Feature Extraction for Classification of Microcalcifications and Mass Lesions in Mammograms

Rabi Narayan Panda
Associate Professor and HOD
Department of MCA
Krishna Institute of Engineering and Technology
Ghaziabad, Uttar Pradesh, India

Dr. Bijay Ketan Panigrahi
Assistant Professor
Department of Electrical Engineering
Indian Institute of Technology
Delhi, India.

Dr. Manas Ranjan Patro
Professor and HOD
Department of Computer Science
Berhampur University
Berhampur, India

Summary
Mammography is the most contemporary option for the premature detection of breast cancer in women. Nevertheless, the opinion of the radiologist has a remarkable influence on the elucidation of the mammogram. The proposed research intends to develop an image processing algorithm for the recognition of microcalcifications and mass lesions to aid the premature detection of breast cancer. The work proposed deals with a novel approach for the extraction of features like microcalcifications and mass lesions in mammograms for early detection of breast cancer. The proposed technique is based on a three-step procedure: (a) regions of interest (ROI) specification, (b) two dimensional wavelet transformation, and (c) feature extraction based on OTSU thresholding the region of interest for the identification of microcalcifications and mass lesions. ROIs are preprocessed using a wavelet-based transformation method and a thresholding technique is applied to exclude microcalcifications and mass lesions. The method suggested for the detection of microcalcifications and mass lesions from mammogram image segmentation and analysis was tested over several images taken from mini-MIAS (Mammogram Image Analysis Society, UK) database. The implementation of the algorithm was carried out using Matlab codes programming and thus is capable of executing effectively on a simple personal computer with digital mammogram as accumulated data for assessment

Key Words:
Mammogram, Breast Cancer, Microcalcifications, Mass lesions, wavelet transformation, OTSU thresholding,

1. Introduction

Breast cancer is considered as one of the primary causes of women mortality [1]. The mortality rate in asymptotic women can be brought down with the aid of premature diagnosis. Despite the increasing number of cancers being diagnosed, the death rate has been reduced remarkably in the past decade due to the screening programs [2]. Premature detection of breast cancer increases the prospect of survival whereas delayed diagnosis frequently confronts the patient to an unrecoverable stage and results in death [3]. According to the World Health Organization's International agency for Research on Cancer in Lyon, France, every year more than one million women worldwide are affected with breast cancer [4]. Well heeled families carry the greatest risk, having incidence rates of >80 per 100,000 population every year. Numerous etiological factors, including reproductive history (early menarche, late or no pregnancy), and Western lifestyle (high caloric diet, little or no of physical activity) and the like have been recognized to be the causative agents of the epidemic of breast cancer. Breast cancer tops the list of the cancers that American women suffer from. According to the American Cancer Society 215 990 new cases of breast carcinoma has been detected in the United States alone in 2004. It is one of the major reasons of deaths occurring due to cancer in women well under the age of 65 [5]. The Indian metropolises of Mumbai, Calcutta, and Bangalore display 23% of all the female cancers as breast cancers followed by cervical cancers (17.5%). Despite the fact that the incidence of breast cancer in India is comparatively lower than that of the western countries, the issue is highly alarming. Hormone dependent cancers have been found to be caused by the organ chlorines [6]. High quality images and mammographic interpretation are mandatory for the detection of premature and delicate symptoms of breast cancer. Mammogram (breast X-ray) is the medical image essential for the diagnosis of breast cancer and is considered to be the most dependable technique for premature detection.

The widely recognized tool for the early detection of breast cancer in women with no symptoms; and to detect and diagnose breast disease in women experiencing symptoms like a lump, pain or nipple discharge, is mammography. Contemporarily, screening mammography and radiographic imaging of the breast are the most effective tools for premature detection of breast cancer. Screening mammographic assessments are carried out on asymptomatic woman to detect premature, clinically unsuspected breast cancer [7]. Still, studies have proved that all breast cancers that are retrospectively detected on

Manuscript received May 5, 2009
Manuscript revised May 20, 2009

the mammograms are not detected by radiologists [8], [9]. Due to the subtle and complex nature of the radiographic findings related with breast cancer, human factors such as varying decision criteria, distraction by other image features, and simple oversight can be responsible for the errors in radiological diagnosis. Computer assisted schemes that work on image processing and pattern recognition techniques can be utilized to enhance the diagnostic efficiency and for the location and classification of probable lesions and thereby alerting the radiologist to observe these areas with specific attention. Furthermore, the computer-assisted schemes can enhance the performance of the automatic computer-aided diagnosis systems that are capable of serving as a “prereader” to the radiologist and offer him the “second opinion” in the diagnosis [10].

Computer-Aided Detection (CAD) systems [11] can be utilized for the enhancement of the sensitivity of screening mammography. The deployment of computer-aided diagnosis (CAD) schemes to aid radiologists in the analysis of mammograms has been suggested by numerous researches [12]. CAD has been developed greatly during the past few years to enhance the diagnostic accuracy of both the recognition and categorization of masses [13, 14] and clustered microcalcifications [15, 16], symptoms that may signify the existence of breast cancer. Radiologists look out for particular abnormalities on mammograms visually. Some significant signs that radiologists pay attention to are clusters of microcalcifications, masses, and architectural deformations. A space-occupying lesion that is visible at more than one projection is referred to as a mass. Masses are illustrated with the aid of shape and margin features. Tiny deposits of calcium those are visible as minute bright spots on the mammogram are called as calcifications. They are exemplified by their type and distribution characteristics. The existence of microcalcifications is one of the significant and probably the only indication of cancer on a mammogram [17]. A majority of the researches on computer analysis of mammograms have focused on the detection of small abnormalities, precisely the microcalcifications.

1.1. Our Contribution

The intricate appearance of the structures of interest in mammograms is the reason behind the process of working with them being tedious. Despite the fact that a mammogram is a fine picture of the breast it is of scarce sufficiency while looking for minute or complex anatomical parts, like the microcalcifications, masses or curvilinear structures during the process of premature detection of breast cancer. Here, we present a novel system for detecting clustered microcalcifications and Mass lesions in digital mammography. This paper proposes a system designed to perform prescreening of digital mammograms for the presence of microcalcifications and Mass lesions based on wavelet decomposition and Otsu Threshold Method. Interpreting medical images that are used for diagnosing process involves preprocessing and detection of regions of interest. Preprocessing stage deals with image enhancement and noise removal. The enhanced image is then scanned for selected region of interest. Further, the two dimensional wavelet transformations, precisely daubechies wavelet are applied to get the decomposition co efficient. Histogram equalization is one of many techniques that used to enhance mammograms. The next stage is to extract features from the region of interest. Otsu method is one such thresholding method and is frequently employed in various areas to extract the features in digital mammogram. Once the clusters of microcalcification and mass lesions are extracted, these can be categorized as benign or malignant. Clustered microcalcifications and mass lesions are one of the earliest signs of potential cancerous changes in breast tissue.

