Abstract—Deformable models have proven to be particularly adapted to multiple cell segmentation and tracking in biology, thanks to their flexibility, robustness to noise and quantitative interpretability. In the context of 4D microscopy, two concurrent techniques have been developed: an active surfaces model, based on an implicit contour formulation, and an active meshes model, based on an explicit formulation. This paper describes and confronts both approaches in terms of quantitative results and computational cost on simulated and real biological sequences, in order to provide an objective review of their strengths and weaknesses.

I. INTRODUCTION

Automated cell segmentation and tracking in multi-dimensional time-lapse microscopy is a field of growing importance in the biological community, for more and more applications focus, for instance, on the quantitative interpretation of cell motility and deformation through time [1]. Yet, many studies related to cell shape and motility rely still today on manual (though generally computer-assisted) analysis in 2D sequences [2]. This process is cumbersome and prone to user bias and reproducibility issues. Migration to full 3D time-lapse experiments is a necessary step to reliably quantify cell shape information, but induces new challenges in terms of image acquisition and quantification, yielding a need for more robust and fully automated analysis techniques.

Reliable automated techniques in segmentation rely nowadays on efficient PDE-based energy-minimization frameworks to increase robustness against the various disturbances occurring during the image formation process [3]. Among these techniques, active contours have proven their efficiency for cell tracking problems [1], [4]. The principle is to draw a contour near an object of interest, and deform the contour iteratively toward the object boundary. The tracking task is then straightforward: once an object has been segmented in a given sequence frame, the resulting contour can be used to initialize the model on the next frame, with the prior assumption that the object has not moved too much between successive frames. In the present context, the contour deformation is obtained by deriving forces from the minimization of an energy functional, which traditionally comprises two sets of terms: data attachment terms concretely define how the contour is attracted by image features to reach the solution, and regularization terms smooth the deformation by regularizing the frequently ill-posed minimization process, so as to avoid local energy minima.

Fluorescence microscopy images are generally characterized by a high level of noise. Consequently, cells boundaries are usually ill-defined, and intra-cellular structures may yield spurious artifacts, causing edge-based detectors to fail. Alternatively, region-based approaches integrate image information over the whole contour interior. They are less sensitive to such artifacts and are able to detect ill-defined edges efficiently. In this respect, we have been focusing on the piecewise-constant or reduced Mumford-Shah functional for image segmentation [5]. This energy functional is defined for a $D$-dimensional image $I$ by:

$$E_{MS}(c_{1..n}, \Gamma) = \sum_{i=1}^{n} \lambda_i \int_{\mathcal{R}_i} |c_i - I|^2 d\Omega + \mu \mathcal{H}^{D-1}(\Gamma), \quad (1)$$

and aims at reconstructing a piecewise constant image $\hat{I}$ partitionned by a discontinuity set $\Gamma$ into $n$ regions $\mathcal{R}_i$ of average intensity $c_i$. The data attachment terms weighted by $\lambda_i$ minimize the difference between $\hat{I}$ and $I$. The regularization term weighted by $\mu$ minimizes the $D-1$-dimensional Hausdorff measure of $\Gamma$. The interesting property of this functional is to concretely express the discontinuity set between the reconstructed regions, which can be directly represented by a contour or a set of contours.

To express $\Gamma$ using active contours, one may choose between two existing families of models studied in the literature: explicit models, relying on a Lagrangian formulation of the contour, and implicit models, relying on an Eulerian formulation. We recall below the main characteristics of both families.

