

OFFLINE SPATIAL PANORAMIC VIDEO AUGMENTATION FOR VISUAL COMMUNICATION IN THE AEC INDUSTRY

Stéphane Côté, Bentley Systems, Québec, Canada, stephane.cote@bentley.com

Julien Barnard, Dept. of computer science, Université Laval, Québec, Canada

Rob Snyder, Bentley Systems, Lexington, KY, USA

Renaud Gervais¹, Bentley Systems, Québec, Canada

ABSTRACT: Renovation of existing buildings is a type of construction work that requires a very good understanding of the existing built environment. Accurate bidding and proper planning for such construction projects requires detailed information both from the 3D CAD model and of the existing building. The problem is both sources of information are heterogeneous: one is in the form of digital data (2D CAD drawings, 3D CAD models), while the other is the actual physical world. The CAD model and drawings deliver the designer's construction intent, while the physical world provides integration context. A user must do significant mental efforts to merge the 2 types of data and form a mental image of the work that is actually required. In this paper, we present a technique that enables the visualization of augmented scenes, packaged in a visualization application that can be used on site or off site by construction bidders, planners and workers. The building environment is first captured by walking in the area holding a panoramic video camera. The panoramic video stream is then post-processed to align each of its frames with a 3D model of the building. A user can then use the viewer to navigate the pre-recorded photo-realistic scene at any location and orientation along the camera paths, and augment it with the 3D model, to reveal hidden structure represented in the model, or 3D elements showing future portions to be added. Our method was tested on a building for which a detailed 3D CAD model (BIM) was available. Results show that the system could enable a better communication between the designer and the builder by displaying the designed construction intent in its real context, and therefore could allow more accurate bidding and work planning, and generally, could facilitate better understanding of the work to be done.

KEYWORDS: Augmented Reality, environment, panorama, CAD model, BIM, construction, renovation.

INTRODUCTION

Renovations and additions to existing buildings are types of work that are constrained by existing structures. For new structures to fit well with the existing environment, the CAD model needs to be designed with care and detailed knowledge of the existing building. Similarly, for accurate bidding and proper planning of such construction projects, detailed knowledge and understanding of both the CAD model and the existing building is required.

One source of difficulty that might arise is that both sources of information are heterogeneous: one is in the form of digital data (2D CAD drawings, 3D CAD models), while the other is the actual physical world. The CAD model and drawings deliver the designer's construction intent, while the physical world provides integration context. User interpretation involves the effort to merge the 2 types of data and form a mental image of the work required. The problem is one of communication: we need to present the real world and model data together in such a way that the work to be done can be visually understood.

Augmented reality could provide an interesting solution to that communication problem by allowing real time navigation and interaction in an information environment that combines the physical world and the CAD data: a user would walk around, or in, the building, raise a handheld device, and see the surrounding scene augmented with aspects of the 3D model relevant to the planned work. Such an application would enable the visualization of the augmented environment from any location and would facilitate the understanding of the relations between the physical world and the proposed designed world. However, to be of any use, the augmentations would need to be very accurate. Indeed, decisions taken by architects, engineers and construction workers have a direct impact on public safety. Those must therefore be supported by accurate data. Although approximate, low accuracy augmented reality is easy to achieve, live accurate augmentation is still a challenge. In addition, augmented reality would not solve all problems:

¹ Current affiliation : INRIA, Bordeaux, France.

- Augmented reality is used to augment the live environment. That means users have to be physically located on site in order to see that site in an augmented fashion. They actually need to be located at the exact position from which they want to see the world augmented. Although that makes sense and can probably be achieved easily for typical city environments, it may not always be ideal for engineering / construction applications. For instance, problems may arise when a user wants to augment the building from a location where it is unsafe or impossible to stand (ex: nuclear reactor, high voltage area, etc.).
- Remote workers who may want to view the site in an augmented fashion will only be able to do so if they travel to the site and augment it in person. They could remotely view a live augmented session done by a colleague located on site – but that requires someone there to assist them when they need it, and they still need to give them navigation instructions that could be complex to explain verbally on the phone.
- Other users might also want to review an augmentation session they did in the past, to check a measurement for instance, something that live AR cannot provide.
- Sometimes they might want to check some augmentation carefully, without having to hold a handheld device in their hands, in the direction of the augmented area.
- Live augmentation requires the tracking of the position and orientation of the hand held device, something that is hard to achieve precisely and accurately in real time. Pre-recorded environments avoid that problem by post-processing all frames after capture, leaving time to accurately align the model with each frame of the video, ensuring stable and accurate augmentations regardless of users' movements.
- Also, while construction communications are being developed by the design team, pre-recorded photographic environments may indeed become a very important aspect of a combined information environment within which designers author their instructions, anticipating that those communications will be viewed later, onsite, immersively, to help people understand the work they are doing. In this sense, offline pre-recorded video may become an essential part of the information environment within which designers and builders author communications, receive them, and interpret and discuss them.

