Cultural Heritage Omni-Stereo Panoramas for Immersive Cultural Analytics –From the Nile to the Hijaz

Neil G. Smith
Steve Cutchin
King Abdullah University of Science And Technology
neil.smith@kaust.edu.sa
Robert Kooima
Louisiana State University
kooima@csc.lsu.edu

Richard A. Ainsworth
Ainsworth & Partners, Inc.
ainsworth@qwerty.com

Daniel J. Sandin
University of Illinois at Chicago
dan@uic.edu

Jurgen Schulze
Andrew Prudhomme
Falko Kuester
Thomas E. Levy
Thomas A. DeFanti
University of California, San Diego
tdefanti@soe.ucsd.edu

Abstract— The digital imaging acquisition and visualization techniques described here provides a hyper-realistic stereoscopic spherical capture of cultural heritage sites. An automated dual-camera system is used to capture sufficient stereo digital images to cover a sphere or cylinder. The resulting stereo images are projected undistorted in VR systems providing an immersive virtual environment in which researchers can collaboratively study the important textural details of an excavation or historical site. This imaging technique complements existing technologies such as LiDAR or SfM providing more detailed textural information that can be used in conjunction for analysis and visualization. The advantages of this digital imaging technique for cultural heritage can be seen in its non-invasive and rapid capture of heritage sites for documentation, analysis, and immersive visualization. The technique is applied to several significant heritage sites in Luxor, Egypt and Saudi Arabia.

Keywords—spherical panoramas; visualization; gigapixel, virtual reality, stereo imagery; cultural heritage; Saudi Arabia, Taif, Luxor

I. INTRODUCTION

Digital imaging in Cultural Heritage has a problem of accurate reconstruction for analysis and a practical concern to provide visually compelling results for virtual and augmented visualization. Although digital acquisition techniques such as LiDAR scanning produce highly detailed and accurate measurements of cultural heritage sites, the generated point clouds fail to render the same color depth and textural fidelity as high resolution digital imagery. As gigapixel resolution 3D visualization systems become common, the possibility of immersing viewers within cultural heritage sites with stereo 20/20 vision becomes possible. This accentuates the shortcomings of digitally measured point cloud or triangulated datasets. The digital imaging acquisition and visualization techniques described here provide hyper-realistic stereoscopic 360° x 180° 1.8 gigapixel capture of cultural heritage sites. The resultant equirectangular digital image projections are mapped to the displays of VR systems providing an immersive virtual environment in which researchers can collaboratively study the important textural details of an excavation or historical site with the same fidelity as when the images were captured. The dual-camera image capture system was developed to explore the combination of high-resolution digital imagery and stereo panoramas with a variety of VR displays [1]. We found one of its most successful applications has been in cultural heritage documentation and representation.

We use the CalVR visualization framework. This framework is compatible with a wide variety of immersive and non-immersive, 3D and 2D display environments, from desktop computers to CAVE environments such as the StarCAVE or NexCAVE [2, 3, 4]. These environments provide the cyberinfrastructure for interactive visualization of massive cultural datasets on scalable, high resolution stereoscopic systems. We call the data imaging acquisition and visualization technique presented here CAVEcam capture.

The goal of the CAVEcam for cultural heritage documentation is to complement other acquisition technologies such as LiDAR or SfM by providing more detailed textural information that can be used in conjunction for analysis and visualization. It facilitates archival, visualization, and analysis of cultural heritage artwork with the fidelity required for cultural analytics to be performed [5]. Users are able to collaboratively view, interrogate, correlate and manipulate generated imagery at a wide range of scales and resolutions. They are able to analyze architectural construction, diagnose artwork, assess surface damage/erosion of cultural heritage monuments, verify the accuracy of professional renditions of inscriptions, wall paintings, mosaics, and increase their visual capacity to intuitively recognize deep patterns and structures, perceive changes and reoccurrences and locate themselves within time and space of cultural heritage sites. In a broader scope, the CAVEcam provides compelling virtual representations of artwork and cultural heritage for museums, exhibitions, and other related public domains.

