MULTIPLE HYPOTHESIS TRACKING IN MICROSCOPY IMAGES

Nicolas Chenouard[1] *, Isabelle Bloch*[2]*, Jean-Christophe Olivo-Marin*[1]

[1]Unité d'Analyse d'Images Quantitative, Institut Pasteur; CNRS URA 2582, Paris France
[2]Télécom ParisTech, CNRS UMR 5141 LTCI, Paris, France

ABSTRACT

Multiple hypothesis tracking (MHT) is a preferred technique for solving the data association problem in modern multiple target tracking systems. However in bioimaging applications, its use has long been thought impossible due to the prohibitive cost induced by the high number of objects that need to be tracked and the poor quality of images. We show in this paper that this broadly accepted view should change. We propose a MHT algorithm (fMHT) that is fast even when dealing with very noisy images of very numerous targets. We have applied the method to the analysis of two sets of real microscopy images that contain thousands of biological targets. By doing so we prove the benefits of the approach when tracking in very noisy environments such as low-light level fluorescent microscopy images.

Index Terms— Particle tracking, biomedical microscopy, Multiple Hypothesis Tracking (MHT)

1. INTRODUCTION

Time-lapse microscopy combined with automatic particles tracking is opening new ways to study and understand intracellular processes by looking directly at dynamic systems at the nanometric scale. The standard paradigm of particle tracking is a detection step of the objects of interest, followed by an association procedure between the measurement set and the active track set. This last procedure is made difficult by the corruption of the set of detections in two ways: (1) some artifacts originating from the acquisition noise are detected as objects of interest, (2) some detections are missed due to the temporary disappearance of targets in low image quality conditions or due to fluorophore extinction events. The Bayesian framework has been broadly adopted in the bioimaging community because it allows modeling these two sources of corruption of the detection process. Standard methods [1, 2] rely on an iterative selection of the most likely association at the considered frame. However, in dense target environments, associations are difficult to discriminate. Moreover, numerous spurious detections make false associations and false tracks creation a major issue with which instantaneous association methods have difficulties to deal with since they consider only one frame at a time.

Multiple hypothesis tracking (MHT) was proposed almost thirty years ago [3] to cope with the limitations of instantaneous tracking. The principle of the MHT is to delay the association step to a later time when the decision is made easier by the knowledge of future frames. In practice it relies on building all possible associations between tracks and detections for a number of successive frames and comparing them. Since it takes advantage of temporal information, the MHT is generally accepted as the method of reference for solving the data association problem in modern multiple target tracking systems. However, despite many improvements [4, 5], it has not been used for tracking biological targets in multidimensional microscopy images because its computational cost is generally prohibitive for this specific domain in which targets and false detections are numerous. Recently, *particle filtering* techniques, which try to reproduce the principle of a delayed decision step as proposed by the MHT, have gained popularity [6]. One main feature of the particle filtering approach is the relaxation of the competition of tracks for detections, which alleviates the time consuming step of enumerating all possible associations. However, in noisy images and numerous targets conditions, we argue that the competition principle between tracks is a key advantage for discarding false measurements and detecting tracks initiation and termination events. So in contrast with the particle filtering approach, we do impose a strict competition between tracks for detections.

We show in this paper that the broadly accepted view that the MHT technique is not applicable for tracking numerous targets in multidimensional microscopy images should be revisited. We propose a MHT formulation (fMHT) that integrates the notion of target perceivability [7], which is the capacity of a target to generate measurements in the future, and prove its impact on the overall performance of the algorithm. The model favors indeed good quality tracks and allows the early detection of a track termination and the initiation of tracks corresponding to real targets only. By doing so, the algorithm is very robust to false detections, and the complexity is reduced.

We introduce an algorithm design which allows us to define efficient pruning strategies and to implement it by taking full advantage of parallel computing technologies. The features of the proposed approach are studied by processing two image sequences in 2D and 3D fluorescent microscopy. While standard methods do not produce satisfactory results because of the poor image quality and the presence of thousands of objects of interest, the proposed algorithm is able to compute very good results in a short time.

We introduce the proposed tracking method in Section 2. In Section 3 comparative experimental results are presented, and the benefits of the MHT technique are proved for processing noisy 2D and 3D fluorescent images of numerous particles.

