

# Distributed User Tracking Concepts for Augmented Reality Applications

Gudrun Klinker, Thomas Reicher and Bernd Brügge

*TU Munich, Dept. of Computer Science, Chair for Applied Software Engineering, D-80333 Munich  
{klinker, reicher, bruegge}@in.tum.de*

## Abstract

*Augmented Reality (AR) constitutes a new user interface paradigm. Using light headsets and hand-held or worn computing equipment, users can roam their daily working environment while being continuously in contact with their computer systems. For AR to work properly in a large factory, new system architectures have to be designed with consideration to the special requirements imposed by AR. In particular, augmented reality requires real-time facilities to track the user's position and viewing direction. In the past, various carefully calibrated sensing devices have been used for this purpose including magnetic trackers and active LED-systems. Research is now focussing on computer vision based methods. It is our hypothesis that in the future, the most successful indoor approaches will combine local (user-worn) vision-based tracking methods with global user tracking schemes from fixed (wall-mounted) sensors, using mobile wireless networking technology to allow the user-worn system to communicate with the globally available sensing infrastructure in an intelligent building. We propose the concept of "AR-ready" intelligent buildings which provide built-in tracking services via different sensing modalities*

## 1. Introduction

Augmented Reality (AR) constitutes a new user interface paradigm. Using light headsets and hand-held or worn computing equipment, users can roam their daily working environment while being continuously in contact with the dynamically changing virtual world of information provided by today's multi-media networks. In many ways, AR is the logical extension to wearable computing concepts, integrating information in a more visual and three-dimensional way into the real environment than current text-based wearable computing applications. Adapted to the user's current location, task, general experience and personal preferences the information is visualized three-dimensionally and mixed with views of the real world.

For AR to work properly in a large factory, new system architectures have to be designed with consideration to the special requirements imposed by AR. In particular, augmented reality requires real-time facilities to user's

position and viewing direction. It also requires real-time information on the current three-dimensional structure of the dynamically changing world surrounding a visitor. Finally, it requires up-to-date information regarding the details of the currently pursued application. Such information must be visualized suitably and inserted into the user's current view of the scene.

It is unlikely that a local, user-worn system will be able to gather and analyse all relevant information independently by itself. In particular, globally consistent positioning – as achieved by GPS outdoors – is likely to be a complex issue indoors. Within buildings, we thus expect mobile AR-systems to become embedded in the architecture of the overall IT infrastructure. In this paper, we discuss architectural concepts of distributed, wearable AR applications. This analysis serves the purpose of laying out the decisive factors for a reasonable software architecture within which distributed AR components can be implemented and refined over time.

## 2. A Distributed Tracking Concept

To augment a user's view with computer-provided synthetic information, an AR-system must know very precisely where the user is and where he is looking. The virtual objects have to be rendered from the same perspective. In the past, various carefully calibrated sensing devices have been used for this purpose including magnetic trackers and active LED-systems [1, 2, 6, 20]. But the precision and working range of such devices is insufficient for most AR applications.

Research is now focussing on computer vision based methods. Currently, the most successful ones track optical markers in indoor demonstrations, using video input from a mobile camera to determine the user's position [12, 16, 18]. Yet, marker-based approaches are unlikely to be extensible to AR scenarios requiring users to operate in very large environments, e.g., a big industrial plant. New approaches are now exploring vision-based approaches which track naturally occurring landmarks in the scene [19, 16], thus getting by without requiring major modifications of the environment. But such systems require exponentially growing computing power when supplied with increasing numbers of natural features to be tracked. The initial question "where am I?"

requires that a set of currently seen landmark features be matched in a combinatorial search against an ever-growing database of registered features. It is our hypothesis that in the future, the most successful indoor approaches will combine local (user-worn) vision-based tracking methods with global user tracking schemes from fixed (wall-mounted) sensors, using mobile wireless networking technology to allow the user-worn system to communicate with the globally available sensing infrastructure in an intelligent building.

