Codebook-Based Background Subtraction to Generate Photorealistic Avatars in a Walkthrough Simulator

Anjin Park¹, Keechul Jung², and Takeshi Kurata¹

¹Center for Service Research, National Institute of Advanced Industrial Science and Technology, Japan
²Department of Digital Media, Soongsil University, Korea
{anjin.park,t.kurata}@aist.go.jp, kcjung@ssu.ac.kr

Abstract. Foregrounds extracted from the background, which are intended to be used as photorealistic avatars for simulators in a variety of virtual worlds, should satisfy the following four requirements: 1) real-time implementation, 2) memory minimization, 3) reduced noise, and 4) clean boundaries. Accordingly, the present paper proposes a codebook-based Markov Random Field (MRF) model for background subtraction that satisfies these requirements. In the proposed method, a codebook-based approach is used for real-time implementation and memory minimization, and an edge-preserving MRF model is used to eliminate noise and clarify boundaries. The MRF model requires probabilistic measurements to estimate the likelihood term, but the codebook-based approach does not use any probabilities to subtract the backgrounds. Therefore, the proposed method estimates the probabilities of each codeword in the codebook using an online mixture of Gaussians (MoG), and then MAP-MRF (MAP: Maximum A-Posteriori) approaches using a graph-cuts method are used to subtract the background. In experiments, the proposed method showed better performance than MoG-based and codebook-based methods on the Microsoft DataSet and was found to be suitable for generating photorealistic avatars.

1 Introduction

Before constructing buildings, details such as the suitability of the floor layout and navigation signs and whether users will feel comfortable in the buildings should be considered. Virtual reality techniques are used to investigate virtual structures in detail. However, since information is generally displayed through monitors and a keyboard or mouse is used to navigate the structure, it is difficult to evaluate the relationship between the details of the structure and the sense of absolute direction of the user. Therefore, we are developing simulation environment, referred as to Walk-through Simulator (WTS), in order to enable subjects to navigate virtual constructions from the perspective of the customer. The hexahedral-shaped device shown in the center of Fig. 1 is the WTS. The virtual building is displayed inside the device using multi-projectors, as shown on the right-hand side of Fig. 1.

In some buildings, such as public institutions, guides provide instructions to customers or visitors to help them reach their destination. In the virtual building, guides are displayed as avatars. In the present study, we use a Photorealistic Avatar, in which the appearance of an actual person is used as CG texture, as a guide. The
The present paper focuses on displaying the photorealistic avatar in virtual buildings. The image of the person that is used to create the photorealistic avatar is extracted by the camera in front of a remote computer connected via a network with the WTS, as shown in Fig. 1(A), and the photorealistic avatar is displayed in a fixed location in the virtual world inside the simulator, as shown in Fig. 1(B).

Fig. 1. Schematic diagram of the WTS: (A) photorealistic avatar extracted from the modeled background, (B) photorealistic avatar integrated into the virtual building, and (C) multi-projectors

In the present paper, it is assumed that the moving foreground in front of the fixed camera of the remote computer is an individual whose image will be used to generate the photorealistic avatar, and the actual person who will take a role of the guide can stand on any places, for example, a room with complex backgrounds. The present paper uses background subtraction to extract the appearance of the guide from images captured by a camera. There are four requirements for the background subtraction: 1) extraction must be performed in real time, 2) memory consumption must be limited, 3) the image must be extracted with little noise, and 4) the boundaries of the avatar must be clear.

The proposed method integrates a codebook-based approach, which helps to perform extraction in real time and reduces the required memory, and an edge-preserving MRF model, which can eliminate noise and generate clear boundaries. Although the codebook-based algorithm [10] can model an adaptive and compact background over a long period of time with limited memory, it cannot be used as the likelihood term in the edge-preserving MRF, because the similarity (rather than the probability) is used to compare input pixels with the modeled background. Therefore, online mixture of Gaussians (MoG) is used to estimate the probabilities for all codewords in the codebook. In addition, the proposed method models the prior term using the codebook-based method in order to substantially reduce extraction errors caused by high-contrast edges in cluttered backgrounds, thereby reducing errors on the boundaries of extracted foregrounds.