ARIMA

Computerized detection and smoothing of contour in mammograms

Cyrille Feudjio[†‡] — Alain Tiedeu[†] — Mihaela Gordan[§] — Samuel Domngang[¶] — Aurel Vlaicu[§]

[†] LETS, National Advanced School of Engineering
PO Box: 8390 Yaoundé – Cameroon
cyrille.feudjio@gmail.com, alain_tiedeu@yahoo.fr

[‡] LAIA, FOTSO-Victor University Institute of Technology
PO Box: 134 Bandjoun–Cameroon

[§] CTMED, Technical University of Cluj-Napoca
400020 Cluj-Napoca–Romania
Mihaela.Gordan@bel.utcluj.ro, Aurel.Vlaicu@com.utcluj.ro

[¶] Faculty of Science, University of Yaoundé I,
PO Box 814, Yaoundé–Cameroon
sdomngang@yahoo.com

ABSTRACT. This paper deals with an approach of detecting and smoothing breast contour in mammograms. Breast contour is a precious primitive for region description, registration, classification and length estimation. In most of the computerized aided detection (CADe) algorithms for breast cancer, breast contour is a needed component. In these applications, a good detection and smoothing of the contour is needed. In mammograms, breast contour detection is a difficult task to achieve due to the fact that its boundary is not clearly visible. The approach developed in this paper is based on global segmentation using Otsu's method and smoothing is based on Fourier descriptors. The results yielded from this approach show the effectiveness of the method on detecting and smoothing contours.

RÉSUMÉ. Cet article présente une approche de détection et lissage de contour du sein dans les mammographies. Il est une primitive importante pour la description de région, le recalage, la classification et l'estimation des longueurs. Il est utilisé dans plusieurs algorithmes de détection assisté par ordinateur (CADe) du cancer du sein. Dans ces applications, une bonne détection et lissage du contour est nécessaire. La détection de contour du sein dans les mammogrammes est difficile à réaliser car sa frontière n'est pas clairement visible. L'approche développée dans ce papier est basée sur la segmentation globale d'Otsu et le lissage à l'aide des descripteurs Fourier. Les résultats de cette approche montrent l'efficacité et la précision de la méthode à la détection et au lissage du contour.

KEYWORDS : mammograms contour detection, smoothing, Fourier descriptors

MOTS-CLÉS : mammogramme, détection de contour, lissage, descripteurs de Fourier

1. Introduction

Mammography is the standard method used in early breast cancer detection. Nowadays, the amount of mammograms to be analyzed by radiologist is continually increasing whereas the number of expert in this area remains weak particularly in developing countries. CADe systems have been introduced to relieve radiologist's workload by providing to them a first aided opinion. Usual steps in CADe systems are image enhancement, segmentation, registration and classification [12, 2]. Registration is important for a better integration and interpretation of information contained in images and for decision making. It uses pattern like shape, area, and contour.

Contour is needed in several applications like nipple detection and registration [13, 10] which are first steps of CADe. The contour is thus a precious attribute of the region's information. Smoothing is performed to attenuate noise and quantization error along the region's boundary [8]. Commonly, after segmentation, the contour of an object in the image is crisp due to quantification or extended due to noise in the boundary which makes the object unreliable for registration and classification tasks.

The aim of this work is to look for a suitable segmentation which consists of determining the area of the region of interest and extract and smoothing the contour which implies contour curve reconstruction. Several authors have reported various methods of segmentation that preserve object boundary in the image by using rational wavelets, gradient, and recursive filters [8, 16]. These methods are iterative, time consuming and sensitive to noise in non identifiable edge areas and yield abnormal shifts on the edge from its true position during segmentation. It is therefore necessary to smooth the edge curve in order reduce these effects.

The basic idea of smoothing contour consists of predicting the curve that should best fit the true edge of the object. Some authors [16, 3] have addressed this problem by building up iterative algorithms which slightly change the position of pixels on segmented contour until a convergence. Smoothing the pixels' contour which are far from the true boundary becomes hard or may require more iteration.

Active contours present advantage of providing a smooth contour during segmentation. This method is often dependent to initialization and in some applications, a prior segmentation must be done to initialize the active contours. In contour detection in mammograms, this method may fail to detect the true boundary due of poor contrast and edge which is not clearly visible [15]. Furthermore, it is an iterative algorithm and obtaining a good results is time consuming which can impose limitation in practical applications [5].

