Quasi–Mixed Reality in Digital Cultural Heritage. Combining 3D Reconstructions with Real Structures on Location—The Case of Ancient Phalasarna

Gunnar Liestøl and Elpida Hadjidaki

Abstract Traditional Mixed Reality as it is currently experienced on popular mobile devices has its obvious limitations in the context of Cultural Heritage. Neither the sensor fusion approach nor the pattern recognition solutions are precise or stable enough to provide a satisfactory visual match between the live video feed and the graphical layer of digital information. The fundamental incompatibility between 2D live video and dynamic 3D graphics also makes this a short–lived solution in a long–range perspective. While we are waiting for sustainable solutions for real time 3D capture and display on mobile and wearable devices it is pertinent to employ and evaluate transitional alternatives for effective use on location . In the research and development reported here we created a static 3D version of the current archaeological site based on photogrammetry. This is done in order to test how it may serve as an intermediate level of representation for increased precision when combining the real present with the reconstructed past. In this chapter we present and discuss the experiences we have gained exploring this type of Indirect Augmented Reality, which we have named 'Quasi–Mixed Reality', on location at the archaeological site of Ancient Phalasarna on western Crete.

Keywords Quasi–mixed reality · Indirect augmented reality · Mixed reality · Situated simulations · Sitsim · Ancient Phalasarna · Photogrammetry

1 Introduction

After a decade of mixed and augmented reality applications available for smartphones and tablets we still have no enduring and satisfactory solution for precise and close coupling of the real and the virtual when visiting cultural heritage sites. Neverthe-

G. Liestøl (✉)
Department of Media & Communication, University of Oslo, Oslo, Norway
e-mail: gunnar.liestol@media.uio.no

E. Hadjidaki
Association for Ancient Phalasarna, Kisamos, Phalasarna, Crete, Greece
e-mail: hadjidaki@gmail.com

H. Kremers (ed.), *Digital Cultural Heritage*,
https://doi.org/10.1007/978-3-030-15200-0_29

less, games like Monster Buster and Pokémon GO, as well as applications for finding restaurants, tube stations and much more, are in abundant supply, if not always in wide–spread popular use. Such applications normally combines a live video feed from the phone's physical camera with an added digital 3D graphics layer. These services may work well in pointing out approximate direction and distance to a place of interest in the vicinity of the user. However, in the context of cultural heritage and on site dissemination and explanation their value and usability are more limited.

Archaeological remains and artifacts may be difficult to understand for the untrained eye when exploring the residual structures of an ancient city or settlement on location. While cultural remnants may be relatively easy to distinguish from the natural environment, it will often be completely impossible for the visiting layman to understand their original shape, use and significance. Positioned physical means, like information boards and plaques, may be helpful in providing explanations and interpretations close to the object, but they can never furnish the object with rich and deep contextual information and knowledge without cluttering the environment and thus disturb the cultural experience one initially intends to improve.

The challenge and goal of close coupling in the mixing of digital depiction and physical reality will most probably find sustainable solutions in near future employments of real time 3D capturing and display [1]. Only then can live digital representation of the real environment be fully integrated with and augmented by additional digital representations of various kinds. Meanwhile, we must look elsewhere for interim solutions, and regarding these we should aim to achieve two things: both solve the immediate challenge for better precision, as well as prepare for the forthcoming interfaces, conventions and potential of Augmented and Mixed Reality based on real time 3D capture and display.

In the current situation we may distinguish between two main types of augmented and mixed reality applications on popular mobile devices (smartphones and tablets):

(1) computer vision–based registration/pattern recognition, and
(2) sensor–based registration (aka 'sensor fusion').

The first solution requires a visual mark the pattern recognition software can interpret in order to place the digital 3D object in the right position on the screen and in accordance with the video information (we should here also include emerging camera–based solutions like Apple's ARKit and Google's ARCore). In the second solution, the case of 'sensor fusion', the hardware sensors (GPS, magnetometer, accelerometer and gyroscope) work together to determine the position, orientation and movement of the mobile device, so that the graphics layer can be placed on top of the video feed as precisely as possible [2].