Explicit models [6], [7] express the contour as a parametric curve or surface, concretely implemented by mapping a spline function of arbitrary degree onto a set of control points. Contour deformation is obtained by applying forces directly to the control points. 2D tracking can be found in [8] and adaptation to the Mumford-Shah functional has been studied in [9]. Nonetheless, such models are strongly dependent on the initial parametrization. An alternative solution called “active meshes” was proposed for motion tracking in 2D natural scenes [10]. The parametric contour is replaced by a discrete triangular mesh that is placed on the image such that the vertices track various points of interest over time to quantify relative objects movements. This idea has been extended to
3D images [11]–[14], where the parametric surface is replaced by a closed triangular manifold evolving in 3D coordinates-space. 3D mesh models are simpler to manipulate than parametric surfaces, and perform efficient self-reparametrization of the surface during its deformation. Some of these models additionally implement global topology operators, allowing semi and fully automated mesh splitting and merging [14], [15], which is a known limitation of spline-based approaches. While most of the available segmentation methods are based on edge information or iso-level extraction, none of these models has focused on the implementation of the Mumford-Shah functional. We will discuss this aspect in section III.

Implicit models [16] embed the contour as the zero-level of a higher-dimensional time-varying Lipschitz function called level set function. The deformation of the contour is thus implicitly directed by the deformation of the level set, obtained by applying the theory of front propagation in fluid dynamics [17]. Implicit models solve most of the problems inherent to classical parametric models. Firstly, the level set function is defined over the whole image domain, therefore no surface parametrization is required. Secondly, topology changes of the contour are implicitly performed as the level set deforms, allowing straightforward segmentation of multiple isolated objects simultaneously. Finally, the level set formalism is naturally multi-dimensional, and has thus been widely used for both 2D and 3D segmentation and tracking [18], [19]. The first level set implementation of the reduced Mumford-Shah functional is due to Chan and Vese for 2D segmentation [20]. This model further restricts the problem to the two-phase case (objects + background), and the discontinuity set $\Gamma$ is represented by a single level set function $\varphi$ that is able to detect all the objects in the image, provided they are spatially isolated and have similar intensity. In our context however, this model fails to distinguish cells with different intensity, as well as cells that touch over time, where their identity is lost due to uncontrolled front merging. A common solution to these problems is to use multiple level sets simultaneously, and couple them through the energy functional. This idea has been developed for 2D images in [21]–[23]. In [24], we developed a 3D extension of [23], where we enhanced contact localization via a volume conservation term. An extension based on local image feature detection was recently published in [25]. This model is further described in section II.

The outline of the paper is as follows: We describe the active surfaces model in section II, and the active meshes model in section III. In section IV, we compare both methods in terms of segmentation quality and computational cost. This comparison is discussed in V, suggesting that mesh models offer significant advantages for the problem at hand.

II. ACTIVE SURFACES

In this section we describe the level set-based model developed for multi-cell segmentation and tracking using a level set formalism. We recall here the key ideas of the model, and refer the reader to [24] for further implementation details.

A. Model

The Chan-Vese-Mumford-Shah model is a two-phase restriction of the problem (objects + background), where a single level set function is used to segment spatially isolated objects. To resolve contact events in a sequence, and to account for inter-cellular intensity variations, the model presented in [23] extends this model to the 2D multi-phase case by assigning a single level set function per cell, and by minimizing the overlapping region between these level sets. We extended this model to 3D images, and we incorporated a volume conservation term to enhance the frontier localization between touching cells, based on the assumption that cells keep a relatively constant volume throughout their displacements. The final energy functional is given by:

$$E(c_{0..n}, \varphi_{1..n}) = \int_{\Omega} \left( \sum_{i=1}^{n} \left[ \lambda_{in} H_{e}(\varphi_{i}) |c_{i} - I|^{2} + \mu \delta_{e}(\varphi_{i}) |\nabla \varphi_{i}| + \gamma \sum_{j=i+1}^{n} H_{e}(\varphi_{i}) H_{e}(\varphi_{j}) + \eta \left( \int_{\Omega} H_{e}(\varphi_{i}) \, d\Omega - V_{i,0} \right)^{2} \right] + \lambda_{out} \prod_{i=1}^{n} (1 - H_{e}(\varphi_{i})) |c_{0} - I|^{2} \right) \, d\Omega, \quad (2)$$

where $\lambda_{in}, \mu, \gamma, \eta$ are non-negative empirical parameters weighting the external, internal, coupling and volume conservation terms, respectively. $H_{e}$ and $\delta_{e}$ are smoothed versions of the Heaviside and Dirac functions (see [20] for more details on these functions), and $V_{i,0}$ expresses the reference volume of object $i$ computed from the first image of the sequence. For genericity purposes, we also implemented an edge detector term (not shown here), to capture edge information when available. In fluorescence microscopy however, such information is rarely available due to the high level of image noise.

