A Bi-Directional Integrated Model for Non-Rigid Motion Analysis

Chunyen Liu
Department of Computer and Information Sciences
University of Delaware
Newark, DE 19716, U.S.A.
cliu@cis.udel.edu

Abstract

To be able to understand the motion of deformable objects, techniques in image processing and computer vision are essential for non-rigid motion analysis in this active research area. We have developed an integrated model that combines the advantages of the boundary-based and region-based approaches and avoids problems caused by each stand-alone approach, e.g. overshrinking, oversegmenting, noise sensitivity. This image segmentation model further iteratively improves each submodel in both directions until it satisfies predefined criteria. Different frame-to-frame prediction methods, naive, inflation, and optic flow, are developed and evaluated. Comparison between our model and other models is studied and illustrated by examples.
Further improvements on our motion tracking model are possible by evaluating new attraction functions and prediction methods. Some important notes on future work are given for incorporating other information (e.g. morphometrics) into our integrated model in several aspects. This is expected to improve our existing model and makes it a better research tool for scientists to answer questions more accurately in a variety of application areas, e.g. biomedial motion analysis, navigation.

1 Introduction

There has been a great deal of research interest in motion tracking [12] [1] [22] because of its increasing popularity and applicability in a wide variety of applications, e.g. biomedical analysis, target tracing, automated navigation. Non-rigid motion interprets a even richer set of objects in the real world. Unlike rigid motion, no spatial relationships between fixed feature points on objects can be utilized as a priori knowledge to calculate object depth and surface structure. Therefore, non-rigid motion is more difficult by its nature. The goal of this research is to study the non-rigid motion tracking problem with our integrated approach. This report will focus on the these issues: (1) How to get the optimized image segmentation result on a single image by integrating boundary-based and region-based approaches? (2) How to achieve better tracking results by making reasonable redictions from previous contours? (3) Comparison between traditional and our approaches. Applying our research in cell motion is mainly motivated by the followings: Cells move and change their shape simultaneously and their displacements are usually

BMVC 1998 doi:10.5244/C.12.39

too small to be perceivable by human eyes - manual tracking is almost impossible and it is an error-prone process. Since cells are small, any algorithmic defects will play a major role of propogating overall errors. Furthermore, most cells have different intensity levels and very often unnoticeable boundaries as in Fig 1. These challenging intuitions motivate our research that incorporates techniques in image processing, computer vision, and numerical analysis [9] [17] for this non-rigid motion analysis. Most of these ill-posed problems [4] [21] require regularization techniques and calculus of variations to find their solutions. Firstly, we study and develop a promising image segmentation approach that combines the advantages of region-based and boundary-based models and avoids common problems in those models. Unlike most integrated models, it is bi-direcrtional, i.e. both submodels iteratively refine each other. With this new model, we can more precisely identify objects of interest in the image and greatly reduce errors that might propagate throughout motion sequences. We also compare and describe why we choose some specific submodels over others. Secondly, predicting the starting contour shape and location is investigated and evaluated for a better selection of prediction methods. Thirdly, our error estimation for comparison between traditional and our approach is based on experiments from tracking both real and synthetic motion sequences. In the future, we will later start exploring the possibility of incorporating information from morphometric analysis [3] into the integrated tracking model to analyze motion and improve tracking.

2 Background

2.1 Image Segmentation

2.1.1 Boundary-Based Approach

To achieve image segmentation, traditional boundary-based approaches find object boundaries by locating intensity discontinuities and linking meaningful edges. Therefore, they tend to be sensitive to local noise and do not have the dynamic behavior as active contour models, also known as *snakes* [13]. To account for motion tracking, active contour models deform their contours with time. Thus, active contour models are the most appropriate choice among boundary-based algorithms. [16] [2] A snake is defined as an energy-minimizing spline under the influence of image forces and external constraint forces. The internal forces serve to impose a piecewise smoothness constraint. The image forces push the snake toward salient image features like edges. The external constraint forces are responsible for putting the snake near the desired local minima. Representing the position of a snake by $v(s) = (x(s), y(s))$, the energy functional of the snake model is defined as follows:

$$E^*_{snake} = \int_0^1 \left[E_{int}(v(s)) + E_{image}(v(s)) + E_{ext}(v(s))\right] ds \tag{1}$$

