

CELL NUCLEI SEGMENTATION BY LEARNING A PHYSICALLY BASED DEFORMABLE MODEL

Marina E. Plissiti, Christophoros Nikou

Department of Computer Science, University of Ioannina, Ioannina, Greece
marina@cs.uoi.gr, cnikou@cs.uoi.gr

ABSTRACT

In this work, we present an efficient framework for the training of active shape models (ASM), representing smooth shapes, which is based on the representation of a shape by the vibrations of a spring-mass system. A deformable model whose behavior is driven by physical principles is used on a training set of shapes. The boundary of the regions of interest of the elements of the training set is detected with the convergence of the physics-based deformable model and attributes of the shapes of interest are expressed in terms of modal analysis. Based on the estimated modal distribution, we develop a framework, similar in spirit to ASM, to detect and describe an unknown new shape. The main difference with respect to standard ASM is that the modal amplitudes of the learnt model are used instead of the 2D landmark points and the cost function to be minimized is accordingly modified. The proposed method is evaluated using cytological images of conventional Pap smears, which contain 44 nuclei of squamous epithelial cells.

Index Terms— Active shape models, physically based deformable model, modal analysis, shape priors, nuclei segmentation.

1. INTRODUCTION

Active shape models (ASM) [1] are parametric deformable models which are based on the construction of a statistical model of the global shape variation from a training set of shapes. They are extensively used for the recognition and localization of objects that follow the same geometric form of a sample of well known shapes. In the last years, several methods based on ASM have been proposed and they concern, among others, face detection [2], biomedical image segmentation [3] and handwritten character recognition [4].

Image segmentation with ASM requires the representation of the shape of the object of interest by a set of points. Based on this representation, a deformable model is iteratively deformed to fit to an instance of the object of interest in an unknown image. The model is constrained by the Point Distribution Model (PDM) [1] in order to vary

only in ways that are learnt in a training set of labelled examples.

In this work, we present a novel method for the segmentation of an image with ASM, which is based on the representation of an object using modal analysis [5]. Thus, a physical model is adopted in the training phase, in which the parameters to be learnt are the variations of the modes of the model [5], instead of the variations of the landmark points describing the shape of the model [1]. Therefore, a more compact description of the shape model is obtained.

Our method has been tested on a data set of 44 nuclei images of conventional Pap smear slides, acquired through an optical microscope. The segmentation of these images has been extensively studied by several researchers [6-9], as the nucleus is the structural part of the cell that presents significant changes when the cell is affected by a disease. The proposed method presents high performance compared to the ground truth.

The paper is organized as follows: In section 2, the physical model adopted in the training phase is described in detail. In section 3, the learning procedure of the physical model is introduced. In section 4, experimental results and a discussion of the proposed method are presented and the derived conclusions are drawn in section 5.

2. PHYSICAL MODEL

In order to obtain a more compact representation of the shape of the nuclei boundary, we adopt the physical deformable model proposed by Nastar and Ayache [5]. Thus, in our experiments, a physics based deformable model is used, whose behaviour is controlled by the governing equations of motion. More specifically, the physical model consists of N virtual masses located at points V_1, V_2, \dots, V_N . The motion of the physical model towards the border of the object of interest is expressed by a finite element formulation and is estimated by solving a $2N$ -dimensional differential matrix equation (for the horizontal and vertical direction):

$$\begin{cases} \mathbf{M}\ddot{\mathbf{U}}_x(t) + \mathbf{C}\dot{\mathbf{U}}_x(t) + \mathbf{K}\mathbf{U}_x(t) = \mathbf{F}_x(t) \\ \mathbf{M}\ddot{\mathbf{U}}_y(t) + \mathbf{C}\dot{\mathbf{U}}_y(t) + \mathbf{K}\mathbf{U}_y(t) = \mathbf{F}_y(t) \end{cases} \quad (1)$$

where \mathbf{M} , \mathbf{C} and \mathbf{K} are $N \times N$ matrices describing the mass, the damping and the stiffness of the model. Moreover, \mathbf{F}_x and \mathbf{F}_y are vectors containing the image force at the nodes locations and \mathbf{U}_x , $\dot{\mathbf{U}}_x$, $\ddot{\mathbf{U}}_x$ and \mathbf{U}_y , $\dot{\mathbf{U}}_y$, $\ddot{\mathbf{U}}_y$ are the vectors of displacement, velocity and acceleration of the model in the horizontal and vertical direction respectively.

