Quick and Dirty: Streamlined 3D Scanning in Archaeology

Jarrod Knibbe[1], Kenton P. O'Hara[1,3], Angeliki Chrysanthi[2], Mark T. Marshall[1],
Peter D. Bennett[1], Graeme Earl[2], Shahram Izadi[3], Mike Fraser[1]

[1]Interaction & Graphics Group, Department of Computer Science, University of Bristol, UK
{ jarrod.knibbe, mark.marshall, peter.bennett, mike.fraser}@bristol.ac.uk

[2]Archaeological Computing Research Group, University of Southampton, UK
{ac14g09, graeme.earl}@soton.ac.uk

[3]Microsoft Research, Cambridge, UK
{v-keohar, shahrami}@microsoft.com

ABSTRACT

Capturing data is a key part of archaeological practice, whether for preserving records or to aid interpretation. But the technologies used are complex and expensive, resulting in time-consuming processes associated with their use. These processes force a separation between ongoing interpretive work and capture. Through two field studies we elicit more detail as to what is important about this interpretive work and what might be gained through a closer integration of capture technology with these practices. Drawing on these insights, we go on to present a novel, portable, wireless 3D modeling system that emphasizes 'quick and dirty' capture. We discuss its design rational in relation to our field observations and evaluate this rationale further by giving the system to archaeological experts to explore in a variety of settings. While our device compromises on the resolution of traditional 3D scanners, its support of interpretation through emphasis on real-time capture, review and manipulability suggests it could be a valuable tool for the future of archaeology.

Author Keywords

Archaeology; Interpretation; 3D scanning

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

Data capture has long been an integral part of archeological practice. Along with other disciplines, archaeology has embraced new technologies on site digs that enable an ever richer set of data to be assembled for the documentation and interpretation of archaeological discovery. A traditional separation exists in archaeology between the practices of excavation in themselves and the post excavation construction of narrative for dissemination [1]. For those working within the constructs of this division, data

gathering and capture has been conceived as something that is simply done prior to stages of post-excavation interpretation [21]. That is, the act of archaeological interpretation is seen as something that is separable from processes occurring on site. This in turn has led to particular ways in which capture technologies are viewed both in their design and in their manifestations in the context of practice.

More recently, however, the separation of excavation and interpretation has come under scrutiny from the academic community and it is becoming more widely accepted that interpretation is 'insidiously ubiquitous' [17]. Within this perspective, the very act of excavation is to transform physical deposits into interpreted archaeological features [7]. This has created a drive towards 'reflexive archaeology' with the intention of recombining excavation and interpretation [6]. A primary feature of the organization of archaeology is that it is a physically irreversible process; that once a process is decided upon, based on interpretation, other lines of enquiry cannot be pursued. Those in favor of reflexivity argue that destroying the relationship between artefacts and their setting means that the moment of excavation is always the best opportunity to explore alternative interpretations [6]. In the context of these arguments, the traditional methods of trying to capture data "objectively" for later interpretation inherently constrain future generations, since they can neither participate in ongoing interpretive practices on-site nor support the subsequent reconstruction of the thinking behind an excavation in order to reconsider processes or conclusions [1].

In considering the support of on-site interpretation, our attentions are shifted to a variety of different factors. Archaeological expertise plays a significant role; for example in differentiating between different types of soil, recognizing plant matter and determining the age of pottery. In his seminal work on archaeological practice, Goodwin [11] argues that *professional vision* – the articulation of skillful seeing that demonstrates and reifies expertise – plays a key role in the production of archaeological interpretation, crucially even in situations where knowledge is objectively codified in processes or artefacts. Physical tacit knowledge also plays a large role and helps inform

interpretation such as feeling changes in the texture of the floor, the cold touch of metal or the sharp edge of a cutting stone [6]. The sharing of this interpretation with other members of the excavation team through talk and body draws on a wider range of previous experience and background knowledge, helping the shared development of interpretation [10]. As a result, there is an increasing focus on collaborative process as well as resulting product, on 'documenting the documenting' [6], so that the patina of the interpretation can be retraced and the unfolding tacit site knowledge is not lost.

The interpretive work being performed and the practices surrounding this work are intimately bound up with the broad range of capture technologies available to the archeologists [14]. These include the more traditional paper-based artefacts such as finds recording forms, and excavator diaries as well a variety of digital capture technologies. For example, digital cameras are now an important method of capture within archaeology providing good two-dimensional records of finds and their surroundings, as well as depicting things such as soil color changes [32]. Likewise, video recording is used on some sites as a means of documenting key features of the excavation either through ad hoc recording or performed by a dedicated site videographer [5]. Increasingly too, more sophisticated 3D scanning and capture technologies are being adopted on archeological dig sites. These methods include geophysics [9], photogrammetry [32], spectrophotometry [31] and laser scanning [25] that all produce highly detailed 3D representations of on-site scenes and artefacts.

What is important about these capture technologies is that they have come to be appropriated in particular ways in the context of archeological practice that subsequently shape and constrain their relationship with the different locales of interpretation on- and off-site. The concerns go beyond what the different representational outputs of these capture technologies enable the archeologists to record, see and analyse. Rather there are a broad set of factors and practical considerations arising in the preparation work and production work involved in different forms of capture that come to bear on the relationship these technologies have to the locales of archeological interpretation. Indeed, even those capture technologies that on the face of it might be regarded as lightweight can often have heavyweight methodologies associated with their use on site. Consider for example the onsite use of digital cameras for capture. In producing the "right" kind of image there are a variety of issues with lighting, resolution and so on that affect how things such as color will be manifest in the captured image. This in turn will affect the extent to which an image may or not be amenable to particular forms of expert inspection and differentiations from surrounding materials. As such there is a significant amount of time consuming work done to "set-up" the captured scene to best highlight the observed and agreed phenomenon [11]. The time spent cleaning, lighting, shading, and measuring prior to this photographic

capture then serves to situate such capture in the context of particular forms of post dig interpretive work that is separate from the ongoing interpretive work at the point of excavation.

