A Multi-ring Color Fiducial System and An Intensity-invariant Detection Method for Scalable Fiducial-Tracking Augmented Reality

Youngkwan Cho Jongweon Lee Ulrich Neumann
{ykcho | jonlee | uneumann}@graphics.usc.edu
http://tracker.usc.edu/~ykcho/star.html
Computer Science Department
Integrated Media Systems Center
University of Southern California, USA

Abstract

In Augmented Reality (AR), a user can see a virtual world as well as a real world. To avoid the registration problem between the virtual world and the real world, the user's pose in both worlds should be exactly the same. Fiducial tracking AR is an attractive approach to the registration problem, but most of the developed fiducial tracking AR systems have very limited tracking ranges and require carefully prepared environments, especially lighting conditions. To provide for wide views and detailed views in large-scale applications, an AR system should have a scalable tracking capability under varying light condition.

In this paper, we propose multi-ring color fiducial systems and a light-invariant fiducial detection method for scalable fiducial tracking AR systems. We analyze the optimal ring width, and develop formulas to obtain the optimal fiducial set with system specific inputs. We present a light-invariant circular fiducial detection method that uses relations among fiducials and their backgrounds for segmenting regions of an image. Our work provides a simple and convenient way to achieve wide-area tracking for AR.

Keywords

Augmented Reality, concentric fiducial, edge detection, fiducial-tracking, multi-ring fiducial, region segmentation, rule-based detection

1 Introduction

In Virtual Reality, where all scenes are computer-generated images, a virtual world could be explored by flying or steering treadmills without the same physical movements. In Augmented Reality (AR), a user sees a virtual world as well as a real world. To avoid the misalignment between two worlds, the user's pose in the real world is directly related to the user's pose in the virtual world and AR requires the same amount of movement in both worlds. To apply AR to large-scale applications, wide-area tracking is essential.

An overlapped virtual world contains virtual objects to help understand a real world [Azuma95]. To make the virtual world useful, the virtual objects should

be aligned properly with the real objects. Registration is one of the major issues in AR [Azuma94][Bajura95][Tuceryan95]. The registration problem requires high accuracy or error-correction mechanisms in tracking. Fiducial tracking has been gaining interest as a solution to the registration problem [Mellor95] [Neumann96][State96][Cho97].

1.1 Motivation

The tracking range of a fiducial tracking AR system is confined by the detectability of fiducials in input images. Most of the developed systems use their own single-size fiducials. A single-size fiducial might help fast fiducial detection, but the system will have a narrow tracking range because all fiducials have the same detection range.

In a large-scale application with details, an AR system should provide wide views as well as detailed (zoom in) views of interesting regions. The developed AR systems, which use single-size fiducials, do not seem to provide such scalability. Different size fiducials will have different detection ranges, and by combining a series of different detection ranges from multi-size fiducials, a whole tracking range can be extended seamlessly.

Fiducial tracking AR systems detect fiducials using image processing or computer vision techniques: boundary detection, color segmentation, watershed, and template matching techniques. These methods require thresholds to segment an image and to detect fiducials, and often can not be used under different lighting conditions without changing thresholds. Since most of the developed fiducial tracking AR systems adopt these detection methods, they seem to require carefully prepared environments including light conditions, and their usability are very limited. To apply AR to a large-scale application, we need a detection method that works under any lighting conditions without manually controlling any parameters, including uneven lighting condition in one image.

1.2 Contribution

We add the multi-ring multi-size concept to the concentric circular fiducials, and introduce a fiducial system concept. This fiducial system concept provides great scalability to fiducial tracking AR. The fiducial systems introduce a much larger number of unique fiducials than a single-size fiducial system, and make fiducial identification much easier. We develop formulas to calculate the optimal fiducial set for any size of applications with some system specific parameters. Users conveniently get the optimal fiducial set by just plugging in those parameters according to their systems.

We develop a robust fiducial detection method under varying lighting conditions, using light-invariant relationships among homogeneous regions instead of thresholds for segmenting regions. We also develop rules and membership functions to detect fiducials.

We present a simple and low-cost way to achieve wide-area tracking, and we hope that this work trigger many research activities in large-scale applications.

