Vol.2014-MBL-72 No.13
Vol.2014-CDS-11 No.9
2014/8/28

Augmented Reality System to use 3-dimensional Marker for 360-degrees Graphic Rotations Seamlessly in Hand-Motions

MAKOTO USAMI[†1] KYOHEI MIURA[†1]
MASAO ISSHIKI[†1]

Abstract: Conventional 2D Augmented Reality (AR) markers enable to display and move 3D graphic images on the 2D plane surface by hand-motions. However the graphic images cannot be flipped seamlessly and smoothly in conjunction with AR markers in the front and the back sides on the plane surfaces. Authors devised a 3-dimentional marker and achieved the AR system to rotate graphic images in 360-degrees seamlessly and continuously. Since the 3D marker is coded, it enables to associate the AR marker with data and to display relevant graphic images on the coded marker. Since the AR system tracks multiple markers at the same time, relevant multiple graphic images can be displayed on multiple markers in left and right hands. The AR system expands application fields and contributes to create attractive Human Computer Interactions.

1. Introduction

Augmented Reality (AR) is a technology that involves the overlay of computer graphic images on the real world. AR system [1] shown in Figure 1 needs to determine the position and orientation of the camera. With a camera, the AR system is then able to render virtual objects in the correct place by calculating the relative pose of a camera in real time. Those are fundamental components of AR.

Nowadays AR is getting popular in entertainment and engineering purposes. However AR systems often use planar type 2-dimensional (2D) markers. The 2D markers are located on the walls, paper magazines and card games in many cases.

The Figure 2 shows an example of the 2D marker with the coordinate system in Rotation (X, Y, Z) based on ARTookKit [2]. We considered to devise a 3-dimenstonal (3D) marker and to expand the application fields. If 3D markers can be used in hand-motion type AR systems, we believe the AR system achieves more attractive human computer interactions (HCI) [3], and especially it is suitable to work with head-mount displays (HMD) [3]. Since HMD enables hands-free operations, 3D markers can be manipulated by hand-motions interactively. When there are 2 of 3D markers, we can use left and right hands to rotate 3D markers, and achieve attractive user experiences and HCI in the AR system. For example, when we watch a 3D game character image associated with a 2D marker such as ARTookKit and ARTag [4] on the plane surface such as a paper card, it is impossible to see the 3D character image in 360-degrees view by flipping the card in front and back side planes continuously and smoothly even though an additional marker is located in back side of the card. Because of the imaging system loses recognition visually of the 2D marker in high angle of views when the 2D marker is extremely rotated away as shown in the Figure 3. We developed the 3D marker enables to display 3D graphic images seamlessly by rotating the 3D marker in 360-degrees views especially in the hand-motion range so that the system can enable the attractive HCI and expand application fields of AR systems. Since this 3D marker is coded, it is possible for the AR system to identify each marker and to display different graphic images on associated 3D markers. The multiple display capability with the coded 3D maker enables AR operations on left and right hands dynamically. We can compare 2 graphic images by different angle of views at the same time.

In order to achieve the purpose, we describe the challenges to overcome below.

- To display overlay graphic images by rotating markers in 360-degrees in Rotations seamlessly and smoothly within hand-reach ranges in marker motions.
- To consider holding the markers by left and right hands for each. The AR system must recognize multiple markers.

In order to resolve challenges above, we utilize 1-dimensional (1D) color-code called Colorbit (CB) [5] [6] as shown in the Figure 4. We deformed the 1D code to 2D code to 90-degrees angled shape shown in the Figure 5. Furthermore, we devised a 3D marker from the 2D code by forming a cylindrical shape and printing RGB colors on the cylindrical object shown in the Figure 6. Since the 3D AR marker is coded, the AR system enables to display associated graphic images by referring to the database. In this paper, we describe the structure of 3D AR Marker, and the method of the new AR system so that the system can display 360-degrees seamlessly rotated graphic images on the marker in hand motions.

