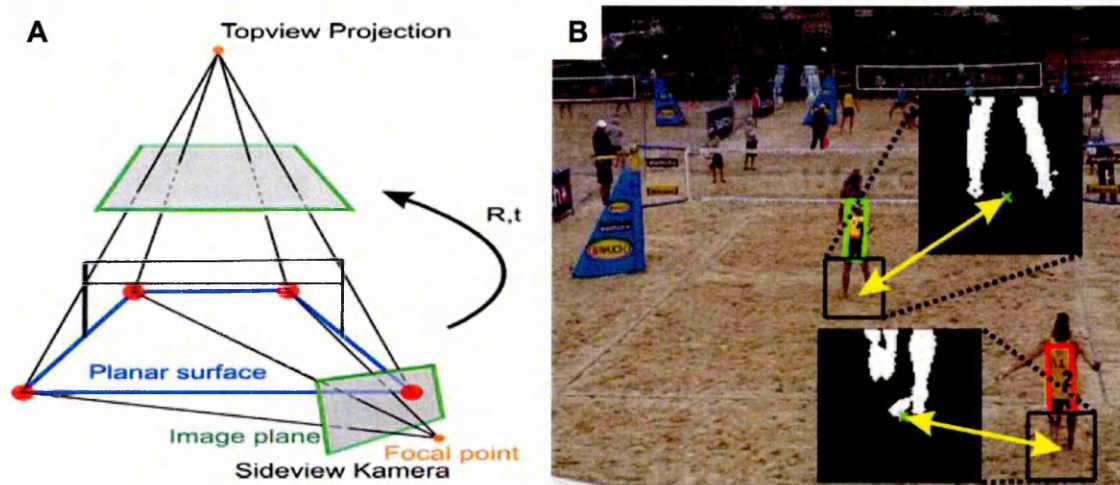


Figure 2.13. A: Perspective camera filming location. B: Player foreground segmentation and identified player location (green cross) used for player position tracking (adapted from Mauthner *et al.*, 2008).

Mauthner *et al.* (2008) applied the algorithm to twelve sequences of competitive beach volleyball and compared results to manually digitised estimates. Mauthner *et al.* (2008) reported tracking errors between 225 and 350 mm concluding that player tracking accuracy was suitable for sport science applications. However player position estimates were dependent on the perspective projection of lower-limbs, e.g. multi-directional nature of volleyball (Mauthner *et al.*, 2008). Furthermore, player jumps (out-of-plane projection error) resulted in unrealistic player trajectories, e.g. Figure 2.14. However Mauthner *et al.* (2008) highlighted that jumping errors could be detected and corrected due to the parabolic nature of position errors.

Mauthner *et al.* (2008) demonstrated that player position could be tracked to an acceptable level of accuracy using a single camera during competition. Camera elevation was low, e.g. views typical of stadia seating (Figures 2.11 and 2.13), suggesting that a similar approach could be applied to match-play tennis. It was not the aim of Mauthner *et al.* (2008) to perform player step analysis. However segmentation was considered reliable enough for lower-limb features to be used to identify player location. This indicates that it might be possible to extract lower-limb images of sufficient quality for feature extraction and gait measurement (described in section 2.3).

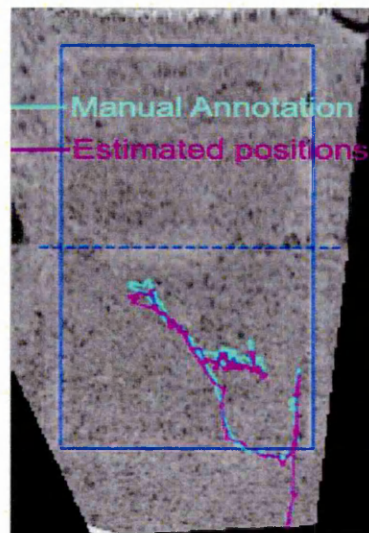


Figure 2.14. Manual and automatic estimates for player position: a spike in estimated position (bottom right) reflects a jumping motion (adapted from Mauthner *et al.*, 2008).

2.2.4. Conclusion

Motion analysis aims to provide objective measurements of position. Motion analysis techniques that interfere with players during match-play violate the rules of tennis; only non-intrusive motion analysis techniques can be used. Non-intrusive motion analysis technologies are predominantly camera based; many multi-camera systems have been developed to track player position for the purpose of match analysis and coaching. However it is impracticable to install multi-camera systems at every match-play tennis event; single camera filming represents a flexible approach to filming in sport stadia. Non-intrusive, single camera approaches are capable of extracting ball and player location as well as key tennis events. Furthermore, player position can be measured to an acceptable level of accuracy using low elevation, single camera footage of

competitive sport. However, player position was dependent on the perspective projection of the lower-limbs. Furthermore, feature extraction for the purpose of gait measurement in sport has not been described; image requirements for measuring gait in match-play tennis are therefore not known.

2.3. *Gait measurement techniques*

Gait recognition developed from the works of Kozlowski and Cutting (1977). Kozlowski and Cutting (1977) demonstrated that people could identify others using information derived solely from a person's gait, i.e. point-light display (Figure 2.15).

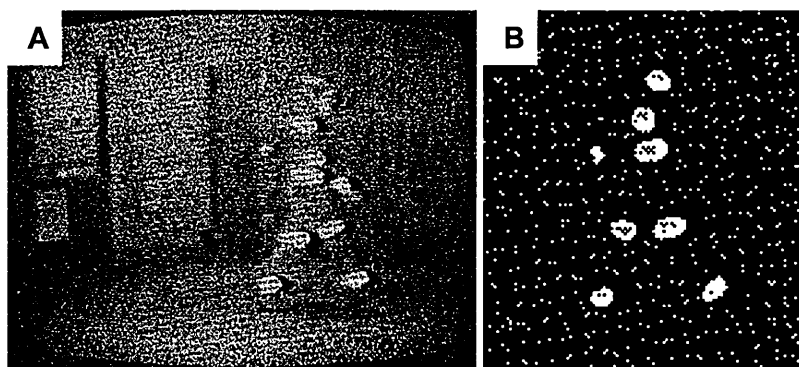


Figure 2.15. A: Reflective tape on anatomical landmarks is floodlit. B: Point-light display (adapted from Kozlowski and Cutting, 1977).

Cunado *et al.* (1999) extended this work and highlighted the extraction and use of gait as a biometric signature. The extraction of gait features can be classified by two main approaches: holistic and model-based feature extraction. Model-based approaches model either the person or walk of the person whereas holistic approaches model the shape or motion of a person as they walk (Johnson and Bobick, 2001).

2.3.1. *Holistic approaches*

Holistic approaches to gait recognition extract feature vectors that provide a descriptive account of a gait sequence. A central advantage to holistic approaches is the application to gait sequences without requiring a model of the person (Boulgouris, Hatzinakos and Plataniotis, 2005). As such, holistic approaches can be applied directly to an image sequence. Numerous classes of holistic gait recognition exist including:

- Optical flow, e.g. Little and Boyd (1998)

- Silhouette contour, e.g. Wang *et al.* (2003), width, e.g. Kale *et al.* (2003) and radius, e.g. angular transform (Boulgouris, Plataniotis and Hatzinakos, 2004)
- Self-similarity, e.g. BenAbdelkader, Cutler and Davis (2004)
- Space-time saliency, e.g. Gorelick *et al.* (2007)
- Frequency transformation of feature vectors, e.g. Lee and Grimson (2002)
- Dimensionality reduction, e.g. principal component analysis (Kale *et al.*, 2003)

However, competitive tennis includes movements such as walking, sidestepping, lunging and running (Hughes and Meyers, 2005). Holistic gait recognition techniques, which derive feature vectors, e.g. Figure 2.16, have predominantly been applied to walking. This is due to the periodic nature of gait, allowing for feature extraction, e.g. double support. Boulgouris, Hatzinakos and Plataniotis (2005) highlight that disadvantages to holistic gait recognition can include high algorithm complexity, low robustness, coarse feature representation (low resolution) and difficult movement phase determination. It would therefore be difficult to apply holistic gait recognition to match-play tennis footage. Furthermore, the measurement of player step strategy using a holistic approach is unclear. Therefore holistic gait recognition approaches are unlikely to be suitable given the context of the current study.

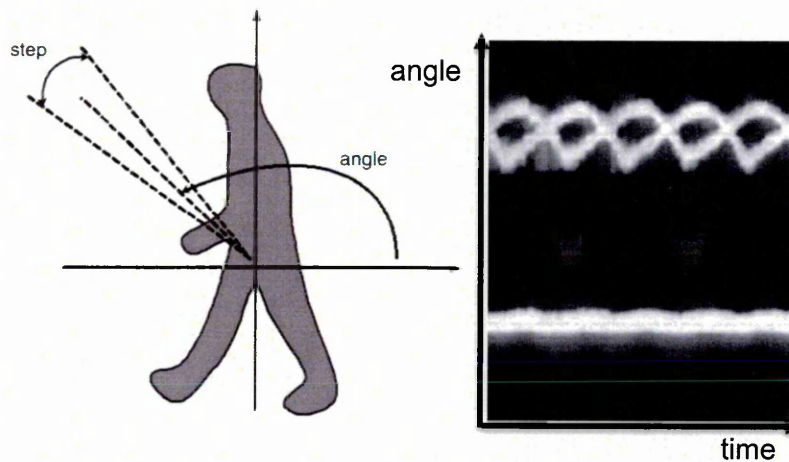


Figure 2.16. Holistic gait recognition features (angular transform) derived from a silhouette (adapted from Boulgouris, Plataniotis and Hatzinakos, 2006).

2.3.2. Model-based approaches

In general, model-based approaches for gait analysis are view and scale invariant (Boulgouris, Hatzinakos and Plataniotis, 2005). This is advantageous because gait sequences are unlikely to be captured from the same camera perspective or distance. For fixed camera footage, large and unpredictable tennis player movements favour the use of a model-based approach. Johnson and Bobick (2001) identified static body parameters by segmenting a silhouette. Johnson and Bobick (2001) measured height, head-to-pelvis distance, pelvis-to-foot distance (both feet) and inter-foot distance. Measurements were updated on each subsequent frame and applied to a variety of camera views. This yielded a view-independent method for determining foot separation distance and calculating gait parameters, e.g. Figure 2.17.

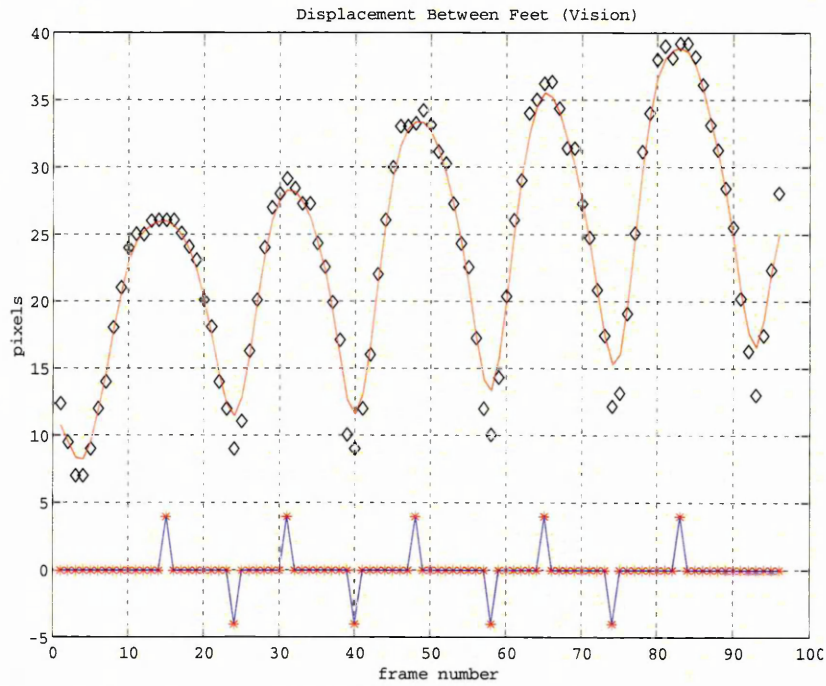


Figure 2.17. Inter-foot distance (diamonds: top) and tertiary signal (asterisks: bottom) highlighting inter-foot distance maxima (positive) and minima (negative). Distance maxima indicate dual-stance (adapted from Johnson and Bobick, 2001).

However Johnson and Bobick (2001) used silhouette height as a scaling factor to convert from image to world measurements. Low-level photogrammetry, e.g. direct image-to-world scaling, limits the accuracy of identified stride parameters. Indeed,

camera calibration represents a central limitation to many model-based, gait measurement techniques (Boulgouris, Hatzinakos and Plataniotis, 2005).

Goffredo *et al.* (2008) used a different approach to segment walker silhouettes. Based on anatomical studies, Goffredo *et al.* (2008) determined that the hip, knee and ankle joints for a human silhouette were located at 50, 25 and 10% of silhouette height respectively (from lowest silhouette pixel). Goffredo *et al.* (2008) refined silhouette segmentation for oblique-sagittal gait measurement by computing the angle at which straight line walking occurred relative to the camera, e.g. Figure 2.18. Goffredo *et al.* (2008) subsequently computed hip and knee joint angles during walking using six unique camera perspectives (azimuth: 0, 20 and 40°, elevation: 0 and 15°). Hip and knee joint angle data were then corrected using a perspective transform (determined from walking direction) to yield sagittal plane data, regardless of camera perspective.

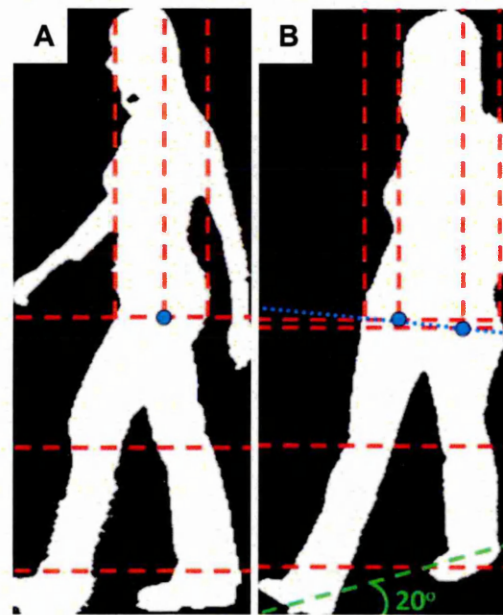


Figure 2.18. Anatomical silhouette segmentation for sagittal (A) and oblique-sagittal (B) walking (adapted from Goffredo *et al.*, 2008). Red and green dashed lines illustrate anatomical segmentation and walking direction respectively.

Goffredo *et al.* (2008) reported good accuracy, citing mean RMSE for joint locations as 1.4% of image resolution. Furthermore, hip and knee joint angle data were correlated (greater than 0.9) to corresponding joint angle data obtained from markers. However to enable the perspective transform of extracted data, Goffredo *et al.* (2008) assume that

travel occurs in a straight line for a minimum of four steps. Tennis movement consists of multi-directional manoeuvres that vary in step number. It is therefore unlikely that extracted joint angle data in tennis could be appropriately transformed to perform analyses. Furthermore, foot region segmentation might be problematic for camera perspectives greater than 40° to the players' direction of travel and result in tracking failure during tennis. Figure 2.18 (B) illustrates that using the lowest coordinate to segment feet regions was not appropriate for the rear foot, e.g. red horizontal dashed line. Foot segments might therefore require individual segmentation.

The extraction of gait features using model-based approaches has also been performed on running. Yam, Nixon and Carter (2002) applied a simple edge detection algorithm to walking and running. An initial Sobel edge detector retrieved edges about the lower-limbs, which were subsequently sorted in relation to the images horizontal coordinate, i.e. leading edge for right-to-left motion. A predefined model was fitted to leading edge data to extract corresponding hip and knee joint angle data for left and right legs, e.g. Figure 2.19.

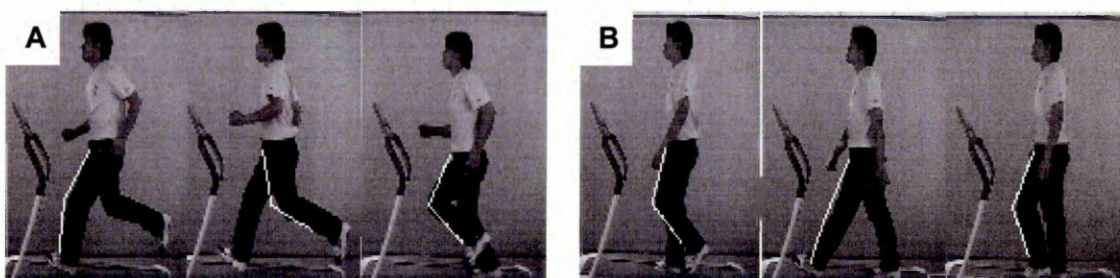


Figure 2.19. Model-based recognition of running (A) and walking (B), joint angle data are derived using the leading edge (highlighted in white) of lower-limbs (adapted from Yam, Nixon and Carter, 2002).

However, the model-based approach used by Yam, Nixon and Carter (2002) only considered sagittal camera views (Figure 2.19). It is therefore unlikely that the application of Yam, Nixon and Carter's (2002) approach to tennis would be suitable. Using model-based approaches, gait features can be automatically extracted from video sequences for walking, e.g. Johnson and Bobick (2001); Goffredo *et al.* (2008); Yam, Nixon and Carter (2002), and running, e.g. Yam, Nixon and Carter (2002). Furthermore, gait features can be extracted independently of camera view, e.g. Johnson and Bobick

(2001); Goffredo *et al.* (2008). However the view-independent reconstruction of joint angle data is complex, e.g. Goffredo *et al.* (2008). The aim of current work is to quantify player step strategy; this can be achieved by the measurement of heel strikes.

2.3.3. Model-based heel strike detection

Bouchrika and Nixon (2006) presented a model-based method for extracting joint locations. Heel strike locations were also extracted to estimate gait periodicity and present joint angle data in relation to gait cycle, e.g. Figure 2.20. Bouchrika and Nixon (2006) used an adaptive background subtraction model (described by Stauffer and Grimson, 1999) to extract foreground silhouettes. To extract heel strikes, Bouchrika and Nixon (2006) collated Harris corners from foreground silhouettes throughout walking image sequences. Harris corners were accumulated into a two-dimensional proximity matrix; due to walking periodicity, cluster centres identified heel strikes, e.g. B (Figure 2.20). When compared to manually identified heel strikes, Bouchrika and Nixon (2006) reported that mean identification error was 0.52% of body height; data were not converted to real-world measurements. The approach presented by Bouchrika and Nixon (2006) would enable the calculation of basic spatial parameters of gait, e.g. step length. However, the sequential accumulation of Harris corner data removed time domain information. As such, temporal parameters of gait, e.g. step rate, could not be calculated. Furthermore, heel strikes occurring in the same location cannot be differentiated from previous or subsequent heel strikes, e.g. walking repeatedly along the same path.

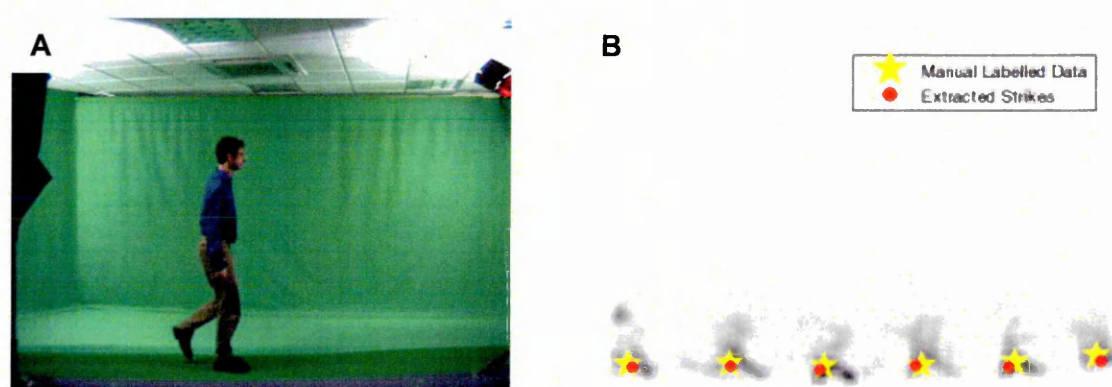


Figure 2.20. A: Sagittal camera view of walking. B: Corner proximity image for heel-strike extraction (adapted from Bouchrika and Nixon, 2006).

Jung and Nixon (2013) developed a novel approach for detecting heel strikes. Using a background image, foreground pixels were extracted using a pixel intensity and colour differencing technique described by Cheung, Kanade, Bouguet and Holler (2000). Pixel colour differencing accounted for shadow present in images. Foreground pixels were accumulated throughout walking image sequences and filtered using a low pass filter, e.g. A (Figure 2.21). Key heel strike frames were identified using the sinusoidal peaks of head trajectory data. A region of interest was identified about the foot and Gradient Descent used to identify heel strike candidates. Jung and Nixon (2013) used the Direct Linear Transformation (DLT), originally described by Abdel-Aziz and Karara (1971), to solve the two-dimensional homography and identify heel strike candidates intersecting the ground plane, i.e. $z = 0$. Using oblique-frontal perspective images captured in a biometric tunnel (research facility for non-contact gait recognition), Jung and Nixon (2013) reported heel strike identification rates of 95.6% with position accuracy of ± 10 cm in three-dimensions. Furthermore, Jung and Nixon (2013) reported identification rates of 93.4 and 93.7% for random direction walking using the PETS (2006) and CAVIAR (2004) databases respectively (train station surveillance images). Due to missing camera calibration parameters for PETS (2006) and CAVIAR (2004) database images, heel strike identification accuracy was reported as ± 5 pixels (two-dimensions) for both databases.

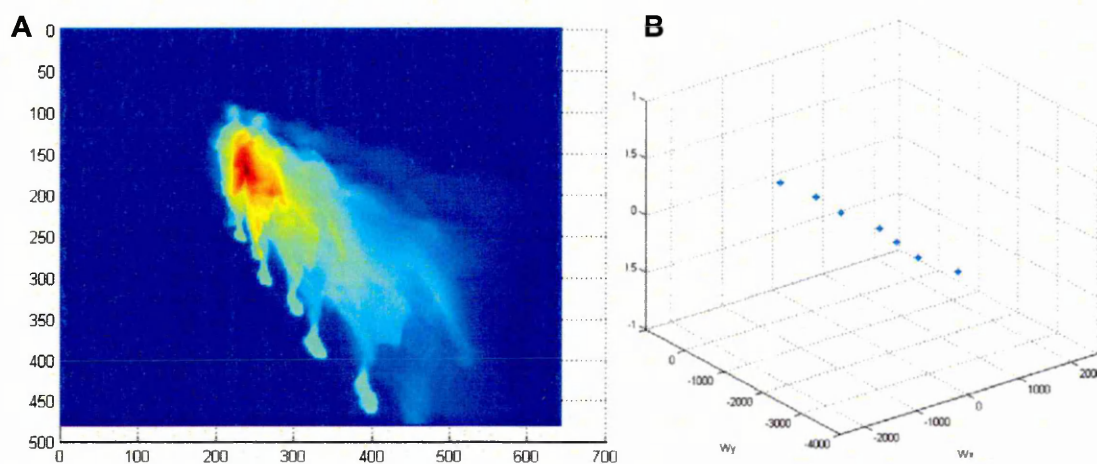


Figure 2.21. A: Accumulator map for extracted foreground pixels. B: Filtered heel strike candidates (adapted from Jung and Nixon, 2013).

Heel strike extraction demonstrates that gait features necessary for calculating step parameters can be automatically identified from different camera views. Low elevation

camera views used by Jung and Nixon (2013) indicate that heel strike extraction could be applied to footage obtained from sport stadia, e.g. Mauthner *et al.* (2008). Furthermore, heel strike data satisfy coplanarity assumptions of two-dimensional planar analysis and can be easily converted to real-world measurements using existing techniques, e.g. 2D-DLT (Walton, 1989 cited by Kwon, 1999). However current approaches, e.g. Bouchrika and Nixon (2006); Jung and Nixon (2013), have only been applied to walking. Furthermore, although Jung and Nixon (2013) retrieved temporal heel strike information, the approach was limited to walking gait. Finally, the sequential accumulation of image data for feature extraction (both approaches) is computationally exhaustive. The additional impact of other image processes, e.g. background modelling etc., might be problematic for feature extraction using high-resolution images, e.g. large filming area required to film tennis.

2.2.4. *Background modelling*

Different approaches to background modelling exist; presented heel strike extraction methods, e.g. Bouchrika and Nixon (2006); Jung and Nixon (2013), used adaptive, e.g. Stauffer and Grimson (1999), and combined pixel intensity and pixel colour differencing, e.g. Cheung, Kanade, Bouguet and Holler (2000), respectively. The lowest level at which background segmentation can occur is the pixel level (Gonzalez and Woods, 2002). Segmentation for the purpose of object or motion detection is typically performed with reference to a background model (Hassanpour, Sedighi and Manashty, 2011). To detect moving objects in video frames, the current image is subtracted from a background model; the difference determines objects that have moved. Background modelling can be classified by statistical and non-statistical approaches (Hassanpour, Sedighi and Manashty, 2011). Statistical approaches, e.g. Gaussian mixture model, estimate the probability function of pixels belonging to the background. Non-statistical approaches, e.g. time independent model, moving average model etc., assume the background model is an image; the subtracted difference infers object motion and deemed the image foreground.

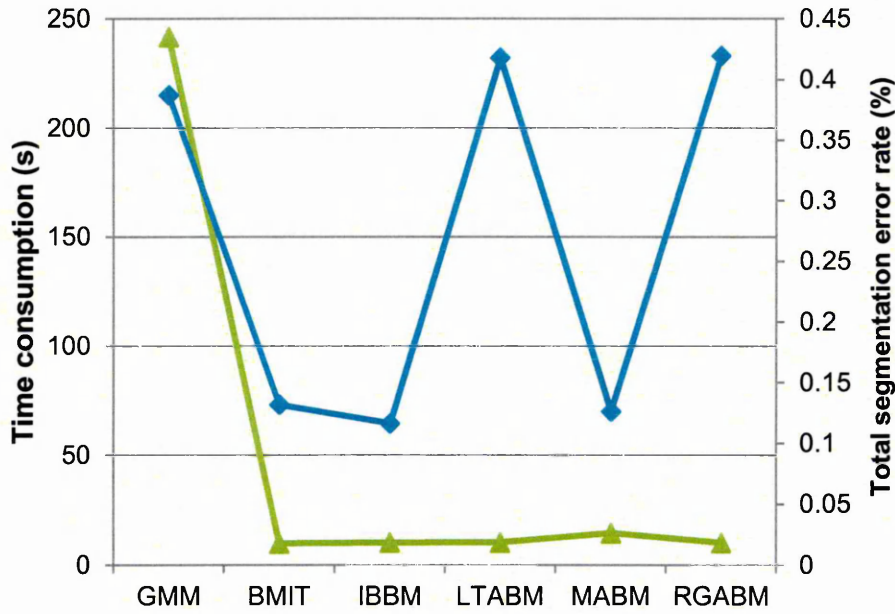


Figure 2.22. Evaluation of six background models using 100 sample images. Analysis time (green) and total segmentation error rate (blue) for indoor applications (adapted from Hassanpour, Sedighi and Manashty, 2011).

A trade-off exists between computational speed and segmentation accuracy. Hassanpour, Sedighi and Manashty (2011) evaluated the performance of statistical, e.g. Gaussian mixture model (GMM), and non-statistical background models, e.g. static or time independent model (BMIT), improved basic model (IBBM), long-term average model (LTABM), moving average model (MABM) and running Gaussian average (RGABM) for indoor and outdoor applications, e.g. Figure 2.22. The Gaussian mixture model demonstrated the largest computational demand as well as high segmentation error rates (Figure 2.22). For non-statistical background models, the time independent model, e.g. first or last image of a sequence (BMIT), demonstrated the lowest computational demand (Hassanpour, Sedighi and Manashty, 2011). Total segmentation error (sum of false negative and false positive errors) was lowest for the improved background model, e.g. iteratively updated model (IBBM). However, total segmentation error for the simplest background model, i.e. BMIT, was also among the lowest for evaluated models (Figure 2.22).

Differentiating foreground objects from shadow is an important segmentation task; shadows can distort the shape of detected foreground objects (Nghiem, Bremond and

Thonnat, 2008). Shadows can be cast on the foreground (self-shadows) and background (cast shadows); for the purpose of foreground segmentation, it is important that only cast shadows are identified (Benedek and Szirányi, 2007). Most approaches to shadow removal assume that shadows do not change object texture and chromaticity, e.g. Hue and Saturation components of HSV colour space (Nghiem, Bremond and Thonnat, 2008). Therefore shadow detection is predominantly performed by colour filtering rather than geometry based approaches (Benedek and Szirányi, 2007).

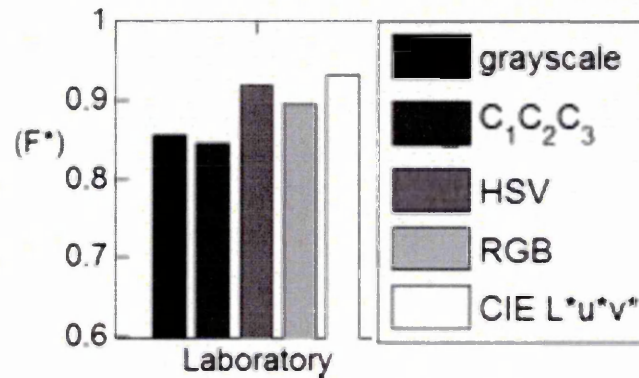


Figure 2.23. Evaluating of shadow removal by colour space. F^* is an efficiency measure quantifying misclassified foreground and background pixels (adapted from Benedek and Szirányi, 2007).

Off-the-shelf cameras typically represent colour in the RGB colour space (Nghiem, Bremond and Thonnat, 2008). However images can be represented in many other colour spaces, e.g. HSV, CIE and CMYK. Recently, Benedek and Szirányi (2007) assessed colour space selection for the removal of cast shadows. Benedek and Szirányi (2007) reported CIE L^*u^*v as the most effective colour space for removing shadow for various applications, e.g. Figure 2.23. However the conversion from RGB to CIE L^*a^*b or CIE L^*u^*v colour spaces has been reported as time consuming (Nghiem, Bremond and Thonnat, 2008). For indoor applications, HSV colour space has been reported as a fast and effective method for shadow removal (Benedek and Szirányi, 2007; Nghiem, Bremond and Thonnat, 2008).

2.2.5. Conclusion

Holistic gait recognition techniques can be applied to gait sequences without requiring a model of the person. However they are disadvantaged by high algorithm complexity, low robustness and difficult movement phase determination. Model-based gait

recognition techniques are view and scale invariant; this is advantageous for an application to tennis due to large and multi-directional player movements. Gait features can be automatically extracted from video sequences for walking and running. However view-independent gait feature extraction has only been performed for walking due to the complexity of joint angle data normalisation.

Model-based heel strike extraction demonstrated that gait features necessary for the calculation of step parameters can be automatically identified from different camera views. Due to data coplanarity, heel strike data can be easily converted to real-world measurements using existing techniques. Furthermore, heel strike extraction has been performed using low elevation camera views and could be applied to footage obtained from sport stadia. However current heel strike extraction methods are limited; current approaches have only been applied to walking and temporal heel strike information is limited to the analysis of walking. Furthermore, the sequential accumulation of image data is exhaustive; the efficiency of other image processes, e.g. background modelling and shadow removal, was considered. Static or time independent background modelling was demonstrated to yield low segmentation error rates and low computational demand. Furthermore, the HSV colour space was identified as a fast and effective method for shadow removal.

2.4. Chapter findings

Newly developed, motion analysis tools should provide measurements that are of practical use to the end-user, i.e. ITF, coaches and practitioners. A review of literature has highlighted that a suitable method for tracking players and identifying foot-surface contacts during match-play tennis does not exist. However separately, methods for tracking players during competition and for extracting heel strikes during walking have been described.

2.4.1. Player tracking

Many non-intrusive motion analysis techniques exist to track players during competition. However it is impracticable to install multi-camera systems at match-play tennis events. Non-intrusive, single camera methods are capable of extracting player location to an acceptable accuracy level (Mauthner *et al.*, 2008). Such methods are highly portable and do not interfere with competition, thus adhering to the rules of

tennis (ITF Rules of Tennis, 2013). Mauthner *et al.* (2008) extracted the lower-limbs to identify player position, indicating the extraction of gait features during competition might be possible. It was not the aim of Mauthner *et al.* (2008) to extract gait features, rather to identify and track player position. However to this end, player position was limited by its dependence on the perspective projection of the lower-limbs.

2.4.2. *Heel strike extraction*

Gait recognition techniques, either holistic or model-based, did not satisfy the aims of this project. However, model-based heel strike extraction demonstrated that gait features necessary for the calculation of step parameters can be automatically identified from different camera views. Heel strike extraction was performed using low elevation camera views that could be obtained when filming at sport stadia. Furthermore, heel strike data can be easily converted to real-world measurements using existing techniques due to data coplanarity. However current heel strike extraction methods are limited; spatio-temporal parameters of heel strikes during walking and running have not been extracted. Furthermore, spatio-temporal parameters of gait, e.g. step length, have not been quantified using single camera footage.

2.4.3. *Conclusion*

Knowledge of tennis player interactions with tennis court is limited (Miller, 2006). The literature review highlighted that the quantification of tennis player step and movement characteristics during match-play was necessary. This would improve current knowledge and inform future field- or laboratory-based research. Separately, methods for tracking players during competition as well as methods for extracting heel strike information have been described. Currently, their application to match-play tennis, particularly heel strike extraction, would be limited. However, combining a player tracking and foot-surface contact identification algorithm based on approaches presented might allow the quantification of tennis player step and movement characteristics.

3 Player step and movement characterisation at the 2011 Roland Garros Qualifying Tournament

3.1. Introduction

As a precursor to developing a single camera system to measure gait parameters during match-play tennis, it was necessary to determine the characteristics of play at a suitably representative tournament: the Roland Garros tournament was chosen for this purpose. The 2011 Roland Garros Qualifying Tournament was held from 16 – 20th May 2011 in Paris, France. Played on clay court surfaces, the qualifying tournament awards winning players a place in the main competition, one of the prestigious ‘Grand Slam’ tennis tournaments. Notational analyses have reported men’s singles rallies to consist of higher intensity movements than women’s singles rallies (O’Donoghue and Ingram, 2001). Indeed, physiological profiles of match-play tennis support notational analyses (Ferrauti, Weber and Wright, 2003; Fernandez, Mendez-Villanueva and Pluim, 2006). However, information regarding player-surface interaction, i.e. step strategy, is currently limited (Miller, 2006). This is because no existing research has quantified step strategy in match-play tennis. The Roland Garros Qualifying Tournament was an opportunity to characterise of player step and movement strategy and assess filming constraints associated match-play tennis.

3.2. Aim and Objectives

The aim of this chapter is to use the Roland Garros 2011 Qualifying Tournament to measure tennis player step and movement strategy and develop data collection methods for match-play tennis. This will inform the development of subsequent motion analysis tools and relates to boxes B and C of the development stage diagram (Figure 1.2).

Objectives:

1. Collect video footage of match-play tennis.
2. Measure player step and movement strategy.
3. Assess data collection techniques.

3.3. Measuring player step and movement strategy at the 2011 Roland Garros Qualifying Tournament

3.3.1. Single camera calibration

The literature review identified that a vision-based system should be developed to minimise interference with match-play tennis. Photogrammetric systems require calibration, therefore filming and calibration at sport stadia must to be considered. Stereo camera calibration requires camera images to be synchronised. Furthermore, images should be of an appropriate resolution (Choppin, 2008) and have appropriate spatial separation to yield sufficient 3D reconstruction accuracy, i.e. camera axis intersection angle of 40 – 140° (Chen and Davies, 2000).



Figure 3.1. Overview of Roland Garros; singles tennis matches performed on court seven (yellow ring) were filmed.

Choppin (2008) compared the accuracy of Zhang's (1999) planar checkerboard calibration method for different resolution images by comparing extracted and reprojected checkerboard coordinates. Choppin (2008) commented that, when converted into millimetres, calibration accuracy decreased as checkerboard resolution decreased. Previous field-based stereo photogrammetry has been performed in small motion capture volumes. In tennis, Choppin (2008) calibrated a $2 \times 2 \times 2$ m volume and in football, Driscoll (2012) calibrated $1.5 \times 1.5 \times 1.0$ m volume; checkerboards filled calibration images thus maximising resolution. For current work, the literature review highlighted that it is necessary to film very large motion capture volumes, i.e. players

travel 8 – 12 m during average tennis rallies (Fernandez, Mendez-Villanueva and Pluim, 2006).

For competitive sport environments, large motion capture volumes, restricted camera positioning and physical checkerboard size limit the use of stereo photogrammetry. Furthermore, the literature review highlighted that to record synchronised camera images, 20 - 80 m of synchronisation and data transfer cable would be required to link stereo cameras for typical sport stadia, i.e. camera intersection angle of 40 – 140° (Chen and Davies, 2000). For calibration of a stereoscopic system, a calibration object, i.e. checkerboard, is required to be positioned in the motion capture volume. However, for single camera calibration, there is no requirement to position a calibration object within the motion capture volume as no corresponding camera view is required, i.e. stereo photogrammetry.

Direct Linear Transformation (DLT) (Abdel-Aziz and Karara, 1971) is an extensively used, three-dimensional photogrammetric technique. A planar modification of DLT, termed 2D-DLT, calculates eight DLT coefficients necessary to reconstruct the 2D position of a point on a plane (Walton, 1981 cited by Kwon, 1999):

$$u = \frac{L_1x + L_2y + L_3}{L_7x + L_8y + 1} \quad [3.1]$$

$$v = \frac{L_4x + L_5y + L_6}{L_7x + L_8y + 1} \quad [3.2]$$

where (u, v) are image coordinates and $L_1 - L_8$ are DLT parameters for a horizontal plane (x, y) .

Therefore, assuming coplanarity, two-dimensional position for image coordinates in relation to a calibration plane, can be reconstructed. 2D-DLT is a popular method for two-dimensional photogrammetry because the camera's optical axis is not required to be perpendicular to the plane of motion (Kwon, 1999); non-perpendicular camera views are typical for filming in sport stadia. Single camera calibration thus represents a flexible approach to sport stadia filming as camera images can be calibrated without the necessity of gaining access to the tennis court to perform a calibration.

3.3.2. *Footage collection in sport stadia*

Permission to film and relevant accreditations were obtained from the Fédération Française de Tennis. Furthermore, approval for all procedures was obtained from the Research Ethics Committee of the Faculty of Health and Wellbeing, Sheffield Hallam University (Appendix 1). For single camera filming, initial requirements were outlined in order to guide filming:

- Full-court field-of-view, including baseline and sideline areas
- Player position for rally movements
- Foot-surface contact position for forehand groundstroke movements

A high-definition video camera (Everio GZ-HD40EK, JVC, Japan), operating at 25 Hz (50 fields / second) with a resolution of 1920×1080 pixels (single CMOS sensor), was mounted on a tripod in an elevated location (Figure 3.2) to obtain a full-court field-of-view. Camera focal length was set manually and subsequently locked. Camera shutter speed was set to $1/250$ s however camera aperture was set automatically due to varying ambient light, i.e. outdoor filming. The video camera was equipped with an on-board disk drive and mains power input (mains power supply located under stadia seating, bottom right of Figure 3.2). This enabled the continual filming of tennis matches and minimised interference with match-play i.e. changing video cassettes etc. As a precautionary measure, all mains power cables and extension drums were waterproofed and tidied away from public access.



Figure 3.2. Elevated camera setup at the 2011 Roland Garros Qualifying Tournament.

3.3.3. *Elevated calibration plane*

As discussed within the literature review, definitions used to reconstruct player position vary. Previous approaches have defined player position by reconstructing the centre of a player (Jiang *et al.*, 2009) or the mean horizontal and maximum vertical coordinate of a bounding box (Mauthner *et al.*, 2008) projected onto a ground level calibration plane. Such approaches do not accurately reflect player position as definitions either neglect the effects of camera perspective or do not identify the centre-of-mass (COM). Further, the manual identification of COM projection onto a ground calibration plane can exacerbate random error due to subjective digitising (Glazier and Irwin, 2001). It was considered that an improvement to current approaches would be to introduce an elevated calibration plane that was of corresponding dimension to the ground plane, but elevated to 914 mm, e.g. net height. The elevated plane reduced out-of-plane distance to player COM and was thus assumed to reduce out-of-plane reconstruction error resulting from camera perspective.

A rigid object, set to 914 mm (net height) and held vertically (using a spirit level), was placed at four court locations corresponding to the ground calibration plane, e.g. singles court line intersections. Ground level and elevated calibration plane locations were then manually digitised (red and yellow filled circles respectively: Figure 3.3), providing the

necessary information to calibrate images using 2D-DLT. As noted, a calibration plane elevated to net height is limited by its assumption that net height corresponds to player COM, i.e. coplanarity requirement for 2D planar analysis. However, without an additional camera to solve collinearity, the introduction of an elevated calibration plane was considered a good compromise.

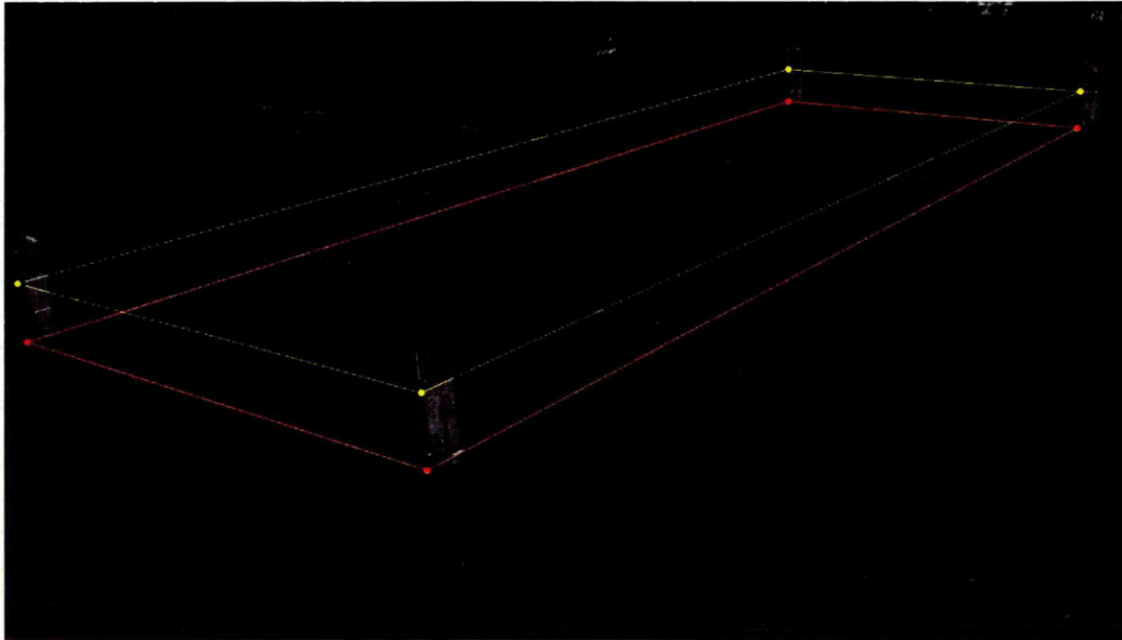


Figure 3.3. Digitised ground and elevated calibration plane locations (red and yellow filled circles respectively) highlighted in a combined image.

3.3.4. Player step and movement definitions

Fernandez, Mendez-Villanueva and Pluim (2006) reported the duration of tennis rallies to be 5 - 10 s. This corresponds to 250 - 500 frames of video footage (50 Hz) for each tennis rally. The use of simplified movement parameters was therefore necessary to enable footage of tennis rallies to be manually digitised. Based on movement definitions by Robinson and O'Donoghue (2008), player displacement was defined as the distance between individual rally movement endpoints, i.e. 'acceleration from stationary' to a 'sharp path change'. For subsequent rally movements, the player was considered to be stationary, thus a rally was approximated as a series of start-stop movements. Player COM was subjectively determined and provided a simple method for quantifying player displacements. Player movements were quantified with respect to the baseline and centreline and proportions of small, medium and large rally movements were quantified to reflect previous characterisations (Fernandez *et al.*, 2006; Hughes and Meyers, 2005).

Foot-surface contacts were defined as the perceived location of peak force application at mid-stance and were digitised during forehand groundstroke movements. Step length and step frequency were calculated to quantify step parameters indicated as important to movement strategy in tennis (Hughes and Meyers, 2005). Finally, ball-racket contacts were also digitised to determine contextual information, i.e. number of strokes etc.

3.3.5. Analysis software

Following data collection, a bespoke analysis system (programmed in Microsoft Visual Studio using the .NET framework) was developed to quantify coronal plane player step and movement strategy as well as contextual information (Figure 3.4). This system allowed player and foot-surface contact data to be reconstructed with ground and elevated calibration planes respectively. Footage was analysed of five right-handed male and five right-handed female players (refer to Appendix 2 for player age, mass and stature detail).

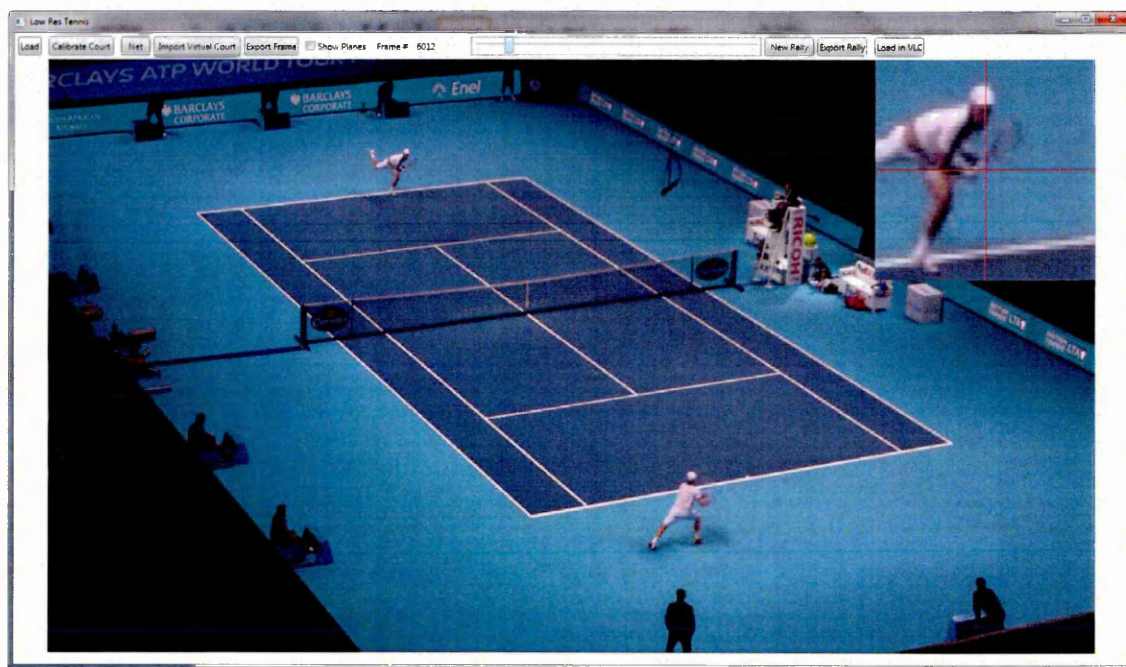


Figure 3.4. Screenshot of analysis system used to manually digitise and quantify player step and movement strategy.

The software was used to load and deinterlace (bob and expand) high-definition video footage to 50 Hz. Deinterlaced video fields could be advanced or regressed by ± 1 , 10, 25 or 250 frames via hotkeys, allowing footage navigation. The program recorded manual user inputs as described in section 3.3.4. Each input event was denoted by

specific hotkeys, setup to assign individual mouse-clicks to relevant events. The analysis software was written specifically to analyse tennis rally events and reconstructed position using 2D-DLT, a typical method for obtaining player position data in sports analyses (Barros *et al.*, 2007). Data were sorted into feet and player COM coordinate data and passed to 2D-DLT with ground and elevated calibration plane coordinates respectively (image and real world). Finally, XY coordinate (Figure 3.5) and time data were exported for analysis in MATLAB (R2013b, The MathWorks, MA, USA). A custom analysis script (Appendix 2) identified and exported player step and movement characteristics as well as descriptive rally information, i.e. number of strokes, duration, etc. The reference frame for player displacement data was changed to reflect baseline and centreline movements, i.e. absolute player displacement in X and Y directions respectively (Figure 3.5), to simplify player movement characteristics.



Figure 3.5. Camera image with calibration points (C1 - C4) and global coordinate system plotted.

In total, 20 men's rallies and 20 women's rallies were digitised according to definitions provided in section 3.3.4. Output data were grouped by gender and a one-way ANOVA (SPSS 16.0, Inc., Chicago, IL, USA) was performed to identify differences (corrected using Welch's F ; $\alpha = 0.05$) between men's and women's rallies. Between factor effect sizes (ES_B : equation 3.3) were also calculated to assess effect size magnitudes. Effect

size magnitudes were interpreted as described by Cohen (1988), e.g. effect sizes of 0.2, 0.5 and more than 0.8 represent small, medium and large differences respectively.

$$ES_B = (\bar{x}_1 - \bar{x}_2) / s_c \quad [3.3]$$

where \bar{x} is the group mean and s_c is the cohort standard deviation.

3.4. Results

Data was obtained from footage of 10 players (five male and five female), over 6 matches. In total, 40 rallies were manually digitised to characterise player step and movement strategy in match-play tennis. Table 3.1 presents descriptive characteristics of men's and women's rally movements performed at the 2011 Roland Garros Qualifying Tournament. Table 3.1 demonstrates similar movement characteristics for men's and women's rallies on a clay surface. A small effect size, e.g. $|ES_B| > 0.2$, highlights a trend of a greater number of steps taken in men's forehand movements; however present data cannot confirm a generalisable difference in relation to this.

Table 3.1. Men's and women's tennis rally characteristics on clay (N = 40).

Variable	Men	Women	ES_B	P Value
	$\bar{x} \pm s$ [Range]	$\bar{x} \pm s$ [Range]		
Forehand movements (n)	1.55 \pm 0.83 [1.00 – 4.00]	1.45 \pm 0.69 [1.00 – 3.00]	0.12	0.679
Forehand movement steps (n)	5.35 \pm 2.54 [3.00 – 10.00]	4.75 \pm 2.61 [2.00 – 10.00]	0.24†	0.466
Rally strokes (n)	8.35 \pm 2.39 [7.00 – 17.00]	8.50 \pm 2.52 [7.00 – 15.00]	-0.06	0.848
Rally duration (s)	8.73 \pm 4.32 [3.68 – 20.70]	9.21 \pm 4.17 [3.96 – 19.02]	-0.11	0.722

† Small effect size ($|ES_B| > 0.2$).

Table 3.2 presents step and movement characteristics for the men's and women's forehand movements. Table 3.2 demonstrates different step strategies for men's and

women's rally movements. Women's forehand movements consisted of higher step frequencies ($P = 0.016$, $ES_B = -6.24$; Table 3.2). However, step lengths did not reveal any differences ($P = 0.267$, $ES_B = 0.19$; Table 3.2). A small trend for greater baseline player displacement, i.e. movement between sidelines, was observed for men's forehand movements ($ES_B = 0.30$; Table 3.2). However, centreline player displacement did not illustrate any differences between men's and women's movements between the baseline and net ($P = 0.918$, $ES_B = 0.03$; Table 3.2). Men's rallies consisted of a lower proportion of small rally movements ($ES_B = -0.43$; Table 3.2) and a corresponding higher proportion of large rally movements ($ES_B = 0.49$; Table 3.2). However, present data could not confirm generalisable differences in relation to these trends ($P = 0.162$ and 0.075 for small and large rally movements respectively; Table 3.2).

Table 3.2. Men's and women's step and movement characteristics for tennis rallies on clay (N = 40).

Variable	Men	Women	ES_B	P Value
	$\bar{x} \pm s$ [Range]	$\bar{x} \pm s$ [Range]		
Step length (m)	1.02 ± 0.36 [0.55 – 1.64]	0.95 ± 0.34 [0.51 – 1.60]	0.19	0.267
Step frequency (Hz)	5.32 ± 0.25 [5.13 – 5.51]	6.88 ± 0.33 [6.63 – 7.12]	-6.24**	0.016*
Baseline player displacement (m)	2.30 ± 1.34 [1.08 – 4.18]	1.90 ± 1.13 [0.65 – 3.35]	0.30†	0.089
Centreline player displacement (m)	1.13 ± 0.82 [0.39 – 2.28]	1.11 ± 0.90 [0.36 – 2.51]	0.03	0.918
Small ($d < 2.5$ m) rally movements (%)	50.61 ± 30.57 [0.00 – 100.00]	63.69 ± 27.37 [20.00 – 100.00]	-0.43†	0.162
Medium ($2.5 \text{ m} \leq d \leq 4.5$ m) rally movements (%)	36.08 ± 30.96 [0.00 – 100.00]	30.39 ± 23.77 [0.00 – 70.00]	0.18	0.518
Large ($d > 4.5$ m) rally movements (%)	13.31 ± 15.16 [0.00 – 50.00]	5.92 ± 9.57 [0.00 – 33.33]	0.49†	0.075

* Significant difference ($P < 0.05$). ** Large effect size ($|ES_B| > 0.8$). † Small effect size ($|ES_B| > 0.2$). Small ($d < 2.5$ m), medium ($2.5 \text{ m} \leq d \leq 4.5$ m) and large ($d > 4.5$ m) rally movements (Fernandez *et al.*, 2006) represent proportions of a rally.

Table 3.3 presents first, second and third quartiles for player displacement (resultant direction) during men's and women's forehand movements. Resultant displacements reflect higher proportions of large rally movements for men in Table 3.2.

Table 3.3. Player displacement (resultant direction) for men's and women's tennis rallies on clay (N = 40).

Resultant direction player displacement (m)	Quartile (%)		
	25	50	75
Men	2.19	2.55	3.27
Women	1.92	2.19	2.91

3.5. Discussion

3.5.1. Rally characterisation

Descriptive parameters presented in Table 3.1 demonstrate that the men's and women's rallies analysed were similar in terms of duration, number of strokes and number of forehand movements. A small trend ($ES_B = 0.24$, Table 3.1) indicated that men's forehand movements might consist of a greater number of steps. This was supported by larger player displacements during these movements (Table 3.3), despite similar step lengths ($P = 0.267$, $ES_B = 0.19$; Table 3.2). This indicates different gait strategies for men's and women's forehand movements on clay, highlighted by greater step frequencies in women's forehand steps ($P = 0.016$, $ES_B = -6.24$; Table 3.2). Such findings support the notion that forehand movements in men's rallies are of a higher intensity than women's rallies (O'Donoghue and Ingram, 2001). However, current data likely suffer from a low sample size, i.e. application of Welch's F correction, and it would therefore be inappropriate to generalise specific findings.

Player travel proportions, i.e. $d < 2.5$ m, $2.5 \text{ m} \leq d \leq 4.5$ m and $d > 4.5$ m for small, medium and large distances (d) respectively, reveal that men's rallies consist of a lower proportion of small movements ($ES_B = -0.43$; Table 3.2) and higher proportion of large movements ($ES_B = 0.49$; Table 3.2). The higher proportion of large movements again reflects greater movement variation observed in men's singles point strategy (O'Donoghue and Ingram, 2001). Furthermore, trends of larger movements along the

baseline ($ES_B = 0.30$; Table 3) highlight the direction of movements that players perform and would be of value to practitioners. However, as previously noted, current data do not support generalisable findings. This is likely due to a low sample size.

Findings demonstrate consistencies with previous notational analyses and go some way to characterising step and movement strategies used in men's and women's rallies during competition. However, large standard deviations highlight multiple phases within forehand movements, i.e. acceleration, constant speed and deceleration phases. As such, movement phases should be identified to segregate step and movement data; clearly defined parameters would improve the ability of subsequent analyses to yield generalisable findings. Determining movement phases clearly requires frame-by-frame analysis, i.e. relation to player movement velocity. As such, an automated player tracking system, capable of determining foot-surface contacts is required to develop work presented in this chapter. An automated system would also enable larger quantities of tennis footage to be analysed, again improving generalisability and usefulness of findings to practitioners.

3.5.2. Filming and analysis

Single camera filming was preferred due to the large field-of-view required and difficulty of synchronising multiple cameras in competitive sport environments. As previously noted, the use of an elevated calibration plane is limited by its assumption that net elevation corresponds to player COM, thus satisfying coplanarity requirements for 2D planar analysis. The use of an elevated calibration plane can only be justified if the reconstruction of player COM position is considered to be beneficial, i.e. improved estimates for player position. With current data, it is not possible to assess whether reconstruction accuracy using an elevated calibration plane represents an improvement on the use of a ground calibration plane. It is therefore clear that this must be addressed to warrant the future use of elevated calibration planes.

Filming conditions at competitive sport environments, in comparison to laboratory settings, are typically restrictive. Due to the large required field-of-view, matches were filmed with a wide filming angle inducing large image distortions (up to 40.7 pixels). Two-dimensional DLT is a popular method for reconstructing the position of a point on a plane. However, Dainis and Juberts (1985) reported that DLT reconstruction error at

the outer 10% of an image was 100% greater than at the image centre. Although modified versions of DLT account for symmetrical lens distortion, e.g. Hatze (1988), current implementations of 2D-DLT do not account for image distortions. The accuracy of reconstructed position data, based on the assumption of image linearity, is therefore limited given the magnitude of image distortion required to film at a competitive sport event, e.g. Figure 3.4. Therefore, reconstruction accuracy for 2D-DLT in competitive sport environments should be assessed and alternative methods of camera calibration devised if necessary.

3.6. Conclusion

The chapter above describes the development of a programme to characterise player step and movement strategies during rallies performed in match-play tennis. Forty rallies, the equivalent of 18,000 frames, were manually digitised. The use of simplified movement parameters was therefore necessary to enable the characterisation of tennis rallies. Findings highlight some differences between men's and women's step and movement strategy in match-play tennis, however generalised conclusions based on current data are not suitable. The use of simplified step and movement definitions was a centrally limiting factor; a frame-by-frame analysis of step and movement strategy is required to define movement phases accordingly. This would require an automated approach to rally analysis, due to the large volume of footage that would be required to elicit meaningful findings.

Filming with a high-definition camera was necessary to provide a suitable resolution of the large capture volume. The camera was set up off-court, in a manner that would not interfere with match-play, i.e. adherence to rules of tennis (ITF Rules of Tennis, 2013). Further, the hard-drive based camera minimised user intervention and all analyses were performed *post-hoc*. However, due to large image distortions induced by wide angle filming, i.e. large field-of-view, the validity of 2D-DLT as a method for position reconstruction must be assessed. Position reconstruction accuracy should be assessed in similar filming conditions prior to future use. The elevated calibration plane was introduced to minimise erroneous player position estimates resulting from manually estimating the ground plane projection of player position. However, the elevated calibration plane elevation was assumed to intersect player COM. Therefore player position reconstruction accuracy must be assessed prior to future use.

This chapter presents a simple system for characterising player step and movement strategy in match-play tennis. The system represents a highly portable system, i.e. single camera, that could be used for a variety of applications, i.e. identifying step and movement strategy as a function of surface type, gender, weather, season (indoor / outdoor), rule change, etc. In relation to the overall project aim, this chapter has highlighted that a single camera system can identify differences between men's and women's step and movement strategy in match-play tennis. However to address current limitations, the measurement of player step and movement strategy must be automatic and derived position data must be validated.

4 Single camera position reconstruction

4.1. Introduction

Chapter 3 highlighted that, for the context of this study, single camera filming was a suitable approach to obtain player and foot-surface contact position data. A method to determine real world position is required. There are many photogrammetric techniques that derive position data in reference to a global coordinate system. The simplest method is linear scaling. However, the technique requires that the camera is horizontally levelled and perpendicular to the plane of motion, i.e. an elevation angle of 90° (Brewin and Kerwin, 2003). Incremental changes to camera elevation, ranging from $88 - 96^\circ$, were shown to result in large reconstruction errors that increased linearly with deviation from the perpendicular position (Brewin and Kerwin, 2003). As such, scaling would be inappropriate for non-perpendicular camera views. Alcock, Hunter and Brown (2009) presented a line fitting method, based on the equation of a line, i.e. $y = mx + c$, for reconstructing two-dimensional position of points on a plane for non-perpendicular camera views. However, for a large field-of-view camera, 2D-DLT yielded more accurate position reconstruction.

DLT (Abdel-Aziz and Karara, 1971) and its planar modification 2D-DLT (Walton, 1981 cited by Kwon, 1999) are used extensively in sport biomechanics to reconstruct position data from images. However, image distortions due to the lens can affect position reconstruction accuracy for DLT methods (Dainis and Juberts, 1985). Wide filming angles, which are often required to film competitive sport events, induce image distortions due to the lens. Current implementations of 2D-DLT do not account for image distortions due to the lens. Therefore filming conditions experienced at competitive sport events can impair the accuracy of position reconstruction using 2D-DLT.

It is important to model the camera accurately and to account for image distortions due to the lens, prior to position reconstruction. Numerous camera calibration models exist; Tsai (1987) and Zhang (1999) present two of the most popular and accurate methods. Due to flexibility and the suitability of the planar calibration for dynamic filming environments (Sun and Cooperstock, 2005), Zhang's (1999) planar checkerboard technique for camera calibration was adopted. Bouguet (2010) presents a useful

MATLAB based camera calibration toolbox, based on Zhang's (1999) calibration method; subsequent work was developed using this toolbox.

4.2. Aim and objectives

The aim of this chapter was to develop and assess a photogrammetric method to reconstruct real world, planar position data derived from images obtained at competitive tennis events. This relates to box D of the development stage diagram (Figure 1.2).

Objectives:

1. Identify a camera-plane model based on Zhang's (1999) planar calibration technique.
2. Develop a method for reconstructing real world position of image coordinates that are coplanar with a physical calibration plane.
3. Assess the accuracy of the position reconstruction method (objective 2) in relation to existing reconstruction methods.

4.3. Monocular photogrammetry

4.3.1. Camera model

Photogrammetry requires a camera model to calibrate derived metrics. Colour images, which are typically comprised of red, green and blue channels, are two-dimensional, i.e. image plane (Figure 4.1). The most basic camera model is that of the ideal pinhole camera (Hartley and Zisserman, 2003). For a pinhole camera, camera aperture is assumed to be a point; no lenses are required to focus rays onto the image sensor. As such, for a perpendicular, pinhole camera model, only a scaling factor is required to transform between image and real world coordinates:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad [4.1]$$

where XYZ and uv are the real world (3D) and image plane coordinates (2D) of a point respectively, s is a scale factor and 1 represents the projection of the image plane coordinate to infinity.

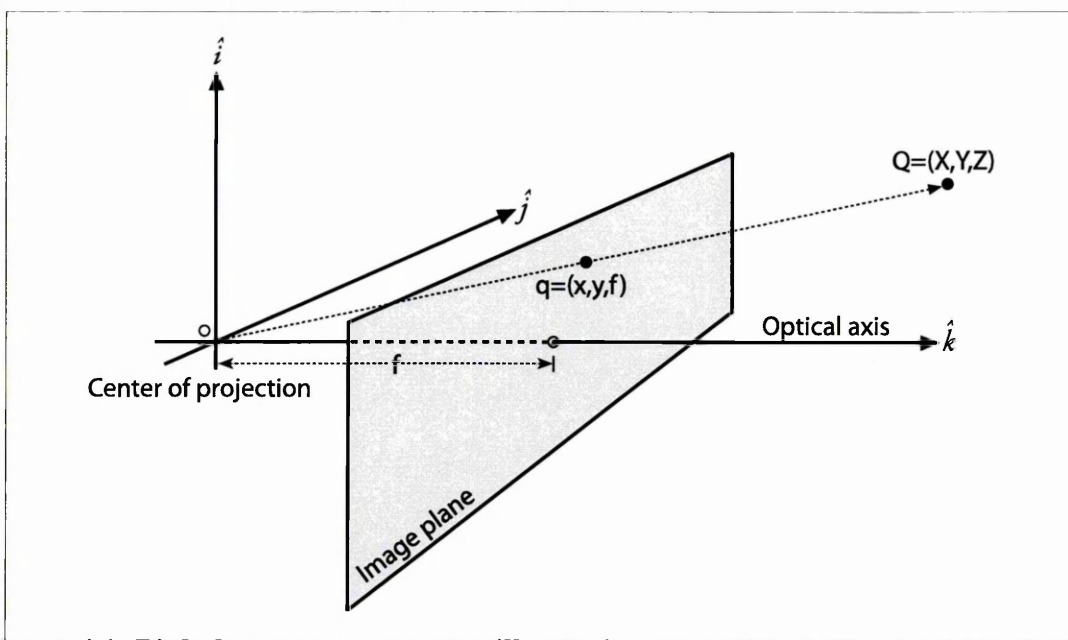


Figure 4.1. Pinhole camera geometry illustrating the real world point Q projected as q in the image plane (adapted from Bradski and Kaehler, 2008).

However, most cameras include lenses. The purpose of a lens is to enable a range of camera viewing angles, i.e. telephoto to wide (far to near scenes respectively), and focal length adjustment when working distance is changed, i.e. improve image sharpness. As such, the projection of a point in the camera coordinate system to the image plane is given as:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & c & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad [4.2]$$

where α and β are focal lengths in u and v image axes respectively (expressed in pixel-related units), c describes pixel skew (assumed to be zero), u_0 and v_0 are horizontal and vertical coordinates of the principal point respectively.

Mass produced spherical lens systems (typical of off-the-shelf cameras) introduce radial and tangential lens distortions to images (Bradski and Kaehler, 2008). For radial distortions, the distortion at the camera's optical centre is zero and increases toward the lens periphery, i.e. spherical nature of lens. Tangential lens distortions can arise due to manufacturing defects, i.e. non-parallel alignment of the lens and sensor.

Image distortions, due to the effects of radial and tangential lens distortions, can be calculated. Bouguet's (2010) 'Camera Calibration Toolbox for MATLAB, is a modified implementation of Zhang's (1999) camera calibration technique. The toolbox processes multiple views of a planar calibration object, i.e. checkerboard pattern (high contrast and known geometry). Locations of square corners are extracted on a semi-automated basis to calculate the planar homographies between the camera and checkerboards (Bouguet, 2010). Bouguet (2010) adopts Heikkilä and Silvén's (1997) intrinsic camera model due to the inclusion of both radial and tangential lens distortion coefficient terms. Collectively, the camera model terms α , β , c , u_0 , v_0 (identified in equation 4.2) and lens distortion term kc , a 5×1 vector, are called the intrinsic camera parameters.

Figure 4.2 illustrates an image captured by a camera with a regular lens using a wide filming angle, i.e. zoomed-out. Image distortions (radial and tangential) were calculated and applied to the z (imaginary) axis of the original camera image to illustrate image distortions induced by the lens system. Photogrammetric techniques that rely on direct image-to-world mapping, i.e. 2D-DLT (Kwon, 2012), are limited when applied to images containing lens distortion. Calibration coordinates, i.e. court-line intersections, manually identified in a distorted camera image (A: Figure 4.2), would be a composite of image plane coordinate and image distortion magnitude, i.e. B (Figure 4.2). Therefore direct image-to-world mapping would be inappropriate.

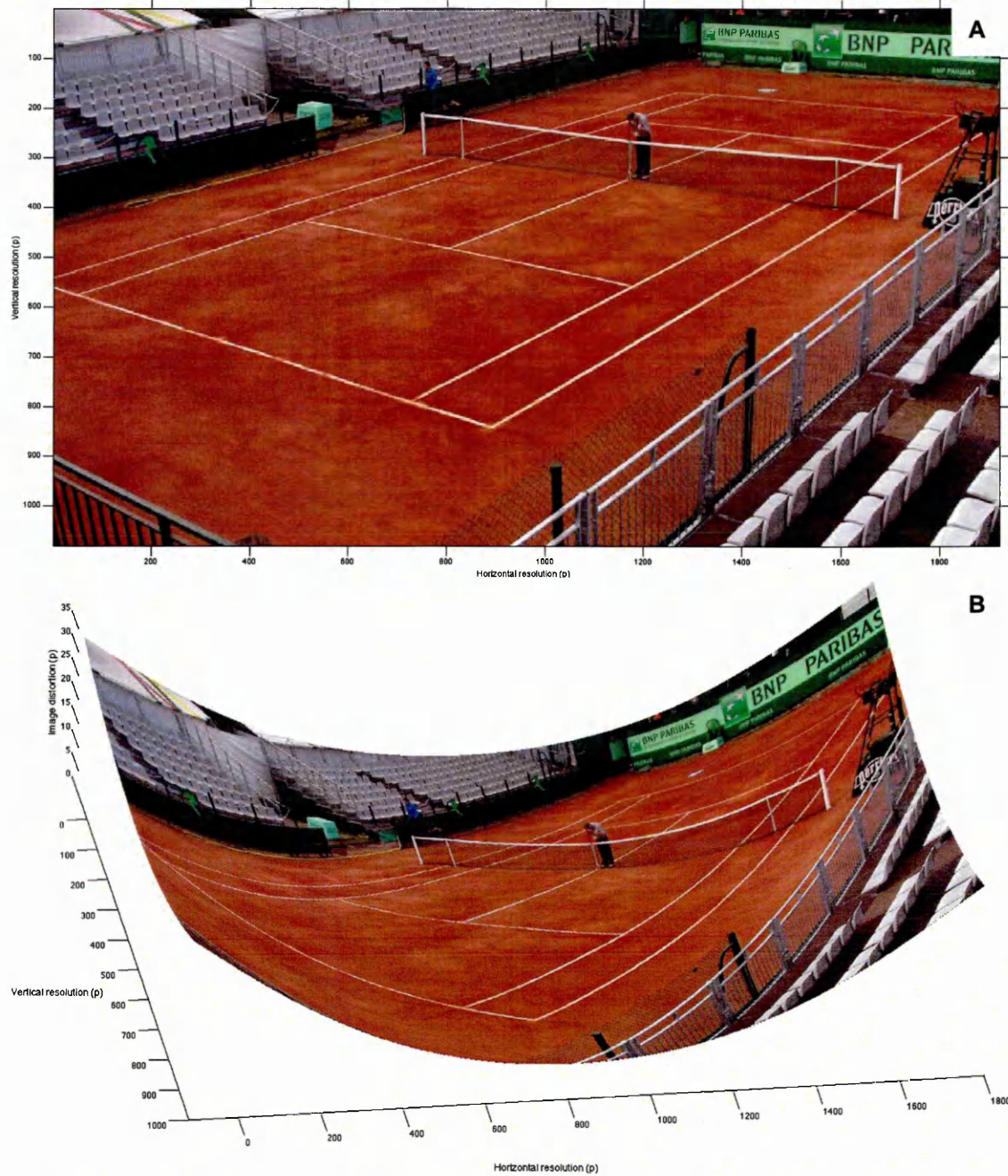


Figure 4.2. Sample image (original format; A) and with distortions (radial and tangential) applied to the z (imaginary) axis (B), illustrating the spherical effect of the lens system.

Following the calculation of intrinsic camera parameters, any uv coordinate in the image plane can be normalised to the camera coordinate system with the following steps:

Subtract principal point and divide by focal length:

$$[x_n, y_n]^T = [\frac{u-u_0}{\alpha}, \frac{v-v_0}{\beta}]^T \quad [4.3]$$

Remove skew:

$$[x_n, y_n]^T = [x_n - c \times y_n, y_n]^T \quad [4.4]$$

Correct distortion:

$$r = x_n^2 + y_n^2 \quad [4.5]$$

$$k_{radial} = 1 + (kc_1 \times r) + (kc_2 \times r^2) + (kc_5 \times r^3) \quad [4.6]$$

$$\Delta x = (2kc_3 \times x_n \times y_n) + kc_4(r + 2x_n^2) \quad [4.7]$$

$$\Delta y = kc_3(r + 2y_n^2) + 2kc_4 \times x_n \times y_n \quad [4.8]$$

$$[x_n, y_n]^T = [\frac{x_n - \Delta x}{k_{radial}}, \frac{y_n - \Delta y}{k_{radial}}]^T \quad [4.9]$$

where x_n and y_n are normalised horizontal and vertical image coordinates respectively, $kc_{1:2,5}$ are radial distortion coefficients and $kc_{3:4}$ are tangential distortion coefficients. The converged normalisation of the coordinate $[x_n, y_n]^T$ is obtained by applying equations 4.5 – 4.9 in a recursive loop of 20 iterations (Bouguet, 2010).

4.3.2. Camera-plane model

At competitive tennis events, large camera distances and non-perpendicular field-of-views result in perspective projected images, i.e. Figure 4.2. Figure 4.3 illustrates the relationship between points q (image plane) and Q (real world) and corresponding

camera and world coordinate systems. The rotation matrix R and translation vector t (collectively known as camera extrinsic parameters) identify the homography between camera and world coordinate systems.

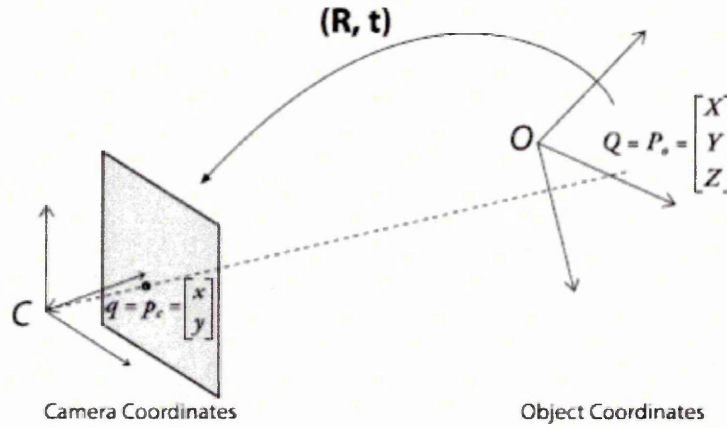


Figure 4.3. Point p_c (camera coordinate) is related to point P_o (global coordinate) by applying the rotation matrix R and translation vector t (adapted from Bradski and Kaehler, 2008).

Assuming intrinsic camera parameters are known, an orthogonal plane, defining image coordinates relating to the world coordinate system, determines camera extrinsic parameters by the following relationship:

$$p_c = R \times P_o + t \quad [4.10]$$

where p_c is a normalised image coordinate and P_o is the corresponding world coordinate, R is a 3×3 matrix of direction cosines and t is a 3×1 translation vector, i.e. $[t_x, t_y, t_z]^T$.

The derived 3×3 matrix of direction cosines correspond to the X , Y and Z rotations about each axis required to align coordinate systems. The 3×1 translation vector defines X , Y and Z translation between coordinate systems. The definition of intrinsic and extrinsic camera parameters allows coordinates to be transformed from image to world coordinate systems and vice versa, i.e. pixel reprojection. This defines the camera-plane model (Figure 4.4) that is required to obtain position information from images, i.e. photogrammetry.

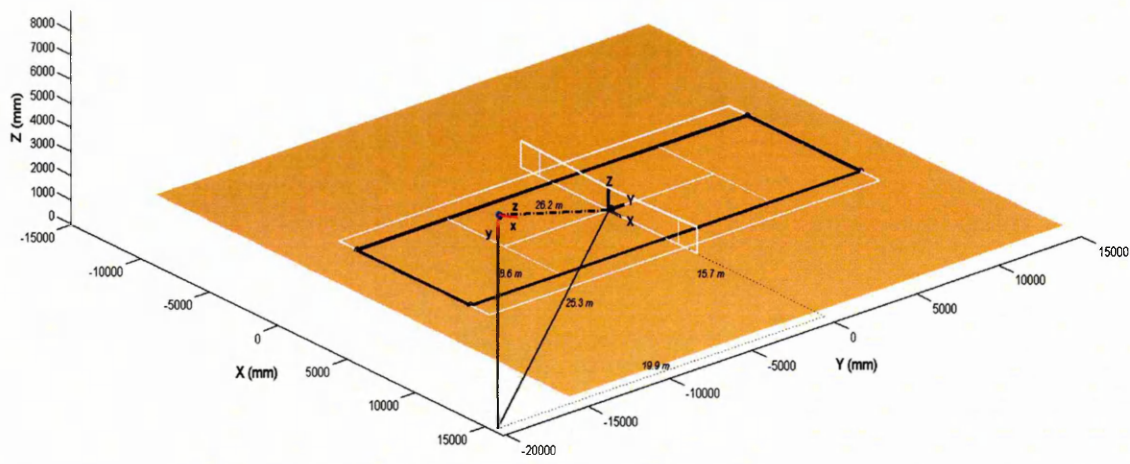


Figure 4.4. Camera-plane model describing extrinsic camera parameters (position and orientation) at the 2011 Roland Garros Qualifying Tournament.

4.3.3. Position reconstruction

To determine the location of an image coordinate in relation to the world coordinate system, i.e. tennis court (Figure 4.4), the two-dimensional (2×1), normalised image coordinate presented in equation 4.9, is modified to create a three-dimensional, 3×1 coordinate vector, where the z component is projected to infinity:

$$[x_n, y_n, 1]^T \quad [4.11]$$

The 3×1 coordinate vector (equation 4.11) is hereafter referred to as the camera ray. It is assumed that the camera ray will intersect the world plane, i.e. tennis court (Figure 4.4). This is true by virtue of extrinsic camera parameters. Subsequently, the location of a normalised image coordinate on the world plane is given by the magnitude of the camera ray's z component and is determined by line-plane intersection geometry. The normalised axes I and J define X and Y directions of the world plane respectively, i.e. tennis court (Figure 4.4), and are transformed into the camera coordinate system using rigid motion transformation, e.g. equation 4.10.

$$n = I \times J \quad [4.12]$$

Equation 4.12 yields the world plane normal vector, n . The dot product of two perpendicular vectors is zero. At the point of intersection, the vector between a camera

ray and world plane origin is perpendicular to the world plane normal vector. The magnitude of z at the point of intersection can thus be expressed by:

$$z = n \cdot (P - C) / n \cdot ([x_n, y_n, 1]^T - C) \quad [4.13]$$

Camera ray length, i.e. z , is substituted into the normalised image coordinate:

$$[x, y, z]^T = [x_n, y_n, z] \quad [4.14]$$

The resulting camera system coordinate is then inversely transformed into the world plane or tennis court coordinate system:

$$\begin{bmatrix} X \\ Y \\ 0 \end{bmatrix} = [R]^T \begin{bmatrix} x & - & t_x \\ y & - & t_y \\ z & - & t_z \end{bmatrix} \quad [4.15]$$

where n is the world plane normal vector, P is the world plane origin and C is the camera origin.

Assuming coplanarity, any normalised point in the camera image can be reconstructed with reference to the tennis court. Pixel-position reconstruction and position-pixel reprojection complete the camera-plane model: image coordinates can be identified in the world reference frame and world coordinates identified in the image reference frame. Model validity was assessed by reprojecting and reconstructing known coordinates in the world reference frame. A grid of world coordinates (X , Y and Z), approximating the dimensions of a singles tennis court, i.e. $8400 \times 24000 \times 0$ mm, spaced at 10 cm intervals, was reprojected into a camera image obtained from the 2011 Roland Garros Qualifying Tournament using corresponding intrinsic and extrinsic camera parameters (refer to Chapter 3). The world positions of image coordinates ($n = 20485$) were reconstructed as described by equations 4.11 – 4.15 and residuals between original and reconstructed coordinates calculated. Residual root-mean square error was 0.00, 0.00 and 0.00 mm for X , Y and Z directions respectively. The maximum residual was 3.2969×10^{-12} , reflecting a negligible effect of passing coordinate data through world-to-image-to-world reference frames and the validity of the camera-plane model.

4.4. Ground plane position reconstruction assessment

Agreement between real world position and camera-plane model photogrammetry must be assessed. Competitive sport environments typically restrict camera field-of-view, i.e. non-perpendicular images, influencing the accuracy of reconstructed coordinates and derived metrics (Brewin and Kerwin, 2003; McLean *et al.*, 2004). Therefore, accuracy assessment should be performed in restricted filming conditions experienced at competitive sport events. Further, a comparative measure is required. 2D-DLT is a popular method for reconstructing planar coordinates (Alcock, Hunter and Brown, 2009; Brewin and Kerwin, 2003; McLean *et al.*, 2004). Therefore position reconstruction using the camera-plane model will be assessed in relation to an existing implementation of 2D-DLT (Meershoek, 1997) for filming conditions experienced at two competitive tennis environments.

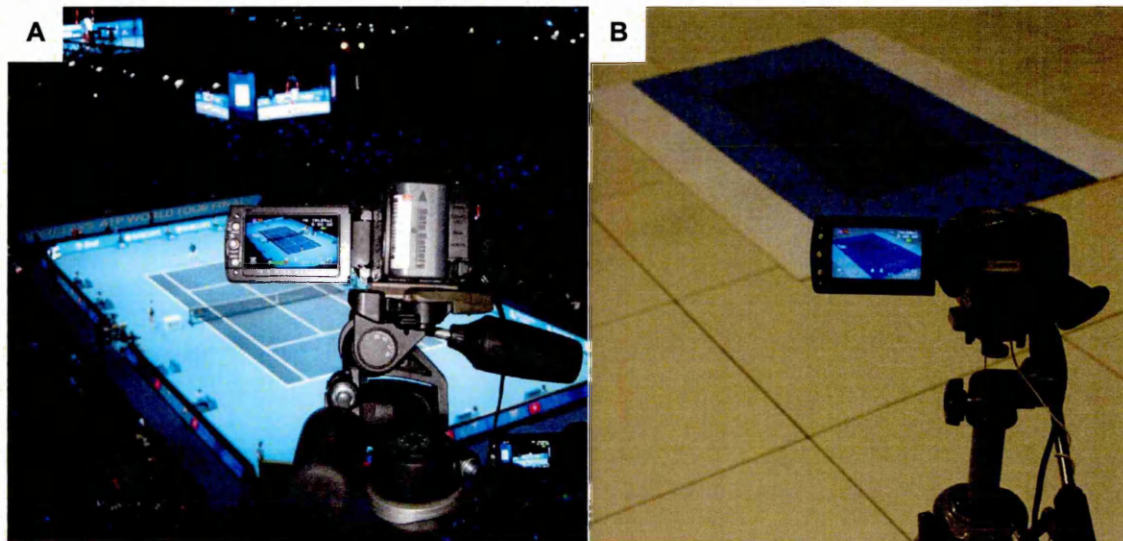
4.4.1. Competitive tennis environments: data collection and modelling

Stereoscopic checkerboard calibrations are typically performed within the motion capture volume (Choppin, 2008; Driscoll, 2012). In competitive tennis environments, i.e. A (Figure 4.5), camera calibration within the motion capture volume can be impractical due to restricted access to tennis courts. Furthermore, the literature review (section 2.2.3) highlighted that between 20 and 80 m of cabling, placed around sport stadia, would be required for stereoscopic filming, e.g. power, data transfer and generator locking. For publically accessible sport stadia, this would be impracticable to implement and potentially unsafe. For single camera calibration, there is no requirement to perform camera calibration within the motion capture volume, i.e. fixing camera position and orientation. This is because camera position relative to the motion capture plane can be defined independently: there is no requirement for a corresponding camera view, i.e. stereo photogrammetry. Furthermore, single cameras are independent: only a single, local power supply is required to film match-play tennis.

