A Model-Based Approach for Automated In Vitro Cell Tracking and Chemotaxis Analyses

Olivier Debeir,1 Isabelle Camby,2,3 Robert Kiss,2,3 Philippe Van Ham,1 and Christine Decaestecker2,3*

1Department of Logical and Numerical Systems, Faculty of Applied Sciences, Université Libre de Bruxelles, Brussels, Belgium
2Laboratory of Toxicology, Institute of Pharmacy, Université Libre de Bruxelles, Brussels, Belgium
3Fonds National de la Recherche Scientifique (FNRS), Brussels, Belgium

Received 3 September 2003; Revision Received 10 February 2004; Accepted 3 March 2004

Background: Chemotaxis may be studied in two main ways: 1) counting cells passing through an insert (e.g., using Boyden chambers), and 2) directly observing cell cultures (e.g., using Dunn chambers), both in response to stationary concentration gradients. This article promotes the use of Dunn chambers and in vitro cell-tracking, achieved by video microscopy coupled with automatic image analysis software, in order to extract quantitative and qualitative measurements characterizing the response of cells to a diffusible chemical agent.

Methods: Previously, we set up a videomicroscopy system coupled with image analysis software that was able to compute cell trajectories from in vitro cell cultures. In the present study, we are introducing a new software increasing the application field of this system to chemotaxis studies. This software is based on an adapted version of the active contour methodology, enabling each cell to be efficiently tracked for hours and resulting in detailed descriptions of individual cell trajectories. The major advantages of this method come from an improved robustness with respect to variability in cell morphologies between different cell lines and dynamical changes in cell shape during cell migration. Moreover, the software includes a very small number of parameters which do not require overly sensitive tuning. Finally, the running time of the software is very short, allowing improved possibilities in acquisition frequency and, consequently, improved descriptions of complex cell trajectories, i.e. trajectories including cell division and cell crossing.

Results: We validated this software on several artificial and real cell culture experiments in Dunn chambers also including comparisons with manual (human-controlled) analyses.

Conclusions: We developed new software and data analysis tools for automated cell tracking which enable cell chemotaxis to be efficiently analyzed. © 2004 Wiley-Liss, Inc.

Key terms: cell motility; chemotaxis; cell tracking; image processing; active contours

Some methods employed to monitor cell cultures and determine cell parameters are based on the use of microscopes equipped with a numerical image acquisition system for taking a set of pictures of the cell cultures, a digitizer board, and a computer that processes images to evaluate cell tracking. This kind of method is referred to as computer-assisted video-microscopy. We previously described such a method, which, coupled with an image analysis software, is able to monitor the motility of a large number of living cells present in a microscopic field during a given time period (1-3). This software, founded on watershed-based segmentation algorithms, enables the cell trajectories to be established and characterized by quantitative descriptors (3). The major problems encountered with this method concern its lack of robustness with respect to variability in cell morphology in different cell types and dynamical changes in cell shapes during cell migration. In fact, the tracking software includes a number of parameters that require relatively sensitive tuning in order to take into account variations in cell morphology and acquisition conditions. Finally, the running time of the software does not allow very high acquisition frequency and, consequently, makes it difficult to track complex cell trajectories, such as those including cell division and cell crossing. The aim of this study is to put forward a new image analysis method based on active contours (4–5) to cope with the problems described above. Furthermore, this new method is able to generate new cell trajectory

Contract grant sponsors: Fonds de la Recherche Scientifique Médicale (FRSM), Belgium; Fondation Yvonne Boël, Belgium.
*Correspondence to: C. Decaestecker, Ph.D., Laboratory of Toxicology, Institute of Pharmacy, Université Libre de Bruxelles CP 205/1, Bld du Triomphe, 1050 Brussels, Belgium.
E-mail: cdecaes@ulb.ac.be
Published online 9 June 2004 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/cyto.a.20040
features, providing useful information for chemotaxis study.

In order to obtain different validations of our approach, three different aspects were considered in our study: 1) tracking on synthetic image sequences where all features are controlled; 2) making a comparison between automatic tracking and human-controlled tracking; and, finally, 3) carrying out experiments with real cell cultures in Dunn chambers in the presence of a potentially chemoattractive agent.

MATERIALS AND METHODS

Cells, Culture Media, and Compounds

The human umbilical veins endothelial cell (HUVEC) lines were established to grow in vitro as monolayers. These lines came from different human umbilical cord veins obtained from the Department of Gynecology at Erasmus University Hospital. The monolayers were obtained by means of a method adapted from the procedure described by Gimbrone et al. (6), as detailed elsewhere (7–8). The HUVEC cells were harvested by centrifugation, plated on gelatin-coated dishes and amplified at 37°C in endothelial cell growth medium 2 (ECGM2; ImTec Diagnostics, Antwerp, Belgium). In all our experiments, we made use of the first three passages of each HUVEC pri-moculture that was established in our laboratory.