2 Previous works

2.1 Fiducial tracking

There are several developed fiducial tracking AR systems. Most of them have used solid or concentric circular fiducials [Mellor95][Neumann96] [State96][Cho97], and a few of them have used the corners of big rectangles as fiducials [Kutulakos96][Koller97]. All of them have adopted single-size fiducials or rectangles, and their tracking areas are limited to just around a desktop or the opposite wall of a room (Fig. 1). None of them seems to support full room (~30x30 feet) tracking from arm-length distance to wall-to-wall distance.

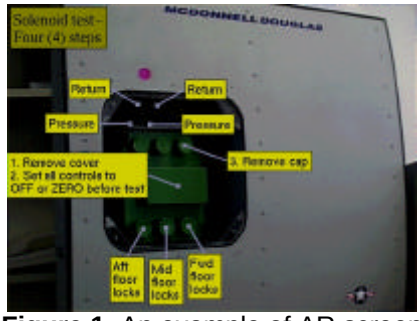


Figure 1. An example of AR screen

2.2 Fiducial detection

Several AR systems have been developed using vision-based techniques. Some authors mentioned their approaches that deal with different lighting conditions while others do not give any information relating the environment restrictions under which their systems work.

Madritsch, Leberl, and Gervautz [Madritsch96] developed a camera-based beacon tracking system. They use red LEDs as beacons and accept lighting restrictions for their system.

State, Hirota, Chen, Garrett, and Livingston [State96] developed a hybrid system which combined a vision-based tracking system and a magnetic tracking system. The basic algorithm to detect fiducials is based on the ratios of RGB component values. They used the system under restricted lighting conditions, and mentioned that the fiducial analyzer's performance diminished with changing lighting conditions despite the use of adaptive brightness evaluation for each landmark.

Uenohara and Kanade [Uenohara95] mentioned the appearance of objects varies in many ways depending on pose and illumination change. They require users to locate objects at the initial recognition step. Normalized correlation and precaptured images are used to operate the system under varying lighting conditions. The complexity of the detection algorithm increases as objects become more complex and more reference images are required.

As we see from the above three systems, vision-based AR systems need a fiducial detection algorithm, and the performance of detection algorithms diminishes when lighting conditions vary. We develop a rule-based algorithm to detect fiducials with varying lighting conditions. The idea behind the algorithm is to use relations among homogeneous regions instead of using threshold values for segmenting regions.

3 Multi-ring color fiducial system

The 2D image location (u,v) of a 3D point (x,y,z) is determined by the following perspective projection.

$$\begin{bmatrix} U \\ V \\ w \end{bmatrix} = \begin{bmatrix} f_u & 0 & U_0 & 0 \\ 0 & f_v & V_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{T} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} U / w \\ V / w \end{bmatrix}$$

where f_u and f_v are the effective focal lengths of a camera in the u and v direction, respectively, and (U_0,V_0) is the image center. **R** is the rotation matrix and **T** is the translation vector for the camera pose. The major axis length d of the ellipse is

$$d = Df / w,$$

where D is the diameter of the fiducial, f $(= f_u = f_v)$ is the effective focal length, and w is the depth of the fiducial in the camera coordinate system.

When a camera is too far or too close to a fiducial, the projected image of the fiducial in the input image is too small or too large to detect it correctly. Therefore, an AR system with single-size fiducials has a very limited tracking range. Although each fiducial has a fixed detectable range, the whole tracking range could be extended by combining different detectable ranges of different size fiducials.

Multi-ring color fiducials have different number of rings at different size levels. The first level fiducial has one core circle and one outer ring. As the level goes up, one extra ring is added outside of the previous level fiducial. The number of rings in a fiducial tells the fiducial level where it belongs. The core circle and rings are painted with six colors (red, green, blue, yellow, magenta, and cyan). It introduces many unique fiducials and, it makes fiducial identification easier.

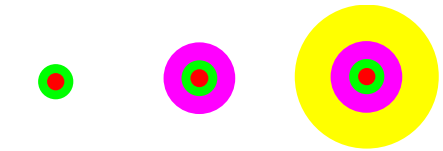


Figure 2. Multi-ring multi-size color fiducials

In the proportional width ring fiducial system, the size ratio c between adjacent levels is the same.

$$D_i = cD_{i-1} \quad (c > 1)$$
$$= c^{i-1} D_1$$

where D_i is the diameter of the level i fiducial. Because the outer ring is wider than the inner ring, the outer ring is more easily detectable from a distance. The higher level fiducials have a farther detectable range. By combining those detectable ranges, we can extend the whole tracking range. This fiducial system is good for wide-area tracking with arbitrary camera movements.

