

## Multiple objects tracking in fluorescence microscopy

Yannis Kalaidzidis

Received: 24 July 2007 / Revised: 2 January 2008 / Published online: 14 May 2008  
© The Author(s) 2008. This article is published with open access at Springerlink.com

**Abstract** Many processes in cell biology are connected to the movement of compact entities: intracellular vesicles and even single molecules. The tracking of individual objects is important for understanding cellular dynamics. Here we describe the tracking algorithms which have been developed in the non-biological fields and successfully applied to object detection and tracking in biological applications. The characteristics features of the different algorithms are compared.

**Keywords** Motion tracking · Object detection · Image segmentation · Fluorescence microscopy · Subpixel resolution

**Mathematics Subject Classification (2000)** 68W99

### 1 Introduction

Many processes in cell and developmental biology are connected to the movement of materials and signals in the form of compact entities: intracellular organelles, vesicles, and even cells themselves in context of the organism. The regulation of

---

**Electronic supplementary material** The online version of this article (doi:[10.1007/s00285-008-0180-4](https://doi.org/10.1007/s00285-008-0180-4)) contains supplementary material, which is available to authorized users.

---

Y. Kalaidzidis (✉)  
Max-Planck Institute of Molecular Cell Biology and Genetics,  
Pflötenhauerstrasse 108, 01307 Dresden, Germany  
e-mail: kalaidzi@mpi-cbg.de

Y. Kalaidzidis  
A.N. Belozersky Institute of Physico-Chemical Biology, Building "A",  
Moscow State University, Moscow 119899, Russia

these movements, their characteristics, and their physiological meaning are still open questions. Many mechanisms are involved in the movements mentioned above: molecular motors run over the actin and microtubule cytoskeletons, actin polymerization/depolymerization forces cell processes to grow, microtubule assembly/disassembly move and separate chromosomes during cell division. In addition, cell adhesion, molecular diffusion, and liquid flow all take part in biological trafficking.

For example, in a single cell nutrients and signaling molecules are taken up by small vesicles and later delivered to intracellular sorting compartments in a molecular motor-dependent manner. Sorting compartments known as early endosomes are 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, etc.). The “mature” early endosome changes its pattern of motility and undertakes conversion to a late endosome [54]. Recycling cargo is removed from early endosomes by a set of heterotypic fission events, where tubules bud off and are passed onto recycling endosomes [45]. The secretion of signals, hormones, etc. is mediated by active vesicular transport to the plasma membrane [22,66]. All of the events mentioned above require the movement of vesicles as an essential part of their function and regulation.

In intracellular microscopy, fluorescent markers allow for objects to be followed that are much smaller than the diffraction limit of light. With molecular biology it is possible to genetically incorporate a fluorophore with given spectral characteristics (generally the Green Fluorescent Protein (GFP) and its derivatives) as a tag for the protein-of-interest. This is a powerful tool for the visualization of specific compartments *in vivo* and allows the researcher to follow them over a relatively long time period. This is a rich source of information concerning the organization and regulation of the intracellular vesicular transport machinery. Computerized microscopes are able to 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 generally inversely proportional to the exposure time.

The high information density of live cell recordings makes them barely amenable to qualitative analysis: only the most drastic alterations of motility patterns can be scored by eye, for example “movement” or “no movement”, extreme redistribution of intracellular objects, etc. Uncovering more subtle, but highly informative 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 most commonly used approach to provide such data. However, besides being extremely time-consuming, manual analysis is prone to systematic errors due to the unconscious pre-selection of vesicles which satisfy the researcher’s non-formalized criteria of “good data”. 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 the automated simultaneous tracking of hundreds of vesicles over thousands of frames, encompassing virtually all vesicles within the image. The resulting data set provides 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 [34,49,54,58,69].

This quantitative analysis includes vesicle detection and tracking. These two procedures can be considered independent in most cases, although some algorithms use the information from the tracking procedure to guide object detection [85]. For intracellular fluorescent object detection a low signal-to-noise ratio and a Poisson noise distribution are typical. The same problem of spot-like object detection in the presence of Poisson noise originally came about in the fields of astronomy and radar readout automation. Object tracking algorithms were originally developed by the aerospace and military sectors to track satellites, aircraft, and ships on the basis of noisy data from sources such as radars, sonars, and telescopes. Starting from the 1960s a lot of effort has been put into this field [6, 7]. The first algorithms were implemented in the analog and semi-analog computers of the military tracking equipment of those times. Later, multi-particle tracking approaches were applied to analyze the movement of marker particles in hydro- and aerodynamic studies (velocimetry) [20, 74]. The development of new approaches in computer vision, street surveillance systems, facial recognition, road tracking and other fields currently provide further applications for object tracking algorithms.

Intracellular live-cell microscopy is a field which has adopted the results and methodology of the above research. Different approaches for the analysis of the movement of intracellular objects have been developed in the last 20 years [2, 16, 49, 57]. Predictably, the algorithms that were developed for one application are sub-optimal or even useless in other applications. Firstly, street surveillance systems are essentially constrained by models of possible shape changes for the objects of interest [50]. 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. It is worth mentioning that intracellular object tracking algorithms are less sophisticated than those used in the radar/sonar tracking field. This can be explained partially by the huge number of objects that have to be tracked, which makes it hard and sometimes impossible to implement an algorithm where complexity grows quickly with the number of simultaneously tracked objects, as is the case with the many thousands of vesicles in the cell interior.

A common feature of satellite/aircraft/ship tracking algorithms is the reliance on well-defined physical models for the object's movements (e.g., inertia, minimum radius of turn, maximal acceleration/deceleration; [7, 9, 28, 33, 44]). These models provide a basis for the branch of filter-like tracking algorithms based on the kinematic model of movement [14, 33]. 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. These regulatory parameter estimations are updated on the basis of the new (noisy) measurements. The result of this model is a trajectory that follows a route which is not the encompassing 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 is 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. In addition, the high viscosity of the cytosol, Brownian motion, and

the unknown mechanisms for switching molecular motors on and off make kinematic models inapplicable to microscopic objects.

Velocimetry measurements use the high level of correlation between directionality and the speeds of closely spaced particles to provide an additional support for probabilistic approaches and to follow not individual particles but groups of particles simultaneously [1, 36, 73]. Again, analogous scenarios are rarely observed in the movement of intracellular organelles.

In this paper I will review the different approaches to small point-like object tracking and specifically those which were used for intracellular vesicle and single molecule tracking.

It can be mentioned, that the transition from 2D to 3D compact object detection/tracking is straightforward in most cases. The additional dimensionality does not require any fundamental adaptations of the tracking algorithms. Even more so-especially when compared to tracking of maximum-projected pseudo-3D-data sets, the additional spatial dimension may decrease the spatial density of targets and thus improve the tracking quality of almost all tracking algorithms. Likewise, object detection in 3D is fundamentally similar to 2D object detection. However, the considerably larger size of 3D image sets in comparison with 2D-data of the same resolution may require implementation on new generation computers with increased memory/computing capacity. Since almost all algorithms considered in this review were originally developed for 2D image analysis, I will mostly restrict myself to 2D-scenarios.

In the text, the terms object and target are used interchangeably to denote the real objects, the signal and alarm denote the result of an object detection algorithm, which can arise from real objects as false objects.

## 2 Object detection

The algorithms for intracellular object detection are logically divided into two main categories: single object searching and multiple object searching. Objects with known a priori shapes are very common in light microscopy, since this category includes both objects with known shape and all objects whose size is smaller than the resolution limit of the microscope. The resolution limit of microscope in lateral direction is defined by light diffraction and depends on the wavelength and the numerical aperture of the microscope lens. For many practical reasons, the resolution limit of epifluorescence microscopy in the  $XY$ -image plane can be estimated as  $\frac{0.6\lambda}{NA}$ , where  $\lambda$  is the wavelength and  $NA$  the numerical aperture. In the case of confocal microscopy, the point-spread function (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  $\frac{0.4\lambda}{NA}$ . The axial resolution of confocal microscopy in the limit of small pinhole is  $FWHM_z = \frac{0.64\lambda}{(n - \sqrt{n^2 - NA^2})}$ , where  $\lambda$  is the wavelength,  $NA$  the numerical aperture, and  $n$  is the refraction index of immersion liquid.