Gastrin (pGlu-Gly-Pro-Trp-Leu-Glu-Glu-Glu-Glu-Ala-Tyr-Gly-Trp-Met-Asp-Phe-NH₂) purchased from Sigma (Bornem, Belgium) was tested as a potential chemoattractant for these cells.

Dunn Chemotaxis Chamber

Chemotaxis was assessed by direct observation and recording of cell behavior in stable concentration gradients of gastrin using the Dunn chemotaxis chamber (Weber Scientific International Ltd., Teddington, UK). As detailed elsewhere (9–10), this is a modified Helber counting chamber slide with two concentric (inner and outer) wells separated by an annular platform which lies 20 μm below the upper face of the chamber, thus creating a diffusion gap when the chamber is covered with a coverslip. When the inner well is filled with control medium and the outer with medium containing potential chemoattractant (such as gastrin, in the present study), a radially-directed linear diffusion gradient becomes established in the diffusion gap within 30 min and has a half-life of 10–30 h (9–10). Cells were cultured on the coverslip, which was then inverted onto the chamber slide. Cells that rested over the annular platform were observed under phase-contrast microscope and recorded automatically by the acquisition system (described below).

Algorithm

This section presents an algorithm that achieves automatic, accurate, and robust tracking of cultured cells.

Images were acquired as previously described (1–3). Briefly, the images were obtained under a phase-contrast microscope (Model IX50; Olympus, Tokyo, Japan; magnification ratio 10:1), with a charged-coupled device (CCD) video camera (Model KP-M1E/K-S10; Hitachi Densihi, Leeds, UK), coupled with an acquisition board (32-bits Matrix Vision, Oppenweiler, Germany PC-GRAB-GI frame grabber). The acquisition typically started about 30 min after the chamber assembling, and was carried out at constant time lapses of 4 min during a number of hours (typically 24 h, i.e., 360 images), in order to fall within the conditions of a linear diffusion gradient in the Dunn chamber (see above). The image dimensions are 700 × 500 pixels discretized into 256 gray levels (8 bits per pixel). The uncompressed images were stored on a computer hard disk.

As detailed below, our algorithm includes three phases. Briefly, a first preprocessing phase was applied to each acquired image in order to enhance image contrast, normalize gray level histograms, and avoid illumination problems that could occur during the acquisition period. A rapid manual second phase consists of selecting the first image of the sequence, in which the user had to identify the cells to be tracked and the border position of the outer well of the Dunn chamber to determine the gradient direction. Finally, the third and essential phase is based on an active region algorithm and analyzes each consecutive image in order to track all the marked cells until the last image of the sequence.

The last part of this section describes the trajectory features extracted from the tracking results with a special focus on chemotaxis characterization.

Preprocessing. We applied image preprocessing steps to avoid illumination problems (such as vignetting and illumination changes) during the experiment and to enhance image contrast. These steps, which were applied to the complete image sequence, essentially included image background detection, background masking, and local gray level histogram equalization (as illustrated in Fig. 1). The combination of these image-preprocessing steps ensures a good contrast between the inner part of the cells (darker areas) and the surrounding region (brighter areas) (Fig. 1D). This also gives good results on all the treated sequences (i.e., ensuring robustness, data not shown). It is interesting to point out that an optimized code enabled an image to be preprocessed in less than 3 s on a standard 1.4 MHz Pentium IV computer.

Cell marking. A human operator is asked to manually select the cells to be tracked on the first frame of the sequence by marking them with the computer mouse. This task can be achieved quickly (i.e., without a high level of accuracy) as a part of the tracking algorithm described below, which was also used to automatically adjust these marks to the initial cell centroid positions. The operator also had to select three points belonging to the well border of the Dunn chamber. It is considered that these three points define the tangent of the well border. This made it possible to determine the radial gradient direction in the experiment, which is perpendicular to the tangent as illustrated by the arrow in Figure 1A. This direction, called the gradient direction (GD), was used as
reference angle during the analysis of the tracking feature values (see below).

Tracking.

General algorithm. In previous works (1–3), we proposed a segmentation-based method using marked watershed transform. This approach segments each frame and tries to reconnect each segmented object at time $t$ with its ancestor at time $t - 1$. Although useful results have been reported (11–13), this kind of method faced two major drawbacks: relatively long computing time and the difficulty in handling dividing objects (e.g., mitoses) and overlapping objects, both frequent in 2D cell cultures such as in Dunn chambers. The new method proposed here uses a model-based approach. Model-based methods, such as active contour (4) or region (5), optimize a model shape by adjusting its parameters to fit the model to a targeted object (i.e., each cell, in this case). In this application, the results in image at time $t - 1$ are used to initialize the iterative processing in the image at time $t$.