Let the desired tracking range be $Z_{near} \sim Z_{far}$, and the camera focal length be f. Let w be the minimum detectable ring width in an input image. w depends on the camera, the digitizer, and fiducial detection algorithm.

3.1 Number of required levels

Let the tracking range of a level i fiducial be $Z_{near,i} \sim Z_{far,i}$ with the conditions $Z_{near} = Z_{near,i}$ and $Z_{far,n} = Z_{far}$. The largest detectable fiducial size in an image is d_{near} ($\geq D_i f / Z_{near,i}$), and the smallest detectable fiducial size is d_{far} ($\leq D_i f / Z_{far,i}$).

To combine the detectable ranges smoothly, there should be no gap between adjacent work ranges.

$$0 \leq Z_{far,i} - Z_{near,i+1} \leq \frac{D_i f}{d_{near}} \left(\frac{d_{near}}{d_{far}} - c \right)$$

$$c \leq \frac{d_{near}}{d_{far}} \quad and \quad c \leq \frac{Z_{far,i}}{Z_{near,i}}$$

The required levels of fiducials can be expressed as a function of the size ratio c (Fig. 3).

$$D_n = \frac{Z_{far}}{f} d_{far}, \; D_1 = \frac{Z_{near}}{f} d_{near}$$

$$\frac{Z_{far}}{Z_{near}} = c^{n-1} \frac{d_{near}}{d_{far}} \geq c^n$$

$$n(c) \leq \frac{\log(Z_{far} / Z_{near})}{\log c}$$

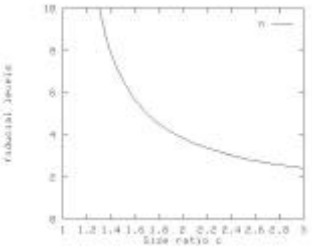


Figure 3. Fiducial levels

3.2 Fiducial size at each level

When the camera is at $Z_{far,1}$ from a level n fiducial, the major axis length of the level *1*, *2*, and *j* rings in an input image are

$$d_{far,1} = \frac{2w}{c\text{-}1} c \geq d_{far}$$

$$d_2 = \frac{2w}{c\text{-}1} c^2 \leq d_{near,2} \leq d_{near}$$

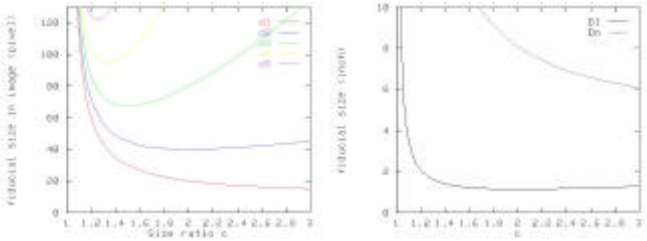
$$d_j = \frac{2w}{c\text{-}1} c^j, (1 \leq j \leq n)$$

The diameters of the level *1* and *n* fiducials are

$$D_n = \frac{Z_{far}}{f} d_{far} = \frac{Z_{far}}{f} \frac{2w}{c-1} c$$

$$D_1 = \frac{Z_{near}}{f} d_{near} \geq \frac{Z_{far}}{f} \frac{2w}{c-1} c^{2-n}$$

Figure 4 shows the major axis lengths of some fiducial levels, and the minimum and maximum fiducial sizes as a function of *c*.



(a) Fiducial sizes in an image (b) Real fiducial sizes

Figure 4. Fiducial sizes

3.3 Fiducial distribution

When a camera is close to fiducials, the camera can see only a small region in the real world, and only small size fiducials can be detected. When the camera is far from fiducials, it can see a large area, and only large size fiducials can be detected. Therefore, lower level fiducials have dense distributions and higher level fiducials have sparse distributions.

Fiducials are usually distributed around interesting regions without any regular pattern. For analysis purposes, we use a regular grid distribution. We assume that the input image has WxH resolution ($W \geq H$).

