Mobile Computing and Industrial Augmented Reality for Real-time Data Access

Xiang Zhang Yakup Genc Nassir Navab

Siemens Corporate Research
Imaging and Visualization Department
755 College Road East
Princeton, NJ 08540, USA
{Xiang.Zhang, Yakup.Genc, Nassir.Navab}@scr.siemens.com

Abstract – **Augmented Reality (AR) is a technology that enhances the user's view of real world with virtual objects. AR is a relatively new field and seeks applications in a wide range of areas from medicine to industry. Applying AR with mobile computing, we present a new method for information access in industrial environments.**

We, in this paper, present a framework that is applicable to different scenarios in larger industrial settings for spatial data access, navigation, and image augmentation in real-time.

In the core of our framework lies a mobile computer equipped with a camera. The camera observes the environment for visual coded markers. These markers are registered to a global coordinate system through available drawings or floor plans. The tracker software processes images coming from the camera at moderate frame rates (typically 12fps) and estimates the location of the user. The system then can guide the user through the environment, provide the user with location relevant data from a spatial database, and augment the view of the user through the camera. Data exchange is done via a wireless network and the user interface allows the user to access various types of data without considerable effort.

I INTRODUCTION AND BACKGROUND

Augmented reality (AR) enhances the user's view of the real world with the virtual objects. The virtual objects displayed to the user spans from simply texts to three-dimensional models with elaborate shading and illumination effects. AR is a relatively new field and seeks applications in a many areas ranging from medicine to industry ([1, 2, 3, 4, 5, 6, 8]).

A typical AR system includes a display and a tracker with associated software. The software reads the tracker events to know the position of the display and renders the virtual objects on the display. In order to render correctly, the virtual objects and the real world need to be registered. This registration implies that the geometry of the virtual camera where the augmentation takes place is known with respect to the real world. In many cases, the tracker should provide very accurate pose of the real-world view obtainer to ensure seamless rendering of the virtual objects over the real-world view.

In industrial operations, a factory database usually needs to be accessed depending on the user's location in the environment. For instance, real-time operation data, maintenance information and documentation are usually stored in a remote server, which is accessible at the site via a wireless network with a mobile computer. The same can be observed in a remote or automated maintenance assistant or training scenario. Of course, in the large environments, the user also needs to be guided to the interested areas.

We present a framework, which is general enough to be applicable to different kinds of scenarios in large industrial environments for image augmentation, navigation, and data interaction in large industrial environments. In the core of our framework lies a mobile computer equipped with a camera. The camera observes the environment for visual coded markers. These markers are registered to a global coordinate system through available drawings or floor plans. The tracker software processes images coming from the camera at moderate frame rates (typically 12fps) and estimates the position and orientation of the user. The system then can guide the user through the environment, provide the user with location relevant data from a spatial database, and augment the view of the user through the camera. Data exchange is done via a wireless network and the user interface allows the user to access various types of data without considerable effort.

Our framework makes use of the latest advances in the mobile computing and wireless communication technologies. With the recent technology advances, powerful computers with substantial storage capacities are be-

0-7803-7241-7/01/$10.00 (c)2001 IEEE

coming available for industrial applications. Examples of existing mobile computers are the Sony Vaio PictureBook, and the Xybernaut MAIV wearable computer. Figure 1 shows two commercially available mobile computers. These computers can either be worn, or hold in hand by the users and connected to a computer network through wireless communication. These types of computers find more and more application areas in information technology(IT)-driven maintenance, training, localization, environment navigation, and spatial data management. It is expected that eventually, mobile computers will become an integrated part of factory automation and industrial administration.

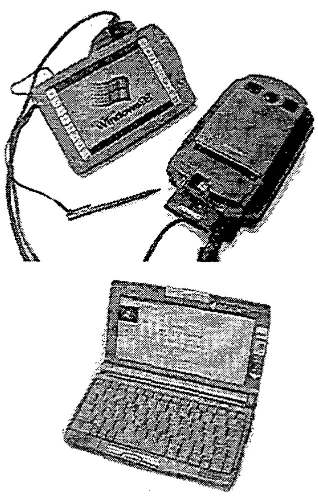

Figure 1: Mobile computers. *Top:* Xybernaut wearable computer. *Bottom:* Sony VAIO PictureBook.

