Diagnosis of Breast Cancer and Clustering Technique using Thermal Indicators Exposed by Infrared Images

Hossein Ghayoumi Zadeh [1], Iman Abaspur Kazerouni [1], Javad Haddadnia[1]

[1.] Sabzevar Tarbiat Moallem University/Department of Electrical Engineering, Sabzevar, Khorasan Razavi Iran
h.ghayoumizadeh@gmail.com

Abstract: In this paper we proceed breast cancer detection through thermal indicators at infrared images to be taken from the patient. The work is notably important according to non application of harmful radiations where are used to produce mammography images for instance. In this method it has been tried to provide people in general, easy detection of breast cancer by using image processing techniques along with computer artificial intelligence tools. In this paper we proceed Half technique to detach breast region out of thermal image and then we cluster detachments using Fuzzy K- Means method. The presented method is highly important in breast cancer detection through which, while applying the technique, there would be possibility to diagnose the cancerous region and cut it away within few seconds. In a better word, there have been used three assimilate procedures of asymmetry analysis, thermogrphy development and K- means clustering to minimize error occurrence. Since breasts with malignant tumors have higher temperature than healthy breasts and even breasts with benign tumors, in this study, we look for detecting the hottest regions of abnormal breasts which are the suspected regions.
[Hossein Ghayoumi Zadeh, Iman Abaspur Kazerouni, Javad Haddadnia. **Diagnosis of Breast Cancer and Clustering Technique using Thermal Indicators Exposed by Infrared Images**. Journal of American Science 2011;7(9):281-288]. (ISSN: 1545-1003). http://www.americanscience.org.

Keywords: Breast cancer; Half Circular; k-means; Thermograph

1. Introduction

Breast cancer is becoming malignant of the main tissue cells of the breast where they produce milk for infant nutrition. The following factors cause cancerous cells to generate heat: (1) higher metabolic activity of cancerous cells compared to normal cells (2) Angiogenesis; a cancer tumor starving for nutrients so produces a chemical that promotes the development of blood vessels that supply the tumor and also causes normal blood vessels to dilate to provide more blood in tumor growth.[1-2] Mamography is one of the most important and effective tools to detect breast cancer. But concerning the matter of ionization risk caused by X ray radiation and its danger to patients, it bears the most important risk while working with mammography. Diagnosis of cancerous tumors by the thermal signs they suggest is one of the important reasons that can be imply to cluster of thermal images taken from breast region. Since all driven characteristics are not useful enough to distinguish targeted samples from non- targeted ones and that sometimes they might happen to worsen the results, it is necessary then the optimum characteristics to be selected in other way.

In this research and in order to select primary composition of characteristics, a sample group of 150 people, consist of 3 images for each, has been chosen. To cluster cancerous areas best, it is initially necessary to extract the breast region out of the image and then image processing operation to be applied. All objects in the universe emit infrared radiation as a function of their temperatures [3]. The higher an object's temperature, the more intense infrared radiation it emits [4].The human body emits infrared (IR) radiation mainly in the range of 2–20 μm wavelength; the application of Plank's equation and Wien's law shows that approximately 90% of the emitted IR radiation in humans is in the longer wavelengths (6–14 μm) [5].In 1961 Williams and Handley [6] demonstrated the possibility of breast cancer detection from a temperature increase in the overall global image of the breast (at that time the thermal sensitivity was ±2°C and an image required an exposure time of many minutes). Later studies have shown how the thermal image might help to detect breast cancer: Pariski et al [7], after a 4-year clinical trial of 769 subjects, reported that infrared imaging offers a safe non-invasive procedure that would be valuable as an adjunct to mammography in determining whether a lesion is benign or malignant; Arora et al [8] conducted a prospective clinical trial of 92 patients and concluded that digital infrared thermal imaging is a valuable (97% sensitivity) adjunct to mammography and ultrasound, especially in women with dense breast parenchyma. Kennedy [9] shows that a combination of thermography and other diagnostic methods may boost both sensitivity and specificity.

The latest most controversial recommendations by the United States Preventive Services Task Force conflicting with recommendations from the American Cancer Society and other medical groups are that women in their 40s should not routinely have mammograms and that woman aged 50 to 74 should have mammograms every 2 years instead of annually. This recommendation is based on an analysis showing that every-other-year screenings could provide 80% of the benefits of annual screening while cutting the risks almost in half. The panel concluded that the benefits outweighed the risks among those over 50, but not in the younger group [10].The use of the thermography is based on the principle that metabolic activity and vascular circulation in both precancerous tissue and the area surrounding a developing breast cancer are almost always higher than in normal breast tissue [11, 12].