2. MHT FOR SWITCHING STATES TARGETS

2.1. Tracks likelihood model

We adopt a Bayesian framework in which we aim at building the set of tracks that has the greatest likelihood $\mathcal{L}(\Theta^l)$, which is defined as the probability of the tracks given the measurements from the se-

N.Chenouard is funded by C'Nano IdF.
Corresponding authors: N. Chenouard and J.C. Olivo-Marin,
email: {nicolas.chenouard, jcolivo}@pasteur.fr

quence of l images. We note respectively Z^l and Θ^l the set of measurements and tracks from time 1 to l. In particle tracking applications the set of measurements $Z(k)$ at time k generally corresponds to the set of detections in the k^{th} frame: $Z(k) = \{z_i(k)\}_{i=1..m_k}$, where each measurement is the vector of the detection coordinates $z_i(k) = [x_i(k), y_i(k), z_i(k)]^T$. Similarly, the set of tracks is composed of n elements $\Theta^l = \{\theta_j\}_{j=1..n}$, where each track θ_j is a succession of the assumed positions of the j^{th} target over time.
In the Bayesian formalism we decompose the likelihood $\mathcal{L}(\Theta^l)$ in the following way:

$$\mathcal{L}(\Theta^l) = P\{\Theta^l|Z^l\} = P\{\Theta^l, Z^l\}/P\{Z^l\} \propto P\{\Theta^l, Z^l\} \quad (1)$$

At each frame k, the set $Z_0(k)$ contains detections remaining unassigned to any track. These measurements are considered as being spurious detections coming from the sensor noise, which does not depend on the targets presence. Moreover, by assuming that target measurement and motion do not depend on other targets, we then write:

$$\mathcal{L}(\Theta^l) \propto \prod_{k=1..l} P\{Z_0(k)\} \prod_{j=1..n} p(\theta_j^{I_j}, z_{t_j}^{I_j})\} \quad (2)$$

The couple $(\theta_j^{I_j}, z_{t_j}^{I_j})$ corresponds to the set of positions of the target j and its associated measurements during the time interval I_j of its presence in the images.
In order to model accurately targets appearance and disappearance we define a two states model of targets perceivability. A target is *perceivable*, i.e. is in the state s^1, when it can be detected. The state s^0 corresponds to a *non perceivable* target which does not produce any measurement in future frames. In microscopy images, such a particle may have physically disappeared, bleached to a low level of intensity, or left the surveillance volume, hence the corresponding track has to be ended. Transitions between states occur with fixed probability between two frames: we note $\pi_{i,j}$ the probability of transition from state i to state j. In the following we consider that a non perceivable target cannot become perceivable again, hence $\pi_{0,1} = 0$.
The model leads us to write the measurement probability $p(\theta_j^k, z_{t_j}^k)$ by taking the target perceivability into account:

$$p(\theta_j^k, z_{t_j}^k) = p(\theta_j^k, z_{t_j}^k, s_j^0(k)) + p(\theta_j^k, z_{t_j}^k, s_j^1(k))$$

For each perceivability state $s_j^i(k)$ at time k we derive the following probability:

$$\begin{aligned} &p(\theta_j^k, z_{t_j}^k, s_j^i(k)) \triangleq \xi_j^i(k) \\ &= p(z_{t_j}(k)|z_{t_j}^{k-1}, \theta_j^k, s_j^i(k)) p(\theta_j(k)|z_{t_j}^{k-1}, \theta_j^{k-1}, s_j^i(k)) \\ &\cdot p(s_j^i(k)|z_{t_j}^{k-1}, \theta_j^{k-1}) p(\theta_j^{k-1}, z_{t_j}^{k-1}) \end{aligned} \quad (3)$$

where it can be shown using the Bayes' rule that:

$$\lambda_j^i(k) \triangleq p(s_j^i(k)|z_{t_j}^{k-1}, \theta_j^{k-1}) = \frac{\pi_{0j}\xi_j^0(k-1) + \pi_{1j}\xi_j^1(k-1)}{\xi_j^0(k-1) + \xi_j^1(k-1)}$$

which is the predicted probability of target perceivability at time k for $i = 1$. The measurement probability $p(z_{t_j}(k)|z_{t_j}^{k-1}, \theta_j^k, s_j^i(k))$ takes two forms in Equation (3), depending on whether it is a real detection or not. Indeed, when no measurement is assigned to a target, we associate it to a predicted measurement. We therefore derive the following probabilities:

$$p(z_{t_j}(k)|\theta_j^k, z_{t_j}^{k-1}, s_j^1(k)) = \begin{cases} p_d p_g \gamma_j^k & \text{if } z_{t_j}(k) \text{ is real,} \\ 1 - p_d p_g & \text{else.} \end{cases} \quad (4)$$

$$p(z_{t_j}(k)|\theta_j^k, z_{t_j}^{k-1}, s_j^0(k)) = \begin{cases} p_{fd}/V & \text{if } z_{t_j}(k) \text{ is real,} \\ 1 & \text{else.} \end{cases} \quad (5)$$

Here γ_j^k is the probability that the measurement $z_{t_j}(k)$ originates from the target j under the assumption that it exists up to the time k. In Equation (4), p_d and p_g are respectively the probability of detecting a target and the probability that the target position falls within its search gate. p_{fd} is the probability that the measurement $z_{t_j}(k)$ is a false detection, and V is the volume of the search gate.
In Equation (3), the probability $p(\theta_j(k)|z_{t_j}^{k-1}, \theta_j^{k-1}, s_j^1(k))$ corresponds to the probability of target position evolution. In the case of biological targets, whose motion can change abruptly, we use the Interacting Multiple Models filter [8, 1] that allows to accurately model switching type of movements by maintaining an adaptive mixture of motion models. When the perceivability state is s^0 this probability is set to 1 as the target position does not exist anymore.
It is worth pointing out that since the probability of perceivability of targets are integrated in the computation of the original score function (1) perceivable tracks are favored for association selection. Thus results are robust to associations with false measurements.

2.2. Fast MHT design

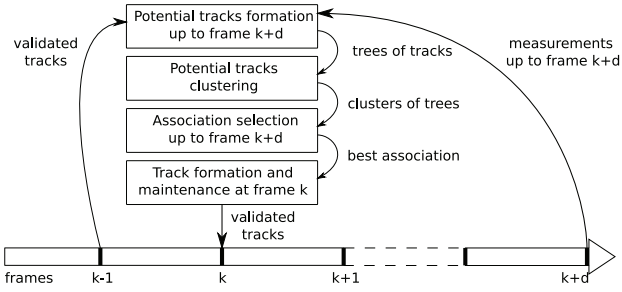

Fig. 1. MHT flow chart

The MHT technique builds iteratively the set of tracks $\Theta^\star(k)$ that maximizes the likelihood $\mathcal{L}(\Theta^{k+d})$ instead of maximizing $\mathcal{L}(\Theta^k)$ as instantaneous association algorithms do. At frame k the set of tracks $\Theta^{\star,k-1}$ has to be extended during $d+1$ frames with the set $Z^{k:k+d}$ of measurements up to time $k+d$. For each frame we adopt a four steps procedure summarized in Figure 1.
First, from each track $\theta_j^{\star,k-1}$ we build Γ_j^{k+d}, the set of potential associations with detections from $Z^{k:k+d}$. We model this process as the construction of a tree of potential tracks. Indeed the track $\theta_j^{\star,k-1}$ may give birth to a set of potential tracks Γ_j^k by association with the subset of measurements that fall into its search gate at frame k. Tracks in Γ_j^k may in turn create the potential tracks Γ_j^{k+1}, and so on. Moreover, we create a tree from each detection in $Z^{k:k+d}$ in order to model the possibility for new targets to appear.
We have taken advantage of the tree structure of the procedure to implement a fast track construction procedure that is massively parallel and allows the use of multithreading computing technologies.
During the nodes formation process we label the potential tracks according to their probability of perceivability. A track θ_j^t at frame t is *confirmed* if $\exists t' \leq t$ such that $\lambda_j^1(t') \geq p_c$, and *terminated* if $\lambda_j^1(t) \leq p_t$. (Details on the computation of the confirmation and termination thresholds, p_c and p_t, will be provided elsewhere.) It is useless to continue a terminated track because of its low probability of perceivability, so we stop the association in this case. For the same reason we consider only confirmed potential tracks. These two exclusion techniques reduce significantly the size of the association problem.
The second step consists in dividing the global association problem

into a set of smaller tasks by clustering the trees that are concurrent for at least one measurement.