To determine the camera pose, three or more non-collinear fiducials are required in an input image [Linnainmaa88][Horaud89][Haralick94]. Figure 5(a) shows an input image of level *i* fiducials at distance $Z_{near,i}$. With this configuration, any camera pose can see three or more fiducials in the valid tracking range of level *i*. The inter-fiducial distance for level *i* fiducials L_i is

$$L_i(c) = \frac{H - d_{near}}{2} \cdot \frac{D_i}{d_{near}}$$

$$= \frac{H(c\text{-}1) - 2c^2 w}{4c^2 w} D_i$$

(a) Input image of the level i fiducials at distance $Z_{near,i}$ (b) fiducial density f(c)

Figure 5. Fiducial distribution

The ratio of the fiducial diameter over the inter-fiducial distance shows the fiducial density.

$$f(c) = \frac{D_i}{L_i} = \frac{4c^2 w}{H(c-1) - 2c^2 w}$$

Figure 5(b) shows the inter-fiducial distances and fiducial density in percentage form.

3.4 Optimal size ratio *c*

All results come out as functions of the size ratio *c*. Which *c* gives the optimal results? We can think of the optimality in two aspects: system performance and fiducial sizes.

For high system performance, fiducial detection process should be fast. We can concentrate on the performance of finding the smallest detectable ring in a fiducial, because the larger rings could be found easily by predicting their boundaries with *c*. The fiducial detection algorithm should look for rings whose diameters are in the range from $d_{far}(=\frac{2w}{c-1}c)$ to $d_{near}(\geq \frac{2w}{c-1}c^2)$. To minimize the processing time for fiducial detection, $f(c)=(d_{near} - d_{far})d_{far}$ should be minimized. $(d_{near} - d_{far})$ is the search range, and d_{far} is the search size. The bigger d_{far} is, the more processing time for boundary detection.

$$f(c) = d_{far}(d_{near} - d_{far}) = \frac{4w^2}{c-1}c^2$$

$$\frac{df(c)}{dc} = \frac{4w^2 c(2-c)}{(c-1)^2}$$

$f(c)=(d_{near} - d_{far})d_{far}$ has minimum when $c = 2$.

The other aspect is fiducial size. If fiducials are too big, it is not easy to find proper places to put them on. They also occupy a large area in the input images. So, we can get the optimality by minimizing the fiducial density.

$$\frac{df(c)}{dc} = \frac{4cwH(c-2)}{(H(c-1) - 2c^2 w)^2}$$

$f(c)$ has the minimum $16w/(H - 8w)$ when $c=2$.

From both aspects, the optimal size ratio *c* is the same value 2.

4 Fiducial detection

We are only interesting in fiducials in an image, which occupy very small areas of an image, so we apply a multi-resolution approach to reduce execution time of our detection algorithm. The detection algorithm is divided into two stages, coarse detection and finer detection (Fig. 6). Coarse detection quickly skims through an image and finds potential regions, and more expensive method, finer detection, is applied to detect possible fiducials. Then shape and color tests are applied to distinguish false fiducials from real fiducials. Coarse and finer detection methods share two steps, assigning membership values and creating line segments. They differ in two ways. One is the coarse detection finds potential regions using only sampled horizontal and vertical scan lines, but the finer detection uses all horizontal and vertical scan lines in small selected regions to detect possible fiducials. Second difference is the coarse detection clusters cross points created with horizontal and vertical line segments, but the finer detection clusters line segments segmented along horizontal and vertical scan lines. Clustering used at coarse and finer detection are done based on locations and colors of cross points or line segments. The bases of the algorithm are rules and fuzzy algorithms. Rules are used for detecting transition areas between a fiducial and its background, and fuzzy algorithms are applied to localize the position of edges. Detail of the algorithm is presented at following

