What is the best fiducial?

Charles B. Owen, Fan Xiao, Paul Middlin
Media and Entertainment Technologies Lab
Media Interface and Network Design Lab
Michigan State University
3115 Engineering Building
East Lansing, MI 48824
cbowen@cse.msu.edu

Abstract—**Fiducial images are a common method for supporting vision-based tracking in augmented reality systems. This paper addresses the question: what is the best fiducial? A set of criteria that are desirable in an optically tracked fiducial are presented and a new fiducial image set is designed that meets these criteria. The images in this set utilize a square black-border pattern with a 15% border width and an interior image that supports orientation determination and unique identification. The interior image is constructed from orthogonal DCT basis images chosen to minimize the probability of misidentification and to be robust to noise and occlusion. We describe how this image can be integrated into an AR software system such as ARToolKit.**

Keywords: **Augmented reality, fiducials, motion tracking**

I. INTRODUCTION

Fiducial images, often referred to as markers, are a popular element of many of the vision-based tracking systems utilized in augmented reality applications, including the ubiquitous ARToolKit. However, images used as fiducials are often selected on an ad hoc basis. This paper addresses the question: what is the best fiducial? Should a fiducial image be round, square, or even triangular? What are the best colors to utilize? Should the image be a bar code or some other blocky pattern? Clearly, there are tradeoffs, so a set of criteria are specified in the paper that make it easier to answer that question for a wide variety of applications.

Fiducial images are images placed in a physical environment in support of tracking, alignment, and identification. Train cars have bar codes that allow machinery to automatically identify and route them through stations. Circuit boards have fiducials that allow masks to be aligned from layer to layer and allow the position of the board in a jig to be precisely measured so that robotic machinery can properly insert components. In augmented reality systems, fiducials are generally used for tracking elements in the environment. They may be placed in the fixed, physical environment so that the location of a moving camera can be identified or they may be placed on moving objects or people so that a location relative to either a fixed or moving camera can be computed. AR systems commonly rely on tracking to determine head position and orientation in support of rendering graphics registered with the surrounding environment.

Proposed fiducials have been as simple as small dots in a pattern or as complex as bar-coded circular or square images. The most relevant examples to this conference are the ARToolKit markers, square fiducial images with a fixed, black band exterior surrounding a unique image interior [1]. Figure 1 is an example ARToolKit fiducial. The outer black band allows for location of a candidate fiducial in a captured image and the interior image allows for identification of the candidate from a set of expected images. The four corners of the located fiducial allow for the unambiguous determination of the position and orientation of the fiducial relative to a calibrated camera.

Figure 1- Example ARToolKit Fiducial

Additional examples of fiducial images include the TRIP (Target Recognition using Image Processing) system, the nested multi-resolution colored ring system, and CyberCode. TRIP is a vision-based sensor using visual markers in the form of rings [2]. One ring is used as an ID code with a very large space. The other rings provide for synchronization information and support for the POSE_FROM_CIRCLE algorithm [3]. Cho, Lee, and Neumann utilize nested colored rings for fiducial images [4]. The nesting allows the rings to work over a wide range. CyberCode is a bit-based fiducial similar to a 2D bar code [5]. Many simple approaches using fixed color squares, circles, or cross patterns have been demonstrated. Most projects approach the problem either from the standpoint of selecting a set of images (as in ARToolKit) or choosing a way to encode data into images (as in CyberCode).

We have taken a constructive approach to the problem. We specify and justify a set of criteria that should be used to construct a fiducial image. This is a summary of the chosen criteria:

- An ideal fiducial image should support the unambiguous determination of position and orientation relative to a calibrated camera.
- The image should not favor some orientations over others.
- The image must be a member of a set of images that are unlikely to be confused such that a large space or set of objects can be uniquely marked.
- The image must be easy to locate and identify using fast and simple algorithms.
- Images must function over a wide camera capture range.

Some of these criteria require quantification. We have chosen to require a minimum of two hundred fiducials, which would allow marking of the ceiling in a medium-size room. We have also decided that we will require a minimum image size in the captured frame of 25 pixels across in the narrowest dimension.

