QR-code based Online Robot Augmented Reality System for Education

Jung Pil Park
The School of Computer Science and Engineering,
Kyungpook National University
1370 Sankyuk-dong, Buk-gu
Daegu 702-701, South Korea
jppark@vr.knu.ac.kr

Min Woo Park
The School of Computer Science and Engineering,
Kyungpook National University
1370 Sankyuk-dong, Buk-gu
Daegu 702-701, South Korea
mwpark@vr.knu.ac.kr

Soon Ki Jung †
The School of Computer Science and Engineering,
Kyungpook National University
1370 Sankyuk-dong, Buk-gu
Daegu 702-701, South Korea
skjung@knu.ac.kr

ABSTRACT
In this paper, we propose an efficient, augmented reality system for education utilizing scenarios acquired from the metadata of quick response (QR) code. Our system enables interaction using an assembled robot in a virtual environment. First, we build a virtual environment around the robot according to the metadata. In addition, we create and manage a wide variety of content using metadata. We estimate the camera pose by using both marker-based and markerless methods. Lastly, we use a Bluetooth module and metadata of QR codes to interact with objects in the virtual environment.

Categories and Subject Descriptors
H.5.1 [Information Interfaces and presentation]: Multimedia Information Systems – artificial, augmented, and virtual realities

General Terms
Human Factors, Algorithms

† Dr. Jung is a corresponding author.

Keywords
Augmented reality, QR code, Metadata, Feature tracking, Marker tracking, Bluetooth

1. INTRODUCTION
1.1 Background
Modern society has developed with computers and various machines, providing a ubiquitous environment accessible everywhere and from anywhere using various multimedia devices and networked community resources. We use a lot of multimedia content from various fields in the environment, along with augmented reality (AR), which has been continually researched for about 40 years and which takes skill when combining virtual objects with the real environment. In particular, the augmented reality technique offers good effects in education and industry research. In the field of industry, we always use the augmented three-dimensional (3D) object to visualize information in advertisements, cards, and special information from networks in real time. On one hand, we can use medical devices or various items in a virtual environment for games and media. Virtual reality is a computer-simulated environment that can simulate a physical presence in both real and imagined worlds. At the same time, we can experience spatial and temporal virtual environments and situations that are the same as the real world using our five senses. And we can interact with objects in virtual environments using actual devices and instructions.

Mixed reality is a combination of augmented reality and virtual reality. When we use mixed reality, we interact with augmented 3D virtual objects in a real world that is the same as virtual reality. Therefore, through immersion, we can get used it. Through this technique, users can engage in more efficient study for their education, and the latest information or knowledge can be added to the contents of the scenarios by using networks.

We propose education combining a quick response (QR) code–based assembly robot and a virtual environment. QR code attached to the robot is like saving data in a 2D barcode. We build the virtual environment using metadata of the QR code. This paper presents a novel education system using interaction between users and the environment, controlling the robot via Bluetooth with a virtual environment built from metadata scenarios.

1.2 Related Work
Research into augmented reality began 40 years ago, and developed very quickly, especially in the field of education. Research by Mariano demonstrated a variety of multimedia content according to the age of the learner. Therefore, it was a very efficient way for learners to study [1]. However, repeated study decreases efficiency when simple information is visualized. Therefore, to solve this problem, researchers have many study methods for interaction between augmented reality and users. Research by Kaiser shows interaction with content by simply using touch by users [2]. And Hagbi developed interaction with various sketches of users and makes a virtual object of content in

SAC'14, March 24-28, 2014, Gyeongju, Korea.