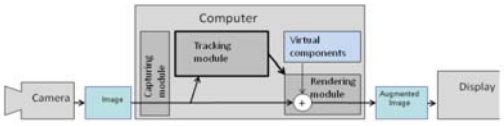


Figure 1: A Simple AR System [1]

†1 Kanagawa Institute of Technolog

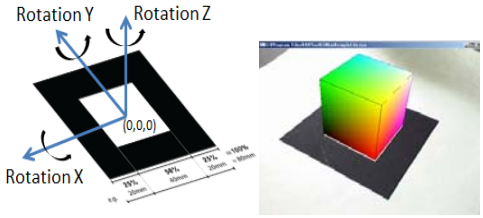


Figure 2: ARTookKit as an Example of 2D Marker

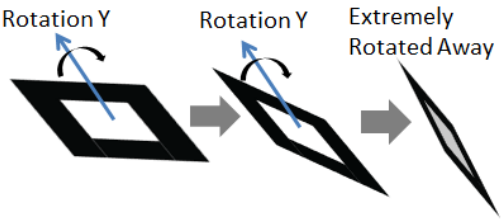


Figure 3: 2D Marker Extremely Rotated Away

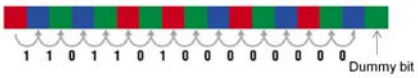


Figure 4: 1D-Colorbit Code

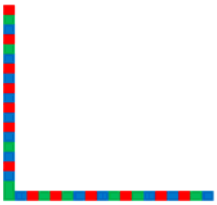


Figure 5: 2D 90 degrees angled Colorbit

We describe the structure of the work in this paper below.

- Chapter 2 describes comparisons to conventional methods.
- Chapter 3 describes the design of 3D AR marker and the method of the AR system in detail.
- Chapter 4 describes implementations and experimental results.
- Chapter 5 describes the conclusion and future works.

2. Comparison to Conventional Methods

In the case of conventional 2D planar type markers such as ARToolKit and ARTag, there is a problem below.

- Since the 2D marker is printed on the 2D plane surface, there is a limitation for an imaging system to recognize the 2D marker in extremely rotated away as shown in the Figure 3. Even though there are printed markers on front and back surfaces, it is impossible to display graphic images seamlessly without a chasm when the AR marker plane is flipped in Rotation X and / or Y.

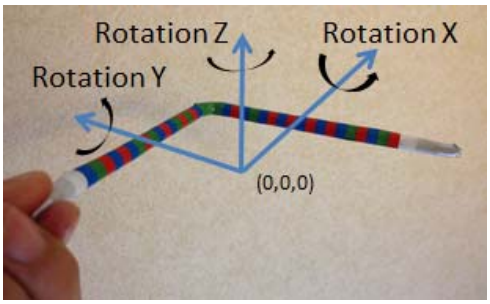


Figure 6: Color Coded 3D Marker by CB

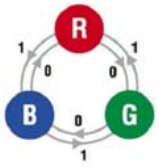


Figure 7: Sequence Law of CB

Figure 8: Deformable CB code

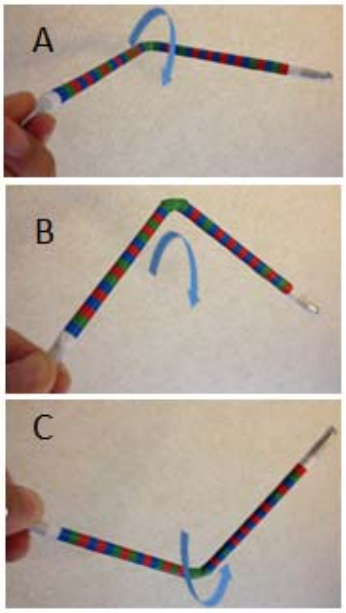


Figure 9: 360-degrees rotations by hand motions

3. Design of 3D Marker and AR system

3.1 Design of 3D maker

3.1.1 Theory of Colorbit (CB) code.

We use CB to structure the 3D AR marker. A standard 1D CB code as shown in the Figure 4 is visually coded by color transitions of red, green and blue as the combination sequences of three color cells. CB generates binary codes based on the law in the Figure 7. The code defines the first bit by a red cell in an edge of CB, and there is a dummy bit is located in the other edge as the stop bit by a green or a blue cell. If N is the total number of cells, CB counts the number of $2^{(N-2)}$. Since a combination of cells as a code is related to the marker profile in the database on the computer, the AR system is able to display the graphic images based on the profile information. Some cells can be assigned for error correction bits. So long as cells are connected as one stroke, CB can be deformed to any shapes as a deformed example shown in the Figure 8. However it is prohibited to overlap multiple CBs and/or branch the code shape. In another word, any CB code must be one-stroke shaped.