Figure 2A shows an example of a preprocessed image, overlaid with the pie-chart shaped region that we used as a cell model. The gray level profile across a single cell (illustrated in Fig. 2B) characterizes cells as a black objects (low gray-level values) surrounded by a lighter halo. To identify a cell, we thus designed a model which fits a black object (i.e., each cell, in this case). This model is fitted at time $t$ to the central location. The radius values are adapted during the iterative process (see below). The two nested polygons obtained defined two gravity centers (or centroids) labeled $G_{\text{int}}$ and $G_{\text{ext}}$. As a matter of fact, these centroids do not coincide with the initial cell center ($x_0$, $y_0$), and are used to adjust by interpolation this initial center to a new position $c_1 = (x_1, y_1)$. The procedure is iteratively processed until a convergence criterion is fulfilled. In each image $t$, the final cell center position obtained is noted as $c^*(t)$. This means that $c^*(t) = c_n(t) = (x_n(t), y_n(t))$ (obtained after $n$ iterations). In practice, few iterations (less than five) are needed to converge to the central location.

Considering that our model is fitted at time $t$ to a cell, i.e., $c^*(t)$ is determined, the shape and the position of this cell model at time $t + 1$ have to be computed. First, the new initial cell position at time $t + 1$, i.e., $c_n(t + 1) = (x_n(t + 1), y_n(t + 1))$, is predicted by extrapolation on the basis of the displacement observed between $c^*(t - 1)$ and $c^*(t)$ (i.e., the final cell centers in image $t - 1$ and $t$, respectively). This extrapolation assumes that the cell keeps its speed constant (in direction and value) in the next time interval. More formally:

$$c_n(t + 1) = c^*(t) + (c^*(t) - c^*(t - 1))$$

This predicted value enables the fitting process to be initiated with a more adequate value and to speed up the convergence.
New radii $r^{\text{ext}}$ and $r^{\text{int}}$ are also determined by extrapolation in order to be able to take into account progressive scale modifications in our model, i.e., progressive shape deformation of the cell being tracked.

The automatic step-invalidation procedure. In order to eliminate the trajectory parts that could contain tracking errors, an automatic procedure has been implemented. This algorithm can tackle two kinds of problems.

The first problem is that of cells leaving the observed microscope field. To overcome this, a minimum distance (typically 20 pixels) from the image border has been defined. The mark representing a cell position (i.e., a center $c^*$) is frozen (i.e., its consecutive positions are invalidated) when it penetrates this border area.

The second problem appears when two marks (that initially followed two different cells) become closer than a certain minimum distance (typically 10 pixels), which is explained by the fact that both marks are following the same cell because one of them has left its previously tracked cell. In other words, two centers $c^*$ are now attracted by the same cell, and thus the two corresponding active contours are trying to match the same cell, even if they remain slightly different during a certain time (e.g., this may occur when two cells cross each other, as is frequently encountered in the Dunn chamber experiments on HUVEC cells reported below). In this case, one of the conflicting marks is frozen and all its consecutive positions are invalidated in order to avoid a double-tracking of the same cell.

Computed attributes. The tracking procedure is applied for each initially marked cell for every image $t$ until the end of the sequence ($t = 1, 2, \ldots, N$). The successive locations of the model center $c^*(t)$ of a cell along the successive images ($t = 1, 2, \ldots, N$) constitute the cell trajectory.

As illustrated in Figure 3A, each trajectory can be thus split into small displacements, which are labeled step vectors, given by the following equation.

$$v(t) = (c^*(t) - c^*(t - 1))$$

Each of these vectors is characterized by its length ($|v(t)|$) also labeled “instantaneous speed,” and its angular direction, $\theta v(t)$, labeled “instantaneous direction.”

The trajectories may show residual noise, which is mostly due to cell shape deformations occurring without any real cell displacement and generating little change in the current $c^*(t)$ cell position. These effects may be
smoothed by means of a facultative parameterized low-pass filtering step (standard window average filtering).

Global nonoriented features. As defined in our previous work (2,3), for each cell trajectory, we computed the average speed (AS), which is the mean distance covered by a cell per time unit and the maximum relative distance to origin (MRDO) parameter. The MRDO parameter refers to the greatest linear distance between the original position of the cell (at time \( t = 0 \)) and the farthest position reached by the cell in its trajectory. MRDO is in fact this maximal distance, normalized by the observation time for the cell analyzed (to be able to compare cell trajectories corresponding to different observation times). This gives synthetic information on directional cell motility, e.g., while a low MRDO value indicates that the cell does not move far from its initial position, a high MRDO results from a high displacement amplitude.

For each cell trajectory, we also computed a third parameter, the total displacement speed (TDS), which is defined as the linear distance (as the crow flies) between the original to the final position of the cell, normalized by the observation time for the analyzed cell. It should be noted that TDS = MRDO in the case where the final cell position is the farthest position that the cell reached in its trajectory, and TDS < MRDO in the other cases.

