EMPHASIZED LANDMARKS FOR MENTAL CALIBRATION IN A MOBILE AUGMENTED REALITY APPLICATION

Helge Fredheim
Student, Department of Media and Communication, University of Oslo - Gaustadalléen 21, 0349 Oslo, Norway

ABSTRACT

Landmark identification both in the real and virtual worlds is a common method used for registration and calibration in augmented reality. In this report, the term 'mental calibration' is defined as calibration taking place in the mind of the user. 'Emphasized landmarks' are suggested as a means to support the corresponding of landmarks taking place prior to this kind of calibration. A situated simulation of the Forum Iulium in ancient Rome was created. It was then tested by 11 students. They reported emphasized landmarks to be good helping tools. Not all participants understood or registered the concept, albeit they reported a good relationship between the real and virtual. Emphasized landmarks are considered beneficial, although a simulation shouldn't rely exclusively on this particular type of landmarks.

KEYWORDS

Augmented reality, situated simulation, landmarks, mental calibration, registration

1. INTRODUCTION

A "situated simulation" (sitsim) is a prototype to a digital genre closely related to augmented reality (AR). The prototype takes advantage of a 3G broadband smartphone (e.g. an iPhone 3GS) with substantial graphics capabilities, GPS positioning features and an accelerometer, and is suggested to be used as a didactic tool. (Liestøl, 2009b) The phone displays a 3D representation on its screen, and allows a user to navigate in the virtual environment (VE) by navigating in the real. Objects not present in the real are presented on the screen. Virtual objects may be presented together with real objects, but these objects are, as opposed to many other AR/mixed reality varieties, presented only as pre-made 3D representations.

A user of the simulation can correspond the representation on the screen to reality. However, there isn't a perfect registration (alignment between virtual and real), because of GPS positioning errors. Precise alignment between virtual and reality is considered to be one of the most important technical aspects of AR systems. (Azuma, 1999) It's assumed this makes it slightly difficult for the user to relate the representation to reality and vice versa. "Without good registration, many AR applications will not be accepted." (ibid.) This paper addresses this issue in a way which, to the author's knowledge, is new. Two concepts are presented – mental calibration and emphasized landmarks.

Landmarks are often used in visual navigation. (Vinson, 1999; Steck et al., 2000) When a user relates a landmark on a map with a landmark in real life, reality can be associated to the map and vice versa. Landmarks are also used for automatic registration and calibration within AR. (Dey et al., 2000; Satoh et al., 2001; Thomas et al., 2002)

The contribution of this paper is twofold. First, it explains how certain kinds of sitsims may need a better registration than others. As an alternative to traditional (technical) calibration, **mental calibration** is introduced as a cognitive alignment (registration) process. **Emphasized landmarks** are suggested as a means to support this kind of calibration. Second, it explains how a sitsim was designed to facilitate mental calibration. It describes how the sitsim was tested, and the results are presented and analyzed.

The test was done within a learning context, and the cognitive aspects are related to learning, mental models and how to acquire information.

2. SITSIM

The sitsim genre prototype has emerged from the Inventio Project at the University of Oslo (http://inventioproject.no/sitsim). Several simulations have been made and tested. A simulation has typically consisted of the following:

- The 3D representation itself.
- Sound from the environment, for example noise from the street, the sound of a brook etc.
- Hypertext; where a link is a label (balloon) connected to a spot in the representation. The user can press the balloon, and accesses node which can be a web URL, a voice-over, an image, a detail view of a 3D object, etc.

2.1 Case: Forum Iulium

During the spring of 2010, a sitsim displaying a 3D representation of the Forum Iulium in ancient Rome was created and tested.

At the northwest corner of the forum, the Temple of Venus Genetrix was raised by Julius Caesar. Only fragments of the temple remain today, including a podium onto which the temple was built, and three dominant columns on top of it, with an entablature fragment on top of them. The temple was the main feature of the simulation.

Exploitation of the fact that parts of the temple remain, while others aren't real anymore, was attempted in order to address the calibration issue mentioned above.

