Pattern Recognition Letters 32 (2011) 1720–1734

Contents lists available at SciVerse ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Foreground prediction for bilayer segmentation of videos

Zhen Tang *, Zhenjiang Miao, Yanli Wan, Forrest F. Jesse

Institute of Information Science, Beijing Jiaotong University, Beijing 100044, PR China

ARTICLE INFO

Article history:
Received 7 August 2010
Available online 6 August 2011
Communicated by S. Aksoy

Keywords:
Bilayer segmentation
Foreground prediction
Opacity propagation
Sudden illumination changes
OPSIC

ABSTRACT

Given a segmentation result (an alpha matte or a binary mask) of the former frame, foreground prediction is a process of estimating the probability that each pixel in the current frame belongs to the foreground. It plays a very important role in bilayer segmentation of videos, especially videos with non-static backgrounds. In this paper, a new foreground prediction algorithm which is called opacity propagation is proposed. It can propagate the opacity values of the former frame to the current frame by minimizing a cost function that is constructed by assuming the spatiotemporally local color smoothness of the video. Optical flow and probability density estimation based on a local color model are employed to find the corresponding pixels of two adjacent frames. An OPSIC (opacity propagation with sudden illumination changes) algorithm is also proposed which is an improvement of our proposed opacity propagation algorithm because it adds a simple color transformation model. As far as we know, this is the first algorithm that can predict the foreground accurately when the illumination changes suddenly. The opacity map (OM) generated by the opacity propagation algorithm is usually more accurate than the previously used probability map (PM). The experiments demonstrate the effectiveness of our algorithm.

1. Introduction

In recent years, several bilayer segmentation algorithms for videos with non-static backgrounds have been proposed (Pham et al., 2009; Bai et al., 2009; Ahn and Kim, 2008; Tang et al., 2010; Criminisi et al., 2006; Yin et al., 2007). In most of these algorithms (Pham et al., 2009; Bai et al., 2009; Ahn and Kim, 2008; Tang et al., 2010), the foreground prediction algorithms are employed. Given a segmentation result (an alpha matte or a binary mask) of the former frame $t-1$, foreground prediction is a process of estimating the probability that each pixel in the current frame t belongs to the foreground. In the previous work (Pham et al., 2009; Bai et al., 2009; Ahn and Kim, 2008; Tang et al., 2010), the prediction result is usually represented as a probability map. If the probability map is of high quality, then bilayer segmentation can be done accurately. Most current probability map estimation algorithms employ color models, for example global color model (Ahn and Kim, 2008; Pham et al., 2009; Tang et al., 2010), or local color model (Bai et al., 2009). In order to discriminate the foreground and background more clearly, other cues such as temporal coherence strips (Ahn and Kim, 2008) and shape model (Bai et al., 2009) are also integrated. In (Bai et al., 2010), a dynamic color flow model which combines the color modeling and motion estimation in a single probabilistic framework is proposed. It achieves good results in many cases.

Two problems still exist in the previous algorithms. The first problem is that in all of the previous algorithms, the probability that one pixel belongs to the foreground is estimated individually by the color model or the combined color and shape model. The affinities between neighboring pixels are not considered. This leads to probability maps of low quality especially for the difficult cases where the foreground and background have similar (un-separable) colors. Actually, if one pixel belongs to the foreground, neighboring pixels that have similar colors also have a high probability of belonging to the foreground. In other words, more accurate foreground prediction may be achieved if affinities of pixels are exploited and employed. The other problem is that, the color model based algorithms cannot deal with the sudden illumination changes between frames. That is because if the illumination changes, the color distribution of the pixels also changes. Thus the color models built using the pixel colors of former frame is no longer suitable to estimate the probability densities of pixels in the following frames.

In this paper, a novel foreground prediction algorithm which is called opacity propagation is presented. We call it "opacity propagation", because we take the alpha matte or the binary mask of the former frame as an opacity map, and our algorithm can propagate the opacity values (0–1 in an alpha matte or 0, 1 of a binary mask) from the former frame to the current through the affinities between pixels. The opacity propagation result is represented as a "opacity map". Opacity propagation is an extension of the closed-form matting (Levin et al., 2008) from 2D to 3D. It is based on the local smoothness assumptions of the foreground and background

* Corresponding author. Tel./fax: +86 10 51683149.
E-mail addresses: 06120376@bjtu.edu.cn (Z. Tang), zjmiao@bjtu.edu.cn (Z. Miao), 07112067@bjtu.edu.cn (Y. Wan), jesse@jesse.org (F.F. Jesse).

