

3D Automated Nuclear Morphometric Analysis Using Active Meshes

Alexandre Dufour^{1,3}, JooHyun Lee², Nicole Vincent³,
Regis Grailhe², and Auguste Genovesio¹

¹ Image Mining Group, Institut Pasteur Korea

² Dynamic Imaging Platform, Institut Pasteur Korea

³ Intelligent Perception Systems (SIP-CRIP5) team, Paris Descartes University
{alexandre.dufour, agenoves}@ip-korea.org

Abstract. Recent advances in bioimaging have allowed to observe biological phenomena in three dimensions in a precise and automated fashion. However, the analysis of depth-stacks acquired in fluorescence microscopy constitutes a challenging task and motivates the development of robust methods. Automated computational schemes to process 3D multi-cell images from High Content Screening (HCS) experiments are part of the next generation methods for drug discovery. Working toward this goal, we propose a fully automated framework which allows fast segmentation and 3D morphometric analysis of cell nuclei. The method is based on deformable models called Active Meshes, featuring automated initialization, robustness to noise, real-time 3D visualization of the objects during their analysis and precise geometrical shape measurements thanks to a parametric representation of each object. The framework has been tested on a low throughput microscope (classically found in research facilities) and on a fully automated imaging platform (used in screening facilities). We also propose shape descriptors and evaluate their robustness and independence on fluorescent beads and on two cell lines.

1 Introduction and Related Efforts

The combination of microscopy and robotics enables to perform 2D visual cell based experiments in parallel and in a fully automated fashion. As a consequence, the exponential increase of images to analyze has motivated the development of fully automated frameworks. However, 2-dimensionality has some limitations, in particular for objects that are heterogeneous along the depth axis such as cell nuclei. Much more information can be obtained by acquiring depth-stacks of images, which allows to analyze the entire 3D structure of cellular or sub-cellular compartments [1].

The cell nuclear morphology constitutes a good start for such a study. A large array of biological functions is accompanied by major changes in the geometry of the nucleus [2]. Determining exactly how geometric characteristics relate to cellular function requires accurate 3D morphological information.

In addition to quantitative measurements, visual observation is also a key aspect of scene interpretation and understanding. Yet, visualizing a 3D scene

during its analysis remains a challenging task. Most methods employ a 3D reconstruction algorithm (e.g. the Marching Cubes [3]) to produce an intuitive rendering of the scene. These algorithms are time-consuming and suffer from surface approximation errors, therefore real-time visualization remains an issue.

The analysis of 3D fluorescent stacks is not trivial. Indeed, fluorescent images generally suffer from many disturbances induced by the imaging protocol (medium autofluorescence, acquisition noise etc.). However, one of these disturbance factors, namely the convolution with the microscope PSF, has a different impact in 2D and 3D. The PSF is not constant along the depth, and has a much stronger blurring effect on slices below and over the focus plane, yielding very fuzzy boundaries along the depth axis (cf. Fig. 1), causing most algorithms to fail detecting the edges correctly in 3D.

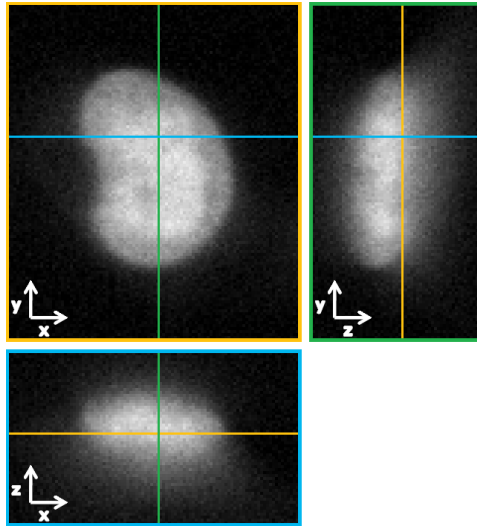


Fig. 1. Axis-based view of a 3D image of size $100 \times 115 \times 60$ pixels and resolution $0.28 \times 0.28 \times 0.5 \mu m$. Center: XY plane view. Right: YZ plane view. Bottom: XZ plane view. The YZ and XZ planes emphasize the blurring effect of the microscope PSF on the lower and higher Z planes of the volume.

