ARCHFIELD: A DIGITAL APPLICATION FOR REAL-TIME ACQUISITION AND DISSEMINATION – FROM THE FIELD TO THE VIRTUAL MUSEUM

Neil Smith¹ and Thomas Levy²

¹Visual Computing Center, King Abdullah University of Science and Technology (KAUST), Saudi Arabia
²Department of Anthropology and Qualcomm Institute, University of California, San Diego, USA

Received: 30/11/2013
Accepted: 17/07/2014
Corresponding author: Neil Smith (neil.smith@kaust.edu.sa)

ABSTRACT
The lack of efficient digital data processing tools during field excavations is a major bottleneck affecting the delay between data collection and dissemination in archaeology. In this paper, we outline the fundamental methodology of ArchField, an integrated digital field recording solution developed to overcome this bottleneck and translate field excavations to virtual museums in real-time. ArchField records sub-centimetre accurate three-dimensional coordinates from Total Stations and RTK GPS units. Recorded field data and measured 3D coordinates are digitally processed to produce auto-generated daily GIS top plans. The processing pipeline enables the generation of publishable online maps from the first day of excavation to the last. It is interoperable with many different GIS viewers and stores data in an online PostGIS database. Digitization of archaeological data in the field is streamlined to facilitate standardization, redundancy and storage that can be immediately made accessible online to the digital community. Consequently, ArchField integrates features such as synchronization, data formatting, re-projection, dynamic labeling and symbolization. It provides immediate online accessibility of field excavations for virtual museums of the future. ArchField enables any archaeological project to inexpensively adopt real-time 3D digital recording techniques in their field methods.

KEYWORDS: GIS, real-time, southern Jordan, LiDAR, SfM.
1. INTRODUCTION

As the humanities moves towards a digital domain of data sharing and analysis, the long delay between field excavation and public dissemination becomes an increasing problem. A primary bottleneck is insufficient data processing, vetting, and database tools to manage the vast amounts of data as it is recovered on a daily basis from field excavation. Instead of reducing the workload of excavations, digitization without the proper data handling tools in place can easily become almost intractable in size and complexity (see Petrovic et al. 2011). The end result is a significant portion of time spent during post excavation to process data into a workable form that only then can be adequately analyzed and visualized. Errors made during digital field recording may not be caught until much later and at this point cannot always be easily resolved. Since archaeologists have only one chance to excavate an area and the only data that can be analyzed are what was properly recorded, it is essential that development of efficient and comprehensive digital recording and processing tools are developed.

In this paper, we present ArchField as a computational solution for efficiently recording, analyzing, and modeling the various sources of data recovered on a daily basis from field excavations. Fundamental to the methodology behind ArchField is a focus on automated digital processing to streamline and speed up the archaeological recording process while providing automatic and user-informed data vetting tools while still in the field. Archfield provides a unified software to combine high precision spatial recordings (Survey and LiDAR/SfM) with supervisor’s observations and digital spreadsheets. Integrated databases are seamlessly synced between the field excavations and lab analyses to enable raw data from the field to be immediately visualized as 3D top plans and queryable field reports in real-time. It is a digital application that serves as a bridge between field excavations and spatially oriented visualization and virtual presentation. ArchField enables any archaeological project to inexpensively adopt real-time 3D digital recording techniques in their field methods. It is field tested having undergone three excavation seasons of development and beta testing in southern Jordan (Smith and Levy 2012). We present its current developments and its interoperability with different archaeological recording schemas, excavation methodologies, measurement instruments and other 3D digital acquisition tools such as LiDAR and Structure-from-Motion (SfM).

