AERIAL AND SATELLITE IMAGE SEGMENTATION BASED ON SHAPE-CONSTRAINED GEODESIC ACTIVE CONTOURS

Konstantinos Karantzalos, PhD student,
National technical University of Athens, Remote Sensing laboratory, Athens, Greece
Haris Papasaika, PhD student,
ETH, Swiss Federal Institute of Technology, Institute of Geodesy and Photogrammetry, Zurich, Switzerland
Nikos Paragios, PhD student,
MAS - Ecole Centrale de Paris, Chatenay-Malabry, Paris, France
karank@central.ntua.gr
haris@geod.baug.ethz.ch
nikos.paragios@ecp.fr

ABSTRACT

In this paper, we propose a novel mathematical model to address segmentation from an aerial and satellite imagery. Segmentation is performed through the consistent recovery of the zero iso-surfaces of a level set function towards image's foreground (desired object for extraction) and background discrimination. The level set method is embedded in a geodesic active contours variational formulation. Geodesic active contours technique is an advancement of the classical snakes and active contours that can handle the limitation to change their topology. The proposed mathematical model consists of a functional with two basic energies; one that forms a region-based energetic module for the evolving interface in the level set space and an another one that takes into account the a priori knowledge of the geometry of the desired for extraction objects. With such a functional shape information is embedded into the level set based segmentation scheme. Shape information can cope with missing or misleading information in the input images due to noise, clutter and occlusion. The shape priors were developed to segment objects of familiar shape in a given image. Promising results demonstrate the potentials of our approach, like in cases where the desired object for extraction were buildings.

INTRODUCTION

Variational image segmentation models are usually trying to generate segments by locally optimizing appropriate cost functionals defined on the space of contours (Paragios et al., 2005). The respective functionals are designed to maximize certain criteria regarding the low-level information such as edge consistency or (piecewise) homogeneity of intensity, color, texture, motion, or combinations thereof. After the introduction of active contours (snakes) by the pioneering work of Kass et al. (1987) semi-automatic methods on linear object extraction and road extraction were presented (Cohen, 1991; Gruen and Li, 1997). Later on Zafiropoulos and Schenk (1998) tackled the problem of embedding colour image information, coming from different channels in deformable models of contour type for the extraction and localization of road structures of small width and Jeon et al. (2000) used snakes to extract roads accurately. Other, recent, applications of the standard or more advanced deformable models can be found in (Laptev et al., 2000; Agouris et al., 2001; Ruther et al., 2002; Rochery et al., 2003; Bailloeul et al., 2005). The main limitation of deformable models, was their incapacity to change their topology (Paragios et al., 2005). The initial contour(s) can not split or merge.

The geodesic active contours along with the use of the level set technique were employed in order to solve the fixed topology problem of the classical snakes. Recently, level set method has been applied to satellite imagery for image segmentation (Samson et al., 2001; Ball and Bruce, 2005; Cao et al., 2005; Besbes et al., 2006). Karantzalos and Argialas (2006) proposed, also, the extraction of buildings from aerial and satellite imagery based on a level set image segmentation energy functional.

In all above cases, the variational energy functionals are aiming at maximizing certain criteria regarding the low-level information, but in practice the imposed models only roughly approximate the true intensity, texture or motion of specific objects in the image. Intensity measurements may be modulated by varying and complex lighting conditions.

Moreover, the observed images may be noisy and objects may be partially occluded. In such cases, algorithms which are purely based on low-level properties will invariably fail to generate the desired segmentation.

In recent years, it was suggested to enhance variational segmentation schemes by imposing (the desired for detection) object specific shape priors. Given one or more silhouettes of an object of interest, one can construct shape priors which favor objects that are in some sense familiar. Prior-based segmentation methods incorporate a representation of a reference shape within the energy functional. Thus, the recovered object boundary should resemble the expected contour, in addition to being constrained by length, smoothness and compatibility with the image gray levels and gradients. In order to introduce prior shape knowledge and a given group of transformations in the level-set formulation, a shape dissimilarity measure should be provided.

