Annotation of Features in Outdoor Augmented Reality Environments

M. Rubio[1], R. Quirós[2], E. Pulido[2], G. Fabregat[2], J. Huerta[2]
[1] Universidad de Oriente Santiago, CUBA
monica@csd.uo.edu.cu
[2] Universitat Jaume I, Castellón, SPAIN
quiros, pulido, fabregat, huerta @uji.es

SUMMARY

Augmented Reality technologies offer new possibilities to the exploration of spatial distributed information. In this work we describe the main issues in the development of an Augmented Reality system that allows the annotation of objects or features in the real world captured by a camera. The general operation of this kind of systems is described, as well as the main problems that need to be solved. Finally, as an example of case of study, we present an application for the annotation of information relative to containers in a marine terminal

KEYWORDS: *Augmented Reality, Spatial annotation, Ubiquitous computing*

INTRODUCTION

The Human-Computer Interaction techniques have experienced a great development during the last years, impelled, to a great extent, by the recent advances in the technologies related to Virtual Environments and Augmented Reality. In the field of Augmented Reality (AR), virtual objects are integrated interactively in real environments (Azuma, 1997). Augmented reality can be considered a variation of the Virtual Environment (VE) concept. AR technologies allow the user to see the real world mixed with virtual objects superimposed to or composed with it.

The main objective of an AR system is to enforce the perception of the users in its interaction with the real world, complementing it with virtual objects that seem to coexist in the same space. The virtual objects show information that the user cannot directly detect with his senses, improving the user ability to perform some kind of tasks. AR has many applications, both indoor and outdoor. Indoor environments are generally closed rooms, sometimes prepared to ease the integration of the virtual objects into the scene. Outdoor environments are much more restricted by the impossibility to use any kind of auxiliary tool that facilitates the real-virtual combination.

In the first years of development of AR technology, most of the applications developed were indoor applications: medical visualization (Bajura & Neumann, 1995) (State et al., 1996), maintenance and repair (Feiner et al., 1997), robot operation planning (Milgram et al., 1993) entertainment (Maes, 1995) (0hshima, 1999)) and space annotation (Wanstall, 1989). Since late nineties a lot of work has been done in three new areas: outdoor mobile AR systems, collaborative AR and the development and deployment of commercial applications (Azuma, 2001).

Partially supported by Spanish Government grants TIC2002-4166-C03 and TIC2001-2416-C03

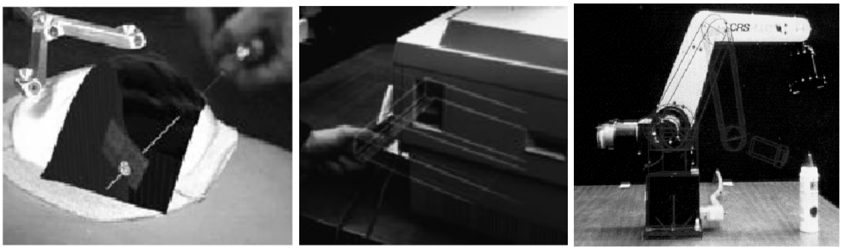

Figure 1. Classical AR indoor applications

The advances in computer science and tracking technologies experienced during last years have allowed the exploration of outdoor environments with mobile systems. These systems offer new possibilities in the use of geographic information. The interest in developing AR systems that provide an exact registry outdoors relies in the possibilities offered to experiment in new areas of application. Among these new areas of application, AR can be used to annotate objects and environments with public or private information, assuming that the data bases that store that information are available (Rekimoto & Katashi, 1995). This possibility has been demonstrated showing windows associated to a standard user interface on specific locations in the real world (Feiner ET al., 1993) (figure 2, left).

Figure 2. Outdoor AR applications

Although outdoor applications offer tempting possibilities, only a few AR prototypes of this kind have been constructed. A prototype of AR system has been developed (figure 2, centre) that shows the user information associated with buildings as it walks around the campus of the University of Columbia (Feiner ET al., 1997). The system includes tracking (compass, inclinometer and a differential GPS), a mobile computer with a 3D graphics board and a see-through Head Mounted Display (HMD). The system shows the user information of the real world (like the name of the buildings), models of demolished buildings (Figure 2, right), the best route to reach to different destinations, or presents videos showing historical facts in the place where they happened (Höllerer ET al., 1999). Applications with laptops for outdoor maintenance systems have also been experimented. None of these systems achieves a good accuracy in the alignment of the real and virtual objects for a wide range of locations (Azuma, 1999).

