

# Chapter 5

## Motion Segmentation and Matting by Graph Cut

Jiangjian Xiao

Ningbo Industrial Technology Research Institute, P.R. China

### ABSTRACT

*Given a video sequence, obtaining accurate layer segmentation and alpha matting is very important for video representation, analysis, compression, and synthesis. By assuming that a scene can be approximately described by multiple planar or surface regions, this chapter describes a robust approach to automatically detect the region clusters and perform accurate layer segmentation for the scene. The approach starts from optical flow field or small corresponding seed regions and applies a clustering approach to estimate the layer number and support regions. Then, it uses graph cut algorithm combined with a general occlusion constraint over multiple frames to solve pixel assignment over multiple frames to obtain more accurate segmentation boundary and identify the occluded pixels. For the non-textured ambiguous regions, an alpha matting technique is further used to refine the segmentation and resolve the ambiguities by determining proper alpha values for the foreground and background, respectively. Based on the alpha mattes, the foreground object can be transferred into the other video sequence to generate a virtual video. The author's experiments show that the proposed approach is effective and robust for both the challenging real and synthetic sequences.*

### INTRODUCTION

Layer-based motion segmentation has been investigated by computer vision researchers for a long time (Adiv 1985, Wang 1994, Ayer 1995, Patras 2001, Ke 2004, Xiao 2004). Once motion segmentation is achieved, a video sequence can be efficiently represented by different layers. Given a

video sequence, motion segmentation consists of two major steps: (1) layer clustering, which is to determine the number of layers in the scene and the associated motion parameters for each layer; (2) dense layer segmentation, which is to assign each pixel in the image sequence to the corresponding layer and identify the occluded pixels. Currently, a number of approaches have been proposed for

DOI: 10.4018/978-1-4666-1891-6.ch005

layer clustering problem, which have achieved good results, such as linear subspace (Ke 2001, Ke 2002, Ke 2004), GPCA (Vidal 2004), K-means (Wang 1994), and hierarchical merging (Wills 2003, Xiao 2004).

However, once the initial layer clustering is achieved, how to correctly assign the pixels to different layers is a difficult problem (Ayer 1995, Ke 2004, Khan 2001, Xiao 2004) as shown in Figure 1. Particularly, if the images contain some non-textured regions such as the blue or white regions corresponding to the sky in Figure 1a, the pixels in these regions may satisfy different layer motion parameters. Hence, the segmentation only using the motion cue may not provide an accurate layer boundary for those regions due to the motion ambiguities.

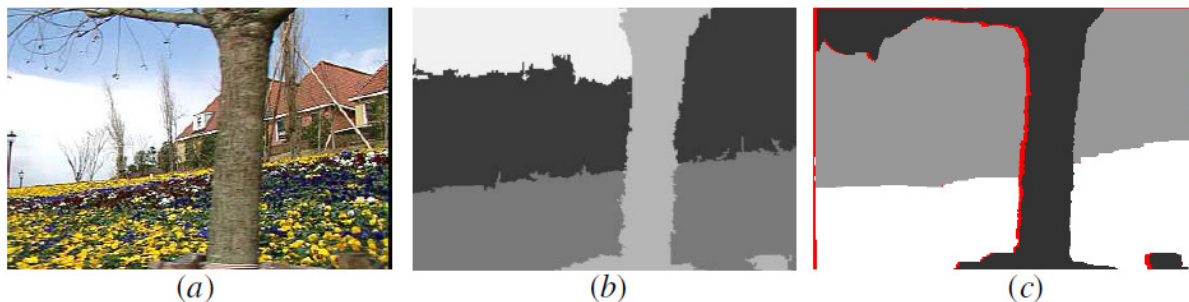
In digital matting, the observed color of the pixels around layer boundaries can be considered as a mixture of foreground and background colors, which is formulated as  $C = \alpha F + (1 - \alpha)B$ , where  $C$ ,  $F$ , and  $B$  are the observed, foreground, and background colors, and  $\alpha$  is the pixel's opacity channel. For single image matting, once a trimap (unknown, definitely foreground, and definitely background regions) is manually specified, the alpha values, foreground, and background colors of the unknown regions can be estimated under certain constraints (Rother 2004, Sun 2004, Ruzon 2000, Chuang 2001). Typically,

the alpha matting techniques are more suitable for smooth regions since they strongly rely on an assumption that the color of the estimated background and foreground should smoothly change in the unknown areas. Given a cluttered background, the performance of the existing alpha matting approaches tends to deteriorate. Compared to the traditional motion segmentation problem, pulling alpha mattes between two overlapping layers can be considered as a refinement step of the segmentation particularly for those ambiguous, smooth regions.

