Digital Mammography, Cancer Screening: Factors Important for Image Compression

Laurence P. Clarke¹, G. James Blaine², Kunio Doi³, Martin J. Yaffe⁴, Faina Shtern⁵, and G. Stephen Brown⁶; NCI/NASA Working Group on Digital Mammography
Daniel L. Winfield; Research Triangle Institute, Research Triangle Park, NC 27709-2194
Maria Kallergi⁶; Department of Radiology, University of South Florida, Tampa, FL 33612-4799

Abstract. The use of digital mammography for breast cancer screening poses several novel problems such as development of digital sensors, computer assisted diagnosis (CAD) methods for image noise suppression, enhancement, and pattern recognition, compression algorithms for image storage, transmission, and remote diagnosis. X-ray digital mammography using novel direct digital detection schemes or film digitizers results in large data sets and, therefore, image compression methods will play a significant role in the image processing and analysis by CAD techniques. In view of the extensive compression required, the relative merit of "visually lossless" versus lossy methods should be determined. A brief overview is presented here of the developments of digital sensors, CAD, and compression methods currently proposed and tested for mammography. The objective of the NCI/NASA Working Group on Digital Mammography is to stimulate the interest of the image processing and compression scientific community for this medical application and identify possible dual use technologies within the NASA centers.

1. Introduction

Mammography is widely accepted today as the most effective method of screening women for breast cancer [1,2,3]. Recent studies indicate that approximately 14,000 women's lives were saved in 1992 through mammographic screening programs. Digital mammography promises significant improvements over currently used screen-film mammography. Its aims are to integrate (a) solid state sensors for digital localized or full-view breast imaging to improve image quality and acquisition process; (b) computer algorithms for image enhancement and extraction of features such as calcifications and masses to assist the radiologist and improve screening and diagnosis; (c) image transmission and storage techniques for telemammography to improve patient care and efficient use of professional expertise [4,5]. The technical challenges in digital mammography are similar to those confronted in other scientific areas, e.g. space-based sensors and space-generated information, indicating significant possibilities for cross fertilization of ideas and dual use technologies. The NCI/NASA Working Group on Digital Mammography has been...
established to explore application opportunities of such dual use technologies from aerospace to medicine [4]. In the following sections, the general requirements are presented in greater detail followed by a summary of the technologies identified until now by this Working Group within NASA Centers and Federal Laboratories as relevant to digital mammography and with a potential of having an impact to the problem.

2. Digital Mammography Systems

Conventional x-ray film-screen mammography is currently an accepted imaging modality for breast cancer detection. Nevertheless there are several technical factors limiting its performance [6]. These are the tradeoffs between contrast and exposure range inherent in the film-based system, the influence of film grain on image noise and the inefficiency of conventional methods for rejecting scattered x-rays. A direct method of mammographic image acquisition in digital form can overcome these limitations and provide improved detection of breast cancer. The target specifications of a digital mammography detector are: (1) efficient absorption of the incident radiation beam; (2) linear response over a wide range of incident radiation intensity; (3) low intrinsic noise; (4) high spatial resolution; 50 μm maximum; (5) high dynamic range; 4,000:1 minimum; (6) should accommodate at least an 18 x 24 cm field size; (7) allow an acceptable imaging time (1-7 s) and heat loading of the x-ray tube; (8) display capabilities: pixel matrices of 2k x 2k for soft copy and 4k x 4k for hard copy.

Various configurations for image acquisition have been considered such as area, point, line and multiline systems. Each approach involves compromises between factors such as spatial resolution, imaging time, readout time, detector dynamic range and sensitivity, cost, susceptibility to artifacts, efficiency of scatter reduction, and available detector size.

Photostimulable phosphors with laser readout is an approach which has been successfully developed for general radiography and it is possible to extract the information from such systems in digital form. Currently, however, this technology does not provide adequate spatial resolution for mammography [7] and, because of inefficiencies in signal collection, may suffer from excessive image noise since the detector may not be x-ray quantum limited at mammographic energies. Interesting developments in area detectors are currently underway. They include large area charge-coupled devices (CCDs) [8], silicon or amorphous silicon [9,10], amorphous selenium [11], and improvements on photostimulable phosphors. One or more of these could play a future role in digital mammography.

