Factors affecting the design and tracking of ARToolKit markers

Dawar Khan *, Sehat Ullah, Ihsan Rabbi

Department of Computer Science and IT, University Of Malakand, Pakistan

A R T I C L E   I N F O

Article history:
Received 12 May 2014
Received in revised form 1 January 2015
Accepted 8 February 2015
Available online 16 February 2015

Keywords:
Fiducial markers
ARToolKit
Augmented Reality
Vision based tracking
Robust recognition

A B S T R A C T

The quality of ARToolKit markers plays a vital role in the performance of Augmented Reality applications, but currently there is no algorithm or quantitative measure to guide users for designing high quality markers and their reliable tracking. This paper presents eleven factors that are important for designing and tracking ARToolKit markers. The effect of each factor on the quality (accuracy, detection speed, and inter-marker confusion) of marker tracking system is investigated and the optimal value(s) of these factors are found. Using the optimal values of these factors one can achieve the goal of reliable and robust tracking of fiducial markers.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Augmented Reality (AR) is the technology in which computer generated information such as graphs, text, objects or audio are superimposed over some part of the real world environment in order to enhance the user perception of the real world environment [1]. Augmented Reality, also known as Mixed Reality, aims to mix virtual world in the real world so that the viewer perceives that virtual objects are part of the real environment [2,3]. AR has many applications like in teaching and learning media [4,5] and many other applications such as in navigating a mobile robot [6] and in video based AR conferencing [7]. It is used in medical, entertainment, military, engineering design, robotics and tele-robotics, manufacturing, maintenance and repairing [8]. A marker based tracking system is used in various Augmented Reality (AR) applications to determine the camera pose by detecting one or more fiducial markers [9]. It is mostly used in indoor vision based applications [3]. Fiducial markers are images placed in a real world scene, to be detected by a video camera by comparing their similarities with their pre-stored templates and are also used for matching two real world images [10]. Pose can be determined with accuracy and low cost because fiducial markers use only a camera as an additional hardware. Some examples of ARToolKit markers are shown in Fig. 1(a–d); where Fig. 1(a), is showing a screenshot in which virtual solid teapots are overlaying the detected markers.

For designing various marker based applications, different toolkits such as ARTag, ARToolKit and ARToolKit Plus are used. ARToolKit is open sourced, easy to configure, well-documented and widely used in AR applications. It has comparatively less execution time than ARTag and ARToolKit Plus [11,12]. Although ARToolKit is a simple toolkit, its users still have several problems in their attempt to achieve high quality and robust tracking of the markers. Throughout the process of the ARToolKit marker tracking, users have no knowledge about the effect of different factors on the quality of the marker tracking. Although example patterns for fiducial markers are available, still users have no knowledge about various properties of a marker. For example, they may not decide whether to keep the black border thinner or thicker. Similarly, they fail to decide whether to create lesser or more number of black objects in the interior white region of the marker. In this paper we present eleven factors that need to be considered while designing and tracking ARToolKit markers. Some of these factors are concerned with the marker shape, some of these are concerned with additional hardware used while other are concerned with programming parameters. Various tradeoffs of these factors are also studied. We experimentally test the effect of each factor on the marker quality (i.e., detection speed, accuracy and inter-marker confusion) and statistically analyze the results to identify the optimal value for each factor. Our work will encourage the ARToolKit marker designers in designing high quality fiducial markers for various applications and will achieve the robust tracking of these markers. The rest of the paper is organized as: in Section 2, related work about fiducial markers, quality of fiducial markers and existing problems are discussed. In Section 3, we present all the eleven factors that we have identified, and the experimental evaluation of each factor. In Section 4, conclusion and future work are described.

2. Related works

Various studies have been conducted in the field of marker based Augmented Reality on the quality and applications of fiducial markers. This section covers the literature review on fiducial markers, their
Fiducial markers are images placed in the environment and are detectable by a video camera [6]. Marker based tracking is one of the popular methods in AR where predefined markers are placed in the real scene, from which their position can be calculated [2]. An ideal fiducial marker should contain at least four points in an approximately squared fashion [13]. However, in some marker tracking systems the shape of the marker is not necessarily square [13]. ARToolKit uses square markers [13]. The square shape is better for easy pose calculation [14]. Black and white markers are robust in various light intensity levels and are easy to print [15] whereas color markers allow a comparatively large number of distinct markers [16]. Markers of ARTag are also square in shape where the marker is divided into 10 x 10 segments [17].