In this paper, a region based level set formulation, which was proposed for the detection of buildings (Karantzalos and Argialas, 2006), was extended with a building shape prior. This knowledge about the appearance of buildings was directly combined with clues given by the image data in order to cope with typical difficulties, like noise and occlusions.

REGION-BASED GEOMETRIC LEVEL SET SEGMENTATION

With the advantage of being implicit, intrinsic and parameter-free, level sets track moving interfaces through either model-free (Paragios and Deriche, 2002) or either model-based (Cremers, 2003) methods.

In this paper a model-free region-driven level set technique was implemented, similar to the energy functional proposed by (Paragios and Deriche, 2002; Karantzalos and Paragios, 2005; Karantzalos and Argialas, 2006), as their segmentation results was promising. The essence of this approach is to optimize the position and the geometric form of the curve by measuring information along that curve, and within the regions that compose the image partition. To this end, one can assume without loss of generality that objects and the background are uniform.

To this end, based on a motion equation that dictates the propagation of a closed structure, one can construct a structure of a higher dimension and define a corresponding flow such that its zero level set yields always to the position of the input structure. A step further is to consider the definition of the problem and the objective function (Zhao et al., 1996) directly on the space of level set representations. For a given open region Ω with smooth boundary, it is assumed the existence of a level set function $\phi(x,y)$ which is Lipschitz continuous. Towards this end, one can define the approximations of Dirac and Heaviside (Zhao et al., 1996) distributions:

$$\delta_a(\phi) = \begin{cases} 0 & , \quad |\phi| < a \\ \dfrac{1}{2a}\left(1+\cos\left(\dfrac{\pi\phi}{a}\right)\right) & , \quad |\phi| > a \end{cases}$$

$$H_a(\phi) = \begin{cases} 1 & , \quad \phi > a \\ 0 & , \quad \phi < -a \\ \dfrac{1}{2a}\left(1+\dfrac{\phi}{a}+\dfrac{1}{\pi}\sin\left(\dfrac{\pi\phi}{a}\right)\right) & , \quad |\phi| < a \end{cases} \qquad \text{Equation 1}$$

and use them to introduce an image partitioning objective function.

Boundary attraction as well region-consistency terms can be defined based on an evolving function ϕ. The geodesic active contour (Caselles et al., 1997) can be used for example to perform boundary extraction.

$$E_{boundary}(\phi) = \underbrace{\iint_{\Omega} \delta_a(\phi) b\left(|\nabla I|\right)|\nabla\phi|\, d\Omega}_{boundary\ \ module} \qquad \text{Equation 2}$$

where $b:\mathbb{R}^+ \to [1,0]$ is a monotonically decreasing function. The lowest potential of this functional corresponds to a minimal length geodesic curve attracted by the boundaries of the structure of interest.

Regional/global information can improve performance of boundary-based flows (Paragios and Deriche, 2002) that suffer of being sensitive to the initial conditions. The central idea behind a region-driven functional is to use the evolving interface to define an image partition that is optimal with respect to some grouping criterion. Within the level set representation such partition is natural according to the sign of the embedding function. The Heaviside function can be considered to define such partition:

$$E_{region}(\phi) = \iint_{\Omega} \left[H_a(\phi) \cdot r_{in}(I) \right] d\Omega + \iint_{\Omega} \left[\left(1 - H_a(\phi)\right) \cdot r_{out}(I) \right] d\Omega \qquad \text{Equation 3}$$

according to some region descriptor functions $r_{in}:\mathbb{R}^+ \to [1,0]$, $r_{out}:\mathbb{R}^+ \to [1,0]$ that are monotonically decreasing functions like:

$r_{in}(I) = \dfrac{\left(\mu_{in} - I\right)^2}{\sigma_{in}^2}$, μ_{in} : mean of the inside object's area, σ_{in} : covariance of the inside object's area

$r_{out}(I) = \dfrac{\left(\mu_{out} - I\right)^2}{\sigma_{out}^2}$, μ_{out} :mean of the background, σ_{out} : covariance of the background

Such descriptors measure the quality of matching between the observed image and the expected regional properties of the structure of interest and the background.

