Mixed Reality on a Virtual Globe

Zhuming Ai and Mark A. Livingston
3D Virtual and Mixed Environments
Information Management and Decision Architectures,
Naval Research Laboratory
Washington
USA

1. Introduction

Augmented reality (AR) and mixed reality (MR) are being used in urban leader tactical response, awareness and visualization applications (Livingston et al., 2006; *Urban Leader Tactical Response, Awareness & Visualization (ULTRA-Vis)*, n.d.). Fixed-position surveillance cameras, mobile cameras, and other image sensors are widely used in security monitoring and command and control for special operations. Video images from video see-through AR display and optical tracking devices may also be fed to command and control centers. The ability to let the command and control center have a view of what is happening on the ground in real time is very important for situation awareness. Decisions need to be made quickly based on a large amount of information from multiple image sensors from different locations and angles. Usually video streams are displayed on separate screens. Each image is a 2D projection of the 3D world from a particular position at a particular angle with a certain field of view. The users must understand the relationship among the images, and recreate a 3D scene in their minds. It is a frustrating process, especially when it is a unfamiliar area, as may be the case for tactical operations.

AR is, in general, a first-person experience. It is the combination of real world and computer-generated data from the user's perspective. For instance, an AR user might wear translucent goggles; through these, he can see the real world as well as computer-generated images projected on top of that world (Azuma, 1997). In some AR applications, such as the battle field situation awareness AR application and other mobile outdoor AR applications (Höllerer et al., 1999; Piekarski & Thomas, 2003), it is useful to let a command and control center monitor the situation from a third-person perspective.

Our objective is to integrate geometric information, georegistered image information, and other georeferenced information into one mixed environment that reveals the geometric relationship among them. The system can be used for security monitoring, or by a command and control center to direct a field operation in an area where multiple operators are engaging in a collaborative mission, such as a SWAT team operation, border patrol, security monitoring, etc. It can also be used for large area intelligence gathering or global monitoring. For outdoor MR applications, geographic information systems (GIS) or virtual globe systems can be used as platforms for such a purpose.

2. Related work

On the reality-virtuality continuum (Milgram et al., 1995), our work is close to augmented virtuality, where the real world images are dynamically integrated into the virtual world in real time (Milgram & Kishino, 1994). This project works together closely with our AR situation awareness application, so it will be referred as a MR based application in this paper.

Although projecting real time images on top of 3D models has been widely practiced (Hagbi et al., 2008), and there are some attempts on augmenting live video streams for remote participation (Wittkämper et al., 2007) and remote videoconferencing (Regenbrecht et al., 2003), no work on integrating georegistered information on a virtual globe for MR applications has been found.

Google Earth has been explored for AR/MR related applications to give "remote viewing" of geo-spatial information (Fröhlich et al., 2006) and urban planning (Phan & Choo, 2010). Keyhole Markup Language (KML) files used in Google Earth have been used for defining the augmented object and its placement (Honkamaa, 2007). Different interaction techniques are designed and evaluated for navigating Google Earth (Dubois et al., 2007).

The benefit of the third-person perspective in AR was discussed in (Salamin et al., 2006). They found that the third-person perspective is usually preferred for displacement actions and interaction with moving objects. It is mainly due to the larger field of view provided by the position of the camera for this perspective. We believe that our AR applications can also benefit from their findings.

There are some studies of AR from the third-person view in gaming. To avoid the use of expensive, delicate head-mounted displays, a dice game in a third-person AR was developed (Colvin et al., 2003). The user-tests found that players have no problem adapting to the third-person screen. The third-person view was also used as an interactive tool in a mobile AR application to allow users to view the contents from points of view that would normally be difficult or impossible to achieve (Bane & Hollerer, 2004).

AR technology has been used together with GIS and virtual globe systems (Hugues et al., 2011). A GIS system has been used to work with AR techniques to visualize landscape (Ghadirian & Bishop, 2008). A handheld AR system has been developed for underground infrastructure visualization (Schall et al., 2009). A mobile phone AR system tried to get content from Google Earth (Henrysson & Andel, 2007).

The novelty of our approach lies in overlaying georegistered information, such as real time images, icons, and 3D models, on top of Google Earth. This not only allows a viewer to view it from the camera's position, but also a third person perspective. When information from multiple sources are integrated, it provides a useful tool for command and control centers.