The above equations describe the equilibrium between internal and external forces of the system. The internal forces are expressed by the definition of the virtual masses of the model and the interaction between them, while the external forces are usually defined as the intensity or the gradient of the image at the pixels where the nodes of the model are located.

The system (1) can be solved by setting the initial values of displacement and velocity equal to zero and then using an explicit Euler scheme, which requires the calculation of the inverse matrix \mathbf{M}^{-1} :

$$\begin{cases} \ddot{\mathbf{U}}(t) = \mathbf{M}^{-1}(\mathbf{F}(t) - \mathbf{C}\dot{\mathbf{U}}(t) - \mathbf{K}\mathbf{U}(t)) \\ \dot{\mathbf{U}}(t + \Delta t) = \dot{\mathbf{U}}(t) + \Delta t \ddot{\mathbf{U}}(t) \\ \mathbf{U}(t + \Delta t) = \mathbf{U}(t) + \Delta t \dot{\mathbf{U}}(t + \Delta t) \end{cases} \quad (2)$$

where Δt is the time step. However, instead of solving directly the equilibrium equation (1), we can use a frequency based technique called modal analysis, which describes the motion of the model in terms of the free vibrations of the system.

More specifically, at a first step the following change of basis is used [10]:

$$\mathbf{U} = \mathbf{\Phi}\tilde{\mathbf{U}} \quad (3)$$

where $\mathbf{\Phi}$ is a square non-singular matrix and $\tilde{\mathbf{U}}$ is the vector of the generalized displacement. The columns of the matrix $\mathbf{\Phi}$ are selected to be the eigenvectors of the generalized eigenproblem:

$$\mathbf{K}\phi = \omega^2 \mathbf{M}\phi \quad (4)$$

where ϕ is the i -th mode and ω_i its frequency. This is an effective way for the expression of the displacement vector \mathbf{U} in terms of modal superposition, that is:

$$\mathbf{U} = \mathbf{\Phi}\tilde{\mathbf{U}} = \sum_{i=1}^N \tilde{u}_i(t) \phi \quad (5)$$

where \tilde{u}_i the amplitude of the i -th mode. It can be shown

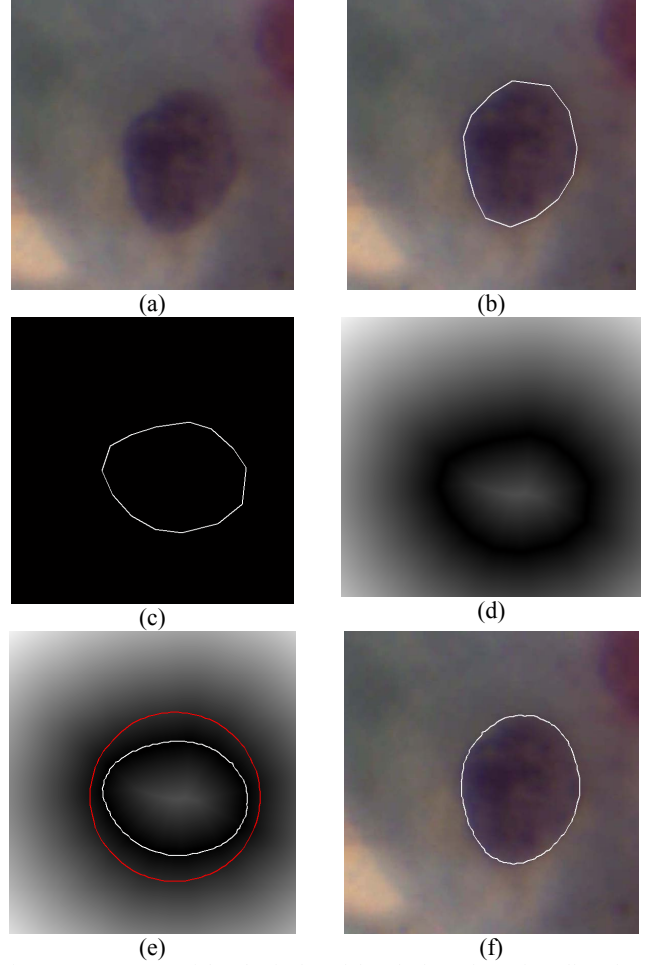


Fig. 1: Convergence of the physical model to the boundary of a cell nucleus in the training phase. (a) The initial image, (b) the manually traced nucleus boundary, (c) registration of the boundary to a reference shape oriented horizontally, (d) the distance transform of (c), (e) the initial (red) and the final (white) position of the physical model, (f) the detected nucleus boundary in the initial image rotated appropriately.

