E-Cadherin Radial Distribution Characterization for Mutation Detection Purposes*

Sofia Esménio[1], Joana Figueiredo[2], Raquel Seruca[2], and J. Miguel Sanches[1,**]

[1] Institute for Systems and Robotics
Department of Bioengineering,
Instituto Superior Técnico, Technical University of Lisbon
jmrs@ist.utl.pt
[2] IPATIMUP - Institute of Molecular Pathology and Immunology of the University of Porto, Portugal
http://www.isr.ist.utl.pt/ jmrs

Abstract. Structural and mechanical properties of the tissues are dependent on the physical linkage between cells. E-Cadherin is a key component on this adhesion mechanism and mutations on its coding gene may produce dysfunctional molecules that compromise the cell-cell linkage and increase the risk of cancer.

The stationary distribution of E-Cadherin is characterized by a clear increased concentration at the membrane where it plays its adhesion role. However, for mutated molecules, the traffic dynamics of E-Cadherin is disturbed and different distributions of E-Cadherin across the cell are observed.

In this work a computational tool is proposed to semi-automatically help in the segmentation of cells from microscopy images of fluorescence with tagged E-Cadherin and to compute an image of radial profiles of the molecule distribution from the center of the cell toward the membrane.

The image of radial intensity profiles of E-Cadherin distribution depend on the location of the nucleus and on the specific geometry of each cell which is not related with the functional role of the molecules.

In this paper the radial profiles are geometrically compensated, to cope with shape and size differences among cells, and a representative profile of the tissue is obtained for mutation detection purposes.

Examples with real microscopy images of fluorescence of epithelial cells of the stomach are presented to illustrate the method.

1 Introduction

In epithelia, cell-cell adhesion is achieved by the establishment of homophilic interactions between two adjacent cells. One of the pivotal molecules to attain this homeostatic interaction between cells is the correct localization and function of the transmembrane protein E- cadherin and its interaction with other members of the adhesion complex [1].

* This work was supported by the FCT project [PEst-OE/EEI/LA0009/2011].
** Corresponding author.

J.M. Sanches, L. Micó, and J.S. Cardoso (Eds.): IbPRIA 2013, LNCS 7887, pp. 173–180, 2013.
© Springer-Verlag Berlin Heidelberg 2013

In cancer, somatic changes in the expression or function of E-cadherin have been implicated in all steps of tumour progression, including detachment of tumour cells from the primary site, invasion of the adjacent tissue, intravasation into the blood stream, extravasation into distant target organs, and formation of secondary lesions or metastasis [2].In hereditary forms of gastric cancer germline E-cadherin mutations are causing events [3] In this lethal cancer disease patients develop highly invasive isolated carcinomas that still pose a very important clinical problem since no effective tools of image screening (endoscopy, MRI and PET) are available yet.

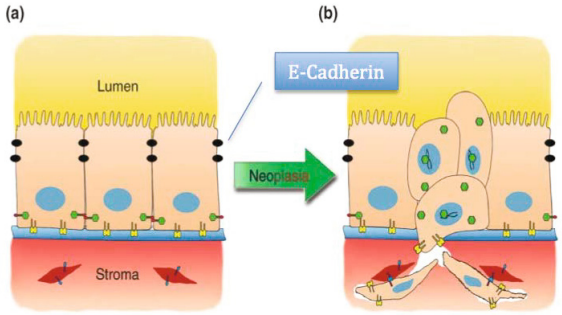

Fig. 1. a)Healthy tissue (WT E-cadherin) b)Neoplasic tissue (mutated E-cadherin)

The distribution of WT E-Cadherin molecules in cells is mainly observed at the membrane where they play its role in cell-cell adhesion. By this, in cells with this type of functional E-Cadherin the intensity of imunofluorescence images is stronger at the membrane. Additionally, the distribution of the molecules in normal cells at the cytoplasm present a characteristic uniform distribution related with the stable traffic pattern of the non mutated molecules. On the contrary, the E-cadherin in mutated cells follow a different distribution. Large concentrations in the cytoplasm or lack of E-cadherin signals in some locations are common features for almost all mutations that are associated with loss of cell-cell adhesion.This distribution, depending on the functional characteristics of the molecule, is also dependent on the specific shape of the cells and on the position of the nucleus within the cell, that is not always at the center of the cell.