The paper is organized as follows: Section 2 presents a review of existing techniques for mammographical feature analysis. Section 3 presents the brief description of mammographic abnormalities. Section 4 detailed the proposed methodology for microcalcification and mass lesion detection. Section 5 describes results and discussion. Conclusion is summed up in Section 6.

2. Literature Review

An algorithm comprising of numerous phases to attain automatic detection of clusters was developed by Cairns et al. [18]. Initially, the following presumptions are made in order to represent microcalcifications in digital mammograms: microcalcifications are small in size, generally linear or round in shape, they are typically brighter than the neighboring pixels, their brightness value is relatively unvarying across their surface, and all of them possess well-defined edges. Ultimately they also projected that microcalcifications are vital only if they occur in groups or clusters. Upon this presumption, an algorithm comprising of the subsequent phases was utilized: edge detection, contour hue generation, location of potential microcalcifications with the aid of graph searching, feature extraction, categorization of the potential microcalcifications, and cluster detection. They were capable of achieving a classification rate of 91.75% for single microcalcifications. Besides, they also attained 100% true-positives with 0% false positives using the re-substitution method, and 98% true-positives with 0% false positives with the aid of the leave one- out method for clustered microcalcifications.

A computerized methodology for the automatic identification of clustered microcalcifications was proposed by Nishikawa et al. [19].Their methodology comprises of three phases: First the signal-to noise ratio of the microcalcifications is improved by filtering the image to decrease the normal background structure of the mammogram. Second, signals (potential microcalcifications) are recognized with the aid of global gray-level thresholding, morphological erosion and subsequently by a local adaptive gray-level thresholding. Third the number of falsely identified signals is brought down by 1) probing the power spectrum of individual signals, 2) computing the spatial distribution of the signals, and 3) examining the relationship between sizes, shape and background pixel value of microcalcifications.

A technique for the adaptive thresholding of the gray-level mammographical images was proposed by Zhao et al [20]. This approach merges the filtering operations with a rule-base. The extraction of the apprehensive areas from a mammogram and proffering the location information on certain microcalcifications of predefined shapes and sizes to radiologists for future assessment are the motives behind the proposed research. An adaptive threshold function was derived by the authors from the morphological operations. : The granular form, casting form, microcalcification size, and microcalcification density were some of the relevant factors for the derivation of an adaptive threshold function. The index number in the skeleton of shapes that denote the microcalcifications in mammograms control the threshold set. The parameters of the adaptive threshold sets are acquired through the interpretation of umbra shadows from an image function.

A microcalcification detection algorithm that works on digital mammograms by merging morphological image processing with arithmetic processing was proposed by Mascio et al. [21]. The algorithm starts with the application of two high-frequency assessments to the original digital mammogram. The first assessment highlights any detail in the image that transforms sharply in intensity and is greater than several pixels in size. Details that are minute and textured are highlighted by the second assessment. Areas common to both the assessments are saved for thresholding. This resulted in the identification of microcalcifications and apprehensive areas.

Barman et al. [23] utilized a low-pass filter to detect microcalcification while analyzing digital mammogram. Besides the fact that their system is still under development the preliminary results were satisfactory yet with further modifications still to be carried out. Methodologies based on wavelet are image-processing techniques that can be utilized in the identification of microcalcifications in digital mammograms.

The KNN algorithm was tailored by Woods et al. [24] postulating that an unidentified test pattern is consigned to a particular class if at least of the KNNs is in that class. The NN rule will be highly responsive to the microcalcification detection and less responsive to non microcalcifications. Bayesian approaches to classification have been employed productively in their application for the diagnosis of microcalcifications [22], [25]. Its decision making process works on basis of opting for the most probable class when a certain feature vector is provided. A likelihood of class membership is determined and utilized for the classification of an area or object.

A system that was built on fuzzy logic that comprises of image fuzzification enhancement, irrelevant structure removal segmentation, and reconstruction was projected by Cheng et al. [26]. Chan et al. [27] examine a convolution neural network based approach that is efficient of bringing down false positive detections. Three feature analysis methods based on rules, artificial neural networks, and a combination of both were compared by Nagel et al. [28]. They report that the combined method performs best because each filter eliminates different types of false positives in mammogram analysis. McGarry and Deriche [29] use a hybrid model combining knowledge of mammographic imaging process and anatomical structure and Markov random fields.

A novel method for the automatic recognition of malignant clusters that works on the adoption of a Multiple Expert System (MES) was projected by Cordella et al. [30]. Their method was found to be particularly suitable for identification of malignant clusters since it permitted the decision about the malignancy of a cluster to be finilized on the basis of the confirmation arriving from both the microcalcifications and the entire cluster. The experimental evaluation was done on a standard database of 40 mammographic images and the approach yielded satisfactory results which confirmed its effectiveness.

A CAD system for the recognition and categorization of microcalcification in digital mammograms was put forth by Brijesh Verma and John Zakos [7]. Fourteen feature extraction techniques, neural-network settings, and a FL detection algorithm were studied and assessed by them. They altered some conventional characteristics and established that a combination of their three modified characteristics like entropy, standard deviation, and number of pixels, was the finest combination of features to differentiate a benign microcalcification pattern from one that is malignant. Additionally, they have also proved that a threshold of 0.3 could further enhance the system. Their system obtained promising results, with 88.9% being the finest.

The application of SVM for detection of MCs in digital mammograms was proposed by Issam El-Naqa et al [32]. In their technique, an SVM classifier was trained with the aid of supervised learning to examine at every location in a mammogram for the presence or absence of MC. The principle of structural risk minimization formed the basis for the formulation of SVM learning. The decision function of the trained SVM classifier is computed in terms of support vectors that were recognized from the examples throughout the training. Results illustrated that the SVM classifier attains low generalization error upon application for the classification of samples that were excluded in training. Additionally, their proposed SEL system was capable of leading to the enhancement of the performance of the trained SVM classifier

An ensemble classifier for the computer-aided diagnosis of breast microcalcification clusters that are tedious to characterize according to the radiologists and computer models alike was proposed by Joseph et al. [33]. The primary intent of the study was to aid radiologists in to recognize if doubtful calcification clusters are benign or malignant, so that they may potentially suggest fewer needless biopsies for actually benign lesions. 292 cases from a publicly available mammography database were utilized by them.

