Available online at www.sciencedirect.com
ScienceDirect
Procedia Social and Behavioral Sciences 8 (2010) 640–647
Procedia Social and Behavioral Sciences
ELSEVIER

International Conference on Mathematics Education Research 2010 (ICMER 2010)

Segmentation of Masses from Breast Ultrasound Images using Parametric Active Contour Algorithm

Abdul Kadir Jumaat[a,*], Wan Eny Zarina Wan Abdul Rahman[b], Arsmah Ibrahim[c], Rozi Mahmud[d]

[a]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Negeri Sembilan, Malaysia
[b,c]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Malaysia, Shah Alam, Malaysia
[d]Faculty of Medicine, Universiti Putra Malaysia

Abstract

The active contour (Snake) is a computer generated curve that can trace boundaries of images. As a method which applies the computer technology in mathematics, Snake is computationally formulated based on controlled continuous splines and adopts the mathematical concept of energy minimization. This paper presents the application of Snake for the segmentation of masses on breast ultrasound images. The images used are taken from Malaysian population. The boundaries of the masses identified may be used in classification of cancers or non-cancerous masses. Specifically the Balloon Snake is applied in segmenting the masses in the breast ultrasound images. Comparison on the masses areas segmented by the Balloon Snake is done against the areas traced by an expert (radiologist). It is found that from forty-five masses tested, the average percentage area difference of Balloon Snake is 4.47%. This implies that the accuracy of segmentation results for the Balloon Snake is 95.53%.

Keywords**:** Active Contour Algorithm; Balloon Snake; Segmentation; Breast Ultrasound Image

1. Introduction

Breast cancer is a type of cancer of the glandular tissue which typically begins in the milk duct of the breast. It starts with mutations or changes in the normal cell's genome or Deoxyribonucleic Acid (DNA). The mutant cell behaves abnormally by dividing without control and form a mass of cells. A mass can be benign which a non cancerous condition is or malignant which is a cancerous condition. A benign mass do not spread to other parts of the body but still may need to be removed because the local tissue may be damaged. On the other hand, a malignant mass can destroy neighbouring tissues and spread to other parts of organ or body (Breast Cancer, 2009; MEDtropolis, 2008).

According to *Pusat Sumber Payudara* (PSP), College of Radiology, Academy of Medicine of Malaysia in Universiti Malaya (PSP, 2007), breast cancer is the most common type of cancer in Malaysian women among all races from the age of 20 years. Early detection for breast mass and getting an appropriate medical attention can save lives. Unfortunately, nearly 40% of the new cases of breast cancer identified each year in Malaysia were already in the critical stage (Pride, 2009). Tabar et al. (2001) have demonstrated that survival rate is greatly improved if breast

* Corresponding author. Tel.: +6-017-303-9601; fax:+6-06-484-2521.
E-mail address: abdulka@ns.uitm.edu.my

doi:10.1016/j.sbspro.2010.12.089

abnormalities such as breast masses are detected at the earliest stage which is concurred by Andersson and Ryden (2001).

Among the screening modalities that can be used to detect breast masses at the earliest stage and to diagnose breast cancer are Magnetic Resonance Imaging (MRI), Mammography, and Ultrasound. Currently, mammography is considered as the “gold standard” or the first line of defence in early breast cancer detection as reported by Rim (2002), American Cancer Society (2007) and Hamid (2008). Ultrasound and MRI are supplementary tools for detecting breast masses or other breast abnormalities that are difficult to detect on mammography. Even though mammography technique is considered as a “gold standard”, it is not effective for women whose breasts are too dense with glandular tissues because abnormalities such as masses may be hidden behind these tissues (Radiologi, 2006). Furthermore, mammography is not suitable for women with silicone breast implants as the x-ray used in this technique is not transparent to the breast implants that can block anything behind them. In addition to this, mammography is also not suitable for pregnant women because the x-ray radiation from mammography can harm the foetus (Ikeada, 2004; Radiology, 2006; Cancer Research UK, 2007). Ultrasound is cheaper compared to MRI technique, gives clear pictures of soft tissue for dense breasts and breast implant. Besides it uses no ionizing radiation and overcome the limitation of mammogram (Stanford Cancer Center. 2006). Ultrasound has been successfully used to distinguish malignant from benign masses as reported by Cho, K.R., et al 2004.

