

Image processing techniques to evaluate mammography screening quality

Técnicas de processamento de imagem para avaliar a qualidade de exames de mamografia

Clara Quintana^{1,2}, Germán Tirao^{1,2} and Mauro Valente^{1,2}

¹ Consejo Nacional de Investigaciones Científicas y Técnicas, Ciudad Autónoma de Buenos Aires, Argentina.

² Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba, Córdoba, Argentina.

Abstract

Mammography imaging has proved to be the best noninvasive method for breast cancer diagnosis, but it requires that irradiation parameters are set within Protocols recommendations (minimal dose delivering). This work presents an investigation on mammography image formation by means of validated Monte Carlo simulations along with further image analysis and mathematical processing. Several image processing methods have been suitably introduced and investigated according to their capability for micro-calcification detection and quality evaluation. The obtained results suggest the feasibility of all the proposed methods. Furthermore, it was possible to characterize the reliability of each and to infer the corresponding advantages or disadvantages, obtaining an image quality evaluation as a function of several parameters configurations.

Keywords: X-ray imaging, image processing, quality evaluation.

Resumo

A imagem por mamografia é comprovadamente o melhor método não-invasivo para o diagnóstico do câncer de mama, mas requer que os parâmetros de irradiação sejam estabelecidos de acordo com as recomendações de protocolos (dose mínima). Este trabalho apresenta uma investigação completa sobre a formação da imagem pela mamografia por meio de simulações Monte Carlo validadas juntamente com uma análise da imagem e o processamento matemático. Diversos métodos de processamento da imagem foram apropriadamente introduzidos e investigados de acordo com a capacidade deles em detectar microcalcificações e avaliar a qualidade. Os resultados obtidos sugerem a facilidade de todos os métodos propostos. Além disso, foi possível caracterizar a confiabilidade de cada método e deduzir as vantagens ou desvantagens correspondentes, obtendo uma avaliação de qualidade da imagem como uma função de várias configurações dos parâmetros.

Palavras-chave: imagem por raios X, processamento de imagem, avaliação da qualidade.

Introduction

Breast cancer is the most frequent cancer in women all over the world. The main established strategies for breast cancer control are based on primary prevention along with early diagnosis. In this sense, breast imaging plays an outstanding role for the screening and diagnosis of symptomatic women. Mammography units may differ by X-ray beam characteristics, breast compressing plate system, and X-ray detector¹. Breast and micro-calcification (μCa) composition material absorption properties show a strong dependence on X-ray spectrum. Furthermore, specific patient breast characteristics, lesion shape and size as well as examination exposure may influence the final image quality. This paper presents different mathematical algorithms with the aim of evaluating the mammographic

image quality, based on dedicated image processing techniques and devoted to μCa detection. In order to assess its feasibility, the integral system was applied to a wide range of clinical situations including different spectral characteristics, breast composition and thickness, μCa shape, size, and composition.

Materials and Methods

Mammography image formation

A typical mammography facility has been modeled by means of dedicated and validated Monte Carlo sub-routines², which are capable of performing absorption contrast images. This simulation toolkit, based on the PENELOPE v. 2008 main code³, has been applied to investigate the

dependence of image quality upon irradiation configurations, according to different parameters: anode and accelerating voltage, breast tissue composition and thickness, and compositions, shapes and sizes of μCa , taking into account typical situations. Fifty-four different combinations have been studied, each containing nine μCa s of different shapes and sizes, in different positions within the breast.

Image processing

Different image processing techniques along with the corresponding numerical algorithms have been developed with the aim of assessing suitable and almost automatic methods for μCa detection. The first approach consists on true-false threshold segmentation methods, which assigns 1 or 0 value to each pixel when its intensity is greater or lower than a certain preestablished threshold value. The second approach implements the Canny algorithm for edge detection, which finds edges by looking for gradient (as the Gaussian filter derivative) local maxima within the original image. The method incorporated two thresholds, devoted to detect strong (T1) and weak (T2) edges, including weak edges in the output only if they were connected to the strong ones. The third approach consists on a template matching routine, designed for searching a user predefined pattern within the original image. The matching process moves the template image to all possible positions in the source image and computes a numerical index (correlation), which indicates how well the template matches the image in that position, resulting in a correlation map with peaks located at positions of μCa .

