

Registration for Outdoor Augmented Reality Applications Using Computer Vision Techniques and Hybrid Sensors *

Reinhold Behringer
Rockwell Science Center (RSC)
1049 Camino Dos Rios, Thousand Oaks, CA 91360
reinhold@rsc.rockwell.com

Abstract

Registration for outdoor systems for Augmented Reality (AR) cannot rely on the methods developed for indoor use (e.g., magnetic tracking, fiducial markers). Although GPS and the earth's magnetic field can be used to obtain a rough estimate of position and orientation, the precision of this registration method is not high enough for satisfying AR overlay. Computer vision methods can help to improve the registration precision by tracking visual clues whose real world positions are known. We have developed a system that can exploit horizon silhouettes for improving the orientation precision of a camera which is aligned with the user's view. It has been shown that this approach is able to provide registration even as a stand-alone system, although the usual limitations of computer vision prohibit to use it under unfavorable conditions. This paper describes the approach of registration by using horizon silhouettes. Based on the known observer location (from GPS), the 360 degree silhouette is computed from a digital elevation map database. Registration is achieved, when the extracted visual horizon silhouette segment is matched onto this predicted silhouette. Significant features (mountain peaks) are cues which provide hypotheses for the match. Several criteria are tested to find the best matching hypothesis. The system is implemented on a PC under Windows NT. Results are shown in this paper.

1. Introduction and context

1.1. Augmented Reality

In recent years, augmented reality (AR) has gained significant attention [2] due to rapid progress in several key areas (wearable computing, virtual reality, rendering hardware) [7]. By bridging the gap between virtual reality and the real world, it occupies a central position in the reality-virtuality continuum [18] by exploiting technologies and methods developed in the virtual reality domain. Although often associated with visualization (starting with the first head-mounted display by Sutherland [31]), augmentation

can also occur in the aural domain [3] [9] [21]. AR technology provides means of intuitive information presentation for enhancing the situational awareness and perception by exploiting the natural and familiar human interaction modalities with the environment. Examples of utilizing these familiar modalities are augmenting paper drawings [16], a desk environment [25], and pen and notebook [32].

There is a wide field for possible applications of AR systems: *Virtual prototyping*, as employed in architecture [39] and interior design [1], is a natural candidate for mixing virtual prototypes with the real world. AR methods in the factory environment can be applied for training [26], for analyzing [20], and for enhancing [19] the manufacturing processes. For tele-presence tasks, AR can provide means for human-robot interaction [17], e.g. for mining [23], nuclear reactor inspection [10] or remote dam inspection [11]. Medical AR can provide means of advanced visualization as for ophthalmic treatment [5], spine surgery [6], ultrasound data [27], MRI and CT [13] visualization and overlay onto the patient's body [35]).

In outdoor scenarios, AR can provide an aid for navigation in an urban setting [12]. We are developing a system that can enhance the situational awareness of users operating in hazardous outdoor conditions, as fire fighters or dismounted soldiers. Information can be overlaid into the field of view of the user by a see-through head-mounted display.

1.2. Registration for AR

An outdoor system can utilize GPS information for obtaining user location. Orientation can be obtained by a digital magnetic compass. Useful are also inclinometers which provide roll and tilt angle information. However, the precision of these devices is less than would be required for true visual merging of rendered world and real world perception. Therefore, many AR systems employ computer vision technology for achieving the required high registration precision, (e.g., [15]). The approach in most cases is to detect visual features which have a known position, and recover camera orientation and position through a matching or pose recovery process. Such methods have been developed for autonomous robotics systems (e.g. [33]).

*This work was supported in part by the Federated Laboratory Advanced Displays and Interactive Displays Consortium (ARL) under Grant DAAL01-96-2-0003

In indoor scenarios, these methods can utilize artificial markers affixed to objects or placed in the environment [8] [29]. In the case of an outdoor scenario, the lack of fiducial markers requires the utilization of visual features provided by natural [24] or man-made objects, such as buildings or characteristic terrain features. We have begun to investigate using terrain horizon silhouettes for visual registration of an AR system. Digital terrain elevation maps provide the necessary data based for the feature matching process.

1.3. Silhouettes as visual clues

There has been research effort in utilizing visual silhouettes for reconstruction of curved object shapes [38] and simultaneous motion estimation [14]. In an outdoor scenario, silhouettes occur at man-made structures (buildings) and from the natural terrain shape.

Natural horizon silhouettes are the visual boundaries between the terrain and the sky. In a well-structured terrain, they provide clues for human visual orientation and navigation. Although not necessarily unique, they can be used to determine position and orientation of the observer (solving the so-called “drop-off problem” [34]), if the area is confined to a given region. Stein [28] demonstrated this using digital elevation maps to match visual and predicted horizon silhouettes by structural indexing.

This paper describes an outdoor registration approach which is being developed by the Rockwell Science Center. It is intended to improve the registration precision in an outdoor scenario by using visual terrain horizon silhouettes as cues for determining the camera orientation [4], if the observer position is known (e.g., from GPS). The silhouette in the video image is extracted by a standard computer vision edge detection technique (Sobel operator). This approach requires the presence of well-structured terrain within the field of view of the camera. The approach is based on matching the location of mountain peaks with data from digital elevation maps and does not require camera calibration.