In this context, deformable models (also known as “active contours”) have shown to be efficient thanks to their handiness and robustness to noise [4]. The idea is to deform an initial contour under the influence of various forces until it fits the target structure. These forces are usually computed from the minimization of a so-called *energy* functional describing the characteristics of the structure, and defined to be minimal when the model coincides with the target. Deformable models also offer a semantic interpretation of each object, allowing independent and precise object measurements (e.g. size, shape, resemblance to a reference model etc.) rather than global image measurements.

Two main families of deformable models can be distinguished, depending on the mathematical representation of the contour: in explicit models (known as *snakes* in 2D [5]), the boundary is represented by a parametric function, and in implicit models, the contour is defined as the zero level of a higher dimensional scalar function (called *level set function*) [6]. Each family has advantages and drawbacks, the choice thus mostly depends on the application. We briefly summarize the main advantages and drawbacks of both approaches in table 1.

Table 1. Brief comparison of the different advantages of 3D explicit and implicit deformable models. We focus on the aspects that concern our applicative context.

	Explicit	Implicit
Topology handling	-	+
Implementation	-	+
Memory consumption	+	-
Real-time visualization	+	-
Shape description	+	-

Implicit models handle contour splitting and merging implicitly, thus they are well suited to segment an unknown number of objects with a single contour. They are easy to implement in any dimension, however they manipulate a heavy data structure (of the size of the image), easily reaching hundreds of megabytes in 3D. Biological applications can be found in [7][8], however visualization is achieved using a 3D reconstruction algorithm, hence real-time visualization is not possible. Finally, geometrical measurements in a voxel-type structure is dependent on the resolution and thus yields approximation errors.

Explicit models perform faster, but are complex to implement in 3D. More and more methods therefore work directly with the discrete form of the surface (often called *polygonal mesh*) consisting of a set of connected points forming a closed polygonal manifold [9]. This representation enables the introduction of geometric rules that can handle surface splitting and merging [10]. Also, geometrical measures can be computed directly from the mesh in a simpler and more precise manner [11]. Since polygonal meshes rely on the same data structure as conventional computer graphic cards, 3D rendering is available with no additional time-cost, allowing real-time visualization. More popular in medical imaging [12], this approach has been recently applied to automated cell segmentation in fluorescence microscopy (the Active Mesh framework [13]).

In this paper, we propose a fully automated framework for nuclear shape segmentation and analysis based on the Active Mesh framework and propose a set of shape descriptors that can be used to discriminate different phenotypes of a given cell line, showing how this framework is suitable for 3D HCS applications. In section 2, we describe the biological experiment and present the analysis framework. Then we evaluate the method as well as the shape descriptors in section 3. Section 4 concludes the paper and discusses pending applications for the proposed framework.

2 Material and Methods

2.1 Biological and Imaging Protocol

A first experiment was conducted on two cell lines: HEK-293 (Human Embryonic Kidney) and Hela (Henrietta Lack), and a second was performed on $10\mu\text{m}$ fluorescent beads (96-Whatman without skirt, Evotec, Germany). All cells were grown on 96-well optical bottom plates, black (Greiner) under same culture conditions (DMEM with 10% FBS). Nuclei were labeled using DNA-specific DRAQ5 fluorescent dye (Biostatus, UK) following the instructions of the manufacturer.

Images were acquired at room temperature using 633nm excitation wavelength with 650nm long pass emission filter. The Z-stacks were obtained, for HEK-293 nuclei, on a confocal line-scanning microscope equipped with a oil-immersed plan apochromat 63x lens of NA 1.4 (LSM 5 Live, Zeiss, Germany), and for both Hela nuclei and fluorescent beads, on an automated Nipkow-disk confocal microscope (Opera, Evotec, Germany) equipped with a water-immersed plan apochromat 40x lens of NA 0.9 (Olympus, Japan).

2.2 Quantitative Analysis Method

In this section we describe the principal components of the nuclei analysis work flow, from segmentation to quantitative analysis. We start by describing the characteristics of the core segmentation method (the Active Mesh model), and then present each step of the final analysis work flow.

