Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Feature Selection, Matching, and Evaluation for Subcellular Structure Tracking

Quan Wen , Jean Gao, and Kate Luby-Phelps

Abstract— **Understanding the motility of subcellular particles like organelles, vesicles, or mRNAs is critical to understand how cells regulate delivery of specific proteins from the site of synthesis to the site of action. The goal of this paper is to present a framework of feature selection, matching, and evaluation for the segmentation and tracking of green fluorescent protein (GFP) labeled subcellular structures. To select stable and distinctive features for small-sized subcellular particles, a grid-based minimum variance (GMV) feature selection method is proposed. To robustly keep tracking of the selected features, we propose a mean minimum to maximum ratio (MMMR) similarity measure for feature matching. In order to quantitatively evaluate the proposed methods, we define two evaluation criteria, feature convergence rate (*FCVR*) and feature consistence rate (*FCSR*), which conform with the proximity and similarity properties of Gestalt visual perception theory. Our technique was validated on real confocal video data with comparison to traditional feature selection and matching methods.**

I. Introduction

The introduction of high resolution electronic imaging devices and the development of methods to tag proteins of interest by green fluorescent protein (GFP) have been the drivers to live cell study. Understanding the motility of these subcellular particles like organelles, vesicles, or mRNAs is critical to understand how cells regulate delivery of specific proteins from the site of synthesis to the site of action at subcellular level. The knowledge of regulation and how it is deranged in various diseased or malfunctioned states will eventually lead to a better understanding of such diseases as diabetes, hypercholesterolemia, and many viral infections. At present, biologists either laboriously track a few vesicles by hand, or use commercially available particle tracking programs whose performance is far below expectations for various demands.

Certain body of work has addressed the segmentation and tracking of cell movement. Methods diverse from simple model-free heuristic intensity thresholding [1] to model-based active contour tracking [2]. Relatively little effort has been focused on subcellular tracking. Cell tracking is much easier than subcellular tracking due to restricted cell movement and larger object size. The main technical challenge for subcelluar structure mobility analysis comes from small object size and nearly homogeneous object appearance. This makes the commonly developed wealth of object tracking techniques can not be directly applied here. For instance, conventional template-based feature matching methods [1], [3], work well in a sparse object environment and are restricted to non-deformable object tracking. Furthermore, large inter-frame displacement and noise mislead the curve/surface evolution of the active contour/surface-based tracking method [4].

To develop tracking methods that are robust in differentiating objects from background, one of the important issues is the selection of stable and distinctive features. Such features may include shape corners, edges, contours, spectral parameters, color histograms, or intensity of objects. While extensive efforts have been devoted to better and robust tracking framework, limited work has been done on distinctive feature selection for tracking process. To tackle feature selection in a low-resolution and noisy image, recent endeavor has been presented by Collins *et al.* [5]. In the paper, features that best distinguish object from background are defined to be the best for tracking. The method selects features by linear combination of RGB color components and by maximizing the contrast between foreground and background. However, the unique properties of GFP particles limit the extension of this approach being applied here. First, GFP image is gray scale (pseudo color) and cannot provide color space to be mixed for new features. Second, since the objects in GFP image are similar and close to each other, it is difficult to define a background near the object of interest. This causes the variance ratio criterion to fail to find highly discriminable features, due to the size of the background required by the moving speed of the object and the amount of similar distracters are both large. Last, the method is not suitable for handling larger shape deformations due to the limitation of bounding rectangular box.

Based on the properties of the GFP image sequences, we present a framework of feature selection by grid-based minimum variance (GMV) and feature matching by mean minimum to maximum ratio (MMMR) for the segmented particle of interest. The tracking process is carried out by our previously-developed particle filter region tracking approach [6]. To evaluate the performance of the proposed feature selection and matching techniques, two evaluation criteria based on the Gestalt visual perception theory are introduced. The first evaluation criterion, called feature convergence rate (*FCVR*), conforms with proximity law of Gestalt psychology to reflect nearness of grouped features. The second criterion, termed as feature consistence rate (*FCSR*), reveals the similarity property of Gestalt theory.

