A comparative analysis of binary patterns with discrete cosine transform for gender classification

RODRIGUES, Marcos <http://orcid.org/0000-0002-6083-1303>, KORMANN, Mariza and TOMEK, Peter

A Comparative Analysis of Binary Patterns with Discrete Cosine Transform for Gender Classification

Marcos A Rodrigues, Mariza Kormann and Peter Tomek

***Abstract*—This paper presents a comparative analysis of binary patters for gender classification with a novel method of feature transformation for improved accuracy rates. The main requirements of our application are speed and accuracy. We investigate a combination of local binary patterns (LBP), Census Transform (CT) and Modified Census Transform (MCT) applied over the full, top and bottom halves of the face. Gender classification is performed using support vector machines (SVM). A main focus of the investigation is to determine whether or not a 1D discrete cosine transform (DCT) applied directly to the grey level histograms would improve accuracy. We used a public database of faces and run face and eye detection algorithms allowing automatic cropping and normalisation of the images. A set of 120 tests over the entire database demonstrate that the proposed 1D discrete cosine transform improves accuracy in all test cases with small standard deviations. It is shown that using basic versions of the algorithms, LBP is marginally superior to both CT and MCT and agrees with results in the literature for higher accuracy on male subjects. However, a significant result of our investigation is that, by applying a 1D-DCT this bias is removed and an equivalent error rate is achieved for both genders. Furthermore, it is demonstrated that DCT improves overall accuracy and renders CT a superior performance compared to LBP in all cases considered.**

***Index Terms*—Image processing, feature extraction, gray-scale, image texture analysis, pattern recognition, discrete cosine transforms, support vector machines**

Marcos A Rodrigues and Mariza Kormann are with the GMPR–Geometric Modelling and Pattern Recognition Research Group at Sheffield Hallam University, Sheffield, UK. Email {*m.rodrigues, m.kormann*}@shu.ac.uk

Peter Tomek is with ATEKNEA, Budapest, Hungary. Email peter.tomek@ateknea.com

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 315525.

I. Introduction

REAL-time gender classification is a requirement for marketing applications where legal and ethical constraints do not allow the saving of images either locally or remotely for later processing. The ADMOS project [1] is funded by the European Union and aims to develop a real-time gender classification and age estimation to be used in private spaces of public use, such as shopping malls, fairs and outdoor events. The main computing operations on an image within the time frame of live capture include face detection, gender classification and age estimation. We are investigating a number of methods that have the potential to be fast, accurate and robust. In this paper we report on a combination of techniques involving LBP–Local Binary Patterns, CT–Census Transform, MCT–Modified Census Transform, DCT–Discrete Cosine Transform and SVM–Support Vector Machines. It is shown that DCT can remove LBP's bias towards higher accuracy for male subjects and that it renders CT a superior technique when compared to LBP.

LBP is a non-parametric method used to summarise local structures of an image and have been extensively exploited in face analysis for gender, age, and face recognition [2], [3], [4], [5], [6], [7]. Normally, LBP are employed in local and holistic approaches and a number of extensions have been demonstrated in the literature (e.g. [4]) in connection with linear discriminant analysis and support vector machines. The Census Transform is similar to LBP; the main difference lies on how bits are concatenated together. Although this is a seemingly small difference, it has significant bearings on the final grey level scale histograms and thus, on the texture descriptors in various regions of an image.

The Census Transform has not been extensively exploited in face analysis as LBP; some previous work include [8], [9]. LBP and DCT have been used together in connection with face recognition and gender classification (e.g. [10], [11], [12]). However, it is important to note that when DCT is used, it is invariably in connection with a 2D-DCT. Normally a DCT is performed over the entire input image using different block sizes. In a similar fashion, LBP is normally applied over regions and over the entire image and such histograms are concatenated into a combined one.

Here we explore these techniques aiming at fast processing for real time applications. We only use a single pass, non-optimised LBP, CT or MCT over the input image, followed by a 1D-DCT applied to the resulting histograms. The purpose is to investigate whether or not the 1D-DCT would improve gender classification. The approach is demonstrated by using a public database from which the various regions of interest are automatically selected by face and eye detection algorithms. It is shown that 1D-DCT improves gender classification in all cases considered.

The method is described in Section II, experimental results are presented in Section III with conclusions and further work in Section IV.