GOALS AND POTENTIAL APPLICATIONS

The goal of the construction of AR systems is to obtain an ideal AR portable system, as light, small and powerful as possible, that allows the user to explore any non-prepared environment, indoor or outdoor, without any restriction. Our more real objective is to design a prototype of mobile system for the annotation of objects or features in outdoor environments. The system is made up of the following devices:

- One or several cameras that capture the scene.
- A display device (standard display or Head Mounted Display).
- Tracking devices to locate the position of the user and the viewing direction.
- Some communications system.
- A portable computer, Tablet PC or PDA.
- Interaction devices.

Along with the mentioned hardware components, the system will use an available database containing the information, that it will be showed as annotations, and a geometric model of the environment to be explored. The system allows the user to pick some object in the environment and to query for its available information. The desired functionality of the system is showed in figure 3.

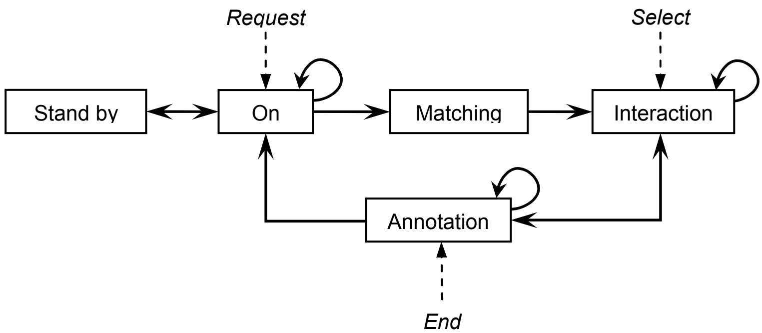

Figure 3. Scheme of operation of the portable system

The system changes between the different states by means of actions that the user makes on the interaction devices. The states that the system can take are the following:

- Stand by: the system is powered off.
- On: the system captures the real environment with a camera and shows the image to the user without performing any operation.
- Matching: this state is achieved when the user requests an interaction. In the matching step, the correspondence between the coordinate systems of the real world and the virtual one is made. After this step we can show virtual information aligned with the objects and elements of the real world.
- Interaction: in this phase the user selects an object in the real world that has associated information available.
- Annotation: the information available for the object is superimposed to the object selected by the user in the interaction phase.

The system thus defined allows the user to observe the environment through the built-in camera. At any time the user can request for information associated to some object or feature in the environment through a simple interaction with devices like buttons or joysticks. The virtual information that will be superimposed to the real image is simple, and is formed by text and graphics as lines or curves.

The potential applications of the system are many. The system can be used to show relevant information relative to any existing spatial object in the environment. The technique can be applied in
fields like:

- Navigation: the system can be used to add relevant information (traffic, outstanding incidences, etc.) for different paths or feasible routes. Also the system can be used to annotate information relative to services, like restaurants, gas stations, etc.
- Tourism: the system can be used to guide visitors in tourist places, showing them the way to follow, the main places of interest, foregoing historical facts, etc.

- Maintenance and repair: the system can show superimposed to the environment the location of electrical lines, water pipes or any other hidden elements in the real environment.
- Monitoring of industrial processes: this is one of the main fields of application of Augmented Reality. In a later section an example of this kind of applications will be proposed as case of study.

Finally, some classical indoor applications could be ported outdoors to facilitate the accomplishment of tasks or to extend the application domain of the technology.

PROBLEM CHARACTERIZATION

The development of the proposed system presents many problems to be solved. Next, we will describe the operation of the system, analyzing the functionality of each one of its components. For each component, we show the detected problems and their possible solutions.

The block diagram of the system is shown in figure 4.

The camera installed in the mobile unit captures the real environment and generates a discreet image, usually of low resolution (640x480 or 800x600 pixels). The tracking system returns a position (x, y, z) and a viewing direction (α, β, γ), both estimated at the moment of the capture. This information is then used in the matching step together with the CAD model of the environment to determine the real position and real viewing direction in the image. This calculation is the core of all the process, since it guarantees the correct alignment of the real objects and the virtual ones. Once that the real parameters are determined, the system combines the image acquired with appropriate texts and graphics and displays the result to the user.