The approach developed in this paper is oriented towards a segmentation method based on global thresholding and accurate smoothing using Fourier descriptors. This method is not iterative and the problem of discrepancy to the true edge is then solved since the smoothing of contour is based on the points of the segmented contour. The rest of the paper is organized as follows: segmentation process, Fourier descriptors and contour smoothing are developed in section 2. Section 3 presents results of the approach and some error estimates. Section 4 concludes the paper.

2. Contour extraction and smoothing

In mammogram analysis, breast contour detection is the main prerequisite used in nipple detection, breast thickness estimation and breast deformation modeling [6, 17] which are different steps of breast cancer CADe. Contour contains the main information on shape and location of objects in the image. Edge defines object's contour and its extraction is based on features such as gray, color or texture discontinuities.

Most edge detection algorithms such as Prewitt, Sobel and Roberts are based on gradient value. As the threshold value in these methods are often empirically determined, over estimation occurs and it is possible to lose some edges. Because of noise and slight variation of gradient along the boundary as in biomedical images, the previous mentioned methods do not give satisfactory results. Estimating contour in these cases is difficult to achieve since accurate smoothing heavily depends on edge detected. Therefore good segmentation which takes into account the noisiness of the boundary without losing much information is needed.

The goal of segmentation is to convert the input grayscale image into a two levels of representation: foreground and background. Amongst the segmentation techniques, the most popular used to address this task are global and local binarization methods. Local binarization methods [1, 14] try to overcome these problems by computing a threshold for each pixel by using information from the local neighborhood of the reference pixel. These methods are able to achieve good results for degraded images but they are slow since the computation is conducted from the local neighborhood of each image pixel. A global binarization method like that of Otsu [11] tries to set a single threshold value for the whole image. This method is fast to compute and the segmentation results allow contour extraction without losing the boundary.

2.1. Global thresholding segmentation

Given an image in which the intensity of a pixel at the position (i, j) is $I(i, j)$ with its value in the range $[0, 255]$. The global thresholding technique aims to compute a single threshold t for all pixels contain in image such that:

$$O(i,j) = \begin{cases} 0 & \text{if } I(i,j) \leq t \\ 255 & \text{otherwise} \end{cases} \qquad [1]$$

The global binary segmentation technique like that of Otsu [13], finds the threshold by minimizing the within group variance of two group of pixels. The within-group variance is defined as followed.

$$\sigma_w^2(t) = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t) \qquad [2]$$

$q_1(t)$ and $q_2(t)$ are the probability of each group to have values lower and greater than t and defined as followed:

$$\begin{array}{l} q_1(t) = \sum_{k=1}^{t} P(k) \\ q_2(t) = \sum_{k=t+1}^{L} P(k) \end{array} \qquad [3]$$

where k is the gray level value contains in the set L of gray level of the image and $P(k)$ the histogram probability of the observed gray value k; $\sigma_1(t)$ and $\sigma_2(t)$ are the variance of each group and are defined as followed:

$$\begin{array}{l} \sigma_1^2(t) = \sum_{k=1}^{t} (k - \mu_1(t))^2 P(k)/q_1(t) \\ \sigma_2^2(t) = \sum_{k=t+1}^{L} (k - \mu_2(t))^2 P(k)/q_2(t) \end{array} \qquad [4]$$

$\mu_1(t)$ and $\mu_2(t)$ are mean of each group defined by:

$$\begin{array}{l} \mu_1(t) = \sum_{k=1}^{t} kP(k)/q_1(t) \\ \mu_2(t) = \sum_{k=t+1}^{L} kP(k)/q_2(t) \end{array} \quad [5]$$

The best threshold t, is found when $\sigma_w^2(t)$ is minimal. The result of this global segmentation method is presented in figure 1. After segmentation, noisy points along the boundary are removed by performing morphological operations which are dilation followed by the erosion to avoid any displacement of the boundary with a 5x5 squared structuring element. At this level the contour is crisp or jagged due to the digital nature of the image and sometimes may be extended due to noise along the boundary. Extracting the real contour of the object generally requires some additional operations. This can be done by smoothing the extracted contour. We have used Fourier descriptors to carry out this task.