As is described in Figure 3A and its legend, the three parameters above can be defined on the basis of the step vectors \( \mathbf{v}(t) \) and two other vectors characterizing a cell trajectory (labeled \( T \) and \( M \) in Fig. 3A). Figure 3B-D schematically illustrate different types of cell trajectories (assumed to be established during the same observation time) presenting the following different features. B: Two trajectories with equal AS and equal TDS, but presenting different MRDO. C: Two trajectories with equal AS and MRDO, but different TDS. D: Two trajectories with equal TDS and MRDO, but different AS.

FIG. 3. A: Illustration of the vectors which constitute the bases of the different features used to characterize a cell trajectory. Each reconstituted trajectory comprises a set of step vectors \( \mathbf{v}(t) \) \( (t = 1 \) to \( N, \) with \( N = 13 \) in this case). \( T \) (Total) is the sum of the \( \mathbf{v}(t) \); \( M \) (Maximum) is the vector linking the origin to the farthest point of the trajectory. To be able to compare different trajectories observed during different time intervals (i.e., with different values for \( N \)), we defined the following quantities (nonoriented features): AS (average speed) = \( \sum |\mathbf{v}(t)|/N \); MRDO = \( |\mathbf{M}|/N \); and TDS = \( |\mathbf{T}|/N \). For the oriented (normalized vector) features, we considered the \( \alpha \text{MRDO} \) vector, which is defined as \( \mathbf{M}/N \) (i.e., \( |\alpha \text{MRDO}| = \text{MRDO} \)), and similarly \( \alpha \text{TDS} \), which is \( \mathbf{T}/N \). B–D: Illustrations of different types of cell trajectories (assumed to be established during the same observation time) presenting the following different features. B: Two trajectories with equal AS and equal TDS, but presenting different MRDO. C: Two trajectories with equal AS and MRDO, but different TDS. D: Two trajectories with equal TDS and MRDO, but different AS.
and the graphic presentation of the oriented features described below.

We extended the concept of MRDO to an oriented vector version, \( \alpha \text{MRDO} \), a vector on the direction between the origin of the trajectory to the farthest-reached point (i.e., the \( \text{M} \) vector illustrated in Fig. 3A–D), whose length is MRDO (i.e., \( |\alpha \text{MRDO}| = \text{MRDO} = |\text{M}|/N \), where \( N \) is the number of step vectors constituting the cell trajectory). Figure 4A and C illustrate two different sets of \( \alpha \text{MRDO} \) vectors in which all the original cell positions are set to the axis origin and the vector directions are reported with respect to GD arbitrarily set to 0°. For each of these (relative) distributions, we computed \( \alpha \text{MRDO}_{\text{mean}} \), which is the mean vector of the \( \alpha \text{MRDO} \) vectors (represented by the large gray arrows in Fig. 4A and C). This indicates the general (mean) direction of the farthest parts of the trajectories, if any. The length of this mean vector \( (|\alpha \text{MRDO}_{\text{mean}}|) \) is indicative of the possible presence of a privileged direction followed by the \( \alpha \text{MRDO} \) vectors. Indeed, \( |\alpha \text{MRDO}_{\text{mean}}| \) is near zero if the vector distribution is relatively uniform (see Fig. 4A) and becomes larger as an increasing number of vectors take a similar direction (see Fig. 4C), or when a group of vectors having a similar direction have particularly large lengths. Similarly, each cell trajectory can also be characterized by its \( \alpha \text{TDS} \) vector, which is the sum of the small displacement vectors constituting the trajectory (i.e., the \( \text{T} \) vector illustrated in Fig. 3), normalized by the number of steps (see Fig 3A and its legend). However, for the same reasons than those described above with MRDO, the \( \alpha \text{MRDO} \) vector emerged as a more interesting oriented feature.

**Instantaneous oriented features.** The parameters and vectors defined in the two previous subsections characterize the individual cell trajectories taken in their entirety. However, it is also interesting to analyze local or instantaneous features characterizing each cell displacement step. This is why we analyzed the distribution of the step vectors, \( \text{v}(t) \), with respect to their angular directions, \( \alpha \text{v}(t) \) (relative to the GD direction). To this end, we arbitrarily defined eight angular bins, each \( \pi/4 \) (45°) wide, centered on the following directions \( 0, \pi/4, \pi/2, 3\pi/4, \pm \pi, -\pi/2, -\pi/4 \), relative to the GD direction of 0. Two histograms, which independently pooled all the step vectors (without taking into account the cell trajectories that they constituted), were computed. The first one, labeled \( H_I(D) \) and plotted in gray pie slices in Fig. 4B and D, indicates the distance proportions covered in each direction \( D \) (of the eight defined above), when all the step vectors are considered. This means that each gray slice in Fig. 4B and D represents the sum of the lengths of all the \( \text{v}(t) \) vectors (i.e., \( ||\text{v}(t)|| \) having the corresponding direction, normalized by the sum of all the vector lengths.

