Views, Alignment and Incongruity in Indirect Augmented Reality

Gunnar Liestøl *
Department of Media & Communication, University of Oslo, Norway

Andrew Morrison **
Centre for Design Research, Oslo School of Architecture and Design

ABSTRACT

Alignment between the real and the virtual has been a defining quality of mixed and augmented reality. With the emergence of Indirect Augmented Reality the problem of alignment is no longer primarily concerned with the relationship between a live video feed and a 3D graphics layer at the level of the screen, but with the relationship between the visual information on the display and the real world perspective outside the display of the device. Experiments show that users easily connect the perspective into the 3D virtual environment on the full screen with their parallel perspective in the real world. It also turns out that although congruence and alignment between the two perspectives is fundamental to the user experience, in certain contexts it may be transcended. This paper describes and discusses applications of Indirect Augmented Reality where we explore how the discrepancy between the virtual and the real perspectives in a variety of ways can be used to improve the user experience. We call these features *views*. The views will be exemplified with several Indirect Augmented Reality applications: reconstructions of Augustus' Forum and the Republican Forum in Rome and a preconstruction of the planned National Museum in Oslo. The applications have been tested with users on location, and their feedback and evaluation is included in the discussion. Finally, we relate the experiential value of the views to some epistemological and pedagogical perspectives.

Keywords: Indirect Augmented Reality, Situated Simulations, sitsim, Views, situated learning

Index Terms: H.5.1[Information Interfaces and Presentation]: Multimedia Information Systems – Artificial, augmented, and virtual realities. J.5 [Arts and Humanities]: Arts, fine and performing.

1 INTRODUCTION

Alignment between the real and the virtual has in various ways always been a defining feature of Augmented Reality (AR). This paper describes and discuss applications of Indirect Augmented Reality (IAR) where we explore how the discrepancy between the virtual and the real perspective can be used to improve the user experience, with what we call *views*.

These views make it possible for the user to uncouple the virtual camera and its perspective in the graphical environment from the real perspective of the user, and they include: normal view, zoom, bird's view, map view, sub view, fly–in, free view as well as the double view provided by the snapshot function, a feature in the application for combining and documenting the virtual and the real perspectives in a single image.

*gunnar.liestoel@media.uio.no, **andrew.morrison@aho.no

IEEE International Symposium on Mixed and Augmented Reality 2013
Arts, Media, & Humanities Proceedings
1 - 4 October 2013, Adelaide, SA, Australia
978-1-4799-2945-0/13/$31.00 ©2013 IEEE

Figure 1: Using the snapshot feature in the *Roman Forum* sitsim, combining the real and the virtual view in front of the Temple for the Deified Julius Caesar.

The views are exemplified with several augmented reality applications that we have developed and studied in what has been termed sitsims (situated simulations) [1], [2], see also www.sitsim.no. This refers to the type of Mobile Augmented Reality system known as Indirect Augmented Reality [3]. The cases we refer to cover historical reconstructions from ancient Rome as well as preconstructions of a future urban building in downtown Oslo. These several applications have been tested qualitatively with various user groups on location, and their feedback and evaluation are considered in discussions of the variety of views in IAR. Finally, some epistemological consequences of these views are discussed with respect to the broader perspective of educational uses and applications, primarily situated learning [4].

In the practice of mixed reality [5] and with typical current solutions – predominantly on smartphones and tablets – a live video-feed from the device's camera is combined with a graphics overlay. The mixing then takes place at the level of the screen and the alignment is supported by sensor fusion. Such a solution is suitable for augmenting present 'live' reality by means of annotation: models, icons and labels floating on top of the live video providing positioned information (verbal as well as nonverbal) and access to services in the immediate and near environment (e.g. platforms such as Layar).

More complex representations, however, are still problematic. The 2D video and the 3D graphics environment are two incompatible forms of representations and cannot be fully integrated. This creates perceptual paradoxes of forefront and background: what is 'spatially' closest to the user's perspective and position in the video is always behind all the information in the graphics layer. Occlusion may reduce such problems to some extent, but the fundamental difference between the flatland of video on the one hand and the dynamic 3D worlds on the other, remains.