3. Methods

Our approach is to partially recreate and update the live 3D scene of the area of interest by integrating information with spatial georegistration and time registration from different sources on a virtual globe in real time that can be viewed from any perspective. This information includes video images (fixed or mobile surveillance cameras, traffic control cameras, and other video cameras that are accessible on the network), photos from high

altitude sensors (satellite and unmanned aerial vehicle), tracked objects (personal and vehicle agents and tracked targets), and 3D models of the monitored area.

GIS or virtual globe systems are used as platforms for such a purpose. The freely available virtual globe application, Google Earth, is very suitable for such an application, and was used in our preliminary study to demonstrate the concept.

The target application for this study is an AR situation awareness application for military or public security uses such as battlefield situation awareness or security monitoring. An AR application that allows multiple users wearing a backpack-based AR system or viewing a vehicle mounted AR system to perform different tasks collaboratively has been developed(Livingston et al., 2006). Fixed position surveillance cameras are also included in the system. In these collaborative missions each user's client sends his/her own location to other users as well as to the command and control center. In addition to the position of the users, networked cameras on each user's system can stream videos back to the command and control center.

The ability to let the command and control center have a view of what is happening on the ground in real time is very important. This is usually done by overlaying the position markers on a map and displaying videos on separate screens. In this study position markers and videos are integrated in one view. This can be done within the AR application, but freely available virtual globe applications, such as Google Earth, are also very suitable for such a need if live AR information can be overlaid on the globe. It also has the advantage of having satellite or aerial photos available at any time. When the avatars and video images are projected on a virtual globe, it will give command and control operators a detailed view not only of the geometric structure but also the live image of what is happening.

3.1 Georegistration

In order to integrate the video images on the virtual globe, they first need to be georegistered so that they can be projected at the right place. The position, orientation, and field of view of all the image sensors are needed.

For mobile cameras, such as vehicle mounted or head mounted cameras, the position and orientation of the camera are tracked by GPS and inertial devices. For a fixed-position surveillance camera, the position is fixed and can be surveyed with a surveying tool. A calibration process was developed to correct the errors.

The field of view and orientation of the cameras may be determined (up to a scale factor) by a variety of camera calibration methods from the literature (Hartley & Zisserman, 2004). For a pan-tilt-zoom camera, all the needed parameters are determined from the readings of the camera after initial calibration. The calibration of the orientation and the field of view is done manually by overlaying the video image on the aerial photo images on Google Earth.

3.2 Projection

In general there are two kind of georegistered objects that need to be displayed on the virtual globe. One is objects with 3D position information, such as icons representing the position of users or objects. The other is 2D image information.

To overlay iconic georegistered information on Google Earth is relatively simple. The AR system distributes each user's location to all other users. This information is converted from the local coordinate system to the globe longitude, latitude, and elevation. Then an icon can be placed on Google Earth at this location. This icon can be updated at a predefined interval, so that the movement of all the objects can be displayed.

Overlaying the 2D live video images on the virtual globe is complex. The images need to be projected on the ground, as well as on all the other objects, such as buildings. From a strict viewpoint these projections couldn't be performed if not all of the 3D information were known along the projection paths. However, it is accurate enough in practice to just project the images on the ground and the large objects such as buildings. Many studies have been done to create urban models based on image sequences (Beardsley et al., 1996; Jurisch & Mountain, 2008; Tanikawa et al., 2002). It is a non-trivial task to obtain these attributes in the general case of an arbitrary location in the world. Automated systems (Pollefeys, 2005; Teller, 1999) are active research topics, and semi-automated methods have been demonstrated at both large and small scales (Julier et al., 2001; Lee et al., 2002; Piekarski & Thomas, 2003). Since it is difficult to recreate 3D models in real time with few images, the images on known 3D models are projected instead at least in the early stages of the study.

To display the images on Google Earth correctly, the projected texture maps on the ground and the buildings are created. This requires the projected images and location and orientation of the texture maps. An OpenSceneGraph (*OpenSceneGraph*, n.d.) based rendering program is used to create the texture maps in the frame-buffer. This is done by treating the video image as a rectangle with texture. The rectangle's position and orientation are calculated from the camera's position and orientation. When viewing from the camera position and using proper viewing and projection transformations, the needed texture maps can be created by rendering the scene to the frame-buffer.

The projection planes are the ground plane and the building walls. This geometric information comes from a database created for the target zone. Although Google Earth has 3D buildings in many areas, including our target zone, this information is not available for Google Earth users and thus cannot be used for our calculations. Besides, the accuracy of Google Earth 3D buildings various from places to places. Our measurements show that our database is much more accurate in this area.