We call this framework ArLoc, from "AR with localization". For example, ArLoc can be applied for accessing the 3D technical documentation created by the as-built reconstruction software Cylicon ([7]). Such 3D documentation is a database including reconstructed 3D virtual objects, images, and CAD data, etc.

Our framework can be implemented using only the underlying technologies of AR, i.e., motion tracking, camera calibration, registration, and virtual object superimposition. In our implementation, we use a set of specially designed coded visual markers as shown in Figure 2 for motion tracking. Our currently implementation detects the markers in images use a watershed transformation algorithm ([6, 10]), which cannot detect partially occluded markers. We use a homography based camera calibration algorithm ([11]) to estimate pose for registration. The accuracy of 3D position is highly dependent on the quality and resolution of the image, the distance between the camera and the markers, and the calibration algorithm. With images from good quality CCD camera, the accuracy of pose estimation results is in centimeters ([9, 10]).

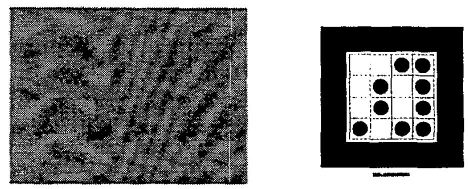

Figure 2: Example of coded markers, *Left*: `4095b`, *Right*: `1365a`.

In the rest of the paper, we present the overview of the ArLoc framework in Section II, some of the implementated applications in Section III. We conclude in the last section with a general discussion and suggestions for the future working directions.

II THE FRAMEWORK

Many people have such an experience: No matter how many times you have been to a shopping mall, especially a very large one like the Franklin Mills Mall in Pennsylvania, you'll still need a floor-map with a red dot that says "You Are Here" to find your way easily. Large scale industrial environments are often much more complicated and the persons working inside need to access more information than just a floor-map. The idea of our AR and Localization platform is to provide the mobile computer user with the floor-map as well as the interaction with the spatial data over the factory network, as shown in Figure 3.

The navigation and data interaction are achieved by applying AR technologies together with wireless LAN connection. Figure 4 depicts the framework of our augmentation (AR), navigation, and data interaction platform. In this system, the real-time video of the environment captured by the attached camera is sent to the tracker for marker detection and pose estimation. Coded visual markers are placed in the environment and their positions and orientations are mapped onto the floor-map. Instead of directly mapping regions on the floor-map with hyperlinks, we provide data access by attaching the markers with hyperlinks to files, websites, and databases. Therefore, the user can interact with various data resources through the network.

Besides visualizing the localization on the floor-map and accessing spatial data through a web browser, we aug-