We will first consider the searching procedure for objects with a known shape. The known shape of the object is encoded in the “template” image. The correlation between the template image and the searched image are calculated for all possible shifts of the

template relative to the analyzed image. The first image in the sequence, a microscope PSF for a sub-resolution object, or a standard image stored in a database can be used as a template. When the template matches the object, the correlation peaks. Gelles et al. have used this algorithm to track plastic beads attached to kinesins and moving along microtubules in vitro with nanometer accuracy [8, 16, 30]:

$$C(x, y) = \sum_{i=-w}^w \sum_{j=-h}^h I(x+i, y+j) \cdot (T(i+w, j+h) - \langle T \rangle) \tag{1}$$

where  $I(x, y)$  is the intensity of image in pixel  $(x, y)$ ,  $T(x, y)$  the intensity of template in pixel  $(x, y)$ ,  $w$  and  $h$  the characteristic size of template image, and  $\langle T \rangle$  is the mean value of template intensity.

Since shifts of the template are calculated on a discrete pixel grid, the accuracy of the determination of the object’s position is one pixel. In the original work of Gelles [30] the pixel size was  $54 \times 43$  nm. The authors have used some modification of the “centroid” method to improve the localization accuracy to 1.7 nm. The threshold value  $Th$  was selected and the coordinates of the center of the bead are calculated by the formula:

$$x_c = \frac{\sum_{x=x_0-w}^{x_0+w} \sum_{y=y_0-h}^{y_0+h} x \Phi(C(x, y) - Th)}{\sum_{x=x_0-w}^{x_0+w} \sum_{y=y_0-h}^{y_0+h} \Phi(C(x, y) - Th)}; \tag{2}$$

$$y_c = \frac{\sum_{x=x_0-w}^{x_0+w} \sum_{y=y_0-h}^{y_0+h} y \Phi(C(x, y) - Th)}{\sum_{x=x_0-w}^{x_0+w} \sum_{y=y_0-h}^{y_0+h} \Phi(C(x, y) - Th)}$$

where

$$\Phi(x) = \begin{cases} x, & \forall x \geq 0 \\ 0, & \forall x < 0 \end{cases},$$

summation comes over the template size  $(2w + 1, 2h + 1)$  around the peak position  $(x_0, y_0)$  of the covariation image  $C$ .

Another way of increasing the accuracy of object localization is an approximation of the covariation image  $C$  by a 2D parabolic function

$$C(x, y) = ax^2 + by^2 + dx + ey + f \tag{3}$$

in the vicinity of its maximum. The maximum of the parabola can be found with sub-pixel accuracy [16].

It is obvious that formula (1) provides not a correlation but a convolution of the image with the template. As a result, the brightest part of the non-uniformly illuminated image will give a global peak on matrix  $C$  even if there is poor geometrical similarity between the matched area and the template. In addition, formula (1) is applicable only to cases where the background in the image is either uniform or carefully removed. These conditions hold for the plastic beads used in the in vitro experiments of Gelles et al., but it is hard to satisfy them in in vivo microscopy. This drawback can be compensated by calculating normalized correlation coefficients [8, 13, 16, 47, 61]:

$$C(x, y) = \frac{\frac{1}{(2w+1) \cdot (2h+1)} \sum_{i=-w}^w \sum_{j=-h}^h I(x+i, y+j) \cdot T(i+w, j+h) - \langle I \rangle_{wh} \cdot \langle T \rangle}{\sqrt{D(I)_{wh} \cdot D(T)}} \quad (4)$$

where  $I(x, y)$  is the intensity of image in pixel  $(x, y)$ ,  $T(x, y)$  the intensity of template in pixel  $(x, y)$ ,  $w$  and  $h$  the dimensions of template image,  $\langle T \rangle$  the mean value of template intensity,  $\langle I \rangle_{wh}$  the mean value of image intensity in the area overlapping with the template,  $D(T)$  the variance of template intensity, and  $D(I)_{wh}$  is the variance of image intensity in the area overlapping with the template.

This method can be easily generalized for a multiple object search algorithm by searching in matrix  $C$  for multiple local maxima above some predefined threshold. If the image noise is known, then the noise variance of matrix  $C$  can be calculated. A reasonable threshold value can be estimated, for example, as  $4\sigma$ , where  $\sigma$  is one standard deviation of  $C$ . The value of  $2\sigma$  is generally too small and results in too many false positive signals, since the probability to overcome this threshold by chance is high, given the millions of pixels in a typical image. The correlation algorithm in the form of (3) can be applied to the image without background subtraction. The only essential limitation of this correlation method is the requirement of a fixed and known shape of the object being searched for. It is notable that, according to [36], the relative error of subpixel position detection of the correlation method (Particle Image Pattern Matching in the terminology of Huang et al.) is almost twofold smaller than the respective error of the covariation method (Cross-Correlation method in the terminology of Huang et al.).

A close relative of correlation/covariation algorithms is the method of Sum-Absolute Differences (SAD). In this method, the sum of absolute differences between the image and the template is calculated at all possible shifts of the template [8, 21, 79]:

$$\text{SAD}(x, y) = \sum_{i=-w}^w \sum_{j=-h}^h \|I(x+i, y+j) - T(i+w, j+h)\| \quad (5)$$

where  $I(x, y)$  is the intensity of image in pixel  $(x, y)$ ,  $T(x, y)$  the intensity of template in pixel  $(x, y)$ ,  $w$  and  $h$  the dimensions of template image.

The minimum of SAD corresponds to the best agreement between the template with the image. All of the above-mentioned concerns of the correlation method, regarding accuracy, parabolic interpolation, etc., are applicable to this method too. An additional drawback of this method, in comparison to the correlation method, is its sensitivity to intensity scaling of the image and the template. This can cause problems since the fluorescent marker can bleach during the course of acquisition. Therefore, if the labeling level (intensity) is point of concern for the process under investigation [54], this method is not applicable.

Another problem with correlation/covariation methods arise from the simple fact that the template image is a rectangular matrix and includes pixels which belong to the tracked object and the background pixels. If the template image is selected from the same experiment, for example, a region of interest from the first frame of the sequence, then the peak for the covariation matrix could be the result of correlation of

background pixels of the template with background pixels of the image under study. This problem is common for correlation, covariation and SAD methods. Cho and Yun have introduced “selective attention” for the pixels as a cure for the false peaks. A weight function is included in the summing to pay more “attention” to the pixels that presumably belong to the target. If the target has some specific moments of brightness (mean, variance, etc.), the probability for a given pixel to belong to the target can be used as a weight. The respective probability can be estimated from the template. Examples from the military field demonstrate that this is a robust tracking method in the case of imperfect matching of the template to the image under study [21].

These methods can be used for a search for multiple objects. The threshold is selected in such a way that peaks in the correlation/covariation image correspond to the objects of interest [25]. Then an accurate object position is calculated for every peak.

The simplest single object search procedure without a fixed and known shape of object is the centroid method [13, 16]. The direct implementation of this algorithm calculates the center of mass of an image:

$$\begin{aligned}x_c &= \frac{\sum_i \sum_j x_i \cdot I_{i,j}}{\sum_i \sum_j I_{i,j}}; \\y_c &= \frac{\sum_i \sum_j y_j \cdot I_{i,j}}{\sum_i \sum_j I_{i,j}}\end{aligned}\tag{6}$$

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

Summing is performed over the whole image. This method is applicable to images where only one object persists and the background is removed so that the integral of the background intensity over the whole frame is a small fraction of the integral intensity of the object of interest. If these conditions are violated, the centroid algorithm is prone to give the coordinates of the center of the image.

There is a straightforward generalization of this method for application to multi-object images with background. The image of interest is smoothed to remove high frequency noise, followed by binarisation by a threshold in such a way that all the objects of interest are preserved but spaces between them are zeroed. After this, the centroids are calculated separately for every connected set of non-zeroed pixels and the resulting values are considered as the centers of the objects of interest [3, 13, 31, 66, 68].

Crocker and Grier used some modifications to this method [25]. After an initial guess concerning the center of the object, the centroid is calculated by summing not only in the mask area but also inside the circle with a radius  $R$  over the non-masked image:

$$\begin{aligned}x_0^k &= \frac{\sum_{i,j} x_i \cdot I_{i,j}}{\sum_{i,j} I_{i,j}}; \\y_0^k &= \frac{\sum_{i,j} y_j \cdot I_{i,j}}{\sum_{i,j} I_{i,j}}\end{aligned}\tag{7}$$

where  $k$  is the index of iteration.

Summing is performed over pixels  $(i, j)$  which belong to the cycle  $(x_i - x_0^{k-1})^2 + (y_j - y_0^{k-1})^2 \leq R^2$  centered on the object position on iteration  $k - 1 (x_0^{k-1}, y_0^{k-1})$ .

If, after calculation, the position of the object shifts more than 0.5 pixels, the iteration (7) is repeated with the position of the newly found center [25, 27, 57].

