

# Top-Down Segmentation of Histological Images Using a Digital Deformable Model

F. De Vieilleville<sup>1</sup>, J.-O. Lachaud<sup>1</sup>, P. Herlin<sup>2</sup>,  
O. Lezoray<sup>3</sup>, and B. Plancoulaine<sup>2</sup>

<sup>1</sup> Laboratoire de Mathématiques, UMR CNRS 5127  
Université de Savoie, 73776 Le-Bourget-du-Lac, France  
francois.de-vieilleville@univ-savoie.fr,  
jacques-olivier.lachaud@univ-savoie.fr

<sup>2</sup> GREYCAN, Centre François Baclesse  
Avenue du Général Harris 14076 Caen cedex 5, France  
p.herlin@baclesse.fr, b.plancoulaine@baclesse.fr

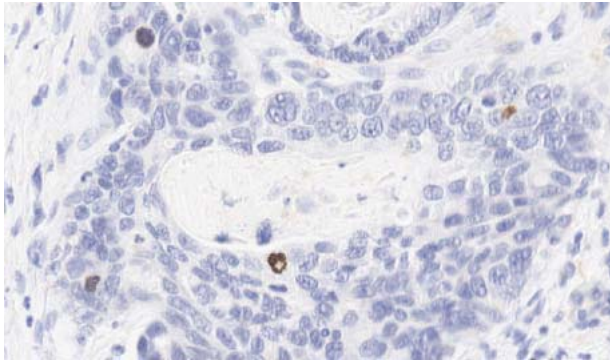
<sup>3</sup> GREYC 6 Boulevard du Maréchal Juin  
14050 Caen Cedex, France  
olivier.lezoray@unicaen.fr

**Abstract.** This paper presents a straightforward top-down segmentation method based on a contour approach on histological images. Our approach relies on a digital deformable model whose internal energy is based on the minimum length polygon and that uses a greedy algorithm to minimise its energy. Experiments on real histological images of breast cancer yields results as good as that of classical active contours.

## 1 Introduction

Breast cancer may be one of the oldest known forms of cancerous tumors in humans. Worldwide, breast cancer is the second most common type of cancer after lung cancer and the fifth most common cause of cancer death. Prognostic and diagnosis largely depend on the examination of stained tissue images by expert pathologists, which is time consuming and may lead to large variations. Therefore, it is essential to develop Image Decision Guided Systems to assist prognostic, diagnostic and early detection of cancer by automatically analyzing images of pathological tissue samples. One important prognostic factor for pathologist is the assessment of cellular proliferation by calculation of a mitotic grade [1]. To that aim, histological slides are stained by Immunohistochemistry that stains cells in proliferation in brown and other cells in blue. To establish an accurate mitotic grade, one has, in addition to mitosis detection, to properly detect cancer cells clusters (clusters of tumoral cells) to evaluate the mitotic grade only in tumor areas of the tissue. Fig 1 presents cancer cell clusters with mitotic and non mitotic cells.

In this paper, we focus on the segmentation of cancer cell clusters in breast histological images. In literature, such a task is performed with machine learning



**Fig. 1.** Cancer cell clusters with cells in proliferation stained in brown

methods [2,3]. Alternatively, we propose a top-down contour based approach using a digital deformable model [4]. On digital curves, using a direct analog of the classic energies of [5] is difficult for the expression of the geometrical quantities, mainly length and curvature estimation, suffer many drawbacks, see discussion in [6] about the curvature estimation. As a result, the internal energy term which usually monitors the smoothness of the curve does not behave as expected. Even when considering digital estimators that are asymptotically convergent on digital curves (see [7]), they often lead the evolution process to non-significant local minimum. As a result we here consider a digital deformable model that benefits of an internal energy based on the minimum length polygon (see [8,9]) yielding a convex functional. As a result, a descent on the internal energy ensures a global minimum. On an open 4-connected simple path this global minimum is very close to the simplest digital straight segment linking the endpoints of the path.

The paper is written as follow, we first recall the definition of the digital deformable model (Section 2). Later-on we elaborate on our Top-Down approach on histological images (Section 3). Experiments illustrate the use of our digital deformable model on real histology images, in particular we compare our model with another internal energy and a with one of the classical active contour formulation using a greedy algorithm for the energy minimisation (Section 4). Eventually, we conclude on the benefits of the proposed approach and possible future works.

