Available online at www.sciencedirect.com
ScienceDirect
European Journal of Cell Biology 86 (2007) 569–578
European Journal of Cell Biology
www.elsevier.de/ejcb

Intracellular objects tracking

Yannis Kalaidzidis[a,b,*]

[a] *Max-Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstrasse 108, D-01307 Dresden, Germany*
[b] *A.N. Belozersky Institute of Physico-Chemical Biology, Building A, Moscow State University, Moscow 119899, Russia*

Received 22 March 2007; received in revised form 22 May 2007; accepted 23 May 2007

Abstract

The tracking of intracellular organelles and vesicles is becoming increasingly important for understanding cellular dynamics. Originally, the development of tracking algorithms was mainly pursued in other fields, e.g. aerospace/military/street surveillance. However, most of this algorithm is not directly applicable to live cell microscopy data. Here we describe the algorithms that have been successfully applied to object detection and tracking specifically in vivo and in vitro motility assays. The characteristics advantages and disadvantages of the different approaches are compared.

Keywords: Motion tracking; Image segmentation; Fluorescence microscopy

Introduction

A hallmark of eukaryotic cells is their content of intracellular organelles. Many of their functions are intricately linked to motor-dependent movement. For example, nutrients and signal molecules are taken up by small vesicles and later delivered to intracellular sorting compartments in a molecular motor-dependent manner. The sorting compartments known as early endosomes are themselves motile and after a cascade of homotypic fusion/fission events accumulate degradative cargo (e.g. LDL, EGF, etc.) rather than recycling cargo (e.g. transferrin). The "mature" early endosome changes its motility pattern and undertakes conversion to a late endosome (Rink et al., 2005). Recycling cargo is removed from early endosomes by a set of heterotypic fission events, where tubules bud off and are passed to recycling endosomes (Maxfield and McGraw, 2004). Secretion of signaling molecules, i.e. hormones and others are mediated by active vesicular transport to the plasma membrane (Collins, 2003; Steyer and Almers, 1999). All the above-mentioned events require vesicle movement as an essential part of their function and regulation.

With the technique of GFP labeling, in vivo microscopy provides a rich source of information concerning the intracellular vesicular transport machinery organization and regulation. Computerized microscopes easily generate sequences of thousands of frames with frame rates spanning the interval from 0.01 to 100 Hz. The quality of the images varies widely and is inversely proportional to the exposure time.

The high information density of live-cell recordings makes them hardly amenable to qualitative analysis: only the most drastic alterations of motility patterns can be scored by eye, for example "movement" or "no movement". Uncovering the subtle, but highly informative

*Corresponding author. Max-Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstrasse 108, D-01307 Dresden, Germany.
E-mail address: kalaidzi@mpi-cbg.de.

0171-9335/$ - see front matter © 2007 Elsevier GmbH. All rights reserved.
doi:10.1016/j.ejcb.2007.05.005

phenotypes resulting from alterations in the regulation of organelle movements requires a quantitative analysis approach.

Manual tracking of vesicles across successive frames of live-cell recordings is the commonly used approach to provide such data. However, besides being extremely time consuming, manual analysis is prone to systematic errors due to unconscious pre-selection of vesicles that satisfy the researcher's non-formalized criteria of "good" ones. This pre-selection is an inevitable step in any manual analysis, which remains nevertheless restricted to double-digit counts of vesicles over 100–200 frames. These problems can be overcome by automatic simultaneous tracking of hundreds of vesicles over thousands of frames, encompassing virtually all vesicles within the image frame. The resulting data sets provide statistically reliable and non-biased results, such as speed distributions, frequency of changes in directionality, processivity, sub-diffusion patterns, and intracellular positioning along with many other parameters (Qian et al., 1991; Tamura et al., 2002; Selmeczi et al., 2005; Rink et al., 2005).

Object detection and tracking algorithms were originally developed by the aerospace and military sectors to track satellites, aircrafts, and ships on basis of noisy data from sources such as radars, sonars, and telescopes. During the last quarter of the 20th century a lot of effort has been put into this field (Bar-Shalom et al., 2001). Later, multi-particle tracking approaches were applied to analyze the movement of marker particles in hydro and aerodynamic studies (velocimetry) (Chetverikov, 2001). The development of computer vision approaches, street surveillance systems, facial recognition, road tracking and other fields currently provides further applications for object tracking algorithms.

Live-cell microscopy is a field that also adopted these researches. Different approaches for the analysis of the movement of intracellular objects have been developed in the last 20 years (Qian et al., 1991; Anderson et al., 1992; Cheezum et al., 2001). It is worth mentioning that the number of publications on intracellular object tracking algorithms is less than in other fields and the algorithms are less sophisticated.

It is easy to predict that algorithms that were developed for one application are sub-optimal or even useless in other applications. First, street surveillance systems are essentially dependent on models of possible shape change constraints of the objects of interest. This method is inapplicable to the point-like objects produced by sonar/radar systems. Likewise, the shape of intracellular organelles e.g. endosomes, often has no known constraints, encompassing point-like vesicles, vacuolar structures and tubular elements.

Second, a common feature of satellite/aircraft/ship tracking algorithms is the reliance on the existence of well-defined physical models of the object movements (inertia, minimum radius of turn, maximal acceleration/deceleration; (Fortmann et al., 1983; Barniv, 1985; Bar-Shalom et al., 2001; Logothetis et al., 2002)). These provide a basis for the branch of filter-like tracking algorithms based on the kinematic model of movement (Hashiro et al., 2002; Cerveri et al., 2003). The kinematic model describes the motion of an object by the polynomial dependency of time on regulatory parameter(s) (i.e. acceleration), which is (are) themselves dependent on time in an unknown manner. The regulatory parameter estimations are updated on the basis of the new (noisy) measurements. The result of this model is a trajectory, which follows a route that is not the same as the noisy measurement points, but closer to the real (unknown) trajectory. This is the reason why tracking, in this context, is called "filtering" and tracking algorithms are also known as "filters". This branch of tracking algorithms are applicable if the uncertainty of the parameter estimations is less than the possible parameter values. In the case of intracellular organelles, the time interval between two sequential frames is too large to make any assumptions concerning their possible acceleration, trajectory smoothness, etc. The high viscosity of the cytosol and Brownian perturbation of motion, as well as the unclear mechanism of switching on and off molecular motors make the kinematic model inapplicable to microscopic objects.

