

A Motion Field Reconstruction Scheme for Smooth Boundary Video Object Segmentation

Jean Gao*, Ninad Thakoor**, and Sungyong Jung**

*Computer Science and Engineering Department

**Electrical Engineering Department

The University of Texas, Arlington

Texas, 76019, USA

Abstract—Motion segmentation is a classic and on-going research topic which is an important pre-stage for many video processes. The reliability of the motion field calculation directly determines either success or failure of such segmentation. Due to the aperture property of optical flow calculation, motion estimation at moving object boundary remains a challenging ill-posed problem. In this paper, we propose a reliable optical-flow estimation method from the view of missing data reconstruction. Furthermore, to overcome the occlusion problem at motion boundaries, we apply a motion segmentation scheme which integrates with spatial segmentation. Experimental results of proposed method and other's are presented.

I. INTRODUCTION

Motion segmentation is the process of categorizing pixels based on the homogeneities of their motions. It has been used for, but not limited to, layered representation of video clips, video object definition to improve compression efficiency, video surveillance, etc. As an example of designing an automatic video object detection and recognition system, knowledge of object shapes, which is the by-products of motion segmentation, plays an important role in the object classification or recognition.

During the past decades, extensive research has been done on motion segmentation. The most classic one would go to Wang and Adelson's affine-model based motion field clustering [1], in which adaptive K-means clustering was carried out for individual pixels to minimize the squared distance between the synthesized and estimated flow vectors. Borshukov et al. [2] combined affine clustering approach with dominant motion approach of Bergan et al. [12]. This approach proposed use of residual error over affine motion model and actual motion as criterion for a multistage merging procedure. Altunbasak et al. [4] used region based affine clustering with color information. This approach yields motion boundaries which match with color segmentation boundaries. These affine model based clustering algorithms are computationally less expensive and give comparable results with other approaches such as

motion segmentation by Hough transform analysis [5, 6], MRF modeling [7, 8], mixture modeling [9], simultaneous motion estimation and segmentation [10], dominant motion estimation [11, 12, 13], change-detection-mask analysis with temporal estimation [14,15] and normalized cuts [16].

For the optical flow based motion segmentation, it is clear that the success of the segmentation is tightly related to the accuracy of the optical flow estimation, whereas the latter suffers from errors due to occlusion and the aperture problem. Occlusion introduces uncertainty of optical flow calculation and results in possible miscalculation of motion vectors at motion boundaries. On the other hand, aperture problem comes from the fact that at each point only one motion constraint equation is given while two unknowns of motion vector exist. Due to the aperture problem it is not possible to calculate correct motion in the regions of image which are textureless. Failure to obtain correct motion estimation leads to errors in clustering during motion segmentation. There is no quantitative method for evaluating correctness of optical flow estimation. However, based on the texturedness one can choose pixels for which motion can be calculated reliably. But such calculation will generate sparse optical flow for textureless regions. On the other hand, exact object boundaries can be obtained only if we have reliable optical flow. This paper proposes methods to achieve both goals: reliable dense optical flow calculation and smooth boundary motion segmentation.

The majority of motion segmentation algorithms mentioned above are based on dense optical flows which are usually not reliable at motion boundaries because of occlusion and smoothness constraints. However sparse optical field based moving object segmentation will lead to clusters biased towards objects comparing with higher optical flow density due to the missing information in the flow field. In this paper, we will propose a motion segmentation algorithm which is an integration of good features selection, thus reliable optical flow estimation, motion vector clustering, and spatial color segmentation to achieve segmentation with as precise as possible moving object boundary description.

This paper is organized as follows. In Section II, good features are first selected based on texturedness and thus optical flow is calculated for these selected points. To overcome the sparseness problem of the optical flow field produced by good feature selection, in Section III we will propose a motion vector reconstruction scheme from the view of missing data evaluation. Section IV will describe an adaptive K-means procedure used for clustering of motion vectors. This motion vector clustering processing combines with spatial segmentation to achieve a smooth and visually meaningful segmentation result. Finally in Section V we will present our experimental results comparing to other researcher’s work, and we will close our paper by conclusion remarks.

II. GOOD FEATURURES BASED MOTION ESTIMATION

For the current various motion estimation techniques, they can be categorized as gradient based [24], correspondence-based, spatiotemporal energy-based [25] and phase-based methods [3]. Performance of these techniques is evaluated by Barron *et al.* [17].