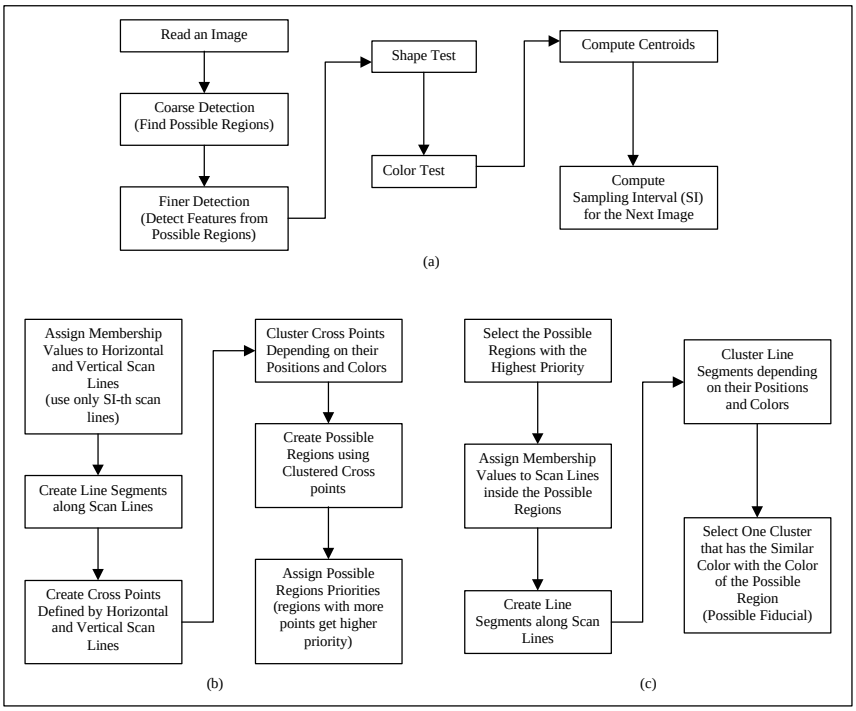


Figure 6. (a) Overview of the algorithm (b) coarse detection (c) finer detection

sections.

4.1 Coarse detection

Potential regions are detected by the coarse detection procedure. The coarse detection procedure utilizes that cross points are created by horizontal and vertical line segments inside a fiducial if a sampling interval is carefully designed. Clusters of cross points create potential regions for a finer detection procedure.

$$s = \frac{d}{\sqrt{2}}, \quad where\ d\ is\ the\ size\ of\ minor\ axis\ of\ the\ smallest\ fiducial$$

$$Sampling\ Interval(SI) = \frac{s}{2}$$

The sampling interval is the most important parameter used in the coarse detection procedure. Small sampling interval increases computation time, and large sampling interval may miss few fiducials. We create a potential region with a cluster that contains at least one cross point, but we use sampling intervals that create at least four cross points for each fiducial without any noise on an image. This improves robustness even if it increases computation time. Intersections among sampled horizontal and vertical lines form square grids, and one square grid contains four cross points. If grids are contained within the inner square (d equals the size of minor axis) of the expected minimum fiducial, there are at least four cross points inside every fiducials (Fig. 7). We can derive the size of the inner square of the smallest fiducial of an image from the minor axis of the smallest fiducial. Then the optimal interval can be computed using a following equation. A sampling interval is adaptively changed according to the size of minor axis of the smallest fiducial of the previous image.

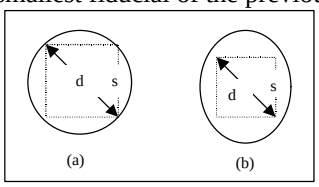


Figure 7. Maximum Sampling Interval (a) circle (b) ellipse

For every sampled horizontal and vertical lines, we assign a membership value to every pixel of them. All lines are segmented according to membership values of individual pixels, and these line segments create cross points if two line segments intersect and have the same color. We cluster cross points depending on their positions and colors, and each cluster defines one potential region for a finer detection procedure. A priority is assigned to each potential region according to the number of cross points in a cluster after eliminating clusters with a larger size than the maximum fiducial size.

4.2 Finer detection

Select the region with the highest priority from the possible regions. We assign a membership value to every pixel of all horizontal and vertical lines inside the selected region, and create line segments along all scan lines based on membership values. These line segments are clustered according to their locations and color, and clusters of line segments are possible fiducials. Sub-procedures of both detection steps are presented at following sections.

4.3 Rules and membership functions

Edge detection techniques using a fuzzy logic method usually use S and π functions as a membership function. These techniques require threshold values and prior knowledge about an input image such as a number of regions on the image to segment regions of an image so we can not apply general membership functions to achieve our goal. We develop the membership function to find the best edge position inside the transition region that is found by applying rules. This is possible because we are looking for the specific fiducials.

