First Order Statistical Feature for Breast Cancer Detection Using Thermal Images

Oky Dwi Nurhayati, Thomas Sri Widodo, Adhi Susanto, Maesadji Tjokronagoro

Abstract**—**In this paper, we report our experimentation to reveal the abnormalities of breast thermograms with tabulated the first order statistics method which are the mean values, standard deviation values, skewness values, and kurtosis values with thermal camera Fluke as a tool for capturing images, after the applications of wiener filter and histogram equalization to enhance the images and region of interest to obtain the specific object. We have used statistical features method to classify types of thermograms after image processing. The results show that the method are promising to detect the abnormality on the breast thermogram images. The normal breast thermograms have minimum standard deviation value and skewness value which differ from those abnormal thermograms in the early stage of breast cancer and the significantly from the advanced of breast cancer.

Keywords—breast thermogram, wiener filter, histogram equalization and region of interest methods, first order statistics method.

I. Introduction

TEXTURE are identifying features or characteristics owned object in a big region and naturally these characteristics could recur in this area. The small region when compared with elements of the available texture inside, could not show the texture itself. Actually, the same texture when saw at different scale would be seen like two different textures, when it was has big scale difference.

The other definition texture was the regularity of certain patterns that were formed from the structure pixel in the digital image. A surface had a texture information, when it region was enlarged without changing the scale, then the characteristics of the expansion surface had resemble the origin surface (regular pattern). The other word, they were emerge repeatedly in the interval of distance and certain direction. The texture information used to distinguish the surface characteristics of object in the image were connected with rough and soft, specific characteristics from coarseness and surface refinement, that completely free from the colour surface.

Oky Dwi Nurhayati is with the Department of Computer System, Faculty of Engineering, University of Diponegoro and Ph.D. Candidate in the Department of Electrical Engineering and Information Technology, Faculty of Engineering, Gadjah Mada University, Yogyakarta. INDONESIA. (e-mail: okydn@undip.ac.id).

Adhi Susanto is with the Department of Electrical Engineering and Information Technology, Faculty of Engineering, Gadjah Mada University, Yogyakarta. INDONESIA. (e-mail: adhi_susanto@jmn.net).

Thomas Sri Widodo is the Department of Electrical Engineering and Information Technology, Faculty of Engineering, Gadjah Mada University, Yogyakarta. INDONESIA. (e-mail: thomas@mti.ugm.ac.id).

Maesadji Tjokronagoro, is with the Department of Medical, Faculty of Medicine, Gadjah Mada University, Yogyakarta, INDONESIA.

The statistical method was used to extract the texture feature of an image. Image characteristics such as the arrange pixel intensity and texture feature namely contrast, entropy and homogeneity with other feature was counted from the image intensity.

Texture feature could be analysed from an image to identificate the characteristics of this image surface. Then, texture feature was count from a certain area and the values produced by the calculation was viewed or compared with the reference values. The measure of feature object was done after image processing with Wiener filtering, histogram equalization, and region growing methods to statistical feature of the thermograms In this research we were used four feature viz mean value, standard deviation value, skewness value, and kurtosis value and plotting them in two dimention graph to classify types of the breast thermogram images.

II. The Objective of the Research

To probe further the breast cancer detection from breast thermogram images, we use some standard image processing such as wiener filter, histogram equalization and region of interest to enhance their overall visibility filtering to reduce the unwanted interference as well as to increase the noticeability of residing objects of interest.

In this case, first order statistics applied after image processing of the thermograms to obtain the interest object and measure it. Objects classification to separate normal and abnormal thermograms for breast cancer detection.

Therefore our research objective is to detect breast cancer with applying the appropriate image processing techniques and first order statistics.

III. The Underlying Theory

Several methods in image processing are Wiener filtering, histogram equalization, and region growing.

A. Wiener Filtering

The most important technique for removal of blur in images due to linear motion or unfocussed optics is the Wiener filter. From a signal processing standpoint, blurring due to linear motion in a photograph is the result of poor sampling. Each pixel in a digital representation of the photograph should represent the intensity of a single stationary point in front of the camera. Unfortunately, if the shutter speed is too slow and the camera is in motion, a given pixel will be an amalgram of intensities from points along the line of the camera's motion.

B. Histogram Equalization

In an image processing context, the histogram of an image normally refers to a histogram of the pixel intensity values. This histogram is a graph showing the number of pixels in an image at each different intensity value found in that image. For an 8-bit grayscale image there are 256 different possible intensities, and so the histogram will graphically display 256 numbers showing the distribution of pixels amongst those grayscale values. The exact output from the operation depends upon the implementation and it may simply be a picture of the required histogram in a suitable image format, or it may be a data file of some sort representing the histogram statistics. The histogram of a digital image with gray values : r_0, r_1, r_2,…r_{L-1} is the discrete function:

$$p(r_k) = \frac{n_k}{n}, k = 0,1,..., L-1$$

n: total Number of pixels in the image

The function $p(r_k)$ represents the fraction of the total number of pixels with gray value r_k. Histogram used to give an information about pixel intensity, number of pixel, and changes of the view by image processing operation.