To create the texture map of the wall, an asymmetric perspective viewing volume is needed. The viewing direction is perpendicular to the wall so when the video image is projected on the wall, the texture map can be created. The viewing volume is a frustum of a pyramid which is formed with the camera position as the apex, and the wall (a rectangle) as the base.

When projecting on the ground, the area of interest is first divided into grids of proper size. When each rectangular region of the grid is used instead of the wall, the same projection method for the wall described above can be used to render the texture map in the frame-buffer.

The position and size of the rectangular region are changing when the camera moves or rotates. the resolution of the texture map is kept roughly the same as the video image regardless of the size of the region, so that the details of the video image can be maintained while the memory requirement is kept at a minimum. To calculate the region of the projection on the ground, a transformation matrix is needed to project the corners of the video image to the ground:

$$M = P \times T \times R$$

where R and T are the rotation and translation matrices that transform the camera to the right position and orientation, and P is the projection matrix, which is

$$P = \begin{bmatrix} d & 0 & 0 & 0 \\ 0 & d & 0 & 0 \\ 0 & 0 & -d & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

where d is the distance between the camera and the projection plane (the ground).

While the camera is moving, it is possible to keep the previous textures and only update the parts where new images are available. In this way, a large region will be eventually updated when the camera pans over the area.

The zooming factor of the video camera can be converted to the field of view. Together with the position and orientation of the camera that are tracked by GPS, inertial devices, and pan-tilt readings from the camera, we can calculate where to put the video images. The position and size of the image can be arbitrary as long as it is along the camera viewing direction, with the right orientation and a proportional size.

3.3 Rendering

The rendering of the texture is done with our AR/MR rendering engine which is based on OpenSceneGraph. A two-pass rendering process is performed to remove part of the views blocked by the buildings.

In the first pass, all of the 3D objects in our database are disabled and only the camera image rectangle is in the scene. The rendered image is grabbed from the frame-buffer. Thus a projected image of the video is obtained. In the second pass the camera image rectangle is removed from the scene. The grabbed image in the first pass is used as a texture map and applied on the projection plane (the ground or the walls). All the 3D objects in the database (mainly buildings) are rendered as solid surfaces with a predefined color so that the part on the projection plane that is blocked is covered. The resulting image is read from the frame-buffer and used as texture map in Google Earth. A post-processing stage changes the blocked area to transparent so that the satellite/aerial photos on Google Earth are still visible.

3.4 Google Earth interface

Google Earth uses KML to overlay placemarks, images, etc. on the virtual globe. 3D models can be built in Collada format and displayed on Google Earth. A Google Earth interface module for our MR system has been developed. This module is an hyper-text transfer protocol (HTTP) server that sends icons and image data to Google Earth. A small KML file is loaded into Google Earth that sends update requests to the server at a certain interval, and updates the received icons and images on Google Earth.

4. Results

An information integration prototype module with the Battlefield Augmented Reality System (BARS) (Livingston et al., 2004) has been implemented. This module is an HTTP server implemented in C++ that sends icons and image data to Google Earth. The methods are tested in a typical urban environment. One user roams the area while another object is a fixed pan-tilt-zoom network surveillance camera (AXIS 213 PTZ Network Camera) mounted on top of the roof on a building by a parking lot. This simulates a forward observation post in military applications or surveillance camera in security applications. The command and control center is located at a remote location running the MR application and Google Earth. Both the server module and Google Earth are running on a Windows XP machine with dual 3.06 GHz Intel Xeon CPU, 2 GB RAM, and a NVIDIA Quadro4 900XGL graphics card.

Fig. 1. Video image of the parking lot and part of a building from a surveillance video camera on the roof top.

The testing area is a parking lot and some buildings nearby. Figure 1 is the video image from the roof top pan-tilt-zoom camera when it is pointing to the parking lot. One of the parking lot corners with a building is in the camera view. Another AR user is on the ground of the parking lot, the image captured by this user in shown in Figure 2 which shows part of the building.

Google Earth can display 3D buildings in this area. When the 3D building feature in Google Earth is enabled, the final result is shown in Figure 4. The images are projected on the buildings as well as on the ground and overlaid on Google Earth, together with the icon of an AR user (right in the image) and the icon representing the camera on the roof of the building (far left in the image). The parking lot part is projected on the ground and the building part

Fig. 2. Image from a AR user on the ground.

Fig. 3. Image of the target zone on Google Earth.

Fig. 4. Recreated 3D scene viewed with 3D buildings on Google Earth. The two field operator's icons and the video image are overlaid on Google Earth.