We propose the concept of “AR-ready” intelligent buildings which provide built-in tracking services via different sensing modalities. Long passage ways and halls through which people typically just transfer from one work site to another are equipped with low precision tracking devices, such as active badges [21], infra-red light barriers, and computerized room access control systems. At work sites where high precision tasks are performed, such as object assembly or quality assurance, stationary cameras are positioned at strategic places, providing the user-worn AR-system with precise location data regarding the user’s current head position.

The user-worn AR-system communicates with the intelligent building, receiving tracking data as well as other information related to pursuing the current task. The user position – and especially the current head orientation – is typically hard to compute precisely from a wall-mounted camera. The user-worn system thus fine-tunes the tracking data further with its own local tracking system, using the global tracking data for initialization. From such initialization, it can prune much of the search space of all registered landmarks in the entire building. After predicting the approximate location of a few, well selected landmarks in the head-mounted camera, it finds and tracks their exact location.

### 3. Distributed Tracking Algorithms

In a distributed user tracking approach, several stationary sensors in the environment each detect the user and communicate the tracking data to other system components. A joint analysis of all findings then determines the best estimate of the current user location.

#### 3.1. Separate User Detection and Tracking Processes of Individual Sensors

In our distributed tracking scenario, we expect various stationary sensors to be spread out throughout the plant. Each such sensor has been calibrated off-line, i.e., its position, orientation, and its sensing parameters (e.g., focal length, center and aspect ratio) and precision have been obtained with respect to a plant-wide global coordinate system. Calibration records of all sensors are

available to all computing devices on the company network.

When some of the stationary sensors detect a mobile AR-system (e.g., by detecting a well-defined tag on the user’s HMD), they make such information available within the computer network. For example, a camera can indicate that it has detected mobile unit  $u$  at time  $t$  at its image at pixel  $(x,y)$  with confidence  $c$ . Other sensing modalities, such as active badges, light barriers, or an electronic access control (key card) system, may provide other kinds of information, such as the detected existence of a mobile unit  $u$  at time  $t$  within an area of specified center  $(x,y,z)$ , radius  $r$  or extent  $(dx,dy,dz)$  with confidence  $c$ . As part of an ongoing tracking process, the mobile unit  $u$  itself also continuously transmits its own hypothesis as to its own position  $(x,y,z)$  and viewing orientation  $(dx,dy,dz)$  with a specified level of confidence  $c$  at time  $t$ .

Such information is determined and maintained at one or more processors specifically assigned to interpreting the raw data input from one or more of the sensing devices. Information is propagated within the distributed processing system as described below.

#### 3.2. Combined 3D Interpretation of Sensor Data

When sensing data is concurrently available from different sensors, the information can be combined to compute a coherent 3D hypothesis describing the most likely current position and orientation of a mobile unit  $u$ . To this end, information from two or more calibrated cameras can be combined via stereo triangulation into a consistent 3D [5]. Sensor fusion techniques provide filtering schemes for adaptively combining more or less consistent information from different sensors with different sensing statistics [8, 15]. The result is a best guess of the mobile user’s current location  $(x,y,z)$  and viewing direction  $(dx,dy,dz)$ .

#### 3.3. Fine-Tuning on the Local Unit

The hypothesized user position is then forwarded to the mobile AR-unit for further fine-tuning, according to the unit’s own tracking capabilities (local camera, gyroscopes, etc). For optical fine-tuning, the wearable unit must rely on appropriate descriptions (3D models) of near-by landmarks.

#### 3.4. Feedback from the Mobile Unit to the Stationary System

Optionally, the locally refined user position can be fed back to the stationary system and to the individual sensor units, thereby providing them with data to analyze their own performance and retune themselves.

Further feedback from the mobile unit involves user reactions and commands related to work progress. Using various input modalities, such as speech, stylus or a wearable mouse, the user can control the augmentations and relevant components of the synthetic data. Appropriate messages are transmitted back to the stationary system.

### 3.5. Issues to be Addressed in the System Architecture

When designing a software architecture to handle distributed tracking, several issues need to be resolved. Which computer should perform the sensor fusion tasks, coordinating the sensor input from all devices? Since hundreds of sensors might be spread throughout very large buildings as well as many mobile units, it might not be feasible to assign all coordination tasks to a single, dedicated computer. Rather, the intelligent building might be broken up into cells within which coordination is dedicated to a special system. Another solution might consider performing sensor fusion work on the wearable AR-systems. That solution would be independent of the number of AR-systems co-located in close vicinity. Yet, it requires more wireless communication bandwidth.

