Proceedings of the 9th WSEAS International Conference on Applied Mathematics, Istanbul, Turkey, May 27-29, 2006 (pp301-308)

Skin-Cancer Region of Interest Detection using Jensen-Shannon Divergence

C. Atae-Allah, A. Atae-Allah, I.Rojas, F.Rojas,O.Valenzuela, J.F.Gómez Lopera
Dpto. Arquitectura y Tecnología de Computadores. Universidad de Granada
Periodista Daniel Saucedo Aranda, s/n, E-18071. Granada. Spain

Abstract.- .- Frequent screening of suspicious skin pigmentations is of paramount importance since, at an early stage; skin cancer has a high cure rate and, in most cases, requires a simple treatment. In this paper, we present a newly methodology for early detection of skin cancer based on the used of image segmentation. This method is based on the use of the Jensen-Shannon divergence of the histograms taken from two sub-window in a sliding window over the image. The advantages of using this entropic edge-detection method are shown; it is especially suitable when the source image is affected by any kind of noise, blur or low contrast effect. Results reveal a better performance than other methods used in this fields.

Key Words: Skin lesions**;** Pigmentation; Image Segmentation; Jensen-Shannon Divergence; Edge detection; Edge linking, Edge thinning.

1. INTRODUCTION

Frequent screening of suspicious skin pigmentations is a very effective approach for detecting skin cancers before they become lethal, since early changes in a malignant nevus typically consist in the development of an irregular pigmentation pattern. However, visual detection of melanomas by a dermatologist has an average diagnostic accuracy of only 58% , but it can be improved using imaging techniques. In general, these imaging techniques rely on delineating the boundary of a lesion, which is then analyzed for certain skin pigmentation characteristics, such as shape, size, symmetry, color, and texture. To this effect, a number of image segmentation techniques have been proposed in the past several years with varying degrees of success.

Skin cancer is the most common of all cancers and represents about one half of all new cancers detected [11]-[13]. Approximately 1.2 million new cases of skin cancer are being detected in the United States each year. About 80% of all skin cancers are basal cell carcinomas, 16% are squamous cell carcinomas, and 4% are melanomas. Melanoma is the deadliest of the skin cancers, accounting for over 7300 cancer deaths per year in the United States. The incidence of melanoma in the United States increased by 3% between 1997 and 1998 and approximately 44,200 new cases of malignant melanoma were detected in 1999. Early detection of skin cancer is of paramount importance [12]. If detected at an early stage, skin cancer has one of the highest cure rates, and in most cases, the treatment is quite simple and involves excision of the lesion. Moreover, at an early stage, skin cancers are very economical to treat, while at a late stage, cancerous lesions usually result in near fatal consequences and have an extremely high costs associated with the necessary treatment. For example, melanomas, the deadliest form of all skin cancers, account for 75% of all cancer deaths; however, when detected at an early stage, the cure rate is higher than 95%.

This paper is structured as follow. Section 1 is the introduction. Section 2 contains the technique of detection, by using the novel edge-detection method proposed. In Section 3 results are presented. Finally, the conclusions are in Section 4.

2. EDGE DETECTION

The taken images of pigmented areas of the skin have, besides the digital noise.
1. Colour variations for the skin type, even among different parts of the body.
2. Body hair, which hinders the separation of the affected zone of the rest of the image.
3. Pores which can be produce a mistake (can be considered as the affected area).
4. Tumours are not always obvious, especially where they are subtle or extremely subtle under the glandular tissue, which makes the task of interpretation difficult even for the radiologists themselves.

Image analysis is a challenging task due to poor illumination and high noise levels in the image that can vary up to 10-15% of the maximum pixel intensity. This is a problem because the image enhancement process may undesirably enhance noise component in the process. The image always seems cluttered, and the background varies greatly between different breasts. Even the worst abnormalities appear quite subtle and irregular.

