Spatio-temporal image segmentation using optical flow and clustering algorithm

Saša Galić and Sven Lončarić*
Ericsson Nikola Tesla
Krapinska 45, 10000 Zagreb, Croatia
*Faculty of Electrical Engineering and Computing, University of Zagreb
Unska 3, 10000 Zagreb, Croatia
e-mail: Sasa.Galic@etk.ericsson.se, Sven.Loncaric@fer.hr

Abstract

Image segmentation is an important and challenging problem in image analysis. Segmentation of moving objects in image sequences is even more difficult and computationally expensive. In this work we propose a technique for spatio-temporal segmentation of medical image sequences based on clustering in the feature vector space. The motivation for spatio-temporal approach is the fact that motion is a useful clue for object segmentation. Two- dimensional feature vector has been used for clustering in the feature space. The first feature is image brightness which reveals the structure of interest in the image. The second feature is the Euclidean norm of the optical flow vector. The optical flow field is computed using a Horn-Schunck algorithm. By clustering in the feature space, it is possible to detect a moving object in the image. Experiments have been conducted using a sequence of ECG-gated magnetic resonance (MR) images of a beating heart. The method is also tested on images with moving background. The experiments have shown encouraging results.

Keywords: spatio-temporal image segmentation, clustering, optical flow, image analysis

1. Introduction

Segmentation is a challenging field of image analysis. In particular, medical image segmentation has become very important with development of complex medical imaging modalities which are capable of producing a large quantity of high-resolution two-dimensional (2-D) and three-dimensional (3-D) images. The problem of image segmentation has been studied extensively and there is a large number of methods described in the literature [1].

In many applications including medical image analysis it is necessary to analyze an image sequence of the same scene but at different time moments. In such a case, the obtained images often consist of a static background and moving object(s) of interest. An example of such a sequence is ECG-gated magnetic resonance (MR) image of the heart taken at different time moments. For example, in ECG-gated MR scanning one may acquire sixteen 3-D MR volumes during a single heart beat cycle. Manual analysis of such images by physicians is difficult and time consuming. Computer assisted methods are capable of faster and more accurate quantitative measurements required for medical treatment.

In this work we propose a new technique for spatio-temporal segmentation based on clustering [3] in the feature vector space. Motion is a useful clue for object segmentation and that is the main motivation for the proposed spatio-temporal segmentation method. Computed features include image brightness and optical flow vector.

The paper is organized as follows. In Section 2 a short overview of optical flow algorithm is presented. In Section 3 segmentation using clustering method is presented. Experimental results are shown in Section 4, and conclusion is given in Section 5.

2. Optical flow field computation

In this work the algorithm proposed by Horn and Schunck [2] has been used. The algorithm determines the optical flow as a solution of the following partial differential equation:

$$\frac{\partial E}{\partial x}\frac{dx}{dt}+\frac{\partial E}{\partial y}\frac{dy}{dt}+\frac{dE}{dt}=0 \tag{1}$$

The solution of Equation 1 is obtained by numerical procedure for error function minimization. The error function E is defined in terms of spatial and time gradients of optical flow vector field and consist of two terms shown in Equations 2a-2e.

$$E=\iint(\alpha^2 E_c^2+E_b^2)dxdy \tag{2a}$$

$$E_b=E_x u+E_y v+E_t \tag{2b}$$

$$E_c^2=(\frac{\partial u}{\partial x})^2+(\frac{\partial u}{\partial y})^2+(\frac{\partial v}{\partial x})^2+(\frac{\partial v}{\partial y})^2 \tag{2c}$$

$$E_x=\frac{\partial E}{\partial x} \qquad E_y=\frac{\partial E}{\partial y} \qquad E_t=\frac{dE}{dt} \tag{2d}$$

$$u=\frac{dx}{dt} \qquad v=\frac{dy}{dt} \tag{2e}$$

E (x, y, t) represents luminance of the image in point *(x, y)* at time moment *t*. To solve the minimization problem a steepest descent method is used which is based on computation of gradient to determine the direction of search for the minimum. The optical flow algorithm has two main phases. In the first phase, gradient coefficients E_x, E_y, E_t are computed from input images. The coefficients represent estimates of the image gradients in space and time, and they are defined by Equation 2d. In the second phase, the optical flow vectors u and v defined by Equation 2e are computed. Depending on image content, the number of iteration steps must be chosen. If the moving object in the image has a big uniform area, the number of iterations must be large. If the image consists of small moving objects, the number of steps can be small.

3. Clustering-based segmentation

Most of the classical image segmentation techniques rely only on a single frame to segment the image. However, motion is a very useful clue for image segmentation. The main idea of this work is to develop a spatio-temporal image segmentation technique for image sequences. In this approach segmentation is not done on a simple frame-by-frame basis but utilizes multiple image frames to segment the objects of interest. For this purpose we extract features both from the actual image that has to be segmented and from neighboring image frames in the sequence. The extracted feature vectors are clustered using a clustering algorithm to determine the characteristic image regions. The research of various feature vectors is underway and here we present the currently used features. The first feature is image brightness which is useful for segmentation because the heart regions of interest are bright while the background is mostly dark. The second feature is the Euclidean norm of the optical flow vector defined by Equation 4.