Permission and relevant accreditations to film at the 2011 ATP World Tour Finals were obtained via the Lawn Tennis Association. Furthermore, approval for all procedures was obtained from the Research Ethics Committee of the Faculty of Health and Wellbeing, Sheffield Hallam University (Appendix 3). The same high-definition video camera used in Chapter 3 (Everio GZ-HD40EK, JVC, Japan), was positioned in an elevated location (fourth level of O2 Arena), i.e. A (Figure 4.5). The camera filmed a

full-court field-of-view at 25 Hz (50 fields / second) with a resolution of 1920×1080 pixels (single CMOS sensor). The same filming procedure used at the 2011 Roland Garros qualifiers (RG) was used at the 2011 ATP World Tour Finals (ATP). In addition to the filming procedures described in Chapter 3, planar checkerboard calibrations were performed at both RG and ATP events.

Figure 4.5. Camera setup at the real (A) and scale model (B) of the 2011 ATP World Tour Final.



The desired camera field-of-view was first set by manually adjusting camera position, orientation and internal camera settings, i.e. zoom, focal length etc. By panning the camera and not altering internal camera settings, i.e. zoom, focal length etc., the checkerboard calibration was then performed, i.e. A (Figure 4.6). The checkerboard was filmed in different positions (less than 4 m) and orientations relative to the camera. The camera was subsequently replaced to the desired field-of-view; all subsequent analyses assumed that internal camera parameters did not change and the camera did not move. Calibration was performed *post-hoc* using Bouguet's (2010) 'Camera Calibration Toolbox for MATLAB'. Checkerboard square corners were extracted on a semi-automated basis to calculate intrinsic camera parameters. Extrinsic camera parameters were defined by passing four manually digitised image coordinates i.e. singles court line intersections (B: Figure 4.6), with their corresponding and world dimensions to the camera calibration toolbox.

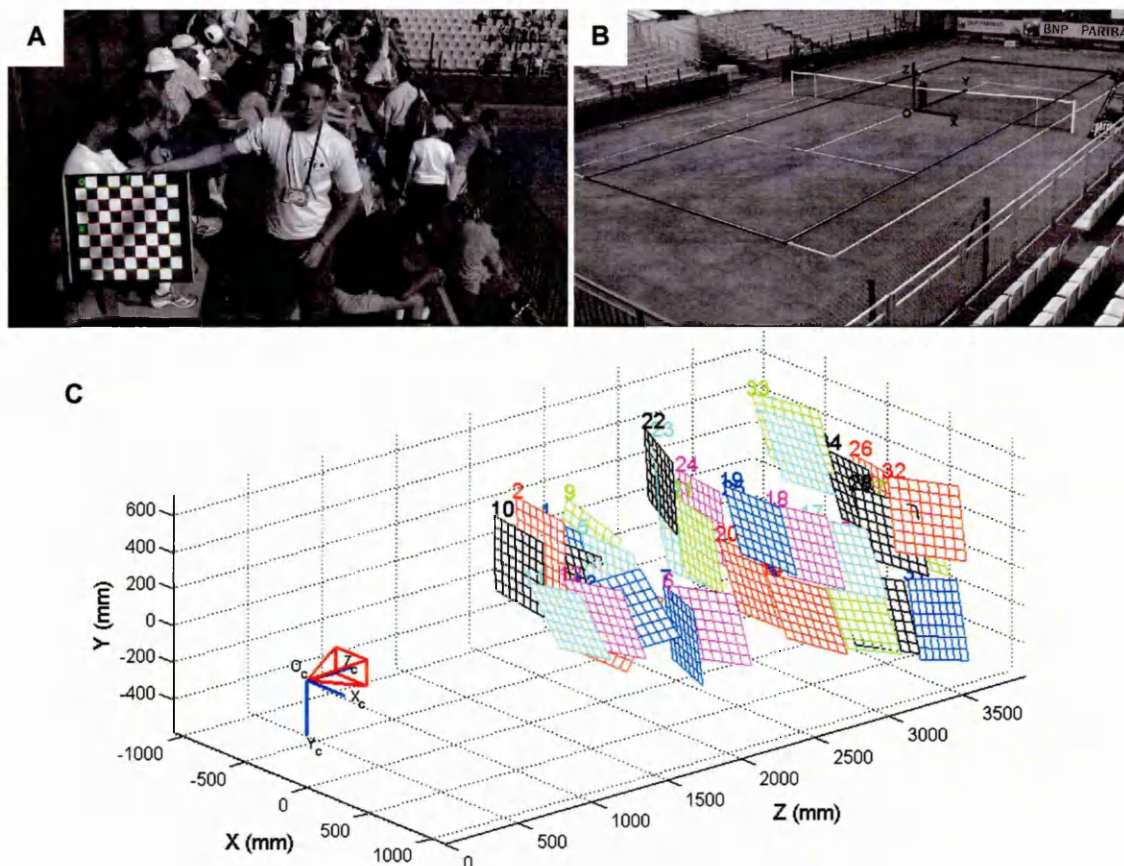


Figure 4.6. Checkerboard calibration at the 2011 Roland Garros Qualifiers: checkerboard corner extraction (A), tennis court field-of-view (B) and extrinsic checkerboard parameters (C).

Figure 4.6 (A) illustrates that for single camera calibration, the image of the checkerboard can be maximised within the camera image. Further, the checkerboard can be held in a wide variety of positions and orientations, covering the entire camera field-of-view (C: Figure 4.6). This helps to improve camera calibration (Choppin, 2008). Single camera calibration thus represents a flexible approach to sport stadia filming because intrinsic, i.e. focal length and lens distortion etc., and extrinsic camera parameters, i.e. position and orientation, can be calculated without accessing tennis courts or compromising public access to sport stadia.

Intrinsic and extrinsic camera calibration parameters were calculated for both RG and ATP filming setups. For the 2011 Roland Garros Qualifying Tournament, maximum image distortion was 40.7 pixels. Extrinsic camera parameters identified that camera elevation was 8.6 m, resultant translation (camera distance to tennis court origin) was 26.2 m and camera azimuth was 52.8° to the court's positive X axis (A: Figure 4.8). For

the 2011 ATP World Tour Finals, maximum image distortion was 3.3 pixels. Extrinsic camera parameters identified that camera elevation was 21.1 m, resultant translation was 63.0 m and camera azimuth was 117.2° to the court's positive X axis (Figure 4.7).

As previously noted, competitive tennis courts are typically inaccessible. Further, manually positioned reconstruction markers, i.e. criterion measure for image photogrammetry, will incur measurement error due to manual positioning. A 1:30 scale model of a tennis court (Appendix 4) was created using CorelDRAW (Graphics Suite 12, Corel, USA), printed on size A0 paper and affixed to a level, planar surface (B: Figure 4.5). Reconstruction markers ($n = 162$) and calibration markers (court line intersections; $n = 21$) were printed on the paper (represented by crosshairs) to aid manual digitising. Extrinsic camera parameters, obtained from the 2011 Roland Garros Qualifying Tournament and 2011 ATP World Tour Finals, were used to position and orientate the same camera in relation to scale models. Still images were then downloaded from the camera and compared to images of the real event, confirming camera position and orientation, i.e. Figure 4.7.

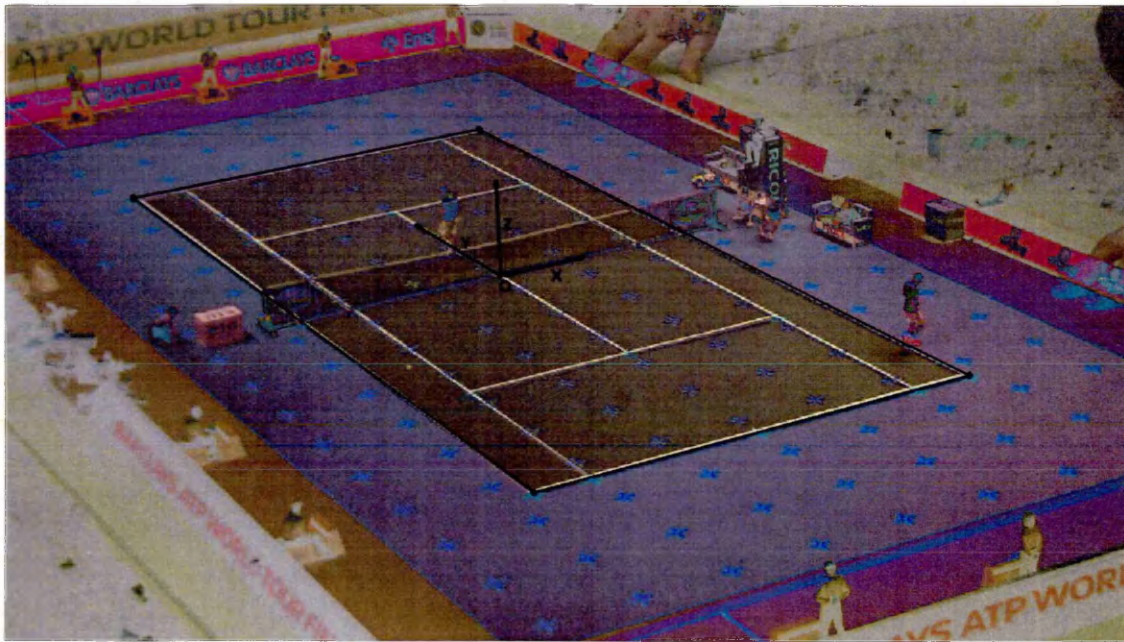


Figure 4.7. Combined camera image of real and scale model tennis courts of the 2011 ATP World Tour Finals. Crosshairs provide a criterion measure for photogrammetric assessment.

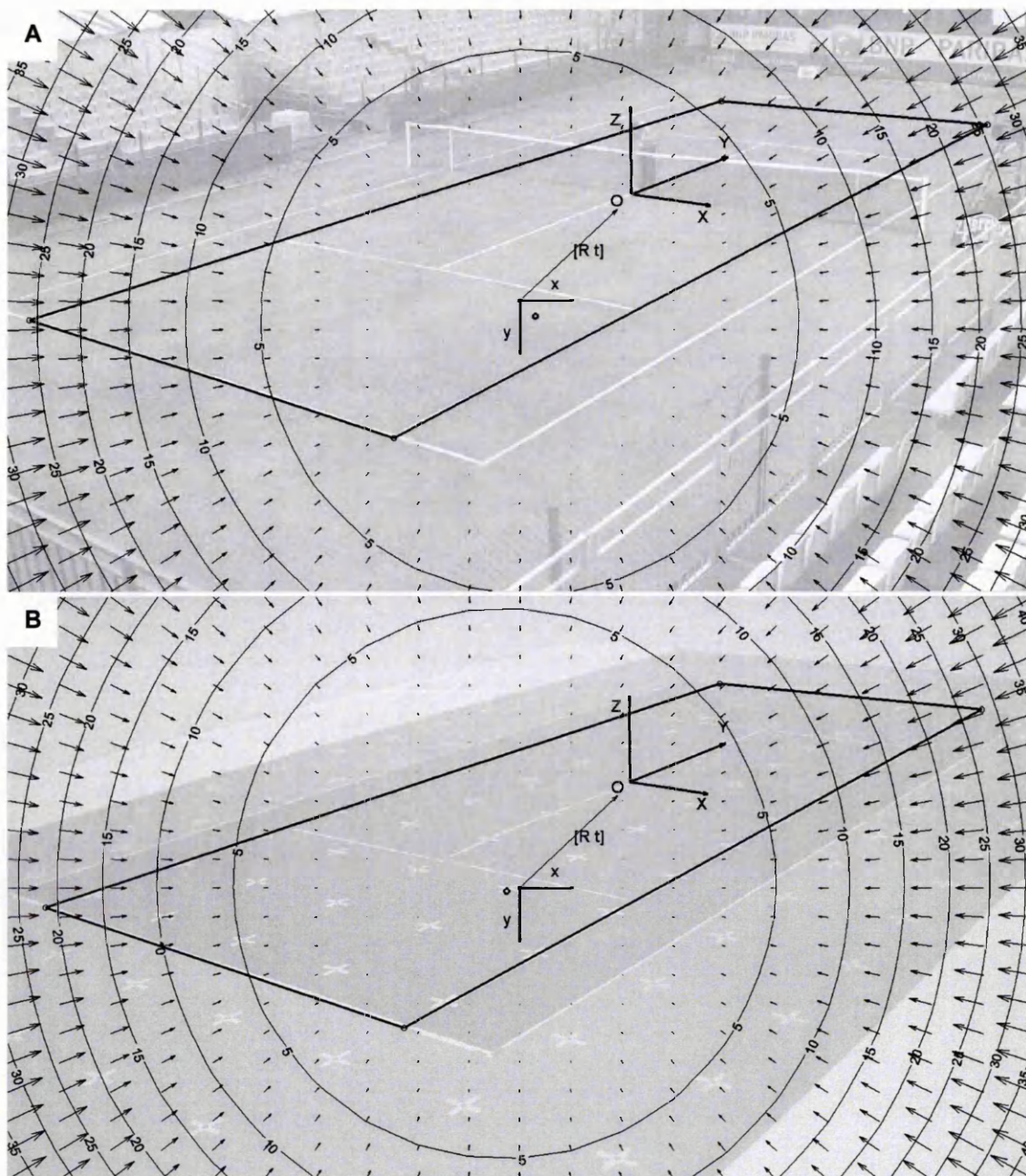


Figure 4.8. Real (A) and model (B) camera perspectives of the 2011 Roland Garros Qualifying Tournament. Images illustrate perspective projection and lens distortion (arrows and rings, values in pixels). $[R \ t]$ illustrates the homography between court (XYZ) and camera (xy) coordinate systems.

Camera shutter speed, aperture and focal length were set manually and then locked; the model tennis court, i.e. B (Figure 4.5) was then filmed for 5 s for both RG and ATP camera perspectives. Camera calibration was then performed to determine intrinsic and extrinsic camera parameters for RG and ATP model camera perspectives. Extrinsic camera parameters for the RG model identified that camera elevation was equivalent to

8.8 m, resultant translation was equivalent to 26.8 m and camera azimuth was 51.8° to the court's positive X axis (B: Figure 4.8). Further, maximum image distortion was 43.2 pixels. Extrinsic camera parameters for the ATP model identified that camera elevation was equivalent to 22.0 m, resultant translation was equivalent to 63.6 m and camera azimuth was 117.9° to the court's positive X axis, e.g. Figure 4.7. Further, maximum image distortion was 1.5 pixels. Similar camera position, field-of-view and image distortion magnitudes for both models, i.e. Figures 4.7 and 4.8, illustrate the efficacy of the scale model approach.

Reconstruction and calibration markers within the camera field-of-view were manually digitised at a sub-pixel resolution on five occasions. Standard error of the mean was up to 0.2 and 0.1 pixels for u and v image coordinates respectively for both RG and ATP models. Raw image coordinates of reconstruction points were reconstructed using existing 2D-DLT algorithms (Meershoek, 1997). The number of calibration points passed to the 2D-DLT algorithm was incremented, due to the impact of additional calibration coordinates on 2D-DLT reconstruction accuracy (McLean *et al.*, 2004). Calibration points were incremented from four to 15 (RG model) and from four to 21 (ATP model). The maximum number of calibration points was constrained by the number of observable court line intersections. The same image coordinates were then passed to the planar reconstruction method using only the initial four calibration coordinates, i.e. Figure 4.8.

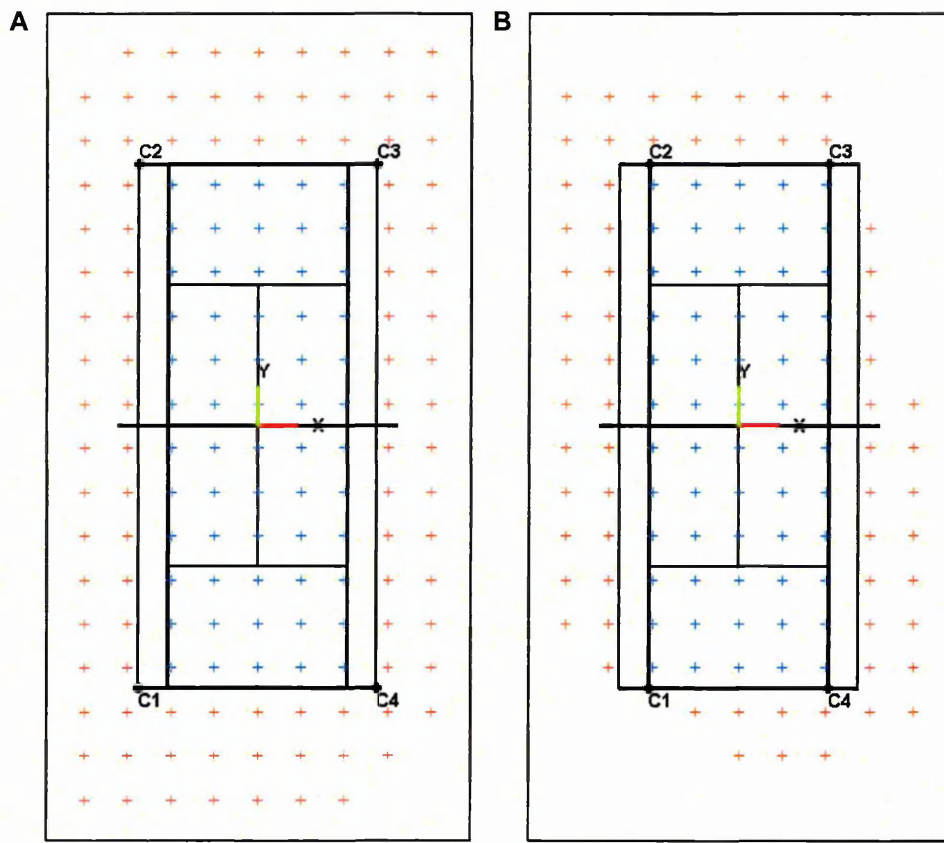


Figure 4.9. Internal (blue pluses) and external (red pluses) reconstruction points and calibration points (C1 - C4) for ATP (A) and Roland Garros (B) models.

For planar reconstruction, image coordinates were normalised to the camera coordinate system (Bouguet, 2010) and reconstructed as described by equations 4.11 – 4.15. All reconstructed coordinates were sorted to identify coordinates located inside or outside of tennis court markings, i.e. internal and external coordinates respectively (Figure 4.9). Root-mean square error (RMSE) between world and reconstructed coordinates was calculated for the X , Y , i.e. net and centreline directions respectively, e.g. Figures 4.7 and 4.8, and resultant (R) directions with the following:


$$RMSE = \sqrt{\sum_{i=1}^N (X_{iR} - X_{ir})^2 / N} \quad [4.16]$$

where X_{iR} is the world coordinate, X_{ir} is the reconstructed coordinate and N is the number of coordinates used.

4.5. Results

Reconstruction data were derived from footage of a 1:30 scale model of a tennis court, filmed with camera perspectives that correspond to real, competitive sport environments. Reconstruction coordinates, i.e. observable court line intersections, were reconstructed using 2D-DLT and planar reconstruction methods. For the Roland Garros model (Table 4.1), RMSE in the R direction ($RMSE_R$) inside tennis court markings was 81.4 to 166.1 mm using 2D-DLT compared to 67.7 mm using planar reconstruction. For points located outside of tennis court markings, $RMSE_R$ was 121.3 to 166.5 mm using 2D-DLT compared to 78.0 mm using planar reconstruction (Table 4.1).

Table 4.1. RMSE (mm) for X , Y and R directions using 2D-DLT (incremented calibration points) and planar reconstruction for a scale model of the 2011 Roland Garros Qualifying Tournament.

		2D-DLT						Planar	Court locations
		4	6	8	12	14	15		
Roland Garros Internal (blue: $n = 60$)	X	61.0	45.7	39.7	40.7	40.2	39.8	55.9	
	Y	154.5	81.0	80.4	80.7	76.3	71.0	38.1	
	R	166.1	93.0	89.7	90.3	86.3	81.4	67.7	
Roland Garros External (red: $n = 64$)	X	75.3	70.0	67.9	69.5	69.5	71.6	69.7	
	Y	148.5	116.5	114.4	99.4	104.0	105.8	34.9	
	R	166.5	135.9	133.0	121.3	125.1	127.8	78.0	

For 2D-DLT, RMSE in the Y axis ($RMSE_Y$) represented the largest component of error. Conversely, $RMSE_Y$ was the smallest component of error for planar reconstruction. For 2D-DLT, $RMSE_Y$ was 38.2 to 116.4 and 64.5 to 113.6 mm higher than planar reconstruction for internal and external reconstruction points respectively (Table 4.1). For the ATP model (Table 4.2), $RMSE_R$ inside tennis court markings was 37.1 to 42.0 mm using 2D-DLT compared to 56.6 mm using planar reconstruction. For points located outside of tennis court markings, $RMSE_R$ was 85.8 to 106.3 mm using 2D-DLT compared to 57.6 mm using planar reconstruction. For both 2D-DLT and planar reconstruction, $RMSE_Y$ represented the largest component of error. For 2D-DLT, $RMSE_R$ was 14.6 to 48.7 mm lower and 28.2 to 46.8 mm higher than planar reconstruction respectively (Table 4.2).

Table 4.2. RMSE (mm) for X , Y and R directions using 2D-DLT (incremented calibration points) and planar reconstruction for a scale model of the 2011 ATP World Tour Finals.

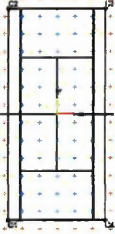
		2D-DLT								Planar	Court locations
		4	5	10	12	14	18	20	21		
ATP Internal (blue: $n = 60$)	X	17.3	15.5	15.5	15.8	15.8	16.6	16.3	16.2	17.6	
	Y	36.1	37.8	33.7	36.5	35.8	38.6	38.3	38.5	53.8	
	R	40.0	40.9	37.1	39.8	39.1	42.0	41.6	41.8	56.6	
ATP External (red: $n = 98$)	X	34.7	36.9	31.6	34.1	34.4	36.5	36.3	36.2	25.9	
	Y	98.5	99.6	79.7	82.2	79.3	81.9	83.5	83.2	51.5	
	R	104.4	106.3	85.8	89.0	86.5	89.7	91.1	90.8	57.6	

Figure 4.10 presents RMSE for 2D-DLT (red) and planar reconstruction (green) RMSE for both internal and external reconstruction points of the RG model. Reconstruction error in the X axis was lower for 2D-DLT when using more than four calibration points (Figure 4.10). However as previously noted, reconstruction error in the Y axis was the largest component of resultant direction reconstruction error for 2D-DLT. The inclusion of more than 12 calibration points did not substantially affect reconstruction error for 2D-DLT; incremented calibration points reduced reconstruction error (R direction) by 0.1 and 1.1 mm for 12 – 14 and 14 – 15 calibration points respectively (Figure 4.10). However, reconstruction error in the resultant direction for 2D-DLT was greater than planar reconstruction irrespective of the number of calibration points passed to the algorithm.

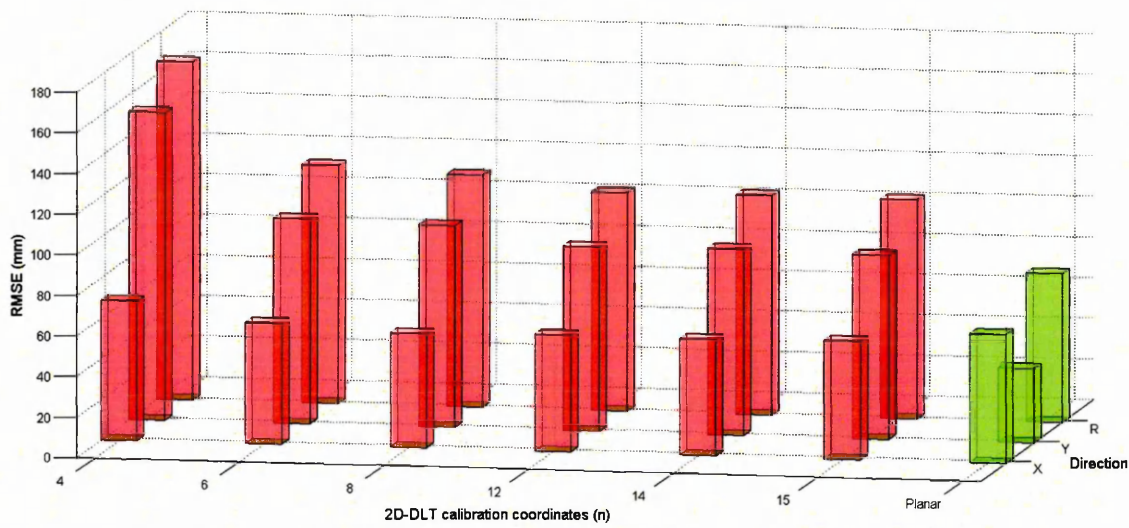


Figure 4.10. Internal and external ($n = 124$) RMSE (mm) for 2D-DLT (red: incremented calibration coordinates) and planar reconstruction (green) in X , Y and R directions for the RG model.

Figure 4.11 similarly presents RMSE for 2D-DLT (red) and planar reconstruction (green) RMSE for both internal and external reconstruction points of the ATP model. Figure 4.11 demonstrates that for 2D-DLT, $RMSE_Y$ represented the major component of $RMSE_R$, reflecting trends exhibited by the RG model (Figure 4.10). Figure 4.11 illustrates that for internal and external position reconstruction, 2D-DLT yielded greater RMSE magnitudes than planar reconstruction, irrespective of the number of calibration points passed to 2D-DLT. The inclusion of more than five calibration points reduces reconstruction error for 2D-DLT; however the use of 18 or more calibration points increases 2D-DLT reconstruction error (Figure 4.11).

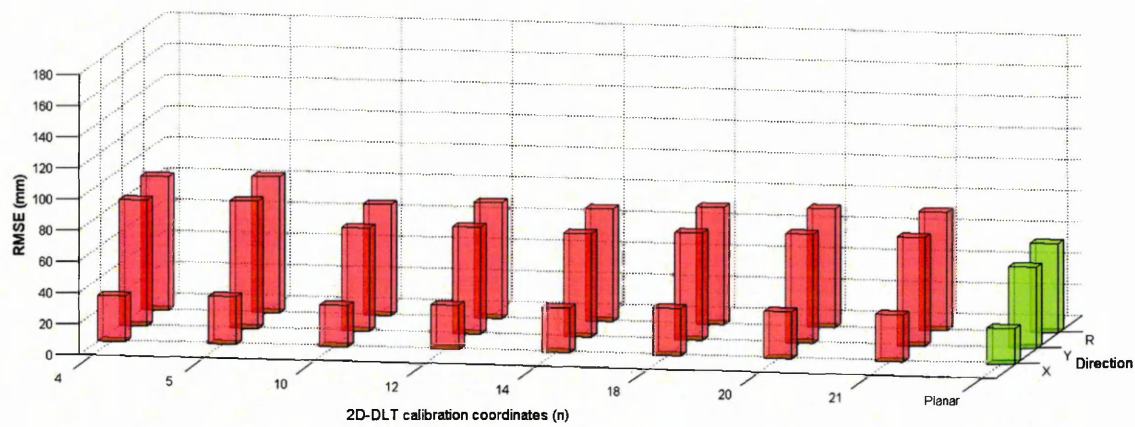


Figure 4.11. Internal and external ($n = 158$) RMSE (mm) for 2D-DLT (red: incremented calibration coordinates) and planar reconstruction (green) in X , Y and R directions for the ATP model.

Table 4.3 presents calibration plane pixel-scale, i.e. camera-plane estimate of individual pixel dimension across the tennis court model, and calibration point distortion, i.e. composite of radial and tangential distortion. Pixel-scale estimates suggest a similar resolution for X and Y directions of the RG model, however Y direction pixel-scale for the ATP model was approximately five times higher than in the X direction. Calibration points for the ATP model exhibit low distortion magnitudes, i.e. 0.1 – 0.6 pixels (Table 4.3). Calibration points for the RG model exhibit high distortion magnitudes that were unevenly distributed across the calibration plane, i.e. 1.3 – 27.8 pixels (Table 4.3). Finally, Table 4.3 presents residual r^2 , i.e. correlation between the absolute differences and the measurement mean (Nevill and Atkinson, 1997), for planar reconstruction and 2D-DLT (using four calibration points). Residual correlation coefficients identified a strong negative correlation for planar reconstruction in the X direction for the RG model.

Table 4.3. Mean calibration plane pixel-scale (mm; *X*, *Y* and *R* directions), calibration point image distortion (pixels; C1 – C4) and residual r^2 (*X* and *Y* directions) for ATP and RG models.

		Pixel-scale (mm)	Calibration point distortion (pixels)				Residual r^2 (Calibration points: C1 - C4)	
			C1	C2	C3	C4	2D-DLT	Planar
ATP	X	8.7			-		-0.14	-0.16
	Y	48.8			-		0.41	0.41
	R	35.0	0.2	0.3	0.1	0.6	-	-
RG	X	36.6			-		0.16	-0.79
	Y	40.9			-		0.07	0.41
	R	38.9	21.7	6.1	27.8	1.3	-	-

4.6. Discussion

Position reconstruction accuracy for two different camera perspectives of competitive sport environments was compared. Camera field-of-view at the 2011 ATP World Tour Finals was unique, providing high camera elevation and large resultant camera translation, i.e. 21.1 and 63.0 m respectively. As such, camera field-of-view was zoomed-in, i.e. narrow filming angle. The effects of lens distortions were thus minimised (Bradski and Kaehler, 2008). This was evident by low image distortion: maximum image distortion was 3.3 pixels. However, ATP camera elevation represents large stadia, i.e. 20 – 30 m elevation (John, Sheard and Vickery, 2007), and are not typical of competitive tennis events. John, Sheard and Vickery (2007) suggest typical stadia elevations range between 7.3 – 13.0 m with typical viewing distances of 30.0 m. Camera field-of-view at the 2011 Roland Garros Qualifying Tournament was within this range; camera elevation and resultant translation were 8.6 and 26.2 m respectively. Due to close range filming and low camera elevation, camera field-of-view was zoomed-out, i.e. wide filming angle. This induced large image distortions: maximum image distortion was 40.7 pixels. At smaller sport events, it is conceivable that camera elevation and viewing distance would be lower than the 2011 Roland Garros Qualifying Tournament, thus requiring wider filming angles. This might require greater image distortion. When filming a football pitch, Alcock, Hunter and Brown (2009) reported camera elevation of 7 m for a camera located 3 m from the halfway-touchline intersection. Such close range filming would require a wide filming angle and would

have induced image distortions due to the lens. Alcock, Hunter and Brown (2009) reported 2D-DLT reconstruction errors of 0.35 ± 0.27 m; approximately twice the magnitude of 2D-DLT reconstruction error for the RG model (RMSE_R using four calibration points; Table 4.1). This highlights the impact of filming location on photogrammetry accuracy.

For the RG model, RMSE_X magnitudes were lowest using 2D-DLT when using six or more calibration points (Figure 4.10). However it is likely that for most practical filming applications, markedly higher RMSE_R magnitudes for 2D-DLT would indicate planar reconstruction to be the most appropriate reconstruction method. Magnitudes of RMSE_X and RMSE_Y for 2D-DLT and planar reconstruction are dichotomous; position reconstruction in the *X* direction was indicated to be more accurate using 2D-DLT than planar reconstruction and vice-versa in the *Y* direction. Table 4.3 demonstrates that for the RG model, pixel-scale of the calibration plane was similar, i.e. 36.6 and 40.9 mm per pixel in *X* and *Y* directions respectively. As such, poor image resolution in the *Y* direction was unlikely to have contributed to higher RMSE_Y for planar reconstruction. Table 4.3 also indicates that calibration points were distorted, i.e. 1.3 – 27.8 pixels. Further, the distribution of distortion was uneven across calibration points, skewing the calibration plane to a non-orthogonal shape. 2D-DLT is dependent on the direct mapping of image to world coordinates, i.e. image linearity (Kwon, 2012). Different distortion magnitudes and non-orthogonality would potentially violate the underlying assumptions of 2D-DLT and partly explain larger RMSE for 2D-DLT in the *Y* direction.

Residual r^2 (Table 4.3) indicates a strong negative correlation for planar reconstruction in the *X* direction. This indicates that position reconstruction residuals for coordinates to the left of the tennis court origin, i.e. negative *X* (Figure 4.9), were greater than residuals to the right, i.e. positive *X* (Figure 4.9). It is likely that position estimates were contaminated by manual digitising error in the highly distorted image. Due to camera perspective, crosshair centres on the negative *X* of the model would have been more difficult to identify and digitise, i.e. lower pixel resolution. In an internal report (refer to Appendix 5), Whyld (2004) noted that because standard DLT algorithms allow axes to be non-orthogonal, residuals categorised into orthogonal reference frames provide no guidance to the true level of accuracy. Therefore it is likely that planar reconstruction provides a truer indication of orthogonal reconstruction error due to image

normalisation and forced orthogonality (Whyld, 2004). Due to the potential violation of orthogonality for 2D-DLT, orthogonal estimates for 2D-DLT reconstruction error, i.e. $RMSE_X$ and $RMSE_Y$, will be inaccurate as standard DLT algorithms merely fit data to parameter equations. 2D-DLT will thus yield seemingly accurate results for inherently inaccurate scenarios (Whyld, 2004). Therefore, only $RMSE_R$ should be used to infer accuracy for 2D-DLT. When considering $RMSE_R$ for internal and external tennis court areas (Figure 4.10), planar reconstruction yielded lower magnitudes for RMSE regardless of the number of calibration points passed to 2D-DLT. This highlights that, for lens distorted images, position reconstruction accuracy using 2D-DLT can be limited, as lens distortion cannot be considered in linear equation solving (Tsai, 1987).

For the ATP model, $RMSE_Y$ was the largest component of reconstruction error for both internal and external reconstruction points for both 2D-DLT and planar reconstruction methods (Table 4.2). For the ATP model, it is likely that 2D-DLT assumptions, i.e. image linearity and orthogonal calibration plane, were not violated. This was supported by low image distortion magnitudes, i.e. up to 3.3 pixels, and low distortion magnitudes for calibration points, i.e. 0.1 – 0.6 pixels (Table 4.3). Table 4.3 indicates that pixel-scale, i.e. physical dimension of an individual pixel, was markedly greater in the Y direction, i.e. 8.7 and 48.8 mm per pixel for X and Y directions respectively. This highlights that image resolution was markedly lower in the Y direction, e.g. Figures 4.7 and 4.8. Inaccuracies in manual digitising would have a corresponding impact to position reconstruction for both methods.

For the ATP model, 2D-DLT yielded lower position reconstruction errors for internal reconstruction points when compared to planar reconstruction (Table 4.2). However for external reconstruction points, 2D-DLT yielded higher position reconstruction errors. Higher reconstruction error using 2D-DLT for points located outside of calibration points has previously been reported (Brewin and Kerwin, 2003). 2D-DLT assumes that reconstruction points lay within calibration points (Kwon, 2012). Extrapolating calibration parameters to reconstruct image coordinates outside of calibration points violates the underlying assumptions of 2D-DLT. However for match-play tennis, reconstructing position outside of calibration points, i.e. tennis court markings, is a common and necessary practice. For planar reconstruction, $RMSE_R$ was 56.6 and 57.6 mm for internal and external reconstruction points respectively (Table 4.2). Further,

Figure 4.11 demonstrates that reconstruction error (internal and external) using planar reconstruction was lower than 2D-DLT irrespective of the number of calibration points passed to the algorithm. This reflects that location of position reconstruction is not a limiting factor for planar reconstruction and represents an important consideration when filming competitive sports.

McLean *et al.* (2004) demonstrated a general reduction of reconstruction error for 2D-DLT as the number of calibration points used was increased. However, reconstruction accuracy did not always improve upon each increment. For the ATP model, 2D-DLT $RMSE_R$ initially falls but subsequently increases if more than 10 calibration points are passed to the algorithm (exception of 14 calibration points; Table 4.2 and Figure 4.11). McLean *et al.* (2004) cite a constructed calibration board with physical marker position errors of up to 0.1 cm from which reconstruction accuracy estimates were made. Given the context of competitive sport filming environments, it is likely that manually digitising additional calibration points exacerbated random digitising error, yielding non-orthogonal calibration points (Whyld, 2004).

Standard DLT algorithms allow axes to be non-orthogonal. This can yield seemingly accurate results for inherently inaccurate scenarios (Whyld, 2004). In contrast, planar reconstruction requires the definition of a plane; the cross-product of the two axes defines a normal vector. Whyld (2004) deemed planar reconstruction to be a more robust and versatile camera calibration method. Therefore, given the context of competitive sport environments, i.e. non-perpendicular and lens distorted images, it is likely that manually digitising additional calibration points to reduce reconstruction error will give rise to other factors that compromise 2D-DLT reconstruction accuracy. For the context of current study, the planar reconstruction method represented the most accurate and flexible method for two-dimensional photogrammetry.

4.7. Conclusion

A camera-plane model, capable of calculating world-to-image and image-to-world coordinate transformations was developed. Photogrammetric accuracy was assessed for camera perspectives that correspond to filming conditions experienced at two unique competitive sport events using a popular method for planar position reconstruction. When using the planar reconstruction method, reconstruction errors demonstrate

improved reconstruction accuracy, i.e. $RMSE_R$, for filming conditions typical of competitive sport events.

For match-play tennis, accurate reconstruction of position outside of tennis court markings is as important as accurate position reconstruction inside tennis court markings. In this regard, position reconstruction using the planar reconstruction method yielded similar $RMSE_R$ for locations inside and outside of court markings. This highlights that photogrammetric error incurred by the relative location of position reconstruction, is not a limiting factor for planar reconstruction method. Findings demonstrate the versatility of the planar reconstruction method for reconstructing position data from footage obtained in restricted, competitive sport events. The presented method will be useful when filming is restricted by camera location, access to the activity plane is impermissible, a limited number of calibration points exist and when position reconstruction is required outside of calibration points.

5 Single camera player position reconstruction using an elevated calibration plane

5.1. Introduction

Chapter 4 demonstrated the planar reconstruction method to be an accurate and flexible method for two-dimensional photogrammetry in match-play tennis. However, reconstructing player position has not been addressed. The literature review highlighted that a variety of methods to define player position in sport exist. Previous approaches have defined player position as the centre of a player silhouette (Jiang *et al.*, 2009) or as the mean horizontal and maximum vertical silhouette coordinates (Mauthner *et al.*, 2008). Both approaches reconstruct player coordinates using a ground level calibration plane. However, these approaches either neglect the effects of camera perspective or do not identify player COM. In sport, quantifying player COM motion yields important information about player movement strategy (Bartlett, 2007). The disparity in definitions used to identify player position reflects the difficulty of reliably identifying image features that correspond to player COM using a single camera without reference points. For single camera footage of match-play tennis, a player COM definition must satisfy the following:

- Be determined without markers and unduly affected by camera perspective.
- Be suitable for dynamic activities, i.e. standing, running, etc.
- Be applicable to filming constraints of match-play tennis, i.e. off-court analysis.

5.2. Aim and objectives

The aim of this chapter was to develop and assess the elevated calibration plane used to reconstruct player position (section 3.3.3) using simulated camera perspectives of competitive sports events. This relates to box E of the development stage diagram (Figure 1.2).

Objectives:

1. Identify a method for reconstructing player position.
2. Assess position reconstruction for different camera perspectives and activities.
3. Assess position reconstruction in relation to an existing method.

5.3. Out-of-plane error

For 2D planar analyses, calibration and reconstruction coordinates are assumed to be coplanar. Figure 5.1 indicates that the magnitude of reconstruction error (Camera A) for the out-of-plane coordinate (X_m, Y_m) would be equal to $(|X_m^a - X_m|, |Y_m^a - Y_m|)$.

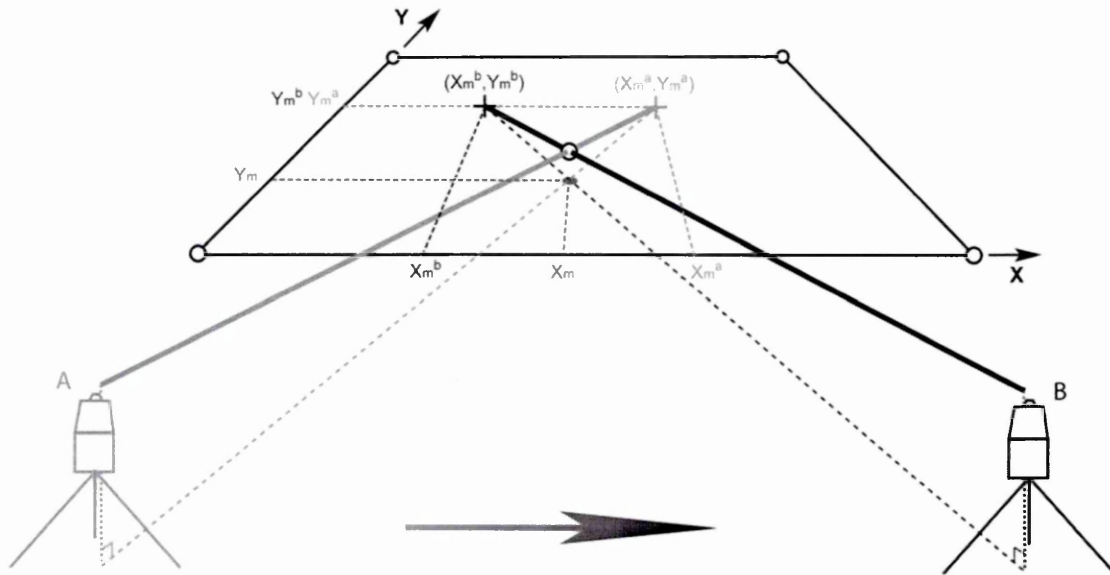


Figure 5.1. Reconstruction error incurred when the reconstructed point and reconstruction plane are not coplanar (adapted from Holden-Douilly *et al.*, 2011).

Sih, Hubbard and Williams (2001) highlighted that out-of-plane error is related to camera distance and perspective. Further, Hinrichs *et al.* (2005) highlighted that non-perpendicular camera angles substantially increase out-of-plane reconstruction error. Therefore, given the context of sport stadia filming, out-of-plane error will be a significant factor for player position reconstruction accuracy, e.g. player COM approximately 1 m out-of-plane. For single camera views that are not perpendicular to the calibration plane, it is inappropriate to reconstruct out-of-plane coordinates, i.e. player COM, using a ground level calibration plane, i.e. Jiang *et al.* (2009). Mauthner *et al.* (2008) highlighted this limitation and reported ground level player position by reconstructing a location about the feet. Using this definition, Mauthner *et al.* (2008) reported mean differences of 0.3 m for player position tracking in relation to manually digitised data. However, the validity of player position was not explicitly tested.

Two-dimensional out-of-plane error can be corrected. Holden-Douilly *et al.* (2011) presented an image-based method to minimise out-of-plane errors using 2D-DLT for

equine gait analysis. For a camera moving perpendicular to the calibrated plane, i.e. camera A \rightarrow B (Figure 5.1), Holden-Douilly *et al.* (2011) demonstrated that for a marker known to be static for a time period, i.e. during stance, vertical out-of-plane distance could be estimated. Minimising the length of an arc yielded by out-of-plane position reconstruction, i.e. camera translation relative to static marker, an estimate for out-of-plane distance was derived. For calibration markers ± 10 cm out-of-plane, Holden-Douilly *et al.* (2011) reported corrected marker position residuals of less than 0.55%. However, the method relied on camera translation relative to a static marker. This is not realistic for match-play tennis. Sih, Hubbard and Williams (2001) presented an out-of-plane correction factor based on non-image information, i.e. physical measurement of out-of-plane distance. However, as previously noted, physical measurements are impractical for the context of match-play tennis. In the absence of physical measurement, an estimate for out-of-plane distance could be implemented. However, the correction factor presented by Sih, Hubbard and Williams (2001) is systematic. The application of a systematic correction to oscillatory motion, i.e. player COM motion, would potentially exacerbate position reconstruction error. Further, the method presented by Sih, Hubbard and Williams (2001) assumes image linearity. As demonstrated in Chapter 4, the use of wide filming angles can violate image linearity assumptions, i.e. image distortions due to the lens.

Chapter 3 used an elevated calibration plane to reconstruct player COM position. The elevated calibration plane was assumed to intersect player COM and to reduce out-of-plane distance. However calibration points were manually identified using a calibration object and it was not possible to quantify player position reconstruction accuracy (refer to Chapter 3). The following sections detail the development of the method presented in Chapter 3 and assessment of player position reconstruction accuracy for filming conditions experienced at two competitive tennis events.

5.4. *Simulating sport stadia camera views*

Chapter 4 presented a camera-plane model that allowed the reconstruction of image coordinates in a world reference frame and the reprojection of world coordinates in an image reference frame. The camera-plane model can therefore be used to define calibration points required to construct an elevated calibration plane. Filming and digitising a real world object that corresponds to reprojected coordinates can assess the

accuracy of reprojected coordinates, i.e. Technical Error of Measurement (TEM; Gore, 2000). As described in section 3.3.3, a rigid object, set to net height (914 mm) and held vertically (using a spirit level) was placed at four court locations corresponding to the ground calibration plane, e.g. singles court line intersections (Figure 3.3). Following calibration of intrinsic and extrinsic camera parameters (using court line intersections), a calibration plane, of corresponding dimension to the ground plane but elevated to 914 mm, was reprojected into the camera image. Reprojected coordinates for the elevated calibration plane were compared to manually digitised coordinates, e.g. yellow filled circles (Figure 3.3). TEM ($\bar{x} \pm s$) was 1.44 ± 0.69 and 1.16 ± 0.87 pixels for u and v coordinates respectively. Agreement between reprojected and real world coordinates support the calculation of an elevated calibration plane in competitive tennis environments using the camera-plane model. However, reconstructing player position using an elevated calibration plane must be assessed.

Camera perspective should not unduly affect player position reconstruction. To assess player position reconstruction using an elevated calibration plane, camera perspectives of 3D point cloud data of a human participant were simulated. To simulate sport stadia camera perspectives, intrinsic and extrinsic camera parameters calculated from the 2011 Roland Garros Qualifying Tournament and 2011 ATP World Tour Finals were used (refer to Chapter 4). The literature review highlighted the relationship between camera and global coordinate systems (tennis court) when considering sport stadia, i.e. equation 2.1. Intrinsic camera parameters, identified by equation 4.2 (Chapter 4), can be incorporated to describe the infinite projection of an image coordinate into world space, in relation to camera azimuth, elevation and radius:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & c & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(az) & \sin(az) & 0 \\ -\sin(el) \times \sin(az) & \sin(el) \times \cos(az) & \cos(el) \\ \cos(el) \times \sin(az) & -\cos(el) \times \cos(az) & \sin(el) \end{bmatrix} + \begin{bmatrix} r \times \cos(el) \times \cos(az) \\ r \times \cos(el) \times \sin(az) \\ r \times \sin(el) \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad [5.1]$$

where α , β , c , u_0 and v_0 are intrinsic camera parameters described in equation 4.1 (section 4.3.1), az , el and r are camera azimuth, elevation and radius respectively.

Equation 5.1 is defined by camera and sport stadia parameters and approximates camera perspective. It is a simplified approximation because it does not account for rotations about the camera's optical axis. Furthermore, camera focal length and working distance, i.e. radius, are related. Therefore, for a given set of intrinsic and extrinsic camera parameters, only rotation matrices were modified, i.e. R (equation 4.10; Chapter 4). Thus intrinsic camera parameters, i.e. focal length etc., remain valid for camera translation. A caveat to this approach is that for sport stadia, camera radius will change in relation to camera azimuth and / or elevation angle. However, accounting for changes to camera radius would require the recalculation of intrinsic camera parameters, i.e. focal length. Whilst this is possible to resolve, sensor dimensions are required to calculate focal length for a given working distance. Further, the calculation of appropriate image distortions due to the lens, i.e. narrow or wide filming angle, as a function of camera working distance, would not be trivial. Therefore a fixed working distance model of sport stadia was adopted.

Using the camera-plane model, real world XYZ coordinates can be reprojected into a camera image. NEVA Electromagnetics (2013) provide freely accessible XYZ point cloud data of human participants in various postures. NEVA Electromagnetics (2013) collected point cloud data using a 3D Model WB4 Laser Scanner (Cyberware, CA, USA). Point cloud data, consisting of 3084 and 2752 data points for standing and running postures respectively, were downloaded for a single, 30 year old male participant (Figure 5.2). For illustrative purposes, 3D point cloud data presented in Figure 5.2 have been meshed (Vollmer, Mencl and Mueller, 1999); meshing was not required for analysis. The assumption of uniform density for body segments can be used to estimate inertial parameters and COM location for body segments (Ackland, Henson and Baily, 1988). As such, a body's volume can be used to estimate COM location. Filled red circles in Figure 5.2 indicate mean XYZ position of point cloud data and as such, approximate the player COM. Further, filled green circles represent the ground plane (XY) projection of player COM, i.e. $Z = 0$. The two-dimensional, XY projection of player COM was used as a criterion measure to assess photogrammetric position reconstruction accuracy.

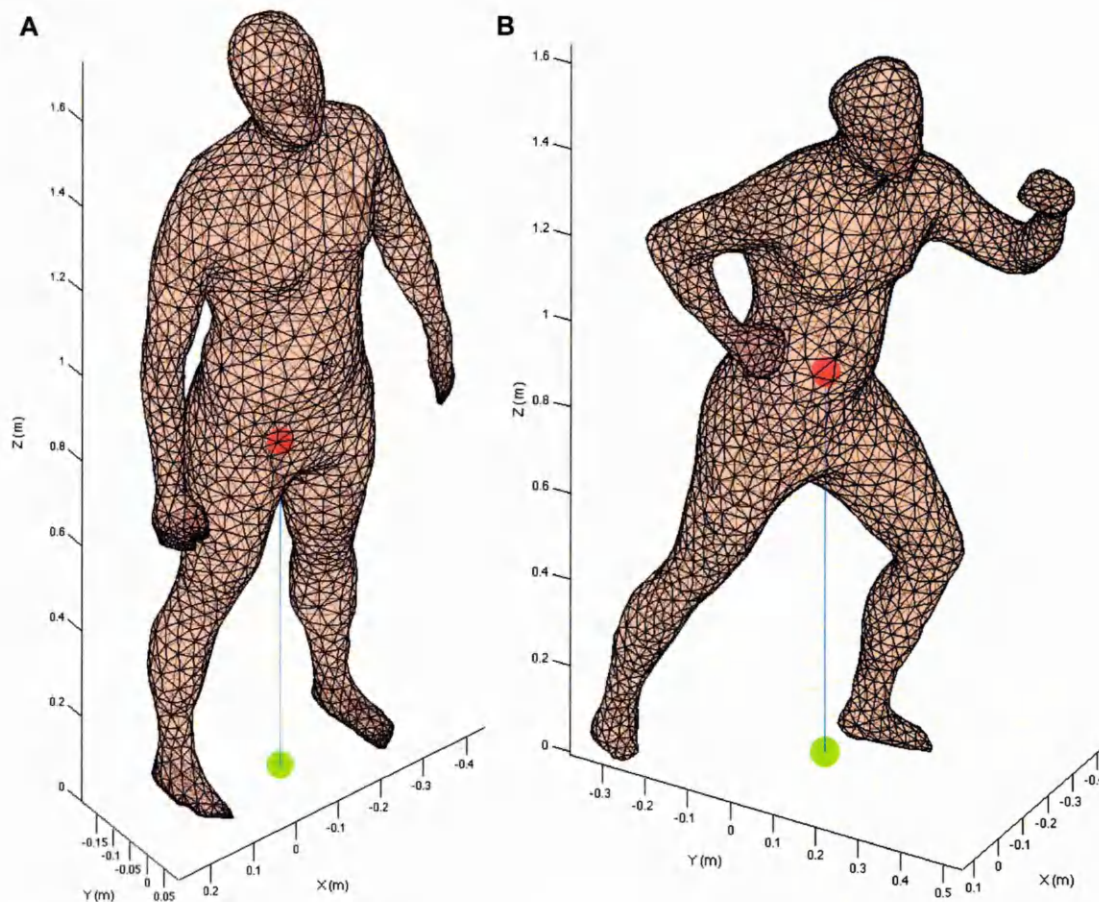


Figure 5.2. Coarsely meshed point cloud data for a human participant in standing (A) and running (B) postures.

Camera-plane models were based on camera calibrations obtained when filming the 2011 Roland Garros Qualifying Tournament and 2011 ATP World Tour Finals (described in Chapter 4). To simulate camera perspective, camera azimuth and elevation angles were incremented from $0 - 360^\circ$ and $10 - 50^\circ$ respectively (where 90° elevation represents an overhead view); each increment was 5° . This yielded 657 unique camera perspectives, i.e. Figure 5.3 ($73 \text{ azimuth} \times 9 \text{ elevation angle combinations}$), for which corresponding direction cosine rotation matrices were calculated. Camera translation, i.e. radius, and intrinsic parameters, i.e. focal length, lens distortion, etc., were not modified.

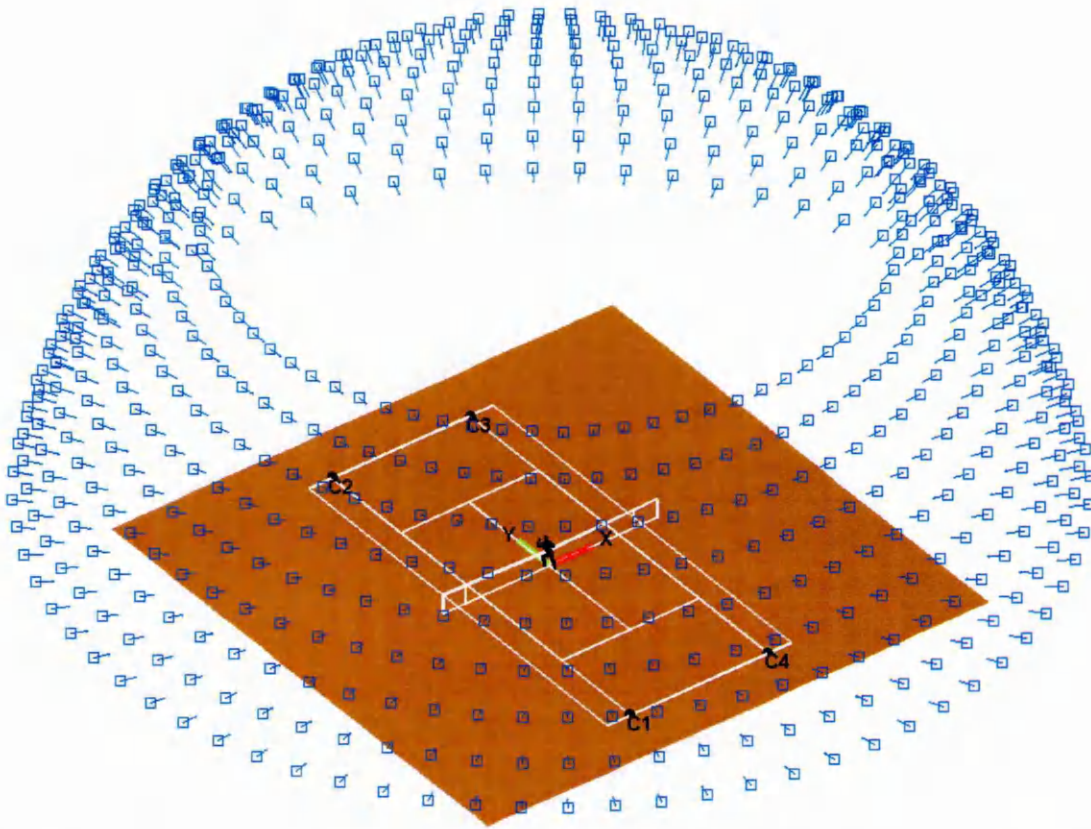


Figure 5.3. Simulated ($n = 657$) camera locations (blue squares: arrows indicate optical axis) relative to running point cloud data (black figure at court centre) for RG.

For each simulated camera perspective, real world XYZ point cloud data were reprojected into a blank camera image using the modified extrinsic camera parameters. The resulting binary image was then morphologically processed; images were dilated using a 5×5 structuring element and then filled. Figure 5.4 presents a processed binary image (cropped); filled red and green circles represent the image reprojection of COM and ground plane COM coordinates obtained from point cloud data, i.e. red and green filled circles presented in Figure 5.2.

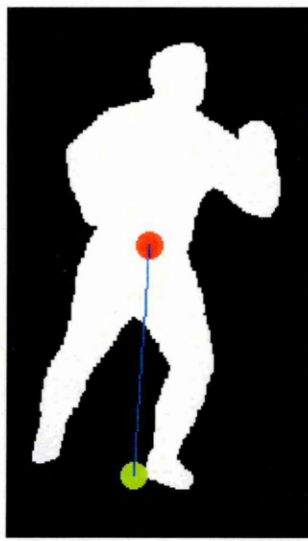


Figure 5.4. Processed camera image of reprojected point cloud data and reprojected COM (red) and ground plane COM (green) locations.

5.4.1. Player position reconstruction

Chapter 3 detailed a calibration plane elevated to 914 mm for reconstructing manually digitised estimates of player COM. Plane elevation was based on net height and as such, was an arbitrary reference for player COM. Ackland, Henson and Baily (1988) demonstrated that, by adopting the uniform density assumption, body volume can be used to estimate COM location. For a binary image, the centroid is equivalent to the COM. Therefore, for each simulated camera perspective image, the binary image centroid was used to identify COM, e.g. blue diamond (A: Figure, 5.5). Furthermore, based on the assumption of uniform density, calibration plane elevation was defined as 50% of participant stature. Stature was determined as the maximum Z coordinate for standing posture data (stature = 1739 mm; Figure 5.2): elevated calibration plane coordinates were reprojected into images with an elevation of 869.5 mm, e.g. blue reconstruction plane (B: Figure 5.5). For each simulated camera image, the image centroid was reconstructed using the elevated calibration plane, e.g. blue diamond (A: Figure 5.5) and blue reconstruction plane (B: Figure 5.5).

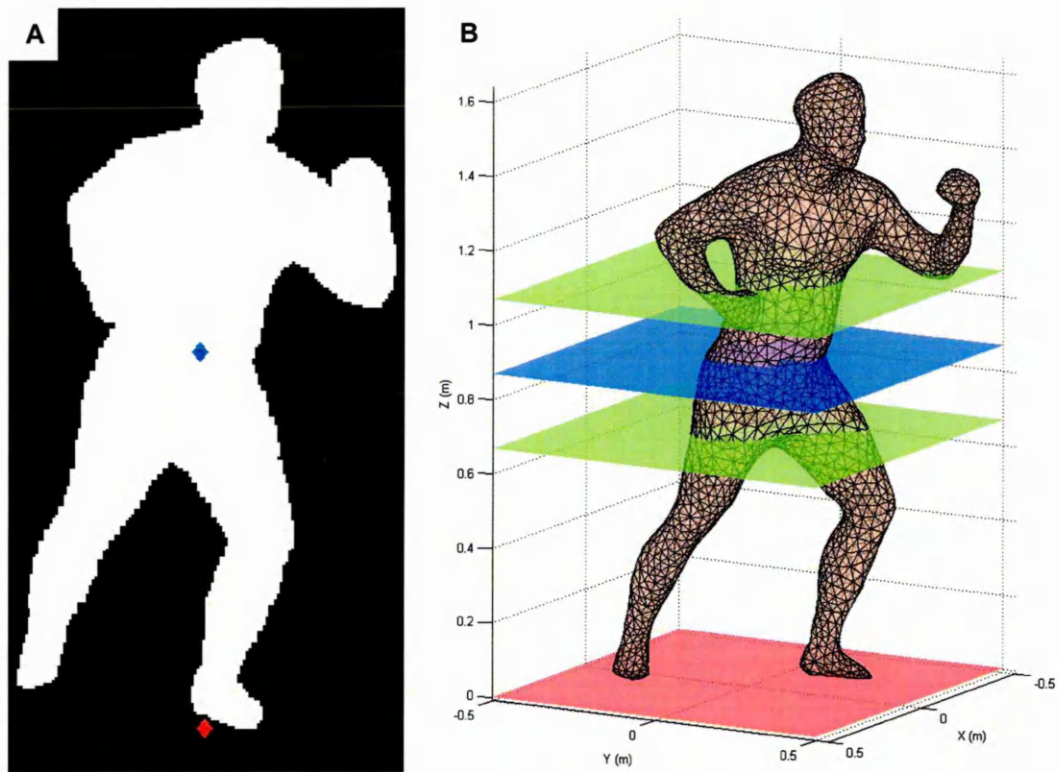


Figure 5.5. A: Sample image features used for ground and elevated plane position reconstruction (red and blue diamonds respectively). B: Representation of ground (red), elevated (blue) and out-of-plane, e.g. ± 200 mm (green), reconstruction planes.

To provide a comparative method for player position reconstruction, the player position definition described by Mauthner *et al.* (2008) was used. For each simulated camera image, the mean horizontal and maximum vertical silhouette coordinate was identified and reconstructed using the ground level calibration plane, e.g. red diamond (A: Figure 5.5) and red reconstruction plane (B: Figure 5.5). To reconstruct player position estimates for both ground and elevated calibration planes, extrinsic camera parameters for ground level and elevated calibration planes were required for each simulated camera perspective. Therefore uv coordinates for ground level and elevated calibration planes were calculated and corresponding extrinsic camera parameters computed for each camera perspective. This enabled player position reconstruction as described in Chapter 4.

Out-of-plane player motion, e.g. vertical COM oscillation in running, will affect player position reconstruction accuracy. To estimate the impact of out-of-plane motion to horizontal plane reconstruction, additional 'out-of-plane' calibration planes were

calculated, e.g. green reconstruction planes (B: Figure 5.5). For each simulated camera perspective, out-of-plane calibration plane coordinates were reprojected into the image at ± 200 mm relative to the elevated calibration plane, i.e. 50% stature. Subsequently, extrinsic parameters for out-of-plane calibration planes were also calculated and used to reconstruct player COM position as described above. In effect, this moves the calibration plane vertically in relation to the image coordinate being reconstructed. This is the equivalent of the image coordinate, or player COM, moving in relation to the calibration plane. Vertical oscillation magnitude has been reported as 84.4 ± 10.0 mm for running (Dallam *et al.*, 2005). Therefore the simulated out-of-plane-error, i.e. ± 200 mm, was considered a suitable magnitude.

Resultant direction differences between the ground plane projection of point cloud COM and reconstructed player position estimates were calculated. Reconstruction differences in X and Y directions were not reported due to changing camera azimuth angle. RMSE in the resultant direction was calculated with the following:

$$RMSE_R = \sqrt{\sum_{i=1}^N (X_{iR} - X_{ir})^2 / N} \quad [5.2]$$

where X_{iR} is the point cloud coordinate, X_{ir} is the reconstructed coordinate and N is the number of coordinates used.

5.5. Results

Data were derived from three-dimensional point cloud data of a single male participant in standing and running postures. Filming conditions experienced at two competitive sport events, i.e. 2011 Roland Garros Qualifiers and 2011 ATP World Tour Finals (described in Chapter 4), formed the basis for simulated camera perspectives. Camera perspectives of point cloud data were systematically generated and image features for elevated and ground plane position reconstruction, described in section 5.4.1, were identified. Further, horizontal reconstruction error as a result of out-of-plane motion was estimated within the limits of ± 200 mm. All position estimates were reconstructed using the planar reconstruction method described in Chapter 4.

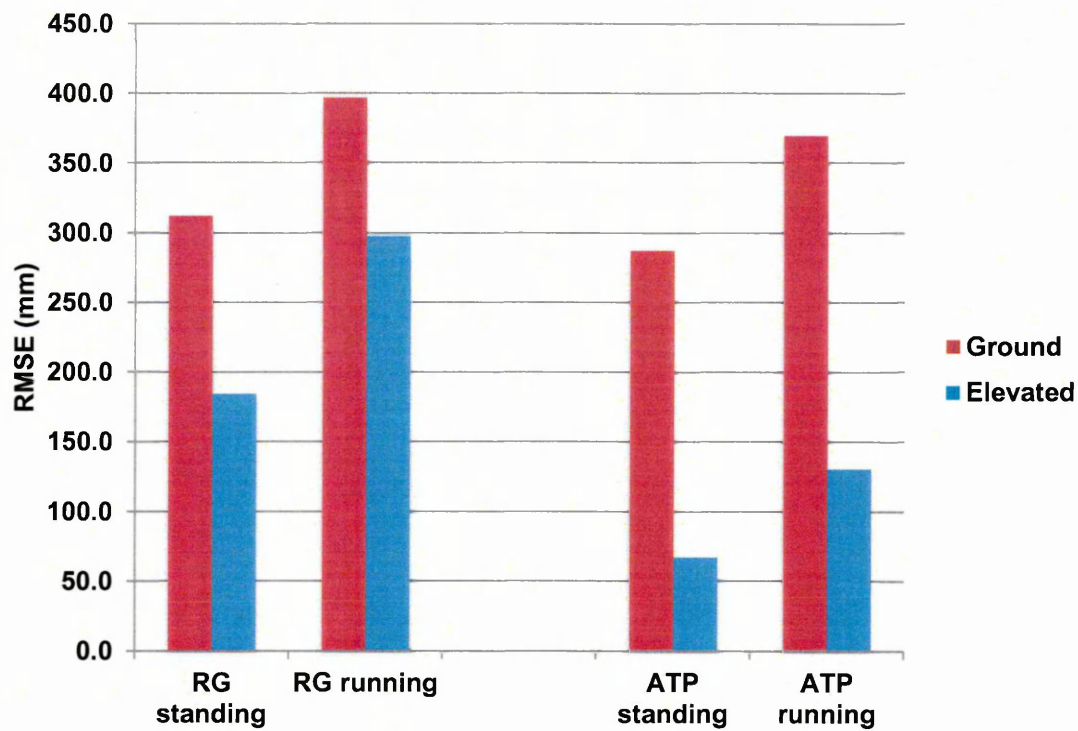


Figure 5.6. Resultant direction RMSE (mm) for player COM position using Ground and Elevated plane definitions.

Figure 5.6 presents player position $RMSE_R$ for all simulated camera perspectives ($n = 657$) of standing and running posture data for RG and ATP. Figure 5.6 indicates that $RMSE_R$ was lower for ATP position reconstruction when compared to corresponding RG simulations. Further, it is apparent that $RMSE_R$ for running posture data was greater for both position reconstruction methods for both RG and ATP simulations. Figure 5.6 demonstrates that elevated plane $RMSE_R$ was lower than ground plane $RMSE_R$ for corresponding camera perspectives of standing and running posture data. When compared to ground plane reconstruction, elevated calibration plane $RMSE_R$ was 128.1 and 99.7 mm lower for standing and running postures respectively for RG simulations and 219.7 and 238.8 mm lower for standing and running postures respectively for ATP simulations.

Table 5.1 presents mean and standard deviation differences (resultant direction) for all simulated camera perspectives ($n = 657$). Table 5.1 indicates that mean reconstruction error magnitudes for elevated plane position reconstruction were lower than ground plane position reconstruction. However Table 5.1 also indicates high standard

deviations for RG elevated plane position reconstruction when compared to ground plane position reconstruction.

Table 5.1. Mean \pm standard deviation (mm) resultant direction (R) player reconstruction error for all camera perspectives (n = 657).

		Standing posture		Running posture	
		Ground	Elevated	Ground	Elevated
RG	R	292.9 \pm 108.9	95.8 \pm 157.7	359.8 \pm 168.6	154.8 \pm 254.3
ATP	R	279.2 \pm 67.3	49.7 \pm 45.6	341.5 \pm 141.6	101.5 \pm 82.6

Figure 5.7 presents standing and running posture data (A and B respectively) position reconstruction differences for RG simulated camera perspectives. Black vertical lines (solid, dash-dot and dots) illustrate camera elevation angles for RG, ATP and maximum recommended stadia elevation angles respectively (John, Sheard and Vickery, 2007). Data represent mean and standard deviation position reconstruction differences for multiple 0 - 360° azimuth angles (n = 73) at each camera elevation angle (n = 9). For ground plane position reconstruction (red: Figure 5.7), standing and running data reconstruction error reduced as camera elevation increased.

For elevated plane position reconstruction (blue: Figure 5.7), comparatively large mean and standard deviation reconstruction errors were evident for low camera elevation angles, i.e. lower than 15°. However reconstruction error for standing and running posture data reduced as camera elevation increased. For standing posture data, mean elevated plane reconstruction error (blue) was 116.4 mm lower for 10° camera elevation and more than 200 mm lower for camera elevation angles greater than 10°, when compared to ground plane reconstruction (red). For running posture data, mean elevated plane reconstruction error (blue) was 63.3 mm greater for 10° camera elevation and more than 200 mm lower for camera elevation angles greater than 10°, when compared to ground plane position reconstruction (red).

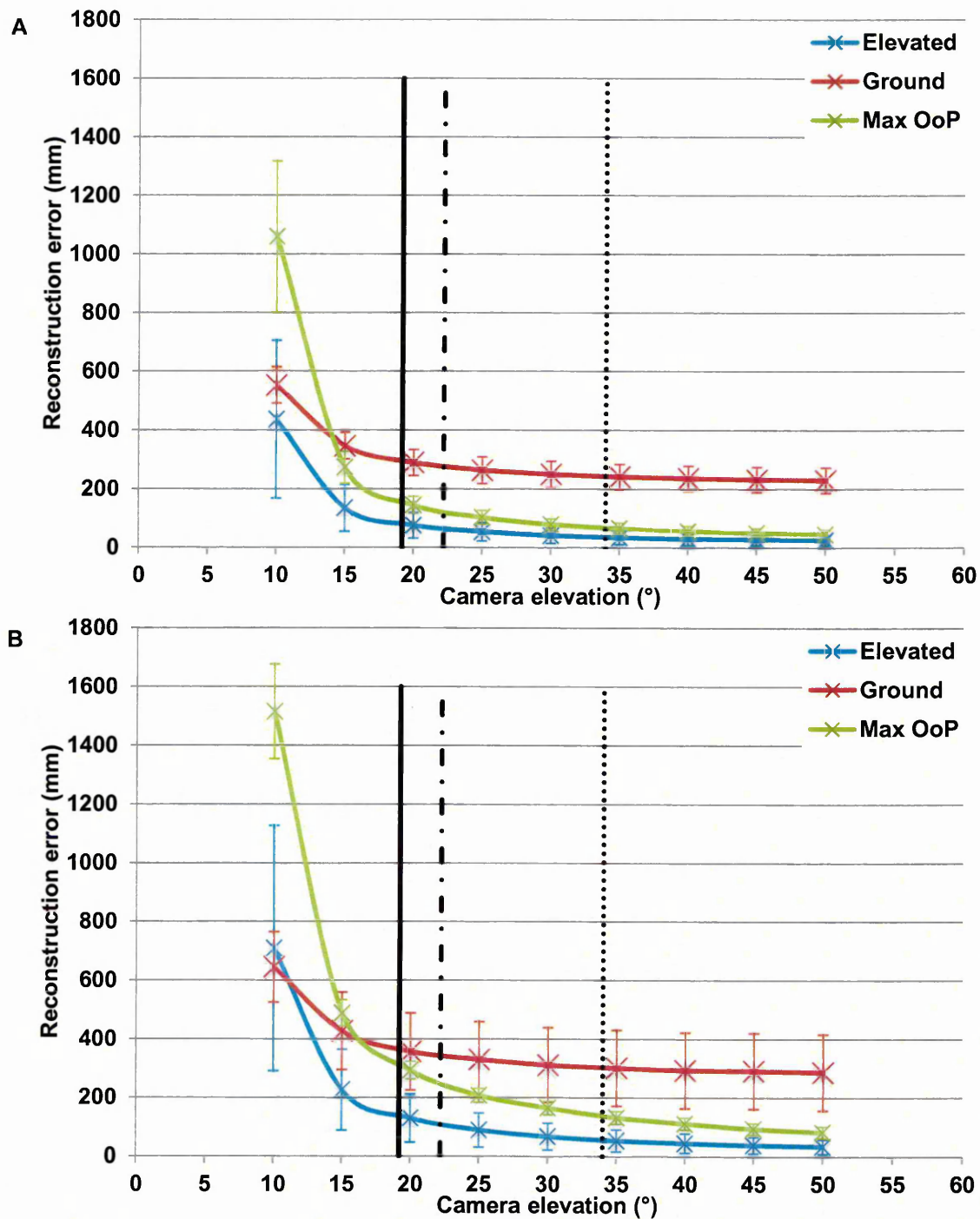


Figure 5.7. Mean (cross) and standard deviation (error bars) player position reconstruction error for RG using standing (A) and running (B) posture data. Black vertical lines illustrate RG, ATP and maximum stadia elevation angles (solid, dash-dot and dots respectively).

The impact of vertical, out-of-plane player motion (± 200 mm) on horizontal plane position reconstruction was assessed. Out-of-plane simulations (green: Figure 5.7) indicate that, for standing posture data, mean out-of-plane reconstruction error was

506.9 mm greater for 10° camera elevation and more than 72 mm lower for camera elevation angles greater than 10° when compared to ground plane position reconstruction (red). For running posture data, out-of-plane simulations (green) indicate that mean out-of-plane reconstruction error was 870.3 and 60.9 mm greater than ground plane reconstruction for 10 and 15° camera elevation respectively. However for camera elevation angles greater than 15°, out-of-plane reconstruction errors were more than 64 mm lower than ground plane reconstruction error.

Figure 5.8 presents standing and running posture data (A and B respectively) position reconstruction differences for ATP simulated camera perspectives. For ground plane position reconstruction (red: Figure 5.8), reconstruction error again reduced as camera elevation increased. In contrast to RG simulations, elevated plane position reconstruction error for ATP simulations (blue: Figure 5.8) were not dramatically greater for low camera elevation angles, i.e. lower than 15°. This reflects lower standard deviation magnitudes presented in Table 5.1. Further, reconstruction error for standing and running posture data reduced as camera elevation increased.

For ATP standing posture data (A: Figure 5.8), mean elevated plane reconstruction errors were more than 200 mm lower than corresponding ground plane reconstruction, regardless of camera elevation angle. For simulated running posture data (B: Figure 5.8), mean elevated plane reconstruction errors were more than 221 mm lower than corresponding ground plane reconstruction, regardless of camera elevation angle. Out-of-plane simulations (green; Figure 5.8) indicate that, for standing posture data, mean out-of-plane reconstruction errors were more than 158 mm lower than the ground plane reconstruction, regardless of camera elevation angle. For running posture data, out-of-plane simulations indicate that mean out-of-plane reconstruction error was 7.3 mm greater than ground plane reconstruction for 10° camera elevation. However for camera elevations greater than 10°, out-of-plane reconstruction errors were more than 80 mm lower than ground plane reconstruction error.

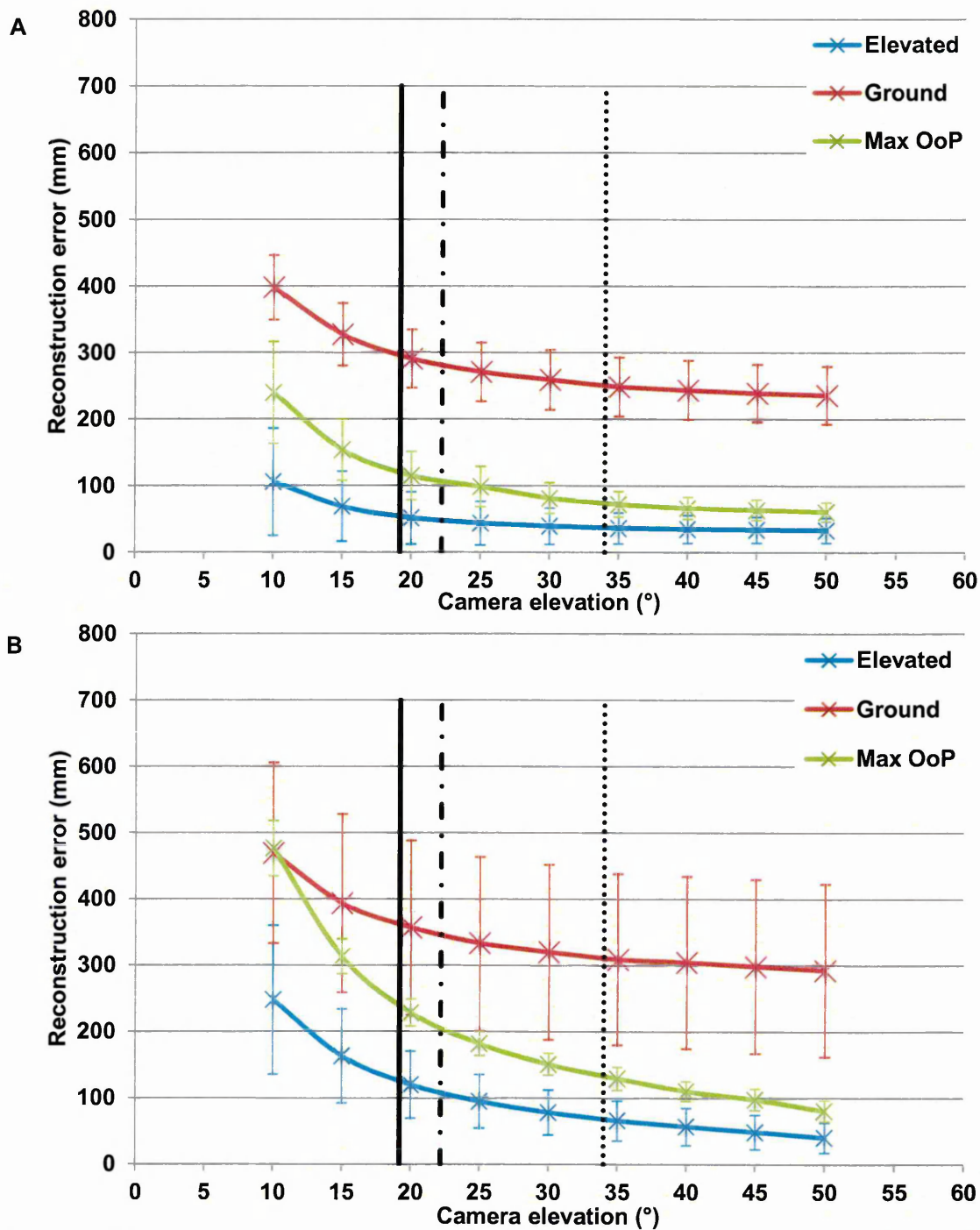


Figure 5.8. Mean (cross) and standard deviation (error bars) player position reconstruction error for ATP using standing (A) and running (B) posture data. Black vertical lines illustrate RG, ATP and maximum stadia elevation angles (solid, dash-dot and dots respectively).

Figure 5.9 presents maximum player position reconstruction error observed for all simulated camera azimuth angles (0 - 360°) for each simulated camera elevation angle. Standing and running posture data are presented as solid and dash-dot lines respectively

for RG (A) and ATP (B) simulations. Maximum reconstruction errors, as a result of changing player orientation, were lower for elevated plane reconstruction for RG camera elevation angles equal to or greater than 15° (A: Figure 5.9). Further, maximum reconstruction errors were lower for elevated plane reconstruction for all ATP simulated camera elevation angles (B: Figure 5.9).