The aim of the association selection procedure is then to find a subset of potential tracks in Γ^{k+d} that has the greatest likelihood. To achieve this task, we propose a solver that takes benefit of the tree structure of the tracks formation process in two ways: building only a very limited number of associations by pruning huge sets of solutions, and a massively parallel computing.

We begin by selecting a track θ_{t_1} in the first tree, so for now the selected set of tracks is: $\Theta = \{\theta_{t_1}\}$. The second tree is then considered. Its root compatibility $\theta_{t_2}^{k-1}$ is checked against measurements used by Θ. If a common measurement exists, the association is abandoned since each detection should be associated to one track at most, else the process is repeated for each node linked to the root. This procedure is applied until the end of the branch is reached. Then a Θ duplicate is extended with the track with which we ended θ_{t_2}: $\Theta = \Theta \cup \theta_{t_2}$, and another tree is selected. An association Θ is valid if it contains a potential track for every non terminated track in $\Theta^{\star,k-1}$, but still it can be extended with tracks corresponding to appearing targets.

We have also designed a *branch and bound* technique that discards association hypotheses with low score at an early stage. The track probability product in Equation (2) is decreasing with each addition of a track to Θ. Hence, each time the product falls beyond the best likelihood found until the current step, the best score cannot be attained anymore with the addition of other tracks. We therefore stop the association process for Θ in this case.

The association selection procedure has been designed in a recursive way such that parallel computing can be used: each time a node is considered a new thread is launched to perform the association task. In the final step the best association is built by merging the association found for each cluster: $\Theta^{\star k+d} = \cup_i \Theta_{c,i}^{\star k+d}$. On this basis, validated tracks are either continued or ended while a number of new tracks are validated.

3. APPLICATION TO MICROSCOPY IMAGES

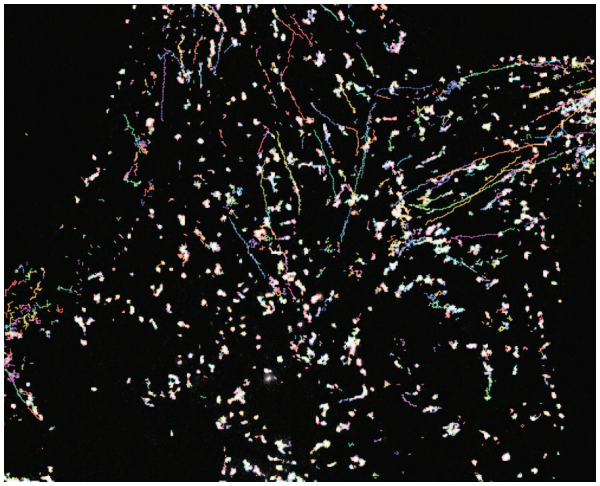

Fig. 2. Trajectories of Golgi units built by the fMHT

Trajectories of Golgi units in Chinese Hamster Ovary cells (CHO) are of first interest for studying the trafficking of a caveolin protein. The studied lineage of CHO expresses the Green Fluorescent Protein fused with caveolin-1, so thousands of Golgi units

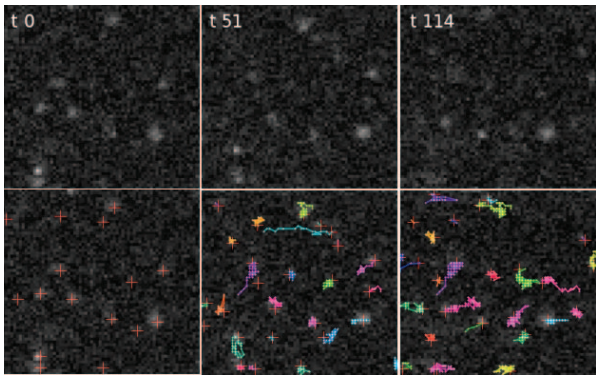

Fig. 3. Golgi units in a 2D restricted area at three times (first row). Second row shows the corresponding tracking by the fMHT.

containing the protein are visualized in disk scanning confocal microscopy as bright spots in an uneven background. 462 time points of a 504 by 405 pixels confocal slice of have been imaged, one of them is shown in Figure 2. The tracking task is challenging due to the numerous targets with various motions and the poor quality of the images. Indeed, the Peak Signal to Noise Ratio (PSNR = amplitude/noise standard deviation) is below 5 for low intensity particles. In order to assess tracking performances, an expert has manually identified Golgi unit trajectories in a 72 by 64 pixels crop during 128 time steps. In this small area 214 target trajectories were labeled.