II. Method

The approach to gender classification adopted in this paper has been described in our previous paper [13] summarised as follows:

- Define a set of measurements or features $m_i (i = 1, 2, \ldots, N_1)$ over an input image and build a vector $M_j = (m_1, m_2, \ldots, m_{N_1})^T$, with $j = 1, 2, \ldots, N_2$ characterising the selected features;
- Build a matrix Ω_k of vectors M_j where the index of k points to the identity of the input vector: $\Omega_k = (M_1, M_2, \ldots, M_s)^T$ where s is the total number of vectors for class k;
- Define a method to estimate the closest distance to a given vector M to the most similar vector in the database. The class of closest vector(s) will point to the most likely class of M.

Arguably the most critical step is feature selection. In [13] we used LBP in connection with eigenvector decomposition to determine class membership. Only raw data were used with no feature optimisation. The purpose of that study was to determine which ROI–region of interest would be more appropriate for robust gender classification. Tests were reported on using a larger image comprising the head with portions of the neck; this invariably also included large chunks of background and, as expected, did not yield robust results. Further experiments on cropped regions of the face used the full, top and bottom halves of the face and various combinations of these. It was shown that, for the dataset used, the best region was the top half of the face. Gender classification accuracy of 88% was reported on non-optimised data and non-optimised classification method. It was pointed out that other feature selection techniques could be used and that a literature review pointed to SVM as a robust technique in connection with LBP.

In this paper we propose a new method for gender classification with a comparative analysis of performance. The method uses the LBP, Census Transform and Modified Census Transform for feature extraction, discrete cosine transform for feature transformation and support vector machines for classification. In particular we are interested in determining whether or not a similar technique to LBP, the Census Transform and its modified version would yield better, worst or indifferent results. Furthermore, whether or not the discrete cosine transform can be effectively used for gender classification in connection with such feature extraction techniques. The steps in the proposed method are depicted in Figure 1 described as follows:

1) Apply LBP on input images and build training and test sets for both female and male subjects. The size of the kernel window is 3×3;
2) Apply the Census Transform to the same images and build training and test sets for both classes, with kernel window size of 3×3;
3) Apply a Modified Census Transform to the

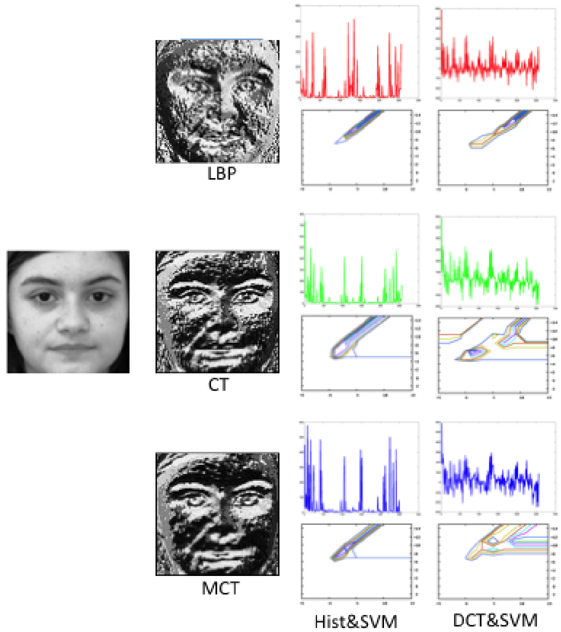

Fig. 1. The proposed method: histograms from LBP, CT and MCT are transformed by DCT. Both original histograms and their DCT are compared through SVM.

same images building training and test sets for both classes, with kernel window size 3×3;

4) Apply the discrete cosine transform to the outputs from steps 1–3;
5) Use SVM for training and testing all data from steps 1–4.

The proportion of data that is used for training and testing can vary considerably; in this paper we use 70–30 (70% of all data for training and 30% for testing).

A. The Census Transform – CT

The Census Transform has been proposed in [14] as a grey-level operator over a local neighbourhood. It applies to an image kernel of size $m \times n$:

$$CT_{m,n}(x,y) = ||_{i=-n/2}^{n/2} ||_{j=-m/2}^{m/2} f(I(x,y), I(x+i,y+j)) \quad (1)$$

where the operator $||$ is a bit-wise concatenation of $f(u,v)$ which is defined as

$$f(u,v) = \begin{cases} 0 & \text{if } u \leq v, \\ 1 & \text{otherwise.} \end{cases}$$

Various modifications have been suggested to the original CT transform such as centre-symmetric weighted kernel [15] and a modified CT using the mean of the centre pixel [16]. Typical window sizes are 3×3, 5×5 and 9×7 as their concatenated binary results fit into 8, 32 or 64 bit. Experiments have shown (e.g. [17]) that using a kernel window of 5×5 is a good compromise between speed and accuracy.

