

2D Mapping of Strongly Deformed Cell Nuclei Based on Contour Warping

Jonas De Vylder, Koen Douterloigne, Winnok De Vos, Wilfried Philips

Abstract—The dynamics of genome regions are associated to the functional or dysfunctional behaviour of the human cell. In order to study these dynamics it is necessary to remove perturbations coming from movement and deformation of the nucleus, i.e. the container holding the genome. In literature models have been proposed to cope with the transformations corresponding to nuclear dynamics of healthy cells. However for pathological cells, the nucleus deforms in an apparently random way, making the use of such models a non trivial task. In this paper we propose a mapping of the cell nucleus which is based on the matching of the nuclear contours. The proposed method does not put constraints on the possible shapes nor on the possible deformations, making this method suited for the analysis of pathological nuclei.

I. INTRODUCTION

The internal organisation of the human cell nucleus in space and time is essential for its function. Within the limited space of the nucleus the entire genome as well as many proteins are accommodated in a non-random manner. Of particular interest are telomeres, the ends of chromosomes, which show a spatiotemporal behaviour that is functionally relevant to the cell and organism. Telomeres are arranged in distinct patterns and display mobility regimes at different time scales. Diverse biological processes such as telomere maintenance, cell ageing and apoptosis, i.e. cell death, are associated with specific telomere movements [1]. Likewise, altered telomere dynamics and by expansion nuclear protein dynamics are associated with specific diseases such as laminopathies and cancer [2], [3]. Therefore, quantitative studies of nuclear dynamics may help reveal novel mechanisms of dysfunction and disease.

However, analyses of dynamics in time-lapse microscopic image data sets are hampered by global cell motion and deformation. This superimposes a motion on the submicron dynamics that needs to be removed. Gladilin et al. have proposed a mapping based on a spherical model [4]. This model can cover for most of the cellular and nuclear displacements (translation, rotation and small affine transformations) that occur in normal cells during interphase. However, during cell division and in certain pathological conditions, such as laminopathies, the nuclear shape alters dramatically [3]. In Fig. 1 an example of such a pathological nucleus is

shown. As can be seen, a spherical model is no longer valid. Hence a method is required that allows reliable motion measurements of subnuclear features in a deformable volume without imposing any shape constraints.

In this paper we propose a new mapping technique which maps two nuclei between consecutive time points without imposing any constraint such as a spherical shape. Instead we will extract the contours out of segmented micrographs. These contours will be matched in such a way that we retrieve a point to point correspondence. Based on this matching a mapping of the full nucleus is calculated using polyharmonic splines. This mapping predicts the location of a telomere based on the deformation of the nucleus, this predicted location can be compared to the real location of the telomere in order to study telomere dynamics.

This paper is arranged as follows: in the next section we describe a simple segmentation technique, which can be used to extract the nucleus out of a micrograph. Section III provides a detailed description of the contour matching algorithm. The interpolation using polyharmonic splines is described in section IV. In section V the validation results are explained and discussed. Section VI recapitulates and concludes.

II. SEGMENTATION

Since the proposed method estimates the topological changes based on the deformation of the nucleus, we first need to segment the nucleus. In this paper we start from fluorescent 3D micrographs captured by a confocal microscope. We project the 3D micrograph on a 2D image. This is done by taking the average in the z direction. Depending on the microscope system it is of course possible to start immediately from 2D micrographs such as e.g. widefield images. In order to minimize the influence of noise, the 2D image is filtered using a Gaussian kernel. The resulting image is then thresholded using Otsu's thresholding [5]. This method calculates a threshold in order to minimize the variance of the foreground and the variance of the background. Due to remaining noise, not all desired pixels are considered to be foreground, i.e. part of the nucleus, whereas some background pixels are considered to be foreground. This problem is solved by using a morphological opening and closing. Finally the border delineating the segment is filtered with a Gaussian kernel, i.e. we assume the nucleus has a smooth surface. This segmentation algorithm was developed for a specific dataset which consists of isolated cell nuclei which have good contrast compared to the background. This of course depends on the application and the used fluorescent

Jonas De Vylder is funded by the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT).

De Vylder J., Douterloigne K. and Philips W. are with the department of Telecommunications and Information Processing, IBBT - Image Processing and Interpretation Group, Ghent University jonas.devyllder@telin.ugent.be

De Vos W. is with the department of Molecular Biotechnology, Bio-imaging and Cytometry Unit, Ghent University.