Similar arguments are at play with the use of more intensive capture equipment, in particular, the various 3D scanning technologies adopted on dig sites. Such equipment is typically very expensive, of the order of tens of thousands of pounds which places financial restrictions on their use and availability on site. Their complexity also requires considerable training and skill to operate, requiring the presence of specialist individuals on site, which again leads to certain restrictions on when and how they can be deployed during excavation. Such techniques, even when they have an automated processing pipeline (e.g. [1]) typically take a significant amount of time to process and produce the actual 3D reconstructions [30, 3], which often occurs after the dig. As such they are not necessarily available for instantaneous feedback on site, making it difficult to integrate results with the on-site interpretation work. The equipment is also cumbersome and time consuming to set up which imposes limitations on the ease with which it can be moved around the site. Finally, the use of such techniques requires significant preparation of the site of scanning. Significantly, in the context of our concerns with on-site interpretation, the scanning area needs to be cleared of the people involved in the excavation so to avoid interference with signals and technologies. So while these techniques serve to produce high quality representations, the factors surrounding their production situate their use in the locale of post-dig interpretation work and inhibit opportunities to juxtapose the scans with on-site archaeological interpretation.

What we see then, as Hodder argues [14], is that despite the value of these capture technologies, the pragmatics of their use still renders them at odds with the broader changes within the discipline that are seeking to erode the division between excavation, analysis and interpretation. As Dellepiane et al argue [8], an erosion of these divisions requires a greater exploration of lightweight technologies combined with lightweight processes for capture. These arguments are not intended to be dismissive of the current capture approaches and their value but rather serve to highlight these unfulfilled niches of interpretation work which appropriately designed capture technologies could facilitate. With these arguments in mind, Dellepiane et al set out a set of key principles that capture technologies would need to adhere to in order to fulfill niches of on-site interpretation: repeatability, comparability, low-cost (in terms of money, time and skill), workflow compatibility and presentability. Based on these principles, they go on to explore a lightweight approach to 3D capture using digital cameras. While their approach addresses a number of the factors that contribute to heavyweight capture processes, they were unable to avoid the multiple hours of post-processing before the material becomes available. As such

there remained key aspects of use that preserved some of the separation between excavation and capture.

The arguments and intentions of Hodder [14] and Dellepiane et al [8] provide an important launch point for the work we present in this paper. Here our concerns are with the exploration and development of a lightweight 3D scanning technology and process for archeological settings that can be situated within the locale of on-site collaborative interpretation work. Such explorations of the "lightweight" remain non-trivial in the sense that all sorts of potential trade-offs in the design space are at play that will come to bear on how such a technology can be situated in on-site interpretive practice. For example issues such as cost, portability, speed of processing, what can be captured and the quality of what is captured are all factors that can potentially be manipulated and traded off against each other in different ways; where one factor is emphasized, another may be compromised.

While there may be no singular right answer here regarding the quiddity of "lightweight-ness" of 3D scanning, what is important is that the design decisions and trade-offs embodying this essence are understood and made in the context of on-site interpretive practice. To this end, then, we begin our explorations with a more detailed inquiry in the form of two field studies carried out at two different archeological dig sites. The aim of these studies is to elicit key features of on-site interpretative practices and how they come to be collaboratively organized in particular ways. In particular, we concern ourselves with some of the temporal and spatial features of these practices and the factors that impact on the pace and scope of activities. Furthermore, our aim is to elicit more detail as to what is important about this interpretive work, what is being lost with existing capture practices and what might be gained through a closer integration of capture technology with these practices.

Drawing on the insights derived from the fieldwork, we go on to present a novel portable, battery-powered and wireless 3D modeling system and discuss its design rationale in relation to these insights. In particular, how by changing parameters such as the temporal and spatial possibilities for capture, the system might enable capture to be situated in new ways in the context of collaborative interpretation. We evaluate this rationale further by giving the system to archaeological experts to explore and critique, with a view to informing directions for future development. Our approach to this work fits within the long CSCW tradition of using ethnography to inform systems design, but we are also mindful of the longer term potential of 3D capture as a technology to enhance collaboration across a range of domains beyond archaeology.

OBSERVATIONS FROM THE EXCAVATION SITES
In order to further understand on site archaeological practices, the work of interpretation and relationship to potential opportunities for capture, we carried out a series of field studies at two archeological dig sites: Clatford in the UK and Çatalhöyük in Turkey.

Clatford
Clatford is a Neolithic site in Wiltshire UK, near to the Stonehenge World Heritage Site. 2 years previously, geophysical data had shown evidence of another 'sarsen stone' henge monument, including further stones that may have been destined for the incomplete outer ring at Stonehenge. The site had also been identified as the possible location of a prehistoric river crossing and is situated within close proximity of a number of important burial sites. Around 15 archaeologists were participating in the excavation, including site supervisors, experienced excavators, geophysical and location mapping specialists and a number of archaeology students. As well as observing on-going work practices of the excavation, we conducted informal interviews with the various participants throughout the dig to provide additional context and explanations for particular organizational features of the work.

Temporal pressures on the organization of work
An overriding feature of the excavation at Clatford was the adverse weather conditions. Given that the archeological team will be on site for a bounded period of time (2 weeks in this case), the weather comes to play a key role in determining the organization and sequencing of particular activities. While we were on site, rain halted digging at various stages, but in doing so it also served to expedite the work conducted during clear moments and highlighted some of the features and pressures of archaeological practice. Key here in the context of the weather was an uncertainty surrounding the availability of time periods where particular tasks would or wouldn't be possible. The consequence is that the site leader and the rest of the team reactively organize the work to make use of current and prospective clear periods. The uncertainty, then, led to many different work streams being conducted simultaneously over different parts of the site: photographic capture, feature excavation and location measurement.

Figure 1: Various parallel activities at the Clatford dig site.

Implications for capture and interpretation
The temporal pressures and resulting distribution of tasks across the site, then had some important consequences for the organization of certain capture activities which in turn impacted on opportunities for interpretation during excavation. To illustrate, let us consider an episode in

which the site supervisor was keen to capture photographs and laser scan measurements of an uncovered feature of the site prior to the next rainfall-caused interruption. In preparing the capture area, the excavators were asked to shift their excavation activities to a different part of the site. Even throughout the increased pressure of the dig, the supervisor proceeded to spend considerable time (half a day) cleaning, framing and measuring his target area prior to capture. This also involved moving a vehicle to the edge of the trench, such that he could achieve a higher vantage point for capture.