In this chapter, we introduce a novel approach, which combines the merits of motion segmentation and alpha matting technique together, to extract accurate layer boundaries and alpha mattes simultaneously from a video sequence. Our algorithm is implemented in two stages. The first stage is layer clustering and the second stage is accurate layer segmentation and matting.

In the first stage, we determine seed correspondences over a short video clip (3-5 frames). Then, we gradually expand each seed region from an initial square patch into an enlarged support region of an arbitrary shape to eliminate the overfitting problem and detect the outliers. This is achieved using a graph cuts approach integrated with the level set representation. After that, we employ a two-step merging process to obtain a layer description of the video clip.

*Figure 1. Previous results for flower-garden sequence. (a) One frame from the original sequence. (b) Result of Ayer and Sawhney (Ayer 1995). (c) Result of Ke and Kanade (Ke 2004). (c) Result of (Xiao 2004), where the red pixels are occluded between the neighboring frames.*



21 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:  
[www.igi-global.com/chapter/motion-segmentation-matting-graph-cut/69072](http://www.igi-global.com/chapter/motion-segmentation-matting-graph-cut/69072)

## Related Content

---

### Study of Various Image Segmentation Methodologies: An Overview

Abahan Sarkarand Ram Kumar (2017). *Applied Video Processing in Surveillance and Monitoring Systems* (pp. 1-27).

[www.irma-international.org/chapter/study-of-various-image-segmentation-methodologies/168898](http://www.irma-international.org/chapter/study-of-various-image-segmentation-methodologies/168898)

### Line Parameter based Word-Level Indic Script Identification System

Pawan Kumar Singh, Supratim Das, Ram Sarkarand Mita Nasipuri (2016). *International Journal of Computer Vision and Image Processing* (pp. 18-41).

[www.irma-international.org/article/line-parameter-based-word-level-indic-script-identification-system/171129](http://www.irma-international.org/article/line-parameter-based-word-level-indic-script-identification-system/171129)

### Detailed JavaScript API for Google Earth Engine to Enhance Monitoring of Surface Water Dynamics in the Drought-Affected Mediterranean Region

Adil Moumane, Jamal Al Karkourianand Batchi Mouhcine (2025). *Remote Sensing and GIS Techniques in Hydrology* (pp. 105-132).

[www.irma-international.org/chapter/detailed-javascript-api-for-google-earth-engine-to-enhance-monitoring-of-surface-water-dynamics-in-the-drought-affected-mediterranean-region/360444](http://www.irma-international.org/chapter/detailed-javascript-api-for-google-earth-engine-to-enhance-monitoring-of-surface-water-dynamics-in-the-drought-affected-mediterranean-region/360444)

### Genetic Algorithms and Other Approaches in Image Feature Extraction and Representation

Danilo Avola, Fernando Ferriand Patrizia Grifoni (2009). *Artificial Intelligence for Maximizing Content Based Image Retrieval* (pp. 1-19).

[www.irma-international.org/chapter/genetic-algorithms-other-approaches-image/4149](http://www.irma-international.org/chapter/genetic-algorithms-other-approaches-image/4149)

### Implementation and Evaluation of a Computational Model of Attention for Computer Vision

Matthieu Perreira Da Silvaand Vincent Courboulay (2013). *Developing and Applying Biologically-Inspired Vision Systems: Interdisciplinary Concepts* (pp. 273-306).

[www.irma-international.org/chapter/implementation-evaluation-computational-model-attention/72033](http://www.irma-international.org/chapter/implementation-evaluation-computational-model-attention/72033)