2. About digital terrain maps

Digital elevation maps provide terrain elevation data in a grid pattern of points. From the elevation data, the view of the terrain from any point can be rendered. Visual features can be predicted similarly as this is done by CAD models, and these features can then be matched with features extracted from real video images.

Digital elevation maps are available in various resolutions and file formats. The most common file formats are denominated as Digital Elevation Model (DEM) maps and Digital Terrain Elevation Data (DTED) maps. DEM data are available from US Geological Survey (USGS) [37]; DTED maps are available from the National Imagery and Mapping Agency (NIMA) [22].

2.1. Data format

The elevation data in digital elevation maps are aligned in profile lines. DEM maps, which have been used in the described work, are available in two formats: 1 degree and 7.5 minute maps. Details about the file formats are published by USGS [36].

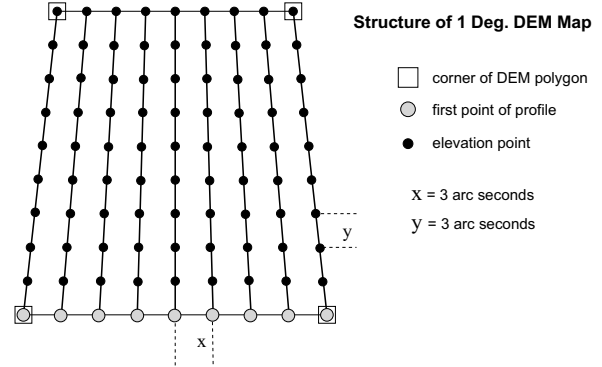


Figure 1. Structure of 1 degree map.

The 1 degree format covers elevation data in a region with an extension of 1-deg by 1-deg. The elevation data sample points are aligned to the geographic coordinate system along the latitude ϕ and longitude ψ axes in a 3 arc sec raster (see Fig. 1).

In order to compute a horizon silhouette from any point within the map, the location of each data point has to be converted into a cartesian coordinate system with the current viewpoint as the origin. After linearizing the mapping equations around this point $P(\psi_p, \phi_p)$, the transformation of a point $P_1(\Delta\psi, \Delta\phi)$ is approximately given by:

$$\Delta x_p \approx r_e \Delta\psi_p (\cos \phi_p - \Delta\phi \sin \phi_p) \quad (1)$$

$$\Delta y_p \approx r_e (\Delta\phi + (\Delta\psi)^2 (\sin \phi_p + \Delta\phi \cos \phi_p) / 4) \quad (2)$$

with:

$$\Delta\psi = \psi - \psi_p$$

$$\Delta\phi = \phi - \phi_p$$

$$r_e = 6,370,997 \text{ m (radius of earth)}$$

Depending on the latitude ϕ , the grid can be assumed to be cartesian, if only a small vicinity around P is to be considered. In general, the grid distance of the data points depends on the latitude ϕ :

$$\Delta y = 92.66 \text{ m}$$

$$\Delta x = \Delta y \cos \phi$$

The smaller 7.5 minute maps cover a higher density grid which is a cartesian coordinate system in UTM coordinates. The grid spacing here is a constant 30 m.

2.2. Natural horizon silhouette

A silhouette contour in a terrain is formed by those terrain points which have a surface normal vector perpendicular to the viewing vector (seen from the observation point P_0 – Figure 2).

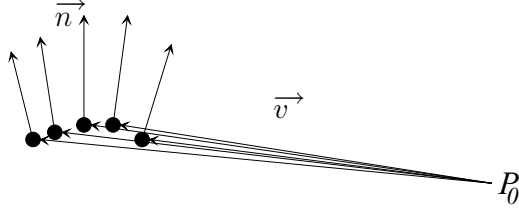


Figure 2. Normal vectors \vec{n} and viewing vectors \vec{v} .

One way to compute silhouette contours from a digital elevation map is to calculate the scalar product of the vector normal \vec{n}_{ij} and the viewing vector \vec{v}_{ij} at each terrain point P_{ij} and test it for being zero (or for having neighboring points where this scalar product has opposite signs). In order for P_{ij} to be a silhouette point, the curvature H_{ij} in direction of \vec{v} at P_{ij} must be positive. The complete surface contour silhouette S_s is the set of points that fulfill this condition:

$$\begin{aligned} \vec{v}_x &= \overrightarrow{P_{i-1,j}P_{i+1,j}} \\ \vec{v}_y &= \overrightarrow{P_{i,j-1}P_{i,j+1}} \\ \vec{n}_{ij} &= \vec{v}_x \times \vec{v}_y \\ S_s &= \{P_{ij} | (\vec{v}_{ij} \cdot \vec{n}_{ij} = 0) \wedge (H_{ij} > 0)\} \end{aligned} \quad (3)$$