Two further experts spent the day making detailed measurements of the site using laser range finding. This heavy, specialist equipment required interruption free, line-of-sight recording, at times leading to further disruption of the excavation work. Significantly, the scanning base stations were frequently subject to marginal movements – for example, to allow for better alignment or to simplify subsequent measurements. This leads to potential issues with the standardization of measurement accuracy across the different scans taken, but the temporal pressures of the dig mean it is not always possible to repeat all the measures from scratch. The complex nature of these capture activities and processes again demanded that the supervisor was there to oversee and direct what was happening, and thereby was away from the activities of the rest of the excavation team who focused on a small ditch running across the site.

Highlighted in this episode are a number of issues relating to our concerns. First of all, the capture process is hugely time consuming in its setup and execution. The equipment is cumbersome to assemble and demands significant expert resources to operate. Of further significance is how this complexity and time to operate can inhibit efforts to repeat or redo measures in an accurately standardized way when trying to account for various pragmatic constraints imposed by the site. Importantly too, the process removes the excavators from the point of capture. The demands of the technology and the process mean that the capture is thus conducted only as an aid to post-excavation interpretation. The temporal pressures imposed by the weather mean that excavation work continues elsewhere in parallel with these capture activities. Significantly, this excavation work continues in the absence of the site supervisor who is overseeing the capture process. Again the time demands of the heavyweight capture process means that he is removed from the point of excavation for long periods of time therefore acting to distance himself from the developing interpretation during excavation work. While it was possible for him to be 'brought up to speed' later, the heavyweight procedure associated with photographic capture prevents this work from being conducted by a less experienced member of the team, in turn freeing the supervisor to continue overseeing the wider work. Finally, as a consequence considerable work was conducted on the excavation, based on interpretation of the materials being unearthed but without any capture being conducted – both the supervisor's absence and the complexity of the capture

equipment being potential factors at play here. As such much of the textural and wider contextual details surrounding the excavation work and in situ finds are left unrecorded and consumed into interpretation without availability for later re-inspection.

What this begins to point to are the variety of ways in which key points of separation between capture and excavation based interpretation are introduced that relate to the incompatibility of capture time demands with the temporal demands of excavation work. This then suggests that significant benefits in the organisation of these practices could be achieved with capture techniques that significantly reduce: time to capture, complexity of capture and required expertise to capture. In order to develop these insights and elicit more detailed understanding of interpretive capture requirements and settings we undertook further fieldwork at a second site: Çatalhöyük in Turkey.

Çatalhöyük
The Neolithic site of Çatalhöyük, Turkey is one of the first urban centres of the world (7400BC) and now a World Heritage Site. The research project at Çatalhöyük has played an important role in both revealing new information about the people who lived in the settlement and developing some of the most influential theoretical, methodological and technological changes in the field of archaeology.

Our study took place over six days of the dig and again we were able to observe and informally interview key members of the archaeological team. We also wanted to capture key episodes of interpretive practice during excavations. The aim here was to understand in more detail how the episodes come about, what kinds of things are oriented to in conversation, how they are collaboratively organised, the relationship of the interpretive work to the artefacts and setting, and how such interpretive work relates to opportunities for capture. In order to capture the events in close detail we employed a lightweight 'Looxcie' ear-mounted digital video camera to be worn by a member of the archeological team. Over the course of the study, seven members of the archeological team wore the camera during their excavation work providing us with a broad coverage of different archaeological processes and conversations as they unfolded. At the end of each day participants would review the captured footage and discuss with us their perspectives on the captured footage. The archaeological work at Çatalhöyük is conducted in English so no translation was required.

Following the deployment, the captured videos were catalogued, and watched to highlight moments of interest in understanding the working and capturing. We followed a methodology based on interaction analysis employed in workplace studies in CSCW [13, 25] that emphasizes the ways that the archaeologists themselves responded to one another's actions. Selected moments were transcribed (using Jeffersonian notation) and subjected to repeated viewing to gauge the ways in which the participants worked to create interpretation in the trenches. Of these sequences,

an illustrative subset is presented that aids the further exploration of the key requirements for streamlined capture.

Pre interpretation and the capture of special finds

The majority of uncovered finds are commonplace and homogeneous, illustrative of known issues and therefore less of a concern for capture. Of greater interest are "special finds" that have greater significance to the archeological narrative of the site. Here it is important that detailed information about 'good' finds is accurately recorded; their condition, context in-situ, exact location and the surrounding area. However, this is not as straightforward as it may at first appear. Indeed the demarcation as a special find and consequently what is important for capture is an emergent product of close collaboration interpretation on site, with an emphasis on how reconstruction for later interpretation is possible. Let us consider the following sequence in which the bone handle of a tool, presumed to be an axe, is found by Earl. The field director Beth looks at the find while Amos, another excavator, focuses on finding any extra missing pieces.

Amos (Excavator): Did you find the rest of the rock that fits in here
Beth (Field Director): (.hhh) (hhh.) (what) a fresh err break of the umm of the stone. And it is cut marked too (.) this is really cool you should walk around the trench and show everybody (.) OK we have the find of the day slash season slash year slash century slash millennium
Amos: you need the (.) you [need]
Beth: [you will see:::]
Amos: You need the
Dave (Excavator): [[inaudible]]
Amos: (inaudible) [[whistle]] Oi whats that
Earl: Let me see a second°
Beth: [Err:::]
Amos: [You] need (.) (Eugen) you need to find the rest that was sticking in [here]
Beth: [yeah:::]
Amos: [It's a long] shaped stone
Beth: [Is it?]
Earl: [(inaudible) something else inside of it?]
Beth: Is it already in the sieve?
Earl: we ok (.) hold up one [second ((raises finger))]
Charlie: [Looks like an axe] wow::: ma:::n
Beth: But its got a
Charlie: mm::: (.) find of the day
Amos: its got a cut up the [[incisional bulb]]
Beth: [um:::]
Earl: this was next to it but this is not it
Amos: you need a stone that fits into here
Beth: did you xfind that one and you xfind that one too right? umm::: did you already send the soil to sieve that might contain the rest of that?
Earl: [yeah]
Amos: [just scratch that thing] on a bit of wood and its like an axe or an axe head

