Spatial Augmented Reality by Using Depth-Based Object Tracking

Daisuke KOBAYASHI, Naoki HASHIMOTO
Graduate School of Informatics and Engineering, The University of Electro-Communications*

Figure 1: *Results of our proposed object tracking for interactive SAR.*

1 Introduction

Recently, spatial augmented reality (SAR) is actively used for attractive projection effects by changing appearance of real objects, like a projection mapping. SAR gives image contents a presence and interactivity extracted from real objects on which the images are projected. So SAR has a high possibility of achieving an novel interaction between computers and us. However, in actual SAR, all objects' shape and posture have to be previously measured, and be static during the projection. This limitation reduces interactive applications of SAR, and also obstructs the wide spread of SAR techniques into our everyday life. Although, in previous methods, a position sensor[Kondo et al. 2008] and camera-based object tracking was proposed, they could not be performed properly on SAR with a difficulty of sensor embedding, and influence of projected light.

2 Our proposal

We use a contactless depth sensor for position tracking of moving targets. This sensor has no influence on the target's appearance, and captured depth value is not affected by projected lights. Then by estimating targets' position and posture with that tracking results, appropriate images for changing their appearance are projected for achieving interactive SAR.

Our tracking technique is composed of off-line and on-line processes. In the off-line process, we measure the whole 3D shape of a target object as a reference by using the depth sensor. And we previously detect feature points for posture estimation used in the next on-line process. In our implementation, we used fast point feature histogram (FPFH). In the on-line process, at first, we estimate initial position and posture of the target by using feature point matching. And we further apply ICP algorithm for much more precise estimation. After this initial estimation, we can tract the target by continuously using ICP algorhtim in theory, like recent trends.

*e-mail: {kobayashi@ims., naoki@}cs.uec.ac.jp

SIGGRAPH 2014, August 10 – 14, 2014, Vancouver, British Columbia, Canada.

ACM 978-1-4503-2958-3/14/08

However, in practical tracking, that approach can not perform successfully, because ICP algorithm supposes that targets are placed closely to their reference. In the interactive SAR system, users can move the target object freely, and then, the position and posture of the target are widely changed in a measuring interval. As the result, ICP algorithm takes lots of computational time for its repetitive processing, and sometimes results wrong posture.

Therefore, in our proposal, we use a particle filter for rough and rapid estimation before ICP algorithm. In the particle filter, each particle represents possible position and posture, and is initialized according to simple uniform linear and circular motion models, including a few random factors. The depth of the target is measured as a point cloud, and compared with all of the candidates estimated by the particles. In that comparison, for each candidate, we calculate the sum of the difference between the measured point and the nearest point of that candidate, as a weight factor for the particle representing that candidate. And then, with the weighted average of the particles, we can roughly estimate rotation and translation between present and previous target position. In this condition, the actual target and that estimated are closely placed, so ICP algorithm performs very well for highly accurate and fast tracking. After that, we sample again the particles representing possible position at the next tracking, from the highly weighted particles.

We implemented a simple SAR system with our proposed method, as shown in Figure 1. In this system, we use a mannequin head of styrene foam as a projection target, and also use a Kinect for measuring its depth. Users can move the mannequin freely, and that position and posture are correctly obtained at 15 fps, only with a CPU (Intel Core i7-4770K). Based on that result, we can project a face texture of the mannequin naturally. In additional implementation, we can interactively edit the face like a make-up or hair-design simulator. Much more faster tracking will be achieved by using GPU implementation and machine learning-based prediction for reducing the particles, now we are developing.

This work was supported by JSPS KAKENHI Grant Number 25330230.

References

KONDO, D., SHIWAKU, Y., AND KIJIMA, R. 2008. Free form projection display and application. *Proc. of the 5th ACM/IEEE International Workshop on Projector camera systems (PROCAMS)*, 8:1–8:2.