Moving Object Tracking Method in Video Data Using Color Segmentation

Jae-Ho Lee, Soo-Hyun Cho, Whoi-Yul Kim
Image Engineering Laboratory, Hanyang University
E-mail : jhlee@vision.hanyang.ac.kr

칼라 분할 방식을 이용한 비디오 영상에서의 움직이는 물체의 검출과 추적

이 재 호, *조 수 현, 김 회 율
한양대학교 영상공학 연구실, *전자통신 연구원
전화 : 02-2290-0561 / 핸드폰 : 018-377-2383

Abstract

Moving objects in video data are main elements for video analysis and retrieval. In this paper, we propose a new algorithm for tracking and segmenting moving objects in color image sequences that include complex camera motion such as zoom, pan and rotating. The proposed algorithm is based on the Mean-shift color segmentation and stochastic region matching method. For segmenting moving objects, each sequence is divided into a set of similar color regions using Mean-shift color segmentation algorithm. Each segmented region is matched to the corresponding region in the subsequent frame. The motion vector of each matched region is then estimated and these motion vectors are summed to estimate global motion. Once motion vectors are estimated for all frame of video sequences, independently moving regions can be segmented by comparing their trajectories with that of global motion. Finally, segmented regions are merged into the independently moving object by comparing the similarities of trajectories, positions and emerging period. The experimental results show that the proposed algorithm is capable of segmenting independently moving objects in the video sequences including complex camera motion.

I. Introduction

Moving object segmentation and tracking technique became more important issue for the video contents analysis and retrieval. So a number of algorithms for segmenting moving objects have been proposed and developed. Conventional moving object segmentation algorithms include motion computation algorithms such as optical flow method, block matching method and affine motion method. Until recently, however, motion vector estimation has been computed by pixel-based or block-based operation. However, these pixel-based or block-based methods were known to have expensive computational complexity not to mention the accuracy in estimation especially in the homogeneous region. To avoid these problems region-based motion estimation method was proposed to replace pixel-based or block-based methods. Region-based motion estimation method computes motion vectors using regions that were pre-segmented by the feature similarity, i.e. color feature to segment regions in an image.

The steps of segmentation of independently moving objects in video sequences are illustrated in Fig. 1 and summarized below:

1. Similar color regions are segmented from the current frame and the next frame.
2. Each segmented region is matched between two frames by comparing features of segmented region such as color, position, area and shape of regions.
3. When a correspondence is established for the region, the motion vector of the region is then estimated.
4. Once all motion vector of regions are estimated, the global motion vector of the current frame is estimated by summing all motion vectors of all matched regions.
5. Repeat steps from 1 to 4 for all frame, adding global motion vector to make trajectory of global motion. Each region will have its trajectory by the same operation.
6. Independently moving regions are extracted by comparing their trajectories with the trajectory of the global motion.
7. The extracted independently moving regions are merged into the objects using the similarity of their trajectories, positions and emerging periods.

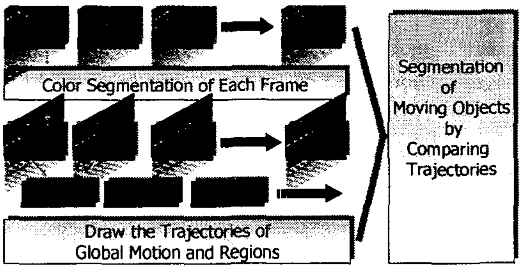


Fig. 1 Procedure of Moving Object Segmentation

II. Color Image Segmentation

To divide an image into a set of regions, color segmentation method by means of Mean-shift algorithm[1] is employed. Mean-shift algorithm proposed by Fukunaga and Hostetler[2], and recently generalized by Cheng[3] is a simple non-parametric procedure for estimating density gradients.

