Eurographics Italian Chapter Conference (2011)
A. F Abate, M. Nappi, and G. Tortora (Editors)

Augmented Reality Aided Maintenance for Industrial Applications

F. Narducci, S. Ricciardi

VRLab, University of Salerno, Italy

Abstract

This paper presents an augmented reality system suited for maintenance training and support in industrial environments. We describe how augmented reality can help in reducing the effort of performing sequences of maintenance tasks in complex systems. The proposed approach is aimed to facilitate items localization in technical equipment by means of context-based instructions, virtual labels, 2D and 3D graphics, and animated virtual tools. These virtual contents are transformed to be visualized, co-registered to the real world, through a see-through Head-Mounted Display (HMD). To the aim of enhancing tracking reliability, we report some effective techniques to address drift errors in head tracking. The interaction paradigm is thought to reduce the keyboard usage to make the system less intrusive during user's activity. Deterministic finite automaton (DFA) equivalent representation is exploited to regulate maintenance assistance operations, providing a versatile mean to define either simple or complex procedure in a easy, verifiable and readable way. We also propose a collection of XML tags which enable the conversion of DFA in XML files, ensuring high extensibility and ease of understanding. First experimental results, targeted to augmentation of industrial racks, show a measurable improvement in performing maintenance operations with regard to the time needed to complete the task.

Categories and Subject Descriptors (according to ACM CCS): H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities

---

1. Introduction

Thanks to continuous advances in technology and research [ABF*01], AR interfaces are becoming increasingly effective in providing users with a richer level of interactivity. AR supplements the real world with virtual (computer-generated) objects that appear to coexist in the same space as the real world. Over the years, researchers and developers find more and more areas that could benefit from augmentation. Design, assembly, and maintenance are typical areas where AR may prove useful. These activities may be augmented in both corporate and military settings.

Fiorentino et al. [FAS02] introduced the SpaceDesign MR workspace (based on the StudierStube framework) that allows for instance visualisation and modification of car body curvature and engine layout. Volkswagen is working on using AR for comparing simulated and actual crash test imagery. The MR Lab used data from Daimler-Chrysler‟s cars to create Clear and Present Car, a simulation where one can open the door of a virtual concept car and experience the interior, dash board layout and interface design for usability testing [TSK*05], [VR00]. In the assembly context, Pentenrieder et al. [PBDM07] shows how Volkswagen use AR in construction to analyse interfering edges, plan production lines and workshops, compare variance and verify parts. An extra benefit of augmented assembly and construction is the possibility to monitor and schedule individual progress in order to manage large complex construction projects. An example by Feiner et al. [FMH99] generates overview renderings of the entire construction scene while workers use their HMD to see which strut is to be placed where in a space-frame structure. Distributed interaction on construction is further studies by Olwal and Feiner [OF04].

AR is proving useful in the area of maintenance for instance in providing “x-ray vision” or automatically probing the environment with extra sensors to direct the users attention to problem sites. Klinker et al. [KBCW03] present an AR system for the inspection of power plants at Framatome ANP (today AREVA). Friedrich [FRI02] show the intention to support electrical troubleshooting of vehicles at Ford, while also Honda and Volvo adopted Nomad Expert Vision Technician systems to assist their technicians with vehicle history and repair information.

A further privileged field to employ AR is training/learning [BRI91]. Along a conceptually similar

DOI: 10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2011/067-072

line, related frameworks can be used for the simulation of production sequences [DFG*05], or even for the rapid evaluation of prototypes before production [BKFT00].

