A Tracking Method for 2D canvas in MR-based interactive painting system

Sandy Martedi[1*], Mai Otsuki[2*], Asako Kimura[2*], Fumihisa Shibata[2*], Maki Sugimoto[1*], and Hideo Saito[1*]

[1*]Keio University, [2*]Ritsumeikan university

ABSTRACT

We have proposed a mixed reality (MR) based painting system. In this paper, we tackle a problem that exists in the conventional method that fully relies on magnetic sensor that is attached on the canvas; the users needed to detach and attach a sensor on the canvas during painting when they want to switch the canvas. Therefore, we aim to automatically detect the shape of the canvas for registration purpose using vision-based tracking method. Using a region detection method such as MSER, we detect and track the shape on the canvas. We then compute the camera pose for virtually overlaying the painting result. Finally, we can generate equivalent results compared to the result of using the sensor.

Keywords: tracking, mixed reality, augmented reality, interaction.

Index Terms: H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities.

1 INTRODUCTION

We have developed the mixed reality (MR) based painting system [7] which enables the user to do a digital painting toward real objects while directly holding the object. It makes digital painting more closely resemble real-world painting, because the brush shaped device provides the similar sensation of an actual painting brush using tip bending.

We have used the magnetic sensors to track the position and orientation of HMD (user's head), brush device, and the canvas (real object). Thus, we needed to detach one sensor and attach it to the other canvas every time, otherwise, prepare magnetic sensors as many as the number of canvases, that can be cumbersome.

To avoid this problem, we attempt to use vision-based tracking instead of physical sensors (Figure 1). We developed a vision-based tracker that utilizes the outline of shapes that is printed in the canvas or the shape of the canvas itself that is robust against occlusions from the brush device and user's hands.

2 RELATED WORK

A virtual painting system with the tangible interaction using a projector to paint real objects is initially introduced in Dynamic shader lamps [1]. Recently, vision-based virtual painting interaction has also been demonstrated

1* sandy@hvrl.ics.keio.ac.jp

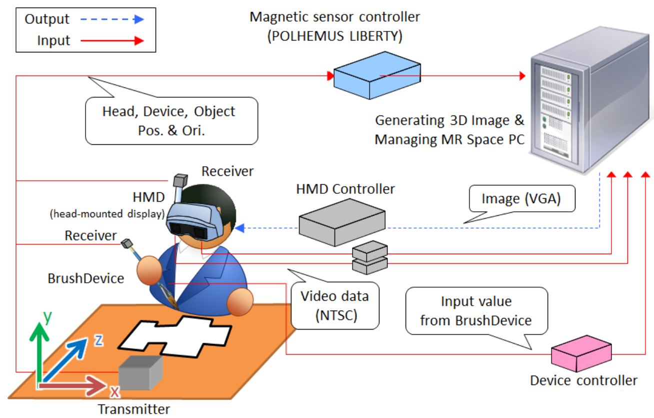


Figure 1: System configuration

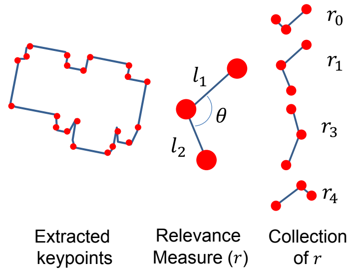
using Kinect Fusion [8]. Users can draw and paint virtually over the reconstructed scene using Kinect. However, in the case using Kinect Fusion, only 3D surfaces can be identified and used as a canvas. For vision-based tracking, although ARToolKit [5] and, ARToolKitPlus [9] are widely used, in our painting system, such marker interrupts user's creativity. In this paper, we use shapes as the cue for tracking instead of marker. Similarly, some systems are capable to recognize hand-drawn 2D shapes in real time [2] [4]. Donoser et al. also showed impressive results in shape tracking using shape concavity [3]. In this paper, we aim to develop a generic representation of shapes regardless its concavity.

3 SETUP

The system configuration is shown in Figure 1. We use a PC that runs on Microsoft Windows XP OS, Intel Core i7 2600 CPU, 3062 MB of RAM and NVIDIA GeForce GTS 250. We use a binocular video see-through HMD (Canon VH-2002). The HMD is connected to a video capture card (ViewCast Osprey-440). The positions and orientations of the HMD and BrushDevice are tracked using Polhemus LIBERTY, a six-DOF magnetic sensor-based tracking system. A transmitter is also used as a reference point for the sensors. The brush device is connected to the main PC through an input/output (I/O) box that retrieves data from the devices and sends them to the main PC. All the codes in the system are written in C++/CLI in .NET Framework. We used OpenGL and the OpenGLUT for the graphics API.