The rest of the paper is organized as follows. In Section

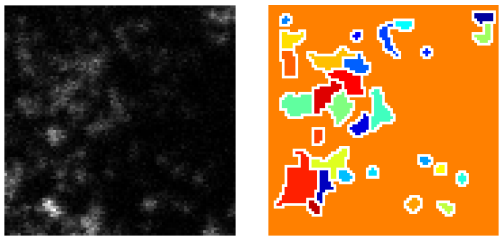
Q. Wen and J. Gao are with the Department of Computer Science and Engineering, The University of Texas at Arlington, Arlington, TX 76109 {qwen,gao}@cse.uta.edu

K. Luby-Phelps is with the Department of Cell Biology, The University of Texas Southwestern Medical Center, Dallas, TX 75390 kate.phelps@utsouthwestern.edu

II, we describe the grid-based minimum variance (GMV) feature selection method. The mean minimum to maximum ratio (MMMR) feature matching is presented in Section III. Feature convergence rate (*FCVR*) and feature consistence rate (*FCSR*) are defined in Section IV. In Section V, comparison results are presented. Finally, conclusion and future work are given in Section VI.

II. Grid-Based Minimum Variance Feature Selection

Before applying the feature selection method, the marker-guided watershed method [7] is applied to segment out the particle of interest, as shown in Fig. 1.



(a) Original image (b) Segmented image

Fig. 1. Result of watershed segmentation method.

From the segmented particle of interest, intuitively all the pixels of the whole region would be utilized for the consequent particle tracking. However, not every pixel conveys useful and robust information for the tracking process. The usage of the whole region is catastrophic for image sequences acquired under undesirable conditions. Therefore, region tracking problem is often relaxed to the process of sub-region tracking stipulated by robust features. There are several methods used in literature to represent region features, such as texture [8], histogram [9], salient points [10]. Nonetheless, the characteristics of GFP image are unlike common images well studied. High frequency features such as edges and corners are not reliable in the GFP tracking due to large shape deformation and dense object environment. Furthermore, the segmentation results are not always ready for direct application of the above methods.

Based on the Gestalt law of visual perception, we define four criteria for robust feature selection in GFP images : 1) homogeneity; 2) stability; 3) majority; 4) proportion. For homogeneity, the features of GFP object are defined as the pixels with the smallest intensity variance within its local support L. The intensity variance σ_L^2 of a feature pixel with its local support is calculated as:

$$\sigma_L^2 = \frac{1}{N_L}\sum_{p\in L}(I(p)-\mu_L)^2 \qquad (1)$$

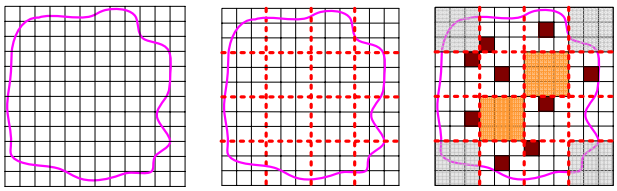
where $I(.)$ is the intensity of the pixel, N_L is the number of pixels in L, and μ_L is the intensity mean within L and calculated as: $\mu_L = 1/N_L \sum_{p\in L} I(p)$. Considering GFP objects illuminate themselves, stable feature point and its local support should have higher intensity than the background and other darker area of the same GFP object. This intensity stability threshold is denoted as I_h. Majority means that besides feature point itself, larger proportion of local support, say larger than the majority threshold T, should belong to the GFP object. The forth criterion is from the histogram point of view, and implies that the larger region should have more feature points than the small region, if the first three criteria are satisfied. The proposed grid-based minimum variance (GMV) method takes into consideration of both the region segmentation and reliability of each sub-region to be tracked, and can be summarized as follows:

Algorithm of GMV Feature Selection

1) Estimate the stability threshold I_h, majority threshold T, and the size of each $n \times n$ sub-region.
2) Find the minimum bounding rectangle (MBR) of the segmented object. If the length or the width of the MBR is not a multiple of n, extend it to be multiple of n by expanding the MBR towards right or bottom of the image with the minimum increase, called as extended minimum bounding rectangle (EMBR).
3) Uniformly divide MBR or EMBR into sub-regions by squares each with size $n \times n$.
4) Repeat to all the sub-regions of MBR or EMBR
   a) Construct candidate features set Ω by selecting the object pixel satisfying criteria 2 and 3.
   b) If Ω is empty, go to step 4) for next sub-region.
   c) If Ω is not empty, apply (1) to each pixel in Ω, and select the one with smallest σ_L^2 as feature point. For duplicate candidates, we select the one with larger μ_L. If the previous condition does not work, we select one with larger number of object pixels in its local support.
5) End Repeat.