2. RELATED WORK

Several joint computer science and archaeological projects have sought in the past to develop software to digitally process data after excavations have occurred. For example, DATARCH (Fabricatore and Cantone 2007) developed in 2006 functioned as an image management system where different media, primarily digital photos, could be stored and connected to a relational database. ArchaeoloGIS (Montesinos et al. 2010) and ETANA-GIS (Gortan et al. 2006) were designed as open-source GIS map servers that could take basemaps (generated in ArcGIS© from traditional excavation techniques) and database tables (recorded in Access© or other spreadsheet programs) and serve them on online. 3D-Murale is another application that stores data in SQL tables, plans to develop tools for on the field recording, and provide a full visualization package including the use of SfM (Green et al. 2002; Van Gool et al.
2002; http://dea.brunel.ac.uk/project/murale/). Unfortunately funding ended for the project in 2003 and only the database for Sagalassos, Greece was implemented. REVEAL an NSF funded computer vision project (Gay et al. 2010) for archaeology is a recording tool that combines plan reports with continuous video recording of excavations and multi-view camera captures of important artifacts and features. The future goal of the project is to orient surfaces and artifacts in 3D space using techniques such as multi-view stereo.

Similarly many projects have adopted off-the-shelf proprietary GIS software (e.g. ArcGIS) or open source GIS programs (e.g. MapInfo, GRASS, OSSIM, QGIS) to digitize their data and publish studies in scholarly journals or in online databases (e.g. Harrrower 2010; Al-Kheder et al. 2009; Ross et al. 2005). Others have used not only surveying tools but also LiDAR or SFM (e.g. Al-Kheder et al. 2009; Allen et al. 2004; Forte 2013; Pollefeys et al. 2003) to document excavations. However, what all these projects lack is software to facilitate integration of their data entry with survey instruments while still in the field.

The development of software to record detailed provenience and descriptor data in the field that directly communicates with high precision recording equipment has remained a project of the commercial sector. Two of the more well-known proprietary data entry software that archaeologists have used are Solofield developed by TDS and ArcPad developed by ESRI. Both of these programs have been developed for surveying and thus are not as easily adaptable to archaeological recording. The major drawback for archaeologists is that there is still little data processing occurring during excavations. The recorded results must still be downloaded, imported into GIS software and then manually processed to create top plans and final maps. They fall short of providing to the field of archaeology a fully implemented solution for high precision data recording, data organization, visualization, and analysis.

3. METHODOLOGY

ArchField builds upon a decade of previous methods of digital field recording (Levy and Smith 2007) and has now undergone several major revisions over the past three years as it has been thoroughly field tested in Southern Jordan. Its evaluation has led us to address in the software’s recording methods four main requirements of digital field excavation: 1) Precise survey instruments must seamlessly integrate with digital excavation data entry in the same application; 2) The software should intelligently facilitate the recording and storage of data with high redundancy, reduction of user intervention, standardization, and remote accessibility; 3) Various vetting tools, automated labelling, and real-time visualization of the on-going excavations must be provided to archaeologists while in the field; and 4) Diverse datasets such as 3D scanning and aerial mapping must be automatically integrated for visualization and digital dissemination.

3.1 Integration of High Precision Survey Tools with Archaeological Data Entry

The primary functionality of ArchField is to facilitate and streamline the procedures to properly record the provenience of artifacts and loci in 3D space and directly link them with their detailed field descriptions. Data entry for artifacts consists of 3D recording (x, y, and z) their unique locations using survey tools and simultaneously storing associated metadata (i.e. basket identifier, provenience, date, classification, description, etc.).

In contrast to most excavations that still rely upon imprecise survey recording methods such as tape measures and periodic dumpy level elevation readings to

---

1 For a detailed description of all the excavation methods applied please see Levy and Smith 2007.
2 Elevation readings can be precise up to 2.0mm when a digital auto level is used. However, not all excavations take a reading for every artifact but assign the elevation of the locus. The
plot artifact or loci onto graph paper, ArchField uses a Leica total station as the primary surveying tool. It takes precise 2.0mm + 2 ppm readings of artifacts and loci that are immediately stored and displayed on a digital Top Plan. When the user presses record on the Total Station, the measurements are directly read into ArchField. The raw (x,y,z) distance measurement is combined with the known position of the total station and a spatial reference system to project the 3D measurement into a coordinate that can be located on a map or in GIS software. The spatial reference system used to store the coordinates is Universal Transverse Mercator (UTM). Within the ArchField software it is possible to convert on-the-fly the coordinate into any other spatial reference system such as WGS84 Lat/Long as used by Google Earth. When the coordinate is received from the total station, the supervisors’ entered information is combined and stored together in ArchField’s database.