4 METHOD

Firstly, the region of the canvas and its contour are extracted using MSER (maximally stable external regions) detector [6]. We then compute the relevance measure r that is defined as



Hash id	(reg id, pt id)
:	
$f(r_0, \boldsymbol{r_1}, r_2)$	(0,**1**)
:	:
$f(r_1, \boldsymbol{r_2}, r_3)$	(0,**2**)
:	:
$f(r_{n-1}, \boldsymbol{r_0}, r_1)$	(1,**0**)
:	

Hash table

Figure 2: Shape registration

Figure 3: Painting results

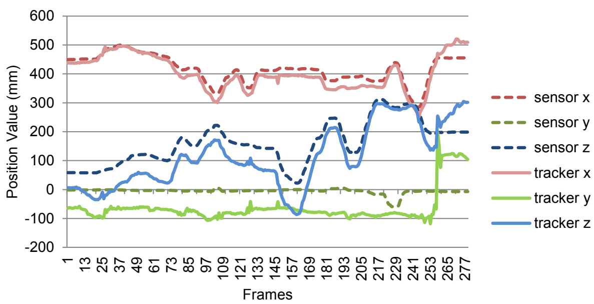


Figure 4: Comparison of the magnetic sensor and the visual tracker using proposed method

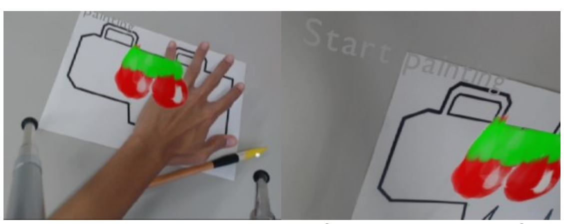


Figure 5: The result in the case of partial occlusion (left), or going out of frame (right).

$$r = \frac{\theta l_1 l_2}{{l_1}^2 + {l_2}^2} \quad (1)$$

where l_1 and l_2 are the length of two connected lines and θ is the angle between l_1 and l_2. Sequences of r are then stored into a hash table. We use three neighboring r values to represent a keypoint of a shape (r_0, r_1, r_2 represents keypoints with id=1). During the on-line shape registration (see Figure 2), the keypoints and sequences of r are computed from the input frame and the corresponding tuple (region id, keypoint id) is looked up in the hash table. This process yields keypoints correspondences between a shape captured in the camera and the template shape. The camera pose is then estimated using homography using keypoints correspondences. The last step is unifying the camera coordinate system and the sensor coordinate system by multiplying the view matrix of the HMD and the rotation and translation matrix from the homography. Finally, the virtual painting can be superimposed as illustrated in Figure 3.

5 Evaluation

We attached a magnetic sensor on the origin of the canvas and we plotted the positions that obtained from the sensor and the visual tracker (Figure 4). The position of the sensor was shifted in the y axis and can be compensated by applying a local transformation on the camera coordinate system. We also recorded the processing time for a shape containing 33 points. The tracking and the painting requires 23.74 msec and 11.39 msec respectively. The processing time is sufficient for real time purpose. The drawback of our system is the tracking will fail due to extreme occlusions, however, in the case of partial occlusions that were occurred by hand or brush device, and a part of marker goes out of the frame can be tolerated (Figure 5)

6 Conclusion

We proposed a MR-based painting system that uses a visual tracker as the canvas. By using a visual tracker, an object can be detected without attaching sensor. We can switch the various canvases freely without detaching and attaching the physical sensors. We achieved to realize real-time and robust against occlusion tracking method.

In this paper, we used the asymmetry 2D object as the tracking marker; the canvas, however, in the future, we extend our method to 3D objects and track and identify the multiple canvases simultaneously. In addition, we targeted to track one canvas in this research; we are planning to implement the ID to each canvas in the case using multiple canvases at the same time.

References

[1] Bandyopadhyay, D. Raskar, R., and Fuchs, H. "Dynamic shader lamps: Painting on movable objects." Proc. ISMAR 2001, pp. 207 - 216. 2001.

[2] Bergig, O., Hagbi, N., El-Sana, J., and Billinghurst, M. "In-place 3d sketching for authoring and augmenting mechanical systems.". Proc. ISMAR 2009, pp. 87 - 94, 2009.

[3] Donoser, M., Kontschieder, P., and Bischof, H. "Robust planar target tracking and pose estimation from a single concavity." Proc. ISMAR 2011, pp. 9 - 15, 2011.

[4] Hagbi, N., Bergig, O., El-Sana, J., and Billinghurst, M. "Shape recognition and pose estimation for mobile augmented reality." IEEE Trans. Visualization and Computer Graphics, Vol. 17, No. 10, pp. 1369 - 1379, 2009.

[5] Kato, H. and Billinghurst, M. "Marker tracking and HMD calibration for a video-based augmented reality conferencing system," Proc. IWAR (1999), pp. 85 - 94.

[6] Nistér, D, and Stewénius, H. "Linear time maximally stable extremal regions." *Computer Vision–ECCV 2008*. Springer Berlin Heidelberg, pp. 183-196, 2008.

[7] Otsuki, M, Sugihara, K., Kimura A., Shibata, F., and Tamura, H. "MAI painting brush: An interactive device that realizes the feeling of real painting." Proc. UIST2010, pp. 97 - 100. 2010.

[8] Shahram, I., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, D., Kohli, P., Jamie Shotton et al. "KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera." Proc. UIST 2011, pp. 559 - 568, 2011.

[9] Wagner, D. and Schmalstieg, D. "ARToolkitPlus for pose tracking on mobile devices," Proc. CVWW'07, pp. 139 - 146. 2007.