A novel scheme for the microcalcification recognition on basis of fuzzy logic and scale space techniques was proposed by Cheng et al. [34]. Initially the fuzzy entropy principal and fuzzy set theory were employed to fuzzify the images. Subsequently, the images were enriched. Ultimately scale-space and Laplacian-of-Gaussian filter techniques were employed to identify the sizes and locations of microcalcifications. Performance evaluation was done with the aid of free-response operating characteristic curve. The primary benefit of their scheme was its ability to identify microcalcifications even in the mammograms of very dense breasts.

Bottigli et al. [35] presented a comparison of some classification system for massive lesion classification. An algorithm based on morphological lesion differences was used to extract the features. The two classes (pathological or healthy ROIs) were differentiated by utilizing the features. A supervised neural network was employed to check the discriminating performances of the algorithm against other classifiers and the ROC curve was used to present the results. In comparison with the other recent studies [38, 39, 40]; the results of the new representation applied are comparable or better, owing to its better ability to distinguish pathological ROIs from the healthy ones.

Masala et al. [36] presented a comparison of different algorithms used in a Computer Aided Detection (CAD) system for classification of masses. Pre-processing and segmentation, region of interest (ROI) search, and feature extraction and classification is the three step procedure based on which the CAD system was designed. Testing was performed on a very large mammographic database. The classification step of their algorithm starts from eight features extracted from a co-occurrence matrix in which second-order spatial statistics information on ROI pixel grey levels is present. Multi Layer Perceptron, Probabilistic Neural Network, Radial Basis Function Network and the K-Nearest Neighbors' classifiers are the algorithms they implemented and tested as classifiers. Radial Basis Function and a Multi Layer Perceptron provide the best results and outperform other classification algorithms in terms of the area under the ROC curve by almost 8%.

A new application of SVM in breast MR image classification was presented by Chuin-Mu Wang et al. [37]. The idea of hyper-plane classifier is the basis of the SVM that seeks for the hyper-plane that maximizes the margin between two classes. Hence, SVM-based classifier can be tailored for employing in MR image classification. The experimental results prove that SVM outperforms CM in the results of real MR image classification and the classification results, a basis of diagnosis and judgment for the patient condition can be provide to the doctor.

3. Mammographic Abnormalities

Numerous characteristics that may signify a probable clinical problem, including asymmetries between the breasts, architectural distortion, confluent densities associated with benign fibrosis, calcifications and masses and the like are identified with the aid of Mammography. The two most customary characteristics associated with cancer are clusters of micro calcifications and masses both of which are discussed subsequently. The detection of micro calcification has been explored by various groups of researchers. Small (sub 15mm), low contrast masses are considered more critical than microcalcifications, since they are more difficult to detect [41]. Of chief concern are the masses that are not accompanied by micro calcifications since they are tumors that develop drastically. Unlike micro calcifications that are well apparent as bright spots, the masses merge with the breast structure in such a way that boundaries are indistinct, and can often be completely hidden from vision if the breast is dense.

3.1. Calcification

Tiny deposits of minerals (calcium) that appear like localized high-intensity regions (spots) in the mammogram are known as calcifications. Calcifications are one of the significant and widespread finding that are frequently apparent in a mammogram. Micro-calcifications and macro calcifications or coarse calcifications are the two common categories of calcifications. Macro-calcifications are coarse calcium deposits that are spread about the breast. Commonly such deposits are accompanied by benign conditions and hardly necessitate a biopsy. The benignity or malignancy of the tumor is indicated by the number calcifications that comprise a cluster. Micro-calcifications are minute (less than 1/50 of an inch or ½ of a millimeter) spots of Calcium deposits that may exist in an area of rapidly dividing cells [42]. They are possibly intramammary, inside and around the ducts, inside the lobules, in vascular structures, in interlobular connective tissue or fat. The onset of cancer might be indicated by the presence of micro-calcifications in a cluster. Almost half of the cancers identified through mammography come into sight as cluster of micro-calcifications. Generally ductal carcinoma in situ (an early cancer confined to the breast ducts) is identified when micro-calcifications become apparent through mammography. The morphology of calcifications is considered to be the most important indicator in differentiating benign from malignant [43].

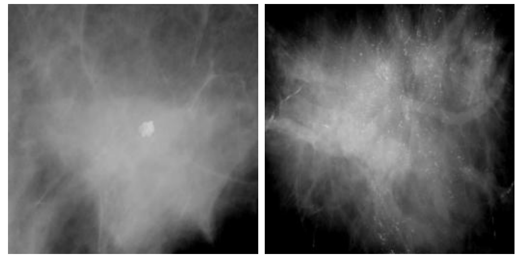

Fig. 1. (a) A mammogram sample showing two isolated coarse calcifications; (b) a microcalcification cluster.

Three categories of calcifications have been identified by the "The American College of Radiology (ACR) BIRADS" (Table: 1):

(a) Typically benign
(b) Intermediate concern
(c) High probability of malignancy

Table 1 Summary of BIRADS Classification of Calcifications

	Type of calcification	**Characteristics**
Typically benign	Skin	Typical lucent center and polygonal shape
	Vascular	Parallel tracks or linear tubular calcifications that run along a blood vessel
	Coarse or pop-corn like	Involuting fibroadenomas
	Rod-shaped	Large rod-like structures usually > 1mm
	Round	Smooth, round clusters
	Punctuate	Round or oval calcifications
	Spherical or lucent centered	Found in debris collected in ducts, in areas of fat necrosis
	Rim or egg-shell	Found in wall of cysts.
	Milk or calcium	Calcium precipitates
	Dystrophic	Irregular in shape but usually large > 0.5mm in size
Intermediate concern	Indistinct or amorphous	Appear round or flake shaped, small and hazy uncertain morphology
High risk	Pleomorphic or heterogenous	Cluster of these calcifications irregular in shape, size and < 0.5mm raises suspicion
	Fine, linear or branching	Thin, irregular that appear linear from a Distance

3.2. Mass Lesions

Breast cancer is characterized with the presence of a mass accompanied or not accompanied by calcifications [44]. There is a possibility of a cyst, which is non-cancerous collection of fluid to resemble a mass in the film. The identical intensities of the masses and the normal tissue and similar morphology of the masses and regular breast textures makes it a tedious task to detect masses in comparison with that of calcifications [45].