While users are moving about in a building, their distance to stationary sensors varies. New sensors need to be dynamically included in the tracking process when users approach them. Others need to be discarded. When users moves into other cells, the responsibility for sensor fusion activities must migrate with them to other sensor fusion server.

At the same time, the system needs to degrade gracefully when users reach boundaries within which a high precision tracker produces reliable information. The augmentations should not exhibit a sudden jump when a camera tracker is replaced by a simpler tracking modality. Thus, the tracking confidence needs to be computed dynamically and conveyed via appropriate visualizations to the users (e.g., by showing blurry, positionally less precise visualizations close to sensing boundaries).

Depending on the available IT infrastructure (available bandwidth and computing power) of an application, different approaches may prove most suitable. The overarching software architecture must allow us to explore alternatives.

### 4. More Complex Tracking Concepts: Dynamic Scene Understanding

In order to augment real worlds with virtual objects, the virtual objects need to be integrated seamlessly into the environment. They have to behave in physically plausible manners: they occlude or are occluded by real objects, they are not able to move through other objects, and they

cast shadows on other objects [4, 7, 8]. To this end, AR applications require a very accurate model of the real environment [14].

Yet, the real world is not likely to remain static while users are working in it. Actions or instructions issued by the computer cause users to perform actions which change the real world and, in turn, prompt the computer to generate new, different augmentations [13]. Thus, to fully exploit the AR paradigm, the computer must not only augment the real world but also track and analyse how it changes. Interactions between computers, humans and the real world can then be viewed as a circular loop rather than a linear back-and-forth between users and computers (Figure 1).

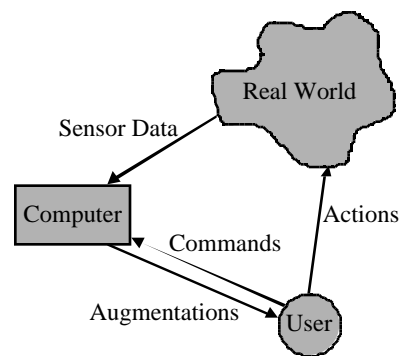


Figure 1: Computer - Human - Real World Interaction

Tracking humans and objects in a dynamically changing world is a challenging problem. First prototypes demonstrating such concepts on stand-alone AR-systems have recently been shown [13,22]. It is particularly hard to track moving objects from a mobile camera. Within an intelligent environment, stationary cameras can use simpler algorithms, such as image differencing, to quickly focus on moving objects while ignoring static scene components. Such approaches have been used successfully in demonstrations, such as the Kids' room [3], allowing children to experience fascinating adventures in a reactive environment. So far, however, such systems are using only wall-mounted displays that present storylines independently of the children's current positions. Showing some of such phantasies as personalized presentations on tracked HMDs is the next step.

#### 4.1. Change Detection from Several Stationary Cameras

From a calibrated, stationary camera, moving objects can be detected by comparing the current image with a reference image taken at a time when no mobile objects were there. Significant image differences are interpreted as scene changes due to object motion. They are collected in

an image mask. Despite some principle limitations of the approach<sup>1</sup>, image differencing is widely used because it works very fast. Image masks from several calibrated stationary cameras can be used to cast a volume in the scene that forms a hull around the mobile object from several of its silhouettes.

The usefulness of such object hulls depends on the positioning of the cameras. An overhead camera in combination with one up and behind a worker's head, may provide very useful localized information on the user's hand movements during an assembly task, thereby allowing the AR-system to model a user's manipulation of real and virtual objects with his hands quite realistically.

When such image differencing algorithms are running on several computers in a distributed environment, computers associated with each camera need to compute the image masks in real-time. All difference images need to be sent to a coordinator which, in turn, computes a 3D volume or a polygonal approximation to it. The result must be transmitted to the mobile system generating the scene augmentations such that proper occlusion relationships between virtual and mobile real objects are taken into account.