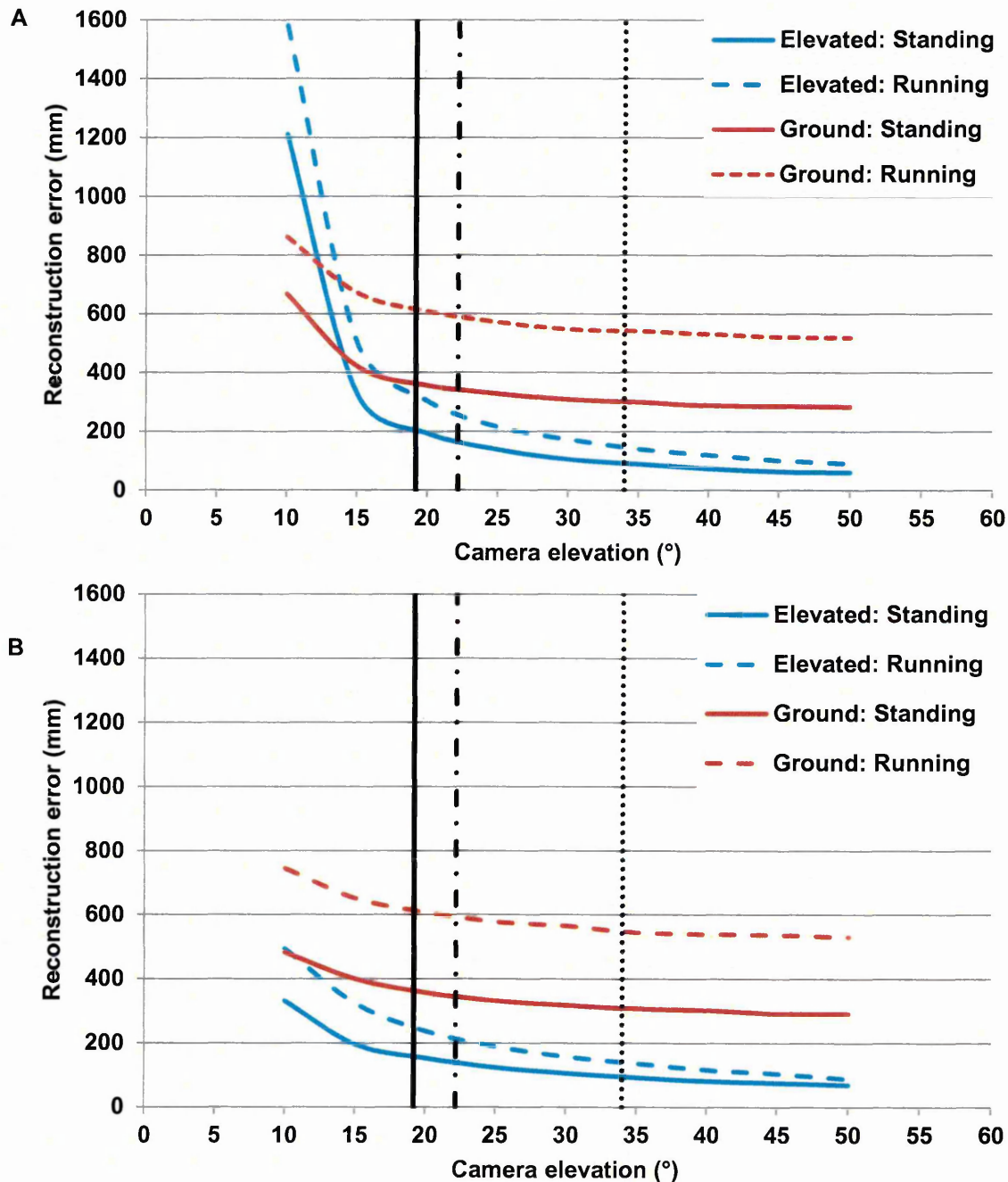


Figure 5.9. Maximum position reconstruction error for all azimuth angles at each elevation angle for RG (A) and ATP (B). Black vertical lines illustrate RG, ATP and maximum stadia elevation angles (solid, dash-dot and dots respectively).

5.6. Discussion

The accuracy of player position reconstruction was assessed using standing and running posture point cloud data, reprojected into simulated camera perspectives of two competitive tennis events. This provided a mechanism for assessing two player position reconstruction methods: a ground plane approach (Mauthner *et al.*, 2008) and an elevated plane approach. Further, due to the impact of vertical, out-of-plane motion on two-dimensional player position, position reconstruction error for ± 200 mm out-of-plane motion was simulated.

For simulated RG and ATP camera perspectives, lower RMSE_R for elevated plane reconstruction of both standing and running posture data (Figure 5.6) indicate that reconstructing ground plane features to infer player position can be limited. Using ground plane features to automatically identify player position, Mauthner *et al.* (2008) reported automatic and manually annotated (ground truth) player position errors of 200 – 400 mm during competitive volleyball. For current data, RMSE_R for ground plane player position reconstruction (described by Mauthner *et al.*, 2008) ranged between 287.1 – 397.3 mm (Figure 5.6). Data presented by Mauthner *et al.* (2008) were derived from a fixed camera location at competitive volleyball matches. Mauthner *et al.* (2008) did not quantify camera elevation; however changing player orientation relative to the camera would have affected position reconstruction accuracy, i.e. Figures 5.7 - 5.9. The current study does not replicate work presented by Mauthner *et al.* (2008). However, comparable position reconstruction errors obtained using the method presented by Mauthner *et al.* (2008) demonstrates the efficacy of the current approach for assessing player position reconstruction.

Table 5.1 indicated that for simulated RG camera perspectives, elevated plane reconstruction error was more variable, i.e. larger standard deviation, than ground plane position reconstruction. Figure 5.7 demonstrated that camera elevation angle affected position reconstruction accuracy for elevated plane reconstruction to a larger extent than ground plane reconstruction. However this reflects a greater reduction in elevated plane position reconstruction error for camera elevation angles $10 \rightarrow 15^\circ$ (Figure 5.7). For all player position reconstruction data presented in Figures 5.7 and 5.8, the lowest camera elevation angles yielded the lowest position reconstruction accuracy. Furthermore, position reconstruction accuracy improved as camera elevation angle increased. This is

consistent with previous observations (Brewin and Kerwin, 2003; Hinrichs *et al.*, 2005). The optimal elevation angle for position reconstruction on a plane is 90° , i.e. perpendicular camera. Position reconstruction accuracy for non-perpendicular camera views has been evaluated for camera elevation angles greater than 30° (Brewin and Kerwin, 2003; McLean *et al.*, 2004). However, camera perspectives at competitive sport events are typically constrained by sport stadia. John, Sheard and Vickery (2007) suggested typical stadia viewing angles were $14 - 26^\circ$, whilst 34° is the maximum recommended viewing angle, i.e. large stadia.

Current image perspectives were simulated to a minimum camera elevation angle of 10° , reflecting minimum sport stadia viewing angles, i.e. 14° (John, Sheard and Vickery, 2007). However camera elevation angles, i.e. Figures 5.7 - 5.9, were calculated in relation to the ground plane, i.e. sport stadia. Using an elevated plane to reconstruct position will effectively reduce camera elevation and thus camera elevation angle. For elevated plane reconstruction, camera elevation angles were the equivalent of 1.9 and 0.8° lower for RG and ATP simulations respectively, than elevation angles presented in Figure 5.7 - 5.9. However, camera elevation angle relative to the ground plane was retained as a convention to enable the comparison of methods in relation to sport stadia. Marginally lower camera elevation angles for elevated plane reconstruction might have exacerbated position reconstruction error at low elevation angles, particularly for RG simulations, i.e. Figure 5.7. However, it is apparent that for both position reconstruction methods, reconstruction accuracy begins to rapidly diminish for camera elevation angles lower than 15° .

Player position reconstruction accuracy will diminish as camera elevation angle reduces to zero. At this point the camera and reconstruction plane will be coincident: a camera ray (projected to infinity) will not intersect the reconstruction plane. For low camera elevation angles, inaccuracies in player COM identification will exacerbate position reconstruction error (Brewin and Kerwin, 2003). This was demonstrated by simulated out-of-plane errors (green: Figures 5.7 and 5.8). The use of an elevated plane to reconstruct position gives rise to reconstruction error resulting from vertical out-of-plane motion (Holden-Douilly *et al.*, 2011). For RG, out-of-plane simulations (green: Figure 5.7), standing and running posture data simulated using 10° camera elevation angles yielded large position reconstruction errors. However, for out-of-plane standing

posture data simulated with 15° camera elevation, elevated plane reconstruction yielded mean position reconstruction errors 72.6 mm lower than ground plane reconstruction (Figure 5.7). For out-of-plane running posture data simulated with 15° camera elevation, elevated plane reconstruction yielded mean position reconstruction errors 60 mm greater than ground plane reconstruction, whilst camera elevation angles greater than 15° yielded out-of-plane reconstruction errors that were lower than that of ground plane reconstruction (Figure 5.7).

For ATP out-of-plane simulations (green: Figure 5.8), with the exception of running posture data simulated at 10° camera elevation, elevated plane reconstruction yielded mean position reconstruction errors that were lower than ground plane position reconstruction for both standing and running posture data. Dallam *et al.* (2005) reported that during running, vertical oscillation of a neck marker (proxy for COM motion) was 84.4 ± 10.0 mm. Tennis is well known to consist of dynamic movements (Fernandez, Mendez-Villanueva and Pluim, 2006). However, to the author's knowledge, a thorough analysis of vertical player COM movement in competitive tennis has not yet been performed. Thus ± 200 mm of simulated out-of-plane error was considered to be suitable margin. As such, current simulations of standing and running posture data support player position reconstruction using an elevated plane for camera elevations equal to or greater than 15°. Further, the approach is supported when out-of-plane errors are incorporated into player position estimates.

Changing player orientation relative to the camera affected player position reconstruction accuracy, i.e. Figures 5.7 - 5.9. Figure 5.9 presents maximum player position reconstruction error observed for all simulated camera azimuth angles (0 - 360°) at each camera elevation angle. Figure 5.9 demonstrates that, regardless of elevation angle, both player position reconstruction methods were sensitive to changes in player orientation relative to the camera. However, with the exception of 10° camera elevation for RG, maximum position reconstruction error for different player orientations was lower for elevated plane reconstruction for both standing and running posture data. This reflects the nature of image features being reconstructed. Ground plane reconstruction, i.e. central location about the feet, does not adequately reflect the ground plane projection of player COM for different player orientations (Figure 5.9).

Further the application of ground plane reconstruction to dynamic movement images, i.e. running posture data, exacerbates position reconstruction error.

For elevated plane reconstruction at RG, 10° camera elevation demonstrates a critical limit to player reconstruction accuracy. Factors such as camera calibration, i.e. high lens distortion images (Chapter 4), might contribute to higher reconstruction errors. However, inaccuracies in player identification will yield large errors for position reconstruction at low camera elevation angles. As noted, position reconstruction accuracy for non-perpendicular camera views has only been evaluated for camera elevation angles greater than 30° (Brewin and Kerwin, 2003; McLean *et al.*, 2004). Further, current position reconstruction estimates were not obtained from systematically identifiable control points: camera perspective will have affected image coordinates identified as player COM. Therefore it would be inappropriate to support player position reconstruction using an elevated plane for camera elevations lower than 15°. For the context of current work, minimum sport stadia viewing angles have been reported as 14° (John, Sheard and Vickery, 2007). It is unlikely, in the case of sport stadia, that greater elevation angles would not be available. For current work, RG represented a small sport stadium, inducing high lens distortions due to a restricted field-of-view. However, it was possible to obtain a camera elevation angle of 19.2°. Therefore, based on current data, reconstructing player position using an elevated reconstruction plane, where elevation is 50% of player stature, is appropriate for footage obtained in sport stadia.

5.7. Conclusion

Photogrammetric estimates for player position during match-play tennis should not be unduly affected by player orientation or activity. The accuracy of two player position reconstruction methods was assessed for simulated camera perspectives of two competitive sport events. Simulated data suggest that reconstructing ground plane features to infer player position can be limited. However, for camera elevation angles equal to or greater than 15°, simulated data support the reconstruction of player position using an elevated plane, even with out-of-plane motion incorporated into player position estimates, e.g. Figure 5.9.

For match-play tennis, accurate reconstruction of tennis player position is important to accurately quantify player movement and velocity. The current chapter developed and assessed a novel method for reconstructing player position. The current method was demonstrated to yield lower reconstruction errors than an existing method for player position reconstruction when camera elevation angle was equal to or greater than 15° . In relation to the overall project aims, this chapter has demonstrated that player position can be reconstructed with greater accuracy when using an elevated reconstruction plane, e.g. blue reconstruction plane (B: Figure 5.5). Further, the camera-plane model developed in Chapter 4 demonstrates that the current approach can be applied to match-play tennis footage without the necessity to gain access to the tennis court or interfere with tennis play.

6.1. *Introduction*

Chapter 3 described a manual approach for identifying foot-surface contacts to identify gait parameters. However, manual identification is user intensive and requires the subjective identification of both the position and time of foot-surface contacts. Furthermore, a stationary camera was used for analysis: changing player position and orientation in relation to the camera would result in player images of varying size and perspective. This will exacerbate random error associated with manual digitising due to the subjective nature of analysis (Glazier and Irwin, 2001). An automated method to identify foot-surface contacts without markers and without user input was required. The literature review (section 2.3) highlighted that research reporting foot-surface contact identification with a single camera is limited. Furthermore, such research has only been applied to walking; to the author's knowledge, no research has identified foot-surface contacts in running or other activities. Empirical evidence is required to identify an algorithm to measure foot-surface contact position and time using single camera footage. A key function of an algorithm would be to identify foot-surface contacts independently of gait, e.g. walking and running, and camera view, e.g. image perspective and size. This will address the fundamental constraints of foot-surface contact identification in match-play tennis, i.e. multi-modal and multi-directional gait.

6.2. *Aim and objectives*

The aim of this chapter is to develop an algorithm to automatically identify foot-surface contacts using single camera video of walking and running, for player images that vary in size and view. This relates to box F of the development stage diagram (Figure 1.2).

Objectives:

1. Identify a suitable vision-based approach for identifying foot-surface contacts.
2. Develop an algorithm to automatically identify foot-surface contacts using single camera sequences of player images that vary in size and perspective.
3. Apply the algorithm to single camera image sequences of different gaits.

The algorithm was developed in MATLAB using the Image Processing toolbox.

6.3. Vision-based foot-surface contact identification

6.3.1. Single camera methods

Bouchrika and Nixon (2006) presented an automatic method for identifying heel-strikes and reported detection accuracy as 0.52% of participant height. However the algorithm removed temporal information of heel strikes and accuracy estimates were not converted into real-world measurements. Jung and Nixon (2013) extended this work by creating an accumulation map of every silhouette pixel throughout a walking image sequence. Key heel strike frames were identified using the sinusoidal peaks of head trajectory. Jung and Nixon (2013) reported detection rates of 95.6% with position accuracy of ± 100 mm. However, although the method was applied to different camera perspectives, the method relied on walking gait to define heel-strikes. Furthermore, the accumulation of image sequence data is an exhaustive approach.

Periodic based gait measurements are not appropriate for different modes of gait, i.e. walking and running. For example, the inversion of gait mode during running, i.e. single stance and dual float, would invalidate the assumptions of algorithms defined by the periodicity of walking gait. A 'gait mode' definition might address the periodicity issue, however it would likely be ambiguous during walk-run transitions and thus not appropriate for the multi-modal nature of tennis movement (Robinson and O'Donoghue, 2008).

To detect motion when tracking people, the use of inter-frame differencing is a common practice when a suitable background image isn't available (Martinez-Martin and Pobil, 2012). However previous work has likened resulting inter-frame motion to error: Figure 6.1 illustrates the 'double image' effect (Zhang, Zhou and Zhu, 2010). Figure 6.1 was obtained from an image sequence where the rear and front feet were in stance and swing respectively. The stance foot is evident by fewer active pixels about the rear foot, i.e. low inter-frame difference. In contrast, more active pixels about the front foot reflect motion, i.e. high inter-frame difference. Therefore inter-frame differencing can provide a mechanism for identifying motion within specific regions. The following sections describe an image processing approach to quantify inter-frame differences within specific regions about the feet.



Figure 6.1. Inter-frame difference image of walking results in a 'double image' (adapted from Zhang, Zhou and Zhu, 2010).

6.3.2. Logical image processing

There are three models for computing motion within an image sequence: Observation, Motion Field and Region Models (Reed, 2004). This work adopts a Region Model approach: it was necessary to quantify motion within specific regions that can change size and move independently within an image sequence. Reed (2004) highlights that whilst such approaches allow meaningful descriptions of motion, independent object segmentation can increase algorithm complexity. As such, basic logical image processing operations, used in subsequent sections, are presented to aid algorithm clarity.

Figure 6.2 illustrates the extraction of basic image properties using MATLAB. Figure 6.2 is a 300×300 pixel logical or binary image; pixel values are either 0 or 1, i.e. black or white respectively. A filled circle with a radius of 50 pixels was created at the image centre. The information presented in the top left are basic object properties obtained directly from the image, i.e. area, major axis length, major axis orientation and centroid (COM) location. Figure 6.2 also graphically illustrates the objects' local coordinate system. The local coordinate system is defined by the objects' bounding box, i.e. the smallest rectangle that can contain the object.

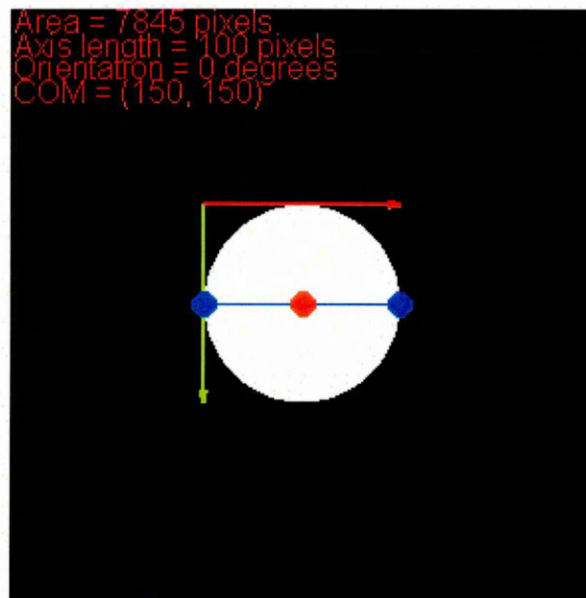


Figure 6.2. Filled circle with basic image properties highlighted.

A local coordinate system is used when performing morphological operations to specific objects; object coordinates can be transformed back into the original image coordinate system. Figure 6.3 illustrates basic morphological operations used in subsequent sections. Image A (Figure 6.3) presents a logical image of a tennis racket and ball. Image B (Figure 6.3) illustrates the removal of objects with less than 100 connected pixels: the ball (comprising of 81 pixels) is therefore removed. This provides a simple step for removing small objects, i.e. image noise. Images C and D (Figure 6.3) illustrate image dilation and erosion respectively using a 10×10 (square) morphological structuring element. Dilation adds pixels to the logical image periphery, i.e. image C (Figure 6.3) whilst erosion removes pixels, i.e. image D (Figure 6.3). Figure 6.4 illustrates the effect of image dilation and erosion at the pixel level.

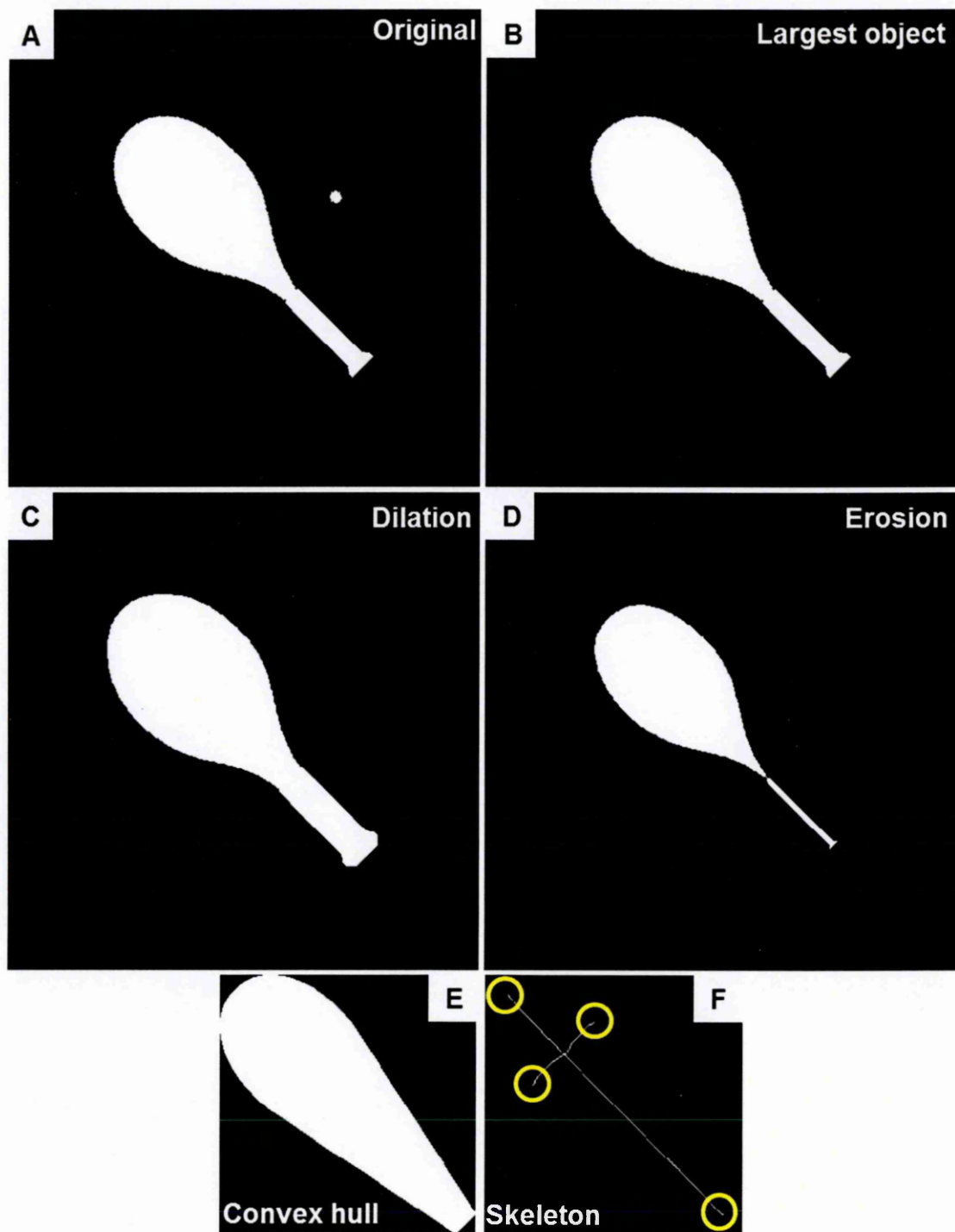


Figure 6.3. Logical image of tennis racket and ball (A), removal of small objects (B), dilation (C), erosion (D), convex hull (E), skeletonised image (F) and skeletonised image endpoints (yellow rings).

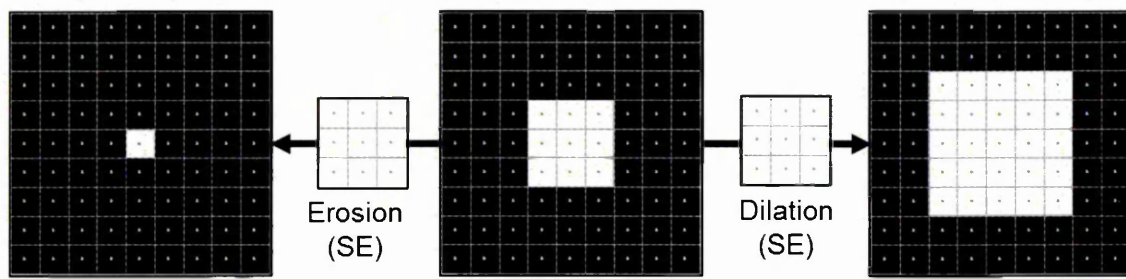


Figure 6.4. Erosion (left) and dilation (right) of a 3×3 logical image (centre), using a 3×3 structuring element.

Image E (Figure 6.3) illustrates the convex hull of image B. The convex hull is defined as the filled, smallest convex polygon that contains foreground pixels. It should be noted that the convex hull and derivative images, i.e. skeletonised image and skeletonised image endpoints (image F: Figure 6.3), are computed within the smallest containing image, i.e. bounding box. Accordingly, the bounding box defines a local coordinate system for which convex hull coordinates can be transformed back into the image coordinate system. Image F (Figure 6.3) illustrates the skeleton of image E. Skeletonisation removes boundary pixels (infinite number of iterations) without allowing the object to break apart. Finally, endpoints of the skeletonised image are identified (highlighted by yellow rings). Endpoints are defined as pixels that have a single 8-connected neighbourhood connection. All of the basic, logical image processing operations presented above were required to develop the foot-surface contact identification algorithm.

6.3.3. Synthetic walking data

Goffredo *et al.* (2008) used synthetic data to test algorithm performance for images corrupted with Gaussian noise. Due to the lack of single camera research describing foot-surface contact identification, this work used synthetic walking data to develop an initial algorithm. The development of an algorithm to be applied to standard colour images of different perspective and gait will be addressed in subsequent sections.

Visual 3D (v3.79, C-Motion, MD, USA) is a powerful three-dimensional motion analysis program. Sample marker position data from C-Motion's freely available lower-body, walking gait tutorial (C-Motion, 2012) were used to create synthetic images of walking, i.e. Figure 6.5. Visual 3D renders conical skins to marker position data to visualise a basic walking silhouette; no conventional images were used.

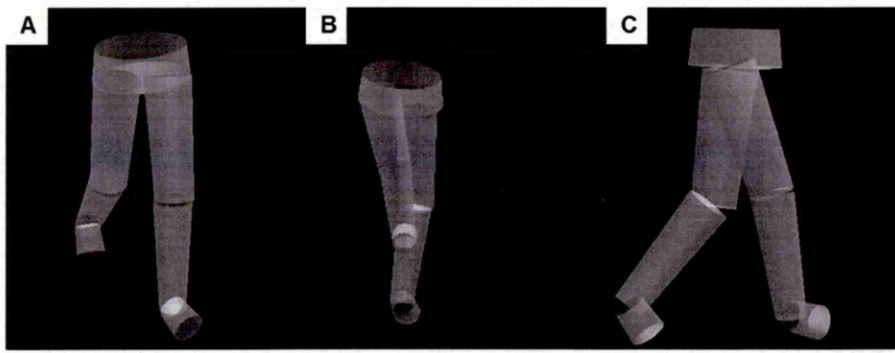


Figure 6.5. Sample images rendered in Visual 3D of walking from oblique frontal (A), oblique rear (B) and sagittal (C) perspectives.

The literature review highlighted that the use of a static or unchanging background model was suitable for current work, due to low background segmentation error rates and low computational demand (Hassanpour, Sedighi and Manashty, 2011). Accordingly, image silhouettes were obtained by background-image differencing (equation 6.1), i.e. absolute difference image between current and reference images (hereinafter referred to as ADI).

$$ADI = |Frame_n - Frame_{Ref}| \quad [6.1]$$

where $Frame$ is a video frame and subscript suffixes n and Ref are current and reference (background) frames respectively.

ADI segmentation is given by equations 6.2 – 6.7 such that the binary image (hereinafter referred to as ADI_B) consists of foreground and background pixels, e.g. pixel values of 1 or 0 respectively. Full-colour images have multiple image planes, e.g. each pixel within an RGB image consists of red, green and blue values.

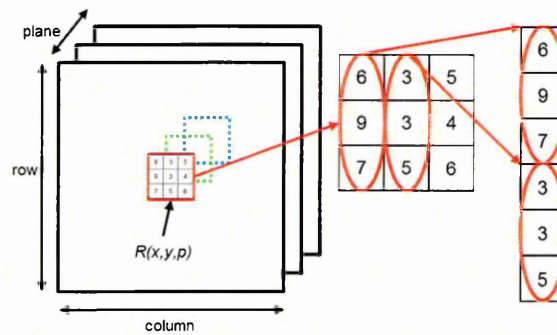


Figure 6.6. Column-wise concatenation of the image plane region $R(x,y,p)$.

Spatial image processing simplifies the segmentation of colour images: each image plane can be processed individually (Gonzalez and Woods, 2002). Image segmentation is therefore performed iteratively for N image planes. The two-dimensional image plane region $R(x,y,p)$ is comprised of i rows and j columns, e.g. Figure 6.6, and can be concatenated (column-wise) to form the one-dimensional vector R for N image planes:

for $p = 1, \dots, N$

$$R_p = \begin{bmatrix} \vec{v}_{i1} \\ \vdots \\ \vec{v}_{ij} \end{bmatrix} \quad [6.2]$$

where \vec{v} is a column-wise vector of pixel intensities for the image plane region, R_p .

The sum of the average (mean) and standard deviation of pixel intensities yield the threshold $level_p$:

$$\bar{x}_p = \frac{1}{n} (\sum_{i=1}^n R_p) \quad [6.3]$$

$$s_p = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (R_p - \bar{x}_p)^2} \quad [6.4]$$

$$level_p = \bar{x}_p + s_p \quad [6.5]$$

where n is the number of elements in R_p .

For corresponding image planes, pixel intensities greater than the threshold $level_p$ determine foreground pixels. Image planes are then combined and converted to logical:

$$ADI_{1,\dots,p} = ADI_p > level_p \quad [6.6]$$

$$ADI_B = (ADI_1 + \dots + ADI_p) > 0 \quad [6.7]$$

Due to the nature of synthetic images, i.e. Figure 6.5, no further image or morphological processing steps were required, e.g. image noise removal, etc.

6.3.4. Foot-region segmentation

Foot-regions must be tracked and segmented in order to quantify inter-frame differences relating to foot-surface contacts. Based on anatomical studies, i.e. Dempster and Graughran (1965), Goffredo *et al.* (2008) defined the lower 10% of a silhouette as the foot and ankle region, i.e. y'_{ankle} (Figure 6.7). Goffredo *et al.* (2008) identified foot locations using lower-limb orientation. However it is clear that for oblique camera views, horizontal segmentation based on silhouette height is inappropriate for segmenting individual feet, i.e. unequal front and rear foot area (Figure 6.7). Therefore, due to the camera perspective, feet must be identified and segmented individually.

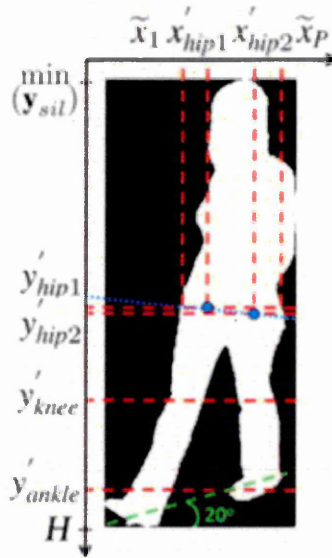


Figure 6.7. Silhouette segmentation based on silhouette height (H ; adapted from Goffredo *et al.*, 2008). Red and green dashed lines illustrate anatomical segmentation and walking direction respectively.

The major axis length of the binary image silhouette, i.e. ADI_B , was multiplied by two to approximate silhouette height, e.g. use of lower-body images (Figure 6.5). Subsequently, for each image within an image sequence, the following morphological processes were used to define individual foot-regions:

1. Image pixels above the silhouette COM row coordinate were set to false, i.e. A (Figure 6.8). The resulting binary image approximates the lower-limbs, i.e. knees, shins and feet (Goffredo *et al.*, 2008).

2. The convex hull, COM and bounding box of the lower-limb image, i.e. A (Figure 6.8), were computed: the bounding box defined a local coordinate system.
3. The lower-limb convex hull was skeletonised and skeleton endpoints computed. Skeleton endpoints superior of the convex hull COM were discarded. Skeleton image and skeleton endpoint coordinates were then transformed into the image coordinate system and are highlighted in relation to the lower-limbs in blue and red respectively (A: Figure 6.8).
4. Lower-limb perimeter pixels were computed, i.e. green outline in B and C (Figure 6.8). The nearest (Euclidean distance) perimeter pixel to remaining skeleton endpoint coordinates, hereafter referred to as SPP (skeleton-perimeter pixel), were identified. This was necessary as endpoint coordinates are a product of skeletonised convex images and not necessarily foreground ADI_B pixels. For images B and C (Figure 6.8), SPP coordinates are highlighted in blue.
5. Each SPP pixel was iteratively dilated using a 5×5 structuring element and multiplied by ADI_B to yield $MaskADI_B$, i.e. foot-region mask (B: Figure 6.8). The major axis length of $MaskADI_B$ was computed and dilation terminated if major axis length was equal to or greater than 10% of silhouette height. This was performed independently for each SPP.

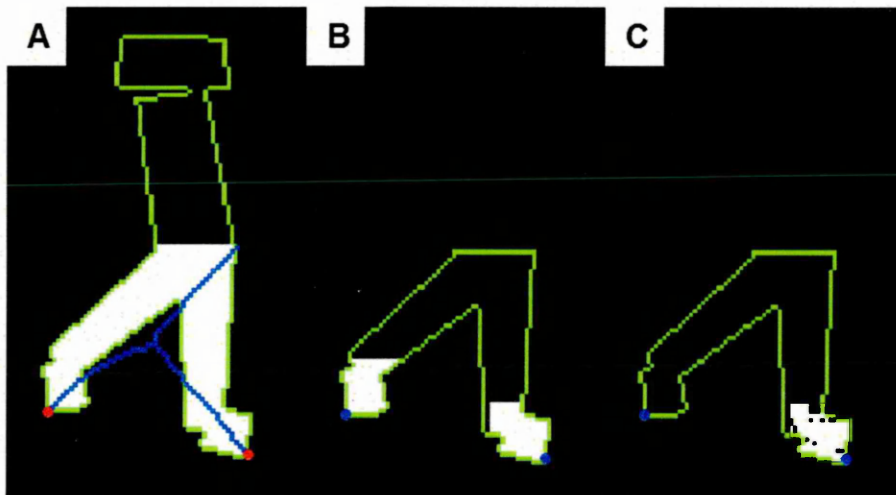


Figure 6.8. Binary images of lower-limbs (A), foot-region masks (B) and foot-region inter-frame differences (C). Connected perimeter and convex-hull skeleton pixels are highlighted in green and blue respectively. Skeleton endpoint and skeleton-perimeter pixels are highlighted in red and blue filled circles respectively.

6.3.5. Foot-region inter-frame motion

The independent definition of foot-region masks, i.e. $MaskADI_B$ (B: Figure 6.8), enables motion to be assessed in regions that can change size and move independently within an image sequence. Inter-frame differences illustrate motion. Inter-frame differencing was performed to yield a frame difference image (hereinafter referred to as FDI). FDI was computed with the following:

$$FDI = |Frame_n - Frame_{n-1}| \quad [6.8]$$

The binarisation of FDI to FDI_B follows equations 6.2 – 6.7. However, FDI_B is inverted: static objects thus become active or foreground pixels (equation 6.9).

$$FDI_B = 1 - FDI_B \quad [6.9]$$

$$MaskFDI_B = MaskADI_B \times FDI_B \quad [6.10]$$

Individual objects within $MaskADI_B$ were identified and labelled: equation 6.10 gives $MaskFDI_B$, i.e. foot-region inter-frame differences (C: Figure 6.8). Finally, the area and COM for corresponding objects within $MaskADI_B$ and $MaskFDI_B$ were computed. To determine whether foot-regions were static, a threshold was used, e.g. Figure 6.9. For initial work, a generic threshold was set to 50% of foot mask area:

$$th_{FSC} = \sum MaskADI_B \times 0.5 \quad [6.11]$$

$$FSC = \sum MaskFDI_B \geq th_{FSC} \quad [6.12]$$

For objects where FSC was true, the corresponding coordinates of $MaskADI_B$ COM and frame number were recorded as a foot-surface contact. The algorithm was applied iteratively to sagittal, oblique frontal and oblique rear perspective image sequences of synthetic walking data, i.e. Figure 6.5. Figure 6.9 illustrates collated area data for identified $MaskFDI_B$ objects during a sagittal walking image sequence. Large and small $MaskFDI_B$ areas, i.e. inter-frame differences, reflect stance and swing phases respectively during a left-to-right walking sequence. The sequence mean for th_{FSC} is plotted in black to illustrate the identification of foot-surface contacts.

Red and blue curves illustrated in Figure 6.9 represent $MaskFDI_B$ area for the first and second objects within a logical image respectively. The algorithm does not differentiate left and right feet. Individual objects in a logical image are labelled iteratively in relation to columns that objects occupy, i.e. left-to-right. As such, both red and blue curves consist of left and right foot data. However peaks are clearly paired. This is the result of the left-to-right walking sequence and logical labelling of foot-region masks. During stance, i.e. highlighted peaks (Figure 6.9), the stance foot is initially labelled as mask two (blue) because it is located to the right of the swing foot. As the swing foot passes the stance foot, the stance foot is located to the left of the swing foot and subsequently labelled as mask one (red).

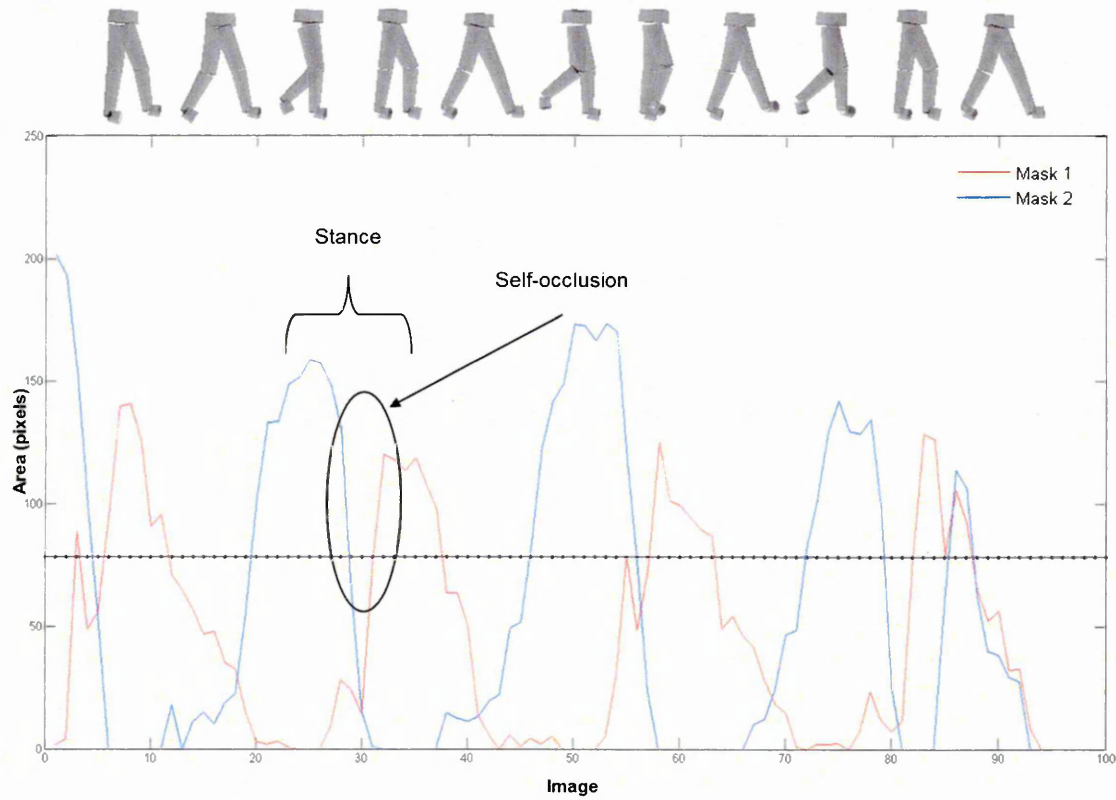


Figure 6.9. $MaskFDI_B$ areas for foot-region masks one and two (red and blue respectively) during a walking sequence. Areas exceeding mean th_{FSC} (black horizontal line) indicate foot-surface contacts.

The decrease (mask two: blue) and subsequent increase (mask one: red) in mask area during the highlighted stance period reflects self-occlusion by the swing foot. Assuming a fixed camera and straight line walking, self-occlusion will only occur for alternate foot-surface contacts. However, Figure 6.9 demonstrates a fall in area for each stance

phase. This was due to translucent conical skins exported using Visual 3D (Figure 6.5) and is only an issue for current, synthetic images.

Figure 6.10 illustrates foot-surface contact data (red crosses) collated throughout sagittal, oblique frontal and oblique rear walking image sequences. Furthermore, transformed skeleton perimeter pixels, e.g. blue filled circles (Figure 6.10), illustrate foot-region identification from different camera perspectives. SPP coordinates (blue filled circles: Figures 6.8 and 6.10) illustrate the flexibility of foot-region identification (steps 1 – 5) for different camera perspectives. Furthermore, Figures 6.9 and 6.10 illustrate that quantifying inter-frame differences within specific foot-regions is a suitable method for identifying foot-surface contacts. The current approach was able to identify (visual correspondence) spatial (Figure 6.10) and temporal (Figure 6.9) parameters of foot-surface contacts from three different camera perspectives. Furthermore, the algorithm imposed no *a priori* assumptions regarding gait mode: foot-region masks were identified and inter-frame motion quantified independently. Therefore gait mode, i.e. single stance (running) or dual stance (walking), is irrelevant to the current algorithm. Finally, the algorithm operated iteratively when applied to an image sequence: for any current image, only a reference and preceding image is required for analysis.

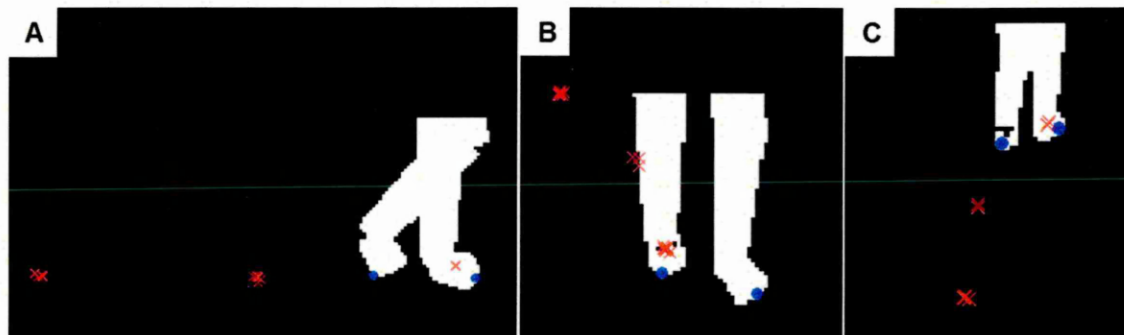


Figure 6.10. Collated foot-surface contact data (red crosses) for sagittal (A), oblique frontal (B) and oblique rear (C) camera views.

Figures 6.9 and 6.10 highlight four important of considerations to the current algorithm. First, it is clear that estimates of stance time will be underestimated due to the application of a foot-surface contact threshold. Obtaining an accurate estimation of stance time is a key issue for any automated, kinematic method of gait event detection

(Hreljac and Marshall, 2000; O'Connor *et al.*, 2007). Appropriate foot-surface contact thresholds will be addressed by subsequent sections. Second, whilst self-occlusion does not limit the current algorithm, self-occlusion (particularly for different camera perspectives) might reduce stance time estimates for occluded feet. The validity of stance time estimates must therefore be addressed. Third, the spatial location of a foot-surface contact is defined as the COM of a foot-region mask. This does not necessarily represent the ground plane projection of a foot-surface contact. The image feature identified by the algorithm will be addressed in subsequent sections and the validity of foot-surface contact position assessed in Chapter 7. Finally, synthetic image sequences represent 'perfect images'. Images were not contaminated, i.e. image noise, shadow, etc., and conical skins are not realistic shapes. Furthermore, the algorithm was not able to verify foot-surface contact data, i.e. spatial location relative to foot and body. Therefore the current approach must be tested with colour images of a real participant and be able to differentiate true and false positive foot-surface contacts.

6.4. Development of an algorithm to measure foot-surface contacts

Initial work in section 6.3 demonstrated that, for synthetic images, evaluating inter-frame differences about feet-regions could identify foot-surface contacts, i.e. Figures 6.9 and 6.10. Synthetic data were used as a 'workbench' to develop an initial algorithm. However, synthetic data do not provide realistic operating conditions, i.e. simplistic shapes, translucent images and no image noise, etc. Therefore the algorithm must be applied to real, colour image sequences. For clarity, Figure 6.11 provides an overview of image collection and analysis steps described in following sections.

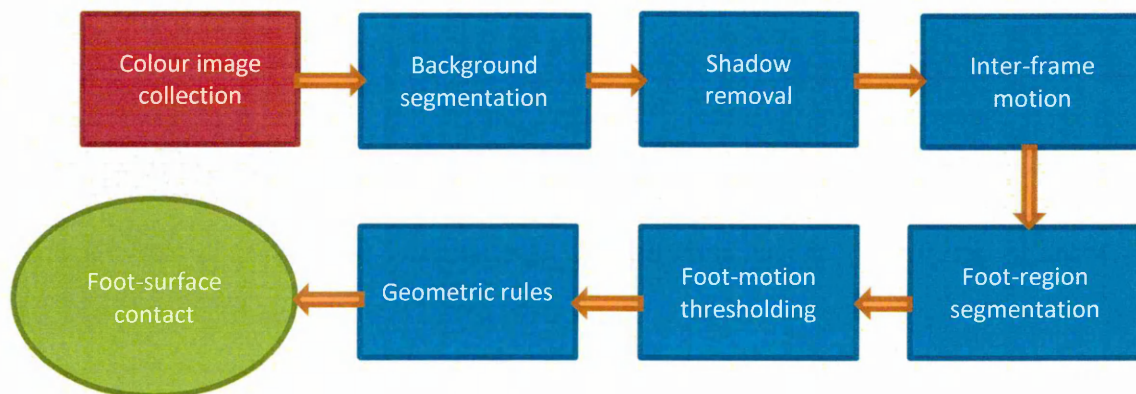


Figure 6.11. Image collection and analysis steps for foot-surface contact algorithm.

6.4.1. Colour image collection

A pilot study was performed. Six male participants (age = 26.8 ± 2.9 years; stature = 1.80 ± 0.08 m; mass = 76.2 ± 10.0 kg) were recruited. Participants were appropriately briefed to aid the completion of the proposed tasks and written informed consent was obtained. Approval for all procedures was obtained from the Research Ethics Committee of the Faculty of Health and Wellbeing, Sheffield Hallam University (Appendix 6).

Participants were asked to walk and run in barefoot and in shod (own trainers) at a self-selected pace, through a 4.0×1.5 m motion capture volume within a carpeted laboratory (Figure 6.12). Four networked cameras (AXIS M1104, AxisTM Communications, Sweden), streaming RGB colour images (1280×720 p) to a data collection computer at 25 Hz, recorded images from frontal, sagittal and two oblique frontal perspectives (Figure 6.12). Prior to motion trials, the participant was required to wait behind a blanking wall whilst images streamed to the data collection computer were buffered (Figure 6.12). This enabled image sample rates to stabilise and provided suitable background images for background modelling (Hassanpour, Sedighi and Manashty, 2011). For brevity, sample images of a single participant obtained from a sagittal perspective of barefoot running, i.e. NCam 4 (Figure 6.12), are used to illustrate algorithm development.

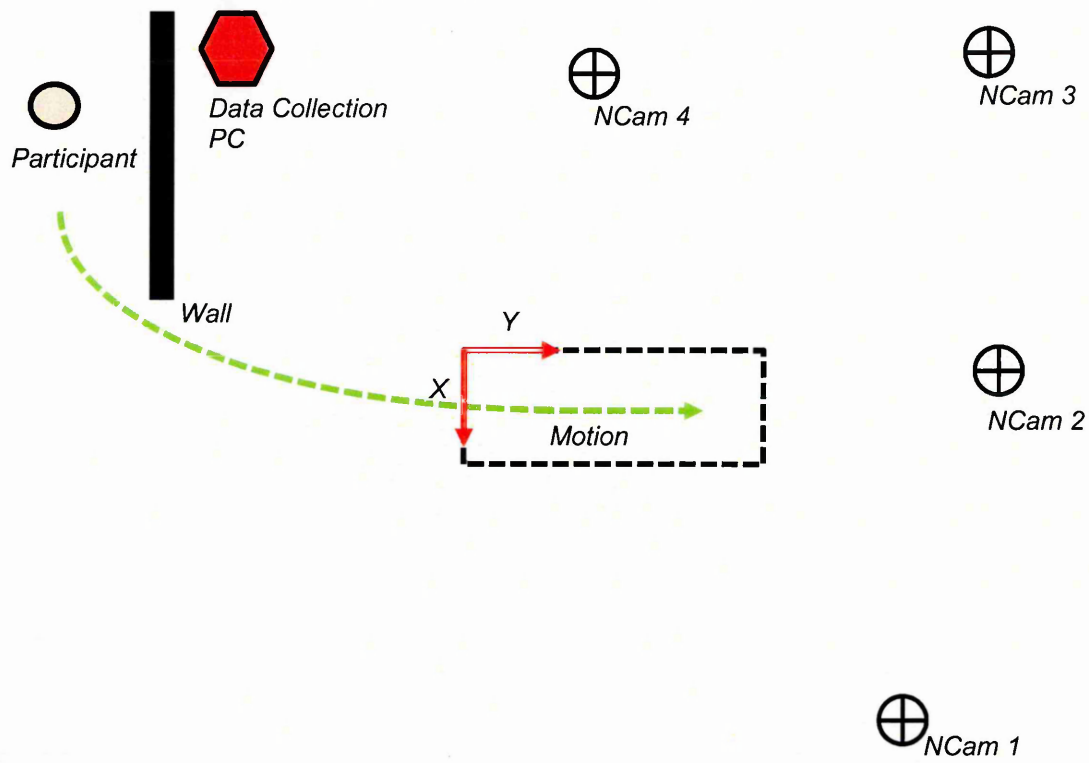


Figure 6.12. Schematic of experimental setup: dashed rectangle represents the motion capture volume filmed by four network cameras, e.g. NCam 1 - 4.

6.4.2. Background segmentation

Let $I_{RGB}(x, y, t_n)$ be the RGB frame (size $R \times C$ pixels) at time $t = [t_2, t_3, \dots, t_n, \dots, t_F]$ for an image sequence of length F , where (x_1, y_1) is the top-left corner of the image, i.e. B (Figure 6.13). Further, let $I_{RGB}(x, y, t_1)$ be the RGB reference frame used for background modelling, i.e. A (Figure 6.13). By applying the absolute image difference method, i.e. equation 6.1, the threshold $level_{RGB}(t_n)$ can be determined for the three components of $ADI_{RGB}(x, y, t_n)$, i.e. C (Figure 6.13), using equations 6.2 - 6.6. Equation 6.7 yields the binary image $ADI_B(x, y, t_n)$, i.e. D (Figure 6.13).

Connected components within $ADI_B(x, y, t_n)$ with an area less than 92 pixels or 0.01% of image resolution, i.e. $R \times C$, are assumed to be noise and are removed. Remaining pixels are closed using a 5×5 structuring element, i.e. E (Figure 6.13). Finally, the largest connected component within $ADI_B(x, y, t_n)$ is retained, i.e. F (Figure 6.13).

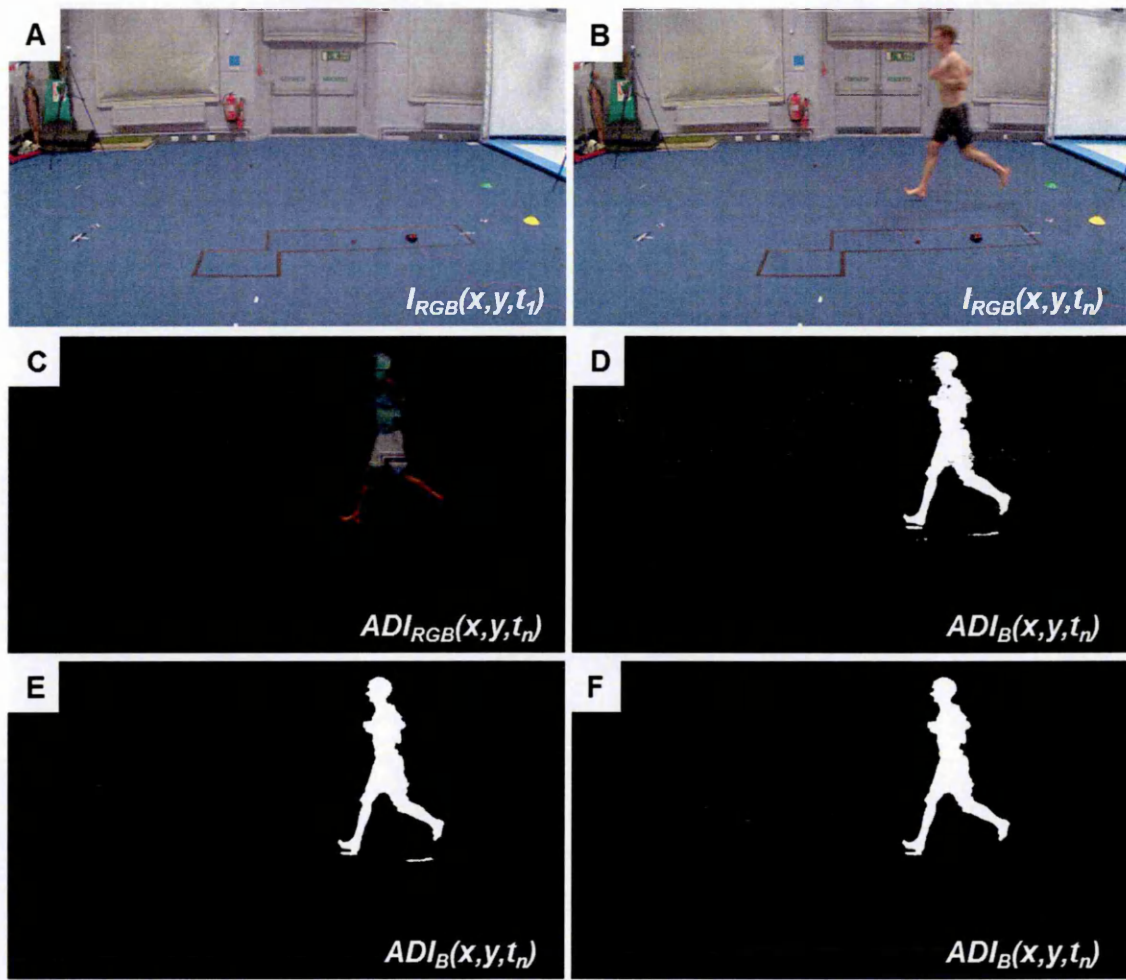


Figure 6.13. Absolute image differencing (A - C), binary image extraction (D) and morphological operations (E - F).

Silhouette height and COM coordinate, i.e. (H, t_n) and (COM_x, COM_y, t_n) respectively, are obtained. Image rows above COM_y are set to false, i.e. A (Figure 6.14). Subsequently the convex hull, i.e. $CHullADI_B(x, y, t_n)$, and bounding box (local coordinate system) of $ADI_B(x, y, t_n)$ are computed. $CHullADI_B(x, y, t_n)$ is then transformed into the image coordinate system, i.e. B (Figure 6.14).

$$ADI_B(1:x, 1:COM_y, t_n) = 0 \quad [6.13]$$

6.4.3. Shadow removal

The presence of shadow is clearly evident under the participants' right foot, i.e. A (Figure 6.14). Shadow removal is important due to the regional assessment of motion. Benedek and Szirányi (2007) identified the HSV colour space as a computationally fast and effective method for suppressing the impact of shadow for indoor scenarios.

However the HSV colour space is sensitive to image noise (Benedek and Szirányi, 2007). As such, shadow removal was limited to the region of interest, i.e. lower-body.

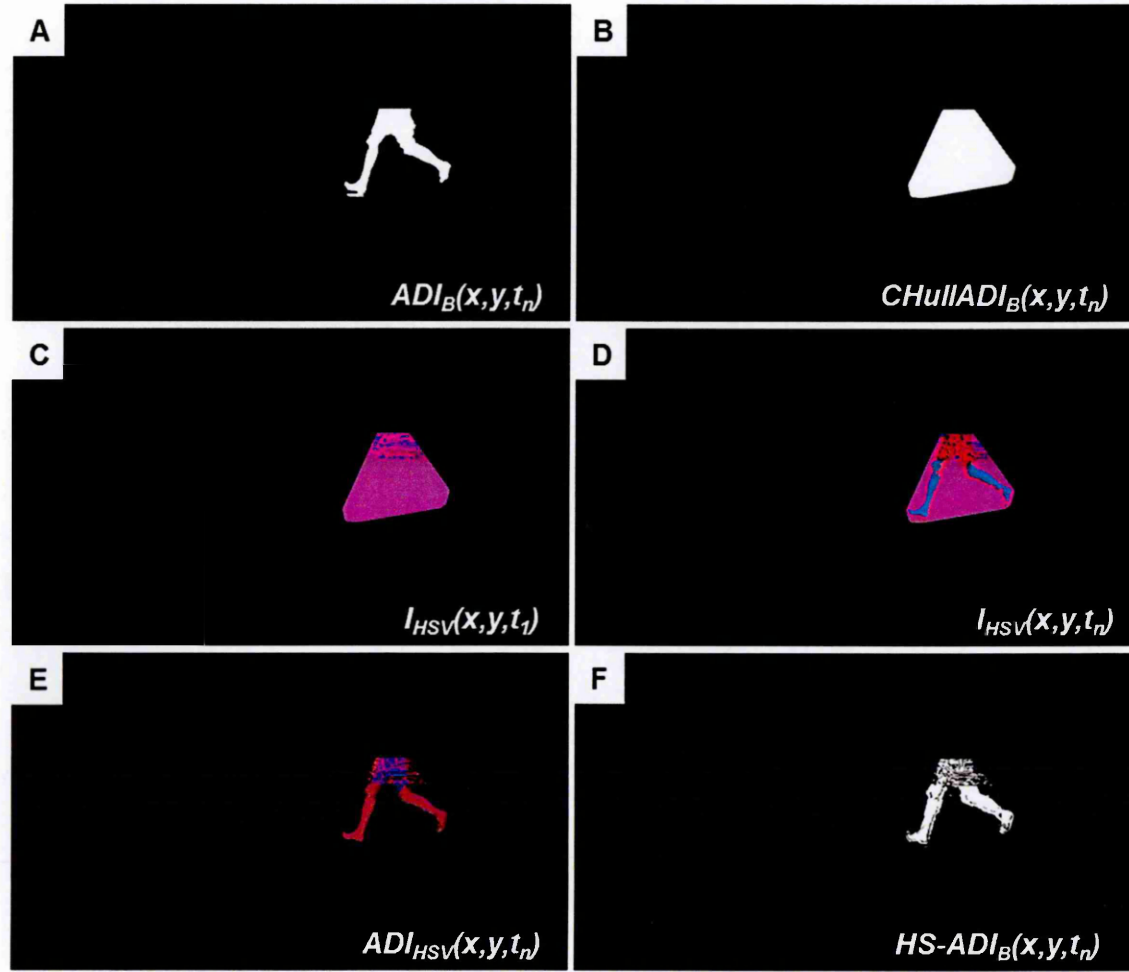


Figure 6.14. Lower-body (A), lower-body convex hull (B), HSV absolute image differencing (C - E) and binary image extraction (F).

$I_{RGB}(x, y, t_l)$ and $I_{RGB}(x, y, t_n)$ are converted to HSV colour space to yield $I_{HSV}(x, y, t_l)$ and $I_{HSV}(x, y, t_n)$ respectively. $CHullADI_B(x, y, t_n)$ is concatenated to yield the three-plane binary image $CHullADI_{BBB}(x, y, t_n)$: Figure 6.15 illustrates the three-plane concatenation of a binary image. $I_{HSV}(x, y, t_l)$ and $I_{HSV}(x, y, t_n)$ are then multiplied by $CHullADI_{BBB}(x, y, t_n)$ to crop HSV images to the convex hull, i.e. C - D (Figure 6.14).

$$I_{HSV}(x, y, t) = I_{HSV}(x, y, t) \times CHullADI_{BBB}(x, y, t_n) \quad [6.14]$$

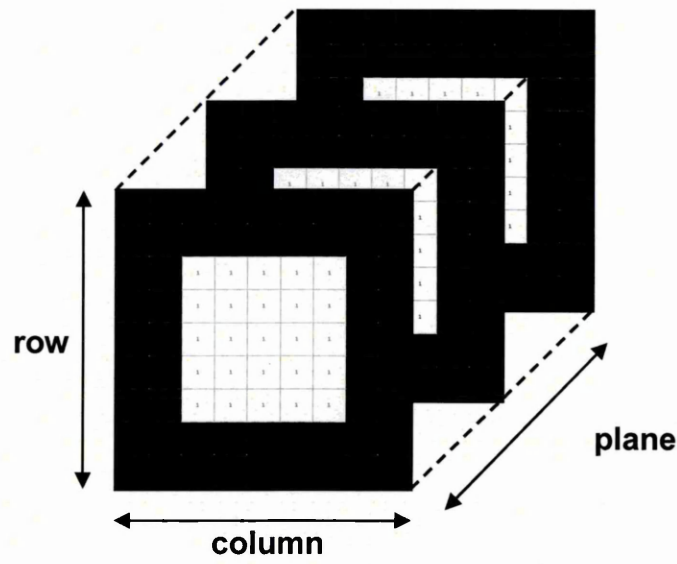


Figure 6.15. Three-plane binary image concatenation.

By applying the absolute difference image method (equation 6.1), the threshold $level_{HSV}(t_n)$ can be determined for the three components of $ADI_{HSV}(x, y, t_n)$, i.e. E (Figure 6.14), using equations 6.2 - 6.6. The third component of $ADI_{HSV}(x, y, t_n)$, i.e. Value plane, is removed. The binary images of Hue and Saturation components are extracted (equation 6.7) using the thresholds $level_{HS}(t_n)$. Extracted binary images are subsequently summed and converted to logical, yielding $HS-ADI_B(x, y, t_n)$, i.e. F (Figure 6.14). The same morphological operations applied in section 6.4.2 are subsequently applied to $HS-ADI_B(x, y, t_n)$, i.e. A (Figure 6.16). $HS-ADI_B(x, y, t_n)$ forms the basis for subsequent foot-region segmentation.

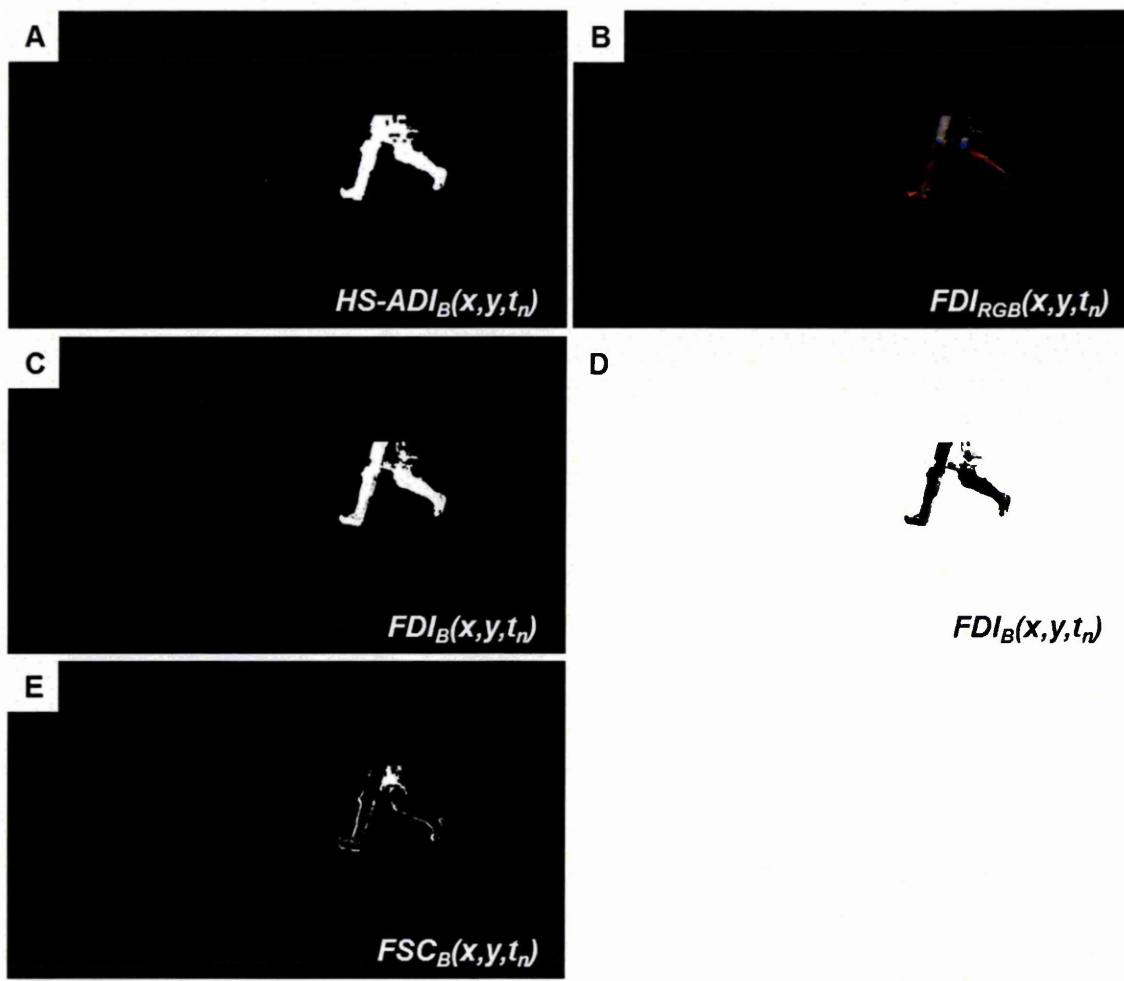


Figure 6.16. Lower-limb binary image with shadow suppressed (A), cropped inter-frame difference image (B) and binary image extraction (C - E).

6.4.4. Inter-frame motion

By applying the frame difference image method (equation 6.8), $FDI_{RGB}(x, y, t_n)$ can be obtained. $HS-ADI_B(x, y, t_n)$ is then concatenated to yield the three-plane binary image $HS-ADI_{BBB}(x, y, t_n)$. The product of $FDI_{RGB}(x, y, t_n)$ and $HS-ADI_{BBB}(x, y, t_n)$ is computed (equation 6.15) to crop $FDI_{RGB}(x, y, t_n)$ to the lower-body, i.e. B (Figure 6.16).

$$FDI_{RGB}(x, y, t_n) = FDI_{RGB}(x, y, t_n) \times HS-ADI_{BBB}(x, y, t_n) \quad [6.15]$$

$$FSC_B(x, y, t_n) = FDI_B(x, y, t_n) \times HS-ADI_B(x, y, t_n) \quad [6.16]$$

The threshold $level_{RGB}(t_n)$ can then be determined for the three components of $FDI_{RGB}(x, y, t_n)$ using equations 6.2 - 6.6. Equation 6.7 yields the binary image $FDI_B(x, y, t_n)$, i.e. C (Figure 6.16), and equation 6.9 subsequently inverts $FDI_B(x, y, t_n)$, i.e. D (Figure 6.16). Finally, $FDI_B(x, y, t_n)$ is multiplied by $HS-ADI_B(x, y, t_n)$ (equation 6.16): foreground

pixels within $FSC_B(x, y, t_n)$ reflect regions about the participant's lower-limbs that exhibit low inter-frame differences, i.e. E (Figure 6.16). For the current example the participant is in flight, thus inter-frame differences about the feet are low.

6.4.5. Foot segmentation

Foot segmentation follows a similar process to that outlined in section 6.3. The convex hull, i.e. $CHullHS-ADI_B(x, y, t_n)$, and bounding box of $HS-ADI_B(x, y, t_n)$ are computed, e.g. A (Figure 6.17). Subsequently the convex hull COM coordinate, i.e. $(CHullCOM_x, CHullCOM_y, t_n)$, is obtained. $CHullHS-ADI_B(x, y, t_n)$ is subsequently skeletonised, i.e. $SkelHS-ADI_B(x, y, t_n)$, and endpoint coordinates calculated. Skeletonised endpoint coordinates above $(CHullCOM_y, t_n)$ are discarded, i.e. B (Figure 6.17).

Remaining skeleton endpoint coordinates are transformed into the image coordinate system, i.e. $(Skel_x, Skel_y, t_n, j)$, where j represents the logical label. Perimeter pixels of $HS-ADI_B(x, y, t_n)$ are extracted, i.e. C (Figure 6.17), and the nearest skeleton-perimeter pixel (SPP) for each transformed endpoint coordinate identified, i.e. (SPP_x, SPP_y, t_n, j) . As previously noted, this is necessary as transformed skeleton endpoint coordinates are the product of $SkelHS-ADI_B(x, y, t_n)$ and not necessarily foreground $HS-ADI_B(x, y, t_n)$ pixels. SPP coordinates, i.e. red filled circles (C: Figure 6.17), serve as reference points for foot segmentation.

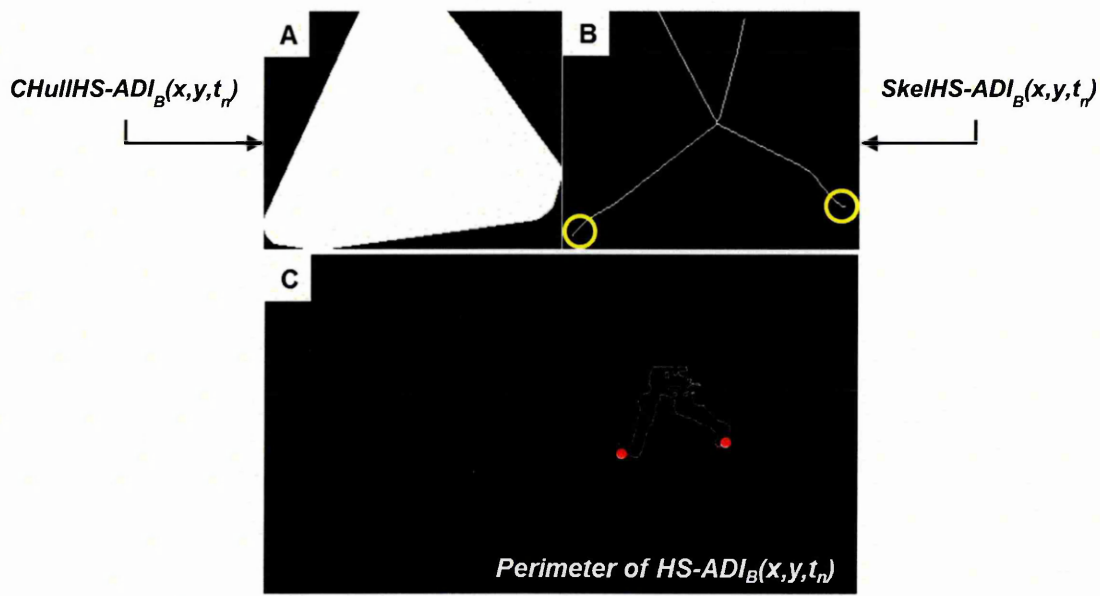


Figure 6.17. Convex hull of $HS-ADI_B$ (A), skeletonised image of $HS-ADI_B$ convex hull with inferior endpoints highlighted (B) and $HS-ADI_B$ perimeter pixels with skeleton endpoints transformed into the image coordinate system (red circles: C).

Foot segmentation is performed initially by creating j circular masks located at (SPP_x, SPP_y, t_n, j) with a radius equal to 20% of height, i.e. (H, t_n) . This yields $InitialMask_B(x, y, t_n, j)$, i.e. A (Figure 6.18). With reference to image A (Figure 6.18), $j = 1$. This indicates that the coordinate used to create the mask was the first logically labelled component within the image. Equation 6.17 yields the initial foot-region $InitialFoot_B(x, y, t_n, j)$, and is illustrated by B (Figure 6.18).

$$InitialFoot_B(x, y, t_n, j) = InitialMask_B(x, y, t_n, j) \times HS-ADI_B(x, y, t_n) \quad [6.17]$$

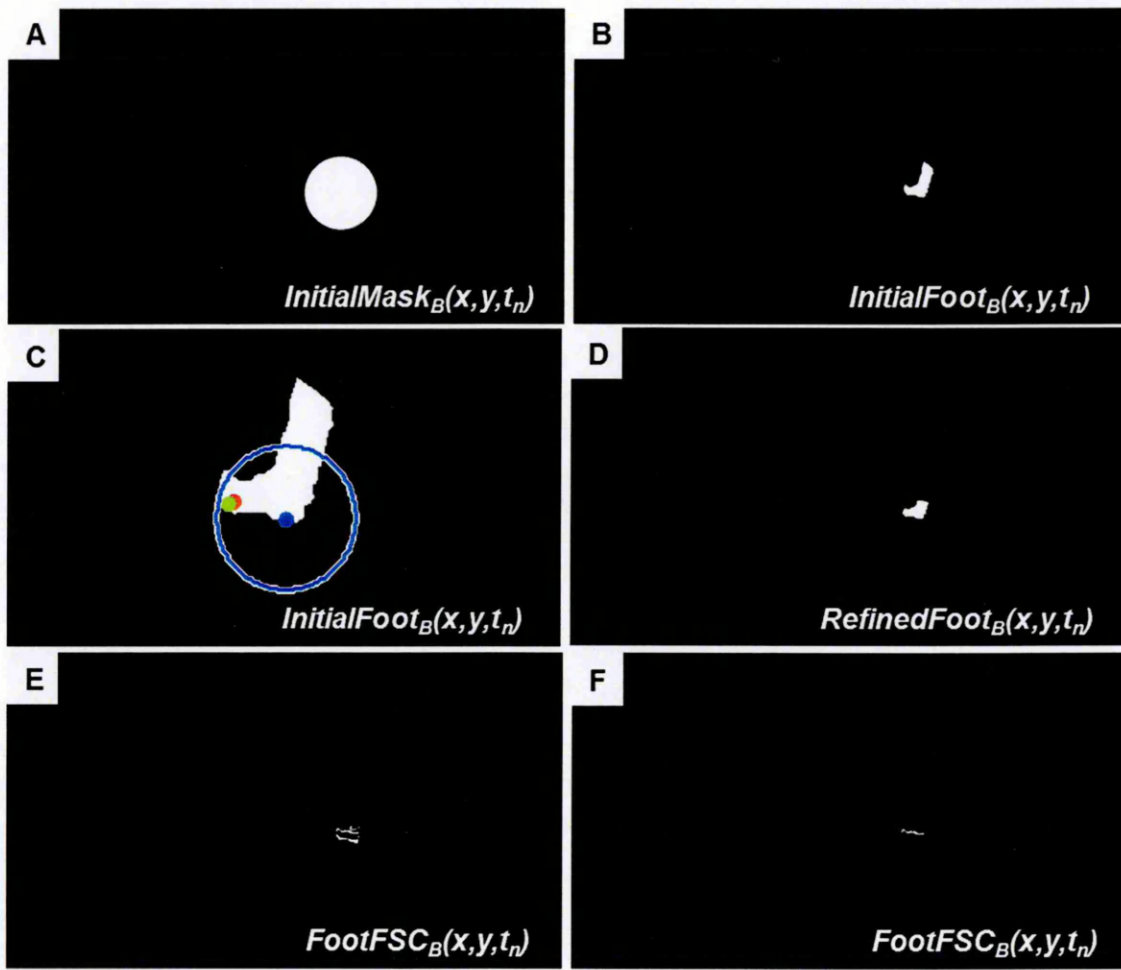


Figure 6.18. Initial circular mask (A), initial foot-region (B), initial foot-region (zoomed) with identified coordinates and refined mask highlighted (C), refined foot-region (D), product of refined foot mask and FSC_B (E) and largest connected component (F).

For j initial foot-regions, j refined foot-regions are then created. Due to the perspective dependent shape of the foot, the reference coordinate (RF_x, RF_y, t_n, j) is identified:

$$(RF_x, RF_y) = \frac{\sum InitialFoot_B(x, y, t_n, j)}{n} \quad [6.18]$$

$$RF_y = \max(InitialFoot_B(RF_x, y, t_n, j)) \quad [6.19]$$

Equations 6.18 and 6.19 identify the lowest pixel within the mean horizontal column, i.e. blue filled circle (C: Figure 6.18). Subsequently, j circular masks, located at (RF_x, RF_y, t_n, j) with a radius equal to 10% of (H, t_n) , yield $RefinedMask_B(x, y, t_n, j)$. For illustration, image C (Figure 6.18) presents a close-up of $InitialFoot_B(x, y, t_n, j)$, with the footprint of $RefinedMask_B(x, y, t_n, j)$ highlighted (large blue circle). Furthermore, the

transformed skeleton endpoint coordinate $(Skel_x, Skel_y, t_n, j)$, skeleton-perimeter pixel (SPP_x, SPP_y, t_n, j) and refined foot reference coordinate (RF_x, RF_y, t_n, j) are plotted as red, green and blue filled circles respectively (C: Figure 6.18).

Equation 6.20 yields the refined foot-region $RefinedFoot_B(x, y, t_n, j)$, i.e. D (Figure 6.18). Finally, equation 6.21 yields $FootFSC_B(x, y, t_n, j)$ and represents inter-frame differences within the refined foot-region, i.e. E (Figure 6.18). Small objects are discarded: only the largest connected component of $FootFSC_B(x, y, t_n, j)$ is retained for analysis, i.e. F (Figure 6.18). Subsequently, the areas of $RefinedFoot_B(x, y, t_n, j)$ and $FootFSC_B(x, y, t_n, j)$ are computed, i.e. D and F respectively (Figure 6.18). Equations 6.22 and 6.23 thus allow the analysis of foot-region, inter-frame differences.

$$RefinedFoot_B(x, y, t_n, j) = RefinedMask_B(x, y, t_n, j) \times HS-ADI_B(x, y, t_n) \quad [6.20]$$

$$FootFSC_B(x, y, t_n, j) = RefinedFoot_B(x, y, t_n, j) \times FSC_B(x, y, t_n) \quad [6.21]$$

$$(AreaRefinedFoot, t_n, j) = \sum RefinedFoot_B(x, y, t_n, j) \quad [6.22]$$

$$(AreaFootFSC, t_n, j) = \sum FootFSC_B(x, y, t_n, j) \quad [6.23]$$

6.4.6. Foot-surface contact threshold

In section 6.3, a generic threshold was applied to inter-frame differences to determine foot-surface contacts for synthetic image data. However the area or resolution of segmented foot-regions will vary depending on camera perspective and distance. Figure 6.19 illustrates the varying size and shape of a participants' foot during stance as a result of camera perspective. Images A - D were obtained from cameras 1 - 4 respectively, i.e. Figure 6.12.

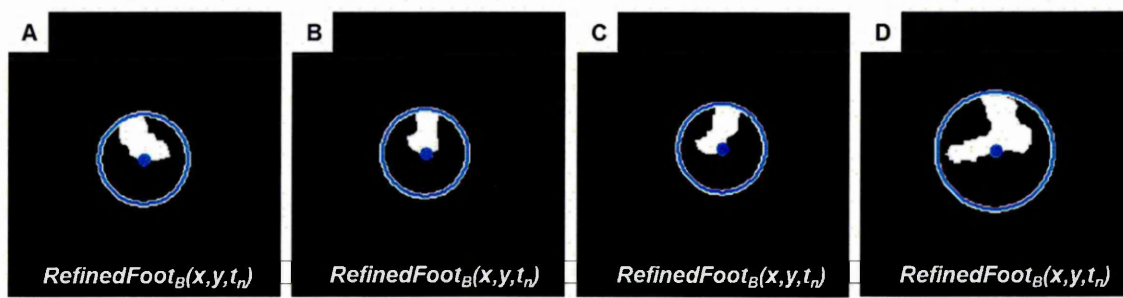


Figure 6.19. Refined foot-regions, i.e. *RefinedFoot_B*, obtained from cameras 1 - 4 (A - D respectively) with refined foot mask footprints highlighted (blue circles).

Therefore a suitable threshold must account for camera distance, perspective variation and inter-individual foot differences. To illustrate threshold identification for different camera distances and perspectives, data obtained from frontal and sagittal camera perspectives are used. Figure 6.20 presents foot-region area data for two segmented feet throughout an image sequence obtained from frontal and sagittal camera perspectives (A and B respectively). For reference purposes, green vertical bars highlight identified foot-surface contacts and the red vertical bar represents threshold initialisation: data obtained during this period are not used.

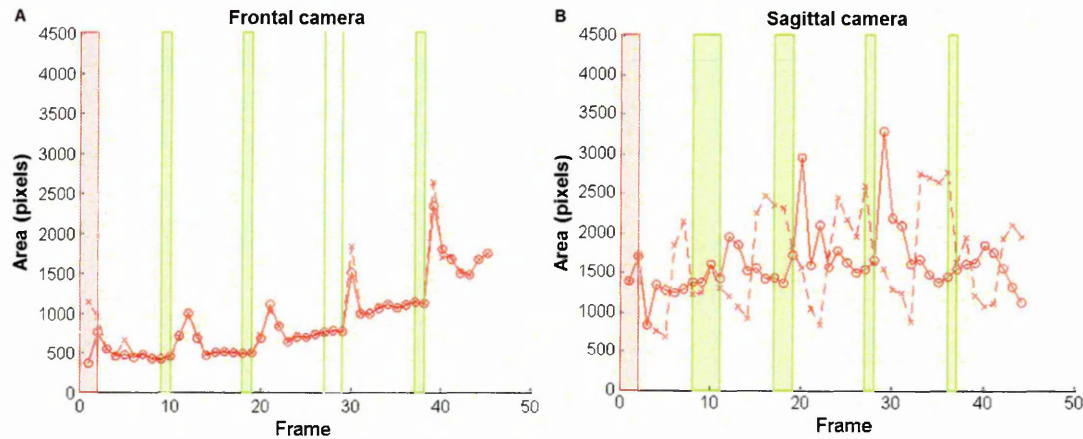


Figure 6.20. Refined foot-region areas (*AreaRefinedFoot*) for frontal (A) and sagittal (B) perspectives. Solid and dashed lines indicate foot-regions one and two respectively.

Foot-surface contacts (green vertical bars) were identified at similar instants despite the distinctly different foot-region areas obtained from frontal and sagittal camera perspectives (Figure 6.20). Foot areas generally increased as the image sequence progressed for frontal perspective data (A: Figure 6.20). This reflects that the participant

ran towards the camera, i.e. increasing resolution of region of interest. As described in section 6.4.5, j foot-regions are created at locations defined by (RF_x, RF_y, t_n, j) . Due to a narrow convex hull base for the frontal perspective of a single stance foot, i.e. A - B (Figure 6.21), the coordinates (SPP_x, SPP_y, t_n, j) used to identify (RF_x, RF_y, t_n, j) are located about the same foot. As such, foot-region areas are derived from similar regions (B: Figure 6.21). Apparent spikes in foot-region area (A: Figure 6.20) reflect the swing foot momentarily passing the stance foot, i.e. C - D (Figure 6.21).

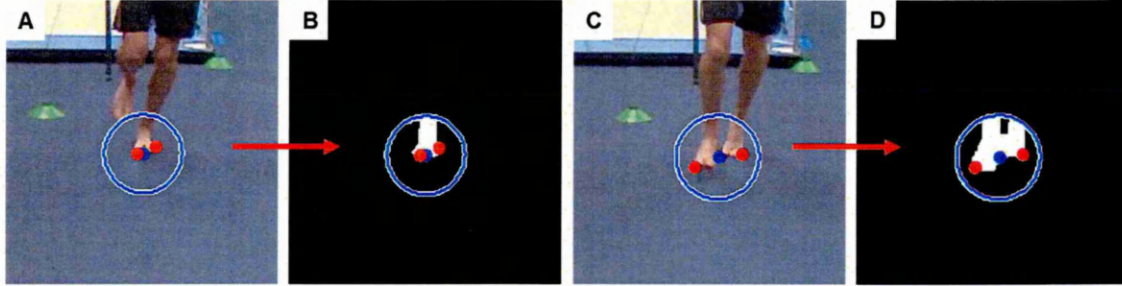


Figure 6.21. Frontal camera perspective of stance with refined foot-regions (blue circles) and skeleton-perimeter pixels (red filled circles) highlighted.