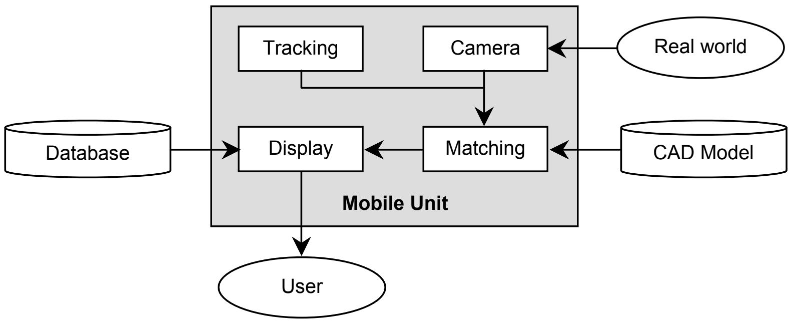

Figure 4: Block diagram of the system

The construction of camera and display components do not present too much difficulties, except that we need to estimate the camera parameters via a previous calibration. The system could use heterogeneous displays (Butz ET al., 1999) and the display will not be necessarily a HMD. The main problems appear in the tracking subsystem and, specially, in the matching subsystem. Next we describe these problems with more detail.

Tracking

The location of the viewer in wide outdoor environments is a more complex task than the location in indoor ones. The latest can be prepared adding to the environment different marks (by means of electromagnetic or optical emitters, for example), allowing the development of highly reliable tracking systems (Feiner ET al., 1997). At the same time, indoor environments can be prepared setting visible marks that improve the estimation obtained by computer vision techniques (Bajura & Neumann, 1995). This last technique is the one used in the popular software *AR Toolkit*.

In outdoor environments, usually we cannot have visible marks or emitters, making the design of functional tracking systems an open problem at the moment. Most of the prototype systems developed up to date are based on GPS and inertial systems (gyroscope and accelerometer), combined with compasses and other elements (Azuma, 1997). Most of these prototypes are based on hybrid approaches, and almost all of them use techniques of Computer Vision to improve the matching between real and virtual world (Azuma, 1999).

Matching

Matching is the core process of the AR system, and the step in which most of the design problems appear. The matching process guarantees the correct alignment between the coordinate system of the real world and the one used to display the 3D CAD model. Figure 5 shows a simple CAD model superimposed to an image of the real world (Julier ET al., 2000).

The sequence of images of the real world captured by the video camera is used to correct the errors dynamically, assuring the correct alignment between virtual and real objects (Azuma, 1999) (State, 1996). During the last years, relevant advances in the registry techniques have been achieved, but these advances do not grant that the problem is solved. AR systems, especially those that lean in the image to improve the registry, need carefully controlled environments. The user cannot freely walk around and browse the environment. The system works only with some specific objects that can be seen from a reduced set of viewpoints. When restrictions of this kind are settled down, the problem is easier to solve, but the flexibility of the system is also restricted, making harder the construction of effective AR systems without the knowledge of an expert. This fact is shown by the lack of commercial systems (Azuma, 1999).

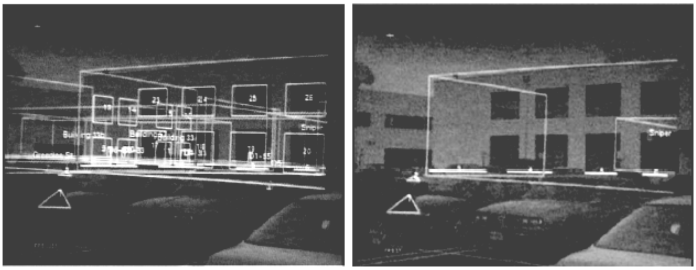

*Figure 5***.** CAD model superimposed over real imagery

In many AR applications the registry is reinforced locating reference points (*fiducials*) in the environment. These fiducials can be LEDs (Bajura & Neumann, 1995) or special marks (Neumann & You, 1999). The location of the reference points is assumed to be well known. We apply a process to the image that detects the fiducials and uses them to correct the registry. These routines assume that one or more of the points remain visible throughout the process. If none of the points are visible the registry can fail. On the other side, when all the fiducials are visible, the registry can be exact. In outdoor applications, given the impossibility to locate marks in the environment, the fiducials must be obtained applying techniques of feature detection in the real image. These techniques allow the extraction of corners (vertices) and edges from the image, that later must be paired with vertices and edges of the CAD model.

