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

Tracking Endothelial Cells using Multiframe Point Correspondence

Fabrizio Lamberti, *Member*, *IEEE*, Bartolomeo Montrucchio, *Member*, *IEEE*

Abstract—**This paper investigates the problem of tracking individual circular cell nuclei of endothelial cells during their migration. The purpose is to extract statistical information to support research activities devoted to reaching a better understanding of the underlying biological phenomena. This goal is achieved by exploiting a general tracking framework based on a multiframe point correspondence approach and capable of translating the matching problem into a maximum path cover problem over a graph based representation of moving objects. The application to phase contrast microscopy images containing hundreds of cells aggregating to form capillary tubes required to design an ad-hoc weight function capable of properly handling cells occlusion and aggregation.**

I. INTRODUCTION

UNDERSTANDING how blood vessels form in the organism has been a central issue in biomedical research during the last decade. The ability to form networking capillary tubes is a cell autonomous property of endothelial cells (ECs). In in-vitro essays it can be observed that randomly seeded ECs migrate over distances which are an order of magnitude larger than their radius and aggregate when they get in touch with one of their neighbors. In a time of the order of ten-twelve hours they form a continuous multicellular network. Today, many efforts are devoted to investigate the biological principles regulating ECs migration according to chemotaxis phenomena. Research activities require an extensive analysis of experimental data. However, the quantity as well as the complexity of microscopy images obtained through laboratory experiments renders manual analysis unreasonably time-consuming. Therefore, the availability of computerized techniques capable of extracting quantitative information from image data would constitute an invaluable support. In this paper we present the work that have been carried out toward the design of an automatic techniques for the analysis of large amount of experimental data, consisting of phase contrast video sequences showing ECs in Matrigel in the earlier phases of the process leading to vascular network assembly.

II. BACKGROUND

Many algorithms for tracking of moving objects have been presented in the literature. Among these, a key role is played by *active contours* of *snakes*, which proved to produce outstanding results also in many cell tracking applications. Nevertheless, such techniques show marked difficulties in handling disappearing objects that can be very common because of motion characteristics as well as of optical registration technique. Furthermore active contours do not perform well when objects overlap or aggregate [1],[2]. Problem above can be effectively addressed though combinatory approaches like point correspondence capable of taking decisions based on global motion information. A large number of correspondence methods have been proposed in recent years. Most of these methods first define a motion model and use some optimization technique to maximize (minimize) a gain (cost) function based on that motion model. These methods differ by the choice of motion model, optimization technique and/or gain function. Most of them are only capable of working on consecutive frames, do not allow to handle occluding cells neither cells entering and exiting the observation area and often rely on the knowledge of statistical information or precise initialization conditions [3]. Recently a novel algorithm exploiting multiframe information to establish moving point correspondences suitable for fine tracking of generic particle systems has been designed [4]. Nevertheless, the proposed approach relies on the availability of information related to mechanism regulating objects motion that are not available in the specific case being considered. In this paper, we present an extension of this general framework to handle tracking of in-vitro migrating and aggregating ECs.

III. THE MULTIFRAME POINT CORRESPONDENCE APPROACH

The output of a correspondence algorithm is a set of tracks, where each track ideally corresponds to a unique point or a object in the real world and specifies its position in every frame of the sequence. The problem can be formulated as follows. Let's define a sequence of n frames F_{t_i} each corresponding to a time instance t_i with $1 \le i \le n$. Let's call $X_i = \left\{ x_1^i, x_2^i, \ldots, x_r^i \right\}$ the set of r points belonging to frame F_{t_i}. We define a track T of length m to be a sequence of m points $\left\langle x_{a_1}^{i_1}, x_{a_2}^{i_2}, \ldots, x_{a_m}^{i_m} \right\rangle$ such that $1 \le i_i \le i_2 \le \ldots n$ and $1 \le a_j \le \left| X_{i_j} \right|$. The backward correspondence of a point $x_{a_j}^{i_j}$ in track T is defined by the point preceding $x_{a_j}^{i_j}$ while the forward correspondence of a point is the point succeeding $x_{a_j}^{i_j}$. The problem is to find a set of tracks $A = \left\langle T_1, T_2, \ldots, T_m \right\rangle$ such that for each $T_i \in A$ each real-world