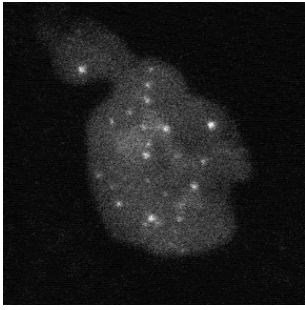


Fig. 1. An example of a fluorescent micrograph of a pathological nucleus.

dyes. The method for nuclei mapping proposed in this paper can also be used with more complex images, but would then require a more suitable segmentation technique such as described in [6], [7].

III. CONTOUR WARPING

In this section contours delineating the nucleus in subsequent frames will be matched, i.e. points belonging to the contour in a frame will be linked to their counterpoints on the contour in the next frame. In order to do so, we calculate a signature out of the contour, and then based on this signature the matching is done.

A. Contour Signatures

Consider the contour $\mathbf{c}(t) = (\mathbf{x}(t), \mathbf{y}(t))$ with $t \in \mathbb{N}$ and $\mathbf{c}(N+1) = \mathbf{c}(1)$, where N is the number of samples. From this contour, different signatures can be calculated:

- a complex signature:

$$\mathbf{s}(t) = \mathbf{x}(t) + i\mathbf{y}(t)$$

where $i^2 = -1$

- a centroid distance signature:

$$\mathbf{s}(t) = \sqrt{(\mathbf{x}(t) - x_c)^2 + (\mathbf{y}(t) - y_c)^2}$$

where (x_c, y_c) is the centroid of \mathbf{c}

- a direction signature, i.e. direction of the tangent line at point t :

$$\mathbf{s}(t) = \cos^{-1} \left(\frac{\mathbf{y}(t+1) - \mathbf{y}(t-1)}{\|\mathbf{c}(t+1) - \mathbf{c}(t-1)\|} \right)$$

- a curvature signature, i.e.

$$\mathbf{s}(t) = \left| \frac{\partial^2 \mathbf{c}(t)}{\partial t^2} \right|$$

The effect of these different signatures on the proposed framework is tested and described in section V. For a more detailed overview of signature functions we refer to [8], [9].

B. Signature Matching

Note that each point on a signature exactly corresponds to one point on the curve. So matching contour points is equivalent to matching points on a signature. For the matching of two signatures, we propose a technique based on *Dynamic Time Warping* or *dog-man distance*. The DTW

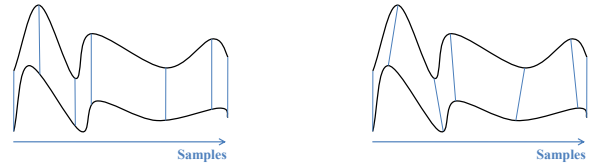


Fig. 2. An example of matching two curves. On the left, a naive approach where the matching is done using the parameterization. On the right, the matching with the minimal cumulative difference between the samples is shown. It is this matching which will correspond with the warping used for DTW.

distance between two signals is analogous of a man and a dog each walking on a different path, i.e. the signals. The dog can walk with a different speed by giving it a longer leash. Fig. 2 shows on the left two paths, where both paths are scanned at the same speed, so man and dog are walking with the same speed. On the right of the figure, different speeds result in a better match, i.e. the cumulative difference between the samples is less. The dog-man distance is then the shortest possible cumulative difference between both paths, by changing the scanning speed, i.e.

$$d_{DTW}(\mathbf{s1}(\cdot), \mathbf{s2}(\cdot)) = \min_{\gamma} \sum_t \|\mathbf{s1}(t) - \mathbf{s2}(\gamma(t))\| \quad (1)$$

where $\gamma(\cdot)$ is a warping function, i.e. any monotonic function mapping $[1, N]$ on $[1, N]$. This can be calculated using dynamic programming:

$$D(i, j) = d(i, j) + \min(\alpha D(i, j-1), \alpha D(i-1, j), D(i-1, j-1)) \quad (2)$$

with $d(i, j) = \|\mathbf{s1}(i) - \mathbf{s2}(j)\|$ and α a real number greater than or equal to one. If $\alpha = 1$, all warping functions are considered equally good. In order to penalize warping functions where too many points of one signal are matched to a single point in the other signal, one can set $\alpha > 1$.