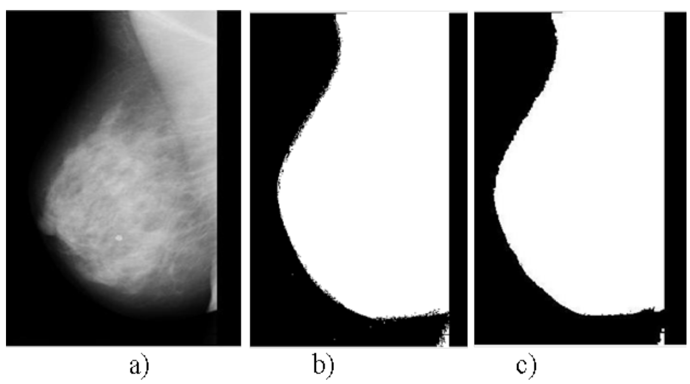

Figure 1. *Mammogram segmentation. a) Initial image, b) Segmented image, c) Spurious pixels removed along the edge of the segmented image*

2.2. Fourier descriptors

The Fourier descriptor is used to describe the boundary of a shape using the Discrete Fourier Transform. In this work N pixels of the contour ordered anti-clockwise are considered. If $Z(x_k, y_k)$ is a set of pixels coordinates describing the contour of an object lying on an xy-plane, the boundary coordinate can be represented as complex vectors of length N representing the number of pixels in the boundary. In this case, each contour element is a complex number represented by Z_k where $Z_k = x_k + jy_k$ for $k = 0, \cdots, N-1$. The advantage of this representation is that it reduces the problem from 2D to 1D. The discrete Fourier transform of Z_k is of the form:

$$C_l = \sum_{k=0}^{N-1} Z_k e^{-j2\pi kl/N} \quad [6]$$

The complex coefficients C_l are called Fourier descriptors of the contour.

The key idea is that applying the Fourier transform to a closed contour is like computing the Fourier transform of a signal. Therefore, Fourier descriptors represent the frequencies in the signal. Figure 2 presents a contour of a mammogram extracted after segmentation process and its corresponding Fourier descriptors. The low frequency components of

Fourier descriptors contain information about the general shape of the contour while the finer details are contained in the higher frequency components.

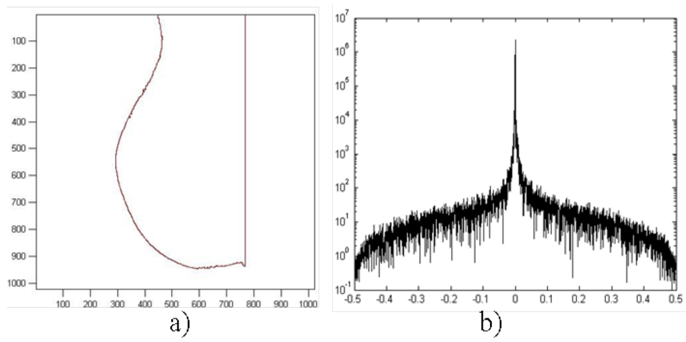

Figure 2. *Extraction of contour, a) Extracted contour, b) Fourier descriptor coefficients*

2.3. Contour smoothing

Shape is the most important low level image feature to human perception. Fourier descriptors have been applied recently in many applications in image processing particularly in matching, retrieving or recognizing objects according to their shape [7, 4]. However, they can also be used for an accurate smoothing of contour.

Segmented contour can be considered as a signal drowned in noise and Fourier descriptors can efficiently remove noise along its curve. Noise appears on the contour as zizag or abnormal shifts on the curvature. Smoothing tends to give to the contour curve a continuous aspect instead of scraps. As the contour information consists of low frequency components of Fourier descriptors, smoothing the contour can be sum-up to the selection of these low frequency components of Fourier descriptors.

A contour containing N pixels on its curve will yield N Fourier descriptors. The difficulty in this approach of smoothing the contour is determining the necessary number of Fourier descriptors that would be needed for an ideal smoothing of the contour curve. For this task when more coefficients are taken; the smoothed curve gets closer to the segmented edge whereas when few coefficients are selected the smoothed curve becomes distant to the segmented edge meanwhile computational time and memory are significantly reduced. On another hand the problem of smoothing can be addressed as determining the amount pixels points in the contour to be taken for computing the Fourier descriptors.

The appropriate number of Fourier descriptors and contours points for smoothing was drawn from a study on a sample mammogram. Errors were estimated for different number of Fourier descriptors and contour points. Optimal values of Fourier descriptors and contours points were considered as point were each error curve tends to become constant as represented in figure 4.