In many cases, and for the reasons stated above, such pre-recorded environments may become preferable to augmented reality. For both authoring and viewing, what would be needed is an augmentation technique that enables users to view the augmented environment remotely without needing assistance from anyone on site, without being constrained to navigating only on physically walkable areas, while being sufficiently immersive to give them the illusion that they are physically on site.

In this paper, we present a technique that enables the augmentation of pre-recorded scenes, packaged in a visualization application that can be used on site or off site by designers, construction bidders, planners and workers. The building environment is first captured using a panoramic video camera that is moved along a path in the area to be augmented. The video stream is then post-processed to align each of its frames with a 3D model of the building. This alignment makes the augmentation possible: a user can then use the viewer to navigate the pre-recorded photo-realistic scene at any location and orientation along the camera paths, and augment it with the 3D model. Further, the combined environment then can be transformed in ways that support specific, clear communication. Our method was tested on a building for which a detailed 3D CAD model (BIM) was available. Results show that the system enables a better communication between the designer and the builder by displaying the 3D design in its real context, and therefore would allow simpler, less expensive and more accurate bidding and work planning, and generally allows better understanding of work that is to be done, which may improve the efficiency, and correctness of performed work.

Our main contribution is a system for the capture and visualization of panoramic video environments that overcomes several of the limitations of augmented reality in the construction context, whilst providing an AR-like experience.

RELATED WORK

Over the past few years, the potential of using panoramic imagery for augmenting scenes has been studied (Côté 2011a, 2011b, 2012; Wither et al., 2011). Langlotz et al (2012) demonstrated live augmentation of pre-recorded video from a fixed position. Hill et al (2011) showed a mirror world augmentation system in which pre-captured panoramic images of the environment were augmented when the user stood at approximately their capture position. Static images offer the advantage of providing precise augmentation (since no camera

tracking is required). Augmentation based on panoramic media also has the potential of being much more accurate because of the numerous points of control located all around the camera that can be tracked over long distances (Lemaire and Lacroix, 2007) and increase the chance of capturing areas of the environment that are suitable for tracking (Wither et al., 2011).

While being frozen in time, augmentation based on static images has been used extensively for advertising (ex: Zhu et al., 2004). In such applications, augmentation is achieved on a paper marker, resulting in an augmentation that is often disconnected from the physical reality around the user. Yet, those applications are popular and serve their purpose. Such offline augmentation could very well find applications in the AEC world too, as it offers the advantage of providing users with the benefits of on-location augmented reality, while also providing access to real-world information not normally available when remote (Mankoff et al. 1998).

The use of augmented reality in the engineering world has also been investigated by several teams. For instance, Woodward et al. (2010) present a mobile AR system for viewing BIM data on construction sites. Schall et al. (2010) describe a technique for visualizing underground infrastructure using AR, and Su et al (2013) looked at the uncertainty related with such visualization. Georgel et al. (2007) proposed a technique for industrial discrepancy check using augmented reality. The vast majority of the works related with augmented reality in the engineering world investigate live augmentation, tracking quality being the most common concern. In this work, we chose to augment static imagery, to the benefit of augmentation precision, accuracy, and offline visualization.