In order to assess the quality of the image as a function of physical parameters involved in its formation and the ability to μCa detection, several mathematical processing techniques were applied to the entire set of simulated images. These techniques were based on intrinsic properties calculations and comparison with ideal binary pattern images, defined with the same geometry details (positions and sizes of μCa used in the simulation). In additions, typical techniques based on intrinsic properties were implemented for this paper, such as: Dynamic Range (DR), Signal-to-noise ratio (SNR)⁴, Contrast-to-noise ratio (CNR)⁴ and Entropy (H)⁵. The implemented comparative techniques were developed based on concepts of information theory, namely: Joint Entropy (JH)⁵, Mutual Information (MI)⁵, Normalized Cross Correlation (NCC)⁶, and index Q (Q)⁷.

The DR is the number of gray levels in which the information is distributed. It is expected that the quality of the image may preserve a direct relation with the DR. It was calculated from the differences between maximum and minimum gray levels within the image. The SNR value of image was calculated as the average SNR value of each pixel, defined as the ratio of pixel intensity to its noise standard deviation. Similarly, CNR value of image can be defined as the ratio of differences between DR and noise standard deviation. Therefore, it may be expected that both SNR and CNR values should increase as image quality gets improved.

The entropy can be seen as a measure of uncertainty, since its maximum occurs when all symbols have equal probability of occurrence⁵. Images may also be considered as carriers of information, where instead of probabilities of symbol occurrence, one has the distribution of gray values. It seems intuitive to infer that an image with a nearly uniform distribution of gray tones may have very little information. The image quality and entropy relationship depends on image noise. Since noise can be interpreted as information, if the noise is low, it is expected higher quality image with more information. The entropy can be calculated using the Eq. 1:

$$H = -\sum_i p(i) \log(p(i)) \tag{1}$$

where:

$p(i)$ is the probability distribution of intensity i .

From the entropy, joint entropy is defined as Eq. 2:

$$JH(A, B) = -\sum_{i,j} p(i, j) \log(p(i, j)) \tag{2}$$

where:

$p(i, j)$ is the joint probability defined from the joint histogram;

A represents a simulated image and

B is the ideal binary image.

Then, images of better quality should minimize the JH.

The MI can be defined in different ways and each of these definitions leads to different interpretations⁵. In this paper, the one used is as in Eq. 3:

$$MI(A, B) = H(A) + H(B) - H(A, B) \tag{3}$$

where:

A represents a simulated image and

B is the ideal binary image.

In this sense, MI can be interpreted as the amount of information that is not exclusive to any of the two images. It is remarkable that the maximization of MI is closely related to minimizing the JH.

In signal processing theory, cross-correlation (or sometimes called "cross-covariance") is a measure of similarity between two signals, often used to find relevant features in an unknown signal by comparison with another that is known. NCC has application in pattern recognition and cryptanalysis. NCC of simulated image A and the ideal binary image B were calculated as in Eq. 4:

$$NCC(A, B) = \frac{\sum_{i,j} (A(i, j) - A_{med}) (B(i, j) - B_{med})}{\sqrt{\sum_{i,j} (A(i, j) - A_{med})^2 (B(i, j) - B_{med})^2}} \tag{4}$$

where:

A_{med} (B_{med}) is the average value of the image A (B).

Wang and Bovik⁷ proposed a mathematically defined universal image quality index Q . By “universal”, the authors meant that the measured quality approach did not depend on the images being tested, the viewing conditions, and individual observers. It should be applicable to several image processing applications and provide meaningful comparisons across different types of image distortions. The index Q was calculated using the Eq. 5:

$$Q = \frac{4\sigma_{AB}A_{med}B_{med}}{(\sigma_A^2 + \sigma_B^2)(A_{med}^2 + B_{med}^2)} \quad (5)$$

where:

$$\sigma_X^2 = \frac{1}{N-1} \sum_{i,j} (X(i,j) - X_{med})^2 \quad \text{and } X = A \text{ or } B,$$

$$\sigma_{AB} = \frac{1}{N-1} \sum_{i,j} (A(i,j) - A_{med})(B(i,j) - B_{med})$$

The dynamic range of Q is $[-1,1]$. The value of 1 is achieved when the original image and test image are equal and the worst value when the test image is twice the mean of original image subtracted by the original image.