Figure 1. The application in use. Testing of the Forum Iulium sitsim on-site revealed that GPS positional error made the representation align improperly to reality, as expected. The temple should have been more to the left, in order to align to the real columns

3. THE USE OF LANDMARKS IN AR

The use of **landmarks** is a suggested means to ensure proper alignment between the real and virtual worlds. The system exploits markers that can be identified in both spaces by capturing the markers in the real space and running algorithms in order to identify where these landmarks belong in the virtual space. After this procedure, the system can calculate its position and map the virtual against the real. This process is often referred to as *calibration*. (Whitaker et al., 1995; Holloway, 1997; Kutulakos et al., 1998; Azuma, 1999; Huosheng et al., 2000; Satoh et al., 2001; Thomas et al., 2002) Calibration has been used within several fields, e.g. gaming (Thomas et al., 2002), brain surgery (Dey et al., 2000; Rasmussen et al., 2007) and mobile robot positioning (Borenstein et al., 1994; Huosheng et al., 2000).

3.1 The GPS Positional Error on iPhone 3GS

It has been argued that in AR applications, the GPS positioning error should decrease when a user moves closer to an object observed. (Milgram et al., 1994; Steck et al., 2000; Thomas et al., 2002) This will increase the probability of maintaining the application's acceptance by the user. (Azuma, 1999) In a sitsim, the user might be relatively close to the object being observed (figure 1), and the positioning error should consequently be small. However, testing of the iPhone 3GS has revealed a "horizontal accuracy of 15-30 meters" (Boss, 2010).

Here, it is thought that if a virtual object is placed up to 30 meters away from its position in the reality of the past, present or future, the GPS positioning error will have a detrimental effect on the user experience. Consequently, the sitsim should – in some way – take care of these positioning errors in order to keep the application acceptable.

3.2 Multiple Tense Categorization

Sitsims previously made and tested have been categorized together with simulations which may be created in the future. (Liestøl, 2009a) The categorization is based on whether the given simulation belongs to the past, present or future, and whether it features a static object (without animated events), a dynamic sequence or is participatory (where users can participate and change the inventory of the VE).

While the previously designed simulations belong to either of the categories along the past-present-future axis, the Forum Iulium simulation belongs to both the past and the present categories. The main parts of the temple are not real anymore, so they belong in the past category only. Other parts are real, and belong to the past as well as the present categories. This phenomenon is hereafter called a **multiple tense categorization**[1].

In mixed reality, the real and virtual worlds can be merged using a video display "upon which computer generated images are electronically or digitally overlaid" (Milgram et al., 1994). The difference between a sitsim and mixed reality is that the real world is never mediated by digital technology in the sitsim. (Liestøl, 2009b) The real stays real, and since the device displays nothing but a 3D representation of reality, it can be claimed that there is no similar need for accuracy in a sitsim as there is in mixed reality. However, as real objects are represented as 3D representations, the border between the real and virtual worlds becomes blurred. The real objects are not mediated, but they are still represented on the same screen as the virtual objects. Therefore, it can be argued that when a sitsim is multiple tense categorized, the genre prototype becomes closer related to mixed reality.

3.3 Need for Alignment

It's assumed that the multiple tense categorization creates a need for a more proper alignment between the reality and the VE, than what has been seen in earlier sitsims. It's possible to imagine how the virtual viewpoint should have looked, by comparing the positions and perspectives of real objects (serving as **landmarks**) with the positions and perspectives of the same virtual objects. For this reason, it's also easily possible to reveal a lack of accuracy, which may lead to lack of acceptance. (Azuma, 1999)

3.4 Mental Calibration

As outlined above, a solution to this may be to let the system itself exploit landmarks, and then calibrate. However, when the Forum Iulium simulation was created, this wasn't possible, due to practical circumstances. An alternative concept was therefore constructed, hereafter called **mental calibration**, which is explained as a calibration taking place in the user's mind.

As there is a need for a more proper alignment, and this need cannot be satisfied by the software, an attempt was made to take advantage of mental calibration. Through a technique explained in the following, it was aimed towards *helping* the user to:

1. Identify landmarks in the virtual world

[1] This phenomenon is also present when a simulation features objects which are available in the present and the future categories.