METHOD

Data

Our method is based on the augmentation of pre-recorded panoramic video streams captured along a path in the area to be augmented. Panoramic streams are captured using a Ladybug panoramic video camera from Point Grey Research. For outdoor scenes, the camera is installed on top of a tripod and carried by a user who also carries a laptop using a shoulder laptop holder. For indoor scenes, the tripod is installed on a dolly onto which the laptop is also installed, and that is moved inside the building. The laptop runs a program developed in-house using the Ladybug SDK that basically captures the individual video streams obtained from the 6 individual cameras sensors, merges them into one single equirectangular panoramic video stream, and stores them on disk in real time as 6 faces of a cube.

Panoramas are augmented using a detailed 3D CAD model that was imported in our augmentation program. In addition to the usual elements (walls, floor, etc.), the model also contains numerous elements representing hidden objects like structure, pipes, etc. The augmentation is used to render those objects that are normally invisible in the physical world because of wall surfaces.

Publication process

The first step in the augmentation process is to align the CAD model to the panorama. This is done manually using the technique described in (Poirier, 2011): the pre-recorded panorama stream is loaded by the program, its first frame displayed on screen. The CAD model is then overlaid on top of it. Through a manual user interface, the user selects a set of control points on the model, and the corresponding locations of those points in the panoramic image. The camera pose is then calculated using a panoramic pose calculation algorithm, as described by Poirier. Since each frame of the video needs to be aligned with the model, we use a panoramic image tracking algorithm described in (Côté et al, 2013): features are identified in the first image and their 3D position is estimated by projection onto model faces from the camera position. These features are then matched with those in the second frame. Their 3D position is inferred from that correspondence, and the new camera pose is calculated using the same panoramic pose estimation algorithm. This process is repeated for each frame, resulting in a stream of camera poses. In other words, the process results in the calculation of the camera position for each individual frame of the panoramic stream. Drift in the tracking process can occur, but its extent is minimized through the use of key frames. Remaining drift manifests itself by a misalignment between the model and the panoramic image. A new set of correspondence points between the model and the panorama is then used to fix the alignment and pursue tracking.

Once the alignment is complete, the set can be exported into an "augmentation package" that consists of an augmentation program, a set of panoramic streams, the camera pose for each frame of those streams, and an

augmentation model.

Navigation and augmentation

The augmentation program included in the package basically consists of a panorama visualization tool: the first frame of the panorama stream is displayed on screen in a first person perspective (see Figure 1, left, top part) called the “augmented view”. A virtual camera is centered in that panorama, and renders it in the view. The user can rotate the view to look around and see the whole environment around the camera by dragging in the top view with a mouse pointer. Since the panorama stream is actually captured along a path, going to the next frame actually means moving the virtual camera to the next position along that path. This can be done by clicking on the virtual arrows displayed above the ground in the augmented view or on a dedicated key on the computer keyboard, which enables the user to move forward and backward along the camera path and see the environment from different vantage points.

To help the user orient himself in the virtual world, a map is also optionally displayed. The map is composed of the 3D building model, the camera path, and the current virtual camera position (see bottom part of Figure 1, left, and Figure 1, right). The user can change the position of the virtual camera by dragging its icon along the displayed path.

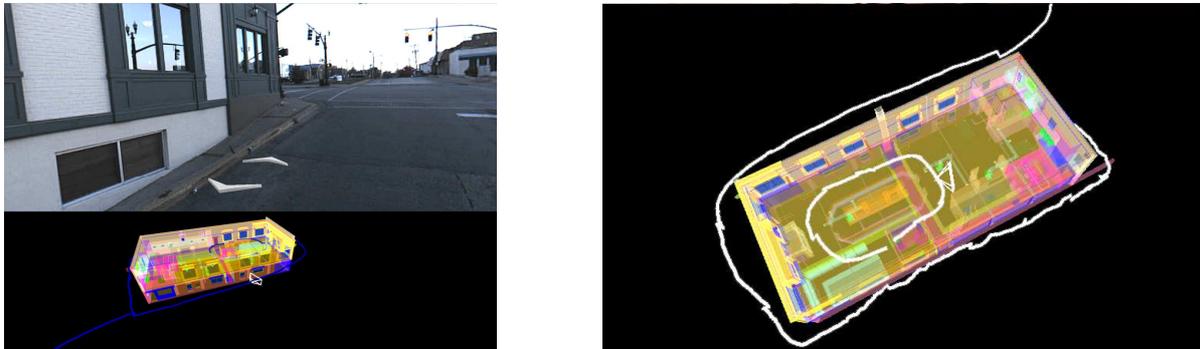


Figure 1: Prototype view showing map, paths, camera position, translation arrows, and scene view (left). Model, seen from the top, showing 2 camera paths and virtual camera frustum (right).