II. RELATED WORK

Our technique brings together several different aspects of related work on panoramic gigapixel image capture, omni-stereoscopic panoramas, and immersive visualization. The initial documented pipeline for generating gigapixel panoramas [6] and a streamlined online visualization of these panoramas [7] has led to new advances in digital image acquisition and visualization. Recent work has focused on augmenting or

imbedding other data for visual analytics such as video in the gigapixel images [8] or generating time-lapsed gigapixel video [9]. The availability of competitively priced camera rotation mounts such as Gigapan™ and stitching software suites has led to the explosion of gigapixel imagery on online venues by amateur and professional photographers. Gigapixel digital image capture has directly impacted documentation of cultural heritage sites and has become an integrated acquisition technique in many projects [see 10, 11, 12].

Related to the advances in ultra-high resolution panoramic acquisition and motorized camera mounts is the development of automatic omni-directional stereo cylindrical panoramas [13,14,15], and stereo spherical panoramas[16]. These 3D panoramas provide a new level of stereographic immersion in digital image visualization. In particular, [16] generates spherical panoramic stereoscopic images using a single camera with a fish-eye lens rotated along a distance from the camera's nodal point. The advantage of this technique is the automated stitching process that produces from a single camera the stereo pair of a spherical panorama. The stereo panorama is visualized using a spherical stereo display such as the iDome [16].

In the domain of cultural heritage documentation, [17, 18] use mono GigaPan generated panoramas for photogrammetric evaluation and texturing. Reference [19] integrates tripod based panoramic capture in conjunction with range finders to create background skydomes to provide greater context to the reconstruction of an excavation. In line with the visualization goals of our technique, [12] renders acquired omni-stereographic panoramas of the Place-Hampi cultural heritage site in a cylindrical stereoscopic projected display system called the Advanced Visualization Interactive Environment (AVIE).

The CAVEcam builds upon these related works, in particular, by combining the multi-gigapixel resolution achieved through the use of a GigaPan and stereoscopic capture of spherical panoramas. Our method is not limited to specific types of VR display systems and provides an efficient solution for rendering gigapixel size stereo panoramas in diverse environments. Our methods are applied in the field to address the specific challenges of cultural heritage documentation.

III. METHODOLOGY

A. Acquisition

A fully automated dual-camera system is used to capture sufficient stereo digital images to cover a complete sphere or rectangle (Fig. 1). A GigaPan EPIC Pro Robotic Controller is automates the capture process. This programmability allows repeat image capture and the option for HDR (high dynamic range) processing and other options. This robotic system can be programmed to accommodate any number of images, matching the configuration of VR and other display technologies. Currently we mount on the GigaPan two Panasonic LUMIX® GF-1 cameras that offer the full feature set, resolution, capability and flexibility of the best digital SLR systems while maintaining a small profile. These cameras offer several essential features, including 12.1-megapixel

resolution and side-by-side mounting at the distance needed for stereo separation. In this configuration, the interocular distance is 70mm. The Ainsworth CC-1 Dual-Camera Controller was developed specifically for this application, and provides an interface between the robotic unit and the dual cameras.[1] This unit is necessary because the Panasonic GF-1 camera exposure and focus are activated by varying current levels, rather than by simple switch closure. The controller accepts triggering output from the GigaPan Epic Pro system and supplies synchronous current pulses to the cameras.

Fig. 1. The dual-camera system is fully automated and can capture any number of stereo images covering a complete sphere or rectangle. Both nadir and zenith can be included, with the exception of a small footprint below the tripod (Photo courtesy of Dick Ainsworth).

The amount of inherent distortion created by our dual-camera system is a direct function of the camera offset, or the distance from the zero parallax point to the center of rotation. Limiting this offset to approximately 35mm or less also limits the distortion effects of this technique to what can be compensated for later in the stitching process, provided that objects are no closer than approximately five feet from the point of rotation. Offsetting each camera by this amount provides a consistent 70mm interocular distance at all points in the resulting sphere. For the spherical imaging with a focal length of 40mm (35mm equivalent) and 360° x 180° capture, the stereo panorama typically requires 150 images and runs 1.8 gigapixels. After blending with the stitching software, the two final spherical images provide 800 megapixels of stereo information. For rectangular displays, the focal length is extended to 70mm (35mm equivalent), creating rectilinear or cylindrical display images covering 120° x 60° at 267 megapixels per eye (see [1]).

Multiple acquisitions sessions are taken at every capture position. Since the GigaPan controller provides precise repeated rotations across the span of the total image capture, multiple sessions allow alternative images to be selected for each of the ca. 75 positions recorded for each eye. The option to select which image to use for each position allows elimination of individuals moving through the scene (Fig. 2), correction of unnecessary shadows, and selection of the best depth of field (focus), f-stop and aperture, during the sorting phase (see below).