Section II details the criteria we have selected for a "good" fiducial and indicates some of the conclusions we reach in that examination. Section III discusses the design of the interior image, the critical element of the system that identifies a specific fiducial. Section IV describes our conclusions and future work we have proposed.

II. Criteria for a Good Fiducial

Clearly there are tradeoffs among the criteria for a good fiducial image. This paper approaches the problem by asking several questions and proposing answers consistent with many applications in augmented reality and commonly available hardware. The questions addressed in this section are:

- What is a good fiducial shape?
- What colors should be utilized in a fiducial image?
- How should a specific fiducial be located in an image?
- How should a specific fiducial be identified?
- Over what range of sizes should the fiducial be identified?

A. Fiducial Shape

The purpose of a fiducial image is to provide automatic correspondences between points in a camera frame and points in a captured image. Clearly, any visual feature can be used as a fiducial if its location is known (or can be computed) and it can be automatically identified. Indeed, tracking systems designed for use in unprepared environments have been proposed that use regions, lines, and other natural environmental features [6, 7]. However, most applications for fiducial images assume a prepared space with specific images placed in the environment, with the assumption that the relative transformation between a camera frame and frames indicated by the fiducials needs to be determined. In tracking terminology, the position and orientation (six degrees of freedom) of the frame marked by fiducials needs to be identified relative to the camera. This problem is also commonly referred to as *pose estimation*.

Determination of position and orientation of a physical object relative to a camera frame requires the correspondence of at least four non-linear points. As an example, estimating the pose of a camera relative to a physical environment will require the identification of four 2D points in the camera image and knowledge of their 3D coordinates in the world coordinate system. It is possible to compute pose from only three points. However, the result is ambiguous, generally emitting two, and often three or four, solutions [8]. Hence, any ideal fiducial solution supporting 6DOF pose estimation should always provide a minimum of four points. Additional points can be used to compute least-square solutions that can average out errors and increase the estimate's accuracy. Many fiducial methods utilize a single, typically very simple, fiducial image such as a ring or disk with the requirement that multiple fiducials must be simultaneously tracked [4]. Since the location of fiducials in camera images will always be permuted by noise and quantization error, there is a clear advantage to tracking additional points, so fiducials that emit multiple tracking points seem advantageous. Also, many applications require tracking of styli, independent marked locations, or multiple users, where placement of a large number of fiducial images is prohibitive.

We reach the conclusion that an ideal fiducial image should emit at least four points. Beyond that, it is clear that the points should approximate a square. The size of the fiducial equates to resolution in the capture image. Four points not in the form of a square will result in some elements of the image presenting a lesser resolution to the camera than others, thereby decreasing tracking accuracy in corresponding orientations.

This does not necessarily imply that the fiducial image itself must be square. Any image that can emit four points would suffice. However, there are clear computational advantages to simplicity, and a square fiducial image is the simplest possible fiducial emitting four points. The straight edges of a square can be used to compute best-fit lines allowing corners to be computed with greater, potentially sub-pixel, accuracy. Indeed, the ARToolKit standard fiducial image is a square image.

It should be noted that a circular marker can be used to determine pose if a point on the circle can be determined. The POSE_FROM_CIRCLE algorithm provides a robust solution given circle edge points [3]. However, an interior image for identification is more difficult to implement and cannot be represented in a rectangular array. Most implementations based on post estimation from circles are based on barcodes (or, more precisely, ringcodes) [2].

B. Fiducial Color

The question of fiducial color is much more difficult to address. Clearly, choosing a color fiducial as opposed to monochrome increases the possible set of fiducial images. However, there are some clear technical reasons to favor a monochrome fiducial:

- Varying chroma resolution in camera systems
- Decreased image representation
- Higher-performance localization algorithms

The spatial frequency sensitivity of the human visual system for luminance components is much greater than for chrominance components [9]. Unfortunately, many imaging systems designed for computers mimic this characteristic, transmitting chrominance information in lower bandwidth channels or representing chrominance information with lower resolution. This necessarily decreases the detection resolution for color fiducials. Even if an RGB color presentation is captured at full resolution, the resulting color image will increase the memory usage and, consequently, the analysis time, by a factor of three (or four).