Various reasons can be discussed here for application of thermal infrared camera in breast cancer diagnosis; [6]

- To be non-invasive
- indirect contact
- absence of radiation
- lower costs of operation comparing to other methods
- possibility to specify tumor's characteristics

To discuss further the reasons for application of the thermal camera particularly at Image processing and development, there should be added the followings;[8,9]

- Producing newer generation of the cameras possessing greater facilities
- Our enhanced ability to better processing of images
- More profound perception to the thermo indications of body physiological activity

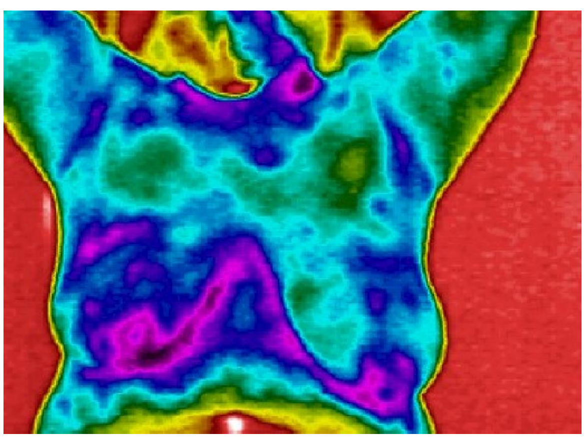


Figure 1. The early thermo image of the patient which shows any part of body by different color according to regional temperature each part produce.

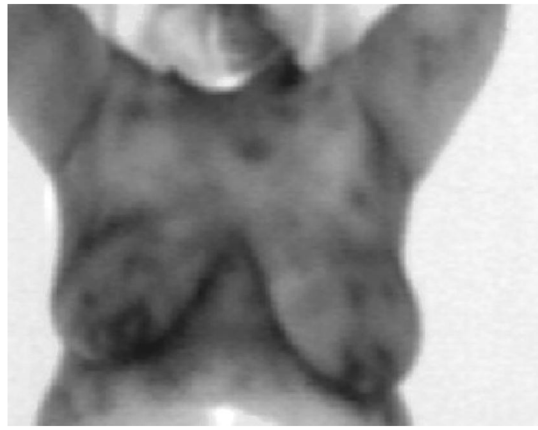


Figure 2. The gray level image, processed from early image

2. Detachment of Breast Region

T To detach breast region we use half circular and half parabola techniques respectively by sequence. Half circular algorithm has been used to diagnose breast area within the early images taken but soon after detection, we seek half parabola technique to completely extract breast region out of the image. The first step to do so is to change color thermo images (fig.1) taken from patient to rgb2gray format.(fig-.2)

3. Half Circular Alteration

The Hough transform can be described as a transformation of a point in the x,y-plane to the parameter space. The parameter space is defined according to the shape of the object of interest. A straight line passing through the points (x_1,y_1) and (x_2,y_2) can in the x , y-plan be described by:

$$y = ax+b$$

This is the equation for a straight line in the Cartesian coordinate system, where a and b represent the parameters of the line. The Hough transform for lines does not using this representation of lines, since lines perpendicular to the x-axis will have an a-value of infinity. This will force the parameter space a, b to have infinite size. Instead a line is represented by its normal which can be represented by an angel θ and a length p

$$p = x \cos(\theta)+y \sin(\theta)$$

The parameter space can now spanned by θ and p, where θ will have a finite size, depending on the resolution used for θ. The distance to the line r will have a maximum size of two times the diagonal length of the image. [13]

The circle is actually simpler to represent in parameter space, compared to the line, since the parameters of the circle can be directly transfer to the parameter space. The equation of a circle is

$$r^2 = (x-a)^2+(y-b)^2$$

As it can be seen the circle got three parameters, r, a and b. Where a and b are the center of the circle in the x and y direction respectively and where r is the radius. The parametric representation of the circle is

$$x = a+r\cos(\theta)$$
$$y = b+r\sin(\theta)$$

We can increase or decrease scales of circles to be appear on the image by changing min or max scale of their radius. As you see in figure 3 , to delete circles which are appearing out of the body , we can primarily erase the edges of the image by applying SOBEL algorithm [14](figure4)and then delete the circles which have been appeared out of our edge erasing format.