2.1. THE JENSEN SHANNON DIVERGENCE

Jensen-Shannon divergence (hereinafter JS), proposed by Lin [10], has proved to be a powerful tool in the segmentation of digital images [1]-[5]. It is a measure of the inverse cohesion of a set of probability distributions having the same number of possible realisations.

$$JS_{\pi}\left(P_1, P_2, ..., P_r\right) \equiv H\left(\sum_{i=1}^{r} \pi_i P_i\right) - \sum_{i=1}^{r} \pi_i H\left(P_i\right) \tag{1}$$

where:

$P_1, P_2, .., P_r$ are discrete probabilities distributions $P_i = \left\{P_{i,j} \,/\, j = 1, ..., n\right\},\ i = 1, ..., r$.

$\pi_1, \pi_2, ..., \pi_r$ are the distribution weights for P_i $\pi \equiv \left\{\pi_1, \pi_2, ..., \pi_r \,/\, \pi_i > 0, \sum_{i=1}^{r} \pi_i = 1\right\}$.

$H\left(P_i\right) = -\sum_{j=1}^{n} P_{i,j} \log P_{i,j}$ is the Shannon entropy.

Divergence grows as differences between its arguments (the probability distributions involved) increase. And vanishes when all probability distributions are identical. In this paper only two probability distributions are used. So, the final expression of Jensen-Shannon divergence is

$$JS\left(P_1, P_2\right) \equiv H\left(\frac{P_1 + P_2}{2}\right) - \frac{1}{2}\left[H\left(P_1\right) + H\left(P_2\right)\right] \tag{2}$$

The application of divergence to detection of ROI in the digital mammography is based on a three-step structured procedure, as follows:

a) Calculation of divergence and directions matrices.

In this step the divergence and direction matrices associated to the image are calculated. The divergence matrix is composed by real numbers, and is similar to that obtained with the gradient operator for edge detection. The direction matrix contains the estimated edge direction for all image pixels.

b) Edge pixels selection.

Edge pixels are selected by means of a local maxima selection criterion from the divergence matrix, resulting in a binary image with the image edges.

c) Edge linking.

The final stage is an edge prolongation procedure that hopes to connect sets of unconnected edge pixels in the binary image previously obtained.

2.2. CALCULATION OF DIVERGENCE AN DIRECTIONS MATRICES

Let us consider a window made up of two identical subwindows and sliding down over a straight edge between two different textures. It has been shown [9] that in such conditions, the JS between the normalised histograms of the subwindows reaches maximum value when each subwindow lies completely within one texture.

According to the previous procedure, it is possible to assign a JS value to each pixel in the image. Hence, pixels with high JS have high probability to be edge pixels, and vice-versa. If, unlike the example shown in [9], the window-to-edge angle is not right (90º), the divergence maximum will be low and even undetectable, while JS inside a given texture it will be close to zero or to the base value. This then means trying several window orientations for each pixel. Only four orientations are, however, technically possible: vertical, horizontal, and two diagonals. Thus, the values JS_1, JS_2, JS_3, JS_4 are calculated for the fixed window orientation $0, \pi/4, \pi/2, 3\pi/4$. In this work we have used a square-shape sliding window, with a user defined size.

Now the question is how to obtain from these four values an estimation of the direction that maximises the divergence, and the value of this maximum, JS_{max}. For a given pixel, divergence value is a π-periodic function of window orientation over the image. It reaches its maximum value for a given orientation, b, and a minimum in b + p. Then, a theoretical model that describes this periodic function can be the following:

$$JS(x) = a + b\cos(\beta + 2\pi x), \quad x \in [0,1] \tag{3}$$

where a and b are constants determining amplitude of divergence oscillation, and $\beta \in [0, \pi)$ is the edge direction in this pixel. Divergence direction, x, is normalised in interval [0,1] to simplify calculations. According with the trigonometric relation, a theoretical model equivalent to (3) is

$$JS(x) = c + m\,\mathrm{sen}(2\pi x) + n\cos(2\pi x), \quad x \in [0,1] \tag{4}$$

where c, m and n are constants. Nevertheless, due to the computational effort needed calculating trigonometric functions, they can be replaced by other functions with similar properties, such as quadratic splines:

$$\mathrm{sen}(2\pi x) \approx f(x) \equiv \begin{cases} -16x^2 + 8x & x \in [0,1/2] \\ 16x^2 - 24x + 8 & x \in [1/2,1] \end{cases}$$

$$\cos(2\pi x) \approx g(x) \equiv \begin{cases} -16x^2 + 1 & x \in [0,1/4] \\ 16x^2 - 16x + 3 & x \in [1/4,3/4] \\ -16x^2 + 32x - 15 & x \in [3/4,1] \end{cases} \tag{5}$$