The number of different light intensities in an image often does not use the whole available spectrum and mostly accentuate a narrow spectrum. Images with such poor intensity distributions can be helped with a process known as histogram equalization. Histogram equalization is a method for "spreading" the histogram of pixel levels more evenly.

Histogram equalization can be used to improve the visual appearance of an image. Peaks in the image histogram (indicating commonly used grey levels) are widened, while the valleys are compressed.

C. Statistical Feature Extraction

The n^{th} moment of the (normalized) gray level histogram is given by:

$$\mu_n = \sum_{i=1}^{L} (k_i - m)^n p(k_i)$$

where:
k_i = gray value of the i^{th} pixel
m = mean gray value of the pixel set
L = the number of distinct gray levels
$p(k_i)$ = normalized histogram (probability density function of the pixel set).
Note that the mean is given by:

$$m = \sum_{i=1}^{L} k_i p(k_i)$$

Thus: $\mu 0 = 1$; $\mu 1 = 0$; $\mu 2 = s^2$ = variance
Variance is a square of standard deviation.
$\mu 3$ = skewness ; $\mu 4$ = kurtosis

First-order statistics are limited as texture descriptors since they carry no information about the relative position of pixels with respect to one another.

D. Region Growing

The main goal of segmentation is to partition an image into regions. Region-based segmentation is a technique for determining the region directly. The basic formulation for region-based segmentation is:

$$\cup_{i=1}^{n} R_i = R$$

where R_i is a connected region, i = 1, 2, ...,n

$R_i \cap R_j = \phi$ for all i = 1,2,…

$P(R_i)$ = true for i = 1,2,…,n.

$P(R_i \cup R_j)$ = false for any adjacent region R_i and R_j

$P(R_i)$ is a logical predicate defined over the points in set $P(R_k)$ and Φ is the null set. Explanation for region-based segmentation are means that the segmentation must be complete; that is, every pixel must be in a region; requires that points in a region must be connected in some predefined sense; indicates that the regions must be disjoint; deals with the properties that must be satisfied by the pixels in a segmented region. For example $P(R_i)$ = TRUE if all pixels in R_i have the same gray level; indicates that region R_i and R_j are different in the sense of predicate *P*.

IV. THE RESULTS OF THE RESEARCH

Breast thermography reveals the infrared heat emittion distributions differently from normal and abnormal breasts. The abnormalities can be due to the existence of cysts, infections, and malignant tumors or other fibrocystic desease. A proper image processing scheme is to be developed to enhance the readability of the thermograms to ease the doctor's diagnosis. The results of our experiments are shown in Figure 1, where the object is convert to grayscale image and filtering by wiener filter.

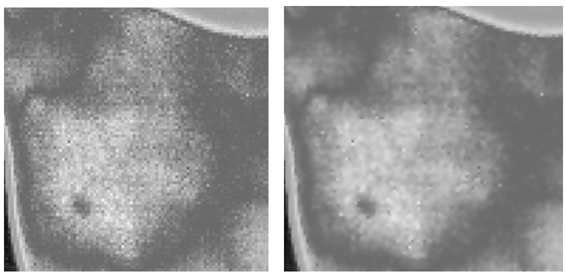

Figure 1 Original thermogram and after wiener filtering

Fig. 2 shows histogram equalization and region growing methods of the breast thermogram.

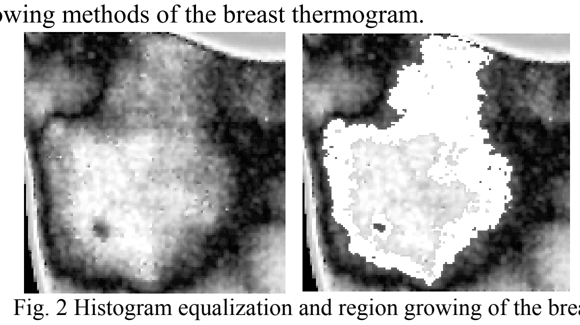

Fig. 2 Histogram equalization and region growing of the breast thermogram

Fitted curve of skewness value as a function of mean value shows in Figure 3.

Figure 4 shows fitted curve of standard deviation value as a function of mean value.

Figure 5 shows fitted curve of entropy value as a function of mean value.

The normal breast thermograms have the smallest standard deviation value and skewness value which differ from those abnormal thermograms in the early stage of breast cancer and the significantly from the advanced of breast cancer. Meanwhile they have the biggest of mean value which differ from those advanced and early breast thermograms.