This method produces all silhouette segments formed by hilly terrain. Visibility criteria have to be applied in order to extract the visible silhouette contour segments in case of occlusion. The horizon silhouette S_h is the subset of S_s that is formed by the terrain points seen from the observer location P_0 under the largest elevation angle for a given azimuth α (see Figure 3). Given the relative height h_{ij} and the distance d_{ij} between P_0 and P_{ij} , the elevation angle is calculated by:

$$\theta_{ij} = \arctan(h_{ij}/d_{ij}) \quad (4)$$

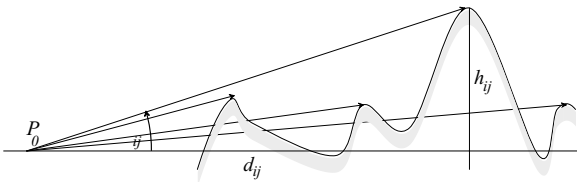


Figure 3. Several silhouette points at the same azimuth angle in various distances.

The quantization of the elevation data grid makes it necessary to find the maximum elevation angle within an azimuth segment of finite width $\Delta\alpha$ (Figure 4). The closer the points are to P_0 , the smaller is the area covered by this segment. Eventually, close neighboring points may not be within the segment. Therefore, the analysis of the elevation angle must also include neighboring points which are not within the current segment $\Delta\alpha$.

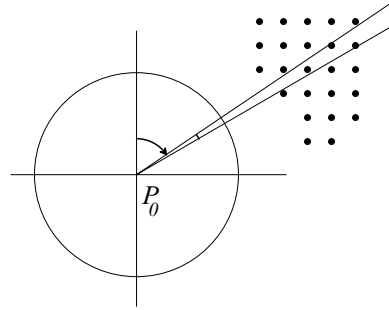


Figure 4. Discretization of silhouette.

3. Concept for silhouette registration

Visual registration can be achieved by trying to match features, which are extracted from reality (e.g., through video image analysis), with features which are predicted from a model or a database of the environment. Digital elevation maps provide the model to predict visual clues like terrain contour silhouettes, mountain peaks, or other remarkable terrain features for visual matching. Of course, a registration system based solely on visual terrain feature extraction is not suitable, because not in every instance such features are available (e.g., ocean or flat plain in desert). Using this approach to obtain the observer location without any other means requires a lot of computational power, which is often not available in an outdoor situation. However, this method can be used to calibrate a registration system which otherwise uses conventional registration methods (inclinometer, magnetometer). Under the boundary condition that the terrain is well-structured and the position of the observer is known, this approach can increase the precision of these registration methods.

We have developed an approach for using DEM maps as database for obtaining the camera orientation in a well-structured terrain by visual horizon silhouette matching. The position is assumed to be known through GPS measurements. The approach is outlined as follows:

- The visual silhouette is extracted from a single video image by using edge detection techniques.
- Based on the known location of the camera/observer, the horizon silhouette is computed for the complete 360 degree surrounding from a digital elevation map.
- Both the visual silhouette and the predicted 360 degree silhouette are evaluated for extrema (peaks and dips).

- Hypotheses are generated for the correspondence of these singularities. These hypotheses define pitch and yaw angle and camera calibration parameters.
- The hypothesis with the minimal distance between the visual silhouette segment and the silhouette from the digital elevation maps is then assumed to be the optimal match. The hypothesis for pitch and yaw angle is fine-tuned by matching discrepancies.

This approach does not assume an initial orientation or a calibration of the camera. However, the computation cost is reduced greatly, if an initial “rough” estimate of the orientation can be given in order to reduce the number of hypotheses to be tested. The method is intended to be used as an initialization for registration. If registration has been achieved after this method has been applied, another approach should be taken to “track” the camera orientation without starting the silhouette matching process again. This method could be a visual feature tracking approach or similar. The current concept, however, deals solely with the initial registration.

3.1. Visual silhouette extraction

The horizon silhouette in a video image is usually characterized by a steep grayscale gradient. This edge can be detected by applying an edge detector, which is scanning the image from top to bottom. In our approach, a Sobel operator is used to detect the mainly horizontal edge. In the case of a hybrid system, an inclinometer could indicate the tilt and roll angle and set the parameters for the Sobel search region and edge angle appropriately. Usually, the operator will detect several edge candidates for each search path. After the image scan has been completed, constraints are applied in order to select only one horizon point per search path. One constraint is that the silhouette must form a piece-wise continuous contour. Another constraint is that the sky area is usually more uniform than the ground area. This could be used to develop a horizon detection algorithm which is based on areal segmentation (texture, color) rather than on edge detection. Also, a horizon edge is very sharp, as opposed to fuzzy edges caused by trees.

The occlusion of the horizon is problematic; if nearby trees or buildings cover the horizon edge, the horizon cannot be detected. Using a stereo vision system or other means to detect scenery depth could help separate foreground and background. But the problem of far away obstructions would remain. Therefore, the constraints must be applied strictly, to flag areas where a horizon edge cannot be identified. The resulting silhouette is an array of piece-wise contiguous points in the image plane.

3.2. Formation of hypotheses

Using the horizon silhouette as visual cue for registration, the problem of determining the camera orientation is reduced to finding the optimal match of two arrays (visual silhouette segments and digitally predicted silhouette). The first step is to detect maxima and minima in both silhouettes.

