Concurrent Segmentation and Recognition with Shape-Driven Fast Marching Methods

Abdulkerim Capar, Muhittin Gokmen

Istanbul Technical University, Computer Engineering Department
34469 Ayazaga, Istanbul Turkey
capar@itu.edu.tr

Abstract

We present a variational framework that integrates the statistical boundary shape models into a Level Set system that is capable of both segmenting and recognizing objects. Since we aim to recognize objects, we trace the active contour and stop it near real object boundaries while inspecting the shape of the contour instead of enforcing the contour to get a priori shape. We get the location of character boundaries and character labels at the system output. We developed a promising local front stopping scheme based on both image and shape information for fast marching systems. A new object boundary shape signature model, based on directional Gauss gradient filter responses, is also proposed. The character recognition system that employs the new boundary shape descriptor outperforms the other systems, based on well-known boundary signatures such as centroid distance, curvature etc.

1. Introduction

In this study, a multitask system which is capable of segmenting and recognizing objects concurrently is developed. The proposed system is applied to license plate and handwritten character segmentation and recognition problems. Object segmentation is the initial part of any image recognition, understanding and tracking system. Level Sets is a promising implicit contour based segmentation method, introduced by Osher and Sethian [1]. Fast marching method is also proposed by Sethian [2] to overcome the high computational requirements of level sets. Although ordinary level sets has the complexity of O(N^3) ( O(N^2) for narrow band) fast marching method can be implemented with O(N) complexity [3].

One of the challenges in the field of image segmentation is the incorporation of prior knowledge on the shape of the segmenting contour [4, 5]. Integration of statistical shape variation into the level set methods was first proposed by Leventon et.al. [6] and then several researchers were worked on this area [7,8,12]. Cremers et al. [9,10] first present a variational integration of nonlinear shape statistics into a Mumford–Shah [11] based segmentation process. There are following differences between these studies and the proposed system:

- In these studies, the evolving front is always forced to have the prior shape. However, we stop the front near object boundaries
- It is stated that, the proposed method does not work when the number of prior object classes is more than one [12]. However, our system is capable to segment and recognize different class of characters.
- Previous researchers obtained the shape statistics from the whole map of level set values; however we employ only the front itself for shape description.
- Previously proposed systems need high calculation power because they have two optimization stages, one is for minimization of image energies, and other is for minimizing shape similarity energies. On the other hand, our system has one optimization step for minimizing both energies.

In this paper, a promising local front stopping scheme is represented for fast marching that integrates the image based and shape based information. Another contribution of the study is to develop a character boundary shape description model that utilizes the directional Gauss gradient filter responses.

2. Fast Marching Representation

The goal of the Level Set method is to track the motion of the interface as it evolves with a known speed [13]. Fast Marching (FM) method is very fast versions of the level set methods with some limitations, that is curve propagation speed F must be of constant sign and the curve must be evolve in one direction. Since FM method guarantees that one image element is passed only one time by the front, evolving equation can be formulated in terms of the time values as

\[ \nabla T \big| F = 1, \]

where T(x,y) is the time function, indicating the time value at which the front passes the point (x,y). We can slow down and even stop the front around the object boundaries by adjusting the speed value F. The approximate solution of Eq. 1 on a 2D grid is given as in [2]

\[ \begin{align*}
\max(D_{ij}^x T, 0)^2 + \min(D_{ij}^x T, 0)^2 + \\
\max(D_{ij}^y T, 0)^2 + \min(D_{ij}^y T, 0)^2 \bigg] = \frac{1}{F_{ij}^2}
\end{align*} \] (2)
where $D_{ij}^{+}\frac{T_{i+1,j}-T_{i,j}}{\Delta x}$ and $D_{ij}^{-}\frac{T_{i-1,j}-T_{i,j}}{\Delta x}$

Fast marching algorithm [2] iterates on a 2D grid which consists of “accepted”, “narrow band trial” and “far away” points. Fast marching method allows only modeling the speed function $F$ which can depend on local or global forces. We have chosen image gradient and front curvature as local forces, and shape representation of the front as global forces.

3. Shape Description Model

Most of the systems that integrate prior shape statistics with level set methodology employ the level set distance map for shape description. This approach is not suitable with our system, since fast marching iterations are achieved near the evolving front and also because of the time consumption. We designed a new boundary based shape description system which has high discrimination power among character classes with respect to other boundary based character recognition methods. The system utilizes directional Gauss gradient responses and Fourier Descriptors for shape representation.