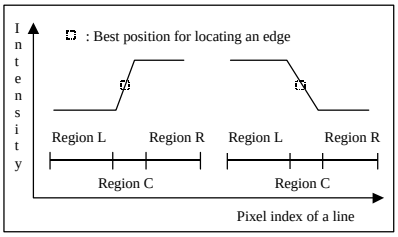


Figure 8. Step edge model

The characteristics of fiducials are the minimum distance between two fiducials, the minimum and maximum size of a fiducial, and the shape of a fiducial. We also know fiducials are placed on solid backgrounds. This means interesting edges are located between two homogeneous regions (Figure 8). We can extract few rules and membership functions from relations existing between homogeneous regions. These rules eliminate nearly all false edges in an input image, and membership functions locate the best positions for edges. The best location for an edge is a pixel location whose intensity is the same as the average value of left and right segments of an edge. Rules and membership functions are as follows.

1. Monotonic average rule
   - Avg(R) > Avg(C) > Avg(L) or Avg(L) > Avg(C) > Avg(R), where Avg(J) is an average intensity value of a region J.
2. Distribution rule
   - Max(C) – Min(C) > Max(R) – Min(R) & Max(C) – Min(C) > Max(L) – Min(L), where Max(J) and Min(J) indicate the maximum and minimum intensity values of a region J.
3. Overlapping rule
   - Max(R) < Min(L) when Avg(R) < Avg(L)
   - Max(L) < Min(R) when Avg(L) < Avg(R)

4. Membership functions

$$m = m1 \times m2$$

$$m1 = \frac{2 \times Min\left(\left|Avg\ (R) - Avg\ (C)\right|, \left|Avg\ (L) - Avg\ (C)\right|\right)}{\left|Avg\ (R) - Avg\ (L)\right|}$$

$$m2 = \frac{Min\ (R) - Max\ (L)}{Max\ (R) - Min\ (L)},\ Avg\ (R) > Avg\ (L)$$

or

$$m2 = \frac{Min\ (L) - Max\ (R)}{Max\ (L) - Min\ (R)},\ Avg\ (L) > Avg\ (R)$$

The membership function contains two parts, m_1 and m_2. m_1 indicates a grade of closeness to the median intensity value between two regions. $m_1 = 1$, if A(C) = |A(R) – A(L)| / 2, and $m_1 < 1$ for other cases. m_2 indicates the grade of closeness to the ideal edge. If an edge is the ideal edge, (Min(R) – Max(L)) = (Max(R) – Min(L)) or (Min(L) – Max(R)) = (Max(L) – Min(R)), and $m_2 = 1$. m_2 is less than 1 for other cases. Therefore the membership function m is used to find a position of an edge that is close to the ideal edge.

4.4 Detect line segments

Collect pixels from each horizontal scan line that pass all rules listed above. Group these pixels which are connected without crossing a pixel which does not pass all rules. From each group select one pixel with the highest membership value that indicates the best location of an edge. Then connect selected pixels to create line segments.

4.5 Cluster line segments

Each fiducial has a solid color, so we can find a position of a fiducial by clustering line segments with the same color. Unfortunately, however, every image contains noise. For example, a region with a solid color has pixels with different color values. This introduces a similarity measure. Traditionally similarity is measured with a distance metric and a threshold. Possible distance metrics used for a color similarity include absolute distance (e.g., Manhattan distance, Euclid distance), 1-norm distance, 2-norm distance, ∞-norm distance, angle between colors in the RGB color cube, and square of cosine of color angle.

These metrics require thresholds to decide whether two color values are similar. Defining a threshold that works for varying lighting conditions is difficult because color values are changed depending on lighting conditions. Therefore we developed a similarity measure which uses the probability theory and which utilizes local information existing on line segments. A uniform probability density function is created for each line segment, and two line segments are considered having the same color when two uniform probability density functions are overlapped. This is possible because a color similarity is

checked when two line segments are next to each other, and fiducials and their background have different colors. The distribution of a region is defined by the minimum, average, and maximum values of a region. Find D, MIN(|A(J) – Min(J)|, |A(J) – Max(J)|), and create a uniform distribution by A(J) – D and A(J) + D. Since we choose MIN(|A(J) – Min(J)|, |A(J) – Max(J)|) as D, effects of noise pixels can be eliminated, and it is used to define a uniform density function to represent an intensity distribution of a region. This simple density function works well for our detection algorithm, and it reduces the computation time of clustering procedures.