---

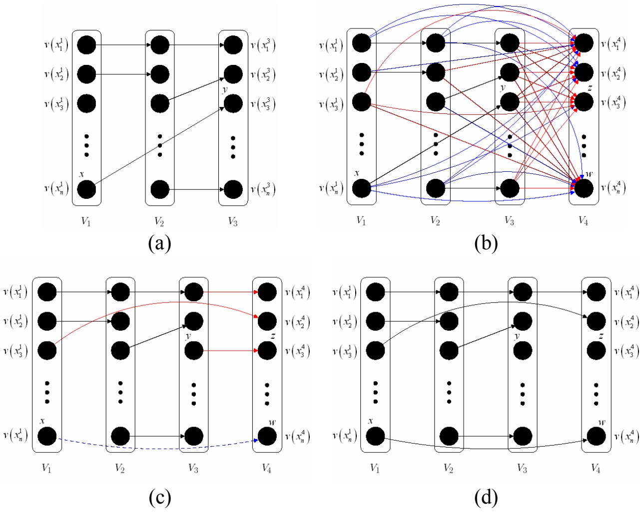
Manuscript received April 24, 2006.

F. Lamberti is with the Dipartimento di Aut. ed Informatica, Politecnico di Torino, C.so Duca degli Abruzzi, 24, 10129 Torino, Italy (phone: 39-011-5647068; fax: 39-011-5647099; e-mail: fabrizio.lamberti@ polito.it).

B. Montrucchio is with the Dipartimento di Aut. ed Informatica, Politecnico di Torino, C.so Duca degli Abruzzi, 24, 10129 Torino, Italy (e-mail: bartolomeo.montrucchio@ polito.it).

point has exactly one track associated with it and each track is associated to exactly one world point. An additional condition is introduced to take into account the presence of points associated to image noise. For this, it must also be that each track is either composed of points corresponding to real-world objects or it is composed of points occurring from noise only. It can be demonstrated that the problem can be reformulated using the graph theory. In particular, a two-frame correspondence problem can be viewed as finding a *maximum weight path cover* of a bipartite directed weighted graph $D=(V,E)$ where the partite sets V_1 and V_2 correspond to the set of points X_1 and X_2 detected in frames F_{t_1} and F_{t_2}. Nevertheless, tracking require to solve a multiframe correspondence problem. It can be demonstrated that under suitable conditions matching techniques like the Hungarian matching algorithm can be used to find a maximum path cover of the graph characterized by a complete association. The weight function must be defined so that it guarantees that the total weight is maximized only if all the edges of T are in the path cover and penalizes the choice of a shorter track when a longer valid track is present.

However, the general problem of multiframe point correspondence is NP-hard for three or more frames. In [4], a greedy algorithm exploiting a look-ahead techniques based on a sliding window and capable of solving the above problem with a polynomial complexity is presented. The algorithm above also provides support to handle occlusions, missed detections and false positives for generic moving objects. The algorithm assumes point correspondences to be already established for $k-1$ frames F_{t_1}, F_{t_2}, . . . , $F_{t_{k-1}}$ on the basis of information available up to time t_{k-1}. Such information are updated each time a new frame (succeeding the actual one) is added to the sequence. Modifications to correspondences at time t_{k-1} can be simple extensions to existing tracks toward points belonging to frame F_{t_k} (that is the most common situation) or they can be corrections that allow to identify a better match with the newly added points (this is the basis for the management of occlusions). The problem of extending the actual set of correspondences with points in frame F_{t_k} can be illustrated as in Fig. 1. Each track at time t_{k-1} is modeled as the set of vertices $v\in V_{t_j}$ and defined as $T=\left\langle v\left(x_{t_1}^{i_1}\right),\ldots,v\left(x_{t_j}^{i_j}\right),\ldots,v\left(x_{t_p}^{i_p}\right)\right\rangle$, $1\le p\le k-1$. Between two consecutive vertices an edge $e\in E$ is placed. Edges representing correspondences before the addition of a new frame are called *old edges*. The extension of each track in frame $F_{t_{k-1}}$ to the points in frame F_{t_k} is modeled by the introduction of new edges linking all the points in the new frame. These edges are defined *extension edges*. Finally, the possibility of correcting a previous error due for example to occlusion or disappearance is represented by the introduction of *correction edges*. The resulting graph is defined *extension graph*.