## 4.2. Real-Time Stereo Vision Processors

Stereo vision systems that operate in real-time have recently become commercially available [17]. Such systems integrate information from a small box containing three calibrated cameras into a 2.5D depth map of the scene according to the view from one of the three cameras. When several such trinocular vision boxes are set up at calibrated positions in a scene, the depth information from each of them can, in principle, be merged into a rough<sup>2</sup> volumetric description of the scene for every instant in time at which the stereo systems are generating depth maps [10]. An analysis of volume descriptions over time can distinguish mobile objects from stationary ones and provide appropriate volumetric masks [11].

When such technology is used in an intelligent building, 2.5D masks have to be transferred from each sensor to a coordinating unit. The unit then needs to send volumetric information to the mobile AR-unit.

---

<sup>1</sup> This approach works well when mobile objects contrast well against the background. When parts of an object happen to look similar to the occluded background, they are not detected, causing the mask to show holes. Shadows of mobile objects tend to cause problems, generating wrong entries in the image mask.

<sup>2</sup> The degree of roughness depends on the number of stereo vision systems that are used. The more stereo systems, the more precise the volumetric description.

## 5. Architectural Concepts for Augmented Reality

Designing and implementing an efficient distributed AR-system is a complex task. Many distribution and information transmission issues need to be explored and optimized, requiring a flexible overall architecture within which alternate approaches can be investigated. Figure 2 shows the components of a distributed AR-system in a diagram. It forms the basis for formulating some of the open questions.

### 5.1. Stationary Scene Analysis

The left rectangle of Figure 2 represents the AR-related stationary components of the IT infrastructure of an intelligent "AR-ready" building. It consists of many stationary sensors, processors, storage devices and displays connected by a high-bandwidth network. As discussed in sections 3 and 4, the sensors record changes in their local environments at various levels of precision. Stationary tracking programs on suitable local processors use the sensor data to determine user motion in form of simple tracks of an AR-unit (section 3). Information from several sensors is combined by a sensor fusion module into a coherent motion model. The resulting motion hypothesis is forwarded via the wireless network to the respective mobile AR-unit.

If more complex 3D scene and action analysis is required in an application, it is computed by additional modules (section 4) and forwarded in form of interpreted user gestures and dynamic 3D scene descriptions to data repositories. 3D scene descriptions are integrated with the current 3D scene model to describe the current physical configuration of the real world – to be forwarded to the mobile AR-unit for physically correct three-dimensional scene augmentations. Gestures are used to modify internal computer data (e.g., to indicate progress in the state model of a machine repair protocol) and to control the selection and visualization mechanisms which specify which information should be presented to the user in what form. The visualizations are rendered on stationary displays and/or forwarded to mobile AR-units for personalized viewing on user-worn displays.

### 5.2. Mobile Tracking

The right box of Figure 2 shows one of potentially many mobile AR-systems. As a minimum configuration, it needs a wireless network connection to communicate with the intelligent building, one or more sensing devices, a processor and a display. Optionally, it also disposes of a suitably sized local cache as well as multi-media data storage capacity, such as a CD-Rom drive, to have local access to information that does not change. Tracking information received via the mobile network connection is merged in the local sensor fusion process with tracking

