Computer vision-based registration techniques for augmented reality

William A. Hoff, Khoi Nguyen
Colorado School of Mines, Division of Engineering
1500 Illinois St., Golden, CO   80401

Torsten Lyon
Surgical Navigation Technologies
530 Compton St., Broomfield, CO   80020

ABSTRACT

Augmented reality is a term used to describe systems in which computer-generated information is superimposed on top of the real world; for example, through the use of a see-through head-mounted display.   A human user of such a system could still see and interact with the real world, but have valuable additional information, such as descriptions of important features or instructions for performing physical tasks, superimposed on the world. For example, the computer could identify objects and overlay them with graphic outlines, labels, and schematics. The graphics are registered to the real-world objects and appear to be "painted" onto those objects. Augmented reality systems can be used to make productivity aids for tasks such as inspection, manufacturing, and navigation.

One of the most critical requirements for augmented reality is to recognize and locate real-world objects with respect to the person's head. Accurate registration is necessary in order to overlay graphics accurately on top of the real-world objects. At the Colorado School of Mines, we have developed a prototype augmented reality system that uses head-mounted cameras and computer vision techniques to accurately register the head to the scene. The current system locates and tracks a set of pre-placed passive fiducial targets placed on the real-world objects. The system computes the pose of the objects and displays graphics overlays using a see-through head-mounted display. This paper describes the architecture of the system and outlines the computer vision techniques used.

Keywords: augmented reality, registration, computer vision, pose estimation, fiducials, head-mounted displays

1. INTRODUCTION

Augmented reality (AR) is a term used to describe systems in which computer-generated information is superimposed on top of the real world[1].   Unlike virtual reality (VR), in which the person is immersed in a completely virtual world, AR involves enhancing or augmenting the user's perception of the real world. The enhancement could take the form of textual labels, virtual objects, or shading modifications. The graphical overlays can be generated using various technologies, such as stationary monitors or head-mounted displays (HMD). A fundamental feature of AR is the juxtaposition of real objects and virtual data, registered in 3-D. Ideally, the virtual and real objects appear to co-exist in the same space, and merge together seamlessly. Even as the user moves his or her head around, the graphics remain aligned to the real objects and appear to be "painted" onto those objects. Therefore, one of the key technical issues in AR is accurately registering the real and virtual worlds. The angular accuracy of registration must be particularly tight. Errors of a few pixels are detectable in modern HMD's, corresponding to a few tenths of a degree[2].

Augmented reality enhances a user's perception of the real world. The virtual objects can show information that the user cannot directly perceive with his or her own senses. For example, an AR system could amplify human sensory capability. Sensors need not be restricted to visual, but could include infrared and ultrasound. Raw data and processed results from these sensors could be displayed, co-registered with the actual view the person is seeing[3]. In another vein, computer processing can detect features that would go unnoticed by a person. For example, a moving target indicator (MTI) could continually monitor the scene for movement of a small object and alert a security officer.

Augmented reality has the potential for a large number of useful applications. Several groups have explored the use of AR for medical visualization and training aids, including visualization of a fetus inside the womb[4], and registration of MRI data with a patient's head[5]. The latter was done using computer vision to locate the patient's head with laser range data. In the area of manufacturing and maintenance, a group at Boeing is developing AR technology to guide a technician in building a wire harness for an airplane[6]. Feiner at Columbia has demonstrated a system for laser printer maintenance[7]. The emphasis of that work was on choosing the appropriate information to display, taking into account information about the user, task, and the position of objects. A group at the European Computer-Industry Research Center (ECRC) has developed a monitor-based AR system that features a hand-held pointer (tracked magnetically)[8]. The pointer is used to designate known points on an object in order to register it to the camera.

Two broad classes of augmented reality systems are those that (1) generate graphical overlays on video using monitors and (2) those that use head-mounted displays (HMD). Monitor-based systems are non-immersive and give the user a view of the world "through a window" as seen from the viewpoint of a remote camera. They have been used in teleoperation and supervisory control of remote robots. HMD-based systems, on the other hand, are immersive and let the user see the world directly surrounding him or her, augmented with additional graphical information. In this class of AR systems, the user is "on-site" and may interact with the real world directly without the need for a robot. Combining real and virtual objects can be done optically, using a optically transparent "see-through" graphics display, or digitally, by mixing the graphics with live video coming from cameras mounted on the head.

Computer vision techniques have the potential to provide the accurate registration data needed by AR systems. A recent survey of AR was done by Azuma[9], which found that the biggest single obstacle to building effective AR systems is the lack of accurate, long range sensors and trackers that report the locations of the user and the surrounding objects in the environment. Magnetic and ultrasonic trackers commonly used in VR do not provide the accuracy and portability needed in AR. Computer vision, on the other hand, can potentially recognize and locate objects in the environment, by measuring the locations of features in the world and tracking them over time, as the user moves his or her head. It is possible to provide accurate data at long ranges. By using miniature head-mounted cameras and belt-worn computers, a system could be developed that was truly portable.