For artifacts, the most critical data entry is the provenience information that describes to which basket, locus, square, and area an artifact belongs. In our excavations the basket number (e.g. 5000...5001...5002) is a sequential number assigned to each artifact. The basket number becomes a unique identifier otherwise known as a ‘key’ in the relational database that can be used to retrieve all the tabular data associated with that artifact, including its coordinate location (see 3.2 for a description of the DBMS architecture and schema).

Loci are digitally recorded as three-dimensional polygons with the ability at every change in depth to create a new representative polygon and updates to the metadata. A locus is defined here as a distinguishable layer of soil deposit in which artifacts and other features are found. Loci are demarcated in excavations by the volumetric space of their depositional layer. A locus in three dimensions can be represented as a polyhedron, but typically it is drawn on graph paper as a boundary of the layer’s extent. In order to digitally record a locus, we use the total station to take multiple position readings along the physical boundary of the locus. The readings from the total station are automatically connected together in ArchField as vertices of a three-dimensional polygon. As its excavated depth increase we can generate a polyhedron to represent its three-dimensional nature (fig. 2). Each locus is assigned a unique number similar to the basket number assigned to artifacts. The extensive tabular data collected on a locus (e.g. sediment composition, density, types of artifacts, associated features, stratigraphic relationships, excavation strategies) are all linked to this unique locus number.

The supervisor is assisted in taking detailed information on each artifact or locus using ArchField’s data entry form. It allows the supervisor to enter all the pertinent information of the artifact or locus and have it automatically combined with its 3D location using the integrated total station surveying tool. To streamline the data entry interface and reduce possibilities of mis-entered data, the application is designed to auto-complete as much information as possible. Once the first artifact/locus is recorded the data entry form remains populated with information that does not change from one artifact to the next. After each artifact is recorded the

---

*Image* Figure 2 Loci extruded as color coded polyhedrons representing opening and closing elevations (visualized in ArtifactVis2).
basket number is incremented to reflect the following unique basket number. Drop down buttons are automatically populated with supervisors’ agreed upon descriptions of artifacts or locus characteristics.

The general descriptions of the artifacts remain standardized across multiple excavation areas and save the supervisor time from manually re-entering the same description for artifacts. The user only needs to change on a regular basis one or two fields saving time in the long run and preventing typical data entry mistakes. Whenever an artifact or locus is recorded a table appears below the data entry form showing all the pertinent features recorded. This serves as a final check that is used to confirm that all the information was entered correctly prior to moving on to the next recording and allows immediate edits to be carried out if mistakes are found. In this manner, ArchField streamlines the user’s entry of data and simplifies user-assisted correction.

Currently, three different versions of ArchField have been designed to address the evolving landscape of emerging computer hardware and operating systems. In 2010, ArchField was designed as a web based version using the combination of HTML, PHP, and Javascript languages (fig. 3).

Figure 3 ArchField web version Top Plan with OpenLayers integration.

The main advantage of this approach is that it can be run on any operating system with a web browser and the code can be easily changed without a need to recompile. The web based version has the most minimal hardware requirements. It can be deployed on an Atom or ARM based processor with 1GB of ram and only consumes 100mb of file storage. Essentially, it can be run on any computing device that can serve a webpage. This means the web based version can run on a tablet, smart phone, imbedded device, netbook or standard laptop.