The Modified Census Transform (MCT) as used in this paper is similarly defined as in equation 1. The difference is that instead of using the grey level intensity of the centre pixel, the average of the kernel window intensity is used.

B. Local Binary Patterns – LBP

Local binary patterns [2], [18], [19], [20] are grey-scale operators defined over local neighbourhood pixels. It was originally defined using a 3×3 array of pixels, but many implementations consider larger radii. The value of the centre pixel is compared with its neighbours and the result (greater or smaller) expressed as a binary number and concatenated over all pixels considered. The concatenated array of binary numbers is normally converted to grey scale from which histograms are produced. LBP can be expressed over P sampling points on a circle of radius R where the value of the centre pixel (x,y) is expressed as:

$$\text{LBP}_{P,R} = \sum_{p=0}^{P-1} T(I_p - I_c) 2^P, \quad (2)$$

where $I_p - I_c$ is the difference of pixel intensity in the grey level between a current pixel and centre pixel of the kernel window. P is the number of pixels on a circle of radius R, and T is a thresholding function defined as:

$$T(.) = \begin{cases} 1 & \text{if } (I_p - I_c) \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

In order to improve the discriminating power of LBPs, images are normally defined in blocks from which individual LBPs are calculated and then concatenated into a single histogram. The analysis of such histograms can be used to differentiate texture patterns. A number of variants to LBP have been proposed in the literature (e.g. [3], [4], [5]). In the experiments reported in this paper we only consider the original LBP definition.

C. The Discrete Cosine Transform – DCT

The DCT transform and its variants have been used in a number of contexts most notably in image and video compression (e.g. [21], [22], [23]). DCT is a close relative to the discrete Fourier transform as it defines a sequence of data in terms of the sum of the cosine functions at different frequencies. There are many versions of the DCT and here we use the unitary discrete cosine transform as defined in Matlab [24]. The DCT transform of a one-dimensional signal z (in our case z is an image histogram) is expressed as:

$$y(k) = w(k) \sum_{n=1}^{N} z(n) \cos\left(\frac{\pi(2n-1)(k-1)}{2N}\right) \tag{3}$$

for $k = 1, 2, \ldots N$ where N is the length of the signal and

$$w(k) = \begin{cases} 1/\sqrt{N} & \text{for } k = 1, \\ \sqrt{2/N} & \text{for } 2 \leq k \leq N. \end{cases} \tag{4}$$

The length of the coefficients y is the same as the original signal z. A useful property of the DCT is that normally it is only necessary a few coefficients to reconstruct the signal; most signals can be described with over 99% accuracy by using only a handful of coefficients. Here we choose to use all coefficients for improved accuracy.

D. Support Vector Machines – SVM

In pattern recognition tasks, algorithms for linear discriminant functions can be used either over the raw or original data features or in a transformed space that can be defined by nonlinear transformations of the original variables (e.g. DCT applied to the LBP and CT histograms as proposed in this paper). Support vector machines are algorithms that implement a mapping of pattern vectors to a higher dimensional feature space and find a 'best' separating hyperplane between the data set. The best hyperplane is commonly referred to as the maximal margin hyperplane, as it is defined where the closest points between classes are at maximum distance [25].

Given a set of M training samples $(l_i, \mathbf{x}_i)$ where l_i is the associated class label ($l_i \in \{-1, 1\}$) of vector $\mathbf{x}_i$ where $\mathbf{x}_i \in R^N$, a SVM classifier finds the optimal hyperplane that maximises the margin between classes l_i:

$$f(x) = \sum_{i=1}^{M} l_i \alpha_i \cdot k(\mathbf{x}, \mathbf{x}_i) + b \tag{5}$$

where $k(\mathbf{x}, \mathbf{x}_i)$ is a kernel function, b is a bias and the sign of $f(x)$ is used to determine the class membership of vector $\mathbf{x}$. For a two-class problem (e.g. the case of gender classification) a linear SVM might suffice. In this case, the kernel function is a dot product in the input space.

III. EXPERIMENTAL RESULTS

A public database as described in [13] is used here, details of the database can be found in [26]. It contains 2,779 images of even balanced number of male and female subjects captured with large variations in pose and illumination. In [13] a set of 50 male and 50 female subjects were used; although all subjects were not in strictly frontal pose, there was not much variation in illumination in that dataset. Here we expand the dataset by including images with uneven illumination. We used the entire set of 99 male and 99 female subjects (one subject was removed from each original set of 100 as they appear to be repeated). The selected data allows the testing of algorithms in a realistic scenario of uneven illumination. Examples of selected images from the database are depicted in Figure 2.