Third, velocimetry measurements use the high level of correlation between directionality and the speeds of closely spaced particles to provide additional support for probabilistic approaches (Veenman et al., 2003). Again, analogous scenarios are rarely observed in the movement of intracellular organelles.

In this paper I will review the different approaches to intracellular vesicle and single molecule localization and 2D movement tracking in order to compare their respective strengths and weaknesses. The words "vesicle" and "object" will be used interchangeably and refer to the fluorescent entity to be tracked over time in the cell interior.

Background subtraction

The tracking problem consists of two logically distinguishable tasks: (a) vesicle identification on every single frame and (b) connecting identified objects in sequential frames into a chain, which belong to the same physical entity. The first procedure is commonly called the object/feature point detection, while the second is called the tracking or the trajectory linking.

In the field of radar/sonar systems, a dynamic approach was developed in which object identification is coupled with the tracking procedure to facilitate the finding of objects where the signal is close to the noise level (Barniv, 1985). To the best of my knowledge, this

approach has never been used in microscopy. Object identification was always separated from tracking. Accordingly, we first consider the object detection procedure in more detail.

The problems of unsupervised intracellular vesicle identification are the low signal-to-noise ratio of the in vivo images and the non-uniform background. Phototoxicity, which is the ability of a fluorophore to produce free-radicals by non-radiative energy transfer from their excited state to the surrounding macromolecules, forces the researcher to use a low excitation light intensity, in turn causing low signal-to-noise ratio. The major sources of background fluorescence are a cytosolic pool of the fluorescent labels that are not placed on the vesicles of interest, the autofluorescence of biological specimens, and out-of-focus objects, which produce scattered light. Another source of non-homogeneous background is uneven illumination in the microscope and different sensitivity of the CCD camera along the field of view. The non-zero dark current of CCD creates a constant offset, which can be easily subtracted.

The scanning confocal microscope essentially decreases the out-of-focus fluorescence and increases the vesicular contrast. The price of this approach is short exposure time per pixel and relatively large time per frame in comparison to epifluorescence microscopy. Another approach with reduced background is total internal reflection fluorescence (TIRF) microscopy. It suppresses the background fluorescence by decreasing the fluorescence excitation volume. The drawback of TIRF is its ability to follow only objects that are within a hundred nanometers below the cover glass surface. TIRF is a standard technique of single molecule microscopy for in vitro as well as in vivo studies of secretory granules and cortex mesh endocytic vesicles. But even in the TIRF images, the background fluorescence can be not fully eliminated and a special procedure is required to remove it.

The simplest method for background estimation is a floating mean of image (Crocker and Grier, 1996; Goulian and Simon, 2000; Sbalzarini and Koumoutsakos, 2005). The floating mean is calculated as

$$B_w(x,y) = \frac{1}{(2w+1)^2} \sum_{i=-w}^{w} \sum_{j=-w}^{w} I(x+i, y+j), \quad (1)$$

where $I(x, y)$ is an intensity at position (x, y) and w is the size of sliding window.

The size of the sliding window is chosen in such a way that it is larger than a typical object of interest but smaller than the mean distance between objects.

Steyer and Almers (1999) used a slightly more complicated algorithm. They have estimated background by applying to the original image a floating window median filter.

$$B_w(x,y) = \text{median}\{I(\xi,\eta) : \xi \in [x-w, x+w], \eta \in [y-w, y+w]\}, \quad (2)$$

where $I(x, y)$ is an intensity at position (x, y) and w is the size of sliding window.

The window size is again chosen to be larger than objects of interest and smaller than the characteristic distance between the objects.

Both approaches are linear high-frequency path filters, which remove the low frequency features of images. The linear high-frequency path filter can be applied directly by convolving the original image with a Laplacian of Gaussian (Sage et al., 2003). The linear background filtering is a fast, simple and efficient approach. Unfortunately, it cannot handle the situation when the background has a scale of intensity gradients comparable to the objects of interest. This often happens, for example, on the boundary between fluorescently labeled cytosol and nucleus. The relatively smooth changes in intensity of the cytosolically-labeled pool fall down abruptly at the transition to the unlabeled nuclear interior. The non-linear probabilistic approaches could be applied to the background estimation and elimination in such a case (Elgammal et al., 2000; Fischer et al., 2000; Guglielmetti et al., 2004). However, those approaches have never been applied to removing background from fluorescence microscopy images.

Object detection

The algorithms of object (vesicle) searching logically are divided into two major categories: single object searching and multiple object searching. Both categories are subdivided into algorithms of searching objects with known shape and without such knowledge. The objects with known a priori apparent shapes are very common in light microscopy, since both objects with fixed known shape and all objects with a size smaller than the diffraction limit of the microscope are included in this category. The diffraction limit depends on the wavelength and the numerical aperture of the microscope and defines the point spread function (PSF) of photons from the point-size light source (Born and Wolf, 1968):

$$\text{PSF}(r) = \left(\frac{2J_1\left(\frac{2\pi NAr}{\lambda}\right)}{r} \right)^2, \quad (3)$$

where J_1 is the first Bessel function; NA the numerical aperture; λ the wavelength and r is the distance from the center of PSF.