2.2. Applications of fiducial markers

Apart from AR applications, fiducial markers are used in other areas including robot localization, medicine, television, industries, video conferences, and human-machine interactions [16,18]. Markers are used in indoor AR, hand-held objects for showing user pose to the camera, message tags to obtain some feature from marker, or general pose calculation for industrial setting [10]. Fiducial markers are used in a wide range of AR applications, including mobile robot navigation [6], 3D modeling [19], and education and learning aids [4,5]. Music AR is described in [20] as an AR based game for children to enhance their attraction using webcam and to allow them to toggle on/off members of the musical team, or their musical instruments. ASR (Augmented Sound Reality) is another project developed in ARToolKit that mixes AR with a 3D sound environment [20]. Besides Augmented Reality, some types of markers may contain other data, such as phone number, a URL, or a GPS coordinate, and are used in mobile applications, giving URL to web browser and obtaining GPS data from a marker [14].

2.3. Quality of fiducial markers

The best fiducial marker is one that can be detected with an easy and reliable fashion in all situations [14]. The selection of the best fiducial marker depends upon the application [14]. When following a proper designing method, the marker can be constructed with high reliability [10]. The requirements for high quality markers are: the marker should be distinct from the surrounding, unique in an existing library of markers, passive (not coated with electronic substances), quickly detectable and effective in low light and noisy environment using a robust image processing algorithm [10]. The use of black and white markers enable easy detection in various lighting conditions [14], whereas using more colors support a greater number of unique markers [16]. The main challenge in AR applications is to achieve high speed and accuracy with low cost and minimal changes in the applications [3].

Rencheng Sun et al. have designed new kinds of markers (i.e., QMarkers), and have carried out comparison with ARToolKit in terms of the marker recognition rate and its reliable tracking [2]. Similarly, in [17], a new type of markers (called diagonal connected component markers) is introduced and the results are compared with ARToolKit markers in terms of detection time. V. F. da Camara et al. [16] use color based markers to create up to 65,000 unique markers with comparatively greater accuracy. They mentioned that their markers may be misdetected; due to low quality cameras. The paper also concluded that larger quantity of distinct markers can be achieved by increasing their size [16]. However, there is no experimental proof given. In [10] Mark Fiala proposes eleven evaluation criteria for the quality of the marker system, namely: 1. the false positive rate (the rate at which absent markers are falsely reported as present), 2. the inter-marker confusion rate (the rate at which wrong id is reported), 3. the false negative rate (the rate of misdetection of a present marker), 4. the minimal marker size (the minimal size in pixel for reliable detection), 5. the vertex jitter characteristics (the noise in marker), 6. the marker library size (the number of unique markers in the library), 7. immunity to lighting conditions, 8. immunity to occlusion, 9. immunity to photometric calibration, 10. perspective support and 11. the performance speed.

2.4. ARToolKit and alternative marker tracking systems

ARToolKit [21] uses markers to translate and rotate virtual objects over the real world environment [11]. ARToolKit is the first marker tracking system used for marker-based AR applications while ARTag and ARToolKit Plus were later introduced [19]. ARTag is designed by the National Research Council of Canada and is also gaining popularity in the recent AR systems due to its improved performance [10,17]. Minimal inter-marker confusion rate, low false negative rate, and immunity to various lighting conditions and partial occlusion are the key advantages of ARTag over ARToolKit and ARToolKit Plus [19]. The main problem with ARTag is the computational workload which makes it incompatible with modern mobile devices [14]. Compared to ARToolKit,
ARTag has two main limitations: no reduction of vertex jitter (noise at the corner) upon defocusing and longer processing time [12]. Source code unavailability is another disadvantage of ARTag [22]. ARToolKit Plus also has the same limitation of longer execution time [11]. ARToolKit is the only toolkit which provides free source codes. Although some other toolkits are also freely available, they are not open source [11]. ARToolKit is widely used in AR and Human–Computer Interaction (HCI) systems due to its freely available source code [10]. ARToolKit runs on a range of different operating systems, including SGI IRIX, PC Linux and PC Windows, each with a separate version of ARToolKit [21]. In addition, ARToolKit has a short execution time and is most suitable for real time applications [11]. The ARToolKit markers are square having some patterns in their internal white region [23].

2.5. Problems in designing high quality markers

Although ARToolKit is very simple and well documented, marker designing, registration (storing templates of markers), and their tracking are still challenging tasks. When two or more different ways exist to reach a particular destination and you have no knowledge about the best way, an ambiguity will automatically come in mind. ARToolKit users also face the same kind of problem in the marker designing and tracking process. For example different sizes can be used for the black border as shown Fig. 2, so the user is confused whether to choose a thinner border or a thicker border for the marker.