Supposing that a GFP object is already segmented as shown in Fig. 2(a), the object region is uniformly divided into sub-regions, each with size of $n \times n$, say 3 by 3 red dotted square as Fig. 2(b). This way will guarantee criterion 4, since large region will have more sub-regions. Then within each sub-region, we select all the points satisfying criteria 2 and 3, and choose the one with the lowest local support variance σ_L^2 as the feature point to represent the sub-region in order to comply with criterion 1. The selected feature points are shown in Fig. 2(c). Using GMV, it is possible that one sub-region has no feature point extracted as shown in the two orange regions of Fig. 2(c), because the local support intensities of all the pixels in the sub-regions are less than stability threshold I_h and violate criterion 2. This is the desired result and is one of the advantages over simple uniformly sampling of the region, which results in feature points unreliable for tracking. Also, the gray-shaded sub-regions have no feature points due to the numbers of pixels in them are less than the majority threshold $T = 5$. The superior of GMV lies in its consistence and convergency of correspondences for GFP cluster tracking.

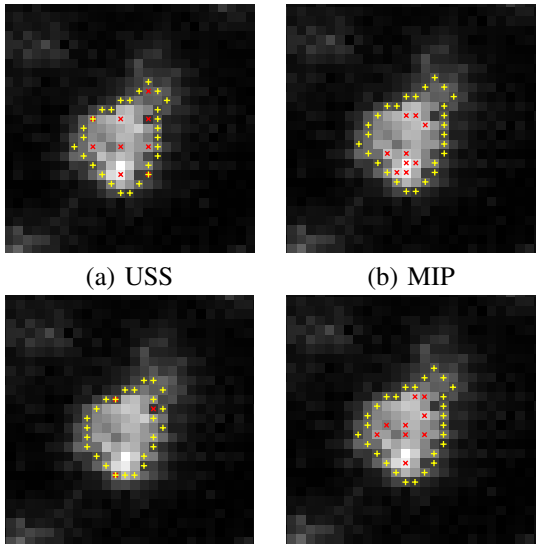
We compare GMV feature selection method with three other often used feature selection methods, namely uniform sub-sampling (USS) method, maximum intensity (MIP)



(a) (b) (c)

Fig. 2. Illustration of GMV feature selection method. (a) Segmented object with its boundary (purple line). (b) The result of applying grid lines (red doted lines) on (a). (c) Feature selection result by GMV method. Brown pixels are the selected feature points .

method, and Harris corner detector [11]. USS is a method of uniformly sub-sampling the region of interest. MIP method uses the pixels with highest intensity as features. The number of feature pixels in MIP is set to be the same as USS method. As can be seen in Fig. 3, the selected feature points of different methods have different patterns. In Fig. 3(a), USS method does not consider the local pixel information such as intensity or gradient, so the selected feature pixels have no obvious patterns except they are uniformly distributed. For MIP method, the feature points aggregate in the high intensity parts, as shown in Fig. 3(b). There are only three feature points detected by the Harris corner detector since it cannot control the exact number of corners. Nonetheless, feature points of GMV trend to represent the intensity composition of the whole region.



(a) USS (b) MIP

(c) Harris corner (d) GMV

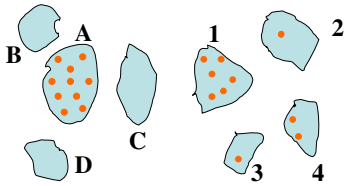
Fig. 3. Comparison of feature selection methods. The red cross and yellow plus signs indicate the positions of the selected feature points and the boundary pixels respectively.