http://dx.doi.org/10.1145/2554850.2555038

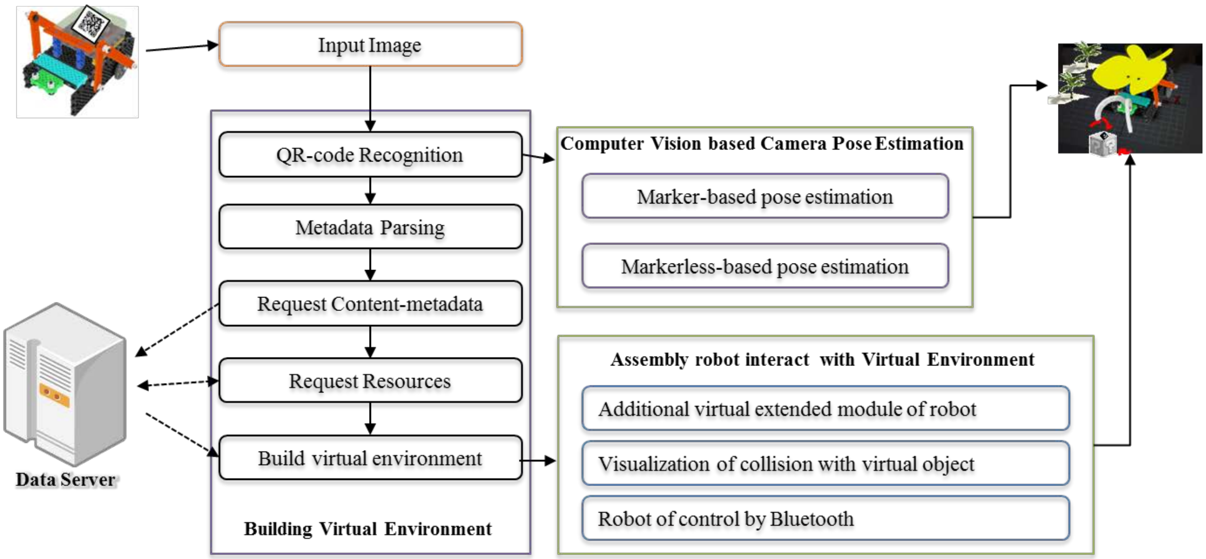

Figure 1. Overview of QR code based robot AR system for education.

real time. Accordingly, we can acquire a lot of information by using interaction with virtual objects [3].

Research has studied marker-based camera pose estimation and markerless camera pose estimation, especially to detect the position of markers. And the goal of tracking is being developed to enable augmented reality [4, 5, 6].

Research by Korah extracts a star-shaped object to estimate camera pose [7]. Research by Hizer offers another method: line-segmentation–based detection of a rectangle. Kan performed a study for augmented reality on the marker of a QR code. It acquires the direction of the marker from a recognized position pattern in the QR code. On one hand, research is being developed into feature-based camera pose estimation methods for augmented reality except in previously mentioned methods. Research by Simon and Taylor computed homography by feature extraction and matching from frames [9, 10]. And other research groups are developing studies to provide an augmented reality library to developers [11, 12].

Developed research into augmented reality is being studied in order to show various types of content using QR code. That research compiled metadata of the QR code in order to use the capacity of mass. And it visualized content information similar to the TAG of metadata for users. Each field of augmented reality has seen actively conducted research. However, for interaction with various types of content for augmented reality, there is a lack of research. Therefore, we propose a method that interacts with a virtual environment using a QR code–based assembly robot.

This paper is organized as follows. Section 2 is overview of proposed system. Section 3 describes the interaction between assembly robot and virtual environment. Section 4 describe metadata of QR code and section 5 describe pose estimation by using image processing and shown the experiments. Section 6 describes conclusion and future works.

2. SYSTEM OVERVIEW

In this paper, we present a QR code–based assembly robot for augmented reality and an overview of a virtual environment system (see Figure 1). This system comprises building a virtual environment and the interaction of an assembly robot with pose estimation. First, we recognize QR code from the camera input stream, then we analyze the metadata and request content metadata. The proposed system is received content resources from a data server to build a virtual environment around the robot and visualize an additional virtual extension module of a robot. Next, we implement initialization and tracking for the assembly robot to estimate the pose of the robot by using a marker-based method. The marker-based method implements QR code pose estimation very quickly, but won't work when the QR code marker disappears partially or on the whole. When pose estimation fails, we compute the homography using a feature extraction and matching method between two frames. Thus, multimedia information is augmented on 3D world coordinates according to the acquired scenario of content metadata from an online server after completion of the pose estimation. And in order to interact with the virtual environment, the assembly robot implements a Bluetooth module. We can control the movement of the assembly robot by using a Bluetooth module if the assembly robot contacts objects in the virtual environment.