Similarly different frame rates of the video camera may be available and user is confused about the optimal frame rate. Another design decision is information complexity. Users have to decide whether a marker containing a small number of black objects in its interior white region is better than the one that contains large and complex information. In addition, printing the markers on the paper causes ambiguity as to whether to use a printer that is deeply toned ink or to use the printer that is slightly toned ink on the paper, or they can use any printer. Another problem is that the ARToolKit provides some patterns of markers but there is no algorithm that checks the marker quality. ARToolKit users can design markers easily but they face difficulties in designing high quality markers. Another problem is that there are different applications of ARToolKit and some applications require high speed while others need high accuracy, and some applications need both accuracy and fast detection speed. So the selection of the best marker is also dependent on the application. Currently there are no clearly defined guidelines or quantitative measurements which can be used for designing high quality markers. Due to all previously discussed challenges, ARToolKit marker tracking has a high rate of false identification (false positive rate), high inter-marker confusion, high rate of miss detections of a present marker (false negative rate) and sometimes detection rate becomes slow.

3. Factors affecting marker quality

We have identified eleven factors that affect the marker quality. The effect of each factor on detection speed, accuracy and inter-marker confusion has been evaluated. Each factor has different levels; we have identified the optimal level for each factor, the optimality has been proved experimentally. However there is a tradeoff among various factors and marker quality, and some factors have different optimal levels for different applications. Now ARToolKit users are encouraged to use the optimal values of all these factors to achieve a reliable and robust tracking. These factors are divided in the four categories as shown in Fig. 3.

To verify the proposed solution, each factor was individually tested. However some of them have the same experimental setup. While
testing one factor all the remaining factors are kept constant. The frame rate in all the experiment was 20 frames per second (except testing the frame rate factor itself). The quality of the marker is measured in terms of some or all of these three parameters: 1) the number of true identifications in a fixed amount of time, 2) the number of false identifications (i.e., inter-marker confusion) and 3) the CF value (i.e., Confidence Factor).

The first parameter (i.e., no. of true identifications in a fixed amount of time) shows average true detection speed, the increase in this parameter means good quality (i.e., quick detection and low miss-detection). In other words, the large numbers of true detections means little miss detections and faster true detection speed and vice versa. The second parameter shows the rate of falsely reporting the ID of an absent marker; the increase in this parameter means bad quality (i.e., high inter-marker confusion and low accuracy). In [10] Fiala has used these two parameters with other names such as ‘false negative rate’ and ‘false positive rate’ respectively. The third parameter (i.e., CF Value) is the value from the range [0 1] which shows how much the marker tracking system is sure that the marker is exactly as the identical one. In [23] the CF value has been used with the alternative name as the ‘degree of similarity’ of the detected marker with its pre-stored template. A lower value means higher ambiguity (i.e., low accuracy) and a higher value means lower ambiguity (i.e., high accuracy). In short, for high quality the first and last parameters (i.e., true detections and CF value) are desired to have higher value and the second parameter (i.e., false detections) is required to have low value. Usually among some of these parameters for measuring the marker quality, there exists a tradeoff. For example, in [16] library size (nos. of markers in library) and accuracy have been increased with the usage of color based markers which leads to reduction in detection rate. The CF values are stored and then we take average of all the values. The marker library was composed of similar markers in order to better analyze the inter-marker confusion.

The experiments were carried out using HP G62 Laptop, having Intel(R) Core(TM) i3, 2.27 GHz CPU, 4GB RAM, an HP webcam 101 and two additional cameras were also used. For measuring light intensity we use a digital light meter and for printer quality we use three different printers.

3.1. Marker shape related factors

These are the factors related to marker designing before printing the markers on the paper. These factors are discussed in the following subsections.