Azuma and colleagues [6] outline a technology independent approach that encompasses various generic AR criteria:

1) combination of real and virtual objects in real environments;
2) interactivity in real time; and
3) registration (alignment) of real and virtual objects.

This definition opens up for solutions not limited to the screen mixing of the virtual and real. Why use screen space to represent reality in limited video resolution when it is available in infinite detail just outside the screen? IAR refers to a blend where the graphics layer or 3D environment occupy the whole screen [3].

The paper proceeds as follows: the next section presents conceptual and theoretical aspects of situated simulation, including the seven sitsim views. Then comes a section on contexts of use, relations to learning and finally, a conclusion.

2 Theory

2.1 Situated simulation

With the emergence of Indirect (and Mobile) Augmented Reality, the problem of alignment is no longer primarily concerned with mixing at the level of the screen [7], but with the relationship between the visual information on the display and the real world perspective outside the display of the device. A Situated Simulation (sitsim) is obviously such a form of Indirect Augmented Reality (see Figure 1).

As the user of a sitsim changes position and moves the device in real space the perspective inside the 3D graphic environment changes accordingly. In the simulated environment spatially distributed hypertext links provide access to additional information: audio narration, detailed 3D objects, sequences of actions and events, online resources via the internal browser, as well as options for adding various modes of information, namely links with written or spoken material and pictures taken with both the real and the virtual camera in combination. In this paper, however, we will focus on the *Views,* and their manipulations of the virtual camera, in order to break with the congruity in the parallel position and orientation of the two perspectives.

The relationship between the real and the virtual perspective has been a key design focus of the sitsim development since its initial conception in 2006. In fact, the non–mixed separate, but congruent relationship between the real and the virtual perspective was directly inspired and informed by Bateson's theorizing of *difference* as a founding principle (and definition) of information [8]. In his discussion Bateson is concerned with "... the bonus of understanding which the combination of information affords" (1988: 72). He provides a series of cases where such an accrual takes place: synaptic summation, synonymous languages, the two sexes and more. The most important in our context, however, is the case of binocular vision.

Combining information from the *different* perspectives of our eyes when looking at an object or area in the visual field common to both of them gives rise to an incremental value. This is new information of a different kind than that found in each individual perspective: information about *depth*. This new level of information is arrived at because of the difference between the individual ones that are combined. This opens up for an important questions in the investigation of Indirect Augmented Reality: To what extent can we exploit the divergence between the real and the virtual perspectives, and how may we differentiate the differences? These are also questions of how we may invent and employ different 'views'.

2.2 Views in Situated Simulations

Earlier experiments have shown that users easily connect the normal view perspective into the 3D virtual environment on the full screen with their parallel perspective in the real world [2]. This is consistent with the findings of Wither, Tsai and Azuma [3]. It also turns out that although congruence and alignment between the two perspectives are fundamental to the user experience, in certain contexts these may be transcended. Differentiating between various views makes it possible to both select positions and perspectives within the simulation environment and user experience as well as to traverse the augmented domain. The views we have categorized, and that have been applied in actual design practice, and in students' user and wider user testing on site, is described in seven variants below.

2.3 Seven *views* in situated simulations

a) *Normal view,* which is stable at an altitude of 1.7m, is actually not a normal view in the sense that it is optimal alignment between the virtual and the real perspective (see Figure 2). During early testing it was soon discovered that the best congruity we were able to achieve between the two perspectives was not good enough. It forced the user to hold the device up high and in a physically exhausting position that, in addition, would stand in the way of the central area in the real view. The solution to this problem was to tilt the virtual camera by 15% so that the device could be lowered and thus neither cover the view nor wear the user out. The normal view in the regular use of the sitsim is thus actually already an intended incongruity.