## 2 Digital Deformable Model

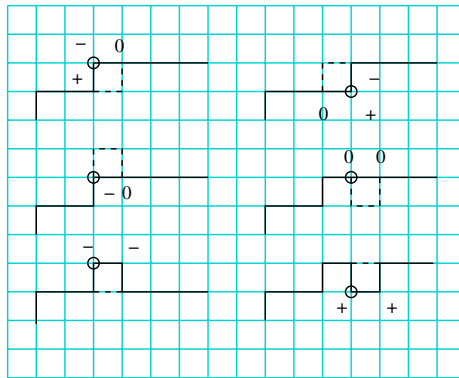
We here recall briefly the definition and main properties of a digital analog of active contours, (see [4]). The geometry of our model is that of an simple oriented digital 4-connected open path  $\Gamma$  with its endpoints being fixed, says  $A$  and  $B$ . The endpoints of the path are not allowed to be moved, while the remaining points of the curve can be deformed using elementary local transformations. Those separates into three types : *flips*, *bumps* and *flats*. All of these can only be applied on points with the proper corresponding geometry such as inside

or outside corners, flat parts or inside or outside bumps. Those features can be easily listed when the contour is read as a succession of straight moves, left moves and right moves. See Fig. 2 for an illustration of possible deformations on some specific contours. The admissible deformations of  $\Gamma$  are chosen such that they preserve its topology when applied and such that they are always reversible.

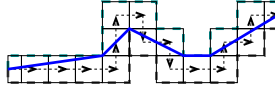
As an analog of active contour, the evolution of the model is monitored and highly dependant of the energy associated to  $\Gamma$ . This energy divides into two terms, one for the geometrical constrain over the curve (internal energy term  $E_{int}$ ) and one for the fit to the underlying datas (image energy term  $E_{image}$ ).

$$E_{DM}^D(\Gamma) = E_{int}^D(\Gamma) + E_{image}^D(\Gamma).$$

In the case of parametrized curves on the Euclidean plane constrain energies are usually based of the length and the integration of the normal and curvature along the curve, that is  $\int_0^1 E_{int}(v(s), v'(s), v''(s))ds$ . In digital space, it is difficult to find good curvature estimators on open contour, although there exists approaches which accurately estimate the curvature on closed digital contour such as the GMC in [10] (which is based on an optimisation scheme), or the Brunet-Malgouyres estimator in [11] (which uses binomial convolutions), these approaches are not suited to make a reliable estimation on the border of open curves. Moreover, as noticed by many authors (see for instance the geometric active contours of [12]), the internal energy is in fact the length of the curve. Our digital analog to  $E_{int}$  is therefore defined as the estimation of the length of  $\Gamma$ . This estimation is based on the euclidean length of the constrained minimum length polygon (CMLP for short) of  $\Gamma$ , that is the minimum length polygon in



**Fig. 2.** Example of some local deformations used by the digital deformable model, solid line for  $\Gamma$  and dashed line for  $\Gamma$  after the local deformation, the points where the deformation is applied are circled. The + symbol stands for a left turn, the - symbol stands for a right turn and the 0 symbol stands for a straight ahead move. From left to right and top to bottom : flip of an outside corner, flip of an inside corner, inside bump of a flat part, outside bump of a flat part, flat of an inside bump and flat of an outside bump.



**Fig. 3.** A digital open path  $\Gamma$  whose Freeman code word is EEEENNESESEENENE. In dashed lines we represent the one pixel band along  $\Gamma$ . The thick line is the constrained minimum length polygon of  $\Gamma$ .

a one pixel wide band along  $\Gamma$  passing through its constrain points. The first and last vertices of the *CMLP* are respectively the first and last point of  $\Gamma$ . See Fig. 3 for an illustration of the CMLP.

Thus our internal energy becomes:

$$E_{int}^D(\Gamma) = \alpha \mathcal{L}_2(CMLP(\Gamma))$$

One important property of this energy is that it is convex (see proof in [4]), that is, a descent on this energy always leads to a global minimum. This minimum is such that the CMLP of  $\Gamma$  is exactly the euclidean straight segment between the endpoints of  $\Gamma$ . The image energy term favors strong gradients and is defined as:

$$E_{image}^D(\Gamma) = \sum_{p \in \Gamma} \left( \max_{c \in Image} (\|\nabla I(c)\|) - \|\nabla I(p)\| \right),$$

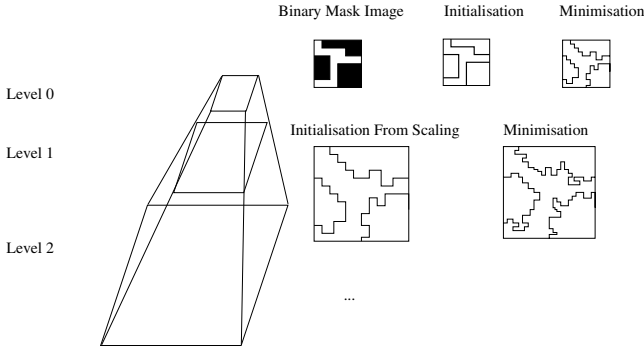
This energy term being positive everywhere, the tuning of the  $\alpha$  parameter is eased with respect to length term.