The distance itself is of little importance to us, but the warping function resulting in the minimal distance is, i.e. $\gamma(\cdot)$. This warping function gives us the correspondence between points from both signatures. Based on the warping function the signatures will be resampled in such a way that there is a unique point to point correspondence. Consider the left case of Fig. 3, where one point, $\mathbf{s}(t)$ of the signature is matched with multiple points in the other signature. First define $\mathbf{s}(t-1)$ and $\mathbf{s}(t+1)$ the last predecessor and first successor of $\mathbf{s}(t)$ with a unique point-to-point correspondence, e.g. $\mathbf{s}'(t-1)$ and $\mathbf{s}'(t+1)$ respectively. The point $\mathbf{s}(t)$ will be replaced by the number of points between $\mathbf{s}(t-1)$ and $\mathbf{s}(t+1)$: for each point $\mathbf{s}'(i)$ between $\mathbf{s}'(t-1)$ and $\mathbf{s}'(t+1)$ a new sample is added on $\mathbf{s}(\cdot)$. These new samples are calculated in such a way that

$$\frac{al(\mathbf{s}'(\cdot), t-1, i)}{al(\mathbf{s}'(\cdot), t-1, t+1)} = \frac{al(\mathbf{s}(\cdot), t-1, i)}{al(\mathbf{s}(\cdot), t-1, t+1)} \quad (3)$$

where $al(\mathbf{k}(\cdot), l, m)$ is the arc length from point l to m over the curve $\mathbf{k}(\cdot)$. The result of this resampling is shown on the right of Fig. 3.

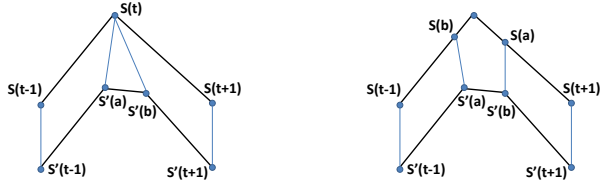


Fig. 3. An example of resampling the curves based on the warping function of DTW

IV. CELL MAPPING

Based on the previous contour matching we get for each point on the contour at frame t a unique point on the contour at frame $(t+1)$. For the mapping of points not on the contour, $\mathbf{p} = (x, y)$, we propose interpolation with polyharmonic splines:

$$\mathbf{p}_{t+1} = a + a_x x + a_y y + \sum_{j=1}^N \mathbf{w}_j \phi(d(\mathbf{p}_t - \mathbf{c}(j))) \quad (4)$$

where a, a_x, a_y and \mathbf{w}_j are a set of weighting coefficients, d is a distance function and $\phi(\cdot)$ is a radial basis function, i.e.

$$\phi(r) = \begin{cases} r^k, & \text{if } k \text{ is odd} \\ r^k \ln r, & \text{if } k \text{ is even} \end{cases} \quad (5)$$

where k is the degree of the basis functions. Note that $k = 2$ is a special case since (4) then results in the *Thin Plate Spline* [10]. However the distance function used in (4) shouldn't be the Euclidean distance as is done by Bookstein [10], since the Euclidean distance does not take the topology of the biological object into account. On the left of Fig. 4 an example of this problem is shown. The Euclidean distance between point A and B is the length of the dotted line. This is unaffected if the cell wall, i.e. the full blue line, intersects the dashed line. By ignoring the topology we overestimate points which might be close in Euclidean distance, but which can be far away in the nucleus topology. Instead we propose the use of the geodesic distance: the length of the shortest path not passing the cell wall is considered to be the geodesic distance between A and B. This is visualized on the right of Fig. 4.

The weighting coefficients in (4) are calculated in such a way that:

- points on the contour are mapped to the corresponding points according to the contour matching
- the weighting vector \mathbf{w} is orthogonal to \mathbf{x} and \mathbf{y} , i.e.

$$\sum_{i=1}^N w_{i,x} x_i = \sum_{i=1}^N w_{i,y} y_i = 0 \quad (6)$$

- the weighting coefficients sum to zero

These weighting coefficients can be found by solving a linear system of equations, which can be done in a fast way as proposed by Beatson et al. [11]. For a more detailed discussion on polyharmonic splines and the influence of the different radial basis functions we refer to [12]. For this