Figure 2: *Normal view* when using the *Augustus' Forum* sitsim.

b) *Zoom* uses the conventions of a telescopic lens to allow adjustable focal length and thus be able to see close up views of objects positioned at a distance (see Figure 3). In zoom view the tilting of the camera angle in the normal view perspective is disabled. A touch–activated ruler on the left side of the screen makes it easy to adjust the focal length. This view has the

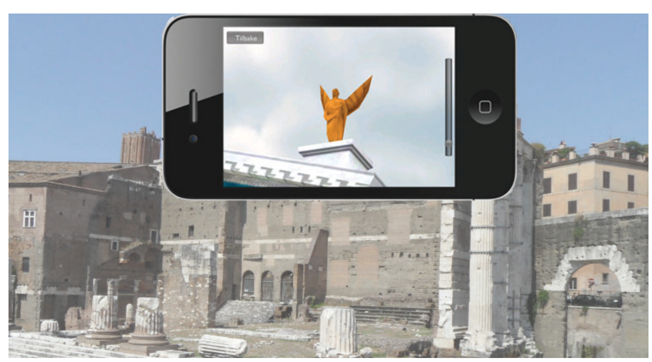

Figure 3: *Zoom view*, the *Augustus' Forum* sitsim. The device is directed towards a statue over the right of the pediment.

disadvantage of a telescope: the more you zoom the more difficult it is to keep the device still. However, an important quality of this view is that the subjective perspective and position remain intact (all the orientation and positioning sensors on the device continue to be active), while the alignment of the two vistas becomes increasingly decontextualized relative to the strength of the zoom.

c) *Bird's view* changes the altitude of the virtual camera to an adjustable height of 8–15m above the user's real position (Figure 4). This function was originally implemented to provide better visual access to objects and details on high buildings, e.g. the metopes of the Parthenon temple on the Acropolis in Athens [2]. However, it soon turned out to be a feature suitable for general use in providing general overview. The switching between normal view and bird's view is continuous and thus avoids disorientation and confusion. All other movements – tilting, panning and continuous change of position – remain active.

Figure 4: *Bird's view* of Caesar's Temple as seen from a position next to the Basilica Julia.

d) *Map view* increases the altitude to about 100m and is intended to give a more general overview of the simulated environment, its' inventory, the user's position, etc. The altitude is sufficient to give a map–like view of the surroundings despite the fact that the environment remains 3–dimensional (Figure 5). This view could serve many of the ways a traditional map presents abstract information by applying various forms of layers and filters for different kinds of data [8]. So far we have tested this view as a tracking device to help and improve navigation and orientation for the user. The idea then is to provide visual answers to the basic questions concerning navigation: Where have I been? Where am I? Where can I go next?

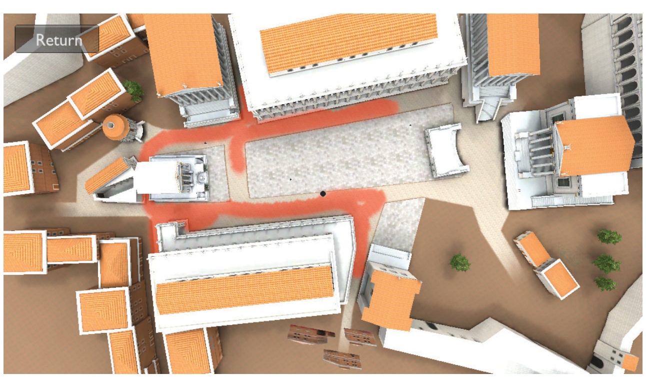

Figure 5: *Map view* of the *Roman Forum*. Red marking indicates covered ground and black dot user's current position; to the left the building covers the user's path indicating that one is still in the 3D model.

e) *Sub view* does the opposite of map and bird's view. The virtual camera is lowered vertically relative to the user's position. This is useful when the simulated terrain is lower than the real, which is the case with Augustus' Forum in Rome, or if one wishes to access subterranean structures like the Cloaca Maxima in the Roman Forum (Figure 6).