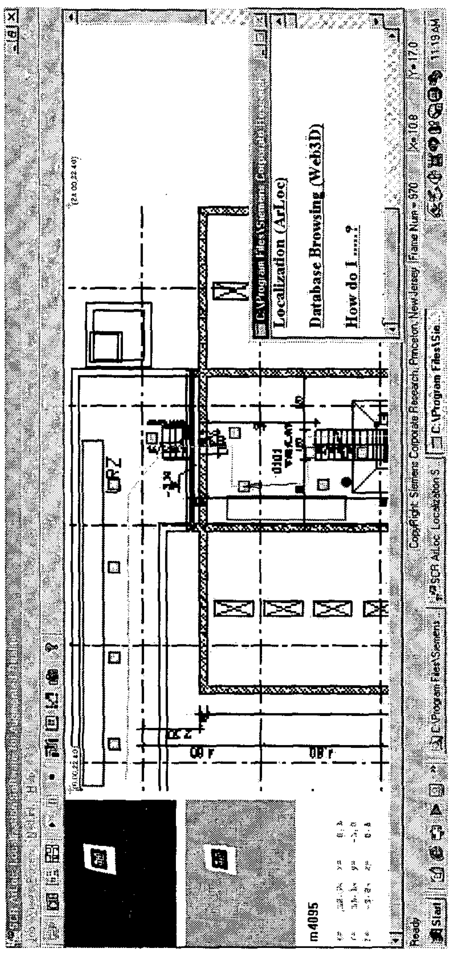

Figure 3: Interface of ArLoc (view in landscape direction): The left side of the main application window is showing the real-time video, marker detection process, code of the marker detected, and the camera local and global coordinates from top to bottom. The floormap is displayed together with guiding path (green line), marker mapping (markers are displayed in purple or bluish colors), and areas to pay attention (red frames). The pop-up window shown at the bottom-right corner is actually a web browser use to automatically update information according to the position of the user. The icon circled by a green line in the window is the wireless network controller.

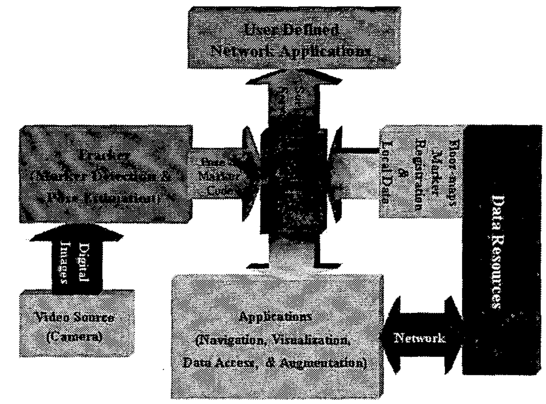

Figure 4: Framework of the AR, localization, and data navigation platform.

ment the real-time video with additional information as the case shown in Figure 8.

In addition, we attach a socket server to our AR and localization platform. This socket server allows network clients to connect to it through the network and get localization information including marker code and the pose of camera in the floor-map defined coordinate system. This component enables other users on the network to communicate with the localization platform and run, for example, cooperative applications, based on the localization result acquired from our localization platform.

III IMPLEMENTED APPLICATIONS

The core applications currently implemented on our AR and localization platform are navigation and data interaction. Using the monitor of the mobile computer as a video-see-through display, we also augment the video captured by the attached camera to provide additional information in real-time. We present more details of these applications in the following with examples.

A *Navigation*

For large site navigation, we set the global coordinate system based on the user loaded floor-map and its given dimensions. Localization is computed from the output of camera calibration and the mapping of the markers on the floor-map.

Figure 5 shows an example of the visualization of the localization result. The scenario is that the mobile computer user is working inside a large industrial building and the user is surrounded by huge containers, pipes, and other equipments all look alike. A floor-map of the working site and the position indicating red dot are displayed on the monitor of his computer. On the floor-map, the markers placed in the working site is also displayed. A

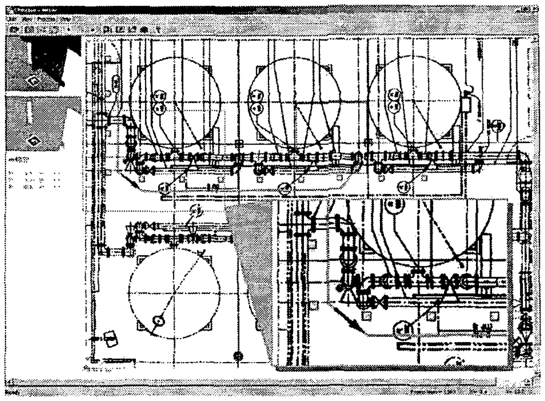

Figure 5: Localization and navigation guidance inside a factory.

green line is drawn on top of the floor-map as a navigation guidance.

Figure 6 shows another case of localization and visualization. When the user points the camera from the floor to a structure, the camera observes some markers that belong to the equipment, then the system automatically switches the display from the floor-map to a industrial drawing of the on-site structure. At the same time, there is a web browser window directly connecting the user to the corresponding data resource to access operation documents as well as real-time operation data such as the current pressure, temperature inside the valve, etc.