Definition of an Active Mesh. An active mesh [13] is a three-dimensional discrete surface 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). To avoid excessive complexity in the manifold structure, a regular sampling is imposed, such that all connected mesh vertices remain within an arbitrary distance interval $[d_{min}, d_{max}]$ from each other. Therefore, as the mesh grows or shrinks, vertices are respectively added or deleted automatically in order to maintain homogeneous edge lengths over the surface. To speed up computation, a multi-resolution approach is chosen, such that the distance interval varies during the evolution: the initial surface has a coarse resolution (vertices are far from each other). Then, as the surface approaches to the solution, d_{min} and d_{max} are progressively reduced, causing a global refinement of the mesh, and so until a suitable resolution is reached. This scheme allows fast and efficient sub-resolution segmentation.

Energy minimization. In our method, we choose to minimize the well-known Mumford-Shah piecewise-smooth functional (or *reduced* Mumford-Shah functional) [14]. This functional reads

$$F(\Gamma, c_1, \dots, c_n) = \lambda \sum_{i=1}^n \left[\int_{R_i} |u_0 - c_i|^2 d\omega \right] + \mu \int_{\Gamma} ds \quad (1)$$

and states that the target regions R_i , described by their mean intensity c_i , should resemble to the original image u_0 (first term), while the boundary set Γ between the regions should be minimal to avoid over-segmentation (second term). λ and μ are non-negative weighting parameters, and $d\omega$ and ds are the elementary volume and surface respectively. This functional has shown to be efficient for cell and nucleus segmentation in both 2D and 3D fluorescence imaging [15][8], since the target entities are fully stained and have very few corners and cusps. One region R_{out} represents the image background, and every other region $R_{i>0}$ represents an object that will be segmented by a specific mesh. The boundary set Γ thus corresponds to the set of meshes that evolve in the image domain, and the equation above can be rewritten as follows:

$$F(\mathcal{M}_1, \dots, \mathcal{M}_n, c_{out}, c_1, \dots, c_n) = \lambda \int_{R_{out}} |u_0 - c_{out}|^2 d\omega + \sum_{i=1}^n \left[\lambda \int_{R_i} |u_0 - c_i|^2 d\omega + \mu \int_{\mathcal{M}_i} ds \right], \quad (2)$$

where R_{out} denotes the background component of the image with mean intensity c_{out} , and c_i is the mean intensity inside the mesh \mathcal{M}_i segmenting the object i . The minimization is done using a steepest gradient-descent method using the Euler-Lagrange equations (see details in [13] and [16]). The final algorithm complexity is $O(N)$ per iteration, where N is the total number of vertices forming the n meshes. The number of iterations depends on the model initialization, as we shall discuss below.

Initialization. Due to the non-convexity of the energy functional in Eq. 2, convergence is only guaranteed to a local minima. Therefore, deformable models perform better and faster when they are initialized close to the solution. To avoid manual initialization, we propose the following automatic scheme:

- a. Blur the original stack with a Gaussian filter,
- b. Threshold the blurred stack using a 2-class K-Means algorithm,
- c. Extract the connected components (number and average diameter),
- d. Eliminate the objects partially visible (i.e. on the image edge),
- e. Initialize each surface by a coarse 3D reconstruction of each component,
- f. Evolve all surfaces simultaneously on the original (non-blurred) stack.

The 3D reconstruction involved in step (e) utilizes the *Marching Tetrahedra* algorithm [17]. This algorithm has the interesting property of using the same data structure as an active mesh. Hence, no data conversion is necessary, and the surface can be directly used as an initialization, that will hence be very close from the target boundary. Although 3D reconstruction algorithms are time-consuming at fine resolution, a coarse (i.e. fast) reconstruction is sufficient in our case since the model handles refinement automatically during the segmentation.

Visualization. Since each active mesh utilizes the same data structure as current graphic cards (typically a set of connected vertices), the rendering is

straightforward and performed on the graphic card parallelly to the main computation, yielding no additional time cost. This feature first allows real-time visual monitoring of the analysis, for instance to tweak the algorithm parameters. Secondly, it allows the method to save the 3D scene corresponding to each stack in a database, in order to provide off-line visual feedback after the analysis.