Both approaches have their limitations. With the sensor fusion method it is difficult to attain the graphical match one would require in a cultural heritage context; for example, if you have an ancient monument, such as the base of a building, and you want to add the shape and look of the original structure on top of the current remains (as presented by the video information at the foot of the screen) the graphical layer representing the reconstruction will not attach itself accurately to the (video representation) of the visible structures. Instead it will continue to move as the sensors struggle to calculate and update position, orientation and movement of the device.

The computer vision–based registration functions under other conditions. It is more stable in the alignment of the video images with the 3D graphics objects, but it is dependent upon a continuous visual connection between the fiducial markers (or other forms of visual pattern) and the video camera. If this connection is broken, for example by the interference of other visitors or lack of light, the positioning collapses. In addition, it is neither very practical nor attractive for a cultural heritage site or an exhibition room to place a number of fiducial markers on valuable antique objects or spread out over a large archaeological site.

These methods have room for improvement, and accuracy is continually rectified. Interesting experiments using live 360° video panoramas mixed with point cloud models have shown good accuracy in AEC settings, both indoor and outdoor [3]. Unfortunately, such solutions still require specialized hardware and do not have the kind of mobility and availability we are concerned with in this context, that is: what can be achieved by means of off–the–shelf hardware, primarily and for the time being, smartphones and tablets, and eventually smartglasses.

Regardsless, there is also a limitation of a more fundamental character involved. The mixing of video and graphics produces a profound problem of compatibility: video is 2D and digital graphics in this context is 3D. Allthough this problem can be partly compensated for, the two levels of representation can never be fully integrated [4]. As a consequence we have looked to, and practiced, another form of mobile AR, know in general as Indirect Augmented Reality (IAR) [5].

2 Indirect Augmented Reality

The notion of 'Indirect Augmented Reality' was launched as an ultimate vision of a live representation of the physical world providing increased complexity by "… a large array of video cameras and other sensors that would capture, in real time, the real environment and permit a perfect reconstruction of that environment, in real time, as seen from any arbitrary viewpoint." [5], similar to our anticipated live 3D capture mentioned above. However, IAR has often been associated with still image–based panorama solutions [6], as static image panoramas [5], or as dynamic environments but with constrained movement between predefined positions [7] (Fig. 1).

In our own experiments with situated simulations (sitsim), as a form of Indirect Augmented Reality [6], we have established the mix (in mixed reality) as a relationship between the full screen 3D graphics environment and the real world outside the device. The advantage with this variant of the sensor fusion method is that the lack of positioning precision is not a pressing problem. This is due to the fact that the graphical congruence is not a question of direct match between objects (and levels) on the screen, but a correspondence between the user's visual perception of the real environment and the user's visual perspective into the 3D graphics environment on the screen. This is a solution that has proved itself well in a series of applications for

Fig. 1 To the left an IAR situated simulation from the '*Roman Forum*' AR–application: The user is physically facing the Capitoline hill and the ruins of the rostra (speaker's platform). On the screen one can see Mark Anthony giving his eulogy for the assassinated Julius Caesar. According to Appian a wax copy of Caesar's beaten body was erected on a pole. The image to the right is from the '*Omaha Beach*' AR–application: The user is physically facing east along the beach. On the screen one can see surviving soldiers crossing the upper part of the tidal flat in the early morning on D–Day before finding temporary cover behind the shingle bank on top of the beach. In both these examples the information on the screen (perspective etc.) is updated continually as the user move. Simultaneously the digital environment include animations to represent actions and events

various cultural heritage sites: the Roman Forum, Acropolis in Athens, the Oseberg Viking ship grave mound in Norway, D–Day on Omaha Beach and more [8–11].

However, a problem arises when you have a reconstruction of an ancient site with considerable visual difference between the current archaeological structures and remains and the reconstructions based on the archaeological evidence. This is the case with parts of the Ancient Phalasarna site on the western tip of Crete.

3 Ancient Phalasarna—History, Excavations and Traditional Dissemination

In 333 BC Ancient Phalasarna was turned into a unique fortified harbour built with Persian funding to be used as the main naval base on Crete against Alexander the Great's forces. From here Spartans, Persians and Greeks hoped to retake the Greek islands and Asia Minor from the Macedonians, cut off Alexander's supply routes, and prevent his attack against the great King Darius of Persia. However, Alexander defeated this master plan. Nevertheless, the city was left with a magnificent military port that harboured the Phalasarnian warships for centuries to come. The ancient city was finally destroyed by the Romans in 69 BC. By that time the inhabitants had turned to piracy and were attacked by Roman forces under the command of general Metellus 'Creticus'.