Then hypotheses are generated for correspondence of these extrema. In order to achieve the highest precision, these features in the visual silhouette should be far apart from each other. The selection of these correspondence points must consider assumptions about the minimum and maximum opening angle of the camera.

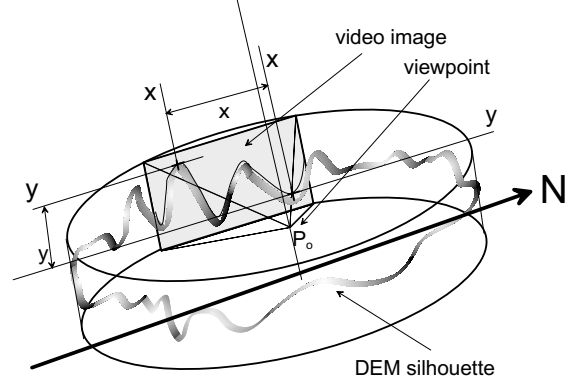


Figure 5. The generation of a hypothesis by assuming correspondence between one maximum and one minimum in both video and DEM silhouette.

Starting with a combination of two extrema in the video silhouette at (x_1, y_1) and (x_2, y_2) , a hypothesis is generated for the corresponding extrema (α_1, θ_1) and (α_2, θ_2) in the 360 degree silhouette. The hypothetical match between these two pairs leads to the assumed camera azimuth α_c and elevation θ_c :

$$\alpha_c = \alpha_2 - \arctan \left(\frac{\sin(\Delta\alpha)}{\cos(\Delta\alpha) - x_1/x_2} \right) \quad (5)$$

$$\theta_c = \frac{\theta_1 * y_2 - \theta_2 * y_1}{\Delta y} \quad (6)$$

It should be noted that the current algorithm is based on the assumption of the camera roll angle being zero. However, when using a third point correspondence, additional equations can be set up for determining the roll angle.

Eq. (5) implicitly considers the non-linear camera “distortion” at the border of a wide angle lens (due to mapping onto spherical image plane). Since the vertical extension $\Delta\theta$ is usually not very large, eq. (6) for the pitch angle is written in the approximate linearized form. Calibration is not required; in fact, the matching hypothesis also provides a hypothesis for the calibration parameters under the assumption of pinhole mapping with the following mapping transformation $(X_c, Y_c, Z_c) \rightarrow (x_c, y_c)$:

$$\begin{bmatrix} x_c \\ y_c \end{bmatrix} = \frac{f}{X_c} \cdot \begin{bmatrix} k_y Y_c \\ k_z Z_c \end{bmatrix} \quad (7)$$

The camera focal length f is included in these calibration parameters $C_{x,y} = f \cdot k_{y,z}$. The hypothesis for the matching parameters can be written as:

$$C_x = \frac{\Delta x}{2 * \cos \frac{\alpha_1 + \alpha_2 - 2 * \alpha_c}{2} * \sin \frac{\Delta \alpha}{2}} \quad (8)$$

$$C_y = \Delta y / \Delta \theta \quad (9)$$

3.3. Error minimization

After the hypothesis has been generated, an error measure is calculated to describe the error of the hypothesis. Registration of the image view with the 3D world is achieved when the x-y-set of video silhouette points can be mapped onto the $\alpha - \theta$ -set of the predicted 360 degree silhouette with the least possible deviation. The error criterion is the “difference” between the DEM silhouette and the visual silhouette segment.

Since the point correspondence is only assumed for the two chosen extrema point pairs, the rest of the silhouettes do not have point correspondences. In order to compute the distance between the silhouettes, for each DEM silhouette point the closest interpolated video silhouette point is selected as a corresponding point. The square sum of all these distances d_i within the field of view of the camera is used as an error measure e for judging the hypothesis (eq. (10)). The hypothesis with the least error e – below a given threshold – is chosen as the valid one. If no hypothesis is matched with an error below the threshold, no matching is assumed to be achieved. This error criterion also provides a measure for fine-tuning the hypothesis: the average horizontal and vertical difference component between the points of both silhouettes ($d_{x,i}, d_{y,i}$) can be used to correct the assumption of camera azimuth α_c and elevation θ_c as well as the calibration factors. The hypothesis based on matching of silhouette extrema may not necessarily be very precise, since the DEM silhouette data are not very dense. By applying a straight-line approximation for $d_{x,i}$ (eq. (11)) and $d_{y,i}$ (eq. (12)), the correction $\Delta \alpha_c$ and $\Delta \theta_c$ of the camera orientation angles and the corrective terms ΔC_x and ΔC_y for the calibration factors can be obtained from the averaging lines $\alpha_i(d_{x,i})$ and $\theta_i(d_{y,i})$:

$$e = \sqrt{\frac{\sum d_i^2}{n}} \quad (10)$$

$$\alpha_i(d_{x,i}) = \frac{\Delta C_x}{C_{x,0}} d_{x,i} + \Delta \alpha_c \quad (11)$$

$$\theta_i(d_{y,i}) = \frac{\Delta C_y}{C_{y,0}} d_{y,i} + \Delta \theta_c \quad (12)$$

The error in camera azimuth and elevation can be computed as the square root of the square sum of each horizontal and vertical difference component:

$$e_x = \sqrt{\frac{\sum d_{x,i}^2}{n}} \quad (13)$$

$$e_y = \sqrt{\frac{\sum d_{y,i}^2}{n}} \quad (14)$$

The pixel errors e_x and e_y can be converted into angle errors.

