CVPR
#***


# Motion From Boundary

Anonymous CVPR submission
Address Line2
Paper ID ***

## Abstract

*Classical motion estimation algorithms are based on unambiguous features with no aperture problem, and thus can not handle textureless examples and may be confused by T-junctions. We propose a boundary-based approach to estimate motion, relying on the energy of a multi-hypothesized graph instead of trusting T-junctions or corners. The hypothesis comes from possible connections between two ends of the graph as illusory boundaries. By minimizing the energy on the graph we obtain reliable estimation for textureless examples even with concolorous occlusions.*

## 1. Introduction

Understanding and analyzing motion from video sequence is one of the key problems in computer vision. The typical algorithms go from low level optical flow, such as Lucas-Kanade [5] and Horn-Schunck [3], to motion segmentation, such as POEM [13]. Although many improvements have been made to these methods, such as combining global and local constraints [1], incorporating sophisticated prior of optical flow field [11], and corner-based layer segmentation [4], the foundation of the current motion estimation algorithms are based on unambiguous features with no aperture problem [12]. There is no ambiguity for the matching of these features, and the matching is propagated to obtain the complete flow field.

These algorithms perform well for images full of texturedness or corners, but poorly for textureless images. A simple example is illustrated in Figure 1 (a), where the gray bar moves to the right and the black to the left. There are in total 12 unambiguous feature points, of which four are T-junctions and eight are corners. The matchings of the eight corners are correct, whereas the matchings of the four T-junctions are spurious. If we trust the matching of the four T-junctions, we shall inevitably introduce error in motion estimation.

This example becomes more tricky as the gray bar becomes black, too, as shown in Figure 1 (b). Even layer-based motion estimation algorithm may "see" this sequence as a stretching polygon with four antennas, but human perception would still see two bars moving separately [9]. This example is not fabricated at all. The occlusion of the dancer's legs in Figure 2 shows the real challenge. This contrast between machine and human perceptions enforces us to reflect on the foundation of the vision algorithms.

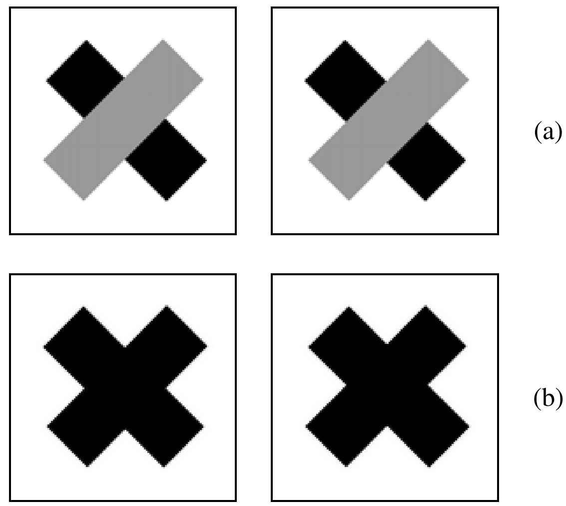


Figure 1. Classical motion estimation methods can hardly solve very simple examples. (a) is an examples from Weiss and Adelson [13]. The results in [13] are poor on the boundary. When the color of the frontal bar becomes the same as the back one, most motion analysis algorithms will fail to analyze that there are two bars moving separately as human perception does [9]

.

A natural way to overcome this problem is to eliminate T-junctions in motion estimation. However, T-junction detection is a difficult task even for human, and sometimes a region as large as 50 pixels in diameter is needed to recognize T-junctions [7]. It is also discovered in [8] that "local junction structure, per se, has relative little explanatory power" in understanding motion.

The recent work in human perception [9] illustrated that in motion estimation "what matters is ... whether illusory contours are introduced when the junction category is changed." This motivates us to modify the foundation of

Figure 2. The toy example in Figure 1 (b) is not fabricated. The intersection of the legs of the dancer is a real challenge for motion analysis.

motion estimation with illusory contours. There are many possible ways to generate hypothesis of illusory contours from a single image, but there is only one, if not more, that can explain two frames perfectly.

There are many problems to solve under this contour-based framework. Though a lot work has been done on boundary tracking and detection from a single image, e.g. [2, 10, 6], it is still a challenging problem to obtain a complete and accurate representation for image boundary. Since boundary detection is outside the interest of this paper, we focus on high-contrast boundaries and propose an algorithm to efficiently locate boundary particles to form a graph...

## 2. Graph Boundary Candidate

## 3. Object Function and Energy Minimization

## 4. Experimental Results

## 5. Conclusion

## References

[1] A. Bruhn, J. Weickert, and C. Schnörr. Lucas/kanade meets horn/schunck: Combining local and global optic flow methods. *International Journal on Computer Vision*, 61(3):211–231, 2005. 1

[2] J. Canny. A computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986. 2

[3] B. Horn and B. Schunk. Determing optical flow. *Journal of Artificial Intelligence*, 17:185–204, 1981. 1

[4] C. Liu, A. Torralba, W. T. Freeman, F. Durand, and E. H. Adelson. Motion magnification. In *Computer Graphics, Proc. SIGGRAPH*, pages 519–526, 2005. 1

[5] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Image Understanding Workshop*, pages 121–130, 1981. 1

[6] D. R. Martin, C. C. Fowlkes, and J. Malik. Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(5):530–549, 2004. 2

[7] J. McDermott. Psychophyiscs with junctions in real images. *Perception*, 33:1101–1127, 2005. 1

[8] J. McDermott and E. Adelson. The geometry of the occluding contour and its effect on motion interpretation. *Journal of Vision*, 4(10):944–954, 2004. 1

[9] J. McDermott and E. Adelson. Junctions and cost functions in motion interpretation. *Journal of Vision*, 4(7):552–563, 2004. 1

[10] P. Pérez, A. Blake, and M. Gangnet. Jetstream: Probabilistic contour extraction with particles. In *Proc. IEEE Int'l Conf. Computer Vision*, pages 524–531, 2001. 2

[11] S. Roth and M. Black. On the spatial statistics of optical flow. In *Proc. IEEE Int'l Conf. Computer Vision*, pages 42–49, 2005. 1

[12] J. Shi and C. Tomasi. Good features to track. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*. 1

[13] Y. Weiss and E. Adelson. Perceptually organized em: A framework for motion segmentation that combines information about form and motion. Technical Report 315, MIT Media Lab Perceptual Computing Section, 1994. 1