The better performance and presence of high confidence measures [17] make Lucas-Kanade method [18] a logical choice for our problem. We use modified Lucas-Kanade method by Shi and Tomasi [19] which extends pure translational model to affine model of motion. To select pixels from image based on the texturedness we use “good features to track” criterion discussed in [19]. Good features are features which satisfy the measurement of appearance *dissimilarity* between two frames. Fig. 1 shows the two original images, and Fig. 2 demonstrates the estimated optical flow field.



Figure 1. Frames 136 and 137 of Calendar and Mobile Sequence.

III. MOTION RECONSTRUCTION

We use a pattern vector \mathbf{X}_n to represent the n^{th} pixel in an image

$$\mathbf{X}_n = (i_n, j_n, V_{xn}, V_{yn})^T \quad (1)$$

where coordinates (i_n, j_n) are position and (V_{xn}, V_{yn}) are optical flow estimates for the n^{th} pixel. Because of the texturedness nature of “good features” selection, as we can see from Fig. 2, the generated motion field tends to be

sparse. To reconstruct or restore the optical field at a reasonable degree, we approach this challenge as missing data problem in pattern recognition [20].

In [21], Dixon described several simple and general techniques for handling missing data. These techniques either eliminate part of the data and estimate the missing values, or compute an estimated value using the distance criterion.

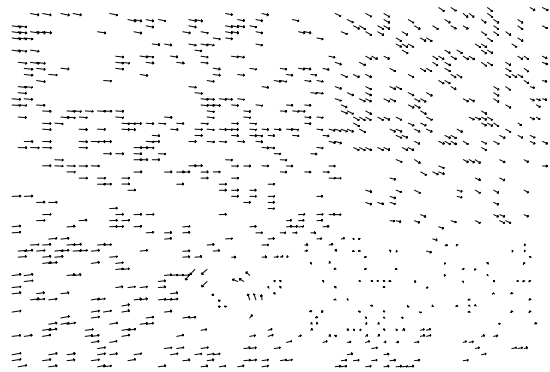


Figure 2. Optical flow calculated by Lucas-Kanade method.

Elimination of data should be used only when number of missing values is very small [21]. In the case of sparse optical flow the number of missing values is significant, hence elimination of data cannot be used. This means that running K-means algorithm directly on the sparse field will produce poor results.

Other technique used is to compute an estimated value based on the distance criterion. To use this technique, it is necessary to be able to calculate component of distance [20]. The distance criterion we chose for K-means algorithm is given by

$$d_{n_m} = (\bar{V}_{xn_m} - V_{xn})^2 + (\bar{V}_{yn_m} - V_{yn})^2 \quad (2)$$

where $(\bar{V}_{xn_m}, \bar{V}_{yn_m})$ are the optical flow estimates for the n^{th} pixel if optical flow model of m^{th} cluster is applied to the pixel. Calculations of these estimates will be discussed in section 4. Components of distance criterion are $(\bar{V}_{xn_m} - V_{xn})$ and $(\bar{V}_{yn_m} - V_{yn})$. None of the distance components can be calculated for sparse optical flow, as V_{xn} and V_{yn} are unavailable. Hence the distance criterion can not be used here.

Estimation of the missing motion vectors is suggested to be done using nearest neighbor information. It is a general observation that pixels belonging to the same spatial segments tend to have similar motions. The motion field reconstruction also uses spatial segmentation result which is a result of *Spedge-and-Medge* algorithm as shown in Fig. 3 [23]. The special segmentation method *Spedge-and-Medge* is an integration of well-known Canny edge detector, edge-linking for broken edge repair, and split-and-merge. As long as the spatial segments have few pixels for which optical

flow values are available, these values can be used to estimate the unavailable optical flow information.