First, we performed a visual comparison of LBP, CT and MCT using a 3×3 kernel window. We proceeded to perform a CT and MCT using a 5×5 window as shown in Figure 3. It is not clear whether or not simply increasing the kernel size would impact on gender classification performance

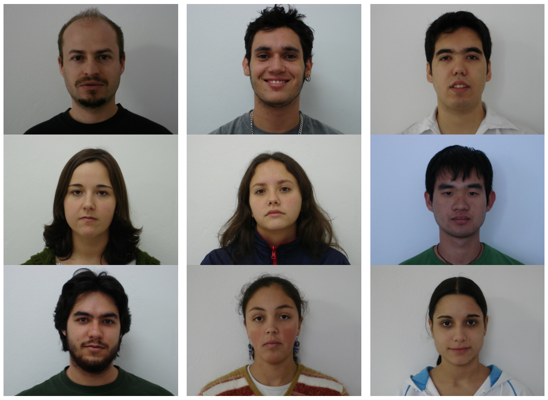

Fig. 2. Examples of image data from the FEI database.

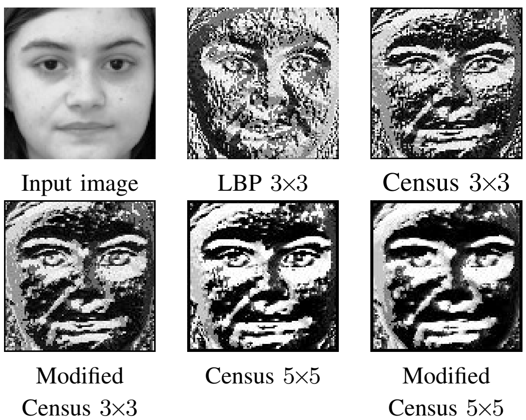

Fig. 3. LBP, Census and Modified Census over an input image.

and this is left for further studies. Here we report on LBP, CT and MCT with constant 3×3 kernel window.

Following results described in [13] we use the face regions labeled as FULL face, TOP and BOTTOM half of the face. Histograms for LBP, CT and MCT were evaluated as described in Section II. From these we also built concatenated histograms for FULL||TOP, FULL||BOTTOM, and TOP||BOTTOM. Furthermore, we also built concatenated histograms of LBP||CT and LBP||MCT for all cases. The histograms were then separated into training and test sets and subject to the SVM discrimination method.

TABLE I
ACCURACY OF HISTOGRAM-BASED CLASSIFICATION (%)

Face ROI & Method	**LBP**	**CT**	**MCT**	**LBP|| CT**	**LBP|| MCT**
FULL					
Female	80.6	87.1	67.7	83.9	77.4
Male	90.3	93.5	83.9	90.3	83.9
TOP					
Female	83.9	87.1	74.7	83.9	80.6
Male	96.8	87.1	77.4	90.3	90.3
BOTTOM					
Female	80.6	74.2	67.7	71.0	74.2
Male	93.5	93.5	93.5	96.8	93.5
FULL||TOP					
Female	83.4	87.1	77.4	77.4	87.1
Male	93.5	93.5	87.1	96.8	87.1
FULL||BOTTOM					
Female	83.9	87.1	64.5	71.0	77.4
Male	90.3	93.5	90.3	96.8	96.8
TOP||BOTTOM					
Female	80.6	67.7	74.2	83.9	80.6
Male	96.8	100.0	90.3	93.5	93.5
Mean	87.9	87.6	79.0	86.3	85.2
STD	6.4	8.9	9.9	9.4	7.3
Mean Female	82.2	81.7	71.0	78.5	79.6
Mean Male	93.5	93.5	87.1	94.1	90.9
Abs difference	11.4	11.8	16.1	15.6	11.3

Results for histogram-based classification are tabulated in Table I. The summary refers to training on 60 data sets (30 female, 30 male) and 60 test sets. The best result is for LBP applied to the TOP half of the face, and this confirms previous results as reported in [13]. Overall, the LBP technique is shown to be superior to CT and MCT and the concatenated combinations with the overall lowest standard deviation of 6.4. Furthermore, it is observed that there is a bias towards more accurate male classification as reported in all papers in the literature; this is the case for all combinations used. The reasons for this behaviour are not yet clear.