An additional element in the choice of color or monochrome is the choice of localization algorithms. High-performance algorithms have been developed for color fiducials, but assume very simple shapes that can be identified by cross-sectional lines [4]. One advantage of color fiducials is the use of color to identify the specific fiducial, as in the multi-ring approach. However, the number of colors that can be uniquely identified varies greatly depending on lighting conditions, and is likely to be small. Specular reflection can also modify the hue of imaged colors. And, the colors must contrast with colors naturally occurring in the scene.

One option for color is to utilize retro-reflective fiducials and infrared illumination [10] or direct imaging of infrared emitters [11]. We will not consider infrared approaches in this paper, as this is a very different technological approach and does not lend itself well to patterned individual fiducials.

C. Locating the Fiducial

The shape and color of a fiducial is directly related to the algorithm utilized to locate it in the camera image. The ARToolKit contains a fiducial tracking system based on work by Kato and Billinghurst [1]. Fiducials in this system are square image with a black border as illustrated in Figure 1. An interior image contained within the border provides identification for the particular fiducial image. It is assumed that the marker will contrast with a surrounding region when converted to a binary image. Typically, this can be achieved by simply ensuring the fiducial is mounted on a white surface or is printed on a larger white sheet of paper. More details of the ARToolKit approach will be included in later sections. Kato and Billinghurst allow for the fiducial corners to be rapidly and accurately located in a camera image. The approach assumes a monochrome fiducial image.

Is this the best fiducial design for localization, the location of the fiducial in an image? There are several distinct advantages to this design. The shape is a square design and yields four corner points for tracking purposes. The edges are straight between the corner points. This allows the corners to be determined by line fitting to the edges, yielding measurements that are less sensitive to noise in the vicinity of the corner and quantization errors. The black border also yields a maximum contrast relative to the background, particularly a white background. Once the corners have been located, the interior can be warped to a common frame of reference (16 by 16 in the ARToolKit approach) for comparison to a database of marker images.

This fiducial approach does not emit an orientation other than through analysis of the interior image; hence, the offset of the interior text in the marker image in Figure 1. Would it be better to design the outline to emit orientation independent of the interior text? This could be accomplished in a variety of ways, including offsetting the interior image, adding an orientation image in addition to the interior image, or using varying colors on the edge. Varying colors is not considered a good choice for the reasons mentioned above and because it would eliminate the homogeneity of the design. Detection performance would be determined by the least common denominator of detection of the two types of borders. Offsetting the image or adding an image component for orientation is equivalent to using a larger interior image and determining orientation from the interior image alone. Figure 2 illustrates this equivalence. When either the interior image is offset or a special orientation pattern is added, the fiducial can be considered equivalent to a simple border with a larger interior image, as indicated by the dotted lines.

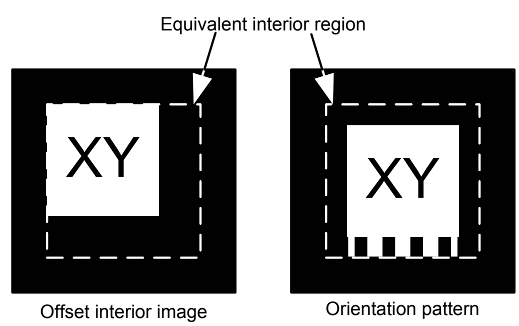

Figure 2- Equivalence of interior images for orientation determination

Given these criteria, the square ARToolKit fiducial outline seems to be a "good" approach. We will adjust the border size and the interior image in this paper, though.

D. Fiducial Identification

The final issue is the identification of an individual fiducial image. The identification of the interior image is simplified if a border has been located. The interior image can then be warped to a square image with a fixed scale.

Clearly, marking a space with identical fiducials would require the analysis of relative placement for identification, so it is advantageous if fiducials are unique. Uniqueness can be accomplished in a variety of ways, including color combinations, bar codes, or patterns. The pattern must be unique and accurately identifiable at a variety of resolutions. We have collected several desirable characteristics for fiducial identification:

- Orientation identification
- Minimal inter-fiducial correlation.
- Resistance to noise or partial obscuring.
- A large identification range.
- A large fiducial identification space.