Further modification of this algorithm, which provides the maximum accuracy of all centroid algorithms, is called the Gaussian mask [70]. Summing in the Gaussian mask algorithm is performed over the whole image but in convolution with a Gaussian kernel:

$$\begin{aligned} x_0^k &= \frac{\sum_{i,j} x_i \cdot I_{i,j} \cdot N_{i,j}^{k-1}}{\sum_{i,j} I_{i,j} \cdot N_{i,j}^{k-1}}; \\ y_0^k &= \frac{\sum_{i,j} y_j \cdot I_{i,j} \cdot N_{i,j}^{k-1}}{\sum_{i,j} I_{i,j} \cdot N_{i,j}^{k-1}} \end{aligned} \quad (8)$$

where  $k$  is the index of iteration and  $I_{i,j}$  is the intensity of pixel  $(i, j)$ ,

$$N_{i,j}^{k-1} = \int_{x_i}^{x_{i+1}} \int_{y_j}^{y_{j+1}} \exp \left( -\frac{(x - x_0^{k-1})^2 + (y - y_0^{k-1})^2}{2R^2} \right) dx dy \quad (9)$$

is an integral of the Gaussian kernel over pixel  $(i, j)$  with the kernel centered in the position  $(x_0^{k-1}, y_0^{k-1})$  of previous iteration  $k - 1$ .

$R$  is the width of the Gaussian.

As in the case of formula (7), the calculation is repeated iteratively, since the value  $N_{i,j}^k$  has to be recalculated after every adjustment of the object center position. This iteration is stopped when the correction to the object's position is below the accuracy level specified by the user.

Another class of sub-resolution object search procedures is the fitting of image intensities by an approximation of the PSF. The shape of a microscope PSF is defined as the Airy disk, but in most cases this is approximated by a Gaussian function [2, 12, 16, 42, 70, 76, 77, 80].

In the absence of noise, the Gaussian center position can be defined by the intensity of 5 pixels in the vicinity of the peak intensity [12, 36]:

$$\begin{aligned} x_c &= x_0 + \frac{1}{2} \cdot \frac{\lg(I(x_0 - 1, y_0)) - \lg(I(x_0 + 1, y_0))}{\lg(I(x_0 - 1, y_0)) + \lg(I(x_0 + 1, y_0)) - 2\lg(I(x_0, y_0))} \\ y_c &= y_0 + \frac{1}{2} \cdot \frac{\lg(I(x_0, y_0 - 1)) - \lg(I(x_0, y_0 + 1))}{\lg(I(x_0, y_0 - 1)) + \lg(I(x_0, y_0 + 1)) - 2\lg(I(x_0, y_0))} \end{aligned} \quad (10)$$

where  $(x_0, y_0)$  is the coordinates of intensity peak.



Unfortunately, this simple approach does not work in the case of noisy images. The robust alternative is a non-linear least square fitting, namely the minimization of:

$$S = \sum_{i,j} \left( I(i, j) - A \cdot \exp \left( -\frac{(x_i - x_0)^2 + (y_j - y_0)^2}{2R^2} \right) - F \right)^2 \tag{11}$$

where  $I(i, j)$  is the intensity of image,  $A$  the amplitude of Gaussian,  $(x_0, y_0)$  the position of the object,  $R$  the width of PSF (suppose to be known in advance), and  $F$  is the background intensity, summing over whole image.

The expansion of (11) to the 3D is straightforward and, beside the additional summation coefficient, introduces the separate scale parameter  $Z$ , to take in account the different width of PSF in axial direction [84]:

$$S = \sum_{i,j,k} \left( I(i, j, k) - A \cdot \exp \left( -\frac{(x_i - x_0)^2 + (y_j - y_0)^2}{2R^2} - \frac{(z_k - z_0)^2}{2Z^2} \right) - F \right)^2$$

The parameters  $(x_0, y_0, A, F)$  which minimize the squared difference between the Gaussian function and the image are taken as the features of the object. If the background was subtracted in advance, then  $F$  is considered zero and excluded from the fitting procedure. As a result of non-linearity, there are only iterative ways to find the minimum of  $S$ . This is computationally expensive, but gives the best accuracy [16,41,42,65,70,77]. A Gaussian fit approach does have a drawback in that it is applicable only to objects with a geometrical size smaller than the diffraction limit of the microscope and with which the intensity distribution can be approximated by one Gaussian.

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

The advantage of having non-coordinate parameters for fitting is an additional way in which to filter out false signals, since the intensity characteristics of the correct signal are either known in advance or can be found by clustering their moments of intensity distribution [25,57].

Different gradient edge detection and thresholding algorithms [29,39,81,82] are used to search for objects with sizes above the diffraction limit and of unknown shapes. They use a variety of ad hoc procedures to verify that selected regions satisfy the criteria of the object of interest. Both categories can produce a binary mask, which is later combined with the centroid algorithm. In this case, the function  $\Phi(x)$  is a mask (see formula (2)).

The threshold algorithms are based on the threshold value  $Th$ :

$$B(x, y) = \begin{cases} 1; & I(x, y) > Th \\ 0; & I(x, y) \leq Th \end{cases} \tag{12}$$

The algorithm constructs a binary image on the basis of the original gray-scale by formula (11). In the simple case of a homogeneous background and high signal-to-noise ratio, the threshold value can be manually selected by the user by examination of a single frame [3]. Different ad hoc approaches were invented to determine the threshold value from the image. For example, Crocker and Grier have used the 30th percentile of brightness of the entire image as the threshold [25], Ku et al. [39] have calculated the heuristic

$$Th = \max(I_{i,j}) \left( \frac{\text{std}(I_{i,j})}{\max(I_{i,j}) - \text{mean}(I_{i,j})} \right), \quad (13)$$

where std means standard deviation, and the calculation is done over all pixels of the image  $(i, j)$ .

A more solid approach based on probability theory is worth mentioning and uses maximum entropy as a criterion for threshold selection [43]. In this approach, the image is considered as a mixture of two sub-images—the sub-image of objects and sub-image of background. The intensities of both sub-images belong to Gaussian distributions with different means and variance. The threshold is selected iteratively, with recalculation of means and variance of distributions, in such a way that it discriminates the sub-images with maximum entropy. Brink did not make assumptions about the specific type of sub-image intensity distributions, but the maximum entropy formula includes the local correlation of pixels in the image (gradient distribution) [11]. For homogeneous organelles with the same size and intensities, e.g. secretory granules, the assumption of samples with a single intensity distribution is reasonable. However, for early endosomes, with high diversity of sizes and intensities, this is not so clear.

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 [81]. After closed boundaries are formed, another set of rules are used to eliminate false objects (i.e. the intensity distribution within the boundary has to satisfy some criteria, e.g. mean value and variance has to fit in a predetermined interval). This approach is quite complicated and has a less solid basis in comparison with intensity fitting by analytical functions, but its implementation could be computationally faster.

An object with a known shape, e.g. rod-like bacteria, can be presented as a sum of blurred segments with constant intensities [78]. In this case the blurring parameters are fixed, but the length, orientation, position and intensity of segment are now the parameters requiring fitting.

Similar approach to objects with arbitrary shape and size above the resolution limit are used by the authors of this work [54]. The objects are fitted by a sum of powered Lorenzians by minimizing:

$$\chi^2 = \sum_{i,j} \frac{1}{\sigma_{ij}^2} \left( I(i, j) - \sum_{k=1}^M \frac{A_k}{1 + \left( \left( \frac{(x_i - x_k) \cos \alpha_k - (y_j - y_k) \sin \alpha_k}{w_k} \right)^2 + \left( \frac{(x_i - x_k) \sin \alpha_k + (y_j - y_l) \cos \alpha_k}{h_k} \right)^2 \right)^{p_k}} \right)^2 \quad (14)$$

where  $(x_k, y_k, A_k, w_k, h_k, \alpha_k, p_k)$  the parameters of  $k$ th Lorentzian,  $M$  the number of Lorentzians, and  $\sigma_{i,j}$  is the standard deviation of noise in the pixel  $(i, j)$ .

Summing is performed over a large enough vicinity of the local maximum.

Any intensity distribution can be presented by a sum of hat-like functions. The Lorentzian has some advantages because the calculation is less expensive than the Gaussian. At the same time, the difference between the Airy function and the squared Lorentzian 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 objects whose size ranges from hundreds of nanometers to a few micrometers. The elongation of the base function along an arbitrary axis with angle  $\alpha_k$  to the axis  $X$  decreases the number of base functions required for accurate object deconvolution. The most probable number of base functions,  $M$ , to fit a given object were selected by calculating the probability  $P(M|I)$ —probability to have  $M$  functions given image  $I$  [63]:

$$P(M|I) = \frac{\text{const} \cdot M!}{V^M} \cdot \frac{(4\pi)^M \exp\left(-\frac{\chi_{\min}^2}{2}\right)}{\sqrt{\text{Det}(\nabla\nabla\chi^2)}} \tag{15}$$

where  $M$  is the number base functions,  $\chi_{\min}^2$  the minimum value of squared residuals (13) at given  $M$ , and  $V$  is the maximum volume of parameter space per one base function.

