# A Segmentation Algorithm for Contrast-Enhanced Images

Junhwan Kim Ramin Zabih
Computer Science Department
Cornell University

## Abstract

*Medical imaging often involves the injection of contrast agents and the subsequent analysis of tissue enhancement patterns. Many important types of tissue have characteristic enhancement patterns; for example, in magnetic resonance (MR) mammography, malignancies exhibit a characteristic "wash out" temporal pattern, while in MR angiography, arteries, veins and parenchyma each have their own distinctive temporal signature. In such image sequences, there are substantial changes in intensities; however, this change is due primarily to the contrast agent rather than the motion of scene elements. As a result, the task of segmenting contrast-enhanced images poses interesting new challenges for computer vision.*

*In this paper, we propose a new image segmentation algorithm for image sequences with contrast enhancement, using a model-based time series analysis of individual pixels. We use energy minimization via graph cuts to efficiently ensure spatial coherence. The energy is minimized in an expectation-maximization fashion that alternates between segmenting the image into a number of non-overlapping regions and finding the temporal profile parameters which best describe the behavior of each region. Preliminary experiments on MR mammography and MR angiography studies show the algorithm's ability to find an accurate segmentation.*

## 1. Contrast-Enhanced Image Sequences

Many medical imaging studies involve the injection of contrast agents in order to visualize structures that cannot otherwise be distinguished. The choice of contrast medium depends upon the imaging modality; in CT the contrast agents are iodine-based, while MR contrast agents use gadolinium [13]. In the resulting contrast-enhanced image sequences the patient is generally stationary, and the substantial changes in intensities that occur are due primarily to the contrast agent. Moreover, different types of tissues often have characteristic enhancement patterns, which form the basis for segmentation. This is especially true in MR, where there is a wide variety of enhancement patterns. For example, in MR mammography, malignancies exhibit a characteristic "wash out" temporal pattern, while in MR angiography, arteries, veins and parenchyma have distinct temporal profiles.

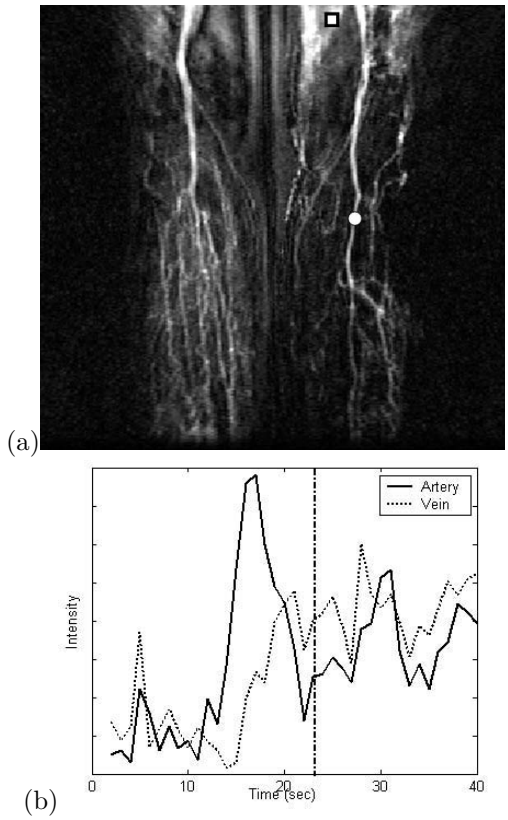



Figure 1: (a) Snapshot from an MR angiography image sequence (b) Intensity profile of selected pixels

Automatic segmentation of contrast-enhanced image sequences poses novel challenges to computer vision. At first glance, motion segmentation techniques such as [6, 21] seem to be applicable. However, these methods are designed for situations where scene elements are moving with respect to a conventional camera. In medical imaging applications, the imaging



Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV'03)
0-7695-1950-4/03 $ 17.00 © 2003 IEEE

IEEE COMPUTER SOCIETY

devices are quite different from traditional cameras. Moreover, in contrast-enhanced image sequences it is common for the patient to be stationary. This is especially common for MR imaging, due to the well-known artifacts that motion produces in this imaging modality [13]. However, it is also often true in CT (for example, when imaging the brain) or lower abdomen. While there is obviously motion present (for example, at the cellular level) even in a stationary patient, the change in intensities observed in contrast-enhanced images is not due to the motion of scene elements.

In order to segment contrast-enhanced image sequences, it is vital to use the information in the sequence. There is simply not enough information in a single image to reliably perform segmentation. This can be easily seen from Figure 1, which shows an example of contrast-enhanced MR angiography of the lower extremities. Figure 1(a) shows a snapshot from the image sequence (to make the arteries and veins visible, the image shown is the difference between a post-contrast image and a pre-contrast image). Figure 1(b) depicts the temporal (intensity) profiles of the selected pixels in Figure 1(a). The black dotted vertical line in Figure 1(b) indicates where the snapshot in Figure 1(a) is taken.