---

[1] The CC-1 controller and complete CAVEcam system are available from Ainsworth & Partners, Inc.

Fig. 2. Multiple capture sessions are run in the same position when there is moving traffic. In this figure, the selection of images from sets of scans generates a clean spherical panorama without unwanted moving/ghosting objects. (TourCAVE at KACST; images courtesy of Dick Ainsworth)

B. Stitiching and Software

Panorama stitching software is designed specifically to overcome minor misalignments when combining multiple images (Fig. 3). This blending process distorts images slightly so that adjacent photographs connect in a seamless manner. This built-in correction capability of stitching programs offers an opportunity to create seamless stereo images from image capture techniques that incorporate small amounts of inherent distortion characteristics. This makes it possible to rotate a camera pair about a central axis, creating two overlapping panoramas that preserve accurate stereo separation.

Fig. 3. Fifteen columns by five rows of images are required to cover the full 360° horizontal and 180° vertical field of view. The composite image results in 75 individual photographs. (Photo courtesy of Dick Ainsworth)

Software converts the two image arrays created by the dual-camera system into two matching stereo views. This software needs to be able to organize the images into collections that match the image capture, adjust individual exposure, white balance, stitch the individual images to create a panorama matching the projection of the intended display, and make final adjustments. There are many comparable stitching software options to perform this task including Microsoft Research ICE stitcher, PTGui and Hugin (http://hugin.sourceforge.net/). Our current workflow includes a combination of several software suites, such as Adobe Bridge® and Adobe Lightroom®, PTGui Pro® and Photo Shop®.

It was found that for stereo panoramas initial sorting and processing of images prior to stitching generates better final results, since disparities in color, focus and orientation are more noticeable when viewed in stereo and can create eye strain. Adobe Bridge or similar software can be used for viewing a large number of images, selecting the correct sequence if several exposures were shot, and matching corresponding left and right pairs. Lightroom is used to make individual adjustments in light level, white balance, and contrast. Arranging the size of the thumbnail images to correspond to the width of the panorama gives a realistic view.

The PTGui Pro® software allows multiple copies to be operated in parallel. This permits easily switching between left and right eye views. After initial alignment, the *Panorama Editor* is selected for each image. Alternating between the two views allows precise adjustment of the roll, pitch, and yaw variables via the *Numerical Transform Editor.* Vertical alignment of the two views is critical, and can be adjusted to within 0.1 degree with this method. The yaw parameter is adjusted to align the picture plane at the preferred distance from the camera. The time required for creating the final equirectangular or cylindrical projected panorama depends on the resolution selected and the capability of the computer used.

C. Visualization

The stereo panoramas are cube mapped onto a cylinder or a sphere in the VR environment (Fig. 4). The radius from the viewer to the image is set at approximately infinity, or at the object of major interest and undergoes the normal VR perspective projection for the left and right eye. Disparities captured in the panorama closer than infinity are effectively subtracted from the disparity at infinity. The two spherical images are displaced horizontally by the interocular distance, as seen from the viewer's perspective. Displacement equal to the interocular distance is maintained perpendicular to a plane aligned with the viewer's direction and perpendicular to a line between the eyes. This presents good stereo separation in whatever direction the viewer looks, even up and down. Objects at infinity will move with the viewer, contributing to an immersive experience. Moreover, the projection into these large immersive environments allows normal perspective to be restored and all straight lines will appear normal when viewed in the VR environment.

Fig. 4. Jurgen Schulze in the Calti2 TourCAVE showing mono version of a 360° CAVEcam stereo photograph of Luxor, Egypt. (Photo by Tom DeFanti)

Once the stereo images are properly stitched and aligned, they are preprocessed into optimized multi-paged TIFF images using software developed alongside the CalVR plugin PanoViewLOD. The equirectangular projected panorama is stored as a spherical cube map. A spherical linear interpolation is used to equalize the solid angle subtended by each image pixel when mapping the texture to a sphere mesh in CalVR. Each face of the spherical cube map is subdivided into a quad-tree resembling a mipmap hierarchy and stored in the multi-paged TIFF file. Spherical cube map pages are enumerated in the TIFF file in breadth-first order, which gathers low-resolution base imagery to the front of the file and provides increasing resolution with increasing file length. To enable seamless linear magnification filtering across this discontinuous image representation, each page is stored with a 1-pixel border. A spherical cube map image of depth d will contain 2^{2d+3} - 2 separate pages. The current renderer has a maximum depth of 7. Given a spherical cube map with page size s and depth d, the effective resolution of the equivalent equirectangular projection is $4 \cdot s \cdot 2^d \times 2 \cdot s \cdot 2^d$. In general, a base page size of 512 and depth of 4 pages (32768 pixel width x 16384 pixel height) is sufficient to accommodate the resolution of the images currently generated from the CAVEcam.