(a) (b)

(c) (d)

Fig. 1. a) Initial correspondences. b) Extension graph with old edges (black), extension edges (red) and correction edges (blue). Not all the edges are shown. c) Path covering. d) False hypotheses removal.

In Fig. 1.a a possible configuration after the third iteration is shown. After the introduction of the fourth frame, the extension graph has been defined (see Fig. 1.b). Matching algorithm selected point *z* for succeeding point *y* (see Fig. 1.c). However, this association relied on the existence of a correspondence between *x* and *y*. Nevertheless, after the matching this correspondence has been replaced by a more favorable correspondence between *x* and *w* (a correction edge reported in blue has been added in Fig.1.c). The reason for this could be a possible occlusion, disappearance or aggregation. The association of *y* and *z* is not consistent (this is defined a false hypothesis). False hypotheses are removed and the path covering algorithm is executed again. As soon as the graph has been updated, a greedy approach is used to solve the maximum path cover problem (for details on the greedy iterative procedure see [4]). The solution is a subset of the edges constituting the extension graph composed by old edges, extension edges and correction edges.

IV. DESIGN OF A WEIGHT FUNCTION FOR TRACKING ECS

In [4], the authors present several applications of their greedy algorithm to tackle the problem of tracking moving objects. However, the considered applications rely on the knowledge of the physical rules governing the motion of tracked points. In the particular scenario being considered the knowledge of the behavior of migrating endothelial cells is actually missing. Therefore, the weight function has to be completely redesigned in order to take into account the specific information available for the motion environment being considered. In the next sub-sections, criteria used to compute an ad-hoc weight function are illustrated.

A. Spatial distance

Since in the particular case under consideration no assumption on the motion of cells can be made, the hypothesis of a *limited* displacement of the involved objects has been considered. In this particular context, the term limited is used to indicate a displacement of the centroid of the cell on the image plane comparable to the dimension of the cell itself. Given the assumption above, the displacement value can be used as key parameter for the computation of the weight of each association hypothesis. In this case a Nearest Neighbor Matching (NRM) approach is used and the edge weight is defined as the Euclidean distance separating a vertex in the graph (that corresponds to a point in the physical) space from any other instance of points in the following frame for which a correspondence is plausible. The remaining of the job will be carried out by the matching function that will select the configuration that maximizes the gain (or equivalently, minimize the overall distance) for all the association hypotheses. Because of the assumption on the limitedness of the displacement in the single unit of time, it is correct to assume that the centroid of a cell moves with a uniform probability at a distance included into a circular area of radius $r_{\max}$ and origin located at the cell centroid. In Fig. 2 a sample configuration allowing to easily understand the matching mechanism conditioned by the distance criteria is illustrated. The instances of two sample cells (colored in black) at frame F_{t_1} are reported. The circular area of radius $r_{\max}$ painted in black and enclosing cell instances C_1^i and C_1^j limits the area of research for the possible associations in the next frames. In frame F_{t_2} each instance has changed its location within the expected region and the new spatial configuration is represented in red (C_2^i, C_2^j). Finally, in frame F_{t_3} the new location of cell instances is depicted in blue. Given such cell configuration, the directed edges of graph D are constructed. Let's define $w_{disp}\left(x_{a_p}^{t_i}, x_{a_q}^{t_j}\right)$ the weight associated to the edge linking vertex $v\left(x_{a_p}^{t_i}\right)$ with vertex $v\left(x_{a_q}^{t_j}\right)$. The weight function can be expressed as:

$$w_{disp}\left(x_{a_p}^{t_i}, x_{a_q}^{t_j}\right) = \begin{cases} \left| x_{a_p}^{t_i} - x_{a_q}^{t_j} \right| / r_{\max} & \left| x_{a_p}^{t_i} - x_{a_q}^{t_j} \right| \le r_{\max} \\ \infty & \left| x_{a_p}^{t_i} - x_{a_q}^{t_j} \right| > r_{\max} \end{cases} \quad (1)$$

From a combinatory point of view the optimization problem that has to be solved can be written as:

$$e_{\min} = \min_{e \in E} \sum_{C_x \in F_{t_i}} \sum_{C_y \in F_{t_j}} w_{disp}\left(C_x^i, C_y^j\right) e\left(C_x^i, C_y^j\right) \quad \forall\, i < j \le n \quad (2)$$

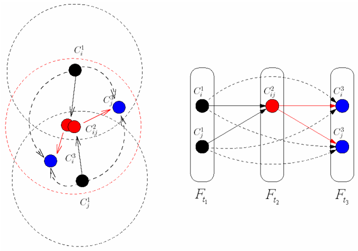


Fig. 2. Spatial distance for a configuration showing aggregating features.