The denominator grows fast with the number of base functions and constitutes an Occam’s razor counterbalance to the trivial fact that increasing the number of base functions will decrease the residual  $\chi_{\min}^2$ . The  $P(M|I)$  behaves better than the Fisher criterion. Last is computationally simpler and can be used in case of known in advance number of possibly overlapped objects [84,85]

The summary of object searching/localization algorithms is presented in the Table 1.

### 3 Tracking algorithms

Tracking is an assignment procedure. A set of sequential measurements which belong to the same physical entity have to be assigned to one track. In other words, the track is a sequence of objects in sequential frames which belong to the same physical entity.

Different scoring systems can be used to determine how good or bad particular assignment is. Scoring systems include the simple Euclidian distance between objects on sequential frames [12]:

$$\text{Score}(k) = \sum_{i=1}^M \alpha_i \left\| p_i^k - p_i^{k-1} \right\|^2 \tag{16}$$

where  $\alpha_i$  is a scaling factor, which compensates for the different units of different parameters and gives them their respective weight,  $p_i^k$  the  $i$ th parameter value at frame  $k$ , and  $M$  is the total number of parameters.

**Table 1** Object detection algorithms

Algorithm purpose	Algorithm description	References
Search for objects with known shape (template matching)	Covariation with template	[8, 16, 30]
	Correlation with template	[13, 16, 61, 47]
	SAD (sum absolute differences)	[8, 21, 79]
	Approximate object intensity by Gaussian	[2, 12, 16, 36, 42, 70, 76, 77, 80, 84, 85]
Search for single object with unknown shape	Centroid	[16]
	Gradient edge detection and thresholding	[25, 29, 39, 43, 81, 82]
Search for multiple objects with unknown shape	Centroid with threshold-based mask	[3, 13, 31, 66, 68]
	Approximate object intensity by sum of powered Lorentzian	[34, 54]
Localization of object with sub-resolution accuracy	Waited centroid	[16, 30]
	Recursive waited centroid	[25, 37, 57]
	Gaussian mask	[70]
	Approximate object intensity by Gaussian	[2, 12, 16, 36, 70, 76, 77, 80]
	Approximate object intensity by sum of powered Lorentzian	[34, 54]

As well as more complex heuristic scores.

Sethi and Jain [19, 50, 57, 59] offered a score function which maps the score to the interval  $[0, 1]$ , where zero corresponds to a perfect fit.

$$\text{Score}(k) = \sum_{i=1}^M w_i \cdot \left( 1 - 2 \frac{\sqrt{p_i^k \cdot p_i^{k-1}}}{p_i^k + p_i^{k-1}} \right) \quad (17)$$

where  $w_i$  is a  $i$ th parameter weight,  $p_i^k$  the  $i$ th parameter value at frame  $k$ , and  $M$  is the total number of parameters.

The score can be a function of object and track state parameters. This means that it can include object parameters like position, intensity, size, shape, etc., as well as track parameters, e.g. velocity, acceleration, trajectory smoothness, etc., or even can be a function of the track's entire history.

The score of the position can be calculated on the basis of only the last object position without calculating track-dependent speeds [56]:

$$\text{Position Score}(k) = \frac{|x_{k-1}x_p - x_{k-1}x_k|}{\sqrt{W^2 + H^2}} \quad (18)$$

where  $x_{k-1}x_p$  is the vector from position at the frame  $k - 1$  to the predicted position at frame  $k$ ,  $x_{k-1}x_k$  the vector from the position at the frame  $k - 1$  to the candidate position at frame  $k$ , and  $W, H$  is the width and height of image, respectively.

The score of trajectory non-smoothness is calculated by:

$$\text{Smoothness Score}(k) = 1 - \frac{\vec{v}_k \cdot \vec{v}_{k-1}}{|\vec{v}_k| \cdot |\vec{v}_{k-1}|} \tag{19}$$

where  $v_k$  and  $v_{k-1}$  is the movement of the object between frames  $k - 2/k - 1$  and  $k - 1/k$ .

Other heuristic scores could be found in [18].

The probability of different assignments can also be considered as a score. Anderson et al. [2] use probability as a score

$$P \propto \exp\left(-\left(\frac{\Delta r}{R_d}\right)^2\right) \cdot \exp\left(-\left(\frac{\Delta I}{I_n}\right)\right) \tag{20}$$

where  $\Delta r$  is the shift of the object between two sequential frames,  $R_d$  the characteristic diffusive radius,  $\Delta I$  the change of the object’s intensity, and  $I_n$  is the characteristic intensity.

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

The probability measure is dependent on the particular value distribution. The different physics of each parameter measurement generates a different error distribution. In the case of jointly normally distributed errors the probability score is [5]:

$$P(X_k|X_{k-1}) = \frac{1}{(2\pi)^{\frac{M}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(X_k - Y_k(X_{k-1}))^T \cdot \Sigma^{-1} \cdot (X_k - Y_k(X_{k-1}))\right) \tag{21}$$

where  $X_k$  is the vector of parameters at frame  $k$ ,  $Y_k(X_k)$  the vector of predicted parameters at frame  $k$ , given  $X_{k-1}$ ,  $\Sigma$  the parameter covariation matrix, and  $M$  is the dimension of parameter vector.

This probability score can be reduced to the Mahalanobis distance [23] by taking  $\text{Score}(k) = \sqrt{-\ln(P(X_k|X_{k-1}))} \propto \sqrt{\frac{1}{2}(X_k - Y_k(X_{k-1}))^T \cdot \Sigma^{-1} \cdot (X_k - Y_k(X_{k-1}))}$ .

Besides different scores, there are two main approaches for object tracking. The first approach is adequate for the situation where the measurement error is relatively small and the measured object parameters are a good approximation of the object state in the context of the problem under study. For example, when one follows the individual endosomes in vivo [54,57,75] or in vitro [4,46], the error of endosome position definition is nonessential in comparison with endosome size. Another extreme case is when the accuracy one needs is much higher than the accuracy of single measurement. This case arises in air/space/ocean surveillance by radar/sonar equipment [5,28,62] and in nanometer-accuracy light microscopy [30,34,41,60,64]. In last case, one is not so interested in connecting the measurements in a chain or track, it even can be a situation where only one measurement exists at every time point (“no-choice” tracking), but the accuracy of the measurement is increased by averaging the information from sequential frames. The information from the track history could increase the resolution

against that of the single frame. These kinds of tracking algorithms are called “filters”. The mixed case uses both—the accurate prediction of the object state on the next frame and object-to-track assignment on the basis of a more accurate score [10,24].

### 3.1 Tracking by measurement assignment

The simplest tracking-by-assignment algorithm is a nearest-neighbor or “no-choice-tracking” algorithm. This assigns a track to the only object that exists in frame  $k$  in the predefined vicinity of the track end in frame  $(k - 1)$ . The predefined vicinity is called the “validation gate,” where the name comes from automated radar tracking systems. If there is no object in the validation gate or there are multiple objects, the track is broken. This algorithm was used for intracellular fluorescent object tracking, for instance, in the works of Goulian and Simon [32]. It is clear that this approach performs well only in cases of slow-moving, sparse objects without false object signals. There are some modifications of this algorithm. The first modification uses only information about the object’s position in the previous frame and the validation gate is centered on this position. The second modification also takes into account the speed and center of the validation gate in the predicted position in frame  $k$ . The third modification uses an expanding gate (“box”), the size of which grows until it either finds an object or reaches the predefined limit.

The straightforward generalization of the “no-choice-tracking” algorithm for multi-target tracking case is a Greedy algorithm [2,31,74]. If the validation gate includes many objects the possible assignments are sorted according to a score and the best assignment is made first. If the subsequent assignment contradicts the previous decisions, then the second best is chosen, etc. In the end all possible assignments are performed in a greedy manner. The advantage of these greedy algorithms is the ability to handle situations of objects disappearing along with temporary occlusions. An upper limit on the possible score can be introduced by adding a dummy object with a fixed maximum score. If there is no better choice, the dummy object is assigned to the track and this causes a track to break. A little more handling is required in order to keep the dummy-marked track for a predefined number of frames and to handle possible occlusions. The disadvantage of greedy algorithms is its tendency to fall into the first local minimum of the search space. If the object density is low and movement is either slow or well organized, such that the score difference between possible candidates is large, then a greedy algorithm is a good choice. In the opposite case of dense, fast-moving objects, the number of errors which are produced by a greedy algorithm becomes unreasonably large.