The harbor region of Ancient Phalasarna has been excavated since 1986 [12, 13]. Notable finds include six fortification towers, fortification walls, quays that ringed the harbor, a channel that provided exit to the sea, a winery, a bath house, a road and a merchant's house. Prior to the development and publication of the app described in this chapter the site has been disseminated by means of various modes and media: The first excavations put Phalasarna as an archaeological site into tourist guides and maps for West Crete. It has been the subject of television documentaries (in Greek), articles in newspapers and popular periodicals (in Greek), and has a website. Brochures in Greek and English are available for distribution, and there is a limited amount of signage at the site. Many artifacts are displayed in local museums. Tourists can visit the site without charge during opening hours, and groups frequently request tours from the excavator.

4 Ancient Phalasarna and Quasi–Mixed Reality—The Circular Tower

In 2010 we developed a simple sitsim prototype reconstructing Ancient Phalasarna as of 333 BC. In 2013 this prototype was extended to also include 69 BC and the Roman attack and destruction of the city. These implementations of Indirect Augmented Reality simulations worked well in showing extensive reconstructions and the city as a whole, but they did little to explain the meaning and function of the majority of structures visible on the site today. The visual and figurative difference was just too substantial between the past and the present. The large fortification wall on the eastern slope of the acropolis is relatively easily identifiable as such. However, it is difficult for the visiting layman to understand most of the exposed structures on the site. Even the remaining base of the excavated circular tower on the south east side of the harbour fortifications, despite its characteristic shape, is demanding for visitors to imagine in its original shape (Fig. 2).

Fig. 2 Most people manage to distinguish between natural and cultural structures in the terrain, but it may be difficult to imagine what these remains represent. In the left photo we see a part of the quay and fortified harbour, which today is positioned 6.5 m above ancient sea level. In the photo to the right is the exposed base of the circular tower located on the south–east corner of the Ancient Phalasarna harbor fortifications. How may monuments like these best be reconstructed by means of Augmented Reality (AR)?

It was obviously need for a middle ground, a mode of representation where the current architectural remains could be provided with partial digital reconstructions without loosing the context of the current conditions. As mentioned above, we had already opted out the Mixed Reality solution, while the loose coupling of our preferred Indirect Augmented Reality approach favored large digital reconstructions. How could we better, and with greater detail and precision connect the physical structure of today with a digital version of the missing parts so that visitors could experience less comprehensible parts of the site as whole objects unifying the physical remains with digital augmentations?

The solution was to create a digital 3D version of the archaeological site as it looks today, including both its untouched and excavated areas as well as the surrounding terrain, all based on photogrammetry. This was conducted in two ways: The general area was photographed from a drone at different altitudes, and the excavated parts were captured manually by ground photography. Accordingly, the general site and surrounding area were modelled in relatively low resolution, while a higher density of triangles was earmarked for the excavated and exposed portions (Fig. 3).

The 3D terrain model then records and represents the shape and look of the Ancient Phalasarna harbor installations at the time of the photo sessions (2015–16). With a 3D–model of the current terrain in place it was possible to add reconstructions immediately onto the digital capture of the current state of the sites surface, for example with the circular defense tower on the south–eastern side and the harbour. The tower is 9 m in diameter and is today preserved to a height of 4.5 m. An impressive feature that is noteworthy on this tower is the carved molding above its base. With the *'quasi–mixed reality'* method it was now possible to show, with high accuracy, the past state of this building in matching combination with its present condition [14] (Figs. 4, 5, 6) .

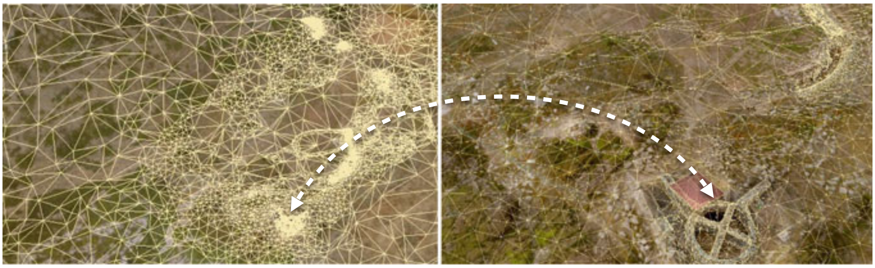

Fig. 3 Different resolution in different parts of the terrain model. In the image to the left one can see the higher density of triangles in certain parts of the site's central area. The image to the right shows a close up of partly the same terrain. The circular tower (at the arrow tips in both pictures) and the excavated quay have a much higher concentration of triangles than the nearby areas, which in this context are of less significance for the purpose of the AR situated simulation