This paper reports on the development of new computer vision-based registration techniques and their use in a working AR system. Our system uses head-mounted cameras and an optically transparent see-through graphics display (Figure 1). The scenario we are working with is that of personal computer (PC) maintenance. Our system can automatically determine the position and orientation (pose) of the PC with respect to the person's head. It continually displays graphical overlays showing the user the location of parts within the PC and guidelines for actions to perform.

Our system is unique in that in combines all of the following capabilities: The system uses completely passive sensors (*i.e.*, video cameras) and passive fiducial markings for landmark target points (*i.e.*, small white and black circles). The system automatically locates and tracks objects, without the need for any human intervention, from a wide variety of initial poses. It overlays graphics registered to those objects on a HMD. The objects and the person's head are both free to move, and are continually registered in real time. The system can recognize more than one object and distinguish between them (specifically, the outside and the inside of the PC).

Section 2 of this paper describes previous related work in registration for AR. Section 3 describes the computer vision techniques used in our registration system. Section 4 describes our overall AR system. Section 5 shows examples of its operation, and section 6 provides a discussion.

Figure 1 AR system using head-mounted cameras and see-through HMD.
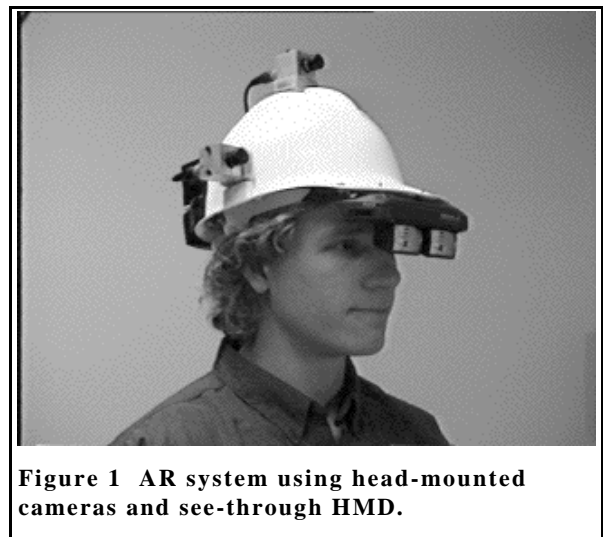
2. PREVIOUS WORK ON REGISTRATION FOR AR

Several AR systems have been developed that use non-vision technology for registration. The AR system developed by Feiner, *et al* used ultrasonic transmitters and receivers, mounted on the user's head and on the objects of interest[7]. The authors state that the registration is not accurate using these sensors. Tuceryan, *et al*, use a hand-held pointer that the operator uses to designate known points on an object in order to register it to the camera[8]. The pointer is tracked magnetically, using a receiver that is mounted on the pointer and a transmitter that is fixed in the environment. The authors report an accuracy of 0.65 cm and 1 degree in the computed object pose. However, other researchers have reported accuracy problems with magnetic trackers due to the presence of metallic objects in the environment[10]. The group at Boeing has also used magnetic sensors in an AR system; however, no specific results on accuracy were given[11].

Of the AR systems that use vision-based registration, most use head-mounted sensors as opposed to externally-mounted sensors. The reason is that it is easier to detect an head orientation change with a head-mounted sensor observing a fixed point in the environment, than with a fixed sensor observing a set of points on the head[12]. Azuma[13] developed a system that used head-mounted optoelectronic sensors to track light emitting diode (LED) beacons mounted on the ceiling. By the use of optical filters, the LED's can easily be detected in the sensors, while excluding other sources of illumination. The LED's were illuminated in sequence, so that there was no ambiguity in identifying which beacon was being illuminated. Azuma combined the optical sensing system with an inertial sensing system, and reports excellent accuracy (0.2 degrees and 2.7 mm average error for the combined system). However, this system requires the ceiling array of beacons to work, and since it does not directly sense the pose of the object of interest, cannot detect any movement of the object.

Bajura and Neumann at U. North Carolina[14] and Janin, *et al*, at Boeing[15] both describe AR systems that use a head-mounted video camera to detect LED's mounted on the object of interest. The LED's were illuminated continuously, and so a set of point features is detected in the camera image. As a result, determining the correspondence between the known LED beacons and the observed illuminated point features is not trivial. To determine correspondence, both systems predict the locations of the features based on the last known pose. Bajura takes the nearest predicted feature to each observed feature to be the correct match. Janin computes an optimal matching for all features using a transportation algorithm. However, both of these systems could fail if the initial pose "guess" was sufficiently far from the true pose, thus potentially causing an incorrect matching between the known beacons and the observed points.

Grimson, *et al* developed a surgical aid that registers a patient's head on the operating table with pre-operative MRI or CT data[5]. The system uses an externally mounted laser range finder to obtain a set of 3-D points on the patient's head. These points are matched to 3-D points from the MRI or CT model, and the pose of the head is determined. Overlays from the MRI or CT model are then projected onto a video image of the head, and displayed on a monitor. Mellor at MIT has developed a variation on this system which uses an alternative technique for the registration[16]. Mellor attaches four small passive fiducial targets to the patient's head, and determines their locations in an image automatically. He then measures the 3-D locations of the fiducials using a laser range finder. From that point on, the head can be registered quickly using only the video data.