4. Half Parabola Technique

Right after extraction of breast region by Half Circular Technique we should next take the Half Parabola Technique. To do so we will primarily make edge erasing of the image. Edge erasing has been done through logarithmic procedure in this project. (Figure 5)

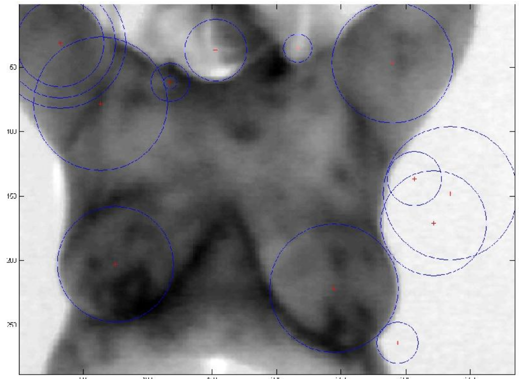


Figure 3. scaling circles inside the image by half circular technique

Speed of processing with half parabola technique is incredibly higher than half circular. We will have to put filter on the images with cut edge if we are about to speed up the operation of half parabola in order to decrease ineffective volume of data. The filter which has been used should be a mask at 6x6 sizes. After using the filter the result as in figure 6, could be easily observed and the fact that data quantities inside the image 5 have been decreased.

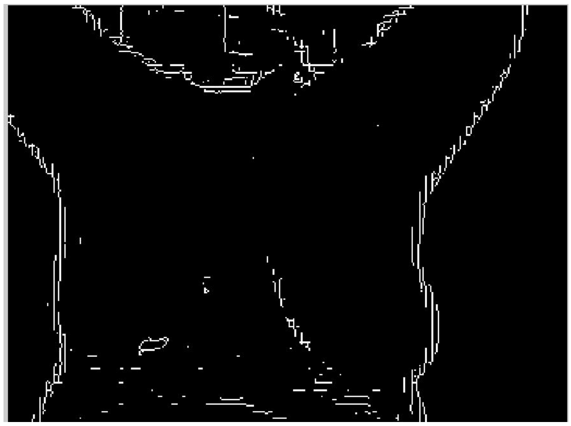


Figure 4 – edge erasing by SOBEL algorithm

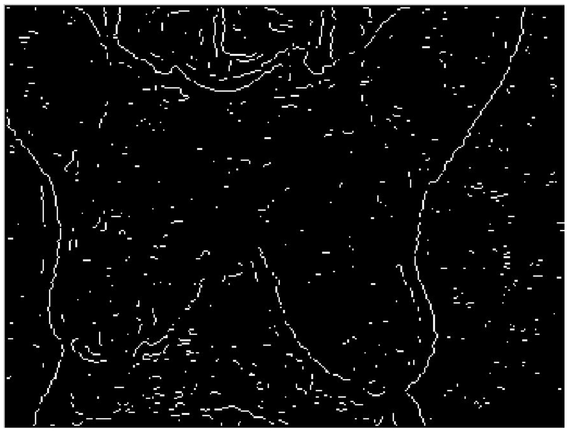


Figure 5. edge erasing of the image by logarithmic procedure

Figure 6. Ineffective data decreasing using filter

Upon functionality of the curves within coordinates of parabola, we will be able to recognize parabola curves inside the real images at whatever direction they might stand (equation 1). By this way, pixels at the edges will be transferred to new coordinates. This shortens computation time. Figure 7 shows the procedure for such a shift.

Coordinates of x-y within rotation angle of θ will take new coordinates of x'-y'. [15]

$$(y^{'} - y_0^{'}) = p(x^{'} - x_0^{'})^2 \quad (1)$$

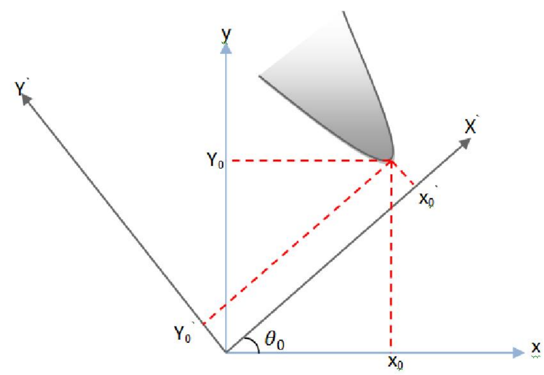


Figure.7.Tilted parabolic curve

To rotate parabola within the image, we can apply the equation 2.