This fragment illustrates what we come to label as the pre-interpretive practices bound up in the capture and recording of finds. Interpretation of these finds is emergent and dependent upon the assembly of other material evidence. These pre-interpretive practices then are expressions of expertise required to anticipate how a record will support later interpretation, and therefore the necessary scope and context of that record. Here, the scope of recording a find is

interwoven with the collaborative interpretation necessary to create a sufficient record for later. The object as a find is worked up into the subject of its importance, and the two are inextricably linked on site. In shaping our thinking about capture technologies, this leads us to consider the importance of how they can be situated within the context of these collaborative pre-interpretive practices with particular attention to their spatial and temporal demands (for example by exploiting faster and more mobile capture possibilities).

Revealing features: colour, texture and context

The excavation at the West Mound area of Çatalhöyük is principally concerned with the mapping and understanding the structural elements of an urban conurbation, much of which involves the delineation of walls and floors. One of the key tasks around the site is to determine where such features were located, and the order in which they were constructed. Observing the excavators at work makes it immediately apparent that differentiation of colour, texture and consistency plays an important role in guiding the direction of excavations.

Figure 2: Images showing an excavator using a water sprayer to exaggerate color details. The excavator can be seen wearing a Looxcie on his ear (left) which captures images such as those on the right.

The most popular method for determining wall layout and building order is to trace a mortar line, a visibly darker line between lighter bricks. If it can be shown that the mortar line continues between sections of walls, then this can be interpreted as a continued wall. An excavator suggests "*When we were digging earlier I saw definite lines, but I am not seeing them now... sometimes they don't come out as well if you don't have the sprayer... with the sprayer they really pop out... See there is the line coming out more. Its proof, we got it.*" The water sprayer is one tool which can be used to exaggerate details of color, both for guiding the direction of excavation and for enhancing color differences for photographic capture (see figure 2 above). Texture proves equally important in differentiation providing the clearest definition between mortar, brick and plaster. On occasion, for example, when attempting to locate the floor, texture plays a very important role in guiding understanding of construction and use.

Anne: get everything out that is (0.5) related to the floor and it probably will ha- ((stutter)) will have a very bumpy base
Bella: mmhmm

Anne: because they might just have had umm room
fill that (.hhh) altogether seems to make a little
bump in the room (.) [in the centre]
Bella: [and then they flattened it] by doing a
floor
Anne: yeah and then they didn't flatten the room
fill [before]
Bella: [yeah]
Anne: they just filled all the bumps in the room
fill with [floor]
Bella: [yeah] so maybe that is why it is thick in
some areas and [thin in others (.) ok]
Anne: [yeah (0.5) yeah]
Bella: yeah
Anne: and we want to expose exactly that so
Bella: yeah ok

Here the discussion centres around exactly what constitutes the floor, and whether the rather loose fill in the bumps actually constitutes the floor, or whether the floor will be exposed upon removing the loose fill to discover a (rather counter-intuitively) bumpy floor. Thus the textural 'thinness' of the filling is found critical to the judgement of the floor's definition. Alongside the need to capture specific colour and textural detail, the excavators were also concerned with capturing this with sufficient surrounding detail to later enable wider understanding.

Losing opportunities for capture
Of particular note in our fieldwork enquiries was how particular aspects of archaeological practice were resulting in key opportunities for capture being missed through disruption, re-arrangement and destruction of material and speed pressures of work. We begin here with a look at how preparations for a particular capture narrative remove and rearrange things of relevance to a different narrative. In the following episode, the site leader interrupts the excavation to prepare it for capture but thereby disrupts the ongoing excavation and interpretive narrative being assembled.

Alan: There is no mortar here (.) there was mortar
above this area but none here (.) this is like one
big compacted block (.5) goes all the way down a
lot thicker
Anne: what you were showing in the photo (.) is
the information which (.) yeah
Alan: ((points)) well this needs to go because it
just looks bad
Anne: So we will never know that it was up to here
((points))
Alan: we could sorta just clean it a bit and leave
it like that (.) I don't know
…
Alan: So we sorta need to think
Anne: Water?
Alan: Water?
Anne: Take one dry photo (.) then we'll wet it
then we'll have colour (2.5) we really have to
take that out ((points)) because even if we leave
it in nobody is gonna know what the hell it is

This process then involves precise trowelling, dusting, tidying and further cleaning before summoning the photographer to capture the scene. On the one hand the fragment highlights the constructed nature of particular capture practices; the need for framing the scene, actively removing or emphasising detail such that specific understanding can be clearly interpreted post-excavation.

But at the same time in these constructed practices, the preparation and cleaning processes remove features of interest to the other narrative leading to Anne's disappointment. Had Anne had more personal control over her own capture, and been able to do it quickly, this narrative evidence might have been recorded.

In a related example we see further evidence of missed or avoided opportunities for capture. Again temporal issues are at play here in which there are discrepancies between speed of capture and speed of excavation. What is of particular significance in this example is the temporal ordering of archeology whereby particular features available at any one moment may need to be destroyed or removed in order to reveal other aspects of relevance to the interpretation. In the example, a senior excavator is describing the process of excavating features in a wall through the differentiation of texture.

Alan: this mud brick just slices (hhh) (.hhh)
(hhh) maybe ((pointing gesture)) (.) maybe
((pointing gesture)) (0.5) I don't know (1.0) not
terribly convincing
Bob: no
Alan: and that's one of the problems is that it's
really (.) difficult to see it until you actually
like cut into it
Bob: yeah
Alan: at which point you kind of lose the (.) edge
a bit

In this case the edge is so poorly defined that gauging the texture requires destructive actions on the surface necessary for guiding and studying the surface itself.