2.3 Statistical Analysis

In order to describe the nuclei shapes as best as possible, we wish to find a set of independent measures in order to compute robust statistics on the objects. In the following, we compute the following criteria from the final mesh: *Surface*, *Volume*, *LongAxis*, *Roughness*, *RadiusCV*, *HullDiff*. While *Surface* and *Volume* are quite self-explanatory, the other criteria are less obvious and detailed below.

- The *LongAxis* measure is the longest distance between two mesh vertices, eventually serving as an object elongation indicator.
- The *Roughness* is a measure outlining the local vibrations of the surface membrane. This measure should be low for convex objects and higher when the surface exhibits local concavities. To compute this value, we start by defining a local curvature measure for each mesh vertex v as the dot product between the outer normal \vec{N}_v (of unit length) and the barycentric normal \vec{B}_v linking v to the center of its neighbor vertices in the mesh (see figure 2). If the vectors have opposite directions (i.e. negative dot product), the surface is locally convex. If the vectors have same directions (i.e. positive dot product), the surface exhibits a local concavity at the given vertex. Finally, the roughness measure is defined as the standard deviation of all the local curvature values. Reference value is 0 for a sphere.
- The *RadiusCV* measure describes how different the object shape is from a sphere. This measure is obtained for each mesh by computing the standard deviation of the distances between the mass center and each vertex, normalized by the mean radius (definition of the coefficient of variation). Reference value is 0 for a sphere.

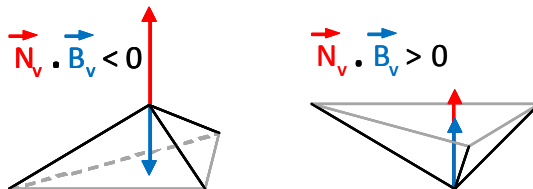


Fig. 2. Description of the roughness measure at a given vertex v . In case of a local convexity (left), the outer normal (red) and barycentric normal (blue) have a negative dot product. In case of a local concavity (right), the dot product is positive.

- The *HullDiff* measure is the difference percentage between the volume of the object and that of its convex hull (i.e. the smallest convex surface that can contain it). This measure will be useful to discriminate bean-shaped objects for instance. Reference value is 0 for a convex object.

3 Experiments and Results

3.1 Segmentation

The automated segmentation and shape measurement protocol was first tested on a set of Z-stacks of HEK-293 cells acquired one by one on a Zeiss LSM 5 *Live* microscope, yielding 22 stacks of size $512 \times 512 \times 60$ voxels and spatial resolution $0.28 \times 0.28 \times 0.5 \mu\text{m}$, totalizing 121 nuclei. Then, the method was applied on Hela cells using an automated imaging platform (Evotec Opera). We used 20 wells of a 96-well plate, and acquired in each well one Z-stack of size $688 \times 520 \times 31$ voxels and spatial resolution $0.327 \times 0.327 \times 0.75 \mu\text{m}$, totalizing 201 nuclei. The computation time ranged from 20 to 40 seconds per stack for all experiments, depending on the number of objects. This time includes: stack loading into memory, initialization (see section 2.2), segmentation and shape measurements of the detected objects.

Figures 3 and 4 present results for the HEK-293 and Hela cells experiments respectively. Left images show a maximum intensity projection (MIP) of one of the Z-stacks. Middle images show a snapshot of the 3D scene taken right after initialization. One can clearly see that cells touching the image edge have been automatically removed, and that the coarse 3D reconstruction using the Marching Tetrahedra are fast and efficient estimates of the nuclei surfaces. Right images show a similar snapshot at the end of the segmentation.