5 Result and discussion

Our implementation has the following configuration:

- SGI Indy 24-bit graphics system with MIPS4400@200MHz
- SONY DXC-151A color video camera with 640x480 resolution, 31.4° in horizontal and 24.37° in vertical FOV, S-video output
- The three level proportional fiducial set with six colors (red, green, blue, yellow, cyan, and magenta). The diameter of the first level fiducials is 1", the second level 2", and the third level 4".

The smallest detectable ring width of our implementation is 7 pixels. We search rings with 24 ~ 56 pixels in diameter. The detection range of the first level fiducial is 1.7' – 3.9', the second level 3.3' ~ 7.7', the third level 6.6' ~ 15.4'. Therefore, the whole detection range is 1.7' ~ 15.4'. Figure 9 shows three snapshots of detection results from typical distances for the three levels under uneven lighting conditions. The detected fiducials are marked with white cross hairs at the centers.

The system performance depends on the number and size of the potential

(a) Distance 3 feet - All three level fiducials are detected

(b) Distance 6 feet - The second and third level fiducials are detected

(c) Distance 12 feet – Only the third level fiducials are detected

Fiducial level	**Diameter** (inch)	**Theoretical tracking range** (feet)	**Snapshot distance** (feet)	**Frame rate** (FPS)
First level	0.8	1.5 – 3.7	3, (a)	1.0
Second level	1.6	3.0 – 7.4	6, (b)	1.5
Third level	3.2	5.9 – 14.8	12, (c)	2.0

Figure 9. Detection results. The detected fiducials have a white cross hair at the center.

fiducials in an image. The current implementation does not use any prediction of fiducial positions, but skims the whole image every time. Even a linear prediction could improve the system performance by reducing search areas for the coarse detection.

As Fig. 9(a) shows, large fiducials could be detected at close distance. Although a whole fiducial is too large in the image, one of the small rings in the fiducial might be detected. After finding one ring, we can predict and find outer rings easily, because the fiducial system uses a proportional ring width. Eventually the whole fiducial could be detected. Therefore, in the multi-ring fiducial system, large fiducials could be used at close distance. With the same reason, the partially visible large fiducials are also detected. These are advantages of the multi-ring fiducial system.

We use two lighting sources (daylight and fluorescent light), and different apertures of a camera to simulate different lighting conditions. The algorithm is tested on images with different backgrounds and apertures, f = 1.8 ~ 8.0. The backgrounds of fiducials are small regions around fiducials (Fig. 10). The algorithm is also tested on live video sequence after it is integrated into the existing AR system. Our detection method is compared with the gradient-based detection method used for the current AR system [Cho97]. The gradient-based detection method detects most fiducials correctly with f = 2.0 ~ 4.0. The presented algorithm detects all fiducials under every aperture settings except a yellow fiducials on white background at f = 1.8 and a green fiducial on black background at f = 8.0 and 5.6 (Table 1). Undetectable fiducials are not easily perceived by human eye either.

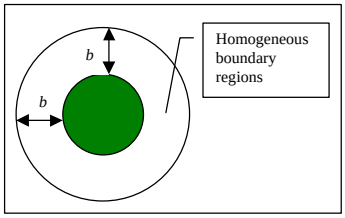


Figure 10. Defines a background of a fiducial. *b* is the size of a homogeneous boundary region considered as a background of a fiducial

This presented fiducial detection algorithm is robust with varying lighting conditions, and it is unique in that it uses rules and membership functions extracted from relations among fiducials and homogeneous backgrounds. The algorithm detects fiducials without the need for any human intervention under varying lighting conditions, including uneven lighting conditions in one image. The algorithm extends the usability of vision-based AR systems quite a bit since the AR system with the presented algorithm can be used under varying lighting conditions.