For a sagittal perspective (B: Figure 6.20), foot areas oscillated about a mean magnitude of ~ 1600 pixels. This reflects that the participant ran perpendicular to the camera. Out-of-phase oscillation reflects foot orientation during foot segmentation. As noted, section 6.4.5 described the creation of j foot-regions at the coordinates (RF_x, RF_y, t_n, j) . For binary foot-regions B and D (Figure 6.22), the calculation of (RF_x, RF_y, t_n, j) yields different foot locations due to foot orientation, i.e. blue filled circles plotted in A - D (Figure 6.22). As such, a circle with the same radius, i.e. 10% of (H, t_n) , will incorporate different foot and shank regions depending on foot orientation, i.e. A - D (Figure 6.22). Furthermore, motion blur can influence foot-region areas, i.e. C - D (Figure 6.22).

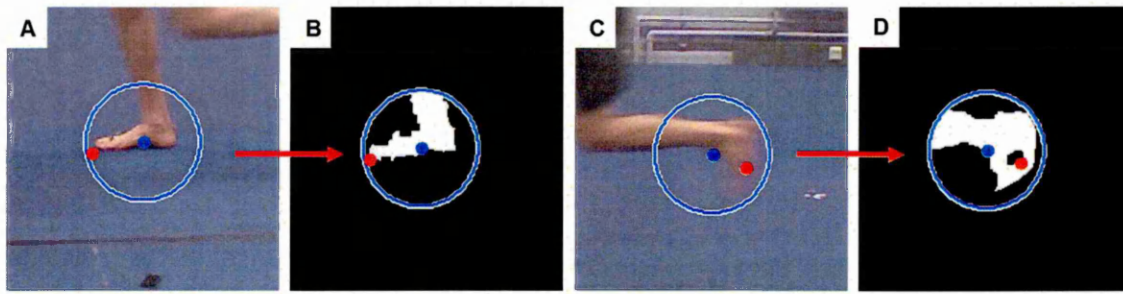


Figure 6.22. Sagittal camera perspective of stance with refined foot-regions (blue circles) and skeleton-perimeter pixels (red filled circles) highlighted.

Figures 6.19 - 6.22 demonstrate that camera perspective and foot orientation influence the area or resolution of segmented foot-regions. To assess segmented inter-frame differences about the feet, a generic threshold would not be appropriate: a dynamic threshold related to foot-region area was required.

Figure 6.23 illustrates a threshold model (M_t) representing the relationship between $(AreaRefinedFoot, t_{2:n}, j)$ and $(AreaFootFSC, t_{2:n}, j)$. Linear functions for $(AreaRefinedFoot, t_{2:n}, j)$ and $(AreaFootFSC, t_{2:n}, j)$ are computed iteratively for an image sequence and yield the univariate polynomials P_1 and P_2 respectively. Subsequently, equation 6.24 defines M_t as the quotient of the mean magnitudes for P_1 and P_2 .

$$(M_t, t_n, j) = \frac{\sum P_1/n}{\sum P_2/n} \quad [6.24]$$

where n is polynomial length.

The computation of a univariate polynomial requires more than one data point: therefore data obtained for $t < 3$ are ignored, i.e. more than two data points (red vertical bars: Figure 6.23). Figure 6.23 demonstrates that M_t exhibits a similar characteristic regardless of camera perspective or foot orientation. This reflects the underlying relationship between FDI and ADI images. Foot-surface contacts (green vertical bars: Figure 6.23) are located in or around troughs of M_t for both frontal and sagittal camera perspectives. This reflects the smaller difference in magnitude between $(AreaFootFSC, t_n, j)$ and $(AreaRefinedFoot, t_n, j)$, i.e. little foot motion during stance.

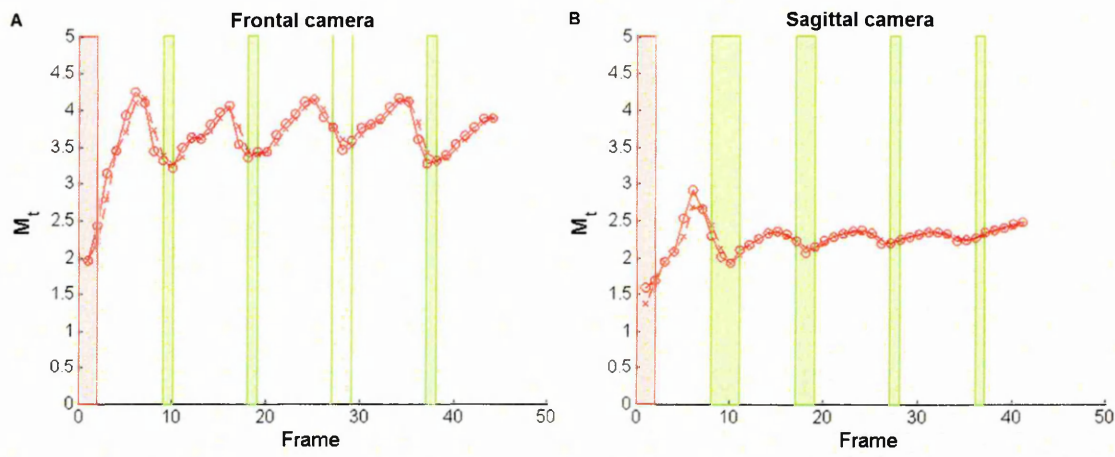


Figure 6.23. M_t for frontal (A) and sagittal (B) camera perspectives. Solid and dashed lines indicate M_t for foot-regions ones and two respectively.

However M_t had generally greater magnitudes for foot-regions of lower resolution, i.e. frontal perspective (A: Figure 6.23). This reflects that image resolution is an important parameter when quantifying inter-frame differences. For j refined foot-regions, the magnitude of $(AreaRefinedFoot, t_n, j)$ is normalised by 75% of (M_t, t_n, j) . This serves to allow the largest magnitudes of $(AreaFootFSC, t_n, j)$ to exceed the normalised $(AreaRefinedFoot, t_n, j)$. The use of the 75th percentile was determined experimentally, i.e. suitable value for all camera perspectives. Therefore, the logical foot-surface contact candidate is given by:

$$(FSC_{Cand}, t_n, j) = (AreaFootFSC, t_n, j) > \frac{(AreaRefinedFoot, t_n, j)}{(M_t, t_n, j) \times 0.75} \quad [6.25]$$

Figure 6.24 illustrates the implementation of equation 6.25: red solid and dashed lines are normalised magnitudes for $(AreaRefinedFoot, t_n, j)$ for j foot-regions respectively. Blue solid and dashed lines are corresponding magnitudes for $(AreaFootFSC, t_n, j)$ for j foot-regions respectively. Green vertical bars (Figure 6.24) highlight foot-surface contacts, i.e. $(FSC_{Cand}, t_n, j) = 1$. Figure 6.24 demonstrates that foot-surface contacts can be identified for camera views of varying size and perspective.

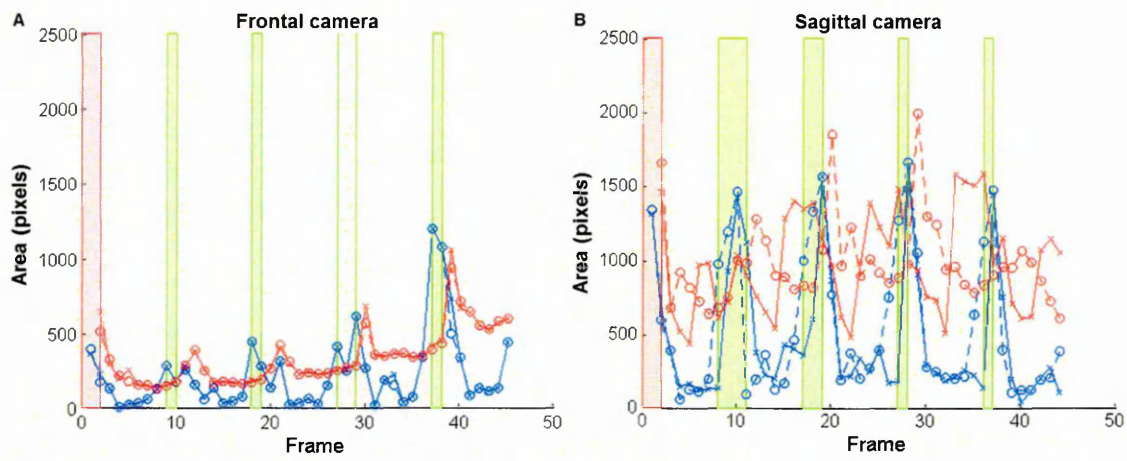


Figure 6.24. Normalised refined foot-region areas (red) and corresponding foot-region inter-frame differences (blue) for frontal (A) and sagittal (B) camera perspectives. Solid and dashed lines indicate foot-regions one and two respectively.

6.4.7. Geometric rules

As indicated in section 6.4.6, (FSC_{Cand}, t_n, j) are foot-surface contact candidates. Geometric rules are required to differentiate true positive and false positive candidates. Inter-frame differences were quantified using the largest connected component within $FootFSC_B(x, y, t_n, j)$. Therefore the corresponding COM coordinate for $FootFSC_B(x, y, t_n, j)$ is obtained, i.e. (fsc_x, fsc_y, t_n, j) . To identify whether (FSC_{Cand}, t_n, j) corresponds to a foot-surface contact, the scalars S_1 and S_2 are formed.

S_1 is given by equation 6.27 and is the vertical distance between (fsc_y, t_n, j) and the inferior-most pixel in the refined foot-region, i.e. equation 6.26. S_2 is given by equation 6.28 and is 50% of the vertical distance between $(Skel_y, t_n, j)$ and $(CHullCOM_y, t_n)$, i.e. vertical coordinates for convex hull skeleton endpoint and COM.

$$(FootMinima, t_n, j) = \max(RefinedFoot_B(x, y, t_n, j)) \quad [6.26]$$

$$(S_1, t_n, j) = (fsc_y, t_n, j) - (FootMinima, t_n, j) \quad [6.27]$$

$$(S_2, t_n, j) = (Skel_y, t_n, j) - ((Skel_y, t_n, j) - (CHullCOM_y, t_n) / 2) \quad [6.28]$$

Subsequently, three geometric rules are applied. The first geometric rule, i.e. R_1 (equation 6.29), identifies whether (fsc_x, fsc_y, t_n, j) is located within the refined foot-region $RefinedFoot_B(x, y, t_n, j)$. Figure 6.25 illustrates the foreground segmentation of a participant during running: RGB pixel intensities have been modified to illustrate

regional segmentation. The convex hull and refined foot-region (stance foot) perimeters are outlined in red and green respectively. Figure 6.25 illustrates that the foot-surface contact candidate (green filled circle and red cross) is located within the refined foot-region.

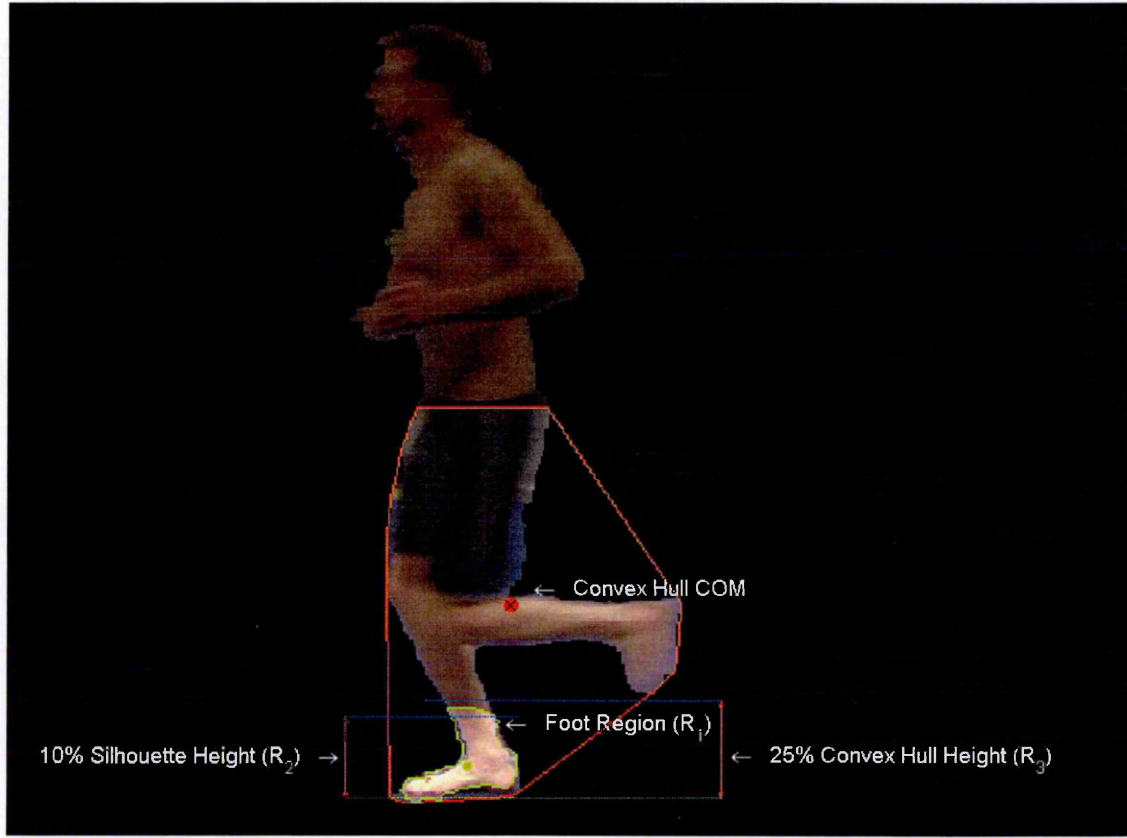


Figure 6.25. Geometric rules applied to a foot-surface contact candidate (green filled circle and red cross).

$$(R_1, t_n, j) = (fsc_x, fsc_y, t_n, j) \times RefinedFoot_B(x, y, t_n, j) \quad [6.29]$$

$$(R_2, t_n, j) = (S_1, t_n, j) \leq ((H, t_n) \times 0.1) \quad [6.30]$$

$$(R_3, t_n, j) = (fsc_y, t_n, j) > (S_2, t_n, j) \quad [6.31]$$

The second geometric rule, i.e. R_2 (equation 6.30), determines whether (fsc_x, fsc_y, t_n, j) is located within the inferior most 10% of the participant silhouette, i.e. foot and ankle region (Goffredo *et al.*, 2008). Figure 6.25 illustrates that the foot-surface contact candidate is located within the foot and ankle region. Finally, the third geometric rule, i.e. R_3 (equation 6.31), determines whether (fsc_x, fsc_y, t_n, j) is located within the inferior-most 25% of $CHullHS-ADI_B(x, y, t_n)$. This is necessary as rules R_1 and R_2 are related to

individual foot-regions: rule R_3 provides a global reference, i.e. the participant, for foot-surface contact candidates. Figure 6.25 illustrates that the foot-surface contact candidate is located within this region. Subsequently, a candidate is deemed a foot-surface contact, i.e. $(FSCx, FSCy, t_n, j)$, if (FSC_{Cand}, t_n, j) and rules R_{1-3} are true:

$$(FSC_{Cand}, t_n, j) \wedge (R_1, t_n, j) \wedge (R_2, t_n, j) \wedge (R_3, t_n, j) \rightarrow (FSCx, FSCy, t_n, j) = (fsc_x, fsc_y, t_n, j) \quad [6.32]$$

6.5. Application to walking and running

Section 6.4 presented a method to automatically identify foot-surface contacts from single camera image sequences. The algorithm was developed using colour image sequences where participant size and perspective varied. The following section demonstrates the application of the current algorithm to different gait modes, i.e. walking and running.

Current markerless gait analysis algorithms incorporate spatio-temporal features associated with walking gait to extract gait parameters, e.g. Johnson and Bobick (2001); Jung and Nixon (2013). As noted in section 6.3, the current algorithm imposes no *a priori* assumptions regarding gait mode: foot-regions and corresponding inter-frame differences are quantified independently. Therefore gait mode, i.e. single stance (running) or dual stance (walking), is irrelevant to the current algorithm. Figure 6.26 illustrates the application of the current algorithm to four camera perspectives of running, i.e. cameras 1 - 4 (Figure 6.12). Figure 6.26 highlights foot-surface contact data identified for each camera perspective (green filled circle and red cross). Figure 6.26 also highlights individually identified foot-region masks: red and blue circles indicate logically identified foot-regions, i.e. $j = 1$ and 2 respectively.

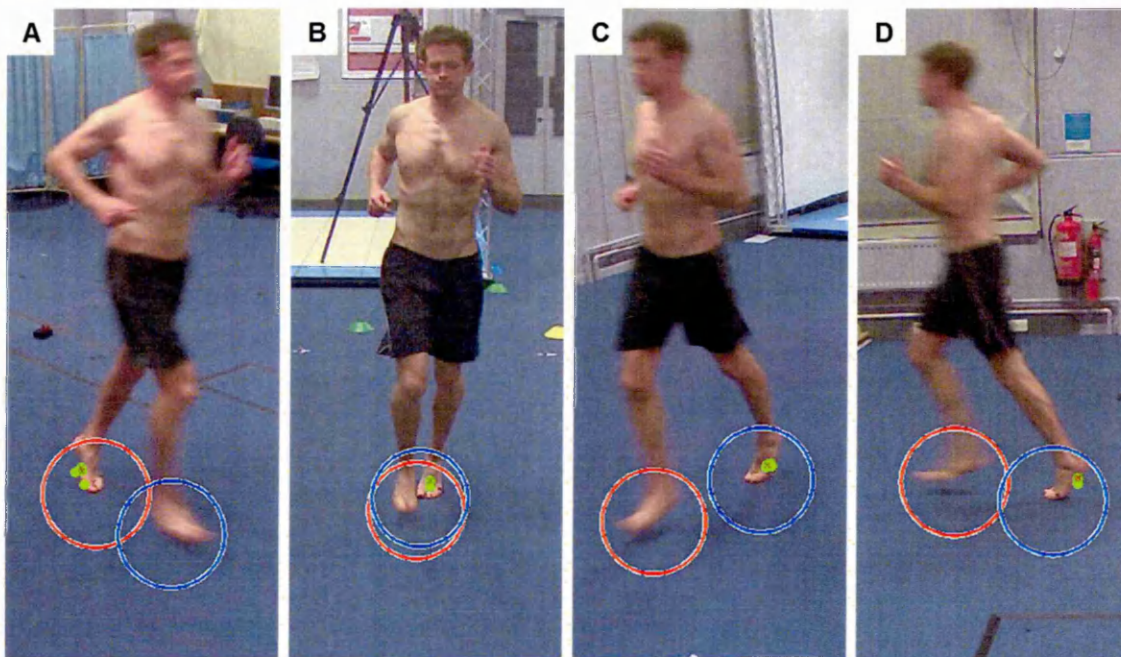


Figure 6.26. Logical foot-region masks (red and blue circles) and foot-surface contacts (filled green circles and red cross) for four camera perspectives.

Figure 6.26 demonstrates that quantifying inter-frame differences can differentiate static and dynamic foot-regions. However, as noted in section 6.3, the algorithm will be limited in its estimation of stance time due to the threshold used to determine foot-surface contacts, i.e. section 6.4.6. Figure 6.27 presents a sagittal perspective of walking and running foot-surface contacts. The stance phase in walking is longer than 50% of the gait cycle (Novacheck, 1998). Therefore at least one foot is always in contact with the ground. For walking, the current algorithm did not identify dual-stance (Figure 6.27). This reflects the threshold applied to inter-frame differences about the foot-region: the algorithm is not restricted in the number of foot-regions it can assess. It is evident that plantar-flexion during the toe-off phase (walking images 13 onwards: Figure 6.27) increased foot-region inter-frame differences such that the foot was not deemed to be static. Therefore stance time estimates are likely to be underestimated.

Walking images (Figure 6.27) demonstrate self-occlusion: the swing foot occludes the stance foot (blue dashed rectangle: images 6 - 8). Self-occlusion can increase detection errors for monocular images due to depth ambiguity (Poppe, 2007). However, due to swing foot motion, the current algorithm does not record data due to high inter-frame motion about the occluded stance foot-region. Therefore the current algorithm is less susceptible to identification errors arising from swing foot occlusion.

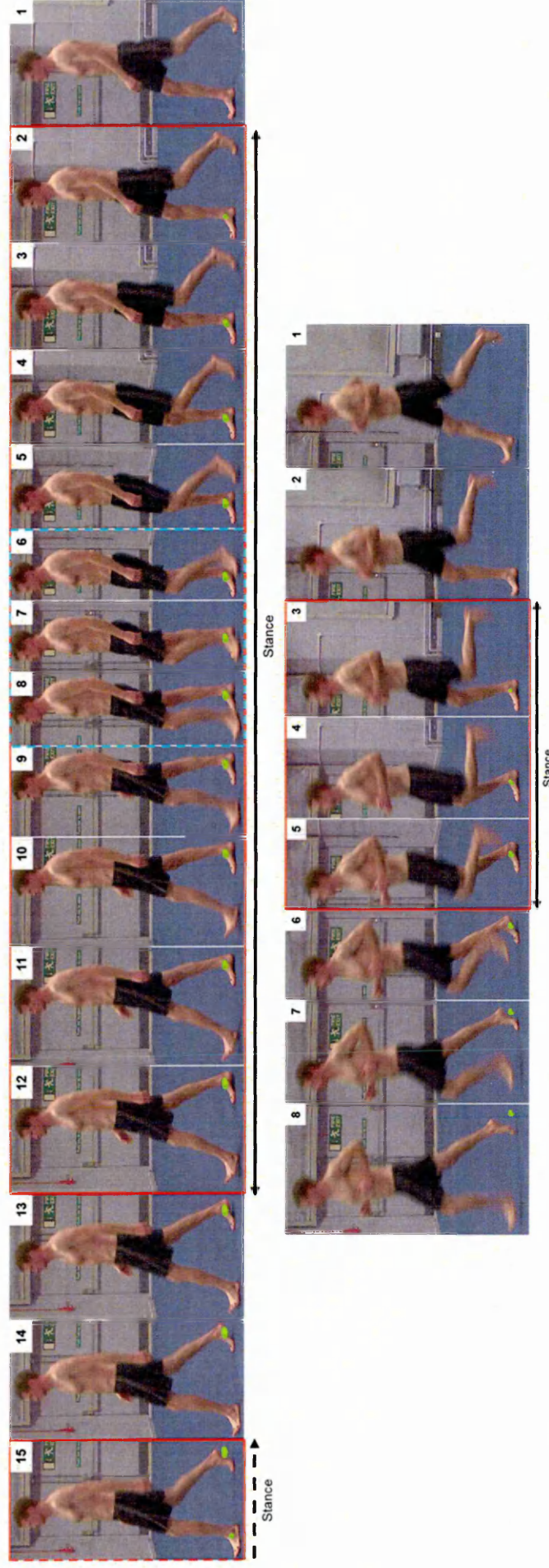


Figure 6.27. Stance (highlighted by red rectangle) is evidenced by accumulating foot-surface contact data points (green filled circles) about the stance foot for both walking (top) and running (bottom) gait (temporal direction is right-to-left). Self-occlusion (walking) is highlighted by blue dashed rectangle (images 6 - 8).

For running, plantar-flexion during toe-off (running images 6 onwards: Figure 6.27) again increased foot-region inter-frame differences such that the foot was not deemed to be static. It is evident that the algorithm identifies the foot-flat phase of stance, reflecting the centroid of low inter-frame differences about the foot. Figure 6.28 demonstrates the application of the current algorithm to walking and running image sequences from four different camera perspectives. Clusters of identified foot-surface contacts (green filled circles) are evident as a result of the algorithm operating on a frame-by-frame basis (Figure 6.28). Figure 6.28 demonstrates that evaluating inter-frame differences about feet-regions can identify foot-surface contacts for different gaits and camera perspectives. However, due to the perspective dependent shape of the foot, camera perspective will influence observed inter-frame motion, i.e. plantar-flexion etc., and thus influence foot-surface contact estimates. An objective assessment of accuracy is required to identify the spatial and temporal validity of foot-surface contact estimates obtained by the current algorithm.

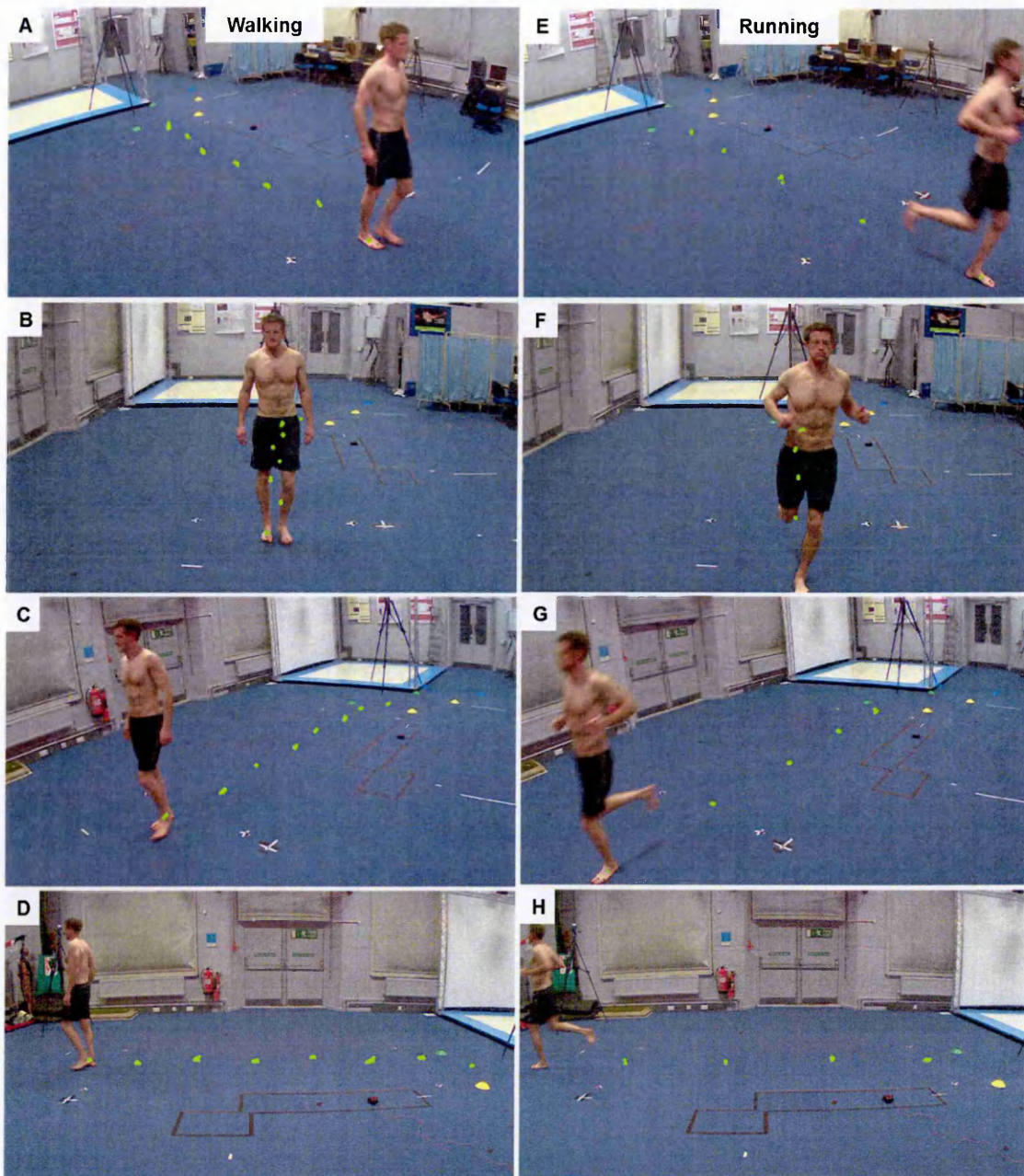


Figure 6.28. Collated foot-surface contact data (green filled circles) superimposed on image t_F of walking (A - D) and running (E - H) image sequences for four camera perspectives.

6.6. Conclusion

This chapter describes the development of a vision-based algorithm for automatically identifying foot-surface contacts from single camera footage. The algorithm was applied to four different camera perspectives of walking and running and returned visually corresponding results. Potential limitations to the current algorithm have been highlighted. An objective assessment is required to identify the spatial and temporal validity of foot-surface contacts identified by the current algorithm.

In relation to the overall project aim, the current algorithm has been demonstrated to identify foot-surface contacts for different gaits, i.e. walking and running, from different camera perspectives and from participant images of different size. In relation to tennis, the current algorithm therefore demonstrates distinct advantages over existing single camera heel-strike identification algorithms which are limited to walking, e.g. Bouchrika and Nixon (2006); Jung and Nixon (2013). The current algorithm is independent of gait mode, i.e. walking and running, and highlights that foot-surface contacts can be identified without *a priori* knowledge of gait. This indicates that the algorithm could be applied to multi-modal gait observed in tennis; however, any application must be evaluated. Finally, the algorithm is objective. The potential reduction in analysis time and random manual digitising error warrant the algorithm's application to match-play tennis.

7 Validation of an automatic technique for identifying foot-surface contacts in walking and running

7.1. Introduction

Chapter 3 described a manual approach for identifying foot-surface contacts and subsequently measure step length. However manual digitising is user intensive. Furthermore, changing player position and orientation will exacerbate random error induced by manual digitising (Glazier and Irwin, 2001). Chapter 6 presented an automatic technique for identifying foot-surface contacts using single camera images. The foot-surface contact identification (hereinafter referred to as FSCi) algorithm identified the time and image coordinates of foot-surface contacts. The algorithm was successfully applied to images of different size and perspective as well as different modes of gait, i.e. walking and running. However an objective assessment is required to identify the validity of foot-surface contact data identified by the FSCi algorithm for walking and running. Existing work has reported the successful identification of heel-strikes in walking videos: 94% of heel-strikes were identified with position errors of ± 100 mm (Jung and Nixon, 2013). This existing work will therefore serve as acceptance criteria for this assessment of the FSCi algorithm.

7.2. Aim and objectives

The aim of this chapter is to validate the automatic technique for identifying foot-surface contacts described in Chapter 6. This relates to box F of the development stage diagram (Figure 1.2). Algorithm performance will be assessed in relation to existing work (identified in section 7.1).

Objectives:

1. Validate the FSCi algorithm for multiple camera perspectives of walking and running.
2. Assess performance of the algorithm with regard to success rate, position error and analysis time.
3. Assess the limitations of the algorithm and its application to match-play tennis.

7.3. Laboratory validation study

In order to validate the FSCi algorithm, a real-world criterion measure must be established. Three-dimensional motion analysis is the industry standard for kinematic analysis. Due to multiple camera views, marker occlusion is not a limiting factor to analyses. Furthermore, multiple foot-surface contacts within a motion capture volume can be collected; this was a central limitation to the use of a force or pressure platform. Accordingly three-dimensional motion analysis was deemed a suitable method for identifying an objective criterion measure.

7.3.1. Participants and procedures

Six male participants (age = 27.9 ± 2.9 years; stature = 1.85 ± 0.05 m; mass = 77.6 ± 8.2 kg) were recruited. Participants were appropriately briefed to aid the completion of the proposed tasks and written informed consent was obtained. Approval for all procedures was obtained from the Research Ethics Committee of the Faculty of Health and Wellbeing, Sheffield Hallam University (Appendix 7). Participants were asked to walk and run at a self-selected pace through a motion capture volume within a carpeted laboratory, e.g. Figure 7.1. Participants were asked to perform three repetitions of these tasks in both barefoot and shod (own trainers) conditions.

7.3.2. Experimental setup

Seven spherical, retro-reflective markers (12.5 mm diameter) were affixed to palpable anatomical landmarks: one marker was placed on the sacrum and three markers were placed on the heel, 2nd and 5th meta-tarsal heads of the left and right feet. Three-dimensional marker position data were recorded using an eight camera motion capture system (Motion Analysis Corporation, Santa Rosa, CA, USA). The motion capture system sampled marker position data at 200 Hz. Cameras were mounted on tripods and wall mounts to ensure the optimal coverage of the motion capture volume measuring $4.0 \times 1.5 \times 1.5$ m in the anterior-posterior, medio-lateral and vertical directions respectively, i.e. MAC 1 - 8 (Figure 7.1). Following calibration, the motion capture system reported resultant calibration residuals of 0.51 ± 0.27 mm. The motion capture system was interfaced with a push-button trigger. This enabled the motion capture system to be triggered to record with an external 5V pulse. The same external pulse was amplified (12V) to illuminate an LED light box when the motion capture system began recording (Figure 7.1).

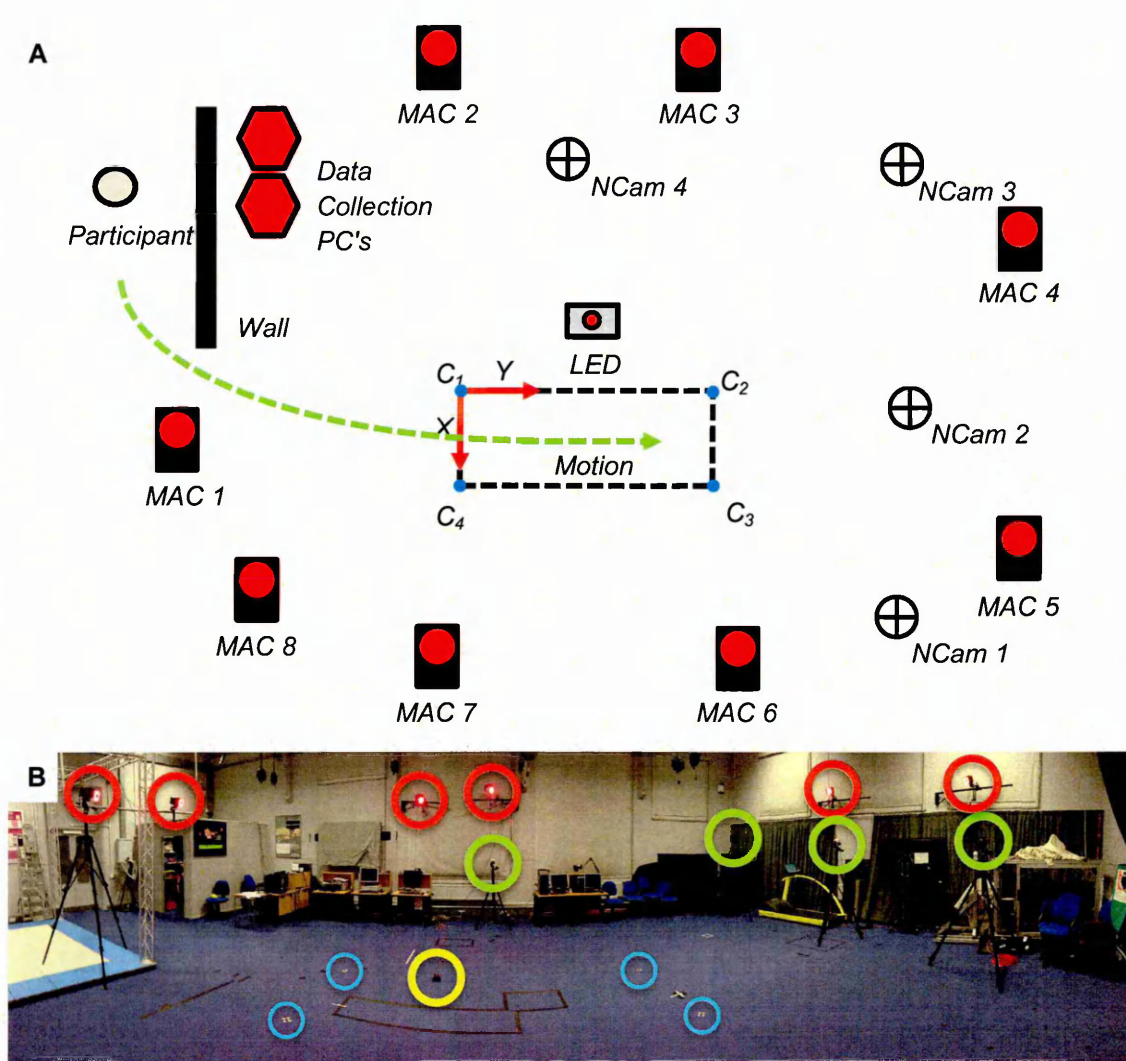


Figure 7.1. Experimental setup: motion analysis cameras, e.g. MAC 1 - 8 (A) and red circles (B), network cameras, e.g. NCam 1 - 4 (A) and green circles (B), filmed the motion capture volume, e.g. $C_1 - C_4$ (A) and blue circles (B), and LED light box, e.g. LED (A) and yellow circle (B).

The maximum sample rate of the commercial, high-definition camera used in section 3.3.2 (Chapter 3) was 25 Hz. Therefore four networked cameras (AXIS M1104, AxisTM Communications, Sweden), streamed RGB colour images (1280×720 pixels) to a dedicated data collection computer at 25 Hz. Network cameras were positioned to record sagittal, frontal and two oblique frontal perspective images of walking and running, i.e. NCam 1 - 4 (Figure 7.1). Following the positioning of network cameras, camera field-of-view and focal length were set manually and locked; no further alterations to camera intrinsic parameters were made. The LED light box was positioned in each camera's field-of-view to provide a time reference for FSCi algorithm data. Prior to motion trials, participants were required to wait behind a blanking wall (Figure 7.1)

whilst images streamed to the data collection computer were buffered. This enabled camera sample rates to stabilise prior to motion trials.

Image sample rates can fluctuate because camera images are streamed to a computer using a network. Sample rate fluctuations can be the result of local network activity or hard-disk write speeds. To enable the quantification of image sample rate, image filenames included CPU clock times. All recorded CPU clock times were assessed. For a single image sequence, the maximum image time interval was 0.5 s for all four network cameras. However this occurred within the initial 5 frames of data collection: image time intervals reduced to 0.04 s after 10 frames. In relation to data analysis, this would have had no impact. Analyses were only performed for images more than 5 s or 125 frames into an image sequence. Therefore, for 288 image sequences, i.e. 6 participants \times 4 cameras \times 3 trials \times 2 activities \times 2 conditions, image time intervals were 0.04 ± 0.00 s intervals, i.e. 25 Hz.

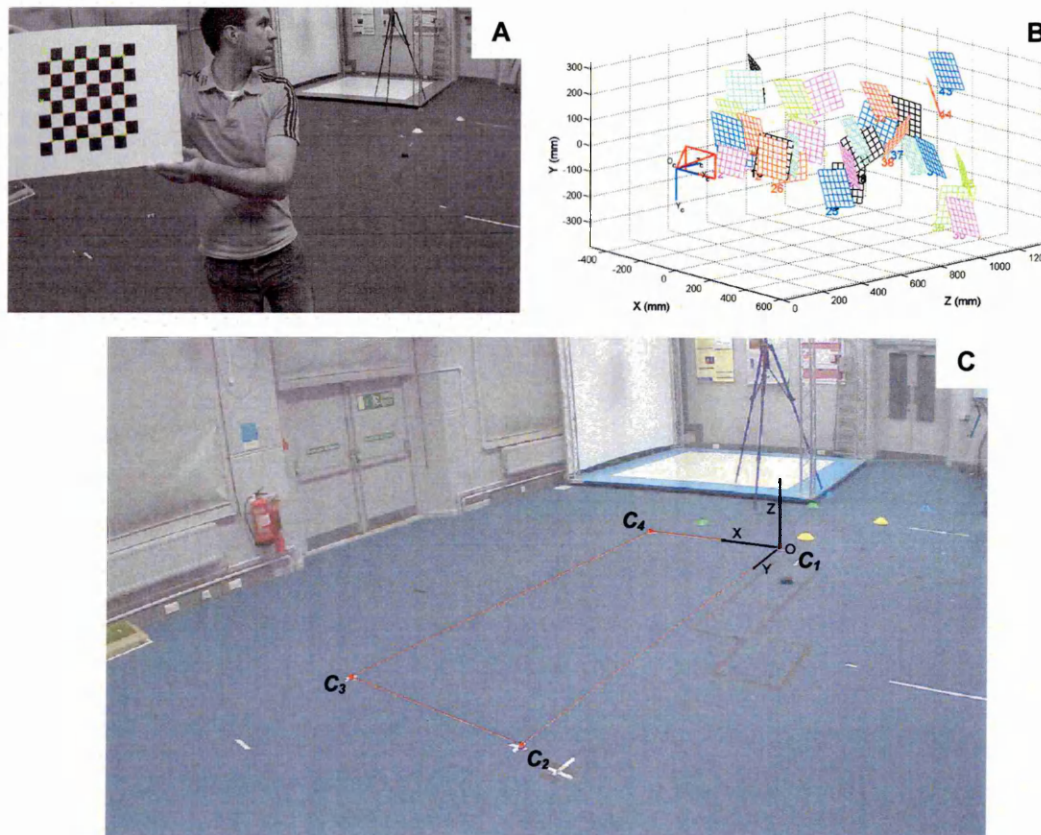


Figure 7.2. Network camera calibration: checkerboard extraction (A), extrinsic checkerboard parameters (B) and network camera image with motion capture volume superimposed (C).

Following the calibration of the motion analysis system, four spherical retro-reflective markers (25 mm diameter) were positioned on the laboratory floor at the corners of the motion capture volume. This defined a reference plane and common coordinate system, e.g. $C_1 - C_4$ (Figure 7.1). Marker locations were measured using the motion analysis system to minimise positioning error and help ensure orthogonality: marker position residuals were 0.92 ± 0.42 mm. For each network camera, single camera calibration was performed. Intrinsic camera parameters were calculated by filming a 6×6 checkerboard of 25 mm squares held in different positions and orientations relative to the camera, e.g. A (Figure 7.2). Checkerboard corners were extracted from camera footage *post-hoc* and processed using the Camera Calibration Toolbox for MATLAB (Bouguet, 2010), e.g. B (Figure 7.2). Finally, a single image from each network camera was used to manually digitise the four spherical markers ($C_1 - C_4$) at a sub-pixel resolution, e.g. C (Figure 7.2). Marker coordinates were digitised on five occasions; standard error of the mean was less than 0.15 pixels for all image coordinates. The mean marker coordinates were used to calculate extrinsic camera parameters for each network camera. Table 7.1 presents mean calibration residuals for network cameras, i.e. differences between real and estimated checkerboard square sizes (Choppin, 2008), and camera azimuth and elevation angle. Camera elevation was calculated relative to the XY plane and camera azimuth relative to the positive Y axis, e.g. Figure 7.1.

Table 7.1. Network camera calibration residuals and extrinsic parameters.

	<i>NCam 1</i>	<i>NCam 2</i>	<i>NCam 3</i>	<i>NCam 4</i>
<i>X (mm)</i>	0.01 ± 0.02	0.00 ± 0.02	0.00 ± 0.02	0.00 ± 0.02
<i>Y (mm)</i>	0.00 ± 0.02	0.00 ± 0.02	0.00 ± 0.02	0.00 ± 0.02
<i>Azimuth (°)</i>	-38.2	-2.4	38.9	89.4
<i>Elevation (°)</i>	24.8	15.2	19.8	15.1

7.3.3. Criterion data treatment

A second order, low-pass Butterworth bidirectional filter was applied to all three-dimensional marker coordinate (MAC) data using cut-off frequencies of 7 and 10 Hz for walking and running data respectively (O'Connor *et al.*, 2007; Queen, Gross and Liu, 2006). All MAC data were subsequently transformed into the common coordinate system to enable direct comparison, e.g. Figure 7.1. A method for determining foot-

surface contacts within MAC data was required. Stance phases were identified within three-dimensional marker trajectory data using the foot-velocity algorithm (O'Connor *et al.*, 2007). Figure 7.3 illustrates heel, 2nd and 5th meta-tarsal head (small circles) marker motion data during stance. Progression from heel-strike to toe-off events is indicated by colour, i.e. blue-to-red. O'Connor *et al.* (2007) defined the foot as the mid-point between the heel and 2nd meta-tarsal head markers.

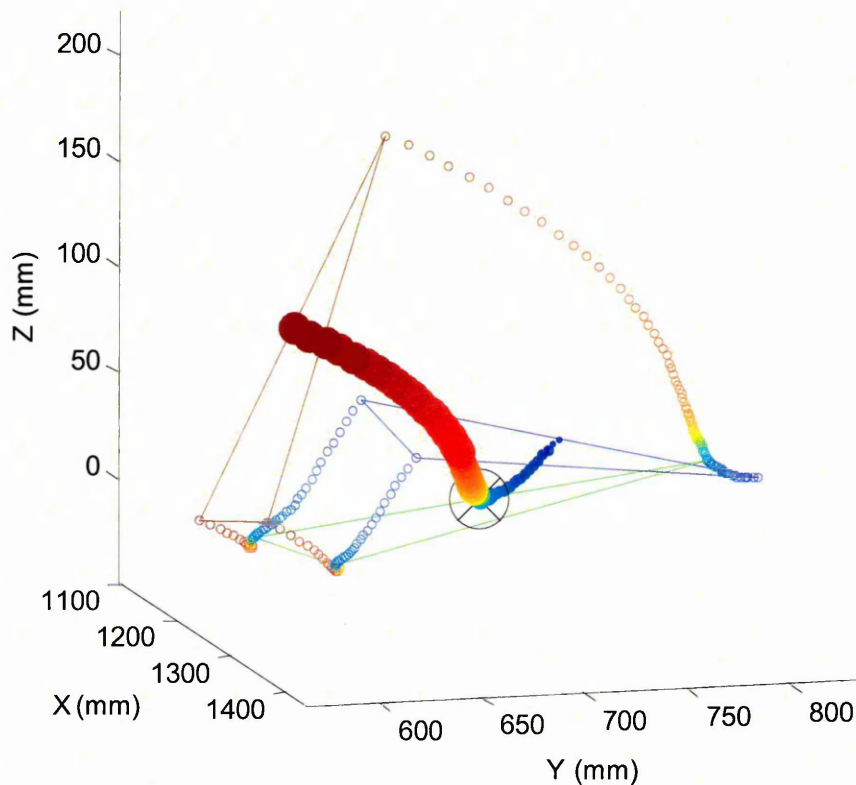


Figure 7.3. Foot marker motion between heel-strike (blue), foot-surface contact (green) and toe-off (red) for walking: progression is indicated by colour (blue-to-red).

The motion of the foot (heel and 2nd meta-tarsal mid-point) is illustrated in Figure 7.3 by the filled circles of increasing size; progression is indicated by colour. For individual stance phases, the time of foot-surface contact was identified using the modal vertical foot position, i.e. black crossed circle (Figure 7.3). Foot position (horizontal plane) and time were recorded as criterion data for foot-surface contacts, i.e. MAC_{XYT} . This was performed for every foot-surface contact inside the calibrated motion capture volume: Figure 7.4 illustrates an example of criterion foot-surface contact data for a walking sequence.

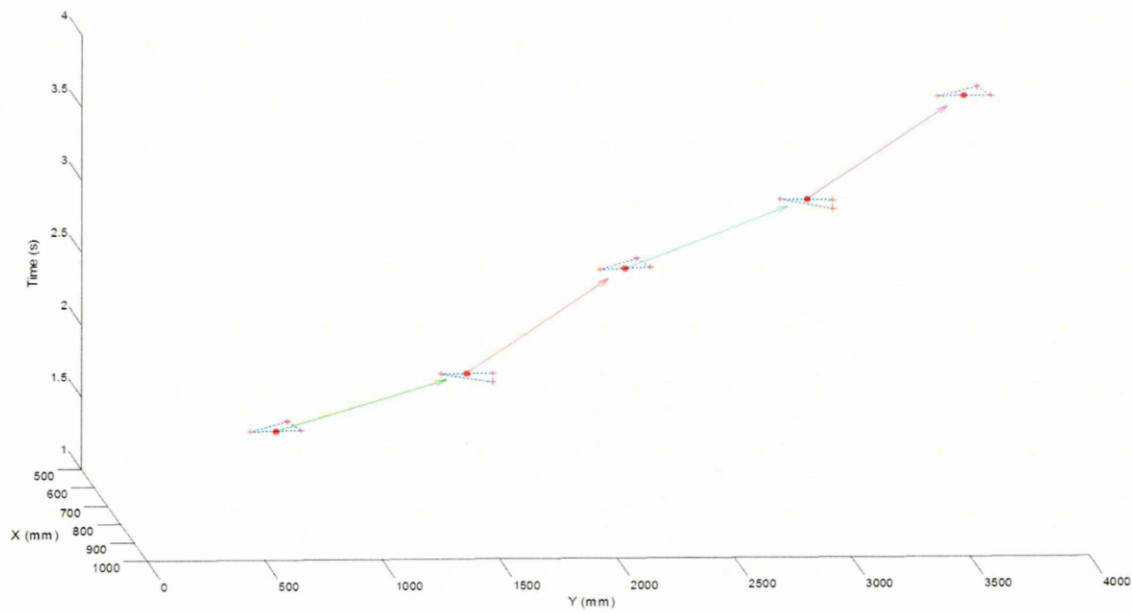


Figure 7.4. Example criterion foot-surface contact data (filled red circles). Foot markers (red pluses and blue dashed lines), resultant step length and step time (coloured arrows) are highlighted for illustration.

7.3.4. FSCi algorithm data treatment

All image coordinate data were measured automatically by the FSCi algorithm. The foot-mask threshold was normalised to 75%, e.g. equation 6.25 (Chapter 6), for all camera image sequences. Network camera image sequences were cropped to the first image where the LED box was illuminated. Foot-surface contact data were clustered in the common coordinate system due to different camera perspectives. Following the calculation of intrinsic and extrinsic camera parameters (described in section 7.3.2), the two-dimensional position of image coordinate data was reconstructed using the planar position reconstruction method, identified by equations 4.11 - 4.15 (Chapter 4).

For each trial, reconstructed data outside of the motion capture volume were discarded, e.g. Figure 7.1. Furthermore, data points considered to be outliers were manually removed. To cluster FSCi position data, adjacent-element distances for FSCi position data were computed. The number of steps for an image sequence was indicated by adjacent-element distances greater than 300 mm. Subsequently, an agglomerative clustering algorithm grouped foot-surface contact data based on the number of steps plus one, i.e. number of foot-surface contacts. To assess the performance of the clustering algorithm, the three-dimensional Delaunay Triangulation of MAC_{XYT} data was computed. Euclidean distances between FSCi position data and MAC_{XYT} data were

computed to identify whether clustering was correct. The median position and time for clustered FSCi data was deemed a foot-surface contact, i.e. $FSCi_{XYT}$.

7.3.5. Reconstruction plane elevation

Figure 6.26 (Chapter 6) illustrated that the FSCi algorithm reconstructs a location about the centre of the foot, i.e. out-of-plane coordinate. Chapter 5 highlighted the use of an elevated reconstruction plane to reduce the impact of out-of-plane error for player position estimates. Therefore, in addition to reconstructing position data using the reference plane, e.g. Figure 7.1, position data were also reconstructed using elevated reconstruction planes. In addition to the pilot study outlined in section 6.4.1 (Chapter 6), three-dimensional motion analysis was performed for walking trials as described above (sections 7.3.2 - 7.3.3). Furthermore, foot-surface contact data obtained by the FSCi algorithm were treated as described above (section 7.3.4). This was performed to assess position reconstruction accuracy in relation to reconstruction plane elevation for walking data.

To allow position reconstruction, image coordinates and extrinsic camera parameters must be calculated for an elevated reconstruction plane, e.g. section 5.4.1 (Chapter 5). Using this method, a reconstruction plane was calculated for 0 - 100 mm elevation (1 mm elevation increments) using the image locations of the four spherical markers defining the original reference plane. Following the reconstruction of $FSCi_{XYT}$ for each elevated reconstruction plane, differences between $FSCi_{XYT}$ and MAC_{XYT} data were calculated and root mean square error (RMSE: equation 7.4) computed. Reconstruction plane elevation that yielded the minimum RMSE was deemed the optimal reconstruction plane elevation. Figure 7.5 illustrates RMSE for a sample sequence of reconstructed foot-surface contacts: 26 mm (blue circle) was the optimum reconstruction plane elevation. For 24 walking trials filmed from four camera perspectives, the mean optimum reconstruction plane elevation was 29 and 35 mm (relative to the reference plane) for barefoot and shod walking respectively. Therefore for current work, position reconstruction for FSCi image coordinate data was performed using reconstruction planes elevated to 29 and 35 mm (relative to the reference plane) for barefoot and shod conditions respectively. Furthermore, optimum reconstruction plane elevation was evaluated for this work as described here (section 7.3.5), to identify suitable reconstruction plane elevations for both walking and running.

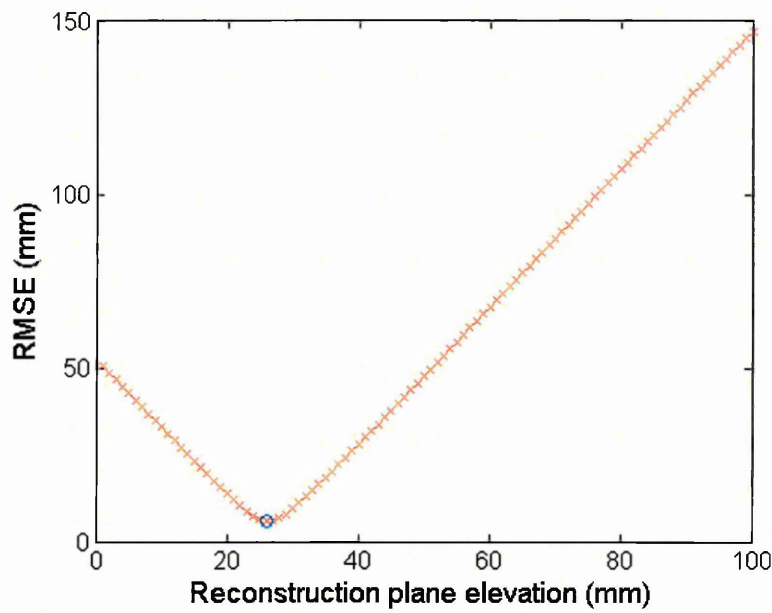


Figure 7.5. Foot-surface contact RMSE (single walking trial) using incrementally elevated reconstruction planes (red crosses).

7.3.6. Data analysis

For individual foot-surface contacts, the following dependent variables were quantified:

- Number of identified foot-surface contacts (n)
- Foot-surface contact position (mm)
- Foot-surface contact time (s)
- Foot-surface contact time relative to stance (%)
- Stance time (s)

Step length and step time was defined as the absolute difference between contralateral foot-surface contact location and time respectively. Direction of progression for walking and running trials is unlikely to be linear (Huxham *et al.*, 2006). Therefore spatial gait parameters such as step distance and step width, e.g. Figure 7.6, were computed using definitions identified by Huxham *et al.* (2006):

$$\text{Step distance} = \frac{b^2 + c^2 - a^2}{2c} \quad [7.1]$$

$$\text{Step width} = \sqrt{b^2 - \text{Step distance}^2} \quad [7.2]$$

$$\text{Step angle} = \frac{\text{atan}\left(\frac{\text{Step distance}}{\text{Step width}}\right)}{\pi} \times 180 \quad [7.3]$$

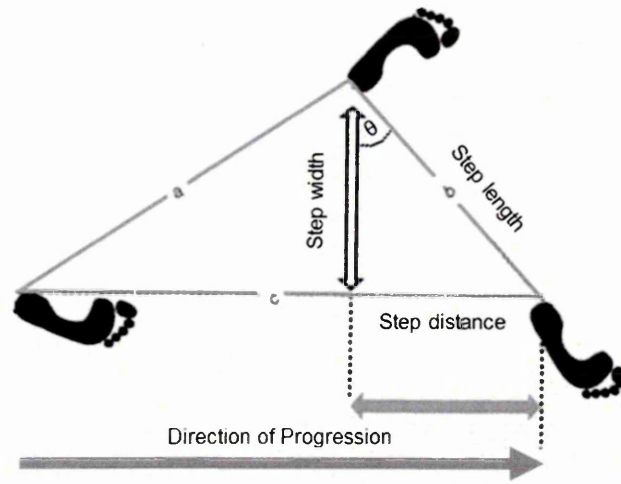


Figure 7.6. Spatial step components, i.e. step length, step width, step distance and step angle (θ), relative to direction of progression (adapted from Huxham *et al.*, 2006).

Step length and step time were calculated for foot-surface contacts where $n \geq 2$. Step distance, step width and step angle (equations 7.1 - 7.3) were calculated for foot-surface contacts where $n \geq 3$. For a sequence of foot-surface contacts, the following dependent variables were quantified:

- Step length (mm)
- Step time (s)
- Step distance (mm)
- Step width (mm)
- Step angle ($^{\circ}$)

Agreement was assessed using Bland and Altman 95% limits of agreement (LOA). In the case of heteroscedastic data distribution, i.e. $|r^2| > 0.1$, ratio LOA was also reported. Furthermore, root-mean square error (RMSE) was calculated with the following:

$$RMSE = \sqrt{\sum_{i=1}^N (X_{iR} - X_{ir})^2 / N} \quad [7.4]$$

where X_{iR} is the criterion, X_{ir} is the estimate and N is the number of data points.

7.4. Results

Data were obtained from six participants walking and running through a motion capture volume at a self-selected pace. Trials were repeated three times and performed in barefoot and trainers, yielding 72 motion trials that consisted of 312 individual foot-surface contacts. Four network camera perspectives meant that the FSCi algorithm analysed 288 image sequences containing a total of 1248 foot-surface contacts. Using a laptop computer (Processor: 2.3 GHz Intel Core i7; Memory: 8 GB RAM), total analysis time for a sample sequence of 123 images was 107.4 s. Analysis time per image was 0.87 ± 0.05 s. For a sample sequence, an analysis of the FSCi algorithm (Appendix 7.4) highlighted that the two most time consuming processes were the infinite skeletonisation of the lower-body (14.5% of analysis time), e.g. Figure 6.17 (B), and image colour space conversion (RGB to HSV: 14.4% of analysis time), e.g. Figure 6.14 (D). All other functions constituted approximately 5% or less of total analysis time.

Table 7.2. Step frequency and resultant velocity for walking and running tasks.

	<i>Step frequency</i>	<i>Resultant velocity</i>
	(Hz)	($m \cdot s^{-1}$)
	$\bar{x} \pm s$	$\bar{x} \pm s$
<i>Walking</i>	1.89 ± 0.11	1.42 ± 0.11
<i>Running</i>	2.69 ± 0.18	3.03 ± 0.52

Table 7.2 presents step frequency and sacrum marker velocity (coronal plane) within the motion capture volume for walking and running. Table 7.3 presents the proportion of accepted and rejected foot-surface contact candidates following the imposition of geometric rules, e.g. section 6.4.7. For both walking and running, the FSCi algorithm accepted a similar proportion of foot-surface contact candidates.

Table 7.3. Accepted and rejected foot-surface contact candidates.

	<i>Foot-surface contact candidates</i>	<i>Accepted</i>	<i>Rejected</i>
	<i>n (%)</i>	<i>n (%)</i>	<i>n (%)</i>
<i>Walking</i>	32710 (100%)	11945 (36.5%)	20765 (63.5%)
<i>Running</i>	17077 (100%)	5252 (30.8%)	11825 (69.2%)

The FSCi algorithm identified data for 1243 of 1248 (99.6%) foot-surface contacts. For 288 image sequences, step analysis was performed automatically for 263 image sequences (91.3%); 25 image sequences (8.7%) required manual intervention, e.g. Figure 7.7. For seven image sequences, a total of 93 data points were manually removed: these data represent 0.5% of data accepted by the FSCi algorithm (Table 7.3). For 20 image sequences, the correct number of foot-surface contacts was manually identified to cluster foot-surface contact data (two image sequences required both operations).

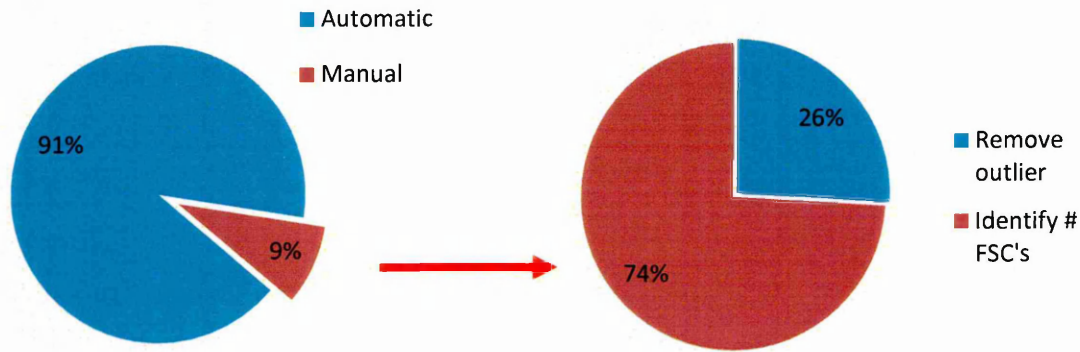


Figure 7.7. Proportion of image sequences requiring manual intervention (left) and manual intervention type (right).

7.4.1. Reconstruction plane elevation

Figure 7.8 presents LOA for all foot-surface contact data in the X direction using reference (A) and elevated (B) reconstruction planes. X direction differences for reference plane position reconstruction (A: Figure 7.8) were heteroscedastic ($r^2 = -0.25$), indicating that errors were related to the measurement, e.g. red regression line. X direction differences for elevated plane reconstruction (B: Figure 7.8) were homoscedastic ($r^2 = -0.07$), indicating that errors were not related to the measurement, e.g. red regression line. For X direction foot-surface contacts reconstructed using an elevated plane, LOA were -10.3 ± 54.7 mm (confidence intervals) and illustrate closer agreement than data reconstructed using the reference plane (A: Figure 7.8).

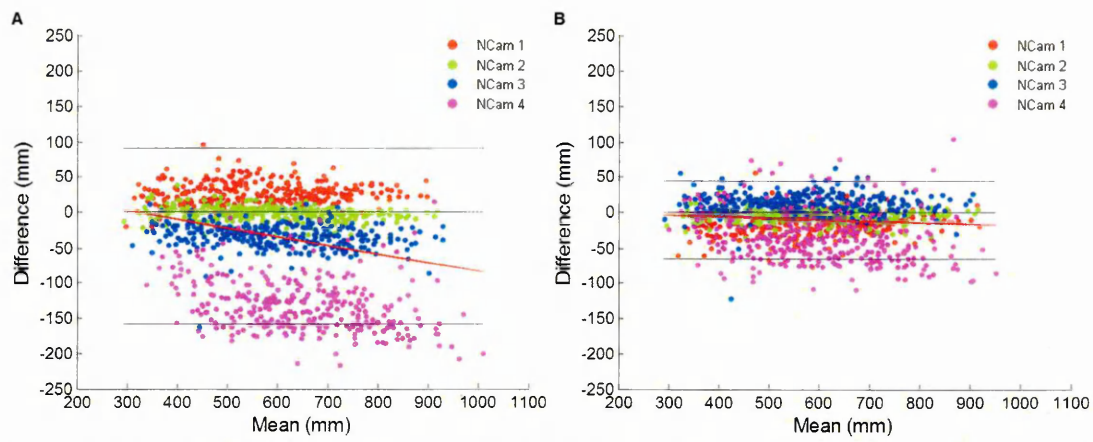


Figure 7.8. LOA (indicated by red and black lines) for all foot-surface contacts ($n = 1243$) in the X direction using reference (A) and elevated (B) reconstruction planes. Red, green, blue and magenta data identify network cameras 1 - 4 respectively.

Figure 7.9 presents LOA for all foot-surface contact data in the Y direction using reference (A) and elevated (B) reconstruction planes. Y direction differences for reference plane position reconstruction (A: Figure 7.9) were also heteroscedastic ($r^2 = -0.27$), as indicated by the red regression line. Y direction differences for elevated plane reconstruction (B: Figure 7.9) were homoscedastic ($r^2 = -0.00$), also indicated by the red regression line. For Y direction foot-surface contacts reconstructed using an elevated plane, LOA were -39.7 ± 106.1 mm (confidence intervals) and illustrate closer agreement than data reconstructed using the reference plane (A: Figure 7.9).

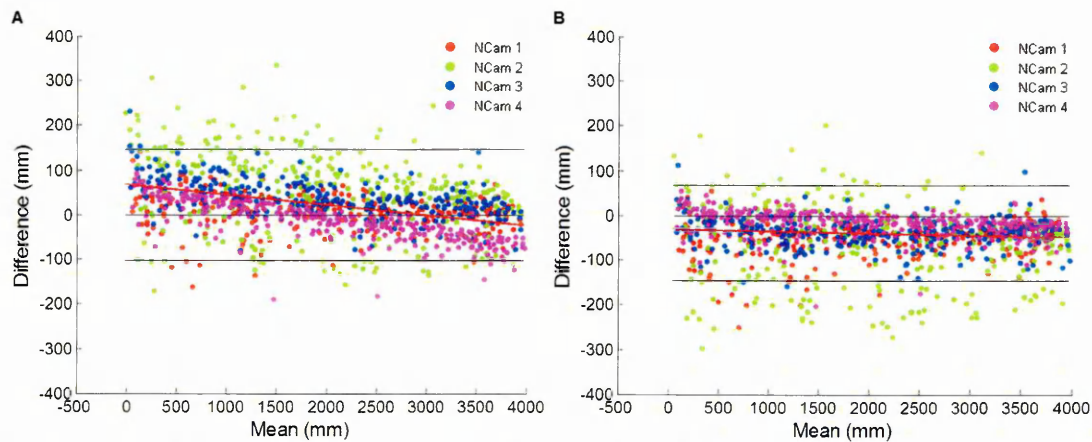


Figure 7.9. LOA (indicated by red and black lines) for all foot-surface contacts ($n = 1243$) in the Y direction for reference (A) and elevated (B) reconstruction planes. Red, green, blue and magenta data identify network cameras 1 - 4 respectively.

Table 7.4 presents RMSE for all foot-surface contacts in X , Y and resultant (R) directions using reference and elevated reconstruction planes. Table 7.4 highlights camera dependent position reconstruction: RMSE for frontal and oblique frontal camera perspectives (network cameras 1 – 3) was lowest in the X direction. RMSE for the sagittal camera perspective (network camera 4) was lowest in the Y direction. For all camera perspectives, i.e. view-independent, RMSE for foot-surface contacts was lower when using an elevated reconstruction plane. Therefore all $FSC_{i_{XYT}}$ data presented hereafter were reconstructed using reconstruction planes elevated to 29 and 35 mm (relative to reference plane) for barefoot and shod conditions respectively.

Table 7.4. RMSE (mm) for all foot-surface contact position data (n = 1243) in X , Y and R directions.

Network camera	Reference plane			Elevated plane		
	X	Y	R	X	Y	R
1	31.6	42.3	52.8	22.8	64.4	68.3
2	10.6	104.9	105.5	10.3	102.7	103.2
3	36.7	56.9	67.7	20.4	48.7	52.8
4	135.2	46.6	143.0	49.9	32.1	59.3
All	72.0	67.4	98.6	29.7	67.1	73.4

7.4.2. Foot-surface contacts

For barefoot and shod walking LOA indicated systematic differences for foot-surface contact position in both the X and Y directions, i.e. FSC_X and FSC_Y respectively (Table 7.5). Furthermore, LOA indicated larger agreement limits for foot-surface contact position in the Y direction (Table 7.5). FSC_X (shod walking) exhibited heteroscedasticity (Table 7.5); ratio LOA indicated that 95% of ratios were between 10.0% of the mean ratio. Positive r^2 indicated increasing estimate errors for increasing criterion measures. RMSE for FSC_X and FSC_Y was similar for both barefoot and shod walking (Table 7.5).

Systematic differences were indicated for the time of foot-surface contacts as well as the duration, i.e. FSC_T and FSC_{Stance} respectively (Table 7.5). Estimates for FSC_T exhibited heteroscedasticity; ratio LOA indicated that 95% of ratios were between 14.6 and 15.6% of the mean ratio for barefoot and shod walking respectively (Table 7.5). Negative r^2

indicated decreasing estimate errors for increasing criterion measures. RMSE for FSC_T was similar for barefoot and shod conditions (Table 7.5). Estimates for stance time, i.e. FSC_{Stance} , also exhibited heteroscedasticity; ratio LOA indicated that 95% of ratios were between 21.0 and 17.8% of the mean ratio for barefoot and shod walking respectively (Table 7.5). Again, negative r^2 indicated decreasing estimate errors for increasing criterion measures. RMSE for FSC_{Stance} was also similar for barefoot and shod conditions (Table 7.5).

Table 7.5. LOA (absolute and ratio), r^2 and RMSE for foot-surface contact (FSC) parameters during barefoot (n = 391) and shod (n = 362) walking.

	Network camera	Condition	Absolute LOA	r^2	Ratio LOA (dimensionless)	RMSE
FSC_X (mm)	All	Barefoot	-11.1 ± 53.7	0.09	-	29.6
		Shod	-14.2 ± 50.2	0.20	$0.98 (\times/\div 1.09)$	29.3
FSC_Y (mm)	All	Barefoot	-27.0 ± 65.4	-0.00	-	42.9
		Shod	-17.3 ± 78.4	-0.09	-	43.0
FSC_T (s)	All	Barefoot	-0.15 ± 0.11	-0.14	$0.93 (\times/\div 1.09)$	0.16
		Shod	-0.14 ± 0.20	-0.19	$0.94 (\times/\div 1.11)$	0.17
FSC_{Stance} (s)	All	Barefoot	0.35 ± 0.45	-0.86	$2.23 (\times/\div 2.82)$	0.34
		Shod	0.35 ± 0.44	-0.85	$2.25 (\times/\div 2.73)$	0.35

For barefoot and shod running, LOA indicated systematic differences for foot-surface contact position in both the X and Y directions (Table 7.6). RMSE for FSC_X and FSC_Y was similar for both barefoot and shod running (Table 7.6). For running, systematic differences for FSC_T and FSC_{Stance} were indicated by LOA (Table 7.6). For barefoot running, FSC_T exhibited heteroscedasticity; ratio LOA indicated that 95% of ratios were between 10.7% of the mean ratio (Table 7.6). Negative r^2 indicated decreasing estimate errors for increasing criterion measures. RMSE for FSC_T was different for barefoot and shod running (Table 7.6). Greater RMSE for FSC_T during barefoot running was the result of four outliers: their removal yields RMSE = 0.15 s and is comparable to RMSE for FSC_T during shod running. However, outlier removal did not affect data distribution; heteroscedasticity remained ($r^2 = 0.24$, ratio LOA were $0.91 \times/\div 1.08$) with 95% of ratios distributed between 15.7% of the mean ratio. Estimates for FSC_{Stance} also

exhibited heteroscedasticity; ratio LOA indicated that 95% of ratios were between 16.0 and 15.7% of the mean ratio (Table 7.6). Negative r^2 indicated decreasing estimate errors for increasing criterion measures. RMSE for FSC_{Stance} was similar between barefoot and shod running conditions (Table 7.6).

Table 7.6. LOA (absolute and ratio), r^2 and RMSE for foot-surface contact (FSC) parameters during barefoot (n = 251) and shod (n = 238) running.

	<i>Network</i>	<i>Condition</i>	<i>Absolute LOA</i>	r^2	<i>Ratio LOA</i>	<i>RMSE</i>
	<i>camera</i>		<i>(dimensionless)</i>			
$FSC_X (mm)$	All	Barefoot	-7.2 ± 61.0	-0.04	-	31.9
		Shod	-6.3 ± 54.3	0.03	-	28.3
$FSC_Y (mm)$	All	Barefoot	-62.1 ± 115.8	-0.05	-	85.7
		Shod	-71.1 ± 136.4	0.02	-	99.4
$FSC_T (s)$	All	Barefoot	-0.16 ± 0.25	-0.16	0.95 ($\times/\div 1.07$)	0.21
		Shod	-0.11 ± 0.06	0.08	-	0.12
$FSC_{Stance} (s)$	All	Barefoot	0.53 ± 0.46	-0.68	3.38 ($\times/\div 2.91$)	0.23
		Shod	0.52 ± 0.46	-0.61	3.34 ($\times/\div 2.89$)	0.24

In relation to individual stance phases, e.g. heel-strike and corresponding toe-off, the FSCi algorithm identified the time of foot-surface contact at $50.1 \pm 9.7\%$ and $64.6 \pm 28.5\%$ of stance for walking and running respectively. Furthermore, Figure 7.10 illustrates the proportion of stance identified by the FSCi algorithm (expressed as a percentage of stance). For walking (A: Figure 7.10), a bimodal distribution reflects self-occlusion, e.g. swing-foot. For walking, distribution peaks were observed at 25.5% and 67.1% of stance (A: Figure 7.10). For running, a positive skew indicates that a smaller proportion of stance was identified, reflecting shorter duration foot-surface contacts in relation to a fixed camera sample rate (B: Figure 7.10). The median proportion of stance identified for running was 26.7%.

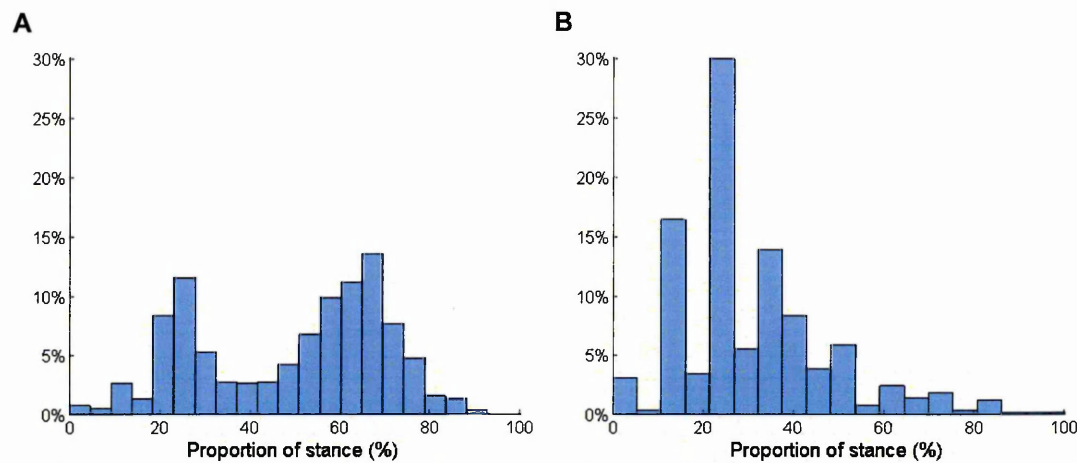


Figure 7.10. Proportion of stance identified for walking (A) and running (B).

Table 7.7 presents mean camera pixel-scale in X and Y directions for the motion capture plane. Camera perspective affects image resolution of a target, i.e. the foot. Cameras perpendicular to the X axis, e.g. NCam 2, had the lowest pixel-scale (highest image resolution) in the X direction (Table 7.7). Similarly, cameras perpendicular to the Y axis, e.g. NCam 4, had the lowest pixel-scale (highest image resolution) in the Y direction (Table 7.7).

Table 7.7. Mean camera pixel-scale (mm) for X and Y directions and schematic of camera locations.

	NCam 1	NCam 2	NCam 3	NCam 4	
X (mm)	3.9	5.8	14.1	16.3	
Y (mm)	16.6	20.1	12.9	6.0	

Figures 7.11 - 7.14 illustrate mean and standard deviation FSCi foot-surface contact location (green filled circles and standard error bars) relative to criterion location (blue filled circles). Position data are presented relative to a normalised foot outline that approximates foot size and orientation. The foot (large red triangle) is 100 mm in width (top) and 250 mm in length and centred about the criterion foot location (heel and 2nd meta-tarsal head midpoint). Images A - D represent network camera perspectives 1 - 4 respectively: the schematic of camera locations (right) highlights the pixel-scale of images, e.g. Table 7.7, used to identify foot-surface contacts.

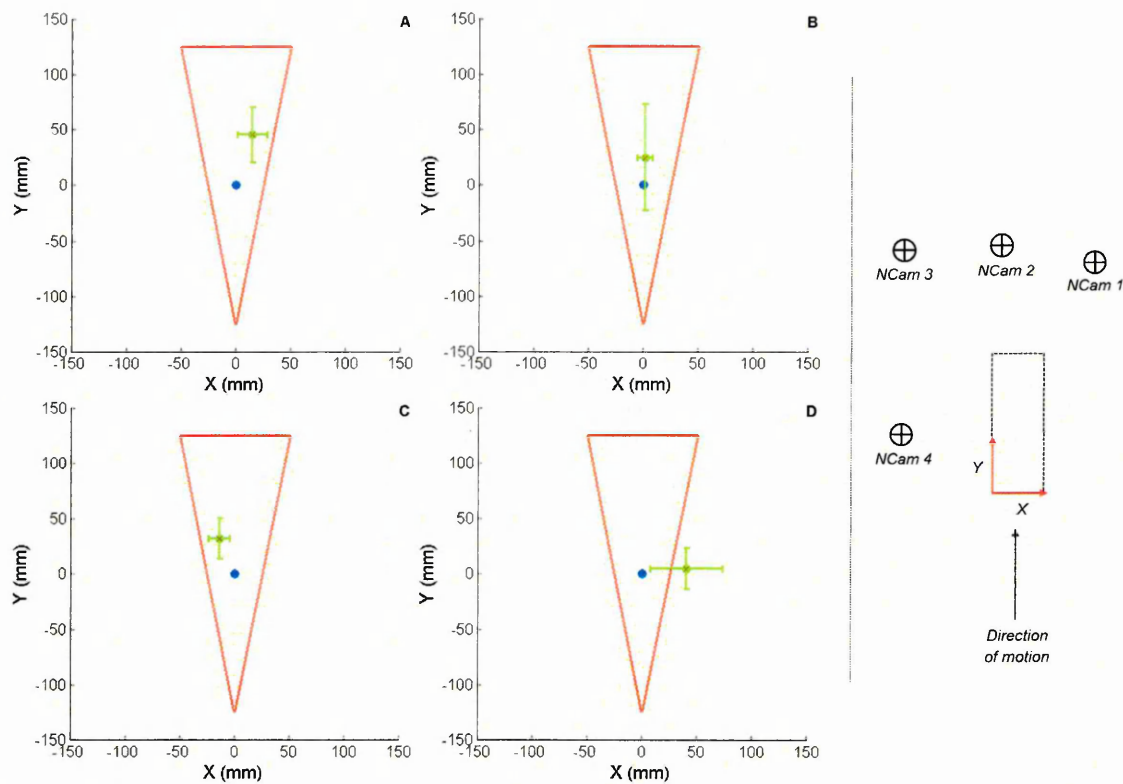


Figure 7.11. Mean (green circles) and standard deviation (green error bars) foot contact locations relative to criterion locations (blue circles) for barefoot walking measured by cameras 1 - 4 (A - D respectively).

Mean and standard deviation foot-surface contacts ($n = 391$) identified during barefoot walking from different camera perspectives illustrate view-dependent position reconstruction (Figure 7.11). Variation in position reconstruction reflects camera perspective of the motion capture plane, i.e. camera locations (right). For barefoot walking, mean foot-surface contact locations were the equivalent of 17.8, 10.4, 4.5 and 11.1 pixels (resultant direction) from criterion locations for cameras 1 - 4 respectively.

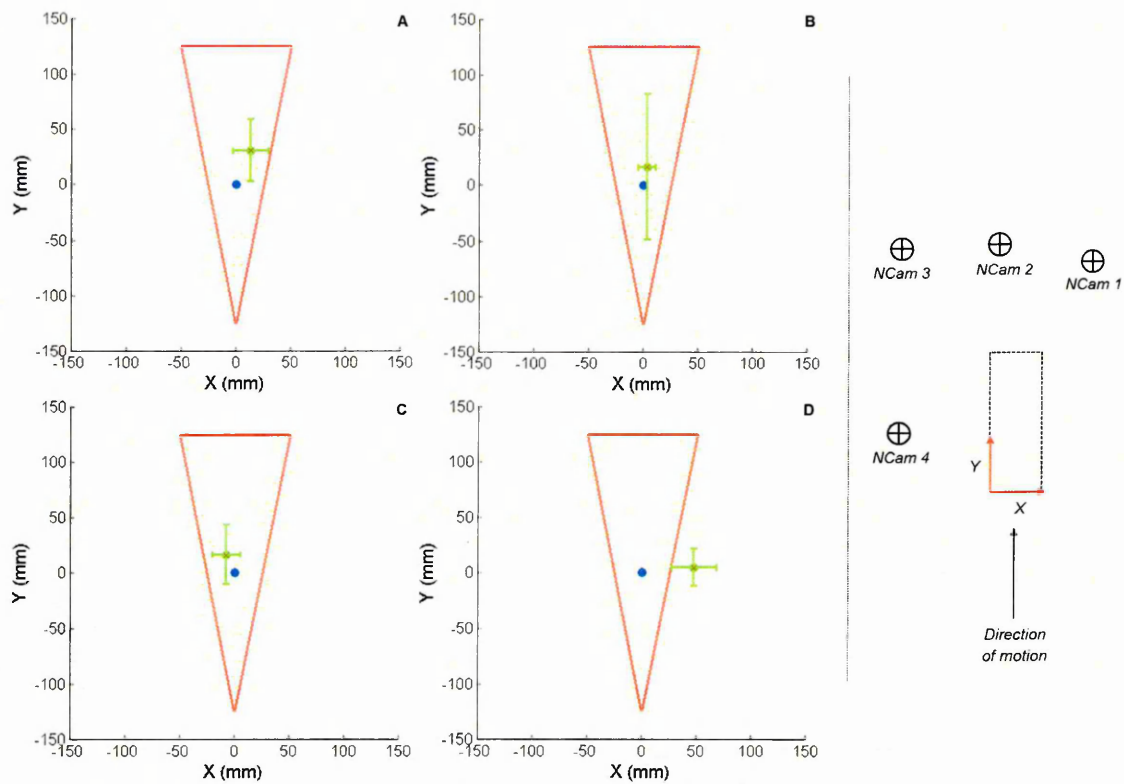


Figure 7.12. Mean (green circles) and standard deviation (green error bars) foot contact locations relative to criterion locations (blue circles) for shod walking measured by cameras 1 - 4 (A - D respectively).