Maintenance and repair operations represent an interesting and opportunity-filled problem domain for the application of augmented reality. The majority of activities in this domain are conducted by trained maintenance personnel applying established procedures to documented designs in relatively static and predictable environments. These procedures are typically organized into sequences of quantifiable tasks targeting a particular item in a specific location. These characteristics and others form a well-defined design space, conducive to a variety of systems and technologies that could assist a mechanic in performing maintenance. AR research faces several technical challenges. Firstly, for any AR application is extremely important that real world and augmented contents are co-registered to ensure augmented information is displayed onto the real world with the correct size, perspective and orientation. This problem involves that tracking system must be reliable, robust, unobtrusive and interference tolerant. Another key aspect for AR fruition is represented by wearable see-through displays, which should ideally feature high resolution, wide field of view, binocular (stereo) view, high luminance and contrast, realistic colour rendition and wide focus depth. However, before AR becomes accepted as part of user's everyday life, just like mobile phones and personal digital assistants (PDAs), issues regarding intuitive interfaces, costs, weight, power usage, ergonomics, and appearance must also be addressed. In this paper, we describe how augmented reality can assist in reducing the effort of performing large sequences of maintenance tasks in complex systems. The proposed system exploits AR to enhance items localization in ordinary maintenance sequences by means of context-based instructions, visual aids, virtual labels, 2D and 3D graphics, and animated virtual tools. These information are transformed to be visualized onto the real world exploiting a tracked head-mounted display (HMD). Our user study demonstrates how technicians performing maintenance sequences within an AR environment are able to locate the items involved in the repair intervention more quickly and can operate more effectively . With regards to tracking stability, we propose a few techniques to reduce drift and loss of position.

The rest of the paper is organized as follows. Section 2 describes the design and performance of the proposed AR system; Section 3 describes the experiments conducted to assess the qualitative insights of these mechanics with regard to the intuitiveness, ease of use, and acceptability of our approach, and Section 4 draws some conclusions.

2. System description

The proposed system consists of three basic software modules, the AR Engine, the Tracking System and the Scene Capturing. These subsystems allow the user to see all required augmented contents according to his point of view. The system architecture is shown in Figure 1.

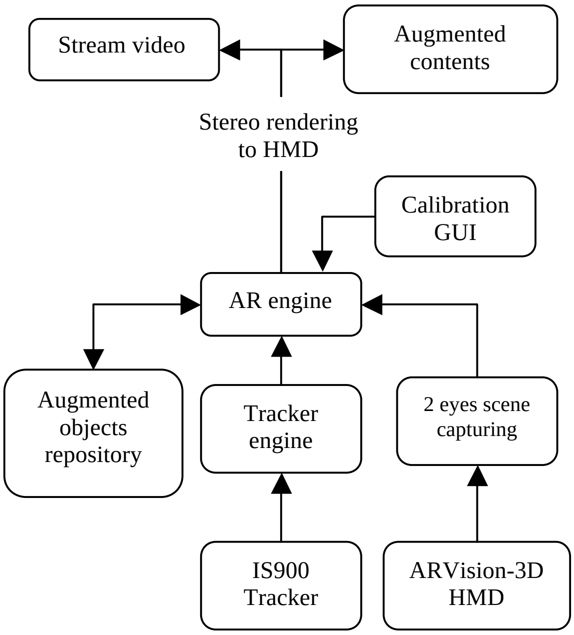


Figure 1 – *The overall system's architecture*

An IS 900-VET ultrasonic wireless tracking system (by Intersense Co.) provides user's head position and orientation within the covered 3D volume. This information is exploited to calculate the user's point of view. From it, the AR engine can load the proper augmented content from the repository. By this way the AR engine can merge virtual contents with video streams provided by two cameras included in the HMD. The rendering obtained is therefore outputted to the two HMD's displays. The calibration Graphical User Interface allows the user to manually correct tracking errors if, on the system start-up, virtual and real world should result not correctly aligned. The main hardware and software components of the system are discussed in the section below.