The circled pixel is on an artery, and its temporal profile shows a sharp increase when the contrast agent arrives. By comparison, the temporal profile of the vein pixel (indicated by the rectangle) shows a steady increase after the contrast agent arrives, which indicates that the pixel is on vein. However, the vein pixel is brighter than the artery pixel in the snapshot shown in Figure 1(a), which would mislead an image segmentation method based on this snapshot image. An image segmentation algorithm based on the temporal profile of each pixel will have much higher chance to correctly classify these pixels as either artery or vein.

### 1.1 Overview of our approach

Our approach is based on the application of time series analysis to pixel temporal profiles. Time series analysis [9] is a set of techniques for analyzing a temporal series of observations. It has primarily been used in the financial community, for example to analyze commodity prices. However, the information contained in any individual pixel's temporal profile is noisy. In addition, a correct segmentation would be spatially coherent (i.e., the pixels that should be labelled artery are not distributed independently at random through the image). This spatial coherence assumption allows us to overcome the noise in a single pixel's temporal profile.

We assume that for each type of tissue that is of interest, there is a model available of that tissue's temporal profile. For example, in MR angiography the model for an artery might be a peak formed by 4 line segments, while the model for a vein might be a gradual rise formed by 2 line segments (see Figure 3). Eventually it might be possible to learn such a model from hand-labelled examples, but we currently assume that such a model already exists. Each model will have a number of parameters, which will need to be estimated. The input to our algorithm will be a set of models, together with an image sequence. The output will be a labelling of the pixels, such that each pixel is associated with a single model, together with an estimate for each model's parameters.

The algorithm itself is very similar in spirit to existing approaches to motion segmentation based on expectation-maximization [11] (we will describe some related work in the next section), and especially to [2]. We alternate between two steps. In the first step, given a set of models, each of which has some parameters, we must label each pixel with a model. In the second step, we estimate the parameters for each model by analyzing the set of pixels that have been labelled with the model. Simple statistical techniques from time series analysis are used in both of these steps.

The rest of this paper is organized as follows. Section 2 summarizes related work. In Section 3, we give a formal definition of our problem. The two alternating steps of our algorithm are described in Section 4 and Section 5. Preliminary experimental results for MR angiography and MR mammography are given in Section 6. We conclude with some suggestions as to how our work can be extended to contrast-enhanced image sequences that also contain significant motion in Section 7.

## 2. Related work

There is, of course, a great deal of prior work on image segmentation, primarily focusing on the segmentation of a single image. However, it does not appear that contrast-enhanced images can be segmented by looking at a single image; as a result, the most relevant related work involves segmentation of image sequences.

Several vision researchers has used time series analysis for various purposes [17, 5, 18, 4], although they in general have not referred to these techniques by that name. Kanade et al. [14] built a VLSI sensor array for a rangefinder that gathers range data in parallel as a scene is swept continuously by a moving stripe. Each cell continuously records the time at which a peak is observed at each sensor pixel, and detects the time at which it observed the peak incident light intensity during the sweep of the stripe. Curless and Levoy [10]



Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV'03)
0-7695-1950-4/03 $ 17.00 © 2003 IEEE

IEEE COMPUTER SOCIETY

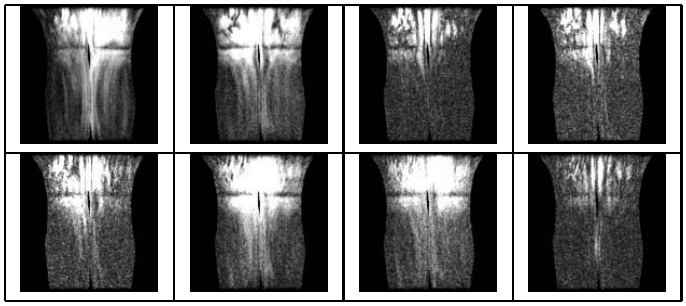


Figure 2: An example of an MR angiography image series: read from left to right and top to bottom.

proposed a ranging method based on analyzing the time evolution of the structured light reflections. They use this information to correct systematic distortions of the range data that stem from curved surfaces, discontinuous surfaces, and surfaces of varying reflectance, thereby attaining significantly higher accuracy.