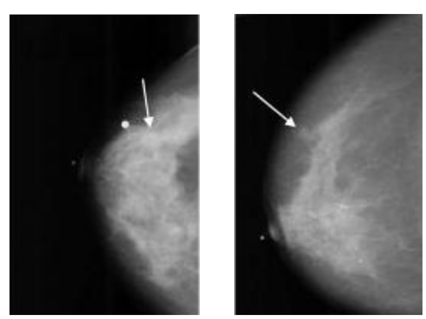

Fig. 2. (a) Dense breast containing a malignant mass. (b) Fatty and glandular breast containing a malignant mass.

The location, size, shape, density, and margins of the mass are highly beneficial for the radiologist to evaluate the probability of cancer. A majority of the benign masses are well circumscribed, compact, and roughly circular or elliptical whereas the malignant lesions are characterized by blurred boundaries, irregular appearances and are occasionally enclosed within a radiating pattern of linear spicules [46]. Nevertheless some benign lesions may also possess spiculated appearances or blurred peripheries.

3.3. Data Sources

It is difficult to access real medical images for experimentation due to privacy issue. The proposed research makes use of the data collection obtained from Mammographic Image Analysis Society (MIAS) [47]. This collection has been employed in numerous other researches intended towards automatic mammography classification as well. The collection comprises of 322 images that fall into one of the following categories: normal, benign and malign. Malign images are regarded as abnormal. Additionally, the malign cases are further classified into six namely: circumscribed masses, spiculated masses, microcalcifications, ill-defined masses, architectural distortion and asymmetry. Every image is digitized at a resolution of 1024x1024 pixels and eight-bit accuracy (gray level). Besides, the locations of the existing anomalies are included as well. Data present in the collection comprises of the location of the abnormality (such as the center of a circle surrounding the tumor), its radius, breast position (left or right), type of breast tissues (fatty, fatty-glandular and dense) and tumor type if present (benign or malign).

4. Methodology

The essential improvement that is required in mammography is the escalated contrast, particularly for dense breasts. There is a possibility for the difference in malignant tissue and normal dense tissue to exist in the mammogram but beyond the threshold of human perception. Likewise, microcalcifications in a dense mass may not be perceptible owing to low contrast. Thus, the definition of features of microcalcifications and mass lesions are tedious. The conventional methodology for the detection of clusters of microcalcifications and mass lesions is a multiple step process that includes: a) Regions of Interest Specification, b) application of Two-dimensional wavelet transform, c) feature extraction based on OTSU threshold method.

4.1. Regions of Interest Specification

The mammograms of miniMIAS database describe diverse areas like the image background, the tissue area, and informative marks. In order to segment the ROIs from breast tissue, it is presumed that pixels that constitute a ROI need to be members of a set of adjoining neighbor pixels with appropriate intensities. The “minimum intensity threshold” and “maximum intensity threshold” are the two thresholds that are utilized for determining the appropriate intensities. According to the observations the diameters of masses fall between upper and lower boundaries [48]. Thus in order to comprehend if a pixel is present in the center region of the ROI, the diameter of the ROI (assuming the pixel in question is the center) needs to be considered foremost. The first step comprises of the manual choice of the number of ROIs for every mammogram. The location selections were made under the supervision of the radiologists involved in the study and this facilitated in obtaining non overlapping ROIs for every mammogram, besides covering most of the breast density. This location also guarantees that we carried out our assessment only with the breast tissue, devoid of the bias brought about by the pectoral muscle or imaging artefacts.

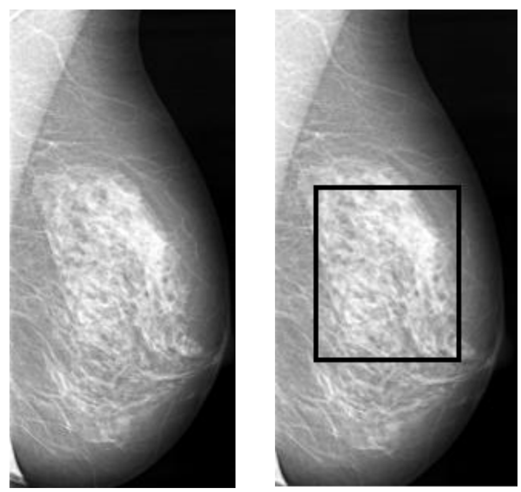

Fig. 3 (a) A mammogram sample showing mass lesion (b) Mammogram with marked ROI

4.2. Application of Two-Dimensional Wavelet Transform

Images are preprocessed with the aid of a wavelet-based spatially adaptive method for mammographic contrast enrichment [49], [50]. Wavelet transform techniques were utilized in the computation of breast ROI. The wavelet transform or wavelet analysis is considered to be the contemporary solution to surmount the drawbacks of Fourier transform. A wavelet can be defined as a waveform of efficiently restricted duration and that has a mean value of Zero [51]. The process of splitting up a signal into and scaled versions of the original (or mother) wavelet is referred to as Wavelet analysis. Daubechies wavelets are the most widely employed orthogonal wavelets [52, 53]. The merit of this family of wavelets is that they are efficiently sustained (finite length) in space. Thus they might tend to have higher correlation to small size structures like the microcalcification and mass lesions that the other wavelets of infinite extents in space. The high frequency components in the resultant image are enriched whereas the low frequency background structure was removed. A global threshold value was applied on the reconstructed image acquired for each mammogram and a binary image providing all the probable points of microcalcifications was formed. The threshold value was set on basis of the values of the wavelet coefficients. Numerous maxima values of the wavelet coefficients were calculated for each line constituting the image in order to obtain the coefficients. The pixel positions analogous to values of the wavelet coefficients those are higher than an imposed restraint value were regarded as points of microcalcifications or mass lesions and were rendered to the binary image.