All these techniques are based on determining the relative projective relation between the objects in the environment and the video camera. The process is complex, because both the detection of characteristics and their relation with the real objects must be executed in real time and must be robust. Usually it is necessary to use hardware and special sensors. Once the projective relation is determined, techniques of visual servoing can be used to keep the worlds registered in real time (Comport ET al., 2003).

ANNOTATION OF CONTAINERS IN A MARITIME TERMINAL

As an example of application of the described techniques, a prototype system of Augmented Reality is being developed. The system allows remote access to alphanumeric information available for the containers stored in a marine terminal. The geometric information referring to the position and distribution of the containers is used as the basis for the interaction. The prototype will work outdoor, and will give the user in the maritime terminal access to the information available for some specific container selected or pointed by him. The associated information is stored in a remote database that is permanently updated and it is assumed to be accessible. The project is funded by the Ministry of Science and Technology of Spanish Government (TIC2002-4166-C03).

At the moment, we are working in two main areas: the construction of our own tracking system capable of operating in the maritime terminal environment, and the definition of Computer Vision based techniques to make an exact registration between the CAD model of the terminal and the obtained image. The investigation in this second area tries to extract from the image a set of characteristics (edges and vertices) that allows the matching with the CAD model.

Using images taken in the maritime terminal we have tried different methods of vertex an edge extraction: the one proposed by Harris (Harris & Stephens, 1988), the SUSAN method (Smith & Brady) and some novel methods proposed by Geusebroek, and van den Boomgaard (Geusebroek, & van den Boomgaard, 2001). Figure 6 shows some images obtained applying these methods.

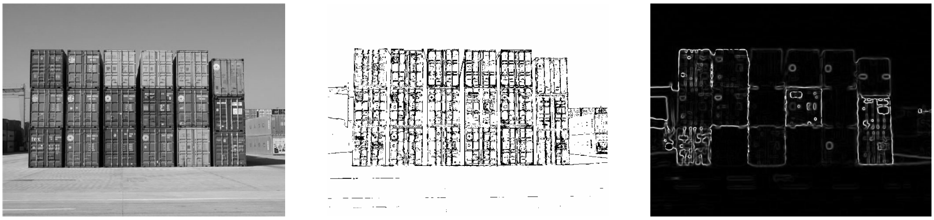

*Figure 6***.** Original image (left) and border detectors. SUSAN (centre), Geusebroek (right)

After extracting the features from the image, the results are processed with a method of edge thinning and suppression of not-maximums. After these methods we can extract the contours of interest of the scene (those corresponding to the limits of each container). Figure 7 shows the final result obtained after applying these processes. The red lines represent the edges detected for the matching.

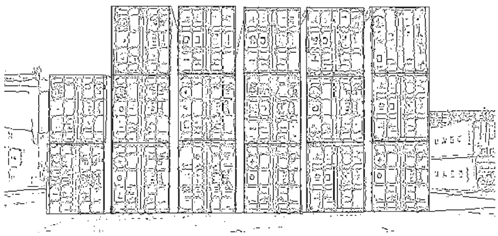

Figure 7. Final matching. Red lines show detected containers

CONCLUSIONS AND FUTURE WORK

Augmented Reality technologies can be useful in outdoor applications, as a visual aid in navigation or maintenance tasks. In this work we have presented a system of spatial annotation that allows to link alphanumeric information to real objects in a sequence of video captured with a video camera. As an example, a prototype of the system is being constructed to annotate containers in a maritime terminal. The first tests made obtain significant features of the scene in less than a second, which allows us to be optimistic and presume that the system could be operative after a brief period of matching, allowing the user to obtain the requested data in a minimum time and to maintain the correct registration in real time.

The future work is directed to construct our own tracking system and an operative mobile unit for the maritime terminal, as well as to extend the field of application of the system for less controlled environments. Specially, we will focus on the study of the feasibility of application of the system in applications of guidance and advising to visitors in tourist places.