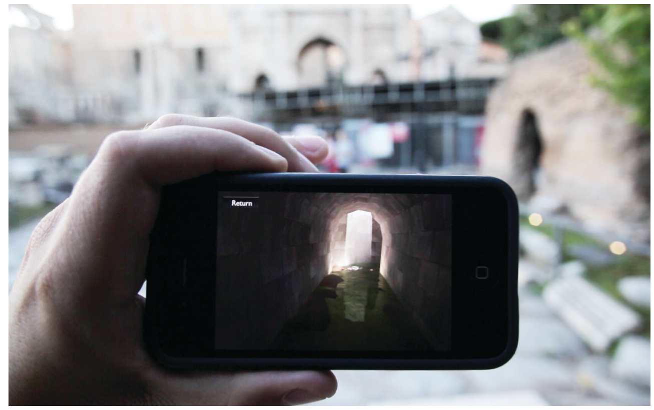

Figure 6: *Sub view* in the *Roman Forum* sitsim. The user is walking on the Via Sacra towards the Capitoline Hill, but the virtual camera has been sent down into the Cloaca Maxima reconstruction a few meters below the user's feet.

f) *Fly–in view* sends the camera on a predetermined route/track, for example inside a reconstructed building, in cases where the user cannot move physically inside the structure (see Figure 7). With fly–in, the GPS–positioning information is overruled, but the other sensors for orientation and movement are still active. When leaving the fly–in view the camera returns to where the user is currently standing. While in the modes of fly–in and fly–out the user may tilt and pan in any direction.

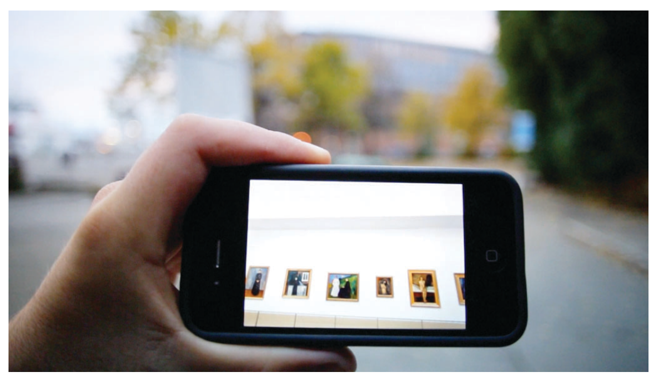

Figure 7: *Fly–in* to the Munch room, planned National Museum in Oslo. The GPS–position is overruled by the track of the moving camera; tilting and panning remain operative movements. The virtual camera is now positioned about 25 m to the west and up on the second floor of the future structure.

g) *Free view* is similar to fly–in, but positioning and orientation of the virtual camera is free for the user to choose while the fixed altitude of the camera at 1.7m remains. This provides a way of exploring the immediate surroundings when physical movement is restricted. Of all the views described here, this is the one that is most open to incongruence, in particular since initially the movement is free and thus can be completely detached from the real subjective perspective of the user. For this reason we decided to constrain the movement by attaching a 'rubber band' to the virtual camera: the further away from the position of the real user the virtual camera is moved (and its perspective detached), the more counteraction is felt in the ability to continue changing the virtual camera's position and perspective. When releasing the control button, the camera is pulled back to the user's real position.

Figure 8: *Free view* in *Augustus' Forum*. As can be seen, the perspective on the device has a different angle on the columns and the stairs than the real perspective of the user.

h) The final view on the list is not really a view, but a view to a combination of the two perspectives. In effect, it is the snapshot function where the app takes a picture with both the virtual and the real camera simultaneously, thus documenting the sitsim in use and combining the two perspectives (see Figure 1 for an example of a snapshot). When this feature is in use the tilting of the virtual camera found in normal view is deactivated. The phone frame and the image taken with the virtual camera may be repositioned and resized so that the two images can be combined as conveniently as possible, thus composing an image analogue to the ones found in the tradition of 'Now & then' books. The various views presented above can be represented schematically relative to the three axes of space (see Figure 9).