Recently ArchField has been ported to run as a native iOS app (fig. 1) and is currently being rewritten to run as an os-independent C++ compiled version to enable more complex features on Windows based tablets (fig. 5). The advantage of ArchField as an os-independent GIS and data entry tool is that it can be deployed on any device rugged enough to be brought out to the field.

3.2 A Database Management System with high redundancy, standardization, and remote accessibility

Archfield integrates PostGIS a SQL relational database management system. PostGIS was chosen due to its broad adoption by various GIS applications (e.g. QGIS, GRASS, ArcGIS) and the Open GIS Consortium (OGC), its rich feature set of GIS functions, and ability to fully serve web based mapping systems. PostGIS provides a robust and efficient query system allowing asynchronous spatial queries across a shared network.

The DBMS architecture is designed to maintain a synchronized dataset across local and remote servers. As data is recorded in the field, it is stored on the field lab’s server computer and on a periodic basis synchronized with an online server hosted back at the university. ArchField handles the synchronization by merging the local and remote DBMS using SQL protocols and checks prior to any merger for conflicts. Any updates made during synchronization are recorded in a separate table with a timestamp allowing a simplified

---

3 With the field lab’s 3G internet connection, it is possible to synchronize all the databases in real-time, but in practice synchronization is conducted only on a periodic basis to save bandwidth.
method of version control to be implemented. This method allows continuous harmonization of the data stored locally and remotely over time. ArchField’s database is a distributed system where multiple users are able to access and change entries in the DMBS without the creation of divergent copies.

The DBMS schema organizes the spatial information and supervisor field entered data into two primary tables: one for artifacts and another for loci. Other tables generated from the analysis of the artifacts or loci are relationally joined to these primary tables (for a description of the complete DBMS schema see Gidding et al. 2013 and Smith et al. 2013). The tabular relationships of the DBMS allow ArchField to later perform SQL based queries across the entire DBMS and render the results as a 2D or 3D map. No matter whether in the field, in the dig lab, or back at the university the server database can provide supervisors and specialists real-time access to data being recorded in the excavations.

ArchField provides a project tailored setup to handle diverse recording methodologies of different archaeological sites. During the creation of a new excavation project in ArchField, users are given the option to setup what field descriptors, unique identifiers, and database attributes (columns) are pertinent to their recording of artifacts and loci. ArchField takes this information to generate a tailored database and data entry interface for the project.

3.3 In the field Vetting, Tagging, and Visualization

An essential aspect of the creation of ArchField is the ability to automatically process incoming data so that it can be viewed in real-time as digital top plans (figure 2). Dynamically changing data can be handled so that a Top plan’s symbology, labels and colors are auto-generated. ArchField is designed to constantly update auto-generated KML\(^4\) files as the PostGIS database receives new entries, changes, and undergoes time (top plans must reflect the current day of excavation rather than all artefacts or loci exposed over the whole season). The symbology (symbols to distinguish different types of artifact), color schemes (unique colors to differentiate soil types) and labeling (key metadata to assist the supervisor) imbedded in the KML is created on the fly whenever a new point is created.

The creation of dynamic KML files allows Archfield to be interoperable with most programs that support KML. These programs become real-time GIS spatial views of that day’s excavation as if several hours were spent preparing a top-plan after the day’s excavation had occurred. This is in contrast to a GIS program simply reading the spatial database, because imbedded in the KML file is not only the spatial and metadata but also specific instructions on how to render each individual artefact and locus.

The latest version of ArchField imbeds OpenLayers, an Open Source pure JavaScript library for mapping data in two-dimensions. Unlike Google Maps, OpenLayers does not require an internet connection to function but like Google Map, MSN Virtual Earth, Bing, etc. when there is internet connectivity it can stream all of the same imagery data. OpenLayers also enables full exploitation of the tablets’ and web based version’s multi-touch interfaces.