Seitz and Dyer [19] provide another example of motion pattern analysis, by determining if an image sequence could have been produced by an object whose motion is periodic in some reference frame. Zhang et al. developed accurate shape acquisition systems using color structured light [23]. They solve the correspondence problem by using a multi-pass dynamic programming algorithm that eliminates the global smoothness assumptions and strict ordering constraints present in many formulations. Recently, they also proposed the "spacetime stereo" that recovers shapes by observing their appearance variations over time using a stereo pair [24]. They define the mean shape over a stereo pair of image sequences as the shape which has minimum matching error over space and time, and the shape variation is modelled as a stochastic linear dynamic system. The method is shown to generate better shape reconstruction than traditional stereo pair.

Bobick and Davis [4] introduced the temporal template representation, consisting of the motion-energy image to represent where there is motion and the motion-history image to indicate when there is motion. Their algorithm and ours share the idea of indicating when an important event occurs (motion in their case, contrast arrival, etc, in our case) for each pixel. However, our algorithm differs from theirs in that (a) we use a comprehensive model to describe the behavior of each pixel, which is necessary for segmentation, and (b) we use spatial coherence as well as the information from each individual pixel.

The body of work that is most closely related to ours is the research on expectation-maximization (EM) based segmentation of image sequences. Most of these methods deal with motion segmentation using layers [1, 21, 22]. The problem we address, however, is not motion segmentation, since there is no motion in the images. Moreover, these methods do not use graph cuts to efficiently label pixels with the appropriate model while preserving spatial coherence. The closest work to ours is the algorithm of Birchfield and Tomasi [2], which is an EM-based method for stereo and motion with slanted surfaces that uses the graph cuts algorithm of [8] to label pixels. Their work, however, does not analyze the temporal profile of individual pixels, and is not straightforwardly applicable to the task of segmenting contrast-enhanced image sequences.

# 3. Problem formulation

The input to our algorithm is a sequence of images $i : \mathcal{P} \times \mathcal{T} \rightarrow \Re$ for each pixel $p \in \mathcal{P}$ and each time $t \in \mathcal{T} = \{0, 1, 2, \cdots, T-1\}$ (see Figure 2 for an example). For convenience, $i_p \in \Re^T$ denotes the temporal profile of the pixel $p$. We represent the output as a labelling $f : \mathcal{P} \rightarrow \mathcal{L}$, where $\mathcal{L}$ is the set of labels, along with temporal profile parameters $\theta_l$ for each label $l$. For instance, $\mathcal{L}$ = {malignant tumor, benign tumor, non-tumor} for a breast image sequence, while $\mathcal{L}$ = {artery, vein/parenchyma} for an MR angiography image sequence.[1] $\theta_l = \{a_l, b_l\}$ when the temporal profile is modelled by a straight line $i_{pt} = a_l t + b_l$. $\theta_l$ will have more elements when we use more sophisticated models.

Our goal is to find a labelling that matches pixels of similar temporal profile while minimizing discontinuities. More specifically, we minimize the following energy functional:

$$E(f) = \sum_{p \in \mathcal{P}} \sum_{q \in \mathcal{N}_p} 2u_{\{p,q\}} T[f_p \neq f_q] + \sum_{p \in \mathcal{P}} g(i_p, p, \theta_{f_p}), \quad (1)$$

where the indicator function $T[\cdot]$ is 1 if its argument is true and 0 otherwise. $\mathcal{N}_p$ is a subset of pixels in $\mathcal{P}$ describing the neighbors of $p$. The first term enforces smoothness by penalizing the discontinuities. $u_{\{p,q\}}$ can be interpreted as a cost of a discontinuity between $p$ and $q$ [8]. The second term is a data-dependent energy term indicating the costs of assigning the labels to the pixel. $g(i_p, p, \theta_{f_p})$ is the cost of assigning the label $f_p$ to the pixel $p$ whose temporal profile is $i_p$.

To minimize the energy functional, we alternate between segmenting the image into disjoint regions by assigning a label to every pixel and finding the temporal profile parameters of each region. We will describe these two steps in the next two sections.

[1] The label vein/parenchyma will be used for veins as well as for parenchyma, which enhances similarly to veins. The portion of the image outside the patient's body can be easily detected automatically and not considered for further processing.



Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV'03)
0-7695-1950-4/03 $ 17.00 © 2003 IEEE

IEEE COMPUTER SOCIETY

# 4. Assigning labels to pixels

Once the temporal profile parameters for each region are available, we calculate the residual for each pixel/label assignment and then assign a label to each pixel so as to minimize the overall residual. First, we calculate the residuals for each pixel to each label:

$$r_{pl} = r(i_p, \theta_l) \tag{2}$$

The way we evaluate the residual depends on how we model the temporal profile.