4. Experimental results

The approach for visual registration, as described in section 3, has been implemented in a development setup for later integration into an outdoor AR system, providing visual aid for navigation and hazard indication. Part of this AR system has been developed for real-time AR using visual servoing [30] as a means of visual registration in an indoor scenario. This system is currently being ported to be operated in an outdoor vehicle, using the registration approach described in this paper, as well as other sensors (inertial, inclinometer, GPS). The goal is to provide real-time registration during camera and observer motion.

4.1. System setup

For the development stage, the system for registration was implemented on a 200 MHz Pentium Pro, running under Windows NT 4.0. The images have been grabbed from a Cohu 2200 CCD camera (1/2“ chip), using an image size of 512×480 pixels. The system is equipped with an Imagenation framegrabber PXC200. For testing, images from a digital still camera (Minolta RD-175) have also been used. The “conventional” registration components include a Cybertrack CT-3.2 magnetometer/inclinometer which is used to obtain magnetic northing, roll, and pitch angle. The software is written in C++, using Microsoft Foundation Classes. The computation time for the various task within the registration system are given in the next sections.

4.2. Silhouette from DEM map

The complete 360 degree silhouette (Figure 6) as seen from the Rockwell Science Center has been computed from the USGS DEM “East L.A.”. The data set contains 9.6 MB of data (1201×1201 data points). The computation of the silhouette, using the complete set, took 40 sec. However, since the map covers an area of 92×111 km, a smaller subset suffices for obtaining the silhouette (range ± 15 km). This reduces the computation time to about 4 sec, which is still too long for real-time performance.

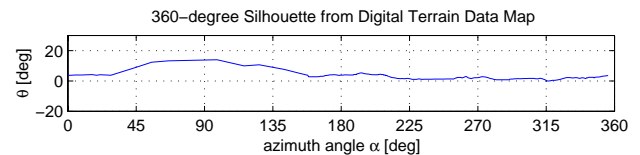


Figure 6. Isoscale plot of DEM silhouette.

The silhouette was computed in azimuth increments of $\Delta \alpha = 0.25^\circ$. The most significant feature is the big hump between 45° and 135° , which comes from a nearby hill behind RSC. However, due to its proximity, it consists only of

a few data points and, therefore is not well-suited to provide a visual cue for registration.

In a non-isoscale plot, more details of the silhouette are revealed (Figure 7). The segment shows mountains which are further away (3-5 km). As a result, the silhouette point array is much denser. Since the location point is not far from the edge of the DEM map, an additional map was used to provide more data for silhouette prediction. This can be seen in Figure 7, where the dotted silhouette segment was computed only from the current map and falsely indicated a deep canyon as part of the horizon silhouette.

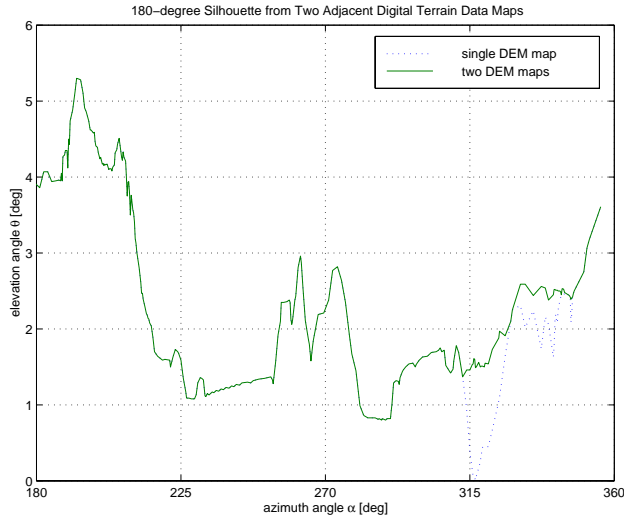


Figure 7. Silhouette computed from two adjacent DEM maps. The elevation axis is stretched to show the effect of the second map.

4.3. Visual silhouette extraction

The video image in Figure 8 was captured by the Cohu camera and was stored as b/w image in 512×480 resolution. Top and bottom image portions have been truncated in Figure 8. The blue Southern California sky made it easy for the system to detect the border line between sky and terrain area. The silhouette extracted from this image is shown in Figure 9.



Figure 8. Video image taken from the Rockwell Science Center, Thousand Oaks (CA).

It should be noted that the approach for visual silhouette extraction was not optimized for efficiency and took around 1 sec. This is because the search paths were not limited by assumptions, and the whole image was evaluated.

4.4. Silhouette matching

In the DEM silhouette there are 60 maxima and 60 minima. Combining these with the 10 maxima and 10 minima of the video silhouette, there are 324000 theoretical combinations for a correspondence hypothesis. After applying the constraints on the selection of possible combinations, only 75 combinations remain to be investigated.