Integration of the boundary and the region-driven term can be considered to perform segmentation (Paragios and Deriche, 2002), namely the geodesic active region model. In the absence of noise, occlusions and corrupted visual information, such method can deal with local deformations. On the other hand, it cannot account for prior shape knowledge, deal with noisy, corrupted and occluded data.

EMBEDDING THE SHAPE PRIOR KNOWLEDGE INTO THE VARIATIONAL FORMULATION

By extending segmentation functionals with a shape prior, knowledge about the appearance of objects can be directly combined with clues given by the image data in order to cope with typical difficulties of purely data-driven image processing caused by noise, occlusion, etc. The design of shape priors strongly depends on ongoing work on statistical shape models [6, 12, 18]. In particular, advanced models of shape spaces, shape distances, and corresponding shape transformations have been proposed recently [36, 15, 31, 3, 19, 29].

Several works, for example (Cremers et al., 2006; Duci et al., 2002; Rousson and Paragios, 2002; Tsai et al., 2003), use the distance function as the level-set and the square difference between level-sets as the shape dissimilarity measure. A symmetric and unbiased modification of this shape distance (called pseudo distance) has been recently suggested by Cremers and Soatto (2003). However, these similarity measures only account for isometric transformations and scaling, since more general transformations (such as non-isotropic scaling or perspectivity) do not preserve the characteristics of distance functions.

The statistical methodology (Chen et al., 2002; Cremers et al., 2003; Huang et al., 2004; Leventon et al., 2000a,b; Rousson and Paragios, 2002; Tsai et al., 2003) accounts for transformations beyond similarity and for small non-rigid deformations by using a comprehensive training set. It characterizes the probability distribution of the shapes and then measures the similarity between the evolving object boundary (or level-set function) and representatives of the training data. It is important to note that there is no distinction in this method between transformation-based and deformation-based shape variation. The modes of variation have to account for both. Moreover, the performance depends on the size and coverage of the training set.

None of the existing methods accounts for projective transformations between the prior shape and the shape of interest. The inability to deal with projective transformations is significant. In the presence of projectivity, neither similarity nor (even) the affine model provide reasonable approximation for the transformation between the prior shape and the shape to segment. The apparent mismatch inhibits the segmentation process and prohibits accurate reconstruction of the missing parts.

Following the ideas of Raviv et al. (2005) a variational approach to prior based segmentation has been employed here, that explicitly accounts for planar projective transformation, using a single reference object. The segmentation process is carried out concurrently with the registration of the prior shape to the shape of interest. The outcomes of the algorithm include the detection of the object of interest and correct extraction of its boundaries. The planar projective transformation between the two object views is accurately recovered through a planar projective homography mapping. In such a homography mapping the corresponding coordinates are deriving from $x' = Hx$, where the planar homography matrix takes the form

$H = R + \frac{1}{d} t n^T$, with $H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \in \Re^{3x3}$	Equation 4

The eight unknowns of *H* (the ratios of its nine entries) can be recovered by solving at least four pairs of equations of the form:

$x' = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}, \quad y' = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}},$	Equation 5

The nine entries of *H* describe the translation (tx, ty, tz) and rotation (α, β, γ) between the image planes, and the scene structure (ξ, ψ, d) and are evaluated by the following equation:

$h_{11} = \cos\beta\cos\gamma - \frac{t_x}{d}\sin\xi, \; h_{12} = \cos\beta\sin\gamma + \frac{t_x}{d}\sin\psi\cos\xi$ $h_{13} = -\sin\beta + \frac{t_x}{d}\cos\psi\cos\xi, \; h_{21} = \sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma - \frac{t_y}{d}\sin\xi$ $h_{22} = \sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma + \frac{t_y}{d}\sin\psi\cos\xi,$ $h_{23} = \sin\alpha\cos\beta + \frac{t_y}{d}\cos\psi\cos\xi, \; h_{31} = \cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma - \frac{t_z}{d}\sin\xi$ $h_{32} = \cos\alpha\sin\beta\sin\gamma - \sin\alpha\sin\gamma + \frac{t_z}{d}\sin\psi\cos\xi, \; h_{33} = \cos\alpha\cos\beta + \frac{t_z}{d}\cos\psi\cos\xi$	Equation 6