Figure 7.12 illustrates mean and standard deviation foot-surface contacts ($n = 362$) for shod walking. Reconstructed foot-surface contact locations were similar when comparing barefoot and shod walking (Figure 7.11). For shod walking, mean foot-surface contact locations were the equivalent of 16.1, 9.9, 4.0 and 11.7 pixels (resultant direction) from criterion locations for cameras 1 - 4 respectively.

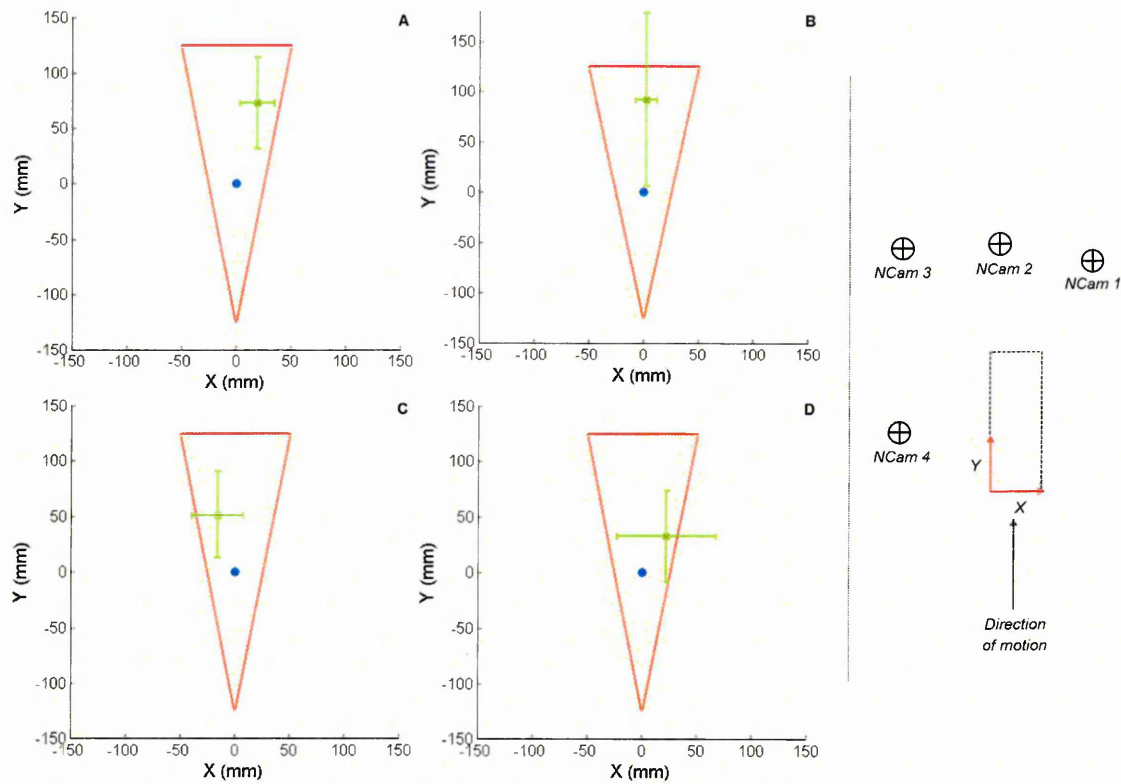


Figure 7.13. Mean (green circles) and standard deviation (green error bars) foot contact locations relative to criterion locations (blue circles) for barefoot running measured by cameras 1 - 4 (A - D respectively).

Figure 7.13 illustrates mean and standard deviation foot-surface contacts ($n = 251$) identified during barefoot running. The Y axis for image B has been rescaled to illustrate data correctly. Reconstructed foot-surface contact locations illustrate greater position reconstruction error along the Y axis (direction of motion). Furthermore, FSCi estimates progressed toward the forefoot during barefoot running. For barefoot running, mean foot-surface contact locations were the equivalent of 20.5, 15.5, 5.1 and 10.0 pixels (resultant direction) from criterion locations for cameras 1 - 4 respectively.

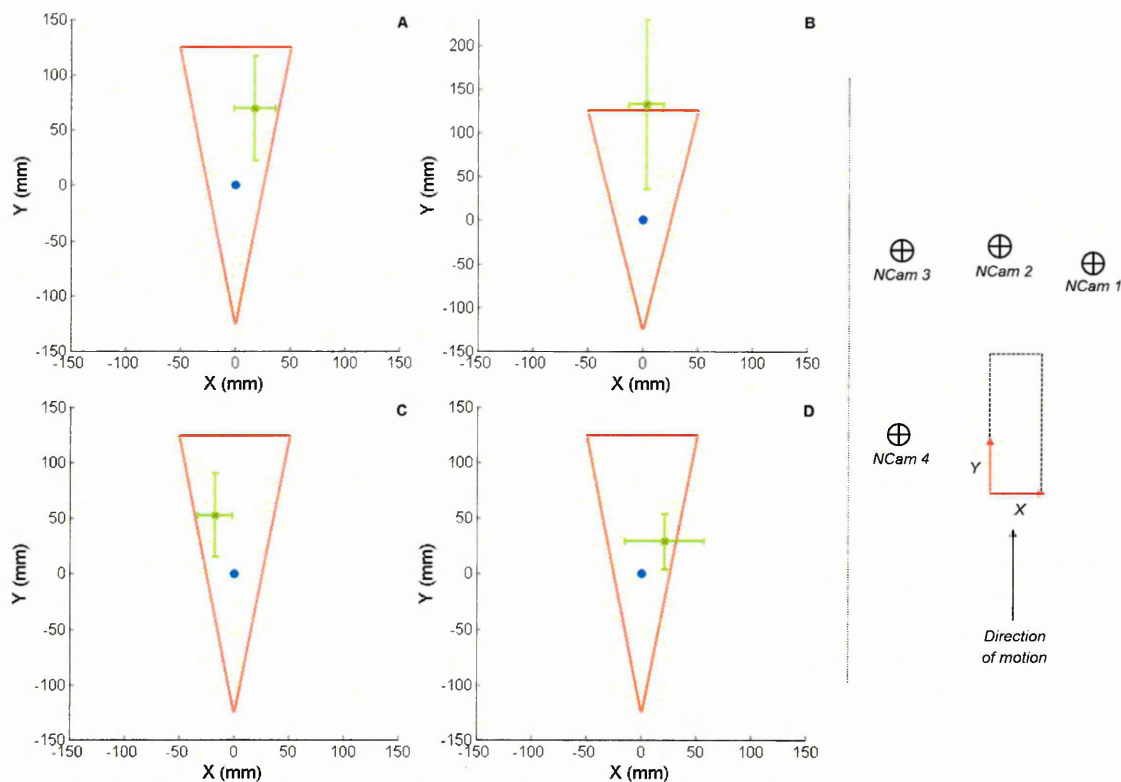


Figure 7.14. Mean (green circles) and standard deviation (green error bars) foot contact locations relative to criterion locations (blue circles) for shod running measured by cameras 1 - 4 (A - D respectively).

Figure 7.14 illustrates mean and standard deviation foot-surface contacts ($n = 238$) identified during shod running. The Y axis for image B has also been rescaled to illustrate data correctly. Reconstructed foot-surface contact locations were similar to barefoot running (Figure 7.13) with the exception of image B: camera 2 exhibited greater position reconstruction error than barefoot running along the Y axis (direction of motion). For shod running, mean foot-surface contact locations were the equivalent of 20.5, 18.7, 5.1 and 9.8 pixels (resultant direction) from criterion locations for cameras 1 - 4 respectively.

7.4.3. Step parameters and reconstruction plane elevation

LOA indicated that shod walking step length estimates were systematically shorter than criterion data by -7.1 mm (Table 7.8). Barefoot walking step length differences were heteroscedastic; ratio LOA indicated that 95% of ratios were between 10.8% of the mean ratio (Table 7.8). Heteroscedastic step length data for barefoot walking was the result of a single outlier: its removal yields $r^2 = 0.04$. Step time estimates for both barefoot and shod walking were similar: a small systematic difference of 0.01 s was

reported (Table 7.8). However, 95% of walking step time estimates were between ± 0.13 and ± 0.22 s of mean step time for barefoot and shod walking respectively. This indicated variation in step time estimates and was reflected by RMSE (Table 7.8).

LOA indicated systematically shorter step distances (-3.7 mm) for shod walking (Table 7.8). Barefoot walking step distance differences were heteroscedastic. Ratio LOA indicated that 95% of ratios were between 9.7% of the mean ratio (Table 7.8). Positive r^2 indicated increasing estimate errors for increasing criterion measures. Step width and step angle differences (barefoot and shod conditions) were also heteroscedastic (Table 7.8). Ratio LOA indicated that 95% of step width ratios were between 30.6 and 26.5% of mean ratios for barefoot and shod conditions respectively. Furthermore, ratio LOA indicated that 95% of step angle ratios were between 2.9 and 1.7% of mean ratios (barefoot and shod respectively).

Table 7.8. LOA (absolute and ratio), r^2 and RMSE for walking step parameters.

	Network camera	Condition	Absolute LOA	r^2	Ratio LOA (dimensionless)	RMSE
Step Length (mm)	All	Barefoot (n = 319)	-3.1 ± 80.4	0.13	1.00 ($\times/\div 1.12$)	41.1
		Shod (n = 291)	-7.1 ± 93.9	0.07	-	48.3
Step Time (s)	All	Barefoot (n = 319)	0.00 ± 0.13	-0.04	-	0.07
		Shod (n = 291)	0.01 ± 0.22	-0.01	-	0.11
Step Distance (mm)	All	Barefoot (n = 247)	-2.5 ± 75.1	0.32	0.99 ($\times/\div 1.10$)	38.3
		Shod (n = 219)	-3.7 ± 91.0	0.04	-	46.4
Step Width (mm)	All	Barefoot (n = 247)	-8.7 ± 38.0	0.32	0.94 ($\times/\div 1.35$)	21.2
		Shod (n = 219)	-12.5 ± 28.8	0.32	0.91 ($\times/\div 1.24$)	19.2
Step Angle (°)	All	Barefoot (n = 247)	0.6 ± 3.0	-0.24	1.01 ($\times/\div 1.04$)	1.6
		Shod (n = 219)	0.9 ± 2.3	-0.35	1.01 ($\times/\div 1.02$)	1.5

For running, step length estimates were similar for both barefoot and shod conditions (Table 7.9). LOA indicated that 95% of estimates were between ± 124.0 and ± 154.8 mm for barefoot and shod conditions respectively: greater variation for shod step length was reflected by RMSE (Table 7.9). Step time estimates for both barefoot and shod running were similar. No systematic differences were indicated for step time for either barefoot or shod running (Table 7.9).

Table 7.9. LOA (absolute and ratio), r^2 and RMSE for running step parameters.

	Network camera	Condition	Absolute LOA	r^2	Ratio LOA (dimensionless)	RMSE
Step Length (mm)	All	Barefoot (n = 179)	1.3 ± 124.0	0.05	-	63.1
		Shod (n = 165)	0.1 ± 154.8	-0.02	-	78.7
Step Time (s)	All	Barefoot (n = 179)	0.00 ± 0.07	0.00	-	0.04
		Shod (n = 165)	0.00 ± 0.07	0.05	-	0.04
Step Distance (mm)	All	Barefoot (n = 108)	-0.4 ± 111.9	-0.04	-	56.8
		Shod (n = 95)	5.0 ± 154.7	-0.03	-	78.6
Step Width (mm)	All	Barefoot (n = 108)	5.6 ± 58.2	-0.09	-	30.1
		Shod (n = 95)	-8.7 ± 54.3	0.21	0.95 (×/÷ 1.96)	28.9
Step Angle (°)	All	Barefoot (n = 108)	-0.4 ± 3.4	-0.03	-	1.8
		Shod (n = 95)	0.5 ± 3.1	-0.25	1.00 (×/÷ 1.03)	1.7

For running, LOA indicated similar barefoot step distances (-0.4 mm) and systematically longer (5.0 mm) shod step distances. LOA also indicated systematically longer step widths (5.6 mm) as well as similar step angles (-0.4°) for barefoot running. Step width and step angle differences for shod running were heteroscedastic. Ratio LOA indicated that 95% of step width and step angle ratios were between 51.7% and 3.1% of the mean ratio respectively.

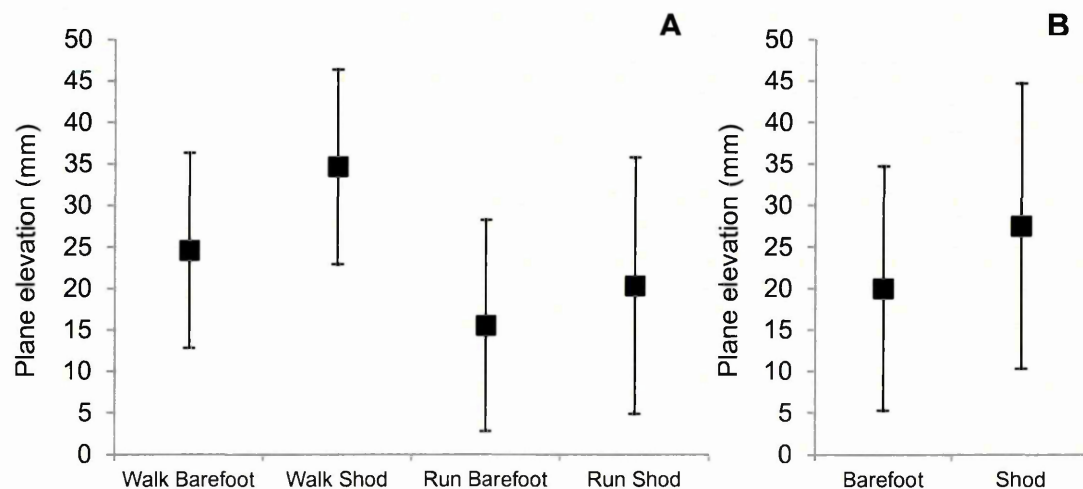


Figure 7.15. Optimal reconstruction plane elevation for task and condition (A) and condition (B) relative to the reference plane for all network cameras.

Figure 7.15 presents mean and standard deviation reconstruction plane elevation (all camera perspectives) that yields the minimum position reconstruction error, i.e. optimal elevation. Figure 7.15 (A) illustrates lower optimal reconstruction plane elevation for running when compared to walking. Furthermore, barefoot conditions illustrate lower optimal reconstruction plane elevation when compared to shod conditions. When optimal plane elevations are grouped across activity, i.e. walking and running (B; Figure 7.15), barefoot conditions require lower reconstruction plane elevation when compared to shod conditions. Optimal reconstruction plane elevation was 20.0 ± 14.7 mm and 27.5 ± 17.2 mm for barefoot and shod conditions respectively.

7.5. Discussion

7.5.1. Analysis time and identification rate

The FSCi algorithm was passed 288 image sequences containing a total of 1248 foot-surface contacts. The FSCi algorithm processed images automatically; data was identified for 1243 (99.6%) walking and running foot-surface contacts from all camera perspectives. Analysis time for a sample image sequence was 0.87 ± 0.05 s per image. Existing single camera heel strike identification algorithms, i.e. Bouchrika and Nixon (2006) and Jung and Nixon (2013), do not present analysis times. This is most likely due to the cumulative nature of algorithms. However Goffredo *et al.* (2010) presented analysis times for a single camera method of extracting lower-limb joint angles. Goffredo *et al.* (2010) reported analysis times of 1.44, 0.97 and 0.37 s per image for image resolutions of 500×490 , 250×245 and 163×163 pixels respectively. The current algorithm processes high-resolution images, i.e. 1280×720 pixels. Analysis times were considered to be acceptable, given the increased memory cost of high-resolution images; however it is inappropriate to make direct comparisons due to different algorithm functions.

An analysis of the FSCi algorithm highlighted that image skeletonisation e.g. Figure 6.17 (B), and image colour space conversion (RGB to HSV), e.g. Figure 6.14 (D), were the most time consuming processes of the algorithm. Furthermore, these standard MATLAB image processes were nearly three times as time consuming as other image processes. Programming language can have a significant impact on analysis time. Matuska, Hudec and Benco (2012) compared image-processing algorithms executed in MATLAB and OpenCV (Open Source Computer Vision Library written in C++).

OpenCV was reported to perform general image processing operations between 4 - 30 times faster than MATLAB. Programming language therefore represents a delimiting factor for FSCi algorithm analysis time.

Many single camera gait analysis algorithms only consider sagittal or oblique sagittal camera perspectives due to resolution changes in frontal perspective walking (Jung and Nixon, 2013). For the FSCi algorithm, data was identified for 99.6% of foot-surface contacts from sagittal, frontal and two oblique frontal camera perspectives of both walking and running. Unidentified foot-surface contacts were predominantly from frontal perspective image sequences: four foot-surface contacts were not identified within frontal perspective image sequences of walking and running. A single foot-surface contact was not identified within a sagittal perspective image sequence (running).

The FSCi algorithm automatically removed false candidates from each image via a self-determined threshold and three geometric rules, e.g. equation 6.32 (Chapter 6). Table 7.3 highlights the acceptance and rejection rate for foot-surface contact candidates. Step analysis was performed automatically for 263 image sequences (91.3%). However 25 image sequences (8.7%) required manual intervention for step analysis to be performed. In total, 93 foot-surface contact data points (81 for walking trials and 12 for running trials) were manually removed. This represents 0.5% of image data accepted by the FSCi algorithm (Table 7.3). Data were predominately removed from frontal perspective image sequences: a single data point was removed from data identified within a sagittal perspective image sequence. Furthermore, the correct number of foot-surface contacts was manually identified for 16 frontal perspective image sequences to enable data clustering. This operation was also required for a further three oblique frontal perspective image sequences. Only a single sagittal perspective image sequence required the correct number of foot-surface contacts to be manually identified.

Unidentified foot-surface contacts and manual interventions for frontal perspective image data reflect the challenges of identifying foot-surface contacts from frontal perspective images (Jung and Nixon, 2013). Furthermore, manually identifying foot-surface contacts highlights a limitation of the current, spatial clustering algorithm. The spatial clustering of foot-surface contact data was necessary to calculate $FSCi_{XYT}$ and

step parameters. It is well known that stance times decrease as gait speeds increase (Mann and Hagy, 1980). The number of data points that comprise a foot-surface contact (measured by the FSCi algorithm) will be related to the duration of stance and could be numerous or even singular. Data clustering involves the statistical classification of multivariate data (Jain, 2010); to this end it is difficult to train a clustering algorithm to group true-positive foot-surface contact data for multiple gait modes.

For current data, the spatial progression of walking and running enabled the step analysis of foot-surface contact data for the majority of cases (91.3%). The FSCi algorithm identifies data for independent feet. This allows foot-surface contact identification irrespective of gait mode, e.g. walking or running. However, foot-surface contacts can exist at two locations at the same time instant, e.g. dual-stance in walking or tennis split-step. Furthermore, foot-surface contacts can occur in the same location at different time instants, i.e. repeated movements about the same location. This can cause the current spatial clustering algorithm to yield an incorrect number of foot-surface contacts. The inclusion of time domain information might help differentiate foot-surface contact data. A time-windowed clustering algorithm would require the identification of relevant clustering parameters however current knowledge of tennis player-surface interaction is limited (Miller, 2006). Further research would be required to identify appropriate clustering parameters. Therefore automatic step analysis for multiple gait modes, i.e. match-play tennis, will currently be limited. Manual confirmation of data clustering will likely be required for step analysis of multi-model gait.

7.5.2. *Foot-surface contact time*

Identifying stance time from marker trajectory data will incur measurement error. The foot-velocity algorithm presented by O'Connor *et al.* (2007) was adopted due to accuracy improvements over existing methods. For walking, O'Connor *et al.* (2007) reported event detection errors of 0.016 ± 0.015 and 0.009 ± 0.015 s for heel-strike and toe-off events respectively (compared to force platform measurements). The current study also applied the foot-velocity algorithm to running data; however the algorithm had only been validated for walking (O'Connor *et al.*, 2007). Maiwald *et al.* (2009) presented a kinematic gait event detection algorithm for running. The approach presented by Maiwald *et al.* (2009) was not adopted because the approach used a different marker set and conversely, had not been validated for walking. Furthermore,

Maiwald *et al.*'s (2009) approach used marker acceleration data; Tirosh and Sparrow (2003) highlighted that gait events identified using acceleration data are highly sensitive to filter cut-off frequency. However, Maiwald *et al.* (2009) did report event detection errors for O'Connor *et al.*'s (2007) foot-velocity algorithm during running: heel-strike and toe-off errors were 0.004 ± 0.008 and 0.064 ± 0.027 s respectively for running speeds of $3.5 \text{ m}\cdot\text{s}^{-1}$.

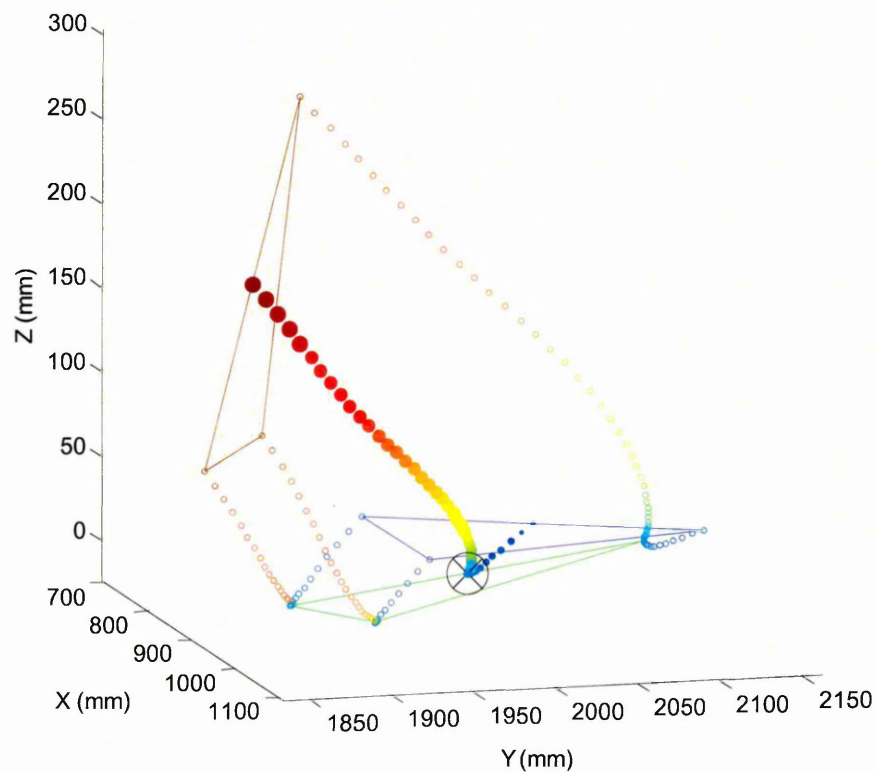


Figure 7.16 Foot marker motion between heel-strike (blue), foot-surface contact (green) and toe-off (red) for running: progression is indicated by colour (blue-to-red).

Running speeds for the current study were $3.03 \pm 0.52 \text{ m}\cdot\text{s}^{-1}$ (Table 7.1). Figure 7.16 illustrates the application of the foot-velocity algorithm to running marker trajectory data; the time of toe-off was clearly overestimated in this case. The current study does not support the quantification of kinematic gait event detection errors. Based on data presented by Maiwald *et al.* (2009), criterion stance time estimates were likely to be overestimated by approximately 0.06 s. However this error was acceptable because no other method was available to assess multiple footfalls.

Tables 7.5 and 7.6 highlight that stance time estimates (FSC_{Stance}) for walking and running were heteroscedastic. The strong negative relationships reflect that stance time estimate errors were greater for shorter duration foot-surface contacts. This reflects the application of a self-determined threshold to identify stance: heel-strike and toe-off events were not identified. The relationship between stance time duration and estimate error, for both walking and running (Tables 7.5 and 7.6 respectively), indicate that stance time is not a suitable parameter for analysis.

Tables 7.5 and 7.6 highlight systematic differences for the time of foot-surface contacts (FSC_T). Furthermore, heteroscedasticity was highlighted for all FSC_T estimates with the exception of shod running. Systematic error will be incurred by event synchronising data: synchronisation will only be accurate to within one image, i.e. 0.04 s. Therefore it is difficult to appropriately interpret FSC_T data presented in Tables 7.5 and 7.6. $FSCi_{XYT}$ was defined as the median of grouped data. Assuming normally distributed stance time data, the median will identify mid-stance, i.e. 50%. For walking, the time of foot-surface contact was identified at $50.1 \pm 9.7\%$ into stance. Therefore the FSCi algorithm was able to correctly identify mid-stance during walking. For running, the time of foot-surface contact was identified at $64.6 \pm 28.5\%$ into stance. Furthermore, overestimated criterion stance times during running indicate that FSCi estimates for the time of foot-surface contact would occur slightly later into stance.

Figure 7.10 illustrates the proportion of stance identified by the FSCi algorithm. Swing-foot occlusion during walking was evidenced by a bimodal distribution: occluded foot-surface contacts yield a smaller proportion of the stance phase (A: Figure 7.10). However, 90% of walking stance data was identified between 18.1 - 76.5% of stance, reflecting that the FSCi algorithm identifies a relatively central proportion of walking stance phases. For running, the proportion of stance identified by the FSCi algorithm was skewed (median = 26.7%). Furthermore, 90% of running stance data was identified between 11.8 - 66.8% of stance, highlighting that the FSCi algorithm did not identify a central proportion of running stance phases. This reflects the application of a self-determined threshold to stance phases where greater foot motion exists, i.e. greater plantar-flexion preceding running toe-off compared to walking (Novacheck, 1998). The foot-mask threshold factor (equation 6.25, Chapter 6) was 75% for current walking and running data. The threshold factor was determined experimentally and as such, might

not be the most suitable for running data. Reducing the threshold factor would cause the FSCi algorithm to accept more foot-surface contact candidates and might identify a central proportion of running stance phases. However reducing the threshold factor might also increase the rate of false positive foot-surface contact identification.

Existing single camera methods for identifying heel-strikes in walking do not report temporal parameters of gait, e.g. Bouchrika and Nixon (2006); Jung and Nixon (2013). Current data demonstrate that stance time estimates were not appropriate for analysis because the FSCi algorithm does not identify heel-strike and toe-off events. However for walking, the FSCi algorithm correctly identified mid-stance; temporal characteristics of walking gait, i.e. step time, can thus be quantified. For running, the FSCi algorithm identified the time of foot-surface contact after mid-stance and highlights a limitation of the FSCi algorithm when applied to short duration stance phases.

7.5.3. Foot-surface contact position

The FSCi algorithm identifies image coordinates about the foot centre; image coordinates are therefore out-of-plane. Figures 7.8 and 7.9 illustrate the effect of reconstructing the same image coordinate data using reference (A) and elevated (B) reconstruction planes for X and Y direction foot-surface contact data respectively. Out-of-plane error affects position reconstruction: reference plane position differences were heteroscedastic (error magnitude related to position). When image coordinate data were reconstructed with an elevated reconstruction plane, position differences were homoscedastic. This indicates that, when an elevated reconstruction plane was used, position error magnitudes were independent of foot-surface contact location. Furthermore, RMSE was lower for elevated plane position reconstruction when all camera perspectives are considered (Table 7.4). Two-dimensional analyses assume coplanarity. Out-of-plane error will still be present in data reconstructed using an elevated plane due to foot vertical motion in relation to a horizontal reconstruction plane. However, data homoscedasticity indicate that it is necessary to reconstruct FSCi image data with an elevated reconstruction plane.

Current work is concerned with the view-independent identification of foot-surface contacts. When considering all camera perspectives, systematic differences for X and Y direction foot-surface contacts, i.e. FSC_X and FSC_Y , were identified for both walking

and running (Tables 7.5 and 7.6 respectively). Systematic differences indicate that the FSCi algorithm did not identify the criterion foot-surface contact location, i.e. mid-point between heel and 2nd meta-tarsal head. Table 7.10 summarises foot-surface contact RMSE in the resultant direction, i.e. resultant of FSC_X and FSC_Y , for walking and running.

Table 7.10. Resultant direction RMSE (mm) for foot-surface contact parameters.

	<i>Walking</i>	<i>Running</i>
	<i>(n)</i>	<i>(n)</i>
<i>Barefoot</i>	52.1 (391)	91.4 (251)
<i>Shod</i>	52.2 (362)	103.4 (238)

Resultant direction RMSE was greater for running when compared to walking (Table 7.10). For running trials, differences for foot-surface contact position were predominantly observed in the *Y* direction (Table 7.6). Furthermore, variation for foot-surface contact position was also greater in the *Y* direction for running trials (Table 7.6). Foot-surface contact position estimates were located closer to the forefoot during running when compared to walking, e.g. Figures 7.11 - 7.14. This reflects the forefoot push off phase as the lowest velocity phase during centre-of-pressure progression in running (De Cock *et al.*, 2008). For an image sequence, this would manifest as low inter-frame differences about the forefoot. The FSCi algorithm would thus be sensitive to changes in foot contact type however criterion data would not.

Shod running foot-surface contacts were located closer to the forefoot when compared to barefoot running (Table 7.6). Furthermore, resultant direction RMSE was 12 mm greater for shod running when compared to barefoot running (Table 7.10). It is unlikely that participants used the same foot contact type during shod and barefoot running (De Wit, De Clercq and Aerts, 2000). Current data do not support the analysis of foot contact type; therefore differences for running FSCi position estimates relative to foot contact type cannot be confirmed. Future assessments should therefore consider centre-of-pressure as a criterion measure.

Camera perspective affected position reconstruction accuracy (Table 7.4). For example, lower RMSE in X and Y directions for network cameras two and four respectively, reflect frontal and sagittal perspectives of participants. Figures 7.11 - 7.14 illustrate the mean and standard deviation position for identified foot-surface contacts for barefoot and shod, walking and running (images A - D represent network camera perspectives 1 - 4 respectively). Perspective dependent estimates reflect individual camera locations relative to the motion capture volume, i.e. Figure 7.1. Furthermore, Figures 7.11 - 7.14 illustrate that variation in foot-surface contact estimates, i.e. standard error bars, was greatest in directions that image resolution was lowest, i.e. pixel-scale (Table 7.7). Position errors highlighted by Figures 7.11 - 7.14 indicate that camera placement affects the accuracy of foot-surface contact estimates and should be considered in future applications.

Camera perspective affects gait feature extraction. Bouchrika *et al.* (2009) reported that gait recognition rates fell from 95.8% (training dataset) to 64.5% when camera views ranging from 36 to 126° were introduced into analyses. For heel-strike extraction, Jung and Nixon (2013) identified 94% of heel-strikes from straight (frontal) and random direction camera perspectives of walking. Jung and Nixon (2013) reported heel-strike identification errors of ± 100 mm. However, identification error was only quantified in real world units for straight line (frontal) walking in a laboratory setting. The FSCi algorithm was applied to four camera perspectives, ranging through 127.6°, i.e. 38.2 to -89.4° (Table 7.1). For 95% of walking foot-surface contact estimates, the resultant of FSC_X and FSC_Y yields ± 84.6 and ± 93.1 mm (barefoot and shod respectively). Furthermore, for 95% of running foot-surface contact estimates, the resultant of FSC_X and FSC_Y yields ± 130.9 and ± 146.8 mm (barefoot and shod respectively). Therefore current foot-surface contact position errors for walking were lower than position errors previously reported by Jung and Nixon (2013). Furthermore, although position errors for running were larger than position errors reported in walking (Jung and Nixon, 2013), no previous research has described the automatic identification of foot-surface contacts in running. Therefore, with regard to acceptance criteria (identified in section 7.1), foot-surface contact data identified by the FSCi algorithm were considered to be acceptable.

7.5.4. Step parameters

Current single camera algorithms that report heel-strike position do not calculate step parameters, e.g. Bouchrika and Nixon (2007); Jung and Nixon (2013). For current data, step analysis was performed automatically for 91.3% of image sequences. Heteroscedastic step length estimates for barefoot walking (Table 7.8) were the result of a single outlier: its removal yields $r^2 = 0.04$. Current data indicated that the maximum random error component for step length estimates was 101.1 and 154.9 mm for walking and running respectively. Furthermore, the maximum random error component for walking and running step time estimates was 0.23 and 0.07 s respectively. When validating the GAITRite® walkway system (GAITRite Gold, CIR Systems, PA, USA), Webster, Wittwer and Feller (2005) reported maximum random error components of 25.1 mm and 0.04 s for walking step length and step time respectively. For estimating step length and step time, current data do not support the FSCi algorithm as a physical walkway replacement. However, for the context of this work, the FSCi algorithm measured basic gait parameters of walking and running without interfering with the activity being observed. Accordingly, the FSCi algorithm could be used to measure basic gait parameters such as step length and step time in match-play tennis.

Owings and Grabiner (2004) demonstrated that step width variability can discriminate the gait of young and old adults. Step distance, step width and step angle were calculated to assess whether the FSCi algorithm could identify parameters relevant to clinical practice. For walking, step distance, step width and step angle data (exception of step distance for shod walking) exhibited heteroscedasticity. For shod running, step width and step angle data also exhibited heteroscedasticity. The relationship between estimate error and estimate magnitude indicated that foot-surface contact position data were not identified accurately enough to derive step distance and step width parameters. However step distance and step width are view-dependent gait parameters, e.g. Figure 7.6. The impact of camera perspective on position reconstruction accuracy has been highlighted, i.e. Figures 7.11 - 7.14. It is therefore likely that view-dependent errors contribute to data heteroscedasticity when view-independent estimates of step distance and step width are calculated (Tables 7.8 and 7.9). View-dependent estimation of step distance and step width might be a viable application for the FSCi algorithm. However for the context of current work, view-independent estimates of step distance, step width and step angle are not suitable for analysis.

7.5.5. Reconstruction plane elevation

Reconstructing foot-surface contact data using elevated reconstruction planes reduced data heterogeneity and position reconstruction error. However reconstruction plane elevation was based on a pilot study. For this study, Figure 7.15 (A) demonstrates that plane elevation was related to both activity, i.e. walking or running, and condition, i.e. barefoot or shod. Optimal reconstruction plane elevation was lower for running conditions. This reflects that the FSCi algorithm identifies a region about the forefoot rather than mid-foot during running. Furthermore, optimal reconstruction plane elevation was greater for shod conditions. This reflects that trainer outsoles tend to elevate the foot when compared to barefoot conditions. When activities are grouped; optimal reconstruction plane elevation was 20.0 and 27.5 mm for barefoot and shod conditions respectively, e.g. B (Figure 7.15). However elevated reconstruction planes were calculated relative to the reference plane which was defined by the centre of four spherical markers 25 mm in diameter, i.e. 12.5 mm elevation. Therefore, when calculated relative to a ground level plane, optimal reconstruction plane elevation was 32.5 and 40.0 mm for barefoot and shod conditions respectively. The application of the current algorithm to the analysis of match-play tennis should therefore reconstruct image coordinate data using a reconstruction plane elevated to 40.0 mm.

7.6. Conclusion

This chapter describes the validation of a technique for identifying foot-surface contacts using single camera images of walking and running. The technique was applied automatically to image sequences obtained from different camera perspectives and does not require markers. The FSCi algorithm identified data for 1243 of 1248 foot-surface contacts (99.6%). Furthermore, step analysis was performed automatically for 91.3% of foot-surface contact data. Manual intervention for the remaining 8.7% of foot-surface contact data enabled step analysis: interventions primarily reflect the limitations of spatial data clustering. While spatial clustering was suitable for step analysis for the majority of walking and running trials, the automatic step analysis for multimodal gait, i.e. match-play tennis, might be limited. Further understanding of tennis gait strategy would be required to develop a suitable, time-windowed clustering algorithm.

For walking, view-independent foot-surface contact position errors were lower than errors reported by existing algorithms. Position errors for running were larger than

errors reported by existing algorithms for walking: no previous research has automatically identified foot-surface contacts in running. Therefore, with regard to identified acceptance criteria (section 7.1), the performance of FSCi algorithm was considered to be acceptable. Using these data, the FSCi algorithm measured basic gait parameters such as step length and step time for walking and running. However, the analysis of more detailed step parameters such as step distance, step width and step angle was not appropriate. Estimates for stance time were not appropriate for analysis because the FSCi algorithm does not identify heel-strike and toe-off events. However the FSCi algorithm did correctly identified mid-stance for walking. For running, the FSCi algorithm identified the time of foot-surface contact after mid-stance and reflects the application of a threshold to identify foot-surface contacts.

Using standard colour images the FSCi algorithm identified foot-surface contacts and measured basic gait parameters of walking and running without interfering with the activity being observed. The FSCi algorithm represents a flexible approach to markerless gait analysis and could be used for *in situ* analyses, i.e. match-play tennis. This is because the FSCi algorithm does not use *a priori* assumptions of gait mode but identifies feet that are stationary (indicative of stance). Furthermore, the FSCi algorithm can be applied to participant images of different size and perspective. In relation to the overall project aim, the FSCi algorithm represents a systematic approach for automatically identifying foot-surface contacts and measuring basic parameters of gait without interfering with play. The FSCi algorithm could be applied to footage of larger filming areas, i.e. match-play tennis: an objective assessment is warranted.

8 A semi-automatic technique for player tracking and foot-surface contact identification at the 2011 ATP World Tour Finals

8.1 Introduction

The Association of Tennis Professionals (ATP) is the organising body for men's professional tennis and provides a ranking system for professional matches played. Performances in ATP World Tour events, including Grand Slam tournaments, contribute to a player's rank. The 2011 ATP World Tour Finals are the season ending championship and are played on a distinctive blue acrylic, indoor hard-court surface (GreenSet Grand Prix). Competitors are the top eight ATP ranked players and thus the top eight tennis players in the world. The 2011 ATP World Tour Final was held in the O2 Arena in London from 20 - 27th November.

A manual system for characterising player step and movement strategy in match-play tennis was developed (Chapter 3). Findings were consistent with previous notational analyses and highlighted gender differences for forehand manoeuvre step frequency. However findings were limited due to a low sample size and movement definitions. Frame-by-frame analysis for player position and foot-surface contact location was required. Due to the large volume of footage required to perform tennis rally analyses, an automated approach was necessary to minimise user input.

An automatic method for the markerless and view-independent identification of foot-surface contacts using single camera footage was developed (Chapter 6). The method identified data for 99.6% of foot-surface contacts during walking and running; step analysis was performed automatically for 91.3% of data. Resultant direction RMSE for foot-surface contact position was 52.2 and 103.4 mm for shod walking and running respectively. Furthermore, RMSE for step length was 48.3 and 78.7 mm for shod walking and running respectively. It would be advantageous to measure player position and foot-surface contacts (to measure basic gait parameters) simultaneously; however, to the author's knowledge, no existing video-based method provides this. A player tracking algorithm, that allows the measurement of foot-surface contacts – described by Chapter 6 – as well as player position, is required. A key function of an algorithm would be to operate using single camera footage of match-play tennis.

8.2. *Aim and objectives*

The aim of this chapter was to develop a player tracking algorithm to measure player and foot-surface contact position for image sequences of match-play tennis rallies. This relates to box G of the development stage diagram (Figure 1.2).

Objectives:

1. Collect video footage of match-play tennis.
2. Develop an algorithm to automatically track tennis player position and extract images suitable for analysis using the FSCi algorithm.
3. Develop a graphical user interface to allow the user to interact with the algorithm and export relevant data.

8.3. *Development of a semi-automatic technique to identify player position and foot-surface contacts*

For clarity, Figure 8.1 provides an overview of image collection and analysis steps described in following sections.

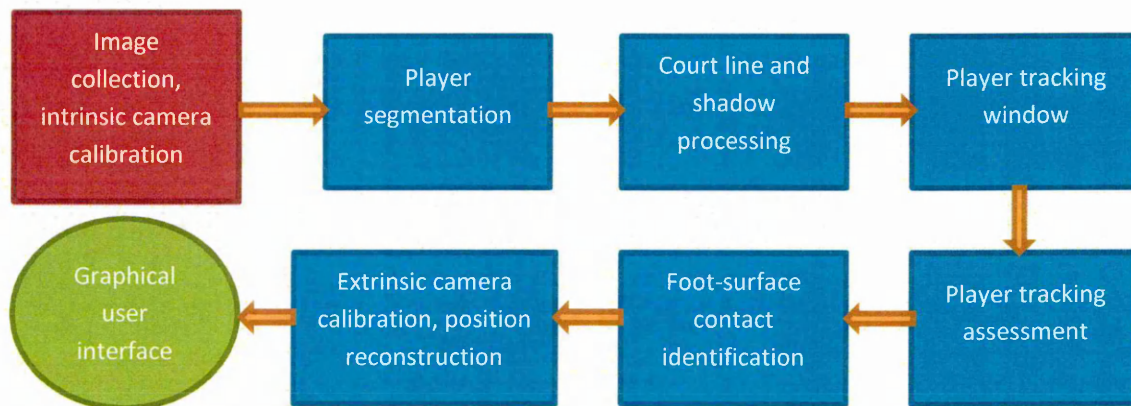


Figure 8.1. Image collection and analysis steps for player tracking and foot-surface contact identification.

8.3.1. *Image collection and intrinsic camera calibration*

Permission and relevant accreditations to film at the 2011 ATP World Tour Finals were obtained via the Lawn Tennis Association. Furthermore, approval for all procedures was obtained from the Research Ethics Committee of the Faculty of Health and Wellbeing, Sheffield Hallam University (Appendix 3). A high-definition video camera (Everio GZ-HD40EK, JVC, Japan), operating at 25 Hz (50 fields / second) with a resolution of 1920×1080 pixels (single CMOS sensor) was positioned in an elevated

location (fourth level of O2 Arena) to obtain a half-court field-of-view, e.g. Figure 8.2. Camera focal length was set manually and subsequently locked. Camera shutter speed was set manually to 1/250 s and locked; camera aperture was set automatically by the camera. The camera was equipped with an on-board disk drive and mains power input (mains power supplies were located on each level of the O2 Arena). This enabled the continual filming of tennis matches and minimised interference with match-play. All mains power cables and extension drums were tidied away from public access.

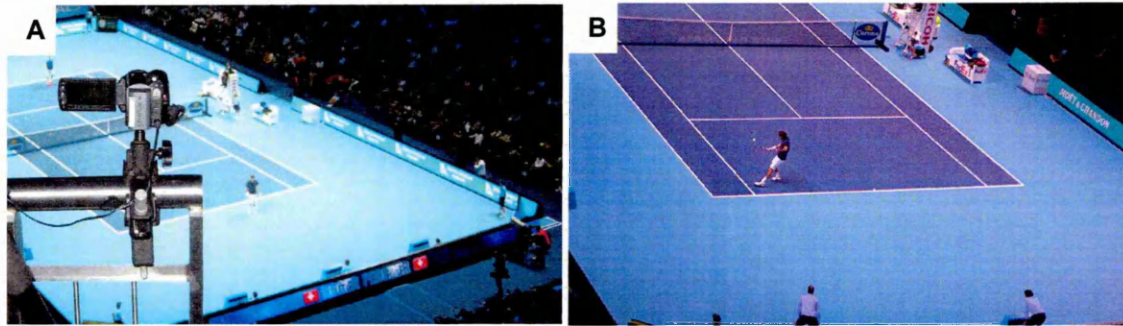


Figure 8.2. Camera setup at the 2011 ATP World Tour Finals (A) and sample camera image (B).

Single camera checkerboard calibration, as described in section 4.4.1 (Chapter 4), was performed. The camera was panned (180°) to film a checkerboard being held in different positions (less than 4 m) and orientations relative to the camera, e.g. Figure 8.3. Internal camera settings i.e. zoom, focal length, etc., were not altered. The camera was subsequently replaced to the desired field-of-view. Camera footage was downloaded to a laptop computer. Intrinsic camera parameters were calculated using deinterlaced (bob and expand) checkerboard images with the Camera Calibration toolbox (Bouguet, 2010).

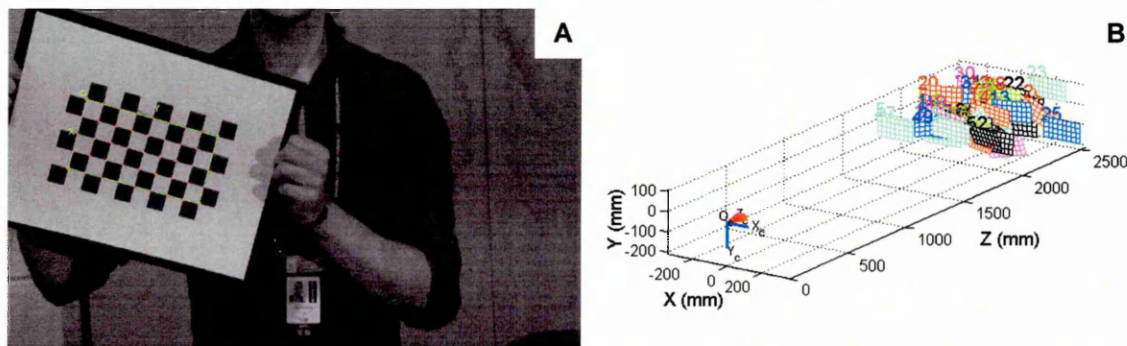


Figure 8.3. Single camera calibration: checkerboard extraction (A) and extrinsic checkerboard parameters (B).

The time codes of individual tennis rallies were identified and image sequences extracted. Figure 8.4 illustrates a 50 Hz sequence of four deinterlaced (bob and expand) player images (image resolution prior to crop was 1920×1080 pixels). Combing artifact was clearly evident about the swing limb, racket and ball. Combing artifact magnitude is related to motion between fields (Lee *et al.*, 2012). Furthermore, combing artifact reflects field scanning. Combing artifact highlights that a deinterlaced frame does not represent a single instant in time. As such, image fields sampled on even rows were discarded using VirtualDub (Lee, 2010). This yields 1920×540 pixel images and a sample rate corresponding to 25 Hz.

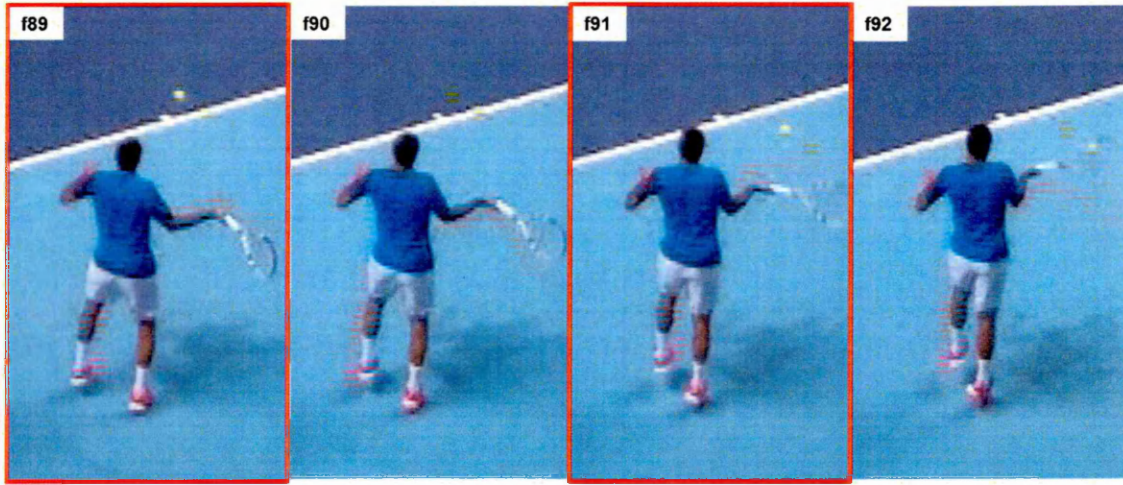


Figure 8.4. Deinterlaced player images (cropped) during a forehand groundstroke.

Extracted fields were subsequently converted to .tiff image files due to the tiled structure of the TIFF image format. Figure 8.5 illustrates extracted fields cropped to the lower-limbs. Combing artifact was still evident within extracted fields, i.e. image A (Figure 8.5). A row-averaging filter can reduce combing artifact (Wang and Farid, 2007). The application of a row-averaging filter (equation 8.1) illustrates reduced combing artifact, i.e. image B (Figure 8.5). Therefore extracted fields used for foreground player segmentation were filtered with a row-averaging filter.

$$I_{RGB}(x, y, t_n) = \left(\frac{1}{2} I_{RGB}(x, y - 1, t_n) \right) + \left(\frac{1}{2} I_{RGB}(x, y + 1, t_n) \right), y \bmod 2 = 1 \quad [8.1]$$

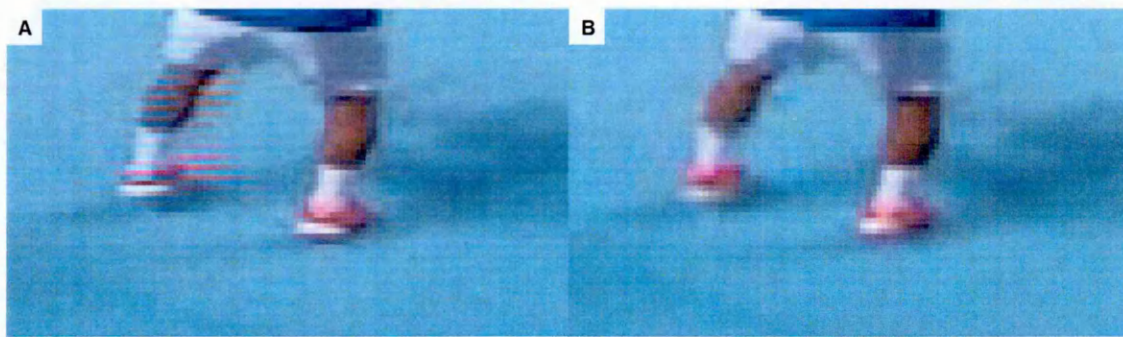


Figure 8.5. Extracted fields without (A) and with (B) a row-averaging filter applied.

Frame-to-background image differencing is a common method to detect motion in image sequences. Frame-to-background differencing was adopted; a static background model was used to reduce computational demand (Hassanpour, Sedighi and Manashty, 2011). Due to the nature of match-play tennis, the static background model $Ref_{RGB}(x, y)$ must be generated because players are on-court at the start and end of tennis rallies.

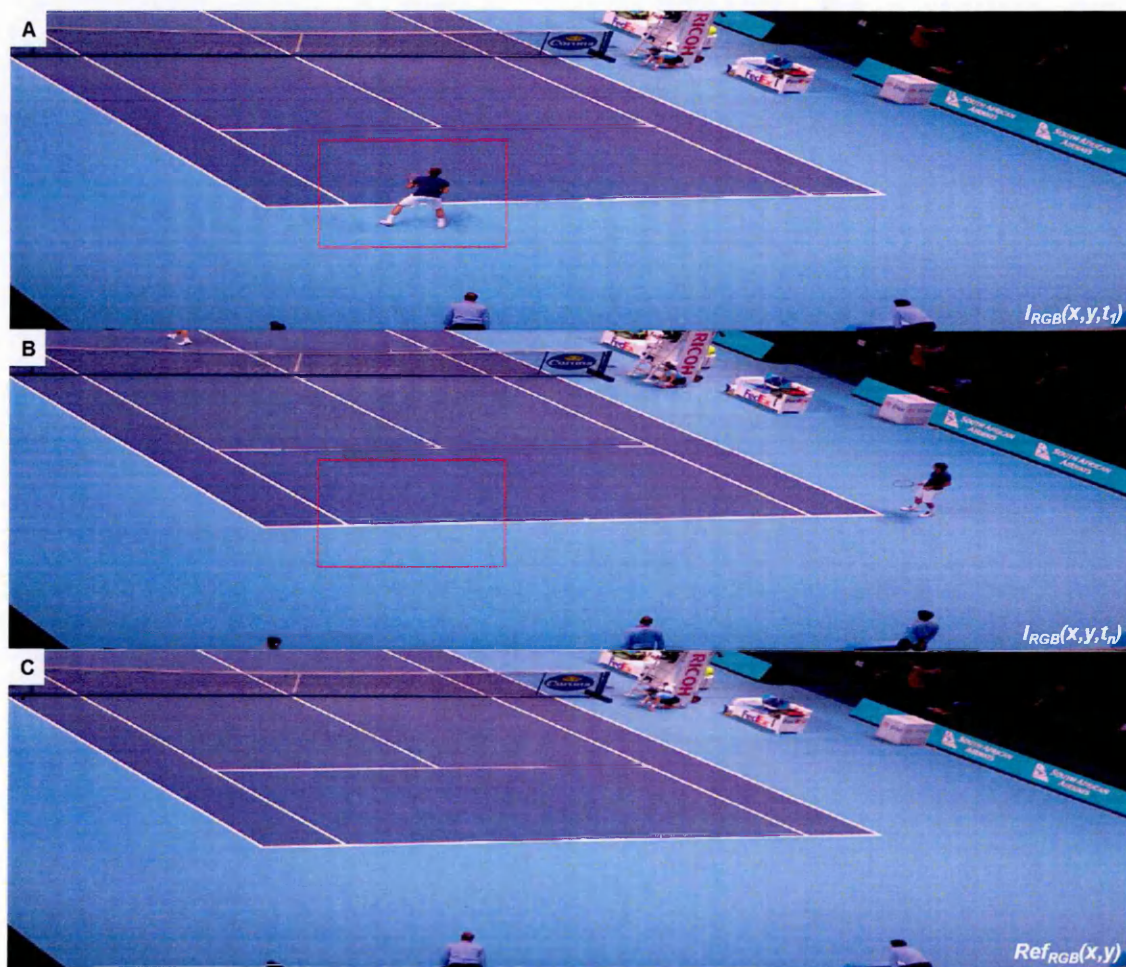


Figure 8.6. Manually defined region (red rectangle) superimposed on image t_1 (A) and image t_n (B) and resultant background image (C).

Let $I_{RGB}(x, y, t_n)$ be the RGB image at time $t = [t_1, t_2, \dots, t_F]$ for an image sequence of length F , where (x_l, y_l) is the top-left corner of the image. The image $I_{RGB}(x, y, t_n)$, where player displacement is maximised relative to $I_{RGB}(x, y, t_1)$, is manually identified. Images A and B (Figure 8.6) illustrate suitable images for background modelling. A rectangular region, encompassing the player and any shadow, is manually defined in $I_{RGB}(x, y, t_1)$, i.e. image A (Figure 8.6). RGB pixel intensities within the rectangular region for $I_{RGB}(x, y, t_n)$, i.e. image B (Figure 8.6), are then substituted into $I_{RGB}(x, y, t_1)$. The resulting background model, i.e. image C (Figure 8.6), is hereafter referred to as $Ref_{RGB}(x, y)$.

8.3.2. Player segmentation

To initiate player tracking, an initial estimate for player position and tracking window dimension is required. Player location is manually identified (single mouse click) within $I_{RGB}(x, y, t_1)$. By applying the absolute difference image method (equation 6.1), the threshold $level(t_1)$ can be identified for $ADI_{RGB}(x, y, t_1)$ using MATLAB's implementation of Otsu's gray-level threshold method (Otsu, 1979). Subsequently, the binary image $ADI_B(x, y, t_1)$ can be extracted. Figure 8.7 (A) illustrates $ADI_B(x, y, t_1)$, cropped to highlight the player. Following the application of basic morphological operations, the centroid and bounding box can be calculated. Player COM, i.e. $(pCOM_x, pCOM_y, t_1)$, is defined as the centroid of the binary image $ADI_B(x, y, t_1)$ and provides an initial estimate for player image location. An initial tracking window, $(tWin_x, tWin_y, tWin_w, tWin_h, t_1)$, is defined as twice the bounding box width and height, centred about $(pCOM_x, pCOM_y, t_1)$. Image B (Figure 8.7) illustrates manually identified and estimated player locations (yellow circle and red triangle respectively) as well as the initial tracking window (red rectangle).

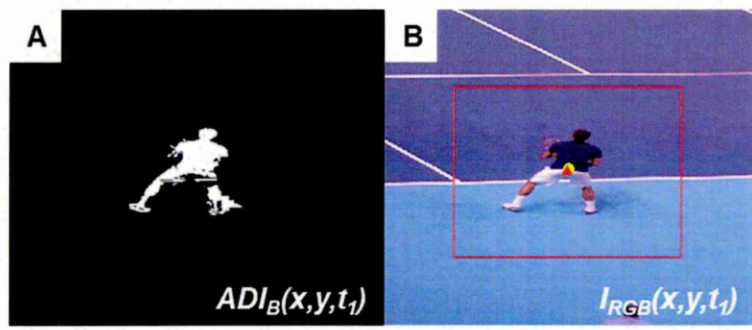


Figure 8.7. Extracted binary image (A) and original image (B) with player position (manual and estimated in yellow and red respectively) and tracking window plotted.

The windowed, player tracking image $Player_{RGB}(x, y, t_1)$ and corresponding background image $PlayerRef_{RGB}(x, y, t_1)$ are then defined using $(tWin_x, tWin_y, tWin_w, tWin_h, t_1)$, i.e. equations 8.2 and 8.3. Windowed images can be imported directly due to the tile structure of the TIFF image format. This is more computationally efficient than importing and cropping full-resolution images to a tracking window dimension. Section 8.3.4 describes the computation of tracking windows for subsequent images, i.e. $t_n > 1$. Applying the absolute difference image method (equation 6.1) to $Player_{RGB}(x, y, t_n)$ and $PlayerRef_{RGB}(x, y, t_n)$ yields $PlayerADI_{RGB}(x, y, t_n)$.

$$Player_{RGB}(x, y, t_1) = I_{RGB}(x : x + w, y : y + h, t_1) \quad [8.2]$$

$$PlayerRef_{RGB}(x, y, t_1) = Ref_{RGB}(x : x + w, y : y + h) \quad [8.3]$$

where x, y, w and h are the corresponding elements of $(tWin_x, tWin_y, tWin_w, tWin_h, t_1)$.

8.3.3. Court line and shadow processing

Teachabarikiti, Chalidabhongse and Thammano (2010) noted that, for tennis performed on acrylic court surfaces, the relative brightness of tennis court lines introduced many false candidates to their ball tracking algorithm. Indeed, Tiarks *et al.* (2003) reported that scrub resistant, high-gloss lacquer paints contain a large proportion of TiO_2 , a pigment that improves paint opacity. TiO_2 has a high refractive index and induces light scattering (Tiarks *et al.*, 2003). The use of high-gloss paint on indoor acrylic tennis court surfaces would support the comments of Teachabarikiti, Chalidabhongse and Thammano (2010). To improve analyses, Teachabarikiti, Chalidabhongse and

Thammano (2010) removed court line regions from images. However for the context of current work, it is important that court line regions are retained.

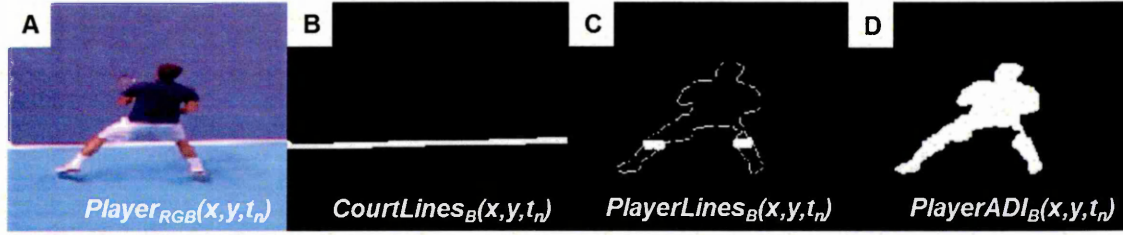


Figure 8.8. Player image (A), identified court lines (B), processed court lines and superimposed player outline (C), combined player and court line silhouette (D).

Image A (Figure 8.8) illustrates the windowed image $Player_{RGB}(x, y, t_n)$. Court line regions are identified within $PlayerRef_{RGB}(x, y, t_n)$ by applying equation 8.4. The binary image $CourtLines_B(x, y, t_n)$ is dilated using a 3×3 structuring element, e.g. image B (Figure 8.8). Court line regions within $PlayerADI_{RGB}(x, y, t_n)$ are identified by first concatenating $CourtLines_B(x, y, t_n)$ into the three-plane binary image $CourtLines_{BBB}(x, y, t_n)$. The application of equation 8.5 yields $PlayerLines_{RGB}(x, y, t_n)$. Court line regions are removed from $PlayerADI_{RGB}(x, y, t_n)$ to reduce image noise, i.e. light scattering (Teachabarikiti, Chalidabhongse and Thammano, 2010). The binary images $PlayerADI_B(x, y, t_n)$ and $PlayerLines_B(x, y, t_n)$ are then extracted individually by applying equations 6.2 - 6.7 (Chapter 6).

$$CourtLines_B(x, y, t_n) = PlayerRef_{RGB} \geq white \quad [8.4]$$

$$PlayerLines_{RGB}(x, y, t_n) = CourtLines_{BBB}(x, y, t_n) \times PlayerADI_{RGB}(x, y, t_n) \quad [8.5]$$

$$PlayerADI_B(x, y, t_n) = (PlayerADI_B(x, y, t_n) + PlayerLines_B(x, y, t_n)) > 0 \quad [8.6]$$

where $white = [0.7, 0.7, 0.7]^T$ for rescaled (max value of 1) double precision images.

Image C (Figure 8.8) illustrates $PlayerLines_B(x, y, t_n)$, i.e. foreground court line regions; the player outline is superimposed for illustrative purposes only. Finally, equation 8.6 combines foreground player and court line regions, i.e. image D (Figure 8.8). Resulting binary player images are morphologically processed. Connected components less than P (equation 8.7) are removed; P represents 1% of mean player area for images $t_2 \dots t_n$, the magnitude of P for t_1 is set to 25 pixels. Remaining pixels within $PlayerADI_B(x, y, t_n)$

are thickened (single iteration) and closed. Pixels connected to the image border are removed.

$$P = \frac{\sum(Area, t_{2...n})}{n} / 100 \quad [8.7]$$

Shadow reduction using the HSV colour space was adopted because it is a computationally fast and effective method for indoor footage (Benedek and Szirányi, 2007). The images $Player_{RGB}(x, y, t_n)$ and $PlayerRef_{RGB}(x, y, t_n)$ are converted to HSV colour space, yielding $Player_{HSV}(x, y, t_n)$ and $PlayerRef_{HSV}(x, y, t_n)$. The application of the absolute image difference method (equation 6.1) yields $PlayerADI_{HSV}(x, y, t_n)$. Court line regions are also processed using HSV colour space images. Substituting $PlayerADI_{HSV}(x, y, t_n)$ into equation 8.5 yields $PlayerLines_{HSV}(x, y, t_n)$. Court line regions are removed from $PlayerADI_{HSV}(x, y, t_n)$ to reduce image noise. Shadow does not change image chromaticity (Nghiem and Thonnat, 2008); therefore the value components of $PlayerADI_{HSV}(x, y, t_n)$ and $PlayerLines_{HSV}(x, y, t_n)$ are removed. Subsequently the binary images $HS-PlayerADI_B(x, y, t_n)$ and $HS-PlayerLines_B(x, y, t_n)$ are then extracted individually by applying equations 6.2 - 6.7 (Chapter 6). Finally, foreground player and court line regions are combined by substituting $HS-PlayerADI_B(x, y, t_n)$ and $HS-PlayerLines_B(x, y, t_n)$ into equation 8.6, yielding $HS-PlayerADI_B(x, y, t_n)$.

Binary images extracted from Hue and Saturation components of $PlayerADI_{HSV}(x, y, t_n)$ are morphologically processed in the same way as binary images extracted from RGB images. Benedek and Szirányi (2007) highlighted that the HSV colour space can be sensitive to image noise. As such, binary images derived from RGB and HS colour spaces (A and B respectively: Figure 8.9) are combined (equation 8.8): this yields the player foreground image $Player_B(x, y, t_n)$. Image C (Figure 8.9) illustrates the combined foreground player image; RGB pixel values and shadow (equation 8.9) perimeter are presented for reference.

$$Player_B(x, y, t_n) = PlayerADI_B(x, y, t_n) \times HS-PlayerADI_B(x, y, t_n) \quad [8.8]$$

$$PlayerShadow_B(x, y, t_n) = PlayerADI_B(x, y, t_n) - HS-PlayerADI_B(x, y, t_n) \quad [8.9]$$

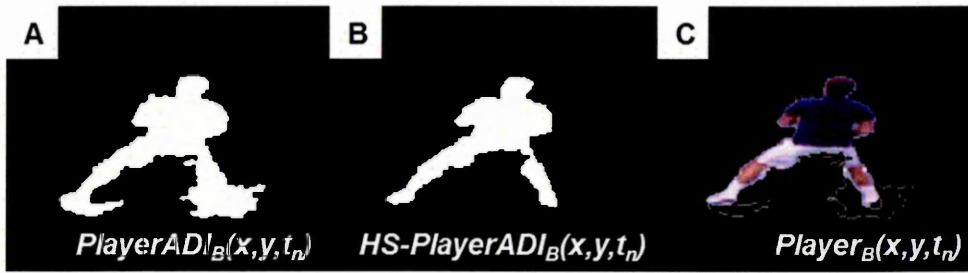


Figure 8.9. Binary images derived from RGB (A) and HS (B) colour spaces. RGB pixel values and shadow perimeter are superimposed on the combined foreground player image for illustration (C).

If more than one component exists within $Player_B(x, y, t_n)$, local components are grouped by fitting ellipses to each component (Lee and Grimson, 2002). The centroid, major and minor axis lengths are computed for each component within $Player_B(x, y, t_n)$. Enlarged ellipses (150% of axis lengths) are constructed about the centroid for each component. A primary ellipse is determined using previous player location, i.e. $(pCOM_x, pCOM_y, t_{n-1})$. Figure 8.10 illustrates a discontinuous binary image, where four individual components exist. Previous player location (red filled circle) and corresponding primary ellipse are plotted in red. Ellipses linked to the primary ellipse via an intersection are retained, i.e. green and blue ellipses (A \rightarrow B; Figure 8.10). Non-intersecting ellipses and corresponding image components are rejected, i.e. magenta ellipse (A \rightarrow B; Figure 8.10).

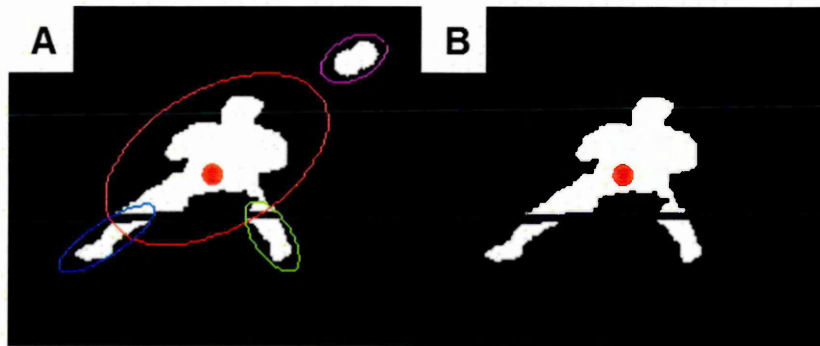


Figure 8.10. Elliptical grouping for a discontinuous binary image, relative to previous player location (red filled circle).

8.3.4. Player tracking windows

To predict the location and dimension of successive player tracking windows, i.e. $t_n > 1$, basic image properties (area, centroid, major axis length and bounding box) are computed from $Player_B(x, y, t_n)$, i.e. $(Area, t_n)$, $(pCOM_X, pCOM_Y, t_n)$, $(Height, t_n)$ and $(Box_X, Box_Y, Box_W, Box_H, t_n)$. The vector $(pVect_X, pVect_Y, t_n)$ defines inter-frame player displacement (equation 8.10). Initial player displacement, i.e. $(pVect_X, pVect_Y, t_1)$, is set to zero pixels. Equation 8.11 yields the tracking window $(tWin_X, tWin_Y, tWin_W, tWin_H, t_{n+1})$ which is defined by twice the width and height of $Box(x, y, w, h, t_n)$, centred about $(pCOM_X, pCOM_Y, t_n)$.

$$(pVect_X, pVect_Y, t_n) = (pCOM_X, pCOM_Y, t_n) - (pCOM_X, pCOM_Y, t_{n-1}) \quad [8.10]$$

$$(tWin_X, tWin_Y, tWin_W, tWin_H, t_{n+1}) = \left(\left(pCOM_X - \frac{w}{2} \right) + pVect_X, \left(pCOM_Y - \frac{h}{2} \right) + pVect_Y, 2w, 2h, t_{n+1} \right) \quad [8.11]$$

where w and h are corresponding components of $(Box_X, Box_Y, Box_W, Box_H, t_n)$.

Figure 8.11 illustrates tracking windows applied to match-play tennis. Image A (Figure 8.11) presents the combination of two images with a 0.04 s time interval (25 Hz). Repositioned and resized player tracking windows (red and green rectangles) relative to the player illustrate effective tracking windows. During match-play tennis, players assume irregular shapes and move at irregular speeds. Image B (Figure 8.11) illustrates player tracking windows for three images during a baseline run: image intervals have been increased to 0.4 s (2.5 Hz) for illustration purposes. Red, green and blue rectangles (B: Figure 8.11) illustrate the various dimensions required for tracking windows to capture tennis player movement.

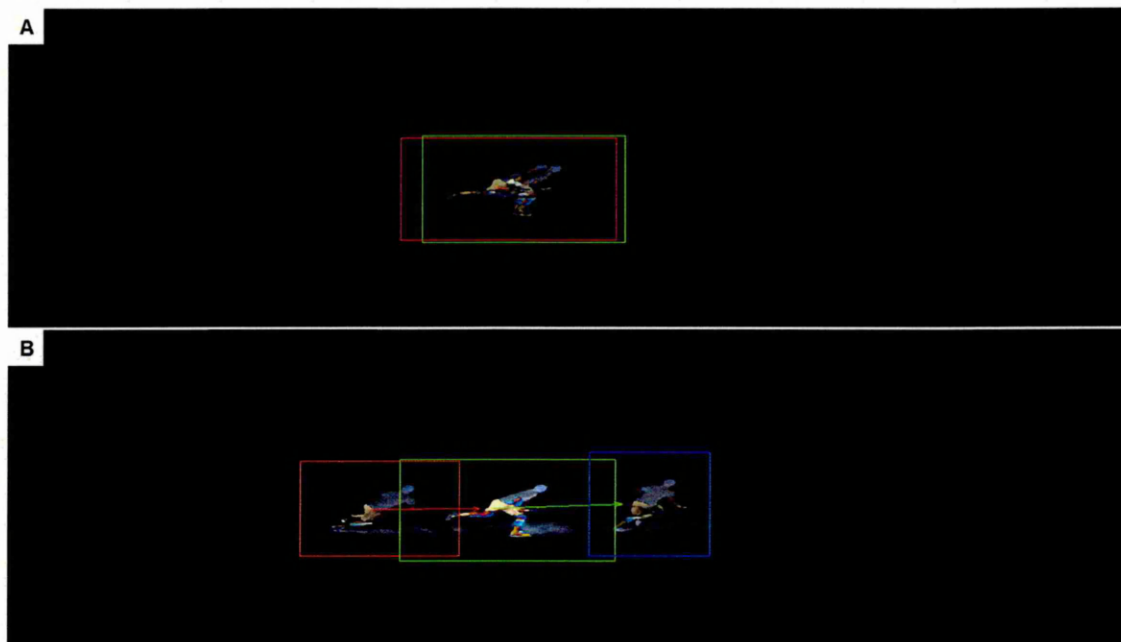


Figure 8.11. Tracking windows superimposed on combined images of a baseline run with original image interval (A) and increased image interval (B).

Tracking window coordinates are assessed in relation to image boundaries. If tracking window coordinates exceed the dimensions of $I_{RGB}(x, y, t_n)$, a default tracking region $1/5^{\text{th}}$ of full-image resolution, i.e. 384×216 pixels, is imposed. Figure 8.12 illustrates a tracking window that has exceeded image boundaries, i.e. red and black rectangles respectively. Infringed boundaries are identified and a default tracking window positioned as close to the tracked window location as possible without exceeding image boundaries (green rectangle).

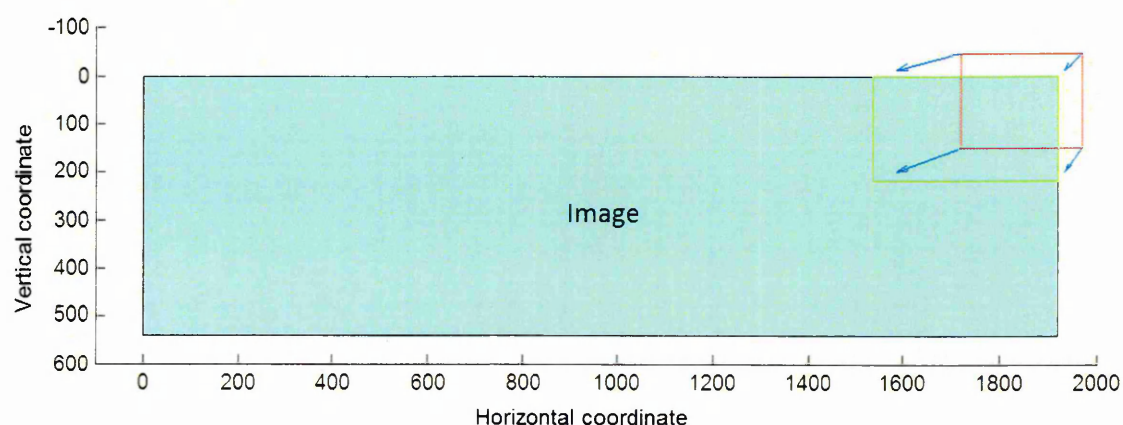


Figure 8.12. Imposition of a default tracking window (green rectangle) if an automated tracking window (red rectangle) exceeds image boundaries (black rectangle).

8.3.5. Assessing player tracking

Algorithms that assess object tracking performance based on motion consistency as well as image shape, area and appearance have been proposed (Wu and Zheng, 2004). However, such approaches are suited to scenarios where moving objects do not change direction or speed dramatically (Li, Dore and Orwell, 2005). This is not the case for match-play tennis. Zhang *et al.* (2012) presented a confidence-level-based particle filter for non-linear object tracking. However particle filtering, based on Monte Carlo simulation, can be computationally expensive (Zhang *et al.*, 2012). For current work a smoothing spline, based on the discrete cosine transform (Garcia, 2010), is applied to horizontal and vertical player coordinate data, i.e. $(pCOM_x, pCOM_y, t_1 \dots t_n)$. An optimal smoothing parameter for tennis player motion was established for a sample tennis rally using residual analysis (Winter, 2005). Player tracking residuals and 95% confidence intervals are calculated relative to horizontal and vertical coordinate smoothing splines to assess player tracking. Figure 8.13 (A) illustrates horizontal player coordinate data tracked throughout an image sequence. Figure 8.13 (B) presents a zoomed-in region of A, illustrating raw and smoothed (black and blue lines respectively) player coordinate data; coordinates that exceed the 95% confidence interval are highlighted by red circles. Player tracking is terminated if ten consecutive player tracking residuals, i.e. 0.4 s, exceed 95% confidence intervals.

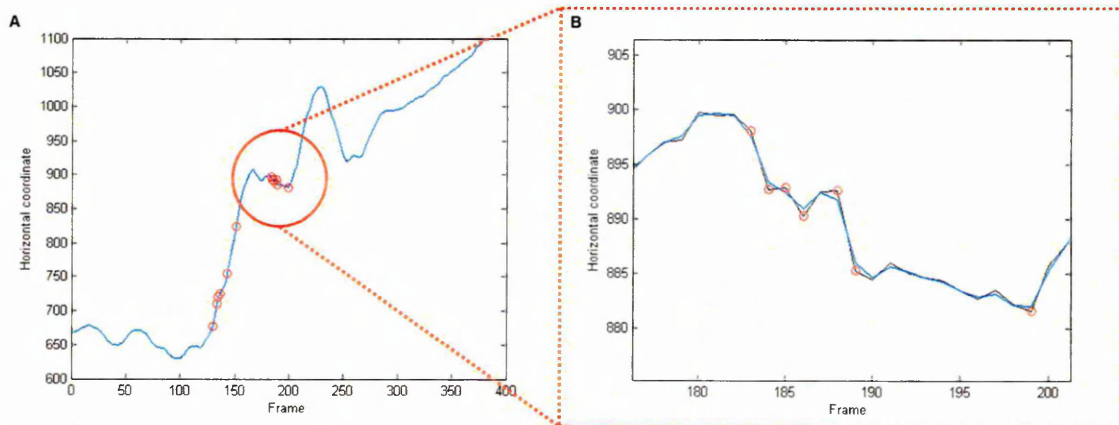


Figure 8.13. Horizontal player coordinate data (black) and smoothing spline (blue). Coordinates exceeding 95% confidence intervals about the smoothing spline are highlighted (red circles).

Player tracking residuals attempt to identify small errors within player trajectory, e.g. incorrect foreground segmentation etc. To detect player trajectory changes indicative of

tracking failure, two logical geometric rules are imposed. The player tracking algorithm assumes that player displacement between video frames will be less than player stature (equation 8.12). Displacements greater than player stature indicate that the tracking algorithm had identified an object that was not the player. Furthermore, player area is assumed to exist within upper and lower bounds (corresponding to $\pm 75\%$) of mean player area (equation 8.13). Player areas exceeding upper and lower bounds would indicate the failure of player segmentation and thus tracking. Player tracking is terminated if either logical rules (PT_{R1}, t_n) or (PT_{R2}, t_n) are false.

$$(PT_{R1}, t_n) = \sqrt{((pVect_x, t_n)^2 + (pVect_y, t_n)^2)} \leq (Height, t_n) \quad [8.12]$$

$$(PT_{R2}, t_n) = \left(\frac{(Area, t_2 \dots t_n)}{t_{n-1}} \times 0.25 \right) \geq (Area, t_n) \leq \left(\frac{(Area, t_2 \dots t_n)}{t_{n-1}} \times 1.75 \right) \quad [8.13]$$

8.3.6. Foot-surface contact identification

The foot-surface contact identification (FSCi) algorithm executes alongside the player tracking algorithm for images $t_2 \dots t_F$. The same image region applied to $Player_{RGB}(x, y, t_n)$ and $PlayerRef_{RGB}(x, y, t_n)$ is applied to $I_{RGB}(x, y, t_n - 1)$, i.e. equation 8.14. This provides the necessary images to quantify inter-frame motion as described in section 6.4 (Chapter 6). Images analysed for inter-frame motion by the FSCi algorithm are not row-average filtered, i.e. equation 8.1.

$$PrePlayer_{RGB}(x, y, t_n) = I_{RGB}(x : x + w, y : y + h, t_n - 1) \quad [8.14]$$

where x, y, w and h are the corresponding elements of $(tWin_x, tWin_y, tWin_w, tWin_h, t_n)$.

Following the identification of basic player image properties, pixels within $Player_B(x, y, t_n)$ above $(pCOM_y, t_n)$ are removed, i.e. $A \rightarrow B$ (Figure 8.14). This follows equation 6.13 as described in FSCi algorithm development, i.e. section 6.4 (Chapter 6). Shadow removal described in section 6.4 (Chapter 6) is not performed as shadow removal is performed during player segmentation (section 8.3.3). As such, the FSCi algorithm is executed as presented in section 6.4 (Chapter 6) from equation 6.15 onwards.

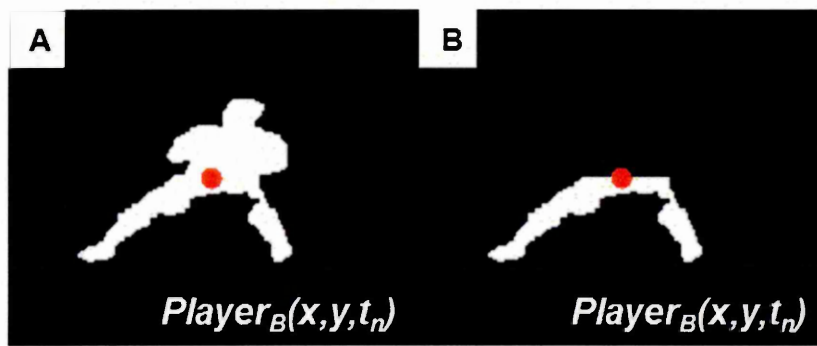


Figure 8.14. Pixels above player COM (red filled circle) are set to zero (A \rightarrow B).

8.3.7. *Extrinsic camera calibration and position reconstruction*

To perform extrinsic camera calibration, intrinsic camera parameters (section 8.3.1) are loaded into MATLAB. The image $I_{RGB}(x, y, t_1)$ is resized (1920×1080 pixels; bicubic interpolation) and manually digitised, i.e. singles sideline-baseline and sideline-serviceline intersections (Figure 8.15). The image is resized to ensure intrinsic and extrinsic camera parameters correspond. To identify extrinsic camera parameters, manually digitised coordinates, i.e. court line intersections 1- 4 (Figure 8.15), and corresponding real-world dimensions are processed using the Camera Calibration toolbox for MATLAB (Bouguet, 2010).

