STRUCTURING ELEMENTS FOLLOWING THE OPTICAL FLOW

Combining Morphology and Motion

Nicolas Laveau[1] and Christophe Bernard[2]

[1]*Centre de Morphologie Mathématique*
35, rue Saint-Honoré, 77305 Fontainebleau, France
laveau@cmm.ensmp.fr

[2]*Let It Wave*
XTEC, École Polytechnique, 91128 Palaiseau, France
bernard@letitwave.fr

Abstract This paper deals with the combination of classical morphological tools and motion compensation techniques. Morphological operators have proven to be efficient for filtering and segmenting still images. For video sequences however, using motion information to modify the morphological processing is necessary. In previous work, iterative frame by frame segmentation using motion information has been developed in various forms. In this paper, motion is used at a very low level, by locally modifying the shape of the structuring element in a video sequence considered as a 3D data block. Motion adapted morphological tools are described and their use is demonstrated on video sequences. Moreover, the features of the motion model best suited to our purpose are also discussed.

Keywords: Mathematical Morphology, Motion compensation, Optical Flow, Structuring Element

Introduction

Mathematical morphology provides very efficient tools for tasks like filtering (alternate sequential filters, levelings, etc) and for segmenting still images. These tools have been also used for 3-dimensional datasets like medical volume imaging where they perform equally well.

Segmentation tools have also been designed to segment video sequences, and to extract video objects ([4], [3]). These segmentation tools mostly dissociate computation performed along the time axis and along the space axes:

C. Ronse et al. (eds.), Mathematical Morphology: 40 Years On, 43–52.

segmentation is done frame by frame, and temporal correlation in the video sequence is exploited by propagating segmentation results along time. For instance, markers are computed from a segmentation of a frame t, and are used to segment frame $t+1$ ([7]). Alternatively, frames are segmented separately beforehand, and the resulting segmentation graphs are matched in a second step ([1], [5]). In these examples, the temporal correlation between the frames is weakly enforced, and as a result the boundaries of the segmented video objects tend to flicker. Another approach is 3D filtering. This approach is more promising because it enforces a stronger temporal correlation between frames.

However either method is limited by time aliasing: when large displacements occur between frames, pixels belonging to the same object in time are not connected together and segmentation can fail. For similar reasons, video sequence filtering is also bound to be less efficient.

Our way around this limitation consists in introducing motion information provided by a motion estimation method, and incorporating this motion information to modifiy locally the grid connectivity of the 3D video volume, so that pixels that belong to the same object stay connected together along the time axis.

In a first part, we will detail the 3D morphological filtering, and why they are not sufficiently efficient. In the second part, new structuring elements following the optical flow will be defined that answer the issue of temporal aliasing, and example and results will be given in a third part. Then we will discuss the issue of the influence of the motion model on the quality of our new structuring elements.

1. 3D morphology on video sequences

3D morphology as already applied for volumic data can be applied to video sequences. In a standard setting, the grid is square and the neighbourhood at (t,x,y) is $(t,x,y)+\{(0,0,0),(\pm 1,0,0),(0,\pm 1,0),(0,0,\pm 1)\}$ and is octaedric. The corresponding structuring element is the set $S=\{(\delta t,\delta x,\delta y):|\delta x|+|\delta y|+|\delta t|\leq 1\}$.

However, the contents of a video sequence does not behave the same way along spatial axes and along the time axis. While the video sequence is at most only slightly aliased along spatial axes, it can be very strongly aliased along the time axis.

As a result, a moving object may not be considered with respect to the resulting connectivity as a single connected component. This is illustrated in fig. 2. On both left and side parts, the grayed squares represent a moving object. On the left figure, an unslanted structuring element is represented. According to the resulting connectivity, the moving object is not a single connected component. If the structuring element is slanted in order to take into account visual