PROPOSED IMAGE SEGMENTATION FUNCTIONAL WITH SHAPE PRIORS

The proposed region-based segmentation functional includes an explicit expression of the projective homography between the prior shape and the shape to segment. Employing the parameterization-free shape description, enabled by the level-set formulation, we gain a significant advantage over landmark-based and template matching techniques that represent shapes by collections of points or features. The suggested distance function between the level-set representations of the matched shapes is well defined and is not depend on shapes sampling. Moreover, transformations applied on the domains of the level-set functions, transform the represented shapes correspondingly. This results in an elegant and powerful mathematical formulation to align the prior and the evolving shape, minimizing their dissimilarity measure with respect to the transformation parameters. The graceful merge of the image data with that of the projectvely registered prior is the essence of the proposed contribution.

Thus, the overall proposed functional for image segmentation is:

$E = E_{region} + \mu \cdot E_{shape}$	Equation 7

where E_{region} is a level set evolution model for the extraction of building from satellite imagery and E_{shape} is calculated by the following equation.

$E_{shape} = \int \left[H(\phi(x,y)) - H\left(T\left(\tilde{\phi}(x,y)\right)\right)\right]^2 dxdy$	Equation 8

For the solution of the functional the energy of the shape prior E_{shape} ,one has to calculate the derivatives for all the 8 parameters of the planar homography transformation.

$\frac{\partial E_{shape}}{\partial \eta} = \int \delta T\left(\tilde{\phi}(x,y)\right)\left[H(\phi(x,y)) - H\left(T\left(\tilde{\phi}(x,y)\right)\right)\right]\frac{\partial T\left(\tilde{\phi},\eta\right)}{\partial \eta} dxdy$ where $\frac{\partial T\left(\tilde{\phi},\eta\right)}{\partial \eta} = \frac{\partial T\left(\tilde{\phi}\right)}{\partial x}\left(\frac{\partial x}{\partial x'}\frac{\partial x'}{\partial \eta} + \frac{\partial x}{\partial y'}\frac{\partial y'}{\partial \eta}\right) + \frac{\partial T\left(\tilde{\phi}\right)}{\partial y}\left(\frac{\partial y}{\partial x'}\frac{\partial x'}{\partial \eta} + \frac{\partial y}{\partial y'}\frac{\partial y'}{\partial \eta}\right)$	Equation 9

The suggested algorithm is demonstrated on real and synthetic examples, in the presence of perspective distortion. The successful segmentation results and the reliable estimation of the transformation parameters suggest this method as a promising tool for various segmentation and registration applications.

In figure 1, a number of shapes describing the boundaries of building are shown. Taking into account the shape of the prior Equation 7 can be solved. In figure 2, the initial IKONOS PAN one meter ground resolution is shown on the left. In the middle the result from the application of the proposed image segmentation geodesic active contours method, which couldn't extracted accurately the boundaries of the building. By taking into account the E_{shape} the result can be much better since it will overcome the limitations of imagetones variation, noise and occlusions

The above described curve evolution level set functional was implemented and tested for the detection of building from an IKONOS PAN image. Before the application of the level set segmentation a pre-processing step for image enhancement and smoothing took place. The applied pre-processing algorithms were described in Karantzalos and Argialas (2006). The pre-processed image was the input to the curve evolution energy and the resulting segmented image was obtained. Finally, certain statistics were calculated for each of the detected segments implying possible buildings: area, perimeter, shape complexity, eccentricity, and orientation. Depending on the above, mainly geometric and shape characteristics, segments that are assumed not to be building were eliminated and thus the final detected buildings were extracted.

The developed scheme has been applied to an IKONOS PAN one meter ground resolution image, from the Agios Stefanos area near Athens, Greece.