In order to determine whether a mass is benign or malignant, radiologists will examine the output images produced by the screening modalities. For a radiologist, an image produced is an information carrier. However, the information may be corrupted by noise or may be tied up with irrelevant information, resulting in difficulties for them to interpret the important information from the image using the conventional method (human eye only). In such cases, a new area called image processing has been developed by combining the computer technology with some mathematical conceptions. Image processing techniques are used to extract information from the image. The first and the most important step in image processing is to segment the image. It is done in order to delineate the foreground (masses region) and the background (Tabrizi, 2003). Tabrizi classified segmentation techniques into low level and high level techniques. The low-level techniques are fast and simple to be implemented compared to high level techniques.

The low-level segmentation techniques are known to be fast and simple, but these methods simply analyze an image by reducing the amount of data to be processed. This problem can result in loss of important information. Moreover, the low-level segmentation techniques may incorrectly identify region or boundary of an object due to the distraction of noise in an image (Mclnerny & Terzopoulos, 1996). The boundary of the abnormality should be identified accurately so that all of the important information required by the radiologist from the object such as shape, margin, and area can be determined. In order for the image to be interpreted accurately, the image must be segmented accurately into regions that correspond to objects or parts of an object.

In order for the ultrasound image to be interpreted accurately, the image will be segmented by the high-level segmentation techniques. High-level techniques imply that a priori knowledge about shape, texture, color, or position of the object in question is included in the segmentation procedure.

The iterative algorithm namely active contours were proven to be the effective high level techniques in line and edge detection, image segmentation, shape modelling, and motion tracking as claimed through research carried out by Kass et al. (1986), Cohen (1993), Tabrizi (2003), and Sandberg (2005). Active contours or Snakes are computer-generated curves that move within an image to find the object boundary. As a method which applies the computer technology in mathematics, Snake is computationally formulated based on controlled continuous splines and adopts the mathematical concept of energy minimization. Snake was originally introduced by Kass et al. in 1986 and the notion of Snake for active contours was inspired by the way snakes move, slithering while minimizing their energy (Kass et al., 1986).

There are two types of active contours namely parametric and geometric active contours (Uppu, 2006). Parametric active contours model are represented explicitly as parametric curves. The original active contour is a parametric model. On the other hand, geometric model is based on the theory of curve evolution and geometric flows. The curve has a unique attribute where it can automatically handle topological changes in an image by splitting and merging during its deformation (Caselles et al., 2004). Parametric active contour is not only a simple method but also easy to interact with users. Moreover, Uppu (2006) notes that the computational complexity of parametric active contour model is lower than that based on geometric active contours.

Parametric active contour (PAC) moves under the influences of the internal energy function within the curve itself and the external energy function which are derived from the image data (Xu & Prince, 1997). There two types of internal energy function apply in PAC, namely the elasticity function and rigidity function (Xu & Prince, 1997). Elasticity energy function is designed to keep the active contours contracted or hold together while rigidity energy function keep the contours smooth by avoiding it from bending too much (Xu & Prince, 1997; Tabrizi, 2003).

Although the original active contour has been found in many applications, it intrinsically has a small capture range and the convergence of the algorithm is mostly dependent on the initial position. Moreover, it also has difficulties in progressing into boundary concavities (Xu & Prince, 1997). Therefore, the Balloon Snake was established in 1991 by Cohen in order to overcome the limitations of the original active contour. Balloon Snake is extensively used for segmentation applications such as segmentation of left ventricle of a human heart from MRI image (Cohen and Cohen, 1993; Xu, 1998), brain (Ivins, 1996), middle ear image from Magnetic Resonance Microscopy (MRM) image (Tabrizi 2003), and teeth images (Garcia et al. 2005). In this paper, the accuracy of Balloon Snake in segmenting masses on breast ultrasound images is evaluated. The images used are taken from Malaysian population. In order to identify the accuracy of the segmentation by the Balloon Snake, masses area of the segmented images by the method are computed and compared with the actual masses area segmented by the expert radiologist.