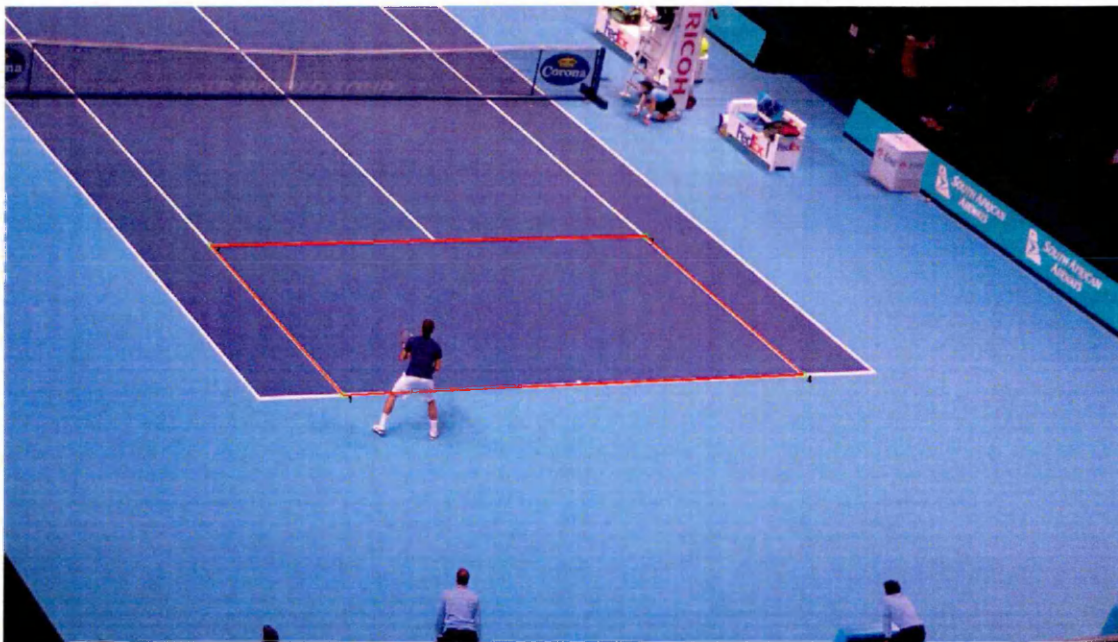


Figure 8.15. Manually digitised sideline-baseline and sideline-serviceline intersections.

Following the computation of intrinsic and extrinsic camera parameters, position reconstruction can be performed. To reduce out-of-plane position reconstruction error, two elevated reconstruction planes are computed. Player and foot-surface contact position data are reconstructed using reconstruction planes elevated to 50% of player stature and 40.0 mm respectively (refer to sections 5.4.1 and 7.5.5 respectively). Real-world coordinates for each elevated reconstruction plane are reprojected into the camera image yielding image coordinates for player and foot-surface contact reconstruction planes.

Image coordinates for player and foot-surface contact reconstruction planes are processed with corresponding real-world dimensions, i.e. court line intersections, using the Camera Calibration toolbox (Bouquet, 2010). This yields extrinsic camera parameters relative to each reconstruction plane. Accordingly, player and foot-surface contact position data, i.e. $(pCOM_x, pCOM_y, t_n)$ and $(FSCx, FSCy, t_n)$ respectively, are reconstructed with corresponding extrinsic camera parameters using the planar reconstruction method, i.e. equations 4.11 - 4.15 (Chapter 4).

8.3.8. Graphical user interface

A graphical user interface (GUI) was developed to provide a simple analysis tool for the player-tracking and foot-surface contact identification algorithms (hereafter referred to as PT-FSCi). The GUI was programmed using MATLAB to enable the use of the Image Processing and Camera Calibration toolboxes. The GUI was required to perform the following:

- Allow a user to analyse match-play tennis footage on a semi-automatic basis.
- Provide interactive controls and visualisation of data analysis.
- Provide method for data export.

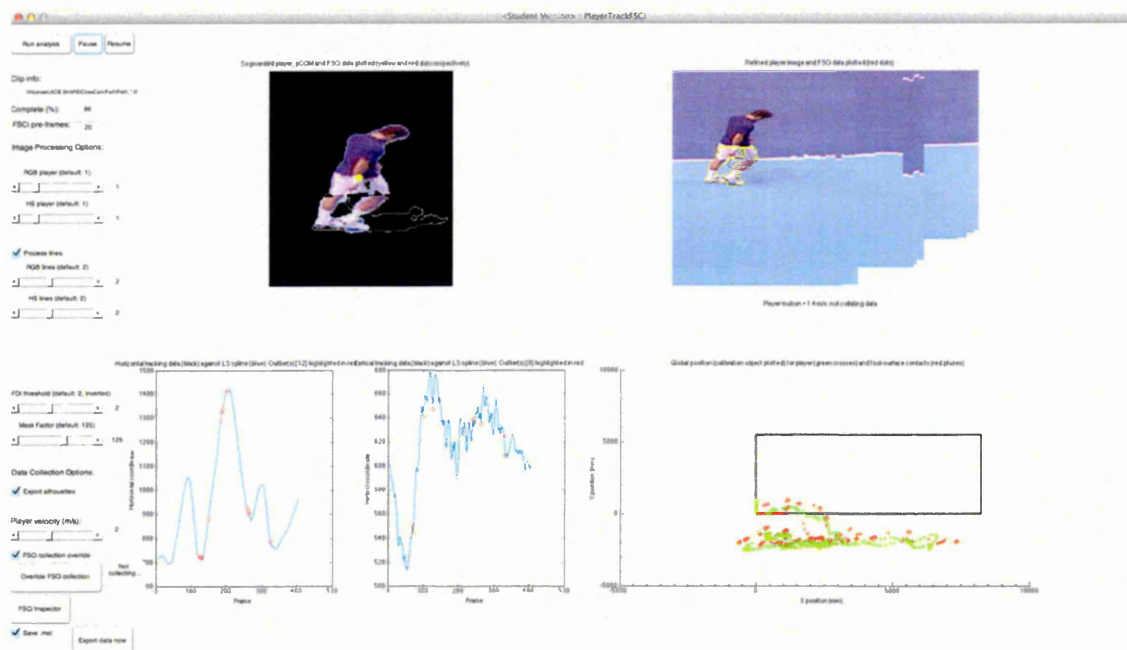


Figure 8.16. Screenshot of main GUI window. User controls are located to the left and data visualisations to the centre and right of the screen.

A full-screen interface window presents the user with analysis input controls and output data visualisations (Figure 8.16). The top-left region of the main window presents the user with buttons to run, pause and resume an analysis, as well as analysis information (A: Figure 8.17). Selecting the 'Run analysis' button will prompt the user to:

1. Select an image sequence for analysis.
2. Perform camera calibration.
 - Intrinsic camera parameters are automatically loaded.
 - User manually digitises four known court locations and inputs corresponding world dimensions.
 - User inputs player stature.
3. Select images to create a suitable background model.
4. Identify player location (single mouse click) at the start of the image sequence.

In addition to creating a background model, the user is asked if it is necessary to create masks (regions removed from analysis). Creating a mask requires the user to identify a region within the background model; identified pixel intensities are substituted with undefined values, i.e. NaN. Following the computation of an absolute difference image, i.e. equation 6.1, pixels values identified as NaN are substituted for zeros. This removes

the mask region from absolute difference images and thus analysis. Following setup, the PT-FSCi algorithm executes automatically.

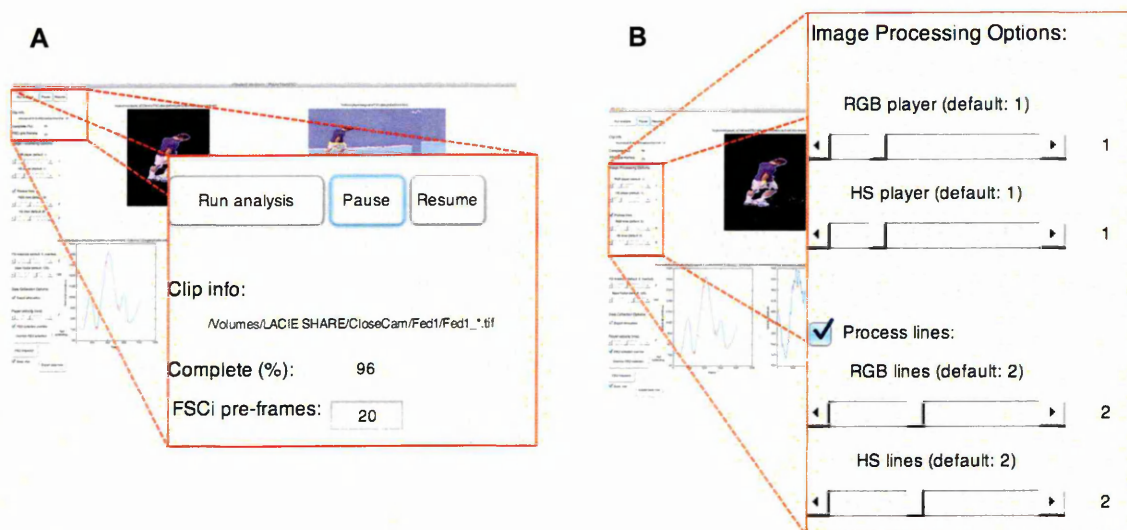


Figure 8.17. Analysis controls and information (A) and threshold sliders (B).

Sliders control thresholds used for foreground player segmentation (B: Figure 8.17). The manipulation of segmentation thresholds is necessary due to large camera field-of-view. For example, player images at the image periphery might be blurred. Equations 6.2 - 6.5 (Chapter 6) indicate that segmentation thresholds are the sum of the mean and standard deviation for an absolute difference image. Blurred or out-of-focus images have lower pixel intensity variation than a corresponding, in-focus image. Sliders therefore multiply image standard deviation magnitude to provide user control over foreground segmentation. Segmentation visualisation enables the user to easily identify an appropriate threshold level. Court line segmentation is optional; segmentation thresholds can also be manually adjusted using sliders if necessary.

Sliders also control parameters passed to the FSCi algorithm. The 'FDI threshold' slider (A: Figure 8.18) enables the user to control inter-frame differences analysed by the FSCi algorithm. The 'Mask Factor' slider (A: Figure 8.18) controls the normalising parameter that determines when an object is deemed to be stationary, i.e. equation 6.25 (Chapter 6). For match-play tennis, this parameter was determined experimentally as 125%. Image sequences can be analysed for foot-surface contacts continuously, i.e. 'FSCi collection override' tick box (B: Figure 8.18) or in relation to estimated player velocity. Player velocity is estimated from reconstructed player position data (resultant

direction) and filtered using a four-point moving average filter. The 'Player velocity' slider (B: Figure 8.18) identifies a minimum player velocity: foot-surface contact data are recorded when estimated player velocity exceeds this threshold.

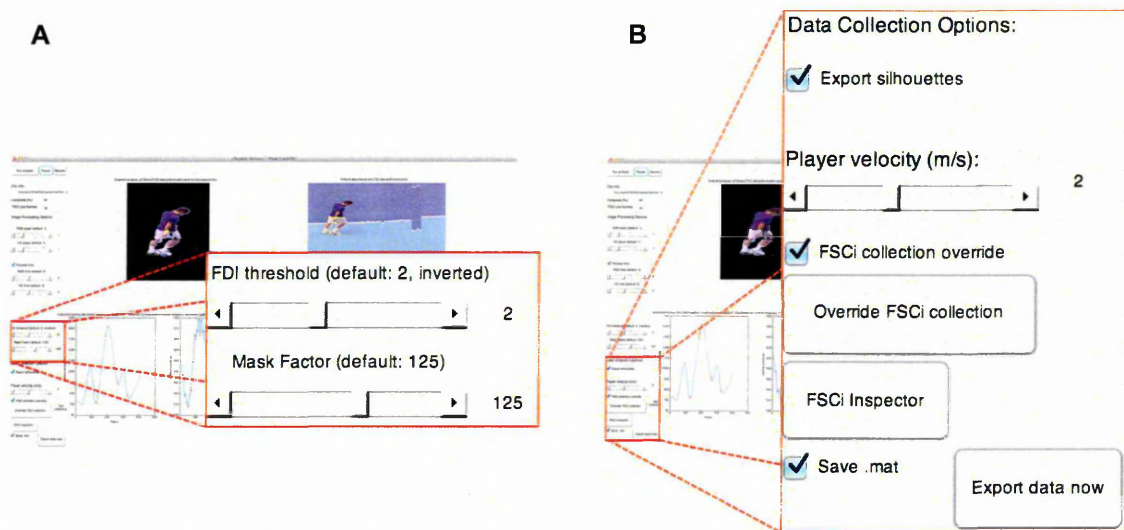


Figure 8.18. Foot-surface contact controls (A) and data collection options (B).

All segmented player images, player position and foot-surface contact data can be visually inspected. Selecting the 'FSCi Inspector' button (B: Figure 8.18) launches a module enabling the user to scroll through segmented player images. Player COM and foot-surface contact data are burnt into each segmented player image (magenta diamonds). Segmented player images where player tracking residuals exceed 95% confidence intervals (section 8.3.5) are also flagged to the user for inspection. By selecting 'Reject' (Figure 8.19), the user can flag images where player segmentation, player position or foot-surface contact location are deemed to be incorrect.

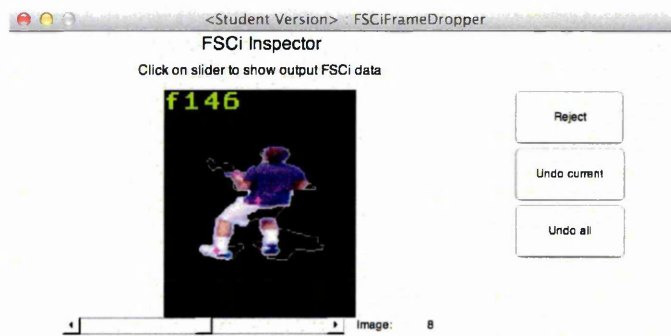


Figure 8.19. Module for reviewing player segmentation as well as player position and foot-surface contact data.

Analysed data are exported when the PT-FSCi algorithm reaches the end of an image sequence or at any point during analysis by selecting the 'Export data now' button (B: Figure 8.18). The following data are exported:

- Segmented player images (.bmp)
- Player COM data
 - *UVT* data (.ascii)
 - *XYT* data (.ascii)
- Foot-surface contact data
 - *UVT* data (.ascii)
 - *XYT* data (.ascii)
- Flagged data
 - Player tracking residual analysis (.ascii)
 - FSCi Inspector analysis (.ascii)
- MATLAB workspace data (.mat)

8.4. *Application to match-play tennis*

Section 8.3 presented a player tracking and foot-surface contact identification algorithm. Furthermore, a graphical user interface was developed to provide a simple analysis tool. Footage of the 2011 ATP World Tour Finals was used to develop the player tracking algorithm for singles match-play tennis. The following section describes the application of the algorithm to sample match-play tennis footage.

Competitors performing in the 2011 ATP World Tour Finals were the top eight ATP ranked players in the world; footage therefore represents elite match-play tennis. Many

methods for player tracking exist in sport, e.g. Connaghan, Moran and O'Connor (2013); Martínez-Gallego et al (2013); Mauthner *et al.* (2008). Current work required a novel player tracking algorithm to enable foot-surface contact analysis, i.e. FSCi algorithm. The context of current work i.e. singles match-play tennis, allowed the simplification of player tracking assumptions. The rules of tennis (ITF Rules of Tennis, 2013) state that playing conditions or anything within the field-of-vision of players cannot be interfered with during match-play. Therefore any moving objects within the tennis court area during play can be assumed to be tennis players. Furthermore, the net segregates the two players and ensures they do not interact or cross paths during play. This reduces the need for a player tracking algorithm to identify individual players or motion paths. Instances where both players are at the net might require player differentiation; however this represents a small proportion of match-play tennis (O'Donoghue and Ingram, 2001). Finally, although court officials might move during rallies, i.e. line judges etc., movement will be small and can be masked if necessary.

Accurate foreground segmentation of tennis players is important, particularly for foot-surface contact identification. Image fields sampled on even rows were discarded due to spatial and temporal asynchrony. Furthermore, a row-averaging filter was applied to remove combing artifact, i.e. Figure 8.5. Wang and Farid (2007) reported that following row-average filtering (full-frame), 100% of pixels were correctly classified as belonging to their spatial neighbours ($r > 0.90$). Combing artifact can therefore be reduced without distorting the appearance of player images. To identify tennis player location, Conaire *et al.* (2009) and Connaghan, Moran and O'Connor (2013) applied a box filter (29×29 pixels) to all background subtracted images obtained from an overhead camera. However this is inappropriate for perspective camera images due to fixed filter size. For perspective camera images of beach volleyball, Mauthner *et al.* (2008) performed relatively few morphological operations to preserve player shape irregularity, i.e. player limbs. For irregular shapes such as tennis players, excessive morphological processing can regularise foreground player image shape. This can incorporate regions that are not foreground player regions and affect estimates for player position and size. Current work performs four basic morphological operations: only two operations (thickening and closing) affect silhouette shape. Player position was defined as the centroid of $Player_B(x, y, t_n)$. Section 5.4 (Chapter 5) demonstrated that player position reconstruction using ground plane features was limited, i.e. irregular shapes of lower-

limbs. Reconstructing player position using the centroid (reconstruction plane elevated to 50% stature) was shown to improve reconstructed player position estimates for camera views typical of sport stadia.

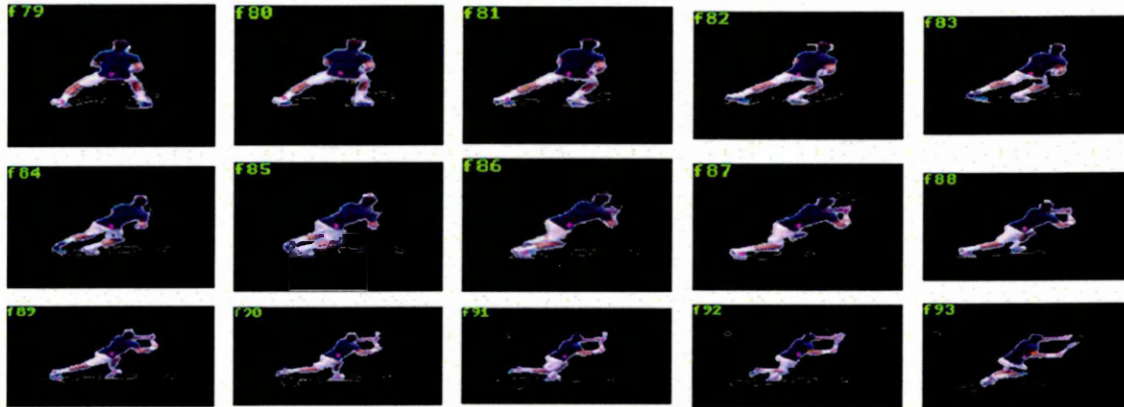
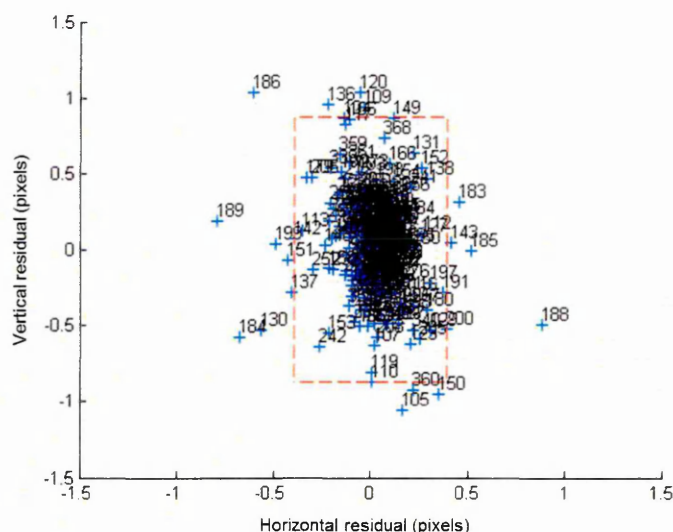


Figure 8.20. Sequence of foreground player images with RGB pixel values and shadow perimeter superimposed for illustration.

Figure 8.20 illustrates foreground player segmentation for a split-step to running forehand groundstroke manoeuvre. This is a high acceleration movement that is typical of match-play tennis (Fernandez, Mendez-Villanueva and Pluim, 2006). The segmented player is centred within cropped images that are tracked and resized appropriately, reflecting the player's changing shape and location relative to the camera. Foreground player segmentation was performed in both RGB and HSV colour spaces to reduce the quantity of shadow present within segmented images. The majority of shadow (shadow perimeter superimposed in grey: Figure 8.20) was removed. However some shadow was still present about the feet (Figure 8.20). Shadow removal using the HSV colour space was successful for laboratory conditions, e.g. laboratory validation of FSCi algorithm (Chapter 7). However match-play tennis performed indoors requires high intensity lighting, i.e. 1076 LUX (ITF, 2013). Shadows, particularly underfoot, will therefore be prominent. The PT-FSCi algorithm allows segmentation thresholds to be adjusted using sliders (B: Figure 8.17). However variation in player movement will make it difficult to manually select an appropriate segmentation threshold. Shadow represents a limitation to the correct identification of foot-surface contacts. Colour filtering using the HSV colour space was used because HSV has been reported as a fast and effective method of shadow removal for indoor applications (Benedek and Szirányi, 2007; Nghiem and Thonnat, 2008). However shadow removal for other applications, e.g. outdoor tennis,

has not been assessed. Colour filtering using the CIE L*u*v colour space might represent a better approach for shadow removal for outdoor applications (Benedek and Szirányi, 2007). However image conversion to CIE L*u*v will also increase computational demand (Nghiem and Thonnat, 2008).



Frame 189 was flagged for inspection (left of cluster centre: Figure 8.21). Figure 8.22 (A) illustrates the corresponding foreground player image with RGB pixel values and

shadow perimeter superimposed for illustration. Image B (Figure 8.22) illustrates a zoomed in region of image A. Player position, i.e. image centroid, is represented by the black and white rings in the centre of the magenta diamond (B: Figure 8.22). The green crosshair and dashed rectangle represent predicted player position and 95% confidence intervals (B: Figure 8.22). Player position, located to the left of the predicted player position, reflects incorrect segmentation of the racket arm. Assessing player tracking data provided a simple method to automatically flag potential player tracking errors to the user. However tracking residuals greater than 95% confidence intervals are not necessarily indicative of player segmentation error. Furthermore the appropriateness of data smoothing is dependent on original data: the application of generic smoothing splines to other tennis player tracking data will be limited. Therefore the user should review all segmented images using the 'FSCi Inspector' module.

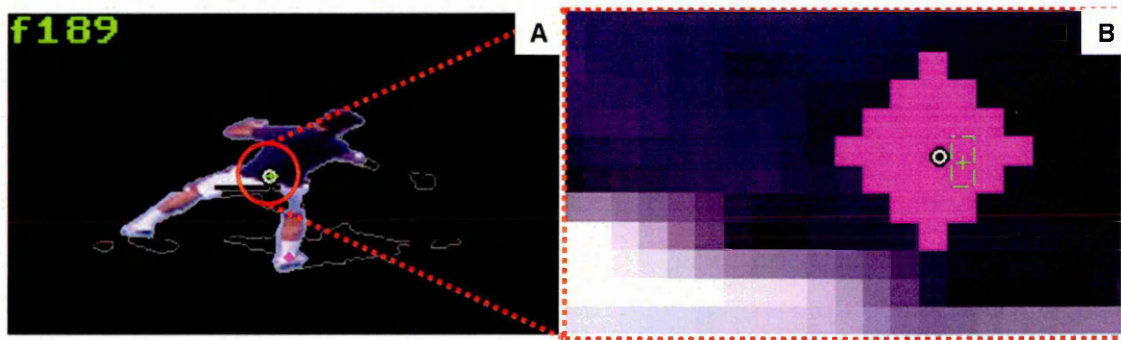


Figure 8.22. Player foreground image (A) and zoomed-in region (B).

For foot-surface contact analysis, the PT-FSCi algorithm applied each player tracking window to the previous image within the image sequence, i.e. $PrePlayer_{RGB}(x, y, t_n)$. This enabled the quantification of inter-frame differences. Accordingly, the FSCi algorithm developed in section 6.4 (Chapter 6) was executed alongside the player tracking algorithm. Shadow removal (described in section 6.4.3) was omitted because shadow removal was performed for player tracking (section 8.3.3). Figures 8.20 and 8.22 demonstrate that the PT-FSCi algorithm identified foot-surface contacts during match-play tennis (magenta diamonds about stance feet). However, due to the nature of the FSCi algorithm, the presence of shadow underfoot will increase the rate of false foot-surface contacts, i.e. identification of low inter-frame motion. Frame 89 (f89: Figure 8.20) illustrates that shadow was identified as a foot-surface contact. Shadow might therefore limit the accuracy foot-surface contacts identified by the PT-FSCi

algorithm. An accuracy assessment is required to identify the validity of both player tracking and foot-surface contact estimates obtained by the PT-FSCi algorithm, using footage of match-play tennis.

8.5. Conclusion

This chapter describes the development of a vision-based method to track tennis player position and identify foot-surface contacts using single camera footage of match-play tennis. Furthermore, a graphical user interface was developed to provide a simple, semi-automatic analysis tool. The tool visualised output data, allowed the user to interact with the algorithm and exported relevant data for further analysis. Potential limitations to the current PT-FSCi algorithm have been highlighted.

In relation to the overall project aim, the PT-FSCi algorithm has been demonstrated to track tennis player position and identify foot-surface contacts without interfering with the activity being observed. The PT-FSCi algorithm segmented and tracked a tennis player during high acceleration manoeuvres that are typical of match-play tennis. Player tracking residuals (relative to a smoothing spline) were less than two pixels for both vertical and horizontal image coordinates. Furthermore, foot-surface contacts were successfully identified during multi-modal gait of match-play tennis; however, success rate and accuracy must be evaluated. The PT-FSCi algorithm allows the *in situ* measurement of tennis player step and movement strategy and represents a novel approach for characterising match-play tennis. Furthermore, the semi-automatic execution of the PT-FSCi algorithm allows objective, frame-by-frame analyses which can reduce analysis time and random error associated with manual digitising. An accuracy assessment is required to identify the validity of tennis player position and foot-surface contact data using match-play tennis footage.

9 Validation of a semi-automatic technique for player tracking and foot-surface contact identification at the 2011 ATP World Tour Finals

9.1. Introduction

A manual system for characterising player step and movement strategy in match-play tennis was developed in Chapter 3. Gender differences for forehand manoeuvre step frequency were highlighted; however findings were limited due to low sample size and movement definitions. A frame-by-frame analysis of match-play tennis was required. A semi-automatic technique for player tracking and foot-surface contact identification (PT-FSCi) using single camera footage of match-play tennis was developed in Chapter 8. Furthermore, a graphical user interface (GUI) was developed to provide a simple analysis tool. Following user-input for initialisation e.g. background modelling, extrinsic camera calibration and one-off player identification, the PT-FSCi algorithm analysed image sequences automatically. However as noted in Chapter 8, an accuracy assessment is required to identify the validity of tennis player position and foot-surface contact data measured by the PT-FSCi algorithm during match-play tennis.

9.2. Aim and objectives

The aim of this chapter is to validate the semi-automatic technique for tennis player tracking and foot-surface contact identification described in Chapter 8. This relates to boxes G and H of the development stage diagram (Figure 1.2).

Objectives:

1. Validate the PT-FSCi algorithm using footage of match-play tennis rallies.
2. Assess the success rate, analysis time and limitations of the algorithm with reference to previous foot contact detection methods.
3. Quantify step and movement parameters for match-play tennis rallies.

9.3. Match-play tennis validation study

Video footage of the 2011 ATP World Tour Finals, e.g. Figure 9.1, captured as described in section 8.3.1 (Chapter 8), was used to validate the PT-FSCi algorithm. Following filming, camera footage was downloaded to a laptop computer and the time codes for 20 tennis rallies / points were identified for an individual player.

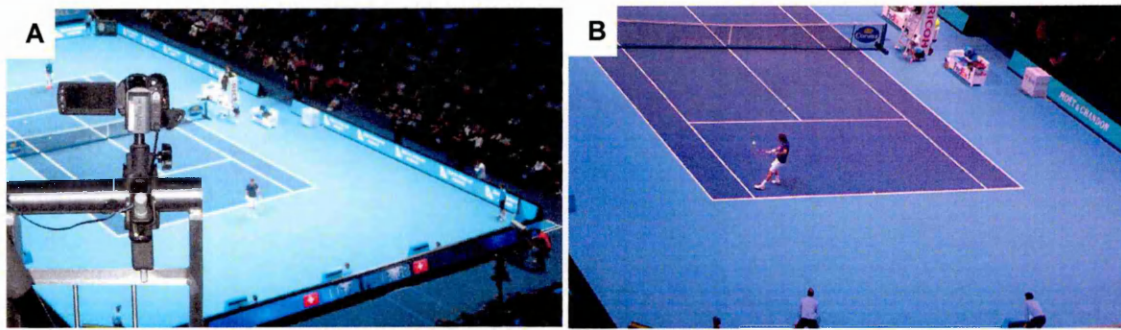


Figure 9.1. Camera setup at the 2011 ATP World Tour Finals (A) and sample camera image (B).

Tennis rally image sequences were extracted and converted to .tiff image files; Table 9.1 summarises the number of images, corresponding duration and number of strokes within extracted tennis rallies.

Table 9.1. Images, duration and strokes for match-play tennis rallies (N = 20).

<i>Images (n)</i>	<i>Time (s)</i>	<i>Strokes (n)</i>
$\bar{x} \pm s$	$\bar{x} \pm s$	$\bar{x} \pm s$
[Range]	[Range]	[Range]
369.9 ± 165.0	14.8 ± 6.6	5.1 ± 2.6
[150 - 700]	[6 - 28]	[2 - 10]

9.3.1. Camera calibration parameters

As described in section 8.3.1, single camera calibration was performed, e.g. Figure 9.2. Checkerboard images were deinterlaced (bob and expand: 1920×1080 pixel images) and processed using the Camera Calibration toolbox for MATLAB (Bouguet, 2010). Prior to calibration, internal camera settings i.e. zoom, focal length, etc., were set manually and were not altered for the duration of match-play tennis filming. Intrinsic camera parameters were therefore valid for all match-play tennis footage.

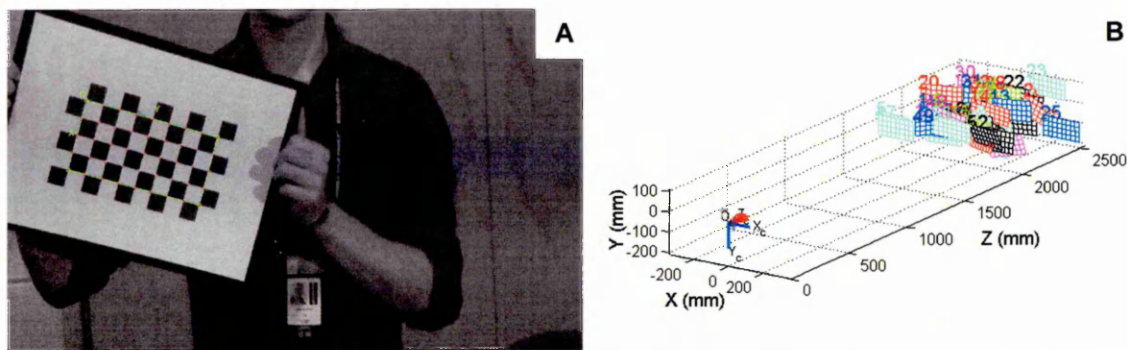


Figure 9.2. Single camera calibration: checkerboard extraction (A) and extrinsic checkerboard parameters (B).

Extrinsic camera parameters were calculated for each rally analysis as described in section 8.3.7. The PT-FSCi algorithm prompted the user to manually digitise four known locations within a camera image (1920×1080 pixel image) and input corresponding real-world dimensions. Furthermore, the user was required to input player stature (1.85 m); player stature was obtained from ATP player profiles (ATP World Tour, 2013). To determine extrinsic parameters, the PT-FSCi algorithm processed manually digitised image coordinates of tennis court markings and corresponding real-world dimensions using the Camera Calibration toolbox for MATLAB (Bouguet, 2010). The PT-FSCi algorithm subsequently calculated two elevated reconstruction planes and corresponding extrinsic camera parameters. For player location and foot-surface contact data, reconstruction plane elevation was 925 mm (50% of stature) and 40 mm (section 7.5.5) respectively.

Table 9.2. Camera calibration residuals and extrinsic camera parameters (n = 20).

	<i>Calibration residuals</i>	<i>Azimuth</i>	<i>Elevation</i>	<i>Resultant Translation</i>	<i>Pixel-scale</i>
	$\bar{x} \pm s$	$\bar{x} \pm s$	$\bar{x} \pm s$	$\bar{x} \pm s$	$\bar{x} \pm s$
<i>X (mm)</i>	0.0 ± 0.1	-	-	-	10.0 ± 0.1
<i>Y (mm)</i>	0.0 ± 0.2	-	-	-	23.2 ± 0.1
<i>R (mm)</i>	-	-	-	50843.8 ± 120.3	-
<i>Angle (°)</i>	-	-121.1 ± 0.2	22.0 ± 0.1	-	-

Table 9.2 summarises camera calibration parameters obtained for the 20 analysed rallies. Furthermore, the camera-plane model (Figure 9.3) illustrates camera position and orientation in relation to the tennis court. Camera azimuth (Table 9.2) was calculated relative to the tennis court's positive X axis (Figure 9.3).

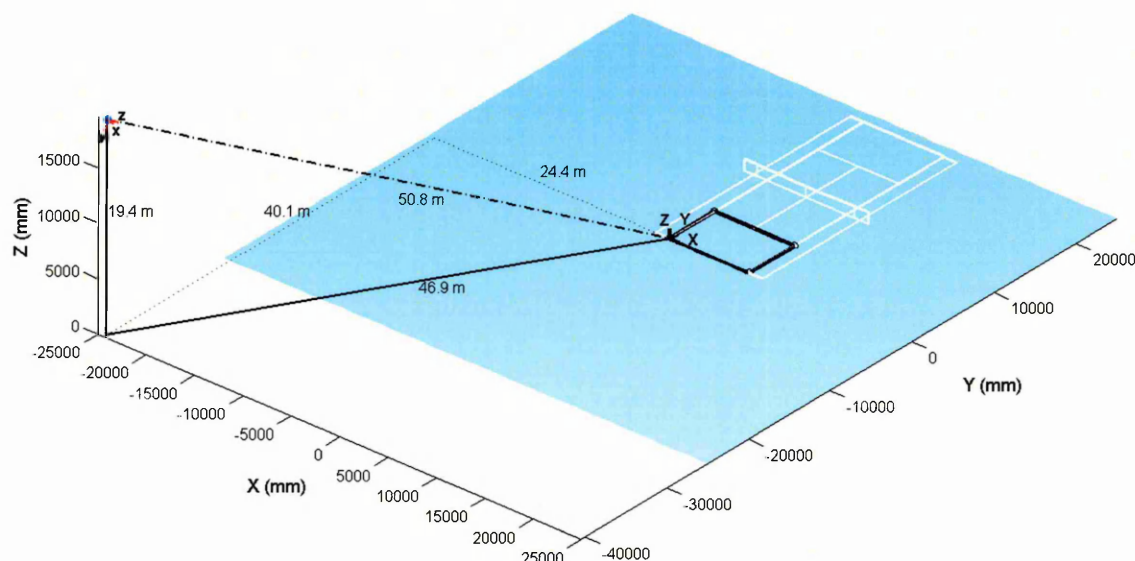


Figure 9.3. Camera-plane model describing extrinsic camera parameters (position and orientation) at the 2011 ATP World Tour Finals.

9.3.2. PT-FSCi tennis rally analysis

Image fields sampled on even rows were removed using VirtualDub (Lee, 2010) yielding 1920×540 pixel images; sample rate was therefore 25 Hz. Following initialisation as described in section 8.3.8 (Chapter 8), the PT-FSCi algorithm executed automatically. The user subjectively assessed player foreground segmentation and adjusted player segmentation thresholds as required. Furthermore, the user reviewed segmented player images and identified data using the 'FSCi Inspector' module. The user flagged images where segmentation or foot-surface contact locations were deemed to be incorrect.

The PT-FSCi algorithm derived player location (PT_{uvt}) and foot-surface contact ($FSCi_{uvt}$) data at a sub-pixel resolution. Position data, i.e. PT_{xyl} and $FSCi_{xyl}$, were reconstructed using corresponding elevated reconstruction planes, i.e. 925 and 40 mm respectively. All image coordinate and world position data were exported to .ascii files. Furthermore, time-codes for player tracking residuals exceeding 95% confidence

intervals as well as user flagged data, i.e. 'FSCi inspector' module, were exported to .ascii files. For reference purposes, segmented player images were exported as .bmp images and all PT-FSCi analysis parameters, e.g. MATLAB workspace data, were exported as .mat files.

9.3.3. Manual tennis rally analysis

All image sequences (1920×540 pixels) were manually digitised at a sub-pixel resolution using Check2D (Centre for Sports Engineering Research, Sheffield, UK). Player COM ($ManPT_{uvi}$) was subjectively determined and digitised on a frame-by-frame basis. Foot-surface contact location ($ManFSCi_{uvi}$) was subjectively determined and digitised at the perceived instant of mid-stance, for all foot-surface contacts. The repeatability of manual digitising was assessed. An image sequence (150 fields) was manually digitised on five separate occasions and standard error of measurement (SEM) quantified. Table 9.3 presents the range for SEM throughout the digitised image sequence. SEM is not reported for player location time due to frame-by-frame digitising. SEM was less than 2 pixels for all player location coordinates and less than 1 pixel for all foot-surface contact coordinates. Furthermore, SEM for foot-surface contact time was equal to or less than 0.05 s.

Table 9.3. Manual digitising standard error of measurement (N = 5).

	u (p)	v (p)	t (s)
	Range	Range	Range
$ManPT_{uvi}$ ($n = 150$)	0.00 - 1.96	0.00 - 1.30	N/A
$ManFSCi_{uvi}$ ($n = 18$)	0.02 - 0.71	0.00 - 0.55	0.00 - 0.05

Manually digitised coordinate data for 20 image sequences were exported to .ascii files. Corresponding camera calibration parameters were used to reconstruct $ManPT_{xyt}$ and $ManFSCi_{xyt}$. The two-dimensional position of image coordinate data was calculated using the planar position reconstruction method identified by equations 4.11 - 4.15 (Chapter 4). Reconstructed $ManPT_{xyt}$ and $ManFSCi_{xyt}$ data were exported to .ascii files.

9.3.4. Data analysis

Manually flagged foot-surface contact data were removed from $FSCi_{xyt}$ data. The spatial clustering algorithm (described in section 7.3.4, Chapter 7) was limited when applied to multi-modal gait such as match-play tennis (discussed in section 7.5.1). Therefore, to assess foot-surface contact data measured by the PT-FSCi algorithm and perform step analysis, the three-dimensional Delaunay Triangulation of $ManFSCi_{xyt}$ data was computed. Euclidean distances between $ManFSCi_{xyt}$ and $FSCi_{xyt}$ were used to cluster $FSCi_{xyt}$ data; the median position and time of clustered $FSCi_{xyt}$ data was deemed a foot-surface contact. For foot-surface contact data, the following dependent variables were quantified:

- Number of identified foot-surface contacts (n)
- Foot-surface contact position (mm)
- Foot-surface contact time (s)

Step length and step time was defined as the absolute difference between successive foot-surface contact position and time data respectively. Erroneous step length and step time estimates, resulting from missing foot-surface contact data, were identified and removed from analyses. The following dependent variables were quantified:

- Step length (mm)
- Step time (s)

Player position estimates were assessed in the X and Y directions. Resultant (R) direction data were also calculated using trigonometry (equation 9.1). Furthermore, the central differencing technique (equation 9.2) was used to calculate player velocity. The following dependent variables were quantified:

- Player position (mm)
- Player velocity ($m \cdot s^{-1}$)

$$R = \sqrt{X^2 + Y^2} \quad [9.1]$$

$$V_i = \frac{X_{i+1} - X_{i-1}}{2\Delta t} \quad [9.2]$$

where V_i is instantaneous velocity, X_i is instantaneous position, t is time and R is the resultant of X and Y direction data.

Agreement was assessed using Bland Altman 95% limits of agreement (LOA). In the case of heteroscedastic data distribution, i.e. $|r^2| > 0.1$, ratio LOA was also reported. Furthermore, root-mean square error (RMSE) was calculated with the following:

$$RMSE = \sqrt{\sum_{i=1}^N (X_{iR} - X_{ir})^2 / N} \quad [9.3]$$

where X_{iR} is the criterion, X_{ir} is the estimate and N is the number of data points.

9.4. Results

Twenty match-play tennis rallies, consisting of 6612 images, were manually digitised and analysed using the PT-FSCi algorithm. Figure 9.4 presents the proportions of foot-surface contact data that were accepted and flagged for removal using the 'FSCi Inspector' module. The 'FSCi Inspector' module was used on 240 occasions (analysis suspended for more than 5 s); foot-surface contact data were removed from 591 of 6612 analysed images (Figure 9.4).

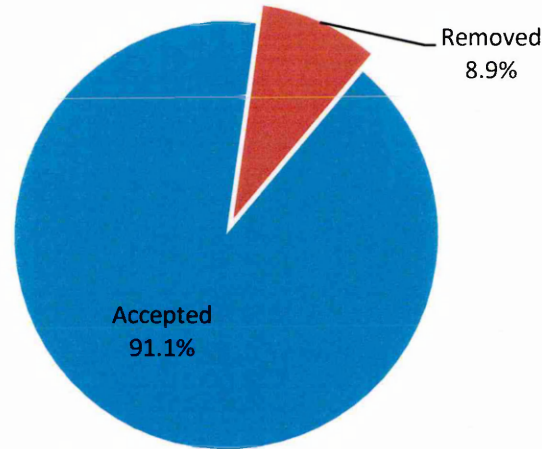


Figure 9.4. Foot-surface contact data (analysed by PT-FSCi algorithm) accepted for analysis or flagged for removal.

Table 9.4 details the proportions of removed foot-surface contact data ($n = 591$). 'Shadow' refers to incorrectly identified foot-surface contacts due to poor shadow removal. 'Flight' refers to foot-surface contacts identified when the foot was not in contact with the surface, i.e. vertex of a jump. 'Misidentification' refers to foot-surface contacts identified about body regions that are not the foot, i.e. lower limb. 'Segmentation' refers to images where inter-frame motion assessment was inappropriate due to missing foot segments.

Table 9.4. Foot-surface contact data removed from analyses ($n = 591$).

	<i>Shadow</i>	<i>Flight</i>	<i>Misidentification</i>	<i>Segmentation</i>
<i>Date removed (%)</i>	28.4	38.6	14.9	18.0

Using a laptop computer (Processor: 2.3 GHz Intel Core i7; Memory: 8 GB RAM), PT-FSCi analysis time per image was 1.23 ± 0.66 s. Manual inspection time (using the 'FSCi Inspector' module) was 14.46 ± 6.15 s. Total PT-FSCi analysis time (6612 images) was estimated as 193.4 minutes and represents analysis time inclusive of manual inspection for all images. Analysis time for manual digitising was 1.96 s per image (average for sample clip of 260 images). Total manual digitising time (6612 images) was estimated as 216.0 minutes.

9.4.1. Player tracking

Figure 9.5 illustrates $ManPT_{uv}$ (mean and standard deviation: green cross and error bars) relative to foreground player coordinates (extracted by the PT-FSCi algorithm) for all analysed images ($n = 6612$). Foreground player coordinates are normalised to the centroid for each foreground player image (colour indicates frequency accumulated in 0.2 pixel^2 bins). Mean and standard deviation $ManPT_{uv}$ was -5.10 ± 10.39 and 4.81 ± 7.31 pixels in the horizontal and vertical directions respectively (relative to foreground player centroid coordinates). Figure 9.5 highlights greater variation in the horizontal direction for manually digitised player location ($ManPT_{uv}$).

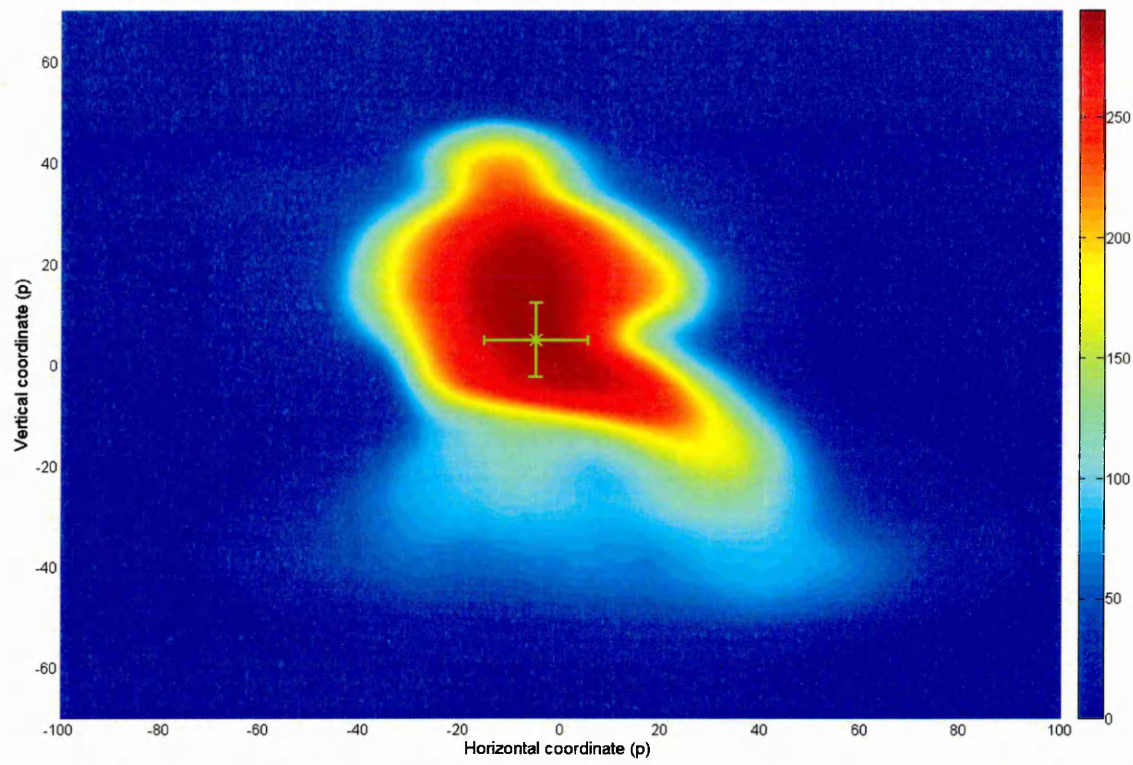


Figure 9.5. Manually digitised player location (green cross and error bars) relative to normalised foreground player coordinates for all analysed images ($n = 6612$). Colour bar indicates foreground coordinate frequency.

For reconstructed player position, LOA indicated systematic differences in the X and Y directions ($Player_X$ and $Player_Y$ respectively: Table 9.5). Variation in player position estimates was similar in both X and Y directions; 95% of estimates were between ± 256.7 and ± 281.3 mm respectively (Table 9.5). For X direction player velocity ($PlayerVel_X$), criterion and estimate differences were heteroscedastic. Log transformation, to assess ratio LOA, yielded complex numbers with imaginary components. It was therefore inappropriate to assess ratio LOA (Bland and Altman, 1986). RMSE for player velocity was larger in the Y direction ($PlayerVel_Y$) than X direction (Table 9.5). For R direction player velocity ($PlayerVel_R$) RMSE was $1.00 \text{ m}\cdot\text{s}^{-1}$ (Table 9.5).

Table 9.5. LOA (absolute and ratio), r^2 and RMSE for player position ($n = 6612$) and player velocity ($n = 6572$) during match-play tennis.

	<i>Absolute LOA</i>	r^2	<i>Ratio LOA</i>	<i>RMSE</i>
	<i>(dimensionless)</i>			
<i>Player_x</i> (mm)	94.0 ± 256.7	0.09	-	161.2
<i>Player_y</i> (mm)	87.6 ± 281.3	0.07	-	168.2
<i>PlayerVel_x</i> (m·s ⁻¹)	-0.01 ± 1.87	-0.11	N/A	0.96
<i>PlayerVel_y</i> (m·s ⁻¹)	0.00 ± 2.80	-0.04	-	1.43
<i>PlayerVel_R</i> (m·s ⁻¹)	0.00 ± 1.96	0.05	-	1.00

9.4.2. Foot-surface contact and step parameters

Figure 9.6 presents the proportion of identified foot-surface contacts following manual data removal and manual clustering (described in section 9.3.4). The PT-FSCi algorithm identified data for 832 of 890 (93.5%) foot-surface contacts during match-play tennis rallies.

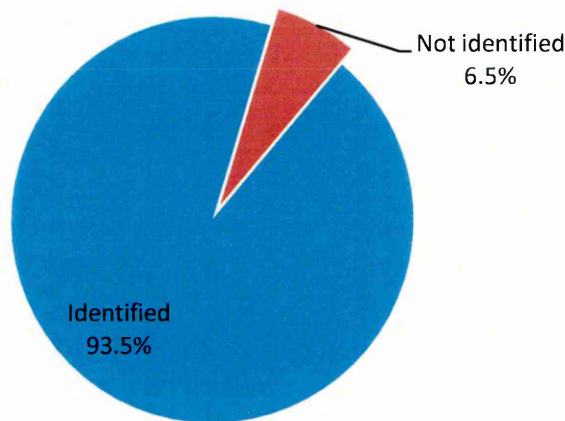


Figure 9.6. Proportion of foot-surface contacts identified during match-play tennis.

For reconstructed foot-surface contact position, LOA indicated small systematic differences in the X (FSC_X) and Y (FSC_Y) directions (Table 9.6). FSC_X and FSC_Y (Table 9.6) indicate that, for the resultant direction, 95% of foot-surface contact estimates were between ± 229.4 mm and RMSE was 121.9 mm. For the time of foot-surface contact (FSC_T), LOA indicated a small systematic difference (0.0019 s) with 95% of foot-surface contact time estimates between 0.15 s (Table 9.6). Furthermore, RMSE for FSC_T was 0.07 s.

Table 9.6. LOA (absolute and ratio), r^2 and RMSE for foot-surface contact parameters (n = 832) during match-play tennis.

	<i>Absolute LOA</i>	r^2	<i>Ratio LOA</i>	<i>RMSE</i>
	<i>(dimensionless)</i>			
FSC_X (mm)	3.1 ± 149.5	0.00	-	76.3
FSC_Y (mm)	-34.2 ± 174.0	0.03	-	95.1
FSC_T (s)	0.00 ± 0.15	0.03	-	0.07

For step length estimates, LOA indicated a small systematic difference (0.08 mm) with 95% of step length estimates between 194.4 mm (Table 9.7). Furthermore, step length RMSE was 99.1 mm. Step time differences were heteroscedastic. Log transformation, to assess ratio LOA, yielded complex numbers with imaginary components. It was therefore inappropriate to assess ratio LOA (Bland and Altman, 1986). RMSE for step time was 0.11 s (Table 9.7).

Table 9.7. LOA (absolute and ratio), r^2 and RMSE for step length and step time during match-play tennis (n = 762).

	<i>Absolute LOA</i>	r^2	<i>Ratio LOA</i>	<i>RMSE</i>
	<i>(dimensionless)</i>			
<i>Step Length</i> (mm)	0.08 ± 194.4	0.06	-	99.1
<i>Step Time</i> (s)	0.01 ± 0.21	0.17	N/A	0.11

9.4.3. Player step and movement characterisation

Twenty match-play tennis rallies of a finals match at the 2011 ATP World Tour Finals in London were analysed. Tennis rallies were 14.8 ± 6.6 s in duration and consisted of 5.1 ± 2.6 strokes (Table 9.1). Figure 9.7 illustrates reconstructed player and foot-surface contact position data for a sample rally, relative to the camera-plane model (Figure 9.3). Player (green dotted line) and foot-surface contact (red filled circles) position were reconstructed using different (925 and 40 mm elevation respectively) reconstruction planes. The accuracy of alignment between player and foot-surface contact position data cannot be quantified for current data. However visual alignment illustrates the efficacy of reconstructing out-of-plane motion with elevated reconstruction planes. Figure 9.7 illustrates that the player started the rally on the deuce court (right of centreline), moving to the advantage court (left of centreline) and returning to the deuce court. The rally was predominantly performed behind the baseline.

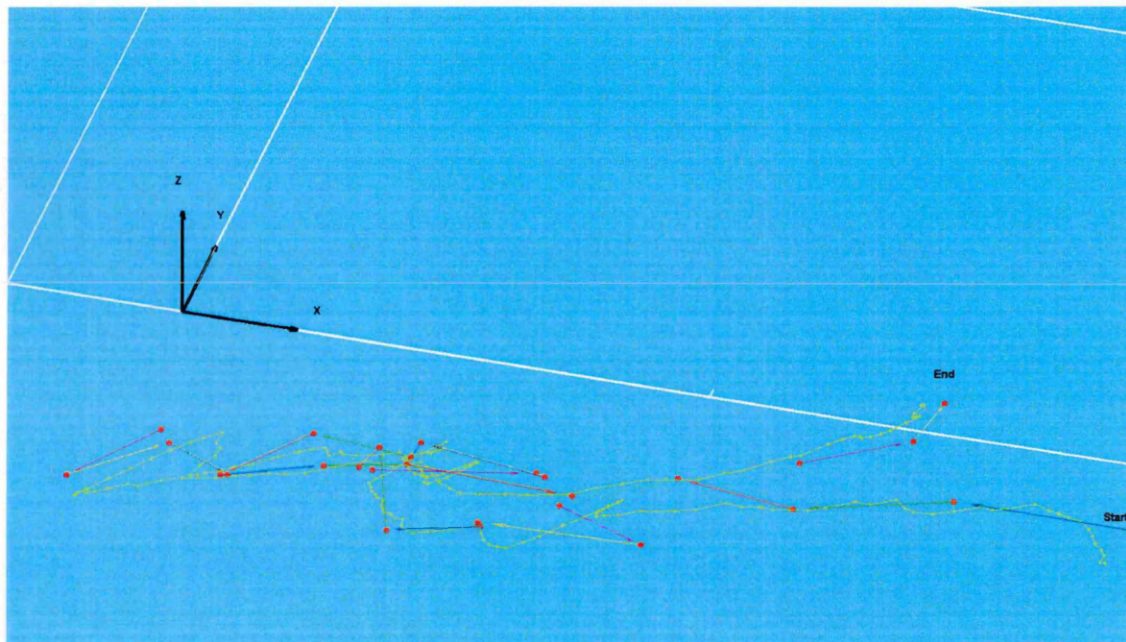


Figure 9.7. Player position (green dotted line) and foot-surface contact (red filled circles) data relative to the camera-plane model (zoomed-in) for a single rally.

Figure 9.8 illustrates player-court occupancy for 20 match-play tennis rallies. Colour indicates occupancy duration (seconds) accumulated within 50 cm^2 bins. Figure 9.8 illustrates that, for the rallies analysed, rallies were predominantly performed on the advantage court (left-hand side). This reflects the manual selection of analysed tennis

rallies, e.g. player was closer to the camera (Figure 9.3). However Figure 9.8 also illustrates that rallies were predominantly performed from behind the baseline

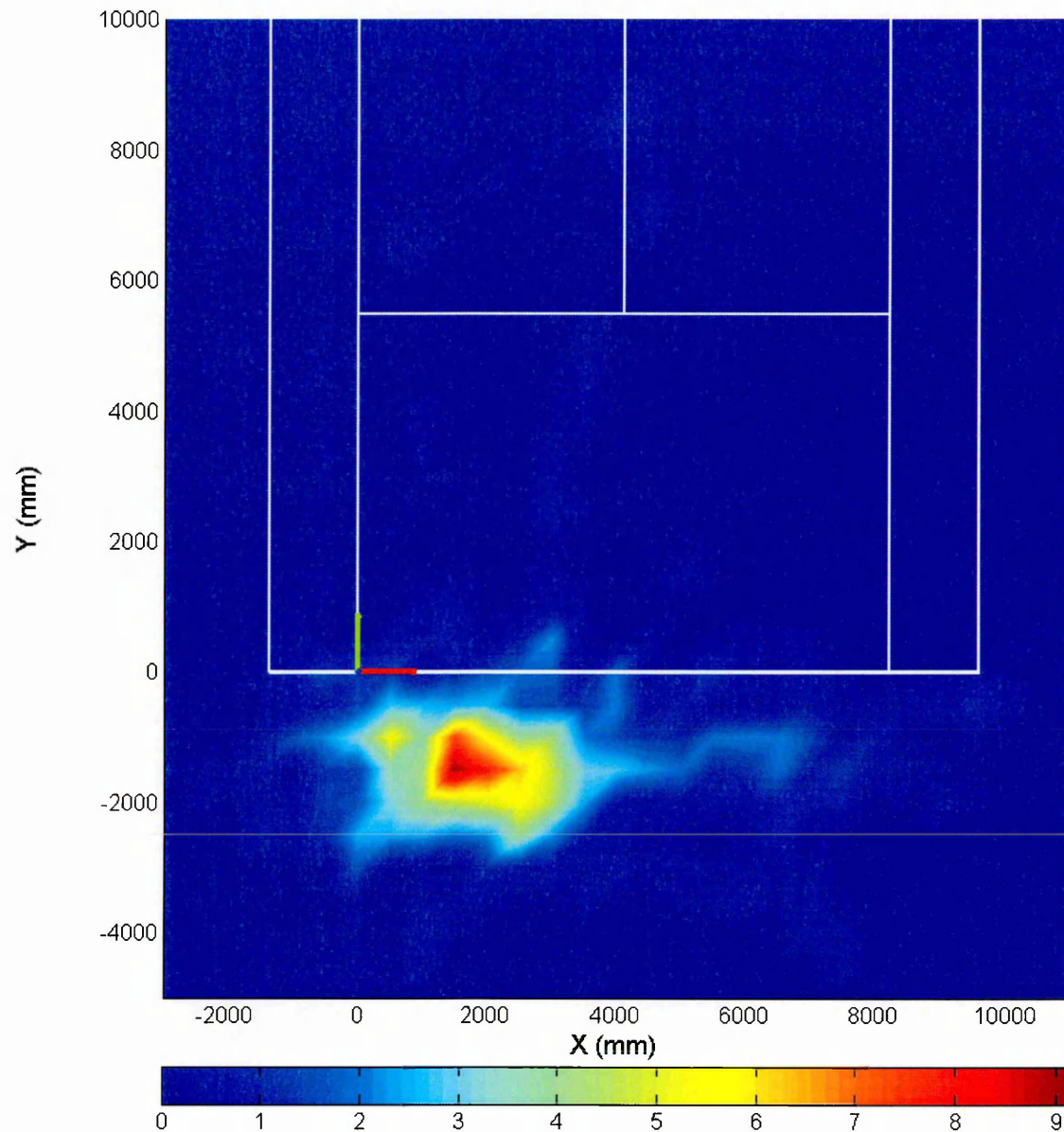


Figure 9.8. Player-court occupancy map (colour bar indicates duration in seconds) for match-play tennis rallies (n = 20).

Table 9.8 presents player travel, speed and step parameters measured during match-play tennis rallies. Tennis player movement (resultant direction) during rallies was characterised by mean player speeds of $1.63 \text{ m}\cdot\text{s}^{-1}$, covering a mean distance of 32.99 m (Table 9.8). Baseline player travel, i.e. X direction (Figure 9.3), was the largest component of resultant direction player travel (Table 9.8). Furthermore, mean player speed along the baseline was also the largest component of resultant direction player

speed (Table 9.8). Tennis player step strategy during rallies was characterised by mean step lengths of 0.87 m at a mean rate of 3.4 steps per second (Table 9.8). Furthermore, the mean number of steps taken during rallies was 38.1 (Table 9.8).

Table 9.8. Mean and standard deviation player travel, absolute player velocity, step number, step length and step rate for match-play tennis rallies (n = 20).

	<i>Player Travel</i>	<i>Player Speed</i>	<i>Steps</i>	<i>Step Length</i>	<i>Step Rate</i>
	(m)	(m·s ⁻¹)	(n)	(m)	(Hz)
	$\bar{x} \pm s$	$\bar{x} \pm s$	$\bar{x} \pm s$	$\bar{x} \pm s$	$\bar{x} \pm s$
X	25.25 ± 12.37	1.88 ± 1.66	-	-	-
Y	20.80 ± 10.38	1.43 ± 1.19	-	-	-
R	32.99 ± 15.55	1.63 ± 1.26	38.1 ± 20.8	0.87 ± 0.39	3.4 ± 6.5

Figure 9.9 illustrates the distribution of player speed (resultant direction). Player speed estimates were positively skewed (skewness = 1.20): 25th, 50th and 75th percentiles for player speed were 0.64, 1.37 and 2.35 m·s⁻¹ respectively. Furthermore, 95% of resultant direction player speed data were less than 4.07 m·s⁻¹.

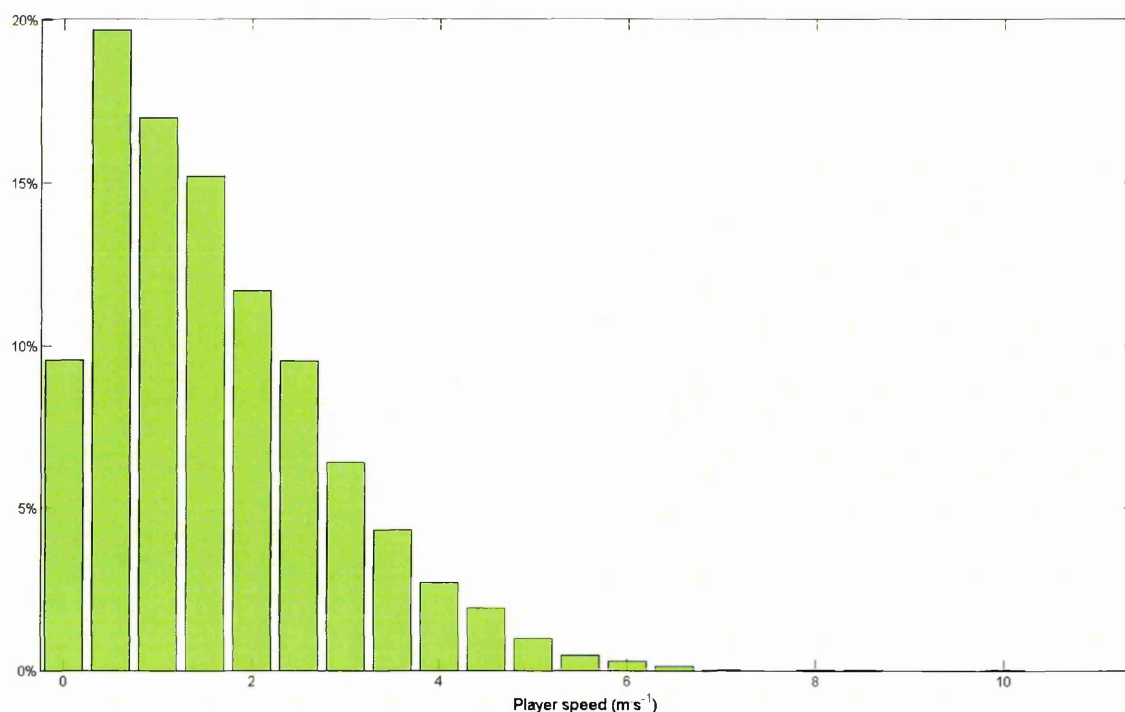


Figure 9.9. Distribution of tennis player speed (resultant direction) during match-play tennis rallies (n = 20).

Figure 9.10 illustrates the distribution of step lengths measured during match-play tennis rallies (n = 20). A small positive skew (skewness = 0.39) was apparent for step length measurements: 25th, 50th and 75th percentiles for step length were 0.61, 0.88 and 1.12 m respectively.

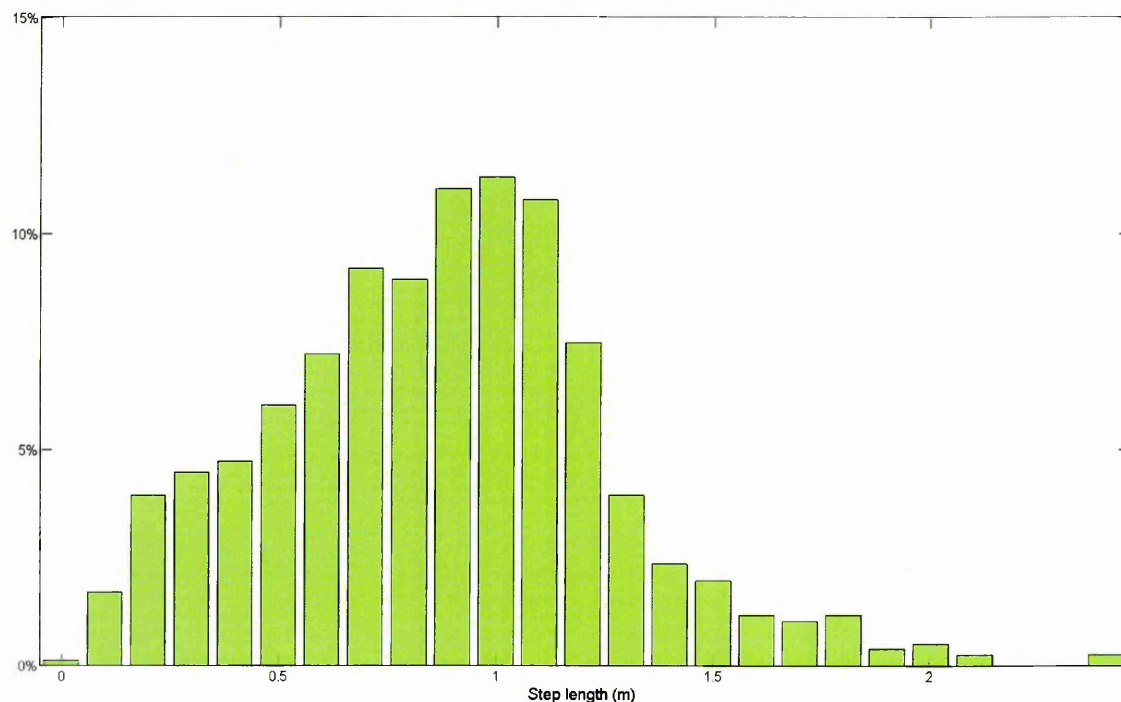


Figure 9.10. Distribution of step length during match-play tennis rallies (n = 20).

Figure 9.11 illustrates player location for step lengths equal to or greater than 1.12 m to provide a spatial representation of the largest 25% of step lengths. Colour indicates frequency (n) accumulated using 50 cm² bins. Figure 9.11 indicates that the largest 25% of step lengths (equal to or greater than 1.12 m) were predominantly located behind the baseline, aligned with both the deuce and advantage court singles sidelines. When considered in relation to Figure 9.8, Figure 9.11 illustrates that larger step lengths were predominantly observed at the extremes of lateral player movements.

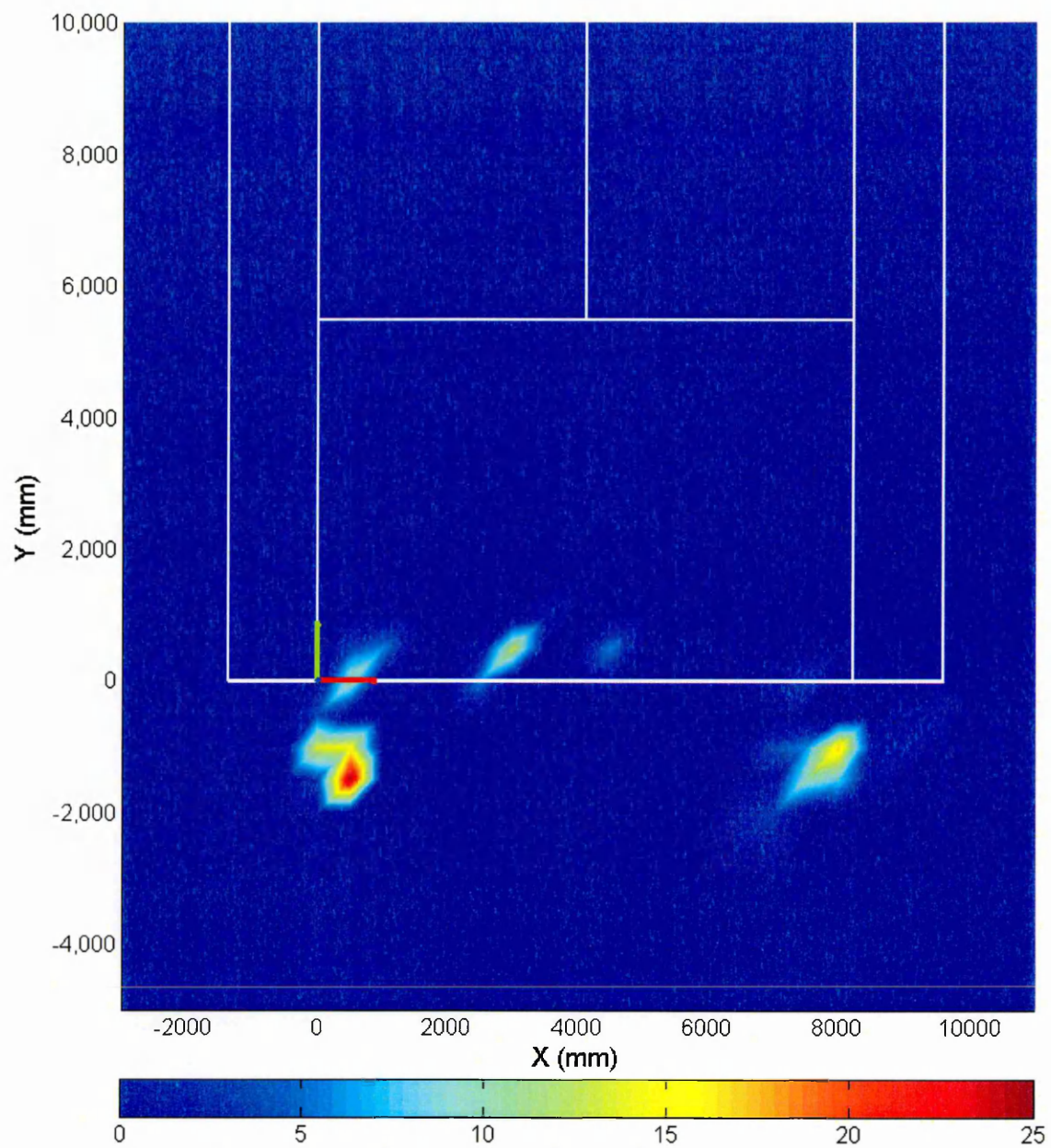


Figure 9.11. Player-court location map for step lengths equal to or greater than 1.12 m (colour bar indicates frequency) for match-play tennis rallies ($n = 20$).

9.5. Discussion

The PT-FSCi algorithm was developed to identify player and foot-surface contact position during tennis rallies without interfering with match-play, i.e. player instrumentation. An accuracy assessment was performed to identify the analysis time and validity of data obtained by the PT-FSCi algorithm.

9.5.1. Analysis time and user intervention

Manual digitising time was 1.96 s per image: total manual digitising time was estimated as 216.0 minutes for all 20 tennis rallies (6612 images). However manual digitising time was based on the average digitising time for a sample clip of 260 images: rally image sequences ranged from 150 - 700 images (Table 9.1). Glazier and Irwin (2001) highlighted that manual digitising drastically increases processing time. Furthermore, the highly repetitive nature of manual digitising will exacerbate random errors associated with manual digitising (Glazier and Irwin, 2001). In comparison, analysis time for the PT-FSCi algorithm was 1.23 ± 0.66 s per image. With the inclusion of manual user inspection time, i.e. 'FSCi inspector' module, total PT-FSCi analysis time was estimated as 193.4 minutes for all 20 rallies. Therefore the PT-FSCi algorithm yields a small time advantage (approximately 32.6 minutes) for tennis rally analyses. As noted in section 7.5.1 (Chapter 7), programming language can have a significant impact on analysis time. Matuska, Hudec and Benco (2012) demonstrated that OpenCV (Open Source Computer Library written in C++) performed image processing operations 4 - 30 times faster than MATLAB; PT-FSCi algorithm analysis time can therefore be reduced further.

PT-FSCi analyses can be performed objectively; this represents a fundamental advantage for the characterisation of match-play tennis. However the PT-FSCi algorithm also allows the user to adjust foreground segmentation thresholds and remove analysed data if necessary. Section 8.3.8 (Chapter 8) noted that pixel intensity variation for blurred or out-of-focus images was lower than sharp or in-focus images. Due to a perspective camera view, i.e. Figure 9.3, player segmentation can vary due to different player-camera distances, i.e. image focus. Incorrect player segmentation can subsequently violate assumptions of the PT-FSCi algorithm and yield erroneous foot-surface contact data. Table 9.4 provides a breakdown of erroneous foot-surface contact data that have been manually removed from analyses. Shadow removal was successful

for laboratory lighting; however match-play tennis requires high intensity lighting, i.e. 1076 LUX (ITF, 2013). Image A (Figure 9.12) illustrates a dark shadow about the lower-limb that has been not been segmented correctly. Shadow can cause the PT-FSCi algorithm to identify false foot-surface contacts due to low inter-frame differences: 28.4% of removed foot-surface contact data were due to shadow. Furthermore, it was noted in section 8.4 (Chapter 8) that shadow removal was developed for indoor applications. Foot-surface contact identification for outdoor applications should therefore be assessed prior to analysis using the PT-FSCi algorithm.

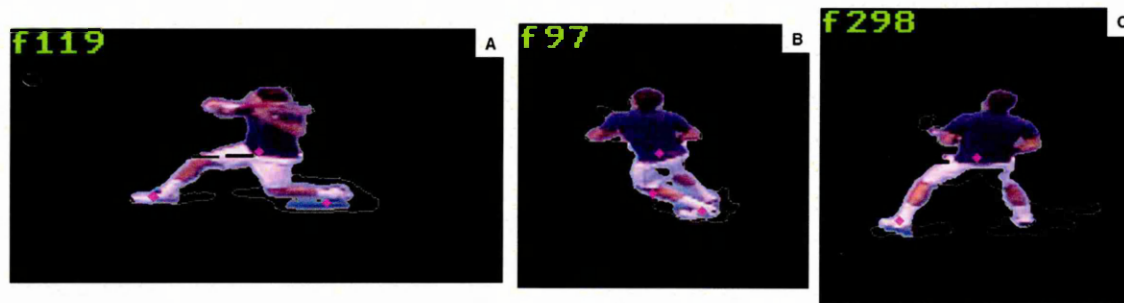


Figure 9.12. Segmented player images and incorrectly identified foot-surface contact data (magenta diamonds about lower-limbs).

'Segmentation' and 'Misidentified' errors (18.0 and 14.9% of removed data respectively) reflect foot-surface contacts identified about the shank, i.e. image B (Figure 9.12). Foot regions can be identified about the shank due to a missing foot segment, i.e. segmentation error, or the failure of geometric rules, i.e. 'Misidentified' error illustrated by image B (Figure 9.12). Misidentified errors highlight that geometric rules are based on vertical image metrics: player geometry in image B (Figure 9.12) was such that the foot-surface contact candidate was not rejected because its location was inferior of lower-body COM (described in section 6.4.5, Chapter 6). Furthermore, the direct application of the PT-FSCi algorithm to match-play tennis derived errors not considered during algorithm development for walking and running. Image C (Figure 9.12) illustrates a 'Flight' error: the feet are momentarily stationary at the vertex of a jump during a split-step manoeuvre and a foot-surface contact is recorded. The split-step is a common manoeuvre in match-play tennis and represented the largest component (38.6%) of foot-surface contact data removed from analyses. False positive foot-surface contacts (Table 9.4) reflect limitations of applying the PT-FSCi algorithm directly to match-play tennis footage. However 91.1% of images analysed by the PT-FSCi

algorithm were accepted for analysis (Figure 9.4): previous single camera heel-strike detection algorithms have only been applied to walking, e.g. Bouchrika and Nixon (2006), Jung and Nixon (2013).

9.5.2. Player tracking assessment

Objective accuracy assessments are difficult to perform during competition. It is impracticable to instrument players during match-play tennis. Furthermore, the accuracy of tracking devices is questionable. For example, O'Conaire *et al.* (2009) noted reliability problems using a UbiSense 3D position tag-tracking system and Duffield *et al.* (2010) reported that GPS devices underestimate distance and speed and exhibit low inter-unit reliability. Manually digitising ground truth data is therefore a common approach for assessing player tracking accuracy in competitive sports, e.g. Mauthner *et al.* (2008); Morais *et al.* (2012). However, markerless identification of player location exacerbates random error due to changing player image size, shape and perspective, e.g. subjective nature of manual digitising (Glazier and Irwin, 2001). The repeated digitising of a sample match-play tennis rally (Table 9.3) revealed SEM was less than 2 pixels for player location coordinate (horizontal and vertical direction). The mean area of foreground player images extracted using the PT-FSCi algorithm, i.e. Figure 9.5, was 4033 ± 564 pixels. Therefore SEM for repeated manual digitising was less than 0.05% of foreground player area and represents an acceptable error margin.

Figure 9.5 illustrates normalised foreground player image coordinates (extracted by the PT-FSCi algorithm) for all analysed tennis rally images. Manually digitised player locations (green cross and error bars) illustrate that subjective estimates of player COM were similar to centroid coordinates for extracted foreground pixels of corresponding images (Figure 9.5). The PT-FSCi algorithm defines player position using this centroid coordinate (section 8.3.7, Chapter 8). Following position reconstruction, error intervals for X and Y direction player position were 94.0 ± 256.7 and 87.6 ± 281.3 mm respectively (Table 9.5). Furthermore, player position differences were homoscedastic, i.e. $|r^2| < 0.1$ (Table 9.5). This indicates that player position estimates were independent of player location on the tennis court.

For the PT-FSCi algorithm, player position RMSE was 161.2 and 168.2 mm in X and Y directions respectively, corresponding to 232.9 mm in the resultant (R) direction.

Mauthner *et al.* (2008) developed a single camera system for tracking volleyball players in competition. Mauthner *et al.* (2008) reported mean position errors of approximately 300 mm (compared to manually digitised ground truth data) and indicated that player tracking errors were acceptable for sport science applications. For the PT-FSCi algorithm, player velocity error intervals were $0.00 \pm 1.96 \text{ m}\cdot\text{s}^{-1}$ in the resultant direction (Table 9.5). Furthermore, resultant direction RMSE was $1.00 \text{ m}\cdot\text{s}^{-1}$ (Table 9.5). Mauthner *et al.* (2008) did not calculate player velocity from player position data due to unrealistic player trajectories resulting from jumps, i.e. perspective projection error. Projection error, a result of out-of-plane motion, limits the validity of player position estimates because the assumption of coplanarity is violated.

The use of an elevated reconstruction plane (50% stature) reduces out-of-plane distance relative to player COM; therefore player COM can be used to estimate player position in perspective projected images. It was demonstrated in section 5.4 (Chapter 5) that the use of an elevated reconstruction plane improved player position estimates when compared to ground plane definitions, e.g. Mauthner *et al.* (2008). Furthermore, for camera elevation angles equal to or greater than 15° , maximum player position reconstruction error (inclusive of $\pm 200 \text{ mm}$ of out-of-plane motion, e.g. Figure 5.8), was lower than for ground plane definitions, e.g. Mauthner *et al.* (2008). This highlighted that player position estimates using an elevated reconstruction plane were not dependent on the perspective projection of the lower-limbs, i.e. player defined by image centroid. PT-FSCi algorithm estimates for player position will not be devoid of out-of-plane error. However the use of an elevated reconstruction plane reduces the impact of out-of-plane motion on player position estimates when compared to existing ground plane definitions. Current data demonstrate agreement with manually digitised ground truth data. Furthermore, position errors are acceptable for sport science applications (Mauthner *et al.*, 2008).

9.5.3. Foot-surface contacts and step parameters

Bouchrika *et al.* (2009) demonstrated that walking direction (relative to camera) affects gait recognition rate for walking. Most research only considers fronto-parallel views of walking (Jung and Nixon, 2013). When applying a heel-strike detection algorithm to random direction walking, Jung and Nixon (2013) reported heel-strike identification rates of 93.4 and 93.7% for PETS (2006) and CAVIAR (2004) databases respectively.

Tennis player movements are both multi-directional and multi-modal; tennis players walk, jog, run and sprint during competition (Hughes and Meyers, 2005). Furthermore, Hughes and Meyers (2005) highlighted that tennis player movements consist of side steps, split steps, skip steps, shuffle steps, cross-overs and lunges. Following the manual removal of false positive foot-surface contact data (described in section 9.5.1), the PT-FSCi algorithm identified data for 832 of 890 (93.5%) foot-surface contacts during match-play tennis.

Due to the multi-modal nature of tennis, the spatial clustering algorithm described in section 7.3.4 (Chapter 7) was not used. This was because foot-surface contacts during tennis rallies can exist at two locations at the same time instant, i.e. split-step or lunge etc., or at the same location at different time instants, i.e. shuffle steps (rapid and repeated steps about the same location for the purpose of balance). Manual intervention for data clustering limits the PT-FSCi algorithm as an automated tennis rally analysis tool. Current knowledge of tennis player-surface interaction is limited (Miller, 2006). Developing a clustering algorithm, suitable for all foot contact types, would require the characterisation of tennis player step and movement strategy to identify relevant clustering parameters. However, the identification and segmentation of large player movements, e.g. indicative of progression between steps, might enable spatial clustering of foot-surface contact data described in section 7.3.4 (Chapter 7). Player movement segmentation could be performed manually by trimming tennis rally clips or by using player displacement data obtained by the PT-FSCi algorithm.

For manually clustered foot-surface contact data, 95% of estimates were between ± 149.5 and ± 174.0 mm for X and Y directions respectively (Table 9.6), corresponding to ± 229.4 mm in the resultant (R) direction. Current data do not support the interpretation of position error in relation to foot orientation. However image resolution might have affected position error, particularly in the Y direction. Deinterlaced images (even rows discarded) were used for analysis; pixel-scale (Table 9.2) indicates that calibration plane resolution was lower in the Y direction (Figure 9.3). Lower vertical image resolution would have exacerbated corresponding position errors. Image deinterlacing was necessary to extract appropriate foreground player images; however reduced vertical image resolution will limit the accuracy of current position data. Future analyses should therefore use cameras that support progressive scanning to avoid image interlacing.

Table 9.9. Resultant direction RMSE (mm) for foot-surface contacts.