The next improvement is a greedy exchange algorithm [56,59,72,73], which includes the exchange of assignments in an attempt to improve the total score of the tracks. This iterative process runs as follows: for every track, all possible replacements are determined. If this pair-wise replacement improves the total score, the new assignment is performed and the process is repeated until no appropriate substitutions are found. If an exchange that improves the total score is found then the exchange search is repeated. This procedure is better than a greedy algorithm, but it nevertheless does not guarantee the detection of the global optimum. Further improvements can

be achieved by applying dynamic programming for track assignment [37,54,55,71]. 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.

If there are a fixed number of objects, i.e. the restricted case in which no new objects appears and no objects disappear during the measurements, and there are no false signals and no non-detected objects, the global optimum can be found in polynomial time. Given a reasonable scoring for object assignment, 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 a solution [40]. The Hungarian algorithm works on a per-frame basis and provides a global optimum score assignment of objects to tracks. The initial track seeds, in this case, are the objects found on the first frame of the sequence.

Unfortunately, the assumption that a fixed number of objects are found in every given frame is impractical. In microscopy, new objects appear either by the genesis of compartments, vesicles, proteins, fission vesicles, etc, or just by their movement into the field of view. At the same time, existing objects can disappear by moving out of focus, changing identity, or fusing with other objects. In addition to the change in the number of real objects, false object recognition is a major problem in the field. The same scenario is applicable to radar/sonar tracking systems. The targets can move into the surveillance region or move out of it. In addition, false objects (clutter) contaminate virtually every measurement. These could be false targets in radar/sonar measurements or spikes of background noise in fluorescence microscopy. Starting from the 1960s, probabilistic approaches of proper tracking (measurement-to-track assignment) were developed to work in these complicated conditions.

### 3.2 Predictive tracking

Tracking procedures were originally developed like filters. The main goal was not the assignment of a sequence of (objects) measurements to a single track, but to find the “real” track on the basis of noisy measurements. The first tracking systems were developed in the military fields for the aiming of weapons. The signal from a radar station identifies a target’s location with error, and at the same time a moving object has to be targeted with higher accuracy than radar could achieve. The filter algorithm provides the best estimation of the “real” (unknown) target position. Since the first computers that controlled weapons were analog, the formulation of the algorithm was oriented to an analog hardware implementation [38,53]. The simplest case corresponds to single target tracking. The target is characterized by a state, which includes its kinematical parameters – position, velocity and acceleration. The predictive filter can be built on the basis of a motion model, which allows for the prediction of motion based on previous data.

The state for which the discrete time point  $k$  can be calculated on the basis of the previous state  $k - 1$ .

$$X_k = AX_{k-1} + Bu_k + C\eta_{k-1} \quad (22)$$

where  $X_k = \begin{pmatrix} \vec{r}_k \\ \vec{v}_k \\ \vec{a}_k \end{pmatrix}$  - target state, which is characterized by position  $\vec{r}_k$ , velocity  $\vec{v}_k$  and acceleration  $\vec{a}_k$  at time point  $k$ .

$A$  is a constant transition matrix, which in case of constant acceleration is  $A = \begin{bmatrix} 1 & \delta\tau & \frac{1}{2}\delta\tau^2 \\ 0 & 1 & \delta\tau \\ 0 & 0 & 1 \end{bmatrix}$ ,  $\delta\tau$  the time step between sequential measurements,  $u_k$  the control parameters, known in advance,  $B$  the transition matrix for control parameters, and  $\eta_k$  a motion model noise (i.e., the atmosphere heterogeneity and turbulence).

In this algorithm, the noise is assumed to be white noise with a mean of zero; this means that there is no correlation between the noise values at the time point  $k$  and time point  $k + 1$ . If  $\eta_k$  has a non-zero mean by nature, it can be included in the control parameters  $u_k$ . The covariation matrix of the noise is known (measured in advance) and equals  $R_k$ .

The measured state

$$Y_k = DX_k + \varepsilon_k \quad (23)$$

is linearly dependent on the real state and has additive noise  $\varepsilon_k$ . The measurement noise is white noise with a mean of zero and known covariation matrix  $Q_k$ .

The simplest case corresponds to the measurement matrix  $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ ; in real radar/sonar systems, the coordinates  $Y$  are non-Cartesian and the measurement matrix is more complicated.

The use of the linear model of movement and measurement is reasonable for many practical applications. If we have a set of measurements and a model, we can determine the model parameters by fitting the model to the measurements. A fitting objective function could be maximum likelihood, which in the case of Gaussian noise reduces to the least square of deviations of predicted and measured states. Unfortunately, the direct fit of all measurements to this model requires increasing computation time and memory over time. But in the case of Gaussian noise, there is a recursive method for updating the maximum likelihood parameter estimation [38] with each new measurement:

$$\tilde{X}_{k+1|k} = A \cdot \tilde{X}_{k|k-1} + B \cdot u_k + G_k \cdot (Y_k - D \cdot \tilde{X}_{k|k-1}) \quad (24)$$

$$G_k = A \cdot \Sigma_k \cdot D^T \cdot (D \cdot \Sigma_k \cdot D^T + Q_k)^{-1} \quad (25)$$

$$\Sigma_{k+1} = A \cdot \left( \Sigma_k - \Sigma_k \cdot D^T \cdot (D \cdot \Sigma_k \cdot D^T + Q_k)^{-1} \cdot D \cdot \Sigma_k \right) \cdot A^T + C \cdot R_k \cdot C^T \quad (26)$$

where  $\tilde{X}_{k|k-1}$  is the predicted target state at time  $k$  on the basis of  $k - 1$  measurements,  $G_k$  the gain matrix,  $G_k \cdot (Y_k - D \cdot \tilde{X}_{k|k-1})$  gives correction to the model on



the basis of “innovation”, the discrepancy between predicted and measured values at time  $k$ .

$\Sigma_k = \text{cov}(X_k - \tilde{X}_{k|k-1}, X_k - \tilde{X}_{k|k-1})$ - covariation of the state uncertainty,  $\Sigma_0$  encodes the uncertainty of the initial state of the target.

The beauty of this algorithm is constant memory usage and constant calculation time which are independent of the number of measurements. The drawback of this algorithm is that it cannot handle clutter. The obvious solution in a cluttered environment is taking the measurement which is closest to the predicted state. If the density of false signals in the validation gate is high and a false signal could be chosen at multiple times in the sequence, the algorithm will end up far away from the real target. The extension of the Kalman filter to the case of clutter is the Probabilistic Data Association (PDA) algorithm by Bar-Shalom [5].

In the PDA algorithm the single value  $Y_k$  in formula (23) is replaced by

$$y' = \sum_{i=1}^{N_k} p(y_{k,i} | \chi_{k,i}, Y_{k-1}) \cdot y_{k,i} \tag{27}$$

where  $N_k$  is the number of measurements in the validation gate at time  $k$ ,  $y_{k,i}$  the  $i$ th measurement at time  $k$ ,  $Y_k$  the set of all validated measurements at time  $k$ ,  $p(y_{k,i} | \chi_{k,i}, Y_{k-1})$  the probability that measurement  $Y_{k,i}$  comes from the target, and  $\chi_{k,i}$  is the event, measurement  $y_{k,i}$  comes from target.

In short, instead of the nearest-neighbor, the weighted sum of all measurements in the validation gate is taken. The probability that the signal is a signal from the target is used as a weight. The probability is dependent on a model for the clutter in the measurement. In the case of indistinguishable point-like objects and a Poisson distribution of false signals with fixed density and normal distribution of measurement errors, the probability that the  $i$ th signal is correct is defined by formula [5]:

$$p(y_{k,i} | \chi_{k,i}, Y_{k-1}) = \frac{f(y_{k,i} | Y_{k-1})}{b_k + \sum_{i=1}^{m_k} f(y_{k,i} | Y_{k-1})} \tag{28}$$

where  $m_k$  is the number of signals in the validation gate at time  $k$ ,  $V$  the volume of the validation gate,  $b_k = \frac{m_k}{V} \frac{\alpha_1 + \alpha_2 - \alpha_1 \alpha_2}{(1 - \alpha_1)(1 - \alpha_2)}$ ,  $\alpha_1$  the probability that a correct measurement is not in the validation gate,  $\alpha_2$  the probability that a correct measurement is not detected,

$$f(y_{k,i} | Y_k) = \frac{1}{1 - \alpha_1} N(D\tilde{X}_{k|k-1}, \Sigma_k),$$

$N(D\tilde{X}_{k|k-1}, \Sigma_k)$ - normal distribution with mean  $D\tilde{X}_{k|k-1}$  and covariation matrix  $\Sigma_k$ .