Geometric compensation is a common procedure in several image modalities, mainly for registration purposes [4]. The main goal is making it possible the comparison and alignment of objects with a wide range of shapes and sizes.

The general strategy in this type of algorithms consists in the estimation of a geometric transformation, rigid or non-rigid [5], by optimizing a metric of similarity in order to make the objects under alignment as similar as possible from shape and size points of view [6].

Geometric compensation in images of microscopy is mainly related with segmentation and tracking [7,8]

In this paper a specific geometric compensation task is addressed in the scope of the problem of the E-Cadherin distribution characterization and quantification in the intracellular space, specially near the membrane, for mutation detection and discrimination purposes. A method for geometric deviation compensation of the cells shape from the perfect centred circular shape is proposed. The main goal is averaging a large number of radial profiles extracted from different cells from different plaques of each type of mutation and WT cells and obtain a typical E-cadherin radial profile to characterize each mutation.The method is performed in three main steps:i) Manual cell selection by the biologist and automatic estimation of the centres of the selected cells, ii) profile extraction along different angles (see Fig. 3.b)) and profile map building with these profiles as shown in Fig. 3.c) and finally iii) geometric compensation, which is the key issue of this paper. Typical images of imunofluorescence images used in this work are

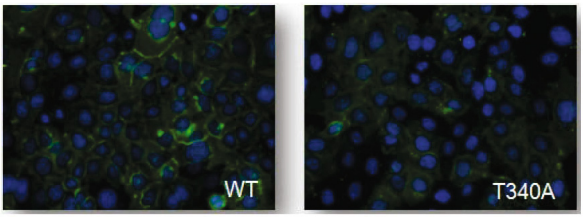

Fig. 2. Typical imunofluorescence images used in this work

displayed in Fig.2. The profile extraction procedure is illustrated in Fig.3 where it is clear the effects of the deviations from the ideal circular shape that would lead to an horizontal straight line representing the intensity at the membrane.

The proposed geometric compensation method is based on a novel approach where the image profiles are modelled as a continuous field in R^2 estimated from the original profiles adjusted along an iterative procedure. The algorithm, contrary to what usually happens, adjusts of observation locations driven by the minimization of an energy function.

2 Problem Formulation

The goal of this work is the creation of a standard intensity profile of the E-Cadherin distribution along radial directions of the cell nucleus center in microscope images of imunofluorescence for functional/dysfunctional discrimination purposes.

In the first step of the pre-processing procedure the relevant cells in each plate (see Fig. 2) are manually selected by the biologist by using a *graphic user interface* (GUI) developed to this project. This process is crucial for the success of the method because cells where the transfection [9] process did not perfectly occurred should not be taken into account for the profile definition.

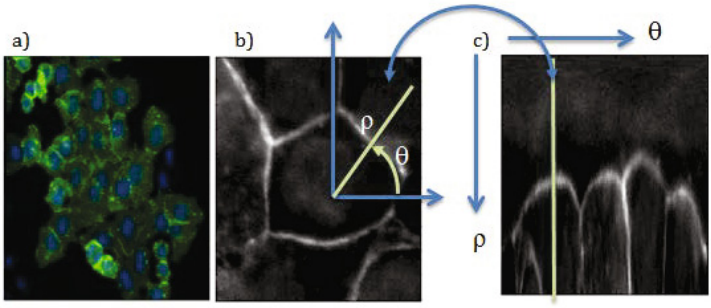

Fig. 3. Pre-processing procedure. a) Original image of imunofluorescence, b) nuclei segmentation and c) Cartesian image of polar profiles.

In a second step, radial profiles, from the centre of the nuclei, for different angles, θ, are extracted from the selected cells, as shown in Fig.3.b) and are organized in columns of a profile image, as shown in Fig.3.c). In an ideal radial distribution of the E-Cadherin, on a spheric cell, this profile image is angle invariant (across columns) presenting only information along the distance direction, ρ.

The third step of the method is the core of this work. In this step, the columns of the image of profiles are non rigidly distorted and adjusted in order to minimize the norms of consecutive columns to make them as similar as possible. This is an ill posed problem [10] and regularization is needed. An iterative algorithm optimizes an energy function where the original locations of the observed pixels are moved under constraints in to stabilize the iterative process and prevent divergence.