	<i>Walking</i>	<i>Running</i>	<i>Match-play tennis</i>
	(<i>n</i>)	(<i>n</i>)	(<i>n</i>)
<i>Barefoot</i>	52.1 (391)	91.4 (251)	-
<i>Shod</i>	52.2 (362)	103.4 (238)	121.9 (832)

Foot-surface contact position differences were homoscedastic (Table 9.6); this indicates that foot-surface contact estimates were independent of player location on the tennis court. For straight line walking, Jung and Nixon (2013) reported heel-strike position errors of ± 100 mm. Foot-surface contact position errors for match-play tennis were therefore larger than an existing walking-based algorithm, e.g. Jung and Nixon (2013). Table 9.9 summarises RMSE for foot-surface contact position in the resultant direction during match-play tennis as well as for walking and running (presented in section 7.5.3, Chapter 7). As noted in section 7.5.3 (Chapter 7), greater RMSE for running compared to walking reflected greater variation in foot contact type. The wide variety of foot contact type observed in tennis (Hughes and Meyers, 2005) will therefore exacerbate foot-surface contact errors measured during tennis rallies. For foot-surface contact time (FSC_T), 95% of estimates were between ± 0.15 s; this corresponds to 3.75 frames. Agreement limits for FSC_T differences highlight that the PT-FSCi algorithm did not identify mid-stance consistently. As noted in section 7.5.2 (Chapter 7), the PT-FSCi algorithm does not identify heel-strike or toe-off events; FSC_X , FSC_Y and FSC_T are the median for measured foot-surface contact data. Therefore variation in foot contact type, as well as self-occlusion during tennis specific manoeuvres, will exacerbate errors for both foot-surface contact time and location estimates.

Variation in FSC_T data was reflected by heteroscedastic step time differences (Table 9.7). Log transformation, to assess ratio LOA, yielded complex numbers with imaginary components. This yields 95% limits of agreement that are too far apart rather than too close, leading to erroneous interpretations (Bland and Altman, 1986). The positive relationship for step time differences ($r^2 = 0.17$; Table 9.7) indicated that the PT-FSCi algorithm underestimated short step times and overestimated long step times. The

relationship between step time duration and step time error reflect both the variety of tennis step type, e.g. short duration shuffle steps to long duration lunge steps. User interaction with foot-surface contact detection thresholds via the GUI (described in section 8.3.7, Chapter 8) might yield more central estimates for mid-stance and subsequently step time. However this might increase the rate of false positive foot-surface contact data. For match-play tennis rallies, the identification of mid-stance and thus step time represents a limitation to the PT-FSCi algorithm.

Step length estimates were homoscedastic and therefore independent of step length and step type. When validating a physical walkway, Webster, Witter and Feller (2005) reported that the maximum random error component for step length during walking was 25.1 mm. For match-play tennis, the maximum random error component for step length was 194.4 mm; current data do not support the PT-FSCi algorithm as a physical walkway replacement. However PT-FSCi step length estimates should be interpreted in relation to the flexibility of measurements obtained. No physical equipment or markers were applied to players; therefore measurements were obtained *in situ* without interfering with tennis play. Furthermore, using a single, off-the-shelf camera, foot-surface contacts were identified within an area (polygonal) equal to 121.7 m² whilst the maximum camera-distance of identified data was 63.6 m. The maximum filming volume for the PT-FSCi algorithm's operation has not been identified. However, in relation to this project's aims and objectives (section 1.2, Chapter 1) the PT-FSCi algorithm represents a tool capable of measuring player position and foot-surface contacts in competitive, match-play tennis.

9.5.4. Rally characterisation

Elite tennis player movements have been characterised at the game level, e.g. Martínez-Gallego *et al.* (2013). Martínez-Gallego *et al.* (2013) highlighted that player movement should be characterised at the rally level to further improve the understanding of tennis strategy. Twenty match-play tennis rallies of a single player (ATP ranking: 4) during a finals match at the 2011 ATP World Tour Finals in London were analysed using the PT-FSCi algorithm. Analysed rallies were 14.8 ± 6.6 s in duration; approximately twice the length of average rally lengths previously reported for competitive matches, e.g. O'Donoghue and Ingram, 2001; Fernandez-Fernandez *et al.*, 2008. Furthermore, Fernandez-Fernandez *et al.* (2008) reported that average rallies (women's tennis)

consisted of 2.6 ± 1.6 strokes; approximately half the number of strokes for analysed rallies (Table 9.1). Therefore analysed rallies are not representative of an entire match.

Figure 9.8 illustrates player-court occupancy for analysed rallies, i.e. player position expressed relative to time. Figure 9.8 highlights a preference for baseline play and corresponds to notational analyses of hard court surface tennis play (O'Donoghue and Ingram, 2001). Rallies were predominantly performed on the advantage court; however this reflects the manual selection of analysed tennis rallies, e.g. minimum camera-player distance. To quantify tennis strategy, Martínez-Gallego *et al.* (2013) defined the defensive court as regions more than 1.5 m behind baseline; the rest of the court (to the net) was defined as the offensive court. For elite tennis players (ATP rank between 5 and 113), Martínez-Gallego *et al.* (2013) highlighted that game winners spent more time in the offensive court; game losers were forced to adopt defensive movement strategies. For current data, movement strategy was predominantly offensive (Figure 9.8); the player spent 69.2% (182.9 s) of rallies in the offensive court and 30.8% (81.5 s) of rallies in the defensive court. However winning and losing players analysed by Martínez-Gallego *et al.* (2013) spent 89.7 and 75.8% of rallies in the offensive zone respectively. For current data, 60% of rallies were won. This indicates that top ranked ATP players can win points using movement strategies that would traditionally be considered as defensive.

Tennis is characterised by short, high intensity manoeuvres (Fernandez, Mendez-Villanueva and Pluim, 2006). Player speed data were positively skewed (Figure 9.9) reflecting that tennis player's walk, jog, run and sprint during rallies (Hughes and Meyers, 2005). Using the SAGIT tracking system (Perš *et al.*, 2002), Martínez-Gallego *et al.* (2013) reported median player speeds of $1.33 \text{ m}\cdot\text{s}^{-1}$ (game winners) for elite tennis players (ATP rank between 5 and 113). Furthermore, Martínez-Gallego *et al.* (2013) reported that game winners travelled 84.2 m (median) during games. For current data, median player speed (resultant direction) was $1.37 \text{ m}\cdot\text{s}^{-1}$, suggesting that player speed was similar. However mean player travel for analysed rallies was 32.99 m (resultant direction). This might indicate that top ranked ATP players travel further during match-play tennis. However generalisable conclusions are not suitable due to a small and selective sample; further match analyses are required.

Large variation in tennis step rate reflects that player movements consist of side steps, skip steps, shuffle steps and lunges (Hughes and Meyers, 2005). Step length data were positively skewed (Figure 9.10) indicating that shorter step lengths were used for the majority of rally movements. However large step lengths (equal to or greater than 1.12 m) were predominantly observed behind the baseline and aligned with both the deuce and advantage court singles sidelines (Figure 9.11). This reflects the extremes of lateral player movement, e.g. baseline rallying, and indicates lunging and turning manoeuvres. No direct link between tennis player step length and shoe-surface friction has been identified. However Starbuck *et al.* (2013) reported a greater attack angle and base of support, e.g. Figure 2.5 (Chapter 2), for forehand turning manoeuvres that resulted in larger slide distances. It is beyond the scope of this work to investigate the biomechanics of step length and sliding. However current data could be used to characterise player step and movement strategy during extreme side-to-side turning manoeuvres that have been associated with sliding and tennis injury (Girard *et al.*, 2007). Furthermore, the PT-FSCi algorithm could be used to assess player step and movement strategy during individual strokes or patterns of play to improve the understanding of competitive tennis strategy.

9.6. Conclusion

This chapter describes the validation and application of a vision-based method to track tennis player position and identify foot-surface contacts during match-play tennis. An off-the-shelf, high-definition camera was used to film the 2011 ATP World Tour Finals. The camera was located in stadia seating, providing an elevated, perspective view of tennis matches. Twenty match-play tennis rallies were analysed using the PT-FSCi algorithm: 91.1% of images analysed were accepted for analysis. Images were analysed in 1.23 ± 0.66 s per image: total analysis time was 193.4 minutes; this represented a reduction of 32.6 minutes when compared to manual digitising (216.0 minutes). Player position errors were independent of player location on the tennis court and were considered suitable for player tracking applications. Data for 93.5% of foot-surface contacts were identified: RMSE was 121.9 mm and errors were independent of their location on the tennis court. However foot-surface contact data were manually clustered to perform step analysis. This reflected the wide variety of step patterns in tennis and highlighted a limitation to the analysis protocol. Future analyses could incorporate player movement data to segment large player movements (indicative of progression

between steps) to allow spatial data clustering. Furthermore, the accuracy of foot-surface contact identification for outdoor applications should be assessed.

Step time estimates were related to step time duration and thus not suitable for analysis. However step length estimates were independent of step length: basic gait parameters could therefore be measured during match-play tennis. The characterisation of tennis rallies indicated that player movement strategy favoured baseline rallying. Furthermore, data might indicate that top ranked players can win points using movement strategies traditionally considered to be defensive. Shorter step lengths formed the majority of step strategy; however the largest 25% of steps were predominantly observed about the extremes of lateral player movements, reflecting lunging and turning tennis manoeuvres.

Using a single, off-the-shelf camera, the PT-FSCi algorithm tracked player position and measured basic parameters of gait during match-play tennis without using markers. Foot-surface contact data were identified within a large filming area, e.g. 121.7 m², demonstrating that the PT-FSCi algorithm can measure foot-surface contacts *in situ*, without interfering with match-play. In relation to the overall project aim, this chapter has validated and used a motion analysis tool that measures player step and movement strategy during match-play tennis. The analysis of the 2011 ATP World Tour Finals improves the understanding of step and movement strategy in match-play tennis as no previous research of this type has been performed. Future analyses could identify player step and movement strategy as a function of surface type, gender, weather, season (indoor / outdoor) or rule change. This will further the International Tennis Federation's (ITF) understanding of player step and movement strategy in match-play tennis and advance the ITF's shoe-surface interaction research.

10 Conclusions

10.1. Introduction

This chapter summarises main findings of previous chapters. Current and future uses of techniques developed within this project will be described as well as potential applications outside of tennis.

10.2. Summary of research

10.2.1. Single camera filming

Filming with a single, high-definition camera was necessary to provide a suitable resolution of large capture volumes, e.g. tennis courts. The camera was mounted on a tripod in surrounding stadia seating. Public stadia seating represents a typical and accessible filming location, increasing the flexibility of future analyses. Single camera filming minimised interference with match-play tennis; there was no requirement to instrument tennis players and camera calibration can be performed remotely. Furthermore, the use of a camera with an on-board disk drive enabled the continual filming of tennis matches; an entire match could be captured without accessing the camera to download footage. This enabled the analysis of player step and movement strategy throughout entire matches.

10.2.2. Manual player step and movement strategy analysis

Tennis player step strategy had not previously been quantified for match-play tennis rallies. Forty men's and women's tennis rallies (equivalent of 18000 frames) filmed at the 2011 Roland Garros Qualifying Tournament were manually digitised to characterise step and movement strategy. Findings were consistent with previous notational analyses of player movement and gender differences for forehand manoeuvre step frequency were highlighted. However, findings were limited due to low sample size and simplified movement definitions (used to enable manual analysis). Furthermore, the validity of position reconstruction using 2D-DLT was limited. Frame-by-frame analysis of player step and movement strategy as well as an assessment of position reconstruction techniques was required.

10.2.3. Two-dimensional position reconstruction

Analyses of the 2011 Roland Garros Qualifying Tournament highlighted that wide angle (zoomed-out) filming induced image distortions due to the lens. Position reconstruction using 2D-DLT was limited because 2D-DLT does not account for image distortion. Furthermore, reconstructing coordinates outside of calibration points, e.g. tennis court markings, violate assumptions of 2D-DLT. A planar position reconstruction method that accounts for lens distortion was developed and position reconstruction accuracy compared to 2D-DLT for filming conditions experienced at two international tennis events. Reduced reconstruction errors (RMSE in resultant direction) demonstrated improved reconstruction accuracy for the planar position reconstruction method. It was concluded that the method was an accurate and flexible method for two-dimensional photogrammetry in filming conditions experienced at international tennis events.

10.2.4. Two-dimensional player position reconstruction

Analyses of the 2011 Roland Garros Qualifying Tournament used an elevated reconstruction plane to estimate player position. However, the accuracy of player position estimates required assessment. Standing and running posture point cloud data were reprojected into simulated camera images ($n = 657$) of two international tennis events. The accuracy of player position estimates using an elevated reconstruction plane (50% stature) was compared to an existing, ground level feature approach. For camera elevation angles equal to or greater than 15° , e.g. sport stadia seating, maximum position reconstruction error was lower using an elevated reconstruction plane. It was concluded that player position accuracy was improved using an elevated reconstruction plane.

10.2.5. Foot-surface contact identification

Manual analyses of player step strategy for the 2011 Roland Garros Qualifying Tournament were time consuming and prone to error, e.g. subjective nature of digitising. An automatic foot-surface contact identification (FSCi) algorithm was developed. Using standard colour images, inter-frame differences about each foot were used to automatically identify foot-surface contacts: no markers were required. The FSCi algorithm identified foot-surface contacts for walking and running obtained from four different camera perspectives. The FSCi algorithm was therefore not limited by gait

mode: this represented a clear advantage over existing heel-strike detection algorithms when considering foot-surface contact identification in tennis.

An objective assessment of the FSCi algorithm was performed to identify the validity of foot-surface contact data. The FSCi algorithm was applied to walking and running images sequences (obtained from four camera perspectives) and compared to three-dimensional motion analysis. Images were analysed in 0.87 ± 0.05 s per image: data for 99.6% of foot-surface contacts were automatically identified. Step analysis was performed automatically for 91.3% of image sequences; manual intervention for remaining image sequences primarily reflected limitations to data clustering. For walking and running, foot-surface contact position RMSE was less than 52.2 and 103.4 mm respectively. Greater position errors for running reflected greater variation in foot contact type. Stance time estimates were not suitable analysis parameters because heel-strike and toe-off events are not identified. Estimates for step length and step time demonstrated agreement: step length and step time RMSE was less than 78.7 mm and 0.11 s respectively. Existing heel-strike detection algorithms have only been applied to walking and do not use data for gait analysis. It was concluded that the FSCi algorithm was a flexible approach for *in situ* measurement of basic gait parameters during walking and running.

10.2.6. Tennis player tracking and foot-surface contact identification

To apply the FSCi algorithm to match-play tennis, a player tracking (PT-FSCi) algorithm was developed. The PT-FSCi algorithm was designed to analyse image sequences automatically following user initialisation, e.g. background modelling, extrinsic camera calibration and one-off player identification. A graphical user interface (GUI) provided control over the PT-FSCi algorithm; output data were visualised online. Furthermore, the user could review output data during analyses, e.g. flag incorrect frames. The PT-FSCi algorithm exported player position and foot-surface contact data to .ascii files for further analysis.

An accuracy assessment was performed to identify the validity of player position and foot-surface contact data using match-play tennis footage. Twenty match-play tennis rallies, filmed at the 2011 ATP World Tour Finals, were analysed using the PT-FSCi algorithm and compared to manually digitised data. The PT-FSCi algorithm analysed

images in 1.23 ± 0.66 s per image. Total analysis time for 20 tennis rallies was 193.4 minutes. This represented a reduction of 32.6 minutes when compared to manual digitising (216.0 minutes) – it was noted that OpenCV (computer vision software library) could further reduce this analysis time. Player position RMSE was 161.2 and 168.2 mm in X and Y directions respectively, and considered suitable for sport science applications. Data for 93.5% foot-surface contacts were automatically identified: foot-surface contact position RMSE was 121.9 mm. However, step analysis (foot-surface contact data clustering) was performed manually. This reflected the wide variety of step patterns in tennis and highlighted a limitation to the analysis protocol. The segmentation of player movements would allow foot-surface contact data clustering in future analyses.

In summary, the PT-FSCi algorithm measured basic gait parameters during match-play tennis *in situ*, without interfering with play. Furthermore, foot-surface contacts were identified in a large filming area, equal to 121.7 m^2 . This highlights the flexibility of the PT-FSCi algorithm for characterising tennis player step and movement strategy; however known limitations should be considered for future applications.

10.2.7. Tennis rally analysis using the PT-FSCi algorithm

Twenty match-play tennis rallies of a single player (ATP ranking: 4) during a finals match at the 2011 ATP World Tour Finals were analysed using the PT-FSCi algorithm. A preference for baseline play was highlighted, reflecting previous hardcourt tennis analyses. Furthermore, median player speed ($1.37 \text{ m}\cdot\text{s}^{-1}$) was similar to previous analyses. The analysed player won 60% of analysed rallies and the match. However, 30.8% of movement was 'defensive', e.g. more than 1.15 m behind baseline. This might indicate that top ranked players can win points using movement strategies traditionally considered to be defensive. Step length data were positively skewed, indicating that shorter step lengths formed the majority of step strategy. However the largest 25% of steps were predominantly observed behind the baseline, aligned with deuce and advantage court sidelines, reflecting lunging and turning manoeuvres at lateral extremes of player movement. The PT-FSCi algorithm enabled a novel analysis of player step and movement strategy during match-play tennis. Furthermore data were obtained using single camera footage filmed from stadia seating. This provides a flexible approach for future analyses. Generalised conclusions based on current data are not suitable due to low sample size; further analyses of match-play tennis are required.

10.2.8. Conclusion

This project has developed a single camera, video-based method for tracking player motion and identifying foot-surface contacts in match-play tennis. The method uses off-the-shelf camera footage that can be obtained from public stadia seating. Footage can therefore be collected *in situ*, without interfering with play. The method enabled a novel analysis of a player's step and movement strategy during a finals match of the 2011 ATP World Tour Finals. This improved the understanding of step and movement strategy during match-play tennis: no previous research of this type has been performed. In tennis, future analyses could identify step and movement strategy as a function of surface type, gender, weather, season (indoor / outdoor) or rule change. This will further the International Tennis Federation's (ITF) understanding of player step and movement strategy during match-play tennis and advance the ITF's shoe-surface interaction research.

10.3. Current and future development

The ITF have conducted filming at the 2013 Davis Cup Finals in Belgrade (Serbia vs. Czech Republic) for analysis and development with the PT-FSCi algorithm. Following analysis, player step and movement data will contribute to the ITF's shoe-surface interaction research.

10.3.1. Analysis procedure

Spatial clustering of multi-modal foot-surface contact data was highlighted as a limitation to the PT-FSCi algorithm. Future analyses could incorporate player movement data to segment large player movements (indicative of progression between steps) to allow spatial clustering of foot-surface contact data.

10.3.2. Programming language

Tennis rally analysis using the PT-FSCi algorithm currently yields a small time advantage when compared manual digitising. However it was noted in section 7.5.1 (Chapter 7) that programming language has a significant impact on analysis time. OpenCV (Open Source Computer Library written in C++) has been demonstrated to perform image processing operations 4 - 30 times faster than MATLAB (Matuska, Hudec and Benco, 2012). Therefore analysis time using the PT-FSCi algorithm could be reduced further.

10.3.3. Outdoor filming

Shadow removal was developed for indoor applications. Foot-surface contact identification should therefore be assessed prior to the analysis of outdoor footage using the PT-FSCi algorithm. Colour filtering using the CIE L^*u^*v colour space might represent a better approach for shadow removal (Benedek and Szirányi, 2007); however image conversion to CIE L^*u^*v will increase computational demand (Nghiem and Thonnat, 2008).

10.3.4. Gait analysis tool

The PT-FSCi algorithm was demonstrated to be a flexible approach for *in situ* measurement of basic gait parameters. Different image resolution, sampling frequency and lens systems might improve accuracy and increase the range of applications that the PT-FSCi algorithm could be applied to. Potential applications include gait analysis for different sports, e.g. sprint running, long jump, triple jump etc., as well as gait monitoring for health care environments, e.g. elderly care homes etc. Furthermore, the PT-FSCi algorithm analyses images iteratively; therefore analyses could be performed using streamed images, e.g. online analyses. Image streams could also be communicated over a network or the internet, enabling remote analyses. Programming language, computer and network performance represent delimitating factors to remote, online gait analyses.

11 References

Abdel-Aziz, Y.I. and Karara, H.M. (1971). Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. In: *ASP Symposium on Close-Range Photogrammetry*. Falls Church, VA, USA, pp. 1 – 18.

Ackland, T.R., Henson, P.W. and Bailey, D.A. (1988). The uniform density assumption: its effect upon the estimation of body segment inertial parameters. *International Journal of Sport Biomechanics*, **4**, 146 - 155.

Alcock, A., Hunter, A. and Brown, N. (2009). Determination of football pitch locations from video footage and official pitch markings. *Sports Biomechanics*, **8**(2), 129-140.

ATP World Tour (2013). Last accessed on 10th September 2013 at: <http://www.atpworldtour.com>.

Barros, R.M.L., Misuta, M.S., Menezes, R.P., Figueroa, P.J., Moura, F.A., Cunha, S.A., Anido, R. and Leite, N.J. (2007). Analysis of the distances covered by first division Brazilian soccer players obtained with an automatic tracking method. *Journal of Sports Science and Medicine*, **6**, 233 - 242.

BenAbdelkader, C., Cutler, R.G. and Davis, L.S. (2004). Gait recognition using image self-similarity. *EURASIP Journal on Applied Signal Processing*, **4**, 572 - 585.

Benedek, C. and Szirányi, T. (2007). Study on color space selection for detecting cast shadows in video surveillance. *International Journal of Imaging and Systems and Technology*, **17**(3), 190 - 201.

Bland, J.M. and Altman, D.G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, **8**(1), 307 - 310.

Bouchrika, I. and Nixon, M.S. (2006). Markerless feature extraction for gait analysis. In: *IEEE SMS Conference on Advances in Cybernetic Systems*, pp. 55 - 61.

- Bouchrika, I., Goffredo, M., Carter, J.N. and Nixon, M.S. (2009). Covariate analysis for view-independent gait recognition. In: *The 3rd International Conference on Advances in Biometrics*, 990 - 999.
- Bouguet, J.Y. (2010). Camera Calibration Toolbox for Matlab [online]. Available from: http://www.vision.caltech.edu/bouguetj/calib_doc/ [Accessed 10 October, 2010].
- Boulgouris, N.V., Plataniotis, K.N. and Hatzinakos, D. (2004). An angular transform of gait sequences for gait assisted recognition. In: *Proceedings of IEEE International Conference for Image Processing*, Singapore, Oct. 2004, pp. 857–860.
- Boulgouris, N.V., Hatzinakos, D. and Plataniotis, K.N. (2005). Gait recognition: a challenging signal processing technology for biometric identification. *IEEE Signal Processing Magazine*, **22**(6), 78 - 90.
- Boulgouris, N.V., Hatzinakos, D. and Plataniotis, K.N. (2006). Gait recognition using linear time normalization. *Pattern Recognition*, **39**(5), 969 - 979.
- Bradski, G. and Kaehler, A. (2008). *Learning OpenCV: Computer Vision with the OpenCV Library*. O'Reilly Media, CA, USA.
- Brewin, M.A. and Kerwin, D.G. (2003). Accuracy of scaling and DLT reconstruction techniques for planar motion analyses. *Journal of Applied Biomechanics*, **19**, 79-88.
- Bylak, J. and Hutchinson, M.R. (1998). Common sports injuries in young tennis players. *Sports Medicine*, **26**(2), 119-132.
- Capel-Davies, J and Miller, S (2007), Evaluation of automated line-calling systems. In: *Tennis Science and Technology 3*, The International Tennis Federation, London, UK, pp 387 - 393, (Blackwell Publishing).
- CAVIAR (2004). Test Case Scenarios. Available from: <http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/>.

Chambers, A.J., Margerum, S., Redfern, M.S. and Cham, R. (2003). Kinematics of the foot during slips. *Occupational Ergonomics*, **3**, 225-234.

Chang, W-R., Grönqvist, R., Leclercq, S., Myung, R., Makkonen, L., Strandberg, L., Brungraber, R.J., Mattke, U. and Thorpe, S. (2001). The role of friction in the measurement of slipperiness, part 1: friction mechanisms and definition of test conditions. *Ergonomics*, **44**(13), 1217 - 1232.

Chapman, A.E., Leyland, A.J., Ross, S.M. and Ryall, M. (1991). Effect of floor conditions upon frictional characteristics of squash court shoes. *Journal of Sport Sciences*, **9**, 33-41.

Chen, X. and Davis, J. (2000). *Camera placement considering occlusion for robust motion capture*. Stanford University Computer Science Technical Report, CS-TR-2000-07.

Cheung, G., Kanade, T. Bouguet, J-Y. and Holler, M. (2000). A real time system for robust 3D voxel reconstruction of human motions. In: *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 1 - 7.

Choppin, S. B. (2008). *Modelling of Tennis Racket Impacts in 3D using Elite Players*. PhD, The University of Sheffield.

Coe, A.O. (2000). The balance between technology and tradition in tennis. In: *Tennis Science & Technology*, The International Tennis Federation, London, UK, pp 3 – 40, (Blackwell Publishing).

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences* (2nd ed). Hillsdale, NJ. Lawrence Erlbaum Associates, Publishers.

Conaire, C., Connaghan, D., Kelly, P., O'Connor, N.E., Gaffney, M. and Buckley, J. (2010). Combining inertial and visual sensing for human action recognition in tennis. In: *Proceedings of the 1st ACM International workshop on Analysis and retrieval of tracked events and motion in imagery streams*, pp. 51 - 56.

- Conaire, C., Kelly, P., Connaghan, D. and O'Connor, N.E. (2009). TennisSense: a platform for extracting semantic information from multi-camera tennis data. In: *DSP 2009 - 16th International Conference on Digital Signal Processing, Greece*.
- Connaghan, D., Moran, K. and O'Connor, N.E. (2013). An automatic visual analysis system for tennis. *Journal of Sports Engineering and Technology*, **0**(0), 1 - 16.
- Cunado, D., Nash, J.M., Nixon, M.S. and Carter, J.N. (1999). Gait extraction and description by evidence-gathering. In: *Proceedings of the Second International Conference on Audio- and Video-Based Biometric Person Authentication*, 43-48.
- Dainis, A. and Juberts, M. (1985). Accurate remote measurement of robot trajectory motion. In: *Proceedings of International Conference of Robotics and Automation*. St. Louis, MO, USA, pp. 92-99.
- Dallam, G.M., Wilber, R.L., Jadelis, K., Fletcher, G. and Romanov, N. (2005). Effect of global alteration of running technique on kinematics and economy. *Journal of Sport Sciences*, **23**(7), 757 - 764.
- Damm, L., Low, D., Richardson, A., Clarke, J., Carré, M. and Dixon, S. (2013). The effects of surface traction characteristics on frictional demand and kinematics in tennis. *Sports Biomechanics*, **12**, 1 - 14.
- De Cock, A., Vanrenterghem, J., Willems, T., Witvrouw, E. and De Clercq, D. (2008). The trajectory of the centre-of-pressure during barefoot running as a potential measure for foot function. *Gait & Posture*, **27**, 669 - 675.
- De Wit, B., De Clercq, D. and Aerts, P. (2000). Biomechanical analysis of the stance phase during barefoot and shod running. *Journal of Biomechanics*, **33**, 269 - 278.
- Dempster, W.T. and Gaughran, G.R.L. (1965). Properties of body segments based on size and weight. *American Journal of Anatomy*, **120**, 33 - 54.

Derrick, T.R. (2004). The effects of knee angle on impact forces and accelerations. *Medicine and Science in Sports and Exercise*, **36**, 832 – 837.

Di Salvo, V., Collins, A., McNeill, B. and Cardinale, M. (2006). Validation of Prozone[®]: A new video-based performance analysis system. *International Journal of Performance Analysis in Sport*, **6**(1), 108 - 119.

Driscoll, H.F. (2012). *Understanding shoe-surface interactions in football*. PhD, Sheffield Hallam University.

Duffield, R., Reid, M., Baker, J. and Spratford, W. (2010). Accuracy and reliability of GPS devices for measurement of movement patterns in confined spaces for court-based sports. *Journal of Science and Medicine in Sport*, **13**, 523 - 525.

Eckl, M., Kornfeind, P. and Baca, A. (2011). A comparison of plantar pressures between two different playing surfaces in tennis. In: *Proceedings of the 29th International Conference on Biomechanics in Sports*, Porto, Portugal, pp. 601 – 604.

Fernandez, J., Mendez-Villanueva, A. and Pluim, B.M. (2006). Intensity of tennis match play. *British Journal of Sports Medicine*, **40**(5), 387-391.

Fernandez-Fernandez, J., Sanz-Rivas, D., Fernandez-Garcia, B. and Mendez-Villanueva, A. (2008). Match activity and physiological load during a clay-court tennis tournament in elite female tennis players. *Journal of Sports Sciences*, **26**(14), 1589 - 1595.

Ferrauti, A., Weber, K. and Wright, P.R. Endurance: basic, semi-specific and specific. In: Reid M, Quinn A, Crespo M, eds. *Strength and conditioning for tennis*. London: ITF, 2003:93–111.

Frey, U. (1969). The relationship between anatomical site of injury and particular sports. *Proceedings of the Royal Society of Medicine*, **62**(9), 917-919.

Garcia, D. (2010). Robust smoothing of gridded data in one or higher dimensions with missing values. *Computational Statistics and Data Analysis*, **54**, 1167 - 1178.

Gillmeister, H. (1988). *Tennis: A Cultural History*. London: Leicester University Press.

Girard, O., Eicher, F., Fourchet, F., Micallef, J.P. and Millet, G.P. (2007). Effects of playing surface on plantar pressures and potential injuries in tennis. *British Journal of Sports Medicine*, **41**(11), 733-738.

Glazier, P. and Irwin, G. (2001). Validity of stride length estimates obtained from optojump. In: *Proceedings of the 19th International Conference on Biomechanics in Sports*, San Fransisco, USA, pp. 98 - 101.

Goffredo, M., Bouchrika, I., Carter, J.N. and Nixon, M.S. (2010). Performance analysis for automated gait extraction and recognition in multi-camera surveillance. *Multimedia Tools and Applications*, **50**, 75 - 94.

Goffredo, M., Seely, R.D., Carter, J.N. and Nixon, M.S. (2008). Markerless view independent gait analysis with self-camera calibration. In: *IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 1 - 6.

Gore, J.C. (2000). *Physiological tests for elite athletes*. Champaign, IL: Human Kinetics.

Gorelick, L., Blank, M., Shechtman, E., Irani, M. and Basri, R. (2007). Actions as space-time shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**(12), 2247-2253.

Gonzalez, R.C. and Woods, R.E. (2002). *Digital Image Processing* (2nd ed). Prentice-Hall Inc.: New Jersey, USA.

Haake, S.J., Allen, T., Chopping, S.B. and Goodwill, S.R. (2007). The evolution of the tennis racket and its effect on serve speed. In *Tennis Science & Technology 3*. Roehampton University, London 2007. ITF, 257 – 271.

- Hardin, E.C., van den Bogert, A.J. and Hamill, J. (2004). Kinematic adaptations during running: Effects of footwear, surface and duration. *Medicine and Science in Sports and Exercise*, **36**, 838 – 844.
- Hartley, R. and Zisserman, A. (2003). *Multiple View Geometry in Computer Vision*, Cambridge University Press, UK.
- Hassanpour, H., Sedighi, M. and Manashty, A.R. (2011). Video frame's background modeling: reviewing the techniques. *Journal of Signal and Information Processing*, **2**, 72-78.
- Hatze, H. (1988). High-precision three-dimensional photogrammetric calibration and object space reconstruction using a modified DLT-approach. *Journal of Biomechanics*, **21**(7), 533 - 538.
- Hawk-Eye (2013). Last accessed on 10th December 2013 at: <http://www.hawkeyeinnovations.co.uk>.
- Heikkila, J. and Silven, O. (1997). A four-step camera calibration procedure with implicit image correction. In: *Computer Vision and Pattern Recognition*, pp. 1106 - 1112.
- Hinrichs, R.N., Morrison, B., Vint, P.F., De Witt, J.K., Mitchell, J. and McLean, S.P. (2005). Predicting out-of-plane point locations using the 2D-DLT. In: *29th Annual Meeting of the American Society of Biomechanics*. Portland, OR, USA, pp. 249 – 251.
- Hjelm, N., Wener, S. and Renstrom, P. (2010). Injury profile in tennis players: a prospective two year study. *Knee Surgery, Sports Trauma, Arthroscopy*, **18**(6), 845 - 850.
- Holden-Douilly, L., Pourcelot, P., Chateau, H., Falala, S. and Crevier-Denoix, N. (2011). A method to minimise error in 2D-DLT reconstruction of non-planar markers filmed with a moving camera. *Computer Methods in Biomechanics and Biomedical Engineering*, 1 - 8.

Hreljac, A. and Marshall, R.N. (2000). Algorithms to determine event timing during normal walking using kinematic data. *Journal of Biomechanics*, **33**, 783 - 786.

Hughes, M. and Meyer, R. (2005). Movement pattern in elite men's singles tennis. *International Journal of Performance Analysis in Sport*, **5**(2), 110 – 134.

Hunter, J.P., Marshall, R.N. and McNair, P.J. (2004). Interactions of step length and step rate during sprint running. *Medicine and Science in Sports and Exercise*, **36**, 261 – 271.

Huxham, F., Gong, J., Baker, R., Morris, M. and Iansek, R. (2006). Defining spatial parameters for non-linear walking. *Gait & Posture*, **23**, 159 - 163.

ITF (2010). *ITF approved tennis balls and classified court surfaces - a guide to products and test methods* [online]. Last accessed 21 July 2010 at: http://www.itftennis.com/shared/medialibrary/pdf/original/IO_46735_original.pdf

ITF Rules of Tennis (2013). *Rules of Tennis* [online]. Last accessed 2nd July 2013 at <http://www.itftennis.com/media/136148/136148.pdf>.

ITF Tennis (2013). *ITF Tennis – Technical* [online]. Last accessed 10 March 2014 at: <http://www.itftennis.com/technical/courts/other/history.aspx>

Jain, A.K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 651 - 666.

Jiang, Y-C., Lai, K-T., Hsieh, C-H. and Lai, M-F. (2009). Player detection and tracking in broadcast tennis video. In: *Proceedings of the 3rd Pacific Rim Symposium on Advances in Image and Video Technology*, pp. 759 – 770.

John, G., Sheard, R. and Vickery, B. (2007). *Stadia: a design and development guide* (4th ed). Oxford, UK: Architectural Press.

- Johnson, A. and Bobick, A. (2001). A multi-view method for gait recognition using static body parameters. In: *Proceedings of the International Conference on Audio- and Video-based Biometric Person Authentication*, pp. 301 - 311.
- Jung, S-U. and Nixon, M.S. (2013). Heel strike detection based on human walking movement for surveillance analysis. *Pattern Recognition Letters*, **34**, 895 - 902.
- Kale, A., Cuntoor, N., Yegnanarayana, B., Rajagopalan, A.N. and Chellappa, R. (2003). Gait analysis for human identification. In: *4th International Conference on Audio and Video-based Biometric Person Authentication*, pp. 706 - 714.
- Kozlowski, L.T. and Cutting, J.E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, **21**(6), 575 - 580.
- Kwon, Y-H. (1999). Object plane deformation due to refraction in two-dimensional underwater motion analysis. *Journal of Applied Biomechanics*, **15**, 396-403.
- Kwon, Y-H. (2012). DLT Method (website). Accessed January 2012 from: <http://www.kwon3d.com/theory/dlt/dlt.html>
- Lee, A. (2010). VirtualDub. Last accessed on 10th May 2010 at: <http://www.virtualdub.org/download.html>.
- Lee, L. and Grimson, W.E.L. (2002). Gait analysis for recognition and classification. In: *Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition*, Washington, DC, pp. 148-155.
- Lee, J-W., Lee, M-J., Lee, H-Y. and Lee, H.K. (2012). Screenshot identification by analysis of directional inequality of interlaced video. *EURASIP Journal on Image and Video Processing*, **7**, 1 - 15.
- Li, Y., Dore, A. and Orwell, J. (2005). Evaluating the performance of systems for tracking football players and ball. In: *Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance*, pp. 632 - 637.

Little, J. and Boyd, J. (1998). Recognizing people by their gait: the shape of motion. *International Journal of Computer Vision*, **14**(6), 83 - 105.

Maiwald, C., Sterzing, T., Mayer, T.A. and Milani, T.L. (2009). Detecting foot-to-ground contact from kinematic data in running. *Footwear Science*, **1**(2), 111 - 118.

Mann, R.A. and Hagy, J. (1980). Biomechanics of walking, running and sprinting. *The American Journal of Sports Medicine*, **8**(5), 345 - 350.

Martínez-Gallego, R., Guzmán, J.F., James, N., Perš, J., Ramón-Llin, J. and Vučković, G. (2013). Movement characteristics of elite tennis players on hard courts with respect to the direction of ground strokes. *Journal of Sports Science and Medicine*, **12**, 275 - 281.

Martinez-Martin, E. and Pobil, A.P. (2012). *Robust Motion Detection in Real-Life Scenarios*. London, UK: Springer.

Matuska, S., Hudec, R. and Benco, M. (2012). The comparison of CPU time consumption for image processing algorithm in Matlab and OpenCV. In: *Proceedings of the 9th ELEKTRO International Conference*, pp. 75 - 78.

Mauthner, T. Koch, C., Tilp, M. and Bishof, H. (2008). Visual tracking of athletes in beach volleyball using a single camera. *International Journal of Computer Science in Sport*, **6**(2), 21 - 34.

McLean, S.P., Vint, P.F., Hinrichs, R.N., DeWitt, J.K. Morrison, B. and Mitchell, J. (2004). Factors affecting the reconstruction accuracy of 2D-DLT calibration. In: *28th Annual Meeting of the American Society of Biomechanics*. Portland, OR, USA, pp. 246 - 247.

Meershoek, L. (1997). Matlab routines for 2-D camera calibration and point reconstruction using the DLT for 2-D analysis with non-perpendicular camera angle [online]. Last accessed 20 January 2011 at: <http://isbweb.org/software/movanal.html>

Miller, S. (2006). Modern tennis rackets, balls, and surfaces. *British Journal of Sports Medicine*, **40**, 401-405.

Morais, E., Goldenstein, S., Ferreira, A. and Rocha, A. (2012). Automatic tracking of indoor soccer players using videos from multiple cameras. In: *25th SIBGRAPI Conference on Graphics, Patterns and Images*, pp. 174 - 181.

NEVA Electromagnetics (2013). Surface Human Body Meshes (website) accessed March 2013 from <https://www.nevaelectromagnetics.com/SurfaceHumanBodyMeshes.html>

Nevill, A.M. and Atkinson, G. (1997). Assessing agreement between measurements recorded on a ratio scale in sports medicine and sports science. *British Journal of Sports Medicine*, **31**, 314 - 318.

Nghiem, A.T., Bremon, F. and Thonnat, M. (2008). Shadow removal in indoor scenes. In: *Proceedings of the IEEE International Conference on Advanced Video and Signal Based Surveillance*, pp. 291-298.

Novacheck, T.F. (1998). The biomechanics of running. *Gait & Posture*, **7**, 77 - 95.

O'Connor, C.M., Thopre, S.K., O'Malley, M.J. and Vaughan, C.L. (2007). Automatic detection of gait events using kinematic data. *Gait & Posture*, **25**, 469 - 474.

O'Donoghue, P. and Ingram, B. (2001). A notational analysis of elite tennis strategy. *Journal of Sport Sciences*, **19**(2), 107-115.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, **9**(1), 62-66.

Owings, T.M. and Grabiner, M.D. (2004). Step width variability, but not step length variability or step time variability, discriminates gait of healthy young and older adults during treadmill locomotion. *Journal of Biomechanics*, **37**, 935 - 938.

Pallis, J. M. (2004). Last accessed 20 December 2013 at:

http://www.tennisserver.com/set/set_04_01.html

Pereria, P., Wells, J. and Hughes, M. (2001). Notational analysis of elite women's movement patterns in squash. In: *Notational Analysis V*, Cardiff: CPA, University of Wales Institute, Cardiff, pp. 223 – 236.

Perš, J., Bon, M., Kvačič, Šibila, M. and Dežman, B. (2002). Observation and analysis of large-scale human motion. *Human Movement Science*, **21**(2), 295 - 311.

PETS (2006). Performance Evaluation of Tracking and Surveillance. Available from: <http://www.cvg.rdg.ac.uk/PETS2006/data.html>.

Pluim, B.M., Staal, J.B., Windler, G.E. and Jayanthi, N. (2006). Tennis injuries: occurrence, aetiology, and prevention. *British Journal of Sports Medicine*, **40**(5), 415-423.

Poppe, R. (2007). Vision-based human motion analysis: an overview. *Computer Vision and Image Understanding*, **108**, 4 - 18.

Queen, R.M., Gross, M.T. and Liu, H-Y. (2006). Repeatability of lower extremity kinetics and kinematics for standardized and self-selected running speeds. *Gait & Posture*, **23**, 282 - 287.

Redfern, M.S., Cham, R., Gielo-Perczak, K., Grönqvist, R., Hirvonen, M., Lanshammar, H., Marpet, M. and Yi-Chung Pai, C. (2001). Biomechanics of slips. *Ergonomics*, **44**(13), 1138-1166.

Reed, T. R. (2004). *Digital Image Sequence Processing, Compression and Analysis*. CRC Press, MA, USA.

Robinson, G. and O'Donoghue, P. (2008). A movement classification for the investigation of agility demands and injury risk in sport. *International Journal of Performance Analysis in Sport*, **8**(1), 127-144.

Sherry, D. and Hawkins, P., Roke Manor Research limited (2001). *Video Processor Systems for Ball Tracking in Ball Games*. [online]. U.K Patent 041884. Patent from WIPO Last accessed 20th December 2013 at: <http://www.wipo.int/pctdb/en/wo.jsp?WO=2001041884>.

Sih, B.L., Hubbard, M. and Williams, K.R. (2001). Correcting out-of-plane errors in two-dimensional imaging using nonimage-related information. *Journal of Biomechanics*, **34**, 257 - 260.

Starbuck, C., Damm, L., Stiles, V., Capel-Davies, J., Clarke, J., Carré, M., Miller, S. and Dixon, S. (2013). The influence of previous experience on clay on player response to tennis surface during a running forehand. In: *BASES Biomechanics Interest Group*, Wolverhampton, UK.

Stauffer, C. and Grimson, W.E.L. (1999). Adaptive background mixture models for real-time tracking. In: *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 246 - 252.

Stiles, V.H. and Dixon, S.J. (2006). The influences of different playing surfaces on the biomechanics of a tennis running forehand foot plant. *Journal of Applied Biomechanics*, **22**(1), 14-24.

Sun, W. and Cooperstock, J.R. (2005). Requirements for camera calibration: must accuracy come at a high price? In *Proceedings of the 7th IEEE workshop on Applications of Computer Vision*, pp. 1 - 6.

Teachabarikiti, K., Chalidabhongse, T.H. and Thammano, A. (2010). Player tracking and ball detection for an automatic tennis video annotation. In: *Proceedings of the 11th International Conference on Control Automation Robotics & Vision*, pp. 2461 - 2494.

Tiarks, F., Frechen, T., Kirsch, S., Leuninger, J., Melan, M., Pfau, A., Ritcher, F., Schuler, B. and Zhao, C-L. (2003). Formulation effects on the distribution of pigment particles in paints. *Progress in Organic Coatings*, **48**, 140 - 152.

- Tirosh, O. and Sparrow, W.A. (2003). Identifying heel contact and toe-off using forceplate thresholds with a range of digital-filter cutoff frequencies. *Journal of Applied Biomechanics*, **19**, 178 - 184.
- Tsai, R. (1987). A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *Journal of Robotics and Automation*, **3**(4), 323-344.
- Vollmer, J., Mencl, R. and Mueller, H. (1999). Improved Laplacian smoothing of noisy surface meshes. *Eurographics*, **18**(3), 1 – 8.
- Vučković, G., Perš, J., James, N. and Hughes, M. (2010). Measurement error associated with the SAGIT/Squash computer tracking software. *European Journal of Sport Science*, **10**(2), 129 - 140.
- Wang, W. and Farid, H. (2007). Exposing digital forgeries in interlaced and deinterlaced video. *IEEE Transactions on Information Forensics and Security*, **2**(3), 438 - 449.
- Wang, L., Tan, T., Ning, H. and Hu, W. (2003). Silhouette analysis-based gait recognition for human identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **25**(12), 1505 - 1518.
- Webster, K.E., Wittwer, J.E. and Feller, J.A. (2005). Validity of the GAITRite[®] walkway system for the measurement of averaged and individual step parameters of gait. *Gait & Posture*, **22**, 317 - 321.
- Winter, D.A. (2005). *Biomechanics of Motor Control and Human Movement* (3rd ed). Hoboken, NJ: John Wiley & Sons.
- Winter, E. and Fowler, N. (2009). Exercise defined and quantified according to the Système International d'Unités. *Journal of Sports Sciences*, **27**(5), 447 – 460.

- Wu, H. and Zheng, Q. (2004). Self-evaluation for video tracking systems. In: *24th Army Science Conference*, pp. 1 - 6.
- Yam, C.Y., Nixon, M.S. and Carter, J.N. (2002). Gait recognition by walking and running: a model-based approach. In: *Proceedings of Asian Conference on Computer Vision*, pp. 1 - 6.
- Yan, F., Christmas, W. and Kittler, J. (2005). A tennis ball tracking algorithm for automatic annotation of tennis match. *British Machine Vision Conference*, 619-628.
- Zhang, Z. (1999). Flexible camera calibration by viewing a plane from unknown orientations. In: *International conference on computer vision*. Corfu, Greece, pp. 666-673.
- Zhang, X., Peng, J., Yu, W. and Lin, K. (2012). Confidence-level-based new adaptive particle filter for nonlinear object tracking. *International Journal of Advanced Robotic Systems*, **9**(199), 1 - 9.
- Zhang, B.F., Zhou, J. and Zhu, J.C. (2010). Research on the three image difference algorithm. In: *IEEE International Conference on Image Analysis and Signal Processing*, pp. 603 - 606.

Personal bibliography

Dunn, M., Wheat, J., Haake, S. and Goodwill, S. (2011). Assessing tennis player interactions with tennis courts. In: *Proceedings of the 29th International Conference on Biomechanics in Sports*, pp. 859 - 862.

Dunn, M., Wheat, J., Miller, S., Haake, S. and Goodwill, S. (2012). Reconstructing 2D planar coordinates using linear and non-linear techniques. In: *Proceedings of the 30th International Conference on Biomechanics in Sports*, pp. 381 - 383.

Dunn, M., Haake, S., Wheat, J., and Goodwill, S. (2014). Validation of a single camera, spatio-temporal gait analysis system. *Procedia Engineering*, **72**, 243 - 248.

Appendix 1

A.1.1. Permission to film and accreditations obtained via the Fédération Française de Tennis.



ITF
M DUNN MARCUS
BANK LANE ROEHAMPTON
SW155XZ LONDON
GRANDE-BRETAGNE

Paris, April 12, 2011

We are pleased to inform you that your credential request has been accepted.

To collect your credential, please proceed directly with this letter to the Welcome Desk, located in the TV Production Area.

→ From Monday May 16th to Friday May 20th, 2011, entrance through:

GATE I - 8, boulevard d'Auteuil - 75 016 Paris, from 10 am to 6 pm.

Hotel reservations will be forwarded to a central booking office, which will send your confirmation directly. For any questions, please contact:

The Accommodation Service: Tel: + 33 1 47 43 40 06 – e-mail: rez-hotels@fft.fr

For any further information, please contact:

David ANSAS - MEDIA DEPARTMENT

Tel: + 33 1 47 43 51 85 - e-mail: dansas@fft.fr

We look forward to welcoming you at Roland Garros.

Yours sincerely,

Sandrine LOPES

Media Department



892329810



Stade Roland Garros - 2, avenue Gordon Bennett - 75016 Paris - Tel. : + 33 1 47 43 48 00 - www.rolandgarros.com

CONFIDENTIAL



**Faculty of Health and Wellbeing
Research Ethics Committee**

Sport & Exercise Research Ethics Review Group

APPLICATION FOR ETHICS APPROVAL OF RESEARCH

In designing research involving humans, principal investigators should be able to demonstrate a clear intention of benefit to society and the research should be based on sound principles. These criteria will be considered by the Sport and Exercise Research Ethics Review Group before approving a project. **ALL** of the following details must be provided, either typewritten or word-processed preferably at least in 11 point font.

Please either tick the appropriate box or provide the information required.

1) Date of application	20 December 2013
2) Anticipated date of completion of project	1 July 2011
3) Title of research	Tennis player step and movement characterisation
4) Subject area	Sports Engineering
5) Principal Investigator Name Email address @ SHU Telephone/Mobile number Student number (if applicable)	Marcus Dunn m.dunn @shu.ac.uk 0114 225 5867 14025464
6) State if this study is: (If the project is undergraduate or postgraduate please state module name and number)	<input checked="" type="checkbox"/> Research <input type="checkbox"/> Undergraduate <input type="checkbox"/> Postgraduate Module name: Module number:
7) Director of Studies/Supervisor/Tutor name	Dr. Simon Goodwill

8) Intended duration and timing of project?	17 – 18 May 2011: film at Roland Garros Qualifying Tournament. June 2011: manual analysis of player step and movement parameters.
9) Location of project If external to SHU, provide evidence in support (see section 17)	Filming at the 2011 Roland Garros Qualifying Tournament. This tournament will be held at Stade Roland Garros, Paris. Development, testing and analysis at SHU.
10) State if this study is:	<input checked="" type="checkbox"/> New <input checked="" type="checkbox"/> Collaborative (please include appropriate agreements in section 17) <input type="checkbox"/> Replication of :
11) Purpose and benefit of the research Statement of the research problem with any necessary background information (no more than 1 side of A4) <p>Current knowledge of tennis player-surface interaction with different tennis court surfaces, particularly during competition, is limited (Miller, 2006). Human centred approaches for the measurement of surface slipperiness highlight that step strategy reflects interactions that occur at the shoe-surface interface (Grönqvist et al. 2001). The measurement of tennis player step strategy would therefore contribute to the understanding of tennis player-surface interactions with different court surfaces.</p> <p>The characterisation of tennis player movement and step strategy during competition is the principle outcome for this work. A manual system for characterising tennis player step and movement strategy during competition has been developed. Footage can be obtained passively, i.e. no markers or researcher intervention is required. Therefore analyses can be performed <i>in situ</i> and provide a novel insight into the characteristics of tennis player step and movement strategy during competition. This will aid the understanding of tennis player-surface interactions with tennis court surfaces.</p> <p>Grönqvist, R., Abeysekera, J., Gard, G., Hsiang, S., Leamon, T.B., Newman, D.J., Gielo-Perczak, K., Lockhart, T.E. and Pai, C, Y.-C. (2001). Human centred approaches in slipperiness measurement. <i>Ergonomics</i>, 44(13), 1167 - 1199.</p> <p>Miller, S. Modern tennis rackets, balls and surfaces. (2006). <i>British Journal of Sports Medicine</i>, 40, 401 - 405.</p>	

12) Participants	
12.1 Number	<p>The project will film singles tennis players competing on a single practice tennis court.</p> <p>Twelve tennis players (three men's singles matches and three women's singles matches) will be filmed over two days.</p>
12.2 Rationale for this number (eg calculations of sample size, practical considerations)	<p>Due to the novel nature of analyses being performed, no previous data exist for the basis of sample size calculations.</p> <p>Sample size represents practical filming and analysis constraints, i.e. filming six matches over two days and time cost of manual analysis.</p> <p>Furthermore, camera footage and data from the competition are desirable for future analysis and development.</p>
12.3 Criteria for inclusion and exclusion (eg age and sex)	<p>The participants will be male and female entrants of the 2011 Roland Garros Qualifying Tournament.</p> <p>See 12.4 for details on research consent.</p>
12.4 Procedures for recruitment (eg location and methods)	<p>The Grand Slam Committee (GCS) outlines the Grand Slam Rules (attached). The Grand Slam Rules state in Article I (section E) that "Each player grants and assigns to the GCS and the management of the events that he enters, the right in perpetuity to make, use and show from time to time and at their discretion, motion pictures, still pictures and live, taped or filmed television and other reproductions of him during said events and in connection with the promotion of said events without compensation for himself, his heirs, devisees, executors, administrators or assigns."</p> <p>The GCS can therefore grant or deny permission to film tennis players during entered competitions, for purposes agreed by the GCS.</p> <p>The International Tennis Federation (ITF) sought permission (email correspondence and accreditation) from the relevant GCS body (Fédération Française de Tennis) to perform tennis match filming for research purposes.</p>

12.5 Does the study have *minors or ‡vulnerable adults as participants?	[] Yes [✓] No
12.6 Is CRB Disclosure required for the Principal Investigator? (to be determined by Risk Assessment)	[] Yes [✓] No If yes, is standard [] or enhanced [] disclosure required?
12.7 If you ticked 'yes' in 12.5 and 'no' in 12.6 please explain why:	

*Minors are participants under the age of 18 years.

‡Vulnerable adults are participants over the age of 16 years who are likely to exhibit:

- a) learning difficulties
- b) physical illness/impairment
- c) mental illness/impairment
- d) advanced age
- e) any other condition that might render them vulnerable

13) Details of the research design

13.1 Provide details of intended methodological procedures and data collection.

(For MSc students conducting a scientific support project please provide the following information: a. needs analysis; b. potential outcome; c proposed interventions).

1. Video

Central to the project is filming of tennis matches at the Roland Garros Qualifying Tournament. In more detail:

- a. A single camera will be used to capture each match. This camera will be situated in the stands and will capture images of the entire tennis court.
- b. Footage will be captured from just before the start of the first match to just after the end of the last match. There are twelve matches: six matches a day for two days.
- c. Footage will be stored on the computers detailed in the Equipment section.
- d. Access to the footage will be restricted. Apart from images used in the thesis and research publications only the principal investigator and the principal investigator's supervisory team (Professor Steve Haake, Dr Simon Goodwill and Dr Jon Wheat) will have access.

2. Data

Data, such as player displacement, step length and step frequency, will be calculated from the footage. In more detail:

- a. Data will be manually digitised from the footage.
- b. Data will be stored on the computers detailed in the Equipment section.
- c. Access to the footage will be restricted. Apart from images used in the thesis and research publications only the principal investigator and the principal investigator's supervisory team (Professor Steve Haake, Dr Simon Goodwill and Dr Jon Wheat) will have access.

3. Equipment

The following computer equipment will be used to capture and store the footage and data:

- a. Laptop:
 - Footage and data will be stored on this computer. The footage and data will be used to for research output.
 - It will be located in A212 Collegiate Hall.
 - Only the principal investigator and the principal investigator's director of studies will have access to this computer.

4. Security

The following steps will be taken to secure footage and data:

- a. The computers detailed in the Equipment section will use the Microsoft Windows 7 Ultimate. Only the principal investigator and the principal investigator's director of studies will have the password.

13.2 Are these "minor" procedures as defined in Appendix 1 of the ethics guidelines?

☒ Yes ☐ No

13.3 If you answered 'no' in section 13.2, list the procedures that are not minor

13.4 Provide details of the quantitative and qualitative analysis to be used

Descriptive statistics and inferential statistics will be used to explore the effect of gender on tennis player step and movement strategy.

A one-way ANOVA will be performed to identify differences between men's and women's tennis rallies. Furthermore, between factor effect sizes will be calculated to effect size magnitudes.

14) Substances to be administered (refer to Appendix VI of the ethics procedures)**14.1 The protocol does not involve the administration of pharmacologically active substances or nutritional supplements.**

Please tick box if this statement applies and go to section 15) [✓]

14.2 Name and state the risk category for each substance. If a COSHH assessment is required state how the risks are to be managed.**15) Degree of discomfort that participants might experience**

Consider the degree of physical and psychological discomfort that will be experienced by the participants. State the details which must be included in the participant information sheet to ensure that the participants are fully informed about any discomfort that they may experience.

Not applicable. The principal investigator will be passively monitoring tennis matches from the stands.



16) Outcomes of Risk Assessment


Provide details of the risk and explain how the control measures will be implemented to manage the risk.

Overall risk is LOW, which is driven by the risk of electrical equipment being present in publicly accessible stands.

The risk is mitigated by the equipment and principal investigator being positioned in the stands.

17) Attachments	Tick box
17.1 Risk assessment (including CRB risk assessment)	✓
17.2 COSHH assessment	N/A
17.3 Participant information sheet (this should be addressed directly to the participant (ie you will etc) and in a language they will understand)	N/A
17.4 Informed consent form	N/A
17.5 Pre-screening questionnaire	N/A
17.6 Collaboration evidence/support correspondence from the organisation consenting to the research (this must be on letterhead paper and signed) See sections 9 & 10.	✓
17.7 CRB Disclosure certificate <u>or</u> where not available CRB application form	N/A
17.8 Clinical Trails form (FIN 12)	N/A

18. Signature Principal Investigator	<p>Once this application is approved, I will undertake the research study as approved. If circumstances necessitate that changes are made to the approved protocol, I will discuss these with my Project Supervisor. If the supervisor advises that there should be a resubmission to the Sport and Exercise Research Ethics Review Group, I agree that no work will be carried out using the changed protocol until approval has been sought and formally received.</p> <p>  Date <u>20/12/13</u> Principal Investigator signature Name <u>MARCUS DUNN</u> </p>
19. Approval Project Supervisor to sign either box A or box B as applicable (refer to Appendix I and the flowchart in appendix VI of the ethics guidelines)	<p>Box A: I confirm that the research proposed is based solely on 'minor' procedures, as outlined in Appendix 1 of the HWB Sport and Exercise Research Ethics Review Group 'Ethics Procedures for Research with Humans as Participants' document, and therefore does not need to be submitted to the HWB Sport and Exercise Research Ethics Review Group.</p> <p>In terms of ethics approval, I agree the 'minor' procedures proposed here and confirm that the Principal Investigator may proceed with the study as designed.</p> <p>  Date <u>20/12/13</u> Project Supervisor signature Name <u>SIMON GOODWIN</u> </p>
	<p>Box B: I confirm that the research proposed is <u>not</u> based solely on 'minor' procedures, as outlined in Appendix 1 of the HWB Sport and Exercise Research Ethics Review Group 'Ethics Procedures for Research with Humans as Participants' document, and therefore <u>must</u> be submitted to the HWB Sport and Exercise Research Ethics Review Group for approval.</p> <p>I confirm that the appropriate preparatory work has been undertaken and that this document is in a fit state for submission to the HWB Sport and Exercise Research Ethics Review Group.</p> <p> Date _____ Project Supervisor signature _____ Name _____ </p>
20. Signature Technician	<p>I confirm that I have seen the full and approved application for ethics approval and technical support will be provided.</p> <p> Date _____ Technician signature _____ Name _____ </p>

	<p style="font-size: 1.2em; font-weight: bold;">Sheffield Hallam University</p> <p style="font-weight: bold;">Faculty of Health and Wellbeing Research Ethics Committee Sport and Exercise Research Ethics Review Group</p> <p style="font-weight: bold;">Risk Assessment Pro Forma</p>
---	---

**Please ensure that you read the accompanying
Risk Assessment Risk Ranking document before completing this form**

Title of research	Tennis player step and movement characterisation
Date Assessed	20 December 2013
Assessed by (Principal Investigator)	Marcus Dunn
Signed	Position
	Principal Investigator

Activity	Risks	Control Measures
Filming	Risk of [physical injury] caused by tripping over cabling. $R1 = C1 \times L1$. LOW RISK	Filming will take place in the stands that have public access. Equipment will not be placed in a thoroughfare. Equipment will not be left unattended. Cables of sufficient length will be used. Cables of will be tidied sufficiently.
Filming	Risk of [physical injury] caused by falling tripod. $R1 = C1 \times L1$. LOW RISK	Filming will take place in the stands that have public access. Equipment will not be placed in a thoroughfare. Equipment will not be left unattended. Cables of sufficient length will be used. Cables of will be tidied sufficiently.
Filming	Risk of [physical injury] caused by transporting equipment to and from venue. $R1 = C1 \times L1$. LOW RISK	Equipment will be kept to the minimum. Equipment will be carried in suitable containers and loaded carefully into vehicles.

Risk Evaluation (Overall)

LOW

General Control Measures

Is a pre-screen medical questionnaire required? Yes [] No [✓]

Emergency Procedures


None required.

Monitoring Procedures

None required.

Review Period

None.

Reviewed By (Supervisor)**Date**

20/12/13

Appendix 2

A.2.1. Age, mass and stature of players analysed at the 2011 Roland Garros Qualifying Tournament.

Table A.2.1. Age, mass and stature of male players.

	Age (years)	Mass (kg)	Stature (m)
Men	31	77.3	1.85
	28	75.0	1.85
	27	84.1	1.88
	30	65.0	1.86
	31	75.0	1.80
Mean	29.4	75.3	1.83
Sd	1.8	6.9	0.05

Table A.2.2. Age, mass and stature of female players.

	Age (years)	Mass (kg)	Stature (m)
Women	21	68.0	1.73
	22	67.3	1.78
	22	67.3	1.83
	28	65.0	1.73
	23	62.3	1.73
Mean	23.2	66.0	1.76
Sd	2.8	2.3	0.04

A.2.2. Custom MATLAB analysis script for manually digitised tennis rally parameters.

```
% File imports exported clipboard data from 'Low Res Tennis'
% Identifies data and performs simple temporal analyses
% Exports data to file C:\Rally.xlsx
clear all
close all
clc
DELIMITER = '\t';
HEADERLINES = 1;

% Import the file
newData1 = importdata('-pastespecial', DELIMITER, HEADERLINES);
```

```

% Create new variables in the base workspace from those fields.
vars = fieldnames(newData1);
for i = 1:length(vars)
    assignin('base', vars{i}, newData1.(vars{i}));
end

frame=data(:,1);
sf = 1/50;
cr = 50;
[i]=(find(isnan(frame)));

service_c=data(1:(i(1)-1),:);
service_t=service_c(1,1);
ball_surface_c=data((i(1)+1):i(2)-1,:);
server_TF=data((i(2)+1):i(3)-1,:); % Formally left foot,now trail foot
server_LF=data((i(3)+1):i(4)-1,:); % Formally right foot,now lead foot
server_com=data((i(4)+1):i(5)-1,:);
server_rbc=data((i(5)+1):i(6)-1,:);
receiver_TF=data((i(6)+1):i(7)-1,:); %Formally left foot,now trail foot
receiver_LF=data((i(7)+1):i(8)-1,:); %Formally right foot,now lead foot
receiver_com=data((i(8)+1):i(9)-1,:);
receiver_rbc=data((i(9)+1):end,:);

for n=1:1%% Find ball trajectory properties of rally
ball_surface_t=ball_surface_c(:,1);
ball_surface_x=ball_surface_c(:,2);
ball_surface_y=ball_surface_c(:,3);
if ball_surface_y(1) > 0 % i.e. value is positive, server at near end
of court
    server_pos = [0, -11.887];
    server_near = 1;
    server_far = 0;
elseif ball_surface_y(1) < 0 % i.e. value is negative, server at far
end of court
    server_pos = [0, 11.887];
    server_near = 0;
    server_far = 1;
end

% Due to service contact being out of plane, assume it is struck at
the baseline T
serve_trajet=[server_pos(1),ball_surface_x(1);server_pos(2),ball_surf
ace_y(1)];
serve_disp=serve_trajet(:,2)-serve_trajet(:,1);
serve_disp_x=serve_disp(1);
serve_disp_y=serve_disp(2);
serve_disp_theta=atand(serve_disp_y./serve_disp_x);
serve_disp_res=sqrt((serve_disp_x.^2)+(serve_disp_y.^2));

serve_time=(ball_surface_t(1)-service_c(1)).*sf;
serve_vel_x = serve_disp_x./serve_time;
serve_vel_y = serve_disp_y./serve_time;
serve_vel_res = serve_disp_res./serve_time;

% Rally ball properties (from ball-surface-contact data)
for n=1:1
ball_d=[];
for p = 2:length(ball_surface_c)
    ball_d{p} = (ball_surface_c(p,:))-(ball_surface_c(p-1,:));

```



```

end

for q=2:length(ball_d)
    ball_dt{q}=ball_d{1,q}(1);
    ball_dx{q}=ball_d{1,q}(2);
    ball_dy{q}=ball_d{1,q}(3);
    ball_dres{q}=sqrt((ball_dx{q}.^2)+(ball_dy{q}.^2));
    ball_d_theta{q}=atand(ball_dy{q}./ball_dx{q});
    ball_vx{q}=ball_dx{q}./((ball_d{1,q}(1)).*sf);
    ball_vy{q}=ball_dy{q}./((ball_d{1,q}(1)).*sf);
    ball_vres{q}=ball_dres{q}./((ball_d{1,q}(1)).*sf);
end

Avg_ball_dres=mean(cell2mat(ball_dres));
SD_ball_dres=std(cell2mat(ball_dres));
Max_ball_dres=max(cell2mat(ball_dres));
Min_ball_dres=min(cell2mat(ball_dres));
Avg_ball_vres=mean(cell2mat(ball_vres));
SD_ball_vres=std(cell2mat(ball_vres));
Max_ball_vres=max(cell2mat(ball_vres));
Min_ball_vres=min(cell2mat(ball_vres));
end
end

assumed_service_c=horzcat(service_t,server_pos); % due to racket ball
contact being so far out of plane, use these data instead of service_c
Rally =
vertcat(assumed_service_c,server_rbc,receiver_rbc,ball_surface_c);
Rally = sortrows(Rally);
Rally_abs_t = (Rally(:,1)-service_t)*sf;

Rally_less_serve = Rally(2:end,:);

for n = 2:length(Rally)
    Rally_rel = (Rally(n,1:3))-(Rally(n-1,1:3));
    Rally_d_cell{n}=Rally_rel;
end
Rally_d=[];
for n=1:length(Rally_d_cell)
    a=Rally_d_cell{n};
    Rally_d=vertcat(a,Rally_d);
end
Rally_d=flipud(Rally_d);
Rally_d(:,1) = Rally_d(:,1)*sf;
Rally = horzcat(Rally,Rally_abs_t);
Rally_d = horzcat(Rally_d,Rally_abs_t(2:end,:));

ball_res=[];
for n = 1:length(Rally_d)
    res=sqrt(((Rally_d(n,2))^2)+((Rally_d(n,3))^2));
    ball_res=vertcat(res,ball_res);
end
ball_res = flipud(ball_res);
Rally_d = horzcat(Rally_d,ball_res);

ball_theta=[];
ball_theta_d=[];
for n = 1:length(Rally_d)
    theta=atan((Rally_d(n,3))/(Rally_d(n,2)));
    theta_d=atand((Rally_d(n,3))/(Rally_d(n,2)));

```

```

        ball_theta=vertcat(theta,ball_theta);
        ball_theta_d=vertcat(theta_d,ball_theta_d);
    end
    ball_theta = flipud(ball_theta);
    ball_theta_d = flipud(ball_theta_d);
    Rally_d = horzcat(Rally_d,ball_theta_d);

    % Finishing point
    if sum(eq(Rally(end,1),server_rbc(:,1))) == 1
        Rally_end = 'Server error';
    elseif sum(eq(Rally(end,1),receiver_rbc(:,1))) == 1
        Rally_end = 'Receiver error';
    elseif sum(eq(Rally(end,1),ball_surface_c(:,1))) == 1
        rally_end_row = eq(Rally(end,1),ball_surface_c(:,1));
        if ball_surface_c((length(rally_end_row)),2) < 0
            if server_near == 1
                Rally_end = 'Server error';
            else
                Rally_end = 'Receiver error';
            end
        elseif ball_surface_c((length(rally_end_row)),2) > 0
            if server_far == 1
                Rally_end = 'Server error';
            else
                Rally_end = 'Receiver error';
            end
        end
    end
end

%% COM trajectory parameters
%Server
for n=1:1
    s_com_d=[];
    for p = 2:length(server_com)
        s_com_d{p} = (server_com(p,:))-(server_com(p-1,:));
        s_com_abs_t{p} = (server_com(p,1)-service_t)*sf;
        if isempty(s_com_abs_t{p})
            s_com_abs_t{p}=(service_t-service_t)*sf;
        else
            end
        end
    end

    for q=2:length(s_com_d)
        s_com_dt{q}=s_com_d{1,q}(1);
        s_com_dx{q}=s_com_d{1,q}(2);
        s_com_dy{q}=s_com_d{1,q}(3);
        s_com_dres{q}=sqrt((s_com_dx{q}.^2)+(s_com_dy{q}.^2));
        s_com_theta{q}=atand(s_com_dy{q}./s_com_dx{q});
        s_com_vx{q}=s_com_dx{q}./((s_com_d{1,q}(1)).*sf);
        s_com_vy{q}=s_com_dy{q}./((s_com_d{1,q}(1)).*sf);
        s_com_vres{q}=s_com_dres{q}./((s_com_d{1,q}(1)).*sf);
    end

    Avg_server_com_dres=mean(cell2mat(s_com_dres));
    SD_server_com_dres=std(cell2mat(s_com_dres));
    Max_server_com_dres=max(cell2mat(s_com_dres));
    Min_server_com_dres=min(cell2mat(s_com_dres));
    Avg_server_com_vres=mean(cell2mat(s_com_vres));
    SD_server_com_vres=std(cell2mat(s_com_vres));
    Max_server_com_vres=max(cell2mat(s_com_vres));
    Min_server_com_vres=min(cell2mat(s_com_vres));

```

```

end

%Receiver
for n=1:1
r_com_d=[];
for p = 2:length(receiver_com)
    r_com_d{p} = (receiver_com(p,:))-(receiver_com(p-1,:));
    r_com_abs_t{p} = (receiver_com(p,1)-service_t)*sf;
    if isempty(r_com_abs_t{p})
        r_com_abs_t{p}=(service_t-service_t)*sf;
    else
        end
end

for q=2:length(r_com_d)
    r_com_dt{q}=r_com_d{1,q}(1);
    r_com_dx{q}=r_com_d{1,q}(2);
    r_com_dy{q}=r_com_d{1,q}(3);
    r_com_dres{q}=sqrt((r_com_dx{q}.^2)+(r_com_dy{q}.^2));
    r_com_theta{q}=atand(r_com_dy{q}./r_com_dx{q});
    r_com_vx{q}=r_com_dx{q}./((r_com_d{1,q}(1)).*sf);
    r_com_vy{q}=r_com_dy{q}./((r_com_d{1,q}(1)).*sf);
    r_com_vres{q}=r_com_dres{q}./((r_com_d{1,q}(1)).*sf);
end

Avg_receiver_com_dres=mean(cell2mat(r_com_dres));
SD_receiver_com_dres=std(cell2mat(r_com_dres));
Max_receiver_com_dres=max(cell2mat(r_com_dres));
Min_receiver_com_dres=min(cell2mat(r_com_dres));
Avg_server_com_vres=mean(cell2mat(r_com_vres));
SD_receiver_com_vres=std(cell2mat(r_com_vres));
Max_receiver_com_vres=max(cell2mat(r_com_vres));
Min_receiver_com_vres=min(cell2mat(r_com_vres));
end

%% Find matching COM data (intra-foot parameters to appear inside COM
manoeuvre loop to create cells of data)
for t=1:1
sct=server_com(:,1);
slft=server_LF(:,1);
stft=server_TF(:,1);
server_com_rows=[];
server_LF_rows=[];
server_TF_rows=[];
init_sl = 1;
for p = 1:init_sl:length(sct)
    for n = 1:length(slft)
        if isequal(sct(p),slft(n))
            [sf_lf]=find(isequal(sct(p),slft(n)));
            server_LF_rows=vertcat(server_LF_rows,n);
        end
        for q = 1:length(stft)
            if isequal(sct(p),stft(q))
                [sf_tf]=find(isequal(sct(p),stft(q)));
                server_com_rows=vertcat(server_com_rows,p); %
creates column of row numbers when starting COM movement frame number
matches LF and (by implication) TF(assuming digitising flow is
correct)
                server_TF_rows=vertcat(server_TF_rows,q); % as
above but trail foot
            end
        end
    end
    init_sl = 2;
end

```

```

        else
        end
    end
end
end
server_LF_rows = unique(server_LF_rows);
server_com_rows = unique(server_com_rows);
server_TF_rows = unique(server_TF_rows);

rct=receiver_com(:,1);
rlft=receiver_LF(:,1);
rtft=receiver_TF(:,1);
receiver_com_rows=[];
receiver_LF_rows=[];
receiver_TF_rows=[];
init_rl = 1;
for p = 1:init_rl:length(rct)
    for n = 1:length(rlft)
        if isequal(rct(p),rlft(n))
            [rf_lf]=find(isequal(rct(p),rlft(n)));
            receiver_LF_rows=vertcat(receiver_LF_rows,n);
        end
        for q = 1:length(rtft)
            if isequal(rct(p),rtft(q))
                [rf_tf]=find(isequal(rct(p),rtft(q)));
                receiver_com_rows=vertcat(receiver_com_rows,p); %
                % creates column of row numbers when starting COM movement frame number
                % matches LF and (by implication) TF (assuming digitising flow correct)
                receiver_TF_rows=vertcat(receiver_TF_rows,q); % as
                % above but trail foot
            end
        end
    end
end
init_rl = 2;
else
end
end
end
receiver_LF_rows = unique(receiver_LF_rows);
receiver_com_rows = unique(receiver_com_rows);
receiver_TF_rows = unique(receiver_TF_rows);
end

%% Inter-foot parameters
% Loops collate lead foot, trail foot and com data into their
% respective
% forehand movements, i.e. feet positions have been digitised and
% sorted
for q=1:1
for n=1:length(receiver_LF_rows)
    if n>=length(receiver_LF_rows)
        p=length(receiver_LF);
        receiver_LF_data{n}=receiver_LF(receiver_LF_rows(n):p,:);
    else
        receiver_LF_data{n}=receiver_LF(receiver_LF_rows(n):receiver_LF_rows(n
+1),:);
        receiver_LF_data{n}(length(receiver_LF_data{n}),:)=[];
    end
end
for n=1:length(receiver_TF_rows)
    if n>=length(receiver_TF_rows)
        p=length(receiver_TF);

```



```

        receiver_TF_data{n}=receiver_TF(receiver_TF_rows(n):p,:);
    else

receiver_TF_data{n}=receiver_TF(receiver_TF_rows(n):receiver_TF_rows(n
+1),:);
    receiver_TF_data{n}(length(receiver_TF_data{n}),:)=[];
    end
end
for n=1:length(receiver_com_rows)
    if n>=length(receiver_com_rows)
        p=length(receiver_com);
        receiver_com_data{n}=receiver_com(receiver_com_rows(n):p,:);
    else

receiver_com_data{n}=receiver_com(receiver_com_rows(n):receiver_com_ro
ws(n+1),:);
    receiver_com_data{n}(length(receiver_com_data{n}),:)=[];
    end
end
end

if isempty(receiver_LF_rows)
else
% Receiver step length, frequency contact time, contact displacement
m=length(receiver_LF_data);
for s=1:m
    for l=1:2:length(receiver_TF_data{s})
        if l >= length(receiver_TF_data{s})
            p=length(receiver_TF_data{s});
            receiver_sl=(receiver_LF_data{s}(p,:)-
receiver_TF_data{s}(p,:));

receiver_sl=(horzcat(receiver_sl,(sqrt(((receiver_sl(:,2)).^2)+((recei
ver_sl(:,3)).^2)))));
            receiver_slr{l}=receiver_sl;
        else
            receiver_sl=(receiver_LF_data{s}(l,:)-
receiver_TF_data{s}(l,:));

receiver_sl=(horzcat(receiver_sl,(sqrt(((receiver_sl(:,2)).^2)+((recei
ver_sl(:,3)).^2)))));
            receiver_slr{l}=receiver_sl;
        end
    end
    for u=2:2:length(receiver_TF_data{s})
        if u >= length(receiver_TF_data{s})
            f=length(receiver_TF_data{s});
            receiver_slb=(receiver_LF_data{s}(f,:)-
receiver_TF_data{s}(f,:));

receiver_slb=(horzcat(receiver_slb,(sqrt(((receiver_slb(:,2)).^2)+((re
ceiver_slb(:,3)).^2)))));
            receiver_slr{u}=receiver_slb;
        else
            receiver_slb=(receiver_LF_data{s}(u,:)-
receiver_TF_data{s}(u,:));

receiver_slb=(horzcat(receiver_slb,(sqrt(((receiver_slb(:,2)).^2)+((re
ceiver_slb(:,3)).^2)))));
            receiver_slr{u}=receiver_slb;
        end
    end
end

```

```

        end
        t=[];
        for q = 1:length(receiver_slr)
            [r]=(find(~isempty(receiver_slr{q})));
            if r == 1
                t=vertcat(t,q);
            end
        end
        receiver_step_length{s}=receiver_slr(t);
        receiver_step_freq{s}=(length(receiver_step_length{s}))/((((receiver_
        TF_data{s}(end-1,1))-(receiver_TF_data{s}(1,1))/cr)));
        receiver_LF_contacts{s}=(length(receiver_LF_data{s}))/2;
        receiver_TF_contacts{s}=(length(receiver_TF_data{s}))/2;
        rlfct=[];
        rtfct=[];
        rlfd=[];
        rtfd=[];
        receiver_LF_contact_time{s}=[];
        for h=2:2:length(receiver_LF_data{s})
            rlfct=vertcat(((receiver_LF_data{s}(h,1)-receiver_LF_data{s}(h-
            1,1))/cr),rlfct);
            receiver_LF_contact_time{s}=flipud(rlfct);
            rlfd_xy=horzcat((receiver_LF_data{s}(h,2)-receiver_LF_data{s}(h-
            1,2)), (receiver_LF_data{s}(h,3)-receiver_LF_data{s}(h-1,3)));
            rlfd_r=sqrt(((rlfd_xy(:,1)).^2)+((rlfd_xy(:,2)).^2));
            rlfd_res=horzcat(rlfd_xy,rlfd_r);
            rlfd=vertcat(rlfd,rlfd_res);
            receiver_LF_displacement{s}=rlfd;
        end
        for h=2:2:length(receiver_TF_data{s})
            rtfct=vertcat(((receiver_TF_data{s}(h,1)-receiver_TF_data{s}(h-
            1,1))/cr),rtfct);
            receiver_TF_contact_time{s}=flipud(rtfct);
            rtfd_xy=horzcat((receiver_TF_data{s}(h,2)-receiver_TF_data{s}(h-
            1,2)), (receiver_TF_data{s}(h,3)-receiver_TF_data{s}(h-1,3)));
            rtfd_r=sqrt(((rtfd_xy(:,1)).^2)+((rtfd_xy(:,2)).^2));
            rtfd_res=horzcat(rtfd_xy,rtfd_r);
            rtfd=vertcat(rtfd,rtfd_res);
            receiver_TF_displacement{s}=rtfd;
        end
    end
    end
    for q=1:1
        for n=1:length(server_LF_rows)
            if n>=length(server_LF_rows)
                p=length(server_LF);
                server_LF_data{n}=server_LF(server_LF_rows(n):p,:);
            else
                server_LF_data{n}=server_LF(server_LF_rows(n):server_LF_rows(n+1),:);
                server_LF_data{n}(length(server_LF_data{n}),:)=[];
            end
        end
    end
    for n=1:length(server_TF_rows)
        if n>=length(server_TF_rows)
            p=length(server_TF);
            server_TF_data{n}=server_TF(server_TF_rows(n):p,:);
        else
            server_TF_data{n}=server_TF(server_TF_rows(n):server_TF_rows(n+1),:);
            server_TF_data{n}(length(server_TF_data{n}),:)=[];
        end
    end

```

```

end
for n=1:length(server_com_rows)
    if n>=length(server_com_rows)
        p=length(server_com);
        server_com_data{n}=server_com(server_com_rows(n):p,:);
    else
        server_com_data{n}=server_com(server_com_rows(n):server_com_rows(n+1),:);
        server_com_data{n}(length(server_com_data{n}),:)=[];
    end
end
end
if isempty(server_LF_rows)
else
% Server step length, frequency contact time, contact displacement
m=length(server_LF_data);
for s=1:m
    for l=1:2:length(server_TF_data{s})
        if l >= length(server_TF_data{s})
            p=length(server_TF_data{s});
            server_sl=(server_LF_data{s}(p,:)-server_TF_data{s}(p,:));

server_sl=(horzcat(server_sl,(sqrt(((server_sl(:,2)).^2)+((server_sl(:,3)).^2)))));
            server_slr{l}=server_sl;
        else
            server_sl=(server_LF_data{s}(l,:)-server_TF_data{s}(l,:));

server_sl=(horzcat(server_sl,(sqrt(((server_sl(:,2)).^2)+((server_sl(:,3)).^2)))));
            server_slr{l}=server_sl;
        end
    end
    for u=2:2:length(server_TF_data{s})
        if u >= length(server_TF_data{s})
            f=length(server_TF_data{s});
            server_slb=(server_LF_data{s}(f,:)-server_TF_data{s}(f,:));

server_slb=(horzcat(server_slb,(sqrt(((server_slb(:,2)).^2)+((server_slb(:,3)).^2)))));
            server_slr{u}=server_slb;
        else
            server_slb=(server_LF_data{s}(u,:)-server_TF_data{s}(u,:));

server_slb=(horzcat(server_slb,(sqrt(((server_slb(:,2)).^2)+((server_slb(:,3)).^2)))));
            server_slr{u}=server_slb;
        end
    end
    t=[];
    for q = 1:length(server_slr)
        [r]=(find(~isempty(server_slr{q})));
        if r == 1
            t=vertcat(t,q);
        end
    end
    server_step_length{s}=server_slr(t);
    server_step_freq{s}=(length(server_step_length{s}))/((((server_TF_data{s}(end-1,1))-(server_TF_data{s}(1,1)))/cr));
    server_LF_contacts{s}=(length(server_LF_data{s}))/2;

```

```

server_TF_contacts{s}=(length(server_TF_data{s}))/2;
slfct=[];
slfd=[];
stfct=[];
stfd=[];
server_LF_contact_time{s}=[];
for h=2:2:length(server_LF_data{s})
    slfct=vertcat(((server_LF_data{s}(h,1)-server_LF_data{s}(h-1,1))/cr),slfct);
    server_LF_contact_time{s}=flipud(slfct);
    slfd_xy=horzcat((server_LF_data{s}(h,2)-server_LF_data{s}(h-1,2)), (server_LF_data{s}(h,3)-server_LF_data{s}(h-1,3)));
    slfd_r=sqrt(((slfd_xy(:,1)).^2)+((slfd_xy(:,2)).^2));
    slfd_res=horzcat(slfd_xy,slfd_r);
    slfd=vertcat(slfd,slfd_res);
    server_LF_displacement{s}=slfd;
end
for h=2:2:length(server_TF_data{s})
    stfct=vertcat(((server_TF_data{s}(h,1)-server_TF_data{s}(h-1,1))/cr),stfct);
    server_TF_contact_time{s}=flipud(stfct);
    stfd_xy=horzcat((server_TF_data{s}(h,2)-server_TF_data{s}(h-1,2)), (server_TF_data{s}(h,3)-server_TF_data{s}(h-1,3)));
    stfd_r=sqrt(((stfd_xy(:,1)).^2)+((stfd_xy(:,2)).^2));
    stfd_res=horzcat(stfd_xy,stfd_r);
    stfd=vertcat(stfd,stfd_res);
    server_TF_displacement{s}=stfd;
end
end
end
%%% Point descriptors
% Identifies point as 'ace' or 'rally' and number of strokes in point.
% DOES NOT ACCOUNT FOR OVERRULINGS ETC.!
[m,n]=size(ball_surface_c);
if m<= 1
    point_type = 'Ace';
    n_strokes = m; % number of 'successful strokes, i.e. when play
continues
else
    point_type = 'Rally';
    n_strokes = m;
end
% Choose date to append to filename (defined as serve frame number) to
save workspace
choice = questdlg('Choose date (17th or 18th)', ...
    'File save', ...
    '17','18','18');
switch choice
    case '17'
        vids = '17_';
    case '18'
        vids = '18_';
    end
path = 'C:\LowResTennis_CalFiles\Analysis_Output\';
fn=num2str(service_t);
file='.mat';
filename=[path,vids,fn,file];
save(filename);

```


Appendix 3

A.3.1. *Permission to film and accreditations obtained via the Lawn Tennis Association.*



The Lawn Tennis Association
The National Tennis Centre
100 Perry Lane
Rochampton
London SW16 5 SJQ
T 020 8487 7000
F 020 8487 7501
www.LTA.org.uk

13th March, 2013

Ref: Mr Marcus Dunn, PhD candidate – Research at Barclays ATP World Tour Finals
2010 and 2011

To whom it concerns

I am writing to confirm the details relating to the permissions given to Mr Marcus
Dunn (PhD candidate at Sheffield Hallam University) to film matches at the Barclays
ATP World Tour Finals in November 2010 and 2011 held at the O2 Arena,
Greenwich, London.