Then $f(x)$ is obtained as a quadratic spline of class 1, with nodes in points $\left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right\}$, interpolating to $\mathrm{sen}(2\pi x)$ in points $\left\{f(0), f\left(\frac{1}{4}\right), f\left(\frac{1}{2}\right), f\left(\frac{3}{4}\right), f(1), f'\left(\frac{1}{4}\right), f'\left(\frac{3}{4}\right)\right\}$. In the same way, it is obtained $g(x)$ as a quadratic spline of class 1 with the same nodes that $f(x)$, interpolating to $\cos(2\pi x)$ in the points $\left\{g(0), g\left(\frac{1}{4}\right), g\left(\frac{1}{2}\right), g\left(\frac{3}{4}\right), g(1), g'(0), g'\left(\frac{1}{2}\right), g'(1)\right\}$. With a least-squares fit of the divergence model (4) and the modification (5) for points JS_1, JS_2, JS_3, JS_4, the solution is:

$$JS(x)=\frac{JS_1+JS_2+JS_3+JS_4}{4}+\frac{JS_2-JS_4}{2}f(x)+\frac{JS_1-JS_3}{2}g(x) \tag{6}$$

and the direction, *x*, having the maximum divergence values can be obtained from equations:

$$\begin{aligned}
&\text{if } JS_1-JS_3\geq 0, JS_2-JS_4\geq 0 \quad \Rightarrow x=\frac{JS_2-JS_4}{4\left[(JS_1-JS_3)-(JS_2-JS_4)\right]}\in[0,1/4]\\
&\text{if } JS_1-JS_3\geq 0, JS_2-JS_4\leq 0 \quad \Rightarrow x=\frac{4(JS_1-JS_3)-3(JS_2-JS_4)}{4\left[(JS_1-JS_3)-(JS_2-JS_4)\right]}\in[3/4,1]\\
&\text{if } JS_1-JS_3\leq 0, JS_2-JS_4\geq 0 \quad \Rightarrow x=\frac{2(JS_1-JS_3)-(JS_2-JS_4)}{4\left[(JS_1-JS_3)-(JS_2-JS_4)\right]}\in[1/4,1/2]\\
&\text{if } JS_1-JS_3\leq 0, JS_2-JS_4\leq 0 \quad \Rightarrow x=\frac{2(JS_1-JS_3)+3(JS_2-JS_4)}{4\left[(JS_1-JS_3)-(JS_2-JS_4)\right]}\in[1/2,3/4]
\end{aligned} \tag{7}$$

Finally, defining $\delta=\pi x\in[0,\pi)$ as the estimated edge direction, the method described above calculates *x* from (7) (*i.e.*, the estimated direction that maximises the divergence), and the estimated JS_{max}, from (6). In this way, each image pixel is labelled with a pair of values: the estimated edge direction and the estimated JS_{max} placing the sliding window according with the estimated edge direction. Thus, two matrices are built: the divergence matrix (which indicates the probability for a pixel to belong to the image edge), and the directions matrix (which estimates the edge direction for the edge-pixels) .

However, direct application of the above method in some digital mammography does not give rise to good results, because the divergence may be too sensible to any changes in grey levels between regions. Thus, it is better to construct the divergence matrix including not only the histogram information but also by using the following expression:

$$JS^{*}_{i,j}=JS_{i,j}\left(1-\alpha+\alpha W_{i,j}\right) \tag{8}$$

where $W_{i,j}=\frac{\left|N_{w1}-N_{w2}\right|}{N_w}$, being N_{w1} and N_{w2} the average grey levels of subwindows W_1 and W_2 , and N_w the maximum grey level inside the window (normalisation factor). The factor $\alpha\in[0,1]$ is the attenuation factor, which determines the weights of divergence and the grey levels inside the window. This modification lets Jensen-Shannon divergence to be suitable to different kinds of digital mammography, thus transforming our algorithm into a hybrid among texture-based algorithms and grey-level based algorithms (gradient, laplacian, laplacian and gradient of the Gaussian).