### 3 Top-Down Segmentation for Histology Images

The minimisation of the digital deformable model relies on a greedy approach (see Alg. 1) being costly because of it requires to compute many times the CMLP of  $\Gamma$ . As a consequence, a top-down approach seems particularly suited. We here consider three levels of resolution, the size of the image being multiplied by four at the next level. At each level there are two phases: an initialisation and a minimisation based on energy criteria.

At the first level, the initialisation is done via a binary mask image, from which we extract the initialisation paths. In fact we have several digital deformable modelsn to minimize at each level. Although we do not strictly preserve the topology of the initialisation, we prevent the paths to collide with one another. The initialisation at the other levels is obtained by scaling the paths resulting from the minimisation phase at the previous level, see Fig. 4 for an illustration.

The minimisation phase is straightforward: for a given path, we try all the deformations and apply the one that brings the smallest energy, we repeat this scheme as long as smaller energies can be found.



**Fig. 4.** The three levels used in our Top-Down approach

---

**Algorithm 1:** Greedy1 algorithm: extracts the deformation that brings the lowest energy among all possibles ones

---

```

1 Function Greedy1( In  $\Gamma$ , Out ) : boolean ;
   Input:  $\Omega$ : an Interpixel Partition Deformable Model
   Output:  $\Omega'$ : when returning true, elementary deformation of  $\Omega$  with
               $E_{DP}^D(\Omega') < E_{DP}^D(\Omega)$ , otherwise  $\Omega' = \Omega$  and it is a local minimum.
   Data:  $Q$  : Queue of (Deformation, double) ;
2  $E_0 \leftarrow E_{DM}^D(\Gamma)$ ;
3 foreach valid Deformation  $d$  on  $\Gamma$  do
4   |  $\Gamma$ .applyDeformation( $d$ );
5   |  $Q$ .push_back( $d, E_{DM}^D(\Gamma)$  );
6   |  $\Gamma$ .revertLastDeformation();
7 end
8 ( $d, E_1$ ) = SelectDeformationWithLowestEnergy ( $Q$ );
9  $\Gamma' \leftarrow \Gamma$ ;
10 if  $E_1 < E_0$  then  $\Gamma'$ .applyDeformation( $d$ );
11 return  $E_1 < E_0$ ;

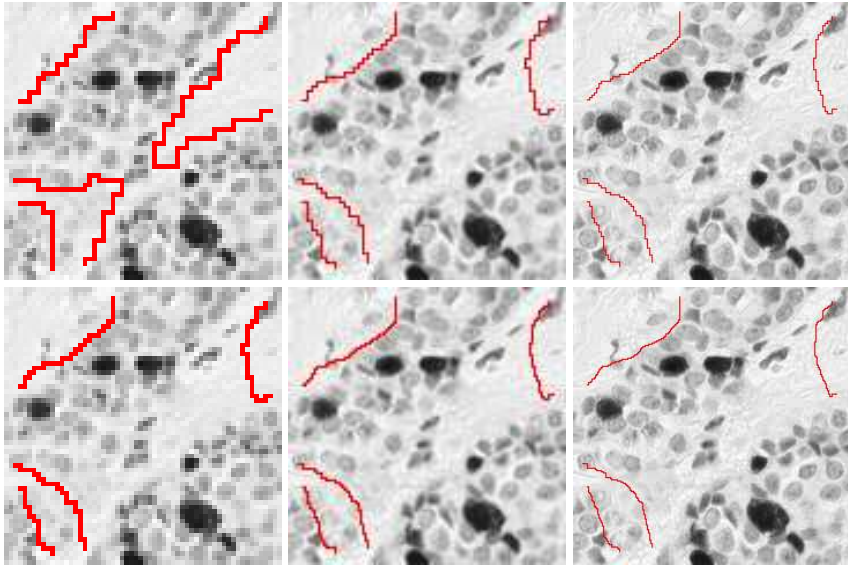
```