3.1.2 Structuring 3D Marker from 1D-CB

When the standard 1D-CB shown in the Figure 4 is deformed as 90-degrees angled shape by utilizing the shape flexible characteristic of CB, the marker becomes a 2D coded maker as shown in the Figure 5. The 90-degrees angled 2D marker enables 3D pose estimations visually if the marker is located on the plane surface as similar to ARToolKit. Since the marker is slender (defining the aspect ratio is 10:1 or larger) in shape, it is possible to achieve the compatibility of a small inhibition area to the plane surface and accurate pose estimations in the AR system. However when the 90-degrees angled marker is flipped over in rotations, discontinuous of the visual recognitions to the marker occur in the front and back side planes the same phenomenon as other 2D markers. In order to resolve the discontinuous problem, we devised the 3D marker by deforming the 2D 90-degrees angled marker to the 3D cylindrical shape so that the 3D marker can be visually recognized in 360-degrees rotations. The Figure 9 of A, B and C shows the continuous visible rotations in 360-degrees by utilizing the 3D marker in hand-motions. Since the visual pose estimation in the AR system provides Rotation (X, Y, Z) and Translation (X, Y, Z) values, the AR system enables to display 3D graphic images based on the related positions to the provided Rotation and Translation values accurately. In this paper, we use the 3D-90 degrees angled CB includes 36-cells and the distance between cells are always 4mm. Since the distances between cells are constant, the distances take a roll of the scale to be used in the measurement of pose estimations.

3.2 Theory of Pose Estimation on the 3D AR Marker

When the AR system displays graphic images precisely related to the pose of markers, the accurate pose estimation is required. In this chapter, we describe the theory of pose estimation techniques. Since the AR system envisions to be used in hand-motions of 3D AR maker, the distance range from the camera to the marker limits within 700mm by considering the typical reach of human hands.

We use the 3D 90-degrees angled-shaped CB (3D marker) shown in the Figure 6 on the related 2D image plain as shown in the Figure 10, and define Rotation (X, Y, Z) and Translation (X, Y, Z) on the 3D world coordinate system related to the principal point of the lens of the camera shown in the Figure 11. The Figure 12 shows the theory how to identify back and front views of 3D marker. In Rotation (-90, -90, 45), since 3D maker is angled to the back side in position, the lengths Yb is larger than Ya. In Rotation (90, 90, 45), since 3D marker is angled to the front side in position, the lengths Ya is larger than Yb. As a result that, the AR system is able to identify the both cases of Rotation (-90, -90, 45) and Rotation (90, 90, 45). If there is no scale by CB mechanism, it is not possible to identify the difference of angled directions in front or back.