Fig. 4 In the left photo the user is positioned just south–west of the circular tower. In the right photo we see the user's view of the real environment and the corresponding 3D-model on the screen. Note the more limited cut in the real photo than on the screen

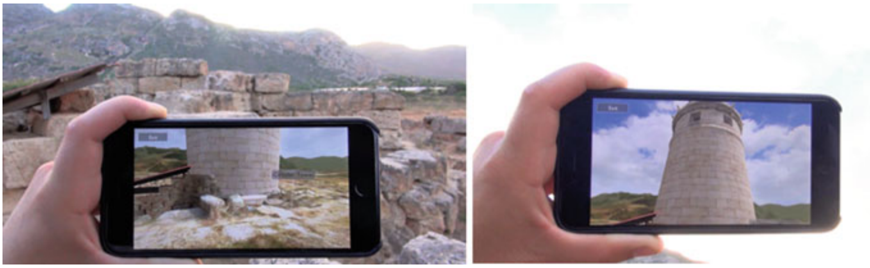

Fig. 5 In the left photo the reconstructed tower appears on top of the base. In the photo to the right the user is tilting the camera (phone) upwards and viewing the top of the tall tower, which in ancient times probably was about 17 m high

Fig. 6 The full sized tower is placed on the digital copy of the real terrain. The sign post in front of the edifice is the spatially positioned hypertext link that activates the graphical reconstruction and its corresponding voice–over. Note the seamless joint between the remaining base and the reconstruction

5 Limits of Quasi–Mixed Reality and Partial Reconstruction—A Ship in a Pond!

Encouraged by the experience with Quasi–Mixed Reality and the partial reconstruction of the circular tower we tried the same approach with another structure: the quay with its characteristic mooring rings. Since the sea level today is 6.5 lower than in ancient times due to an earthquake in western Crete in 365 AD it was important to bring the excavated quay in contact with the sea. We thus moored one of the modelled ships to the 3D–modelled quay and added water around the ship to indicate the ancient sea level and the relationship between quay, mooring, vessel and water. The effect was strikingly informative because at a micro level it brought explanatory context to the excavated structures.

However, it also had a side effect—an almost comic one—which we had not anticipated. The partial reconstruction which worked so well with the tower now, in a wider perspective, had the opposite effect. The limited reconstruction of the ancient sea level in which the moored trireme floated created a visual paradox: it made it look like the opposite of the open connecting ocean—it rather had the appearance of a self–contradiction: a seagoing vessel in a tiny pond! (Fig. 7)

Capturing the real environment of an archeological or cultural heritage site to create a one–frame 3D model of the current surface and terrain represents a promising way of combining existing artefacts and ruins with digital reconstructions. The precision achieved is far superior to the Mixed Reality solutions combining 2D video and 3D graphics. However, the current surrounding is not recorded and represented in real time, so it requires continual updating and revisioning, for example as the excavations unfold and expose new structures.

Our implementation of this Quasi–Mixed Reality mode of combining the real and the virtual—present and past—also shows that it needs to be done with caution. Partial reconstruction proved to be effective in the example with the circular tower,

Fig. 7 In the left photo a partial reconstruction with a ship moored to the quay floating in and surrounded by the water level from ancient times. Although it provides explanatory context by adding the ancient sea level in a limited area the partial reconstruction has a slightly unrealistic side effect: the ship is perceived as floating in a pond without connection to the sea. While the full reconstruction of the harbour as it may have looked in 333 BC, in the wider perspective, provides a much more informative context; the position of the quay in the harbour and its connection to the sea

where the added digital reconstruction completed the existing structure with seamless mending and gave an informative impression of the whole. While on the other hand the quay structure, which was not broken or destroyed, when adding a partial context provided mixed results. In Quasi–Mixed Reality designs it is thus important to understand the different meanings and functions involved in the use of partial and total reconstructions.