2.1 Tracking System

A basic requirement for any AR application is the tracking system. There are a lot of approaches to this topic, based on different technologies. The proposed architecture is based on the ultrasonic/inertial motion tracking system IS-900/VET (from Intersense Co.) as it features a wide scalable coverage volume and a wireless sensors option. Moreover, it's robust to electrical and magnetic fields, providing reliable operation even in situations where other tracking systems may be affected from external environment. The radio technology is based on 2.4 GHz frequency hopping radios that can support multiple separate base/remote pairs. According to technical data the tracking resolution is of the order of 3mm for position and 0.15 degree for rotation. However our study revealed that too often some drifts, vibrations and loss of signal affect tracking reliability. We have tried to address these issues developing some fast techniques to prevent errors. For a more clear comprehension of the adopted solutions is useful to explain how we can simply describe a drift. An ideal tracking solution should track the user movement in a continuous manner with infinite accuracy. The tracking curves resulting will be free of peaks and very smooth. We noted that the drift occurred often during the rest condition. In such situations the movement tracking curves should be nearly linear (only a robot could makes these curves perfectly linear). If, for example, a drift occurred for tracking position while the users is at rest, the relative curve will present a slow and constant progression, either positive or negative, over time (fig. 2). In an AR context the visual impact of this kind of error is very strong. A drift implies a shifting of augmented contents while the user's point of view in the real world is still the same.

Therefore, to prevent a drift it is necessary to avoid this phenomenon. The techniques adopted to detect drifts during the tracking has been obtained from computation of user's movements speed. When such speed is constant and lower than a threshold (generally a very small number) the tracker engine detects a drift. Until the drift detection condition is valid, the tracker engine ignores data stream from tracker and uses as current position the latest rest values. The choice of the threshold is the most important parameter for a correct detection. If this value is too high, the system will detect even voluntary user's movements as drift. It is sufficient that threshold isn't larger than few millimeters per second to grant a good drift detection. This strategy works as human users cannot perform movements with very high precision.

Vibrations and loss of signal are other two problems that undermine the tracking fidelity. A vibration is intended as a swift turbulence on the tracker values either for position and orientation. In presence of vibrations, the movement curve will have a high frequency and small amplitude (fig. 3). Figure 4 clearly depict why even small vibrations of the head tracker may cause relevant augmentation problems, particularly on rotational degrees of freedom.

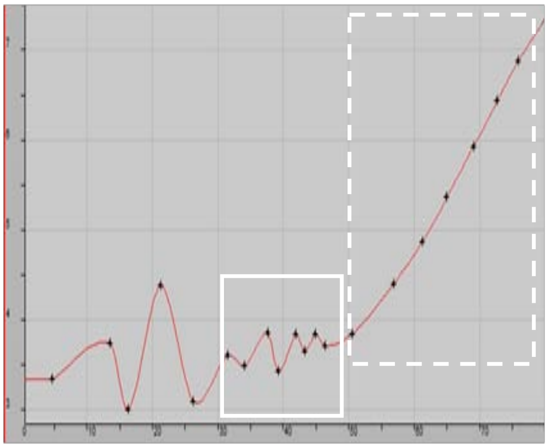


Figure 2 – *An example of tracking curve. After a period of rest (continuous line), occurred a drift (dashed line).*

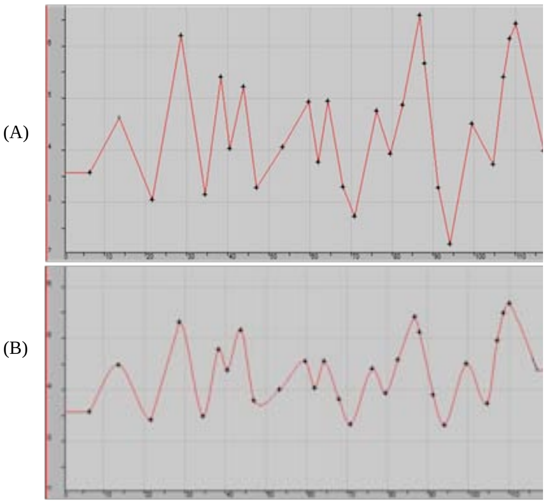


Figure 3 – *An example of movement's raw curve (A) and post-processed smoothen same curve (B).*

Consider the user focusing his attention on an augmentation point in the world when a vibration α occurs on Y rotation data stream. Suppose the observed point is at distance d'=0.50m (meter) or d''=1.00m. In this case the error amount for the same vibration α = 1 degree will be E = d tan(α). For our example E' = 0.008m and E'' = 0.017m. This means that at 50 centimeters of distance, the error amount will be of 8 millimeters, a value already not tolerable for precision applications. At distance of 1 meter the error of 17millimiters becomes unacceptable.