Snake models usually have the problem of overshrinking by their nature of attracting contours to local minima. Therefore, some authors add an additional force to the energy equation and this *balloon force* is defined for inflating or deflating the snake [7]. Suppose we employ the image force as $F = -P$ where the direction of F implies steepest descent in P. Equilibrium is achieved at points where P is a minimum in the direction normal to the curve. The force F equipped with the new balloon force now becomes $F = k_1 n(s) - k_2 \frac{\nabla P}{|P|}$ where $n(s)$ is the normal unitary vector to the contour at point $v(s)$ and k_1 is the amplitude control. The sign of k_1 controls the effect of inflation and deflation. Their variations, dealing with topological change, different dimensions, physically-based approaches, can be found in [15] [16] [19].

2.1.2 Region-Based Approach

To achieve image segmentation, most region-based algorithms mainly depend on pixel statistics and how uniformly distributed intensity levels are in regions of interest. Therefore, for objects with shaded regions, inaccurate segmentation is doomed to occur, e.g. *region growing methods* [17] [9]. Some authors [20] develop a continuous version of Geman's model which can be solved using variational scheme. The functional is defined as follows:

$$E(f,B) = \int\int_R w_1(f-g)^2 dxdy + \int\int_{R-B} w_2 \left|\nabla f\right|^2 dxdy + w_3 \left|B\right| \qquad (2)$$

where R = image domain; B = union of segmented boundaries; g = original image data; f = piecewise smooth image estimate; w_1, w_2, w_3 = defined weights. Image segmentation is performed by finding f and B which minimize this functional. Another image segmentation method is called the *gradient watershed model*.[8] It is described by following the drainage patterns of simulated rainfall on an image that can be used to partition an image into watershed regions called hills and dales. The boundaries of hills correspond to ridge tops and the boundaries of dales correspond to valley bottoms, so multiscale watershed analysis provides an alternative method to study the scale-space behavior of ridges and valleys for image segmentation. The computed boundaries of gradient watersheds correspond closely to the edges of the original image. This is the main reason why we choose the gradient watershed model over other region-based methods.

2.1.3 Hybrid Approach

Hybrid models for image segmentation are developed to avoid those common problems in stand-alone boundary-based and region-based models. Most hybrid approaches [6] [23] [18] are uni-directional, i.e. taking output from one submodel as input and computing the other submodel but not vice versa. Actually, almost all uni-directional approaches start with region-based submodels and then refine boundary-based submodels. One bi-directional approach is performed by integrating two submodels by game theory. [5] Two submodels are treated as two players playing a stochastic non-zero sum game. The game stops when none of the modules can improve their positions. The final solution accounts for the *Nash Equilibrium*. The interacting model is illustrated in Figure 2. Optimization procedures can be viewed as minimizing cost functions using dynamic programming techniques. Our segmentation problem can be formulated in the following structure that minimizes two cost functions F_B and F_R:

$$\begin{aligned} F_B(p_B,p_R) &= f_B(p_B) + \alpha f_{RB}(p_B,p_R), \\ F_R(p_B,p_R) &= f_R(p_R) + \beta f_{BR}(p_B,p_R). \end{aligned} \qquad (3)$$

where f_B, f_R: cost functions of boundary-based and region-based submodels; p_B, p_R: boundary-based and region-based submodels; f_{RB}: cost function of boundary-based influenced by region-based; f_{BR}: cost function of region-based influenced by boundary-based; α, β: control coefficients. We have implemented our integrated model by an enhanced snake model and the gradient watershed model. The interactive influences f_{RB}, f_{BR} on submodels in the new model are developed by our research. This scheme is demonstrated to be a promising approach in terms of efficiency and accuracy.