Fiducials are distributed on the interested objects and/or environments in a real world. The number of fiducials in the real world depends on an application

fiducials in the room. Because of FOV of a camera, the camera can see only a

determine the camera pose, correspondence between the real fiducials and the

Number of detected fiducials		11	12	12	12	12	11	11	12	12	12	11	11
White background	R	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	G	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	B	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Y	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
	C	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	M	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Black background	R	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	G	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	N
	B	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	C	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	M	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Aperture (f)		1.8	2.0	2.8	4.0	5.6	8.0	1.8	2.0	2.8	4.0	5.6	8.0
Lighting source		Day light						Fluorescent light					
R: Red, G: Green, B: Blue, Y: Yellow, C: Cyan, and M: Magenta													

Table 1. Results of fiducial detection algorithm.
(Y indicates detection and N indicates no detection.)

Acknowledgement

We acknowledge support by NSF grant No. CCR-9502830 and the USC Integrated Media Systems Center, an NSF ERC. We thank Anthony Majoros and McDonnell Douglas Aerospace for their invaluable assistance in defining the applications, and loan of the aircraft section model shown in Figure 1.

References

[Azuma94] Ronald Azuma and Gary Bishop, Improved Static and Dynamic Registration in an Optical See-through HMD, SIGGRAPH 1994, pp. 197-203

[Azuma95] Ronald Azuma, A Survey of Augmented Reality, SIGGRAPH 1995 Course Notes #9

[Bajura95] Michael Bajura and Ulrich Neumann, Dynamic Registration Correction in Augmented Reality Systems, VRAIS 1995, pp. 189-196

[Cho97] Youngkwan Cho, Jun Park, and Ulrich Neumann, Fast Color Fiducial Detection and Dynamic Workspace Extension in Video See-through Augmented Reality, Proceedings of the Fifth Pacific Conference on Graphics and Applications, October 1997, pp. 168-177

[Haralick94] R. Haralick, C. Lee, K. Ottenberg, and M. Nolle, Review and Analysis of Solutions of the Three Point Perspective Pose Estimation Problem, IJCV, Vol. 13, No. 3, 1994, pp. 331-356

[Horaud89] Radu Horaud, Bernard Conio, and Olivier Leboulleux, An Analytic Solution for the Perspective 4-Point Problem, CVGIP 47, 1989, pp. 33-44

[Koller97] D. Koller, G. Klinker, E. Rose, D. Breen, R. hitaker, and M. Tuceryan, Real-time Vision-Based Camera Tracking for Augmented Reality Applications, Proceedings of the Symposium on VRST, Lausanne, Switzerland, September, 1997, pp. 87-94

[Kutulakos96] K. N. Kutulakos and J. Vallino, Affine Object Representations for Calibration-free Augmented Reality, VRAIS 1996, pp. 25-36

[Linnainmaa88] Seppo Linnainmaa, David Harwood, and Larry S. Davis, Pose Determination of a Three-Dimensional Object Using Triangle Pairs, PAMI, Vol. 10, No. 5, September 1988, pp. 634-647

[Madritsch96] F. Madritsch, F. Leberl, and M. Gervautz, Camera based Beacon Tracking: Accuracy and Applications, VRST '96, pp.101-108, 1996

[Mellor95] J. P. Mellor, Enhanced Reality Visualization in a Surgical Environment, AI Technical Report 1544, 1995

[Neumann96] Ulrich Neumann and Youngkwan Cho, A Self-Tracking Augmented Reality System, VRST 1996, pp. 109-115, July 1996

[State96] Andrei State, Gentaro Hirota, Davod T. Chen, Bill Garrett, and Mark Livingston, Superior Augmented Reality Registration by Integrating Landmark Tracking and Magnetic Tracking, SIGGRAPH 1996, pp. 429-438

[Tuceryan95] M. Tuceryan, D. S. Greer, P. T. Whitaker, D. Breen, C. Crampton, E. Rose, K. H. Ahlers, Calibration Requirements and Procedures for a Monitor-Based Augmented Reality System, IEEE Transactions on Visualization and Computer Graphics, Vol. 1, No. 3, pp. 255-273, September 1995

[Uenohara95] M. Uenohara and T. Kanade, Vision-Based Object Registration for Real-Time Image Overlay, Proceedings of Computer Vision, Virtual Reality, and Robotics in Medicine '95, 13-22