This iterative procedure is performed in two alternate steps: i) vector of coefficients estimation, $\hat{\mathbf{c}}_k$, describing each column as smoothed continuous functions under similarity constraints between columns and ii) observation location adjustments, $x_k(i)$, in a column wise basis, to the smooth functions estimated in the previous step.

Each column, representing each angle profile, is described by a finite dimension continuous function

$$f_k(x) = \sum_{i=0}^{N-1} c_k(i)\phi_k(x) \tag{1}$$

where $\phi_k(x)$ are triangular interpolating functions and $\mathbf{c}_k = \{c_k(i)\}$ is a vector of coefficients estimated by minimizing an energy function [11]

$$c_k = \arg\min_{c} E(\mathbf{x}_k, \mathbf{y}_k, \mathbf{c}_k) \tag{2}$$

where $E(\mathbf{x}_k, \mathbf{y}_k, \mathbf{c}_k) = E_y(\mathbf{x}_k, \mathbf{y}_k, \mathbf{c}_k) + E_p(\mathbf{c}_k)$. The *data fidelity term*, that pushes the solution toward the data, is $E_y(\mathbf{x}_k, \mathbf{y}_k, \mathbf{c}_k) = \sum_j (f_k(x_k(j)) - y_k(j))^2$

and the *prior* term used to stabilize the iterative process and smooth the solution is $E_p(\mathbf{c}_k) = \alpha \sum_{k,j} (c_k(j) - c_k(j-1))^2 = \alpha \|\mathbf{c}_k\|^2$ where α is the prior hyper parameter. Using matrix notation this energy function is

$$E(\mathbf{x}_k, \mathbf{y}_k, \mathbf{c}_k) = (\Phi\mathbf{c} - \mathbf{y})^T(\Phi\mathbf{c} - \mathbf{y}) + \alpha(\theta\mathbf{c})^T(\theta\mathbf{c}) \tag{3}$$

where Θ is a difference operator. The minimizer of (3) is computed by finding its stationary point, $\nabla_c \mathbf{E} = 0$,that leads to the following solution,

$$\mathbf{c}^0 = (\Phi^T\Phi + \alpha\Theta)^{-1}\Phi^T\mathbf{y}; \tag{4}$$

where $\Theta = \theta^T\theta$.

In the second step of the iterative process a prior term is added to the expression of the energy of the system with the purpose to force similarity between columns. The energy will become: $E(\mathbf{x}_k, \mathbf{y}_k, \mathbf{c}_k) = E_y(\mathbf{x}_k, \mathbf{y}_k, \mathbf{c}_k) + E_p(\mathbf{c}_k) + E_p(\mathbf{c}_k, \mathbf{c}_{k-1})$ where β is the new prior hyper parameter and $E_p(\mathbf{c}_k, \mathbf{c}_{k-1}) = \beta \sum_{k,j} (c_k(j) - c_{k-1}(j))^2$. Using matrix notation this energy function becomes:

$$\begin{aligned} E(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_k (\Phi\mathbf{c} - \mathbf{y})^T(\Phi\mathbf{c} - \mathbf{y}) + \alpha(\theta\mathbf{c}_k)^T(\theta\mathbf{c}_k) + \\ \beta(c_k - c_{k-1})^T(c_k - c_{k-1}) \end{aligned} \tag{5}$$

The minimization of (5), results in:

$$\Phi^T(\Psi\mathbf{c}_k - \mathbf{y}_k) + \alpha\Theta\mathbf{c}_k + \beta(c_k - c_{k-1}) + \beta(c_k - c_{k+1}) = 0 \tag{6}$$

which can be rewritten as:

$$\Sigma(x_k)c_k - \beta\mathbf{C}P_k = \Phi^T(x_k)\mathbf{y}_k \tag{7}$$

where $\Sigma = (\Phi^T\Phi + \alpha\Theta + 2\beta I_N)$, $\beta(c_{k-1} - c_{k+1}) = \beta\mathbf{C}P_k$, $\mathbf{C}$ is the matrix containing in the k^{th} column the actual c_k and P_k the k^{th} column of a shift matrix. Rearranging the terms in order to calculate the c_k of each column we finally obtain:

$$\mathbf{c}_k = \Sigma(x_k)^{-1}(\beta\mathbf{C}P_k + \Phi^T(x_k)\mathbf{y}_k) \tag{8}$$