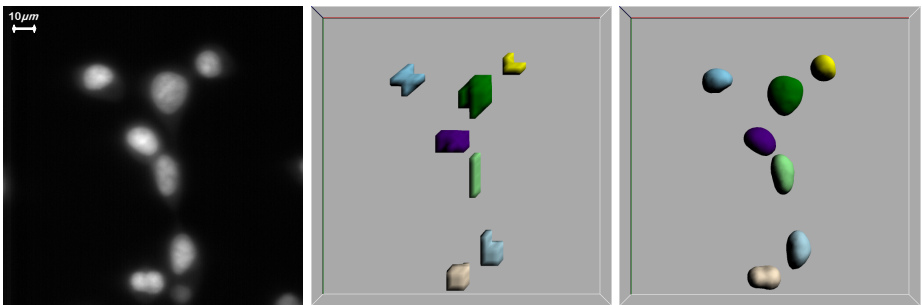


Fig. 3. Segmentation of a HEK-293 cell nuclei Z-stack (size $512 \times 512 \times 60$). Left: maximum intensity projection of the original stack. Middle: snapshot after initialization (coarse 3D reconstruction). Right: snapshot after segmentation.

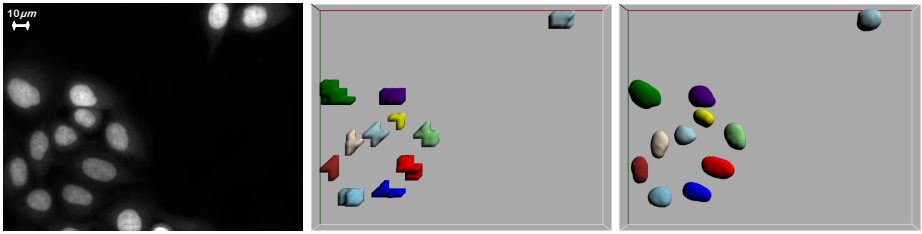


Fig. 4. Segmentation of a HeLa cell nuclei Z-stack (size $688 \times 520 \times 31$). Left: maximum intensity projection of the original stack. Middle: snapshot after initialization (coarse 3D reconstruction). Right: snapshot after segmentation.

3.2 Shape Analysis

The validation contains two steps. First, we check that our shape measures are consistent on fluorescent beads. Then, we check their independence in order to keep a compact set of non-redundant shape descriptors.

Validation on Fluorescent Beads. We have conducted a screening experiment on fluorescent beads following the protocol described in section 2.1. Expected values and average measures over 100 beads are given in table 2. Although all measures are close from the expected values, detected objects seem generally bigger than the real objects (e.g. the *LongAxis* measure is 14% higher). This is due to the growing effect of the microscope PSF along the Z axis. This effect decreases as the objects size increases, therefore this error is expected very low for our real experiments, where nuclei are bigger than the beads.

Table 2. Evaluation of shape descriptors on 10 μm fluorescent beads. Measured values are averaged over 100 beads. Coefficients of variation below 1 indicate low-variance populations.

	Surface	Volume	LongAxis	Roughness	RadiusCV	HullDiff
Expected	314.1	523.5	10	0	0	0
Measured	326.9	546.2	11.4	0.02	0.09	0.002
Coef. Var.	0.230	0.015	0.016	0.058	0.032	0.165

Dispersion. We evaluate the dispersion of each measure by computing its coefficient of variation (CV) on each population, i.e. the standard deviation-to-mean ratio. Results are shown in table 3. All coefficients are below 1, implying stable measures, nonetheless, some measures have a higher value than others. For instance, the *HullDiff* measure has a CV around 0.5 for both populations, therefore care should be taken in its interpretation in a shape comparison context. Same remark applies to the *Volume* measure in the HEK-293 case.