The digital top plan is designed to allow the field supervisor and registrar to visualize and vet their excavations as they would on a traditional paper top plan but with the added benefit of a full GIS. They are able to conduct queries, toggle layers, make quick edits, and zoom in on pertinent features. They see the current day’s top plan as it is

\(^4\) KML is an OSG approved vector format supported among other GIS programs (i.e. Google Earth, QGIS, GRASS, ArcGIS, etc). KML files created in ArchField can be opened in any of these other programs.
constructed and can immediately tell whether a locus or artifact was incorrectly recorded. The real-time top plans allow an archaeological project’s supervisor to catch mistakes while still in the field, being more in-tune with the current process of digital recording, and be freed up to focus on other aspects of documentation after a morning of excavation.

Having real-time and vetted top plans by the end of the excavation enables immediate analysis and a direct transition to final publishable excavation plans. ArchField’s ability to automatically curate spatial data into a comprehensive GIS makes it the essential backbone to later virtual museums and 3D visualization.

One other technique to reduce error and streamline field recording is ArchField’s ability to generate printed labels with barcodes for every basket (Figure 4). Once an artifact or basket elevation is recorded a button appears for label printing. When this button is clicked it pulls the information from the table and prints a label with all the important information on an encoded barcode. The printed labels save time for the registrar, eliminate the chance that they may write the wrong information, and prevent the label from being misread by others.

During laboratory hours, the label and barcode further reduce possible error in the whole process. The barcode stores a unique identifier enabling that specific entry to be located in the database. After a day’s excavation, the lab supervisor uses their barcode reader to quickly ‘check in’ all the artifacts and buckets collected that day from the field. In this way, before the data may be used by other lab specialists it is triple checked to make sure there is no error in the data. Finally, when a find is scanned it is recorded as received from the field and where it is currently stored (e.g. washing, conservation lab, photography lab, 3D scanning lab, pottery lab, storage, etc). When the sorted artifacts are moved to storage in a crate, the label is scanned again to update the database. By this manner, a detailed final list of where all artifacts are stored can be printed at the end of the excavations.

3.4 Integration of 3D Scanning, and Visualization

The application of LiDAR and Structure-from-Motion to archaeology has opened the avenue to easily capture 3D models of site architecture and stratigraphic levels. These techniques are integrated with ArchField to meet the goal of a total 3D documentation of the field excavations. The ability to scan a complete excavation in its full three-dimensions provides a context to visualize the three-dimensional artifact positions and locus boundaries recorded by total-stations and GPS.

In our recording methodology, ArchField is used to generate the ground control points to geo-reference 3D scans (e.g. LiDAR and SfM) so that the resulting point cloud models can be loaded into the same geographic space as the recorded artifacts and loci. SfM is employed to supplement the LiDAR scans where there are occluded or inaccessible areas and provide more frequent capture of locus surfaces and changes as the site is excavated (Figure 5). The initial geo-referenced LIDAR point-cloud is used as the reference for all future SfM recordings so that as new scans are produced they can be registered in proper position.

The latest version of ArchField has been rewritten in C++ and uses OpenSceneGraph to efficiently render these point cloud models onto the recorded Top plan. Although in its initial alpha stage, ArchField C++ is able to render our most dense SfM models and triangulated meshes on Windows Surface Tablets at full 60 frames per second (Figure 5). Every component of the data recorded are immediately availa-
ble for analysis and visualization both in ArchField and our 3D visualization software called ArtifactVis2 (see Figure 7; Smith et al. 2013).

These procedures enable the accurate digital reconstruction of the site’s architecture and significant artifacts. The intention of the reconstruction is not necessarily to show artistically how the site may have looked at one point in time but provide a faithful three-dimensional record of the excavations for on-going analysis. This has allowed us to continually return to the site examining its architecture, spatial distribution of artifacts and stratigraphic layers in a fully immersive 3D environment that is connected to the same GIS server that ArchField updates on a daily basis in the field. We have been able to show other archaeologists the excavations and discuss in detail various aspects of the excavation process and theories on its use from across the globe.