In the third and last step, the observations position will be adjusted after the computation of the vectors of coefficients, c_k. The energy function of this step is $E(\mathbf{x}_k, \mathbf{y}_k, \mathbf{c}_k) = E_y(\mathbf{x}_k, \mathbf{y}_k, \mathbf{c}_k) + E_p(\mathbf{x}_k)$ where $E_p(\mathbf{x}) = \gamma \sum_{j=1}^{M} (x_j - x_{j-1})^2$. In this case the minimization of this expression is with respect to $\mathbf{x}$, $\nabla_x \mathbf{E} = 0$, which leads to

$$2(f_k(x_r) - yr)\dot{f}_k(x_r) + 2(x_r - x_{r-1}) + 2(x_r - x_{r+1}) = 0 \tag{9}$$

Rearranging the terms, $x_k(i)$ can be calculated making:

$$x_k(i) = 1/2 \left(\frac{Z_k(i)}{\gamma} + (X_k^T * \Psi_k^T)^T \right) \tag{10}$$

where $\mathbf{Z}_k(i) = (f_k(x_k(i)) - y_k(i))\dot{f}_k(x_k(i))$, X_k is the actual k^{th} column of the matrix which collects the actual positions of the observations and Ψ_k the k^{th} column of a difference matrix operator. The iterative method of geometric compensation begin with the initialization of C^0 (see (4)) and afterwards alternates the recomputation of C^k forcing similarity between columns and the repositioning of the observations (see (10)).

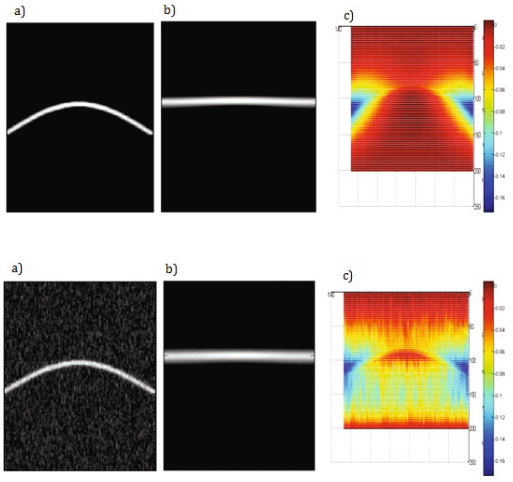

Fig. 4. Results with synthetic data. a) Radial profile, b) interpolated image after geometric compensation and c) displacement maps.

3 Experimental Results

The proposed algorithm was implemented in $Matlab_@$ and synthetic and real data were used for illustrative purposes.

3.1 Synthetic Data

In this experiment two 256×130 gray scale synthetic image containing a half arc of cosine with different levels of noise (see Fig.4 left) are used to illustrate the application of the proposed algorithm. These synthetic images aim to simulate the concentration of E-Cadherin at the membrane of non ideal spherical cells, where the radial profiles would be perfect horizontal lines. The results of geometric compensation applied to this images are shown in Fig. 4 in middle and right columns.

All the profiles were correctly compensated showing that the algorithm is stable for different intensities of background noise. The displacement maps (Fig.4, left column) show higher displacements in areas where the curvature is higher (extremes of the half cosine) as expected. Also the increase of noise increases the number of observations displaced due to local corrections which results in a denoising effect.

3.2 Real Data

The real data are images, displayed in Fig.2, of imunofluorescence of CHO cells, transfected [9] with vectors encoding the wild type E-cadherin, where the distribution of the molecules can be observed in the intra and inter cellular space and mainly at the membrane.

For illustrative purposes a single radial profile created from a region of interest of a fluorescence microscopy image (see Fig. 6 a)).The results are displayed in Fig.5 and Fig.6.

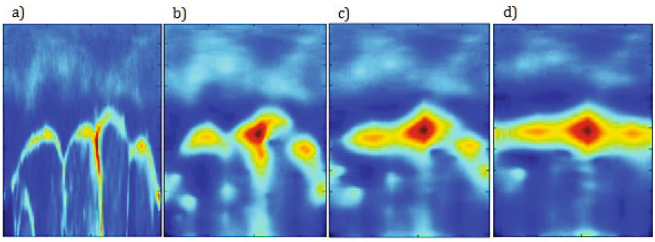

Fig. 5. Geometric compensation algorithm.a) Initial radial profile and compensated image profile after (b) 300, c) 700 and d) 1000 iterations.