2. Approach and Methods

Snake is a vector valued function in the spatial domain of an image and parametrically expressed as $\mathbf{v}(s) = (x(s), y(s))$ where $0 \leq s \leq 1$. A curve consists of *n* vertices *v* connected by straight lines. The parameter *x* and *y* are the coordinate of the vertices, *v* and are functions of the normalized arc length *s*. The Snake has a dynamic behavior that deforms from an initial position and converges to the boundary of the object in the image. It moves through the domain of the image by minimizing its energy function, E_{snake} which is defined as

$$E_{snake} = \int_0^1 [E_{\text{int}}(\mathbf{v}(s)) + E_{ext}(\mathbf{v}(s))]ds \tag{1}$$

The internal energy function is

$$E_{\text{int}}(\mathbf{v}(s)) = \frac{1}{2}\{\alpha \,|\, \mathbf{v}'(s)\,|^2 + \beta \,|\, \mathbf{v}''(s)\,|^2\} \tag{2}$$

The internal energy function, $E_{\text{int}}(\mathbf{v}(s))$ is computed based on the local shape of the curve **v**(*s*), and is responsible in determining the continuity and the smoothness of the curve. The parameter α and β are the coefficient of the internal energy function. The parameter α is the elasticity parameter. For a large value of α , the curve becomes very straight between two points. The parameter β is the rigidity parameter and for a large value of β the curve becomes smooth. On the other hand, the external energy function $E_{ext}(\mathbf{v}(s))$ is derived based on the image information and it drives the curve to the boundary of the object. Different types of Snakes use different type of external energy function. By calculus of variation, Equation (1) is minimized by solving the associate Euler's Equation as follows;

$$-\alpha\mathbf{v}''(s) + \beta\mathbf{v}^{(4)}(s) + \nabla E_{ext}(\mathbf{v}(s)) = 0 \tag{3}$$

The Euler's equation is approximated with finite difference method since it is difficult be computed analytically. Converted to vector notation with $\mathbf{v}_i = (x_i, y_i)$, now Equation (3) becomes

$$\alpha_i(\mathbf{v}_i - \mathbf{v}_{i-1}) - \alpha_{i+1}(\mathbf{v}_{i+1} - \mathbf{v}_i) + \beta_{i-1}(\mathbf{v}_{i-2} - 2\mathbf{v}_{i-1} + \mathbf{v}_i) - 2\beta_i(\mathbf{v}_{i-1} - 2\mathbf{v}_i + \mathbf{v}_{i+1}) + \beta_{i+1}(\mathbf{v}_i - 2\mathbf{v}_{i+1} + \mathbf{v}_{i+2}) + \left(\frac{\partial}{\partial x}E_{ext}(x,y), \frac{\partial}{\partial y}E_{ext}(x,y)\right) = 0 \tag{4}$$

In matrix form

$$A\mathbf{x} + \frac{\partial}{\partial x}E_{ext}(x,y) = 0, \; A\mathbf{y} + \frac{\partial}{\partial y}E_{ext}(x,y) = 0 \tag{5}$$

where $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{bmatrix}$

A is a pentadiagonal banded matrix and the size of the matrix is nxn where n being the number of vertices. For example, for a Snake consist of 5 vertices, matrix A is as follows

$$A = \begin{bmatrix} 2\alpha+6\beta & -\alpha-4\beta & \beta & 0 & 0 \\ -\alpha-4\beta & 2\alpha+6\beta & -\alpha-4\beta & \beta & 0 \\ \beta & -\alpha-4\beta & 2\alpha+6\beta & -\alpha-4\beta & \beta \\ 0 & \beta & -\alpha-4\beta & 2\alpha+6\beta & -\alpha-4\beta \\ 0 & 0 & \beta & -\alpha-4\beta & 2\alpha+6\beta \end{bmatrix}$$

In order to solve the matrix equation, the right hand sides of Equation (5) is set equal to the product of a step size γ and the negative time derivatives of the left hand sides (Kass et al, 1996). Now, Equation (5) becomes

$$\begin{aligned} A\mathbf{x}_t + \frac{\partial}{\partial x}E_{ext}(x_{t-1}, y_{t-1}) &= -\gamma(x_t - x_{t-1}) \\ A\mathbf{y}_t + \frac{\partial}{\partial y}E_{ext}(x_{t-1}, y_{t-1}) &= -\gamma(y_t - y_{t-1}) \end{aligned} \tag{6}$$

where subscript t is the iteration number. By matrix inversion I, Equation (6) becomes

$$\begin{aligned} \mathbf{x}_t &= (A+\gamma I)^{-1}(\gamma\mathbf{x}_{t-1} - k\frac{\partial}{\partial x}E_{ext}(x_{t-1}, y_{t-1}) \\ \mathbf{y}_t &= (A+\gamma I)^{-1}(\gamma\mathbf{y}_{t-1} - k\frac{\partial}{\partial y}E_{ext}(x_{t-1}, y_{t-1}) \end{aligned} \tag{7}$$