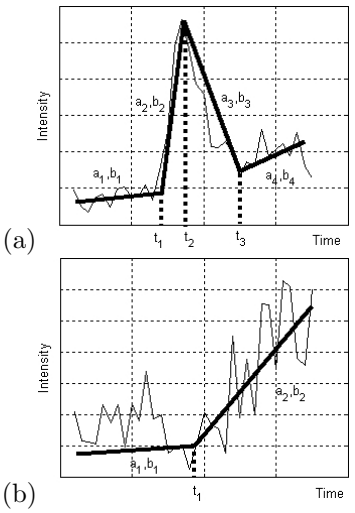



Figure 3: Parameters for MR angiography image series: (a) Artery (b) Vein / Parenchyma

**MR angiography:** Figure 3 depicts the parameters for each label in an MR angiography image series. Note that the number of parameters in $\theta_l$ can be different depending on label $l$. The temporal profile for artery pixels is modelled by 4 line segments. Thus we have 8 parameters ($a_{iA}$'s and $b_{iA}$'s) for the line segments and 3 parameters ($t_{iA}$'s) for the intersection of consecutive line segments, which leads to $\theta_{\text{Artery}} = \{a_{1A}, b_{1A}, t_{1A}, a_{2A}, b_{2A}, t_{2A}, a_{3A}, b_{3A}, t_{3A}, a_{4A}, b_{4A}\}$. Similarly, we have $\theta_{\text{Vein/Parenchyma}} = \{a_{1V}, b_{1V}, t_{1V}, a_{2V}, b_{2V}\}$. The temporal profile of the region whose label is $l$ and whose model is parameterized by $n$ line segments is given by

$$j_{lt} = \begin{cases} a_{1l}t + b_{1l} \text{ if } t < t_1 \\ a_{kl}t + b_{kl} \text{ if } t_{k-1} \le t < t_k, (k = 2, \cdots, n-1) \\ a_{nl}t + b_{nl} \text{ if } t_{n-1} \le t \end{cases}$$

**MR mammography:** According to recent findings [20, 15], malignant lesions are characterized in MR mammography by rapid wash-in and rapid wash-out of a contrast agent. On the other hand, benign lesions

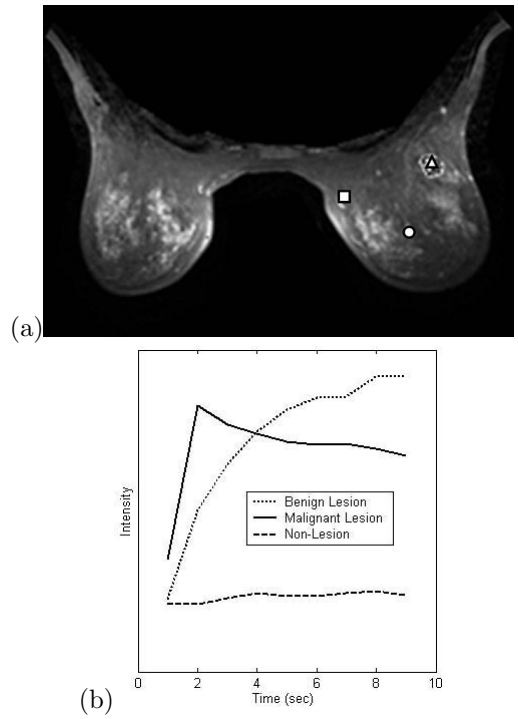



Figure 4: (a) A snapshot from an MR mammography image sequence (b) Temporal profile of the selected pixels

are characterized by rapid wash-in and slow wash-out of contrast agent (see Figure 4). To capture these properties, we model the temporal profile of lesion by two line segments, one for wash-in and the other for wash-out ($\theta_{\text{Malignant-lesion}} = \{a_{1M}, b_{1M}, t_{1M}, a_{2M}, b_{2M}\}$, $\theta_{\text{Benign-lesion}} = \{a_{1B}, b_{1B}, t_{1B}, a_{2B}, b_{2B}\}$). For non-lesions, we use one line as a model of the temporal profile ($\theta_{\text{Non-lesion}} = \{a_{1N}, b_{1N}\}$)(see Figure 5).