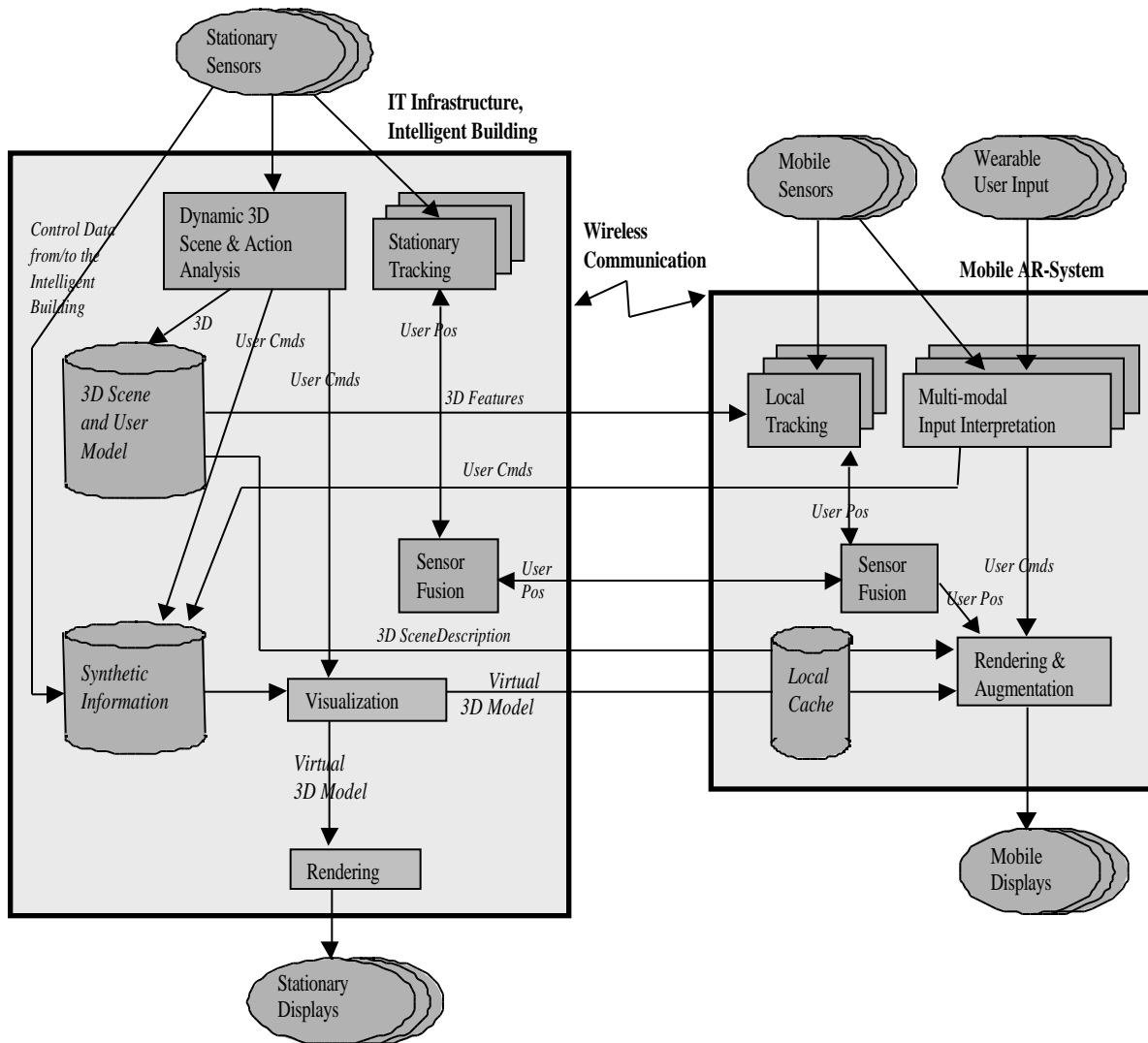


Figure 2: Components of a Distributed AR-Environment

information of one or more mobile sensors. The result is fed back to the sensor fusion model of the intelligent building for further learning. It is also forwarded to the local rendering and augmentation module to show the synthetic information from the same vantage point as assumed by the mobile viewer. The result is shown on the mobile display devices (e.g., an HMD or a hand-held display). According to user reactions, the real world may change. If such real world changes are recorded by local sensors, they are interpreted as multi-modal user input in the wearable unit and forwarded to the intelligent building as input to the virtual data repository.

### 5.3. Wireless data transfer

Figure 2 illustrates several wireless communication paths between the intelligent building and a mobile AR-unit. Most importantly, tracking information must be

transmitted back and forth rapidly. Such information does not require much bandwidth. Yet, transmission lag should be kept to an absolute minimum to support frame-rate (30 Hz.) tracking performance on the mobile unit.