[10] that the matrices \mathbf{K} , \mathbf{M} and \mathbf{C} are simultaneously diagonalized:

$$\begin{cases} \mathbf{\Phi}^T \mathbf{M} \mathbf{\Phi} = \mathbf{I} \\ \mathbf{\Phi}^T \mathbf{K} \mathbf{\Phi} = \mathbf{\Omega}^2 \end{cases} \quad (6)$$

where \mathbf{I} is the identity matrix and $\mathbf{\Omega}^2$ is the diagonal matrix whose elements are the eigenvalues $\omega_i^2, i = 1, \dots, N$.

Premultiplying eq. (1) by $\mathbf{\Phi}^T$ and substituting the displacement vector with its equivalent form of eq. (3) leads to:

$$\ddot{\tilde{\mathbf{U}}} + \tilde{\mathbf{C}}\dot{\tilde{\mathbf{U}}} + \mathbf{\Omega}^2 \tilde{\mathbf{U}} = \tilde{\mathbf{F}} \quad (7)$$

where $\tilde{\mathbf{C}} = \mathbf{\Phi}^T \mathbf{C} \mathbf{\Phi}$ and $\tilde{\mathbf{F}} = \mathbf{\Phi}^T \mathbf{F}$. The above matrix-form equations can be decoupled for each direction into N scalar equations of the form:

$$\ddot{\tilde{u}}_i(t) + \tilde{c}_i \dot{\tilde{u}}_i(t) + \omega_i^2 \tilde{u}_i(t) = \tilde{f}_i(t) \quad (8)$$

The solution of these equations in time t leads to the calculation of the amplitudes $\tilde{u}_i(t)$, $i=1, \dots, N$ and the deformation of the model is estimated using the modal superposition equation (eq. (5)). In each time step, the new positions of the nodes of the model $\mathbf{V}(t)$ are given by:

$$\mathbf{V}(t) = \mathbf{V}(t_0) + \mathbf{U}(t), \quad (9)$$

where $\mathbf{V}(t_0)$ is the vector containing the initial spatial positions of the model and $\mathbf{U}(t)$ is the nodal displacement vector.

In practice, the nodal displacements $\mathbf{U}(t)$ are approximated by $\hat{\mathbf{U}}(t)$ using a fraction of the modes, which present the highest amplitudes, that is:

$$\hat{\mathbf{U}}(t) = \sum_{i=1}^{\ell} \tilde{u}_i(t) \phi_i \quad (10)$$

where $\ell \ll N$. For the choice of the number of modes ℓ the total energy is calculated by:

$$E = \sum_{i=1}^{\ell} \tilde{u}_i^2 \quad (11)$$

and we choose the first ℓ amplitudes conforming a percentage of the total energy.

An issue that must be clarified is the calculation of the eigenvectors and eigenvalues of the generalized problem of eq. (4). From the classical theory of vibration of a crystal lattice, it can be proved that the relationship between spatial (k) and temporal (ω) frequencies is given by:

$$\omega^2(p) = \frac{4K}{M} \sin^2\left(\frac{k(p)\alpha}{2}\right) \quad (12)$$

In (12), due to the periodicity of the closed chain, it is given that:

$$k(p)\alpha = \frac{2p\pi}{N}, p \in \mathcal{B}(N) \quad (13)$$

where $\mathcal{B}(N)$ is the first Brillouin zone [5]:

$$\mathcal{B}(N) = \begin{cases} \left[-\frac{N}{2} + 1, \dots, \frac{N}{2} \right], & \text{for } N \text{ even} \\ \left[-\frac{N-1}{2}, \dots, \frac{N-1}{2} \right], & \text{for } N \text{ odd.} \end{cases} \quad (14)$$