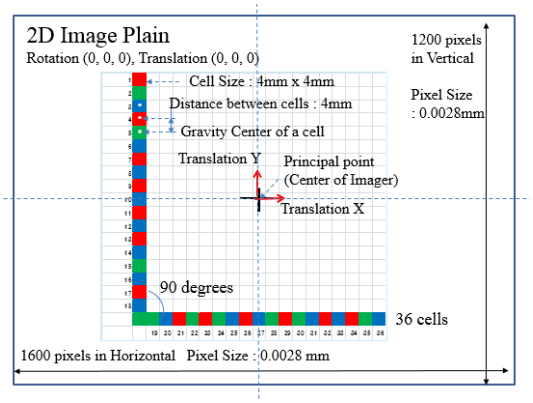


Figure 10: CB Design on 2D Image Plain

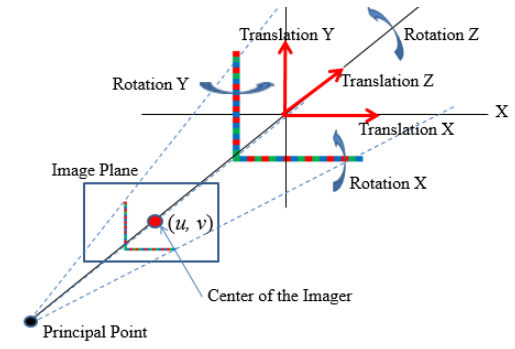


Figure 11: Rotation and Translation in 3D World Coordinate System X, Y, Z

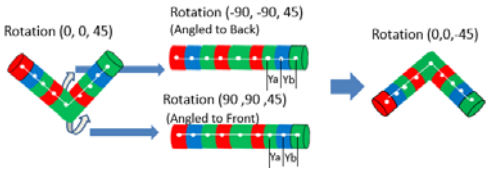


Figure 12: Theory of Front and Back Identifications

3.2.1 Flow and Theory of Pose Estimations

As shown in the Figure 13, we use the flow to obtain Rotation (X, Y, Z) and Translation (X, Y, Z) by utilizing the parameters of the 3D AR marker.

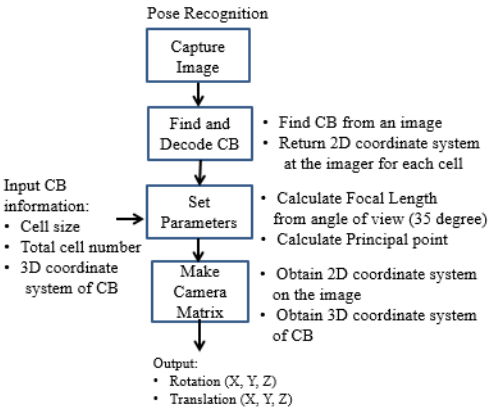
- Camera captures images by 35 degree angle of view.
- Image processing software on the computer (Software) finds the 3D marker from the image.
- Software decodes the 3D marker and calculates the gravity center of each cell, and obtains the 2D coordinate system on the imager. Each cell distance of the 3D marker is an important value to determine the pixel numbers and poses reflected to the imager.
- Software imports the 3D maker parameters such as cell size, total cell number and the 3D marker position of 3D coordinate system shown in the Figure 11.
- 

Figure 13: Flow of Pose Estimation (Recognition)

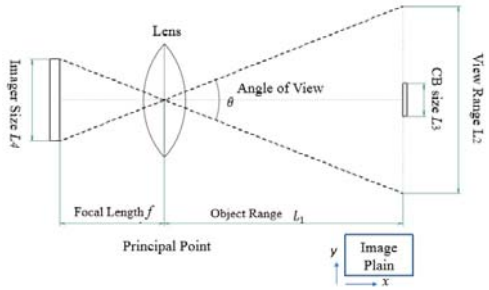


Figure 14: Pinhole Camera Model

The focal length (*fx*, *fy*) is calculated as following functions (1) and (2) The unit by pixels:

$$f_x = \frac{L_4 x}{2} \div \tan(\frac{\theta_x}{2}) \quad (1)$$

$$f_y = \frac{L_4 y}{2} \div \tan(\frac{\theta_y}{2}) \quad (2)$$

The principal point (*Cx*, *Cy*) is defined to the center of the imager as following functions (3) and (4) The unit by pixels:

$$c_x = \frac{L_4 x}{2} \quad (3)$$

$$c_y = \frac{L_4 y}{2} \quad (4)$$

- The system does neither apply chessboard calibrations [7] due to the CB's robustness to image distortions nor require pre-capturing of the marker image. ARToolKit requires complicated calibrations as well as pre-capturing many images.
- Software structures a camera matrix, and outputs Rotation (X, Y, Z) and Translation (X, Y, Z).

The camera matrix: CM can be structured by the equation as following (5).

$$\mathrm{CM} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

We use the function cvFindExtrinsicCameraParams2 of OpenCV1.0 [9]. When the 2D coordinate system on the imager is defined to the *s* (*u*, *v*), the equation is described as following (6). [10]

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (6)$$

X, Y and Z are input values of 3D coordinate system of the 3D marker. The unit is by [mm], and the values are normalized from the 3D coordinate system of the 3D marker.
Matrix r11 to r33 describe the rotation coordinate system.
Matrix t1 to t3 describe the Translation vectors.
We use Rodrigues formula [11] by the function cvRodrigues2 of OpenCV1.0 in order to transform the rotation coordinate system to the rotation vectors.
Finally the pose estimation system provides Rotation (X, Y, Z) by the unit of [radian] and Translation (X, Y, Z) by the unit of [mm] from the center of imager are obtained.
We use OpenGL [12] to display graphic images by utilizing the pose data of Rotation (X, Y, Z) and Translation (X, Y, Z).

