Pattern Recognition 47 (2014) 2266–2279

Contents lists available at ScienceDirect

Pattern Recognition

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

Automatic segmentation of granular objects in images: Combining local density clustering and gradient-barrier watershed

Huiguang Yang*, Narendra Ahuja

Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL 61801, USA

A R T I C L E I N F O

Article history:
Received 8 February 2013
Received in revised form
26 September 2013
Accepted 6 November 2013
Available online 20 November 2013

Keywords:
Image segmentation
Granular object recognition
Local density clustering
Gradient-barrier watershed
Intensity moment
Edge detection
Cell/nuclei segmentation
Nanoparticle segmentation

A B S T R A C T

Blob or granular object recognition is an image processing task with a rich application background, ranging from cell/nuclei segmentation in biology to nanoparticle recognition in physics. In this study, we establish a new and comprehensive framework for granular object recognition. *Local density clustering* and *connected component analysis* constitute the first stage. To separate overlapping objects, we further propose a modified watershed approach called the *gradient-barrier watershed*, which better incorporates intensity gradient information into the geometrical watershed framework. We also revise the marker-finding procedure to incorporate a clustering step on all the markers initially found, potentially grouping multiple markers within the same object. The gradient-barrier watershed is then conducted based on those markers, and the intensity gradient in the image directly guides the water flow during the flooding process. We also propose an important scheme for edge detection and fore/background separation called the *intensity moment approach*. Experimental results for a wide variety of objects in different disciplines – including cell/nuclei images, biological colony images, and nanoparticle images – demonstrate the effectiveness of the proposed framework.

1. Introduction

Recognition and segmentation of blob or granular objects in the image is an important and fundamental task in image processing. This problem has a rich practical background in applications, such as the recognition of biological cells [12,17], cell nuclei [8–10], colonies, and pollen [34,35], as well as nanoparticles [6], and so on. Very large numbers of objects in the image make manual segmentation and counting quite tedious, if it is even feasible, so computer vision is crucial to the task. Given that objects may also vary in shape, size, and intensity and may overlap or cluster, the challenges of recognition and segmentation are in no way trivial.

The recognition of blob objects in an image can be first regarded as detecting clusters of high-density foreground pixel (pixel-of-interest) clouds in the image. For detecting the clusters of pixel-of-interest, local density clustering together with connected component analysis constitutes a good scheme and will be discussed in detail in Section 3. Local density clustering is able to cluster objects of any shape and any size. However, the major drawback of this method is that it cannot separate overlapping objects. All the clusters or objects that overlap will be grouped into the same cluster, since they are connected. To overcome this limitation, more processing is needed, such as making use of clues in the object shape or intensity gradient within a connected component.

There have been many approaches to separating overlapping objects. These include the *watershed algorithm* [7–15], the *gradient or edge detection method* [39], *morphological erosion* [6], the *active contour method* [16–24], the *sliding band filter approach* [25,26], and others. A nice review and comments on some of these approaches can be found in [6]. The gradient or edge method apparently does not work well in cases where there is no obvious intensity difference between the overlapping objects or if the objects are strongly textured. The active contour method is quite computationally demanding, making it unsuitable for a case in which the number of objects is large, which is in fact the most meaningful case for computer-aided segmentation. The sliding band filter approach requires that the range of object size be known beforehand, and it does not work well if the size range is wide. While the watershed method is still an effective and efficient method to separate overlapping objects, improvements can be made to the algorithm.

The advantages of the watershed approach are (i) it can provide the natural growth of the region corresponding to each object independent of object shape and size, and (ii) it automatically provides a closed contour as well as computational efficiency. However, directly applying the watershed algorithm to the image or its gradient can lead to severe over-segmentation due to large numbers of local minima/maxima in the image or its gradient version. Many remedies have been proposed to overcome this issue [9,11,27–31]. *Hierarchical watershed segmentation* aims to

* Correspondence to: 1520 Beckman Institute, University of Illinois, 405 N. Mathews, Urbana, IL 61801, USA. Tel.: +1 217 402 5252.
E-mail addresses: hyang30@illinois.edu, hyang30@uiuc.edu, n-ahuja@illinois.edu (H. Yang).

0031-3203/$ - see front matter
http://dx.doi.org/10.1016/j.patcog.2013.11.004

merge the over-segmentation hierarchically to form meaningful object regions, for example based on the mosaic image transform and associated graph [27] or by multi-scale filtering of the image and segmenting on the filtered and simplified image [30]. Some other studies have proposed using the pattern classification and object model learned from the data to direct the region-merging [9,11,29]. Compared with the methods aiming to conduct a blind watershed on the image first and then merge the over-segmentations afterward, it would be better not to over-segment the image in the first place.