After each time interval new information have to be combined with existing ones in order to construct an environment containing the required level of complexity capable of supporting the NNM approximation on a time interval sufficiently long (sliding window). Thus, it can be observed that in Fig. 2 correspondences between cells in frame F_{t_1} and cells in frame F_{t_3} are established. Referring to the original algorithm in [4], these are defined correction edges and their function is to enable the correct handling of possible occlusion conditions. The application of the covering procedure produces two tracks, one touching all the frames, the other touching just a cell on the first and a cell on the last frame. The occluded cell is not considered in frame F_{t_2} but information on the preceding and following frames are used to estimate its most probable destination.

B. Temporal distance

In the considered scenario it is not possible to assume that for a particular cell the distance from the starting point increases as time passes. In other words, it is possible for the instance x' of a cell in frame $F_{t_{k+1}}$ to be more distant from the corresponding instance x in frame F_{t_k} than from the same instance x'' in frame $F_{t_{k+2}}$. As a consequence, the similarity between the characteristics of features interested by a possible association between two vertices does not justify the unconditional jump of frames along the temporal sequence (especially when features are on distant frames). In fact, in the considered scenario the hypothesis of approximation of NNM would fail without considering the temporal distance between different instances of the same cell and, as a result, it would assign as successor of instance x in frame F_{t_k} instance x'' that is closer from a geometrical point of view (even if located two frames ahead). Errors like the one presented above can be effectively addressed by inserting in the weight function the information related to the chronological distance between instances of cells laying on different frames. Therefore, a distance Δ that can be expressed as the number of frames (that is, levels in the graph) covered by an association edge is defined. Once Δ is introduced, Δ_{off}^2 can be defined as the offset contribution produced by an association in function of the distance between the vertices joined by the association itself. Greater the distance between vertices which participate in the association hypothesis, greater the contribution that prevents the association to take place. In this way the associations that occur between a limited number of frames are implicitly favored, while associations between two remote frames are considered only when no other hypothesis is plausible. The result is that associations between consecutive frames are favored since actual tracks can be completed without the necessity to create new shorter tracks often caused by noise introduced in the segmentation phase.

C. Likelihood

Likelihood is an additional characteristic that can be used to establish a further distinction between features. From an algorithmic point of view, likelihood can be effectively measured through mathematical solutions descending from statistics. One of the most used methods is Principal Components Analysis (PCA). PCA is a technique that can be used to simplify a dataset. It is a linear transformation that chooses a new coordinate system for the dataset such that the greatest variance by any projection of the dataset comes to lie on the first axis (called the first principal component), the second greatest variance on the second axis, and so on. PCA can be used for reducing dimensionality in a dataset while retaining those characteristics that contribute most to its variance, by keeping lower-order principal components and ignoring higher-order ones. The idea is that such low-order components often contain the "most important" aspects of the data. Dataset is represented by a matrix X containing experimental data. To construct X, a two-dimension array containing pixels values corresponding to the bounding box of a particular feature is computed. Each array represents one experimental data. Pixel values are reorganized into one-dimension arrays and are stored as rows of matrix X. The direct application of PCA on such matrix is extremely expensive from the computational point of view, since each row could count up to thousand of pixels because of possible creation of features representing clusters of cells counting hundreds of pixels. The Karhunen-Loève transform represents a simplification of PCA and allows to find the singular value decomposition of a data matrix X with a lower complexity. By finding the eigenvalues and eigenvectors of the covariance matrix $C = X^T X$, the first k eigenvectors with the largest eigenvalues correspond to the dimensions that have the strongest correlation in the dataset. Once X has been computed for all the cells of each frame, a measure of the likelihood between two cells is obtained by projecting the associated image vectors along the k principal components and computing the distance between the points obtained. By defining u as the matrix containing the first k principal components, given the images of two cells properly registered and linearized on two vectors c_1 and c_2, the distance between their projection can be written as

$$d(c_1, c_2) = \left|(c_1 - c_2) \cdot u\right| \tag{3}$$

It is worth remarking that the lack of an upper bound for likelihood makes it necessary to compress its domain in the range $[0,1]$ through a polynomial relationship as follows:

$$w_{like}(d) = 1 - 1/\left(1 + d^2\right) \tag{4}$$

Values of $w_{like}(d)$ close to zero indicate a sharp likelihood, while higher values indicate an increasing dissimilarity.