Assume, p(x) is the probability density function of the p-dimensional feature vectors x. A sphere S_x of radius r, centered on x contains the feature vector y such that $\| y-x \| \leq r$. The expected value of the vector z = y-x, given x and S_x is

$$\mu = E[x \mid S_x] = \int_{S_x} (y-x) p(y \mid S_x) dy = \int_{S_x} (y-x) \frac{p(y)}{p(y \in S_x)} dy \quad (1)$$

If S_x is sufficiently small then the volume of the sphere is computed by :

$$p(y \in S_x) = p(x)\, V_{S_x} \ \ where \ \ V_{S_x} = c \cdot r^p \quad (2)$$

The first order approximation of p(y) is

$$p(y) = p(x) + (y-x)^T \nabla p(x) \quad (3)$$

where $\nabla p(x)$ is the gradient of the probability density function in x. Then Eq. (1) becomes

$$\mu = \int_{S_x} \frac{(y-x)(y-x)^T}{V_{S_x}} \frac{\nabla p(x)}{p(x)} dy \quad (4)$$

since the first term vanishes. The value of the integral is

$$\mu = \frac{r^2}{p+2} \frac{\nabla p(x)}{p(x)} \quad (5)$$

or

$$E[x \mid x \in S_x] - x = \frac{r^2}{p+2} \frac{\nabla p(x)}{p(x)} \quad (6)$$

Thus, the mean shift vector, the vector of difference between the local mean and the center of the window, becomes proportional to the gradient of the probability density at x. The mean shift vector is a reciprocal to p(x) and small $\nabla p(x)$. On the other hand, regions with low density correspond to large mean shifts(also by small p(x) values). The mean moves toward the direction for the probability density maximum. Then the mean shift value is close to zero. This property can be exploited by a simple, adaptive steepest algorithm. The steps for the Mean-shift algorithm are presented below:

1. Choose the radius r of the search window.
2. Choose the initial location of the window.
3. Compute the mean shift vector and translates the search window by that amount.
4. Repeat until mean shift vector becomes close to zero.

To modify the Mean-shift algorithm for the color image segmentation, the Luv color space was chosen for the feature space, which was found satisfactory for image understanding applications[4].

The detail explanation of color image segmentation using the Mean-shift algorithm is addressed the Dorin[1].

The segmentation steps for color image results in the several groups of the significant colors and the region boundaries. To make clean boundaries of regions, morphological operation is used after color image segmentation. Then, small regions are removed because they have no contextual meaning. In Fig. 2, the procedure of color image segmentation is shown.

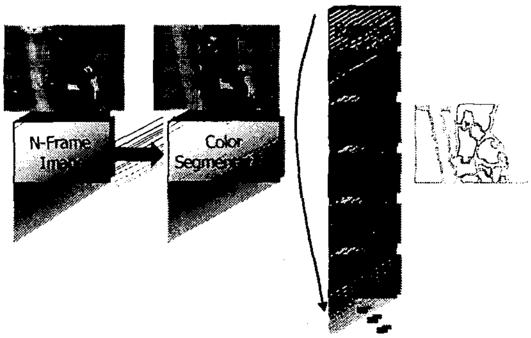


Fig. 2. Color Image Segmentation Procedure

III. Region Matching within Video Sequences

The region matching process is followed by color segmentation process of the current frame and the consecutive frames. Each labeled region is represented by feature values such as color, area, position and shape. The similarities, between labeled regions of the current frame and those of the next frame, are computed by summing distance of each feature.

$$Dist(I,\ J) = \sum_{k=0}^{n} W_k \cdot (I_k - J_k) \qquad (7)$$

where I is a region of the current frame

J is a region of the next frame

n is the feature number

W_k is the weight of each feature

The weights to the features are chosen by experiments. Each region of the current frame is estimated to compute the distances with the all labeled regions of the next frame. Then, one region that has minimum distance of the next frame is selected for the correspondence. Once region is matched, the motion vector of each region is estimated and stored to characterize its trajectory using the method that was presented by MPEG-7[5]. All motion vectors of regions are summed for compute the global motion. The global motion is then used to segment all independently moving objects. Fig. 3 shows the results of the each region matching.