To fix this problem we used a damping function, able to smooth position and rotation curves based on a threshold value. If threshold is high the curve will result more damped. This means that the user perceives a latency in the augmentation of the scene.

As for drift threshold, it's important to define a proper value for vibrations correction. Referring to technical data the damping threshold need to be around 3 millimeter for positions and close to zero for rotations. In the same manner this damping solution prevents the loss of signal too. When the wireless receiver loses the data stream from wireless transmitter, the curve will have a drastic peak to zero.

By using a damping function, in this case, the curve tends to zero more slowly, giving the system the time needed to restore wireless communication from transmitter and receiver. If the communication restoring is rapid, the users will not notice the loss of signal.

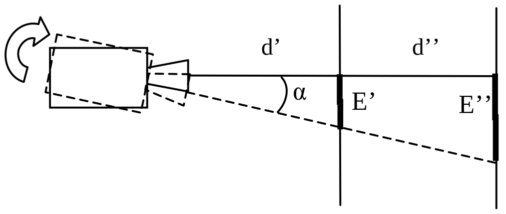


Figure 4 – How camera's *vibrations affect augmentation accuracy.*

2.2 Scene capturing

The surrounding scene is captured by means of two separate video stream outputted by two cameras mounted on the video see-through HMD (Trivisio ARVision-3D) shown in figure 5. Scene acquisition is performed at 30 FPS (Frame Per Second). Every frame from each of the two cameras is sent to the AR Engine which process it to finally visualize the augmented scene it on the relative LCD display.

2.3 The AR engine

The AR engine is the system's main software component. It is built on the Quest3D graphics programming environment. Quest 3D is a commercial authoring environment for real-time 3D applications featuring an "edit-while-executing" programming paradigm. The AR engine exploits the head tracking data to render the proper augmented object (labels, arrows, virtual tool etc), transformed according to the user's viewpoint onto the video streams. More in detail, the tracker samples user's head motion sending six data streams to the AR engine (XYZ position and pitch, yaw, roll rotation). The main task of AR engine is ensuring that virtual objects appear co-registered with real objects, i.e. properly visualized with respect to position, orientation, size and perspective. This non-trivial goal is achieved only if two basic constraints are satisfied: 1) the tracking of user's head position and rotation have to be accurate; 2) real world measurements and augmented contents measurements have to match. The first constraint has been already discussed in section 2.1. Fixed a common reference system and a common unit system, the latter constraint has been satisfied by specific XML descriptors which determine the augmented contents positioning in the 3D space. XML meta-language results very suited to content description as it's very extensible, easy to understand and non-ambiguous.

Two kind of basic XML files are exploited in the proposed system: 1) the *environment file* records the points in the world where the objects to be augmented are placed; 2) the *object file* defines the object's active hotspots, e.g. where augmented info is visualized or interaction is possible for that object.

Figure 5 – *Trivisio ARVision video see-through HMD.*

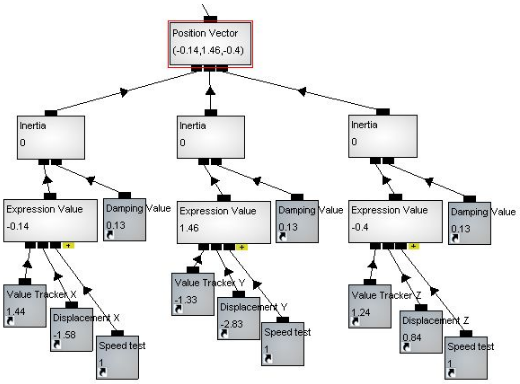


Figure 6 – *A screenshot of AR engine implementation by Quest3D programming environment. The graph provide to allocate a Value Vector with three position coordinates from tracker after drift and vibration correction.*