Following this initial comparison, all histograms were subject to discrete cosine transform, trained and tested with the SVM method as per previous sets. Results are tabulated in Table II. Two important observations can be made: the overall accuracy has improved for all sets and the bias towards higher male accuracy has been removed. Furthermore, the CT is now shown to be a superior technique with the

TABLE II
ACCURACY OF DCT-BASED CLASSIFICATION (%)

Face ROI & Method	**LBP**	**CT**	**MCT**	**LBP|| CT**	**LBP|| MCT**
FULL					
Female	93.5	93.5	80.6	93.5	93.5
Male	87.1	90.3	83.9	87.1	87.1
TOP					
Female	90.3	90.3	71.0	93.5	80.6
Male	83.9	90.3	77.4	90.3	83.9
BOTTOM					
Female	80.6	80.6	71.0	80.6	74.2
Male	77.4	87.1	77.4	77.4	80.6
FULL||TOP					
Female	93.5	93.5	93.5	90.3	93.5
Male	87.1	90.3	80.6	87.1	87.1
FULL||BOTTOM					
Female	90.3	93.5	87.1	90.3	90.3
Male	83.9	87.1	74.1	83.9	83.9
TOP||BOTTOM					
Female	87.1	87.1	93.5	87.1	90.3
Male	87.1	93.5	83.9	87.1	83.9
Mean	86.8	89.8	81.2	87.4	85.7
STD	4.8	3.8	7.6	4.8	5.7
Mean Female	89.2	89.9	82.8	89.2	87.1
Mean Male	84.4	89.8	79.6	85.5	84.4
Abs difference	4.8	0.0	3.2	3.7	2.6

lowest standard deviation of 3.8. With the removal of bias, both female and male classification are equivalent, with absolute difference between their means (“Abs difference”) ranging from 0.0–4.8. This favourably compares to 11.3–16.1 of previous set of experiments. These results demonstrate that the discrete cosine transform can effectively be applied over the grey level histograms for improved gender classification.

IV. CONCLUSIONS

This paper has presented a comparative analysis of gender classification based on LBP, CT and MCT in connection with the discrete cosine transform and support vector machines. The new proposed method is based on evaluating binary patterns and building histograms of various regions of the face including the full face, top and bottom halves, and a combination of these. Histograms are then subject to the discrete cosine transform. Discriminant analysis is performed through support vector machines both on the original and DCT-transformed histograms.

If only histograms are used (including various concatenations as reported here) it is shown that LBP is a more accurate technique than either CT and MCT (LBP is only marginally more accurate than CT, but it has a much lower overall standard deviation making it a more robust technique). It is also observed that the best classification results are obtained over the top half of the face, a region that includes the front and the eyes. This confirms previous results from our research reported in [13]. Furthermore, it is noted that there is a bias towards more accurate classification over male subjects, as reported in the literature. Some explanations for why this is so has been attempted in the literature but it is pointed out (e.g. [15]) that a thorough analysis is required to explain this behaviour.

Experimental results for DCT-transformed histograms demonstrate that the CT method is the more accurate technique. A number of observations can be made: overall accuracy is improved for all cases, the standard deviation is substantially decreased for all cases, and the bias towards higher accuracy for male subjects is removed. We do not yet offer an explanation for this as a detailed mathematical analysis is required, which is left for further studies.

The results reported in this paper clearly show that the discrete cosine transform yields more appropriate data for accurate classification. Further work includes applying the transformations on histograms of concatenated data from sub-regions of the face – the principle that has been shown in the literature is that it seems that the more features are used the better the accuracy. Obviously that there is a limit to the number of features and this is an area for further investigation. Moreover, we intend to apply the techniques described here on other public databases namely FERET and SUMS. It appears that a larger number of different algorithms have been applied to these databases as reported in the literature and this will provide a more direct comparison of performance.

REFERENCES

[1] ADMOS (2013). *Advertising Monitoring System Development for Outdoor Media Analytics*, EC Grant Agreement

31552. [Online] Available at http://admos.eu

[2] M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen (2011). *Computer Vision Using Local Binary Patterns*. Springer.

[3] J, Ylioinas, A. Hadid, M. Pietikäinen (2011). Combining contrast and local binary patterns for gender classification. *SCIA 17th Scandinavian Conference on Image Analysis*, 2011.

[4] C. Shan (2012). Learning local binary patterns for gender classification on real-world face images, *Pattern Recognition Letters* 33 (2012) 431–437.