4.3. Microcalcification and Mass Lesions Feature Extraction

It is necessary to identify the breast tissue so as to extract characteristics related to Microcalcification and mass Lesions. Removal of the background is the foremost step in segmentation. A global threshold is applied on the image to achieve the same. The threshold value is automatically obtained from the grey level histogram with the application of a peak detection method

4.3.1. Histogram Equalization

Features are based on the grey-level histograms from selected regions of the breast. The distances to the skin normalized from 0 to 100 (providing invariance to the size of the breast) are utilized in the construction of the regions. Histogram modeling techniques alter an image in order to ensure that the histogram is of the desired shape. This is beneficial for the elongation of low levels of mammograms with the narrow histograms. Histogram equalization is a conventional histogram modeling methodology. Let us regard the mammogram histogram as a probability distribution. According to the information theory, the uniform distribution attains maximum entropy, which encloses the most information. Thus, the mammogram information needs to be maximized in order to redistribute the gray-levels and achieve a the most uniform histogram

4.3.2. OTSU Threshold

A significant technique for image segmentation that attempts to recognize and extract a target from its background with the aid of the distribution of gray levels or texture in image objects is referred to as Thresholding. The statistics of the one-dimensional (1D) histogram of gray levels and on the two-dimensional (2D) co-occurrence matrix of an image form the basis of a majority of the thresholding techniques. Precisely, the discriminant criterion chooses the optimal threshold in order to maximize the separability of the resultant classes in gray levels. The procedure makes use of only the zeroth- and the first-order cumulative moments of the gray-level histogram and hence is trouble-free. It is possible to extend the method to multithreshold problems in an uncomplicated manner [54].

An image is a 2D grayscale intensity function, and contains N pixels with gray levels from 1 to L . The number of pixels with gray level i is denoted f_i, giving a probability of gray level i in an image of

$$p_i = f_i / N \tag{1}$$

In the case of bi-level thresholding of an image, the pixels are divided into two classes, C_1 with gray levels $[1,\cdots,t]$ and C_2 with gray levels $[t+1,\cdots,L]$. Then, the gray level probability distributions for the two classes are

$$C_1 : p_1/\omega_1(t), \cdots p_t/\omega_1(t)\ and$$
$$C_2 : p_{t+1}/\omega_2(t), p_{t+2}/\omega_2(t), \cdots, p_L/\omega_2(t), \tag{2}$$

Where $\omega_1(t) = \sum_{i=1}^{t} p_i$ and $\omega_2(t) = \sum_{i=t+1}^{L} p_i$ (3)

Also, the means for classes C_1 and C_2 are

$$\mu_1 = \sum_{i=1}^{t} i\ p_i / \omega_1(t) \tag{4}$$

And

$$\mu_2 = \sum_{i=t+1}^{L} i\ p_i / \omega_2(t) \tag{5}$$

Let μT be the mean intensity for the whole image. It is easy to show that

$$\omega_1\mu_1 + \omega_2\mu_2 = \mu_T \tag{6}$$

$$\omega_1 + \omega_2 = 1 \tag{7}$$

Using discriminant analysis, Otsu defined the between-class variance of the threshold image as [31]

$$\sigma_B{}^2 = \omega_1(\mu_1 - \mu_T)^2 + \omega_2(\mu_2 - \mu_T)^2 \tag{8}$$

For bi-level thresholding, Otsu verified that the optimal threshold t^* is chosen so that the between-class variance $\sigma_B{}^2$ is maximized; that is,

$$t^* = Arg\ Max\{\sigma_B{}^2(t)\}, \quad 1 \le t \le L \tag{9}$$

The previous formula can be easily extended to multilevel thresholding of an image. Assuming that there are $M-1$ thresholds, $\{t_1, t_2, \cdots, t_{M-1}\}$, which divide the original image into M classes: C_1 *for* $[1, \cdots, t_1]$, C_2 *for* $[t_1+1, \cdots, t_2], \cdots, C_i$ *for* $[t_{i-1}+1, \cdots, t_i], \cdots, and C_M$ for $[t_{M-1+1}, \cdots, L]$, the optimal thresholds $\{t_1{}^*, t_2{}^*, \cdots, t_{M-1}{}^*\}$ are chosen by maximizing $\sigma_B{}^2$ as follows:

$$\{t_1{}^*, t_2{}^*, \cdots, t_{M-1}{}^*\} = Arg\ \ Max\ \{\sigma_B{}^2(t_1, t_2, \cdots, t_{M-1})\}$$
$$1 \le t_1 < \cdots < t_{M-1} < L \tag{10}$$

Where $\sigma_B{}^2 = \sum_{k=1}^{M} \omega_k(\mu_k - \mu_T)^2$ (11)

With

$$\omega_k = \sum_{i \subseteq Ck} p_i \tag{12}$$

$$\mu_k = \sum_{i \subseteq Ck} i\ p_i / \omega(k) \tag{13}$$

The ωk in Eq. (12) is regarded as the zeroth-order cumulative moment of the k^{th} class C_k, and the numerator in Eq. (13) is regarded as the first-order cumulative moment of the k^{th} class C_k; that is,

$$\mu_k = \sum_{i \subseteq Ck} i\ p_i$$

5. Results and Discussion

The mammograms of our study were analyzed to detect possible clusters of microcalcifications and mass lesions employing the automated detection scheme described above. The method suggested for the detection of microcalcifications and mass lesions from mammogram image segmentation and analysis was tested over several images taken from mini-MIAS (Mammogram Image Analysis Society, UK) database. Figure 4 shows an example of an original image containing a mass lesion, and the results of the detection procedure. Figure 5 shows an example of an original image containing a cluster of microcalcifications, and the results of the detection procedure.