Finally, Kutulakos and Vallino at U. Rochester have developed a system that can overlay 3-D graphics onto live video using a completely uncalibrated camera[17]. The system requires the operator to interactively specify (using a mouse) a set of points in the image, which are then tracked automatically. The operator also specifies the initial location of a virtual object. From then on, the virtual object is transformed into its correct location in subsequent images.

3. VISION-BASED REGISTRATION TECHNIQUE

This section describes a vision-based registration system that we have developed that is completely automatic, passive, and can acquire and track multiple objects starting from a wide variety of initial poses. The goal of our machine vision system was to identify and estimate the poses of objects of interest in the scene. Geometric models of the objects are assumed known. A technique similar to the one described was used by Hoff, *et al* in a computer vision-based teleoperator aid for robotics[18].

In general, model-based object recognition via computer vision typically involves: (a) extraction of features from the image, (b) finding a correspondence between image features and features on an object model, and (c) determining the pose of the object from the resulting correspondence[19]. A fully general, domain independent object recognition system is beyond the state of the art today. Vision systems of today have difficulties when (a) there are a large number of object models, (b) there are a large number of features, or (c) features may come from the background or from unknown objects.

The approach taken in our work was to greatly simplify the object recognition task by placing carefully designed fiducial targets on the object to be recognized. These targets are unique features that can easily and reliably extracted from the images. As a result, we do not have a large number of spurious features - all detected features come from the object of interest. To further simplify the correspondence process, the target points are arranged in a distinctive geometric pattern. These steps, as well as the pose estimation step, are described below.

3.1 Image features

Although significant progress has been made in the field of machine vision, no system exists at present that can identify large numbers of different objects against multiple backgrounds at video update rates. One alternative is to place fiducial targets, that can be recognized at video rates, on the objects. Visual targets which have been used to simplify the object recognition process are summarized by Gatrell[20].

The Concentric Contrasting Circle (CCC) image feature[20, 21] was used as a fiducial target. A CCC is formed by placing a black ring on a white background, or vice-versa. Figure 2 shows an example of the sequence of image processing operations that are performed to find the CCC's. The original image captured by the head-mounted camera is shown in (a). Using a simple thresholding operation, the black and white regions are easily separated, or segmented. Given the large contrast between the two regions, a wide range of threshold values will work. The raw thresholded image is shown in (b). Next, morphological image filtering operations are performed to eliminate small white or black regions. These filtering operations consist of an erosion followed by a dilation to eliminate small white regions, and a dilation followed by an erosion to eliminate small black regions[22]. The filtered thresholded image is shown in (c). Next, a connected component labeling operation is performed[23] to find connected white and black regions, as well as their centroids. The centroids of black regions are compared to the centroids of white regions — those black and white centroids that coincide are CCC's.

This image feature is invariant to changes in translation, scale, and roll; and is only slightly affected by changes in pitch and yaw. The image processing operations are linear with respect to the size of the image and can be reliably extracted from the image rapidly with low cost image processing hardware. The centroid of a circular shape is the most precisely locatable image feature[24]. By examining the color of the pixel located at the feature centroid, the CCC can be classified as either "white" (a white center surrounded by a black ring) or "black" (a black center surrounded by a white ring). Figure 2 (d) shows the original image with cross-hairs overlaid on the detected CCC's — vertical crosshairs on black CCC's and diagonal crosshairs on white CCC's.