For many practical reasons, the resolution limit can be estimated as $0.6\lambda/NA$. In the case of confocal microscopy the PSF in the focal plane is a convolution of the exciting light PSF and the PSF of fluorescence

emission. As a rule of thumb, the resolution limit of confocal microscopy can be estimated as $0.4\lambda/NA$.

We first consider the single vesicle searching procedure with known shape, called correlation search. The correlation between the object image template and the searched image are calculated for all possible shifts of template relative to the analyzed image. Either the first image in the sequence or a microscope PSF for sub-resolution object or a database stored standard image can be used as a template. When the template matches the object, the correlation peaks. Gelles had applied the term correlation to convolution of template with the searched image (Gelles et al., 1988; Cheezum et al., 2001):

$$C(x,y) = \sum_{i=-\alpha/2}^{\alpha/2}\sum_{j=-\beta/2}^{\beta/2} I(x+i,y+j) \times \left[K\left(i+\frac{\alpha}{2},j+\frac{\beta}{2}\right) - \langle K\rangle\right], \quad (4)$$

where $I(x, y)$ is the intensity of image in pixel (x, y); $K(x, y)$ the intensity of template in pixel (x, y); α and β are the dimensions of template image; $\langle K\rangle$ is the mean value of template.

It is easy to see that in formula (4), this gives not a correlation matrix but a convolution of the image with the template. As a result, the brightest part of non-uniformly illuminated image will give a global peak on matrix C even if there is poor geometrical similarity between the matched area and template. In addition, formula (4) is applicable only to cases where the background in the image is either uniform or carefully removed. This drawback can be easily compensated by calculating normalized correlation coefficients (Ngoc et al., 1997; Cheezum et al., 2001; Sintorn et al., 2004; Carter et al., 2005):

$$C(x,y) = \frac{\frac{1}{\alpha\beta}\sum_{i=-\alpha/2}^{\alpha/2}\sum_{j=-\beta/2}^{\beta/2} I(x+i,y+j)K\left(i+\frac{\alpha}{2},j+\frac{\beta}{2}\right) - \langle I\rangle_{\alpha\beta}\langle K\rangle}{\sqrt{D(I)_{\alpha\beta}D(K)}}, \quad (5)$$

where $I(x, y)$ is the intensity of image in pixel (x, y); $K(x, y)$ the intensity of template in pixel (x, y); α and β are dimensions of template image; $\langle K\rangle$ the mean value of template; $\langle I\rangle_{\alpha\beta}$ the mean value of image in the area overlapped with template; $D(K)$ the variance of template and $D(I)_{\alpha\beta}$ is the variance of image in the area overlapped with template.

Since the shifts of template are calculated on a discreet pixel grid, the accuracy of object position determination is one pixel. The accuracy of object localization can be improved by approximating the correlation matrix C by a 2D parabolic function $C(x,y) = ax^2 + by^2 + dx + ey + f$ in the vicinity of its maximum value. The maximum of the parabola can be found with sub-pixel accuracy.

This method can be easily generalized to the multiple object search algorithm by searching in matrix C for many local maxima above some predefined threshold. If the statistics of image noise is known, then the noise variance of matrix C can be calculated. The reasonable threshold value can be estimated, for example, as 4σ, where σ is a standard deviation of C. The value of 2σ is generally too small and reveals too many false signals, since the probability to overcome it by chance is high, given the million of pixels in the typical image. The correlation algorithm in the form of (5) can be applied to the image without background subtraction. The only essential limitation of the correlation method is the requirement of a fixed and known shape of the searched object.

A close relative of correlation algorithm is the method of sum-absolute differences (SAD). In this method, the sum of absolute differences between the image and template is calculated at all possible shifts of the template (Barnea and Silverman, 1972; Bohs et al., 2000):

$$\mathrm{SAD}(x,y) = \sum_{i=-\alpha/2}^{\alpha/2}\sum_{j=-\beta/2}^{\beta/2}\left|I(x+i,y+j) - K\left(i+\frac{\alpha}{2},j+\frac{\beta}{2}\right)\right|, \quad (6)$$

where $I(x, y)$ is the intensity of image in pixel (x, y); $K(x, y)$ the intensity of template in pixel $(x, y)$and α and β are the dimensions of template image.

The minimum of SAD corresponds to the best coincidence of the template with the image. All comments about accuracy, parabolic interpolation, and other aspects of the correlation method mentioned above are applicable to this method, too. An additional drawback of this method, in comparison to the correlation method, is its sensitivity to the intensity scaling of image and template. This can cause problems since the fluorescent marker can bleach during acquisition time. If the labeling level is a necessary part of the process under investigation (Rink et al., 2005), the method is not applicable.

The most common single vesicle searching procedure without fixed and known shape of object is a centroid method (Cheezum et al., 2001; Carter et al., 2005). The direct implementation of this algorithm calculates the center of mass of an image:

$$\begin{pmatrix} C_x \\ C_y \end{pmatrix} = \frac{\sum_i\sum_j \begin{pmatrix} x_i \\ y_j \end{pmatrix} I_{i,j}}{\sum_i\sum_j I_{i,j}}, \quad (7)$$

where (C_x, C_y) are the center position of object and $I_{i,j}$, is the intensity in the pixel (i, j).

The summing is performed over the whole image. The method is applicable to images where only one object persists and background is removed so well that the integral of background intensity over the whole frame is a small fraction of the integral intensity of the object of interest. If this condition is violated, the centroid algorithm is prone to give coordinates of image center. There is a straightforward generalization of this method to multi-object images with background. The image of interest is smoothed to remove the high frequency noise, and then it is binarised by a threshold in such a way that all objects of interest are preserved but spaces between them are zeroed. After that the centroids are calculated separately for every connected set of non-zeroed pixels separately and the resulting values are considered to be centers of the objects of interest (Ghosh and Webb, 1994; Steyer and Almers, 1999; Apgar et al., 2000; Suh et al., 2003; Carter et al., 2005).