2.2 Motion Tracking and Analysis

Apparent motion of in the images can be characterized by observing features or brightness patterns. There are two distinct approaches: (1) Feature-based approaches assume there is only rigid body motion and inter-frame correspondence has been established between features. and are based on extracting a set of relatively sparse, but discriminatory features, such as corners, occluding boundaries. (2) Optic-flow-based approaches are based on computing instantaneous velocities of brightness values in the images. For a general review and framework on motion tracking, see [12] [1] [2] [10] [22]. Here we only focus on (2). *Optic flow* is defined as the apparent motion of some brightness patterns. [11] There are approximately three ways to compute optic flow: (1) gradient-based, (2) matching-based, and (3) frequency domain based methods. Because gradient-based methods are better investigated are commonly used in the literature, we focus on them in more detail than the other two approaches. Gradient-based methods assume brightness pattern will remain the same for a small time interval δt. $I(x, y, t)$ is the image intensity at (x, y, t), and x, y, t are coordinates and time. Expand I around (x, y, t) using Taylor series and discard second- and higher-order terms in δx, δy, and δt. We obtain:

$$I_x u + I_y v + I_t = 0 \tag{4}$$

where $u = \frac{dx}{dt}$, $v = \frac{dy}{dt}$, $I_x = \frac{\partial I}{\partial x}$, $I_y = \frac{\partial I}{\partial y}$, $I_t = \frac{\partial I}{\partial t}$. Equation 4 is called the *optic flow constraint equation*, since it expresses a constraint on the components u and v of the optic flow. This is an under-determined equation since we have only one equation but two unknowns. Again, we solve this equation by imposing an additional smoothness constraint , which is known as *regularization*. The equation is now as follows:

$$F = \int\int((u_x^2 + u_y^2) + (v_x^2 + v_y^2))dxdy + \alpha\int\int(I_x u + I_y v + I_t)^2 dxdy \tag{5}$$

To solve the problem of minimizing a functional F, calculus of variations is employed. The corresponding Euler equations yield the solution. [11] [14]

3 The Integrated Model

3.1 Boundary-Based Submodel

The enhanced active contour model we implement here as our boundary-based model is the generally adopted boundary-based method for tracking. Interactive input through a computer pointing device is most commonly used. One of the most popular functions used as image force is image gradient because of its efficiency. In order to stabilize gradients, two investigated solutions are used. One is normalization, i.e. keeping their range from 0 to 1. This will reduce the contour moving step. The other is smoothing. This will allow a wider range of contour points to move in the right direction, especially for plateau points. Now that we have placed the initial contour, we need to move these contour points to optimal locations according to the energy minimization solution. The balloon force is moving along the contour normal to avoid overshrinking. Most authors in the literature apply fixed number of iterations to manually stop moving contour points owing to their previous experience on certain sets of images. We use a flexible model of checking contour length change, enclosed area change, enclosed gradient change, and/or *steady-support criterion* [16] as in $D(v) = \frac{|\int_v P_{prev} - \int_v P_{new}|}{\int_v P_{prev}}$, where D is the relative difference, P is one of the mentioned properties, e.g. gradient. If one and/or more D's are less than a predetermined value, then the iteration process stops.

3.2 Region-Based Submodel

The major region-based approach investigated by our research for image segmentation is the watershed model. The reasons we choose this model as opposed to other region-based models are mainly because of its excellent visual correspondence of region boundaries and high image gradients. The watershed model divides an image into subregions by following gradient paths of local minima. Watersheds are achieved by assigning each pixel the same subregion number as its local minimum. Region merging in the watershed model is done by detecting difference in mean, variance, adjacent boundary length, and edge strength to see if they satisfy a predetermined threshhold. Visually and noticeably while compared with other approaches, the watershed model provides the remarkable correspondence of boundaries and high image gradients which turn out very helpful because we are interested in the integration of information both from boundary-based and region-based models.