As discussed, we assume that the fiducial is located within a fixed monochrome box that is symmetrical, as in the ARToolKit fiducials and that the orientation, and thereby the correspondence of detected image corners with physical coordinates, is determined by an interior image. Consequently, the image must support determination of a unique orientation. In ARToolKit, fiducials are commonly designed with offset text or blocks that make the orientation unique. Then, a candidate image is compared to the known images in each of the four possible orientations. This necessarily limits what can be selected as a fiducial, particularly if users desire fiducial images with visually perceptible meaning.

A key characteristic of fiducial images is that there is minimal inter-fiducial correlation in all orientations. A variety of methods are possible for comparing images. The mean squared error (MSE) is a common measure of image similarity, particularly when measuring image degradation:

$$c(I,P)=\left(\sum_x\sum_y\left(I(x,y)-P(x,y)\right)^2\right)^{1/2}$$

In this equation, I(x,y) is the candidate image, P(x,y) is the pattern, and c(I,P) is a measure of the dissimilarity between the two. For an MSE measure, small values indicate similarity. This approach is not luminance invariant, however. A better approach is the correlation coefficient. First, the mean and standard deviations for the image and pattern are computed (clearly the pattern data can be precomputed):

$$\mu_I=\frac{1}{xy}\sum_x\sum_y I(x,y)$$

$$\mu_P=\frac{1}{xy}\sum_x\sum_y P(x,y)$$

$$\sigma_I=\left(\sum_x\sum_y (I(x,y)-\mu_I)^2\right)^{1/2}$$

$$\sigma_P=\left(\sum_x\sum_y (P(x,y)-\mu_P)^2\right)^{1/2}$$

Then, the correlation coefficient is computed as:

$$\rho=\frac{\sum_x\sum_y (I(x,y)-\mu_I)(P(x,y)-\mu_P)}{\sigma_I\sigma_P}$$

This is, indeed, the comparison method utilized in ARToolKit. If the coefficient for one image is maximal for the image set and exceeds a fixed threshold (0.5), the image is accepted.

Clearly, no guarantees can be made about inter-fiducial correlations when images are chosen ad-hoc. As an example, consider the fiducial set in Figure 3. The Hiro and Kanji images (first two images in the first row) are standard fiducials included with ARToolKit. The remaining images illustrate an obvious idea of using alphabetic characters as interior images. These images were compared using the correlation coefficient. The Hiro pattern had a worst case correlation to the A pattern of 0.163. The Kanji pattern has a worst case correlation to the A pattern of 0.498, just below the standard threshold. The G pattern correlates to B with 0.637 and C with 0.820, both far above the identification threshold. Obviously, the letters have too little difference to be good choices, but even the Hiro and Kanji fiducials have correlations of 0.204.

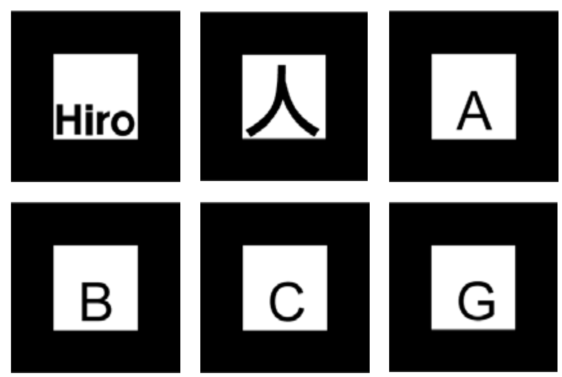

Figure 3- Example images for correlation tests

One issue related to fiducial identification is the question of human identification. Is the pattern expected to be identified by humans as well as by machine? Fiducials can very well serve that purpose, provided choices are made the have small inter-correlation values. However, this paper will concern itself with determining fiducials with low inter-correlation without respect to human identification.