$$\begin{bmatrix} x^{'} \\ y^{'} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2)$$

θ is taken as rotating angle here and considering formula 3, parabola equation will be equal to:

$$(-x\sin\theta + y\cos\theta) - (-x_0\sin\theta + y_0\cos\theta) = p[(x\cos\theta + y\sin\theta) - (x_0\cos\theta + y_0\sin\theta)]^2 \quad (3)$$

It will be deducted from derivative of the equation that:

$$-\sin\theta + \frac{dy}{dx}\cos\theta - 2p[(x\cos\theta + y\sin\theta) - (x_0\cos\theta + y_0\sin\theta)].[\cos\theta - \frac{dy}{dx}\sin\theta] \quad (4)$$

By substitution of equation 3 with 4, we come up with new formula presented in equations 5 and 6 which specifies the parabola's head and direction.

$$Y_0 = \left[\frac{k_1(x\cos\theta_0 + y\sin\theta_0) + (x\sin\theta_0 - y\cos\theta_0)}{(k_1\sin\theta_0 - \cos\theta_0)}\right] - \frac{(k_1\cos\theta_0 + \sin\theta_0)}{(k_1\sin\theta_0 - \cos\theta_0)}x_0 \quad (5)$$

Where:

$$k_1 = \frac{-\sin\theta_0 + \frac{dy}{dx}\cos\theta_0}{2(\cos\theta_0 + \frac{dy}{dx}\sin\theta_0)} \quad (6)$$

Then we shift coordinates of parabola's head into the half circular region. In other words we transfer the points with cut edges from half circular to half parabola coordinates and try to locate our ideal parabola. The ideal parabola could be displayed only when more points from cut edges region of circular intersect with coordinates of parabola. The result can be seen in figure 8.

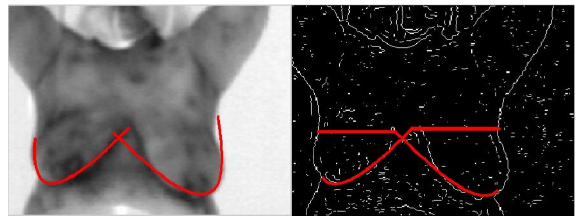
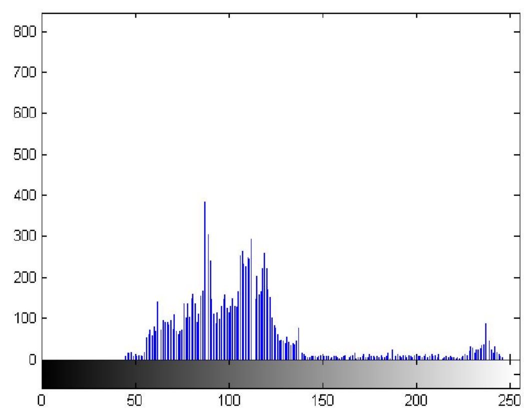
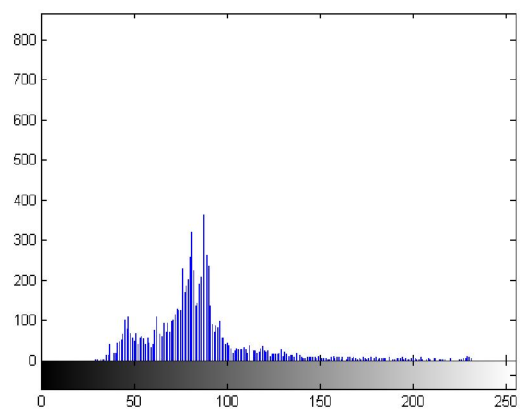


Figure 8.realizing breast region by changing parabola half

Although asymmetry is quite distinguishable apparently through images displayed in figure 8, but in order to do the realization intelligent, the matter can be proven through histogram graphs to be drawn by computer. (fig.9)



(a)



(b)

Figure 9.histogram of left(a) and right(b) breasts

Assuming that histogram might not be able to explain asymmetry tangible and precise well, we may then use other indicators to increase brightness of image. For example we can use "moment" to intensify the image through which valuable statistical information can be presented. This information may include mean, variance, skewness and kurtosis where they have been taken as follow:

(7)

$$\text{Mean } \mu = \frac{1}{N}\sum_{j=1}^{N} P_j$$

$$\text{Variance } \sigma^2 = \frac{1}{N-1}\sum_{j=1}^{N}(p_j - \mu)^2$$

$$\text{Kurtosis} = \frac{1}{N}\sum_{j=1}^{N}\left(\frac{p_j-\mu}{\sigma}\right)^4$$

$$\text{Skewness} = \frac{1}{N}\sum_{j=1}^{N}\left(\frac{p_j-\mu}{\sigma}\right)^3$$

Here pj is taken as the probability density. As an example, values attributed to cancerous and non cancerous images have been tabulated in table1.

Table1: histogram moments

Moments	Left breast cancerous	Right breast cancerous	Left breast non-cancerous	Right breast non-cancerous
Mean	.0011	.0009	.0011	.009
Variance (10-6)	3.06	2.2	2.4	1.6
Skewness (10-6)	3.7	2.2	3.7	3.4
Kurtosis (10-8)	2.1	.1	1.1	1.2

5. Thermal images development

All of substances have electromagnetic radiation in temperatures higher than absolute zero, to be realized as natural or thermal radiation. Amounts of the pixels in thermograms are representing thermal radiation of the temperature by human body. Various tissues and organs in body demonstrate different amounts of radiation. Therefore at a given temperature pattern or in other words within the pixels pattern, we will be able to detect any abnormality may exist. [16, 17]

At initial point we were able to detach breast region away of the image. Of course we can diagnose the case which follows through differentiation of thermo colors based on fuzzy procedures of K-means and C- means. But the aim is expedition in diagnosis procedure and the accuracy.

Based upon studies on various images of the patients (150 people), an experimental relationship has been got. Considering equation of number 8, we would be able to detach the concerning regions if we assume that the pick pixel intensity of thermal image of RGB is equal to T [18]:

$$H(x,y) = \begin{cases} 1 & \text{if} \quad R_{T(x,y)} > 100 \;\&\; G_{T(x,y)} < 20 \;\&\; B_{T(x,y)} < 20 \\ 0 & \text{else} \end{cases} \quad (8)$$

Now at the next stage if we adapt the results with our early images, we will be able to observe the classification well. Some examples of such images have been given in figure 10.

If we are now going to begin to diagnose upon extracted red color areas, we should first put together the segmentation we obtained with already detached area of patient breast image. This synchronization could be observed in Figure 11.

