A critical review: Emergent and best practice of location specific cultural heritage technologies.

Carl Smith
LTRI

Abstract

This literature review will report on the current state of location based technology including mobile augmented reality (MAR) and GPS to inform the EPSRC and English Heritage funded project Representing Re-Formation. The focus will be on its application for use within cultural heritage as an educational and outreach tool. The key questions and areas to be investigated include: What are the requirements for effective digital intervention within the cultural heritage sector? What are the affordances of mixed and augmented reality? What mobile technology is currently being utilised to explore cultural heritage? What are the key projects? The review will conclude with an overview of the main issues to consider when developing location specific cultural heritage technologies alongside some broad design guidelines.

Key words**:** Augmented and mixed reality, interaction design patterns, design guidelines, contextual design research.

Introduction

Traditionally museums and site specific learning environments offer static experiences to their visitors which leave less opportunity for alternative interpretation of those exhibits. According to Cheok (2012), advances in mobile technology are moving us from an era of 'information communication' to 'experience communication'. As a result the role of cultural heritage and cultural institutions is no longer just to exhibit significant objects but to create augmented contexts which consist of rich, interactive and engaging experiences for visitors. Traditional Virtual Reality (VR) creates a world of its own which is usually totally indifferent to the physical context; in contrast MAR (Mobile Augmented Reality) is intrinsically tied to the local environment.

One of the unique affordances of well designed MAR is that it should enable the creation of situations and concepts that could not have been realised with just the physical or just the virtual elements of the context because it unites the strengths, features and possibilities of both. This literature review is concerned with highlighting best practise in how to formulate these situations and concepts. A good example of this is an engineer working with digital overlays to detect potential problems with an engine. See video:
http://learning.londonmet.ac.uk/TLTC/carl/bmw.mp4

What are the requirements for effective digital intervention within the cultural heritage sector?

The central challenge for educational designers is to create contexts that promote effective learning. With the generation and increasing adoption of mobile augmented reality (MAR) and mixed reality techniques we now have the potential to explode the form and complexity of these learning contexts. The core question of this review is

can we develop augmented contexts that are more effective because they take advantage of the affordances of these mixed reality methods and techniques. The majority of mobile learning research and mobile app development creates experiences which tie all the requirements of the user's attention down to and onto a four inch screen. This includes the majority of MAR applications. To avoid this, new interfaces must be created that take advantage of the physical and digital affordances of each learning situation.

Gallagher (2010) defines cultural heritage as being concerned with collections of physical structures and the intangible values that they project about the culture in which they are situated. He believes MAR has the potential to augment these heritage contexts bi-directionally: "Traditionally, cultural heritage studies has explored physical structures as stable entities and the intangible values as contextually fluid; augmented reality attacks this traditional structure and demonstrates that the physical structures themselves, along with the values that they accompany, are in states of constant flux. This flux is interpreted, mediated, and reconstructed in the individual learner."

New forms of contextual representation and engineering can now do real-time interactive justice to the complexity of both the form and function of cultural heritage. However, there are some key issues to consider when deciding on a technological solution. Boyer and Marcus (2011) state that an unfortunate fact of most augmented reality applications is that screenshots of an application give a better impression of the functionality than the actual use. In addition, without thoughtful design digital interventions risk distracting visitors from meaningful engagement with the cultural objects they are actually designed to augment.

What are the affordances of mixed and augmented reality?

In order to achieve the aim of designing and supporting learning across physical and virtual space we need to combine the affordances of the physical with the affordances of the digital. This gives us an opportunity to reinvest value back into the full thick description of physical site specific space and at the same time ensure we are using embodied experience (and not just vision as is common in most AR) to interact with these spaces. In Kevin Slavins mobile Monday presentation (2010) he discusses the importance of peripheral vision in learning situations by illustrating how reality is not actually communicated via a single focus. Reality is the whole world around us and not just what is in front of us. As a result MAR can often make things seem *less* real. Reality is only augmented when it feels different and not just when it looks different.

Those areas which can benefit from the affordances of mixed reality are those that require greater interactivity in the learning process. A well designed augmented environment should allow the users to regularly ask their own questions while analysing the associations between pieces of information, rather than just isolated facts. Therefore the issues of what information should be presented, in what way, and in what order are essential. Arguably any application of this technology can only be justified if the user is able to interact with the object in ways which would not previously have been possible. This alternative form of interaction should significantly alter the user's relationship with the object.