3.2 Initial acquisition of target pattern

Once CCC's have been detected in the image, their correspondence to the model must be determined. In order to simplify the correspondence process, the target points are arranged in a distinctive geometric pattern. Four white CCC's are placed in a flat rectangular pattern on the object to be recognized. A fifth CCC is placed on a side of the rectangle to remove the roll ambiguity. The three colinear CCC's can be found by testing each subset of three points for colinearity. Once these are found, the remaining two points can be identified by their location relative to the first three. The result is that each visible target point is matched to a point on the model. In designing the five point target for a particular object, care must be taken to ensure that all five CCC's will be visible from the expected viewing positions. By offsetting the position of the middle CCC, two distinct target patterns can be created (Figure 3). The measured position of the middle feature point determines which of the two patterns is recognized.
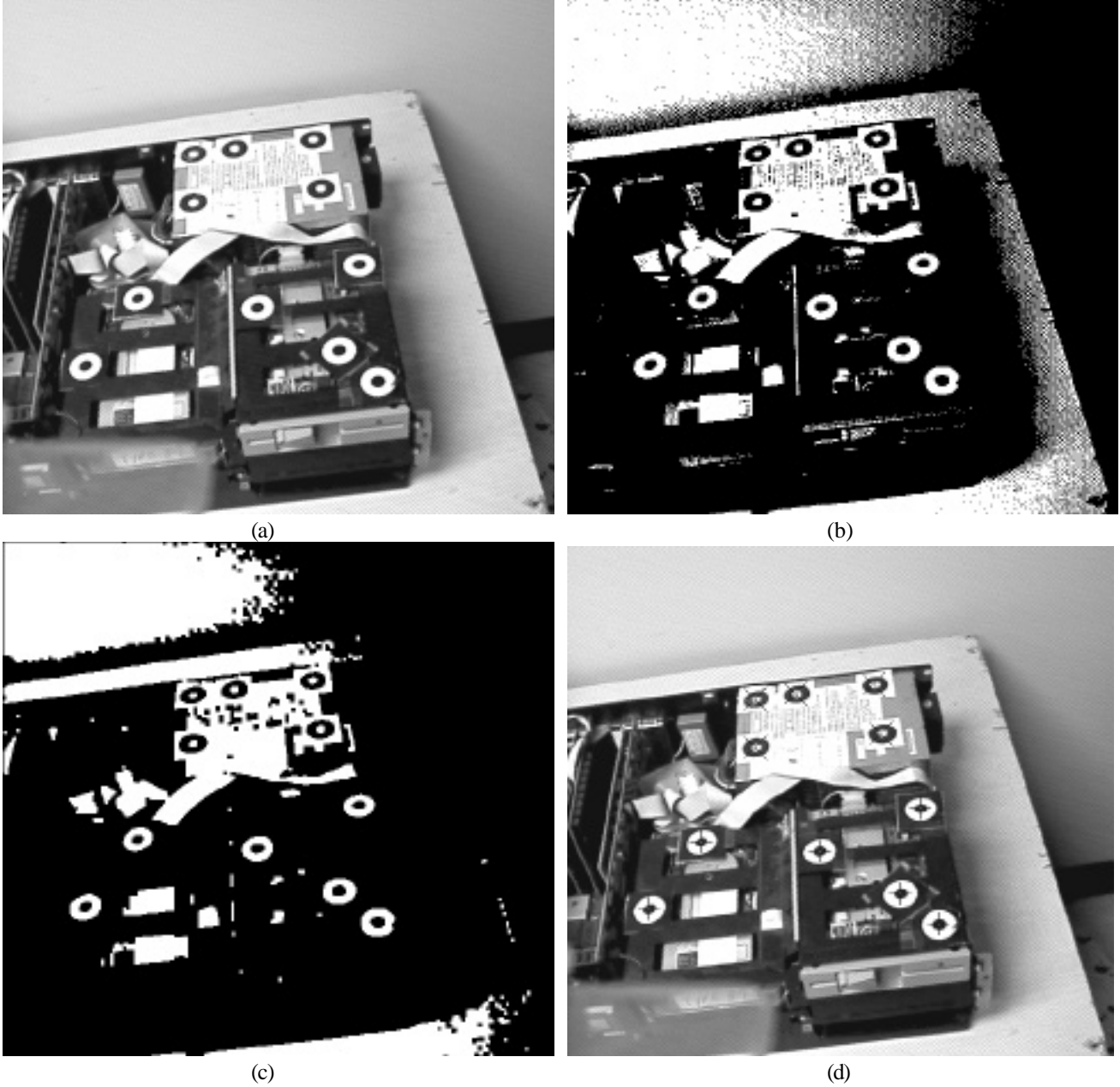
4

(a)

(b)

(c)

(d)

Figure 2 Example of image processing operations to locate target feature points (Concentric Contrasting Circles). (a) Original image, (b) thresholded image, (c) filtered image, (d) detected CCC's.
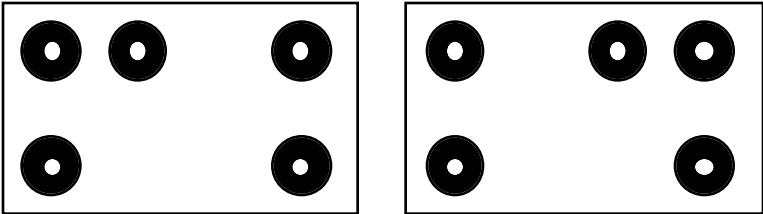

Figure 3 Two distinct target patterns formed by offsetting the position of one of the CCC's.