[5] Y. Guo, G. Zhao, M. Pietikäinen, and Z. Xu (2010). Descriptor learning based on fisher separation criterion for texture classification. *In Proc. ACCV2010*, 185–198, 2010.

[6] H. Lian, B. Lu (2007). Multi-view gender classification using multi-resolution local binary patterns and support vector machines. *Int J Neural Systems* 17 (6), 479–487.

[7] N. Sun, W. Zheng, C. Sun, C. Zou, L. Zhao (2006). Gender classification based on boosting local binary pattern. *In: Int Symp on Neural Networks*, 2006.

[8] R. Verschae, J. Ruiz-del-Solar, M. Correa (2006). Gender Classification of Faces Using Adaboost, *11th Iberoamerican Congress in Pattern Recognition*, CIARP 2006 Cancun, Mexico, November 2006, 68–78.

[9] B. Jun, T. Kim and D. Kim (2011). A compact local binary pattern using maximization of mutual information for face analysis, *Pattern Recognition* Volume 44, Issue 3, 2011, 532–543.

[10] S.A. Khan, M. Ahmad, M. Nazir and N. Riaz (2014). A Comparative Analysis of Gender Classification Techniques, *Middle-East Journal of Scientific Research* 20(1):1–13.

[11] H.F. Alrashed and M.A. Berbar (2013). Facial Gender Recognition Using Eyes Images, *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 2, Issue 6, June 2013.

[12] A.M. Mirza, M. Hussain, H. Almuzaini, G. Muhammad, H. Aboalsamh and George Bebis (2013). Gender Recognition Using Fusion of Local and Global Facial Features. *In G. Bebis et al. (Eds.): ISVC 2013, Part II, LNCS 8034*, pp. 493–502, 2013.

[13] M. Rodrigues, M. Kormann and P. Tomek (2014). ROI Sensitivity Analysis for Real Time Gender Classification, submitted to *CSCC-2014 The 18th Int Conf of Circuits, Systems, Communications and Computers*, July 17–21, 2014, Greece.

[14] R. Zabih, J. Woodfill (1994). Non-parametric local transforms for computing visual correspondence. *In: ECCV. (1994)*, Secaucus, NJ, USA, Springer-Verlag New York, Inc. 151–158.

[15] R. Spangenberg, T. Langner, R. Rojas (2013). Weighted semi-global matching and center-symmetric census transform for robust driver assistance. *In Wilson, R., Hancock, E., Bors, A., Smith, W., eds.: Computer Analysis of Images and Patterns.* Volume 8048 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2013) 34–41.

[16] B. Froba and A. Ernst (2004). Face detection with the modified census transform, *Proc. of IEEE 6th Int Conf on Automatic Face and Gesture Recognition*, 91–96.

[17] S.K. Gehrig, C. Rabe (2010). Real-Time Semi-Global Matching on the CPU. *In: CVPR Workshops*, San Francisco, CA, USA (June 2010) 85–92.

[18] T. Ojala, M. Pietikäinen, and D. Harwood (1994). Performance evaluation of texture measures with classification based on Kullback discrimination of distributions, *Proceedings of the 12th IAPR International Conference on Pattern Recognition (ICPR 1994)*, vol. 1, 582–585.

[19] T. Ojala, M. Pietikäinen, and D. Harwood (1996). A Comparative Study of Texture Measures with Classification Based on Feature Distributions, *Pattern Recognition*, vol. 29, 51–59.

[20] T. Ahonen, A. Hadid, and M. Pietikäinen (2006). Face description with local binary patterns: Application to face recognition. *TPAMI*, 28(12):2037–2041, 2006.

[21] S. Belkasim (2011). Multi-resolution Analysis Using Symmetrized Odd and Even DCT Transforms, *Data Compression Conference (DCC)*, 447pp.

[22] K. Daewon and S. Daekyu (2003). Energy-based adaptive DCT/IDCT for video coding, *International Conference on Multimedia and Expo ICME'03*, Vol 1, 557–560.

[23] S. Gharge and S. Krishnan (2007). Simulation and Implementation of Discrete Cosine Transform for MPEG-4, *International Conference on Computational Intelligence and Multimedia Applications*, Vol 4, 137–141.

[24] Matlab Revision R2012b Documentation (2012). *Mathworks Online Documentation*, [Online] Available at www.mathworks.co.uk/help

[25] A. Webb and K. Copsey (2011). *Statistical Pattern Recognition*, 3rd edition, Wiley, 666pp.

[26] FEI Face Database (2014). [Online] Available http://fei.edu.br/ cet/facedatabase.html