The probability that there is no correct signal at time  $k$  is:

$$p(y_{k,0} | \chi_{k,0}, Y_{k-1}) = \frac{b_k}{b_k + \sum_{i=1}^{m_k} f(y_{k,i} | Y_{k-1})} \tag{29}$$

It was shown that the Kalman filter failed at a clutter density of 0.5–0.75 false signals in the validation gate. The PDA algorithm is much more robust and continues to reliably track at a clutter density of 2–3 false signals per validation gate.

The major drawback of PDA is that it works only for a single target. If there are many targets in the viewfield, but their respective validation gates are not overlapping, then the problem is equal to the set of independent PDA trackers. If the validation gates are overlapping, the Joint Probabilistic Data Association (JPDA) is a way to handle this problem [28]. The JPDA uses a common model PDA, but differs in the manner with which the association probabilities are computed. The JPDA algorithm is the root of many derivative algorithms which utilize different methods for calculating probabilities in order to handle specific circumstances in each case [10, 17, 26, 35].

### 3.3 Multi-variant tracking

Clutter can complicate the tracking of even a single target. The branch of algorithms stemming from the “track splitting” algorithm was developed as an alternative to the PDA/JPDA algorithms [24, 48, 51, 52, 67]. Generally a track splitting algorithm creates a tree of possible track continuations when multiple assignments are feasible. The decision of which track to keep is postponed to a later stage where the difference in the quality of the tracks becomes obvious. This track splitting approach has an exponentially growing number of combinations over time, so it must incorporate some pruning strategy to keep the size of the graph of the tracks reasonable. The Multiple Hypothesis Tracking (MHT) algorithm [24, 51, 52, 67] is a particular branch of track splitting algorithm. The hypothesis in this context means one possible instantiation of the track. A set of hypotheses about all measurements, which were done before the current time point, is used as parents to generate a set of next level hypotheses. Each hypothesis has a full set of labels for every signal: it can come from a known object, a false positive, or a new object appearing in the viewfield. Some probability is assigned for every hypothesis. The probabilities of next-level hypotheses conditioned on the parent hypotheses can be calculated by the Bayesian rule with parent hypothesis probabilities as a prior. This generates a recursive method for the calculation of probabilities. Since the set of hypothesis is exhaustive, the pruning strategy can be based on a threshold cut-off, i.e., keeping the  $n$ -most-probable hypothesis, etc. The conditional probabilities of a hypothesis are formulated in terms of the discrepancy between the predictions of the parent hypothesis and the hypothesis about the last measurement set. This again returns to the predefined model of object behavior which was incorporated into the prediction. In the original work of Reid, the probabilities were calculated on the basis of a linear object motion model of the form (22) and (23). The same kind of model with some variation was used in most aircraft tracking programs.

### 3.4 Tracking intracellular objects

After objects have been detected in each frame of a sequence, tracking is reduced to connecting the information about the objects into a track. In cases where only a single object exists in each 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 in different frames to the same track becomes a non-trivial problem. The main objective is to track objects in the absence of established models of motion or kinematical restrictions of shape changes for microscopic objects. These restrictions are essential for surveillance systems [14, 15, 83], but have almost no practical application in intracellular object analysis. As in aerospace/ship tracking and flow velocimetry, in intracellular object tracking the shape of the object is not predictable, either because of low signal-to-noise ratio along with the limitations of resolution or because it varies too quickly between two sequential measurements. Due to this limitation, point tracking algorithms were used in those fields. At the same time kinematic models like (22) and (23) are not applicable to intracellular microscopy with low-to-middle frame rates. As a result the motion model-based approaches like (PDA/JPPA), despite their substantial development over the last 30 years, have little application to intracellular objects tracking. With MHT, although it is a multi-target tracking algorithm, the number of simultaneously-tracked objects is very limited. This is acceptable in the radar/sonar systems, where the number of targets is generally in the range from 1 to 20. In intracellular microscopy and velocimetry, the number of objects is in the range of 100–10,000. Therefore, less sophisticated algorithms are generally used in this field.

The most popular intracellular tracking algorithm is still the greedy algorithm modified in different ways [29, 30, 41, 57, 77]. But in the case of a crowded environment and a high clutter density, it is impossible to choose the proper assignment on a per-frame basis. Multi-frame analysis can generate a reasonable increase in accuracy. The validation gate, which can be selected on the basis of the maximum possible object velocity, decreases the number of probable track continuations to a manageable value. In this case, a dynamic programming procedure becomes a feasible method of finding the global assignment optimum [54, 55]. Non-deterministic algorithms, i.e. simulating annealing, give a possibility to search the global optimum in case when deterministic algorithms becomes non-efficient [86]. The non-deterministic algorithms do not guarantee the reaching of the global optimum in any given case, but in average they perform well and allow naturally include many biological events such as a new object generation, fusion, fission and decay.

The microtubule-dependent movement of intracellular vesicles can be processive enough to produce the smooth curve-like maximum-intensity projection on a time dimension. In those cases trajectories of vesicles can be found before tracking either manually or by fitting the pieces of straight line to the image [87]. After trajectory was found the 1D+t section of 3D+t image (kymogram) is preformed. The restriction to the known trajectory (kymogram) dramatically decreases the number of possible assignments. As result, the simple algorithms, i.e. “no-choice-tracking”, perform well.

The summary of tracking algorithms is presented in Table 2.

## 4 Conclusion

Much work in cell biology and biophysics has been done in the field of microscopic fluorescent object tracking. From this work, a wealth of information concerning the

**Table 2** Object tracking algorithms

Algorithm purpose	Algorithm description	References
Single object tracking in absence of false signals (clutter)	Tracking by assignment: no-choice-assignment	[24]
Single object tracking in presence of noise (clutter)	Kalman filter (kinematic model has to be known)	[38,53]
	PDA (kinematic model has to be known)	[5,6]
Fixed number of objects tracking in absence of false signals	Hungarian algorithm (method finds the global optimum of assignment)	[40]
Multiple object tracking in presence of false and missing signals, object appearing and disappearing	JPDA (applicable to the small number of objects)	[6,7,10,17,26,28,35]
	MHT (applicable to the small number of objects)	[24,51,52,67]
	Greedy assignment (method fails to find the global optimum in case of crowded environment)	[2,19,29,31,81,74]
	Greedy assignment + exchange (method partially overcome the problem of Greedy assignment in crowded environment)	[56,59,72,73]
	Simulated annealing (method finds the global optimum in average)	[86]
	Dynamic programming (method finds the global optimum, including many frames in simultaneous consideration increases it complexity)	[37,54,55,71]

underlying mechanisms of the molecular machinery of the cell was extracted. The physical properties of molecular motors of the kinesin, myosin and dynein superfamilies were discovered. Despite this success, the average level of sophistication in the algorithms used is still lower than that of other fields. This is partially explained by the more complex input data, where the signal-to-noise ratio is low and the complexity of system is high. At the same time, kinematical models, which are used to provide higher tracking quality, are the subject of continuing research. This precludes the researcher from the use of a model of motion in the tracking algorithm. High clutter density and essentially non-homogenous backgrounds in live-cell fluorescent microscopy have effects, such that, e.g., the number of endosomes which human beings can see in a movie is much higher than the number of endosomes one can see in still images. The current state-of-the-art object searching algorithms, which are based on function fitting procedures, can find virtually all the objects visible on one frame. However, this is considerably fewer than can be seen in a movie. In the case of many fast objects in a crowded surrounding, the human eye is still better than the best software available. Yet lowering the object search threshold can result in an explosion of clutter density that almost prohibits tracking. The track-before-detecting approach [37,85], where the object search is performed in the close vicinity of the predicted track position, can be performed with a much lower threshold, and this looks

promising. Generally, development of probabilistic predictive algorithms, like MHT and JPDA, which are appropriate to the movement of biological objects, is fruitful direction of future research too. Another possible direction is an iterative approach: the initial tracking is performed by greedy algorithm, than the model of object movement is generated and model-based probabilistic approach is used for tracking. On base of second tracking model is refined and process is repeated. The demand to the more reliable tracking of low intensity objects and new development in the microscopy techniques open the wide field for the future research in the intracellular object tracking filed.