(the windows, the door, and part of the walls) is projected on vertical polygons representing the walls of the building. The model of the building is from the database used in our AR/MR system. When the texture was created, the part that is not covered by the video image is transparent so it blended into the aerial image well. The part of the view blocked by the building is removed from the projected image on the ground.

Google Earth supports 3D interaction; the user can navigate in 3D. This gives the user the ability to move the viewpoint to any position. Figure 4 is from Google Earth viewed from an angle instead of looking straight down. This third-person view is very suitable in command and control applications. The projected images are updated at a 0.5 second interval, so viewers can see what is happening live on the ground. It needs to point out that the 3D building information in Google Earth is not very accurate in this area (especially the height of the buildings), but is a good reference for our study.

The result shows the value of this study which integrates information from multiple sources into one mixed environment. From the source images (Figure 1 and Figure 2), it is difficult to see how they are related. By integrating images, icons, and 3D model as shown in Figure 4, it is very easy for the command and control center to monitor what is happening live on the ground. In this particular position, the AR user on the ground and the simulated forward

observation post on the roof top can not see each other. The method can be integrated into our existing AR applications so that each on-site user will be able to see live images from other users' video cameras or fixed surveillance cameras. This will extend the X-ray viewing feature of AR systems by adding information not only from computer generated graphics but also live images from other users in the field.

5. Discussion

The projection errors on the building in Figure 4 are pretty obvious. There are several sources of errors involved. One is the accuracy of the models of the buildings. More serious problems come from camera tracking, calibration, and lens distortion. The lens distortion are not calibrated in this study due to limited time, which is probably one of the major causes of error. This will be done in the near future.

Camera position, orientation, and field of view calibration is another issue. In our study, the roof top camera position is fixed and surveyed with a surveying tool, it is assumed that it is accurate enough and is not considered in the calibration. The orientation and field of view were calibrated by overlaying the video image on the aerial photo images on Google Earth. The moving AR user on the ground is tracked by GPS and inertial devices which can be inaccurate. However in a feature-based tracking system such as simultaneous localization and mapping (SLAM) (Durrant-Whyte & Bailey, 2006), the video sensors can be used to feed Google Earth and accuracy should be pretty good as long as the tracking feature is working.

The prerequisite of projecting the images on the wall or other 3D objects is that a database of the models of all the objects is created so that the projection planes can be determined. The availability of the models of such big fixed objects like buildings are in general not a problem. However there is no single method exist that can reliably and accurately create all the models. Moving objects such as cars or persons will cause blocked parts that can not be removed using the methods that are used in this study. Research has been done to detect moving objects based on video images (Carmona et al., 2008). While in theory it is possible to project the video image on these moving objects, it is not really necessary in our applications.

Google Earth has 3D buildings in many areas; this information may be available for Google Earth users and thus could be used for the calculations. The accuracy of Google Earth 3D buildings varies from place to place; a more accurate model may be needed to get desired results. Techniques as simple as manual surveying or as complex as reconstruction from Light Detection and Ranging (LIDAR) sensing may be used to generate such a model. Many studies have been done to create urban models based on image sequences (Beardsley et al., 1996; Jurisch & Mountain, 2008; Tanikawa et al., 2002). It is a non-trivial task to obtain these attributes in the general case of an arbitrary location in the world. Automated systems are an active research topic (Pollefeys, 2005; Teller, 1999), and semi-automated methods have been demonstrated at both large and small scales (Julier et al., 2001).

6. Future work

This is a preliminary implementation of the concept. Continuing this on-going effort, the method will be improved in a few aspects. This includes registration improvement between our exiting models and the Google Earth images as well as the calibration issues noted above. The zooming feature of the camera has not been used yet, which will require establishing

a relation between the zooming factor and the field of view, another aspect of camera calibration. Other future work includes user studies related to effectiveness and efficiency of the system in terms of collaboration.

Currently when the texture map is updated, the old texture is discarded, it is possible to keep the previous textures and only update the parts where new images are available. In this way, a large region will be eventually updated when the camera pans over a larger area.

There are a few aspects contributing to the error of the system that should be addressed in the future. This will be done in the near future.

7. Conclusion

In this preliminary study, the methods of integrating georegistered information on a virtual globe is investigated. The application can be used for a command and control center to monitor the field operation where multiple AR users are engaging in a collaborative mission. Google Earth is used to demonstrate the methods. The system integrates georegistered icons, live video streams from field operators or surveillance cameras, 3D models, and satellite or aerial photos into one MR environment. The study shows how the projection of images is calibrated and properly projected onto an approximate world model in real time.