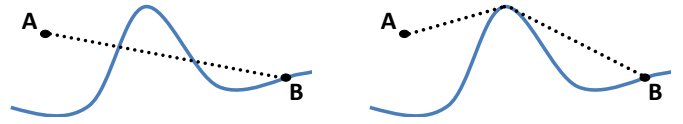


Fig. 4. An example of distances between point A and B, i.e. the length of the dotted line. The blue line represents a nucleus wall. In the left the Euclidean distance is visualised, where on the right the geodesic distance is visualised

application we propose to add an extra constraint to the splines: the displacement should be maximal at the contour of the nucleus, i.e. the derivative of the contour normal to the contour equals zero. Tests showed that without this constraint the displacement within the cell is strongly overestimated for non-convex nuclei. To achieve this constraint we add for each point on the contour an extra point on the normal of the curve at a short distance outside the nucleus.

V. RESULTS

Three data sequences were used for the validation of the proposed method. The datasets contain an isolated fibroblast cell nucleus of a patient with a homozygous nonsense mutation in lamin A/C gene. These are cells where the nucleus can have big deformations [3]. The datasets were captured using confocal controlled light-exposure microscopy [13], resulting in images of 288×288 pixels. These cells were transfected with a construct expressing a fluorescent fusion protein targeting the telomeres (TRF2-mCitrine). Non-bound protein diffuses freely throughout the nucleus and allows for delineating the boundaries for contour identification. For the validation datasets the telomeres were manually tracked by the author using MTrackJ. This annotation is done in 27 frames from three different cells, resulting in the detection of 272 telomeres. The centroids of these tracked telomeres will serve as ground truth for the validation. Note that the telomeres themselves might move locally, but this a small local motion, e.g. 1 to 2 pixels, which is neglectable compared to the big motion introduced by deforming nuclei.

In Fig. 5 an example of the mapping is shown. The red and blue curves are the contours of a nucleus in subsequent time frames. The red and blue dots are the centroids of the telomeres corresponding to respectively the red and blue contours. The green dots are the location of the "red" telomeres mapped with the proposed technique using the complex signature and a radial basis of degree 1. So the green dots are the prediction of the location of the blue dots. As can be seen, the mapping closely approximates the real telomeres' location in the next time frame.

Tests have been done for different signatures and for different radial basis functions used by the mapping. These tests were done using 100 samples for the signatures, tests showed that more samples didn't result in better results. In Table I the results are shown for the full test set. The first row of Table I the used basis function is shown. The first column shows the signature which is used. In this column RAW means the non processed data, i.e. the centroid coordinates

method	1st degree			2nd degree			3rd degree			4th degree		
	av. err.	σ^2	ICP err.	av. err.	σ^2	ICP err.	av. err.	σ^2	ICP err.	av. err.	σ^2	ICP err.
RAW	7.90	4.88	47	7.90	4.88	47	7.90	4.88	47	7.90	4.88	47
Complex	6.23	3.52	31	6.96	3.80	41	6.50	3.52	30	6.52	3.64	29
Centroid distance	5.75	3.99	27	6.30	4.01	36	6.02	3.96	33	6.08	4.19	27
Direction	4.93	2.99	8	5.55	3.06	13	5.45	3.31	14	5.52	3.63	19
Curvature	5.24	3.16	10	5.68	3.38	14	6.10	4.73	20	5.56	3.51	14

TABLE I

ERROR MEASUREMENTS OF THE PROPOSED METHODS USING DIFFERENT SIGNATURES AND DIFFERENT RADIAL BASIS

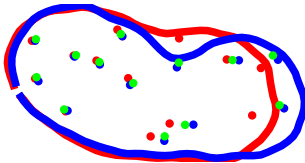


Fig. 5. An example of mapping telomeres from a frame to a subsequent time frame. The red and blue curves represent the nucleus contour in the subsequent frames. The red and blue dots represent the telomere centroids in the frame corresponding to respectively the red and blue curve. The green dots are the mapping of the red "telomeres" based on the warping of the red and blue contour.

are used as they are extracted out of the images, without any mapping. So these values are independent of the radial basis function. For each basis function three error measurements are shown. First, the average Euclidean distance between the telomeres in the one frame and where the telomeres ought to be according to the mapping of the location of the telomeres in the previous frame. This distance or error is expressed in pixels. Second, the standard deviation of this error is shown. As a last measurement the telomeres of a frame were matched with the telomeres in the previous frame using the *Iterative Closest Point (ICP)* method. The amount of false matches for the full test set is shown. As can be seen in the results the biharmonic spline, i.e. radial basis of degree 1, performs best for all signatures. From the four signatures tested the "direction signature" gives the best result, just slightly better than the "curvature signature".