3. INTERACTION SCENARIO

In this section, we describe a scenario with an interactive assembly robot in a virtual environment (see Figure 2). It is a scenario of the life of a dancing butterfly, and we can control the butterfly by using the assembly robot.

First, we visualize a virtual flower and extend a virtual butterfly module on an assembly robot; the other objects in the forest will be added after we recognize the QR code and analyze the metadata. Therefore, the virtual butterfly is, according to the scenario from the metadata, searching for honey on a virtual flower. At the same time, the user can control the assembly robot by using the Bluetooth module. And we can estimate the 3D position of the QR code marker on the assembly robot from the camera.

Thus, if the estimated distance between the assembly robot and the virtual flower is less than a minimum range, we determine that contact has occurred. As a result, this enables many effects using Extensible 3D (X3D) and interacts with the virtual flower by scaling up. And we provide virtual animation of the virtual butterfly on a robot with an additional module. Therefore, we show that a smart environment gives the user a better experience than immersion with an assembly robot in an education scenario under mixed reality.

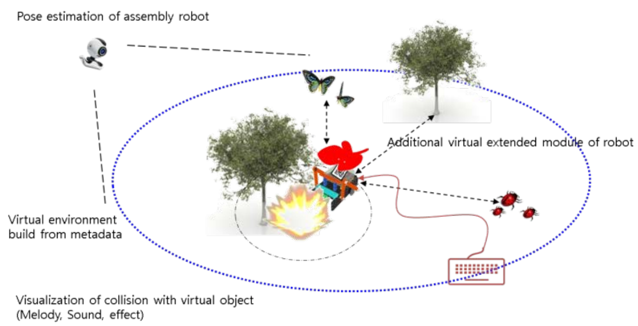

Figure 2. Interactive assembly robot with virtual environment.

4. QR-CODE BASED METADATA

In this section, we describe the method of processing and analyzing the QR code data from the input stream. Section 4.1 describes QR code metadata, and Section 4.2 describes the content of metadata acquired from the QR code. Section 4.3 describes the authoring method for content by using metadata.

4.1 QR Code Metadata

Because QR code can include a various information in the marker, we use it to make various scenarios. Also, it is very easy to manage complicated QR code data. Metadata in this system includes two parts: the metadata of the QR code, and the metadata of detailed content. Metadata also implements management of updated information in the QR code and the content.

Table 1. QR code metadata

TAG	Attribute	Description
Information	-	Start/End
Subject	Type	Contents type
URL	URL	URL of detailed contents metadata
Name	String	Name of Contents
Date	Date	Created date
From	String	author

QR code metadata consists of a creation date, the author, the content type, and a URL path for downloading the content's detailed information from a data server (see Table 1). It uses extensible markup language (XML) to manage the QR code metadata. XML was proposed in 1996 by the World Wide Web Consortium. Therefore, it provides uniform standardization for sending structured documents on the Web. QR code metadata with an XML structure is used by our system to separate the various content classifications. Thus, we acquire content information for visualization by using the network from the path of the detailed metadata.

4.2 Contents Metadata

We visualize content in the virtual environment from the content of detailed metadata loaded from the data server through QR code metadata. At the same time, content detail uses JavaScript object notation (JSON), which is usually used in XML.

Table 2. Contents of detail metadata

TAG	Attribute	Description
Model	URL	Additional Robot's 3D Module URL
Name	String[]	Marker Name of 3D Model
Resources	String[]	Resource Information of 3D Model
Texture	String[]	Texture of 3D Model
Animation	Boolean	Animation state
Material	URL	Material value of 3D Model

JSON decreases packet capacity for data transmission because it has less information than XML. Additionally, JSON is easier to read and write than XML. It is efficient for transmitting the meaning of variance in the programming language because it is saved in an array of data structure. Therefore, we can control a lot of resources with it. We propose metadata for the content detail, as shown in Table 2.