We have applied a wavelet based detector [9] with sensitive settings which allows detecting low intensity targets, but which produces numerous false detections (24.5%). Hence the robustness to spurious measurements is a strong requirement for tracking procedures to succeed. Moreover the detections set contains 86.6% of the manually identified targets positions, hence taking into account missed detections is also an important issue.

The manually identified tracks are used as references for comparing the results obtained by different tracking procedures. Three standard tracking procedures were applied to the set of detections obtained by the wavelet analysis: instantaneous maximum likelihood tracking (IML) [1], Bayesian tracking with gap capping (GC)[2] and a MHT algorithm of depth 3 proposed by Cox and Hingorani in the computer vision field [5]. The fMHT algorithm performance was also studied, four settings were used ($d = 2, 3, 4, 5$) to investigate the influence of the depth of the algorithm.

For performances assessment purposes, we consider that a target has been correctly tracked if more than 75% of its positions are close (distance ≤ 3 pixels) to the measurements of a single track. In this case this track is said *significant*, while a track is said *false* if less than 75% of its measurements are close to the positions of reference trajectories. To quantify the quality of the set of tracks we use the Jaccard similarity index that takes into account both the number of significant and false tracks. Indeed, the index is computed as the ratio of the number of significant built tracks over the sum of the number of reference and the number of false tracks. We summarize tracking results obtained by the investigated methods in Table 1, while tracks obtained by the fMHT for $d = 4$ are shown in Figure 3.

Results show that the corruption of the detection set by spurious measurements leads instantaneous association algorithms such as IML and GC to produce many false tracks. It reveals their lack of robustness to cluttered conditions. The high corruption rate of the tracks set makes the extracted trajectories useless for most downstream analyzes. On the other hand, Multiple Hypothesis tracking algorithms show an improved robustness to spurious detections

Tracking method	Significant tracks	False tracks	Process time	Jaccard similarity
GC	51	30	6.9s	**0.21**
IML	175	142	2.3s	**0.49**
MHT [5] d =3	191	95	25s	**0.62**
fMHT d=2	183	37	8.8s	**0.73**
fMHT d=3	184	35	11.9s	**0.74**
fMHT d=4	177	34	17.0s	**0.71**
fMHT d=5	173	25	25.6s	**0.72**

Table 1. Golgi units tracking results. Top: performances for standard algorithms. Bottom: performances of the fMHT technique.

which is due to the automatic spurious detections exclusion made possible by the processing of both past and future frames.
When comparing MHT techniques, the proposed approach is shown to achieve the best Jaccard similarity index, whatever the value of the depth. The main reasons of the superiority of the fMHT over the MHT from [5] are twofolds: first we incorporate dedicated models of motion for biological particles, and second, the robustness to false detections is improved further, which results in a lower number of false tracks. The robustness in noisy conditions is due to the accurate model of target perceivability which is included in the tracks likelihood computation.
It is worth pointing out that the algorithm depth influences the resulting tracks only to a limited extent. For instance, increasing the depth improves only slightly the robustness to false detections. This characteristic is explained by the short range autocorrelation of diffusive motions, which are specific to biological applications. Indeed, when observing diffusive motions, only the few previous and next frames contain a significant information for the current tracking problem. The fMHT can therefore be used in these applications with a small depth ($d \leq 5$), which reduces the computation time without any loss of performance.
The fMHT procedure with a depth of 4 was also applied to the whole sequence of images. It took only 6 minutes, with a Mac Pro Quad 2.66GHz, to process the 462 frames which is fast compared to the complexity of the scene. In Figure 2 the 4024 resulting trajectories are displayed. It shows the ability of the algorithm to deal with high densities of targets, and to take into account various types of movements, while results obtained by the other methods are not exploitable because of their poor quality. Result movies can be found on line at *http://bioimageanalysis.org/2435/*.
In a second experiment the use of the fMHT technique on a 3D sequence of images was considered. The effect of the over expression of the tau protein on the transport of vesicles was already investigated with tracking tools in [1]. Vesicles were labeled with red fluorescent quantum dots and 3D images of living Hela cells were acquired slice by slice with a disk scanning confocal microscope. We have processed the whole sequence composed of 45 images ($440\times360\times5$) with a wavelet based detector slice by slice, and with the fMHT technique. The tracking task is made difficult both by clutter and by areas with high densities of closely spaced targets. Even in this case we were able to track the quantum dots in a short time: 5 minutes and 30 seconds. The 792 resulting trajectories are presented in Figure 4 (the movie is available on line). A qualitative inspection reveals the good quality of the resulting tracks.