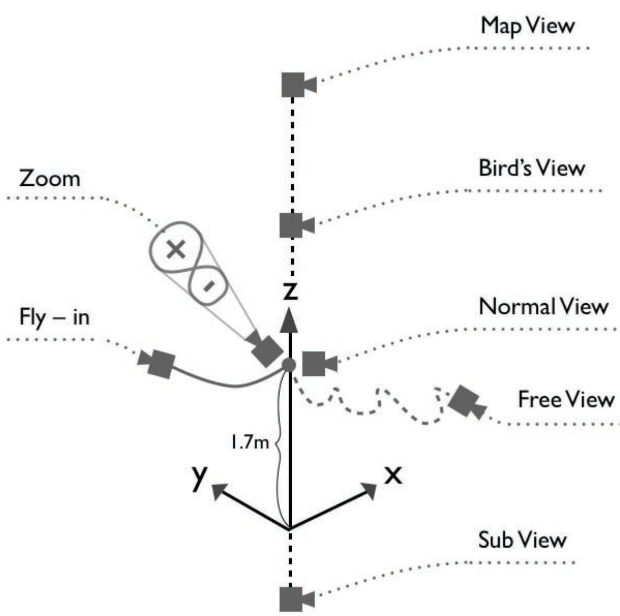

Figure 9: The current set of views in the sitsim solution.

The seven views described here do not add up to a finite list of possibilities. There are several more, which we still plan to implement. This is particularly the case when we combine features from the different views. For example, currently Free view only includes free movement horizontally or relative to a surface (visible or not). It is not possible to change altitude. One could then imagine combinations of horizontal and vertical manipulation included in one 'view', including zoom. A total free navigation is also possible, but then we are bordering on the founding condition of the whole sitsim platform – the situatedness.

3 Context of Use

3.1 Development and Background for Testing

The development of these views has occurred over several years and in relation to a number of different cultural historical locations as well as diverse groups of users. The development was also realized through co-design activities between a programmer, a 3D developer and a media design researcher. This took place in consultation with various parties, such as museums, urban planning authorities, cultural theorists and media design teaching. In the context of this paper, the views were subsequently taken up by a specific group of bachelor level students on site in Rome.

The various views were tested qualitatively with 15 B.A. students at The Norwegian Institute in Rome attending the course 'The Augustan Golden Age'. The simulation in question was a reconstruction of Augustus' Forum just north of the Roman Forum as it may have looked in the year 2 CE. The views where available as buttons in a hide/show dock at the bottom of the screen, except for one 'fly–in' view, which was a regular positioned hypertext link.

The students were given a quick general demo of the application's basic features, but no hands on use, prior to the actual test. The students tested the application themselves individually using Apple's iPad2. They each answered written questionnaires concerning general aspects of using the sitsim as well as on the views. After the test they engaged in group discussions prior to a larger plenary session. These sessions were video-recorded and reviewed.

3.2 Testing and Evaluation

All the students (100%) tried fly–in, 60% tried zoom, 33% took up map view, 20% sub view and 13 % free view. In general, they all responded positively to the views. Asked if the overriding of the real subjective perspective was confusing or not, about 50% found it confusing at first, but quickly adjusted to and enjoyed the flexibility.

One view received some criticism: the zoom feature. The detail provided was low, so there was actually not much added value in the zooming in on 'details'. This is due to the limited processing power on the device and the need to keep the geometric detail (poly count and draw calls) as low as possible [10]. However, these students saw value of the feature and welcomed it as a function with potential.

Drawing on several years design work on IAR and previous related teaching and learning, this small study in Rome highlighted the potential educational relevance of situated simulation and the role of the different views.

4 Perspectives, Views and Epistemology

In this section we return to Bateson's conception of multiple descriptions. When combined different descriptions give rise to new forms of information and knowledge [9]. We then also relate these to wider epistemological and pedagogical perspectives in sociocultural cognition and situated learning.