The syntax of these files is quite simple. An environment XML file starts with tag `<environment>` which follows a tag list of all objects to be augmented. For each is specified the root tag which must contain `<position>`, `<orientation>`, `<size>` tag each of which has its own attribute to specify position, rotation and size of the virtual object in the 3D space. The following mandatory tag refers to the XML object file placed in that 3D point. The syntax of an arbitrary object file consists essentially of a list of `<hotspot>` tags. Any hotspot tag has a *type* attribute aimed to object identification and an `<offset>` child which represents the relative object's position (the object's pivot) with respect to the environment's reference system.

The AR engine, according to the aforementioned descriptors, builds up the virtual scene by means of a DOM XML parser. To find the required data in the application database a XML-based *XPath* query language is used. Combining the information from tracker and from XML database, the AR engine is therefore able to locate in every moment the user in the real world and to extract all the required augmented contents from the repository.

Besides the augmentation task, the engine also performs another crucial task: the maintenance procedure management. A generic maintenance procedure can be represented as a finite deterministic automaton (DFA). According to this approach, a particular state represents a maintenance step and its links define the execution order. DFA result very suited to the maintenance context providing all the elements (states, links) required to represent both simple and complex maintenance procedure in a easy, verifiable and legible way.

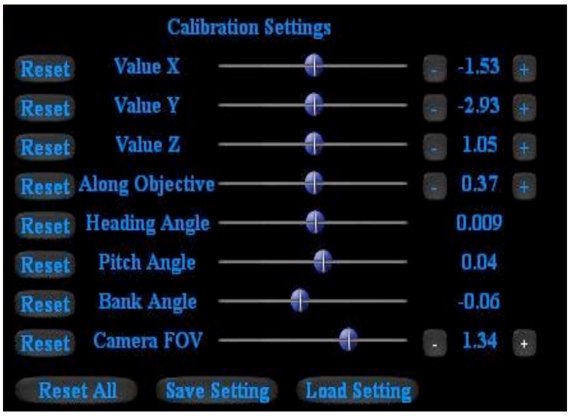
We found useful to convert the DFA representation of a particular procedure in a XML file where a tag `<step>` defines a state. Any possible path through the automaton defines a procedure's file. By this approach a single procedure file XML defines a single execution order of a maintenance procedure. This file is parsed from beginning to the end. In every moment the user can switch to the next or previous task, external tasks or jump to a particular tasks are possible, though not supported yet.

An example of a generic procedure step is shown below. This step refers to the device *Server.xml*. A label informs the user that there are two screws which have to be unscrewed. By parsing *Server.xml* the engine locates the proper screws and highlight them augmenting the real screws by means of blinking circles. As for the current step a screwdriver is required, the relative `<tool>` tag allows to specify it. Thanks to this tag the AR engine can load from the virtual content repository the proper 3D model and render it over the two screws to show how to perform the task.

```
<step equipmentRef="Server.xml">
   <label rgbText="default"
   rgbBackground="default">
      Unscrew the two fixing screws
   </label>
      <hotspot>screw</hotspot>
      <tool>screwdriver</tool>
</step>
```

Co-registration between real and virtual camera

We described in the previous sections a few techniques adopted to improve tracking precision and to prevent measurement errors. Nevertheless, a calibration GUI is necessary to achieve the accurate alignment between real and virtual camera. In figure 7 is shown the GUI. It consists of a set of parameters accessible through a corresponding set of sliders which enable the manual fine calibration of the virtual camera's position, orientation and Field of View (FOV). Apart from the controls for XYZ position and Heading/Pitch/Bank rotation, which are self-explanatory, another position control, referred as *Along Objective*, is available. This particular slider controls the perpendicular axis at the camera's view plane. By moving this slider a zoom in or out is performed as the camera moves closer to or away from the focused view point. The FOV slider allow the user to modify the virtual camera's field of view so that it and the resulting perspective transformation will match the real camera behavior. Though the calibration parameters can be saved and restored, often a new calibration is required when powering up the system after a previous usage. This is partly due to some technological limits of the ultrasonic tracking system which may affect the repeatability of measurements, eventually making the old calibration file useless. Another explanation is that each system's user may have a different height. The tallest users come closer to ultrasonic emitters which results in better tracking fidelity and minimum co-registration error.