Motion of individual spatial segments can be represented using a six parameter affine motion model:

$$V_{xn} = (a_{p,1} - 1)i_n + a_{p,2}j_n + a_{p,3} \quad (3)$$

$$V_{yn} = a_{p,4}i_n + (a_{p,5} - 1)j_n + a_{p,6} \quad (4)$$

where $a_{p,1}, a_{p,2}, \dots, a_{p,6}$ give affine motion parameters for p^{th} spatial segment C_p . Solution of these equations can be obtained using least squares method. But this method is not robust. We use robust least squares solution [22] to evaluate the values of affine parameters. Doing this assures that outliers, if any, are taken care of. After calculating the motion model, it can be applied to spatial segment and the missing optical flow can be reconstructed as

$$V'_{xn} = (a_{p,1} - 1)i_n + a_{p,2}j_n + a_{p,3} \quad (5)$$

$$V'_{yn} = a_{p,4}i_n + (a_{p,5} - 1)j_n + a_{p,6} \quad (6)$$

It is possible that for some of the spatial segments have insufficient or no motion information to build affine motion model. For real images such regions are smaller in proportion and tend to be smaller in size. Elimination of data can be used to deal with such regions. Fig. 4 demonstrates the reconstructed motion field.



Figure 3. Color segmentation result obtained from spedge-and-medge.

IV. CLUSTERING

Once the optical field is reconstructed, we apply an iterative spatial segmentation based K-means clustering for the motion segmentation. The parameters of the K affine motion models A_1, A_2, \dots, A_K are estimated by minimizing

$$\sum_p \sum_{n \in C_p} (\bar{V}_{xn_m} - V'_{xn})^2 + (\bar{V}_{yn_m} - V'_{yn})^2 \quad (7)$$

where

$$\bar{V}_{xn_m} = (a_{m,1} - 1)i_n + a_{m,2}j_n + a_{m,3} \quad (8)$$

$$\bar{V}_{yn_m} = a_{m,4}i_n + (a_{m,5} - 1)j_n + a_{m,6} \quad (9)$$

are estimated motion vectors and m is the labeling of motion segment for each pixel. m can take values from the set $\{1, 2, \dots, K\}$. Solution to this problem can be achieved using the region based affine clustering procedure proposed by Altunbasak et al. [4].

V. EXPERIMENTAL RESULTS AND CONCLUSIONS

Motion segmentation approach discussed in this paper was implemented and tested for variety of sequences. Hierarchical implementation of Lucas-Kanade algorithm with three levels of hierarchy was used to calculate the optical flow of the image. Optical flow was calculated at feature points selected by “Good features to track algorithm”. Images were then segmented using spedge-and-medge color image segmentation algorithm [23]. For sparse individual color regions with optical flow density above a threshold, the motion fields were reconstructed by the proposed method. The reconstructed optical flow later was used for carrying out K-means procedure in which robust least square method was used for the clustering. Fig. 5 shows the segmentation result by the proposed method, while Fig. 6 shows the result by Altunbasak et al. [4].

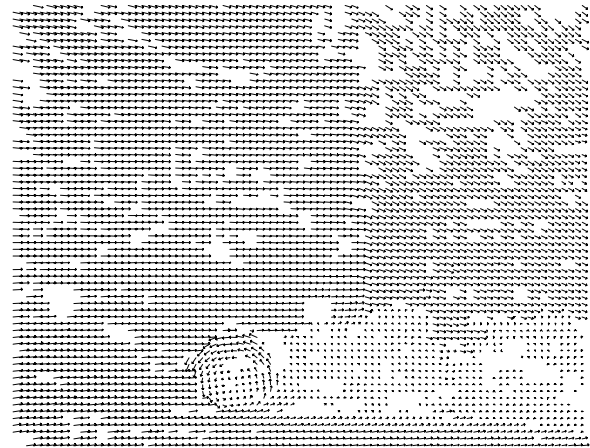


Figure 4. Dense optical field reconstructed by proposed method.

Motion segmentation is a classic and on-going research topic which is an important pre-stage for many video processes. In this paper we proposed a reliable motion field reconstruction method to the further generating of motion segmentation. The idea of sparse optical flow reconstruction by missing data evaluation is innovative and new. Furthermore, to achieve a smooth motion boundary, the temporal segmentation is carried out by assimilating the homogeneity of spatial region information.



Figure 5. Motion segmentation results by the proposed method.



Figure 6. Motion segmentation results by Altunbasak et al. [4] method.

REFERENCES

[1] J. Y. A. Wang and E. H. Adelson, Representing moving images with layers, *IEEE Trans. Image Process.* 3(5), Sept. 1994, pp. 625-638.

[2] G. D. Borshukov, G. Bozdagi, Y. Altunbasak, A. M. Tekalp, Motion segmentation by multistage affine classification, *IEEE Transactions on Image Processing*, Volume 6 Issue 11, Nov. 1997, pp. 1591-1594

[3] A. M. Tekalp, *Digital Video Processing*, Prentice Hall, New York, 1995.