**Acknowledgments** I am grateful to Dr. Gary Brouhard and Dr. Jochen Rink for the help on editing of the manuscript.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. Adrian, R.J.: Particle image tracking for experimental fluid mechanics. *Annu. Rev. Fluid Mech.* **23**, 261–304 (1991)
2. Anderson, C.M., Georgiou, G.N., Morrison, I.E.G., Stevenson, G.V.W., Cherry, R.J.: 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 (1992)
3. Apgar, J., Tseng, Y., Fedorov, E., Herwig, M.B., Almo, S.C., Wirtz, D.: Multiple-particle tracking measurements of heterogeneities in solutions of actin filaments and actin bundles. *Biophys. J.* **79**, 1095–1106 (2000)
4. Bananis, E., Nath, S., Gordon, K., Satir, P., Stockert, R.J., Murray, J.W., Wolkoff, A.W.: Microtubule-dependent movement of late endocytic vesicles in vitro: requirements for dynein and kinesin. *Mol. Biol. Cell* **15**, 3688–3697 (2004)
5. Bar-Shalom, Y., Tse, E.: Tracking in a cluttered environment with probabilistic data association. *Automatica* **11**, 451–460 (1975)
6. Bar-Shalom, Ya., Fortmann, T.E.: Tracking and Data Association. Academic Press Inc., London (1988)
7. Bar-Shalom, Ya., Li, X.R., Kirubarajan, T.: Estimation with Applications to Tracking and Navigation. Wiley, New York (2001)
8. Barnea, D.I., Silverman, H.F.: A class of algorithms for fast digital image registration. *IEEE Trans. Comput.* **21**, 179–186 (1972)
9. Barniv, Y.: Dynamic programming solution for detecting dim moving targets. *IEEE Trans. Aerosp. Electron. Syst.* **21**, 144–156 (1985)
10. Blanding, W.R., Willett, P.K., Bar-Shalom, Y.: Offline and real-time methods for ML-PDA track validation. *IEEE Trans. Signal Processing* **55**(5), 1994–2006 (2007)
11. Brink, A.D.: Using spatial information as an aid to maximum entropy image threshold selection. *Pattern Recognit. Lett.* **17**, 29–36 (1996)
12. Bonneau, S., Dahan, M., Cohen, L.D.: Tracking single quantum dots in live cell with minimal path. *Proc. IEEE Comp. Soc. Conf. Comp. Vis. Pattern Recognit.* 141–148 (2005)
13. Carter, B.C., Shubeita, G.T., Gross, S.P.: Tracking single particles: a user-friendly quantitative evaluation. *Bys. Biol.* **2**, 60–72 (2005)
14. Cerveri, P., Pedotti, A., Ferrigno, G.: Robust recovery of human motion from video using Kalman filters and virtual humans. *Hum. Mov. Sci.* **22**(3), 377–404 (2003)
15. Cham, T.J., Rehg, J.M.: A multiple hypothesis approach to figure tracking. *Comput. Vis. Pattern Recognit.* 239–245 (1999)
16. Cheezum, M.K., Walker, W.F., Guilford, W.H.: Quantitative comparison of algorithms for tracking single fluorescent particles. *Biophys. J.* **81**, 2378–2388 (2001)

17. Chen, B., Tugnait, J.K.: Tracking of multiple maneuvering targets in clutter using IMM/JPDA filtering and fixed-lag smoothing. *Automatica* **37**, 239–249 (2001)
18. Chetverikov, D., Verestoy, J.: Comparative performance evaluation of four feature point tracking techniques. In: *Proceedings of the Workshop of the Austrian Pattern Recognition Group*, pp. 225–263 (1998)
19. Chetverikov, D., Verestoy, J.: Feature point tracking for incomplete trajectories. *Computing* **62**(4), 321–338 (1999)
20. Chetverikov, D.: Particle image velocimetry by feature tracking. *Comput. Anal. Images Patterns* 21–24, 325–332 (2001)
21. Cho, J.S., Yun, B.J.: Selective-attention correlation measure for precision video tracking. *IEICE Trans. Inf. Syst.* E88-D(5), 1041–1049 (2005)
22. Collins, R.N.: Rab and ARF GTPase regulation of exocytosis (Review). *Mol. Membr. Biol.* **20**, 105–115 (2003)
23. Cox, I.I.: A review of statistical data association techniques for motion correspondence. *Int. J. Comput. Vis.* **10**(1), 53–66 (1993)
24. Cox, I.I., Hingorani, S.L.: 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 (1996)
25. Crocker, J.C., Grier, D.G.: Methods of digital video microscopy for colloidal studies. *J. Colloid Interface Sci.* **179**, 298–310 (1996)
26. Ding, Z., Hong, L.: Bias phenomenon and compensation for PDA/JPDA algorithms. *Math. Comput. Model.* **27**(12), 1–16 (1998)
27. Ewers, H., Smith, A.E., Sbalzarini, I.F., Lilie, H., Koumoutsakos, P., Helenius, A.: Single particle tracking of murine polyoma virus-like particles on live cell and artificial membranes. *PNAS* **102**, 15110–15115 (2005)
28. Fortmann, T.E., Bar-Shalom, Ya., Scheffe, M.: Sonar tracking of multiple targets using joint probabilistic data association. *IEEE J. Oceanic Eng.* 8(3), 173–184 (1983)
29. Gasman, S., Kalaidzidis, Y.L., Zerial, M.: RhoD regulates endosome dynamics through diaphanous-related formin and Src tyrosine kinase. *Nat. Cell Biol.* **5**(3), 195–204 (2003)
30. Gelles, J., Schnapp, B.J., Sheetz, M.P.: Tracking kinesin-driven movements with nanometer scale precision. *Nature* **331**, 450–453 (1988)
31. Ghosh, R.N., Webb, W.W.: Automated detection and tracking of individual and clustered cell surface low density lipoprotein receptor molecules. *Biophys. J.* **66**, 1301–1318 (1994)
32. Goulian, M., Simon, S.M.: Tracking single protein molecules within cells. *Biophys. J.* **79**, 2188–2198 (2000)
33. Hashiro, M., Kawase, T., Sasase, I.: Maneuver target tracking with an acceleration estimator using target past positions. *Electron. Commun. Jpn., Part I* 85(12), 29–37 (2002)
34. Helenius, J., Brouhard, G., Kalaidzidis, Y., Diez, S., Howard, J.: The depolymerizing kinesin MCAK uses lattice diffusion to rapidly target microtubule ends. *Nature* **441**(4), 115–119 (2006)
35. Hong, L., Cui, N., Cong, S., Wicker, D.: An interacting multipattern data association (IMPDA) tracking algorithm. *Signal Processing* **71**, 55–77 (1998)
36. Huang, H., Dabiri, D., Gharib, M.: On errors of digital particle image velocimetry. *Meas. Sci. Technol.* **8**, 1427–1440 (1997)
37. Johnston, L.A., Krishnamurthy, V.: Performance analysis of a dynamic programming track before detect algorithm. *IEEE Trans. Aerosp. Electron. Syst.* **38**(1), 228–242 (2002)
38. Kalman, R.E.: A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Eng.* **82**(ser.D), 35–45 (1960)
39. 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.: An automated tracking system to measure the dynamic properties of vesicles in living cells. *Microsc. Res. Tech.* **70**, 119–134 (2007)
40. Kuhn, H.W.: The Hungarian method for the assignment problem. *Naval Res. Logistics Q.* **2**, 83–97 (1955)
41. Kural, C., Kim, H., Syed, S., Goshima, G., Gelfand, V.I., Selvin, P.R.: Kinesin and dynein move a peroxisome in vivo: a tug-of-war or coordinated movement? *Science* **308**, 1469–1472 (2005)
42. Kural, C., Serpinskaya, A.S., Chou, Y.H., Goldman, R.D., Gelfand, V.I., Selvin, P.R.: Tracking melanosomes inside a cell to study molecular motors and their interaction. *PNAS* **104**(13), 5378–5382 (2007)
43. Leung, C.K., Lam, F.K.: Performance analysis for a class of iterative image thresholding algorithms. *Pattern Recognit.* **29**, 1523–1530 (1996)