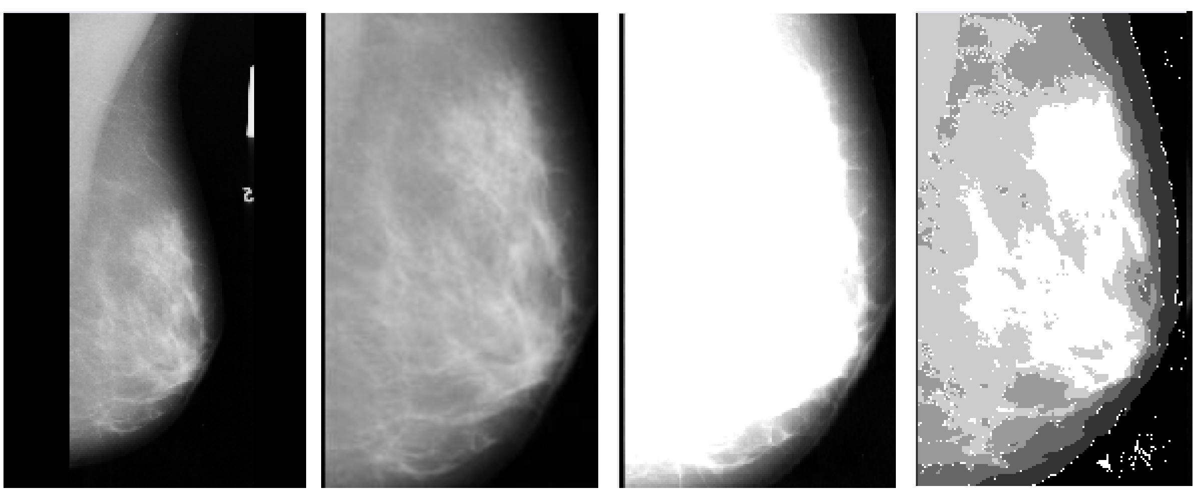

Fig. 4 (a) A mammogram sample showing mass lesion (b) Cropped ROI (c) After wavelet Transformation (d) Mass lesion feature extracted by applying OTSU threshold

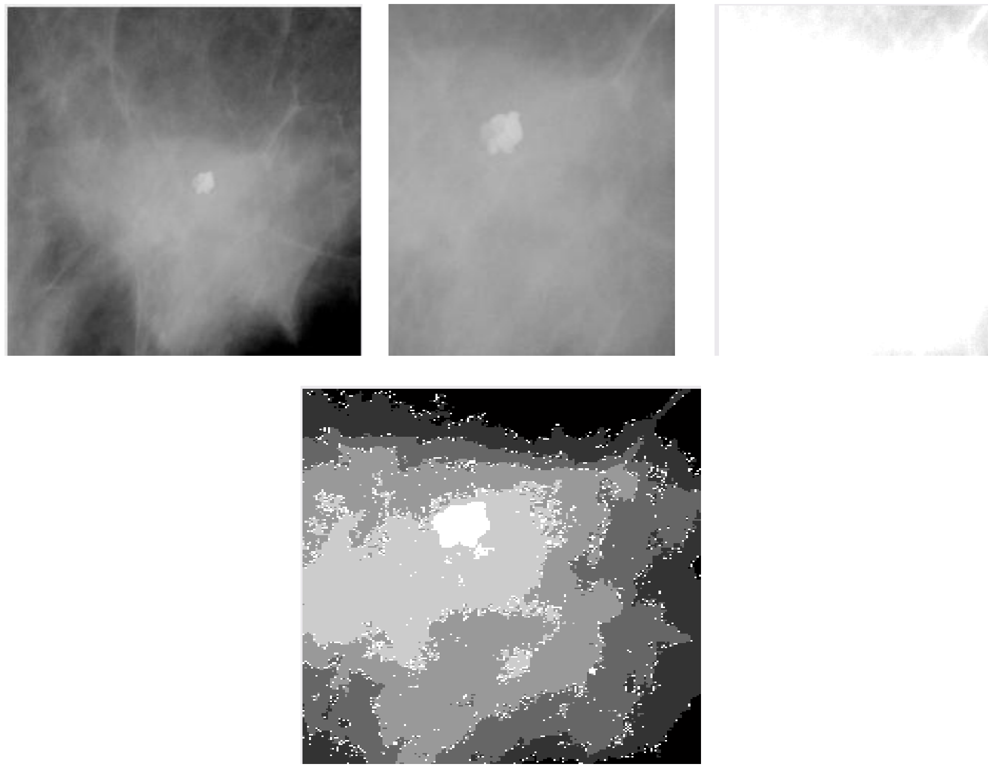

Fig. 5 (a) A mammogram sample showing Microcalcification (b) Cropped ROI (c) After wavelet Transformation (d) Microcalcification feature extracted by applying OTSU threshold

6. Conclusion

The contemporary preference for the premature detection of breast cancer in women is Mammography. Nevertheless, the elucidation of mammograms greatly depends on radiologist's opinion. In this paper, we have presented a novel approach for massive lesion and microcalcification features extraction. The approach depends on an OTSU threshold operator strategy, for the segmentation of mass/microcalcification. Wavelet transform coefficients are obtained from the selected ROIs for differentiating the features. In the proposed work we have assessed an automated detection method for one of the principal signs of breast cancer: clusters of microcalcifications and mass leisions. Experimental results illustrate that the system is capable of aiding the interpretation of radiologists in their daily practice besides enhancing their diagnostic performance. The assessment of the system was done on MIAS dataset and the experimental results demonstrate the accuracy of the system

References

[1] R.A. Smith, "Epidemiology of breast cancer in a categorical course in physics," Technical Aspects of Breast Imaging, 2nd ed., RSNA publication, Oak Book, II, pp.21, 1993.

[2] R. Peto, J. Boreham, M. Clarke, C. Davies., V. Beral, "UK and USA Breast cancer deaths down 25% in year 2000 at ages 20-69 years", THE LANCET, Volume 355, Issue 9217, Page 1822, 20 May 2000.

[3] Ranadhir Ghosh, Moumita Ghosh, John Yearwood, "A Modular Framework for Multi category feature selection in Digital mammography", In Proceedings of the 12th European Symposium On Artificial Neural Networks ESANN'2004, Bruges (Belgium), pp. 175-180, 28-30 April 2004.

[4] Bernard W. Stewart, Paul Kleihues, "WORLD CANCER REPORT", WHO, International Agency for Research on Cancer, IARC Press, Lyon 2003.

[5] Stamatia Detounis, "Computer-Aided Detection and Second Reading Utility and Implementation in a High-Volume Breast Clinic", Applied Radiology, pp: 8–15, 2004.

[6] M.K.J. Siddiqui, M. Anand, P.K. Mehrotr, R. Sarangi, N. Mathur, "Biomonitoring of Organ chlorines in Women with Benign and Malignant Breast Disease", Environmental Research, pp: 1-8, 2004.

[7] Verma, K., Zakos, J., "A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction techniques", IEEE Transactions on Information Technology in Biomedicine, vol. 16, pp. 219–223, 2000.

[8] E. L. Thurfjell, K. A. Lernevall, and A. A. Taube, "Benefit of independent double reading in a population-based mammography screening program," Radiology, vol. 191, pp. 241–244, 1994.