Once we have the model of the temporal profile, the residual for assigning the label $l$ to the pixel $p$ is given by

$$r(i_p, \theta_l) = \min_{\beta_p, \Delta t_p} \frac{1}{N} \sum_{t_{Min} \le t \le t_{Max}} \rho(i_{pt}, \beta_p j_{p(t+\Delta t_p)}),$$

where $t_{\text{Min}} = \max(0, -\Delta t_p)$, $t_{\text{Max}} = \min(T, T - \Delta t_p) - 1$, $1/N$ is a normalization constant, $N = t_{\text{Max}} - t_{\text{Min}} + 1$, and $\Delta t_p$, $\beta_p$ are the time displacement and intensity scale of pixel $p$ respectively. We introduce $\Delta t_p$ and $\beta_p$ to offset the variation in time displacement and intensity scale between pixels. For the distance metric $\rho$, we use $\rho(x, y) = \min(|x - y|, K)$ for some constant $K$ following [3]. Based on the residual, we calculate the cost of assigning a label to each pixel:

$$g(i_p, p, \theta_{f_p}) = \frac{e^{-r_{pl}^2/2\sigma^2}}{\sum_{l' \in \mathcal{L}} e^{-r_{pl'}^2/2\sigma^2}}, \tag{3}$$



Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV'03)
0-7695-1950-4/03 $ 17.00 © 2003 IEEE

IEEE COMPUTER SOCIETY

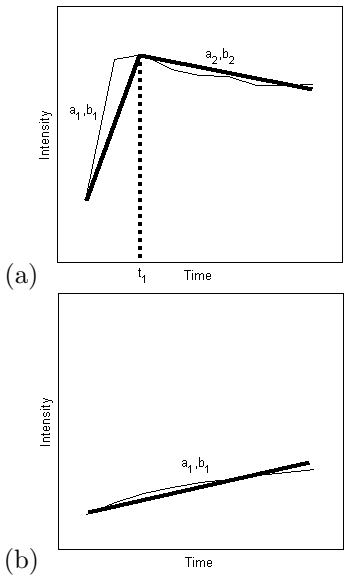



Figure 5: Parameters for MR mammography image series: (a) Lesion (b) Non-Lesion

where $\sigma$ corresponds to the amount of residual expected in the overall data. To find an approximate solution of minimizing $E(f)$, we use the graph cut algorithm of [8]. Since minimizing Eq (1) has been shown to be NP-hard [8], the algorithm is not guaranteed to find the global minimum. Nevertheless, it finds a good local minimum, in the sense that the final energy cannot be lowered by exchanging any subset of pixels having a common label with any other subset of pixels having a common label. After the graph cut algorithm has converged, we find the temporal profile parameter for each label, as explained in the next section.

# 5. Estimating temporal profile parameters

The temporal profile parameters of each label describe how the pixel intensity of a pixel with the label will vary over time. Once the assignment of labels to pixels is known, we can find the average temporal profile of all the pixels with each label. Simply put,

$$\theta_l = \text{ExtractParameters}\left(\sum_{p\in\mathcal{P}, f_p=l} \beta_p i_{p(t+\Delta t_p)}\right), \quad (4)$$

where $\Delta t_p$ and $\beta_p$ are the time displacement and intensity scale of pixel $p$. Again, ExtractParameters depends upon our model of the temporal profile.

**MR angiography:** To estimate the parameters, we first estimate the intersection times ($t_{il}$'s), then estimate the line segment coefficients ($a_{iA}$'s and $b_{iA}$'s) using least squares fitting. More specifically, for the parameters for artery, $t_{1A}$ is the first time index such that intensity at $t_{1A}$ is above $\mu+k\sigma$, where $\mu$ and $\sigma$ are the mean and the standard deviation of the intensities for the time index from 1 to $t_{1A}-1$. We set $k = 2.5$ in our experiments. $t_{3A}$ is estimated similarly except that we examine whether the intensity is above $\mu+k\sigma$ going backwards. $t_{2A}$ is set to be the time index that has the maximum intensity between $t_{1A}$ and $t_{3A}$. For the parameters for veins/parenchyma, $t_{1V}$ is the time index that has the minimum intensity in a predefined time range. Since we have only two labels, we don't actually have to use the multiway-cut algorithm, and can instead use the 2-way cut algorithm from [12], which gives the globally optimal labelling in polynomial time.

**MR mammography:** Similarly to the case of MR angiography data, we use robust least squares fitting to extract parameters for the non-lesion label. We set $t_1$ to be the time index that has the maximum intensity in a predefined range, namely 2 to 4, since it is known that the transition between wash-in and wash-out phases occurs between times 2 and 4 [20, 15].

# 6 Preliminary experiments

In this section, we give some experimental results for our algorithm on MR angiography data and mammography data. In each experiment, we describe how we obtain the data, model temporal profiles, and extract temporal profile parameters, then show the result of our algorithm. An individual image in MR angiography is 2D, while in MR mammography it is 3D. Our approach can handle both 2D and 3D data; however, to simplify matters we will confine our attention to 2D data by looking at single slices in MR mammography.