In our current prototype, augmentation is achieved through a virtual excavation feature, that lets the user see through walls, floor and ceiling. The augmentation technique is similar to the one described by (Schall *et al.*, 2010; Côté, 2011b) for augmenting subsurface utilities. In these projects, a virtual excavation is drawn on the surface of the road, revealing hidden infrastructure (see Figure 2). The technique basically consists of creating a virtual hole in an object, by clipping all but some of the elements composing it. In this project, we used the same technique but applied it on walls, floors and ceilings.

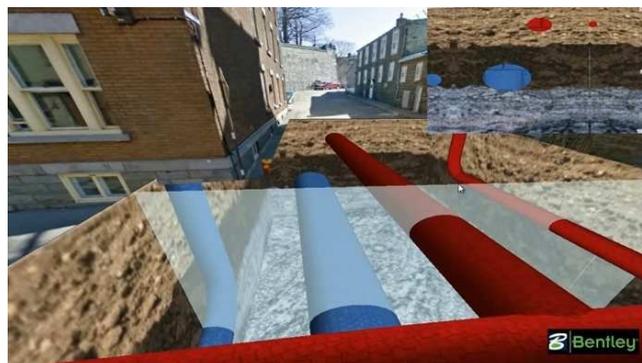


Figure 2: Virtual excavation for subsurface utilities as shown in (Côté, 2011b).

RESULTS

We tested our method in the Paddy Wagon Irish Pub located in Richmond, KY, USA (Figure 3, left). We chose that site because we also had a detailed CAD model (BIM) of that building (Figure 3, right), created by *McKay*

Snyder Architects, James McKay, Architect, using MicroStation® and had permission to use it given by the building owner. We captured 2 panoramic video streams, one inside the pub, the other one outside (see white paths in figure 1, right). Each individual panoramic frame was stored as 6 individual 1024 x 1024 pixel panels forming the sides of a cube. The augmentation package was then created, and the method tested.

The prototype user interface, shown in Figure 1 (left), enables the user to navigate in the environment while seeing his position on a 3D map. When panning the top view, the user actually rotates the virtual camera and can view the environment in all orientations. Clicking on the arrows displayed just above the ground moves the camera along the path. Zoom is also enabled. Manipulation on the bottom part of the view enables model rotation and zoom, as well as camera translation (by dragging the camera frustum icon along the paths).

Augmentation is achieved by right clicking on the top view and dragging to create a rectangle on the surface of a wall. When applied onto the outside walls of the building, it reveals the model interior (see Figure 4). Once the excavation is drawn, the camera can be moved along the path, revealing a different perspective both outside and inside the building. The excavation can also be moved along the wall, revealing other portions of the model.



Figure 3: The Paddy Wagon Irish Pub (left) and its CAD model (right).

Although the excavation is a useful tool, it is often too drastic. If, for instance, one would be interested in viewing the network of pipes hidden inside the wall, a virtual excavation that clips all elements is useless, as it would also clip the pipes that one wants to view. The excavation object is actually a 3D volume attached to a 3D surface (the wall). Clipping away all the elements inside the excavation volume may be too much, as it cuts everything out. What would be desired is a way to select which elements are to be clipped by the excavation. We implemented a selective clipping tool and tested it on our dataset. In Figure 5, the left image shows the excavation applied onto an external wall, the one on the right shows the same excavation, but with 2 pipes (purple and red) not clipped by the excavation.

During our tests, we encountered several issues. For instance, when using the program, it became clear that the loading of sequential panoramic images when navigating along a path was too slow. Each image frame being 6 megapixels large, loading it and updating the display at a speed required by normal navigation in the building environment was slow and made navigation painful. We partly fixed that problem by keeping a buffer of panoramic frames into memory, and by storing a low resolution version of the panoramas. When the user keeps the traveling key or button down, the system simply loads the lower resolution versions of the panoramas and updates them, which produces a much smoother and faster movement. When he releases the key, the system automatically stops camera movement and loads the higher resolution version of the destination panoramic frame. That turned out to provide a much quicker navigation.