3.3 Integration of Boundary-Based and Region-Based Submodels

To overcome previous problems by applying a single boundary-based or region-based approach, the integration of the two models seems to be a reasonable and promising solution in several aspects. Our approach is thus presented to interact these two modules, boundary-based active contour model and region-based watershed model. The output of one module is taken as input of the other for the purpose of optimizing results and vice versa. It finally reaches an equilibrium after satisfying some mathematical constraints. Let us leave out the image preprocessing and attraction function computation steps we have implemented in our program. The sketch of our algorithm is as follows: *Step 1:* Apply the pure region-based watershed model at this initialization step. *Step 2:* Apply the boundary-based active contour model on image segmentation with a balloon force which moves in a direction guided by the output of the region-based watershed model. Determining if the balloon force should be moving along the inward or outward pointing normal is based on how much a subregion is enclosed in our current contour. We compute the overlapping percentage for each subregion against the current contour. If desirable percentage of overlapping is satisfied, then we assign each pixel in the entire subregion with -1, otherwise we mark the entire subregion with $+1$. What this means is when we are optimizing a boundary point in the active contour model, a region-based balloon force will direct this point to maximize the area of the desirable region as well as minimize the energy term in the active contour model. This is illustrated in Figure 3. The current boundary-based contour is drawn with dotted lines. The f_{BR} term in equation 3 is: $-\int\int_{A_{in}} I_r(x,y)dA$, where the negative sign means "minimizing the enclosed area A_{in}" from the segmented regions I_r. *Step 3:* Apply the region-based watershed model on image segmentation with a merging constraint which forces some subregions to be merged in a way influenced by the boundary-based active contour model. Candidate subregions are selected based on the overlapping constraint. We compute the overlapping constraint for each subregion against the output contour from the boundary-based model. If desirable percentage of overlapping is satisfied, then we assign each subregion with a $*$, otherwise we discard the subregion. We merge these subregions into a new region. We thereafter continue our region-based process and test if this new region can be further merged. This is illustrated in Figure 4. All subregions marked with a $*$ are to be merged and this new region will be further merged with other regions according to the watershed model merging rules. The f_{RB} term in equation 3 is: $\sum_{(i,j)\in A_{in}}(I_{i,j}-P_{i,j})^2+\sum_{(i,j)\in A_{out}}(I_{i,j}-Q_{i,j})^2$, where A_{in} represents pixels inside the contour area, A_{out} represents pixels outside the contour area, P and Q are pixel intensities inside and outside the contour area, and I is the mean of the target region. *Step 4:* Repeat Steps 2 and 3 until the defined stopping criteria are satisfied. *Step 5:* Apply a selected frame-to-frame prediction method described in section 3.4 and restart from Step 1. Figure 5 illustrates the results from pure and integrated models. Our integrated model has two outputs, optimized boundaries and regions as shown.

3.4 Frame-to-Frame Prediction Method

The prediction methods we have implemented are used to initialize the starting contour for the next frame by results generated in the current frame. That means as long as we place our initial contour at the first frame, ideally motion tracking for the entire image sequences will be done automatically without further input. Three methods are investigated. The first method is a very naive guess using exactly the optimized contour C_t at frame t as starting contour at frame $t+1$. We then go on and optimize C_{t+1}. The second method uses a similar scheme as in the first, but we expand about the centroid of C_t with an inflation factor and use this new C_t' as starting contour for C_{t+1}. In the third method, we compute image optic flow [11] from frame t to $t+1$ and use this as a constraint to compute C_t', which is again used as starting contour for C_{t+1}. From experiments, we see that second and third approaches perform better among these methods. However, in the second approach some scaling factors make the predicted starting contour too far away from the image gradients and thus it is not as effective in a general sense.

3.5 Error Estimation

To estimate the errors from the above algorithms we have implemented for image motion tracking, we demonstrate by their performance on both real and synthetic image motion sequences which contain randomly generated translation, rotation, scaling, and intensity levels by our program. Our synthetic sequence generating program also outputs the true boundary solutions for comparison purposes. The experimental results are illustrated in Figure 6 and Table 1.

Error Estimation	**Test 1**	**Test 2**	**Test 3**	**Test 4**
average Euclidean (pixels)	1.64	0.73	1.31	0.65
length (percent)	0.90	0.90	25.14	1.00
area (percent)	36.02	2.45	35.10	1.02
max axis angle (percent)	0.02	0.01	15.25	0.02
compactness (percent)	25.00	0.01	12.41	0.01

Table 1: Error estimation after tracking synthetic and real image motion sequences (1) pure active contour model on synthetic sequences (2) integrated model on synthetic sequences (3) pure active contour model on real sequences (4) integrated model on real sequences

4 Conclusion

As we addressed our motivations and concerns earlier, , the following major tasks have been achieved: (1) To achieve desirable image segmentation which combines the advantages of boundary-based and region-based models, we develop a promising bi-directional integrated model. In this model, we simultaneously solve some common problems encountered by using traditional boundary-based and region-based stand-alone models. (2) We further evaluate this integrated model by tracking motion sequences of deformable objects with different frame-to-frame prediction methods. (3) Finally we compare our results with traditional methods on real and synthetic sequences. This work can be used as a tool for tracking and analyzing deformable objects in motion sequences, e.g. biomedical image analysis, automated surveillance and navigation, target tracing, intelligent image editing. Although there is a trade-off between the more complex integrated model and traditional models, accuracy of analyzing deformable objects is always the main concern and this model occasionally outperforms others in efficiency because of good initial estimations for convergence. Future work can be further investigated on several aspects. New attraction functions and prediction methods can be developed because they contribute most of the errors. Morphometric analysis [3] can be experimented to refine the model in prediction and adaptive schemes.