B. Energy minimization

The minimization of eq. (2) is then performed using a classical Euler-Lagrange steepest gradient descent, yielding the following evolution equation for each level set $\varphi_{i}$:

$$\frac{\partial \varphi_{i}}{\partial t} = \delta_{e}(\varphi_{i}) \left( \lambda_{out} |c_{0} - I|^{2} - \lambda_{in} |c_{i} - I|^{2} + \mu \nabla \cdot \frac{\nabla \varphi_{i}}{|\nabla \varphi_{i}|} \right)$$

$$- \gamma \sum_{i \neq j} H_{e}(\varphi_{i}) H_{e}(\varphi_{j}) - 2\eta (V_{i,t} - V_{i,0}) \right) \quad (3)$$

Each equation is discretized using an explicit time-stepping scheme to compute the spatial derivatives of the regularization term $\nabla \cdot (\nabla \varphi_{i}/|\nabla \varphi_{i}|)$, imposing a small time step per iteration. Using an implicit scheme (i.e. larger time steps) would indeed reduce the number of iterations needed to converge [19], but increases the risks of jumping back and forth around the optimal solution. Instead, we managed to considerably reduce...
the number of iteration by providing a good initialization to the method (see [24] for details).

The complexity of the naive implementation is of the order of \( O(N^3) \) per level set per iteration, where \( N \) is the number of voxels along each image dimension. To improve performances, we employed the narrow-band approach to restrict each level set update to a narrow band around its zero-level [26], reducing the complexity to the order of \( O(wN^2) \) where \( w \ll N \) is the width of the narrow band. While this complexity yields reasonable computation times, the main bottleneck of this model resides in its memory consumption. Indeed, for each segmented cell, a level set function of same size as the image is created. This hinders application to large-scale populations.

C. Visualization

Visualization of the deforming surface is achieved by iteratively extracting the zero-level of the function and triangulating it into a 3D mesh model suitable for rendering. To do so we employed one of the most popular 3D reconstruction algorithms called the Marching Cubes [27]. This algorithm creates a triangular mesh from an image iso-level by dividing the image space into a cubic grid of arbitrary size (defining the final mesh resolution), and by generating mesh faces within cubes that intersect the desired iso-level. Unfortunately, such algorithms have two major drawbacks. Firstly, reconstruction times grow exponentially with the mesh resolution (of the order of minutes for a grid size of one voxel for typical fluorescent images). Hence, in order to maintain a reasonable computation time, we performed coarse surface reconstruction during the segmentation, and fine reconstruction only after convergence of the model on each sequence frame. Secondly, the grid decomposition induces surface approximation errors, visible in the form of a staircase effect (cf. fig. (2)-left). Although shading tricks can be employed to flatten this effect, this is undesirable and may affect qualitative interpretation.

III. ACTIVE MESHES

In this section we reformulate the above model using deformable meshes. We start by describing the geometrical properties of the meshes, and then rewrite the active surfaces model using this formalism. The reader may also consult [28] for technical details.

A. Model

3D triangular mesh models are discrete surfaces defined by a list of vertices forming a closed set of oriented triangles, such that the mesh boundary represents at all times the contour of a volumetric object. The deformation of the mesh is driven by that of its vertices, which evolve in a real-coordinates space bounded by the image (i.e. the vertices are not fixed on the image grid, and image values for each vertex are computed by trilinear interpolation of the neighboring voxels).