In order to realise the potential of MAR or mixed reality it is necessary to appreciate the power it provides over traditional methods:

- The mixed reality environment promotes opportunities for the exploration, alteration and manipulation of complex data sets.
- Mixed reality allows the user to compare and contrast objects from a variety of disparate sources within a unified environment and create a representation of the original context for objects and structures which have been removed from their context.
- The augmented model implies an association of information with space. Mixed reality can recreate and reanimate culturally determined points of view which are essential for a more complete understanding of any structure's meaning.
- Mixed reality can provide a multiple point of view of the same object at different times or at different levels of conceptual analysis. The augmented reconstruction can also provide answers to questions of the relevancy of data and highlight discrepancies or inconsistencies in existing data.
- Mixed reality offers layers of information that otherwise could not be displayed due to the aesthetic and space concerns (Thian 2012)

What mobile and location based technology is currently being utilised within cultural heritage applications?

Marcus (2011) notes that the majority of location based AR applications fall into two categories: GPS-based and computer vision-based. Oomen, J. et al (2011) expand on these categories by creating clusters of location-based applications used by GLAMs (Galleries, Libraries, Archives, Museums):

CLUSTER	FEATURE
Location-aware display of content	Mobile applications that use the GPS function to determine the place of the device and display content connected to that location.
Contributing content by end-users	Applications that allow users to contribute content that is linked to a certain place. This can include texts, photos, video, audio.
QR codes	Using QR codes to connect the physical space and related online content.
Browsing using an augmented reality application	The use of a third party augmented reality browser (Layar, Wikitude or Junaio) to display content linked to a geographical location.
Location-based games and Geocaching	The gameplay of a location-based game somehow evolves and progresses with a player's location. Geocaching is the most prominent example with a large community. Typical is a single-player kind of treasure hunt which is usually played using hand-held GPS receivers with user-hidden boxes.

Clusters of location-based applications used by GLAMs

For the purposes of this review a number of projects from each of these categories and clusters will be discussed. An overview of the advantages and disadvantages of the GPS-based and computer vision-based solutions are summarised here:

GPS based solutions:

Advantages:
(i) Cheap and easy to produce (GPS is free and does not require 3G or wireless connectivity)
(ii) Processor friendly and can make additional use of a phone's accelerometer, gyroscope, and other technology to determine the location, heading, and direction of the phone.
(iii) Can be used to identify zones which when entered trigger locally stored media (reducing cost and download times.
(iv) The content can be easily updateable with minimal technical skill.
(v) Commonly available on many phones.

Disadvantages:
(i) Not reliable indoors - imprecise location data, difficulty discerning the heading of the phone - a serious drawback when hoping to create photo overlays.
(ii) Does not easily supply true context sensitivity although there are workarounds for this.

Computer vision: 2D marker and 3D object based solutions:

Computer vision based applications use powerful computer vision libraries to help the computer identify what it is seeing through a digital camera (Marcus, 2011). In many cases, this is done through the creation of a unique 2D symbol that the computer identifies and then uses to launch different forms of interactive media.

However as cultural heritage sites often do not allow obtrusive physical markers to be placed in situ it is important to include the full range of computer vision solutions within this review.

It is now becoming possible to use 2d photographs and 3d objects (of content at the actual site) as markers. A good example of this is 'The Augmented City' from metaio: (http://www.youtube.com/watch?v=ACu6rehFXBM&feature=player_embedded
To achieve the aim of triggering both 2D and 3D content via the actual physical context the following methodology can be adopted: SLAM (simultaneous location and mapping) (http://en.wikipedia.org/wiki/Simultaneous_localization_and_mapping) can be used to find feature points in the scene via the video stream and detect the ones that are consistent from frame-to-frame in order to create a 3D map of points. This map of points will then be used as a tracking model for inserting the content. The most common engine is PTAM (parallel tracking and mapping).

There are a number of natural feature tracking or 3D object tracking libraries which have been made available via the Qualcomm AR SDK (http://developer.qualcomm.com/dev/augmented-reality) and the Metaio Unifeye SDK (http://www.metaio.com/software/sdk/). Also Metaio Junaio Glue (http://www.augmentedplanet.com/2010/07/metaio-change-the-game-with-junaio-glue) and Layar vision (http://devsupport.layar.com/entries/20331366-get-started-with-layar-vision) both allow you to register and track an image on mobile devices.

Advantages:
(i) Greater opportunity for context sensitivity.