## 6.1. Trifurcation

Contrast-enhanced MR angiography has become a routine clinical tool for pre-treatment mapping of vasculature [16]. Figure 2 shows an example of MR angiography imagery of the trifurcation. The data typically consists of 30–45 2-dimensional images.

It is difficult to compare our results with previous work, since to our knowledge no other methods have been published for segmenting contrast-enhanced image sequences. However, there is a comparison that we find instructive, in that it shows the importance of performing segmentation based on the temporal pixel histories in the entire image sequence, rather than on a single image. We obtained several series where an experienced radiologist hand-selected a pair of images (pre-contrast and post-contrast), such that the difference



Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV'03)
0-7695-1950-4/03 $ 17.00 © 2003 IEEE

IEEE COMPUTER SOCIETY

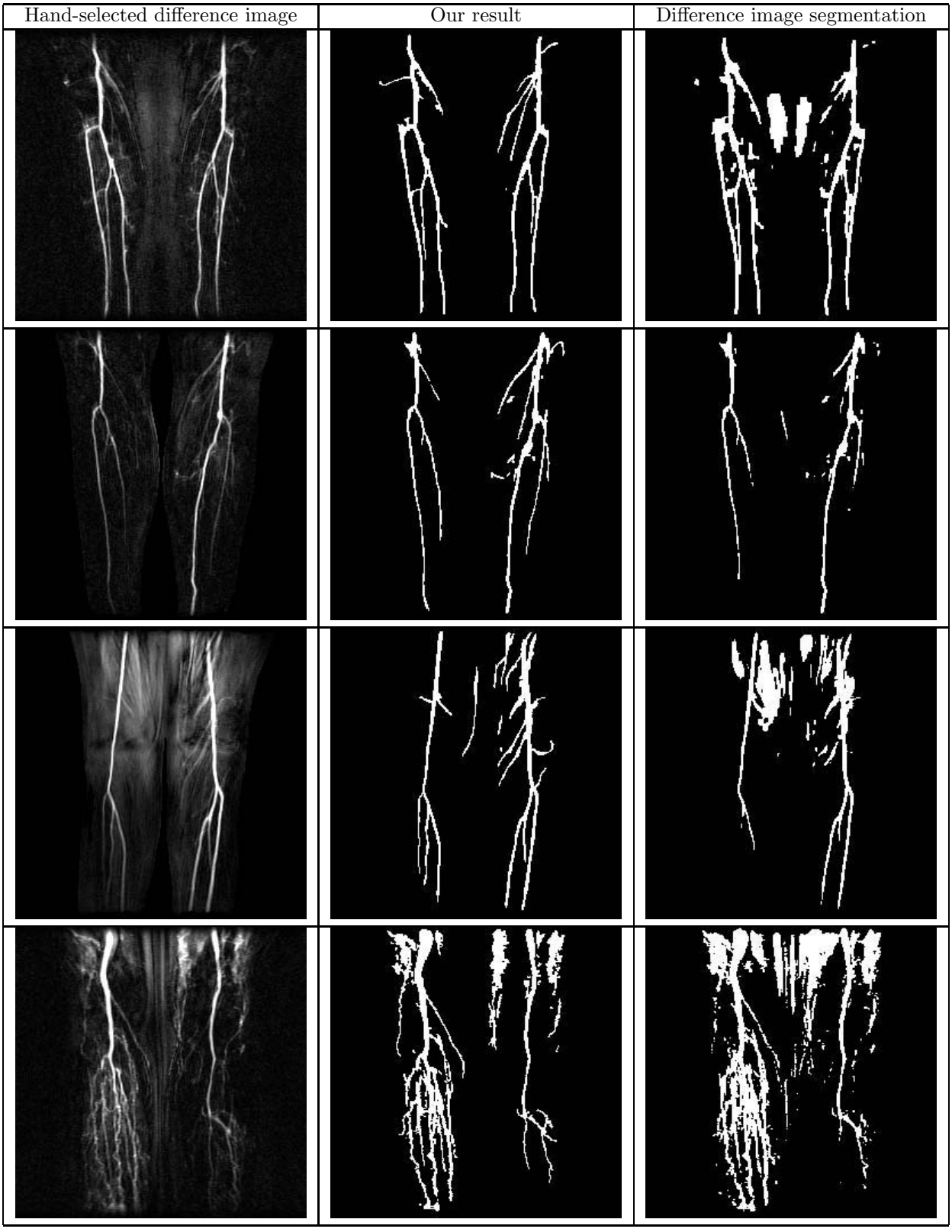



Figure 6: Comparison of our results on MR angiography image sequence with other methods: Left: Best difference image manually selected by an experienced radiologist. Middle: Results from our algorithm. Right: Results from segmentation of the manually selected difference image.



Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV'03)
0-7695-1950-4/03 $ 17.00 © 2003 IEEE

IEEE COMPUTER SOCIETY

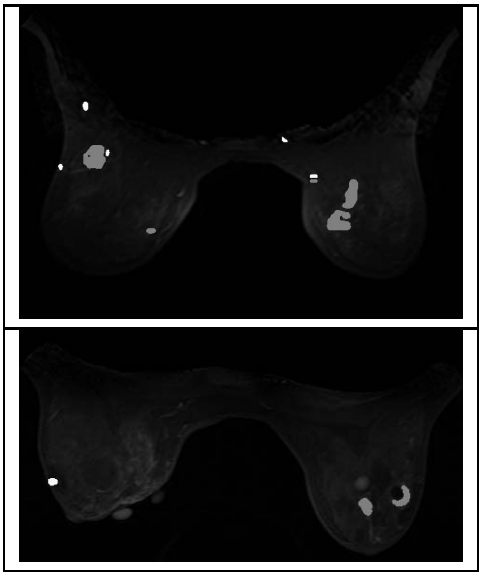


Figure 7: Results of our algorithm. Malignant lesion (in white) and benign lesion (in gray) are overlaid with a darkened input image at time index = 2

image gave the clearest view of the arteries. This manually selected difference image can then be compared to the result of our algorithm. These results are shown in the left and middle columns of Figure 6. Visual inspection by the authors clearly shows that our algorithm gives far less false artery (i.e. a vein/parenchyma pixel misclassified as an artery), which mainly results from patient motion, and far less false vein/parenchyma (i.e. an artery pixel misclassified as a vein/parenchyma), which primarily comes from faint signal because of bad timing. Both of these advantages derive from our algorithm's use of the entire image sequence.

In Figure 6 we have also included an additional segmentation in the rightmost column for the sake of comparison. Here we have applied a binary segmentation algorithm to the hand-generated image in the leftmost column. Since spatial coherence is very important, instead of a simple thresholding scheme we used a simple algorithm based on graph cuts that finds connected groups of bright pixels. We believe that the poor results obtained by this method demonstrates the importance of using the entire image sequence, rather than performing static segmentation on a single image.

### 6.2. Mammography

In MR mammography, a series of 9 images is acquired at approximately 90 second intervals, with a contrast agent injected immediately after the first image acquisition. Figure 7 shows the result of image segmentation overlaid with one input image at a specific time. Currently, we use the second image. Body parts other than breast are removed manually to simplify the parameter estimation process, although it should be relatively easy to automate this removal process given that the image orientation is fixed. The results show that our algorithm avoids being distracted by very bright pixels that occur mainly on arteries. Any image segmentation algorithm based on one single image will have no way to distinguish such pixels from lesion pixels. In these two cases, an experienced radiologist who read the cases without our system believed that the patient suffered from breast cancer; this diagnosis was confirmed by biopsy.

## 7. Extensions

The methods that we have described show promising preliminary results for MR angiography and MR mammography. The experimental results we have described are quite preliminary, and a more careful comparison with the results of other segmentation algorithms should be performed, ideally using some kind of (probably hand-segmented) data as a gold standard.

There are also a number of ways in which these results could probably be improved. For example, the temporal models that we have used in our experiments to date are fairly simple. It is important to note that our algorithm does not require a particular temporal model for pixel behavior, as long as there is some piecewise model where we can pre-specify the number of pieces. There is no restriction that the pieces need to be linear, although piecewise linear models like the ones we have used are obviously the first class that should be explored. It would be interesting to apply more sophisticated temporal models for these problems, such as autoregressive models or pharmacologically derived models.

Our preliminary experiments show promising results for contrast-enhanced image sequences, such as MR angiography and MR mammography, where the patient is usually stationary. However, it would be interesting to extend our approach to image sequences with both contrast and motion, e.g., coronary artery images where the heart is clearly visible and beating as the contrast flows in and out. Currently, our technique is based on the assumption that there is no motion, i.e., the time series for a single pixel corresponds to a single location in the patient. Of course, if we know exactly how each pixel moves, we still can obtain the time series. It may be possible to handle unknown motions by performing the same calculation based on an assumption about the motion of each pixel and then re-estimating the motion. In other words, instead of



Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV'03)
0-7695-1950-4/03 $ 17.00 © 2003 IEEE

IEEE COMPUTER SOCIETY

iterating between pixel labelling and parameter estimation, we can iterate between pixel labelling, parameter estimation, and motion estimation. Of course, the more compactly we can represent the motion (i.e., the fewer motion parameters we have), the more likely it is that this algorithm will converge to the desired solution.