3.3 Pose Estimation

Once the target pattern is recognized, and the correspondence between image features and object features has been established, the pose of the object relative to the camera can be computed by many techniques[25]. We currently use the simple and fast Hung-Yeh-Harwood pose estimation method[26]. The inputs to the pose algorithm are the centers of the four corner CCC's, the target model, and a camera model. The pose algorithm essentially finds the transformation that yields the best agreement between the measured image features and their predicted locations based on the target and camera models. In this work, a relatively simple camera model was used — that of a pinhole camera with an aspect ratio scaling factor. In the future, we plan to use a more sophisticated camera model that incorporates lens distortion.
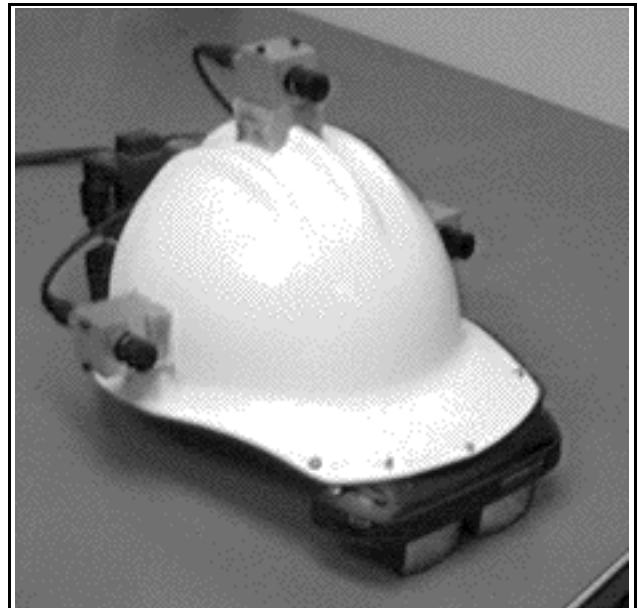
In the work described in the paper, we used a single head-mounted camera. However, we have developed a multi-camera algorithm and have shown that it can provide significantly more accurate pose data than a single camera, especially for small target configurations[27]. We are currently working to incorporate this algorithm into the current system.

4. OVERALL SYSTEM ARCHITECTURE

This section describes the overall architecture of a prototype augmented reality system. Figure 4 contains a picture of the helmet apparatus developed. The helmet incorporates an optical see-through stereo display mounted in front of the user's eyes, three color CCD cameras mounted on either side of and on top of the helmet, and inertial sensors mounted at the rear of the helmet. This proof-of-concept system has been constructed from off the shelf components rather than investing large amounts of time and money in developing custom components. The helmet chassis upon which all components are affixed is simply a hard hat with adjustable head strap. This works well as it allows the head strap of the helmet to be adjusted for each user while maintaining inter-camera and camera-to-display rigid calibrations.

The optical see-through graphics display is a commercially available product called i-glasses™, manufactured by Virtual i-o Corporation. The advantage of this product is its low cost and ease of use. Several different types of video input are accommodated, including VGA, for display on the glasses. Additionally, a stereo display electronics module separates the odd and even fields from the video image and displays only one field for each eye, thus allowing different images to be displayed for each eye. Shifting the graphics overlays presented to each eye provides three-dimensional overlay capabilities. One disadvantage of the product is the attenuation of light from the model scene, however this is an artifact of all optical see-through displays. Another disadvantage is the limited graphics resolution. This does not turn out to be a major factor since the graphics overlays in AR models are relatively simple and do not require extensive graphics rendering. The field-of-view of the i-glasses display system is 30 degrees in each eye.

The video cameras mounted on the CSM AR helmet are micro-head charge-coupled-device (CCD) color cameras manufactured by Panasonic, model number GP-KS162. The cameras are remote head devices which allow the camera heads to be extremely small and light weight, with the camera control units

Figure 4 AR helmet, featuring see-through stereo display, three color CCD cameras (side and top), and inertial sensors (rear).

being placed off-platform from the user. The system incorporates lenses with a nominal focal length of 7.5mm (Panasonic part number GP-LM7R5TB), providing a nominal field of view of 44 degrees. Each camera head incorporates a 1/2" CCD, weighs 14g, and provides 480 horizontal lines of resolution. For the work described in this paper, we only used one camera; however, we are currently extending the algorithms to make use of all three.
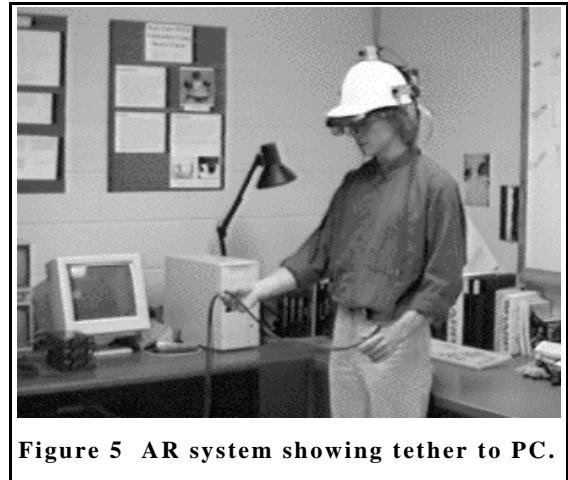
Positioned on the rear of the helmet are inertial sensors consisting of a three-axis gyroscope and three orthogonally mounted single-axis accelerometers. The gyroscope is manufactured by Watson Industries and is capable of measuring angular rates up to $\pm 100°$ per second with resolution limited by instrumentation noise floor. The accelerometers are manufactured by IC Sensors and are capable of measuring full-scale accelerations of $\pm 2g$ with an accuracy of $\pm 1\%$. We have developed a data acquisition and filtering system to process the data from the inertial sensors and use them to estimate short term head motions. We are currently working to integrate this system with the vision system.