4.3 Creation Content using Metadata

Our system for authoring content consists of making QR code, setting the data server, and making metadata. We describe how to make QR code by using XML in this section.

This system proposes insertion of an online address of the metadata in the QR code because it reduces the size of the QR code text type. We add secret code to distinguish other types of QR code. Therefore, we use the term QRAR with the on-line URL for the metadata, which makes it the same size (width and height) for versatility in the QR code. Finally, we generate a QR marker applied to 25mm of the outline and margin for efficient image processing–based tracking.

Table 3. QR code metadata

```
<?xml version="1.2" encoding="utf-8"?>
<INFORMATION>
<SUBJECT>3DModel</SUBJECT>
<NAME>butterfly</NAME>
<URL>http://vr.knu.ac.kr/~jppark/Contents/Information
/butterfly.txt</URL>
<DATE>2013-03-01</DATE>
<FROM>jppark</FROM>
</INFORMATION>
```

We show example code of QR code metadata to load a butterfly contents metadata from data server (see Table 3). When our system is loaded the butterfly contents metadata, the system can parse the XML from contents metadata to visualize the butterfly object. Table 4 shows an example of the contents metadata to visualize the butterfly object.

Table 4. Contents metadata


```
{
"Model":"http://vr.knu.ac.kr/~jppark/Contents/3DModel/butterfly/butterfly.3ds",
"MarkerName": [    "tree1.patt"    "tree2.patt"   ],
"MarkerResources":
["http://vr.knu.ac.kr/~jppark/Contents/3DModel/butterfly/tree1.3ds"
"http://vr.knu.ac.kr/~jppark/Contents/3DModel/butterfly/tree2.3ds" ],
"MarkerTexture":
["http://vr.knu.ac.kr/~jppark/Contents/3DModel/butterfly/tree1.png"
"http://vr.knu.ac.kr/~jppark/Contents/3DModel/butterfly/tree2.png" ],
"Material":"http://vr.knu.ac.kr/~jppark/Contents/3DModel/butterfly/butterfly.material",
"Animation":"false"
}
```


The content to visualize the butterfly module acquires resources from the data server using TAG in the content metadata (e.g., Model, MarkerResources, MarkerTexture). Animation of the TAG exists for users' interaction with the virtual environment. Therefore, we can adjust the scenario by using detailed content metadata.

5. CAMERA POSE ESTIMATION

This section describes image processing–based camera pose estimation. An overall flowchart for this approach is illustrated in Figure 3. First, we estimate the camera pose from the outline and position of the QR code. And we estimate camera pose to be robust by using homography between frames. If we fail with marker detection, we have a failure of marker detection from the field of view of the camera.

This process consists of two steps: marker-based pose estimation using QR code and feature-based pose estimation.

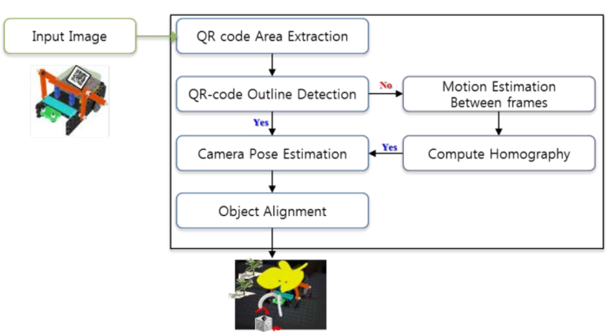

Figure 3. Image processing based camera pose estimation.

5.1 Marker-based Pose Estimation

Markers for augmented reality are very simple for optimization tracking. However, QR code is unsuitable for augmented reality, because it is created in order to save a lot of information. In this section, we describe analysis of the QR code and tracking for pose estimation.