**Acknowledgments**

We thank Martin Prince, MD and Ruth Rosenblatt, MD for numerous discussions on these topics and for providing the images shown. We also thank Amy Gale for proofreading, and the anonymous reviewers for their constructive critiques. This work was supported by NSF grants IIS-9900115 and CCR-0113371 and by a grant from Microsoft Research.

# References

[1] Serge Ayer and Harpreet Sawhney. Layered representation of motion video using robust maximum-likelihood estimation of mixture models and MDL encoding. In *International Conference on Computer Vision*, pages 777–784, 1995.

[2] S. Birchfield and C. Tomasi. Multiway cut for stereo and motion with slanted surfaces. In *International Conference on Computer Vision*, pages 489–495, 1999.

[3] M. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1):75–104, Jan. 1996.

[4] A.F. Bobick and J.W. Davis. The recognition of human movement using temporal templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(3):257–267, March 2001.

[5] Jean-Yves Bouguet and Pietro Perona. 3D photography on your desk. In *International Conference on Computer Vision*, pages 43–52, 1998.

[6] Patrick Bouthemy and Edouard François. Motion segmentation and qualitative dynamic scene analysis from an image sequence. *International Journal of Computer Vision*, 10(2):157–182, April 1993.

[7] Yuri Boykov and Vladimir Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in computer vision. In *Proc. EMMCVPR*, volume 2134 of *Lecture Notes in Computer Science*, pages 359–374. Springer-Verlag, September 2001.

[8] Yuri Boykov, Olga Veksler, and Ramin Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, November 2001.

[9] J. Cromwell, W. Labys, and M. Terraza. *Univariate Tests for Time Series Models*. Sage Pub, 1994.

[10] Brian Curless and Marc Levoy. Better optical triangulation through spacetime analysis. In *International Conference on Computer Vision*, pages 987–994, 1995.

[11] A. Dempster, N. Laird, and D. Rubin. Maximum likelihood from incomplete data using the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39:1–38, 1977.

[12] D. Greig, B. Porteous, and A. Seheult. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society, Series B*, 51(2):271–279, 1989.

[13] John H. Juhl, Andrew B. Crummy, and Janet E. Kuhlman. *Paul and Juhl's Essentials of Radiologic Imaging*. Lippincott Williams & Wilkins, 1998.

[14] T. Kanade, A. Gruss, and L. Carley. A very fast VLSI rangefinder. In *IEEE International Conference on Robotics and Automation*, pages 1322–1329, 1991.

[15] G. P. Liney, P. Gibbs, C. Hayes, M. O. Leach, and L. W. Turnbull. Dynamic contrast-enhanced MRI in the differentiation of breast tumors: user-defined versus semi-automated region-of-interest analysis. *Magnetic Resonance Imaging*, 10(6):945–949, 1999.

[16] M. R. Prince, T. M. Grist, and J. F. Debatin. *3D Contrast MR Angiography*. Springer-Verlag, New York, 1998.

[17] Kari Pulli, Habib Abi-Rached, Tom Duchamp, Linda G. Shapiro, and Werner Stuetzle. Acquisition and visualization of colored 3D objects. In *International Conference on Pattern Recognition*, pages 11–15, 1998.

[18] P. Saisan, G. Doretto, Y. Wu, and S. Soatto. Dynamic texture recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 58–63, 2001.

[19] Steven M. Seitz and Chuck R. Dyer. View-invariant analysis of cyclic motion. *International Journal of Computer Vision*, 25(3):231–251, 1997.

[20] S.Sinha, F.A.Lucas-Quesada, N.D.DeBruhl, J.Sayre, D.Farria, D.P.Gorczyca, and W.Bassett. Multifeature analysis of Gd-enhanced MR images of breast lesions. *Journal of Magnetic Resonance Imaging*, 7:1016–1026, 1997.

[21] John Wang and Edward Adelson. Representing moving images with layers. *IEEE Transactions on Image Processing*, 3(5):625–638, September 1994.

[22] Yair Weiss and Edward Adelson. A unified mixture framework for motion segmentation: Incorporating spatial coherence and estimating the number of models. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 321–326, 1996.

[23] Li Zhang, Brian Curless, and Steven M. Seitz. Rapid shape acquisition using color structured light and multi-pass dynamic programming. In *Symposium on 3D Data Processing, Visualization, and Transmission*, pages 19–21, 2002.

[24] Li Zhang, Brian Curless, and Steven M. Seitz. Spacetime stereo: Shape recovery for dynamic scenes. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 367–274, 2003.



Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV'03)
0-7695-1950-4/03 $ 17.00 © 2003 IEEE

IEEE COMPUTER SOCIETY