Fiducial images need to be robust in the presence of noise and partial occlusion. This is a potential drawback of the ad hoc choices in Figure 3. Were a small part of the G obscured, it would be indistinguishable from the C. This is even more of an issue when bar-codes are applied to fiducials [2, 5]. The TRIP system, for example, requires 15 unique regions in the cross-section of the image center. If reduced to a size of 25 pixels across, most regions are one or two pixels and would be difficult to detect with edge detection algorithms. Small errors will change the code, violating the ringcode parity and rejecting the marker or falsely identifying it.

E. *Fiducial Identification Range*

The identification range of a fiducial depends on the camera resolution and camera parameters. Some systems have been designed to have redundant identifiers at multiple scales, so that larger images become available as the camera moves beyond the range of smaller images [4]. Clearly, the same effect can be had by creating multiple fiducials in the space of varying size and, indeed, size ratios of two are shown by Cho, Lee, and Newmann to be an effective choice.

For the purpose of this paper, we are concerned with how small an image of a fiducial may become and still be reliably recognized. This size determination is primarily dependent on the native size of the identification image. We have chosen a 16 by 16 identification image size. Therefore, the minimum dimension of the identified image in any axis must be 16 pixels. This is not the size of the actual fiducial, but rather the minimum size of the interior image.

Given this criteria, the question arises: how wide should the border be? To ensure reliable outline location, the border must be wide enough to ensure the point spread function of some pixel on the border will cover the region at every point along the border. If the border is too narrow, the border may fall between pixels, leaving the pixels a shade of gray too low to allow edge following. So, the edge must be wider than twice the distance between any two pixels. This is actually 2.83, because the worst case distance between pixels is 1.41 for diagonal lines. Hence, the image must be at least 16+2.83+2.83=21.66 pixels wide in the recognized image. For design purposes, this implies that the border must be at least 13% of the fiducial width. We have conservatively selected a 15% border width. Note that the border width should be kept minimal in order to increase the size of the interior image and allow for a larger recognition range.

F. *A Large Fiducial Identification Space*

The size of a marked space and the number of marked implements in that space is limited by the number of unique fiducials that can be applied to the space. Marking each two foot square ceiling tile in a twenty foot square room will require one hundred unique fiducials. Clearly, a desirable characteristic of fiducials is a large space of identifiers. While some of the bar code solutions claim ranges in the millions, this range is dependent on recognition of a high-resolution code in camera images from varying distances. Consequently, the images must be relatively large.

A 16 by 16 image can have up to 256 patterns that are orthogonal to each other, if the minimal correlation criterion is desired, though the set is easily expanded to 512 if maximum negative correlation is also allowed. Treating that 16 by 16 images as a 256 binary value would massively increase the number of possible fiducial images at the expense of highly correlated images.

G. *Summary of Desirable Characteristics*

This is a summary of the chosen criteria: An ideal fiducial image should support the unambiguous determination of position and orientation relative to a calibrated camera. The image should not favor some orientations over others. The image must be a member of a set of images that are unlikely to be confused such that a large space or set of objects can be uniquely marked. The image must be easy to locate and identify using fast and simple algorithms. Images must function over a wide camera capture range.

Given these criteria, we have supported the design of a square fiducial with a black border 15% of the width of the image and some internal image suitable for identification of the fiducial. The next section will detail the design of a suitable interior image.

III. A "Good" Fiducial Interior Image

A "good" fiducial interior image set will have a large set of images to choose from, a means for accurate orientation determination, and a fast algorithm for identification. The main goal is to select images such that the correlation coefficient of any two images is minimal. The optimum selection, then, would be a set of images wherein correlation coefficients among any two non-equivalent images are null. Other obvious criteria are that the image can be represented (no negative pixel values), and that the intensity be maximal (as bright as possible).