Another issue was related with the image frames' spatial density being higher than necessary: capturing 15 FPS while walking at a speed of about 3 km/h results in frames captured at camera positions only about 5 cm away from each other. For outdoor augmentation, such density is not really necessary, and results in an unnecessarily large augmentation package. Moreover, the spatial frame density varied along the path because the user carrying the camera cannot easily walk at constant speed. To fix that and provide a visualization path with more regularly spaced frames, we implemented a basic downsampling algorithm that automatically selects a subset of the frames that are equally distant based on a user-chosen target spatial density.

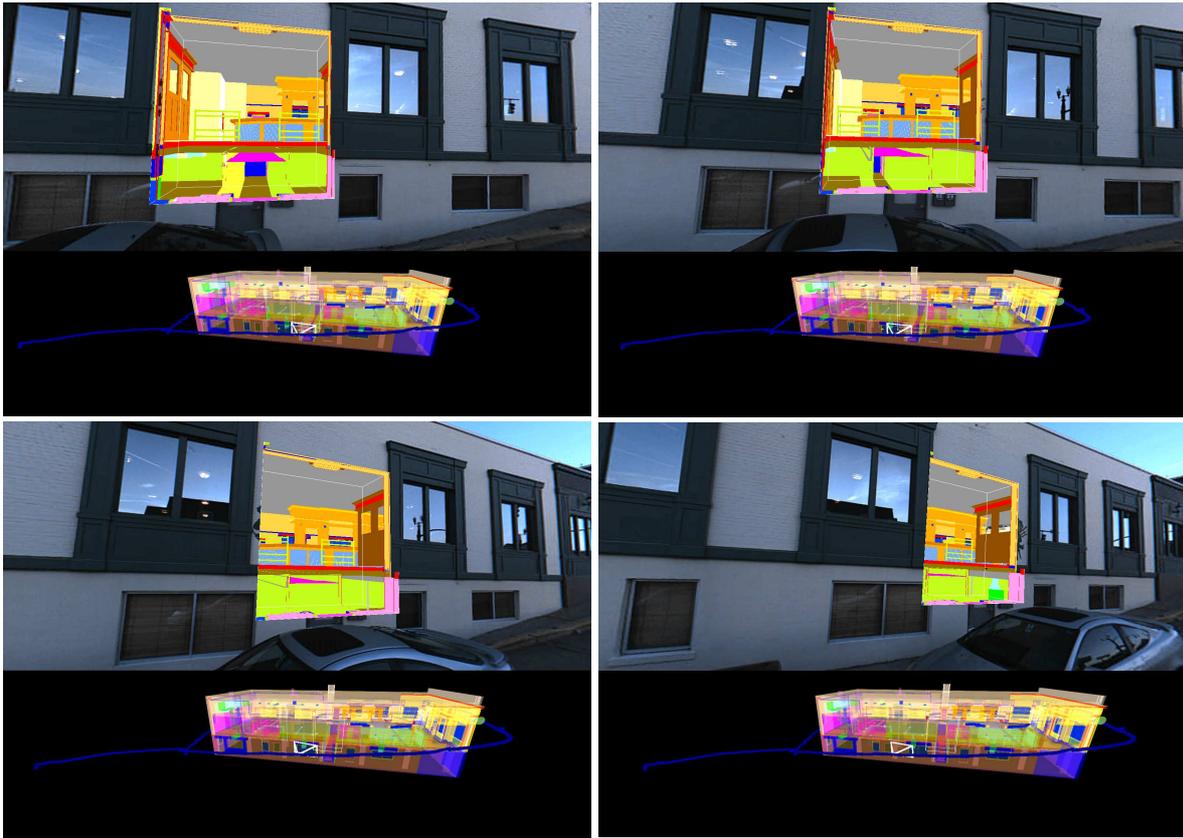


Figure 4: Virtual excavation displayed on external building wall, as the camera is being moved from right to left.



Figure 5: Virtual excavation showing selective clipping applied on selected pipes.