---

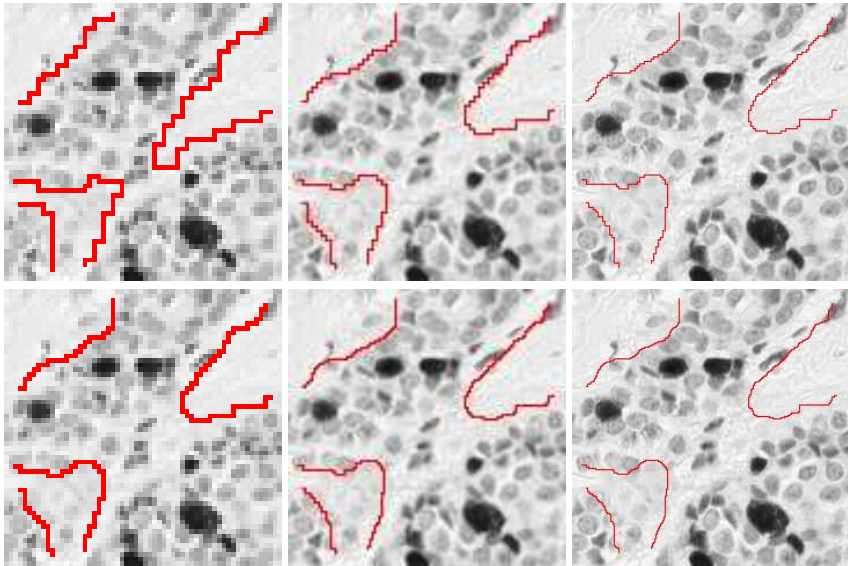
## 4 Experiments

The binary masks at low resolution are obtained with a coarse segmentation algorithm. The latter performs an automatic binary thresholding by entropy on a simplified version of the image with a morphological opening by reconstruction, see [13] for a similar method.

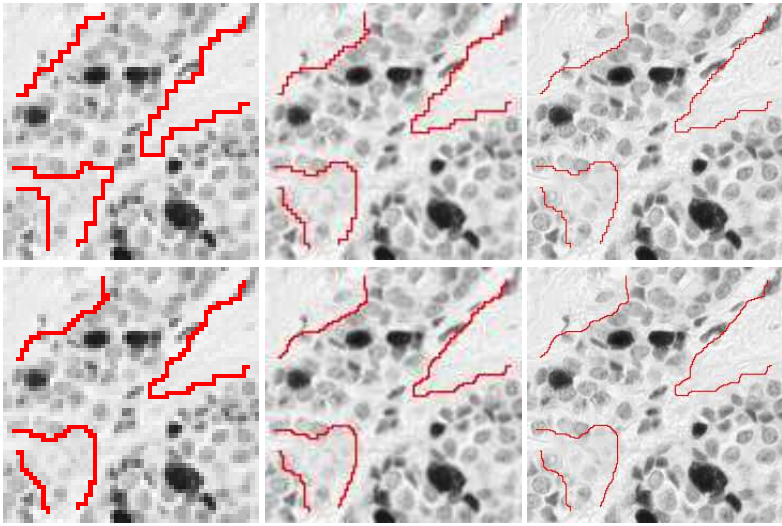
Our first examples use the constrained minimum length polygon as internal energy. As the model explicitly uses a weighting coefficient, we have run our experiments with various values, as shown on Fig. 5,6 and 7. This coefficient is such that the smaller, the less the internal energy monitors the deformation, consequently for  $\alpha$  equal to 400 the length penalisation overcome the data term.



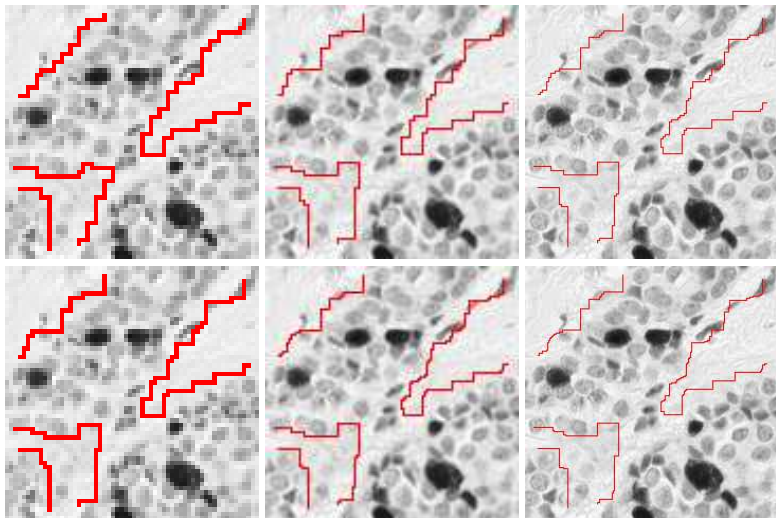
**Fig. 5.** Results of the top-down segmentation process using the value 400 as balance term, the internal energy term is based on the CMLP. From left to right and top to bottom: image at level 0,1 and 2 with initialisation, image at level 0,1 and 2 after minimisation. The high value of  $\alpha$  penalise the length of the contours, smoothing the contours too much, only top-left contour seems to be correctly delineated.