2.3. EDGE PIXEL SELECTION

In this step the procedure selects which pixels from the divergence matrix are edge pixels. Thresholding the divergence matrix is not always useful, since maximum divergence values depend on the composition of adjacent textures, and will thus vary according to texture. Consequently, it would seem more appropriate to use a local criterion. A monodimensional window with an (adjustable) odd size *n* is centred on each pixel *i* and placed perpendicularly to the estimated edge direction. The pixel *i* is declared to be an edge if:

$$JS_i^{A}\cdot\left(JS_i-\frac{1}{n-1}\sum_{\substack{j=1\\ j\neq i}}^{n}JS_j\right)\geq 1 \tag{9}$$

For any other pixel *j* in that particular monodimensional window , where *A*, *B* and *n* are parameters to be adjusted. This local edge point detection method requires a simple divergence matrix pre-processing. Small fluctuations, often due to noise in the original

image or to texture regularity, may introduce a great number of false maximums, although these are usually fairly low. The divergence matrix is thus smoothed out by repeatedly applying a 3μ3 mean filter.

2.4. EDGE LINKING

The two steps described above make extracting the image edge pixels possible. However, it is not always feasible to establish a good compromise between the quality of the obtained binary image and the desired connectivity of the edge pixels. This may be due to the presence of noise in the original image, and the texture composition of the image regions. In order to deal with these problems, a third step may be added: edge pixel linking. It attempts to join the various sets of edge pixels, by using information from the divergence matrix associated with the image, together with knowledge of the direction in which maximum divergence is produced. In broad terms, linking procedure consists on extracting edge-pixels which have not been marked since they did not satisfy the condition, but were close to. It not all the pixels in the image are candidates for filling the gaps, only those classified as neighbour candidates of *end points*.
The definition of end point: includes several variants that may influence the result of the linking process. The present paper uses the definition of end point as a point having one or two marked pixels joined together. Thus, only certain neighbour end points are candidates to prolong image edges.

3. EXPERIMENTAL RESULTS

The database, used in work is compound of 267 fluorescence images (the data-base has been obtain from the Institute of Biophysics, University of Regensburg, Germany), 100 images of actinic keratosis, 100 images of basal carcinoma and 67 psoriasis images. Original images have a very high resolution of 1368x1712 pixels and an approximate size of 1.5 MB. To reduce the processing time, images are resized to 256x320 pixels using bicubic interpolation. These values maintain the original aspect ratio of the image. The hair can be optionally removed, or minimized, by median filtering the image using two structuring elements of size [1x5] and [5x1]. Finally, to suppress large variations within the background and the lesion, and to reduce the effect of different skin color variations, the original color RGB images are transformed into intensity (grayscale) ones. The separate values of the three color channels (R, G, B) are combined to produce an intensity image (Y) using a commonly accepted transformation, namely Y = 0.3*R + 0.59*G + 0.11*B. Figure 1 shows and example of three different skin lession.

In this section, results for noisy images are presented. These experiments are important due to two reasons: they can simulate experimental conditions that commonly appear in practice, and they allow us to evaluate the robustness of edge-detection methods. The original image is contaminated with three synthetic distortions (Blur, Gaussian and impulsive) to simulate experimental problems that may appear in practice.
Different features of the images are measured to determine the behaviour of the algorithm, taking into account that the ROI is a circle (as determined by a human expert). These parameters are as follows:

Centre (1-by-ndims(L) vector): the centre of mass of the region. Note that the first element of Centroid is the horizontal coordinate (or x-coordinate) of the centre of mass, and the second element is the vertical coordinate (or y-coordinate).

Diameter (Scalar): the diameter of a circle with the same area as the region. Computed as:

$$Diameter = \frac{Area}{2r} \tag{10}$$

where r is the ratio of the circle.