2. Identify the corresponding landmarks in the real world
3. Mentally align (calibrate) those landmarks
4. Mentally transfer the whole representation from the virtual world to the real world, properly aligned

4. CREATING THE SIMULATION

A 3D representation of the Forum Iulium as it appears today was constructed using Cheetah3D modeling software. The construction was based on images of the Forum Iulium, observations done on-site, images at *Google Maps Street View* and *Bing Maps*, drawings of the Forum (Ulrich, 1993) and the Rome Reborn project (www.romereborn.virginia.edu). This version was called the *naïve* version.

Additionally, a 3D representation of the Forum Iulium as it might have appeared between 44 BC and 80 AD was constructed based on the resources mentioned above. Today's fragments were used as guidelines, and the rest of the Forum and the temple were built around them.

Now two 3D representations were created; one *original* version and one *naïve* version. Both the representations were transferred to the iPhone via Unity iPhone software.

4.1 Selection of Real Landmarks

The Temple of Venus Genetrix most likely featured 22 columns in total. There were eight columns along the front (south), and eight columns along the right and left side. The rear of the temple was solid. (Ulrich, 1993)

Only three of the 22 columns remain today, located together at the south-western corner of the temple. These columns were considered to be the most significant landmarks; they are easy to spot and they differ from the rest of the Forum area. Landmark salience may influence a user's selection of landmarks. (Steck et al., 2000) Therefore, the three columns were chosen.

4.2 Using the Landmarks

Since the virtual temple in the original representation featured 22 columns in total, and only three of them are real, it was needed to show the user exactly which of the columns in the model the three real columns correspond to. The idea was that by knowing this, the user could *mentally align* or *calibrate* the virtual to the real perspective, and then imagine how the virtual objects could augment the real ones.

The user was helped in two different ways:

4.2.1 Support 1: The Naïve Model

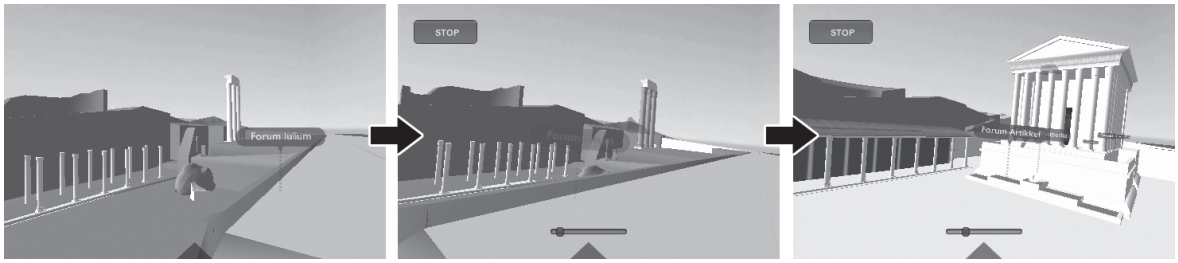

Figure 2. The process of exchanging the naïve model with the original model

At the start of the simulation, the present topography and fragments were displayed in the simulation by viewing the naïve model (figure 2, left). In this way, the intention was to make the user relate the real world to the represented world, and through this understand how reality was simplified (thus 'naïve') in the virtual.

The user then pressed a balloon labeled "Forum Iulium", and the naïve representation was exchanged with the original one. Simultaneously, a fragment of the entablature which both in the real and the naïve model was placed on the Forum, was moved by animation to what could have been its original position, in order to make the user establish a mental link between the naïve and original representations. The columns

were simultaneously emphasized in the representation by the use of a red-colored texture. The red texture on the columns faded out after a while. The exchange process is illustrated in figure 2.