Finally, the system is tethered to allow exchange of data between the AR helmet and the PC platform. The tether includes three video signals from the cameras to the camera control units, inertial sensor information from the gyroscope and accelerometer, and a graphics overlay signal from the PC to the i-glasses see-through display. The tether can be seen in Figure 5.

The computer system includes an IBM PC compatible computer in conjunction with the helmet-mounted hardware already discussed. Figure 6 presents an overview of the flow of information in our AR system. The helmet apparatus at the top of the diagram includes the CCD cameras, inertial sensors and optical see-through display. The general flow of information includes a video signal from each camera as well as angular velocities and accelerations from the inertial trackers to the PC computer. Software algorithms process this data and generate graphic overlays for displaying on the see-through display.

Currently, the CSM AR project does not attempt to process information from inertial sensors. The advantages of incorporating inertial sensors in AR systems can be significant. Azuma[13] discusses a system incorporating head-mounted inertial sensors to predict head motion. Results presented by Azuma suggest prediction with inertial sensors produces errors 2-3 times smaller than prediction without inertial sensors, and 5-10 times smaller than using no prediction at all.

Figure 5 AR system showing tether to PC.

A brief synopsis of the process being performed by the PC follows. Since inertial inputs are not currently incorporated, the only inputs to the computer are video signals from each of the cameras. The Sharp board is used to digitize the video signals, resulting in binary images with discrete pixel resolution of 512 columns by 480 rows. The Sharp board has built-in image processing routines to perform the thresholding, morphological filtering, and connected component labeling steps described in the previous section. Pose estimation is done by the host PC. The resulting pose estimate is used to generate a projection of the overlay model into the user's reference frame. A projection of the model is output to the see-through display for presentation to the user.

This system uses a 100 MHz 80486 IBM compatible PC running QNX version 4.22. The throughput currently achieved with the system described above is approximately 120 ms per update iteration, or 8.3 Hz. This update rate is too slow for many applications but it is sufficient to prove the feasibility of concepts and algorithms. Incorporating inertial sensors could increase the update rate and the accuracy of the graphical overlays.

5. RESULTS FROM EXAMPLE SCENARIO

In this section we show some sample output of the system on an example scenario - that of PC maintenance. We developed a system that could provide informative graphical overlays, to provide directions to a person and assist in identifying parts within the PC. We affixed a set of fiducial targets to the cover of a PC, and also to the interior frame of the chassis. A different five-point target pattern was used on the cover and on the interior, thus enabling the system to distinguish the two.
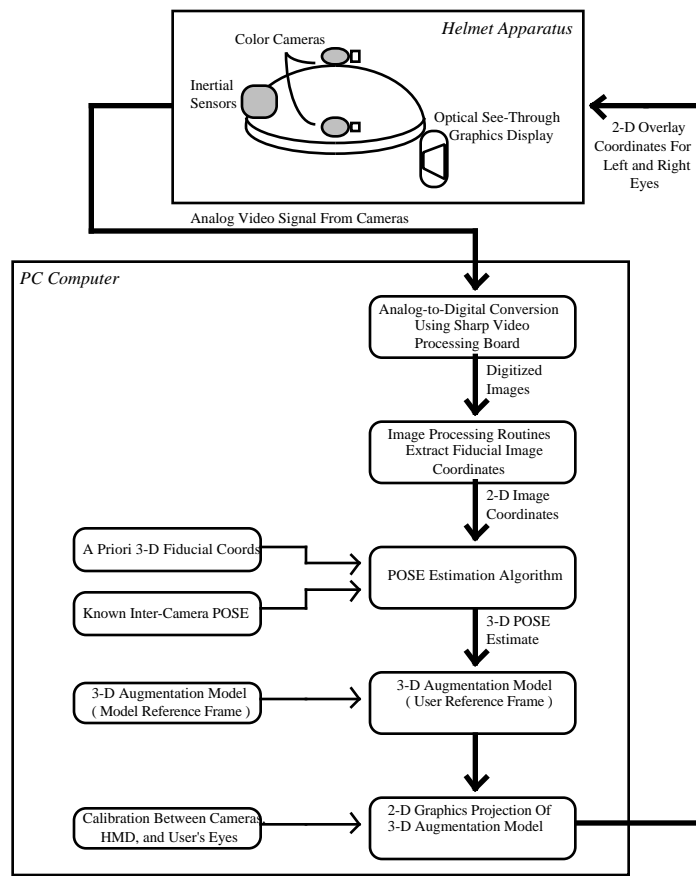
**Figure 6 Overview of processing flow in augmented reality system.**

Figure 7 (a) shows a person wearing the AR helmet, looking at the cover of the PC. Figure 7 (b) shows the view that the person would see through the left eyepiece of the HMD. To obtain this picture, we placed a video camera at the position of the person's left eye, and grabbed a frame from the video output. Visible in the picture is the outside frame of the eyepiece, and the overlays through the center of the eyepiece. The overlays in this particular example are a set of arrows, which show the person how to pull off the cover of the PC.