At an infrequent rate – whenever the user turns to a previously unseen area of the building – new 3D landmarks have to be transmitted to the mobile unit where they can be kept in a local cache. The features describe 3D target descriptions that are likely to be in the current field of view of the user (and a head-worn camera). No large transmission bandwidth is required. Transmission lag can also be accounted for, if a predictive user motion module is used to indicate ahead of time where a user is going and what features are thus likely to become important and which ones in the local cache can be overwritten.

User input, such as control commands and data, that is provided via speech, wearable keyboards or other interfaces needs to be transmitted from the mobile unit to the stationary system such that the data repositories can be updated accordingly. Bandwidth and response requirements depend on the application.

The virtual 3D model is the basis of all augmentations. Depending on the application, it may be constant, such as a planned architectural building or interior design, or it may change dynamically in response to user actions or to production-related measurements in buildings such as chemical or power plants. In the former case, if local storage capacity is large enough, the model can be transmitted once off-line, or made available on non-volatile devices such as CDs. In the latter case, relevant data must be selected and/or updated on the stationary system, cast into appropriate visualizations and then transmitted to the mobile unit for rendering and augmentation purposes. Depending on the application, future visualizations might be predictable (in a machine repair scenario) or not (real-time visualization of processing measurements). Accordingly, the requirements on transmission bandwidth and lag vary.

3D scene descriptions are needed in applications that require high-precision augmentations involving physically correct 3D occlusion handling between virtual and moving real objects, such as the user's hands and colleagues. In such cases, large data volumes need to be transmitted in real-time from the stationary system to the mobile unit.

#### **5.4. Situation-Dependent System Use (Quality of Service)**

Once the first efficiently working system has been built and demonstrated as a "proof of concept", the next critical issue will be to make the system robust against varying levels of quality of service. Transmission bandwidth and speed can drop dramatically during times of high, concurrent usage by several mobile AR-units in close vicinity. Furthermore, sensors or other system components can fail temporarily in an ongoing industrial application.

Mobile AR-units need to be able to degrade their performance gracefully, defaulting to lower quality tracking and augmentation schemes in times of high system contention. Thus, the system architecture needs to provide concepts of different levels of connectivity, including a completely disconnected mode in which the mobile system has to temporarily get by with its own means. To this end, the local system must continuously maintain a minimal level of data and status information in its own cache. What the minimal set will be needs to be explored on an application-dependent basis. The

software architecture must allow for various alternatives to be tested.

#### **5.5. Load Balancing and Object Migration**

Wearable computers are generally less powerful than some of the best stationary systems. If the mobile unit is not able to track and augment fast enough to satisfy the requirements of an AR application, load balancing and object migration concepts have to be considered. How much and which computation should be performed on the mobile AR-unit, how much should be processed in the stationary infrastructure? For example, virtual information can be preprocessed and pre-visualized on powerful stationary graphics servers, such as an SGI Onyx, leaving only the final rendering steps to the mobile computer. Similarly, powerful image processing arrays can help perform some of the optical tracking task of the mobile system – at the cost of transmitting the image data to the stationary computer and the results back to the mobile unit. If the mobile computer is powerful enough by itself, it is better to perform such computations on-board, thereby saving bandwidth and lag, as well as being independent of the number other AR-units that might also request services of the same stationary machines. When stationary machines become overloaded, similar load balancing considerations need to determine dynamically whether some task can migrate to other adjacent computers. The software architecture must allow us to explore trade-offs between various options, depending on a particular application and its system configuration.

Section 3.5 has already mentioned another object migration issue: dynamically changing subsets of stationary sensors track mobile users while they are roaming an intelligent building. In a cellular setup, information coordination and filtering algorithms must follow the user from cell to cell, migrating between suitable processors as needed.

#### **5.6. Adaptive System Operation**

A distributed AR-system needs to be dynamically adaptable to changing requirements. Depending on application demands, the number of users and system failures, it must be able to adjust its tracking and augmentation algorithms to warrant real-time response, showing less information with lower precision rather than lagging behind.

Furthermore, the system must also be adaptable to changing user preferences and to growing user experience. Depending on the context, the time of day, the user's task and education, as well as the location, different kinds of visualization may be appropriate. For example, different levels of detail should be used when viewing a scene at a distance. Augmented reality should not clutter the HMD to a degree that it occludes most of the real reality.