In a final example, we highlight other work and time pressures precluding accurate capture. In this episode an archaeologist has been tasked with lowering an entire section of floor. In the process of this primary task, he excavates a lot of 'find' material (e.g. pottery and bone), but has to delay processing it until later. He create piles of material on an opposite surface, where additional material is then placed. This placement of material, combined with his pick-axe based excavation, leads to a significant amount of contamination and breakage. After considerable work, as the piles of finds become structurally unstable, he breaks to record data surrounding this material. He also makes it known to the capture expert that he requires a 'point' (an accurate, standardized position measurement) '*at some point (.) but no hurry*'. At this time all accurate contextual information surrounding each find and information surrounding its precise location has been lost as considerable excavation has been conducted subsequently.

RECONSIDERING OPPORTUNITIES FOR CAPTURE
With these issues in mind we wanted to consider how new forms of capture technology might offer new ways to recombine with the practices and processes of excavation. Our intention here was not to look to replace other capture practices and technologies but rather, consider niches not currently supported in this broader ecology of capture practice. In particular, we wanted to explore how new technologies could sit within the collaborative practices of

interpretation in terms of their spatial and temporal organization and specific features to which they orient. More specifically we wanted to consider ways of putting capture capabilities into the hands of the excavators in ways that might support and document their excavation and interpretive work. In this regard, we were looking to avoid the need for time consuming clean up and preparation practices in the context of capture; the way this disrupts the excavation, the way this removes the excavators from the point of excavation, the way this demands the dedicated attention of site leader, the ways in which this demands dedicated technical capture expertise, and the ways that this disrupts the material assemblies of the site relating to the interpretive narratives of the excavators. Further our aim was to orient to the inherent time pressures of archeological excavation through which useful capture and review opportunities are lost – in particular where states of artefacts and assemblies at any moment may be permanently lost through the destructive requirements of interpretation.

To this end, we looked at how we could alter particular parameters and characteristics of a scanning technology (and the balance of trade-offs) to fit within the niche outlined above. First of all, then, our emphasis was first on speeding up capture activity to avoid lost opportunities and enable the review of the material in situ in support of ongoing interpretation. Second, we looked to emphasise greater mobility – to have something that could be personal to an excavator and that could be more easily positioned around the site and particular artefacts in the context of excavation based interpretation. Such shifts will entail certain compromises elsewhere, such as, for example, resolution. In making these shifts we wanted to ensure that the resolution of capture still revealed the *sufficient* detail of artifact form and texture oriented to in the development of these interpretive narratives. We move on to consider these shifts in the context of a personal 3D scanner.

A PERSONAL MOBILE 3D SCANNER

Given these requirements we turned to recent developments in real-time, small form factor scanning exemplified by recent advances in Simultaneous Localisation and Mapping (SLAM). With their Kinect Fusion algorithm [20, 28], Izadi et al present a system capable of real time dense model acquisition using a Microsoft Kinect camera. Their approach allows the near real time capture of 3D models of objects, scenes and body movement that can be integrated together or differentiated from one another - providing a bridge between the requirement of personal capture in reflexive archaeology, and the demand for 3D modeling in archaeological capture and interpretation: both capturing for interpretation and capturing interpretation. A related approach is presented in [29], where a depth camera is replaced with a traditional monocular webcam. Such an approach, though, relies heavily on movement to capture a scene, leading to an unnatural capture technique, and textural data is more heavily affected by color variation. Thus we decided to use Kinect Fusion with our device.

Hardware

To develop our mobile scanner (see figure 3), we extracted the cameras, infra-red laser and main camera board from a Kinect and combined this with a wifi dongle, a low-power Gumstix embedded linux computer, a 7.5V camera battery and a custom-made circuit board to route the resulting camera data and to provide the required voltages to power all of the components (which require a mixture of 1.8V, 3.3V or 5V). A fan, heatsinks, on/off LEDs and a charger adaptor were also added. The design reduces the physical size and power requirements of the existing Kinect, facilitating its mobility. Finally, we designed a case to enable one-handed use of the device, supporting interaction with the environment during capture. We have published our open source hardware design [27] for reuse in other applications.
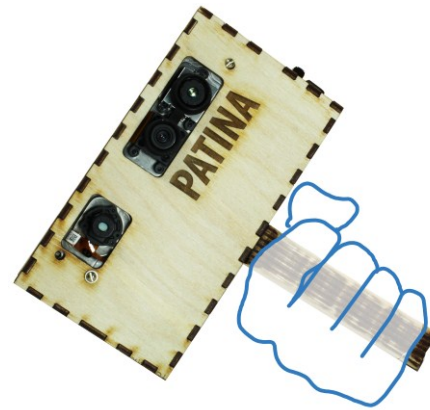

Figure 3: Our prototype personal 3D scanning device

Software

The software for the device is in two parts. Firstly, using open source drivers recompiled for the Gumstix, we built network code that would enable the Kinect to connect to a WiFi network and transmit depth images at roughly 24 frames-per-second. Given this frame rate is sufficient for our purposes, we have not optimized depth data streaming, although there are novel algorithms in this area [22]. Secondly, Microsoft's Kinect Fusion algorithm [see 20 for full description], running on a nearby server is used to construct full 3D models from the data. The algorithm determines the six degrees of freedom transformation of the Kinect between subsequent frames and uses this to integrate all captured points into a 3D voxel volume. This volume can be ray-cast in real time, producing a model of the scene. This process is carried out on the GPU in order to achieve a real-time and interactive frame-rate. The algorithm treats moving pixels in the scene as outliers, discarding them from the model. If these pixels are maintained separately, however, it can be determined when they overlap with the captured model in 3D space. This means that, should a hand enter the scene, any interaction it has with objects in the scene can be recorded.

To facilitate the use of the scanning device on large outdoor sites, we developed an application for an Android Tablet to enable remote control of the device and the software. The

app allows the user to start and stop the capture process as well as control both the volume ($2m^3$ to $20cm^{3)}$) and resolution of the area being captured (to support shifting perspectives of the archaeologists from wide areas, such as an entire wall, to specific points, for example where a piece of pottery lies). The excavators may also be capturing models for a variety of different purposes, for example as a quick reminder or for more detailed examination later. By controlling the temporal properties of the Kinect Fusion algorithm, the archaeologists can control the time taken to capture frames for the model and, thereby, the resolution of the output. This adheres to our design considerations, allowing the user to determine the trade-offs between detail and efficiency for each scan.