An approach particularly attractive for digital mammography is through area detectors. Such an approach is convenient and makes efficient use of the heat loading applied to the x-ray tube. Unfortunately, area detectors which combine adequate spatial resolution and field coverage with good signal-to-noise characteristics do not currently exist. For example, simple coupling of a large-area phosphor to a small-area photodetector via lenses requires a large minification factor (M). Because the efficiency of light transfer is approximately proportional to M^2, this is an inefficient means of imaging which causes the system not to be x-ray quantum limited.

Although areas detectors probably present the most acceptable long range solution to digital mammography, it is not clear how long it will take to overcome the technical challenges to produce a practical clinical system of this type. It is, however, currently possible to meet the specifications described above using a scanned-beam method of image acquisition [12]. The superior efficiency of scatter reduction inherent in a scanning system compared to an area detector can provide advantages in terms of image SNR/radiation dose. For scanning systems,
x-ray tube heat loading is always a concern. Scanned point and single line systems are impractical for this reason.

At the University of Toronto (UT), a slot-beam imaging system for digital mammography, shown schematically in Figure 1, is currently under development [13]. The radiation beam forms a "slot" of dimensions approximately 24 cm by 4 mm. After transmission through the breast, x-rays are incident on a fluorescent phosphor, and the emitted light enters a fiber optic assembly consisting of two fan-shaped tapers. The end of each of the 2X demagnifying tapers is ground to a 45° angle at the detector input surface where the two tapers are fused together in a smooth joint without a line that is parallel to the scan direction. The output surface of each taper is mated to a CCD array. This arrangement provides a pixel size of 50 μm referred to the midplane of the breast.

The image is acquired using time delay integration (TDI), by scanning the fan x-ray beam and the slot detector across the breast in a direction parallel to the short dimension of the detector [14]. The 45° joint of the input surface of the fiber optic tapers and the TDI acquisition solves one of the major problems associated with modular detectors, which is the presence of artifacts at their junctions. The TDI motion will average the variations in signal along detector columns due to detector structure, including that due to the joint between modules, thereby avoiding disturbing artifacts. A secondary advantage of the TDI acquisition is that both dark current and detector uniformity corrections can be made by acquiring "image" data first without x-rays and then with a uniform x-ray exposure to the detector and sorting one offset correction and one gain factor for each of the 4096 detector columns.

Current state-of-the-art does not yet provide adequate "soft-copy" display resolution (> 4096 x 4096 pixels). A high resolution CRT display (2k x 2k pixels) will be provided with the clinical system. This will allow rapid viewing of the mammogram as a complete image at reduced resolution for adjustment of display parameters or with full 50 μm resolution in a region of interest which can be positioned with a trackball over any part of the image. The image output will also be provided as a laser-printed film image (4k x 5k pixels). The radiologist will be able to manipulate the display on the monitor to define the display characteristics of the image to be printed on the film.

The clinical version of the digital mammography system is still under construction. However, some preliminary measurements have been made on a prototype. The system can acquire a mammogram in 3-6 seconds with a dose to the breast of 0.85 mGy or less. Resolution has been measured at 9.5 line-pairs/mm. The dynamic range of the CCD is 5000:1 and with digitization of 12 bits a range of over 100 in x-ray exposure transmitted by the breast can be accommodated with a "worst case" display capability of over 40 shades of gray even in the densest part of the breast, depending on the level of quantum noise.

3. Computer Assisted Diagnosis (CAD)

CAD refers to a diagnosis made by a radiologist who takes into consideration the results of a computerized analysis of radiographic images and uses them as a "second opinion" [4,5]. The goal of CAD is to reduce the screening load and improve the diagnostic accuracy by reducing the number of false negative diagnoses. Preliminary results indicate that computers can aid in recognizing abnormalities and actually point out suspicious findings such as microcalcifications and masses. Several computer techniques have been applied to mammograms including: automatic (operator and image independent) enhancement methods for outlining...
specific features such as normal parenchymal tissues, microcalcifications, and suspicious areas and pattern recognition and image segmentation methods for automatic localization, detection, and classification of suspicious breast lesions or normal parenchymal tissues.

In developing machine-assisted screening and diagnostic methods, one strives for automatic techniques which are both sensitive and specific, since the consequences of false negative interpretation (missed cancers) and false positive (FP) interpretation (traumatic, expensive investigation) are both serious. The goal of the computerized methods is to improve the performance of the radiologist by noise suppression, detail preservation, edge detection, and contrast enhancement and standardize the methods for image interpretation.