3.1.1. Black to white ratio (B/W)

Black to white ratio (B/W) means how much the white region inside the black border is enough. For a fixed size marker if we decrease the B/W (i.e., decrease black border or increase white region) multiple information can be added into the white region to design a number of unique markers, but mis-detection will occur because the smaller border becomes invisible some time due to the orientation of the marker. On the other hand, the marker with larger B/W (i.e., larger black border) is mostly visible in various orientations and has low mis-detections. Similarly, few distinct markers can be created. In short, the markers with larger B/W are optimal for applications which require less number of markers in the library whereas smaller B/W is better when a large number of markers are required in the library. The B/W of a marker can be calculated as: If ‘S’ is one side length of the square marker and ‘b’ is the breadth of the black border then:

\[
\text{Total area} = S^2 \quad (1)
\]

\[
\text{White area} = (S - 2b)^2 \quad (2)
\]

\[
\text{Black area} = S^2 - (S - 2b)^2 \quad (3)
\]

\[
\text{Black area} = 4b(S - b) \quad \text{Now} \quad (4)
\]

\[
\frac{\text{Black area}}{\text{White area}} = \frac{4b(S - b)}{(S - 2b)^2} \quad (6)
\]

The B/W has a direct relation with percent border which can be calculated with respect to the length of one side using Eq. (7):

\[
\text{Percent border} = \frac{b}{S} \times 100. \quad (7)
\]

For example, the marker in Fig. 4 has a length of S = 100 pixels on one side, and border width b = 10 pixels. Now using Eq. (1) the total area = 100 × 100 = 10,000 pixels, using Eq. (2), the white area = 80 × 80 = 6400 pixels and using Eq. (4) the black area = 10,000−6400 = 3600 pixels. Similarly using Eq. (7), the Percent border = 10 and using Eq. (6) the ratio is B/W = 3600/6400 = 36/64. The larger the B/W ratio, the higher the accuracy but not good for creating a large number of distinct markers because the white region decreases with the increase in B/W.

To experimentally identify the optimal range of B/W, we used five different levels of B/W ratio (see Table 1). In order to get the accurate results we used three or four markers in each level and calculated the average results for identification of the optimal level(s). For all markers, the edge sharpness was 93.11% (unwanted pixels 5.89%), whereas the information complexity was 11% black pixels in the interior white region with only 2 objects in it.

We rotated and translated each of these markers for 5 min in front of the camera and counted the number of true and false identifications of each marker and CF values. The frame rate used here was 20 frames/s. The average values for each level are listed in Table 2.

<table>
<thead>
<tr>
<th>Level</th>
<th>B/W ratio</th>
<th>True identification</th>
<th>Average CF</th>
<th>False identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>1.01</td>
<td>2113</td>
<td>0.6892</td>
<td>1309</td>
</tr>
<tr>
<td>Level 2</td>
<td>2.35</td>
<td>2532</td>
<td>0.8129</td>
<td>1024</td>
</tr>
<tr>
<td>Level 3</td>
<td>5.02</td>
<td>3198</td>
<td>0.8375</td>
<td>823</td>
</tr>
<tr>
<td>Level 4</td>
<td>12.97</td>
<td>2977</td>
<td>0.7104</td>
<td>1410</td>
</tr>
<tr>
<td>Level 5</td>
<td>66.10</td>
<td>1432</td>
<td>0.6673</td>
<td>873</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Different marker structures used for testing black to white ratio.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level</th>
<th>Nos. of markers</th>
<th>B/W of each marker</th>
<th>Average B/W</th>
<th>Each marker’s % border width w.r.t its length</th>
<th>Average % border width</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>4</td>
<td>0.64, 0.85, 1.11, 1.42</td>
<td>1.01</td>
<td>11.01, 13.3, 15.6, 17.89</td>
<td>14.45</td>
</tr>
<tr>
<td>Level 2</td>
<td>3</td>
<td>1.81, 2.3, 2.92</td>
<td>2.35</td>
<td>20.18, 22.48, 24.77</td>
<td>22.48</td>
</tr>
<tr>
<td>Level 3</td>
<td>3</td>
<td>3.75, 4.86, 6.42</td>
<td>5.02</td>
<td>27.06, 29.45, 31.65</td>
<td>29.35</td>
</tr>
<tr>
<td>Level 4</td>
<td>3</td>
<td>8.70, 12.20, 18.01</td>
<td>12.97</td>
<td>33.94, 36.23, 38.53</td>
<td>36.23</td>
</tr>
<tr>
<td>Level 5</td>
<td>3</td>
<td>28.70, 51.80, 117.81</td>
<td>66.10</td>
<td>40.82, 43.12, 45.41</td>
<td>43.22</td>
</tr>
</tbody>
</table>
The results in Table 2 show that level 2 and level 3 are the best quality levels. Level 4 is also better (in terms of true detection) but level 4 has a smaller white region so that the possibility of creating a larger number of unique markers is reduced. It is smaller that the smaller white region has no capability of containing more objects. So, while keeping level 4 we cannot create more distinct markers, and the intermediate marker confusion will increase quickly with an increasing number of markers in the library. With a 25 × 25 cm² marker size and keeping the B/W ratio at level 4, we have an average border width of 36.23% on one side length (i.e., 36.23% of 25 cm which is equal to 9 cm). By cutting a 9 cm border from each side (left, right, top, bottom), we get only a 7 × 7 cm² interior white informative region. So with the 7 × 7 cm² interior white informative region (smaller region), we cannot create more distinct markers. In addition, the false identification rate is also comparatively larger for level 4 (see Table 2), so level 4 is not considered as an optimal level and the optimal levels are level 2 and level 3. These optimal levels as shown in Table 1 have B/W ratios in the intervals [1.42 8.70] and percent border thickness in the intervals [17 34].