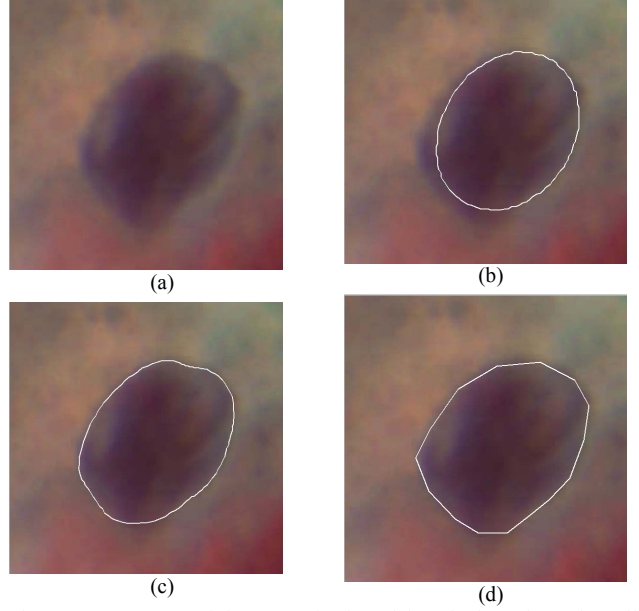


Fig. 2: Convergence of the constrained model to the boundary of a cell nucleus. (a) Initial image, (b) initial position of the mean shape model, (c) final position of the model after the convergence of ASM, (d) the ground truth. The figures are best viewed in colors.

Combining equations (12) and (13) we can calculate the temporal frequencies ω_i^2 , which correspond to the eigenvalues of eq. (4). The corresponding eigenvectors $\phi(p)$ are then given by:

$$\phi(p) = \left[\dots, \cos\left(\frac{2p\pi n}{N}\right), \dots \right]^T \quad (15)$$

Thus, using (15), analytic forms for the eigenvectors are obtained and the motion of the model can be easily expressed in terms of frequency modes.

3. LEARNING THE PHYSICAL MODEL

In our work, we follow the basic steps of the standard ASM algorithm, which aims to match a trained model to a new unknown image. However, instead of describing the object of interest by a set of n labelled landmark points, we focus on the learning of the generalized displacements \tilde{u}_i of the model, in each image of the training set. This is an equivalent representation, since the combination of (3) and (9) results in the spatial coordinates of the shape.

In the training phase, the nuclei boundaries were manually traced by an expert (Fig. 1(b)) in all the images of the training set. An issue that must be taken into account for the correct training of the model is that the shapes in each image must be registered. Given the fact that the nuclei generally follow an ellipse-like shape, we registered all the manually traced shapes with a reference shape having its

major axis horizontally oriented (Fig. 1(c)). Based on this boundary, the distance transform was estimated for every image (Fig. 1(d)). On the resulted image, a physical model was initialized and deformed until convergence (Fig. 1(e)), in order to detect the desired boundary. As a result, an accurate nucleus boundary was obtained (Fig. 1(f)).

From the final shape of the model, the generalized displacement vector $\tilde{\mathbf{U}}$ was estimated and from the entire training set, the mean $\tilde{\mathbf{U}}$ was calculated, which entails in the representation of the mean shape of the nucleus boundary. More specifically, given a set of L vectors $\tilde{\mathbf{u}}_i$, the mean is calculated as:

$$\tilde{\mathbf{u}} = \frac{1}{L} \sum_{i=1}^L \tilde{\mathbf{u}}_i \quad (16)$$

and the covariance matrix as:

$$\mathbf{S} = \frac{1}{L-1} \sum_{i=1}^L (\tilde{\mathbf{u}}_i - \tilde{\mathbf{u}})(\tilde{\mathbf{u}}_i - \tilde{\mathbf{u}})^T \quad (17)$$

Using principal component analysis (PCA), the eigenvectors \mathbf{a}_i of the covariance matrix \mathbf{S} are used for an equivalent representation of the shape, that is

$$\tilde{\mathbf{u}} = \tilde{\mathbf{u}} + \mathbf{A}\mathbf{b}, \quad (18)$$

where \mathbf{A} is the matrix with columns the eigenvectors \mathbf{a}_i and \mathbf{b} is a vector containing the model coordinates in the eigenvector basis, computed by

$$\mathbf{b} = \mathbf{A}^T (\tilde{\mathbf{u}} - \tilde{\mathbf{u}}). \quad (19)$$

If we take into account the j eigenvectors which correspond to the j largest eigenvalues of the covariance matrix, the shape can be approximated by

$$\tilde{\mathbf{u}} = \tilde{\mathbf{u}} + \mathbf{A}_j \mathbf{b}_j \quad (20)$$

where \mathbf{A}_j and \mathbf{b}_j are derived as described above using only the j selected eigenvectors.