5 Acknowledgement

The author would like to thank Dr. John Gauch from the University of Kansas for invaluable discussions and assistance on this topic and Dr. Chandra Kambhamettu from the University of Delaware for his support.

References

[1] J. K. Aggarwal and N. Nandhakumar. On the computation of motion from sequences of images – a review. *Proceedings of the IEEE*, 76(8), August 1988.

[2] Andrew Blake, Rupert Curwen, and Andrew Zisserman. A framework for spatiotemporal control in the tracking of visual contours. *International Journal of Computer Vision*, 11(2), 1993.

[3] Fred L. Bookstein. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press, 1991.

[4] Terrence Boult. What is regular in regularization. In *IEEE Proceedings of the First International Conference on Computer Vision*, 1987.

[5] H. Bozman and J. Duncan. A game-theoretic approach to integration of modules. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(11), November 1994.

[6] Amit Chakraborty, Lawrence H. Staib, and James S. Duncan. Deformable boundary finding influenced by region homogeneity. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1994.

[7] Laurent D. Cohen. Note on active contour models and balloons. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 53(2), March 1991.

[8] John M. Gauch. Image segmentation and analysis via multiscale gradient watershed hierarchies. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, To appear.

[9] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Addison-Wesley Publishing Co., 1994.

[10] Robert M. Haralick and Linda G. Shapiro. *Computer and Robot Vision*, volume II, chapter 15 Motion and Surface Structure from Time-Varying Image Sequences. Addison-Wesley Publishing Co., 1993.

[11] B. K. Horn. *Robot Vision*, chapter 12 Motion Field and Optical Flow. MIT Press, 1986.

[12] T.S. Huang. *Image Sequence Analysis*. Springer-Verlag, 1983.

[13] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4), 1988.

[14] Jiansu Lai. *Tracking Non-Rigid Object Motion in Color Image Sequences*. PhD thesis, Northeastern University, May 1994.

[15] Kok F. Lai and Roland T. Chin. Deformable contours: Modeling and extraction. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1994.

[16] Frederic Leymarie and Martin D. Levine. Tracking deformable objects in the plane using an active contour model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(6), June 1993.

[17] Vishvjit S. Nalwa. *A Guided Tour of Computer Vision*. Addison-Wesley Publishing Co., 1993.

[18] Theo Pavlidis and Yuh-Tay Liow. Integrating region growing and edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(3), March 1990.

[19] Ramin Samadani. Adaptive snakes: Control of damping and material parameters. In *SPIE Vol. 1570 Geometric Methods in Computer Vision*, San Diego CA, July 1991.

[20] Jayant Shah. Segmentation by nonlinear diffusion ii. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1992.

[21] Demetri Terzopoulos. Regularization of inverse visual problems involving discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(4), July 1986.

[22] J. Wang, T.S. Huang, and N. Ahuja. *Motion and Structure from Image Sequences*. Springer-Verlag, 1993.

[23] S. C. Zhu, T. S. Lee, and A. L. Yuille. Region competition. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1995.