Compactness: This parameter is defined by the following expression:

$$Compactness = \frac{Perimeter^2}{4\pi Area} \tag{11}$$

In **¡Error! No se encuentra el origen de la referencia.**, numerical results of centre (vector; the x- and y-coordinates of the centre of mass of the region), Diameter and Compactness of ROI are presented. These parameters are extracted using 3 different algorithms (Sobel, Canny and proposed algorithm) and the result of and human expert. The image is contaminated with three synthetic distortions (blur, and gaussian and impulsive noise) to simulate experimental problems that may appear in practice, Blur, whose properties are similar out-of-focus in image acquisition systems, is achieved recursively applying a 3µ3 mask 5 or 8 times to the original image. Gaussian noise has zero mean, and typical deviation 5, 10 and 15. Impulsive noise is salt-and-pepper noise with amounts 5%, 10% and 15%. Gaussian and impulsive noise are very common in image acquisition and transmission.

In Figure 2, and example of the detection of the ROI using different approaches is presented. It can conclude that the proposed method is a much better ROI detector than methods based on Sobel and Canny.. Furthermore, the proposed proceeding is more robust against noise than other methods.

4. CONCLUSION

Skin cancer is now one of the most common forms of cancer. Digital image has been introduced in several countries over the last few years. The new technology requires new optimising methods considering for instance the increased possibility of changing the absorbed dose and the effect it has on the image processing and noise level. There is no agreement on the noise level that may be accepted and the absorbed doses needed. This paper proposes a newly algorithms based on an entropic measure in recognizing cancer regions on tissue, which is very robust against the presence of noise, low contrast or blurring in the image. This property makes it very suitable in digital image because previous filter is not needed.

Experiments show the performance of the proposed procedure using several images. It can be shown that the algorithm is robust, effective and presents better results than other method presented in the literature.

References

[1] Bovik,A.C., Huang,T.S., Munson, D.C., 1986. “Nonparametric tests for edge detection in noise” Pattern Recognition 19 (3), 209–219.

[2] Huang, J.S., Tseng, D.H., 1988. “Statistical theory of edge detection” Comput.Vision Graphics Image Process. 43, 337–346.

[3] Pratt, W.” Digital image processing” California: John Wiley and Sons, 1991.

[4] Frei, W.; Chen, C.”Fast boundary detection: A generalization and a new algorithm” IEEE Transactions on Computers 26 (1977), 988–998.

[5] Hou, Z., Koh, T.S., 2003. “Robust edge detection” Pattern Recognition 36, 2083–2091

[6] Lim, D.H., Jang, S.J., 2002. “Comparison of two-sample tests for edge detection in noisy images”. J. R. Stat. Soc. Ser. D—Statist. 51 (1), 21–30.
[7] Qu Ying-Donga, Cui Cheng-Songa, Chen San-Benb, Li Jin-Quan “A fast subpixel edge detection method using Sobel–Zernike moments operator” Image and Vision Computing 23 (2005) 11–17
[8] Fabio Boschetti “Improved edge detection and noise removal in gravity maps via the use of gravity gradients” Journal of Applied Geophysics, Volume 57, Issue 3 , April 2005, Pages 213-225.
[9] V.Barranco López, P.Luque Escamilla, J.Martinez Aroza, R.Román Roldán, “Entropic texture-edge detection for image segmentation” Electronics Letters, vol.31,n.11, pp.867-869, 1995.
[10] Lin, J “ Divergence measures based on the Shannon entropy” IEEE Transactions on Information Theory. Vol 37. No 1 pp 145-150. 1991.
[11] Joanne F. Aitken, Monika Janda, Mark Elwood, Philippa H. Youl, Ian T. Ring and John B. Lowe “Clinical outcomes from skin screening clinics within a community-based melanoma screening program” Journal of the American Academy of Dermatology, Volume 54, Issue 1, January 2006, Pages 105-114
[12] Matthias Lauth, Anne Birgitte Unden and Rune Toftgård “Non-melanoma skin cancer: pathogenesis and mechanisms” Drug Discovery Today: Disease Mechanisms, Volume 1, Issue 2, November 2004, Pages 267-272.
[13] Beth R Santmyire-Rosenberger, Cherie Young and Gary L. Peck “Importance of non-melanoma skin cancer as a risk factor for melanoma” Journal of the American Academy of Dermatology, Volume 50, Issue 3, Supplement 1, March 2004