$$E_{of} = \sqrt{(u^2 + v^2)} \qquad (4)$$

By using the above features, we obtain both the spatial and temporal information about the scene. K-means clustering algorithm has been used in this work [3]. The feature space is divided into four characteristic areas corresponding with four image regions. The first region is the static background and the initial cluster center vector $u_1(0)$ for first group is set to (0,0). The second image region represents the moving background and has the initial cluster center vector $u_2(0)$ equal to $(0,M)$, where M is maximum norm in E_{of} matrix computed by Equation 4. The third image region represents static objects with the initial cluster center vector $u_3(0)$ equal to $(0,m)$ where m is the maximum image brightness. And finally the fourth region represents moving objects and has the initial cluster center vector $u_4(0)$ equal to (m, M). The rule for partitioning the feature space is defined by Equation 5.

$$x_i \varepsilon R_k \Leftrightarrow d(x_i, u_k(n)) = \min\{d(x_i, u_j(n))\} \quad j=1...K \qquad (5)$$

where R_k is k-th cluster, K is the number of clusters (in our case $K=4$), and $U_k(n)$ is the center of the k-th cluster. The clustering algorithm works iteratively where in each cluster the membership for each feature vector is determined and then new cluster centers are computed by Equation (6).

$$u_k(n+1) : \sum_{x \varepsilon R_k} d(x_i, u_k(n+1)) = \min\{d(x_i, y)\} \quad k=1...K \qquad (6)$$

The process is repeated until centers stabilize, i.e. until $u_k(n+1)$ becomes equal to $u_k(n)$. The resulting clusters correspond to four characteristic image regions in the segmented image. Computed results must be taken with reserve because it is possible that some pixels are mistakenly classified to cluster three instead of four. This can happen if there is a very large motion in picture. With such a motion the object will be separated in two regions. The first region shows us the portion of the object that has very large motion, and the second region corresponds to the rest of the object. It is also important to mention that before the computation takes place, input features (image brightness and optical flow energy) are normalized to the same range of values (in our case we normalized optical flow energy according to brightness).

4. Experimental results

Experiments have been conducted using ECG-gated MR heart images acquired at sixteen time moments. The procedure consists of two steps. In the first step the optical flow is computed for two input images with the same content but acquired at different time moments. The result, which represents the optical flow field, is saved in a file that is loaded by the program for clustering analysis that is performed in the second step. The result consists of four binary images representing the four characteristic regions of the input image.

The results are shown in Figures 1 to 3. Figure 4 represents the segmented image computed by simple thresholding, and Figure 5 represents the segmented image computed using only the optical flow. Figure 6 shows the position of feature vector in the feature space with centers marked as boxes.

Figure 1 Input MR image.

Figure 2 Optical flow field.

Figure 3 Segmented image using clustering method.

Figure 4 Segmented image using threshold method.

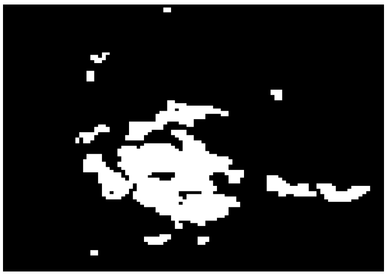

Figure 5 Segmented image using optical flow us input parameter.

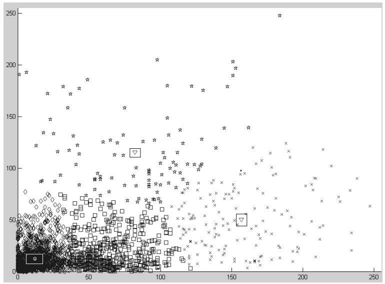

Figure 6 Clusters and centers.

The proposed method works correctly even in the case of moving background. In this case, the optical flow is shown in Figure7, while Figure 8 represents segmented image, which gives the same result as in the above experiment with mostly static background.

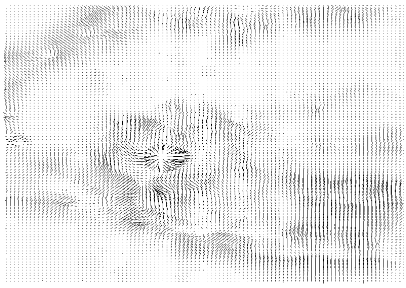

Figure 7 Computed optical flow.

Figure 8 Segmented image using clustering method.

5. Discussion and conclusions

In this work we have presented a method for spatio-temporal image segmentation which is based on clustering and optical flow computation. The experiments have been conducted using MR image sequence of the beating heart and have demonstrated the feasibility of the method. We have compared the results computed by clustering method with the results obtained by simple threshold segmentation method and with the results computed by applying optical flow vector as condition for detecting moving objects in the image. We have proved that the proposed method is robust to background movement what is a useful property. Future work will include investigation of other features and clustering techniques and implementation of the algorithm in 3-D.

Acknowledgement

We would like to thank Professor James Duncan from Yale University for providing the ECG-gated MR cardiac images required for this project.

References

[1] Haralick R. H, Shapiro L. G., Computer and Robot Vision, Vol. I, Addison–Wesley Publishing Company, 1993

[2] Horn B. K. P, Schunck G. B., "Determining Optical Flow", *Artificial Intelligence*, Vol.17, pp.185-203, 1981

[3] Haralick R. H, Shapiro L. G., Computer and Robot Vision, Vol. II, Addison–Wesley Publishing Company, 1993

[4] Prince J. L, Mc Veigh E. R, "Motion Estimation From Tagged MR Images Sequence", *IEEE Trans. on Medical Imaging*, 11 (2), pp. 238-249, 1992

[5] Thompson W. B, "Combining motion and contrast for segmentation", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.2 (6), pp. 543-549, 1980