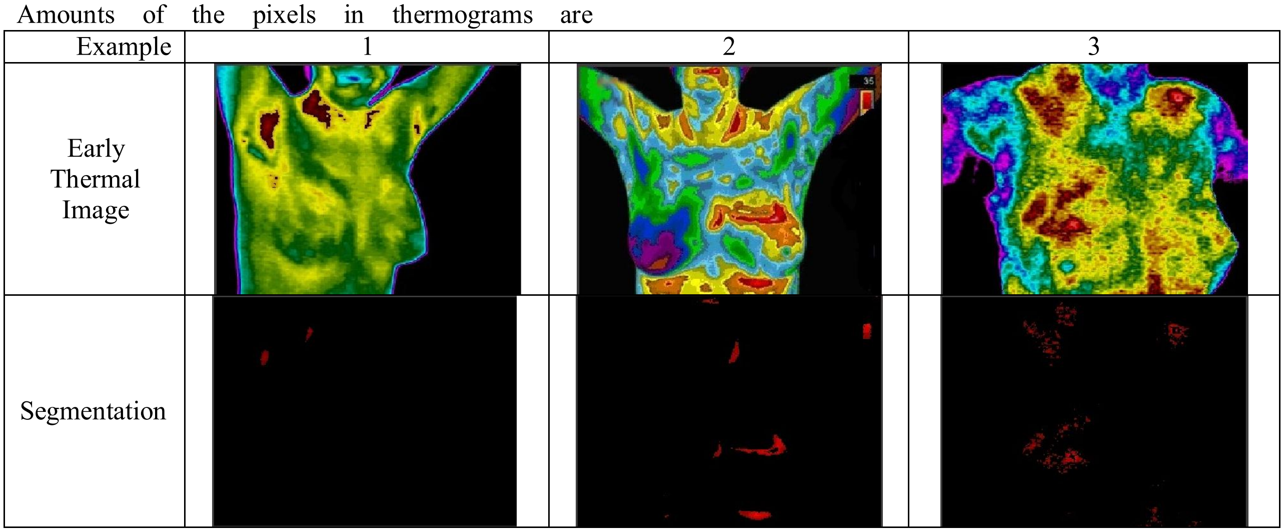
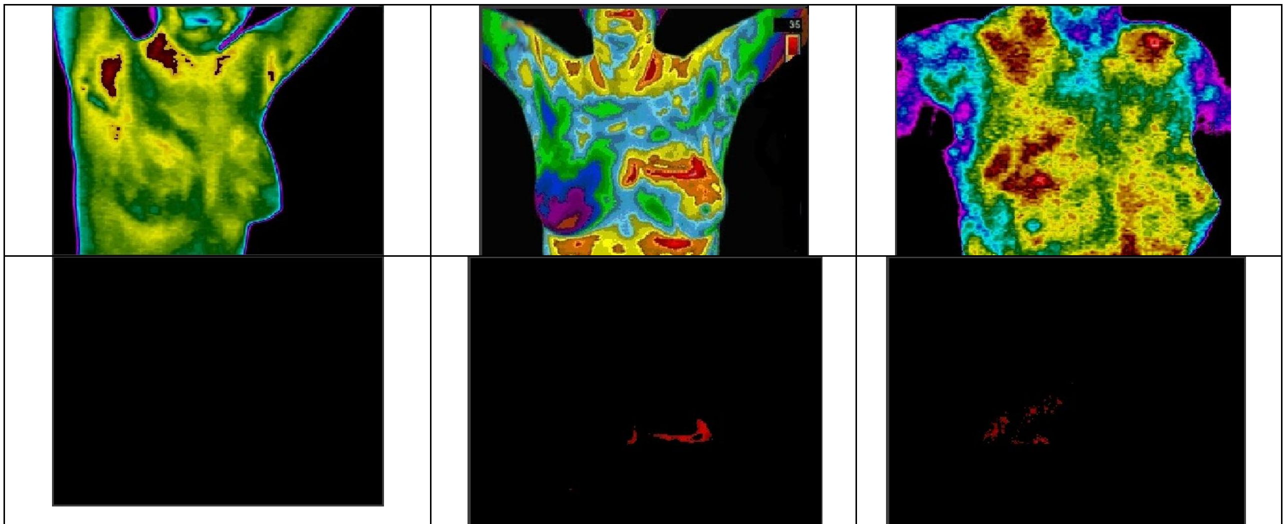


Figure 10. Segmentation of suspicious regions out of primary thermo images



System has not diagnosed any breast cancer	Patient has cancer in her left breast	Patient has cancer in her right breast

Figure 11. Segmentation of breast cancerous area from other regions

Normal uniform non-vascular	Normal uniform vascular	Equivocal (questionable)	Abnormal	Severely abnormal

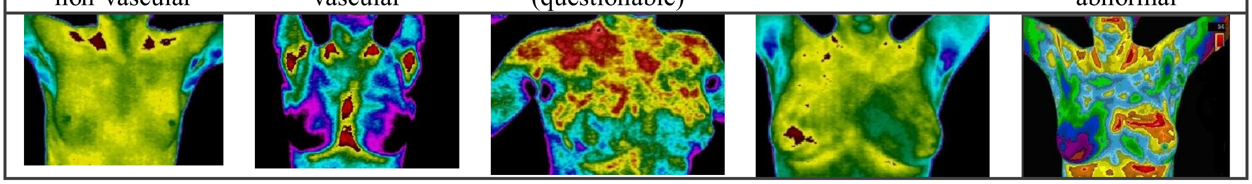


Figure 12. Classification of thermo images

Some points in diagnose of cancer from the images are worth mentioning here:

There exist organs and surface areas in body like under neck carotid where they basically demonstrate higher temperature, comparing to other organs as it can be observed quite clearly in figure 8. If we took K-Means procedure to make the segmentation, suspicious regions would have increased significantly to the level that diagnosis procedure might get into difficulty. We may impose changes to quotation 8 in case that we may need to increase or decrease areas.

In order to increase diagnosis accuracy and to reduce occurrence of any possible error, we have selected neck carotid region as an absolute base criteria region. This area can be used both to calibrate clear imaging and to adjust the diagnosis in a way that the red color segments in breast region, which are smaller than neck carotid region, can be easily filtered through computer intelligent system. According to analysis to be done already on such images, they can be classified to five groups. This classification is presented in figure 12. [19, 20]We use k-means method to make classification properly accurate.