III. Similarity Measure for Feature Matching

One important aspect in feature point based object tracking is how to measure the similarity between the feature and its correspondence. The mostly used similarity measure is normalized cross-correlation coefficient (*NCC*). This method works well for feature points demonstrating translational motion and with less intensity variation, but it is sensitive to feature points undergoing affine motion. To overcome this, a different similarity measure between two feature points p_s and p_t is defined as:

$$S(p_s, p_t) = \frac{1}{|L|} \sum_{p \in L} \frac{\min(I_k(p), I_n(p + p_t - p_s))}{\max(I_k(p), I_n(p + p_t - p_s))} \quad (2)$$

where L is the local support of p_s, p is point in L, $|.|$ is the cardinal operation. I_k, I_n represent intensity of frame k, n respectively. p, p_t, and p_s are in coordinate format. In the case when $\max(x, y) = 0$, we set $\min(x, y)/\max(x, y) = 1$. We call $S(p_s, p_t)$ mean minimum to maximum ratio (MMMR).



(a) Frame t-1 (b) Frame t

Fig. 4. Illustration of feature convergence. Region A in (a) has ten feature pixels indicated by orange dots. Its feature pixels' correspondences scatter in different regions in (b), indicated by orange dots.

IV. Feature Evaluation Methods

As mentioned in Section I, the convergence and consistence of selected features are important in tracking objects having characteristics as GFP cluster. The discrimination and robustness of the features are two major concerns. We interpret the discrimination and robustness of features points selected in region tracking as feature convergence rate (*FCVR*) and feature consistence rate (*FCSR*). The idea of *FCVR* is illustrated in Fig. 4. Suppose there are four regions *A*, *B*, *C*, and *D* in frame $t-1$ as shown in Fig. 4(a), we denote them as R_A, R_B, R_C, R_D respectively. Similar to frame t as shown in Fig. 4(b), its four regions are represented by R_1, R_2, R_3, R_4. The tracked region is R_A with N_{fA} feature points. The correspondence of each feature point of region R_A is searched in frame t and scatters in R_1, R_2, R_3, R_4, with each region having a number of N_{f1}, N_{f2}, N_{f3}, N_{f4} feature points respectively. When the ground truth is that region R_1 is the correspondent region, we define the *FCVR* of the feature points selected at frame $t-1$ for region R_A as:

$$FCVR = \frac{N_{f1}}{N_{fA}} \quad (3)$$

For the example in Fig. 4, $N_{f1} = 6, N_{fA} = 10$, and $FCVR = 0.6$. The higher the *FCVR*, the better the feature selection method. Since *FCVR* weighs more on features grouped together in the same region, it conforms with the proximity property of Gestalt perception criterion.

On the other hand, *FCSR* is defined as the mean similarity value between the correspondent points in the correspondent region and their original feature points of the tracked region. It reflects the similarity law of Gestalt theory. Different from

the correspondence search area of *FCVR*, which includes R_1, R_2, R_3, R_4 as in Fig. 4, we limit that of *FCSR* only in one region, for instance R_1. Thus, the *FCSR* is defined as:

$$FCSR = \frac{1}{N_{fA}} \sum_{p_s \in R_A, C(p_s) \in R_1} S[p_s, C(p_s)] \qquad (4)$$

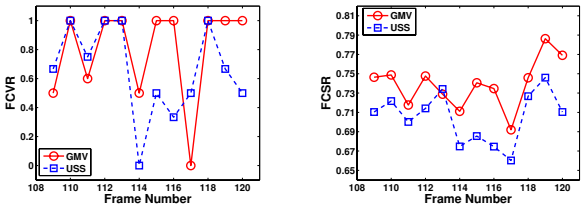
where $S(.)$ is the same as (2), $C(p_s)$ is the correspondence of p_s searched only in region R_1. Same as *FCVR*, the higher the *FCSR*, the better the feature selection method.

As can be seen from the definitions of *FCVR* and *FCSR*, we do not emphasize the discrimination and robustness of each individual feature. Instead, what we are most interested in is the statistical properties of the whole feature set embedded within the region of interest.

V. Experiment Results

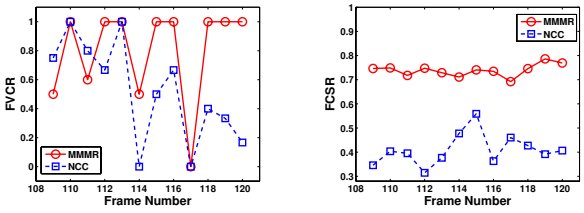
The experiment image sequences is acquired by Leica TCS-SP laser scanning confocal microscope with a 100X objective lens. The goal of this study is to track movement and analyze velocity of caveolin-1-GPF which is a GFP tagged caveolin-1 protein. The video images are of size 512x512 with a time interval of 1.2 sec between two frames. The experiment results are obtained by integrating the feature selection method into a framework of region tracking by particle filter (PF) [6].