The use of the system showed stable augmentations not characterized by the typical jittering observed in typical augmented reality systems. Augmentation was always stable, although some inter-frame jittering was observed, which was probably due to our tracking algorithm that was not sufficiently accurate. Once those tracking errors are removed, the resulting augmentation will be stable both temporally and spatially, and 100% deterministic.

DISCUSSION

Our experiment seems to confirm the potential of such tools for augmentation. The use of pre-recorded media allows prior accurate alignment between the model and the physical world, providing precise augmentation and potentially much more accurate than live augmentation techniques. Moreover, pre-recorded media allowed us to augment the site from a remote place, which could be very useful for planning site visits, or for doing a quick check on something one might have missed during the first visit, but without returning to the site. Finally, such augmentation of pre-recorded media contributes to the authoring of communications that deliver designed construction intent, communications within the information environment that show, explain, clarify, communicate and direct, with regard to work to be done. The use of panoramic video captured along a path

ensures minimal navigation constraints in spite of the fact that the media is pre-recorded. In other words, the user benefits from the immersive aspects of live augmentation, in addition to all the advantages of offline augmentation.

Unfortunately, the system is not perfect. The use of pre-recorded media implies, by definition, that the media is out of date. That means any dynamic event cannot be reproduced in the augmentation. The tool would therefore need to be seen as a package for offline augmentation, and used accordingly. Of course, this is suitable for primary design construction intent communications, where as always, it is understood that the real world may deviate over time from the design communications (like drawings). Those deviations are typically detected and understood by people, and usually do not invalidate the construction intent communications.

We see this prototype as a first step to achieving better communication in AEC. The combination of model data with the physical world is in its infancy, and is likely to develop into a set of tools that will not only facilitate communication, but also will help the construction worker directly in his building tasks.

Several minor improvements could be made to the prototype. Those include:

- Increasing the panorama rendering speed, to improve the navigation experience.
- Integrating other augmentation techniques, such as clicking on objects to identify them and display maintenance history and other element data, displaying plan and section drawings, etc.
- Using the tool for onsite augmentation. The system could automatically track the user's position and display the closest panorama that was captured in the area. Although this would only simulate true augmented reality, the resulting augmentation would have the advantage of being very precise and much more accurate than current outdoor live AR solutions.
- Using the tool for collaborative remote augmentation, where users could augment the area from their respective position and add annotations on this augmented environment, all at the same time.
- Display the camera path in the augmentation window, giving the user a better sense of where he can navigate in the physical world (see Figure 6, left).
- When used outdoor, the camera was carried by a user walking around the building. Although care was taken to walk in a smooth fashion, the camera moved up and down in the process (see Figure 6, right). This movement manifests itself as an apparent "shaking" in the resulting stream, when navigating from frame to frame. We could evaluate downsampling methods to properly select frames that would minimize that movement.
- Another issue we encountered is the fact that the tracking module did not always succeed in finding the exact camera pose. That resulted in temporal misalignment that is very clear when moving the virtual camera during augmentation. Work should be put on a better panoramic tracking method.

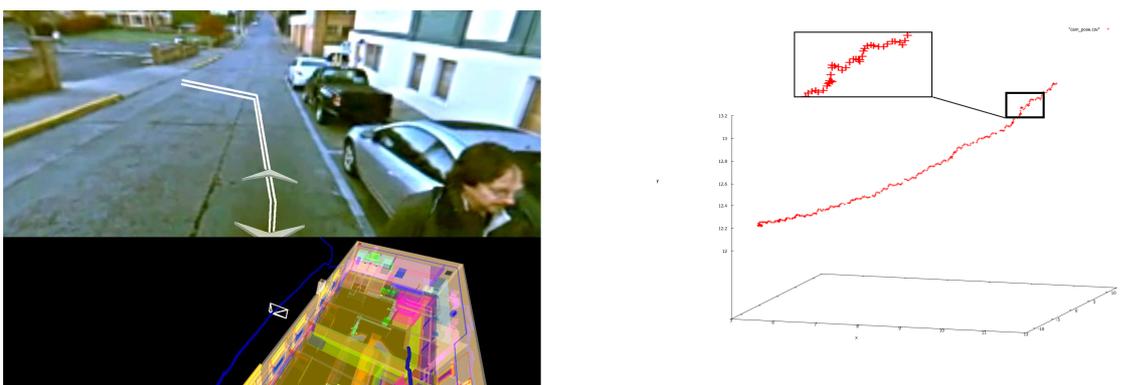


Figure 6: Displaying the path in the panoramic view as a proposed improvement (left). Captured camera 3D positions that show user's movements while walking (right).