6. Fuzzy k-means

The K-means algorithm was proposed in 1960's (Mac- Queen) and has been used for many pattern recognition Problems [21, 22]. It is one of the simplest and the fastest unsupervised learning techniques for clustering. Based on studies being done on the existing functions to segmentation, Fuzzy C-Mean is considered as an appropriate function to denote single registration of targeted and non targeted stimulations. As it is understood by the name of the function, fuzzy clustering is the base benchmark to the procedure and operation of the function does not necessarily require any instructor. Also fuzzy nature of the method will provide possibility to determine membership percentage attributed to each single registration.

It is one of the simplest and the fastest unsupervised learning techniques for clustering. It partitions N data points to K disjoint subsets Sj where j=1,2,...,K. It is based on minimization of the objective function

$$J=\sum_{j=1}^{k}\sum_{n\in s_j}\left\|x_n-\mu_j\right\|^2 \quad (9)$$

Where xn is a vector representing nth data points in Sj and μj is the geometric centeroid of data points. The algorithm Consists of 2 steps: choosing the number of cluster K, and then implying the K-means clustering to the image.The fuzzy clustering algorithm compares the colors in a relative sense and groups them in clusters which are not with crisp boundaries and data point can belong to more than one cluster while in hard clustering (K-means) data is grouped into crisp clusters where each data point belongs to exactly one cluster [23, 24].

In order to test the program, we use thermal feature of extracted region from previous stage (figure 13).

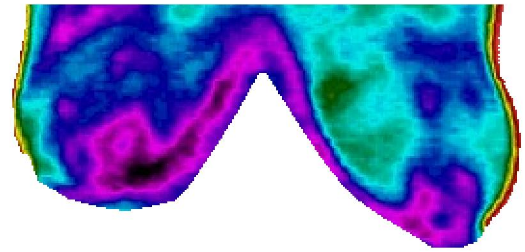


Figure 13. extracting breast region out of thermal image

We now initiate to cluster on the basis of K-Means algorithm. We can increase clusters by number or due to our own style to whatever we wish. If we use a calibrated standard IR camera, we can be sure that relevant imagery and accurate temperatures can be viewed since each color and brightness is associated with a specific temperature. It is important to have an accurate and stable calibration so that to satisfy the temperature sensitivity of the camera. [25]There are also other important factors in order to make thermal images in a standard form.The final results are displayed in figure 14.

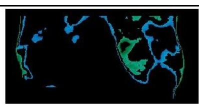
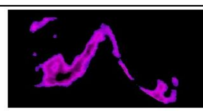
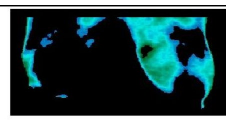
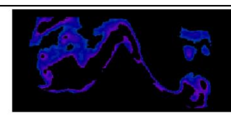
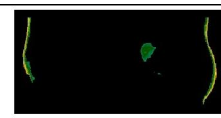
A	B	C	D	E
				

Figure 14. Image clustering to 5 pieces based on K-Means algorithm

Taking into account the clusters in color, we understand that they have been classified by different colors in a way that each cluster is representative of a particular temperature. This is clearly displayed in figure 8-b and demonstrates the suspicious region or in other words the cancerous area.

Conclusion

In order to make a diagnosis well accurate at a considerable prompt time through an intelligent communized system, there should be in place three methods which appear by sequence to complete the whole procedure together and they are:

Asymmetry

Thermal Image Development

Clustering by use of K-Means

In a better word, we might be able to adjust the errors already were exist in one method through putting together second method with previous one that is in line with first procedure.

References

1. Ng, E. Y. K., and Kee, E. C., Advanced integrated technique in breast cancer thermography. Journal of Medical Engineering & Technology. 32:2103–114, 2007 doi: 10.1080/0309190060 0562040.
2. Koay, J., Herry, C. H., & Frize, M. (2004). Analysis of Breast Thermography with Artificial Neural Network, In Proceedings 26th IEEE EMBS Conf., San Francisco, CA, USA, pp: 1159– 1162, Sep. 1–5.
3. Siegel, R., and Howell, J. R., Thermal radiation heat transfer. Hemisphere, Washington, DC, 1992.
4. Jones, B. F., A reappraisal of the use of infrared thermal image analysis in medicine .IEEE Transactions on Medical Imaging. 17:61019–1027, 1998 doi:10.1109/42.746635.
5. Ng, E. Y. K., A review of thermography as promising noninvasive detection modality for breast tumor. Int. J. Therm. Sci. 48:849–859, 2009. doi:10.1016/j.ijthermalsci.2008.06.015.