A. *Deriving the Image*

Setting the equation for the correlation coefficient to zero for two images I_1 and I_2:

$$\frac{\sum_x \sum_y (I_1(x,y)-\mu_{I_1})(I_2(x,y)-\mu_{I_2})}{\sigma_{I_1}\sigma_{I_2}} = 0$$

implies:

$$\sum_x \sum_y (I_1(x,y)-\mu_{I_1})(I_2(x,y)-\mu_{I_2}) = 0$$

This equation will be satisfied if I_1 and I_2 each is the sum of a DC offset and a member of a set of functions such that the dot product of any two non-equivalent basis functions is zero. In other words, a good choice for fiducial interior images is a set of orthogonal basis functions scaled to a peak-to-peak range equal to the pixel intensity range and added to a DC offset sufficient to make the image non-negative.

There are a wide variety of basis function sets available. Most existing 2D linear transforms, including Fourier, Hadamard, Haar, and many other, emit sets of real-valued basis images. Among the 2D sets with real values, a popular set for which a fast transform exists is the set of basis functions for the Discrete Cosine Transform (DCT), specifically that of DCT-II [12]. This N by N 2D basis function set is defined by:

$$B_{u,v}(x,y) = \cos\left(\frac{(2x+1)u\pi}{2N}\right)\cos\left(\frac{(2y+1)v\pi}{2N}\right)$$

In our application, N=16. Alternative sizes could be utilized, though larger fiducial images would be required and the recognition range would be decreased.

One approach would be to construct a fiducial interior image as:

$$I_{u,v}(x,y) = \frac{B_{u,v}(x,y)+1}{2}$$

For simplicity, we are assuming normalized pixel intensities in the range [0,1]. This interior image set supports 256 combinations of (u,v) as fiducial interior images, all of which have intercorrelation values of zero for non-equal interior images.

However, this solution does not satisfy one of our specified requirements: it does not directly include orientation information. In fact, the images with even (u,v) values can be rotated 180 degrees and yield the same image. The odd values could be utilized and do indicate orientation, but that would reduce the set of fiducial images by 75%.

Our approach to constructing the interior image is to consider the image as the sum of three parts: the DC offset value, an orientation image, and an identification image. The orientation image is the (1,0) basis image:

$$B_{1,0}(x,y) = \cos\left(\frac{(2x+1)\pi}{2N}\right)$$

This basis image is a bit less than a half cycle of a cosine wave ($\pi/32$ to $31\pi/32$).

With proper scaling, the fiducial interior image is defined as:

$$I_{u,v}(x,y) = \frac{B_{u,v}(x,y)+B_{1,0}(x,y)+2}{4}$$

When the orientation component and the mean of the image (the DC component) are subtracted, all images are orthogonal to each other, reducing the likelihood of false fiducial identification. An advantage of a DCT basis function as a fiducial image is that the pixels within the fiducial are highly correlated. This makes any correlation-based detection less sensitive to partial occlusions and noise. Whereas some fiducial systems store the information in edge data, particularly barcode-based systems, or within bounded regions as in CyberCode, the DCT basis approach embeds the identification information in the entire interior image gradient.

The interior fiducial image equation assumes the creation of a 16 by 16 image. However, the images used in a room are much larger than 16 by 16. In practice, a fiducial image will be created at a high resolution for printing, sampled by the image capture system, sampled again by the warping, and compared to the basis set. We commonly utilize 3 inch square images printed by a 600dpi laser printer. So, the analysis fiducial image (16 by 16) must be resampled to the printer resolution. The equation for creating a resampled fiducial image of arbitrary size is:

$$\widehat{I}_{u,v}(x,y) = I_{u,v}\left(\frac{xN}{W}-\frac{1}{2}, \frac{yN}{H}-\frac{1}{2}\right)$$

In this equation, x,y are coordinates in a W by H image. This equation is used to create the fiducial image at high resolution. The one half pixel offset ensures that the high resolution image will properly resample if divided into 16 by 16 square regions and sampled in the center of the region. This is important to ensure the fiducial image is not offset.

Figure 4 illustrates several example fiducial images based on this system. The lighter characteristic on the left side is due to the orientation image component. The sinusoidal patterns of the basis functions are clearly visible in the images.