D. *Gain function*

Considering the contribution of the weight criteria reported in the previous subsections, the cost function to be minimized by the cover procedure can be written as

$$w\left(x_{a_p}^{t_i}, x_{a_q}^{t_j}\right) = -\frac{1}{\alpha \cdot w_{disp}\left(x_{a_p}^{t_i}, x_{a_q}^{t_j}\right) + \beta \cdot w_{like}\left(x_{a_p}^{t_i}, x_{a_q}^{t_j}\right) + \Delta_{off}^2} \tag{5}$$

where $\alpha, \beta \in [0,1]$.

V. RESULTS

Results obtained for a sample sequence counting 300 frames using a window of 25 frames and a balancing of α and β favoring the NNM contribution to manage a significant segmentation noise compensated by a reduced migration speed are illustrated in Fig. 3. Histogram in Fig. 4. shows the distribution of errors due to miss detection. For each percentage of error, two values are reported, namely the number of tracks characterized by such error, and the number of tracks that can be tracked up to their final destination (that is, before disappearance or aggregation into final cluster) despite the misses.

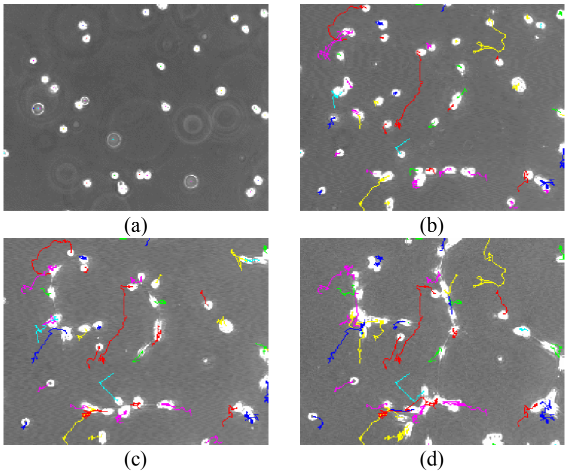


Fig. 3. EC sample microscopy sequence. Frame a) 0, b) 60, c) 120, d) 180.

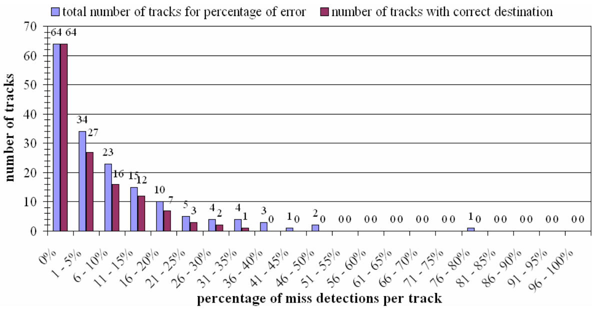


Fig. 4. Distribution of tracking errors for sample video sequence in Fig. 3.

ACKNOWLEDGMENT

The authors wish to thank Marco Bongiovanni for his precious help in the development of the analysis framework.

REFERENCES

[1] C. Zimmer, E. Labruyère, V. Meas-Yedid, N. Guillén, J.C. Olivo-Marin, "Segmentation and tracking of migrating cells in videomicroscopy with parametric active contours: a tool for cell-based drug testing," *IEEE Trans. Med. Imag.*, vol. 21, 2002, pp. 1212-1221.

[2] N. Ray and Scott T. Acton, "Active Contours For Cell Tracking," *Fifth IEEE Symp. on Image Anal. and Interpr.*, 2002, pp. 274-278.

[3] T. Kirubarajan, Y. Bar-Shalom and K. R. Pattipati, "Multiassignment for tracking a large number of overlapping objects," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 37, 2001, pp. 2-21.

[4] K. Shafique and M. Shah, "A noniterative greedy algorithm for multiframe point correspondence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, 2005, pp. 51-65.