The software renderer provides threaded loading of multi-paged TIFF and a texture cache using OpenGL pixel buffers. This enables rapid seamless retrieval of the image data as it is streamed into the environment while preventing overloading gpu memory. Only currently visible regions of the sphere are loaded into the renderer and cached when no longer viewed. It is possible to stream much higher resolution panoramas and quickly switch between sets of stereo panoramas with no delay in load times.

The visualization application allows users to select the radius of the cylinder that the image is to be displayed on. A good value for this radius is 30 found through experimental tests in the multiple cave environments. The horizontal and vertical angle width of the panorama image is dependent on how the panorama was shot. Angles and radius allow the sphere or cylinder to be fully defined for proper rendering.

The generated sphere or cylinder is positioned such that its center is in the user's head position, while its height is user adjustable. As the user interacts with the scene, head tracking is used to adjust the frustrum and maintain proper parallax and perspective.

IV. APPLICATION TO CULTURAL HERITAGE

Application of the CAVEcam to Cultural Heritage presents unique challenges when seeking to archive, analyze and visualize the acquired stereo gigapixel panoramas. Several of these challenges include overcoming limited accessibility to hard to reach or inaccessible areas, poor lighting conditions, graffiti, congested tourist areas, occlusion, and the sheer physical expanse of many cultural heritage sites.

The CAVEcam was first extensively applied to cultural heritage in Luxor, Egypt. On the eve of the Egyptian revolution permission was granted to test the technology at the temple complexes in Luxor (Figs. 5-6). Due to the impending revolution, very few tourists were in Luxor. This allowed capture with minimal interference. The results currently provide the highest resolution image capture of this area where the hieroglyphics can be read with the same clarity as viewing them in person. However, this experiment was too short for a systematic documentation of the extensive area of Luxor. The Luxor monuments are both structural and artistic works of art that encode through hieroglyphics the history of Egyptian life, culture, and events. Therefore, the archival process facilitates not only the user's ability to re-visit these sites within a virtual environment but also to examine the rich textural detail for further epigraphic study, diagnostics, and future restoration.

Fig. 5. Plain Projected CAVEcam imagery of Luxor, Egypt showing full 360 spherical recording (Photo by Tom DeFanti).

Fig. 6. Hallway in Luxor showing straight walls, appearance of depth, and reliefs with high fidelity; despite the large bezels and differing LCD orientation the image remains properly aligned. (KAUST NexCAVE, images courtesy of Tom DeFanti)

A second expedition was conducted in Taif and Mahd ad-Dahab, Saudi Arabia in 2012. Several new techniques were added to address specific challenges that arose in the past. First, the same areas were captured from multiple positions allowing fulle coverage and users to step through the site. Second, at the Late Islamic Taif Fortress graffiti covered the site. By using standard image editing the graffiti could be removed with minimal damage to the stereo effect (Fig. 7). Third, multiple bracketed exposures were taken in the same stations to compensate for poor lighting conditions. This technique was especially important during the capture of the Al-Samlagi Dam, Taif which at ca. 200m long and 30m high cast a considerable shadow affecting lighting conditions (Fig. 8). Fourth multiple sessions on the same positions allowed the removal of individuals from capture.
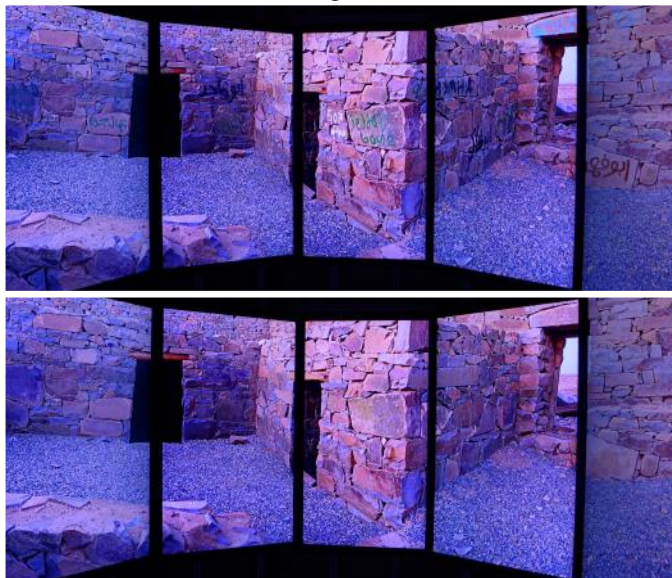