Automatic CAD schemes are the ultimate goal in mammography. Their development faces problems such as high-false positive detection rates, long CPU times, limited databases, low quality display devices. For such a workstation to be successful, one requires: high quality digital mammograms, high speed computers, large databases, efficient image processing techniques, characterization of image features of normal and abnormal patterns, understanding of image interpretation process by radiologists [4,5]. Intelligent radiologic workstations will not only retrieve, display, and process images but will also provide a wide range of tools to help us think more effectively about radiologic problems. This implies the inclusion of case-specific background information, reference images, consultations and new information from the literature.

It is anticipated that various methods with varying degrees of complexity will be required for optimum image enhancement, segmentation and pattern recognition of the mammographic features. The implementation of many of these algorithms will demand extensive computation times. Very large scale integrated (VLSI) circuits and image compression algorithms may provide a fast and cost effective technological solution to these computer vision and image processing areas. Some tasks accomplished until now are described in the following.

A. Automated Detection of Clustered Microcalcifications

At the University of Chicago (UC), a computer program is being developed to automatically locate clustered microcalcifications on mammograms [15,16,17,18]. With this method, a digital mammogram is processed by a linear filter to improve the signal-to-noise ratio of microcalcifications on the image. Gray-level thresholding techniques, which combine a global gray-level thresholding procedure and a locally adaptive gray-level thresholding procedure, are then employed to extract potential signal sites from the noise background. Subsequently, signal-extraction criteria are imposed on the potential signals to distinguish true signals from noise or artifacts. The computer then indicates locations that may contain clusters of microcalcifications on the image.

For 60 mammograms used in the study, the true-positive (TP) cluster detection accuracy of our automated detection program reached 85% at an FP detection rate of 2 clusters per image. An ROC study was performed to determine whether this performance level could result in an improvement in radiologists’ performance when the CAD results were displayed on images. The results of the ROC study, as shown in Figure 2, indicated that CAD does significantly improve radiologists’ accuracy in detecting clustered microcalcifications under conditions that simulate the rapid interpretation of screening mammograms. The results suggested also that a reduction in the computer’s false-positive rate will further improve radiologist’s diagnostic accuracy, although this improvement fell short of statistical significance.

At the Center for Engineering and Medical Image Analysis (CEMIA) at the University
of South Florida (USF), two-channel and three-channel quadrature mirror filters are developed for image decomposition and reconstruction [19,20] and dynamic neural networks are implemented for breast feature detection and extraction [21]. The sensitivity and specificity of detection is very high with these approaches. A preliminary study with 15 mammograms each containing at least one calcification cluster showed a TP rate of 100% with only 0.1 FP clusters per image; application of these methods to a larger data set is currently under way for a fuller evaluation.

B. Automated Detection of Mammographic Masses

Similarly, a computerized scheme is under development in UC for the detection of masses in digital mammograms [22,23,24]. Based on the deviation from the normal architectural symmetry of the right and left breasts, a bilateral-subtraction technique is used to enhance the conspicuity of possible masses. The scheme employs two pairs of conventional screen-film mammograms (the right and left MLO views and CC views), which are digitized. After the right and left breast images in each pair are aligned, a nonlinear bilateral-subtraction technique is employed that involves linking multiple subtracted images to locate initial candidate masses. Various feature-extraction techniques are then used to reduce false-positive detections resulting from the bilateral subtraction. In an evaluation study using 154 pairs of clinical mammograms, the scheme yielded an 85% TP rate at an average of 3 false-positive detections per image.

Alternatively, tree-structured nonlinear filters, quasi-range edge detectors, and wavelets (two-channel quadrature mirror filters) are developed in CEMIA at USF and used for enhancement and edge detection of circumscribed, irregular and stellate masses. Preliminary results on a small number of mammograms show improved performances of these algorithms for noise suppression with simultaneous image detail preservation [25].