[9] J. G. Elmore, C. K. Wells, C. H. Lee, D. H. Howard, and A. R. Feinstein, "Variability in radiologists' interpretations of

mammograms,” The New England journal of Medicine, vol. 331, no. 22, pp. 1493–1499, 1994.
[10] Huai Li, K. J. Ray Liu, and Shih-Chung B. Lo, "Fractal Modeling and Segmentation for the Enhancement of Microcalcifications in Digital Mammograms", IEEE Transactions on Medical Imaging, vol. 16, no. 6, December 1997.
[11] M. L. Giger, “Computer-aided Diagnosis in Medical Imaging – A New Era in Image Interpretation,” Medical Imaging Ultrasound, WMA Business Briefing - Global Healthcare, pp. 75-79, 2000.
[12] Tahoces, P.G., Correa, J., Souto, M., Gómez, L., Vidal, J.J., “Computer-assisted diagnosis: The classification of mammographic breast parenchymal patterns”, Physic in Medicine and Biology, vol. 40, pp. 103-117, 1995.
[13] Yin, F. F., Giger, M. L., Doi, K., Vyborny, C. J., Schmidt, R. A., “Computerized detection of masses in digital mammograms: Automated alignment of breast images and its effect on bilateral subtraction technique,” Medical Physics, vol. 21, pp. 445 - 452, 1994.
[14] Méndez, A.J., Tahoces, P.G., Lado, M.J., Souto, M., Vidal, J.J., “Computer-aided diagnosis: Automatic detection of malignant masses in digitized mammograms,” Medical Physics, vol. 25, pp. 957-964, 1998.
[15] Veldkamp, W., Karssemeijer, N., “Accurate segmentation and contrast measurement of microcalcifications in mammograms: A phantom study”, Medical Physics, vol. 25, pp. 1102-1110, 1998.
[16] Gavrielides, M.A., Lo, J.Y., Vargas-Voracek, R., Floyd, C.E., “Segmentation of suspicious clustered microcalcifications in mammograms,” Medical Physics, vol. 27, pp. 13-22, 2000.
[17] Bassett LW, and Gambhir S, "Breast imaging for the 1990s," Seminars in oncology, vol. 18, pp. 80-86, 30 March 1991.
[18] A. Y. Cairns, I. W. Ricketts, D. Folkes, M. Nimmo, P. E. Preece, A. Thompson, and C. Walker, “The automated detection of clusters of microcalcifications,” in Proc. Inst. Elect. Eng. Colloquium on Applications of Image Processing in Mass Health Screening, pp. 3/1–5, 1982.
[19] R. M. Nishikawa, Y. Jiang, M. L. Giger, K. Doi, C. J. Vyborny, and R. A. Schmidt, “Computer-aided detection of clustered microcalcifications,” in Proc. IEEE Int. Conf. Syst., Man, Cybern., pp. 1375–1378, 1992.
[20] D. Zhao, M. Shridhar, and D. G. Daut, “Morphology on detection of calcifications in mammograms,” in Proc. IEEE Int. Conf Acoustics, Speech, and Signal Processing, San Francisco, CA, pp. 129–132, March 1992.
[21] L. Shen, R. Rangayyan, and J. Desaultels, “Detection and Classification Mammographic Calcifications”, International Journal of Pattern Recognition and Artificial Intelligence. Singapore: World Scientific, pp. 1403–1416, 1994.
[22] K. Bowyer and S. Astley, “The Art of Digital Mammographic Image”, Singapore: World Scientific, vol. 7, 1994.
[23] H. Barman, G. Granlund, and L. Haglund, “Feature extraction for computer- aided analysis of mammograms,” in State of the Art of Digital Mammographic Image Analysis. Singapore: World Scientific, 1994, vol. 7, pp. 128–147.
[24] K.Woods, C. Doss, K. Bowyer, J. Solka, C. Priebe, andW. Kegelmeyer, “Comparative evaluation of pattern recognition techniques for detection of microcalcifications in mammography,” Int. J. Pattern Recog. Artif. Intell., vol. 7, no. 6, pp. 1417–1436, 1994.
[25] R. Nishikawa, M. Giger, K. Doi, C. Vyborny, and R. Schmidt, “Computer- aided detection and diagnosis of masses and clustered microcalcifications from digital mammograms,” in State of the Art of Digital Mammographic Image Analysis. Singapore: World Scientific, vol. 7, pp. 82–102, 1994.
[26] H-D. Cheng, Y. M. Lui, and R. I. Freimanis, ``A novel approach to microcalcification detection using fuzzy logic technique", IEEE Transactions on Medical Imaging, vol. 17, no. 3, pp. 442 -450, June 1998.
[27] H.P. Chan, S.C. B. Lo, B. Sahiner, K.L. Lam, and M.A. Helvie. ``Computer -aided detection of mammographic microcalcifications: Pattern recognition with an artificial neural network,", Medical Physics, vol. 22, no. 10, pp. 1555-1567, October 1995.
[28] R. H. Nagel, R. M. Nishikawa, J. Papaioannou, and K. Doi, “Analysis of methods for reducing false positives in automated detection of clustered microcalcifications in mammograms", Medical Physics, vol. 25, no. 8, pp. 1502-1506, August 1998.
[29] G. McGarry and M. Deriche, ``Mammographic image segmentation using a tissue -mixture model and Markov random fields", IEEE International Conference on Image Processing, ICIP, 2000.
[30] L. Cordella, F. Tortorella, M. Vento, “Combining experts with different features for classifying clustered microcalcifications in mammograms”, International Conference on Pattern Recognition, vol. 4, 2000.
[31] N. Otsu, “A threshold selection method from gray-level histogram,” IEEE Transactions on System Man Cybernetics, Vol. SMC-9, No. 1, 1979, pp. 62-66
[32] I. El-Naqa, Y. Yang, M. Wernick, N. Galatsanos, R. Nishikawa, “A support vector machine approach for detection of microcalcifications”, IEEE Trans. Med. Imag. 21 (12) (2002) 1552–1563.
[33] J. Y. Lo, M. Gavrielides, M. K. Markey, and J. L. Jesneck, “Computer-Aided Classification of Breast Microcalcification Clusters: Merging of Features from Image Processing and Radiologists,” Medical Imaging 2003: Image Processing, vol. 5032, pp. 882-889, 2003
[34] H. D. Cheng, Jingli Wang, Xiangjun Shi, “Microcalcification Detection Using Fuzzy Logic and Scale Space Approaches”, Pattern Recognition, pp: 363–375, v.37, 2004.
[35] U. Bottigli, D.Cascio, F. Fauci, B. Golosio, R. Magro, G.L. Masala, P. Oliva, G. Raso, and S.Stumbo, "Massive Lesions Classification using Features based on Morphological Lesion Differences", Proceedings Of World Academy Of Science, Engineering And Technology Volume 12 March 2006 .
[36] G. L. Masala et al., "Comparative study of feature classification methods for mass lesion recognition in digitized mammograms", IL NUOVO CIMENTO, Vol. 30 C, N. 3 Maggio-Giugno 2007.
[37] Chuin-Mu Wang et al., "Classification for Breast MRI Using Support Vector Machine", IEEE 8th International