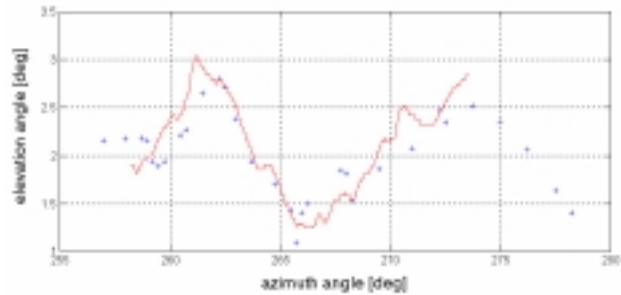


Figure 9. Best match of silhouettes. The contiguous line is the video silhouette, the dots are DEM silhouette points.

The computation of the error criterion was done in less than 0.5 sec. In Table 1 the numerical results of the matching algorithm can be seen. The column denoted “corrected” shows the values after the corrective terms (section 3.3) have been calculated. It should be denoted, that these orientation angles were obtained without using the data of the magnetometer/inclinometer.

	hypothesis	corrected	error
α_c	265.9°	265.9°	±0.3°
θ_c	1.54°	1.62°	±0.2°
C_x	33.1 pix/°	33.3 pix/°	
C_y	21.3 pix/°	21.2 pix/°	

Table 1. Matching results.

4.5. AR overlay

We have begun to develop an outdoor AR system which provides an overlay onto the grabbed video images and later in the user’s field of view. In Fig. 10, an example of a 2D overlay is shown, visualizing a vertical axis, an artificial horizon, ground plane visualization, and pitch and azimuth angles as well as compass directions. It also shows textual information from a geographical information system (GIS) indicating the name of landmarks and the distance to them.

The silhouette, which was computed from the DEM data, is overlaid onto the video image in order to give a reference for the registration quality.

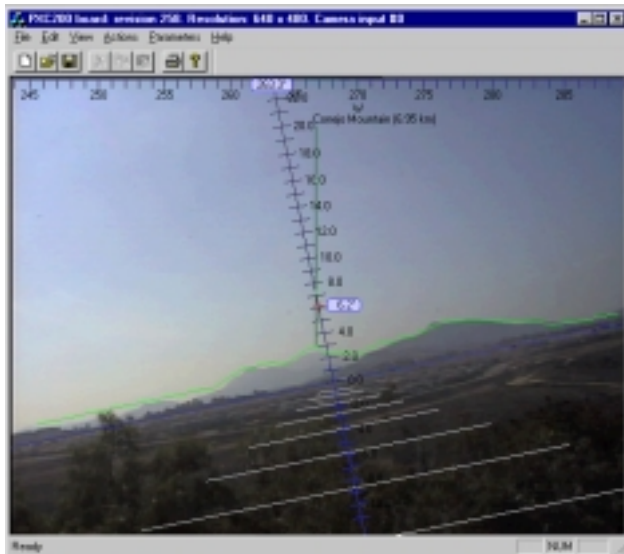


Figure 10. 2D overlay onto video image.

Since the visual calibration has been computed off-line, the AR system is able to run with a framerate of 10 Hz. The angular data (pitch, yaw, roll) are obtained by the CT-3.2, after it has been calibrated with the above described visual silhouette method. This offline calibration could be used at a later time when the mountains in the background were partially disappearing in fog (Figure 11).

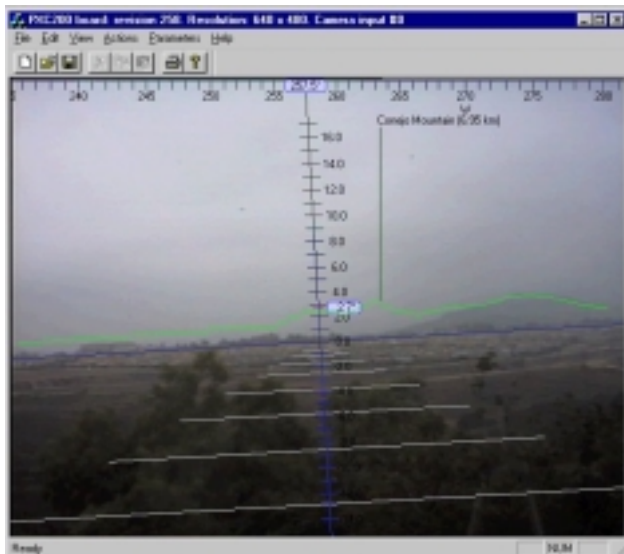


Figure 11. Silhouette overlay during fog.

It is evident here, that the visual registration algorithm could not produce the registration required for this overlay. However, the previously recorded calibration parameter were used to correct the data from the CT-32 and provide a

true augmentation of the video stream. The current overlay is in 2D style only, but we have designed the system to provide 3D rendering by the Sense8 World Toolkit libraries.