Crocker used some modifications of this method. After an initial guess about the object center, the centroid is calculated by summing not in the mask area but inside the circle with radius w over the non-masked image:

$$\begin{pmatrix} C_x^t \\ C_y^t \end{pmatrix} = \frac{\sum_{i,j:\left(x_i - C_x^{t-1}\right)^2 + \left(y_j - C_y^{t-1}\right)^2 \leqslant w^2} \begin{pmatrix} x_i \\ y_j \end{pmatrix} I_{i,j}}{\sum_{i,j:\left(x_i - C_x^{t-1}\right)^2 + \left(y_j - C_y^{t-1}\right)^2 \leqslant w^2} I_{i,j}}, \tag{8}$$

centered on position of initial guess of the object position on iteration $t-1$ $\begin{pmatrix} C_x^{t-1} \\ C_y^{t-1} \end{pmatrix}$. If after a calculation the position of object shifts more than 0.5 pixels, the iteration (8) is repeated with the position of the newly found center (Crocker and Grier, 1996; Sbalzarini and Koumoutsakos, 2005; Ewers et al., 2005).

Another modification of this algorithm, which provides the maximum accuracy over all centroid algorithms, is called Gaussian mask (Thompson et al., 2002). The summing in the Gaussian mask algorithm is performed over the whole image but in convolution with a Gaussian kernel:

$$\begin{pmatrix} C_x^t \\ C_y^t \end{pmatrix} = \frac{\sum_i \sum_j \begin{pmatrix} x_i \\ y_j \end{pmatrix} I_{i,j} N_{i,j}^{t-1}}{\sum_i \sum_j I_{i,j} N_{i,j}^{t-1}}, \tag{9}$$

where $I_{i,j}$, intensity in the pixel (i, j),

$$N_{i,j}^{t-1} = \int_{x_i}^{x_{i+1}} \int_{y_j}^{y_{j+1}} \times \exp\left(- \frac{\left(x - C_x^{t-1}\right)^2 + \left(y - C_y^{t-1}\right)^2}{2w^2} \right) \mathrm{d}x\, \mathrm{d}y \tag{10}$$

is an integral of the Gaussian kernel over pixel (i, j) with the kernel centered in the position $({C_x}^{t-1}, {C_y}^{t-1})$ of previous iteration $t-1$. w, is the width of Gaussian. As in the case of formula (8) the calculation is repeated iteratively, since the value ${N_{i,j}}^t$ has to be recalculated after every adjustment of object center position. The iteration is stopped when the correction of position of object becomes below the accuracy level specified by the user.

The next class of sub-resolution size object searching is fitting of image intensities by PSF. The shape of a microscope PSF is defined as the Airy disk (see formula (3)), but in the most cases it is approximated by a Gaussian (Anderson et al., 1992; Cheezum et al., 2001; Thompson et al., 2002; Yildiz et al., 2003; Li et al., 2004; Bonneau et al., 2005).

$$S = \sum_i \sum_j \left[(I(i,j) - A \exp) \times \left(- \frac{(x_i - x_0)^2 + \left(y_j - y_0\right)^2}{2w^2} \right) - B \right]^2, \tag{11}$$

where $I(i, j)$ is the intensity of image; A the amplitude of Gaussian; (x_0, y_0), position of object; w the width of PSF and B is the background intensity; sum is going over whole image.

The parameters (x_0, y_0, A, B) which minimize the squared difference between the Gaussian and the image are taken as features of the object. If background was subtracted in advance, the B is considered equal to zero and excluded from the fitting procedure. The minimization procedure is non-linear and generally iterative. This fitting approach is the most computationally expensive, but it gives the best accuracy (Cheezum et al., 2001; Thompson et al., 2002; Yildiz et al., 2003; Snyder et al., 2004; Kural et al., 2005). A Gaussian fit approach does have a drawback in that it is applicable only to objects with geometrical size smaller than the diffraction limit of the microscope.

Like in the case of centroid fitting, the Gaussian fit can be easily generalized to the multi-object case. The image is preprocessed by subtracting background and by smoothing. Then the local peaks above threshold are found as the candidate points of possible object localization and the fitting is performed in the local vicinity of every candidate point.

The advantage of having sub-resolution objects is an additional possibility to filter out false signals, since the intensity characteristic of correct signal is either known in advance or can be found by clustering their moments of intensity distribution (Crocker and Grier, 1996; Sbalzarini and Koumoutsakos, 2005).

Different gradient edge detection and thresholding algorithms (Tvarusko et al., 1999; Gasman et al., 2003; Ponti et al., 2003; Ku et al., 2007) are used to search objects with sizes above the diffraction limit and unknown shape. They use a variety of ad hoc procedures to verify that selected regions satisfy criteria of the object of interest. Both categories can produce a binary mask, which is later combined with centroid algorithms.

The threshold algorithms are based on the threshold value, which allows for the conversion of the original gray-scale image to the binary image in such a way that all the objects of interest are marked white, and all other pixels black. In the simple case of homogeneous background and high signal-to-noise ratio the threshold value can be manually selected by the user on basis of the examination of a single frame (Apgar et al., 2000). Different ad hoc approaches were invented to select the threshold value automatically. For example, the 30th percentile of brightness of the entire image was chosen as the threshold in the work of Crocker and Grier (1996). Ku et al. (2007) have used the heuristic $T = \max(I_{i,j})(\mathrm{std}(I_{i,j})/\max(I_{i,j}) - \mathrm{mean}(I_{i,j}))$, where std means standard deviation, and the calculation is done over all pixels of the image (i, j). A more solid approach was proposed on the basis of probability theory. It is worth mentioning the maximum entropy as a criterion for threshold selection (Leung and Lam, 1996). Brink (1996) uses the maximum entropy formula that includes the local correlation of pixels in the image.