INITIAL EVALUATION WORK

To understand our device further within the context of archeological practice we conducted some initial evaluation studies with archaeologists. First of all we gave the system to a team of archaeologists working in the Crypt of Winchester Cathedral. This provided an interesting point of contrast for our evaluation discussion as the team had previously deployed more precise archaeological laser scanning techniques as part of their work there. We also gave our system to a number of different archeological specialists working in Lithics, Maritime, and Osteo-archaeology laboratories. The archeologists had the opportunity to use the system and review the captured outputs in the context of their work. We observed them using the system and discussed their experiences in relation to how particular system characteristics pertained to their interpretive practices.

Granularity of detail

We begin our discussion here with a look at the central issue of granularity of detail captured in the scans. What was apparent here was that the required granularity of detail in the scans was dependent on the area of expertise and, more specifically, the materials of interest. For both the osteo-archaeologists and maritime archaeologists, the concerns of their work are with less durable materials subject to erosion within their natural environment. As such, their focus is often on coarser features of their finds.
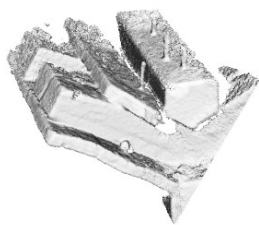

Figure 4: Maritime archaeologist explaining a ship segment (left) and the captured models (right).

For example, while scanning a hull segment of a ship with our device, a maritime archaeologist remarked that the detail captured was of sufficient quality to examine the important features, such as location of the securing bolts, the timber construction techniques used and the location of

the mast (figure 4). Working in heavily constrained environments, these details are the typical focus of maritime archaeologists, as they allude to the manufacturing location and period, and overall size of the vessel. He remarked that the ability to *quickly* revisit and explore this information upon completion of his work would be useful to him. Communication within his work setting is limited, and so interpretation is typically shared immediately upon surfacing. He suggested a model of this resolution would facilitate this process.

For those working with these softer materials then, the resolutions of the scans were suitable for key aspects of their work. For the archaeologists working with stone and other similar materials there were some requirements for the finer detail of their objects. For the lithics specialists (focusing on the fine details of smaller hand axes and blades), they suggested that the capture resolution of Kinect Fusion was not sufficient for some of their analytic requirements. Of further significance in their comments was that while they invest a large amount of time in drawing and capturing data around their finds, it is not just the output of this process that is of interest. Rather, the aim is the interpretation gained by spending time focusing on the minutiae of their research materials. Much of their work is also dependent upon cleaning the finds post-excavation for use with specialist equipment in a controlled environment. As such, it was not felt that our scanner would deal well with the specific site and details that are central to the interpretive work of these specialists.

However, what we found was that the demands of resolution detail can also be dynamic and that models of the same things at different resolutions can play different roles in the interpretations they enable. We see this illustrated by the team working in the Crypt of Winchester Cathedral. During renovation works over the last few centuries, various stones have been moved from their original setting to other parts of the cathedral. Located within the crypt is an example of constructor graffiti, where one of the masons had engraved a message into the stone. The team had previously scanned this using more precise archaeological laser scanning techniques. A comparison of the output of our model against the output of a photogrammetric model can be seen below (figure 5). While it can be seen that the model produced by our device contains significantly less detail, it does show the textural detail of the surrounding wall, including brick outlines and mortar thickness. The scanning expert at the Crypt remarked that *'you managed to capture this bit a little more clearly than here [the higher resolution model],'* showing that, by removing some of the finer details of tool marks and chips, the structure of the material can be more clearly defined.

However, where these concerns really come into play is that interpretive work is not just bound up in static scans but in their additional dynamic transformations and manipulations. Resolution detail here then was not simply about what could be viewed in the static but what it meant for the way the model could be manipulated. The team

articulated that high-resolution models demand a significant computational resource to render and explore. For these reasons, they deliberately invest time to reduce the resolution of their 3D models (or treat them as texturally rich 2D photographs) to support their faster manipulation and explorability, referring to this as *"throwing information away"*. Such a trade-off between resolution detail and manipulability is embodied in our scanner. This was highlighted rather nicely by one of the experts there who felt able to explore the model from our lower resolution scanner more freely than some of the other scanned material. During his manipulations, he turned the model upside down and repositioned the virtual lighting, something he had been previously unable to do. In doing this he realised that the text could in fact be read (as can be seen in figure 5, where the models have been rotated from their original capture orientation).
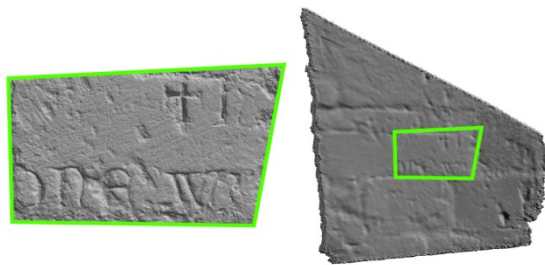

Figure 5: A comparison between a photogrammetric model (left) and our captured model (right) of masonry graffiti at Winchester Cathedral.

Speed of Capture and review

Of further significance in our discussions with the archaeologists was the ability to capture and review the scanned models in a matter of seconds. Where this becomes important is in the constructed nature of the narratives that are bound up in the process of capture – that is how they are designed to render certain features visible. For these reasons, it is important to review the outputs from capture to ascertain whether it appropriately reveals things as intended by the narrative. Such review of the capture materials can actually be problematic with the current high end scanning technologies. As the scanning expert on the team at Winchester Cathedral explained, creating a photogrammetry model takes 1 hour of on-site capture (during which time the capture process cannot be interrupted and no work can be conducted within the capture vicinity) and 6 hours to render on a powerful GPU cluster. During this time there is minimal feedback, resulting in uncertainty surrounding the final output. The total time for model acquisition makes it impractical for small changes to be made and models to be repeated. For this reason, he expressed that a significant benefit of our system lay in the instantaneous nature of our capture. This would allow outputs to be reviewed in near real time to ascertain how they were working in the context of the constructed narrative being produced.