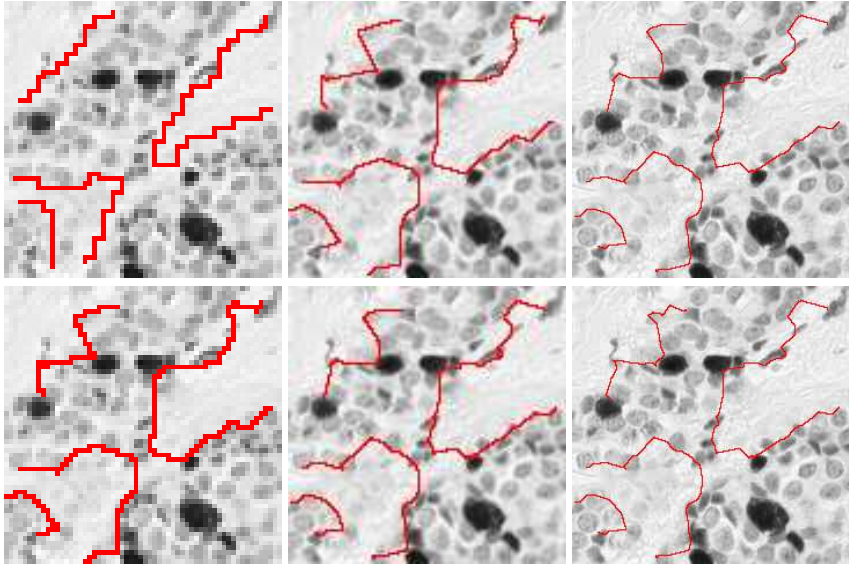
**Fig. 6.** Result of the segmentation using the value 150 as balance term, the internal energy term is based on the CMLP. From left to right and top to bottom: image at level 0,1 and 2 with initialisation, image at level 0,1 and 2 after minimisation. Top-left and bottom-left contours are correctly delineated.



**Fig. 7.** Results of the top-down segmentation process using the value 100 as balance term, the internal energy term is based on the CMLP. From left to right and top to bottom: image at level 0,1 and 2 with initialisation, image at level 0,1 and 2 after minimisation. Global delineation of contours is correct.



**Fig. 8.** Results of the top-down segmentation process using the value 100 as balance term, the internal energy term is based on the length of the freeman code of the path, other positive values for the balance term bring very similar results. From left to right and top to bottom: image at level 0,1 and 2 with initialisation, image at level 0,1 and 2 after minimisation. Although the global delineation seems correct, the obtained contours are not smoothed at all.



**Fig. 9.** Results of the top-down segmentation process using the deformable model described in [6]. Parameters are such that  $(\alpha, \beta)$  equal  $(0.1, 0.45)$ , the neighborhood is chosen as a  $3 \times 3$  square, and minimisation is done until no smaller energy can be found. From left to right and top to bottom: image at level 0,1 and 2 with initialisation, image at level 0,1 and 2 after minimisation. Top-right contour and middle contours are well delineated, top left and bottom left are partially correctly delineated.

In order to compare our results we have also run the same experiments using another internal energy. We have used a simpler internal energy which also leads to a global minimum. This internal energy is simply defined as the sum of the length of the freeman chain code constituting the path, that is the number of points of the path minus one. Using various positive coefficients as  $\alpha$ , results are very similar and the paths are much less smoothed than with the CMLP, as illustrated on Fig. 8.