BIBLIOGRAPHY

Azuma, Ronald T. "A Survey of Augmented Reality. Presence: Teleoperators and Virtual Environments" 6, 4 (August 1997), 355-385.

Azuma, Ronald T. "The Challenge of Making Augmented Reality Work Outdoors". In Mixed Reality: Merging Real and Virtual Worlds. Yuichi Ohta and Hideyuki Tamura (ed.), Springer-Verlag, 1999. Chp 21 pp. 379-390.

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B. Recent Advances in Augmented Reality. *IEEE Computer Graphics and Applications 21*, 6 (Nov/Dec 2001), 34-47.

Bajura, Michael and Ulrich Neumann. "Dynamic Registration Correction in Video-Based Augmented Reality Systems". IEEE Computer Graphics and Applications 15, 5 (September 1995), 52-60.

Butz A. et al., Enveloping Users and Computers in a Collaborative 3D Augmented Reality,. *Proc. 2nd Int.l Workshop Augmented Reality. (IWAR '99).* San Francisco, 20-21 Oct. 1999, pp. 35-44.

Comport A., Marchand E., Chaumette F. A real-time tracker for markerless augmented reality. In *ACM/IEEE Int. Symp. on Mixed and Augmented Reality, ISMAR'03*, Pages 36-45, Tokyo, Japon, Octobre 2003

Feiner, Steven, Blair MacIntyre, Marcus Haupt, and Eliot Solomon. Windows on the World: 2D Windows for 3D Augmented Reality. *Proceedings of UIST '93* (Atlanta, GA, 3-5 November 1993), 145-155.

Feiner, Steven, Blair MacIntyre, and Tobias Höllerer. "A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment". Proceedings of FirstInternational Symposium on Wearable Computers (Cambridge, MA, 13-14 October 1997), 74-81.

Geusebroek, Jan-Mark y Rein van den Boomgaard. "Color Invariance". IEEE transactions on pattern analysis and machine intelligence, vol. 23, no. 12, december 2001

Harris C. and Stephens M., "A Combined Corner and Edge Detector", in *Proc. 4th Alvey Vision Conference*, pp. 189-192, 1988.

Höllerer, T. et al., "Exploring MARS: Developing Indoor and Outdoor User Interfaces to a Mobile Augmented Reality System," Computers and Graphics, vol. 23, no. 6, Dec. 1999, pp. 779-785.

Maes, Pattie. Artificial Life Meets Entertainment: Lifelike Autonomous Agents. *CACM 38*, 11 (November 1995), 108-114.

Milgram, Paul, Shumin Zhai, David Drascic, and Julius J. Grodski. Applications of Augmented Reality for Human-Robot Communication. *Proceedings of International Conference on Intelligent Robotics and Systems* (Yokohama, Japan, July 1993),1467-1472.

Neumann U., You S., Natural Feature Tracking for Augmented Reality, *IEEE Trans. Multimedia*. vol. 1, no. 1, Mar. 1999, pp. 53-64

Ohshima T., et. al., RV-Border Guards: A multi-player mixed reality entertainment,. *Trans. Virtual Reality Soc. Japan*, vol.4, no.4, 1999, pp. 699-705.
Rekimoto, Jun, and Katashi Nagao. The World Through the Computer: Computer Augmented Interaction with Real World Environments. *Proceedings of UIST '95* (Pittsburgh, PA, 14-17 November 1995), 29-36.
Smith S. M. and Brady M.. "SUSAN_ A new approach to low level image processing". International Journal of Computer Visión, vol. 23(1), 45-78, 1997. También aparece: Technical Report TR95SMS1c, Defence Research Agency, Farnborough, England, 1995.
State, Andrei, Mark A. Livingston, Gentaro Hirota, William F. Garrett, Mary C. Whitton, Henry Fuchs y Etta D. Pisano. "Techniques for Augmented-Reality Systems: Realizing Ultrasound-Guided Needle Biopsies". Proceedings of SIGGRAPH '96 (New Orleans, LA, 4-9 August 1996), 439-446.
Wanstall, Brian. HUD on the Head for Combat Pilots. *Interavia 44* (April 1989), 334-338. [A89-39227].