The geometrical properties of our mesh model are globally equivalent to that of Lachaud in [14], although we have adapted some of these properties to our particular context. Firstly, we maintain a regular mesh sampling by setting min and max edge length thresholds (defining the mesh resolution), and merge, split or invert edges that fall out of this length range. We do not want to adapt the mesh resolution according to local image discontinuities as done in [29] because fluorescence microscopy images usually contain spurious noise artifacts that may needlessly increase the mesh resolution. Secondly, we implement an analogous mesh splitting operator to account for eventual cell divisions, but we replace the merging operator by an efficient collision detection algorithm, in order to minimize contour overlap during cell-to-cell contact events (this algorithm is further detailed below).

Let us now describe how the level set-based model described earlier can be implemented using deformable meshes. We start by rewriting the energy functional eq. (2) using a mesh formalism:

\[
E(c_{0...n}, M_{1...n}) = \lambda_{out} \int_{\Omega_0} |c_i - I|^2 d\Omega + \sum_{i=1}^{n} \lambda_{in} \int_{\Omega_i} |c_i - I|^2 d\Omega + \mu \int_{M_i} dS + \gamma \sum_{j=i+1}^{n} \int_{\Omega_i \cap \Omega_j} d\Omega + \eta \left( \int_{\Omega_i} d\Omega - V_{i,0} \right)^2
\]

where \( \lambda_{in}, \mu, \gamma, \eta \) are defined similarly to eq. (2), \( \Omega_0 \) is the image region outside all meshes, and \( \Omega_i \) is the image region bounded by mesh \( M_i \).

B. Energy minimization

By analogy with the level set case, the Euler-Lagrange equations related to the minimization of this functional yield evolution equations (one per mesh), each given by:

\[
\frac{\partial M_i}{\partial t} = -\nabla \left( \mu \kappa_i + \lambda_{in} |c_i - I|^2 - \lambda_{out} |c_0 - I|^2 + \gamma \sum_{j \neq i} \xi_j (M_i) + 2\eta (V_{i,t} - V_{i,0}) \right),
\]

where \( \nabla \) is the outer unit normal vector to \( M_i \), \( \kappa_i \) expresses the local mean curvature of mesh \( M_i \) (thereby minimizing the global surface area), and \( \xi_j \) is an overlapping indicator function such that \( \xi_j = 0 \) outside \( M_j \) and \( \xi_j > 0 \) inside.

Most of the above terms are quite straightforward to implement, except the coupling term \( \sum_{j \neq i} \xi_j (M_i) \), for which we have developed an efficient collision detection strategy. Briefly, for we perform intersection tests only for vertices that are likely to penetrate other meshes. We first check if the mesh bounding spheres intersect, and extract for each mesh the vertices located within the opposite mesh bounding sphere. Only then, the actual intersection test is done by applying
the Jordan curve theorem, stating that a point lies within a contour if any ray traced from this point intersects the curve an odd number of times. The penetration index $\xi$ can be directly obtained if the ray is traced toward the initial mesh center. This algorithm is computational very efficient, although it is slightly more expensive than in the level set case where surface overlaps are permanently available at the voxel level.

The complexity of this model is $O(P)$ per mesh per iteration, where $P \approx N$ is the number of mesh vertices, yielding a drastic reduction compared to its level set counterpart, both in terms of computation time and memory consumption, since meshes are simply represented by a list of vertices and faces. Note that memory consumption is now dependent on the size of the meshes, but is substantially lower than that of the image.

C. Visualization

Mesh models offer the appealing property of being naturally in the form of a graphical model, they are therefore perfectly suited for real-time 3D rendering without having to compute a surface from the underlying model. Moreover, the mesh surface is no longer dependent on a reconstruction grid, solving the aforementioned staircase effect of level set models. Note that efficient real-time visualization also offers various advantages regarding qualitative interpretation, interactive manipulation, intuitive parameter adjustment and high-quality movie production, e.g. for scientific communication purposes.

