論文の内容の要旨

論文題目　　Formulating Color Information for Image and Video Segmentation
（画像・映像セグメンテーションにおける色情報の記述方式に関する研究）

氏　　名　　ファン　ヴェトクォク

Image segmentation is a technique for extracting a foreground region in a picture or photograph from its background. It is a critical early step in most computer vision applications in domains such as object recognition, image editing, surveillance and human-computer interaction. This dissertation focuses on the problems of image and video segmentation based on color information. I consider the segmentation problem as a binary labeling problem of minimizing an energy function, and propose several methods for formulating color information in the energy function. My proposed formulations cover different types of energy terms, including the pixel-wise term, the pairwise term and the domain-wise term. Each contribution is proved to be effective for a certain aspect of segmentation problems.

My first research work addresses the problem of online video segmentation and matting. My method is based on a parallel process of two stages: bilayer segmentation and alpha matting. I propose an accurate bilayer segmentation method for extracting the foreground region from the background, using a new pixel-wise formulation based on spatio-temporal color coherence and a local classifier using the propagation of color likelihoods. The proposed local likelihood proves to be very effective for correcting segmentation errors caused by global models.

My second work targets the problem of segmenting an object of interest from a given bounding layer. My strategy is to optimize the trade-off of making the foreground region as large as possible while keeping the similarity between the foreground and background regions as small as possible. This similarity is formulated by using a new pair-wise term evaluating the similarities of distant pixel pairs. With this formulation, I generate an energy function that is expressed purely in terms of unknown segmentation and can be optimized using only one max-flow calculation.

The third work relates to the problem of image segmentation with a reference distribution. The input distribution does not need to be precise, and is used as a guide to infer the latent distribution and its consistent region. My key observation is that the latent distribution resembles the distribution of the consistent region but is distinct from the distribution of the complement region. I state the problem as the minimization of an energy function consisting of domain-wise terms which evaluate global similarities and dissimilarities using the Bhattacharyya distance. I also propose a novel iterative scheme for jointly optimizing distribution and segmentation.

Differences in performance among these works can basically explain the different characteristics of each type of energy terms. Understanding the different behavior of each type is particularly important for solving similar segmentation problems.