The weighting parameter k is an additional parameter to control the internal and external energy functions. When the right hand side of Equation (7) equals to zero, then the location or position of the vertices in the successive iteration is unchanged. As proposed by Cohen (1991), the Balloon Snake uses the sum of pressure energy and image energy as the external energy function, $E_{ext}(\mathbf{v}(s))$ defines as follows

$$E_{ext} = -k \frac{F_{image}}{\| F_{image} \|} + k_{pressure}\mathbf{n}(s) \tag{8}$$

The parameter $k_{pressure}$ is the pressure weight or pressure energy and its positive or negative sign causes the Snake to inflate or deflate respectively. The image energy F_{image} is the gradient of the image edge map. The parameter k is the image energy weighting. The symbol $\mathbf{n}(s)$ represents the unit normal vector to vertices. In order to determine the vertex position using Balloon Snake, Equation (7) which represents the vertex position using the original Snake is modified as follows:

$$\begin{aligned} \mathbf{x}_t &= (A+\gamma I)^{-1}(\gamma \mathbf{x}_{t-1} \pm k_{pressure}\mathbf{n}(x_{t-1}) - k\frac{\partial}{\partial x}F_{image}(x_{t-1}, y_{t-1})) \\ \mathbf{y}_t &= (A+\gamma I)^{-1}(\gamma \mathbf{y}_{t-1} \pm k_{pressure}\mathbf{n}(y_{t-1}) - k\frac{\partial}{\partial y}F_{image}(x_{t-1}, y_{t-1})) \end{aligned} \tag{9}$$

3. Implementation

Fourty breast ultrasound images from Malaysian population are obtained from Palace of the Golden Horses Screening Center. The images are cropped in Adobe Photoshop CS2 to the size of 64 pixels by 64 pixels. Cropping is done around the region of interest (ROI) which contains breast masses. The ROI images are first preprocessed using median filter to remove speckle noise, and then Histogram stretching method is applied to enhance the contrast of the images. Some parameter values are standardized experimentally for all the images. The image edge map is computed using Canny edge detector with threshold value used is 0.3. This value is suitable for the ROI images because image information will reduce if the value is higher whereas noise will increase if the threshold value is decreased. The number of iterations for Snakes deformation is standardized to 500 which are enough for ROI images of size 64 by 64 pixels. The pressure energy for Balloon Snake is 0.08 whereby positive number indicates inflation of the Snake. As recommended by Xu & Prince (1999), the values of parameters α , β and γ used are 0.15, 0, and 1 respectively. The Matlab R2008a software is used to implement the median filter, histogram stretching, and the deformation of Balloon Snake. In order to identify the accuracy of the edge segmentation, the areas of the segmented masses by the Balloon Snake are compared with the true pixel area value. The true pixel area value is obtained from the average of two pixel area values of a mass traced by expert radiologist in each image. The unit of the area used is in $pixel^2$.

4. Result

Figure 1 shows the four samples of original ROI of real breast ultrasound images that contain breast masses in part (a) and the corresponding segmented images in part (b). The samples are taken from the original ROI number 4,9,18, and 19. True area of the mass in part 4(a) is 628.40pixel^2 while the segmented area by Balloon Snake in part 4(b) is 629.21pixel^2 resulting to 0.13% difference. The percentage difference of masses areas of the remaining images in Figure 1 are listed in Table 1.