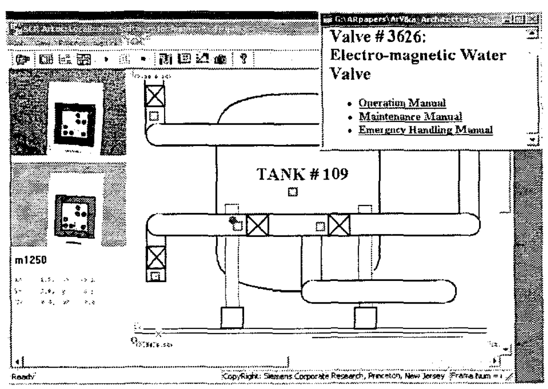

Figure 6: When the camera is pointing to a marker next to a valve, the location of the valve in the industrial drawing is displayed and a web browser automatically connects the user to related data.

B Data Navigation

One of the main applications of this framework is the location based data navigation that means navigating through the data space of the network. The mobile computer is connected to the network with a wireless LAN card. The user can always access spatial data through the hyperlinks mapped with the markers as shown in Figure6.

In addition to the web browser based data interaction, we implemented a socket server on our platform to allow the user(s) on the network to obtain localization results from the mobile computer. In this case, the users on the network can plug in any customer defined applications. The advantage of this is that the users on the network now will be able to see what the mobile computer user is seeing and access the same spatial data. This feature is particularly useful in the situations such as cooperation, training, etc.

Figure 7 shows the interface of Web3D – a network application developed at SCR for interacting with the 3D industrial documentation database that is built using CyliCon ([7]). Web3D can run from any computer in the network, including the mobile computer itself, while the CyliCon database could be placed on any computer. In operation, Web3D obtains the localization results through the socket connection, then extracts corresponding images, 3D virtual structures, and other parameters from the CyliCon database([7]). It automatically updates the 3D visualization of the current environment. The user can click on the items of interest to interact with related data. Shown in Figure 7, a pop-up Internet Explorer window is showing the parameters of a piece of pipe extracted from the CyliCon database.

C Augmentation

We augment the real-time video with text and virtual objects to provide the mobile computer user with further information. For example, we can carpet the floor in the image with the floor map and put arrows on the floor as the navigation guidance, and/or the most important real-time operation data of the nearest equipment, for example the pressure of the closest valve/container. Figure 8 shows the case that in addition to visualization of the localization with the floor-map, a blue arrow and text are superimposed on the image indicating the direction to another room. Such augmented information could be particularly useful, since it provides navigation guidance related to the camera direction.

IV DISCUSSION AND FUTURE WORK

We present a software platform for augmented reality, navigation, and information access in large scale industrial environments with mobile computers. Our implementation can be used to

- Provide reliable localization results including both the position and the orientation in large scale industrial environments.

- Augment the real-time view of the working environment with 3D object or additional information.
- Provide flexible ways for data interaction using a pop-up web browser that automatically connects the user to corresponding spatial data resources.
- Provide a network interface for other users connected to the network to plug in their own applications that obtain and the localization result and make use of it.

Certainly, there is still room to improve our localization platform and broaden the range of its applications. Here are some of future work directions we bear in mind:

- Hands-free operation: using head mounted displays (HMD) to free the hands from holding the display devices, and using voice operation to free the hands from using the mouse or a keyboard.
- Combined localization method: combining the visual and marker based localization method together with other motion-tracking/localization methods. For example, combining with a GPS system, we can create a localization system that can be applied to much larger environments.
- Motion-tracking/localization that are based on natural features: with improved image processing algorithms and much advanced feature extracting system together with carefully designed machine learning system, we hope one day that we will be able to throw away all these coded markers. The localization then will be based on the natural visual features extracted from the environment in real-time.