0167-8655/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2011.07.025

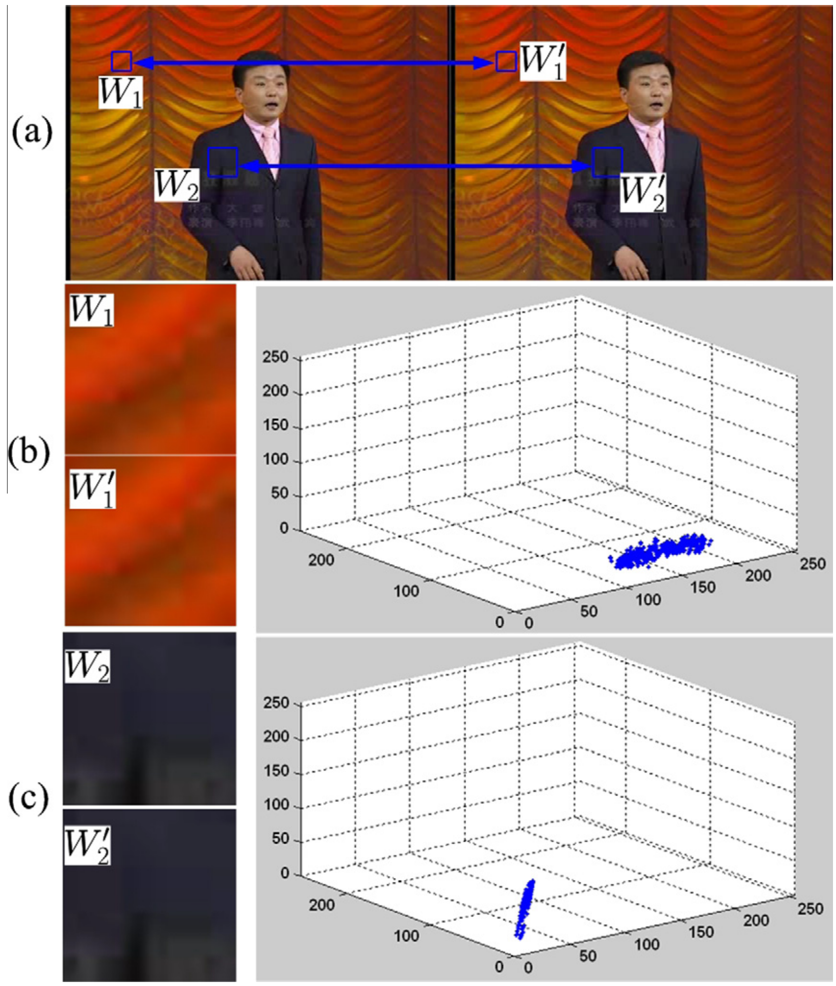

Fig. 1. The color line models of two 3D spatiotemporal windows. (a) Two consecutive frames with two 3D spatiotemporal windows. (b, c) The corresponding color line models of the two 3D spatiotemporal windows.

colors in a small 3D spatiotemporal window. We demonstrate that this local smoothness assumption can also be formulated into a quadratic function that can be minimized by solving a sparse liner system of equations. The optical flow and local color model based probability density estimation is also used to find the corresponding pixels of two adjacent frames. In order to cope with the sudden illumination changes, we incorporate a local color transformation model into our algorithm and propose a OPSIC (opacity propagation with sudden illumination changes) algorithm. The local color transformation model can eliminate the negative effect of the illumination changes under the observation that the illumination changes are locally smoothed. We demonstrate that the transformed colors also satisfy the local smoothness assumptions. An ANCC (Adaptive-Normalized Cross-Correlation) based algorithm (Heo et al., 2011) is employed to find the corresponding pixels between two frames with sudden illumination changes. The experiments show that the opacity map generated by our algorithm is more accurate than the previously used probability maps, and the illumination changes between frames can also be well processed.

The rest of this paper is organized as follows. We first review related work in Section 2 and provide a detailed description of our opacity propagation algorithm in Section 3. Section 4 describes how to obtain our OPSIC algorithm by adding a local color transformation model into the proposed opacity propagation algorithm. We test our approach through several experiments in Section 5, and a discussion and conclusion is presented in Section 6.

2. Related work

PM estimation is the most commonly used foreground prediction algorithm in the previous work (Pham et al., 2009; Bai et al., 2009; Ahn and Kim, 2008; Tang et al., 2010). In these PM estimation algorithms, color is the most commonly used cue for probability estimation. Ahn and Kim (2008), Tang et al. (2010) and Pham et al. (2009) all employed global color models to estimate the probability map. Their global color models are built using all the foreground and background pixels in the already well segmented former frame. In (Ahn and Kim, 2008), the authors built two sample sets M_0 and M_1 using the background and foreground pixels respectively, and then employed the Kernel Density Estimation (KDE) algorithm to estimate the non-parametric pdf's (probability density functions) for the fore and background. In (Tang et al., 2010; Pham et al., 2009), the authors built two histograms for the segmented fore and background layers in the former frame, and then employed the Bayes formula to estimate the foreground probability of each pixel in the current frame. Tang et al. (2010) further employed a Weighted Kernel Density Estimation (WKDE) to refine the probability maps generated by the Bayesian estimation. The color model used for WKDE is built within a dilated foreground mask instead of the whole image, which can be seen as a semi-local color model. Bai et al. (2009) using local color models to estimate the probability map. They dropped a series of overlapped windows along the foreground boundary in the former frame, and then built GMM (Gaussian Mixture Model) color model