44. Logothetis, A., Krishnamurthy, V., Holst, J.: A Bayesian EM algorithm for optimal tracking of a maneuvering target in clutter. *Signal Processing* **82**, 473–490 (2002)
45. Maxfield, F., McGraw, T.: Endocytic recycling. *Nat. Rev.* **5**, 121–132 (2004)
46. Nielsen, E., Severin, F., Backer, J.M., Hyman, A.A., Zerial, M.: Rab5 regulates motility of early endosomes on microtubules. *Nat. Cell Biol.* **6**, 376–82 (1999)
47. Ngoc, S.N., Briquet-Laugier, F., Boulin C., Olivo, J.-C.: Adaptive detection for tracking moving biological objects in video microscopy sequences. In: Proceedings of the International Conference on Image Processing, 26–29 Oct 1997. Washington, DC, USA, vol. 3, pp. 484–487 (1997)
48. Noyes, S.P., Atherton, D.P.: Control of false track rate using multiple hypothesis confirmation, target tracking 2004: Algorithms and Applications, IEE, pp. 115–120 (2004)
49. Qian, H., Sheetz, M.P., Elson, E.: Single-particle tracking: analysis of diffusion and flow in two-dimensional systems. *Biophys. J.* **60**, 910–921 (1991)
50. Rangarajan, K., Shah, M.: Establishing motion correspondence. *CVGIP Image Understanding* **54**(1), 56–73 (1991)
51. Poore, A.B.: Multidimensional assignments and multitarget tracking. Partitioning Data Sets DIMACS Workshop, pp. 169–196 (1995)
52. Reid, D.D.: An algorithm for tracking multiple targets
53. Rhodes, I.B.: A tutorial introduction to estimation and filtering, *IEEE Trans. Automat. Contr.* **16**(6), 688–706 (1971)
54. Rink, J., Ghigo, E., Kalaidzidis, Y., Zerial, M.: Rab conversion as a mechanism of progression from early to late endosomes. *Cell* **122**, 735–749 (2005)
55. Sage, D., Hediger, F., Gasser, S.M., Unser, M.: Automatic tracking of particles in dynamic fluorescence microscopy. *Proc. Int. Symp. Image Signal Processing Anal.* **3**, 582–586 (2003)
56. Shafique, K., Shah, M.: A noniterative greedy algorithm for multiframe point correspondence. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(1), 51–65 (2005)
57. Sbalzarini, I.F., Koumoutsakos, P.: Feature point tracking and trajectory analysis for video imaging in cell biology. *J. Struct. Biol.* **151**, 182–195 (2005)
58. Selmecci, D., Mosler, S., Hagedorn, P.H., Larsen, N.B., Flyvbjerg, H.: Cell motility as persistent random motion: theories from experiments. *Biophys. J.* **89**, 912–931 (2005)
59. Sethi, I.K., Jain, R.: Finding trajectories of features points in a monocular image sequence. *IEEE Trans. Pattern Anal. Mach. Intell.* **9**(1), 56–73 (1987)
60. Shiroguchi, K., Kinoshita, K. Jr.: Myosin V walks by lever action and Brownian motion. *Science* **316**, 1208–1212 (2007)
61. Sintorn, I.-M., Homman-Loudiyi, M., Soderberg-Naucler, C., Borgefors, G.: A refined circular template matching method for classification of human cytomegalovirus capsids in TEM images. *Comput. Methods Prog. Biomed.* **76**, 95–102 (2004)
62. Singer, R.A., Sea, R.G., Housewright, K.B.: Derivation and evaluation of improved tracking filters for use in dense multitarget environments. *IEEE Trans. Inf. Theory* **20**, 423–432 (1974)
63. Sivia, D.S., Carlile, C.J.: Molecular spectroscopy and Bayesian spectral analysis—how many lines are there? *J. Chem. Phys.* **96**(1), 170–178 (1992)
64. Smith, D.A., Steffen, W., Simmons, R.M., Sleep, J.: Hidden-Markov methods for the analysis of single-molecule actomyosin displacement data: the variance-hidden-Markov method. *Biophys. J.* **81**(1), 2795–2816 (2001)
65. Snyder, G.E., Sakamoto, T., Hammer, J.A., Sellers, J.R., Selvin, P.R.: Nanometer localization of single green fluorescent proteins: evidence that myosin v walks hand-over-hand via telemark configuration. *Biophys. J.* **87**, 1776–1783 (2004)
66. Steyer, J.A., Almers, W.: Tracking single secretory granules in the live chromaffin cells by evanescent-field fluorescent microscopy. *Biophys. J.* **76**, 2262–2271 (1999)
67. Streit, R.L., Graham, M.L., Walsh, M.J.: Multitarget tracking of distributed targets using histogram-PMHT. *Digit. Signal Processing* **12**, 394–404 (2002)
68. Suh, J., Wirtz, D., Hanes, J.: Efficient active transport of gene nanocarriers to the cell nucleus. *PNAS* **100**, 3878–3882 (2003)
69. Tamura, K., Hidaka, Y., Yusuf, Y., Kai, S.: Anomalous diffusion and Levy distribution of particles velocity in soft-mode turbulence in electroconvection. *Phys. A* **306**, 157–168 (2002)
70. Thompson, R.E., Larson, D.R., Webb, W.W.: Precise nanometer localization analysis for individual fluorescent probes. *Biophys. J.* **82**, 2775–2783 (2002)

71. Tonissen, S.M., Evans, R.J.: Performance of dynamic programming technique for track-before-detect. *IEEE Trans. Aerosp. Electron. Syst.* **32**(4), 1440–1451 (1996)
72. Veenman, C.J., Reinders, M.J.T., Backer, E.: Resolving motion correspondence for densely moving points. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(1), 54–72 (2001)
73. Veenman, C.J., Reinders, M.J.T., Backer, E.: Motion tracking as a constrained optimization problem. *Pattern Recognit.* **36**, 2049–2067 (2003)
74. Verestoy, J., Chetverikov, D., Marcell, N.: Digital particle image velocimetry: a challenge for feature based tracking. *Mach. Graphic Vis.* **8**(4), 553–569 (1999)
75. Vonderheit, A., Helenius, A.: Rab7 associates with early endosomes to mediate sorting and transport of semliki forest virus to late endosomes. *PLoS Biol.* **3**(7), 1225–1238 (2005)
76. Wilson, K.M., Morrison, I.E.G., Smith, P.R., Fernandez, N., Cherry, R.J.: Single particle tracking of cell-surface HLA-DR molecules using R-phycoerythrin labeled monoclonal antibodies and fluorescence digital imaging. *J. Cell Sci.* **109**, 2101–2109 (1996)
77. Yildiz, A., Forkey, J.N., McKinney, S.A., Ha, T., Goldman, Y., Selvin, P.: Miosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science* **300**, 2061–2065 (2003)
78. Zhang, B., Enninga, J., Olivo-Marin, J.C., Zimmer, C.: Automated super-resolution detection of fluorescent rods in 2D. In: 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro (6–9 April 2006), pp. 1296–1299 (2006)
79. Bohs, L.N., Geiman, B.J., Anderson, M.E., Gebhart, S.C., Trahey, G.E.: Speckle tracking for multi-dimensional flow estimation. *Ultrasonics* **38**, 369–375 (2000)
80. Li, C.H., Bai, L., Li, D.D., Xia, S., Xu, T.: Dynamic tracking and mobility analysis of single GLUT4 storage vesicle in live 3T3-L1 cells. *Cell Res.* **14**(6), 480–486 (2004)
81. Tvarusko, M., Bentele, M., Misteli, T., Rudolf, R., Kaether, C., Spector, D.L., Gerdes, H.H., Eils, R.: Time-resolved analysis and visualization of dynamic processes in living cells. *PNAS* **96**, 7950–7955 (1999)
82. Ponti, A., Vallotton, P., Salmon, W.C., Waterman-Storer, C.M., Danuser, G.: Computational analysis of F-actin turnover in cortical actin meshworks using fluorescent speckle microscopy. *Biophys. J.* **84**, 3336–3352 (2003)
83. Mitiche, A., Faghali, R., Mansouri, A.: Motion tracking as spatio-temporal motion boundary detection. *Rob. Auton. Syst.* **43**, 39–50 (2003)
84. Thomann, D., Rines, D.R., Sorger, P.K., Danuser, G.: Automatic fluorescent tag detection in 3D with super-resolution: application to the analysis of chromosome movement. *J. Microsc.* **208**(1), 49–64 (2002)
85. Thomann, D., Dorn, J., Sorger, P.K., Danuser, G.: Automatic fluorescent tag localization. II. Improvement in super-resolution by relative tracking. *J. Microsc.* **211**(3), 230–248 (2003)
86. Racine, V., Hertzog, A., Jouanneau, J., Salamero, J., Kervrann, C., Sibarita, J.-B.: Multiple-target tracking of 3D fluorescent objects based on simulated annealing. *IEEE, ISBI 2006*, 1020–1023 (2006)
87. Racine, V., Sachse, M., Salamero, J., Fraissier, V., Trubuil, A., Sibarita, J.-B.: Visualization and quantification of vesicle trafficking on a three-dimensional cytoskeleton network in living cells. *J. Microsc.* **225**(3), 214–228 (2007)