Since the performance of both MIP and Harris corner detector is far behind the GMV and USS, we only present the the numerical comparison results between the last two, as shown in Fig. 5. In case one method fails tracking the object correctly, we manually correct it to continue the tracking process. The means of *FCVR* and *FCSR* for GMV are 0.80 and 0.74 respectively. For USS, these values are 0.66 and 0.70. To evaluate our similarity measure MMMR, we



(a) (b)

Fig. 5. Results of feature selection methods comparison.



(a) (b)

Fig. 6. Results of similarity measures comparison.

compare it with the most used normalized cross-correlation coefficient (*NCC*), with the results illustrated in Fig. 6. The advantage of MMMR over *NCC* is obvious.

VI. Conclusion And Future Work

In the paper, we presented a framework of feature selection, matching, and evaluation criterion for tracking GFP labeled subcellular structures. Comparison results are presented for different approaches. The experimental results validated the performance of the proposed techniques in feature selection and similarity measure. The framework conveys the potential to analyze other visually-similar scientific image sequence data such as celestial observation and meteorologic phenomena. In the future work, we plan to extend the current method to deal with situations when a single particle splits into several ones or when multiple particles merge together.

VII. Acknowledgments

The authors are grateful for the support from National Science Founddation (NSF) under grant IIS-0546605.

References

[1] V. Awasthi, K. W. Doolittle, G. Parulkar, and J. G. McNally, "Cell tracking using a distributed algorithm for 3-d image segmentation," *Bioimaging*, vol. 2, no. 2, pp. 98–112, 1994.

[2] C. Zimmer, E. Labruyere, V. Meas-Yedid, N. Guillen, and J.-C. Olivo-Marin, "Segmentation and tracking of migrating cells in videomicroscopy with parametric active contours: a tool for cell-based drug testing," *IEEE Transactions on Medical Imaging*, vol. 21, no. 10, pp. 1212–1221, Oct 2002.

[3] F. Briquet-Laugier, C. Boulin, and J.-C. Olivo-Marin, "Analysis of moving bioligical objects in video mircorscopy sequences," in *Proc. IS&T/SPIE Conference on Capture, Analysis and Display of Image Sequences*, vol. 3642, San Jose, CA, Jan. 1999, pp. 4–11.

[4] A. Dufour, V. Shinin, S. Tajbakhsh, N. Guillen-Aghion, J.-C. Olivo-Marin, and C. Zimmer, "Segmenting and tracking fluorescent cells in dynamic 3-d microscopy with coupled active surfaces," *IEEE Transactions on Image Processing*, vol. 14, no. 9, pp. 1396–1410, Sept. 2005.

[5] R. T. Collins, Y. Liu, and M. Leordeanu, "Online selection of discriminative tracking features," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1631–1643, 2005.

[6] Q. Wen, J. Gao, and K. Luby-Phelps, "Region based tracking of protein molecules," in *IEEE International Symposium on Biomedical Imaging*, April 6-9 2006.

[7] L. Vincent, "Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms," *IEEE Transaction on Image Processing*, vol. 2, no. 2, pp. 176–201, April 1993.

[8] Y. Liu, X. Zhou, and W. Ma, "Extracting texture features from arbitrary-shaped regions for image retrieval," in *IEEE International Conference on Multimedia and Expo*, vol. 3, June 27-30 2004, pp. 1891–1894.

[9] M. J. Swain and D. H. Ballard, "Color indexing," *Internaltional Journal of Computer Vision*, vol. 7, no. 1, pp. 11–32, 1991.

[10] N. Sebe, Q. Tian, E. Loupias, M. Lew, and T. Huang, "Color indexing using wavelet-based salient points," in *IEEE Workshop on Content-based Access of Image and Video Libraries*, June 12 2000, pp. 15–19.

[11] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proceedings of The Fourth Alvey Vision Conference*, Manchester, 1988, pp. 147–151.