3. Experiments and results

This study was carried out on the mammograms database of MIAS (Mammographic Image Analysis Society) [9]. The database consists of 1024x1024 pixels size images scanned with 256 grayscale levels.

3.1. Results of the proposed method

After binary segmentation and extraction of contour, smoothing was performed in order to redress the contour curve. Figure 3 shows the result of smoothing on a mammogram contour.

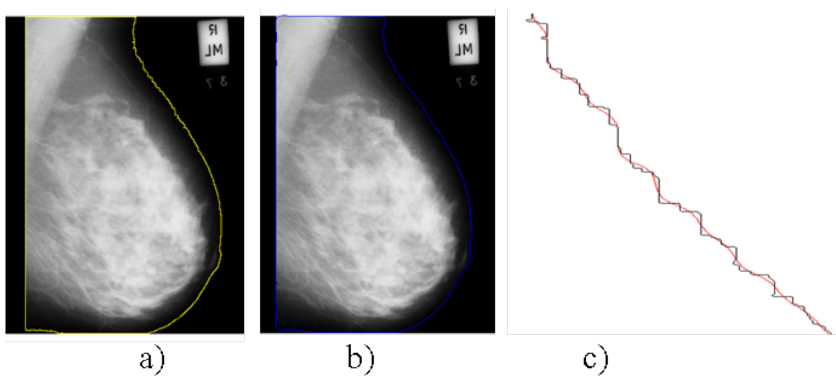

Figure 3. *Smoothing of a breast contour: a) Contour obtained after segmentation, b) Contour smoothed, c) Comparison of smoothed (red) and segmented (black) contour at a noisy area of breast edge.*

The contour appears with some slight displacement after segmentation and smoothing tends to reduce the displacement produced by noise by keeping curvature constant. It is noted that the smoothed curve remains close to the segmented contour. As a result, the contour of the object in the image appears more regular as in its real state.

3.2. Numerical estimates and error rates

The displacements of pixels were evaluated in order to compute the smoothing accuracy. The smoothing will be more accurate if the smoothed contour is closed to the segmented contour. The accuracy can therefore be studied by computing error rates. The evaluation of the error for different numbers of Fourier descriptors and contour points resulted in the error plots shown in Figure 4.

It is noted that the error is high for low number of Fourier descriptors (nFD) and of contour points (nCP). The error falls very fast and becomes almost constant at 75 FDs and 1200 CPs which represents a ratio of 6% of FD used. These previous values were taken in term of percentage as reference values for smoothing contour of selected mammograms in the MIAS database. The numerical estimates obtained are presented in the Table 1. It gives the number of contour points and Fourier descriptors used, the mean and the standard deviation of average errors in each case of mammogram.

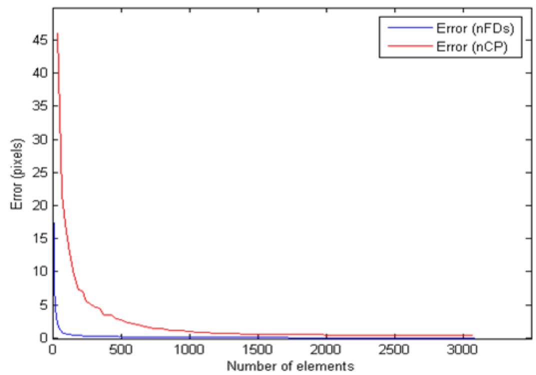

Figure 4. *Contour smoothing error estimates: in terms of selected nFD and nCP*

Case	nCP	nCP used	nFDs	Mean (pix.)	STD (pix.)
mdb013	3188	1275	78	0,712	0,840
mdb015	2430	972	60	0.538	0.537
mdb039	2368	947	58	0.710	0.645
mdb055	2570	1112	68	0.738	0.733
mdb057	3322	132	80	0.739	0.804
mdb071	2954	1181	72	1.052	0.905
mdb072	2782	1112	68	0.825	0.905
mdb073	3654	1061	64	0.655	0.617
mdb074	2730	1092	66	0.621	0.631
mdb088	2996	1198	72	0.880	0.810

Table 1. *Numerical estimates of mammogram contour smoothing*

4. Conclusion

This article has presented an accurate edge detection method using global threshold segmentation and a non iterative smoothing of contour in mammograms using Fourier transform. The method was tested on a set of mammograms and yields a good compromise between noise rejection in boundary detection and fast accurate smoothing. This algorithm is easy to implement as it is based on Fourier descriptors. In applications dealing with contour, smoothing is performed to improve estimation of length or area, and classification or registration task. Minimizing error rates could be done by selecting non-uniformly the points of contour with priority to areas where deformations due to noise are important. On other hand, the selection of nCPs and nFDs for optimal smoothing can be chosen automatically so that they dependent on mammogram considered instead of being fixed as reference from a mammogram. These aspects will be investigated in further work as well as a comparison of the results of this method with a gold standard.