Figure 4. QR code for tracking.

The content of QR code consists of position pattern, direction pattern and data area (see Figure 4). We use position pattern for tracking QR code. And we apply it to the outline of the QR code for more efficient tracking. Depending on the purpose for space recognition of QR code, however, we confuse direction because it is square. Therefore, this method is difficult to implement for augmented reality by using detection of a square. In this paper, we decide on the direction of the QR code by pattern of position and the outline of the QR code.

We calculate 3D coordinates by homography between the input stream and the plane to get the initial pose of the camera. It is a measured marker of position and size in the 3D real world. And we track camera pose via changing camera motion. Camera pose estimation is obtained by decomposing the homography under an assumption as to where QR code exists in the three-dimensional space [6, 15]. The following formula shows the transformation.

$$\begin{aligned} H &= KE = K[r_1, r_2, t], \\ E &= K^{-1}H = [r_1, r_2, t]. \end{aligned} \tag{1}$$

H is homography. K is the intrinsic parameter of the camera. E is the extrinsic parameter of the camera, r_1, r_2 is the rotation vector of the X, Y axes, and t is the translation vector. At the same time, we can calculate the rotation of the Z-axis through the cross product of the X and Y axes.

$$P = K[r_1, r_2, (r_1 \times r_2), t]. \tag{2}$$

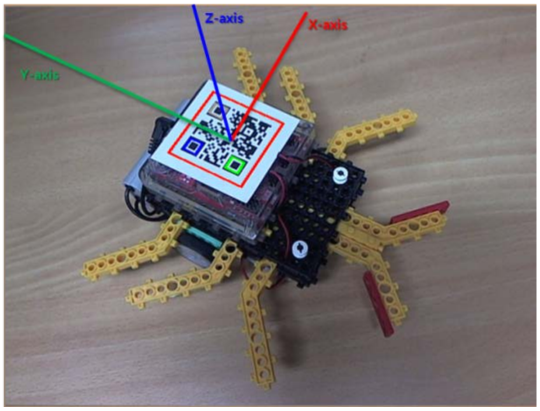

Figure 5. Rectangle of detection and pose estimation.

P is projection of a 3×4 matrix from 3D coordinates to 2D coordinates. Figure 5 shows detection of the rectangle from the QR code and augments the result of the *XYZ*-axes on 3D coordinates.

5.2 Feature-based Camera Pose Estimation

We show the results of fast and exact pose estimation by using a square set on the described QR code from the previous section. However, there is a problem in that the marker is a partially or wholly occluded area when it is apart from the camera.

This paper proposes feature-based camera pose estimation by u sing frames of the input stream when previous methods fail. We c alculated a motion of frames using a sum of squared difference aft er FAST features detection continuously between frames. Feature detection and matching require a lot of computation. Therefore, w e made an image pyramid from the input stream in order to reduce computation. Additionally, we implement feature extraction on a 25% reduction of the original image. The whole image is separate d for feature tracking and matching on an area near the center of t he image. Therefore, we maximize speed of detection and matchin g by selecting a certain feature [16, 17].

We calculate between the frames when feature matching is completed. And we employ RANSAC for robust computation of the homography and elimination of outliers on feature matching. With Formula 3, we calculate the position in 3D world coordinates and image coordinates between QR codes. It is necessary to multiply the homography from sequence frames. Additionally, we apply formulas (1) and (2) for pose estimation by using computed homography.

$$H_c = H_i H_{i-1} \cdots H_0. \quad (3)$$

We show the results of camera pose estimation by computed homography between frame feature detection and matching (see Figure 6).

Figure 6. Pose estimation using homography between frames.

5.3 Contents for Augmented Reality System

We show the content of the results in Figure 7, which builds the virtual environment around the robot with acquired information from the metadata of the QR code. At the same time, we can control the assembly robot through a keyboard by using a Bluetooth module. And the result is a visualized 3D effect towards contact when the estimated distance between virtual objects is smaller than a certain range.