The speed of our device would also make it possible to capture a more varied set of outputs. An example he gave here was that he would be able to capture detail surrounding the engraving, such as how the particular stone fits within the makeup of the wider wall. The ability to quickly capture multiple scans of the surrounding contexts as well as the details of particular features would add significant evidence to the kinds of interpretations they would be able to make.

We saw some related issues here with the Osteo-archaeologists. One of the problems they experience is that logs completed during excavation are often of varying quality and precision with much of the situational information surrounding the find lost. Commenting on a scanned model of a skull using the system she suggested *"if you can get some kind of 3D model in the grave to actually get an idea of bones spatially in relation to one another, and depths [that] would be useful"*. Of particular concern to her though was the degrading of artefacts of interest over time and she pointed to the opportunities opened up for repeated captures over time: "*An issue we have with skeletal remains rather than maybe hand axe stuff, is the fact that it does disintegrate a lot more… it will completely change, the bones will move, the coloration will change and you end up having a different interpretation of what you are looking at, so it is useful to be able to capture things at different stages.*" Here being able to capture models quickly would greatly facilitate repeated capture over time preventing the loss of artifact states important for her interpretive work. The speed of capture of our device would allow her scanning to be conducted in a manner that results in 3D 'time-lapse' style scans, in turn affording her the ability to determine between the effects of excavation and the effects of time. This is one aspect of her archaeology that is currently not possible for her to capture. She also raised the possibility of printing 3D models from the scans that would greatly assist in these aspects of her work.

Pragmatics of use

In observing the archaeologists using the device, it was evident how they referred constantly to the monitor which depicts both what the camera can see in real time and the resultant effect on the created model. In doing this they were constantly verifying that they were focused on the area of relevance and capturing it appropriately for their concerns. Given the particular setup we employed, there were times when people were looking in the opposite direction from which they were capturing and, therefore, having to mentally translate the movements required to better frame their target object or fill holes in the model. In addition, the Kinect requires a distance of roughly 60cm between the camera and the object for depth readings to be made. While this was explained to every participant, there was a natural tendency to move towards the target object, as though 'zooming in' to enhance resolution. This behavior decreased in occurrence with continued use, but reliance on the monitor for feedback continued. This suggests that there may be some usability benefits to be gleaned from adding a small output screen on the back of the scanner – making it easier shift focus between the actual objects and

corresponding capture output. We will look to explore this in future iterations. Another aspect affecting ease of use was the portable form factor of our device. Participants happily carried the device with them and continuing to converse and gesture whilst holding the device. Here we see an indication of its ability to support the interleaving of capture with the interpretive discussions surrounding it.

DISCUSSION

In this paper, we have begun to explore some of the interpretive practices in archeology and the particular ways that they are bound up with certain practices of capture. More specifically, a central concern in our work, has been with understanding some of the separations that happen between excavation, analysis and interpretation and how these relate to the ways that on site work of archeology has come to be organised. Much of this is in response to particular temporal pressures of the work and the contextual constraints at the sites of practice. Furthermore, we have highlighted some key features of current capture technologies and practices within these archaeological settings that do not sit neatly with the ongoing collaborative practices of interpretation during excavation. As a consequence, much of this interpretive practice remains unsupported by existing capture techniques – neither supporting it as it occurs, nor documenting it in aid of post excavation practices. Much of the temporal unfolding of the archaeology, and what is destroyed in the process of revealing remain lost.

In orienting to these separations, our approach is not to be critical of current capture techniques and practices per se. Indeed, these practices remain hugely important to current archaeological methods. Rather than replace these then, our aim has been to understand how new capture technologies might be positioned differently within the broader ecology of these capture practices. In particular, our concern has been with how we can develop new capture technologies that can interleave more closely with the collaborative practices of interpretation during excavation. Here we sought explicitly to investigate how we could alter key parameters of the capture process through technologies that emphasised certain performance attributes at the expense of others. With our 3D scanning system then, the emphasis here was on the portability of a device that could be personally controlled by an excavator during their work and ongoing interpretive practice. Furthermore, the device focused on the '*quick and dirty',* placing particular emphasis on real time capture, review and manipulability of captured data in the context of interpretation, while compromising to an extent on the kinds of image resolution possible with the higher end scanning techniques. This streamlining of capture and excavation aims to be less disruptive to work practices and supports the continued development of interpretation.

In our initial evaluations of the technology, we were able to get some early indications of the ways that such a shift in the relative balance of performance trade-offs might play out – where they might offer some important benefits and where, on the other hand, they would not. Here, for example, we start to see the dependent nature of resolution requirements in a captured scan. For some concerns and some areas of archaeology, the lower resolutions work well - softer objects subject to deformation as a result of their immediate environment and handling are typically considered through their larger features, as exhibited by the examination of bones and ships. By contrast more durable materials such as stone are often examined for the minutiae of their form. But even here, in certain cases, lower resolutions were sometimes used to exaggerate features that remain hidden in more highly detailed scans. Perhaps of greater significance though is that lower resolution can bring certain benefits in terms of the dynamic manipulation of capture models. And it is in these manipulation possibilities that key features of interpretation can take place. Benefits too were to be found in the real time capture and review pointing to enhanced possibilities to monitor and review capture activity in the construction of particular interpretation narratives. This provides opportunity for frequent or repeated scans to re-capture in case of error or reframe to embody a different interpretation. Furthermore, there was a suggestion that by emphasising speed of capture, there would be greater freedom to capture more models. This would enable the documenting of degrading objects and the interpretive significance of these temporal concerns. It would also provide greater freedom to capture a greater selection of models each with different perspectives, foci and surrounding context.

In conclusion, through examination and observation of different aspects of archaeological practice we have identified key features of current capture technologies that do not sit neatly with ongoing interpretation and, more specifically, the drive towards 'reflexive archaeology.' As a result of this, we have investigated the key parameters of the capture process that could be altered to streamline the process with ongoing interpretation during excavation. We have presented a novel 3D capture device that emphasizes 'quick and dirty' capture and examined its use in a variety of settings. By working in the long tradition of ethnographic CSCW, we have shown that devices such as these, designed to support ongoing collaborative interpretation, could play an important role in archaeology in the future, providing a greater freedom to capture a wider range of data. We also believe that our work serves to highlight the longer term potential of 3D capture as a medium for digital collaboration across a range of domains.