The edge detection algorithm consists of a set of rules which discriminate between the candidate pixels on the basis of the characteristics of surrounding pixels (i.e. a pixel is classified as an edge pixel if it has a potential continuation in two directions and if the gradient is maximal between the two neighbors in the direction of the gradient) (Tvarusko et al., 1999), after the closed boundaries were formed; another set of rules were used to eliminate false objects (i.e. the intensity distribution within the boundary has to satisfy some criteria, e.g. mean value and standard deviation). This approach is quite complicated and has a less solid basis in comparison with intensity fitting, but its implementation could be very efficient and fast.

Another approach to objects with sizes above the resolution limit was used by Rink et al. (2005). The objects were fitted by a sum of squared Lorenzians:

$$S = \sum_{i,j} \frac{1}{\sigma_{ij}^2} \left(I(i,j) - \sum_{k=1}^{M} \frac{A_k}{1 + \left(\left(\frac{(x_i - x_k)\cos\alpha_k - (y_j - y_k)\sin\alpha_k}{w_k} \right)^2 + \left(\frac{(x_i - x_k)\sin\alpha + (y_j - y_l)\cos\alpha_k}{h_k} \right)^2 \right)^2} \right)^2, \quad (12)$$

where $(x_k, y_k, A_k, w_k, h_k, \alpha_k)$ are the parameters of kth Lorenzian; M the number of Lorenzians; $\sigma_{i,j}$ is the standard deviation of noise in the pixel (i, j).

Any intensity distribution could be presented by a sum of hat-like functions. The Lorenzian has some advantages because the calculation is less expensive than the Gaussian. At the same time, the difference between the Airy function and squared Lorenzian is small relative to the noise level of typical live-cell images. This approach also has the advantage of accuracy comparable to Gaussian fit algorithms and the ability to find vesicles with size ranges from hundred nanometers to a few micrometers. The elongation of the base function along an arbitrary axis with angle α_k to the axis X decreases the number of base functions required for accurate object deconvolution.

Tracking algorithms

After objects are detected on every frame of the sequence, the tracking is reduced to connecting the information about objects into a chain. In cases where only a single object exists in every frame, this task is trivial. But problems arise when there are many objects or one real object and many false signals (clutter). In this situation the assignment of objects found on different frames to the same track becomes a non-trivial problem. A set of methods was developed on the basis of object shape analysis and possible kinematic restrictions on object shape changes (Vieren et al., 1995; Cham and Rehg, 1999; Cerveri et al., 2003; Mitiche et al., 2003; Chang et al., 2005). This restriction is applicable to surveillance systems and has almost no practical impact on intracellular object analysis. Aerospace/ship tracking, flow velocimetry, and intracellular object tracking share a common situation where object shape is not available, either because of low signal-to-noise ratio and resolution limitation or because it changes too fast in between two sequential measurements. As a result, point tracking algorithms where developed in those fields.

In the simplest case, on every frame there is the same fixed number of objects. This means that consideration is restricted to the case when no new object appears and no object disappears during the measurement. If the reasonable scoring for object correspondence is available, the problem of finding the best object-to-track assignment is reduced to the well-known problem of

optimal resource distribution. The classical Hungarian algorithm provides the solution (Kuhn, 1955). The scoring system can include the features of tracks we consider essential: distance of object from either previously measured or predicted position, trajectory smoothness and some object characteristics like integral intensity. The Hungarian algorithm works on per frame basis and provides a global minimum score assignment of objects to the tracks. The initial track seeds in this case are objects found on the first frame of the sequence.

Unfortunately, a fixed number of objects found in every given frame are an impractical assumption. In microscopy, new objects appear either by genesis of compartments, vesicles, proteins, fission vesicles, etc., or just by their coming into the field of view. At the same time some existing object can disappear by moving out of focus, changing identity, or fusing with another object. In addition to the change in the number of real objects, the false object recognition is an essential problem in the field. The same scenario is applicable to the radar/sonar tracking systems. The targets can come to the surveillance region and leave it. In addition the false measurements (clutter) contaminate virtually every measurement. It could be false targets in military radar/sonar measurements or spikes of background noise in fluorescence microscopy. Starting from the 60 s the probabilistic approaches of proper tracking (measurement-to-track assignment) in these complicated conditions were developed (Bar-Shalom and Fortmann, 1988; Cox, 1993; Bar-Shalom et al., 2001; Veenman et al., 2003). The probabilistic approach for target tracking is based on estimation of the join probabilistic data association (JPDA) (Fortmann et al., 1983). The main idea of JPDA is maximizing probability of particular object-to-track assignment given a model of object movement. The JPDA can be implemented with EM algorithm when the first assignment is done on the basis of join probability. Then, the probability distribution is updated on the basis of each newly done assignment. The alternative approach is multiple hypothesis tracking (MHT) (Reid, 1979; Cox and Hingorani, 1996). In the MHT algorithms, many alternative versions of tracks are kept until a late time point, when the distinctions of scores between alternative trajectories becomes essential enough to make a decision.

Despite the algorithms mentioned above being called multi-target tracking algorithms, the number of simultaneously tracked objects is very limited. Those algorithms belong to the group of NP-hard algorithms, meaning that their complexity grows exponentially with increasing number of tracked objects. This is acceptable in the radar/sonar systems, when the number of targets is generally in a range 1–20. In intracellular microscopy and velocimetry, the number of objects is in range of 100–10000. Therefore less sophisticated algorithms are generally used in this field.