Results and Discussion

Due to the fact that the considered breasts have different absorption paths associated with their ellipsoidal shapes, the threshold segmentation method could not provide a satisfactory performance because of the nonuniform background. Therefore, its suitability is strongly dependent on the image position, as indicated in Figure 1. It was found that the performance of this method showed significant dependence upon the threshold value. However, it should be mentioned that in some cases this drawback may overcome if a suitable background (BG) is first subtracted from the original image.

The subtracted BG was calculated from the original image following these steps: suitable smoothing process is applied to original image, for each point in the longitudinal central axis (around which μ Cas were positioned, as shown in Figure 1). It has been calculated the average along the transverse profile obtaining therefore a suitable BG for each point on the longitudinal central axis, which has been used for μ Ca detection by means of the 1D algorithms. A whole 2D BG can be directly and straightforwardly calculated generalizing the proposed method to 2D dimensions by means of considering repeating the mechanism for all non-central longitudinal axes (image rows).

The implemented edge detection method is required to assess suitable combination of the user-defined threshold

values, in order to achieve a good performance. As expected, greater values for the T2 threshold parameter (focused on weak edges detection) resulted in best performances for detecting image details, but becoming also greater the possibility of “false true” due to the detection of noise as image detail. Therefore, it may be advisable to overcome this risk by means of considering high T1 threshold values (dedicated to strong edges detection), as reported in Figures 2a and 2b. However, for lower T2 values, it may be advisable to employ also lower T1 in order to be able to detect μ Ca edges, as reported in Figures 2c and 2d.

The template matching approach has been applied, obtaining a very good performance. In fact, this was the method that best worked, detecting all μ Ca (9 of 9), while the Canny algorithm only detected 8 μ Ca. The true-false threshold segmentation method detected 7 μ Ca, but with

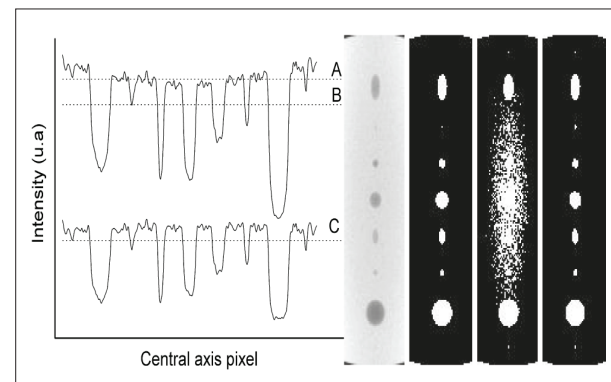


Figure 1. μ Ca detection using image segmentation algorithms. From left to right – Central axis profiles: original (top curve) and BG subtracted (bottom curve) image; Original image, segmented image (threshold B), segmented image (threshold A), BG subtracted image segmented (threshold C). Image parameters: glandular breast tissue with 30 mm thick and 40 kV incident spectrum, calcium oxalate μ Ca composition.

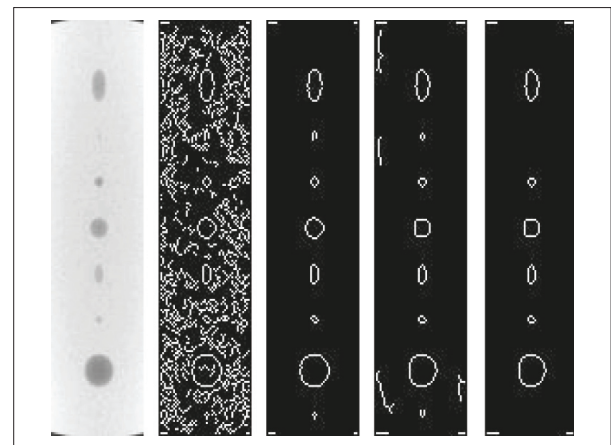


Figure 2. μ Ca detection using edge filter algorithms with different tolerance values. From left to right: a: Original image; b: T1=0.05, T2=0.1; c: T1=0.1, T2=0.1; d: T1=0.05, T2=0.8; e: T1=0.1, T2=0.8. Mammographic parameters: glandular breast tissue with 30 mm thick for 40 kV incident spectrums, calcium oxalate μ Ca composition.

a suitable BG subtraction it can detect all μCa . Another important advantage of this method is that the detection efficiency depends sensitively on the details size of the template, and not its form. This fact is clearly shown in Figure 3, where correlation maps for different templates sizes can be observed, and also how a large template distinguishes more noticeable large (small) μCa .