3.2.2 Simulation of Required Pixel Numbers at Marker Distances

The Figure 24 shows the simulation of allocated pixel numbers on the imager (1600 horizontal x 1200 vertical pixels in the maximum resolution of the camera) depending on the 3D marker distances and Rotation X angle by using the theory of pinhole camera model [8] as shown in the Figure 14 and the 3D marker

shown in the Figure 6. It is important to know the maximum 3D maker distance to be recognized by the AR system in hand-motions on the AR system. Obviously more Rotation angle makes more difficult to recognize the 3D marker, and more distance makes more difficult to recognize the 3D marker.

3.2.3 Occlusion Issue It is inevitable to overcome the occlusion issues of cell-colors overlapped in a pose of 3D marker as shown in the Figure 15 if a visual pose estimation is used. In this paper, we describe the study excluding viewing angles in occlusions. The occlusion issue is one of future scopes to be improved.

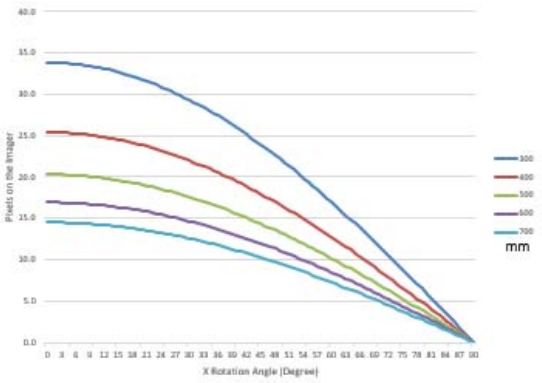


Figure 24: Reflected Pixels at Imager in Rotation X

4. Experimental and Measurement Results

4.1 Measurement Tools and Setting for Pose Accuracy

By using tools with the specs on Table 1 are shown in the Figure 16 including a camera, a lens, a sliding rail, a scale and mechanisms to adjust angles of the 3D marker. Distances and physical rotation angles were measured by the physical scales.

- The camera captures images by 35 degrees angle view of the lens.
- Measurement tools are always covered by the box to stabilize the light condition.
- Lighting condition is constantly adjusted in 500-lux luminance, and in 5,500-Kelvin color temperature.
- Camera is connected via USB2.0 to PC embeds Intel Dual Core-i5 processor.
- An example of recognized a 3D marker image with Rotation and Translation data is shown in the Figure 17. There is a white dot for each cell. Each dot is located in the gravity center at each cell.
- The Focal length 35 degrees angle of view in the standard measurement system is calculated to 7.1mm by the formula (1) at the chapter 3.2.1.

4.2 Measurement of Pose Estimation Accuracy

It is important to estimate positions and orientations of AR markers accurately so that the AR system can estimate the pose of the 3D marker in Translation (X, Y, Z), Rotation (X, Y, Z) and display graphic images precisely. We measure the accuracy of rotations and translations in the distance from 300 to 700mm between the 3D marker and the lens principal point by setting the camera resolution to 1600 horizontal (H) x 1200 vertical (V) pixels in maximum. The measurement ranges of Rotations in X and Y are limited in ±70 degrees due to the mechanical limitation of the measurement tool. The Figure 17 shows the standard pose of the 3D maker in Rotation (0, 0, 0) and Translation (0, 0, 0) on the measurement tool. When the AR system recognizes 3D marker, displays a dot for an each cell by decoding CB. Then the displayed positions of dots represent the gravity centers of the cells. The distance between dots change in rotations of 3D marker.

The Table 2 shows static measurement results of the pose estimation accuracy related to the distance differences on the AR system. Each data is sampled 100 times and averaged. The rotation error s of X axis shows within ±4.3 degrees, Y axis shows within ±3.5 degrees and Z axis shows ±1.9 degrees in the distance up to 700mm. The data of Translation Z describes the distance accuracy between the 3D marker and the lens principal point.