4.6. Further development

The registration approach as presented in this paper will be further developed to perform in real-time. This will include optimizations of the algorithms as well as a change of mode after successful initialization to a tracking phase, using the visual servoing approach [30]. This will provide full 6-DOF real-time registration for outdoor AR during observer motion. The complete AR system is currently being ported to a dual Pentium system (333 MHz), which will be installed in a test van for mobile outdoor testing. The 3D overlay will consist of nothing information and object visualization from a remote database. For fast 3D rendering of overlay objects, a Symmetric graphics adapter with GLINT Open GL accelerator is already used as video output. Simple overlay functionality is provided by the standard Windows API functions, and more sophisticated 3D visualization is managed by the Sense8 World Toolkit.

5. Summary and Conclusion

We have demonstrated that in a well-structured terrain and under good lighting conditions, horizon silhouettes can provide visual clues to improve registration precision significantly. Our approach provides means for obtaining camera azimuth and elevation by using a non-calibrated camera, and even provides the calibration parameters. Further work needs to be done to make this approach also applicable under a variety of lighting and visibility conditions. Principal limitations of computer vision in outdoor scenarios, however, require that a robust registration system include other sensors and is built as a hybrid system. A registration system relying solely on silhouettes, would fail in flat, unstructured terrain or under conditions of reduced visibility. Using other wavelength (IR) may improve visibility under certain conditions, but the problem of feature obstructions has to be addressed.

References

- [1] K. Ahlers, A. Kramer, D. Breen, P.-Y. Chevalier, C. Crampton, E. Rose, M. Tuceryan, R. Whitaker, and D. Greer. Distributed Augmented Reality for collaborative design applications. In *Proc. Eurographics '95*, pages C13–C14, Maasticht, The Netherlands, 1995.
- [2] R. T. Azuma. A survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385, 1997.
- [3] B. B. Bederson. Audio augmented reality: a prototype automated tour guide. In *Proc. of Conf. on Human Factors in Computing Systems (CHI)*, pages 210–211, Denver, CO, May 1995.
- [4] R. Behringer and N. Ahuja. Image registration for computer vision-based Augmented Reality. In *Proc. of 2nd Annual FedLab Symp. on Advanced Displays and Interactive Displays*, pages 44–48, College Park, MD, Feb. 1998.