(a)	(b)

Figure 7 (a) Person wearing AR system. (b) View through head-mounted display, showing overlays indicating direction to remove cover of PC.

Once the cover is off, the person looks at the interior of the PC. Figure 8 (a) shows an image from the head-mounted camera, with overlays. Figure 8 (b) shows the view as seen through the eyepiece of the HMD. In this example, the overlays outline the position of the disk controller board, and also the backup battery. (Note: not shown in these examples is the view from the right eyepiece, which is slightly different to achieve a stereo effect.)

(a)   (b)
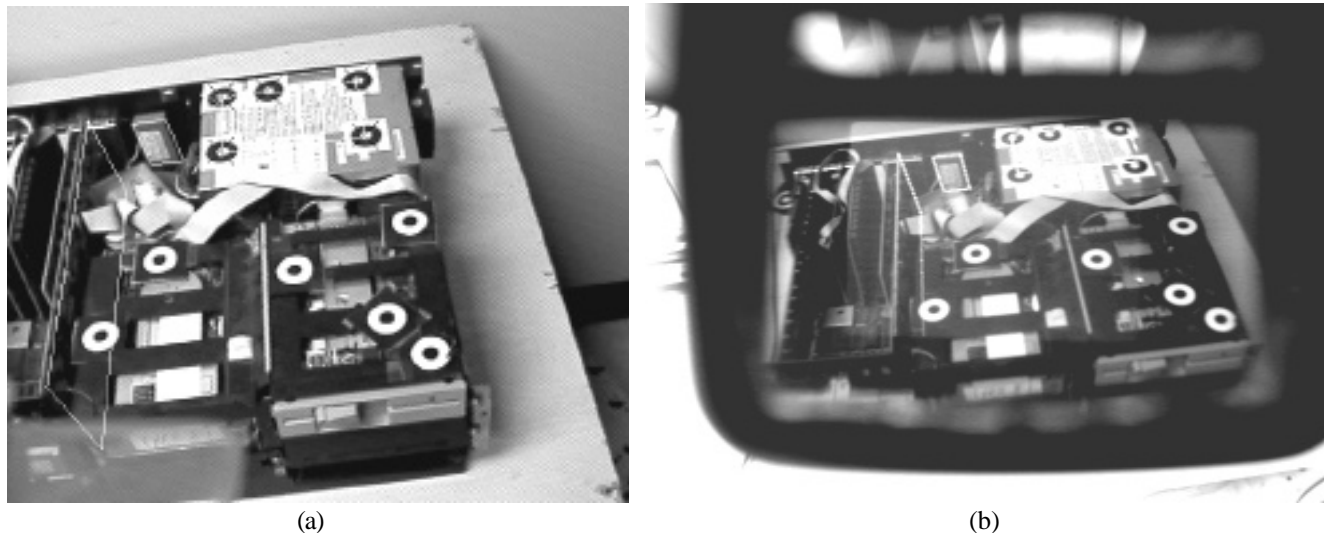
Figure 8 (a) Image of PC interior, from head mounted camera. (b) View through head-mounted display. Overlays in both images outline the positions of the disk drive board and also the backup battery.

6. DISCUSSION

This paper has described a prototype augmented reality system that uses computer vision to recognize an object and register it to the user's head-mounted display. To simplify the object recognition task, fiducial targets were placed on the objects to be recognized. Our system is unique in that it uses completely passive sensors (*i.e.*, video cameras) and passive fiducial markings for landmark target points (*i.e.*, small white and black circles). The system automatically locates and tracks objects, without the need for any human intervention, from a wide variety of initial poses. It overlays graphics registered to those objects on a HMD. The objects and the person's head are both free to move, and are continually registered in real time. The system can recognize more than one object and distinguish between them (specifically, the outside and the inside of the PC).

A principal limitation of the current system is the need to keep the five-point target pattern within the field of view of the head-mounted camera. If the target moves outside the field of view, then the pose of the object cannot be determined and no overlays are drawn. One approach to addressing this limitation is to add additional fiducial points to the object. Figure 2 shows a set of "black" CCC's, which can easily be distinguished from the "white" CCC's used in the five point target. The additional fiducials are visible from a large set of viewing poses. Another approach is to incorporate inertial sensors (gyroscopes, accelerometers) to estimate head pose from dead reckoning when the target pattern is not visible. Finally, more general object recognition techniques would allow the recognition and registration of objects without the need for fiducial targets.

7. REFERENCES

[1]   W. Robinett, "Synthetic experience: A proposed taxonomy," *Presence: Teleoperators and Virtual Environments*, Vol. 1, No. 2, pp. 229-247, 1992.

[2]   R. Azuma, "Tracking requirements for augmented reality," *Communications of the ACM*, Vol. 36, No. 7, pp. 50-51, 1993.

[3] W. Robinett, "Electronic expansion of human perception," *Whole Earth Review*, 1991, pp. 16-21.

