Context-assisted Tracking for Dynamic Target Augmentation

Youngmin Park and Woontack Woo

I. Introduction

Context-awareness in computer vision is expected to bring improvements of computation efficiency by filtering out search targets based on context. The problem is to analyze which contexts are useful for a given problem and determine how to exploit them.

Previous researches showed that spatial context is used effectively. One famous application is to use the proximity of a moving camera to exhibitions in a museum [1]. In that application, only the targets those are near to the camera is considered in the recognition. The effect is that the recognition is performed much faster and the resulting accuracy is improved. Another application of spatial context is to re-use a same marker in different and separated space [2]. By differentiating same markers in a different space, markers can be re-used several times. This reduces the efforts to generate numerous markers and prevents misrecognition arising from similar markers.

We propose a novel method of using temporal context for vision-based tracking. Temporal context describes the state of a dynamically changing target and gives a pre-knowledge of the appearance of the target. In that way, the temporal context works as a detailed filter after that of spatial. Possible applications include tracking and augmenting a video advertisement on a screen.

II. Context-assisted Tracking

A. Context Acquisition

We assume that the spatial context is obtained with user positioning systems and the temporal context is generated by the computer embedded in the target. For example, the target object can display a video sequence and generate timestamp at each frame. The contexts can be shared between the player and tracker through a middleware for ubiquitous computing environment such as [3].

This research is funded by ETRI OCR and the CTI development project of KOCCA, MCT in Korea.

Youngmin Park is with the Gwangju Institute of Science and Technology, Gwangju 500-712, S.Korea (corresponding author to provide phone: +82-62-970-3157; fax: +92-62-970-2204; e-mail: ypark@gist.ac.kr).

Woontack Woo is with the Gwangju Institute of Science and Technology, Gwangju 500-712, S.Korea (e-mail: wwoo@gist.ac.kr).

B. Description Hierarchy

The description hierarchy contains spatial division, several target objects belonging to each space, and temporal information. A wide space and target objects in it can be divided into and described as a tree hierarchy. In the lowest level of the hierarchy, there are specific target objects. Then, temporal context places in each target object to describe the state of each.

C. Tracking Changing Target

For the implementation, we used a library which supports real-time detection of a textured surface having natural features [4]. It works in real-time for a single static surface based on off-line training. However, it might not applicable to detect a video sequence on a surface since it requires matching with decades or hundreds of frames according to the length of the video.

The spatial and temporal contexts are exploited as filtering criteria so that most matching candidate is excluded. Depending on spatial context, at first, only targets within a specific range is selected. Orientation difference between the user and the target is also applied because the user is not likely to look at an object behind. Then, the temporal context support to estimate which video frame is displayed at each moment. Thus, matching target is narrowed down to few frames from decades or hundreds frames.

III. Conclusion

A novel method which uses temporal as well as spatial context for vision-based tracking is presented. The proposed method is efficient when the target is not static but dynamically changing its appearance along time. Possible applications include tracking and augmenting a video advertisement on a screen.

References

[1] Bruns, E., Brombach, B., Zeidler, T. and Bimber, O., "Enabling Mobile Phones To Support Large-Scale Museum Guidance", in IEEE Multimedia, 2007

[2] M. Kalkusch, T. Lidy, M. Knapp, G. Reitmayr, H. Kaufmann, D. Schmalstieg, "Structured Visual Markers for Indoor Pathfinding", in Proceedings of the First IEEE International Workshop on ARToolKit (ART02), 2002

[3] Yoosoo Oh, Woontack Woo, "How to build a Context-aware Architecture for Ubiquitous VR", in ISUVR2007 (submitted)

[4] BazAR, url(http://cvlab.epfl.ch/software/bazar/index.php)