- [5] J. W. Berger and D. S. Shin. Computer-vision-enabled ophthalmic Augmented Reality: A PC-based prototype. In *Proc. of First Int. Workshop on Augmented Reality (IWAR)* '98, San Francisco, CA, 1998.
- [6] P. Brodeur, J. Dansereau, J. de Guise, and H. Labelle. Points-to-surfaces matching technique for the application of Augmented Reality during spine surgery. In *Proc. of IEEE Conf. on Engineering in Medicine and Biology*, pages 1197–1198, Montreal, Canada, Sept. 1995.
- [7] T. P. Caudell. Introduction to Augmented and Virtual Reality. In *Proc. of SPIE Conf. on Telemanipulator and Telepresence Technologies*, pages 272–281, Boston, MA, Oct. 1994.
- [8] Y. Cho, J. Lee, and U. Neumann. A multi-ring color fiducial system and a rule-based detection method for scalable fiducial-tracking Augmented Reality. In *Proc. of First Int. Workshop on Augmented Reality (IWAR)* '98, San Francisco, CA, 1998.
- [9] M. Cohen, S. Aoki, and N. Koizumi. Augmented audio reality: Telepresence/VR hybrid acoustic environments. In *Proc. of Workshop on Robot and Human Communication*, pages 361–4, Tokyo, Japan, Nov. 1993.
- [10] N. R. Corby and C. A. Nafis. Augmented Reality telemanipulation system for nuclear reactor inspection. In *Proc. of SPIE Conf. on Telemanipulator and Telepresence Technologies*, pages 360–365, Boston, MA, Oct. 1994.
- [11] J. Cote and J. Lavallee. Augmented Reality graphic interface for upstream dam inspection. In *Proc. of SPIE Conf. on Telemanipulator and Telepresence Technologies II*, pages 33–39, Philadelphia, PA, Oct. 1995.
- [12] S. Feiner, B. MacIntyre, T. Höllerer, and A. Webster. A touring machine: prototyping 3D mobile Augmented Reality systems for exploring the urban environment. In *Proc. of 1st Int. Symp. on Wearable Computers*, pages 74–81, Cambridge, MA, Oct. 1997.
- [13] W. E. L. Grimson, G. J. Ettinger, S. J. White, T. Lozano-Perez, W. M. W. III, and R. Kikinis. An automatic registration method for frameless stereotaxy, image guided surgery, and enhanced reality visualization. *IEEE Transactions on Medical Imaging*, 15(2):129–140, Apr. 1996.
- [14] T. Joshi, N. Ahuja, and J. Ponce. Structure and motion estimation from dynamic silhouettes under perspective projection. In *Proc. of ICCV '95*, 1995.
- [15] D. Koller, G. Klinker, E. Rose, D. Breen, R. Whitaker, and M. Tuceryan. Real-time vision-based camera tracking for Augmented Reality applications. In *Proc. of VRST '97*, Lausanne, Switzerland, Sept. 1997.
- [16] W. Mackay, D. Pagani, L. Faber, B. Inwood, P. Launiainen, L. Brenta, and V. Pouzol. ARIEL: Augmenting paper engineering drawings. In *Proc. of Conf. on Human Factors in Computing Systems (CHI)*, pages 421–422, Denver, CO, May 1995.
- [17] P. Milgram. Augmented Reality for telerobotic control: A practical approach for today's bandwidth. *Advanced Imaging*, 11(5), May 1996.
- [18] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino. Augmented Reality: a class of displays on the reality-virtuality continuum. In *Proc. of SPIE Conf. on Telemanipulator and Telepresence Technologies*, pages 282–292, Boston, MA, Oct. 1994.
- [19] D. Mizell. Virtual reality and Augmented Reality in aircraft design and manufacturing. In *Proc. of Wescon Conference*, page 91ff, Anaheim, CA, Sept. 1994.
- [20] J. Molineros, V. Raghavan, and R. Sharma. AREAS: Augmented Reality for evaluating assembly sequences. In *Proc. of First Int. Workshop on Augmented Reality (IWAR)* '98, San Francisco, CA, 1998.
- [21] E. D. Mynatt, M. Back, R. Want, and R. Frederick. Audio Aura: light-weight audio Augmented Reality. In *Proc. of ACM UIST '97*, pages 211–12, Banff, Canada, Oct. 1997.
- [22] NIMA. National Imagery and Mapping Agency web site. <http://www.nima.mil/>.
- [23] A. Park and R. N. Katzman. Augmented Reality for mining teleoperation. In *Proc. of SPIE Conf. on Telemanipulator and Telepresence Technologies*, pages 119–129, Boston, MA, Oct. 1994.
- [24] J. Park, S. You, and U. Neumann. Natural feature tracking for extendible robust augmented realities. In *Proc. of First Int. Workshop on Augmented Reality (IWAR)* '98, San Francisco, CA, 1998.
- [25] M. Rauterberg, T. Mauch, and R. Stebler. Digital playing desk: A case study for Augmented Reality. In *Proc. of IEEE Workshop on Robot and Human Communication*, pages 410–415, Tsukuba, Japan, Nov. 1996.
- [26] D. Reiners, D. Stricker, G. Klinker, and S. Müller. Augmented Reality for construction tasks: Doorlock assembly. In *Proc. of First Int. Workshop on Augmented Reality (IWAR)* '98, San Francisco, CA, 1998.
- [27] A. State, M. A. Livingston, W. F. Garrett, G. Hirota, M. C. Whitton, E. D. Pisano, and H. Fuchs. Technologies for Augmented Reality systems: Realizing ultrasound-guided needle biopsies. In *Proc. of SIGGRAPH*, New Orleans, LA, 1996.
- [28] F. J. Stein. *Structural Indexing for Object Recognition*. PhD dissertation, University of Southern California, Apr. 1992.
- [29] D. Stricker, G. Klinker, and D. Reiners. A fast and robust line-based optical tracker for Augmented Reality applications. In *Proc. of First Int. Workshop on Augmented Reality (IWAR)* '98, San Francisco, CA, 1998.
- [30] V. Sundareswaran and R. Behringer. Visual servoing-based Augmented Reality. In *Proc. of First Int. Workshop on Augmented Reality (IWAR)* '98, San Francisco, CA, Nov. 1998.
- [31] I. E. Sutherland. A head-mounted three dimensional display. In *Proc. of Fall Joint Computer Conference*, pages 757–764, Washington, DC, 1968.
- [32] Z. Szalavari and M. Gervautz. The personal interaction panel – a two-handed interface for Augmented Reality. In *Proc. of EUROGRAPHICS '97*, Budapest, Hungary, Sept. 1997.
- [33] R. Talluri and J. K. Aggarwal. Position estimation for an autonomous mobile robot in an outdoor environment. *IEEE Journal of Robotics and Automation*, 8(5):573–584, Oct. 1992.
- [34] W. B. Thompson, H. L. Pick, B. H. Bennet, M. R. Heinrichs, S. L. Savitt, and K. Smith. Map-based localization: The drop-off problem. In *Proc. of DARPA Image Understanding Workshop*, Pittsburgh, PA, Sept. 1990.
- [35] M. Uenohara and T. Kanade. Vision-based object registration for real-time image overlay. In *Proc. of 1st Int. Conf. on Computer Vision, Virtual Reality, and Robotics in Medicine*, Nice, France, Apr. 1995.
- [36] USGS. *Digital Elevation Models - Data Users Guide*. United States Department of the Interior - US Geological Survey, Reston, VA.
- [37] USGS. United States Geological Survey web site. <http://www-nmd.usgs.gov/>.
- [38] B. Vijayakumar. *Recognition and Pose Estimation of 3D Curved Objects with Models Built Using Active Vision*. PhD dissertation, Yale University, May 1996.
- [39] A. Webster, S. Feiner, B. MacIntyre, W. Massie, and T. Krueger. Augmented Reality in architectural construction, inspection, and renovation. In *Computing in Civil Engineering*, pages 913–919, New York, NY, 1996.