[4] M. Bajura, H. Fuchs, and R. Ohbuchi, "Merging virtual objects with the real world: Seeing ultrasound imagery within the patient," *Computer Graphics*, Vol. 26, No. 2, pp. 203-210, 1992.

[5] W. E. L. Grimson, et al, "An automatic registration method for frameless stereotaxy, image guided surgery, and enhanced reality visualization," *Proc. of Computer Vision and Pattern Recognition*, IEEE, pp. 430-436, 1994.

[6] T. P. Caudell, "Introduction to augmented and virtual reality," *Proc. of Telemanipulator and Telepresence Technologies, Proc. SPIE*, Vol. 2351, pp. 272-281, 1994.

[7] S. Feiner, B. MacIntyre, and D. Seligmann, "Knowledge-based augmented reality," *Communications of the ACM*, Vol. 36, No. 7, pp. 53-62, 1993.

[8] M. Tuceryan, et al., "Calibration requirements and procedures for a monitor-based augmented reality system," *IEEE Trans. Visualization and Computer Graphics*, Vol. 1, No. 3, pp. 255-273, 1995.

[9] R. Azuma, "A survey of augmented reality," *Proc. of SIGGRAPH '95 Course Notes*, ACM, Los Angeles, August 8, 1995.

[10] K. Meyer, et al, "A Survey of Position Trackers," *Presence*, Vol. 1, No. 2, pp. 173-200, 1992.

[11] T. P. Caudell and D. W. Mizell, "Augmented Reality: An Application of Heads-up Display Technology to Manual Manufacturing Processes," *Proc. of Hawaii Int'l Conf. on System Sciences*, January, pp. 659-669, 1992.

[12] J.-F. Wang, R. Azuma, G. Bishop, V. Chi, J. Eyles, and H. Fuchs, "Tracking a Head-Mounted Display in a Room-Sized Environment with Head-Mounted Cameras," *Proc. of Helmet-Mounted Displays II*, Vol. 1290, SPIE, Orlando, FL, April 19-20, pp. 47-57, 1990.

[13] R. Azuma, "Improving static and dynamic registration in an optical see-through HMD," *Proc. of SIGGRAPH 94*, ACM, pp. 197-204, 1994.

[14] M. Bajura and U. Neumann, "Dynamic Registration Correction in Augmented-Reality Systems," *Proc. of Virtual Reality Annual Int'l Symposium*, IEEE Computer Society Press, Research Triangle Park, NC, March 11-15, pp. 189-196, 1995.

[15] A. Janin, K. Zikan, D. Mizell, M. Banner, and H. Sowizral, "A Videometric Head Tracker for Augmented Reality Applications," *Proc. of Telemanipulator and Telepresence Technologies*, Vol. 2351, SPIE, Boston, MA, Oct 31 - Nov 4, 1994.

[16] J. P. Mellor, "Enhanced Reality Visualization in a Surgical Environment," MIT, Cambridge, MA, A.I. Technical Report 1544, January 1995.

[17] K. N. Kutulakos and J. Vallino, "Affine Object Representations for Calibration-Free Augmented Reality," *Proc. of VRAIS*, IEEE Computer Society, Santa Clara, CA, March 30 - April 3, pp. 25-36, 1996.

[18] W. Hoff, L. Gatrell, and J. Spofford, "Machine Vision Based Teleoperation Aid," *Telematics and Informatics*, Vol. 8, No. 4, pp. 403-423, 1991.

[19] W. E. L. Grimson, *Object Recognition by Computer*, Cambridge, Massachusetts, MIT Press, 1990.

[20] L. Gatrell, W. Hoff, and C. Sklair, "Robust Image Features: Concentric Contrasting Circles and Their Image Extraction," *Proc. of Cooperative Intelligent Robotics in Space*, Vol. 1612, SPIE, W. Stoney (ed.), 1991.

[21] C. Sklair, L. Gatrell, and W. Hoff, "Optical Target Location Using Machine Vision in Space Robotics Tasks," *Proc. of Advances in Intelligent Systems*, Vol. 1387, SPIE, November, pp. 380-391, 1990.

[22] W. K. Pratt, *Digital Image Processing*, 2nd ed., New York, Wiley & Sons, 1991.

[23] D. H. Ballard and C. Brown, *Computer Vision*, New Jersey, Prentice Hall, 1982.

[24] C. Bose and I. Amir, "Design of Fiducials for Accurate Registration Using Machine Vision," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 12, No. 12, pp. 1196-1200, 1990.

[25] R. Haralick and L. Shapiro, *Computer and Robot Vision*, Addison-Wesley Inc, 1993.

[26] Y. Hung, P. Yeh, and D. Harwood, "Passive Ranging to Known Planar Point Sets," *Proc. of IEEE International Conference on Robotics and Automation*, Vol. 1, St. Louis, Missouri, 25-28 March, pp. 80-85, 1985.

[27] T. M. Lyon, "Three-Dimensional Pose Estimation from Noisy Two-Dimensional Image Observations Using Multiple Cameras," *MS Thesis, Division of Engineering*. Golden, CO: Colorado School of Mines, 1996, pp. 149.