The system must also allow for dynamic on-line integration or removal of sensors. During extended system use, users are likely to detect areas of the building that are too densely or not sufficiently covered by sensors. The system must provide tools to investigate and visualize the current sensor coverage, simulate potential improvements, and also aid the user in actually moving and recalibrating sensors.

### 5.7. Security and Confidentiality Concepts

Allowing mobile systems to communicate with the IT infrastructure of a large industrial plant opens the door to many security questions. In the long run, we expect users to bring their own favorite wearable AR-equipment (similar to today's laptops) into the building, communicating with the building via standardized AR communication protocols to use some of the provided services while also remaining within their own, alien system environment.

AR communication protocols need to establish protection mechanisms granting user-dependent access rights to specified system services, while maintaining a fire wall to protect critical data and services – e.g. AR-tools to operate some of the plant machinery. Vice versa, the protocols also need to protect the users. They may be prepared to share some of the data on their wearable system with specified people in the company, but not all. At the same time, they may be quite interested in showing some of their newest demonstrations, using their mobile unit as a demo-server to show new kinds of algorithms on company-provided test data (e.g., a sales person for new computer-controlled high-precision measurement tools).

Furthermore, track records in themselves constitute confidential personal user information. Concepts how to track AR-units without revealing unnecessary details about their users are quintessential to the general acceptance of AR by workers.

### 5.8. User Safety and Comfort

Another important issue of AR involves user safety and comfort. While wearing AR-equipment, users must remain aware of safety hazards in their real environment (e.g., in exterior construction applications and in medicine). Thus, the user must not be confused by augmentations.

Furthermore, the AR-equipment must be intuitive to use, providing simple, non-distracting user interfaces. This issue reiterates the need for a flexible, adaptive system that allows us to explore different approaches and also allows users to configure the setup according to their preferences.

## 6. Summary

Augmented Reality is an exciting new user interface paradigm, showing much promise for a breadth of industrial applications. In this paper, we have discussed the need for integrating mobile AR-tools into the IT system infrastructure of buildings that are becoming increasingly controlled by computers. AR-systems need to be seen as the user interfaces that provide users mobile access to all synthetic information, presented in particularly spatial, multi-media formats (rather than just as texts and 2D graphs) and embedded into the work environment. To achieve the full potential of AR applications, many software engineering concepts of distributed computer and sensing systems have to be reevaluated and redesigned to address the high-bandwidth, compute-intensive requirements of vision-based AR algorithms. Considering current trends and growth rates of the IT industry, rapid improvements for AR are likely to occur within the next decade, turning AR-devices into commodities and evolving de-facto communication standards for “AR-ready” IT infrastructures.

## References

1. R. Azuma and G. Bishop. Improving Static and Dynamic Registration in an Optical See-through HMD. *Proc. Siggraph'94*, Orlando, July 1994, pp. 194-204.
2. R.T. Azuma. The challenge of making augmented reality work outdoors. *Proc. ISMR'99 (1. International Symposium on Mixed Reality)*, Yokohama, Japan, March 1999, pp. 379-390.
3. A. Bobick, S. Intille, J. Davis, F. Baird, C. Pinhanez, L. Campbell, Y. Ivanov, A. Schütte and A. Wilson, The KidsRoom: A Perceptually-Based Interactive and Immersive Story Environment, MIT Media Lab, Technical Report No. 398, September 1997. <http://vismod.www.media.mit.edu/vismod/demos/kidsroom/info.html>.
4. D.E. Breen, E. Rose, and R.T. Whitaker. Interactive occlusion and collision of real and virtual objects in augmented reality. Technical Report ECRC-95-02, ECRC, Arabellastr. 17, D-81925 Munich, 1995.
5. O. Faugeras. Three-Dimensional Computer Vision: A Geometric Viewpoint. MIT Press, 1993.
6. S. Feiner, B. MacIntyre and T. Höllerer. Wearing It Out: First Steps Toward Mobile Augmented Reality Systems. *Proc. ISMR'99 (1. International Symposium on Mixed Reality)*, Yokohama, Japan, March 1999, pp. 363-377.
7. A. Fournier. Illumination problems in computer augmented reality. *Journee INRIA, Analyse/Synthese d'Images (JASI)*, pp. 1-21, January 1994.
8. Gelb, A. (ed.) (1974). *Applied Optimal Estimation*. MIT Press, Cambridge, MA.
9. K. Ikeuchi, Y. Sato, K. Nishino, and I. Sato. Photometric Modeling for Mixed Reality. *Proc. ISMR'99 (1. International Symposium on Mixed Reality)*, Yokohama, Japan, March 1999, pp. 147-163.
10. T. Kanade, A. Yoshida, K. Oda, H. Kano, and M. Tanaka. A stereo machine for video-rate dense depth mapping and