More formally, let \( S_t(D) \) be the sum of the lengths of all the step displacements covered by any cell in the direction \( D \) (of the eight defined above):

\[
S_t(D) = \sum_{\alpha \text{v}(t) \in D} ||\text{v}(t)||
\]

The \( H_f(D) \) histogram is thus computed as follows.
Thus, \(H_{dl}(D) = S(D)/\sum_D S(D)\)

A high value in this histogram is thus indicative of a direction to which the trajectories of many cells are mostly parallel and/or in which the recorded instantaneous cell speeds (i.e., \(|v(t)|\)) are particularly large (see Fig. 4D).

We also computed a differential version of the \(H_l\) histogram that we labeled \(H_{dl}\) and plotted in white pie slices in Fig. 4B and D. In this case, for each pair of opposite directions \((D, D - \pi)\), we computed the differences between sums \(S(D)\) and \(S(D - \pi)\) defined above, and retained only the positive values after normalization, as detailed below:

\[
\text{if } S(D) > S(D - \pi) \\
H_{dl}(D) = \frac{S(D) - S(D - \pi)}{S(D) + S(D - \pi)} \\
\text{else } H_{dl}(D) = 0
\]

Thus, \(H_{dl}(D)\) belongs to \([0, 1]\) and

- \(H_{dl}(D) = 1\) if of all the cell displacements oriented in direction \(D\) or \(D - \pi\), all are in direction \(D\);
- \(H_{dl}(D) = 0.5 \ (0.33)\) if the sum of the lengths of the displacements in direction \(D\) is three (two) times larger than that of the displacements in direction \(D - \pi\);
- \(H_{dl}(D) = 0\) if the sum of the lengths of the displacements in direction \(D\) equals that of the displacements in direction \(D - \pi\).

This \(H_{dl}\) histogram enhances the unbalance in the displacement distribution across the different directions considered and can thus be particularly useful to visualize chemoattractive effects (as illustrated in Fig. 4).

As illustrated in the Results section, the \(H_l\) and \(H_{dl}\) histograms can also be used to describe the distribution of the \(\text{\texttt{oMRDO}}\) vectors.

**Data analysis.**

*General approach.* To study the impact of a chemoattractive agent on cell displacements, we analyzed the influence of such an agent on the distributions of the different types of vector used to characterize the cell trajectory directions (i.e., step vectors \(v(t)\), and global vectors \(\text{\texttt{oMRDO}}\)). The Rayleigh test and a modified form of it were used to this aim. The Rayleigh test was originally designed to test the null hypothesis that a set of \(N\) independent angles \(\theta_i\) has been drawn from a uniform circular distribution. The rejection of the null hypothesis by the Rayleigh test indicates a significant unimodal clustering of cell directions. The Rayleigh test is based on the \(R\) statistic with:

\[
R^2 = \left[\frac{1}{N} \sum_{n=1}^{N} \cos \theta_n\right]^2 + \left[\frac{1}{N} \sum_{n=1}^{N} \sin \theta_n\right]^2
\]

The observed value of this statistic corresponds to the length of the mean vector computed on the set of the \(N\) unit vectors \((1, \theta_i)\). \(R\) is near or far from zero according to the uniformity or clustering of the angles.

Under uniformity, the statistic \(2NR^2\) follows a \(\chi^2\) distribution. As detailed elsewhere (14), the following variant of the \(2NR^2\) statistic is similarly distributed and benefited of an improved order of error:

\[
S = 2NR^2 - R^2 + \frac{N_1^4}{2},
\]

and a very good approximation to the significance level can be computed as follows:

\[
p = \exp\left(\sqrt{1 + 4N + 4(N^2 - N_1)} - 1 - 2N\right)\]

where \(s\) is the observed value of \(S\) (i.e., computed on the data sample).

A modified form of the Rayleigh test proposed by Moore (15) is able to take the vector lengths \((r_n)\) into account (in the form of ranks) when analyzing the distribution of a set of \(N\) independent vectors, \((r_n, \theta_n)\). The principles are as follows. The vectors are ranked according to their lengths. This means that rank \(r_k = 1\) was attributed to the shortest vector and rank \(r_k = N\) to the longest. The modified Rayleigh statistics, labeled \(R^*\) accordingly to Moore (15), are computed as follows:

\[
(R^*)^2 = \frac{\left[\sum_{n=1}^{N} n \cos \theta_n\right]^2 + \left[\sum_{n=1}^{N} n \sin \theta_n\right]^2}{N^3}
\]

A significance level has to be computed on the data with respect to the null hypothesis that the vector distribution has been drawn from a uniform circular distribution. In the case of \(N \geq 30\) (always the case in our experiments) a very good approximate probability value is given by the following expression:

\[
p = \exp\left(-\frac{6r^2N^2}{(N+1)(2N+1)}\right)
\]

where \(r\) is the observed value of \(R^*\) on the data sample (15).