Disadvantages:
(i) More expensive.
(ii) Large amounts of processing power are required meaning only the newest smartphones are capable of using this method.
(iii) Displaying high-quality AR content requires high frame rates and extremely responsive tracking in varying light levels.
(iv) AR Markers + QR codes can distract attention away from the actual heritage objects under investigation.

Key projects

The ARCHEOGUIDE project (http://archeoguide.intranet.gr/project.htm) provides alternative approaches for accessing information at cultural heritage sites using augmented reality, 3D-visualization, mobile computing, and multi-modal interaction techniques. Particular emphasis in existing case studies is given to the virtual reconstruction of heritage sites.

Streetmuseum is an app that has been released by the Museum of London (www.museumoflondon.org.uk/streetmuseum.htm) which is an augmented reality application that provides access to hundreds of historic images of London as both 2D and 3D overlays. This system also allows users to create their own trails around London.

The British Museum have also released a project called 'the passport to the afterlife' (http://bit.ly/dLSwtH) which has successfully introduced marker based AR into museum galleries. The trail runs in the permanent Egyptian galleries. It reveals insights about the appropriate use of AR for children, the potential of mobile AR for kinesthetic learning and the importance of blending AR interaction with analog techniques to ensure a rewarding learning experience. (Mannion, 2012)

The *Philly*History.org (http://phillyhistory.org/PhotoArchive) project has an augmented reality application which allows users to point their phone at nearby buildings or streets in order to discover historic photos of that location. The application can pull information into the local environment from over 93,000 images in the growing database. All archive photos may be searched by proximity to an Address, Intersection, Place Name, or Neighbourhood. The project team have considered how their research findings and technology experiments can be applied to the larger community of cultural institutions and as a result they plan to release the *Philly*History.org augmented reality application to the public.

A recent mobile app called 'Kew' (http://www.kew.org/visit-kew-gardens/visit-information/garden-guides/mobile-app/index.htm) released by Kew gardens aims to provide a seamless experience through a combination of a GPS enabled customizable map, a QR code reader and an augmented reality browser. Each of these technologies was used for different purposes at different locations, for instance the GPS detection technology was used for general orientation, the QR code reader was used to allow visitors to scan certain pre determined plants and the augmented reality browser was used to prompt discovery and encourage meandering exploration.

Another core intention behind this technologically seamless approach was to ensure that visitors were not wedded to the phone visually but that the app would act on the periphery, as a catalyst for serendipitous learning (Mann, 2012). When evaluating the impact of the app the project team found that the GPS and QR code were more widely used by the visitors than the AR browser which was due to a lack of familiarity.

The Cistercian in Yorkshire archaeological learning environment (Smith, 2009) consisting of the reconstruction of 5 Cistercian abbeys in Yorkshire as part of the CONTSENS project (http://www.ericsson.com/ericsson/corpinfo/programs/using_wireless_technologies_for_context_sensitive_education_and_training) was designed to allow the user to explore the interaction between the virtual and the physical (the reconstructions and the ruins) in order to create a rich research context. The central hypothesis behind the creation of the resource was that the user can learn a great deal more about a specific building or style of architecture if they are able to interact with the material in ways that is not possible, either on site, or via traditional print media. A comprehensive database of visualisations contained (but did not prescribe) the entire set of reasoning that led from the design of the 2D plans to the 3D reconstructions. The process of data-mining hierarchies of evidence (intellectual transparency) in the quest for constituent parts, key narratives and evolutions of form can be explored in depth. This declarative design was essential as the total range of knowledge contained within the application could then be utilised, allowing many diverse and wide-ranging opinions to be tested and unified for further analysis. See a video for a brief synopsis of the project: http://learning.londonmet.ac.uk/TLTC/carl/cistercians_in_yorkshire.mov

Another case study of the CONTSENS project also created a MAR system to support the training of urban design professionals. The project enabled these urban planners to examine past and present representations of school architecture in situ to see how its organisation and restructuring related to educational discourse. The intention was for learners to examine the community from the past, in order to engage, understand, and inform the present, as urban space and society is made and remade. A core mechanism of the project was to use participatory design to allow these urban designers to understand that the physical design of any educational institution has a direct impact on the delivery of the education that takes place and that learning space design shapes our behaviour and influences our thinking. A database of architectural designs and 3D models of pedagogically effective schools was available via the mobile devices to be used and adapted within the MAR system. See a video summary of this case study here: http://learning.londonmet.ac.uk/TLTC/carl/urban_planning_and_education.mov