4.1 Reconsidering double description

Bateson distinguishes between a three basic levels of learning (in principle there are more): the mere reception of information (zero learning); by placing of information in the context of other information one achieves a new comprehension (learning); and finally the comprehension of the diversity of contexts to place that information in constitutes yet another level of learning (learning how to learn). Bateson's conceptions of difference and multiple descriptions are closely related to his theory of levels in learning [11]. In the cases of situated simulations and views described and presented above the multiple relationships between information, perspectives and contexts are crucial.

By means of the sitsim platform and the (possibly innovative) perceptions and experiences it renders possible, the user is probably active on all three levels: both receiving and combining information with other information and seeing the incremental value of those operations into the extensions of diverse contexts to further position that information. Given the focus on real place and position in the sitsim user context, the situatedness itself becomes a key condition and ingredient in the epistemological experience. How does this relate to education and learning in a broader perspective? How may we further develop the sitsim IAR platform in order to make it relevant to various educational contexts, both formal and informal? To tentatively embark on an answer to these questions we find it relevant to have a brief look at the pedagogical theory known as situated learning.

4.2 Situated learning

Having outlined the sitsim views and their accessibility on site and with users, we now turn to their epistemological potential, particularly their possible position within learning as a socio-cultural activity. In this approach to learning, that also increasingly applies to in-and-out-of-school settings, what is primary is that participants to learning events take an active and dialogical role. This is realized in relating to and constituting the settings, contexts and developmental character of the learning as an activity. In short, drawing on the theories of Vygotsky [12], learning is cast as a situated, developmental and dynamic activity.

This approach has been championed by Lave and Wenger [4] as situated learning so that topic, context, task and persons are seen as related to situated purposes and actual activities of the event in progress [13]. Situated cognition further involves the notion of a community of practice [14] that comes into being via shared activity and meaning making, a dialogical or shared yet negotiated process that also includes digitally assisted and blended learning [15] A situated learning approach has been taken into a wider socio-cultural perspective on learning that refers to mediated meaning making. Here tools, technologies and mediated artifacts and articulations all come into play in complex relations and reverberations. These too are enacted in use and via dialogical engagement and have a transformational character

By extension, Mobile IAR contexts, such as sitsims, add an additional mixed and simulated technological layer to the already complex fabric of physical world, literacy practices and communicative tools and cultural conventions. In terms of the developmental transformation that is part of a situated cognition view, sitsim is especially challenging and deceptively straightforward. Learning about past and future settings and events occurs on location, yet situated within these contexts is the mediational and artifactual level of the simulations.

4.3 Situated and simulated learning

In an Indirect Augmented Reality frame, context may be understood as being realized through designing for collaboration and coordination in spatial scale [16]. In our take on Views in IAR, and following Bateson's insights, what is needed is an understanding of the dynamic and dialogical activities that may be played out. In this approach, it is the interplay of the physical and the virtual that is realized through the selection and enactment of a View, and, indeed, through the relations between changing and orienting and navigating the real and the symbolic via shifting these views. We suggest that this might be termed Situated Simulated Learning (SSL).

5 CONCLUSION

In conclusion, in the sitsim views we have developed within Indirect Augmented Reality what is also relevant is that sharing is possible by discussing points of view with others on the same site but also by the inclusion of comments, additions and contributions via our design of links *inside* the simulated environment. Each of these can then also be tailored to reflect on aspects of being in a particular View and how such a perspective contributes to the situatedness of the learning event or locatively mediated information. Importantly, our offering of such views is not intended as taxonomic but rather a mode of exploratory and constructional inquiry in digital design [17].

Our design, seven view categories, contextual experiental use, and placement of these views in relation to Situated Simulated Learning (SSL) together help extend our understanding of the role of position, the dynamics of locational movement, the importance of visual perspective and the enactment of dialogical meaning making in meeting the challenges of alignment, incongruity and views in Indirect Augmented Reality.