A simple extension of this test also enables two vector distributions to be compared by computing vector differences between pairs of vectors randomly picked from both distributions. No significant difference exists between these two vector distributions if the difference vectors follow a uniformly circular distribution (15).

The nonparametric Uniform-Scores test also enabled us to test the equality of two angular distributions (without
taking the vector lengths into account). This is a nonparametric variant of the standard Rayleigh test that does not require random selection of pairs of vectors (14).

It should be noted that the sample independence assumed in the different tests mentioned above has the following consequences. First, with the analysis of the \( \alpha \text{MRDO} \) distribution, the cell trajectories were considered as independent. The possible cell–cell influences were thus not taken into account. Second, with the analysis of the \( v \) distribution, the different step vectors constituting a cell trajectory were also considered as independent.

Cautions with respect to the analysis of the step vector distributions. In the case of the \( v \) vectors, the very large number (\( N \)) of measurements available (at least 25,000 per experimental condition) causes a very slight deviation from uniformity (corresponding to a very low observed value for the \( R \) statistic, which is an \( N \)-independent measurement, see its expression above), which may be associated with a high level of significance by the standard Rayleigh test. This is essentially due to the role played by \( N \) in the determination of the significance level (see its expression above). For instance, given a value of \( R \) fixed at 0.05, the variations in the \( P \)-values with respect to \( N \) are as follows: \( P \approx 0.37, 0.082, 0.0067, \) and \( <10^{-6} \) when \( N = 200, 500, 1,000, \) and 10,000, respectively. This is why, to analyze the distributions of the \( v \) vectors, we preferred to fix a threshold on the observed value of \( R \) in order to decide whether the distribution presented a significant unimodal clustering of directions. This value was chosen as \( R \geq 0.11 \) because this corresponds to \( P < 0.01 \) when \( N = 200 \) (to be comparable with the results obtained on the \( \alpha \text{MRDO} \) vectors, see Results).

Furthermore, we did not use the modified version of the Rayleigh test and its statistic, \( R^* \), because of two essential reasons. First, the lengths of the step vectors (\( |v(I)| \)) presented a higher level of homogeneity than those of the \( \alpha \text{MRDO} \) vectors and thus weighting each angle contribution by the rank of the vector length was less appropriate than it was in the case of the \( \alpha \text{MRDO} \) analysis. Second, the values of the highest ranks increase with \( N \) and thus strongly weight (if \( N \) is very large) the corresponding angle contributions in the computation of the observed value of \( R^* \). This increase is not completely balanced by the normalization factor (\( N^3 \)) used in the expression of \( R^* \) detailed above, as observed in data simulation (data not shown). Consequently, when computed on data from the same distribution, the observed value of \( R^* \) increases with \( N \) (except in the case of perfect uniformity) and thus decreases the associated \( P \)-value.

RESULTS

Validation Through Simulated Cell Trajectories

In order to test the method on perfectly known data, a simple simulator has been realized. The simulated cells are roughly represented by a dark gray area of almost the same size as actual cells, surrounded by a white border (see Fig. 5A). A sequence of such synthetic images has been generated in order to simulate four cell trajectories in response to an assumed biochemical gradient (schematically represented by the arrow in Fig. 5A and set to 0 in Fig. 5B). The shape of the simulated cells has also been modified during the simulated trajectories to bring them closer to biological reality. As illustrated in Figure 5A, three simulated cells took paths in (approximately) the \(-\pi/4\) direction and a fourth one moved roughly in the \(-3\pi/4\) direction with respect to GD. Our algorithms were thus applied on these artificial data to compute the trajectories. The reconstituted trajectories fitted the original simulated ones, as illustrated by superposing them in Figure 5A. The gray pie chart in Figure 5B shows the \( H_l(D) \)
Comparison Between Automatic and Human-Controlled Cell Tracking

A complete experiment made on HUVEC cells during 24 h has been inspected by a human expert using a visualization tool (not detailed here). The expert’s role was to validate or invalidate each step displacement computed by our algorithm on the culture image sequence. To this aim, the visualization tool superposed the computed cell centroids on the original images. Image after image, the expert confirmed or refuted the centroid positions. Each invalidated displacement was discarded and did not contribute to feature computation. In this experiment, only 120 of 21,062 displacements were invalidated (i.e., 0.6%). This was due to the automatic step-validation procedure included in the tracking algorithm. The small number of human-invalidated steps showed the efficiency of this procedure and did not significantly change the distribution values of the different features characterizing the cell movements (data not shown).