6. Lloyd-Williams, K., and Handley, R. S., Infrared thermometry in the diagnosis of breast disease. Lancet. 2:1378–1381, 1961.
7. Parisky, Y. R., Sardi, A., Hamm, R., Hughes, K., Esserman, L., Rust, S., and Callahan, K., Efficacy of computerized infrared imaging analysis to evaluate mammographically suspicious lesions. AJR. 180:263–269, 2003
8. Arora, N., Martins, D., Ruggerio, D., Tousimis, E., Swistel, A. J., Osborne, M. P., and Simmons, R. M., Effectiveness of a noninvasive digital infrared thermal imaging system in the detection of breast cancer. Am. J. Surg. 196:523–526, 2008
9. Kennedy, D. A., A comparative review of thermography as a breast cancer screening technique. Integer. Cancer There. 8:9–16, 2009. Doi: 10.1177/1534735408326171.
10. Mandelblatt, J. S., et al., Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. Ann. Intern. Med. 151:738–747, 2009
11. Zhixiong, G., and Kan-Wan, S., Simulated parametric studies in optical imaging of tumors through temporal log-slope difference mapping. Med. Eng. Phys. 29:1142–1148, 2007.
12. Tan, J. M. Y., Ng, E. Y. K., Acharya, R., Keith, L. G., and Holmes, J., Comparative study on the use of analytical software to identify the different stages of breast cancer using discrete temperature data. J. Med. Syst. 33:141–153, 2009. Doi: 10.1007/ s10916-008-9174-4.
13. Rafael C. Gonzalez and Richard E. Woods. Digital Image Processing. Prentice Hall, 2007. ISBN 0-13-168728-x.
14. Wenshuo Gao , Xiaoguang Zhang , Lei Yang , Huizhong Liu . An Improved Sobel Edge Detection. Computer Science and Information Technology (ICCSIT), pages 67 – 71, 2010 .
15. .M. Z. Jafri and F. Deravi. Enceinte algorithm for the detection of parabolic curves. In Vision Geometry III, volume 2356, pages 53-62. SPIE, 1994.
16. Qi, H., Kuruganti, P. T., and Snyder, W. E., Detecting breast cancer from thermal infrared images by asymmetry analysis, biomedical engineering handbook. CRC, Boca Raton, 2006(ch. 27–1 to 27–14).
17. V. Umadevi , S. V. Raghavan , Dr. Sandeep Jaipurkar .Interpreter for Breast Thermogram Characterization. IEEE EMBS Conference on Biomedical Engineering & Sciences, 2010.
18. B. F. Jones. A reappraisal of the use of infrared thermal image analysis in medicine. IEEE Transactions on Medical Imaging, 17(6):1019-1027, December 1998.
19. J. R. Keyserlingk, P. D. Ahlgren, E. Yu, N. Belliveau, and M. Yassa. Functional infrared imaging of the breast. IEEE Engineering in Medicine and Biology, pages 30-41, May/June 2000.
20. Keith, L. G., Oleszczuk, J. J., and Laguens, M., Circadian rhythm chaos: A new breast cancer marker. International Journal of Fertility and Women's Medicine. 46:238–247, 2001.
21. Bezdek, J. C., Pattern recognition with fuzzy objective function algorithms. Plenum, New York, 1981.
22. http://mathworld.wolfram.com/ K-Means Clustering Algorithm. html (last accessed Aug. 2008).
23. Sowmya, B., & Bhattacharya, S. (2005). Color image segmentation using fuzzy clustering techniques, IEEE Indicon, Conf .Chennai, India, pp: 41–45, Dec.
24. Solvenia , K. R. (2006). Fuzzy C-means clustering and facility location problems. Proceeding of Artificial Intelligence and Soft Computing. Palma de Mallorca, Spain.
25. M. EtehadTavakol, S. Sadri and E. Y. K. Ng . Application of K- and Fuzzy c-Means for Color Segmentation of Thermal Infrared Breast Images, Springer Science + Business Media, LLC 2008.

3/5/2011