Conference on Computer and Information Technology Workshops, 2008.
[38] F. Fauci, S. Bagnasco, R. Bellotti, D. Cascio, S. C. Cheran, F. De Carlo, G. De Nunzio, M. E. Fantacci, G. Forni, A. Lauria, E.Lopez Torres, R. Magro, G. L. Masala, P.Oliva, M. Quarta, G. Raso, A. Retico, S.Tangaro, “Mammogram Segmentation by Contour Searching and Massive Lesion Classification with Neural Network”, Proc. IEEE Medical Imaging Conference, October 16-22 2004, Rome, Italy; M2-373/1-5, 2004.
[39] U. Bottigli, B. Golosio, G. L. Masala, P. Oliva, S. Stumbo, D. Cascio, F. Fauci, R. Magro, G. Raso, R. Bellotti, F. De Carlo, S. Tangaro, I. De Mitri, G. De Nunzio, M. Quarta, A. Preite Martinez, P. Cerello, S. C. Cheran, E. Lopez Torres, “Dissimilarity Application for Medical Imaging Classification”, on proceedings of The 9th World Multi-Conference on Systemics, Cybernetics and Informatics WMSCI 2005, Orlando 10-13 July 2005, vol III pag 258-262, 2005.
[40] G. Masala, B. Golosio, D. Cascio, F. Fauci, S. Tangaro, M. Quarta, S. C Cheran, E. L. Torres, “Classifiers trained on dissimilarity representation of medical pattern: a comparative study”, on Nuovo Cimento C, Vol 028, Issue 06, pp 905-912 , 2005.
[41] Tabar, L., Dean, B., “Teaching Atlas of Mammography”, 2nd edition, Thieme, New York (1985)
[42] Nagel Rufus H, N. R. M., Papaioannou John, Doi Kunio “Analysis of methods for reducing false positives in the automated detection of clustered microcalcifications in mammograms”, Med Phys, vol. 25, no. 8, pp. 1502-1506, August 1998.
[43] Imaginis, “Breast cancer diagnosis”, from http://www.imaginis.com/breasthealth/menu-diagnosis.asp, May 2008.
[44] American Cancer Society, “Cancer Facts and Figures 2003”, Atlanta, GA: American Cancer Society, 2003.
[45] Feig SA, Yaffe MJ, “Digital Mammography, Computer-Aided Diagnosis and Telemammography”, The Radiologic Clinics of North America, Breast Imaging, , vol 33, 6, 1205-1230, January 1995.
[46] Phil Evans W., “Breast Masses Appropriate Evaluation”, The Radiologic Clinics of North America, Breast Imaging, vol. 33, no. 6, pp. 1085-1108, January 1995.
[47] Suckling, J., Parker, J., Dance, D., Astley, S., Hutt, I., Boggis, C., et al., “The mammographic images analysis society digital mammogram database”, Exerpta Medical International Congress Series, vol. 1069, 1994, pp. 375–378. URL: http://peipa.essex.ac.uk/ipa/pix/mias/
[48] Norlia Md Yusof, Nor Ashidi Mat Isa and Harsa Amylia Mat Sakim, "Computer-Aided Detection and Diagnosis for Microcalcifications in Mammogram: A Review" IJCSNS International Journal of Computer Science and Network Security, VOL.7 No.6, June 2007.
[49] P. Sakellaropoulos, L. Costaridou and G. Panayiotakis, “A wavelet-based spatially adaptive method for mammographic contrast enhancement”, Phys. Med. Biol., vol. 48, pp. 787-803, 2003.
[50] L. Costaridou, P. Sakellaropoulos, S. Skiadopoulos and G. Panayiotakis, “Locally adaptive wavelet contrast enhancement”, in Medical Image Analysis Methods, L. Costaridou, Ed. Taylor & Francis Group LCC, CRC Press, Boca Raton, FL., pp. 225-270, 2005.
[51] Kaiser, G., "A friendly guide to wavelets", Boston: Birkhäuser, 1994.
[52] I. Daubechies, "Orthogonal bases of compactly supported wavelets", Comm. Pure and Appl Math., vol. XLI, pp. 909-996, 1988.
[53] I. Daubechies, "Ten Lectures on Wavelets", CBMS-NSF Regional conference series in Applied Mathematics, no. 61, SIAM, Philadelphia, PA, 1992.
[54] Ping-Sung Liao, Tse-Sheng Chen and Pau-Choo Chung, "A Fast Algorithm for Multilevel Thresholding", Journal of Information Science and Engineering, vol. 17, pp. 713-727, 2001.

Rabi Narayan Panda is currently working as Associate Professor and Additional Head of the Department of MCA dept. at Krishna Institute of Engineering and Technology, Ghaziabad, Uttar Pradesh, India. His research interest includes Data Mining, Pattern Recognition, and Medical Image Analysis.

Dr. Bijaya Ketan Panigrahi is currently working as Assistant Professor in Electrical Engineering department at Indian Institute of Technology, Delhi, India. He has received Young Scientist Award for the year 2004 given by, Orissa Bigyan Academy, Department of Science & Technology, Govt. of Orissa. He is having number of publications in Journals and International Conferences to his credit. His area of specialization is Soft computing application to Power System Planning, Operation, and Control.

Dr. Manas Ranjan Patro is currently working as Professor and Head of the Department of Dept. of Computer Science, Berhampur University, Berhampur, India. He has got his Ph.D. degree in Computer Science from the Central University, Hyderabad. His research area includes Agent based Software Engineering, Artificial Intelligence, Distributed systems, Intrusion detection system.