As a co-organiser (and host national association) for this event the Lawn Tennis
Association gave Mr Dunn permission to film singles matches and take
measurements of the court surface for research purposes only on the understanding
that the project would focus on the player court surface interaction in elite men's
professional tennis.

If you have any questions relating to the permission and access given to Mr Dunn
please don't hesitate to contact me directly.

Best wishes,

A handwritten signature in black ink, appearing to read "Karl Cooke".

Dr Karl Cooke
Sports Science Manager
Karl.cooke@lta.org.uk



Patron: Her Majesty The Queen
Honorary President: HRH The Duchess of Gloucester GCMG

President: Peter Braberton
Deputy President: Cathie Sabin
Chief Executive: Roger Draper

CONFIDENTIAL



**Faculty of Health and Wellbeing
Research Ethics Committee**

Sport & Exercise Research Ethics Review Group

APPLICATION FOR ETHICS APPROVAL OF RESEARCH

In designing research involving humans, principal investigators should be able to demonstrate a clear intention of benefit to society and the research should be based on sound principles. These criteria will be considered by the Sport and Exercise Research Ethics Review Group before approving a project. **ALL** of the following details must be provided, either typewritten or word-processed preferably at least in 11 point font.

Please either tick the appropriate box or provide the information required.

1) Date of application	20 December 2013
2) Anticipated date of completion of project	1 July 2013
3) Title of research	Semi-automatic tennis player step and movement characterisation
4) Subject area	Sports Engineering
5) Principal Investigator Name Email address @ SHU Telephone/Mobile number Student number (if applicable)	Marcus Dunn m.dunn @shu.ac.uk 0114 225 5867 14025464
6) State if this study is: (If the project is undergraduate or postgraduate please state module name and number)	<input checked="" type="checkbox"/> Research <input type="checkbox"/> Undergraduate <input type="checkbox"/> Postgraduate Module name: Module number:
7) Director of Studies/Supervisor/ Tutor name	Dr. Simon Goodwill

8) Intended duration and timing of project?	<p>26 November 2011: film at O2 Arena, London.</p> <p>December 2011 - June 2013: development of automatic player tracking and foot-surface contact identification algorithm.</p> <p>June - July 2013: Application and analysis of tennis player movement and step strategy in match play tennis.</p>
9) Location of project If external to SHU, provide evidence in support (see section 17)	<p>Filming at the 2011 ATP World Tour Finals. This tournament will be held at the O2 Arena, London.</p> <p>Development, testing and analysis at SHU.</p>
10) State if this study is:	<p><input checked="" type="checkbox"/> New</p> <p><input checked="" type="checkbox"/> Collaborative (please include appropriate agreements in section 17)</p> <p><input type="checkbox"/> Replication of :</p>

11) Purpose and benefit of the research

Statement of the research problem with any necessary background information (no more than 1 side of A4)

Current knowledge of tennis player-surface interaction with different tennis court surfaces, particularly during competition, is limited (Miller, 2006). Human centred approaches for the measurement of surface slipperiness highlight that step strategy reflects interactions that occur at the shoe-surface interface (Grönqvist et al. 2001). The measurement of tennis player step strategy would therefore contribute to the understanding of tennis player-surface interactions with different court surfaces.

A manual system for characterising tennis player step and movement strategy during competition was developed. Footage was obtained *in situ* and provided a novel insight into gender differences for step and movement strategy during competition (Dunn et al., 2011). However manual analyses are time consuming and prone to subjective digitising error (Glazier and Irwin, 2001). An automatic and objective method for identifying player and foot-surface contact location is required.

An automatic, foot-surface contact identification (FSCi) algorithm has been developed. The FSCi algorithm analyses standard colour video sequences of gait activities, i.e. walking, running etc., and identifies foot-surface contacts automatically based on image processing techniques; no markers or user intervention is required. A tracking algorithm, to track tennis players in competition, will be developed specifically to enable the application of the FSCi algorithm to video footage of tennis (PT-FSCi algorithm). Footage of match play tennis is required to identify the challenges associated with filming and analysing tennis player step and movement strategy with the PT-FSCi algorithm. Furthermore, footage of match play tennis is required to identify the validity of player step and movement data identified using the PT-FSCi algorithms.

The application and analysis of tennis player step and movement strategy using an automated technique will aid the future understanding of tennis player-surface interactions with tennis court surfaces.

Dunn, M., Wheat, J., Haake, S. and Goodwill, S. (2011). Assessing tennis player interactions with tennis courts. In Vilas-Boas, J.P., Machado, L., Kim, W., Veloso, A.P. (Eds.), Proceedings of the 29th International Conference on Biomechanics in Sports, pp. 859 - 862.

Glazier, P. and Irwin, G. (2001). Validity of stride length estimates obtained from Optojump. In Blackwell, J.R. and Saunders, R.H. (Eds.), Proceedings of the 19th International Conference on Biomechanics in Sports, pp. 98 - 101.

Grönqvist, R., Abeysekera, J., Gard, G., Hsiang, S., Leamon, T.B., Newman, D.J., Gielo-Perczak, K., Lockhart, T.E. and Pai, C, Y.-C. (2001). Human centred approaches in slipperiness measurement. *Ergonomics*, 44(13), 1167 - 1199.

Miller, S. Modern tennis rackets, balls and surfaces. (2006). *British Journal of Sports Medicine*, 40, 401 - 405.

12) Participants	
12.1 Number	<p>The project will film singles tennis players competing on a single tennis court.</p> <p>Matches of four male tennis players will be filmed over a single day.</p>
12.2 Rationale for this number (eg calculations of sample size, practical considerations)	<p>Due to the novel nature of analyses being performed, no previous data exist for the basis of sample size calculations.</p> <p>Sample size represents practical filming and analysis constraints, i.e. filming two matches over one day and time cost of manual digitisation (validation of PT-FSCi algorithm).</p> <p>Furthermore, camera footage and data from the competition are desirable for future analysis.</p>
12.3 Criteria for inclusion and exclusion (eg age and sex)	<p>The participants will be male entrants of the 2011 ATP World Tour Finals.</p> <p>See 12.4 for details on research consent.</p>
12.4 Procedures for recruitment (eg location and methods)	<p>The Grand Slam Committee (GCS) outlines the Grand Slam Rules (attached). The Grand Slam Rules state in Article I (section E) that "Each player grants and assigns to the GCS and the management of the events that he enters, the right in perpetuity to make, use and show from time to time and at their discretion, motion pictures, still pictures and live, taped or filmed television and other reproductions of him during said events and in connection with the promotion of said events without compensation for himself, his heirs, devisees, executors, administrators or assigns."</p> <p>The GCS therefore grants or denies permission to film tennis players during entered competitions, for purposes agreed by the GCS (letter of collaboration).</p> <p>The Lawn Tennis Association (LTA) sought permission from the relevant GCS body (ATP Tour) to perform tennis match filming for research purposes.</p>
12.5 Does the study have *minors or †vulnerable adults as participants?	<p>[] Yes [✓] No</p>

12.6 Is CRB Disclosure required for the Principal Investigator? (to be determined by Risk Assessment)	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No If yes, is standard <input type="checkbox"/> or enhanced <input type="checkbox"/> disclosure required?
12.7 If you ticked 'yes' in 12.5 and 'no' in 12.6 please explain why:	

*Minors are participants under the age of 18 years.

‡Vulnerable adults are participants over the age of 16 years who are likely to exhibit:

- a) learning difficulties
- b) physical illness/impairment
- c) mental illness/impairment
- d) advanced age
- e) any other condition that might render them vulnerable

13) Details of the research design

13.1 Provide details of intended methodological procedures and data collection.

(For MSc students conducting a scientific support project please provide the following information: a. needs analysis; b. potential outcome; c proposed interventions).

1. Video

Central to the project is filming of tennis matches at the ATP World Tour Finals. In more detail:

- a. A single camera will be used to capture each match. This camera will be situated in the stands and will capture images of half of the tennis court.
- b. Footage will be captured from just before the start of the first match to just after the end of the last match. There are two matches performed on a single day.
- c. Footage will be stored on the computers detailed in the Equipment section.
- d. Access to the footage will be restricted. Apart from images used in the thesis and research publications only the principal investigator and the principal investigator's supervisory team (Professor Steve Haake, Dr Simon Goodwill and Dr Jon Wheat) will have access.

2. Data

Data, such as player displacement, step length and step frequency, will be calculated from the footage. In more detail:

- a. Data will be manually digitised from the footage.
- b. Data will be analysed using the player tracking and foot-surface contact identification algorithm.
- c. Data will be stored on the computers detailed in the Equipment section.
- d. Access to the footage will be restricted. Apart from images used in the thesis and research publications only the principal investigator and the principal investigator's supervisory team (Professor Steve Haake, Dr Simon Goodwill and Dr Jon Wheat) will have access.

3. Equipment

The following computer equipment will be used to capture and store the footage and data:

- a. Laptop:
 - Footage and data will be stored on this computer. The footage and data will be used to for research output.
 - It will be located in A212 Collegiate Hall.
 - Only the principal investigator and the principal investigator's director of studies will have access to this computer.

4. Security

The following steps will be taken to secure footage and data:

- a. The computers detailed in the Equipment section will use the Microsoft Windows 7 Ultimate. Only the principal investigator and the principal investigator's director of studies will have the password.

13.2 Are these "minor" procedures as defined in Appendix 1 of the ethics guidelines?

☒ Yes ☐ No

13.3 If you answered 'no' in section 13.2, list the procedures that are not minor

13.4 Provide details of the quantitative and qualitative analysis to be used

Bland and Altman 95% limits of agreement will be used to assess agreement between player step and movement data obtained automatically (algorithm) and manually (criterion measure).

Descriptive statistics will be used to explore tennis player step and movement strategy obtained automatically from footage of match play tennis.

14) Substances to be administered (refer to Appendix VI of the ethics procedures)**14.1 The protocol does not involve the administration of pharmacologically active substances or nutritional supplements.**

Please tick box if this statement applies and go to section 15) [✓]

14.2 Name and state the risk category for each substance. If a COSHH assessment is required state how the risks are to be managed.**15) Degree of discomfort that participants might experience**

Consider the degree of physical and psychological discomfort that will be experienced by the participants. State the details which must be included in the participant information sheet to ensure that the participants are fully informed about any discomfort that they may experience.

Not applicable. The principal investigator will be passively monitoring tennis matches from the stands.


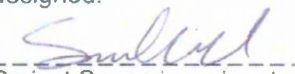
16) Outcomes of Risk Assessment

Provide details of the risk and explain how the control measures will be implemented to manage the risk.

Overall risk is LOW, which is driven by the risk of electrical equipment being present in publicly accessible stands.

The risk is mitigated by the equipment and principal investigator being positioned in the stands.

17) Attachments	Tick box
17.1 Risk assessment (including CRB risk assessment)	✓
17.2 COSHH assessment	N/A
17.3 Participant information sheet (this should be addressed directly to the participant (ie you will etc) and in a language they will understand)	N/A
17.4 Informed consent form	N/A
17.5 Pre-screening questionnaire	N/A
17.6 Collaboration evidence/support correspondence from the organisation consenting to the research (this must be on letterhead paper and signed) See sections 9 & 10.	✓
17.7 CRB Disclosure certificate <u>or</u> where not available CRB application form	N/A
17.8 Clinical Trails form (FIN 12)	N/A

18. Signature Principal Investigator	<p>Once this application is approved, I will undertake the research study as approved. If circumstances necessitate that changes are made to the approved protocol, I will discuss these with my Project Supervisor. If the supervisor advises that there should be a resubmission to the Sport and Exercise Research Ethics Review Group, I agree that no work will be carried out using the changed protocol until approval has been sought and formally received.</p> <p> Date <u>20/12/13</u></p> <p>Principal Investigator signature</p> <p>Name <u>MARCUS DUNN</u></p>
19. Approval Project Supervisor to sign either box A or box B as applicable (refer to Appendix I and the flowchart in appendix VI of the ethics guidelines)	<p>Box A:</p> <p>I confirm that the research proposed is based solely on 'minor' procedures, as outlined in Appendix 1 of the HWB Sport and Exercise Research Ethics Review Group 'Ethics Procedures for Research with Humans as Participants' document, and therefore does not need to be submitted to the HWB Sport and Exercise Research Ethics Review Group.</p> <p>In terms of ethics approval, I agree the 'minor' procedures proposed here and confirm that the Principal Investigator may proceed with the study as designed.</p> <p> Date <u>20/12/13</u></p> <p>Project Supervisor signature</p> <p>Name <u>SIMON GOODWILL</u></p>
	<p>Box B:</p> <p>I confirm that the research proposed is <u>not</u> based solely on 'minor' procedures, as outlined in Appendix 1 of the HWB Sport and Exercise Research Ethics Review Group 'Ethics Procedures for Research with Humans as Participants' document, and therefore <u>must</u> be submitted to the HWB Sport and Exercise Research Ethics Review Group for approval.</p> <p>I confirm that the appropriate preparatory work has been undertaken and that this document is in a fit state for submission to the HWB Sport and Exercise Research Ethics Review Group.</p> <p>_____ Date _____</p> <p>Project Supervisor signature</p> <p>Name _____</p>
20. Signature Technician	<p>I confirm that I have seen the full and approved application for ethics approval and technical support will be provided.</p> <p>_____ Date _____</p> <p>Technician signature</p> <p>Name _____</p>



Sheffield Hallam University

**Faculty of Health and Wellbeing Research Ethics Committee
Sport and Exercise Research Ethics Review Group**

Risk Assessment Pro Forma

****Please ensure that you read the accompanying
Risk Assessment Risk Ranking document before completing this form****

Title of research	Semi-automatic tennis player step and movement characterisation
--------------------------	---

Date Assessed	20 December 2013
----------------------	------------------

Assessed by (Principal Investigator)	Marcus Dunn
---	-------------

Signed	Position
	Principal Investigator

Activity	Risks	Control Measures
Filming	Risk of [physical injury] caused by tripping over cabling. R1 = C1 x L1. LOW RISK	Filming will take place in the stands that have public access. Equipment will not be placed in a thoroughfare. Equipment will not be left unattended. Cables of sufficient length will be used. Cables of will be tidied sufficiently.
Filming	Risk of [physical injury] caused by falling tripod. R1 = C1 x L1. LOW RISK	Filming will take place in the stands that have public access. Equipment will not be placed in a thoroughfare. Equipment will not be left unattended. Cables of sufficient length will be used. Cables of will be tidied sufficiently.
Filming	Risk of [physical injury] caused by transporting equipment to and from venue. R1 = C1 x L1. LOW RISK	Equipment will be kept to the minimum. Equipment will be carried in suitable containers and loaded carefully into vehicles.


Risk Evaluation (Overall)
LOW

General Control Measures
Is a pre-screen medical questionnaire required? Yes [] No [✓]

Emergency Procedures
None required.

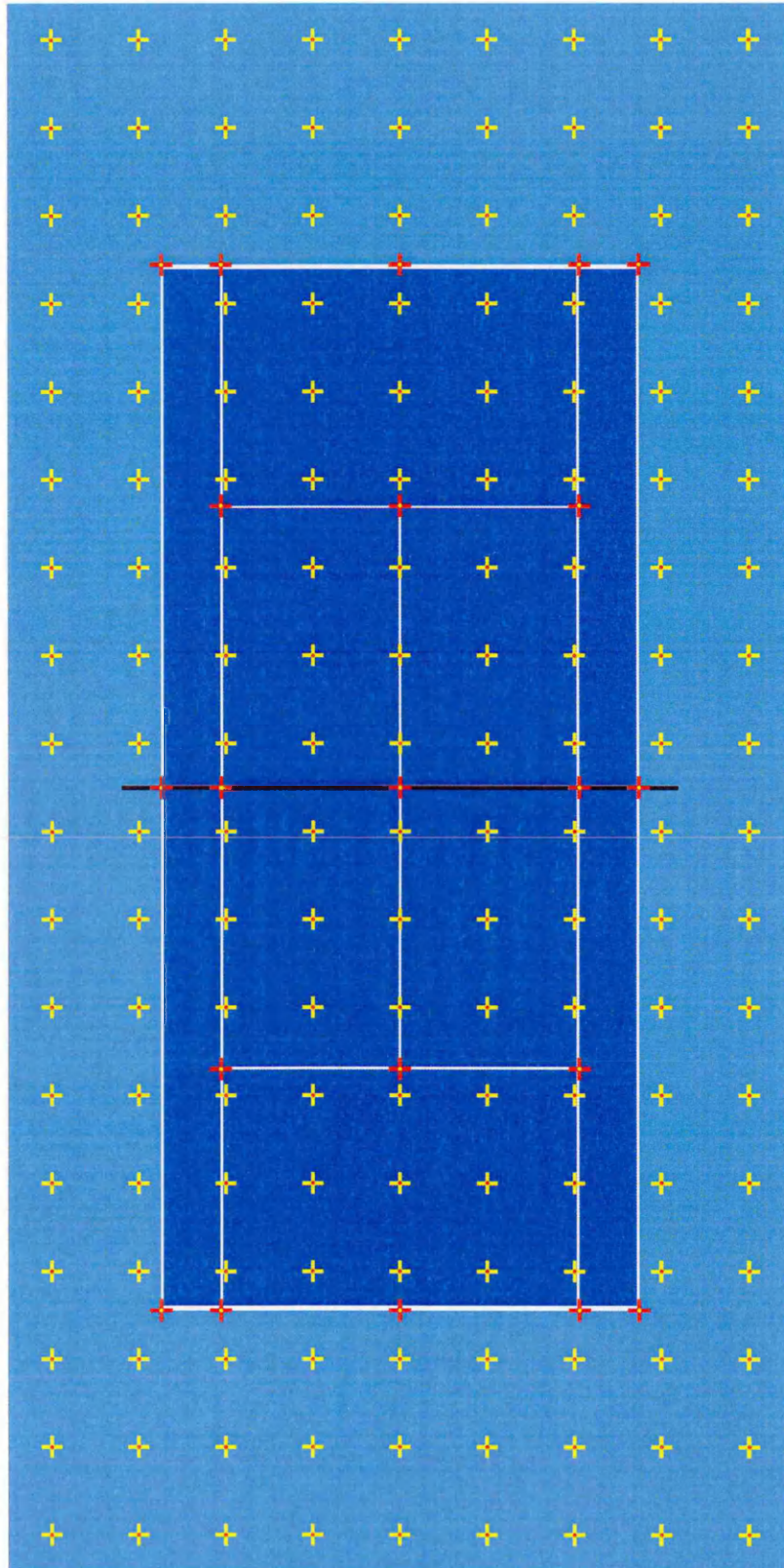
Monitoring Procedures
None required.

Review Period	None.
----------------------	-------

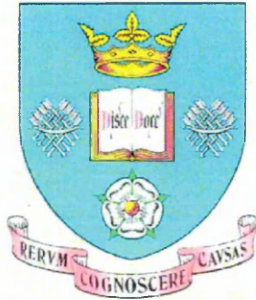
Reviewed By (Supervisor)	Date
	20/12/13

Appendix 4

A.4.1. Scale (1:30) tennis court model.



A. Internal Report: Assessment of Calibration Techniques



The University of Sheffield
Department of Mechanical Engineering
Sports Engineering Research Group

3D Image Reconstruction from a Stereo Camera Pair: Assessment of Calibration Techniques

Neil Whyld
August 2004

Table of Contents

1. INTRODUCTION	1
2. EXPERIMENTAL PROCEDURE	2
2.1. Grid Calibration	2
2.2. Checkerboard Calibration	4
3. RESULTS	5
4. DISCUSSION	8
5. CONCLUSIONS.....	11
Appendix 1	12

1. INTRODUCTION

The use of stereo camera imaging to re-construct points in three-dimensions is well documented. There are several techniques available for camera calibration and position reconstruction. The most suitable method for a given application is very much dependent on the specifics of that situation.

This short report aims to assess the suitability of three different methods of calibration/reconstruction for analysing tennis ball and racket interactions, as well as full scale player testing. It will focus predominantly on accuracy, flexibility and ease of use.

The methods to be assessed are:-

- Grid calibration using:-
 - Standard DLT algorithm (calculates 11 independent camera parameters which are used to relate 2D positions (u,v) in the two camera image planes to the corresponding 3D position (x,y,z) in the global reference frame).
 - Modified DLT algorithm (Development of the Standard algorithm which ensures the three principle axes (x,y,z) are orthogonal).
- Checkerboard calibration (Algorithm uses the image distortions of a checkerboard held in different orientations to calculate the required camera parameters).

2. EXPERIMENTAL PROCEDURE

2.1. Grid Calibration

A planar calibration grid was constructed from “Bosch” extruded aluminium beams. The structure was painted black and 18 highly reflective markers were attached at carefully measured locations. The evenly spaced markers were divided into two groups (numbers and letters) to allow for a variety of different calibration and verification point combinations to be investigated. The x,y and z locations of these markers are given in Appendix 1.

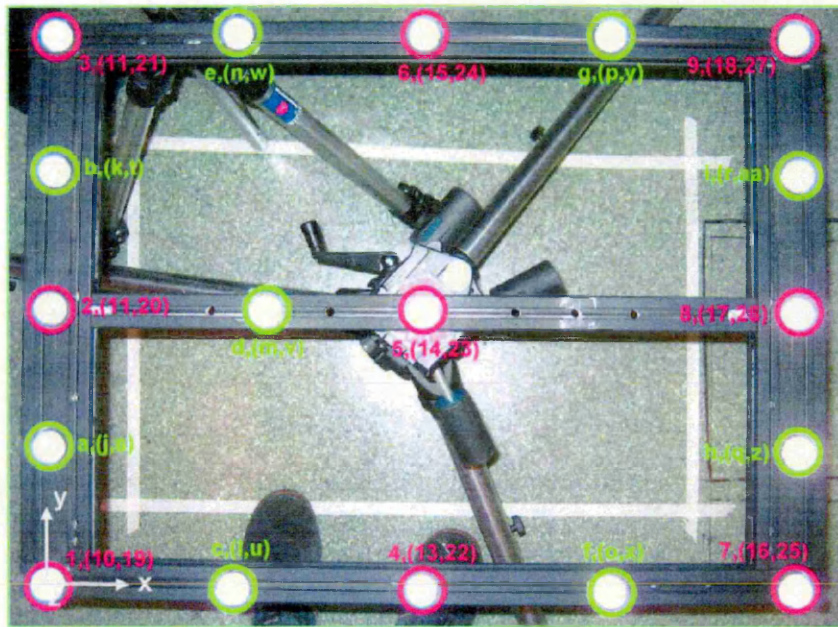


Figure 2.1 Plan view of the calibration grid, showing point labels.

Attaching the calibration grid to a tripod allowed the whole frame to be translated in the z direction. Recording the points at three different heights effectively increased the number of calibration and verification points from two sets of 9 to two sets of 27. The lowest level was taken as a reference height of 0 mm and contained points 1-9 and a-i as illustrated in Figure 2.1. The grid was then raised by 126 mm to produce the second level, containing an identical set of points from 10-18 and j-r. Finally raising the grid to a height of 240 mm above the reference level provided points 19-27 and s-aa.

Taking point 1 (at the reference height) as the origin (0,0,0) the relative position of all other points were calculated by careful measurement from the origin.

Two MotionCorder high speed video cameras were used as the stereo pair. They were orientated at approximately 90° to each other and with the calibration grid in the centre of view for both cameras. See Figure 2.2.

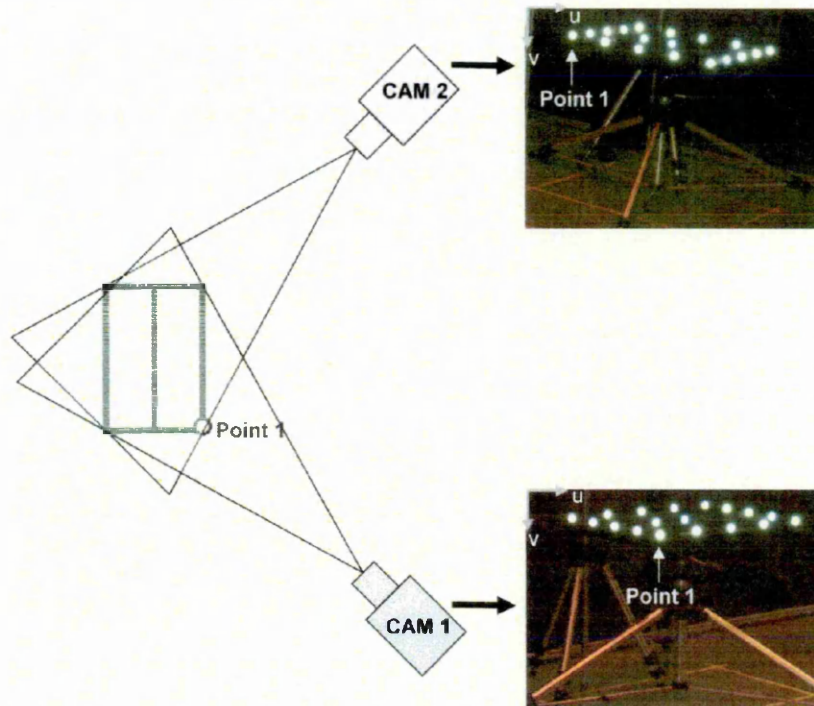


Figure 2.2 Stereo camera setup with views from each camera.

A short amount of footage was taken from both cameras at each of the three grid heights. The footage from each camera was then analysed using Richimas v3.2, and the u,v pixel location of each point (1-27 and a-aa) recorded.

Inputting the u,v data obtained for each calibration point together with their positions relative to the local coordinate system origin (point 1) into routines created in Matlab enabled the calculation of the 11 DLT parameters from both the Standard algorithm and the Modified algorithm. The input of a separate file of u,v data for the verification

points enabled the Matlab routines to reconstruct the positions of the points in the local coordinate system (x,y,z). The difference between the reconstructed positions and the physically measured positions were then compared for accuracy for the two algorithms.

This procedure was repeated using different combinations of calibration and verification points.

2.2. Checkerboard Calibration

A rigid board containing a 14 x 14, 40mm square checkerboard pattern was positioned in different orientations within the calibration area (defined by the previous position of the calibration grid).

The checkerboard calibration was conducted using the same camera setup as that used in the grid calibration. Approximately 20 board orientations were recorded by each camera and analysed. A sample pair of images is shown in Figure 2.3.

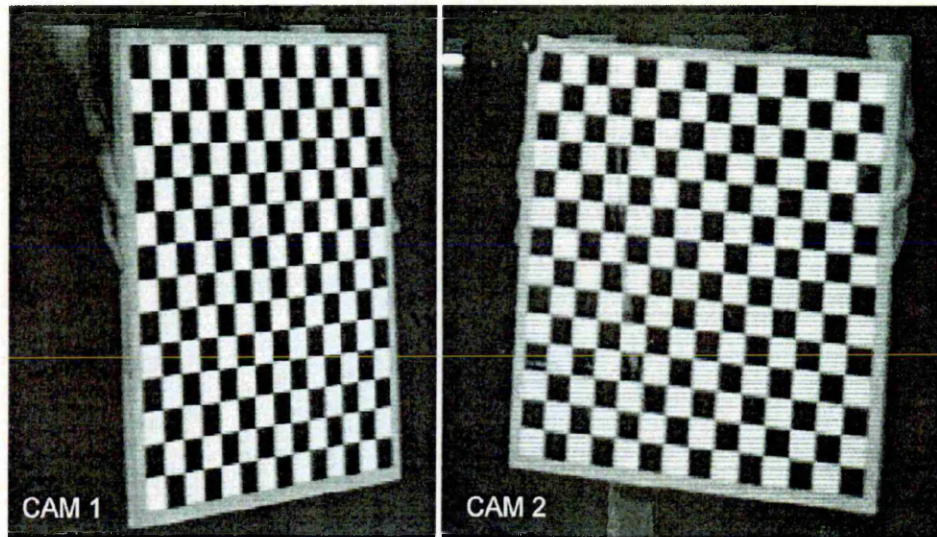


Figure 2.3 Images of the same checkerboard orientation from both stereo cameras

During the checkerboard analysis, it is possible to define a reference plane in any orientation. The planes used were chosen to correspond with the local coordinates set up by the calibration grid.

The first plane was defined using the xy axis and the second plane was defined with the xz plane, as shown in Figure 2.4.

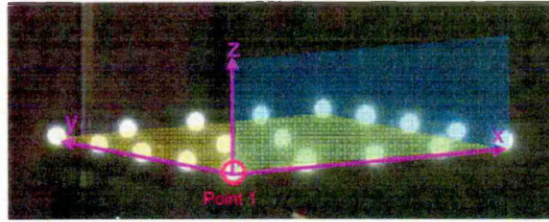


Figure 2.4 Checkerboard reference planes: First plane shown in yellow, second plane in blue.

3. RESULTS

Four combinations of grid calibration and verification points were investigated. Two tests used different sets of points for calibration and verification (i.e. numbers to calibrate the system and letters to check the reconstruction, and vice versa) and the other two sets used the same points for calibration and verification (i.e. numbers to calibrate and also to check reconstruction). These four tests were reconstructed using both the Standard DLT algorithm and the Modified DLT algorithm, and compared to the Checkerboard reconstruction results of the same points.

The results are summarised below:-

Grid calibration

CALIBRATION POINTS	VERIFICATION POINTS	AVERAGE ERROR (mm)		MAXIMUM ERROR (mm)	
		STAND DLT	MOD DLT	STAND DLT	MOD DLT
NUMBER	LETTER	1.07	3.85	2.85	19.50
LETTER	NUMBER	1.26	4.25	2.26	8.20
NUMBER	NUMBER	1.10	4.48	2.87	19.36
LETTER	LETTER	0.92	1.72	2.87	8.18

Table 3.1 Grid calibration results summary.

Checkerboard calibration

VERIFICATION POINTS	AVERAGE ERROR (mm)		MAXIMUM ERROR (mm)	
	XY FIXED	XZ FIXED	XY FIXED	XZ FIXED
NUMBER	2.5	2.04	9.04	7.44
LETTER	1.42		8.86	

Table 3.2 Checkerboard calibration results summary.

The following results focus on the use of number positions for system calibration and letter positions for the checking of re-construction accuracy:-

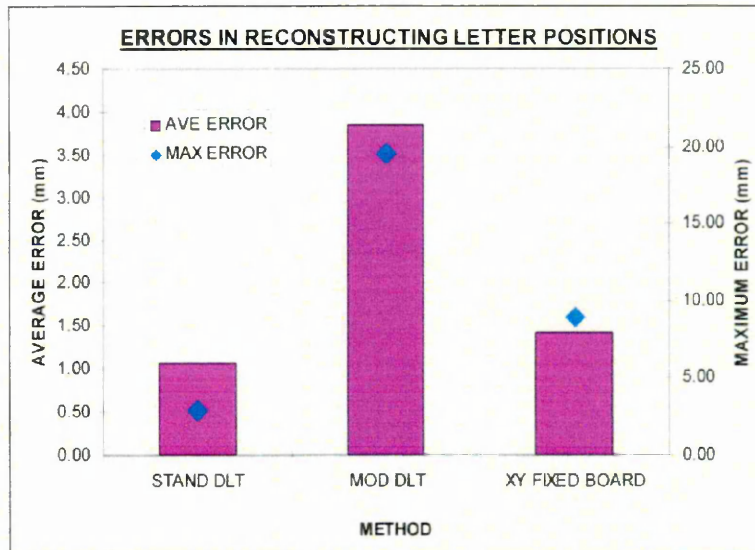


Figure 3.1 Average and maximum letter position reconstruction errors for the three calibration methods, using number positions for system calibration.

Looking at how the error in point reconstruction (in the x, y and z directions) vary with increasing total distance from the origin (point 1), we obtain the following results:-

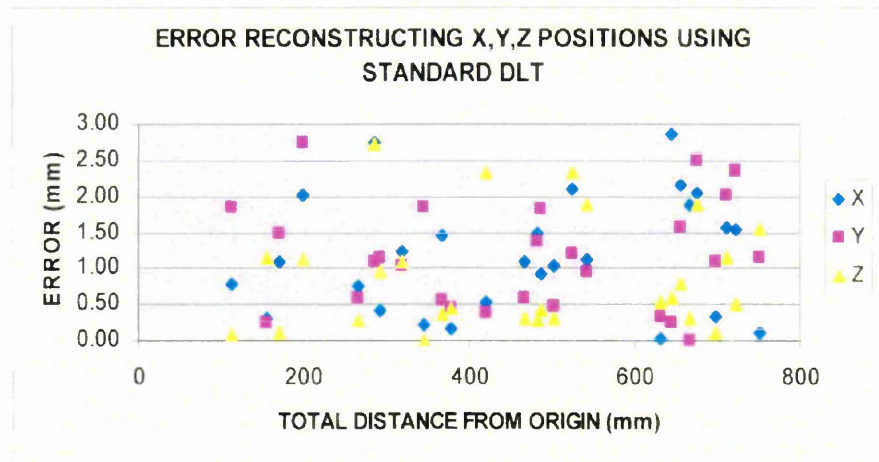


Figure 3.2 Change in reconstruction error with increasing total distance from the origin for the Standard DLT algorithm.

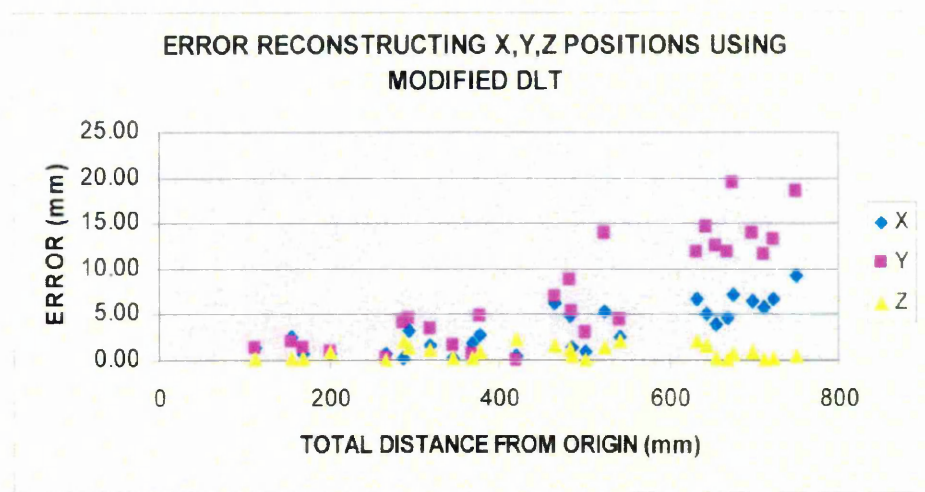


Figure 3.3 Change in reconstruction error with increasing total distance from the origin for the Modified DLT algorithm.

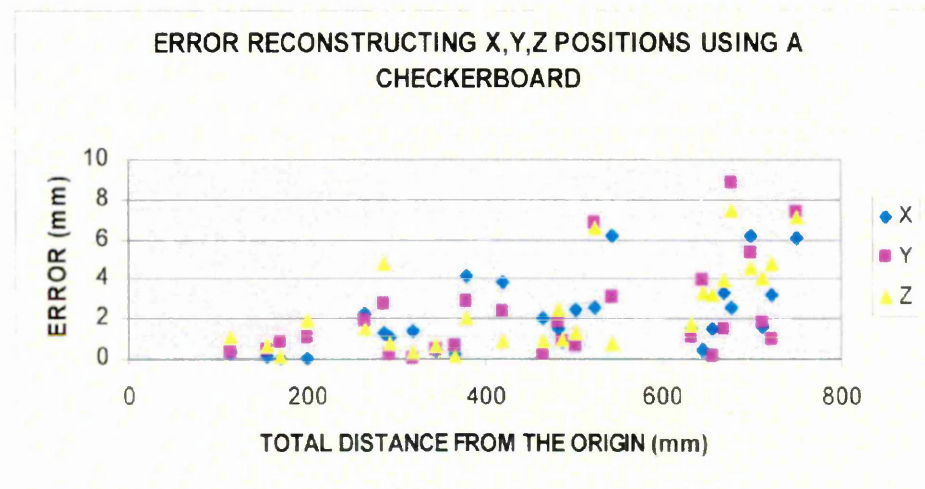


Figure 3.4 Change in reconstruction error with increasing total distance from the origin for the checkerboard method.

4. DISCUSSION

Looking at Table 3.1, we can see that the Standard DLT algorithm is producing reconstruction results of considerably higher accuracy than the Modified algorithm. The average and maximum errors produced by the Standard DLT algorithm remain virtually constant over the four tests performed; indicating that the results are valid. The results from the Modified algorithm are far more erratic with large differences in maximum error between tests.

The results shown in Table 3.2 indicate that in general, the Checkerboard method provided reconstruction accuracies lower than the Standard DLT method but higher than the Modified version. The Checkerboard method performed slightly better with the xz axis fixed compared to when the xy axis was fixed. This finding is likely to be specific only to this investigation however, and suggests that there could be significant errors in the true location of the test markers on the calibration grid. If the grid was perfectly square, with markers located in a truly orthogonal pattern, the Checkerboard calibration would provide identical results irrespective of which axis points were used to define the reference plane.

A skewed pattern of marker locations would explain the discrepancy between the Standard and Modified DLT results. As the Standard DLT algorithm has no constraint over the orthogonality of the axes, it merely fits the data points produced to the required parameter equations, giving very accurate results for what is an inherently inaccurate situation. Therefore if the Modified DLT algorithm is trying to fit non-orthogonal data points to an orthogonal model, we would expect the results to be pushed away from their true values, giving what are essentially low accuracy results for a theoretically high accuracy situation. The Checkerboard provides intermediary results as although it doesn't have the orthogonal constraint over the three principle axes of the Modified model, it does have a reference plane defined by the data points. Although the two axes defining this plane are not required to be at 90° to each other, the third axis (normal to the reference plane) is defined by the cross-product of the two other axes, and, must therefore be orthogonal to the reference plane, when in reality it may not be.

In summary: The Standard DLT algorithm allows for all axes to be non-orthogonal. Modified DLT assumes all axes to be orthogonal. The checkerboard method takes one axis to be orthogonal to the plane defined by the other two axes.

If a non-orthogonal grid is responsible for the results described above, one would assume that for the Modified DLT and the Checkerboard methods, the reconstruction errors should increase with distance from the origin. Additionally, due to the level of constraint of each method, the Modified DLT algorithm should be effected the most and thus produce higher errors than the Checkerboard method. Finally, due to the lack of constraints, the error produced by the Standard DLT algorithm should remain independent of the distance from the origin.

Graphs 3.2 to 3.4 support this theory and illustrate all of the trends mentioned above. The errors produced by the Standard DLT algorithm appear totally independent of the distance from the origin. The variation in values will be due to random errors associated with the manual digitisation of the control points. The errors produced by the Modified DLT algorithm clearly increase with distance from the origin, with the y and x positions being effected more than the z reconstruction. The errors associated with the Checkerboard calibration are also shown to be related to the distance away from the origin, but as predicted the level of error is considerably less than that produced by the Modified DLT algorithm.

For the size of test grid used, it is worth noting that the inaccuracy in grid construction needed to produce the levels of error encountered is extremely small.

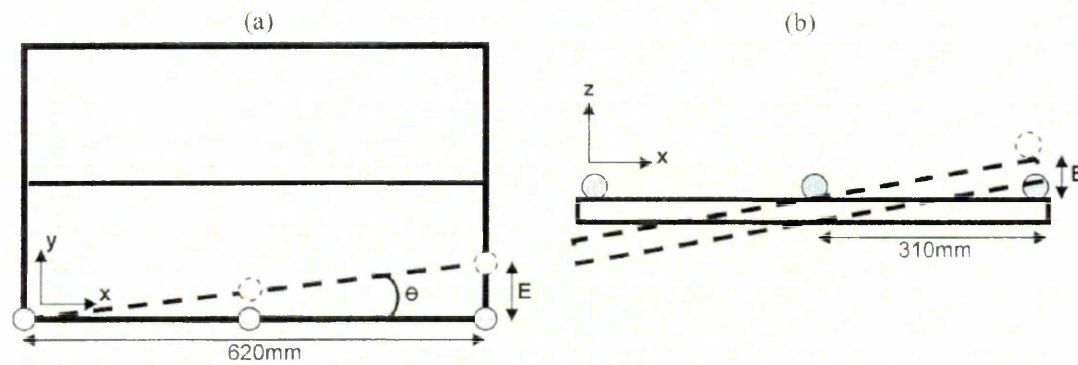


Figure 4.1 (a); Plan view illustrating error, E , produced in the y direction for a misalignment of markers along the x axis, θ degrees from the true alignment.
(b); Side view illustrating error, E , produced in the z direction due to misalignment of the calibration grid on the tripod.

Figure 4.1(a) shows that a very small error in the grid construction will produce very large errors in the actual position of the markers. It only requires the line of markers to be aligned 0.46° from the true orientation to offset the end marker by a distance of 5mm in the y direction. Further a rotation of only 0.92° in the line of the markers will provide an error in the y direction of 10mm.

If the grid is in anyway warped and/or not set absolutely level on the tripod, Figure 4.1(b) illustrates that the error in the z direction can be very significant for only small angle deviations. For an error of 5mm the grid alignment need only be 0.92° from its true orientation.

5. CONCLUSIONS

The Standard DLT algorithm proved to be the most “forgiving” to errors in grid construction as it does not constrain the principle axes to be orthogonal. Although this method accurately reconstructed the positions of test points, the true meaning of the results must be carefully considered before this method is used. If it is essential that the results be split into a truly orthogonal reference frame then this method provides no guidance as to the true level of accuracy.

Due to inaccuracies in the grid construction, the Modified DLT algorithm provided the least accurate results. However, the poor reconstruction results did highlight the fundamental problem of the grid not being accurate enough. If a suitably accurate grid were manufactured then this method would provide very accurate and reliable results. Producing this grid would be expensive and time consuming however.

Producing the levels of accuracy required to manufacture a suitable grid is very difficult, especially as its size is increased. The grids are also delicate and difficult to transport.

A far more versatile method of camera calibration is the Checkerboard technique. Its results were of acceptable accuracy given the test situation. A set of markers stuck to the floor of the test area would ensure that the principle axes remain orthogonal and true to the World reference frame. Perhaps the greatest advantage of this method however is its simplicity. The test object is quick and easy to manufacture and use. It is also relatively small and durable, and hence easy to transport.

Although the relative error associated with the production of a checkerboard is small in comparison to the construction of a calibration grid, it is not zero. Forming the grid from small sections printed on a standard A4 printer allows for the introduction of cutting and alignment errors. These errors can be eliminated relatively cheaply by having a large grid custom printed in one piece.

Appendix 1

Below are details of the measured locations of both sets of marker points. The values are given in millimetres and are measured from the origin (taken as Point 1).

POINT	X	Y	Z
1	0	0	0
2	0	230	0
3	0	461	0
4	311	0	0
5	311	230	0
6	311	461	0
7	622	0	0
8	622	231	0
9	622	462	0
10	0	0	126
11	0	230	126
12	0	461	126
13	311	0	126
14	311	230	126
15	311	461	126
16	622	0	126
17	622	231	126
18	622	462	126
19	0	0	240
20	0	230	240
21	0	461	240
22	311	0	240
23	311	230	240
24	311	461	240
25	622	0	240
26	622	231	240
27	622	462	240

POINT	X	Y	Z
a	0	114	0
b	0	345	0
c	155.5	0	0
d	181	230	0
e	155.5	461	0
f	466.5	0	0
g	466.5	461	0
h	622	114	0
i	622	346	0
j	0	114	126
k	0	345	126
l	155.5	0	126
m	181	230	126
n	155.5	461	126
o	466.5	0	126
p	466.5	461	126
q	622	114	126
r	622	346	126
s	0	114	240
t	0	345	240
u	155.5	0	240
v	181	230	240
w	155.5	461	240
x	466.5	0	240
y	466.5	461	240
z	622	114	240
aa	622	346	240

Tables 5.1 Measured location of the reflective position markers.

Appendix 6

A.6.1. Ethics application form: Validation of an automatic foot-surface contact identification algorithm.

CONFIDENTIAL



**Faculty of Health and Wellbeing
Research Ethics Committee**

Sport & Exercise Research Ethics Review Group

APPLICATION FOR ETHICS APPROVAL OF RESEARCH

In designing research involving humans, principal investigators should be able to demonstrate a clear intention of benefit to society and the research should be based on sound principles. These criteria will be considered by the Sport and Exercise Research Ethics Review Group before approving a project. **ALL** of the following details must be provided, either typewritten or word-processed preferably at least in 11 point font.

Please either tick the appropriate box or provide the information required.

1) Date of application	5-2-13
2) Anticipated date of completion of project	
3) Title of research	Validation of an automatic foot-surface contact identification algorithm
4) Subject area	Sports Engineering
5) Principal Investigator Name Email address @ SHU Telephone/Mobile number Student number (if applicable)	Marcus Dunn m.dunn@shu.ac.uk 07717410501 14025464
6) State if this study is: (If the project is undergraduate or	<input checked="" type="checkbox"/> Research <input type="checkbox"/> Undergraduate

postgraduate please state module name and number)	<input type="checkbox"/> Postgraduate Module name: Module number:
7) Director of Studies/Supervisor/Tutor name	Simon Goodwill, Jon Wheat and Steve Haake
8) Intended duration and timing of project?	One off data collection to be completed within one working day.
9) Location of project If external to SHU, provide evidence in support (see section 17)	Biomechanics lab (A010 Collegiate Hall)
10) State if this study is:	<input checked="" type="checkbox"/> New <input type="checkbox"/> Collaborative (please include appropriate agreements in section 17) <input type="checkbox"/> Replication of :

11) Purpose and benefit of the research

Statement of the research problem with any necessary background information (no more than 1 side of A4)

Current knowledge of tennis player-surface interaction with different tennis court surfaces, particularly during competition, is limited (Miller, 2006). Human centred approaches for the measurement of surface slipperiness highlight that step strategy reflects interactions that occur at the shoe-surface interface (Grönqvist et al. 2001). The measurement of tennis player step strategy would therefore contribute to the understanding of tennis player-surface interactions with different court surfaces.

Computer vision is a field of research that aims to identify and analyse features within video footage to derive metrics, i.e. position data, on an automatic basis. The benefit of this type of approach is the ability to operate without markers, i.e. no participant instrumentation is required (this is an important consideration for match play tennis), and the low time-cost of analysis. Many approaches exist for extracting and tracking human motion features (Wang et al, 2003). Typically, approaches fall into either model-based or non-model based approaches. Model based approaches benefit from the ability to cope well with occlusion and self-occlusion. However, the automated extraction of joint positions for model-based analyses can be difficult due to the wide range of motions exhibited in human motion (Bouchrika and Nixon, 2006).

The location of foot-surface contacts is the principle outcome for this work. Therefore, a non-model based approach has been adopted and an automatic foot-surface contact identification (FSCi) algorithm developed. The FSCi algorithm analyses standard colour video sequences of gait activities, i.e. walking, running etc., and identifies foot-surface contacts automatically based on image processing techniques; no markers or user intervention is required.

The validity of the FSCi algorithm is to be assessed against position data obtained for both feet during three activities, i.e. walking, running and a split-step run and turn. Three-dimensional motion analysis, i.e. industry standard, will be used to validate the FSCi algorithm.

References

- Bouchrika, I. and Nixon, M.S. (2006). People detection and recognition using gait for automated visual surveillance. In: *Proceedings of The Institution of Engineering and Technology Conference on Crime and Security*, London, England, pp. 576 - 581.
- Grönqvist, R., Abeysekera, J., Gard, G., Hsiang, S., Leamon, T.B., Newman, D.J., Gielo-Perczak, K., Lockhart, T.E. and Pai, C, Y.-C. (2001). Human centred approaches in slipperiness measurement. *Ergonomics*, 44(13), 1167 - 1199.
- Miller, S. Modern tennis rackets, balls and surfaces. (2006). *British Journal of Sports Medicine*, 40, 401 - 405.
- Wang, L.A., Hu, W.M. and Tan, T.N. (2003). Recent developments in human motion analysis. *Pattern Recognition*, 36(3), 585 - 601.

12) Participants	
12.1 Number	Six
12.2 Rationale for this number (eg calculations of sample size, practical considerations)	<p>Due to the novel nature of the algorithm being assessed, no previous data exist for the basis of sample size calculations.</p> <p>Based on practical considerations, it is proposed that six participants are recruited for the purpose of validating position data obtained by an automatic foot-surface contact identification algorithm with data obtained by a 3D motion analysis system, i.e. industry standard for motion analysis.</p>
12.3 Criteria for inclusion and exclusion (eg age and sex)	Male 18 - 35
12.4 Procedures for recruitment (eg location and methods)	Local recruitment, i.e. within Collegiate Hall
12.5 Does the study have *minors or †vulnerable adults as participants?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No
12.6 Is CRB Disclosure required for the Principal Investigator? (to be determined by Risk Assessment)	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No If yes, is standard <input type="checkbox"/> or enhanced <input type="checkbox"/> disclosure required?
12.7 If you ticked 'yes' in 12.5 and 'no' in 12.6 please explain why:	

*Minors are participants under the age of 18 years.

†Vulnerable adults are participants over the age of 16 years who are likely to exhibit:

- a) learning difficulties
- b) physical illness/impairment
- c) mental illness/impairment
- d) advanced age
- e) any other condition that might render them vulnerable

13) Details of the research design

13.1 Provide details of intended methodological procedures and data collection.

(For MSc students conducting a scientific support project please provide the following information: a. needs analysis; b. potential outcome; c proposed interventions).

Participants will be asked to perform three repetitions of three activities: walking, running and a split-step run turn. The Principal Investigator will explain and demonstrate these actions if required. Participants will be asked to wear shorts, t-shirt and their own trainers. They will be asked to perform the above activities at a self-selected pace within a predefined motion-capture volume (identified by floor markings). Participants will also be required to perform these activities in barefoot and in trainers (shod). The purpose of shod and barefoot conditions is to compare the algorithm beyond a previous pilot study which was conducted in barefoot only. This is necessary as the most typical application of the FSCi algorithm will be for shod based conditions.

The above procedure will be performed by six participants, all performing three repetitions of the three activities above in shod and barefoot conditions. This is a total of 108 movement trials, i.e. $6 \text{ (participants)} \times 3 \text{ (activities)} \times 3 \text{ (repetitions)} \times 2 \text{ (shod .vs barefoot)} = 108 \text{ trials}$.

The purpose of this study is to validate the FSCi algorithm. The FSCi algorithm will require RGB video footage of these movement trials from different perspectives. This will be obtained from four networked cameras (AXIS M1104, Axis™ Communications, Sweden) streaming images (1280 × 720 p) to a data collection computer at 25 Hz. The FSCi algorithm will be used to analyse these images sequences *post-hoc*, to obtain image coordinates of foot-surface contacts. Global horizontal position data, in reference to the motion-capture volume identified above, will be obtained following a camera calibration process.

Reference position data for both feet will be obtained by three-dimensional online motion analysis (Motion Analysis Corporation, Santa Rosa, CA), i.e. industry standard, sampling at 250 Hz. This will require the application of six (2×3) spherical, retro-reflective markers to each participant's feet. Three-dimensional marker position data will be cropped to stance, i.e. maximum jerk of relevant (minimum vertical position) feet markers, and their horizontal plane position, i.e. centroid of heel, 2nd and 5th metatarsal head markers, recorded. These data will then be compared to the output of the FSCi algorithm and root mean square error computed.

In addition to the comparison of absolute position data between the FSCi algorithm and three-dimensional motion analysis, standard gait parameters, i.e. step length, step width, step frequency and stance time, will also be computed and with root mean square error computed.

13.2 Are these "minor" procedures as defined in Appendix 1 of the ethics guidelines?

☒ Yes ☐ No

13.3 If you answered 'no' in section 13.2, list the procedures that are not minor

13.4 Provide details of the quantitative and qualitative analysis to be used

The following outcome measures will be compared: foot position, step length; step width; step frequency and stance time.

14) Substances to be administered (refer to Appendix VI of the ethics procedures)

14.1 The protocol does not involve the administration of pharmacologically active substances or nutritional supplements.

Please tick box if this statement applies and go to section 15) ☐

14.2 Name and state the risk category for each substance. If a COSHH assessment is required state how the risks are to be managed.

N/A

15) Degree of discomfort that participants might experience

Consider the degree of physical and psychological discomfort that will be experienced by the participants. State the details which must be included in the participant information sheet to ensure that the participants are fully informed about any discomfort that they may experience.

Low. All procedures are sub-maximal activities.

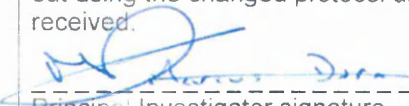
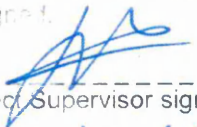

16) Outcomes of Risk Assessment

Provide details of the risk and explain how the control measures will be implemented to

manage the risk.

- Completion of pre-screen medical questionnaire
- Understanding of participant information sheet - verbal clarification given if required
- Completion of informed consent forms
- Strict adherence to test protocol
- First aid and technical assistance will be on-site
- Conduct general laboratory safety procedures, e.g. remove trailing cables, clear obstacles etc.

17) Attachments	Tick box
17.1 Risk assessment (including CRB risk assessment)	✓
17.2 COSHH assessment	
17.3 Participant information sheet (this should be addressed directly to the participant (ie you will etc) and in a language they will understand)	✓
17.4 Informed consent form	✓
17.5 Pre-screening questionnaire	✓
17.6 Collaboration evidence/support correspondence from the organisation consenting to the research (this must be on letterhead paper and signed) See sections 9 & 10.	
17.7 CRB Disclosure certificate <u>or</u> where not available CRB application form	
17.8 Clinical Trails form (FIN 12)	

18. Signature Principal Investigator	<p>Once this application is approved, I will undertake the research study as approved. If circumstances necessitate that changes are made to the approved protocol, I will discuss these with my Project Supervisor. If the supervisor advises that there should be a resubmission to the Sport and Exercise Research Ethics Review Group, I agree that no work will be carried out using the changed protocol until approval has been sought and formally received.</p> <p> Date <u>5/2/13</u></p> <p>Principal Investigator signature</p> <p>Name <u>Mark Dunn</u></p>
19. Approval Project Supervisor to sign either box A or box B as applicable (refer to Appendix I and the flowchart in appendix VI of the ethics guidelines)	<p>Box A:</p> <p>I confirm that the research proposed is based solely on 'minor' procedures, as outlined in Appendix 1 of the HWB Sport and Exercise Research Ethics Review Group 'Ethics Procedures for Research with Humans as Participants' document, and therefore does not need to be submitted to the HWB Sport and Exercise Research Ethics Review Group.</p> <p>In terms of ethics approval, I agree the 'minor' procedures proposed here and confirm that the Principal Investigator may proceed with the study as designed.</p> <p> Date <u>5/2/13</u></p> <p>Project Supervisor signature</p> <p>Name <u>Jon Wheat</u></p>
	<p>Box B:</p> <p>I confirm that the research proposed is <u>not</u> based solely on 'minor' procedures, as outlined in Appendix 1 of the HWB Sport and Exercise Research Ethics Review Group 'Ethics Procedures for Research with Humans as Participants' document, and therefore <u>must</u> be submitted to the HWB Sport and Exercise Research Ethics Review Group for approval.</p> <p>I confirm that the appropriate preparatory work has been undertaken and that the research is in a fit state for submission to the HWB Sport and Exercise Research Ethics Review Group.</p> <p>Project Supervisor signature _____ Date _____</p> <p>Name _____</p>
20. Signature Technician	<p>I confirm that I have seen the full and approved application for ethics approval and technical support will be provided.</p> <p> Date <u>5/2/13</u></p> <p>Technician signature</p> <p>Name <u>S. Scott</u></p>



Sheffield Hallam University

**Faculty of Health and Wellbeing Research Ethics Committee
Sport and Exercise Research Ethics Review Group**

Risk Assessment Pro Forma

****Please ensure that you read the accompanying
Risk Assessment Risk Ranking document before completing this form****

Title of research	Validation of an automatic foot-surface contact identification algorithm
Date Assessed	5-2-13
Assessed by (Principal Investigator)	Marcus Dunn
Signed	Position
	Principal Investigator

Activity	Risks	Control Measures
Sub-maximal walking, running and turning activities in shod and barefoot	1. Muscular injury ($R1 = C1 \times L1$)	LOW: Additional demand on musculoskeletal system is low. Control measure implemented: pre-screening for old or existing injuries plus an opportunity to warm up.
	2. Cardiovascular complications ($R1=C1 \times L1$)	LOW: Additional strain is placed on the cardiovascular system when exercising. Additional load will be light because of short duration. Control measures: Pre-screening questionnaire to assess participants' current level of fitness and health; first aider available in close proximity.
	3. Participants sliding on carpet during turning manoeuvre ($R1=C1 \times L1$)	LOW: Unaccustomed turning on laboratory surface (carpet), particularly in barefoot might cause the participant to slide. Control measures: habituation period prior to the test; reiteration of sub-maximal nature of activities.
	3. Participants tripping equipment, i.e. cables etc., ($R1=C1 \times L1$)	LOW: Participant might trip on cables used to control cameras in experimental setup. Control measures: cables to be tidied away and taped down to floor if necessary.
	4. Participants experiencing skin reaction to adhesive labels used for 3D motion analysis ($R1=C1 \times L1$)	LOW: Participants might experience a skin reaction, i.e. redness, to adhesive labels used to affix spherical markers for 3D motion analysis. Control measure implemented: participant information sheet to make participant aware of risk.

Risk Evaluation (Overall)
Low

General Control Measures

Is a pre-screen medical questionnaire required? Yes [☒] No [☐]

Pre Screening health and fitness questionnaire. Participant Information Sheet provided and agreed. Communication maintained between participant and Principal Investigator at all times.

Emergency Procedures

Emergency First Aid sort from Technical Officers (A016).

Monitoring Procedures

Experiment is monitored at all times by the Principal Investigator.

Review Period

Reviewed By (Supervisor)

Date



6/2/13.

 Sheffield Hallam University	
Faculty of Health and Wellbeing Research Ethics Committee Sport and Exercise Research Ethics Review Group	
Participant Information Sheet	

Project Title	Validation of an automatic foot-surface contact identification algorithm
Supervisor/Director of Studies	Simon Goodwill, Jon Wheat and Steve Haake
Principal Investigator	Marcus Dunn
Principal Investigator telephone/mobile number	07717410501

Purpose of Study and Brief Description of Procedures <i>(Not a legal explanation but a simple statement)</i>
<p>Current knowledge of tennis player-surface interaction with different tennis court surfaces, particularly during competition, is limited. The measurement of tennis player step strategy would contribute to the understanding of tennis player-surface interactions with different court surfaces. Computer vision is a field of research that aims to identify and analyse features within video footage to derive metrics, i.e. position data, on an automatic basis. The benefits of this type of approach are the ability to operate without markers (an important consideration for match play tennis) and the low time-cost of analysis.</p> <p>A non-model based approach has been adopted and an automatic foot-surface contact identification (FSCi) algorithm developed. The FSCi algorithm analyses standard colour video sequences of gait activities, i.e. walking, running etc., and identifies foot-surface contacts automatically based on image processing techniques; no markers or user intervention is required.</p> <p>The purpose of this study is to validate the FSCi algorithm. You will be asked to perform three repetitions of three activities: walking, running and a split-step run turn. The Principal Investigator will explain and demonstrate these actions if required. You will be asked to wear shorts, t-shirt and your own trainers. You will be asked to perform the above activities at a self-selected pace within a predefined motion-capture volume (identified by floor markings). You will also be required to perform these activities in barefoot and in trainers (shod). This will result in a total of 18 movement trials and should take no longer than 1 hour to complete.</p>

The FSCi algorithm will require RGB video footage of these movement trials from different perspectives. This will be obtained from four networked cameras. Reference position data for both feet will be obtained by three-dimensional online motion analysis, i.e. industry standard. This will require the application of seven spherical, retro-reflective markers to your feet (2×3) and sacrum (1). The markers are affixed with adhesive rings which might cause skin irritation. Please notify the Principal Investigator if this is a concern for you. You have the right to withdraw from the study at any time.

It has been made clear to me that, should I feel that these Regulations are being infringed or that my interests are otherwise being ignored, neglected or denied, I should inform Mr David Binney, Chair of the Faculty of Health and Wellbeing Research Ethics Committee (Tel: 0114 225 5679) who will undertake to investigate my complaint.

Appendix 7

A.7.1. Ethics application form: Validation of an automatic foot-surface contact identification algorithm.

CONFIDENTIAL

**Sheffield
Hallam
University**

**Faculty of Health and Wellbeing
Research Ethics Committee**

Sport & Exercise Research Ethics Review Group

APPLICATION FOR ETHICS APPROVAL OF RESEARCH

In designing research involving humans, principal investigators should be able to demonstrate a clear intention of benefit to society and the research should be based on sound principles. These criteria will be considered by the Sport and Exercise Research Ethics Review Group before approving a project. **ALL** of the following details must be provided, either typewritten or word-processed preferably at least in 11 point font.

Please either tick the appropriate box or provide the information required.

1) Date of application	20-3-13
2) Anticipated date of completion of project	
3) Title of research	Validation of an automatic foot-surface contact identification algorithm
4) Subject area	Sports Engineering
5) Principal Investigator	
Name	Marcus Dunn
Email address @ SHU	m.dunn@shu.ac.uk
Telephone/Mobile number	07717410501
Student number (if applicable)	14025464

6) State if this study is: (If the project is undergraduate or postgraduate please state module name and number)	<input checked="" type="checkbox"/> Research <input type="checkbox"/> Undergraduate <input type="checkbox"/> Postgraduate Module name: Module number:
7) Director of Studies/Supervisor/ Tutor name	Simon Goodwill, Jon Wheat and Steve Haake
8) Intended duration and timing of project?	One off data collection to be completed within one working day.
9) Location of project If external to SHU, provide evidence in support (see section 17)	Biomechanics lab (A010 Collegiate Hall)
10) State if this study is:	<input checked="" type="checkbox"/> New <input type="checkbox"/> Collaborative (please include appropriate agreements in section 17) <input type="checkbox"/> Replication of :

11) Purpose and benefit of the research

Statement of the research problem with any necessary background information (no more than 1 side of A4)

Current knowledge of tennis player-surface interaction with different tennis court surfaces, particularly during competition, is limited (Miller, 2006). Human centred approaches for the measurement of surface slipperiness highlight that step strategy reflects interactions that occur at the shoe-surface interface (Grönqvist et al. 2001). The measurement of tennis player step strategy would therefore contribute to the understanding of tennis player-surface interactions with different court surfaces.

Computer vision is a field of research that aims to identify and analyse features within video footage to derive metrics, i.e. position data, on an automatic basis. The benefit of this type of approach is the ability to operate without markers, i.e. no participant instrumentation is required (this is an important consideration for match play tennis), and the low time-cost of analysis. Many approaches exist for extracting and tracking human motion features (Wang et al, 2003). Typically, approaches fall into either model-based or non-model based approaches. Model based approaches benefit from the ability to cope well with occlusion and self-occlusion. However, the automated extraction of joint positions for model-based analyses can be difficult due to the wide range of motions exhibited in human motion (Bouchrika and Nixon, 2006).

The location of foot-surface contacts is the principle outcome for this work. Therefore, a non-model based approach has been adopted and an automatic foot-surface contact identification (FSCi) algorithm developed. The FSCi algorithm analyses standard colour video sequences of gait activities, i.e. walking, running etc., and identifies foot-surface contacts automatically based on image processing techniques; no markers or user intervention is required.

The validity of the FSCi algorithm is to be assessed against position data obtained for both feet during three activities, i.e. walking, running and a split-step run and turn. Three-dimensional motion analysis, i.e. industry standard, will be used to validate the FSCi algorithm.

References

- Bouchrika, I. and Nixon, M.S. (2006). People detection and recognition using gait for automated visual surveillance. In: *Proceedings of The Institution of Engineering and Technology Conference on Crime and Security*, London, England, pp. 576 - 581.
- Grönqvist, R., Abeysekera, J., Gard, G., Hsiang, S., Leamon, T.B., Newman, D.J., Gielo-Perczak, K., Lockhart, T.E. and Pai, C. Y.-C. (2001). Human centred approaches in slipperiness measurement. *Ergonomics*, 44(13), 1167 - 1199.
- Miller, S. Modern tennis rackets, balls and surfaces. (2006). *British Journal of Sports Medicine*, 40, 401 - 405.
- Wang, L.A., Hu, W.M. and Tan, T.N. (2003). Recent developments in human motion analysis. *Pattern Recognition*, 36(3), 585 - 601.

12) Participants	
12.1 Number	Six
12.2 Rationale for this number (eg calculations of sample size, practical considerations)	<p>Due to the novel nature of the algorithm being assessed, no previous data exist for the basis of sample size calculations.</p> <p>Based on practical considerations, it is proposed that six participants are recruited for the purpose of validating position data obtained by an automatic foot-surface contact identification algorithm with data obtained by a 3D motion analysis system, i.e. industry standard for motion analysis.</p>
12.3 Criteria for inclusion and exclusion (eg age and sex)	<p>Male volunteers within the age range of 18 – 35 will be included in this study. Volunteers with current musculoskeletal injuries which could be exacerbated by running (identified by SHU 'Pre-Test Medical Questionnaire') will be excluded.</p>
12.4 Procedures for recruitment (eg location and methods)	Local recruitment, i.e. within Collegiate Hall
12.5 Does the study have *minors or †vulnerable adults as participants?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No
12.6 Is CRB Disclosure required for the Principal Investigator? (to be determined by Risk Assessment)	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No If yes, is standard <input type="checkbox"/> or enhanced <input type="checkbox"/> disclosure required?
12.7 If you ticked 'yes' in 12.5 and 'no' in 12.6 please explain why:	

*Minors are participants under the age of 18 years.

†Vulnerable adults are participants over the age of 16 years who are likely to exhibit:

- a) learning difficulties
- b) physical illness/impairment
- c) mental illness/impairment
- d) advanced age
- e) any other condition that might render them vulnerable

13) Details of the research design

13.1 Provide details of intended methodological procedures and data collection.

(For MSc students conducting a scientific support project please provide the following information: a. needs analysis; b. potential outcome; c proposed interventions).

Participants will be asked to perform three repetitions of three activities: walking, running and a split-step run turn. The Principal Investigator will explain and demonstrate these actions if required. Participants will be asked to wear shorts, t-shirt and their own trainers. They will be asked to perform the above activities at a self-selected pace within a predefined motion-capture volume (identified by floor markings). Participants will also be required to perform these activities in barefoot and in trainers (shod). The purpose of shod and barefoot conditions is to compare the algorithm beyond a previous pilot study which was conducted in barefoot only. This is necessary as the most typical application of the FSCi algorithm will be for shod based conditions.

The above procedure will be performed by six participants, all performing three repetitions of the three activities above in shod and barefoot conditions. This is a total of 108 movement trials, i.e. $6 \text{ (participants)} \times 3 \text{ (activities)} \times 3 \text{ (repetitions)} \times 2 \text{ (shod vs barefoot)} = 108 \text{ trials}$.

The purpose of this study is to validate the FSCi algorithm. The FSCi algorithm will require RGB video footage of these movement trials from different perspectives. This will be obtained from four networked cameras (AXIS M1104, Axis™ Communications, Sweden) streaming images (1280 × 720 p) to a data collection computer at 25 Hz. The FSCi algorithm will be used to analyse these images sequences *post-hoc*, to obtain image coordinates of foot-surface contacts. Global horizontal position data, in reference to the motion-capture volume identified above, will be obtained following a camera calibration process.

Reference position data for both feet will be obtained by three-dimensional online motion analysis (Motion Analysis Corporation, Santa Rosa, CA), i.e. industry standard. This will require the application of seven spherical, retro-reflective markers to each participant; six (2×3) will be applied to the feet and one to the sacrum (to characterise movement speed). Three-dimensional foot marker position data during stance will be identified, i.e. existing gait event identification algorithms (O'Connor et al., 2007), and their horizontal plane position, i.e. centroid of heel, 2nd and 5th metatarsal head markers, recorded. These data will then be compared to the output of the FSCi algorithm and root mean square error computed.

In addition to the comparison of absolute position data between the FSCi algorithm and three-dimensional motion analysis, standard gait parameters, i.e. step length, step width, step frequency and stance time, will also be computed and with root mean square error computed.

References

O'Connor, C.M., Thorpe, S.K., O'Malley, M.J. and Vaughan, C.L. (2007). Automatic detection of gait events using kinematic data. *Gait & Posture*, 25(3), 469-474.

13.2 Are these "minor" procedures as defined in Appendix 1 of the ethics

guidelines?
<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
13.3 If you answered 'no' in section 13.2, list the procedures that are <u>not</u> minor
13.4 Provide details of the quantitative and qualitative analysis to be used
The following outcome measures will be compared: foot position, step length; step width; step frequency and stance time.

14) Substances to be administered (refer to Appendix VI of the ethics procedures)

14.1 The protocol does not involve the administration of pharmacologically active substances or nutritional supplements.

Please tick box if this statement applies and go to section 15) []

14.2 Name and state the risk category for each substance. If a COSHH assessment is required state how the risks are to be managed.

N/A

15) Degree of discomfort that participants might experience

Consider the degree of physical and psychological discomfort that will be experienced by the participants. State the details which must be included in the participant information sheet to ensure that the participants are fully informed about any discomfort that they may experience.

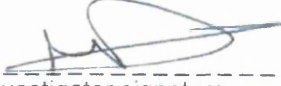
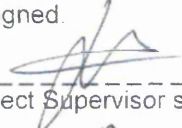

Low. All procedures are sub-maximal activities.


16) Outcomes of Risk Assessment

Provide details of the risk and explain how the control measures will be implemented to manage the risk.

- Completion of pre-screen medical questionnaire
- Understanding of participant information sheet - verbal clarification given if required
- Completion of informed consent forms
- Strict adherence to test protocol
- First aid and technical assistance will be on-site
- Conduct general laboratory safety procedures, e.g. remove trailing cables, clear obstacles etc.

17) Attachments	Tick box
17.1 Risk assessment (including CRB risk assessment)	✓
17.2 COSHH assessment	
17.3 Participant information sheet (this should be addressed directly to the participant (ie you will etc) and in a language they will understand)	✓
17.4 Informed consent form	✓
17.5 Pre-screening questionnaire	✓
17.6 Collaboration evidence/support correspondence from the organisation consenting to the research (this must be on letterhead paper and signed) See sections 9 & 10.	
17.7 CRB Disclosure certificate <u>or</u> where not available CRB application form	
17.8 Clinical Trails form (FIN 12)	

18. Signature Principal Investigator	<p>Once this application is approved, I will undertake the research study as approved. If circumstances necessitate that changes are made to the approved protocol, I will discuss these with my Project Supervisor. If the supervisor advises that there should be a resubmission to the Sport and Exercise Research Ethics Review Group, I agree that no work will be carried out using the changed protocol until approval has been sought and formally received.</p> <p style="text-align: right;"> Date <u>20/7/13</u></p> <p>Principal Investigator signature _____</p> <p>Name <u>MARCUS DUNIN</u> _____</p>
19. Approval Project Supervisor to sign <u>either</u> box A <u>or</u> box B as applicable (refer to Appendix I and the flowchart in appendix VI of the ethics guidelines)	<p>Box A: I confirm that the research proposed is based solely on 'minor' procedures, as outlined in Appendix 1 of the HWB Sport and Exercise Research Ethics Review Group 'Ethics Procedures for Research with Humans as Participants' document, and therefore does not need to be submitted to the HWB Sport and Exercise Research Ethics Review Group.</p> <p>In terms of ethics approval, I agree the 'minor' procedures proposed here and confirm that the Principal Investigator may proceed with the study as designed.</p> <p style="text-align: right;"> Date <u>20/3/13</u></p> <p>Project Supervisor signature _____</p> <p>Name <u>Dr. Jon HITCHEN</u> _____</p>
	<p>Box B: I confirm that the research proposed is <u>not</u> based solely on 'minor' procedures, as outlined in Appendix 1 of the HWB Sport and Exercise Research Ethics Review Group 'Ethics Procedures for Research with Humans as Participants' document, and therefore <u>must</u> be submitted to the HWB Sport and Exercise Research Ethics Review Group for approval.</p> <p>I confirm that the appropriate preparatory work has been undertaken and that this document is in a fit state for submission to the HWB Sport and Exercise Research Ethics Review Group.</p> <p style="text-align: right;">Date _____</p> <p>Project Supervisor signature _____</p> <p>Name _____</p>
20. Signature Technician	<p>I confirm that I have seen the full and approved application for ethics approval and technical support will be provided.</p> <p style="text-align: right;"> Date <u>20/3/13</u></p> <p>Technician signature _____</p> <p>Name <u>S. SCOTT</u> _____</p>

	Sheffield Hallam University
Faculty of Health and Wellbeing Research Ethics Committee Sport and Exercise Research Ethics Review Group	
Risk Assessment Pro Forma	

****Please ensure that you read the accompanying
Risk Assessment Risk Ranking document before completing this form****

Title of research	Validation of an automatic foot-surface contact identification algorithm
Date Assessed	20-3-13
Assessed by (Principal Investigator)	Marcus Dunn
Signed	Position
	Principal Investigator

Activity	Risks	Control Measures
Sub-maximal walking, running and turning activities in shod and barefoot	1. Muscular injury ($R1 = C1 \times L1$)	LOW: Additional demand on musculoskeletal system is low. Control measure implemented: pre-screening for old or existing injuries plus an opportunity to warm up.
	2. Cardiovascular complications ($R1=C1 \times L1$)	LOW: Additional strain is placed on the cardiovascular system when exercising. Additional load will be light because of short duration. Control measures: Pre-screening questionnaire to assess participants' current level of fitness and health; first aider available in close proximity.
	3. Participants sliding on carpet during turning manoeuvre ($R1=C1 \times L1$)	LOW: Unaccustomed turning on laboratory surface (carpet), particularly in barefoot might cause the participant to slide. Control measures: habituation period prior to the test; reiteration of sub-maximal nature of activities.
	3. Participants tripping equipment, i.e. cables etc., ($R1=C1 \times L1$)	LOW: Participant might trip on cables used to control cameras in experimental setup. Control measures: cables to be tidied away and taped down to floor if necessary.
	4. Participants experiencing skin reaction to adhesive labels used for 3D motion analysis ($R1=C1 \times L1$)	LOW: Participants might experience a skin reaction, i.e. redness, to adhesive labels used to affix spherical markers for 3D motion analysis. Control measure implemented: participant information sheet to make participant aware of risk.
Risk Evaluation (Overall)		
Low		

General Control Measures

Is a pre-screen medical questionnaire required? Yes [☒] No [☐]


Pre Screening health and fitness questionnaire. Participant Information Sheet provided and agreed. Communication maintained between participant and Principal Investigator at all times.

Emergency Procedures

Emergency First Aid sort from Technical Officers (A016).


Monitoring Procedures

Experiment is monitored at all times by the Principal Investigator.

Review Period**Reviewed By (Supervisor)****Date**

Dr Jon WHEAT

20/3/13.

 Sheffield Hallam University	
Faculty of Health and Wellbeing Research Ethics Committee Sport and Exercise Research Ethics Review Group	
Participant Information Sheet	

Project Title	Validation of an automatic foot-surface contact identification algorithm
Supervisor/Director of Studies	Simon Goodwill, Jon Wheat and Steve Haake
Principal Investigator	Marcus Dunn
Principal Investigator telephone/mobile number	07717410501

Purpose of Study and Brief Description of Procedures <i>(Not a legal explanation but a simple statement)</i>
<p>Current knowledge of tennis player-surface interaction with different tennis court surfaces, particularly during competition, is limited. The measurement of tennis player step strategy would contribute to the understanding of tennis player-surface interactions with different court surfaces. Computer vision is a field of research that aims to identify and analyse features within video footage to derive metrics, i.e. position data, on an automatic basis. The benefits of this type of approach are the ability to operate without markers (an important consideration for match play tennis) and the low time-cost of analysis.</p> <p>A non-model based approach has been adopted and an automatic foot-surface contact identification (FSCi) algorithm developed. The FSCi algorithm analyses standard colour video sequences of gait activities, i.e. walking, running etc., and identifies foot-surface contacts automatically based on image processing techniques; no markers or user intervention is required.</p> <p>The purpose of this study is to validate the FSCi algorithm. You will be asked to perform three repetitions of three activities: walking, running and a split-step run turn. The Principal Investigator will explain and demonstrate these actions if required. You will be asked to wear shorts, t-shirt and your own trainers. You will be asked to perform the above activities at a self-selected pace within a predefined motion-capture volume (identified by floor markings). You will also be required to perform these activities in barefoot and in trainers (shod). This will result in a total of 18 movement trials and should take no longer than 1 hour to complete.</p>

The FSCi algorithm will require RGB video footage of these movement trials from different perspectives. This will be obtained from four networked cameras. Reference position data for both feet will be obtained by three-dimensional online motion analysis, i.e. industry standard. This will require the application of seven spherical, retro-reflective markers to your feet (2 × 3) and sacrum (1). The markers are affixed with adhesive rings which might cause skin irritation. Please notify the Principal Investigator if this is a concern for you. You have the right to withdraw from the study at any time.

It has been made clear to me that, should I feel that these Regulations are being infringed or that my interests are otherwise being ignored, neglected or denied, I should inform Mr David Binney, Chair of the Faculty of Health and Wellbeing Research Ethics Committee (Tel: 0114 225 5679) who will undertake to investigate my complaint.

A.7.4. Foot-surface contact identification algorithm profile: most expensive processes.

FSCI (1 call, 153.359 sec)

Generated 28-Jul-2014 19:03:41 using cpu time.

script in file [/Users/Marcus/Dropbox/ITE PhD/Write-up/Corrected - final/FSCI.m](#)

[Copy to new window for comparing multiple runs](#)

This function changed during profiling or before generation of this report. Results may be incomplete or inaccurate.







Refresh

- ☒ Show parent functions
 ☒ Show busy lines
 ☒ Show child functions
☒ Show Code Analyzer results
 ☒ Show file coverage
 ☒ Show function listing













Parents (calling functions)

No parent

Lines where the most time was spent

Line Number	Code	Calls	Total Time	% Time	Time Plot
339	<code>Im_bwBS_ConvexSkel = bwmorph(I...</code>	127	22.554 s	14.7%	
247	<code>ImHSV_limb = rgb2hsv(Im);</code>	127	22.020 s	14.4%	
401	<code>[Mask,MX,MY] = RefineFootMask(...</code>	255	7.824 s	5.1%	
420	<code>MaskedFSC_im = imdilate(Masked...</code>	255	5.492 s	3.6%	
217	<code>Im_bwConvexProps = regionprops...</code>	127	4.506 s	2.9%	
All other lines			90.962 s	59.3%	
Totals			153.359 s	100%	

Children (called functions)

Function Name	Function Type	Calls	Total Time	% Time	Time Plot
regionprops	function	1908	23.073 s	15.0%	
bwmorph	function	254	22.705 s	14.8%	
rgb2hsv	function	128	22.193 s	14.5%	
std2	function	1016	12.895 s	8.4%	
imread	function	255	7.659 s	5.0%	
RefineFootMask	function	255	7.651 s	5.0%	
im2double	function	382	6.620 s	4.3%	
imdilate	function	763	6.613 s	4.3%	
immultiply	function	1018	6.598 s	4.3%	
im2bw	function	1016	4.040 s	2.6%	
imabsdiff	function	381	3.762 s	2.5%	
drawCircle	function	255	3.054 s	2.0%	
bwconncomp	function	510	1.685 s	1.1%	