Figure 7. Final Result

6. CONCLUSION

This paper proposes an augmented reality system for education that interacts with a virtual environment by controlling an assembly robot. First, we build a virtual environment from the metadata of QR code. Therefore, we obtained resources, interaction and information about the scenario. We perform estimation of the camera pose by using both marker-based and markerless methods. And we use a Bluetooth module in order to interact with the virtual environment and the robot.

Future work will make various types of content by using the metadata of QR code. And interactions will be developed for a variety of content authoring and creation. We will improve run-time for estimating camera pose more efficiently and will conduct additional studies to make it more quick and robust.

7. ACKNOWLEDGMENTS

This research is supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program (Immersive Game Contents CT Co-Research Center).

8. REFERENCES

[1] Mariano Alcaniz, Manuel Contero, David C. Perez-Lopez and Mario Ortega (2010). Augmented Reality Technology for Education, New Achievements in Technology Education and Development, Safeeullah Soomro

[2] E. Kaiser, A. Olwal, D. McGee, H. Benko, A. Corradini, X. Li, P. Cohen, S. Feiner, "Mutual disambiguation of 3D multimodal interaction in augmented and virtual reality," International conference on Multimodal interfaces, pp. 12-19, 2003.

[3] Hagbi, N. Grasset, Raphael, Begig, Oriel, Billinghurst, Mark, El-Sana, Jihad, "In-Place Sketching for content authoring in augmented reality games," Virtual Reality conference, pp. 91-94, 2010.

[4] R. Azuma, "A Survey of Augmented Reality," Presence-Teleoperators and Virtual Environments, Volume 6, Number 4, pp. 355-385, 1997.

[5] G. Papagiannakis, G. Singh, N. M. Thalmann, "A Survey of Mobile and Wireless Technologies for Augmented Reality Systems," Computer Animation and Virtual Worlds, Volume 19, Issue 1, pp. 3-22, 2008.

[6] V. Teichrieb, JPS do Monte Lima, E. L. Apolinario, TSMC de Farias, M. A. S. Bueno, J. Kelner, Ismael H. F. Santos, "A Survey of Online Monocular Markerless Augmented Reality," International Journal of Modeling and Simulation for The Petroleum Industry, Volume 1, Number 1, 2007.

[7] T. Korah, J. Wither, Y. Tsai, R. Azuma, "Mobile Augmented Reality at the Hollywood Walk of Fame," IEEE Conference of Virtual Reality, pp. 183-186, 2011.

[8] M. Hirzer, "Marker Detection for Augmented Reality Applications," Research Paper at Granz University of Technology, 2008.

[9] G. Simon, A. W. Fitzgibbon, A. Zisserman, "Markerless Tracking using Planar Structures in the Scene,“ IEEE and ACM International Symposium on Augmented Reality, Volume 40, Issue 1, pp. 120-128, 2000.

[10] S. Taylor, E. Rosten, T. Drummond, "Robust Feature Matching in 2.3μs," IEEE CVPR Workshop on Feature Detectors and Descriptors: The State of The Art and Beyond, 2009.

[11] Qualcomm Augmented Reality, https://developer.qualcom m.com/develop/mobile-technologies/augmented-reality

[12] Metaio Mobile AR, http://www.metaio.com/projects/mobile

[13] H.S Yoon, N.Y Park, W.W Lee, Y.K Jang, W.T Woo, "QR Code Data Representation for Mobile Augmented Reality," International AR Standards Meeting, pp. 17-19, 2011.

[14] J.P. Park, M.O. Park, S.G. Soon, "Robust pose estimation for augmented reality system using mobile sensor fusion," Korea Computer Congress, pp. 871-875. 2012.(In Korean)

[15] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge.

[16] Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing, Prentice Hall, ISBN: 0131687288, third edition, 2008.

[17] M. Kim, I. S. Oh, "Estimation of Rapid-motion for Mobile Devices using Temporal Coherence," 23rd International Conference Image and Vision Computing New Zealand, pp. 1-6, 2008.