Figure 5 ArchField C++ with 3D point cloud of Khirbat al-Iraq excavations.

4. FIELD EVALUATIONS

From 2010 through 2012, ArchField was evaluated by site supervisors and staff at five sites in Southern Jordan dating from the Early Bronze to Islamic periods. In this paper, we focus on the most recent excavations at Khirbat Faynan in Southern Jordan. Evaluations of the software discussed below are based upon the application of the software in the field excavations and the written comments of the field supervisors.

In 2012, Khirbat Faynan one of the largest sites in lowland Edom with extensive occupation during the Iron Age, Roman, Byzantine and Islamic periods was excavated using ArchField. During this season, the iOS and improved web version were tested. In contrast to past seasons (2002-2008), where often the supervisors had to stay behind to complete recording artifacts at the site with the Total Station, ArchField allowed the supervisors to keep up with the excavation process and finish on time. Several evaluators specifically noted that the label printing in the field removed the tedious writing out of labels. In addition, top plans were complete at the end of the daily excavation, requiring ca. 10-15 minutes in the lab to synchronize the field database with the main lab server and print paper copies of the Top Plans. A general conclusion found in all the evaluations was that ArchField enabled a new level of efficiency in survey tool integration, top plan generation and lab workflow.

After testing the iOS version of ArchField the staff evaluations provided very useful suggestions of how to adapt and improve on the system in future versions. First, there was an agreement that the lighter weight of the iPod and the ability to use it in sunlight outweighed the benefits of the larger screen on the iPad. Smartphones and other handheld devices are preferable in field data entry. Second, although the purpose of the iPod version of ArchField was to enable the supervisors to become more mobile, the observed practice of the users was that they remained next to the total station and registrar table even though they were untethered.

Finally, evaluations found that a remaining limitation of ArchField was still a need to communicate with the Total Station and label printer through a laptop since neither could be supported directly through the iPad. ArchField requires a somewhat complex initial setup and involves many components (batteries, transformers, and RS232-to-USB converts) to integrate the label printer and total station which when

---

5 In the past (Levy and Smith 2007) an average of 1-3 hours was spent importing artifact and loci recordings into ArcGIS and manually generating the Top Plan.
not communicating properly introduced problems in the field-workflow. The main feature request by the supervisors was a completely wireless solution, with easy setup, and could communicate directly with surveying instruments and the label printer without the need of extra hardware, power sources, or cables.

4. CONCLUSIONS AND FUTURE WORK

The greatest outcome from using ArchField in the last several excavations is how it has allowed us to disseminate our data through scientific visualization with minimal post-processing after excavations. Every component of the data stored in ArchField has been immediately available for analysis and visualization in ArtifactVis2 (see Figure 6; Smith et al. 2013) and ArchaeoStor (Gidding et al. 2013). This has allowed us to continually return to the site examining its architecture, spatial distribution of artifacts and stratigraphic layers. It has enabled us in our own research to digitally present our ongoing research to a large audience of archaeologists on our individual computers, online and in virtual museums. ArchField is a computational solution to translate field excavations to virtual museums and catalogues in real-time.

Figure 6 ArchField artifacts and loci displayed in same geographic space as SFM and LiDAR scans (visualized in ArtifactVis2)

Over the three seasons of ArchField’s use, mobile and scanning technology has significantly changed. Future research will be directed towards reducing components, going fully wireless and integrating more extensively SFM into ArchField with daily recording of surfaces, loci, and in situ artifacts.

ACKNOWLEDGEMENT

This project was funded in part by a NEH Digital Humanities Start-up grant, NSF IGERT, and King Abdullah University of Science and Technology. We are grateful to the Department of Antiquities of Jordan, Dr. Mohammad Najjar, and the ELRAP/L2HE staff and students.

REFERENCES