ACKNOWLEDGEMENTS

We would like to thank the following for their valuable work making the sitsim project and its many views possible: Tomas Stenarson of CodeGrind for excellent programming and co–design and Šarūnas Ledas of Tag of Joy for all the 3D–modeling over the past several years. Thanks also to collegues at the Department of Media & Communication and EngageLab (both at the University of Oslo), and to The Norwegian Research Council and the Arts Council Norway for continued support. Thanks also to Tobias Langlotz and the WARM 2013 event for constructive feedback.

REFERENCES

[1] Liestøl, G. 'Augmented Reality and Digital Genre Design - Situated Simulations on the iPhone' in *Proceedings of ISMAR 2009 IEEE International Symposium on Mixed and Augmented Reality - Arts, Media and Humanities*. 19-22 Oct. 2009, 29-34.

[2] Liestøl, G., Rasmussen, T. & Stenarson, T. 'Mobile Innovation: designing and evaluating situated simulations', in *Digital Creativit*y, 22: 3. 2011. 174–186.

[3] Wither, J., Tsay, Y-T., Azuma R (2011): 'Indirect Augmented Reallity' in *Computers & Graphics* 35 (2011). 810–822

[4] Lave, J. & E. Wenger (1991) *Situated Learning. Legitimate peripheral participation*, Cambridge: University of Cambridge Press

[5] Milgram, P. & F. Kishino, 'A Taxonomy of Mixed Reality Visual Displays', in *IEICE Transactions on Information Systems*, Vol E77-D, No. 12, December 1994. 1321-1329.

[6] Azuma, R., Y. Baillot, R. Behringer, S. Feiner, S. Julier, B. MacIntryre, 'Recent Advances in Augmented Reality' in IEEE *Computer Graphics and Applications*, Vol. 21, No. 6, Nov./Dec. 2001. 34–47.

[7] Almeida, L., Stricker, D., Gleue, T., Christou, I.T., Carlucci, R., and Ioannidis, N. 'Archeoguide: first results of an augmented reality, mobile computing system in cultural heritage sites'. In *Proceedings*

of the 2001 conference on Virtual reality, archeology, and cultural heritage (VAST '01). ACM, New York, NY, USA, 2001. 131-140. http://doi.acm.org/10.1145/584993.585015

[8] Bolter, J. D., M. Engberg & B. MacIntyre (2013) 'Media Studies, Mobile Augmented Reality, and Interaction Design' in *Interactions* Vol. 20, No. 1 pp. 36–45.

[9] Bateson, G. *Mind and Nature. A Necessary Unity*. Bantam Books, 1988.

[10] Liestøl, G., Doksrød, A., Ledas, S. and Rasmussen, T. 'Sensory Media: Multidisciplinary Approaches in Designing a Situated & Mobile Learning Environment for Past Topics' in *International Journal of Interactive Mobile Technologies* (iJIM), Vol. 6 2012. 18–24 (Open access).

[11] Bateson, G. *Steps to an Ecology of Mind. A revolutionary approach to man's understanding of himself*. New York: Ballantine Books, 1985.

[12] Vygotsky, L. *Mind in Society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press. 1978.

[13] Suchman, L. *Plans and Situated Actions*. Cambridge University Press, 1987.

[14] Wenger, E. *Communities of Practice: Learning, meaning, and identity*. Cambridge: Cambridge University Press. 1998.

[15] Fauconnier, G. and Turner, M. *The Way We Think*. Basic Books, 2002.

[16] Barba, E. and MacIntyre, B.. 2011. A scale model of mixed reality. In *Proceedings of the 8th ACM conference on Creativity and cognition (C&C'11)*. ACM, New York, NY, USA, 117-126. DOI=10.1145/2069618.2069640 http://doi.acm.org/10.1145/2069618.2069640

[17] Wagner, I., Bratteteig, T. and Stuedahl, D. (Eds). *Exploring Digital Design*. Vienna: Springer. 2010.