The simplest algorithm is just assignment of the object in a predefined vicinity. If there is no object in the area or there is more than one object, the track is broken. This algorithm was used, for instance, in the works of Ghosh and Webb (Ghosh and Webb, 1994; Goulian and Simon, 2000). This approach works only in cases of slowly moving sparse objects with high signal-to-noise ratio.

Next in range of complexity is a greedy approach (Anderson et al., 1992; Verestoy and Chetverikov, 2000). Scores of different types (probabilistic or ad hoc) are calculated for all possible track continuations. The best score is assigned first, then the next one and so on. If the best continuation for a given track is used by another track, then the second best is chosen. The different characteristics of tracks, which are applicable to the particular case, can be included in the scoring function. The score always includes the position, and then speed, trajectory smoothness, object intensity characteristics and others. Sethi and Jain (Sethi and Jain, 1987; Chetverikov and Verestói, 1999) offered the score function which maps the score to the interval (0, 1), where zero corresponds to the perfect fit.

$$\text{Score} = 1 - 2\frac{\sqrt{x_i x_{i+1}}}{x_i + x_{i+1}}, \tag{13}$$

where x_i and x_{i+1}, are the parameters of the object measured in the frames i and $i+1$, respectively.

Penalty for the trajectory non-smoothness is calculated by:

$$\text{Score} = 1 - \frac{\vec{v}_i \cdot \vec{v}_{i+1}}{|\vec{v}_i| \cdot |\vec{v}_{i+1}|}, \tag{14}$$

where v_i and v_{i+1}, shifts of object between frames $i-2/i-1$ and $i-1/i.$.

The total score is calculated as a weighted sum of scores for all applicable parameters. The best score is the minimal one.

Another approach is a quadratic penalty score when the squares of parameters differences are summed. This approach functionally corresponds to the previous one, but it requires appropriate scaling coefficients for different parameters, i.e. position and intensities:

$$\text{Total score} = \sum_i \alpha_i \left(p_{i,t} - p_{i,t-1}\right)^2,$$

where α_i, is a scaling factor; $p_{i,t}$, ith parameter value at frame t; sum is done over the all parameters.

The scale factor α compensates the different units of different parameters (i.e. distance and intensity).

The probability of different assignment can also be considered as a score. Anderson and coauthors

(Anderson et al., 1992) use probability as a score:

$$P = \exp\left(-\left(\frac{\Delta R}{R_d}\right)^2\right)\exp\left(-\left(\frac{\Delta Z}{Z_n}\right)\right), \quad (15)$$

where ΔR, is a shift of the object between two sequential frames; R_d, the characteristic diffusive radius; ΔZ the change of the object intensity; Z_n is the characteristic intensity.

The algorithm for probability maximization could be easily converted to penalty minimization by taking $-\ln(P)$ as a score.

The advantage of the greedy algorithms is the ability to handle situations of object disappearing and temporary occlusions. The upper limit on the possible score can be introduced by adding a dummy object with fixed maximum score. If there are no better choices, the dummy object is assigned to the track and it causes a track break. A little more handling is required in order to keep the dummy-marked track for a predefined number of frames to handle possible occlusion.

The disadvantage of greedy algorithm is its clear tendency to fall into the first local minimum of search space. If the object density is low and movement is either slow or well organized so that the score difference between the possible candidates is large, then the greedy algorithm is a good choice. In the opposite case of dense fast objects, the number of errors that are produced by the greedy algorithm becomes large.

The next step in improvement of the searching algorithm is the iterative optimization of the results from the greedy algorithm (Sbalzarini and Koumoutsakos, 2005). More clearly, at the first stage, the greedy assignment of objects to tracks is done. Then the iterative process runs: for every track i marked by dummy object the best replacement is searched. If it is found and it is already assigned to another track j, then the second best (not assigned yet) candidate for the track j is searched. If a candidate is found and this double replacement improves the total score, the new assignment is done and process repeated until either all tracks are non-dummy or no appropriate substitution is found. On the per-frame scoring, this approach gives a better result than the greedy algorithm. It has a smaller number of broken tracks, but it does not guarantee a global optimum finding. The search for possible reassignment not only through the dummy-marked (potentially broken) tracks but over the whole track set is impossible because there are an exponentially growing number of replacement combinations.

Another improvement can be achieved by applying dynamic programming for track assignment (Sage et al., 2003). The dynamic programming algorithm, in the case of assigning objects to tracks on a per-frame basis, can be reduced to the classical matrix implementation.

Sometimes one cannot assign the fast moving objects in the crowded environment only on the basis of one frame analysis. The straightforward expansion considers at once many sequential frames. In this case the dynamic programming procedure becomes more complicated but is still possible (Rink et al., 2005).

Conclusion

A variety of intracellular object tracking algorithms have been implemented in the last two decades. Unfortunately, only a few works have compared the accuracy of those different approaches of object localization (Cheezum et al., 2001; Thompson et al., 2002; Ober et al., 2004) and no one has compared different approaches to object tracking. There are clear unsolved problems in microscopy tracking, for example, examination of a movie of live cell endosomal behavior shows that the number of endosomes which human beings can see on the movie is much higher than the number of endosomes one can see on still images. The current state of the art object searching algorithms, which are based on function fitting procedures, can find virtually all objects visible on one frame, but it is considerably less than can be seen on the movie. In the case of many fast objects in a crowded surrounding, the human eye is still better than the best available software. It can be further stressed by the simple fact that manual checking is the 'gold' standard of the tracking algorithm quality. I hope that future developments in the tracking algorithms, specifically applied to intracellular object tracking, will better integrate more mathematically developed procedures from other tracking fields.