Chemotaxis Experiment

In this experiment, we analyzed the chemotaxis potential of gastrin on HUVEC cells. For this, we tested (during 24 h) the effects on the cell trajectories of different concentrations of gastrin (CT = 0, 10⁻⁹, and 10⁻⁸ M) added in the outer wells of Dunn chambers. Figure 6A illustrates the computed trajectories obtained in an experiment in the presence of 10⁻⁹ M gastrin and superposed on the initial image of the sequence (with the arrow indicating the GD). A representative representation (more easily interpretable) of the same trajectories is shown in Figure 6B, where GD is arbitrarily set to 0° and all the original cell positions are set at the axis origin (0, 0). The latter shows that a majority of cells seemed attracted in the direction of the gastrin gradient (i.e., cell directions between -π/4 and π/4). Pooling data from three independent experiments carried out under each condition, Figure 7 shows the histograms (H₁ in black and H₉ in white) describing the distributions of the global oMRDO vectors (upper frames; Fig. 7A, C, and E) and the instantaneous v vectors (lower frames; Fig. 7B, D, and F), in the absence (Fig. 7A, B) or presence of 10⁻⁹ M (Fig. 7C, D) and 10⁻⁸ M (Fig. 7E, F) gastrin. These histograms (taking the vector lengths into account, see Materials and Methods) enabled us to suspect strongly that the HUVEC cells were chemoattracted by gastrin (see Fig. 7C-F). Indeed, the oMRDO histograms (especially the differential ones in white) and the mean vectors (arrows) indicated that the farthest parts of the cell trajectories were covered following directions near the gastrin gradient direction. However, this attraction was stronger in the 10⁻⁹ M condition (Fig. 7C) than in the 10⁻⁸ M condition (Fig. 7E), as indicated by the lengths of the mean vectors (arrows). Comparatively, the control exhibited a clearly more uniform distribution characterized by small contributions to the differential histogram and a mean vector with a very small size (Fig. 7A). The v histograms (Figs. 7B, D, and F) show the same tendencies, which were smoothed by the very large number of vectors analyzed (about 100 times larger than in the case of the oMRDO vectors). Statistical analyses confirmed these observations as follows. Either analyzing the angular directions only, or taking the vector lengths into account, the Rayleigh tests (in their standard or modified versions, respectively) indicated that the oMRDO vectors were uniformly distributed (P > 0.05) in control (Fig. 7A). In contrast, in the presence of gastrin (10⁻⁹ or 10⁻⁸ M) the distributions significantly differed from uniformity (i.e., R > 0.28 with P < 10⁻⁶, and R⁺ > 1.55 with P < 0.001); the oMRDO vectors thus presented a significant unimodal clustering of directions. This means that, in presence of a gastrin gradient, HUVEC cells presented a significant directionality of migration, as strongly suspected when analyzing the histograms. The two-group tests (Rayleigh and Uniform-Scores) confirmed significant differences in the distributions between control and gastrin conditions. It should be noted that gastrin was able to significantly increase the (global) motility level of the HUVEC cells, as evidenced on the global nonoriented feature, MRDO (data not shown).

With respect to the v vectors (Fig. 7B, D, and F), we used the rule R ≥ 0.11 to decide if the distribution presented a significant unimodal clustering of directions, as justified under Materials and Methods. With respect to this rule, the analysis of the data illustrated in Figure 7B, D, and F evidenced the presence of a significant unimodal clustering of the cell instantaneous displacements in the presence of 10⁻⁹ M gastrin (R = 0.12, which is associated with P = 0.003 if N = 200). This was not observed in the case of 10⁻⁸ M gastrin (R = 0.07) or in control, for which the histogram is characterized by the lowest R value (R = 0.02).

DISCUSSION

Chemotaxis, the migration of cells in the direction of a spatial chemical gradient, is important in the progression of different pathologies such as inflammation and cancer, and in the immune response to these problems. The ability to match chemoattractive gradient directions and the corresponding motility patterns is essential to the study of chemotaxis. Different assays exist to measure chemotaxis in vitro. One of the most used is the Boyden chamber. However, this assay does not enable the cell behaviors and their locomotion to be monitored (e.g., in
terms of directionality persistence) in response to a chemical gradient. Only the comparison between the initial and the final cell positions (after a given time fixed in advance) can be used to evaluate chemotaxis. Wilkinson (16) reported considerable uncertainty over the assignment of chemotaxis on the basis of this kind of assay. This is why a number of chemotaxis studies reported in the literature used visual assays based, for example, on time-lapse photograpy (10,17–19). This kind of experiment requires direct-viewing chambers enabling easy image acquisition and especially a stable and linear chemical gradient to be maintained during a sufficient time period. The Dunn chamber designed by Zicha et al. (9) represents a considerable advance by responding to these requirements and was thus used in the above-mentioned chemotaxis studies. A remaining limitation of this kind of study was the
absence of fully automatic methods for establishing cell trajectories and extracting parameters useful for assessing chemotactic cell response. Manual or interactive computer-assisted methods were used in the studies mentioned above. Consequently, the number of cells tracked and the following period were relatively restrained (e.g., 2 or 3 h of tracking, as used in Webb et al (10) and Allen et al. (17), respectively) because this kind of method is time-consuming for the user. While relatively short tracking periods may be convenient in certain applications, others (like cancer or endothelial cell migration) require longer tracking periods to be informative. In any case, tracking of a large number of cells is beneficial for obtaining statistically robust results. Of course, the latter point can be satisfied by multiplying the experiments (and thus also the user time in the case of interactive image analysis). In response to these limitations, the present study proposes a fully automatic system of cell image analysis, enabling numerous cell trajectories to be automatically and quickly established for a relatively long period of time (e.g., 24 or 48 h), where cell culture and gradient conditions allow. The use of an incubator-based device helps in maintaining these conditions, as described elsewhere (2–3). It should be noted that the complete processing time required (on a standard computer such as a 1.4 GHz Pentium PIV) is less than 10 s per image. It is thus reasonable to talk of real-time tracking processing, since the acquisition lapse is thus about 4 min. Moreover, this approach is obviously more robust than a human-interactive tracking method with respect to such a heavy repetitive task. Furthermore, this task may appear complex for inexperienced people if a large number (i.e., dozens) of cells are simultaneously tracked, in particular due to cell deformation, crossing, and division. With regard to this complexity, we included an automatic step-invalidation procedure in our tracking algorithm. This was successfully tested by means of a number of human-supervised experiments (not shown here).