4.2.2 Support 2: Emphasized Landmarks

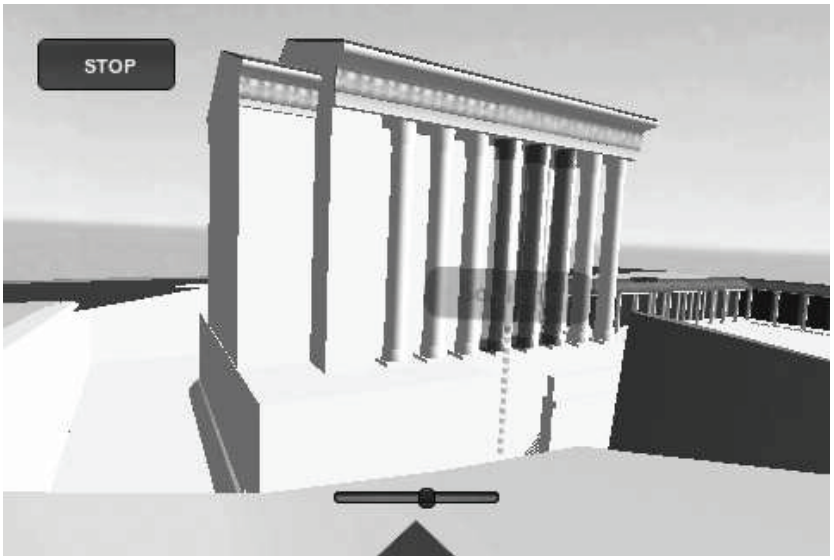

Figure 3. The temple with emphasized landmarks

The columns had a balloon attached to them, available at the rear of the temple. When the user activated it, the columns changed colors to red (as illustrated in figure 3) and a voice-over was played simultaneously. The color change technique is hereafter called **emphasized landmarks**. It's thought that this emphasis would make it easier for a user to identify the landmarks in the virtual world.[2]

5. DATA COLLECTION

Data was collected by performing a user test at the Forum Iulium in Rome, Italy. A total of 11 Norwegian students (6 females, 5 males) who studied art, architecture and archaeology at The Norweigan Institute in Rome participated in the test. 10 of the students had previously tested a sitsim designed for Parthenon in Athens, and consequently had some experience with the genre prototype and how to use the device already.

First, they got a brief introduction in a conference room. They were told where to start the application and where to walk. Afterwards they were given an iPhone with the application installed. Then they went to the Forum. They started the application, and held the device towards the forum to see the representations.

The participants were observed while using the application, and helped if things went wrong. Some of them gave feedback while testing. When they did, follow-up questions were asked and notes were taken. Video recordings and still photos were also collected during the test.

After having tested the Forum Iulium application, the participants went to test another similar application named The Forum of Julius Caesar. When finished, they went back to the conference room. They answered two surveys (one for each application). Three questions related specifically to landmarks.

When all the students were finished, they were asked to discuss three overall questions in two groups. The discussions were recorded on video. At the end, the groups gathered, and they presented the results of their discussions. Follow-up questions were asked when things were unclear.

[2] While the naïve and original representations featured other landmarks as well, such as the Forum with its tabernae and a church as they appear today, it's here paid particular attention to the emphasized landmarks.

6. RESULTS

The results from the survey are summarized in table 1.

Table 1. 90 percent reported that they were satisfied with the spatial relationship between the real and the virtual. 73 percent reported the relation was 'good' or 'very good', while 18 percent reported it as 'OK'. The latter group reported however the emphasized landmarks to work 'well and 'very well' respectively

Sex	*Experience*[3]	*Relationship*[4]	*Landmarks*[5]	*Remarks*
F	Some	Bad	Bad	47 meters GPS error
M	Some	Good	Well	Landmarks should be emphasized earlier
F	None	Good	Bad	Didn't note the emphasis of landmarks
F	None	Good	Well	Landmarks work well for orientation
M	Some	OK	Well	The "Columns" balloon could have been placed in front of the temple
M	A lot	Good	Well	N/A
F	None	OK	Well	N/A
M	None	Good	Very bad	Didn't understand [the concept of emphasized landmarks]
M	A lot	Good	Well	N/A
F	None	Good	OK	N/A
F	Some	Very good	Very well	Good to see their relation to the rest of the temple

7. ANALYSIS

7.1 The Usefulness of Landmarks in General

7.1.1 Landmarks for Context

When discussing the simulation, the participants went thoroughly into the advantage of the columns. One participant (I1) related the Forum Iulium simulation the Temple of Julius Caesar simulation. This simulation didn't include easily identifiable landmarks. The participant said: *"At the Temple of Julius Caesar, it's hard to see the relationships between the fragments on the ground and the reconstruction on the screen."*

Another participant (I2) explained the comment by I1: *"The advantage at the Forum Iulium is the three columns, so you manage to place ... [the representation in the real]."*

A third participant (I3) wished to see more of the surroundings of the Forum and the temple.