REFERENCES

1. Allen, P., Feiner, S. Troccoli, A., Benko, H., Ishak, E., Smith, B., Seeing into the Past: Creating a 3D Modeling Pipeline for Archaeological Visualization., in *Proc. 3DPVT'04*, pp. 751-758, 2004.

2. Andrews, G., Barrett, J. C and Lewis, J. S. C., Interpretation not record: the Practice of Archaeology. *Antiquity*, 74 (285), pp. 525-530, 2000.

3. Balzani, M., Santopuoli, N., Grieco, A., Zaltron, N., Laser Scanner 3D Survey in Archaeological Field: the Forum of Pompeii, *International Conference on Remote Sensing Archaeology*, Beijing, October 2004.

4. Besl, P. and McKay, N., A Method for Registration of 3D Shapes, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14, pp. 239-256, 1992.

5. Brill, D., Video Recording as Part of The Critical Archaeological Process, in Hodder, I (ed), *Towards Reflexive Method in Archaeology: The Example at Çatalhöyük*, Cambridge: MacDonald Institute for Archaeological Research, 2000.

6. Carver, M., On Archaeological Value, in *Antiquity*, 70 (267), pp. 45-56, 1996.

7. Chadwick, A., Post-Processualism, prefessionalisation and archaeological methodologies. Towards reflective and radical practice, in *Archaeological Dialogues* 10 (1): 97-117, 2003.

8. Dellepiane, M., et al, Archaeological excavation monitoring using dense stereo matching techniques, in *Journal of Cultural Heritage*, 2012.

9. Gaffney, C., Detecting trends in the prediction of the buried past: a review of geophysical techniques in archaeology, in *Archaeometry*, 50 (2), pp. 313-336.

10. Goodwin, C., The semiotic body in its environment, in *Discourses of the body*, pp. 19-42, Macmillan.

11. Goodwin, C., Professional Vision, in *American Anthropologist*, 96, pp. 606-633, 1994.

12. Healey, J. and Picard, R.W., StartleCam: A Cybernetic Wearable Camera, in *Proc. Second International Symposium on Wearable Computing (ISWC'98)*.

13. Heath, C., Hindmarsh, J. and Luff, P., Video in Qualitative Research: Analysing social interaction in everyday life, Sage, 2010.

14. Hodder, I. The Archaeological Process: An Introduction, Blackwell Publishers, 1999.

15. Hodder, I. Archaeological Reflexivity and the 'Local' Voice, in *Anthropological Quarterly*, 76 (1), pp. 55-69.

16. Hodder, I. (ed), *Towards Reflexive Method in Archaeology: The Example at Çatalhöyük*, McDonald Institute for Archaeological Research / British Institute of Archaeology at Ankara Monograph, 28, 2000.

17. Hodder, I., Shanks, M., Alexandri, A., Buchli, V., Carman, J., Last, J., Lucas, G. *Interpreting Archaeology*, Routledge, London, 1995.

18. Hodges, S., Williams, L., Berry, E., Izadi, S., Srinivasan, J., Butler, A., Smyth, G., Kapur, N. and Wood, K., SenseCam: a Retrospective Memory Aid, in *Proc. Ubicomp 2006*, pp. 177-193, 2006.

19. Hori, T. and Aizawa, K., Context-based video retrieval system for the life-log applications, in *Proc. 5th ACM SIGMM international workshop on Multimedia information retrieval*, November 2003.

20. Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A. and Fitzgibbon, A., KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera, in *Proc. UIST 2011*, ACM.

21. Joukowsky, M. *A Complete Manual of Field Archaeology*. Prentice Hall, Englewood Cliffs, NJ, 1980.

22. Pece, F., Kautz, J. and Weyrich, T., Adapting Standard Video Codecs for Depth Streaming, in *Proc. of Joint Virtual Reality Conference of EuroVR (JVRC)*, pp. 1-8, Nottingham, UK, September 2011.

23. Lamming, M. and Flynn, M. Forget-me-not: Intimate Computing in Support of Human Memory, in *Proc. Friend 21*, 1994.

24. Levy, T. E., et al, On-site digital archaeology - new developments, delivery and the creation of a data avalanche, in M. Forte (ed.) *Introduction to cyber-archaeology*, pp. 135-153, Archaeopress, 2010.

25. Luff, P., Hindmarsh, J. and Heath, C., Workplace Studies: Recovering work practice and informing system design, Cambridge University Press, 2000.

26. Mann, S., Fung, J., Aimone, C., Sehgal, A. and Chen, D., Designing EyeTap Digital Eyeglasses for Continuous Lifelong Capture and Sharing of Personal Experiences, in *Adjunct Proc. alt.chi 2005*, ACM.

27. http://big.cs.bris.ac.uk/projects/mobile-kinect

28. Newcombe, R, Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A. and Fitzgibbon, A., KinectFusion: Real-Time Dense Surface Mapping and Tracking, in *Proc. ISMAR 2011*.

29. Newcombe, R., Lovegrove, S., Davison, A., DTAM: Dense Tracking and Mapping in Real-Time, *IEEE International Conference on Computer Vision*, 2011.

30. Novo, A., Grasmueck, M., Viggiano, D., Lorenzo, H., 3D GPR in Archaeology: What can be gained from dense Data Acquisition and Processing, in *International Conference on Ground Penetrating Radar,* Birmingham, UK, 2008.

31. Shackley, MS., An Introduction to X-Ray Fluorescence (XRF) Analysis in Archaeology, in *Geoarchaeology*, 2011.

32. Wheatley, D., *High-Dynamic Range Imaging for Archaeology*, 1996.

33. Wood, K., Fleck, R. and Williams, L., Playing with SenseCam, in *Proc. Playing with Sensors: Exploring the boundaries of sensing for playful ubiquitous computing (W3)*, UbiComp 2004, Nottingham, UK