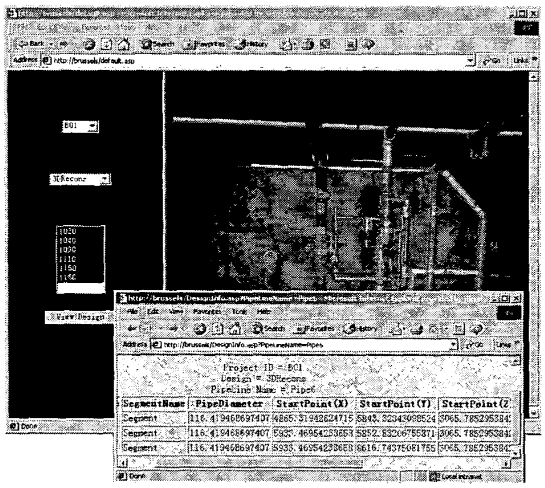

Figure 7: Data Navigation: The web browser is connected to the 3D industrial documentation database. Clicking on a pipe showing in the window, the user can access parameters of the pipe, including the real-time operating data.

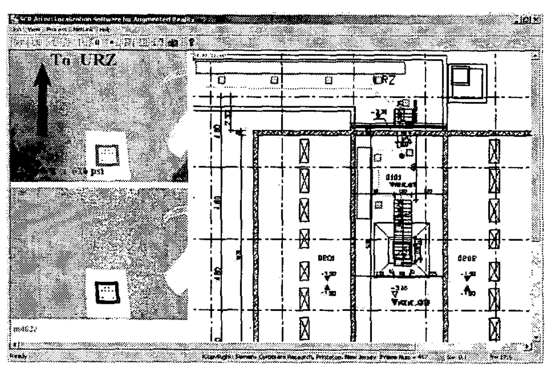

Figure 8: Augmenting the video with a navigation guidance arrow and the text stating the data of the closest important equipment.

References

[1] M. Bajura, H. Fuchs, and Ohbuchi. Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. *Computer Graphics*, pages 203–210, July 1992.

[2] W. Birkfellner, K. Huber, F. Watzinger, M. Figl, F. Wanschitz, R. Hanel, D. Rafolt, R. Ewers, and H. Bergmann. Development of the varioscope AR: A see-through HMD for computer-aided surgery. In *Int. Symp. on Augmented Reality*, pages 54–59, 2000.

[3] T. Caudell and D. Mizell. Augmented reality: An application of heads-up display technology to manual manufacturing processes. In *Proceedings of the Hawaii International Conference on System Sciences*, pages 659–669, 1992.

[4] S. Julier, M. Lanzagorta, Y. Baillot, L. Rosenblum, S. Feiner, and T. Hollerer. Information filtering for mobile augmented reality. In *Int. Symp. on Augmented Reality*, pages 3–11, 2000.

[5] G. Klinker, D. Stricker, and D. Reiners. Augmented reality: A balancing act between high quality and real-time constraints. In Y. Ohta and H. Tamura, editors, *Mixed reality: merging real and virtual worlds*, pages 325–346. Ohmsha, Ltd. Tokyo, 1999.

[6] D. Koller, G. Klinker, E. Rose, D. Breen, R. Whitaker, and M. Tuceryan. Real-time vision-based camera tracking for augmented reality application. In *Symposium of virtual reality software and technology*, 1997.

[7] N. Navab, E. Cubillo, B. Bascle, J. Lockau, K. D. Kamsties, and M. Neuberger. CyliCon: a software platform for the creation and update of virtual factories. In *Proc. of the 7th IEEE Int'l. Conference on Emerging Technologies and Factory Automation*, pages 459–463, Barcelona, Spain, 1999.

[8] F. Sauer, F. Wenzel, S. Vogt, Y. Tao, Y. Genc, and A. Bani-Hashemi. Augmented workspace: Designing an AR testbed. In *International Symposium for Augmented Reality*, pages 165–174, Munich, Germany, October 2000.

[9] X. Zhang and N. Navab. Tracking and pose estimation for computer assisted localization in industrial environments. In *IEEE Workshop on Applications of Computer Vision*, pages 214–221, 2000.

[10] X. Zhang and N. Navab. Design coded visual markers for tracking and camera calibration with mobile computing systems. Technical Report SCR-01-TR-691, Siemens Corporate Research. Inc., April 2001.

[11] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In *Proc. Int. Conf. Comp. Vision*, pages 666–673, 1999.