3.1.2. Edge sharpness

Edge sharpness refers to how the intensity change has occurred in the marker and whether it is an abrupt change from black to white (sharp) or a slow and smooth change. We found that sharp edges are better for all circumstances because there is no ambiguity in such a case. Markers with sharp edges are identified accurately. An example of markers having different levels of edge sharpness is shown in Fig. 5. For an objective measurement of edge sharpness we calculated the noise in the interior white region of the marker using Eq. (9). Here we counted the number of unwanted pixels, i.e., the grey level pixels between black and white. We considered the pixels whose values lie in an interval of [0 20] as black and the pixels whose values lie in an interval of [235 255] as white while the rest (i.e., [21 234]) were considered as unwanted pixels.

\[
\text{Percent value of noise at edges} := \frac{\text{Number of unwanted pixels}}{\text{Total number of pixels}} \times 100. \quad (8)
\]

To experimentally identify the effect of edge sharpness on marker quality, we used three markers, as shown in Fig. 5. The noise at the edges (as calculated by Eq. (8)) in Marker 1 is 6%, in Marker 2 is 23.5% and in Marker 3 is 36.9%. For all the three markers, the B/W ratio was 3.0 and information complexity was 19% black pixels in the interior region with only 2 objects in it. We rotated and translated each of the marker in front of the camera for 5 min and stored the above three measurements (i.e., true identifications, false identification and CF value) and finally, the results were compared. We have identified that the sharper edges are better in detection. Marker 1, Marker 2 and Marker 3 as shown in Fig. 5 are three markers, such that the edge sharpness decreases in the following order: Marker 1, Marker 2 and Marker 3. The results of the five minute experiments are shown in Fig. 6 such that Fig. 6(a) shows the true and false identifications per 5 min and in Fig. 6(b) the CF values are plotted graphically after sorting in ascending order. Here we found Marker 1, being sharper in edges, to be the best one in the number of true identifications as well as in CF value. Although only few (three or four) CF values in the best case of Marker 3 are greater than those of Marker 2 and Marker 1, examining the overall results proves that Marker 1 is the most optimal. In other words the results at the end of each plot line show the best case of the respective marker, so we did not consider the four CF values at the best case of Marker 3 because the overall results of this marker were not good.

3.1.3. Information complexity

Information complexity means the number of black objects in the interior white region of a marker. We calculated the information complexity of a marker using two parameters. These are the number of same sized objects in the interior white region and the percentage of the black pixels in this region (calculated using Eq. (9)). Examples of information complexity are shown in Fig. 7, in which Fig. 7(a) contains simple markers, Fig. 7(b) contains average complexity markers and Fig. 7(c) contains markers of complex information:

\[
\text{complexity} := \frac{\text{Number of black pixels}}{\text{Total number of pixels}} \times 100. \quad (9)
\]

The information complexities of markers given in Fig. 7(a) are 4.2%, 5.7% and 8.0%, while they have 2, 3 and 4 numbers of same sized objects respectively. Similarly, the markers in Fig. 7(b) have 9.9%, 11.5% and 13.0% information complexities, with the number of same sized objects
5, 6 and 7 respectively. The markers in Fig. 7(c) have 16.5%, 20.7% and 34.0% information complexities, with the number of same sized objects 10, 12 and 16 respectively. For all the markers, the B/W ratio and noise at the edges (edge sharpness) were 3.0 and 5.89% respectively. The simple marker is having minimal black objects in its interior white region and it is having minimal intensity changes (number of edges) from black to white and vice versa.

To identify the effect of information complexity experimentally, we used nine markers having different complexity levels as shown in Fig. 7. We divided the nine markers in the three levels of complexity: level 1 contains the markers with simple information as shown in Fig. 7(a), level 2 contains average complexity markers as shown in Fig. 7(b), whereas level 3 contains complex markers as shown in Fig. 7(c).