Value X -1.53
Value Y -2.93
Value Z 1.05
Along Objective 0.37
Heading Angle 0.009
Pitch Angle 0.04
Bank Angle -0.06
Camera FOV 1.34

Figure 7 – *A screenshot of calibration GUI.*

3. First experiments: industrial racks maintenance

To the aim of testing the proposed system in facilitating the maintenance operations while shortening intervention duration, we experimented it in a real applicative scenario: industrial rack maintenance and repair. A typical rack may contain various equipment and each of them may involve led status indicators, buttons, handles, screws etc. The test-bed used for experiments is based on a PC featuring Intel dual core processor and Nvidia Quadro FX-500 series graphics board. This hardware setup has been capable to render in stereo the augmented scene featuring up to hundreds of thousands of polygons at an output resolution of 2x800x600 pixel (HMD's displays native resolution) providing a sustained frame rate always above 30 fps. A stereoscopic view provides a more effective way to show objects and locations in 3D space, particularly in situations where hotspots overlap each other. During the AR assisted intervention the user sees through the HMD the virtual contents rendered in stereo over the captured scene (see fig. 8a-8b). This design allows the user to focus on the relevant items to perform the required tasks. For each equipment involved, the application shows virtual labels explaining the item's reference number and other technical data. This information helps the user in locating defective items, performing functionality testing, removing and replacing items. The software is thought to avoid any usage of mouse and keyboard except for tracking calibration, procedure selection and navigation, though the latter can be easily operated via a wireless device. It has to be remarked that, though the processing hardware requirements are not demanding, the Intersense tracking system and the stereo-capable see-through HMD represent an expensive gear.

4. Conclusions and future enhancements

In this paper, we presented an AR system for maintenance support and training. The proposed architecture exploits an ultrasonic tracking system to provide the user with an augmented work area, featuring additional info, visual aids and context dependent virtual objects. A calibration GUI helps the user during the crucial co-registration process to improve augmentation accuracy. By means of a stereoscopic rendering displayed via a video see-through head mounted display, the user experiments the illusion that virtual and real objects coexist, enhancing his visual scene understanding and improving his confidence during task execution.

The first results obtained so far in an industrial environment shows an improvement in intervention accuracy and a reduction of overall intervention duration.