With this shape representation, the algorithm fits the desired model in the image, driven by the image characteristics described by the image function and the prior training. Thus, given an initial position of the shape of the model in an unknown image and the image based term f (i.e. $f(x, y) = -|\nabla I(x, y)|^2$), the ASM algorithm repeats the following steps until convergence:

1) Compute the gradient of the cost function

$$\frac{df}{d\mathbf{x}} = \left[\frac{\partial f(x_1, y_1)}{\partial x}, \frac{\partial f(x_1, y_1)}{\partial y}, \dots, \frac{\partial f(x_n, y_n)}{\partial x}, \frac{\partial f(x_n, y_n)}{\partial y} \right]^T \quad (21)$$

2) Update frequency vibration modes:

$$\tilde{\mathbf{u}}^{(t+1)} = \tilde{\mathbf{u}}^{(t)} - \tau \sum_{i=1}^N \left(\mathbf{a}_i^T \frac{df}{d\mathbf{u}} \right) \mathbf{a}_i. \quad (22)$$

Notice that the initial value for shape $\tilde{\mathbf{u}}$ is the mean shape obtained with the training of ASM and τ is the time step. Given that

$$\frac{df}{d\tilde{\mathbf{u}}} = \frac{df}{d\mathbf{x}} \frac{d\mathbf{x}}{d\tilde{\mathbf{u}}} = \frac{df}{d\mathbf{x}} \Phi \quad (23)$$

and premultiplying eq. (22) by Φ to the left we get:

$$\Phi \tilde{\mathbf{u}}^{(t+1)} = \Phi \tilde{\mathbf{u}}^{(t)} - \tau \Phi \sum_{i=1}^N \left(\mathbf{a}_i^T \frac{df}{d\mathbf{x}} \Phi \right) \mathbf{a}_i \quad (24)$$

Regarding eq. (3) we get

$$\mathbf{U}^{(t+1)} = \mathbf{U}^{(t)} - \tau \Phi \sum_{i=1}^N \left(\mathbf{a}_i^T \frac{df}{d\mathbf{x}} \Phi \right) \mathbf{a}_i \quad (25)$$

and finally, from eq. (9) the local positions of the landmark points are calculated by:

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{U}, \quad (26)$$

where \mathbf{x}_0 is the initial position of the mean shape of the model in the image.

4. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed method was tested in terms of the accurate determination of the nuclei boundary on a data set of 44 images of single nuclei from conventional Pap stained cervical cell slides. The images were acquired through a CCD camera adapted to an optical microscope. We have used a 40x magnification lens and the acquired images were stored in JPEG format.

The method was tested in the entire image data set following the leave-one-out scheme. More specifically, in each image, we first calculate the mean learnt shape by taking into account the frequency modes of the physical model, when it is applied in all the other images of our data set (i.e. in 43 images). This shape was used as a first approximation of the nucleus border in the remaining image which constitutes the test image.

However, a problem that must be solved is the selection of the initial position of this shape in the image. This is a crucial step, since a prerequisite for the correct deformation of the active shape model, in order to produce accurate

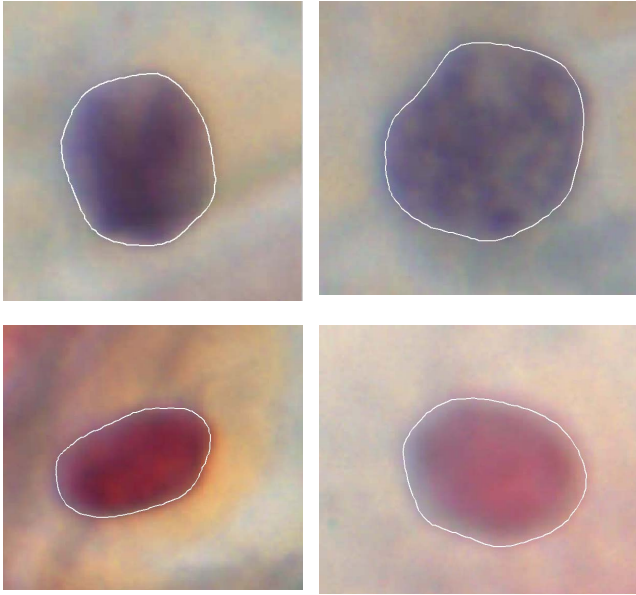


Fig. 3: Resulted nuclei boundaries. The figures are best viewed in colors.

results, is to be placed close to the actual boundaries of the object of interest.

Thus, an edge image was produced from the initial image, using the Canny edge detector in the corresponding grayscale image. In the obtained image, we search for the suitable position of the active shape model with registration by Chamfer matching [11] (Fig. 2(b)).

After the initialization of the physical based model, it was deformed according to equations (25) and (26) until it converged (Fig. 2(c)). This procedure was repeated 44 times for the entire image set, each time using a different training set and test image.