We also developed several trajectory features in order to characterize cell locomotion. These features were extracted from the entirety of each cell trajectory established by our algorithms. This contrasts with the features used by others, e.g. Allen et al. (17) and Zicha et al. (18,19), which were computed from a part of each trajectory (i.e., between its starting point and the point at which it first crossed a virtual horizon line, also called the “horizon method”). We preferred to keep track of each cell (more exactly the part of each cell trajectory detected as valid by our automatic step-invalidation procedure), and to weight its contribution in the chemotaxis effect analysis according to the length of the cell vector taken as reference. However, taking a sole vector as reference to indicate a cell trajectory direction is not that simple. From the various experiments already performed on cell tracking, the MRDO measurement and its $\alpha$MRDO vector version emerged as interesting and robust features. Additionally, we also considered the set of step displacements performed by the different cells analyzed to detect whether a

FIG. 7. The different histograms characterizing the HUVEC cell displacements in the absence (CT) (A, B) or presence of a gastrin gradient at $10^{-8}$ M (C, D) or $10^{-9}$ M (E, F). While the upper frames (A, C, and E) show $\alpha$MRDO histograms, the lower ones (B, D, and F) present the corresponding distributions of the $v$ vectors. The remainder of the legend is similar to that of Figure 4.
majority of steps were performed in the gradient direction. To summarize the different types of information obtained, we used mean vectors, weighted histograms (taking the vector lengths into account), and their differential versions (which enhance the unbalance in the vector distribution across the different directions). As illustrated in our different experiments, these data are able to bring to the fore cell chemotaxis in the presence of a chemical gradient. Appropriate statistical analyses on directional data can confirm these observations, but in some cases these analyses must be interpreted with caution, as explained in Materials and Methods.

The next problem to be faced in a full system for automatic chemotaxis screening is how to obtain maximal information from each Dunn chamber, i.e., how to extend the observation field in order to follow a maximum number of cells. Two complementary approaches may be considered. The first consists of using a computer-controlled microscope stage to acquire multiple images by scanning a large part of the Dunn chamber. This would enable large images to be constructed. However, the robustness of the process should be checked because of possible perturbations of the actual cell locations due to the vibrations induced by the chamber displacement. The second approach (currently under study) is based on the use of a camera with a high level of resolution (i.e., $5 \times 10^6$ pixels in place of the 350,000 pixels used in the present study). The expected observed area (at the same spatial resolution level) is roughly 15 times larger than the current setup, enabling about 15 times more cells to be tracked for a single experiment (giving more accurate statistics). Furthermore, our algorithm presents robustness against resolution decrease (to half the resolution level) is roughly 15 times larger than the current setup, enabling about 15 times more cells to be tracked for a single experiment (giving more accurate statistics). Consequently, decreasing the magnification ratio of the microscope would also be possible in order to increase the observation field.

To conclude, in the present study, we developed new software and data analysis tools for automated cell tracking which enable cell chemotaxis to be efficiently analyzed. Different extensions are possible, such as the use of the data provided by our system to model cell population migration, as recently proposed by Demou and McIntire (20) using Markov chain models. Another possibility is to refine the active contour model presented in this study in order to better approximate the cell membranes of moving cells and thus be able to track the changes in cell shape during migration. With respect to the latter aim, our system seems potentially able to work on gray-level images acquired from standard cell cultures (with sufficiently high resolution and magnifying powers). This would thus exempt us from using fluorescent markers, a method carried out by various authors (e.g., 21).

**LITERATURE CITED**