References

Anderson, C.M., Georgiou, G.N., Morrison, I.E.G., Stevenson, G.V.W., Cherry, R.J., 1992. Tracking of cell surface receptors by fluorescence digital imaging microscopy using a charge-coupled device camera low-density lipoprotein and influenza virus receptor mobility at 4 °C. J. Cell Sci. 101, 415–425.

Apgar, J., Tseng, Y., Fedorov, E., Herwig, M.B., Almo, S.C., Wirtz, D., 2000. Multiple-particle tracking measurements of heterogeneities in solutions of actin filaments and actin bundles. Biophys. J. 79, 1095–1106.

Bar-Shalom, Y., Fortmann, T.E., 1988. Tracking and Data Association. Academic Press Inc., London.

Bar-Shalom, Y., Li, X.R., Kirubarajan, T., 2001. Estimation with Applications to Tracking and Navigation. Wiley, New York.

Barnea, D.I., Silverman, H.F., 1972. A class of algorithms for fast digital image registration. IEEE Trans. Comput. C-21, 179–186.

Barniv, Y., 1985. Dynamic programming solution for detecting dim moving targets. IEEE Trans. Aerospace Electron. Syst. AES 21, 144–156.

Bohs, L.N., Geiman, B.J., Anderson, M.E., Gebhart, S.C., Trahey, G.E., 2000. Speckle tracking for multi-dimensional flow estimation. Ultrasonics 38, 369–375.

Bonneau, S., Dahan, M., Cohen, L.D., 2005. Tracking single quantum dots in live cell with minimal path. Proc. IEEE Comp. Soc. Conf. Comp. Vision Pattern Recogn. (CVPR) 05, 141–148.

Born, M., Wolf, E., 1968. Principles of Optics. Pergamon Press, Oxford-London-Edinburgh-New York-Paris-Frankfurt.

Brink, A.D., 1996. Using spatial information as an aid to maximum entropy image threshold selection. Pattern Recogn. Lett. 17, 29–36.

Carter, B.C., Shubeita, G.T., Gross, S.P., 2005. Tracking single particles: a user-friendly quantitative evaluation. Phys. Biol. 2, 60–72.

Cerveri, P., Pedotti, A., Ferrigno, G., 2003. Robust recovery of human motion from video using Kalman filters and virtual humans. Hum. Mov. Sci. 22, 377–404.

Cham, T.-J., Rehg, J.M., 1999. A multiple hypothesis approach to figure tracking. In: Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR'99), Fort Collins, Colorado, USA, June 23–25, pp. 239–245.

Chang, C., Ansari, R., Khokhar, A., 2005. Multiple object tracking with kernel particle filter. IEEE Comp. Soc. Conf. Comp. Vision Pattern Recogn. (CVPR'05) 1, 566–573.

Cheezum, M.K., Walker, W.F., Guilford, W.H., 2001. Quantitative comparison of algorithms for tracking single fluorescent particles. Biophys. J. 81, 2378–2388.

Chetverikov, D., 2001. Particle image velocimetry by feature tracking. Comput. Anal. Images Patterns 21–24, 325–332.

Chetverikov, D., Verestói, J., 1999. Feature point tracking for incomplete trajectories. Computing 62, 321–338.

Collins, R.N., 2003. Rab and ARF GTPase regulation of exocytosis. Mol. Membr. Biol. 20, 105–115.

Cox, I.I., 1993. A review of statistical data association techniques for motion correspondence. Int. J. Comp. Vision 10, 53–66.

Cox, I.I., Hingorani, S.L., 1996. An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking. IEEE Trans. Pattern Anal. Mach. Intell. 18, 138–150.

Crocker, J.C., Grier, D.G., 1996. Method of digital video microscopy for colloidal studies. J. Colloid Interface Sci. 179, 298–310.

Elgammal, A., Harwood, D., Davis, L., 2000. Non-parametric model for background subtraction. In: Proceedings of the Sixth European Conference Computer Vision, Dublin, Ireland, June 26–July 1.

Ewers, H., Smith, A.E., Sbalzarini, I.F., Lilie, H., Koumoutsakos, P., Helenius, A., 2005. Single particle tracking of murine polyoma virus-like particles on live cell and artificial membranes. Proc. Natl. Acad. Sci. USA 102, 15110–15115.

Fortmann, T.E., Bar-Shalom, Y., Scheffe, M., 1983. Sonar tracking of multiple targets using joint probabilistic data association. IEEE J. Oceanic Eng. OE 8 (3), 173–184.

Fischer, R., Dose, V., von der Linden, W., 2000. Background estimation in experimental spectra. Phys. Rev. E 61, 1152–1160.

Gasman, S., Kalaidzidis, Y., Zerial, M., 2003. RhoD regulates endosome dynamics through Diaphanous-related formin and Src tyrosine kinase. Nat. Cell Biol. 5, 195–204.

Gelles, J., Schnapp, B.J., Sheetz, M.P., 1988. Tracking kinesin-driven movements with nanometer scale precision. Nature 331, 450–453.

Ghosh, R.N., Webb, W.W., 1994. Automated detection and tracking of individual and clustered cell surface low density lipoprotein receptor molecules. Biophys. J. 66, 1301–1318.

Goulian, M., Simon, S.M., 2000. Tracking single proteins within cells. Biophys. J. 79, 2188–2198.

Guglielmetti, F., Fischer, R., Dose, V., 2004. Mixture modeling for background and sources separation in X-ray astronomical images. In: Twenty-fourth International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering.

Hashiro, M., Kawase, T., Sasase, I., 2002. Maneuver target tracking with an acceleration estimator using target past positions. Electron. Commun. Japan – Part I 85, 29–37.

Ku, T.-C., Huang, Y.-N., Huang, C.-C., Yang, D.-M., Kao, L.-S., Chiu, T.-Y., Hsieh, C.-F., Wu, P.-Y., Tsai, Y.-S., Lin, C.-C., 2007. An automated tracking system to measure the dynamic properties of vesicles in living cells. Microsci. Res. Technol. 70, 119–134.