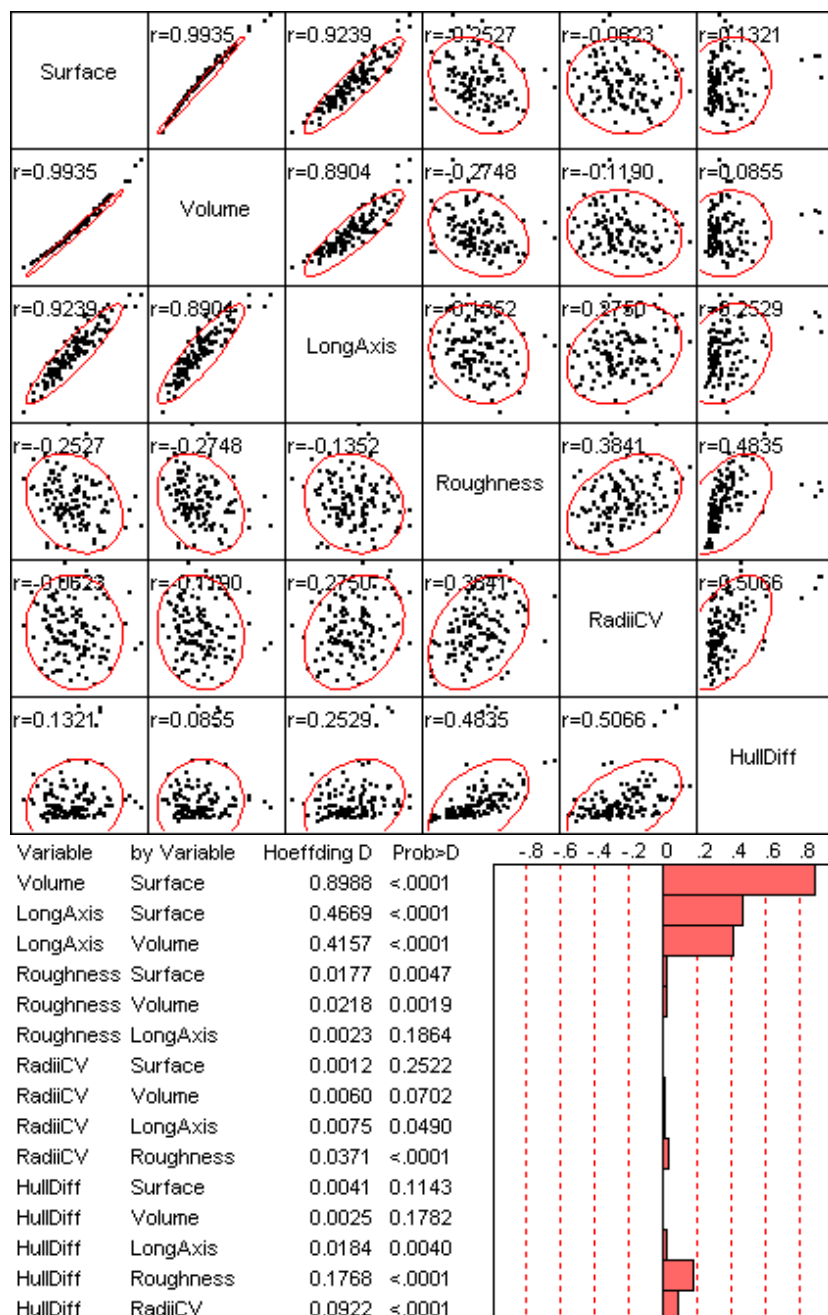


Fig. 5. Statistics on the HEK-293 cell line. Correlation (top) and Hoeffding's D (bottom) measures are given for the criteria presented in section 2.3. D values range from -0.5 to 1 , 1 indicating complete dependence. Red ellipses cover 90% of the population.