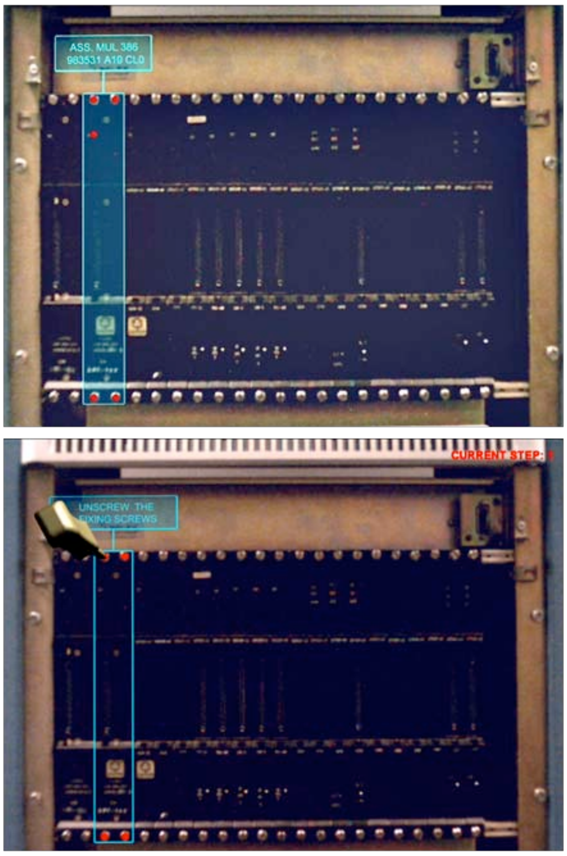


Figure 8a-8b – *A simulation of the AR environment during a procedure maintenance. By means of edges and label the users can understand and distinguish the equipment's assembly. When a tool like a screwdriver is needed, it is rendered over the interested screws.*

Even a user with little knowledge is able to complete simple maintenance task in a limited amount of time without consulting printed documentation. Though the techniques adopted to prevent tracking instability have proved to be effective, the risk of position/orientation drifts or instability could not be totally avoided. For this reasons we are currently evaluating a marker-based approach which could improve system reliability while making it much more affordable.

References

[ABF*01] Azuma, R. Baillot, Y. Behringer, R. Feiner, S. Julier, S. MacIntyre, B. (2001) Recent advances in augmented reality, IEEE Computer Graphics and Applications, Volume: 21, Issue: 6, 2001, pp.34-47

[BKFT00] S. Balcisoy, M. Kallmann, P. Fua, D. Thalmann (2000). A framework for rapid evaluation of prototypes with augmented reality, Proceedings of the ACM symposium on Virtual reality software and technology, Seoul, Korea, pp. 61 – 66, 2000

[BRI91] M. Bricken (1991). Virtual reality learning environments: potentials and challenger, ACM Computer Graphics, Volume 25 , Issue 3, pp. 178 – 184, 1991

[DFG*05] W. Dangelmaier, M. Fischer, J. Gausemeier, M. Grafe, C. Matysczok and B. Mueck (2005). Virtual and augmented reality support for discrete manufacturing system simulation. Computers in Industry, Volume 56, Issue 4, May 2005, Pages 371-383

[FAS02] M. Fiorentino, R. de Amicis, G. Monno, and A. Stork. Spacedesign: A mixed reality workspace for aesthetic industrial design. International Symposium on Mixed and Augmented Reality (ISMAR'02), 2002, pp. 86–318.

[FMH99] S. Feiner, B. MacIntyre, and T. Höllerer. Wearing it out: First steps toward mobile augmented reality systems. Ohmsha (Tokyo)-Springer Verlag, Berlin, 1999, pp. 363–377.

[FRI02] W. Friedrich, "ARVIKA-Augmented Reality for Development, Production and Service," Proc. Int. Symp. on Mixed and Augmented Reality (ISMAR '02), 2002, pp. 3–4.

[KBCW03] K. Kiyokawa, M. Billinghurst, B. Campbell, and E. Woods. An occlusion-capable optical see-through head mount display for supporting co-located collaboration. ISMAR 2003, pp. 133–141

[OF04] A. Olwal and S. Feiner. Unit: modular development of distributed interaction techniques for highly interactive user interfaces. ACM New York, NY, USA ©2004, pp. 131–138.

[PBDM07] K. Pentenrieder, C. Bade, F. Doil, and P. Meier. Augmented reality-based factory planning -an application tailored to industrial needs. In ISMAR'07: Proc. 6th Int'l Symp. on Mixed and Augmented Reality, pp. 1–9, Nara, Japan, Nov. 13-16 2007. IEEE CS Press. ISBN 978-1-4244-1749-0.

[TSK*05] M. Tönnis, C. Sandor, G. Klinker, C. Lange, and H. Bubb. Experimental evaluation of an augmented reality visualization for directing a car driver‟s attention. ISMAR '05 Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality, pp. 56–59.

[VR00] L. Vaissie and J. Rolland. Accuracy of rendered depth in head-mounted displays: Choice of eyepoint locations. In Proc. AeroSense, vol. 4021, pp. 343–353, Bellingham, WA, USA, 2000. SPIE Press.