To identify the effect of information complexity experimentally, we used nine markers having different complexity levels as shown in Fig. 7. We divided the nine markers in the three levels of complexity: level 1 contains the markers with simple information as shown in Fig. 7(a), level 2 contains average complexity markers as shown in Fig. 7(b), whereas level 3 contains complex markers as shown in Fig. 7(c).

The average results of the three markers were calculated and all the three levels were compared as shown in Fig. 8. Here, information complexity is evaluated while keeping the remaining factors unchanged. We identified that markers with simple information are quickly detected, but not effective if we need a larger number of unique markers in the library (i.e., larger library size). The complexity in information is slow in detection and miss detection may occur but we can create multiple unique markers by adding more information to the interior white region of the marker. In general, we can say that information complexity is dependent on the application in which we use it. If we need a larger library size the information must be enough larger to design all the markers distinct in structures. In contrast markers with simple information are optimal if library size is small and quick detection is desired.

In addition, we designed an optimized and a non-optimized marker. The optimized marker has 3.0 B/W ratio, 9% black pixels with a single object in it, while the noise (unwanted pixels) is 5.43%. The non-optimized marker has 1.3 B/W ratio, 29% of black pixels with 17 objects in it, having 15.06% noise. The results are shown in Fig. 9, which reveal that the optimized marker is better in all the three parameters (true detection, false detection and CF values).

### 3.2. Marker placement related factors

The following four factors are concerned with how to place the marker in the real world environment, so that the video camera can easily and accurately detect it.

#### 3.2.1. Light intensity

The luminance of light is very necessary. As human eye, the computer eye (i.e., camera) cannot detect the passive markers in the absence of light and even very bright light. We need optimal light intensity for accurate and quick detection. As we are talking about passive markers (i.e., based on light reflection alone rather than electronic coating), light intensity has an important effect on the marker quality.

To experimentally identify the effect of light intensity on marker tracking, we used two markers in a library of 18 markers. We performed the test using different intensity levels of light. To measure light intensity we used a digital light meter which measures the light intensity in unit of Foot Candle (FC). We tested the markers using fourteen different experiments. The first eight experiments were carried out in a room environment. The next two experiments were carried out in an open shadow environment (i.e., where the sunlight does not directly fall on the marker). The remaining four experiments at intensity values in a range of [3.2 KFC 5.3 KFC] were in an environment open to sunlight. Fig. 10(a) shows the speed of identification (true and false) for 5 min, whereas in Fig. 10(b) the average CF value is given which shows accuracy. The results show that greater light intensity in a shadow environment provides faster and accurate detection and minimizes inter-marker confusion. However in a sunny environment where the sunlight directly falls on the marker with very high brightness is not good for the detection process. The optimal range that we identified for light intensity is [55 FC 2.8 KFC]. This range as shown in Fig. 10 is better in true identifications as well as in CF values; the false identification rate is also not so great.

#### 3.2.2. Light–marker–camera (LMC) angle

It is the angle made between light source, marker and camera. The light source means the origin from where the light is coming. It may be a light generating torch or a bulb or even a window in a room. If normal to the marker surface is anti-parallel to camera and the LMC angle is in the range of [70 120] (as shown in Fig. 11(b)) or [−70 −120] degree (as shown in Fig. 11(c)) the detection is better.
To identify the effect of LMC angle on marker tracking we used two distinct markers (one having complex information and the second one, simple) and both were tested in the three different ranges of the angle. For both markers, the B/W ratio and noise at the edges (edge sharpness) were 3.0 and 5.89% respectively. The simple marker contained one black object with 7% of black pixels while the complex one contained 11 black objects with 26% of the black pixels in the interior region. The library size was 20 markers during these experiments. We performed all the experiments in a closed room where the light was coming from one window (i.e., light source was a window). When the camera was in between the marker and window (i.e., light source), the LMC angle was 0° and moving the marker up and down made the range $[-50, 50]$ (as shown in Fig. 11(a)) which is called range 1 of the LMC angle. Similarly, when the marker and camera were co-linear, such that the coming light falls perpendicular to normal to the marker surface and the camera was anti-parallel to normal to the marker surface, the angle was 90° and marker is moved up and down or left and right in the range $[70, 120]$ or $[-70, -120]$ degree (as shown in Fig. 11(b) and Fig. 11(c)) which is range 2. In the same way, when the marker was in between the light source (i.e., window) and the camera, the LMC angle was 180° and moving the marker up and down or left and right in the range $[130, 230]$ (as shown in Fig. 11(d) which is range 3). These ranges are enough for testing because for the remaining values of the LMC angle the markers are out of the camera sight. We tested each range for 7 min and stored the results for each. The results are shown in Table 3. The results of both the markers under testing were aggregated in each range. The range 2 where the LMC angle is near to 90° was found best in the numbers of identification (i.e., quick detection) as well as in the CF value (i.e., accuracy); false identification is also less in this range.