Fig. 7. Before and after results of removal of graffiti from spherical stereo panoramas. (KACST TourCAVE, images courtesy of Dick Ainsworth)

Fig. 8. Before and after results of accommodating complex lighting situations for cultural heritage using multiple sessions of bracketed photographs.

Imagery of pre-Islamic Al-Samlagi Dam, Taif looking west along the expanse of the ancient dam. (KACST TourCAVE, courtesy of Tom DeFanti)

During the 2012 expedition, every acquisition area was also captured using the Structure-from-Motion technique to generate dense three-dimensional point clouds (Fig. 9). Using GPS acquired ground control points these point clouds are geo-referenced and scaled to their real world locations. In turn, the SfM point cloud model and input images enable the center point of the spherical stereo panorama to be located for visualization within the VR environment. This method allows multiple CAVEcam acquired panoramas to be imbedded within the spatial context of the cultural heritage site. Visualization users are then able to switch back and forth between CAVEcam imagery and the SfM model. The spatial context of the point clouds enables users to fly to different areas of the cultural heritage site quickly and 'jump' back into the CAVEcam imagery acquired at that position.

The advantages of this system for cultural heritage documentation can be seen in its non-invasive and rapid capture of such areas. Within a short period of exploration in the Kingdom of Saudi Arabia we acquired imagery of sites that can now be made available for world heritage. This allows individuals to view sites that they may never gain physical access to. The recent techniques discussed here provide many benefits to implementation in conjunction with other techniques such as LiDAR and SfM. The also provides an increased flexibility in how cultural heritage sites, especially complex and texturally rich sites can be captured and visualized using digital imaging techniques.

Fig. 9. Transition back and forth between CAVEcam (Top) and SfM pointcloud (Bottom). Imagery is from a pre-Islamic well north of the Al-Samlagi Dam, Taif. The panorama is in proper orientation and scale to SfM. The CAVEcam captures further distances and more accurate color rendition than possible with LiDAR and SfM—these contribute to a greater immersive feel in VR. (KAUST NexCAVE, top courtesy of Thomas DeFanti, lower courtesy of Dan Sandin and Neil Smith)

V. Conclusions and Future Work

The digital image processing techniques presented here provides a compelling 3D experience with high-resolution photographic realism for cultural heritage sites. These stereo panoramas have been shown publicly to hundreds of viewers in a variety of VR environments. The content provides texturally rich datasets that can be utilized by cultural historians, archaeologists, epigraphers and others for documentation and analysis. In these terms, we have sought to meet the goals of cultural analytics to create a visualization interface to collaboratively explore, interrogate, and correlate these hyper-realistic immersive digital captures of cultural heritage. The equipment, methodology, and results are accessible and repeatable to a broad audience which we hope will facilitate further adoption by researchers and practitioners working in the area of digital imaging in cultural heritage.

Systematic coverage of entire cultural heritage areas requires further development of the technique in planning, acquisition and processing. In future work, we hope to provide seamless movement between areas without loss of fidelity. A major step in this process will involve further integration of the gigapixel resolution textures with LiDAR and SFM, transitioning between the two different VR techniques and texturing the generated meshes [see 17,18]. Finally, we hope to broaden accessibility of acquired imagery through online servers and mobile applications designed to augment viewing of cultural heritage sites in situ.

References

[1] R. A. Ainsworth, D. J. Sandin, J. P. Schulze, A. Prudhomme, T. A. DeFanti, and M. Srinivasan, "Acquisition of stereo panoramas for display in VR environments," *Proc. SPIE* 7864, Three-Dimensional Imaging, Interaction, and Measurement, 786416, January 27, 2011.

[2] C. Cruz-Neira, D.J. Sandin, T.A. DeFanti, R. Kenyon, and J.C. Hart, "The CAVE: audio visual experience automatic virtual environment," Communications of the ACM, 35(6), 1992, pp. 64-72.