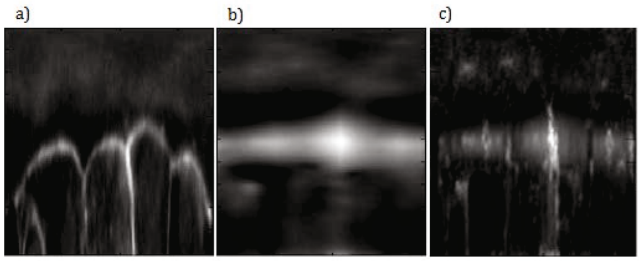

Fig. 6. a) Initial radial profile, b) real image after geometric compensation and c) real image obtained from repositioning of the initial observations, without interpolation

Fig. 5 shows the evolution of the image during the algorythm of geometric compensation. The image is rearranged in order to minimize the distance between observations with similar intensity creating a peak in the middle of the image, as expected. Fig. 6 b) is the grayscale result of the geometric compensation of 6 a) radial profile. The result in 6 c) is the imaged recriated with the displacement of the original obervations (and the interpolation of the empty remained empty pixels). This result also shows a peak of intensity in the middle of the image in coherence with what was expected.

4 Conclusion

In this paper a semi-automatic segmentation procedure of cells from fluorescense microscopy images with tagged E-Cadherin, is described as well as the algorithm to extract the radial intensity profiles of them for distribution characterization purposes.These profiles are affected by geometric differences between cells that are not perfect spheres, and are confound factors in the process of E-Cadherin mutation detection and discrimination.

In this paper an algorithm to compensate for these differences is described. The method is based on the estimation of a continuous field in R^2 based on moving observations driven by the minimization of an energy function containing specific priors that regularize the estimated field along the distance to the centroid of the cell. The presented algorithm of geometric compensation can be applied for instance in cases of biological variety in geometry. A future goal is to denoise the images to increase the stability range of this method.

References

1. Simoes-Correia, J., Figueiredo, J., Lopes, R., Stricher, F., Oliveira, C., Serrano, L., Seruca, R.: E-cadherin destabilization accounts for the pathogenicity of missense mutations in hereditary diffuse gastric cancer. PloS One 7(3) (2012)
2. Paredes, J., Figueiredo, J., Albergaria, A., Oliveira, P., Seruca, R.: Epithelial e- and p cadherins: Role and clinical significance in cancer. Biochimica et Biophysica Acta 1826 (2012)
3. Suriano, G., Seixas, S., Rocha, J., Seruca, R.: A model to infer the pathogenic significance of cdh1 germline missense variants. J. Mol. Med. (84) (2006)
4. Fonseca, M., Manjunath, C.: Registration techniques for multisensor remotely sensed imagery. Photogrammetric Engineering and Remote Sensing 14, 1049–1056 (1996)
5. Zitova, B., Flusser, J.: Image registration methods: a survey. Image and Vision Computing 21(11), 977–1000 (2003)
6. Sanches, J.M., Marques, J.S.: Joint image registration and volume reconstruction for 3d ultrasound. Pattern Recognition Letters 24(4-5), 791–800 (2003)
7. Li, S., Wakefield, J., Noble, J.: Automated segmentation and alignment of mitotic nuclei for kymograph visualisation. In: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, March 30-April 2, pp. 622–625 (2011)
8. Dzyubachyk, O., Essers, J., van Cappellen, W.A., Baldeyron, C., Inagaki, A., Niessen, W.J., Meijering, E.: Automated analysis of time-lapse fluorescence microscopy images: from live cell images to intracellular foci. Bioinformatics 26(19), 2424–2430 (2010)
9. Bonetta, L.: The inside scoop-evaluating gene delivery methods. Nature Methods 2, 875–883 (2005)
10. Vogel, C.R.: Computational methods for inverse problems. Frontiers in Applied Mathematics (2002)
11. Moon, T., Stirling, W.: Mathematical methods and algorithms for signal processing. Prentice-Hall (2000)