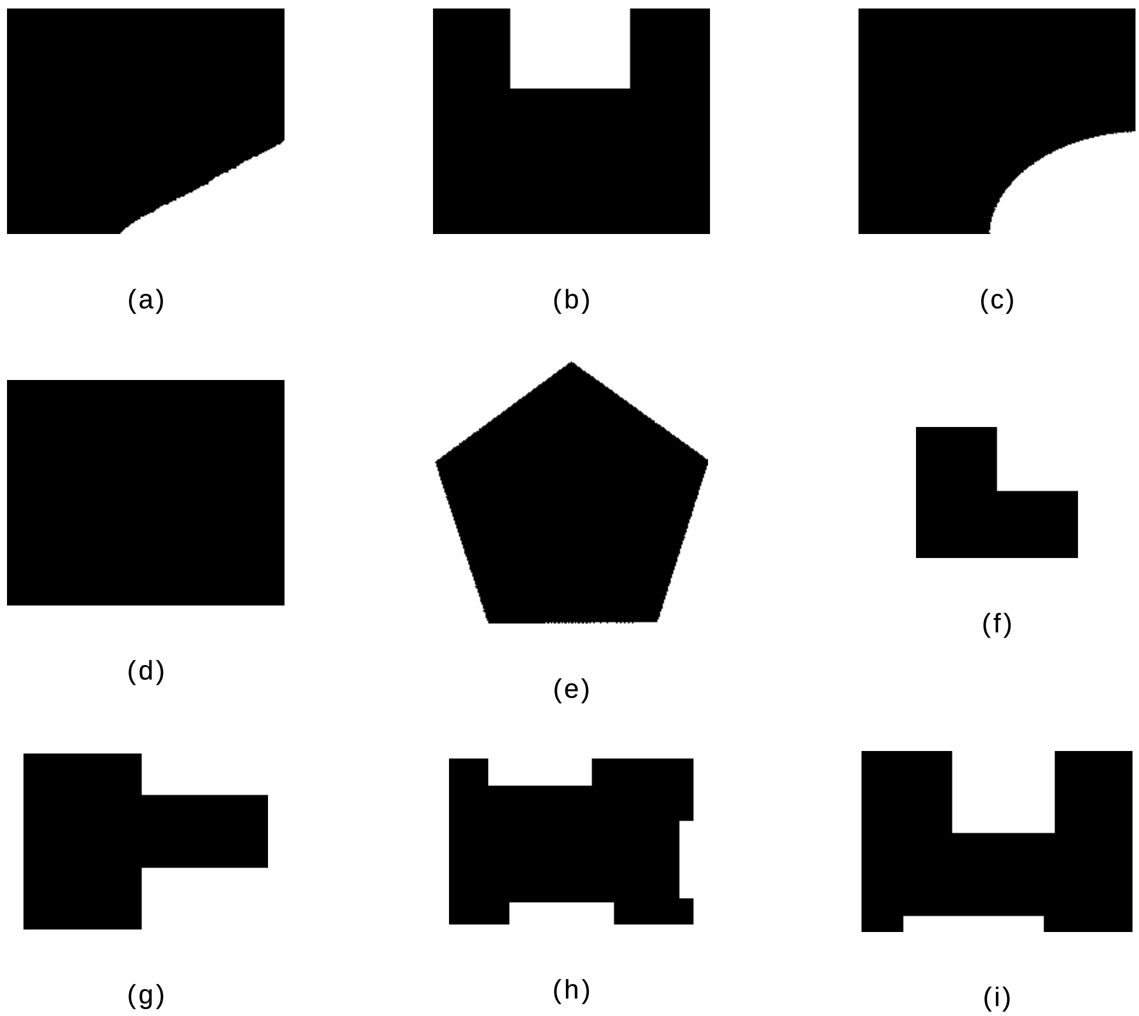

Figure 1. Different building shape priors.