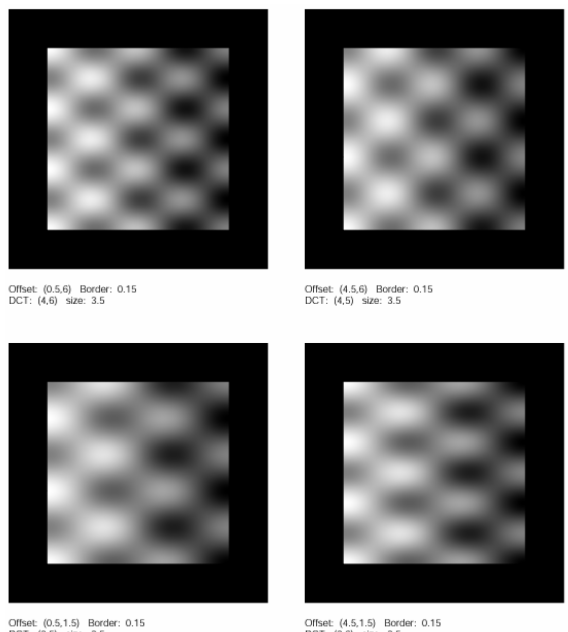

Figure 4- Example fiducial images

B. Detection

The choice of DCT basis functions as components of a fiducial image allows for fast identification using the Discrete Cosine Transform. Fast algorithms exist for the DCT, especially for the 16 by 16 size utilized in the MPEG video compression standard [12]. Computing the DCT performs a simultaneous correlation with all 256 possible basis images.

The 2D DCT-II N by N unnormalized transform is:

$$F(u,v)=\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cos\left(\frac{(2x+1)u\pi}{2N}\right)\cos\left(\frac{(2y+1)v\pi}{2N}\right)$$

The DCT result is linearly dependent on the amplitude of the input signal. This amplitude (effectively μ_I with a fixed scaling from the correlation coefficient equation) can be directly determined by examining the F(0,0) (DC) term of the DCT result. Dividing all other values by F(0,0) normalizes for intensity. (A threshold is used to reject images less than a minimum intensity).

An interesting characteristic of the DCT-II is its behavior under rotation. Let I be an original image and I' the image obtained by rotating I though 90 degrees counter clockwise. Then, letting $\widehat{I}=DCT(I)$ and $\widehat{I}'=DCT(I')$:

$$\widehat{I}(u,v)=(-1)^{v}I(v,u)$$

Applied recursively, it can be seen that the DCT of any orientation can be easily derived from the DCT of any other orientation. The orientation is indicated by the presence of the following DCT terms:

- F(1,0) positive: No rotation.
- F(0,1) negative: 90^{o} rotation.
- F(1,0) negative: 180^{o} rotation.
- F(0,1) positive: 270^{o} rotation.

Once the orientation is determined, identification of the fiducial consists of determining the cell with the maximum absolute value (other than 0,0, 1,0 and 0,1). The cell can then be trivially corrected for the image rotation by exchange of terms and/or negating the correlation result.

Note that the DCT is only performed once. The orientation is determined and then the index of the cell with the maximum absolute value in the current orientation can be translated to the index for the cell in the normal orientation.

C. *Performance and Experience*

The detection was implemented for testing and the implementation from image input to fiducial identification with point correspondence was timed on a 1.7GHz P4 with 1GB of memory. All test images used for timing were 320 by 240 pixels and had four fiducial images. Figure 5 is an example test image. The algorithm execution time was consistently under 2ms, well within the requirements of real-time tracking. It should be noted that the implementation contained an accelerated solution for outline determination that does not create intermediate images and contributed to the increased performance.

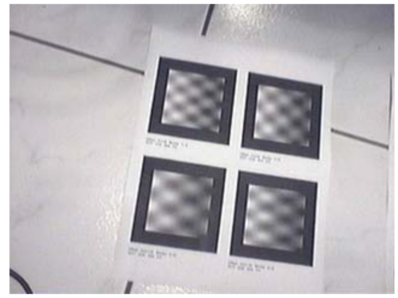

Figure 5- Example test image

Experience with the implementation has led to several interesting discoveries. It is important that the images be adequately illuminated. In under-lit conditions, motion blur can cause errors in outline determination and fiducial identification. This is evidenced by loss of acquisition. We hope to design a detailed test series in the near future that will allow us to characterize the range of conditions and exactly determine thresholds.