The translation error of Z axis shows within 32mm as 5.3% to the distances.

Table 1: Specification and Setup of Tools

Item	Unit	Specifications
Pixel Number (Camera: Microvision VC-4302)	*Pixel*	1600 Horizontal (H) x 1200 Vertical (V)
Imager Size	*mm*	4.48 H x 3.36 V
Pixel Size of Imager	*mm*	0.0028 x 0.0028
Lens Angle of View (Lens:Tamron MV12VM412)	*Degree*	35
Lighting (LED)	*Lux*	500
Color Temperature	*K*	5500

Table 2: Measurement of Pose Estimation Accuracy

	Distance (mm)	Rotation X (Degree)	Rotation Y (Degree)	Rotation Z (Degree)	Translation Z (mm)
		Measurement Range: -60 to 60 degrees		-90 to 90 degrees	
Error	300	+2.0 to +3.2	+0.5 to +3.0	-1.4 to +0.9	-4 to +9
	400	+1.7 to +2.9	+0.9 to +2.3	-0.7 to +0.6	-8 to +8
	500	-0.9 to +2.9	+0.7 to +3.3	-3.2 to +0.0	-13 to +15
	600	+1.8 to +4.3	-3.5 to +2.1	-1.9 to +1.1	-7 to +32
	700	-0.2 to +4.2	-2.6 to +1.6	-0.5 to +1.7	-20 to +7

Figure 15: Occlusion by Color Overlapping

Figure 16: Tools for Measurements

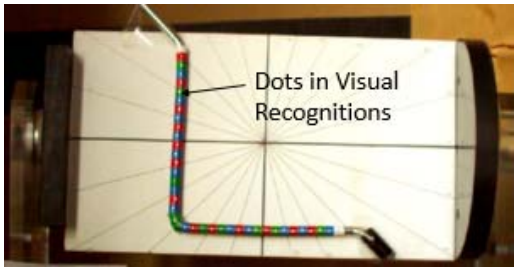


Figure 17: 3D Marker on the Measurement Tool

4.3 Real-Time Measurement of 3D Marker Rotations

4.3.1 Measurement Tool and setting for the Real-Time rotations.

Figure 18: Standard Positioned 3D marker Rotation (0, 0, 45) on the Turn Table.

The Figure 18 shows the 3D marker on the rotation table controlled by a DC power supply. We use the same camera and scaling system as the measurement of pose estimation accuracy in the chapter 4.2. Some differences are below.

- The shutter speed of the camera is set to 1/500 second to avoid motion blur.
- 3D marker is located in the distance 300mm away from the lens principal point by considering typical hand-motion ranges of human.
- Standard position of Rotation (X, Y, Z) = (0, 0, +45) as shown in the Figure 18.
- The turn table rotates in the clockwise constantly

4.3.2 Rotations in Real-Time Motion on Turn Table

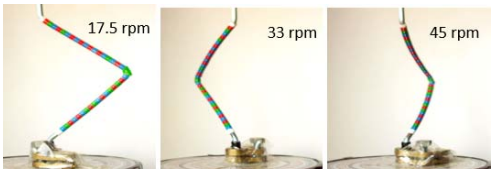


Figure 19: 3D Marker Rotations on Turn Table

In the Figure 19, the images show the recognized 3D markers and rotation shapes in motion on the turn table. We experimented the rotations in 17.5 rpm, 33 rpm and 45 rpm. It is possible to identify the visual recognitions of the 3D marker by shown dots in each cell. In the case of 17.5 rpm, image distortion is not observed. However in 33 and 45 rpm, the distortions in the 3D marker images occurred. The distortions in fast moving objects in imaging were caused by the rolling shutter [13] mechanism of the camera imager. If we use the camera is equipped with a global shutter [13] imager, the distortion can be avoided. However even though in distortions, the 3D marker is recognized continuously. We believe the rotation speed 17.5 rpm is more than enough for practical uses in the hand-motion AR system.