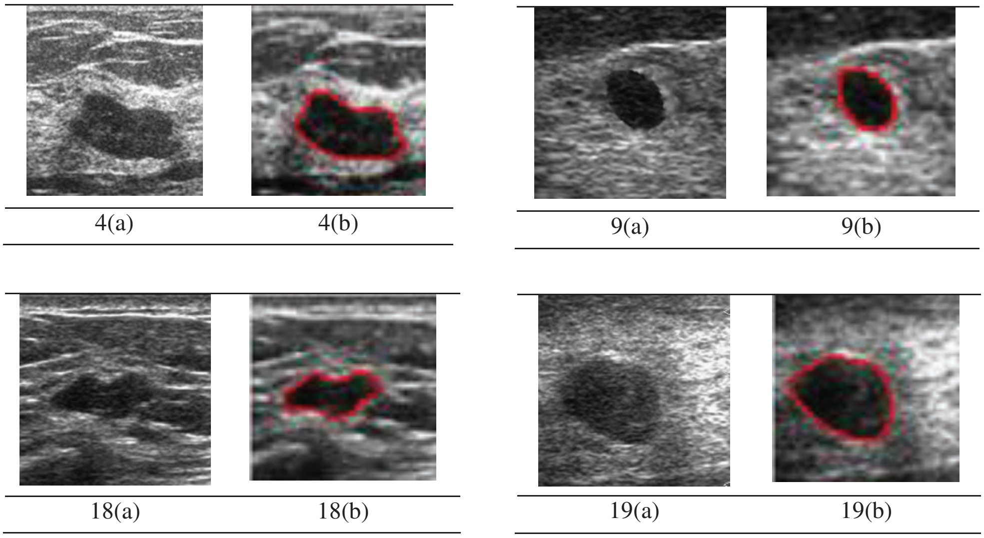

Figure 1: (a) Original ROI (b) Balloon Snake

Table 1. Percentage Difference of Segmented Masses

Image	**True Area** (*$Pixel^2$*)	**Balloon Snake Area** (*$Pixel^2$*)	**Percentage Difference** (%)
9	305.205	307.28	0.68
18	191.94	179.05	6.72
19	293.83	294.73	0.31

We have tested on 45 ROI images and the result is tabulated in the bar chart in Figure 2 as follows

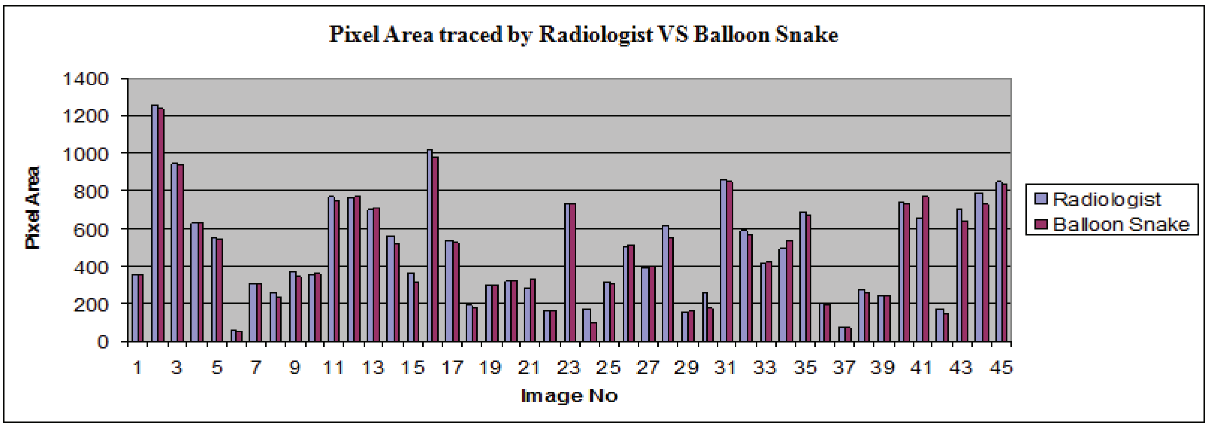

Figure 2: Bar Chart on Pixel Area traced by Radiologist versus Balloon Snake

Based on the bar chart, the average percentage area difference for all 45 images used in this experiment is found to be 4.47%.

5. Conclusion

In this paper, we have attempted to segment masses on the breast ultrasound images using Balloon Snake by combining the mathematical optimization conception together with the computer technology. Segmentation on the images shows that the average percentage area difference of Balloon Snake is 4.47% which mean 95.53% accurate. In the view of the radiologist, the average percentage area differences produced by the Balloon Snake are small resulting to almost accurate boundary segmentation which implies that the Balloon Snake is good in segmenting the masses. The boundary of the masses identified will assist radiologists in locating potentially cancerous cases for further analysis such as the classification of malignant or benign masses, or in tissue sampling in the biopsy procedure.

We recommend to the future researcher to invent a method with higher degree of accuracy in medical image segmentation by bringing more technology in mathematics. It has to be understood that the works in medical image processing generally or medical image segmentation specifically are not trying to substitute the appreciation of the physicians or radiologist, but to introduce approaches in order to improve the diagnosis by giving second opinion to the radiologist that hopefully reduce the requirement of biopsy which is expensive, time-consuming and uncomfortable technique.