4. CONCLUSION

The use of a new multiple hypothesis tracking algorithm (fMHT) is proposed for multidimensional images processing. The algorithm incorporates the probability of perceivability of a target. The evaluation of the probabilistic model for a number of successive frames allows us to automatically favor good quality tracks and to discard associations with false measurements. We have proposed an implementation that makes the procedure very fast even when tracking very numerous targets. This represents a major advance as MHT techniques have been known to be too demanding for such complex problems. The computation is efficient because the algorithm design allows automatic and efficient hypothesis pruning and taking advantage of advanced parallel computing technologies.
Our improved MHT algorithm was shown to outperform standard instantaneous tracking methods even in the case of a difficult particle tracking application. One key point of the method is its robustness to false measurements, which opens the way to the processing of biological images that were previously too noisy to be analyzed automatically.

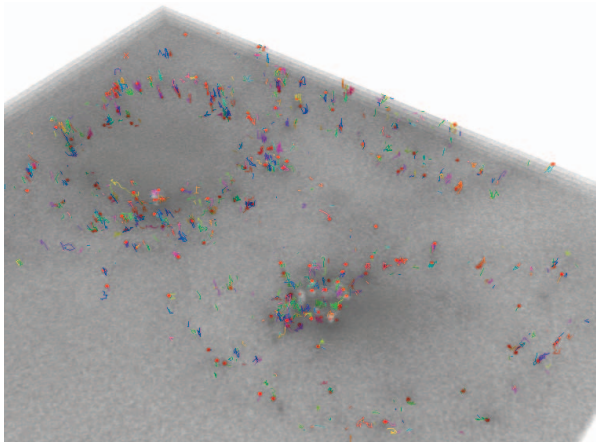

Fig. 4. 3D vesicles tracking in a Hela cell with the fMHT.

5. REFERENCES

[1] A. Genovesio, Liedl T., Emiliani V., Parak W.J., M Coppey-Moisan, and J.-C. Olivo-Marin, "Multiple particle tracking in 3D+T microscopy: methods and application to the tracking of endocytosed quantum dots," *ITIP*, vol. 15, no. 5, pp. 1062–1070, 2006.

[2] K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S. L. L. Schmid, and G. Danuser, "Robust single-particle tracking in live-cell time-lapse sequences.," *Nature methods*, vol. 5, no. 8, pp. 695–702, July 2008.

[3] D. B. Reid, "An algorithm for tracking multiple targets," *ITAC*, vol. AC-24, pp. 843–854, Dec. 1979.

[4] S. S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *ITAES*, vol. 19, no. 1, pp. 5–18, 2004.

[5] Ingemar J. Cox and Sunita L. Hingorani, "An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking," *ITPAM*, vol. 18, no. 2, pp. 138–150, February 1996.

[6] I. Smal, K. Draegestein, N. Galjart, W. Niessen, and E. Meijering, "Particle filtering for multiple object tracking in dynamic fluorescence microscopy images: Application to microtubule growth analysis," *ITMI*, vol. 27, no. 6, pp. 789–804, June 2008.

[7] N. Li and X.R. Li, "Target perceivability and its applications," *ITSP*, vol. 49, no. 11, pp. 2588–2604, Nov 2001.

[8] H. A. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with markovian switching coefficients," *ITAC*, vol. 33, no. 8, pp. 780–783, Aug. 1988.

[9] Olivo-Marin J.-C., "Extraction of spots in biological images using multiscale products," *Pattern Recognition*, vol. 35, no. 9, pp. 1989–1996, 2002.