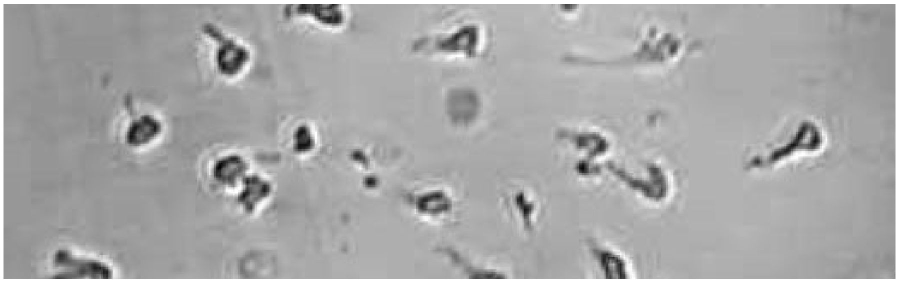

Figure 1: An example of cells in an image

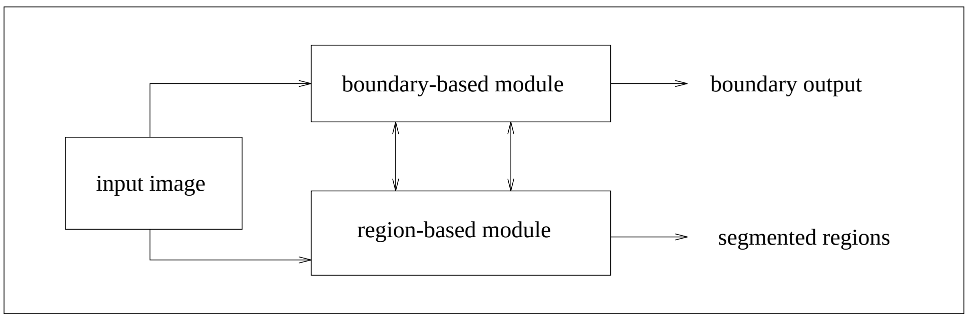

Figure 2: The bi-directional integrated model

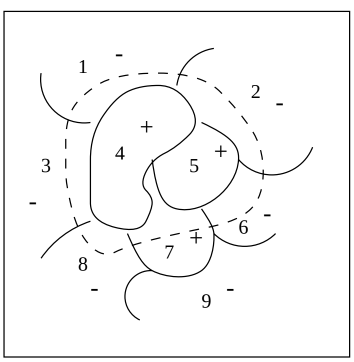

Figure 3: Boundary-based influenced by region-based.

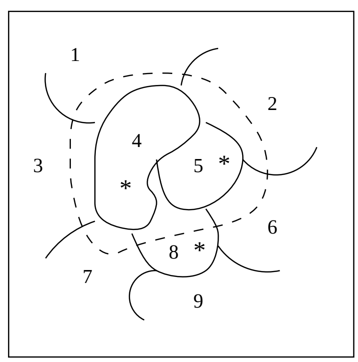

Figure 4: Region-based influenced by boundary-based.

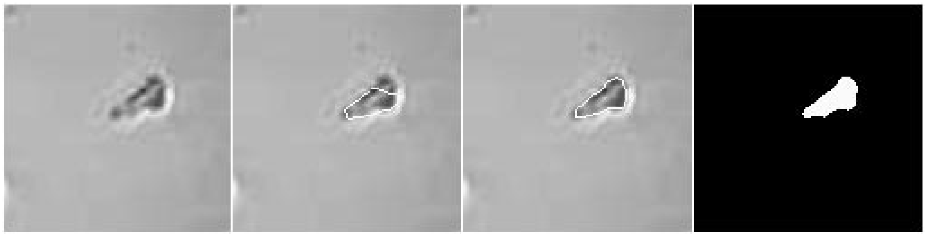

Figure 5: After tracking (left to right): (1) final frame (2) pure active contour model (3) our snake submodel (4) our watershed submodel

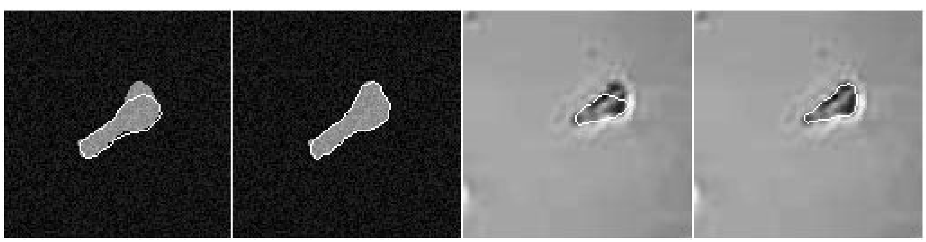

Figure 6: After tracking (left to right): (1) pure active contour model on synthetic sequences (2) integrated model on synthetic sequences (3) pure active contour model on real sequences (4) integrated model on real sequences