A New Framework for Background Subtraction Using Multiple Cues

SeungJong Noh and Moongu Jeon

School of Information and Communication,
Gwangju Institute of Science and Technology
{ivyun,mgjeon}@gist.ac.kr

Abstract. In this work, to effectively detect moving objects in a fixed camera scene, we propose a novel background subtraction framework employing diverse cues: pixel texture, pixel color and region appearance. The texture information of the scene is clustered by the conventional codebook based background modeling technique, and utilized to detect initial foreground regions. In this process, we employ a new texture operator namely, scene adaptive local binary pattern (SALBP) that provides more consistent and accurate texture-code generation by applying scene adaptive multiple thresholds. Background statistics of the color cues are also modeled by the codebook scheme and employed to refine the texture-based detection results by integrating color and texture characteristics. Finally, appearance of each refined foreground blob is verified by measuring the partial directed Hausdorff distance between the shape of a blob boundary and the edge-map of the corresponding sub-image region in the input frame. The proposed method is compared with other state-of-the-art background subtraction techniques and its results demonstrate that our method outperforms others for complicated environments in video surveillance applications.

1 Introduction

One of the most important tasks in automated surveillance systems is designing an accurate algorithm to detect moving objects. In general stationary camera based systems, moving objects can be detected by a background subtraction method [1–4] that consists of several modules among which the background modeling most greatly influences the detection performance. In some sense, the background modeling can be considered as a technique to collect fixed scene information in a pixel-wise manner, and there are several well-known obstacles in precisely and robustly constructing background models such as bootstrapping, moving background, moved background, camouflage, light changes and shadows [5]. In this work, we try to deal with only the first five challenging problems.

Many features are developed and available for background modeling, but the most useful and widely used features are (1) pixel color, (2) local image structure, and (3) combination of (1) and (2).

With the color feature, the background model is constructed from consecutively observed temporal color samples under the assumption that each pixel in

K.M. Lee et al. (Eds.): ACCV 2012, Part III, LNCS 7726, pp. 493–506, 2013.

the scene is independent of each other. The popular idea to estimate background color statistics is utilizing multi-modal distributions such as mixture of Gaussians [6–9] or kernel density functions [10, 11]. However, the former has problems with rapidly changing background and the latter has a drawback of the high spatial complexity [12]. To solve these problems, codebook [12, 25] and K-means algorithm [13] based approaches were proposed, and recently, L. Maddalena et al. presented self-organizing maps (SOMs) based background modeling schemes [14, 15]. Although these techniques show remarkable performance in most of backgrounds, they cannot handle the camouflage condition properly.

To overcome this drawback, several researchers focused on other types of features such as intensity gradient or texture information which can represent the structural information of local image regions. Especially, texture features have been widely adopted recently because of their discriminative property and simple calculation. A texture feature captures local image information based on intensity differences between a center pixel and its regularly predefined neighborhood. Local binary pattern (LBP) [16], local ternary pattern (LTP) [17] and scale invariant local ternary pattern (SILTP) [18] are representative operators to produce such texture features. In general, the approaches based on the local spatial feature can detect foregrounds more robustly than color-based methods if there is sufficient amount of textures either in backgrounds or in moving objects even though their colors are very similar. However, at the same time, they cannot detect changes in sufficiently large uniform regions if the moving object is also uniform [20].

Some researchers tried to utilize both color features and local spatial features. Javed et al. used gradient statistics to verify temporary foreground blobs detected only based on a color background model in the pixel level processing [19]. However, this method still causes problems in the camouflaged object detection. Yao et al. proposed a layer-based background subtraction method exploiting both color information and texture information to statistically represent complex scene context [20]. However, their approach is also not sufficient for precise moving object segmentation because it extracts texture features based on the LBP operator which is not robust to small image noise in uniform image regions [17, 18].

In this paper, we present a novel framework using the texture model and the color model to segment moving foreground regions from an input frame under the diverse challenging background conditions. In our work, three main contributions are included. First, we propose a new scene adaptive LBP (SALBP) operator, which produces more consistent and correct codes by empolying scene adaptive multiple thresholds than existing texture operators (i.e., LBP, LTP and SILTP) employing inadequate thresholding mechanisms in their texture-code calculation [21]. Second, we propose a method of taking advantages of both color and texture features. Instead of using a global constant-valued weight indicating the contribution of the texture information to the final decision [20], we estimate local binary-valued weight per pixel depending on regional background conditions. Finally, we propose an effective algorithm to verify each foreground