A fourth participant (I4) related the Forum Iulium simulation to the done for Parthenon mentioned above. This simulation featured the temple and objects inside of it. I4 explained how the Parthenon *"became kind of abstract"* because he couldn't see the relationship between the temple and its surrounding buildings. This participant didn't focus on emphasized landmarks, but on landmarks in general.

7.1.2 Landmarks for Mental Calibration

I2 explained that without landmarks, the device needed a 100 percent accuracy of the GPS in order to be useful. He explained: *"If the GPS has an error, you will be mis-situated. So […] [by the help of landmarks] you have something to relate it to, and you see that you are situated on a given spot in the real, and you have to wait to let the device update [its GPS positions]."*

The participants suggested that the application could feature an overview map telling a user where she is, where she can go and where she has been. A fifth participant (I5) did however argue *"[n]ot everyone is used to maps […], so it may be a good idea just to see [...] [the landmarks]. Simple communication!"*

[3] Question: Have you experience with touch telephones or handheld controllers? (E.g. Iphone, Ipod Touch, Playstation Portable, Nintendo DS)
[4] Question: How was the relationship between real and virtual?
[5] Question: We tried to mark out the columns in order to make it easier to relate the real to the virtual. How did it work?

7.1.3 Summary

From the results mentioned above, it can be extracted that the virtual world doesn't need to be 100 percent properly aligned to reality in order to make the user relate the two, when landmarks are available in the representation. Landmarks were considered to **support** mental calibration.

This challenges the notion that "without good registration, many AR applications will not be accepted" (Azuma, 1999), as mental calibration seemed to compensate for registration problems (as illustrated on figure 1). A combination of *acceptable* registration and landmarks may keep an AR application accepted.

7.2 The Usefulness of Emphasized Landmarks

Although the vast majority of the students reported satisfaction with the relationship between the real and virtual worlds, it's questionable whether the emphasized landmarks *helped* to create this satisfying relationship, or whether the relationship between both visualizations was independent of landmarks.

7.2.1 Relationship between Real and Virtual Landmarks

Four users gave negative feedback to the use of emphasized landmarks, albeit three of the four did report the relationship between real and virtual to be 'good'.

This aspect may suggest that the relationship between real and virtual may be perceived as 'good' **independently** of the emphasized landmarks. Additionally, it points to both technical and cognitive pitfalls related to landmarks, and that an application shouldn't rely on this exclusively.

7.2.2 Early Emphasized Landmarks

Two participants wanted the landmarks to be emphasized earlier in the simulation.

The landmarks were emphasized during the exchange event at the beginning of the simulation. This event lasted for about 15 seconds (explained at 4.2.1). When finished, the emphasis faded away. Probably, this event wasn't captured by all the users, and it's likely they didn't have a good opportunity to imagine a calibration for this short period of time. After the event, they had to walk through nearly the whole simulation before they could observe and activate the "Columns" balloon placed behind the temple and take advantage of the emphasis.

First, the suggestion of early emphasized landmarks points to the fact that the emphasized landmarks couldn't have effect until the "Columns" balloon was activated, close to the end of the simulation. This may indicate that the relationship between the virtual and real was considered to be 'good' **independently** of the emphasized landmarks. It is questionable whether the relationship was considered to be 'good' throughout the whole simulation, or only at the end of it, albeit it's likely the general impression of the relationship wouldn't have been 'good' if it was 'good' only at the end of it.

Second, the suggestion points to the notion that the concept of emphasized landmarks for mental calibration was well received among these participants, as they wanted them to be more available. This indicates that emphasized landmarks were considered **supportive**.

Third, it implies that landmarks in a simulation designed in the future should be emphasized either early or throughout the whole simulation. Those suggesting early emphasized landmarks constituted a minor group of 18 percent of the total number of participants. Even so, when two people come up with the same idea in an experiment like this, it can be argued their idea should be taken into account and tested in a future experiment.

8. CONCLUSION

This paper suggests that emphasized landmarks support mental calibration in a sitsim which can be multiple tense categorized. In the tested application, emphasized landmarks were well received by the students who used it, although they were not understood or identified by all of them. The latter shows that a sitsim shouldn't rely exclusively on emphasized landmarks, but also could take advantage of landmarks in general.