ClayVision (http://www.youtube.com/watch?v=bULl71EfEG4) reengineers the design conventions of Augmented Reality. ClayVision uses computer vision and image processing techniques to dynamically transform the appearance of building structures in real time. The digital data is not overlaid in the usual 'bubble metaphor' of traditional AR but actually becomes 'fused together with the urban environment'. This is achieved through the use of two techniques referred to as 'default texture' and 'diminished background'. This project is an example of computer vision based localisation for mobile devices which does not require the use of markers. The basic idea of their localization technique is simple: each frame of the real-time video feed is

compared to a collection of photos, shot from the same location using the same device beforehand (Takeuchi & Perlin, 2012).

Conclusion

Bell (2010) asks whether learning designers of location based and mixed reality experiences are following old museum models rather than exploring augmented reality as a unique new form of interpretation. AR (especially vision-based AR) however is still in its infancy but has the potential to supply the kinds of interactions that initiate sense-making, construction of personal and social meaning, dialogue, and emotional responses (Thian, 2012). The impact of AR technology is still under researched and as a result visitor experiences, or the social and educational interaction that is shaped by its characteristics and objects involved are not well documented.

A crucial decision that needs to be made by the project team is whether mobile devices will be supplied on site or whether users will be required to use their own devices. The BYOD (bring your own device) movement is much greater than it was even a year ago but in reality the actual cost of developing for multiple mobile platforms may prove more expensive than actually supplying a set of devices.

The main utopian power of MAR is that we are capable of creating a new immersive reality completely beyond our known limits, and that it can be embedded not in a blog, a device or a computer, but in the world (Baraona, 2012). A good example of this new immersive reality is being developed in Japan by Professor Michitaka Hirose and goes by the name of "diet goggles" (http://www.youtube.com/watch?v=spk-2EuZ3hk). The goggles are designed to trick dieters into eating less by digitally enlarging food whilst they are actually eating it. This is an example of context engineering and highlights the extent to which MAR can subvert our perception of the physical world. The context becomes available for the learner to 'get to everything, add to everything, keep track of everything, and tie everything together (Waltham, 1989).

The inception app uses MAR (http://inceptiontheapp.com) with sound by synchronizing an augmented layer of music and ambient noise with the world around you. The app manages to augment context by forcing the user to get into a real world context that is producing a certain sound. Unless that sound is present then the content is not released by the 'listening' app. The app actually acts as a trigger or catalyst for real world activity. This allows the inception 'environment' to become a world where a lot of other stories can take place. This is in stark contrast to the usual predefined and prescribed content of traditional non context dependent apps (which also, invariably do not tap into existing analogue skill sets).

Some potential design guidelines could offer advice on the following: What is the most suitable digital content for this context? How should it displayed? What is the most appropriate interaction method? For how long will the augmentation be present? This final question is important for learning because most instances of visualisation are used to act as scaffolding which is designed to be removed to ensure that the learning has been internalised. Is there also scope in the intervention for users to adapt and manipulate the content elements of the mixed reality system? Have the essential properties of the context been identified, analysed and utilized?

The manipulation of context is a core function of any mixed reality system. Designing these systems extends the scope of the context to involve the use of physical space and other objects and people in the user's physical surroundings, not necessarily relevant for design previously. The following are general design guidelines that may be referenced whilst designing the digital intervention:

1) The Digital Intervention must Justify its Existence: Has the designer provided evidence that the intervention allows the user to engage with the situation on many levels that were not previously possible? Is there perhaps another method which does not require MAR that could achieve the same result?

2) Take Advantage of Real World Interaction: In order for users to be engaged usability needs to be intuitive. It should be easy for the user to learn how to use a system without prior training. Users should be able to accomplish a task with a minimum of interaction steps. In addition the system should provide direct and immediate feedback so that the user can adjust their actions accordingly.

3) Context Modelling: Designing MAR systems extends the scope of the context to involve the use of physical space and other objects and persons in the user's physical surroundings, not necessarily relevant for design previously. The manipulation of context is therefore a core function of MAR. Have the essential properties of the context been identified, analysed and utilized?

4) Participatory Design: Is there scope in the digital intervention for users to adapt the content elements of the MAR system?

5) Technology Narrative: What is the ratio of push vs pull within the system in regard to content acquisition? Does the digital intervention allow the user to remain active in their learning process?