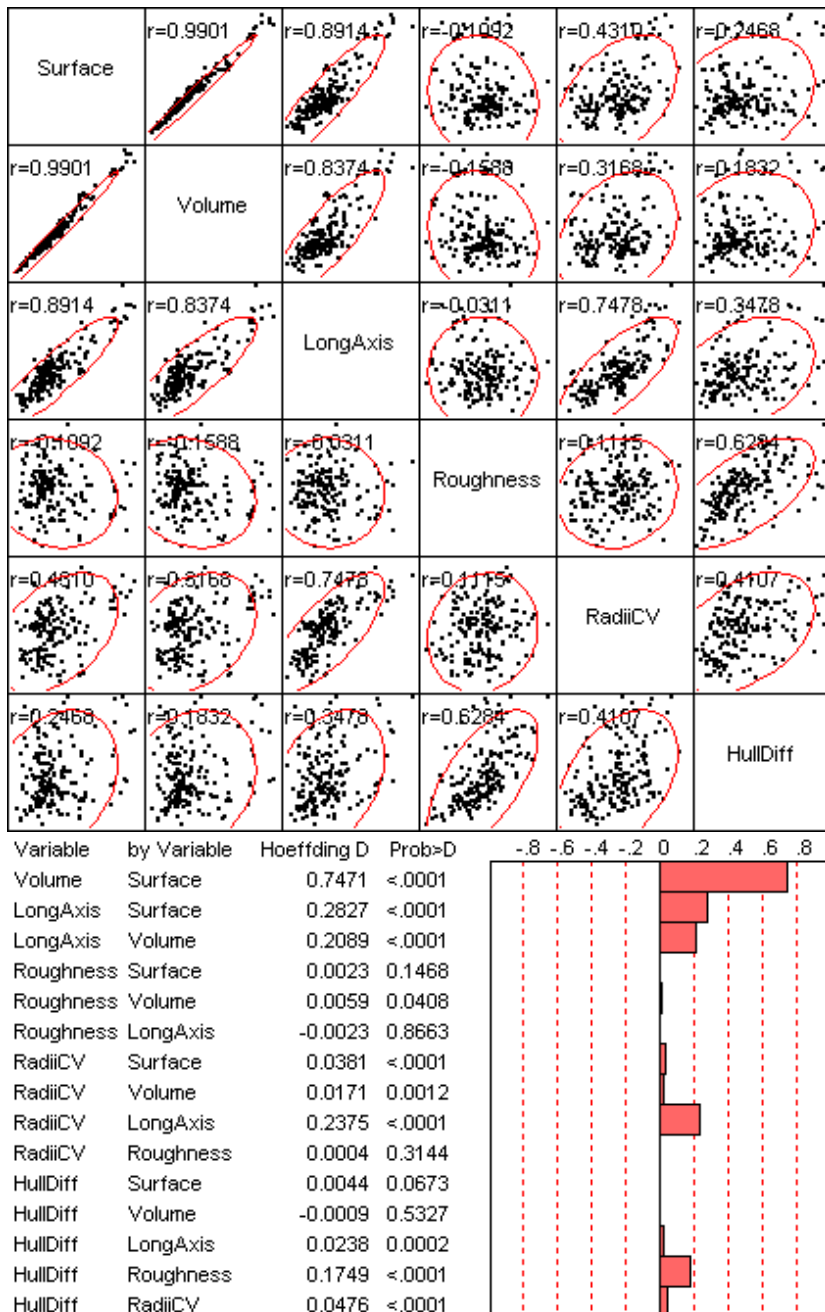


Fig. 6. Statistics on the HeLa cell line. Correlation (top) and Hoeffding's D (bottom) measures are given for the criteria presented in section 2.3. D values range from -0.5 and 1 , 1 indicating complete dependence. Red ellipses cover 90% of the population.

Table 3. Coefficient of variation of each measure on each population

	Surface	Volume	LongAxis	Roughness	RadiusCV	HullDiff
HEK nuclei	0.280	0.425	0.145	0.108	0.281	0.544
Hela nuclei	0.196	0.278	0.135	0.128	0.264	0.455

Robustness. Finally, we determine the robustness of our criteria by computing two correlation measures: the classical correlation and the Hoeffding measure of dependence D [18]. Results are given in Figures 5 (HEK cell line) and 6 (Hela cell line). Figures were obtained using the multiple correlation analysis tool of JMP software (SAS Institute, 1994). The strong correlation between the *Volume* and *Surface* measures, as well as with the *LongAxis* measure, coincides with the fact that these three measures are closely linked for any convex object. Another interesting observation is the relation between the *HullDiff* and the *Roughness* measures. This is due to the fact that a surface concavity at a given point creates a volume gap with the convex hull at that point. However, due to its local nature, the *Roughness* measure is not suited to detect large but smooth concavities such as for bean-shaped objects, for which *HullDiff* is much more efficient.