Fig. 9. Comparison of an optimized and a non-optimized marker.

Fig. 10. Results of experiments on the effect of light intensity.
3.2.3. Surface smoothness

The surface upon which the marker is mounted needs to be smooth. The smoother the surface the better the detection will be and vice versa.

To identify the effect of surface smoothness on marker tracking we used two markers and the library was composed of 20 markers. Here we tested both markers in three surfaces of different smoothness. One marker at three different surfaces is shown in Fig. 12. These markers were printed on similar papers and then mounted over these different surfaces. Surface 1 is very smooth as shown in Fig. 12(a), surface 2 has an average smoothness as shown in Fig. 12(b), and surface 3 is non-smooth (rough) as shown in Fig. 12(c). For an objective measurement of surface smoothness, we stored the image pattern (patt1) and then we took the image of the marker after mounting over the surface (patt2). We used image correlation [24], to calculate the degree of similarity between the two images (patt1 and patt2). The maximum degree of similarity means low distortion in the shape and high smoothness of the surface. In other words, the roughness of the surface creates distortion in the image which makes the marker difficult for detection. The degree of similarity of the stored pattern with marker on surface 1 is 0.99 (i.e., 0.01 distortion), on surface 2 it is 0.86 (i.e., 0.14 distortion), while on surface 3 it is 0.66 (i.e., 0.34 distortion). Here we used the same camera for distortion measurement of the three surfaces.

Table 3: Results of the experiments on LMC angle.

<table>
<thead>
<tr>
<th>LMC angle (in degree)</th>
<th>Range 1</th>
<th>Range 2</th>
<th>Range 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMC angle (in degree)</td>
<td>[−50 50]</td>
<td>[70 120] and [−70 −120]</td>
<td>[130 230]</td>
</tr>
<tr>
<td>True identified (per 7 min)</td>
<td>2865</td>
<td>3674</td>
<td>333</td>
</tr>
<tr>
<td>False identified (per 7 min)</td>
<td>1098</td>
<td>1007</td>
<td>17</td>
</tr>
<tr>
<td>Average CF value</td>
<td>0.661766</td>
<td>0.75096</td>
<td>0.64349</td>
</tr>
<tr>
<td>Remarks (quality)</td>
<td>Good</td>
<td>Best</td>
<td>Worst</td>
</tr>
</tbody>
</table>

The combined results of both markers in each surface are shown in Fig. 13 which shows that marker tracking with smoother surface is better in accuracy and speed and gives less inter-marker confusion. The CF values are also given in Fig. 13 for each surface. As the marker library contains twenty markers with slight differences and of which only two are under testing. It is why in surface 2 the number of false identifications is greater.

3.2.4. Physical movement of the camera

In most of the applications, either the camera or the marker itself is in motion. The marker detection also depends upon this speed. The lower the speed, the accurate and better the detection will be. In the case of higher speed the frame rate of the video camera needs to be greater.

To identify the effect of the camera’s physical movement on marker tracking, we used two markers and the library size was 18 markers. Here we tested both the markers in three levels of speed (i.e., slow, average and fast). By physical speed of the camera, we mean that the camera and/or the marker is in movement. Hence instead of moving the camera we moved the markers in front of the camera. The
Markers were moved in 20, 45, and 65 cycles per minute for slow, average and fast speed respectively. While one cycle was 1 m long. Each experiment’s duration was 7 min. The frame rate was 20 frames/s and all the remaining factors were constant. The marker size was 25 × 25 cm² and the distance between the camera and the marker was in a range of [0.80 m 1.20 m]. The results of the experiments are shown in Fig. 14. Here we identified that the faster speed of the camera causes less number of detection (i.e., slow detection rate) and less accuracy.

3.3. Programming parameter related factors

3.3.1. Frame rate

The number of frames that a video camera captures in 1 s is called the frame rate. The lower the frame rate the more accurate the detection will be. However, if the physical speed of the camera(marker) is greater, a higher frame rate is desired. For most of the normal applications a range of [10 20] frames/s is optimal as identified below.