References

American Cancer Society (2007). Mammograms and Other Breast Imaging Procedures. Atlanta, Ga: Retrieved 1st March 2008 from http://www.cancer.org

Andersson, I., and Ryden, S. (2001). *Early Detection and Prevention: Benefits, Costs and Limitations of Screening*. Breast Cancer: New Horisons in Research and Treatment, Tobias, J. S.; Houghton, J.; and Henderson, I. C. eds. London: Arnold. 105-117.

Breast Cancer (2009). Retrieved 27 April 2009 from www.breastcancer.org.

Cancer Research UK (2007). Breast Cancer Test. United Kingdom. Retrieved on 1st March 2008 from http://www.cancerhelp.org.uk

Caselles V, Catte F, Coll T & Dibos F (1993): "A geometric model for active contours," Numerical Mathematic, (66), 1-31

Cho, K.R., et al (2004). "A comparative study of 2D and 3D ultrasonography for evaluation of solid breast masses", European Journal of Radiology 54 (2005) 365-370

Cohen,L.D. (1991). On Active Contour Models and Balloons, Computer Vision, Graphics, and Image Processing: *Image Understanding*. 211-218

Cohen, L.D., & Cohen, I. (1993): "Finite-element Methods for Active Contour Models and Balloons for 2-D and 3-D images," *IEEE Trans. on PAMI*, 15 (11), 1131-1147.

Garcia.R.L., Estepar.R.S.J., Lopez.C.A., (2005). Pull-Push Level Sets: A new term to encode prior knowledge for the segmentation of teeth images. *Medical Imaging 2005: Image Processing.* University of Valladolid, Valladolid, Spain

Hamid R. (2008): "Pattern Recognition, Computer Vision, Medical Imaging, Machine Intelligence", Faculty of Engineering University of Waterlo. Retrieved on 1st March 2008 from (http://pami.uwaterloo.ca/tizhoosh/index.html)

Ikeada, D. M. (2004). *Breast Imaging*. California: Stanford University School of medicine.

Ivins.J.P., (1996). Statistical Snakes: Active Region Models. Artificial Intelligence Vision Research Unit, Department of Psychology, University of Sheffield, England.

Kass M, Witkin A & Terzopoulos D (1986): "Snakes: Active Contour Models," *International Journal of Computer Vision*, 3, 321-331.

McInerney T & Terzopoulos D (1996): "Deformable models in medical image analysis," *IEEE Proceedings of the Workshop on Mathematical Methods in Biomedical ImageAnalysis*, 171-180.

MEDtropolis (2009). What is cancer. Retrieved 28 April 2009 from http://medtropolis.com/

Pride foundation pink ribbon deeds statistics. (2009). Retrieved 29 April 2009 from http://www.pride.org.my/index.php

PSP (2007). Kanser Payudara-Fakta & Statistik.). Pusat Sumber Payudara, College of Radiology, Academy of Medicine of Malaysia, Universiti Malaya. Retrieved on 1st March 2008 from http://www.radiologymalaysia.org

Radiology (2006). "Breast Ultrasound", Radiological Society of North America. Retrieved 1st March 2008 from http://www.radiologyinfo.org

Rim,A.S. (2002). "Breakthroughs in Early Diagnosis and Prevention", WebMD University. Retrieved 1st March 2008 from (http://www.medicinenet.com)

Sandberg.K.(2005). Visualizing Calculus: The Use of the Gradient in Image Processing. University of Colorado, Boulder

Stanford Cancer Center (2006), Diagnostic Radiology Services, Stanford University, School of Medicine. Retrieved on from www.cancer.stanfordhospital.com

Tabar, L.; Vitak, B.; Chen, H-HT.; Yen, M-F.; Duffy, S. W.; and Smith, R. A. 2001. Beyond Randomized Trials: Organized Mammographic Screening Substantially Reduces Breast Carcinoma Mortality. Cancer 91:1724 –1731.

Tabrizi.J.H. (2003). Using Active Contours for Segmentation of Middle-Ear Images. Jan 2003, McGill University, Quebee

Uppu,S.K. (2006). A Study of Parametric Active Contours for the Application of MR Axial Abdominal Images. August 2006, Madurai Kamaraj University, Madurai, India

Xu,C.,&Prince,J.L. (1997). Gradient Vector Flow: A New External Force for Snakes. *Proc. IEE Conf. on Computer Vision.* 66-71, The John Hopkins University, Baltimore.

Xu.C.,(1999): Deformable Models With Application to Human Cerebral Cortex Reconstruction from Magnetic Resonance Image. The John Hopkins University, Baltimore.