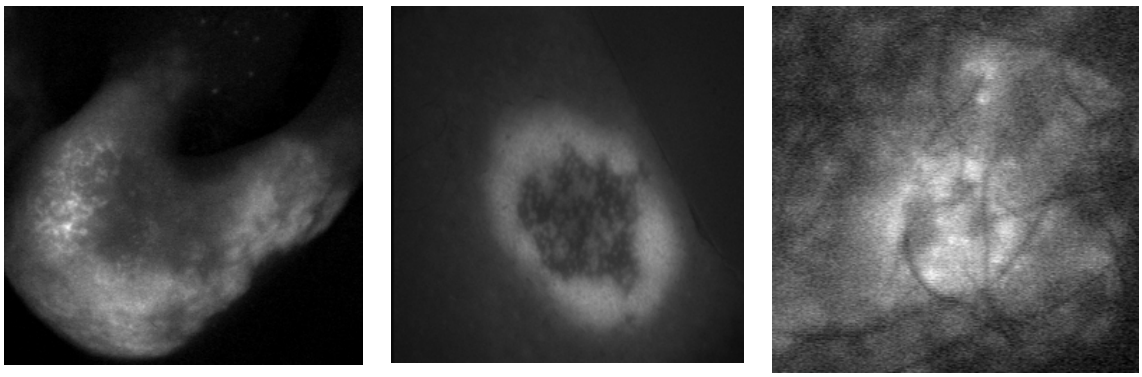
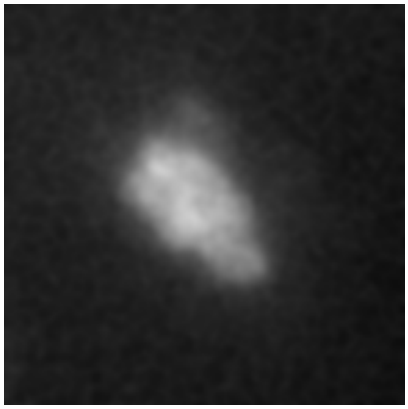


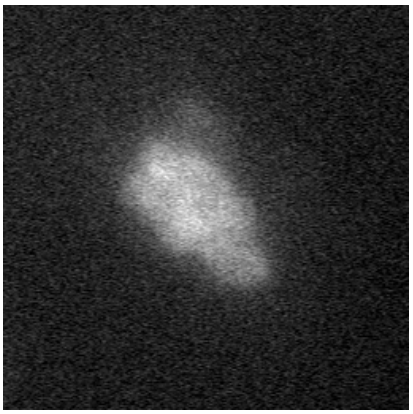
Fig 1. (a) Basal Cell Carcinoma. (b) Psoriasis (c) Actinitic Keratosis.

PARAMETER	PROPOSED METHOD	SOBEL	CANNY
Centroid	[125.1768 133.3813]	[73.2000 123]	[73.2500 116.8750]
BoundingBox:	[78.5000 87.5000 89 91]	[72.5000 120.5000 2 5]	[70.5000 111.5000 5 12]
MajorAxisLength:	138.9695	5.8877	14.1608
MinorAxisLength:	81.4446	1.6007	5.7348
Eccentricity:	0.8103	0.9623	0.9143
Orientation:	-47.9832	-78.2507	84.6718
ConvexArea:	5131	8	53
FilledArea:	4722	5	16
EulerNumber:	0	1	1
EquivDiameter:	22.4545	2.5231	4.5135
Solidity:	0.0772	0.6250	0.3019

Table 1: Comparison of ROI detection using different approaches



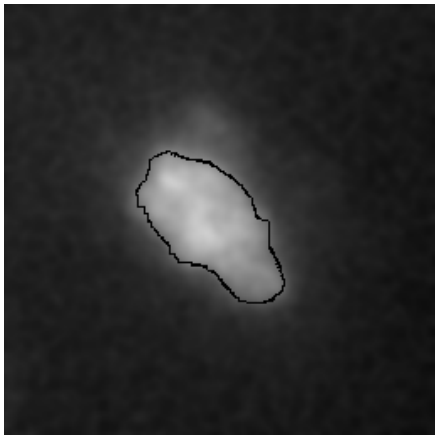
a)



b)

c)

d)



e)

Figure 2. Results of edge detection in blurred images; a) Original skin lesion; b) Image blurred 5 times; c) Results of segmenting images with Sobel edge detector; d) applying Canny edge detector; e) Proposed edge detector.