8. References

Azuma, R. T. (1997). A Survey of Augmented Reality, *Presence* 6: 355–385.
    URL: *http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.30.4999*
Bane, R. & Hollerer, T. (2004). Interactive Tools for Virtual X-Ray Vision in Mobile Augmented Reality, *ismar* 00: 231–239.
    URL: *http://dx.doi.org/10.1109/ISMAR.2004.36*
Beardsley, P. A., Torr, P. H. S. & Zisserman, A. (1996). 3D Model Acquisition from Extended Image Sequences, *ECCV (2)*, pp. 683–695.
    URL: *http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.29.4494*
Carmona, E. J., Cantos, J. M. & Mira, J. (2008). A new video segmentation method of moving objects based on blob-level knowledge, *Pattern Recogn. Lett.* 29(3): 272–285.
    URL: *http://dx.doi.org/10.1016/j.patrec.2007.10.007*
Colvin, R., Hung, T., Jimison, D., Johnson, B., Myers, E. & Blaine, T. (2003). A dice game in third person augmented reality, *Augmented Reality Toolkit Workshop, 2003. IEEE International*, pp. 3–4.
    URL: *http://dx.doi.org/10.1109/ART.2003.1320416*
Dubois, E., Truillet, P. & Bach, C. (2007). Evaluating Advanced Interaction Techniques for Navigating Google Earth, *Proceedings of the 21st BCS HCI Group Conference*, Vol. 2.
Durrant-Whyte, H. & Bailey, T. (2006). Simultaneous localization and mapping: part I, *IEEE Robotics & Automation Magazine* 13(2): 99–110.
    URL: *http://dx.doi.org/10.1109/MRA.2006.1638022*
Fröhlich, P., Simon, R., Baillie, L. & Anegg, H. (2006). Comparing conceptual designs for mobile access to geo-spatial information, *MobileHCI '06: Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, ACM, New York, NY, USA, pp. 109–112.
    URL: *http://dx.doi.org/10.1145/1152215.1152238*

Ghadirian, P. & Bishop, I. D. (2008). Integration of augmented reality and GIS: A new approach to realistic landscape visualisation, *Landscape and Urban Planning* 86(3-4): 226–232.
URL: *http://dx.doi.org/10.1016/j.landurbplan.2008.03.004*

Hagbi, N., Bergig, O., El-Sana, J., Kedem, K. & Billinghurst, M. (2008). In-place Augmented Reality, *Mixed and Augmented Reality, 2008. ISMAR 2008. 7th IEEE/ACM International Symposium on*, pp. 135–138.
URL: *http://dx.doi.org/10.1109/ISMAR.2008.4637339*

Hartley, R. & Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*, 2 edn, Cambridge University Press.
URL: *http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/0521540518*

Henrysson, A. & Andel, M. (2007). Augmented Earth: Towards Ubiquitous AR Messaging, *Artificial Reality and Telexistence, 17th International Conference on*, pp. 197–204.
URL: *http://dx.doi.org/10.1109/ICAT.2007.48*

Höllerer, T., Feiner, S., Terauchi, T., Rashid, G. & Hallaway, D. (1999). Exploring MARS: Developing Indoor and Outdoor User Interfaces to a Mobile Augmented Reality System, *Computers and Graphics* 23(6): 779–785.

Honkamaa (2007). Interactive outdoor mobile augmentation using markerless tracking and GPS, *In Proc. Virtual Reality International Conference (VRIC)* .
URL: *http://virtual.vtt.fi/multimedia/publications/aronsite-vric2007.pdf*

Hugues, O., Cieutat, J.-M. & Guitton, P. (2011). GIS and Augmented Reality : State of the Art and Issues, *in* B. Furht (ed.), *Handbook of Augmented Reality*, chapter 1, pp. 1–23.
URL: *http://hal.archives-ouvertes.fr/hal-00595205/*

Julier, S., Baillot, Y., Lanzagorta, M., Rosenblum, L. & Brown, D. (2001). Urban Terrain Modeling for Augmented Reality Applications, *in* M. Abdelguerfi (ed.), *3D Synthetic Environments Reconstruction*, Kluwer Academic Publishers, Dordrecht, pp. 119–136.