Finally we have run the experiments using one of the classic active contour method in the literature, [6]. In this the curve is approached by a polygon whose vertices have integer coordinates. The minimisation process uses a greedy algorithm, which is closer to our optimisation scheme than a variational method. The quantity being minimized by this approach is:

$$E = \int (\alpha E_{cont} + \beta E_{curv} + \gamma E_{image}) ds.$$

At each iteration, and for each vertex a fixed-sized neighborhood around the point is considered and the point giving the smallest value becomes the new vertex. The energy terms are such that small values of  $E_{cont}$  causes the polygon points to be more equidistant and  $E_{curv}$  is the estimation of the curvature using



finite differences. The  $\gamma$  coefficient is chosen equal to one and  $E_{image}(p)$  is chosen as  $\max(\|\nabla I\|) - \|\nabla I(p)\|$  to ease the comparison with our approach. Between each level we also double the number of points of the model, obtained results are good considering that this method allows the end points of the polygon to move according to the minimisation process, as illustrated on Fig. 9.

## 5 Conclusion

In this paper we have used a top-down straightforward approach using a digital deformable model to segment cancer cluster cells in histological images. The behaviour of our digital deformable model was shown to behave as the classic continuous active contours and shown to give similar results. Future work will focus on integrating our digital deformable model with topological map (following [14]) so as to ensure that no topological changes may occur during the deformation of the contours. Once this step is achieved, top-down approaches on irregular pyramids such as [15] would be able to mix region and contour based segmentation process.

## References

1. Cheryl, E.: Assessment of cellular proliferation by calculation of mitotic index and by immunohistochemistry. In: *Metastasis Research Protocols. Methods in Molecular Medicine*, vol. 57, pp. 123–131. Humana Press (2001)
2. Doyle, S., Agner, S., Madabhushi, A., Feldman, M.D., Tomaszewski, J.: Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features. In: *ISBI*, pp. 496–499 (2008)
3. Signolle, N., Plancoulaine, B., Herlin, P., Revenu, M.: Texture-based multiscale segmentation: Application to stromal compartment characterization on ovarian carcinoma virtual slides. In: Elmoataz, A., Lezoray, O., Nouboud, F., Mammass, D. (eds.) *ICISP 2008* 2008. LNCS, vol. 5099, pp. 173–182. Springer, Heidelberg (2008)
4. de Vieilleville, F., Lachaud, J.O.: Digital deformable model simulating active contour. In: *Proceedings of 15th International Conference on Discrete Geometry for Computer Imagery* (accepted, 2009)
5. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. *International Journal of Computer Vision* 1, 321–331 (1988)
6. Williams, D.J., Shah, M.: A fast algorithm for active contours and curvature estimation. *CVGIP: Image Underst* 55, 14–26 (1992)
7. Lachaud, J.O., Vialard, A., de Vieilleville, F.: Fast, accurate and convergent tangent estimation on digital contours. *Image and Vision Computing* 25, 1572–1587 (2007)
8. Sloboda, F., Stoer, J.: On piecewise linear approximation of planar Jordan curves. *J. Comput. Appl. Math.* 55, 369–383 (1994)
9. Klette, R., Yip, B.: The length of digital curves. *Machine Graphics Vision* 9, 673–703 (2000); Also research report CITR-TR-54, University of Auckland, NZ (1999)
10. Kerautret, B., Lachaud, J.-O.: Curvature estimation along noisy digital contours by approximate global optimization. *Pattern Recognition* (in Press, 2009) (Corrected Proof)

11. Malgouyres, R., Brunet, F., Fourey, S.: Binomial convolutions and derivatives estimation from noisy discretizations. In: Coeurjolly, D., Sivignon, I., Tougne, L., Dupont, F. (eds.) DGCI 2008. LNCS, vol. 4992, pp. 370–379. Springer, Heidelberg (2008)
12. Caselles, V., Catta, F., Coll, T., Dibos, F.: A geometric model for active contours. *Numerische Mathematik* 66, 1–31 (1993)
13. Elie, N., Plancoulaine, B., Signolle, J., Herlin, P.: A simple way of quantifying immunostained cell nuclei on the whole histological section. *Cytometry* 56 A, 37–45 (2003)
14. Dupas, A., Damiand, G.: First results for 3d image segmentation with topological map. In: Coeurjolly, D., Sivignon, I., Tougne, L., Dupont, F. (eds.) DGCI 2008. LNCS, vol. 4992, pp. 507–518. Springer, Heidelberg (2008)
15. Goffe, R., Brun, L., Damiand, G.: A top-down construction scheme for irregular pyramids. In: Proceedings of the Fourth International Conference On Computer Vision Theory And Applications (VISAPP 2009), pp. 163–170 (2009)