IV. QUANTITATIVE COMPARISON

A. Segmentation

Ground-truth data is rarely available in biological imaging. In 2D, one may rely on manual segmentation produced by experts, but manual analysis of 3D images is a tedious and unprecise task. A more reliable approach is to evaluate the performance on simulated data, where ground-truth is available. We have thus generated synthetic 3D images containing two random-sized non-touching ellipses imitating cells, and disturbed these images to simulate a typical fluorescence microscopy system: background auto-fluorescence, poisson distribution of detected photons, gaussian noise induced by the instrument, and anisotropic voxel resolution (lower along the depth axis). We have generated three sets of 100 images with varying signal-to-noise ratio (SNR) by changing the standard deviation of the final Gaussian noise: 5 voxels for a high SNR, 10 voxels for a medium SNR and 15 voxels for a low SNR. A sample of medium and low SNR images is shown in fig. (1), alongside with a real image serving as a visual reference.

Segmentation results are given as a discrepancy measure in voxels between the result and the original binary object. Tables I and II present three error measures for each model: an absolute error giving the exact ratio of misclassified voxels, and two errors where misclassified voxels lying within one and two voxels of the real boundary are discarded.

Results show that both methods perform robust and equivalent segmentations, although active meshes tend to yield better results on higher SNR. In both cases, errors are confined within a small area around the real object surface, which is a satisfying behavior. The main difference between the models is their computational load: active meshes are much lighter and faster than active surfaces.

B. Tracking

In this section we evaluate tracking performance of both methods. For this experiment we have generated 100 binary sequences of 3 frames each, where two elliptic objects are separated in the first frame, touch in the second and separate in the third. During contact, the objects overlap by a small amount of voxels. Note that we have not employed noisy data here as we do not want to modify the evaluation with an additional disturbance factor that has already been studied in the previous section. Results are presented in table III.

<table>
<thead>
<tr>
<th>Tracking performance</th>
<th>Discrepancy (avg over 100 sequences)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active Surfaces</td>
<td>Active Meshes</td>
</tr>
<tr>
<td>before contact</td>
<td>0%</td>
</tr>
<tr>
<td>during contact</td>
<td>0%</td>
</tr>
<tr>
<td>after contact</td>
<td>0%</td>
</tr>
</tbody>
</table>

TABLE III

Tracking performance on simulated contact events. Errors are given as the sum of the absolute cell-background and cell-cell voxels misclassification ratios.
As expected, no segmentation error occurs before or after contact, showing that the coupling term correctly preserves the objects identity. During contact, both methods perform equivalently although active meshes have a minor advantage.

C. Real data

Figure 2 presents results on the final frame of a real biological sequence containing 4 cells. As expected, segmentation quality appears equivalent, however the 3D rendering of the active surfaces clearly depicts the staircase effect due to the grid dependance of the reconstruction algorithm, while active meshes match the cells surface in a smoother, yet more accurate manner.

Fig. 2. Analysis of a real fluorescence image of size 200 × 200 × 17 and resolution 1.5 × 1.5 × 5μm. (a) Result using active surfaces. Segmentation took 20 seconds and 3D reconstruction 30 seconds. (b) Result using active meshes. Segmentation took 8 seconds (rendering is performed in real-time).

V. Conclusion

We have presented two 3D active contour models implementing the reduced Mumford-Shah functional for multi-cell segmentation and tracking in fluorescence microscopy: active surfaces, based on a level set formalism, and active meshes, based on a discrete triangular mesh formalism. Although quantitative evaluation yields comparable efficiency, we have shown that active meshes clearly outperform active surfaces in terms of computational load and complexity, while offering real-time rendering of the surface during its deformation. We are currently working on applying this method to automated cell tracking and deformation studies, where robust quantitative measures will be of great importance to study the mechanisms of cell deformation in 3D environments.

REFERENCES