In order to estimate the performance of the proposed method, the Hausdorff distance of the final position of the deformable model and the ground truth was calculated in each image, and it was compared to the corresponding distance of the initial position of the model and the ground truth. Thus, the initial Hausdorff distance is 28.15 ± 14.42 (mean \pm std), and the final distance is 9.29 ± 3.55 . This implies that the resulted segmentation is closer to the manually traced nuclei. In Fig. 3, the results of our method in several images are depicted and as it can be observed the boundaries of the nuclei are accurately determined.

It must be noted that the parameters of the steps of the method were selected after several tests. Thus, for the calculation of modal vibrations (eq. (12)), we selected $K=5$ and $M=1$. Furthermore, in eq. (8) the value for all $\tilde{c}_i, i=1 \dots N$ is 1. For the initialization of the physical model, a circle of radius 85 pixels and 120 points in its circumference was used.

From eq. (10), we have calculated that the first 14 $\tilde{u}_i, i=1, \dots, 14$ constitute more than 99% of the total energy in each image of the training set. Thus, only 14 parameters

(instead of 120 landmark points) are sufficient for the accurate representation of the desired shape. Furthermore, after the application of PCA in these learnt parameters, only 4 eigenvectors which correspond to the highest eigenvalues lead to an almost exact shape representation, as they represent the 99,9% of the total energy. This clarifies that the proposed segmentation method provides a more compact shape representation which results in the reduction of the parameters to be learnt. Thus, the obtained nucleus boundary after the application of the ASM model is smooth and accurate.

5. CONCLUSION

We have developed a segmentation method combining the physically based model, which provides the compact representation of the shape of the object of interest, and the active shape model, which takes advantage of the *a-priori* knowledge of the expected shape. The method has been tested in terms of the accurate segmentation of the nuclei borders in images from Pap smear slides, and as it was verified by the results it presents a high performance.

6. REFERENCES

- [1] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Gaham. "Active shape models - their training and application", *Computer Vision and Image Understanding*, Vol. 61, No 1, pp. 38-59, 1995.
- [2] K. W. Wan, K. M. Lan and K. C. Ng, "An accurate active shape model for facial feature extraction", *Pattern Recognition Letters*, Vol 26, No 15, pp. 2409-2423, 2005.
- [3] B. van Ginneken, A. F. Frangi, J. J. Staal, B. M. ter Haar Romeny and M. A. Viergever, "Active shape model segmentation with optimal features", *IEEE Transactions on Medical Imaging*, Vol 21, No 8, pp. 924-933, 2002.
- [4] D. Shi, S. R. Gunn, R. I. Damper, "Handwritten Chinese radical recognition using nonlinear active shape models", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No.2, pp. 277-280, 2003.
- [5] C. Nastar and N. Ayache, "Frequency - based nonrigid motion analysis: application to four dimensional medical images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 18, No 11, pp. 1067-1079, 1996.
- [6] P. Bamford and B. Lovell, "Unsupervised cell nucleus segmentation with active contours", *Signal Processing*, Vol. 71, No. 2, pp. 203-213, 1998.
- [7] S. F. Yang-Mao, Y. K. Chan, Y. P. Chu, "Edge enhancement nucleus and cytoplasm contour detector of cervical smear images", *IEEE Transactions on System, Man and Cybernetics -Part B: Cybernetics*, Vol. 38, No 2, pp. 353-366, 2008.
- [8] M. E. Plissiti, C. Nikou and A. Charchanti, "Accurate localization of cell nuclei in Pap smear images using gradient vector flow deformable models", *Proceedings of 3rd International Conference on Bio-inspired Signals and Systems (BIOSIGNALS 2010)*, pp 284-289, 2010.

- [9] M. E. Plissiti, C. Nikou and A. Charchanti, "Watershed-based segmentation of cell nuclei boundaries in Pap smear images", *Proceedings of the 10th IEEE International Conference on Information Technology Applications in Biomedicine (ITAB'10)*, 3-5 November 2010, Corfu, Greece.
- [10] C. Nikou, G. Bueno, F. Heitz and J. P. Armspach, "A joint physics-based statistical deformable model for multimodal brain image analysis", *IEEE Transactions on Medical Imaging*, Vol. 20, No. 10, pp. 1026-1037, 2001.
- [11] G. Borgefors, "Hierarchical chamfer matching: a parametric edge matching algorithm", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 6, pp. 849-865, 1988.