We believe the best way to conduct watershed segmentation is not directly from the original image or its gradient version. It is better to first find the marker corresponding to each object in the image and then to conduct the watershed based on those markers. This approach gives a much better guarantee of object counts and approximate locations in the image. Therefore *automatic detection of markers* is the most critical step in using watershed segmentation. There are several approaches to detecting markers, including the *distance transform*, *morphological erosion*, and the *gradient transform*. Under the appropriate condition, such as for convex object, the first two approaches can be shown to be essentially similar, as the final result of morphological erosion is also the local maximal distance region [6]. Both approaches are based on purely geometrical information and require the overlapping objects to display a bottleneck region as the hint for the location of separation. Both also require the individual object to be more or less convex in shape, and may lead to over-segmentation when this requirement is violated. An alternative way to detect markers is the gradient transform. It is based on the assumption that the inter-object gradient is larger than the intra-object gradient, and connected low gradient regions are detected as markers. However, this method is very sensitive to image noise and often leads to over-segmentation. Therefore using the distance transform as the basic framework and combining gradient information into the system would be a good option.

To combine gradient information into the watershed process based on the distance transform framework, one study uses the gradient-weighted distance transform [9] to alter the "distance" at a certain pixel regarding its gradient. Such a method is free of parameter tuning, but the incorporation of the gradient into the geometric framework is based on heuristics. So it is not immediately apparent where the watershed or boundary will be, or whether it will be accurate. Therefore in this study we propose an alternative version of watershed, the *gradient-barrier watershed*, in which the flooding process is still carried out based on the distance transform framework, but the image gradient directly guides the water flow in the watershed process.

In addition to the mainstream watershed techniques, we propose an important method for edge detection and foreground/background separation, which is essential for the object detection and connected component analysis. The method is based the concept we propose in this paper called the *intensity moment*, which will be detailed in Section 2.

The structure of the paper is as follows. In Section 2 we discuss the concept of intensity moment and intensity moment scheme for foreground/background separation of the image. In Section 3 we discuss the local density clustering method for object detection and delineation. The entire methodology and approach are discussed in detail in Section 4. The experimental result is presented in Section 5 and finally we conclude in Section 6.

2. Foreground/background separation: the intensity moment scheme

Given the image, the first step in recognition is typically finding the foreground pixels or the pixels-of-interest. The commonly used foreground/background pixel classification methods include *intensity thresholding* (for example, the Otsu method [32] for automatic threshold detection), the *gradient or edge detection method*, and so forth. Each method has its strengths and weaknesses, and no single method can perfectly handle all images. In this study, we propose another scheme called the *intensity moment approach*. In essence, the intensity moment approach tries to find the imbalance of the intensity distribution within the local patch of a certain scale around each pixel. In an analogy with the force moment, the intensity moment is calculated via the vector summation of the product between pixel intensity and vector distance to the patch center for each pixel within the patch (Fig. 1).

$$\vec{M}(i_0,j_0) = \sum_{(i,j) \in D} I(i,j)\vec{L}(i,j) \tag{1}$$

In above D is the local patch or domain centered at (i_0,j_0), $I(i,j)$ and $\vec{L}(i,j)$ are the pixel intensity and the vector distance to the patch center (i_0,j_0) for each pixel (i,j), respectively, and $\vec{M}(i_0,j_0)$ is the intensity moment of the patch centered at (i_0,j_0).

If the intensity has variation but the overall distribution is balanced or uniform in the scale of the patch, such as is the case with local textures, then the intensity moment has only a weak response at that point. However, if there is a salient edge in the patch (the edge between the object and background), then the intensity variation is in no sense balanced, repeating, or uniform in the patch scale. Rather, there are two distinct halves in the patch, so the response of the intensity moment will be strong at that point. Therefore, by the value of the intensity moment, we can locate the salient edge between the object and the background in the image, while ignoring unwanted details. Since the intensity moment approach takes into account the balance of intensity distribution within a certain scale, it is much better than gradient edge detection at finding the salient edges or the "main structure" of the image, as shown in Fig. 2. Here, what is most important is not the property of a single pixel or the gradient at that point, but rather the behavior of the local patch up to a certain scale, specifically whether it is uniform or has distinct parts. Finally, we can further classify the foreground/background pixels based on the detected object outlines.

3. Local density clustering for object detection

Besides the intensity moment approach mentioned above, some other methods can also be applied to differentiate foreground/background of the image (such as the Otsu thresholding [32]) and each method might be suitable for some case. Given we

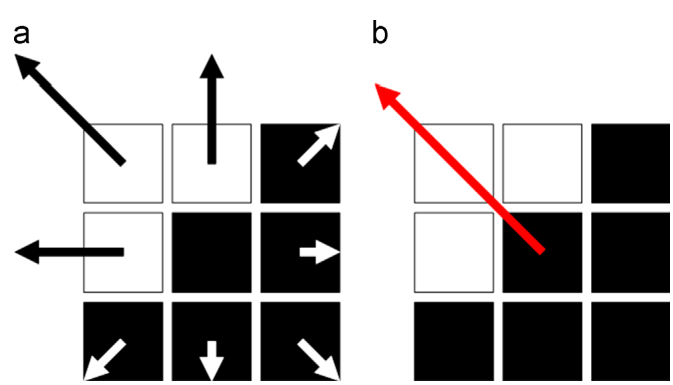

Fig. 1. Schematic illustration of the concept of intensity moment. Each block represents a pixel, dark block has lower intensity value than bright block. (a) Intensity moment and its direction of each pixel in the region with respect to the center pixel. (b) Total intensity moment and its direction of the entire region (or of the center pixel).