4. Telemammography

Telemammography faces all the challenges associated with the acquisition, storage, transmission, processing and display of large amounts of data. The resolution and dynamic range currently required for digital representation of chest and musculoskeletal radiography (image size 2k x 2k to 4k x 4k pixels, pixel intensity encoded in 10 to 12 bits) stress both storage capacities and bandwidth capabilities of existing picture archiving and communications systems (PACS). Data compression studies and applications have been reported in the medical literature for over 10 years. Fidelity criteria are currently based on observer performance studies using selected case material and board certified radiologists as observers. Both lossless and lossy techniques have been offered in commercial systems [26]. Lossy techniques are generally used in cases not requiring primary interpretation from the lossy data set. Although standards are in the process of being developed (American College of Radiologists - National Electrical Manufacturers Association (ACR-NEMA) joint committee to develop a Standard for Digital Imaging and Communications in Medicine), the struggle between lossless and lossy compression techniques continues [27].

A screening mammography test consists of at least four images with each digital image ranging from 16 Mpixels to 64 Mpixels with dynamic ranges of 10 to 16 bits per pixel. Such large data sets and the fidelity requirements of mammography challenge the storage and bandwidth capabilities of existing communication systems. Cost effective storage of these images
and responsive image delivery via telecommunications channels can be facilitated using data compression technologies. Factors of 2 to 3 for lossless storage and transmission may be supported by existing encoder/decoder implementations and require no compromise in image fidelity. Significant cost savings in both research and clinical database storage is likely to result. Lossy compression approaches, offering higher gains, will need to be evaluated against observer performance for visual presentation and primary interpretation of mammographic data and evaluated for applications involving additional image processing and CAD.

Most of the results reported for lossless compression achieve a factor of 2 to 3 in compression ratio. A number of carefully constructed observer performance studies have reported successes with a "visually lossless" presentation of the data using lossy compression techniques. Block oriented discrete cosine transforms (16 x 16 to full frame) coupled with various adaptive encoding strategies produced compression results in the range of 20-to-30 :1 with no statistically significant differences in radiologist performance [28,29,30,31]. Compression ratios of 2-3 have been reported with lossless methods, e.g. tree-based codes.

Recent mammogram compression studies using wavelets also show promise with ratios up to 70:1 depending on the image [32]. Figures 3 and 4 present examples of mammogram compression with Haar wavelets at different rates. The original images (Figs. 3(a) and 4(a)) contain clustered microcalcifications and are compressed at 25:1 (Fig. 3(b)) and 50:1 rates (Fig. 4(b)). Although some image detail is lost with this type of wavelet compression, the appearance of the microcalcification clusters is not significantly affected. Furthermore, processing the original and the compressed reconstructed images with two channel wavelets results in similar segmentation of the calcification cluster from either image despite the losses during compression [25]. These results indicate that image processing of the compressed data could partially compensate for the information loss and encourage the acceptance of "visually lossless" compressed images.

5. Dual-Use Technologies

The results presented in the previous sections are representative of current preliminary work in digital mammography. It is anticipated that alternative approaches could be identified or developed which may be more successful and worthy of further study. In this context, a survey of technologies currently used in NASA Centers and Federal Research Laboratories was undertaken and has identified several projects that are promising and may have an impact in mammography. These projects span all the areas of interest and include: (a) scanning slot detectors using glass or plastic scintillating micro-fiber plates as the x-ray converting material and fiber-optic coupling to a CCD camera [33], silicon or amorphous-silicon arrays, and other advanced digital sensors for x-ray imaging; (b) software packages and algorithms such as neural networks, wavelets, and Bayesian classifiers used for target or object detection; (c) lossless and lossy compression algorithms for handling large amounts of space image data, real-time software and systems for telemetry applications; (d) storage devices and local area networks to transmit real-time voice and video traffic with simultaneous transmission of computer data; (e) VLSI circuits suitable for implementing wavelets and neural networks for pattern recognition and compression problems in real time; and (f) telerobotic developments with potential applications to stereotactic mammography procedures. The idea of technology transfer is, therefore, realistic, and is expected to receive increasingly enthusiastic response.
References


Figure 1. Schematic diagram of scanned-slot digital mammography system.

Figure 2. Comparison of ROC curves for two reading conditions (with and without computer output) for observer performance study in detecting microcalcification clusters in mammograms.
Figure 3. (a) Section of original digitized mammogram with clustered microcalcifications and (b) Haar wavelet compressed and reconstructed image at a rate of 25:1.
Figure 4. (a) Section of original digitized mammogram with clustered microcalcifications and (b) Haar wavelet compressed and reconstructed image at a rate of 50:1