[4] Y. Altunbasak, P. E. Eren, and A. M. Tekalp, Region-based parametric motion segmentation using color information, *Graphical models and image processing*, January 1998, pp. 13-23.

[5] G. Adiv, Determining three-dimensional motion and structure from optical flow generated by several moving objects, *IEEE Trans. Pattern Anal. Machine Intell.* 7, 1985, pp.384-401.

[6] M. Bober and J. Kittler, On combining the Hough transform and multiresolution MRF's for the robust analysis of complex motion, in *Proc. Second Asian Conf. on Computer Vision (ACCV)*, Dec. 1995.

[7] D. W. Murray and B. F. Buxton, Scene segmentation from visual motion using global optimization, *IEEE Trans. Pattern Anal. Machine Intell.* 9(2), Mar. 1987, pp. 220-228.

[8] J. M. Odobez and P. Bouthemy, Direct model-based image motion segmentation for dynamic scene analysis, in *Proc. Second Asian Conf. on Computer Vision (ACCV)*, Dec. 1995.

[9] Y. Weiss and E. H. Adelson, A unified mixture framework for motion segmentation: Incorporating spatial coherence and estimating the number of models, in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, June 1996.

[10] M. M. Chang, A. M. Tekalp, and M. I. Sezan, An algorithm for simultaneous motion estimation and scene segmentation, in *Proc. IEEE Int. Conf. Acoust. Speech Sign., Adelaide, Australia*, April 1994.

[11] M. Hotter and R. Thoma, Image segmentation based on object-oriented mapping parameter estimation, *Signal Process.* 15, 1988, pp. 315-334.

[12] J. R. Bergen, P. J. Burt, and K. Hanna, Dynamic multiple-motion computation, in *Artificial Intelligence and Computer Vision* (Y. A. Feldman and A. Bruckstein, Eds.), pp. 147-156, Elsevier, Amsterdam, 1992.

[13] M. Irani and S. Peleg, Motion analysis for image enhancement: Resolution, occlusion, and transparency, *J. Visual Commun. Image Repr.* 4(4), Dec. 1993, pp. 324-335.

[14] R. Mech and P. Gerken, Automatic Segmentation of Moving Objects, *Doc. No. ISO/IEC JTC1/SC29/WG11 MPEG 96/1188*, Sept. 1996.

[15] G. Russo and S. Colonnese, Automatic Segmentation Techniques, *Doc. No. ISO/IEC JTC1/SC29/WG11 MPEG 96/1181*, Sept. 1996.

[16] S. Jianbo and J. Malik, Motion segmentation and tracking using normalized cuts, *Proceedings of Sixth International Conference on Computer Vision*, 1998, pp.1154-1160.

[17] J. L. Barron, D. J. Fleet, S. S. Beauchemin, T.A. Burkitt, Performance of optical flow techniques, *Proceedings Computer Vision and Pattern Recognition*, 1992, pp.236-242.

[18] B. Lucas and T. Kanade. An iterative image registration technique with application to stereo vision, *Proc. Image understanding workshop*, 1981, pp. 121-130.

[19] S. Jianbo and C. Tomasi, Good features to track, *Proc. IEEE Comput. Soc. Conf. Computer Vision and Pattern Recognition*, 1994, pp. 593-600.

[20] A. K. Jain and R. C. Dubes, *Algorithms for clustering data*, Prentice Hall, 1988.

[21] J. K. Dixon, Pattern recognition with partly missing data, *IEEE Trans. on Systems Man and Cybernetics*, Vol. SMC 9, 1979, pp. 617-621.

[22] P. J. Huber, *Robust statistics*, Wiley, New York, 1981.

[23] J. Gao and A. Kosaka, "A color image segmentation editor by the integration of edge linking, and energy minimization," *IEEE Proceedings of ICIP*, 1999.

[24] W. Enkelman, "Investigations of multigrid algorithms for the estimation of optical flow fields in image sequences," *Computer Vision and Graphical Image Processing*, vol. 43, 150-177, 1988.

[25] D. J. Heeger, "Optical flow using spatiotemporal filters," *Int. J. Computer Vision*, vol. 1, 279-302, 1988.