Acknowledgements

This work was partially supported by the Romanian "Eugen Ionescu" research funding, grant established by AUF (Agence Universitaire de la Francophonie).

5. References

[1] BERSEN J., "Dynamic Thresholding of Gray Level Images", *in Proc. Intl. conf. on Pattern Recognition*, 1986.

[2] BOZEK J., MUSTRA M., GRGIC M., "A Survey of Mammographic Image Processing Algorithms for Bilateral Asymmetry Detection", *51st International Symposium ELMAR-2009*, 28-30 September 2009, Zadar, Croatia.

[3] DONGGANG Y., HONG Y., "An Efficient Algorithm for Smoothing Binary Image Contours", *Proceedings of IEEE International Conference on Pattern Recognition*, 1996.

[4] EKOMBO P. L. E., ENNAHNAHI N., OUMSIS M., MEKNASSI M., "Application of affine invariant Fourier descriptors to shape-based image retrieval", *Intl. Journal of Computer science and network Security*, vol. 9, num. 7, 2009.

[5] FERRARI R. J., FRÈRE A. F., RANGAYYAN R. M., DESAUTELS J. E. L., BORGES R. A., "Identification of the Breast Boundary in Mammograms Using Active Contour Models", *Medical and Biological Engineering*, vol. 42, 2004.

[6] HIGHNAM R., BRADY M., SHEPSTONE B., "Estimation of Compressed Breast Thickness During Mammography", *The British Journal of Radiology*, vol. 71, 1998.

[7] LARSSON F., FELSBERG M., FORSSÉN P.-E. "Patch Contour Matching by Correlating Fourier Descriptors", *Digital Image Computing: Techniques and Applications*, vol. 17, 2009.

[8] LEGAULT R., SUEN C. Y., "Optimal Local Weighted Averaging Methods in Contours Smoothing", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, num. 8, 1997.

[9] MINI-MIAS DATABASE OF MAMMOGRAMS " http://piepa.essex.ac.uk/ipa/pix/mias"

[10] OLSEN C., GEORGSSON F., " Problem Related to Automatic Nipple Extraction", *Lecture Notes in Computer Science*, 2005.

[11] OTSU N., "A Threshold Selection Method from Gray Level Histograms", *IEEE Trans. Systems, Man and Cybernetics*, vol. 9, num. 1, 1979.

[12] RICHARD F. J. P., COHEN L. D., "A New Image Registration Technique with Free Boundary Constraints: Application to Mammography", *Computer Vision and Image Understanding*, vol. 89, 2003.

[13] RUITER N. V., STOTZKA R., MULLER T. O., GEMMEKE H., REICHENBACH J. R., KAISER W. A., "Model-based Registration of X-ray mammograms and MR Images of the Female Breast", *IEEE Transactions on Nuclear Science*, vol. 53, num. 1, 2006.

[14] SAUVOLA J., PIETIKAINEN M., "Adaptative Document Image Binarization", *Pattern Recognition*, vol. 33, num. 2, 2000.

[15] WEI L., SHENG L. Y., YI R. X., PENG D., "A New Contour Detection in Mammogram Using Sequential Edge Linking", *IEEE second Intl. Symposium on Intelligent Information Technology Application*, 2008.

[16] YU L., MA F., JAYASURIYA A., SIGELLE M., PERREAU S., "A New Contour Detection Approach in Mammogram Using Rational Wavelet Filtering and MRF Smoothing", *Digital Image Computing Techniques and Applications*, vol. 67, 2007.

[17] ZHANG O., QIU A., GOLDGOF B., SARKAR S., LIHUA L., "3D Finite Element Modeling of Non-rigid Breast deformation for Feature Registration in X-ray and MR Images", *IEEE Workshop on Applications of Computer Vision*, 2007.