9. FUTURE WORK

A possible next step in the evaluation of landmarks in sitsims could be to create a simulation in which the emphasized landmarks are visible early or throughout the whole simulation.

ACKNOWLEDGEMENTS

Without the input of Professor Gunnar Liestøl at Department of Media and Communication at the University of Oslo and the engeneering expertise of Tomas Stenarson at Mementio AB, this study could never have been conducted.

Thanks to Ingeborg Marie Dehs Thomas, Vegard Fleischer Orkelbog and Øystein Reistadbakk the fellow students, Mathilde Liestøl for coordinating the experiment, Magnus Liestøl for taking pictures and Dr. Inge Rasmussen and Joseph Smith for proofreading.

REFERENCES

Azuma, R. T., 1999. The Challenge of Making Augmented Reality Work Outdoors. In Mixed Reality: Merging Real and Virtual Worlds.

Borenstein, J., H. R. Everett, et al., 1994. Where am I? Sensors and Methods for Autonomous Mobile Robot Positioning. University of Michigan, Michigan.

Boss, S. K. (2010). Potential of the iPhone 3Gs as a Field Mapping Instrument. North-Central Section (44th Annual) and South-Central Section (44th Annual) Joint Meeting. Boulder, USA, The Geological Society of America.

Dey, D., P. Slomka, et al., 2000. Mixed Reality Merging of Endoscopic Images and 3-D Surfaces. Springer Berlin, Berlin Heidelberg

Holloway, R. L., 1997. Registration Error Analysis for Augmented Reality. In Presence: Teleoperators and Virtual Environments Vol. 6, No. 4, pp. 413-32.

Huosheng, H. and G. Dongbing, 2000. Landmark-based navigation of industrial mobile robots. In Industrial Robot: An International Journal Vol. 27, No. 6, pp. 458-67.

Kutulakos, K. N. and J. R. Vallino, 1998. Calibration-Free Augmented Reality. In IEEE Transactions on Visualization and Computer Graphics Vol. 4, No. 1, pp. 1-20.

Liestøl, G., 2009a, Augmented reality and digital genre design - Situated simulations on the iPhone. IEEE International Symposium on Mixed and Augmented Reality - Arts, Media and Humanities, 2009. ISMAR-AMH 2009, pp. 29-34

Liestøl, G., 2009b. Situated Simulations: A Prototyped Augmented Reality Genre for Learning on the iPhone. In International Journal of Interactive Mobile Technologies (iJIM) Vol. 3, No. SI1, pp. 24-8.

Milgram, P. and F. Kishino, 1994. A Taxonomy of Mixed Reality Visual Displays. In IEICE Transactions on Information Systems Vol. E77-D, No. 12.

Rasmussen, I. A., F. Lindseth, et al., 2007. Functional neuronavigation combined with intra-operative 3D ultrasound: Initial experiences during surgical resections close to eloquent brain areas and future directions in automatic brain shift compensation of preoperative data. In Acta Neurochirurgica Vol. 149, No. 4, pp. 365-78.

Satoh, K., M. Anabuki, et al. (2001). A Hybrid Registration Method for Outdoor Augmented Reality. Proceedings of the IEEE and ACM International Symposium on Augmented Reality (ISAR'01), IEEE Computer Society**:** 67.

Steck, S. D. and H. A. Mallot, 2000. The Role of Global and Local Landmarks in Virtual Environment Navigation. In Presence: Teleoper. Virtual Environ. Vol. 9, No. 1, pp. 69-83.

Thomas, B., B. Close, et al., 2002. First Person Indoor/Outdoor Augmented Reality Application: ARQuake. In Personal Ubiquitous Comput. Vol. 6, No. 1, pp. 75-86.

Ulrich, R. B., 1993. Julius Caesar and the Creation of the Forum Iulium. In American Journal of Archaeology Vol. 97, No. 1, pp. 49-80.

Vinson, N. G. (1999). Design guidelines for landmarks to support navigation in virtual environments. Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit. Pittsburgh, Pennsylvania, United States, ACM**:** 278-85.

Whitaker, R. T., C. Crampton, et al., 1995. Object Calibration for Augmented Reality. In Computer Graphics Forum Vol. 14, No. 3, pp. 15-27.