Dynamic Spatial Augmented Reality with a Single IR Camera

Naoki Hashimoto, Daisuke Kobayashi
The University of Electro-Communications*

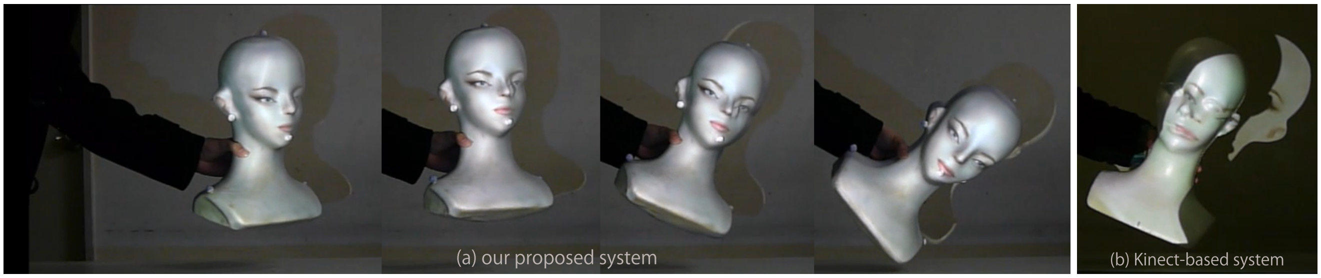

Figure 1: *(a) A result of our dynamic SAR for a moving mannequin-head. Spherical markers are used to obtain Ground Truth with an optical tracker. They are not auxiliary factors, but rather disincentives for our method.* *(b) The same situation with our Kinect-based system.*

Abstract

We propose a dynamic spatial augmented reality (SAR) system with effective machine learning techniques and edge-based object tracking. Real-time 3D pose estimation is the significant problem of projecting images on moving objects. However, camera-based feature detection is difficult, because most targets have a textureless surface. Image projection and projected images also interfere with detection. Obtaining 3D shape information with stereo-paired cameras [Resch et al. 2016] is still a time-consuming process, and using a depth sensor with IR [Koizumi et al. 2015] is still unstable and have a fatal time-delay for the dynamic SAR. Therefore, we quickly and robustly estimate the 3D pose of the target objects by using effective machine learning with IR images. And by the combined use of high-speed edge-based object tracking, we realize a stable and low-delay SAR for moving objects.

Keywords: spatial augmented reality, machine learning, tracking

Concepts: •**Human-centered computing** → **Mixed / augmented reality;** *Virtual reality;*

1 3D pose estimation with machine learning

Our 3D pose estimation is based on a machine learning technique with Hough Forests, which can provide robust detection for noisy and occluded objects. These features are indispensable for dynamic SAR, where targets are often handled and are portable. We generate training image patches by rendering 3D models of target objects from various viewpoints. In order to detect small pose changes, which are significant for precise image projection, we accumulate tiny variations of oriented gradients with the pose changes in each of the patches. These perspectively-robust patches reduce the total amount of the patches and perform fast and precise estimation.

In addition, we use a Random Ferns classifier for fast voting of the patch classification process. Because common Random Ferns are insufficient for precise 3D pose detection, we propose a two-layered Random Ferns classifier. The first layer is for rough classification of the pose which is divided into 8 directions. And the second layer is for precise detection of the pose in the direction detected at the first layer. This architecture reduces the estimation time of the classifier and enlarges recognition accuracy.

The classifier is trained with IR images (2040 positive images and 40 negative images) which reduce the color effect of the target surface, and achieves 96 % recognition accuracy for position estimation, and 76 % for rotation estimation. These are fully compensated with the following tracking, and used as the initial position.

2 Edge-based object tracking

We use edge-based object tracking with a high-speed IR camera for low-delay and smooth image projection in parallel to the pose estimation. A general projector has fatal 2 or 3 frame delays, which cause a large gap between the target and the projected image. Therefore, we adopted motion prediction by using high-speed edge-based tracking with the high-speed IR camera (170fps). The latest pose is tracked by minimizing the edge-based gap from the last frame within 12msec. The estimated pose with the machine learning classifier is mutually compensated with this tracking.

We implemented a dynamic SAR demonstration with a moving mannequin-head as shown in Fig.1. An expression texture is projected with a standard 60 fps projector. The length of the motion in Fig.1 is about 2 seconds. Compared with the system using Kinect [Koizumi et al. 2015], our system can obviously reduce the misalignment of projected images. It also achieves low-delay and responsive SAR on the moving target. This work was supported by JSPS KAKENHI Grant Number 16K00267.

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

SIGGRAPH '16, July 24-28, 2016, Anaheim, CA,
ISBN: 978-1-4503-4371-8/16/07
DOI: http://dx.doi.org/10.1145/2945078.2945083

References

KOIZUMI, R., KOBAYASHI, D., AND HASHIMOTO, N. 2015. Acceleration of dynamic spatial augmented reality system. *CyberWorlds2015*, 50–53.

RESCH, C., KEITLER, P., AND KLINKER, G. 2016. Sticky projections - a new approach to interactive shader lamp tracking. *ISMAR2014*, 151–156.