3.1. Fourier Descriptors

Fourier Descriptors (FDs) are mostly employed for boundary shape description. Zhang and Lu [14] compared shape retrieval using FDs derived from different shape signatures and from different Fourier invariants in terms of computation complexity, robustness, convergence speed and retrieval performance. They reported that the centroid distance shape signature outperforms other signature methods in terms of above criterions.

For a given shape defined by a closed curve $C$ which in turn is represented by a one dimensional periodic function $u(t)$, called shape signature. The discrete Fourier transform is given by

$$a_n = \frac{1}{N} \sum_{t=0}^{N-1} u(t) \exp(-j2\pi nt)$$

(3)

where $a_n$, $n = 0, 1, 2, ..., N-1$ are the Fourier coefficients which are used to derive Fourier descriptors. Scale invariance is achieved by dividing the magnitudes by the DC component, i.e., $a_0$. The normalized Fourier coefficients are called FDs. Selecting the shape signature $u(t)$ is the most critical step for FDs. Various signature models were proposed in the literature such as complex boundary coordinates, centroid distance and boundary curvatures. The performance of these signature models is not sufficient when they are employed with a shape driven active contour system. Therefore, we proposed a new shape signature model based on the response of directional Gauss gradient filters.

3.2. FDs with Directional Gauss Gradient Filters as Shape Signature

We setup special directional gradient filter kernels and integrate the responses on the boundary points as shape signature for FDs. The directional signature $DS^k$ on the $i^th$ front point $(i,j)$ is calculated as

$$DS^k(t) = \sum_{x=-N/2}^{N/2} \sum_{y=-N/2}^{N/2} \zeta^k(x,y) (I(i+x,j+y))$$

(4)

where $I$ is the image intensity map and $\zeta^k$ is the $k$th rotational Gauss gradient filter. They are defined as

$$\zeta^k = R_{\theta_k} \nabla G_{\sigma}, \nabla G_{\sigma} = \left[\frac{\partial G_{\sigma}}{\partial x}, \frac{\partial G_{\sigma}}{\partial y}\right]^T$$

(5)

where $R_{\theta_k}$ is the affine rotation matrix. Directional signatures $DS^k$ are normalized with their mean and variance values.

![Fig. 1. Directional Gauss gradient filter responses. (a) Input image, (b) gradient filter responses, rotated by (b) 0°, (c) 90°, (d) 135°, (e) 45°.](image1)

We have selected totally four filter kernels with the rotation angles of 0, 45, 90 and 135 degrees. Filter responses are illustrated in Fig. 1. Obtained FDs are combined and fused to form the character feature vector and applied a dimension reduction process with Principal Component Analysis (PCA). In classification phase, we utilized two different classifiers - nearest neighbor (KNN), support vector machine (SVM) and compare their performances.

![Fig. 2. Proposed segmentation & recognition system](image2)
\[
F = \frac{F^*}{1 + \alpha \sqrt{G^* I}} 
\]  
(6)

where \( F^* \) denotes a curvature-related term in order to keep the propagating curve as smooth as possible [15]. Algorithm at this step is as follows:

1) Initialize the front at the borders of the image (First image in Fig. 3)
2) Perform Fast Marching iterations with speed function in Eq.6. At each iteration
   a) Evolve the front
   b) Find 8-connected front components, examine size of each blob whether it is a character
   c) Stop the character-like front components (red colored regions in Fig. 3) and send them to the next step
   d) Finish iterations, if whole front has stopped.

Fig. 3. Coarse segmentation results

4.2. Fine Segmentation & Recognition

![Diagram of Fine Segmentation & Recognition System]

Fig. 4. Fine segmentation & recognition system

Purpose of this step is to locate fine character boundaries and classify them with evolving active contours in the fast marching scheme. As seen in Fig. 4, fast marching iterations are processed to minimize three types of energies:

\[ E_{\text{gradient}}, E_{\text{smoothness}}, E_{\text{shape}} \]

Gradient and smoothness energies are minimized with fast marching speed function (Eq. 6). The speed function is inversely proportional with the local image gradient strength and proportional with the curvature. Therefore, evolving front slows down around image regions with high gradients, and when the evolving contour has small curvatures. Shape information is embedded into the front stopping conditions. We built up an algorithm to stop the evolving front by the help of these local and global measurements as the following:

1) Initialize the front at the borders of the character image
2) Make Fast Marching iterations with speed function in Eq. 6. At each iteration
   a) Evolve the front
   b) Use “Narrow Band Trial Points” to form the contour and find global shape similarity confidence (SC) with the help of shape descriptors. If SC < T1 that means it is not trained shape go to (a).
   c) Start the front stopping process on “Narrow Band Trial Points”. We have two conditions to stop a point
      Condition 1. If the Gradient Magnitude (GM) of the node is greater than T2 and this GM is a min/max value in the direction of front normal.
      Condition 2. If the number of “Stopped Trial Point” in the 8-neighbourhood of the point is greater than two. (This condition supplies the smoothness of stopped curve.)
      If one of these conditions is fulfilled than mark the point as “Stopped Trial Point”. These stopped points are still in the list of trial points but they have no chance to march anymore (See gray colored points in Fig. 5).
   d) Record the classification result.
   e) If 95% of the trial points are stopped then finish iterations. Else go to (c).
3) Evaluate the recognition records and determine the resulting class id.

Fig. 5. Fine segmentation iterations. Black points are the trial points and gray points are the stopped trial points.

5. Experimental Work and Results

Performance of the proposed segmentation and recognition system is measured on license plate (LP) and handwritten character datasets.

License Plate database is collected from European customs gates. Plates are detected and skew normalized before character segmentation. There are totally 10904 gray level numeric license plate characters in the database, 6000 of them are used for training and others for testing. No normalization is applied to the characters. The MNIST database of handwritten digits, available in LeCun home page, http://yann.lecun.com/exdb/mnist/, has a training set of 60000 examples, and a test set of 10000 examples. The digits have been size-normalized and centered in a fixed-size image. The characters in these databases are tested by our variational segmentation & recognition system. We employed three types of shape signatures for FDs; curvature, centroid distance and the directional filter signatures. K-nearest neighbors (KNN) and support vector machines (SVM) are used as classifiers.

Table 1 demonstrates the recognition performances on the LP database on the test set. “Curvature” and “centroid” signatures represent similar percentages but the new “directional” signature outperformed them drastically. The
results are not very different in binary MNIST database (Table 2). Again directional signature has the best performance for both classifiers. As seen, the curvature signature is better than centroid signature for handwritten character database, but wise versa for LP database. As classifier, SVM outperformed KNN in all experiments.

<table>
<thead>
<tr>
<th>Table 1. Recognition percentages on LPR database</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>KNN</td>
</tr>
<tr>
<td>SVM</td>
</tr>
</tbody>
</table>

Table 2. Recognition percentages on MNIST database

|      | Curvature | Centroid | Directional |
| KNN  | 82,78     | 76,48    | 94,34       |
| SVM  | 85,17     | 81,23    | 96,89       |

The proposed system is also capable of segmenting broken characters, which is very difficult task for any boundary based character recognition system. Fig. 6 represents some examples of broken characters and the segmentation results while Fig.7 illustrates the segmentation results on some number plate images.

6. Conclusion

We aimed to develop a multitask system that segments and recognizes the objects concurrently. We established a shape driven active contour model, which utilizes a variational fast marching algorithm, and applied it to the license plate and handwritten character segmentation & recognition problems. The system works in two steps, first rough location of each character is found by an ordinary fast marching technique and then exact boundaries and class IDs of characters are determined by means of a special fast marching methodology, which depends on gradient, curvature and shape similarity information.

Shape similarity statistics were embedded into fast marching method for stopping the evolving front when the front resembles one of the trained shapes, instead of enforcing the front to get a priori shape by another optimization algorithm. We formulated a promising local front stopping scheme for fast marching that integrates the image based and shape based information together to provide blocking the contour near real edges and when the contour gets the desired shape. We also develop a promising boundary based shape descriptor. The recognition performance of this shape descriptor is compared with other well known shape signature models on different character databases and it is experimentally demonstrated that the descriptor outperforms others significantly.

The proposed system is also capable to capture the object boundaries under some corruptions such as object breaking, physical degradations and noise distortions, which are very hard problems for a segmentation system. Segmentation on occluded objects is a limitation of the system, since the interested object shape is not known at startup and the system does not try to enforce the contour to get a prior shape.

7. References