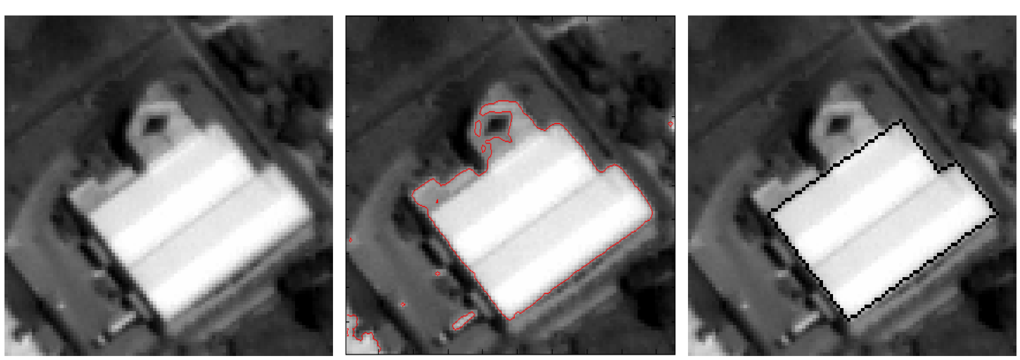

Figure 2. Level set image segmentation based on shape priors.

CONCLUSIONS AND FUTURE WORK

In this paper a region-based segmentation functional includes an explicit expression of the projective homography between the prior shape and the shape to segment. Employing the parameterization-free shape description, enabled by the level-set formulation, a significant advantage over landmark-based and template matching techniques that represent shapes by collections of points or features. The suggested distance function between the level-set representations of the matched shapes is well defined and is not depend on shapes sampling. Moreover, transformations applied on the domains of the level-set functions, transform the represented shapes correspondingly. This results in an elegant and powerful mathematical formulation to align the prior and the evolving shape, minimizing their dissimilarity measure with respect to the transformation parameters. The graceful merge of the image data with that of the projectively registered prior is the essence of the proposed contribution.

ACKNOWLEDGMENTS

The project is co - funded by the European Social Fund (75%) and National Resources (25%) - Operational Program for Educational and Vocational Training II (EPEAEK II) and particularly the Program PYTHAGORAS.

REFERENCES

A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, E. Grimson, and A. Willsky. Model–based curve evolution technique for image segmentation. In Comp. *Vision Patt. Recog*., pages 463–468, Kauai, Hawaii, 2001.

Agouris P., Gyftakis S., Stefanidis A., 2001a. Dynamic Node Distribution in Adaptive Snakes for Road Extraction, *Vision Interface 2001*, pp. 134-140, Ottawa, Canada.

Bailloeul T., Prinet V., Serra B., Marthon P., Chen P. and Zhang H., 2005. Urban building land use change mapping from high resolution satellite imagery, active contours and Hough voting, *ISPMSRS*, Beijing, China.

Ball, J. E., L. M. Bruce. 2005. Level Set Segmentation of Remotely Sensed Hyperspectral Images. *Proceedings of the IEEE Geoscience and Remote Sensing Symposium*. CDRom. 5638 - 564.

Besbes, O., Belhadj, Z., Boujemaa, N., 2006. Adaptive Satellite Images Segmentation by Level Set Multiregion Competition, INRIA, Research Report-5855.

Cao Guo, Yang Xin, Mao Zhihong, 2005. A Two-Stage Level Set Evolution Scheme for Man-Made Objects Detection in Aerial Images. *CVPR* (1): 474-479.

Caselles V., Kimmel R., and Sapiro G., 1997. Geodesic Active Contours. *IJCV*, 22:61–79.

Cohen, L.D., 1991. On active contour models and balloons. *CVGIP Image Understand*. 53, 211–218.

Cremers. D., 2003. A Variational Framework for Image Segmentation Combining Motion Estimation and Shape Regularization. In *IEEE CVPR*, pages 53–58.

D. Cremers, F. Tischhauser, J. Weickert, and C. Schnorr. Diffusion Snakes: Introducing statistical shape knowledge into the Mumford–Shah functional. *Intl. J. of Computer Vision*, 50(3):295–313, 2002.

Gruen, A., E. Baltsavias, and O. Henricsson, 1997. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*. Birkhaeuser, Basel.

Gruen, A., Li, H., 1997. Semi-automatic linear feature extraction by dynamic programming and LSB-Snakes. *Photogrammet. Eng. Remote Sensing,* 63, 985–995.