4.3.3 Accuracy of Pose Estimations in Real-Time Motion

It is important to estimate pose of AR makers accurately so that the AR system can display graphic images in precise positions. The Figure 20 shows the measurement data in the 17.5 rpm rotation speed in pose estimations of the 3D marker by following the time flow. The Figure 21 shows the position setting of 3D marker and rotations by defining the starting point and the standard position and pose of the 3D marker on the turn table. As shown in the Figure 20, rotation angles are mostly accurate. However we observe larger errors in the angle at around 90 and -90 degrees positions.

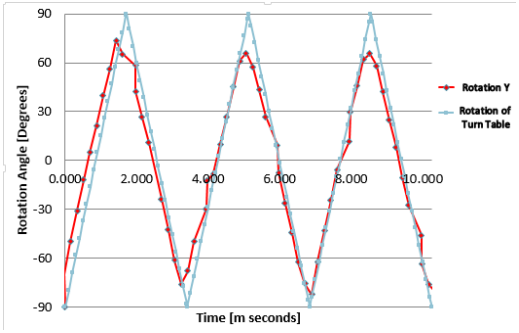


Figure 20: Time Flow and Rotation Angle in Motion (17.5 rpm)

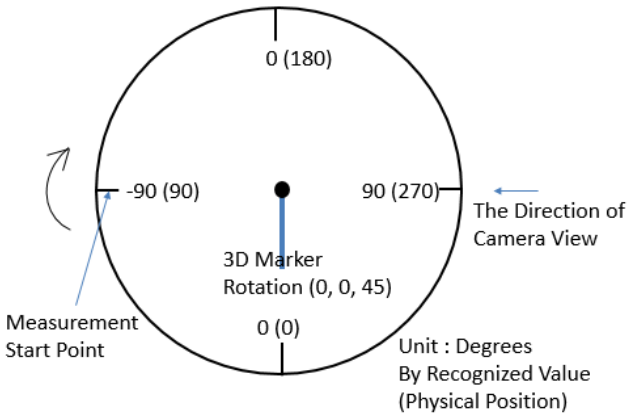


Figure 21: 3D Marker Positioning on the Turn Table

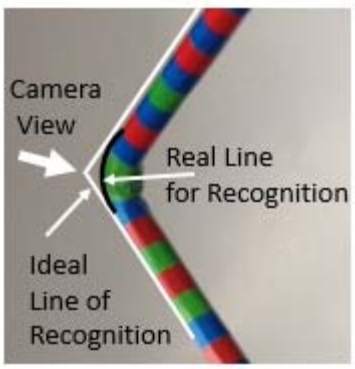


Figure 22: Reason of Error in the Angled Part

The Figure 22 shows the angled part of the 3D marker, and describes the ideal line for pose estimation and real line to be recognized by the AR system. When we develop the 3D marker, aluminum cylindrical sticks were used.
As a result that the angled part was bended carefully by avoiding distortions occurred, and caused larger errors in the round shaped part.

4.3.4 Discussion of Multiple 3D markers

In order to use the AR system in practical applications, multiple markers shall be recognized and display multiple graphic images simultaneously. In this method, since CB is utilized for the scaling purpose of pose estimations as well as coding to identify each 3D marker. We confirmed the AR system to recognize the multiple 3D markers and display associated graphic images for each 3D marker in real-time hand-motions. This paper does not describe the maximum numbers of recognition and display capabilities of the 3D marker in the AR system. However since it is important to know the CPU load dependency based on the number of 3D markers on the AR system, the experiment shall be done in the next step.

4.3.5 Demonstration of AR Display

Although we observed the accuracy to achieve 360 degrees seamless rotations by data measurements, it is important to experience and confirm the AR system by displaying graphic images as the feasibility demonstration level at least. The Figure23 shows the implemented AR demonstration by using the AR system. In order to display the graphic images, OpenGL [12] is used by getting Rotation and Translation values from the pose estimated results of OpenCV. We achieved 360-degrees rotations seamlessly by showing the graphic images associated with the 3D marker by hand-motions.