- its new applications. *Proc. 15th IEEE Computer Vision and Pattern Recognition Conference (CVPR)*, 1996.
11. T. Kanade, P. Rander, S. Vedula and H. Saito. Virtualized Reality: Digitizing a 3D Time-Varying Event As Is and in Real Time. *Proc. 1<sup>st</sup> International Symposium on Mixed Reality (ISMR)*, Yokohama, Japan, March 1999, in book Y. Ohta, H. Tamura (eds.), *Mixed Reality – Merging Real and Virtual Worlds*, Springer Verlag 1999, pp. 41-57.
  12. G. Klinker, D. Stricker, and D. Reiners. Augmented Reality: A Balance Act between High Quality and Real-Time Constraints. 1. International Symposium on Mixed Reality (ISMR'99), Y. Ohta and H. Tamura, eds., "Mixed Reality – Merging Real and Virtual Worlds", March 9-11, 1999.
  13. G. Klinker, D. Stricker and D. Reiners. An Optically Based Direct Manipulation Interface for Human-Computer Interaction in an Augmented World. *Proc. Eurographics Workshop on Virtual Environments (EGVE'99)*, Vienna, June 1999.
  14. G. Klinker. Augmented Reality: A Problem in Need of Many Computer Vision-Based Solutions. *Proc. NATO Advanced Research Workshop "Confluence of Computer Vision and Computer Graphics"*, Ljubljana, Slovenia, August 1999.
  15. D. Koller, G. Klinker, E. Rose, D. Breen, R. Whitaker, and M. Tuceryan. Automated camera calibration and 3D egomotion estimation for augmented reality applications. *7<sup>th</sup> Int'l Conference on Computer Analysis of Images and Patterns (CAIP'97)*, Kiel, 1997.
  16. U. Neumann, S. You, Y. Cho, J. Lee and J. Park. Augmented Reality Tracking in Natural Environments. *Proc. ISMR'99 (1. International Symposium on Mixed Reality)*, Yokohama, Japan, March 1999, pp. 101-130.
  17. Point Grey Research Inc. Color Triclops Stereo Vision System, 1999. <http://www.ptgrey.com/>.
  18. A. State, G. Hirota, D.T. Cheng, W.F. Garrett, and M.A. Livingston. Superior augmented reality registration by integrating landmark tracking and magnetic tracking. *Proc. SIGGRAPH*, pp. 429-438, New Orleans, Aug 4-9, 1996. ACM Press.
  19. D. Stricker, G. Klinker, and D. Reiners. A fast and robust line-based optical tracker for augmented reality applications. 1<sup>st</sup> International Workshop on Augmented Reality (IWAR'98), San Francisco, 1998.
  20. M. Tuceryan, D. Greer, R. Whitaker, D. Breen, C. Crampton, E. Rose and K. Ahlers. Calibration requirements and procedures for a monitor-based augmented reality system. *IEEE Transactions on Visualization and Computer Graphics*, 1, September 1995, pp. 255-273.
  21. R. Want, A. Hopper, V. Falcao, and J. Gibbons. The Active Badge Location System. *ACM Transactions on Information Systems*, Vol. 10, No. 1, January 1992, pp. 91-102.
  22. N. Yokoya, H. Takemura, T. Okuma and M. Kanbara. Stereo Vision Based Video See-through Mixed Reality. *Proc. ISMR'99 (1. International Symposium on Mixed Reality)*, Yokohama, Japan, March 1999, pp. 131-141