Some part of the remaining error might be due to the 2d mapping of a 3d deformation. However, big deformations of the nuclei mainly occur in the x and y direction and not in the z direction. So there might be a remaining error due to a 2d mapping instead of a 3d mapping, but the proposed mapping will cope with the most important, i.e. biggest, deformation.

It took on average 4.18s to calculate the mapping of 13 telomeres embedded in a contour of 100 samples. Approximately 90% of this time was spend on the calculation of the geodesic distance between points. There is no significant difference in computational time depending on the type of signature or on the type of basis function used. These time measurements were done using a matlab implementation run on an intel i7 processor with 4Gb memory.

VI. CONCLUSION

In this paper a new mapping algorithm was proposed for the mapping of cell nuclei. This mapping does not impose

any shape constraints, which allows it to map cell nuclei of pathological cells. The proposed technique works in two steps. First the nucleus contours are matched. Four different matching techniques are defined and tested. Based on this contour matching the full nucleus is mapped using polyharmonic interpolation, radial basis up to degree four have been tested. The proposed method was tested on three real datasets showing significant improvement over the unprocessed data. The best results were achieved using a polyharmonic spline of degree one, in combination with the "direction" signature.

REFERENCES

- [1] V. Raz, B. J. Vermolen, Y. Garini, *et al.*, "The nuclear lamina promotes telomere aggregation and centromere peripheral localization during senescence of human mesenchymal stem cells," *Journal of Cell Science*, vol. 121, no. 24, pp. 4018–4028, 2008.
- [2] S. F. Louis, B. J. Vermolen, Y. Garini, *et al.*, "c-myc induces chromosomal rearrangements through telomere and chromosome remodeling in the interphase nucleus," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 27, pp. 9613–9618, 2005.
- [3] W. H. De Vos, F. Houben, R. A. Hoebe, *et al.*, "Increased plasticity of the nuclear envelope and hypermobility of telomeres due to the loss of a-type lamins," *Biochimica et Biophysica Acta (BBA) - General Subjects*, vol. In Press, Corrected Proof, pp. –, 2010.
- [4] E. Gladilin, S. Goetze, J. Mateos-Langerak, *et al.*, "Shape normalization of 3d cell nuclei using elastic spherical mapping," *Journal of Microscopy-Oxford*, vol. 231, no. 1, pp. 105–114, 2008.
- [5] N. Otsu, "A threshold selection method from gray level histograms," *IEEE transactions on Systems, Man and Cybernetics*, vol. 9, pp. 62–66, 1976.
- [6] P. K. Yan, X. B. Zhou, M. Shah, and S. T. C. Wong, "Automatic segmentation of high-throughput rna fluorescent cellular images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 1, pp. 109–117, 2008.
- [7] F. Cloppet and A. Boucher, "Segmentation of complex nucleus configurations in biological images," *Pattern Recognition Letters*, vol. In Press, 2010.
- [8] V. V. Kindratenko, "On using functions to describe the shape," *Journal of Mathematical Imaging and Vision*, vol. 18, no. 3, pp. 225–245, 2003.
- [9] D. S. Zhang and G. J. Lu, "Review of shape representation and description techniques," *Pattern Recognition*, vol. 37, no. 1, pp. 1–19, 2004.
- [10] F. L. Bookstein, "Principal warps - thin-plate splines and the decomposition of deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 6, pp. 567–585, 1989.
- [11] R. K. Beaton, M. J. D. Powell, and A. M. Tan, "Fast evaluation of polyharmonic splines in three dimensions," *IMA Journal of Numerical Analysis*, vol. 27, pp. 427–450, 2007.
- [12] N. Arad, N. Dyn, D. Reissfeld, and Y. Yeshurun, "Image warping by radial basis functions: applications to facial expressions," *CVGIP: Graph. Models Image Process.*, vol. 56, no. 2, pp. 161–172, 1994.
- [13] W. H. De Vos, R. A. Hoebe, G. H. Joss, *et al.*, "Controlled light exposure microscopy reveals dynamic telomere microterritories throughout the cell cycle," *Cytometry Part A*, vol. 75A, no. 5, pp. 428–439, 2009.