4 Conclusion

In this paper, a fully automated framework has been proposed for efficient 3D segmentation and morphometric analysis of cell nuclei, in live cells. We have found five independent 3D shape descriptors to describe our cell lines: *Volume*, *LongAxis*, *Roughness*, *RadiusCV*, and *HullDiff*. These measures will be used to study changes in cell phenotypes under challenging conditions. The method is robust and particularly well adapted to 3D fluorescence microscopy. We further plan to implement a larger array of shape descriptors, in order to enable better discrimination. Although this is not the case of nuclei, it is clear that in some applications the objects of interest may be touching and would need to be separated before their analysis. We are thus working on automated separation algorithms, in order to provide a robust and generic analysis tool for shape analysis in 3D HCS.

References

1. Vonesch, C., Aguet, F., Vonesch, J.L., Unser, M.: The colored revolution of bioimaging. *IEEE Signal Processing Magazine* 23(3), 20–31 (2006)
2. Leman, E., Getzenberg, R.: Nuclear matrix protein as biomarkers in prostate cancer. *Journal of Cell Biochemistry* 86(2), 213–223 (2002)
3. Lorensen, W., Cline, H.: Marching cubes: a high resolution 3D surface construction algorithm. In: *SIGGRAPH'87. 14th annual conference on Computer graphics and interactive techniques*, pp. 163–169. ACM Press, New York (1987)

4. Zimmer, C., Zhang, B., Dufour, A., Thebaud, A., Berlemont, S., Meas-Yedid, V., Olivo-Marin, J.C.: On the digital trail of mobile cells. *Signal Processing Magazine* 23(3), 54–62 (2006)
5. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. *International Journal of Computer Vision* 1, 321–331 (1988)
6. Sethian, J.A.: *Level set methods and fast marching methods*, 2nd edn. Cambridge University Press, Cambridge (1999)
7. Malpica, N., de Solorzano, C.O.: Automated Nuclear Segmentation in Fluorescence Microscopy. In: *Science, Technology and Education of Microscopy: an Overview. Microscopy Book Series, Formatex*, vol. 2, pp. 614–621 (2002)
8. Dufour, A., Shinin, V., Tajbaksh, S., Guillen, N., Olivo-Marin, J., Zimmer, C.: Segmenting and tracking fluorescent cells in dynamic 3d microscopy with coupled active surfaces. *IEEE Transactions on Image Processing* 14(9), 1396–1410 (2005)
9. Delingette, H.: General object reconstruction based on simplex meshes. *International Journal on Computer Vision* 32, 111–146 (1999)
10. Lachaud, J., Montanvert, A.: Deformable meshes with automated topology changes for coarse-to-fine three-dimensional surface extraction. *Medical Image Analysis* 3(2), 187–207 (1999)
11. Zhang, C., Chen, T.: Efficient feature extraction for 2D/3D objects in mesh representation. In: *International Conference on Image Processing, Thessaloniki*, pp. 935–938 (2001)
12. Zhukov, L., Bao, Z., Guskov, I., Wood, J., Breen, D.: Dynamic deformable models for 3D MRI heart segmentation. In: *SPIE Medical Imaging*, vol. 4684, pp. 1398–1405 (2002)
13. Dufour, A., Vincent, N., Genovesio, A.: 3D Mumford-Shah based active mesh. In: *Martínez-Trinidad, J.F., Carrasco Ochoa, J.A., Kittler, J. (eds.) CIARP 2006. LNCS*, vol. 4225, pp. 208–217. Springer, Heidelberg (2006)
14. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure App. Math.* 42, 577–684 (1989)
15. Zhang, B., Zimmer, C., Olivo-Marin, J.C.: Tracking fluorescent cells with coupled geometric active contours. In: *International Symposium on Biomedical Imaging, Arlington*, pp. 476–479 (2004)
16. Zimmer, C., Olivo-Marin, J.C.: Coupled Parametric Active Contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(11), 1838–1842 (2005)
17. Gueziec, A., Hummel, R.: Exploiting triangulated surface extraction using tetrahedral decomposition. *IEEE Transactions on Visualization and Computer Graphics* 1(4), 328–342 (1995)
18. Hoeffding, W.: A class of statistics with asymptotically normal distribution. *The Annals of Mathematical Statistics* 19(3), 293–325 (1948)