6) Effective Choice of Media: Does the choice of media within the MAR system work well for the chosen target audience?

7) Multiple Ways of Seeing: Has the MAR environment incorporated multiple ways of seeing? Whether utilising microscopic, xray or macroscopic all these augmented ways of seeing are available within MAR systems and should be taken advantage of if appropriate.

8) Design for Social Interaction: Has the designer incorporated the use of real time social data into the system?

9) Build in Adaptability: Has the MAR designer created an appropriate level of challenge neither too easy nor too difficult which will keep the learner stimulated and motivated? Can the user themselves decide on the level of difficulty involved within the system?

10) Interdisciplinary Media Production: Creating MAR experiences draws on skills from a broad range of creative disciplines including scriptwriting, sound design, interaction design, location sourcing, production management, software development

and testing. Although some of these disciplines may not always be called upon it is important for the design team to be aware of them.

References

Baraona, E. (2012). From line to hyperreality. Domus.
http://www.domusweb.it/en/architecture/from-line-to-hyperreality

Bell, T. (2010). Literature Review of Technology Used in Interpretation of Sensitive Cultural Heritage Sites. University of Oregon.
http://traceyjbell.com/Literature%20Review%20Tracey%20Bell.pdf

Boyer, D. and J. Marcus (2012) Implementing Mobile Augmented Reality Applications for Cultural Institutions . In J. Trant and D. Bearman (eds). Museums and the Web 2011: Proceedings. Toronto: Archives & Museum Informatics. Published March 31, 2011. Consulted July 2, 2012.
http://conference.archimuse.com/mw2011/papers/implementing_mobile_augmented_reality_applications

Cheok, A. (2012). Keynote multi modal sensory human. ICALT.
http://www.adriancheok.info/post/26135544227/icalt2012-keynote-talk-multi-modal-sensory-human

Gallagher,S.M. (2010). The Flaneur Was Here: Mobile Augmented Reality and Urban Cultural Heritage Learning in Lower Manhattan. Elearning and Digital Cultures. University of Edinburgh.
http://michaelgallagher.files.wordpress.com/2010/12/ededcfinalassignment1.pdf

Hirose, M. (2012). Future Of weight loss diet goggles.
http://www.youtube.com/watch?v=spk-2EuZ3hk

Mannion, S. (2012) Beyond Cool: Making Mobile Augmented Reality Work for Museum Education. The British Museum, London
http://www.museumsandtheweb.com/mw2012/programs/beyond_cool_making_mobile_augmented_reality_

Mann, C. (2012) A study of the iPhone app at Kew Gardens: Improving the visitor experience. The Susie Fisher Group. Electronic Visualisation and the Arts (EVA 2012) London, UK, 10 - 12 July 2012. http://ewic.bcs.org/category/17061#1

Oomen, J. et al., (2011) Picture War Monuments: Creating an Open Source Location-Based Mobile Platform. In J. Trant and D. Bearman (eds). Museums and the Web 2011: Proceedings. Toronto: Archives & Museum Informatics. Published March 31, 2011. Consulted July 3, 2012.
http://conference.archimuse.com/mw2011/papers/picture_war_monuments_creating_open_source_location_based_mobile_platform

Slavin, K. (2010) Mobile Monday presentation http://www.slideshare.net/momoams/kevin-slavin-reality-is-plenty-thanks

Smith (2009). WP5 Art Gallery and Museum Education. Evaluation Report. CONTSENS: Using Wireless Technologies for Context Sensitive Education and Training.
http://www.ericsson.com/ericsson/corpinfo/programs/using_wireless_technologies_for_context_sensitive_education_and_training/products/london_wp5_evaluation_report.pdf

Takeuchi, Y. & Perlin, K., (2012). ClayVision: The (Elastic) Image of the City. In Proc. of CHI. Available at: http://www.tinylab.me/mobile/publications/chi12cv.pdf

Thian, C. (2012). Augmented Reality—What Reality Can We Learn From It? Asian Civilisations Museum, Singapore. Museums and the Web 2012: Proceedings. Toronto: Archives & Museum Informatics. http://www.museumsandtheweb.com/mw2012/papers/augmented_reality_what_reality_can_we_learn_fr

Waltham (1989) cited in Robert E. Horn, Mapping Hypertext: Analysis, Linkage, and Display of Knowledge for the Next Generation of On-Line Text and Graphics,p. 259.