Jurisch, A. & Mountain, D. (2008). Evaluating the Viability of Pictometry Imagery for Creating Models of the Built Environment., *in* O. Gervasi, B. Murgante, A. Laganà, D. Taniar, Y. Mun & M. L. Gavrilova (eds), *Lecture Notes in Computer Science*, Vol. 5072, Springer, pp. 663–677.
URL: *http://dblp.uni-trier.de/db/conf/iccsa/iccsa2008-1.html#JurischM08*

Lee, J., Hirota, G. & State, A. (2002). Modeling Real Objects Using Video See-Through Augmented Reality, *Presence: Teleoperators & Virtual Environments* 11(2): 144–157.

Livingston, M. A., Edward, Julier, S. J., Baillot, Y., Brown, D. G., Rosenblum, L. J., Gabbard, J. L., Höllerer, T. H. & Hix, D. (2004). Evaluating System Capabilities and User Performance in the Battlefield Augmented Reality System, *Performance Metrics for Intelligent Systems Workshop*, Gaithersburg, MD.

Livingston, M. A., Julier, S. J. & Brown, D. (2006). Situation Awareness for Teams of Dismounted Warfighters and Unmanned Vehicles, *Enhanced and Synthetic Vision Conference, SPIE Defense and Security Symposium*.

Milgram, P. & Kishino, F. (1994). A Taxonomy of Mixed Reality Visual Displays, *IEICE Transactions on Information Systems* E77-D(12).
URL: *http://vered.rose.utoronto.ca/people/paul_dir/IEICE94/ieice.html*

Milgram, P., Takemura, H., Utsumi, A. & Kishino, F. (1995). Augmented Reality: A Class of Displays on the Reality-Virtuality Continuum, *Proceedings of the SPIE Conference on Telemanipulator and Telepresence Technologies*, Vol. 2351 of *Proceedings of SPIE*, Boston, Massachusetts, USA, pp. 282–292.

OpenSceneGraph (n.d.). `http://www.openscenegraph.org/projects/osg`.

Phan, V. T. & Choo, S. Y. (2010). A Combination of Augmented Reality and Google Earth's facilities for urban planning in idea stage, *International Journal of Computer Applications* 4(3): 26–34.

Piekarski, W. & Thomas, B. H. (2003). Interactive Augmented Reality Techniques for Construction at a Distance of 3D Geometry, *7th Int'l Workshop on Immersive Projection Technology / 9th Eurographics Workshop on Virtual Environments*, Zurich, Switzerland.

Pollefeys, M. (2005). *3-D Modeling from Images*, Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Regenbrecht, H., Ott, C., Wagner, M., Lum, T., Kohler, P., Wilke, W. & Mueller, E. (2003). An augmented virtuality approach to 3D videoconferencing, *Mixed and Augmented Reality, 2003. Proceedings. The Second IEEE and ACM International Symposium on*, pp. 290–291.
URL: *http://dx.doi.org/10.1109/ISMAR.2003.1240725*

Salamin, P., Thalmann, D. & Vexo, F. (2006). The benefits of third-person perspective in virtual and augmented reality?, *VRST '06: Proceedings of the ACM symposium on Virtual reality software and technology*, ACM, New York, NY, USA, pp. 27–30.
URL: *http://dx.doi.org/10.1145/1180495.1180502*

Schall, G., Mendez, E., Kruijff, E., Veas, E., Junghanns, S., Reitinger, B. & Schmalstieg, D. (2009). Handheld Augmented Reality for underground infrastructure visualization, *Personal Ubiquitous Comput.* 13(4): 281–291.
URL: *http://dx.doi.org/10.1007/s00779-008-0204-5*

Tanikawa, T., Hirota, K. & Hirose, M. (2002). A Study for Image-Based Integrated Virtual Environment, *ISMAR '02: Proceedings of the 1st International Symposium on Mixed and Augmented Reality*, IEEE Computer Society, Washington, DC, USA, p. 225.

Teller, S. (1999). Automated Urban Model Acquisition: Project Rationale and Status, *Bulletin de la SFPT* (153): 56–63.
URL: *http://citeseer.ist.psu.edu/111110.html*

Urban Leader Tactical Response, Awareness & Visualization (ULTRA-Vis) (n.d.). http://www.darpa.mil/ipto/Programs/uvis/uvis.asp.

Wittkämper, M., Lindt, I., Broll, W., Ohlenburg, J., Herling, J. & Ghellal, S. (2007). Exploring augmented live video streams for remote participation, *CHI '07 extended abstracts on Human factors in computing systems*, ACM, New York, NY, USA, pp. 1881–1886.
URL: *http://dx.doi.org/10.1145/1240866.1240915*