Kuhn, H.W., 1955. The Hungarian method for the assignment problem. Naval Res. Logistics Quart. 2, 83–97.

Kural, C., Kim, H., Syed, S., Goshima, G., Gelfand, V.I., Selvin, P.R., 2005. Kinesin and dynein move a peroxisome in vivo: a tug-of-war or coordinated movement? Science 308, 1469–1472.

Leung, C.K., Lam, F.K., 1996. Performance analysis for a class of iterative image thresholding algorithms. Pattern Recogn. 29, 1523–1530.

Li, C.H., Bai, L., Li, D.D., Xia, S., Xu, T., 2004. Dynamic tracking and mobility analysis of single GLUT4 storage vesicle in live 3T3-L1 cells. Cell Res. 14, 480–486.

Logothetis, A., Krishnamurthy, V., Holst, J., 2002. A Bayesian EM algorithm for optimal tracking of a maneuvering target in clutter. Signal Process. 82, 473–490.

Maxfield, F., McGraw, T., 2004. Endocytic recycling. Nat. Rev. Mol. Cell Biol. 5, 121–132.

Mitiche, A., Faghali, R., Mansouri, A., 2003. Motion tracking as spatio-temporal motion boundary detection. Robot. Auton. Syst. 43, 39–50.

Ngoc, S.N., Briquet-Laugier, F., Boulin, C., Olivo, J.-C., 1997. Adaptive detection for tracking moving biological objects in video microscopy sequences. Proc. Int. Conf. Image Process. 3, 484–487.

Ober, R.J., Ram, S., Ward, E.S., 2004. Localization accuracy in single-molecule microscopy. Biophys. J. 86, 1185–1200.

Ponti, A., Vallotton, P., Salmon, W.C., Waterman-Storer, C.M., Danuser, G., 2003. Computational analysis of F-actin turnover in cortical actin meshworks using fluorescent speckle microscopy. Biophys. J. 84, 3336–3352.

Qian, H., Sheetz, M.P., Elson, E., 1991. Single-particle tracking: analysis of diffusion and flow in two-dimensional systems. Bipohys. J. 60, 910–921.

Reid, D., 1979. An algorithm for tracking multiple targets. IEEE Trans. Autom. Control AC- 24, 843–854.

Rink, J., Ghigo, E., Kalaidzidis, Y., Zerial, M., 2005. Rab conversion as a mechanism of progression from early to late endosomes. Cell 122, 735–749.

Sage, D., Hediger, F., Gasser, S.M., Unser, M., 2003. Automatic tracking of particles in dynamic fluorescence microscopy. In: Proceedings of the Third International Symposium on Image and Signal Processing and Analysis (ISPA'03), Rome, Italy, September 18–20, pp. 582–586.

Sbalzarini, I.F., Koumoutsakos, P., 2005. Feature point tracking and trajectory analysis for video imaging in cell biology. J. Struct. Biol. 151, 182–195.

Selmeczi, D., Mosler, S., Hagedorn, P.H., Larsen, N.B., Flyvbjerg, H., 2005. Cell motility as persistent random motion: theories from experiments. Biophys. J. 89, 912–931.

Sethi, I.K., Jain, R., 1987. Finding trajectories of feature points in a monocular image sequence. IEEE Trans. Pattern Anal. Mach. Intell. 9, 56–73.

Sintorn, I.-M., Homman-Loudiyi, M., Soderberg-Naucler, C., Borgefors, G., 2004. A refined circular template matching method for classification of human cytomegalovirus capsids in TEM images. Comp. Meth. Prog. Biomed. 76, 95–102.

Snyder, G.E., Sakamoto, T., Hammer, J.A., Sellers, J.R., Selvin, P.R., 2004. Nanometer localization of single green fluorescent proteins: evidence that myosin V walks hand-over-hand via Telemark configuration. Biophys. J. 87, 1776–1783.

Steyer, J.A., Almers, W., 1999. Tracking single secretory granules in live chromaffin cells by evanescent-field fluorescence microscopy. Biophys. J. 76, 2262–2271.

Suh, J., Wirtz, D., Hanes, J., 2003. Efficient active transport of gene nanocarriers to the cell nucleus. Proc. Natl. Acad. Sci. USA 100, 3878–3882.

Tamura, K., Hidaka, Y., Yusuf, Y., Kai, S., 2002. Anomalous diffusion and Levy distribution of particles velocity in soft-mode turbulence in electroconvection. Physica A 306, 157–168.

Thompson, R.E., Larson, D.R., Webb, W.W., 2002. Precise nanometer localization analysis for individual fluorescent probes. Biophys. J. 82, 2775–2783.

Tvarusko, W., Bentele, M., Misteli, T., Rudolf, R., Kaether, C., Spector, D.L., Gerdes, H.H., Eils, R., 1999. Time resolved analysis and visualization of dynamic processes in living cells. Proc. Natl. Acad. Sci. USA 96, 7950–7955.

Veenman, C.J., Reinders, M.J.T., Backer, E., 2003. Motion tracking as a constrained optimization problem. Pattern Recogn. 36, 2049–2067.

Verestoy, J., Chetverikov, D., 2000. Experimental comparative evaluation of feature point tracking algorithms. In: Proceedings of the Workshop on Evaluation and Validation of Computer Vision Algorithms, pp. 183–194.

Vieren, C., Cabestaing, F., Postaire, J.-G., 1995. Catching moving objects with snakes for motion tracking. Pattern Recogn. Lett. 16, 679–685.

Yildiz, A., Forkey, J.N., McKinney, S.A., Ha, T., Goldman, Y., Selvin, P., 2003. Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. Science 300, 2061–2065.