To experimentally identify the effect of frame rate on marker tracking we used two markers and a marker library size of 18 markers. Here we tested both the markers in six different frame rates (i.e., 5, 10, 15, 20, 25, 30 frames/s). Although professional cameras are capable of greater than 30 frames/s, they are not considered due to additional hardware cost. For each frame rate, both the markers were moved in front of the camera in an average speed (i.e., 45 cycles per minute with a 1 m long cycle) for 5 min. The results are shown in Fig. 15 which show that frame rates in the range of [10 20] are optimal. As we have identified, this range is better in detection speed (no. of true identifications) and accuracy (CF value). In addition, the false identification rate is also lower in this range. However it depends upon the speed of the camera and for higher speed a higher frame rate is desired.

3.3.2. Marker segmentation

While storing the pattern we have an option of segmentation. The marker is segmented into 4 × 4, 8 × 8, 16 × 16, 32 × 32 or 64 × 64 segments and intensity information is stored for each segment. The higher the number of segments the higher the accuracy will be but detection will be slower and usually miss detection will occur. The lower the number of segments the quicker the detection will be but inter-marker confusion will be greater and proxy will occur.

To identify the effect of marker segmentation on marker tracking, we used one marker and a library size of five markers. We stored templates of all the five markers in 16 × 16 segments and performed the

![Fig. 13. Results of experiments on surface smoothness.](image1)

![Fig. 14. Results of experiments on the camera's speed.](image2)

![Fig. 15. Results of experiments on the frame rate.](image3)

![Fig. 16. Results of experiments on marker segmentation.](image4)
testing and then we tested it after saving the templates in 32 × 32 segments. In the remaining segmentations i.e., the 4 × 4 and 8 × 8 segments were impractical due very low accuracy and high inter-marker confusion. In other words the markers with these segmentations give up to 50% false identifications. The detection rate of the template with 64 × 64 segmentations was neglected due to very slow detection rate (i.e., just once per 5 min). The reason behind this slow detection rate is high accuracy of the 64 × 64 segmented template, so we neglected these from testing consideration. The results of the two experiments are shown in Fig. 16 which show that segmentation at 32 × 32 has low inter-marker confusion and it is most accurate (i.e., less nos. of false detection and high CF value) but very slow (i.e., less nos. of true detections); in contrast the segmentation of 16 × 16 is faster but it has high inter-marker confusion and is less accurate. As most of the applications need high speed and real time interaction so we suggest the 16 × 16 segments as the optimal value of segmentation however if accuracy is objective then 32 × 32 is optimal for use.

3.4. Additional hardware related factors

3.4.1. Resolution of the camera used

A better resolution camera will lead to accurate and quick detection and vice versa. To identify the effect of the resolution of the camera on marker tracking, we used two markers and total eighteen markers in the library. The two markers were tested three times using the three different cameras as shown in Table 4.

The results of the experiments on the three cameras are shown in Fig. 17, which show that camera 1 is the best in this case. We found camera 1 to be the best one in quick detection (i.e., high nos. of true identifications), low inter-marker confusion (i.e., less false identifications) and high accuracy (i.e., high CF value).

3.4.2. Quality of printer used

The quality of the printer used for printing out the marker into hard copy format also has a great effect upon the marker detection process. If the printer quality is better the marker will be printed on paper without noise and thus the detection will be better.

To experimentally identify the effect of the printer’s quality we print four markers of the marker library using three different printers as shown in Table 5. Two markers were tested in the library of four markers. All the markers were printed on each printer and results were compared. We found printer 1 (i.e., which prints with low noise and deep ink) to be the best one in quick detection and accuracy. The results of the 7 min experiments are shown in Fig. 18.

4. Conclusion and future work

We have identified eleven factors that affect the quality of ARToolKit markers. Some of these factors are related to the shape of the markers, some are concerned with the context and placement of the markers while others are concerned with programming parameters and additional hardware used. We changed each factor and checked its effect on the marker quality (i.e., accuracy, time and inter-marker confusion). We then identified the optimal value(s) of each factor. While keeping all these factors in the optimal range, one can design optimal markers and use the ARToolKit marker tracking system in a reliable fashion. In the future we aim to design a GUI application that will automate the design of high quality ARToolKit markers using the optimal value for various factors like B/W ratio, edge sharpness and information complexity.

References


[14] S. Siltanen, Theory and Applications of Marker Based Augmented Reality, P.O. Box 1000 (Vuorimiehentie 5, Espoo), FI-02044 VTT, Finland 2012.