Jeon, B.K., Jang, J.H., Hong, K.S, 2002. Road detection in spaceborne SAR images using a genetic algorithm. *IEEE Trans. Geosci. Remote Sensing* 40 (1).

Karantzalos K. , Argialas D., 2006. "An image segmentation level set method for building detection", International Symposium of Remote Sensing, Korea, pp.610-614.

Karantzalos K. and N. Paragios, 2005. Implicit free-form-deformations for multi-frame segmentation and tracking, *IEEE VLSM-ICCV05*, China.

Kass, M., Witkin, A., Terzopoulos, D., 1987. Snakes: Active contour models. *Internat. J. Comput. Vis*., 321–331.

Laptev, I., Mayer, H., Lindeberg, T., Eckstein, W., Steger, C., Baumgartner, A., 2000. Automatic extraction of roads from aerial images based on scale space and snakes. *Machine Vision Applicat*. 12 (1), 23–31.

Li, H., 1997. Semi-automatic road extraction from satellite and aerial images. Ph.D. Thesis, Report no. 61, Institute of Geodesy and Photogrammetry, ETH-Zurich, Switzerland.

M. Leventon, W. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In Proc. CVPR, pages 316–323. *IEEE Comp. Soc*., 2000.

M. Rousson and N. Paragios. Shape priors for level set representations. In A. Heyden et al., editors, *Proc. of the Europ. Conf. on Comp. Vis*., volume 2351 of LNCS, pages 78–92, Copenhagen, May 2002. Springer, Berlin.

Osher S. and Paragios N., 2003. *Geometric Level Set Methods in Imaging Vision and Graphics*, Springer Verlag, ISBN 0387954880.

Osher S. and Sethian J., 1988. Fronts propagation with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations, *J. Comput. Physics*, vol. 79, pp. 12–49.

Paragios N. and Deriche R., 2000. Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:266–280,.

Paragios N. and Deriche R., 2002. Geodesic Active Regions: A New Framework to Deal with Frame Partition Problems in Computer Vision, *Journal of Visual Communication and Image Representation*, vol. 13, pp. 249–268.

Paragios N., Chen Y. & Faugeras O., 2005. *Handbook of Mathematical Models of Computer Vision*, Springer, ISBN 0387263713.

Rochery M., Jermyn I. H., and Zerubia J.. 2003. Higher order active contours and their application to the detection of line networks in satellite imagery. In *Proceedings of the IEEE VLSM / ICCV03*, Nice, France.

Ruther, H., Martine, H.M., Mtalo, E.G., 2002. Application of snakes and dynamic programming optimisation technique in modeling of buildings in informal settlement areas. *ISPRS J. Photogrammet. Remote Sensing* 56 (4), 269–282.

Samson, C , Blanc-Feraud, L , Aubert, G , and Zerubia, J , 2001. Two variational models for multispectral image classification," Energy Minimization Methods in Computer Vision and Pattern Recognition, Lecture Notes in *Computer Science*, vol. 2134 , pp. 344 -356.

T. Riklin-Raviv, N. Kiryati and N. Sochen. 2004. Unlevel Sets: Geometry and Prior-based Segmentation. In *Proceedings of the 8th European Conference on Computer Vision* (LNCS 3024, Part IV), Prague, Czech Republic, May.

T. Riklin-Raviv, N. Kiryati and N. Sochen. 2005. Prior-based Segmentation by Projective Registration. In *Proceedings of the International Conference on Computer Vision*, Beijing, China.

W. Mio, A. Srivastava, and X. Liu. Learning and bayesian shape extraction for object recognition. In *European Conf. on Computer Vision*, volume 3024 of LNCS, pages 62–73, Prague, 2004. Springer.

Zafiropoulos, P., Schenk, T.F., 1998. Extraction of road structures with color energy models. In: Sabry, F., El-Hakim, Gruen, A. (Eds.), *Videome-TRICS VI*., Proc. SPIE, vol. 3641, pp. 276–290.

Zhao H-K., Chan T., Merriman B., and Osher S., 1996. A variational Level Set Approach to Multiphase Motion. *Journal of Computational Physics*, 127:179–195.