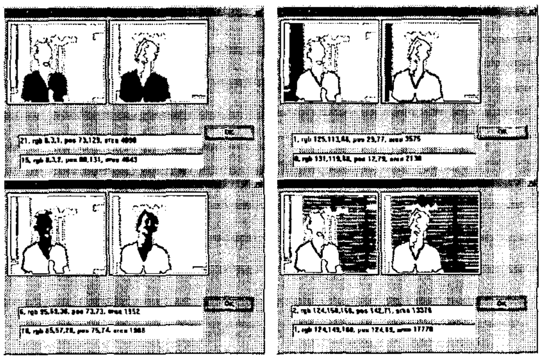


Fig. 3. The results of each region matching between i frame and i+k frame, where k is a time interval.

The color segmentation and region matching process are not applied in every frame to save the computation time. Time interval k is selected to reduce the operation complexity as well as to obtain resonable matching results.

IV. Moving Object Segmentation

After region matching process for the all frame, the global motion trajectory and trajectories of all regions are estimated and stored. For segmentation of independently moving objects, trajectory information are utilized. First step for object segmentation is the extraction of regions that have dissimilar motion trajectories with the global motion trajectory. The distance between the motion trajectory of each region with the global motion trajectory is computed by :

$$Dist(X, Y) = \frac{\sum_{n=s}^{f}(X_n - Y_n)}{f-s} \quad (8)$$

where

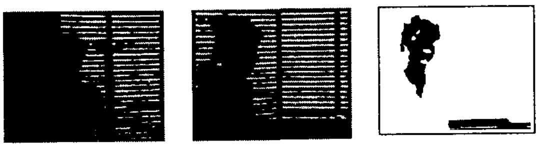
X_n is a trajectory of the region motion

Y_n is a trajectory of the global motion

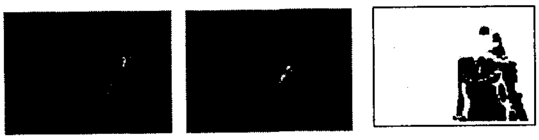
s is the first frame number of region appeared

f is the last frame number of region disappeared

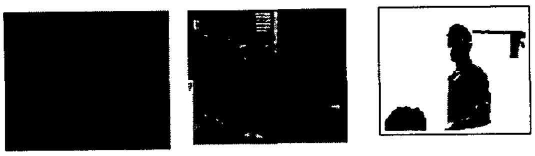
Once the independently moving regions are extracted, a region merging technique is followed to construct the object image. Region merging technique is a kind of classifying procedure of the similar trajectory region within the extracted regions. Computing similarity distance is similar to Eq. (8) by replacing i as the first frame number of common emerging period of compared regions and j as the last frame number of that period. Region merging procedure utilizes the estimated similarity distance to classify whether a region is in the same object or not. The results of independently moving object segmentation are shown in Fig. 4.



(a)



(b)



(c)

Fig. 4, Results of Independently Moving Object Segmentation,

In Fig. 4 images on the 1st column are the start frames, images on the 2nd are the end frames and images on the 3rd are the results. In Fig. 4, (a) camera is moving left while woman is moving right, in (b) camera is moving left while woman is moving and in (c) camera is moving left while man is moving left. The result images are captured by one frame within the result sequences for just display. But moving objects can be displayed (tracked) in every sequences that include moving objects because the trajectories of the regions are stored in a database.

V. Conclusion

In this paper, we proposed a new moving object segmentation method using Mean-shift color segmentation algorithm and stochastic criterion for region matching method. The proposed method solves the problem of inaccuracy estimation of motion vector using region-based matching. The proposed method using color feature can segment the independently moving object even with the complex moving background environments such as zoom, rotate and tilting. However, the propose method has still some problems in detecting accurate boundaries of objects and in classifying regions into exact objects. These problems will be studied further for more accurate segmentation.

References

[1] D. Comaniciu, P. Meer, "Mean Shift Analysis and Applications," IEEE Int'll. Conf. Comp. Vis., 1999.

[2] K. Fukunaga, "Introduction to Statistical Pattern Recognition," Academic Press, Boston, 1990.

[3] Y. Cheng, "Mean shift, mode seeking, and clustering," IEEE. Trans, Pattern Anal., Machine Intell., 1995.

[4] G. Wysjecki, W. S. Stiles, "Color Science: Concepts and Methods, Quantitative Data and Formulae," Willey, New York, 1982.

[5] "Motion Trajectory," MPEG-7 Document XM 8.0, 2000.