The Ancient Phalasarna situated simulation AR app was recently updated and is now connected to a tracking application that records user sessions every time the application is activated and used on location. Feedback from this tracking will be important for future updates of the system.

Acknowledgements Thanks to Tomas Stenarson founder of CodeGrind AB, Šarūnas Ledas cofounder of Tag of Joy, as well as the students at MEVIT3810, University of Oslo, for making this version of the Phalasarna sitsim possible.

References

1. Dai, A., Niesser, M., Zollhöfer, M., Izadi, S., Theobalt, C. (2017) ‘BundleFusion: Real–Time Globally Consistent 3D Reconstruction Using ON–the–Fly Surface Reintegration’ in *ACM Transactions on Graphics*, Vol. 36, No. 3, Article 24, pp. 24:1–18. Publication date: May 2017.
2. Nurminen, A., Järvi, J., and Lehtonen, M. (2014) ‘A Mixed Reality Interface for Real Time Tracked Public Transportation’, paper presented at the 10th ITS European Congress, Helsinki, Finland 16–19 June, (2014).
3. Côté, S., Trudel, P., Desbiens, M. A., Giguère, M. & Snyder, R. (2013). ‘Live mobile panoramic high accuracy augmented reality for engineering and construction.’ In: N. Dawood and M. Kassem (Eds.), *Proceedings of the 13th International Conference on Construction Applications of Virtual Reality*, 30–31 October 2013, London, UK.
4. MacIntyre, B., Lohse, M., Bolter, J. D. & Moreno, E. (2001) ‘Ghosts in the Machine: Integrating 2D Video Actors into a 3D AR System’ in *Proceedings of the International Symposium on Mixed Reality*, pp 189–202.
5. Wither, J., Tsai, J.-T. and Azuma, R. (2011) ‘Indirect augmented reality’, *Computer and Graphics*, 35(4): 810–822.
6. Bolter, J. D., Engberg, M. & MacIntyre, B. (2013.) Media studies, mobile augmented reality, and interaction design. *Interactions* 20, 1 (January 2013), 36–45. DOI=http://dx.doi.org/10.1145/2405716.2405726.
7. Madsen, J. B. & Madsen, C. (2013). ‘An interactive visualization of the past using a situated simulation approach.’ in *Proceeding of Digital Heritage Congress,* pp. 307–314.
8. Liestøl, G. & Morrison, A. (2013) ‘Views, Alignment and Incongruity in Indirect Augmented Reality’ in Proceedings of ISMAR 2013, *IEEE International Symposium on Mixed and Augmented Reality - Arts, Media and Humanities*. 1–4 Oct. 2013, pp 23–28.
9. Liestøl, G., Rasmussen, T. and Stenarson, T. (2011) ‘Mobile Innovation: Designing & Evaluating Situated Simulations’ in *Digital Creativity* 2011, Vol. 22, No. 3, pp. 172–184. Abingdon: Routledge, Taylor & Francis Group (2011).
10. Liestøl, G. (2014) ‘Along the Appian Way. Storytelling and Memory across Time and Space in Mobile Augmented Reality’, in M. Ioannides et al. (Eds.) *Progress in Cultural Heritage: Documentation, Preservation, and Protection*. Berlin Heidelberg: Springer Verlag, pp 248–257.
11. Liestøl, G. (2018) ‘Storytelling with mobile augmented reality on Omaha Beach: Design considerations when reconstructing an historical event in situ.’ MW18: MW 2018. Published February 16, 2018. Consulted April 28, 2018. https://mw18.mwconf.org/paper/

storytelling-with-mobile-augmented-reality-on-omaha-beach-design-considerations-when-reconstructing-an-historical-event-in-situ/.
12. Hadjidaki, E. (1988). 'Preliminary Report of Excavations at the Harbor of Phalasarna in West Crete'. *American Journal of Archaeology, 92*(4), 463–479. https://doi.org/10.2307/505244.
13. Hadjidaki, E. (2001) The Roman destruction of Phalasarna. In Higham, N.J. (ed.) *Archeology of The Roman Empire, A tribute to the life and works of Professor Barri Jones*. BAR International Series, vol. 940, pp. 155–166.
14. A video of the system in use is available at. https://youtu.be/lphW5YbQaEo.