5. Conclusion and Future Work

In this paper, we described the development of the 3D AR marker to enable 360-degrees seamless rotations of graphic images on the 3D marker by deforming the 1D color-code Colorbit to the 3D marker on the AR system. We confirmed the practical level accuracy of pose estimations in the hand-motion range and the capability to display graphic images on the 3D marker in 360-degrees seamless rotations. Since multiple 3D markers can be recognized by the AR system, we can watch and compare different graphic views associated with 3D markers in right and left hands. As a next step, we study the CPU and other loads dependency related to the number of AR markers, and feasibility of the AR system can be achieved by generic smartphones or tablets. As of other objective to overcome, there are occlusion issues and the improvement of pose estimation accuracy at the angled part of 3D marker. Therefore we also plan to study the improvement of 3D marker structure by deforming it to many different shapes in order to overcome occlusion and accuracy issues.

We believe the 3D marker in the AR system is suitable to be used in engineering purposes as well as entertainments such as games, and especially combinations of HMD and the AR system with 3D marker contribute to expand HCI application fields in the near future due to the attractive and intuitive hand-motion interactions base on the result of this paper. In this paper, we limited to describe the hand-motion method to control the 3D AR markers and display graphic images on the 3D markers. In the future work, we have a plan to study the method that the 3D marker can be fixed in places such as hanging from the ceiling and multiple persons can see the 3D graphic images from different 360-degrees views by the persons are equipped with HMDs are moving around the 3D marker.

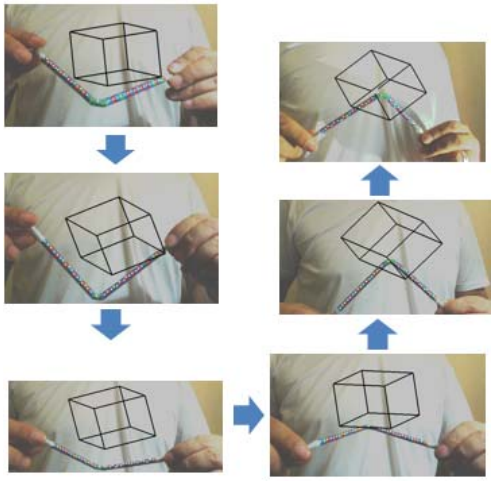


Figure 23: Demonstration of Seamless 360-dgrees Rotations to Display AR Graphics Associated with 3D Marker

Reference

1) Sanni Siltanen, Theory and applications of marker-based augmented reality Espoo 2012. VTT Science 3. 1-198pp
2) ARToolKit http://artoolkit.sourceforge.net/ (Accessed June 22, 2014)
3) Wouter Alexander de Landgraaf, Interaction between users and Augmented Reality systems: Human-Computer Interaction of the future an essay for HCI 2003/2004, Vrije Universiteit Amsterdam Stdnr: 1256033 1-18pp
4) Mark Fiola, ARTag, An Improved AR Marker System Based on ARTookKit, National Research Council Canada, July 2004 1-36pp
5) Kimura et al. United States Patent No: US 8,113,432 B2 Feb.14,2012
6) B-Core Inc. Colorbit Reference May 2005. www.colorbit.jp/en/wp-content/uploads/2010/05/colorbit_reference_EN2.pdf (Accessed in June 22, 2014)
Colotbit is the trademark of B-Core Inc.
7) Arturo De la Escalera, Jose María Armingol Automatic Chessboard Detection for Intrinsic and Extrinsic Camera Parameter Calibration Special Issue Instrumentation, Signal Treatment and Uncertainty Estimation in Sensors pp. 2027-2044 March 15, 2010
8) Derek Hoiem, Projective Geometry and Camera Models, Computer Vision CS 543 / ECE 549 University of Illinois Jan, 20, 2011 1-63pp
9) Gary Bradski, Adrian Kaehler Learning OpenCV Computer Vision with the OpenCV Library By Publisher: O'Reilly Media Released: September 2008
10) The OpenCV Reference Manual Release 2.3 October 15, 2011 www.opencv.org (Accessed in June, 22, 2014)
11) Rebecca M. Brannon Rotation: A review of useful theorems involving proper orthogonal matrices reference to three-dimensional physical space. pp. 1-167 May 9,2002
12) OpenGL http://www.opengl.org/ (Accessed in July 21, 2014)
13) Dr. Joachim Linkemann, Björn Weber, Global Shutter, Rolling Shutter - Functionality and Characteristics of Two Exposure Methods (Shutter Variants) pp. 1-6, January 2014