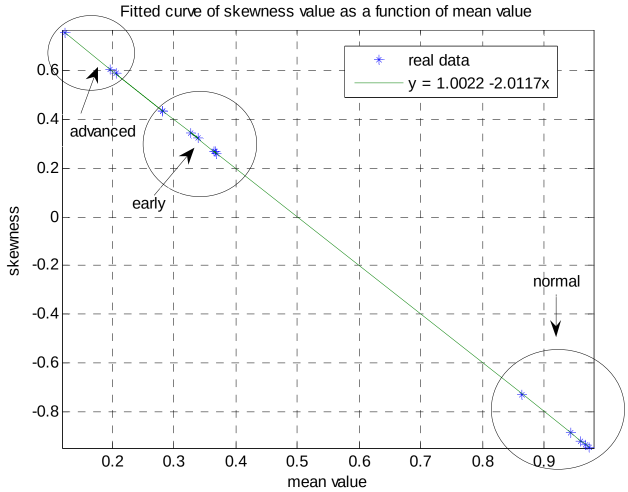

Fig. 3 Plot classification of mean and skewness value

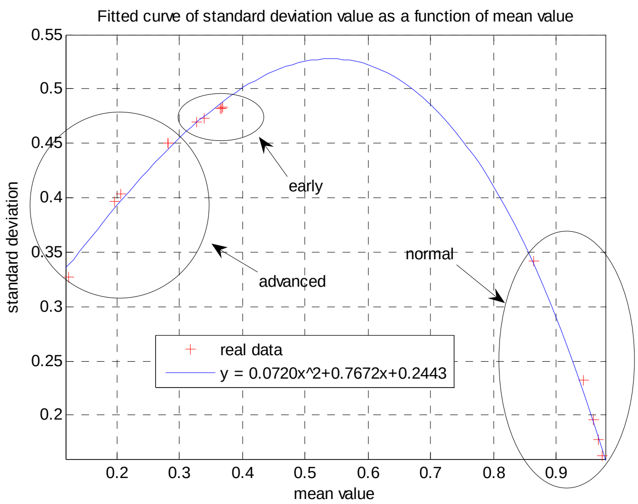

Fig. 4 Plot classification of mean and standard deviation value

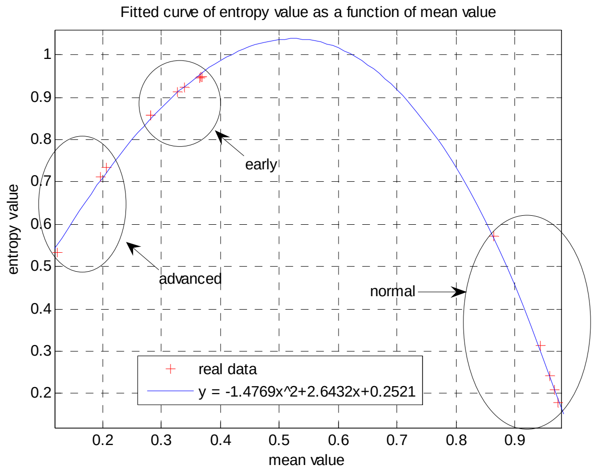

Fig. 5 Plot classification of mean and entropy value

V. Conclusion

A. The experimental results show that first order statistics method are promising to detect the abnormality on the breast thermogram images.

B. Further experiments should be coupled with spectral and structural methods for analysis.

References

[1] Jain, A. K., Fundamentals of Digital Image Processing, Prentice-Hall, Inc., A Division of Simon & Schuster Engelwood Cliffs, New Jersey, 1989.

[2] Gonzalez, R.C., Richard E. Woods, Digital Image Processing, Prentice-Hall, Inc., Upper Saddle River, New Jersey, 2008.

[3] Keyserling, J.R, P.D.Ahlgren, E.Yu, and N.B.Belliveau, Overview of functional infrared imaging as part of a multi-imaging strategy for breast cancer detection and therapeutic monitoring, Proc. 2nd Joint IEEE EMBS/BMES Conf., Houston, TX, pp. 1126-8, 2002.

[4] Haralick, R. M., Shanmugam, K., and Dinstein, I., Textural features for image classification, IEEE Trans. Syst., Man, Cybern., SMC, 1973

Oky Dwi Nurhayati : Semarang, 2 October 1979. Graduate from the Department of Telecommunication Engineering, Faculty of Engineering, Telecommunication Engineering Institute, Indonesia on 2002. Master from Post Graduate Program, Department Electrical Engineering, Gadjah Mada University, Indonesia on 2008. 2008—now is Ph. D. Candidate in the Department of Electrical Engineering, Faculty of Engineering, Gadjah Mada University, Yogyakarta, Indonesia.

From 2009—now is Lecturer in the Department of Computer System, faculty of Engineering, Diponegoro University, Indonesia. Current and interest research is in biomedical image processing.

Oky Dwi Nurhayati, S.T, M.T., membership in IAENG (International Association of Engineer, IMAVIS (International Journal on Image and Vision Computing).