However, the method could also be pushed one major step further. In this experiment, we showed that an immersive pre-recorded environment that can be browsed freely using a virtual excavation and packaged in a viewer that can be used offline shares many of the benefits of true live AR. Although such a tool is likely to be useful to the AEC industry, it is interesting to note that the purpose of design authoring of communications is to show, explain, and clarify, to communicate directly, about specific work to be done. In that context, a

browsing tool aimed at showing and revealing specific directive aspects and communications might be preferred. We envision that a directed browsing tool, showing users what to look at, understand, and do, would nicely complete the set of functionalities described in this paper, and would likely find more uses in the AEC world.

ACKNOWLEDGEMENTS

Many thanks to Mathieu Giguère and Marc-Antoine Desbiens for software development work, and to NSERC for financial support via the Industrial Undergraduate Student Research Award program.

REFERENCES

- Côté S. (2011a). Augmented reality for infrastructure: a first step. Published electronically on *BE Communities*.
- Côté S. (2011b). Augmented reality for underground infrastructure: the problem of spatial perception. Published electronically on *BE Communities*.
- Côté S. (2012). Augmented reality for building construction and maintenance: augmenting with 2D drawings. Published electronically on *BE Communities*.
- Côté S., Trudel P., Desbiens M., Giguère M., and Snyder R. (2013). Live mobile panoramic high accuracy augmented reality for engineering and construction. Proceedings of the Construction Applications of Virtual Reality (CONVR), London, England, Oct. 2013.
- Georgel P., Schroeder P., Benhimane S., Hinterstoisser S., Appel M., Navab N., 2007. An Industrial Augmented Reality Solution For Discrepancy Check. Proceedings of the 6th International Symposium on Mixed and Augmented Reality (ISMAR), Nara, Japan, Nov. 2007
- Hill A., Barba E., MacIntyre B., Gandy M., Davidson B. (2011). Mirror Worlds: Experimenting with Heterogeneous AR. 2011 International Symposium on Ubiquitous Virtual Reality, Jeju-si, Republic of Korea.
- Lemaire T. and Lacroix S., (2007). SLAM with panoramic vision. *Journal of Field Robotics*, Vol 24.
- Langlotz T, Zingerle M., Grasset R., Kaufmann H., Reitmayr G., 2012. AR Record&Replay: Situated Compositing of Video Content in Mobile Augmented Reality . OzCHI '12 Proceedings of the 24th Australian Computer-Human Interaction Conference.
- Mankoff J., Somers J. and Abowd G.D., 1998. Bringing People and Places Together with Dual Augmentation, Proceedings of the Collaborative Virtual Environments conference CVE'98.
- Poirier S. (2011). Estimation de pose omnidirectionnelle dans un contexte de réalité augmentée. M.Sc. Thesis, Université Laval, 2011.
- Schall G., Junghanns S. and Schmalstieg D. (2010). VIDENTE - 3D Visualization of Underground Infrastructure using Handheld Augmented Reality, in *GeoHydroinformatics: Integrating GIS and Water Engineering (CRC press)*.
- Su X., Talmaki S., Cai H., Kamat V.R. (2013). Uncertainty-aware visualization and proximity monitoring in urban excavation: a geospatial augmented reality approach. *Visualization in engineering*, Vol. 1.
- Witther J., Tsai W-T., Azuma R., 2011. Indirect augmented reality. *Computers & Graphics* 35, 810–822.
- Woodward, C., Hakkarainen, M., Korkalo, O., Kantonen, T., Aittala M., Rainio, K., Kähkönen, K., 2010. Mixed reality for mobile construction site visualization and communication. Proceedings of the Construction Applications of Virtual Reality (CONVR), Sendai, Japan, Nov. 2010.
- Zhu W., Owen C.B., Li H., Lee J.-H. (2004). Personalized In-store E-Commerce with the PromoPad: an Augmented Reality Shopping Assistant. *Electronic Journal for E-commerce Tools and Applications*.