D. *Extra Images for Free*

Using the DCT basis set (or any 16 by 16 basis set) yields over 200 fiducials available for use. We typically avoid the lowest frequency components due to a tendency to match image content. We utilize (u,v) values of 3 or greater.

It has been assumed that an orthogonal basis set is the best choice for this application. However, the set can be supplemented using the negatives of the basis functions. Negative correlation is, in fact, as good as zero correlation in an identification system. The only correlation in the augmented set is between basis functions and their negatives. This negative correlation is easily identified in the DCT result. Adding the negative basis functions does not in any way decrease performance and effectively doubles the set size for free.

IV. Conclusion and Future Work

This paper has presented a design for fiducial images based on a black square outline containing an image constructed from DCT basis functions. The goal has been to create a set of fiducial images that is optimal relative to criteria we have set forth in the paper. The result is a fast fiducial design and associated detection algorithm that supports a large set of unique fiducials with good inter-correlation characteristics.

The choice of the DCT was a somewhat obvious starting point once the criteria were set. However, it is not clear that it is the "best" basis set. The higher frequency components are more sensitive to errors in the outline detection process and image warping and cameras have decreased high frequency content. In addition, image blurring impacts high frequency content more than low frequency content. In preliminary experiments we have constructed several custom basis image sets that may exhibit better high frequency and frequency spreading characteristics. The trade-off is the lack of a fast transform for identification. However, as the transform is separable, it may be possible to construct a custom transform that runs fast enough for this application.

V. Acknowledgements

The work was supported by grant number ITR-0082743 from the National Science Foundation and an MSU Intramural Research Grant. Acknowledgement should also be given to the students in CSE 891, Augmenting Reality, who contributed to discussions on the design of this solution.

VI. References

1. Kato, H. and M. Billinghurst. *Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System*. in *2nd IEEE and ACM International Workshop on Augmented Reality (IWAR 99)*. 1999. San Francisco, CA.
2. López de Ipiña, D. *TRIP: a Low-Cost Vision-Based Location System for Ubiquitous Computing*. in *2001 Workshop on Perceptive User Interfaces*. 2001.
3. Forsyth, D., et al., *Invariant Descriptors for 3-D Object Recognition and Pose.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991. **13**(10): p. 971-991.
4. Cho, Y., J. Lee, and U. Neumann. *Multi-ring Color Fiducial Systems and An Intensity-Invariant Detection Method for Scalable Fiducial Tracking Augmented Reality*. in *IEEE International Workshop on Augmented Reality*. 1998.
5. Rekimoto, J. and Y. Ayatsuka. *CyberCode: Designing Augmented Reality Environments with Visual Tags*. in *DARE 2000*. 2000.
6. Strickler, D., G. Klinker, and D. Reiners. *A Fast and Robust Line-Based Optical Tracker for Augmented Reality Applications*. in *International Workshop on Augmented Reality*. 1998.
7. Jiang, B. and U. Neumann. *Extendible tracking by line auto calibration*. in *IEEE and ACM International Symposium on Augmented Reality*. 2001. New York, NY.
8. Wolfe, W.J. and D. Mathis, *The Perspective View of Three Points.* IEEE Transactions on PAMI, 1991. **13**(1).
9. Wandell, B.A., A.E. Gamal, and B. Girod, *Common Principles of Image Acquisition Systems and Biological Vision.* Proceedings of the IEEE, 2002. **90**(1).
10. Ribo, M., A. Pinz, and A. Fuhrmann. *A new Optical Tracking System for Virtual and Augmented Reality Applications*. in *IEEE Instrumentation and Measurement Technology Conference IMTC*. 2001.
11. Welch, G., et al. *The HiBall Tracker: High-Performance Wide-Area Tracking for Virtual and Augmented Environments*. in *Symposium on Virtual Reality Software and Technology*. 1999.
12. Yip, P. and K.R. Rao, *Discrete Cosine Transform: Algorithms, Advantages, and Applications*. 1997: Academic Press.