[3] T.A. DeFanti, G. Dawe, D. J. Sandin, J. P. Schulze, P. Otto, J Girado, et al., "The StarCAVE, a third-generation CAVE and virtual reality OptIPortal," Future Generation Computer Systems, 25(2), 2009, pp. 169-178.

[4] J. P. Schulze, A. Prudhomme, P. Weber, and T. A. DeFanti, "CalVR: an advanced open source virtual reality software framework," Proc. of IS&T/SPIE Electronic Imaging, The Engineering Reality of Virtual Reality 2012, San Francisco, CA, February 4, 2013.

[5] S. Yamaoka, L. Manovich, J. Douglass, and F. Kuester, "Cultural analytics in large scale visualization environments," IEEE Computer, cover feature for the special issue on computers and the arts, 2012.

[6] J. Kopf, M. Uyttendaele, O. Deussen, and M. F. Cohen, "Capturing and viewing gigapixel images," ACM SIGGRAPH 2007 papers *(SIGGRAPH '07)*, New York, NY, USA, Article 93, 2007.

[7] S. E. Chen, "QuickTime VR: an image-based approach to virtual environment navigation," Proceedings of the 22nd annual conference on Computer graphics and interactive techniques (SIGGRAPH '95), Susan G. Mair and Robert Cook (Eds.). ACM, New York, NY, USA, 1995, pp. 29-38.

[8] S. Pirk, M. F. Cohen, O. Deussen, M. Uyttendaele, and J. Kopf, "Video enhanced gigapixel panoramas," SIGGRAPH Asia 2012 Technical Briefs (SA '12). ACM, New York, NY, USA, Article 7 , 2012.

[9] R. Sargent, C. Bartley, P. Dille, J. Keller, I. Nourbakhsh, and R. LeGrand, "Timelapse gigapan: capturing, sharing, and exploring timelapse gigapixel imagery," Fine International Conference on Gigapixel Imaging for Science, 2010.

[10] K. Kwiatek and M. Woolner, "Embedding interactive storytelling within still and video panoramas for cultural heritage sites," *15th International Conference on Virtual Systems and Multimedia, 2009. VSMM '09*, 9-12 Sept. 2009, pp. 197,202.

[11] Z. Bilá and K. Pavelka, "Possible use of GigaPan for documenting cultural heritage sites," XXIIIrd International CIPA Symposium, Prague (Czech Republic), Sep. 2011.

[12] S. Kenderine, "The irreducible ensemble: place-Hampi," Proceedings of the 13th international conference on Virtual systems and multimedia (VSMM'07), Theodor G. Wyeld, Sarah Kenderdine, and Michael Docherty (Eds.), Springer-Verlag, Berlin, Heidelberg, 2007, pp. 58-72.

[13] H.C. Huang and Y.P. Hung, "Panoramic stereo imaging system with automatic disparity warping and seaming," Graphical Models and Image Processing, vol. 60, no. 3, May 1998, pp. 196–208.

[14] S. Peleg, and M. Ben-Ezra, "Stereo panorama with a single camera," IEEE Conference on Computer Vision and Pattern Recognition, pp 395-401, Ft. Collins, Colorado, June 1999.

[15] M. Ben-Ezra, Y. Pritch, and S. Peleg, "Omnistereo: panoramic stereo," Imagine, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 3, pp. 279-290, March 2001.

[16] P. Bourke, "Capturing omni-directional stereoscopic spherical projections with a single camera," 2010 16th International Conference on Virtual Systems and Multimedia (VSMM), Oct. 2010, pp.179,183, 20-23.

[17] E. d'Annibale, "Image based modeling from spherical photogrammetry and structure for motion, The case of the Treasury, Nabatean Architecture in Petra," XXIIIrd International CIPA Symposium, Prague (Czech Republic), Sep. 2011.

[18] C. Pisa, F. Zeppa, and G. Fangi," Spherical photogrammetry for cultural heritage—San Galgano Abbey and the Roman Theater, Sabratha," J. Comput. Cult. Herit. 4, 3, Article 9 (December 2011), 15 pages.

[19] P. Allen, S. Feiner, A. Troccoli, H. Benko, E. Ishak, and B. Smith, "Seeing into the past: creating a 3D modeling pipeline for archaeological visualization," Proc. International Symposium on 3D Data Processing Visualization and Transmission 3DPVT, 2004, pp. 751-758.