Regarding mathematical processing techniques to evaluate image quality and μCa detection, which are both the intrinsic properties calculations as parameters for comparison, it was noticed that almost all the studied parameters, except for JH, showed the expected tendencies with respect to the physical parameters involved in the image formation: accelerating voltage, breast tissue composition and thickness, and μCa composition. The JH parameter did not reflect the expected behavior, inferred according to image formation processes and physical parameters dependence. This fact may be explained because very different images were compared. Despite the functional relationship of MI and JH, MI showed the correct behav-

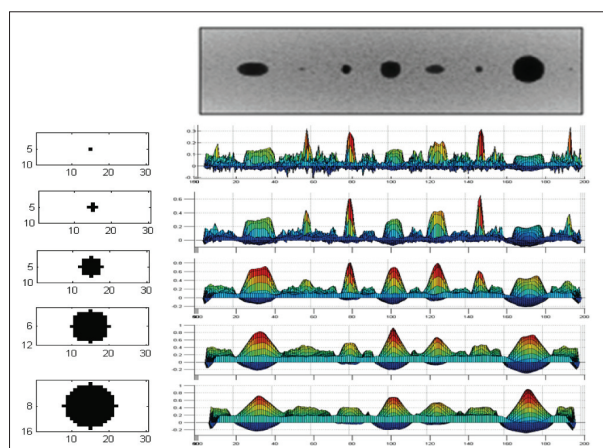


Figure 3. Correlation profiles for different templates sizes.

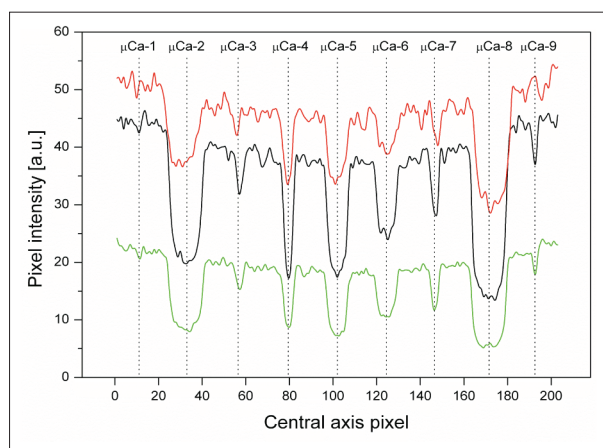


Figure 4. Central axis profiles for simulated images. The black, red and green lines correspond to 30, 70 and 30 mm thickness of glandular breast tissue, and 34, 40 and 24 kV incident spectrums, respectively. μCa composition was calcium oxalate.

iors because JH is several orders of magnitude (about 3) smaller than H.

Besides, the studied parameters were very useful to establish a ranking of the images and thus to assess a quality criterion. The ranking is set taking into account the relationship of the parameter with the image quality, *i.e.*, a maximum, minimum, or an appropriated value. In this sense, the DR, CNR, H, and Q establish almost the same ranking. These parameters measure the overall image contrast compared to the other studied parameters. The obtained ranking does not reflect the feasibility of μCa detection and therefore the image quality. For example, Figure 4 shows the central axis profiles for three different simulated images, the black, red and green lines correspond to 30, 70 and 30 mm thickness of glandular breast tissue, and 34, 40 and 24 kV incident spectrums, respectively. For these cases, μCa composition was calcium oxalate. In Figure 4, the μCa -9 was detected only for the black and green axis profiles. Then, the image corresponding to the red line presents a lower quality than the one corresponding to the green line, and therefore should be in a worse position in the ranking. But the obtained ranking with this parameter was sixth, seventh, and eight for the black, red and green lines, respectively.

However, the ranking obtained using SNR values was very similar to those obtained from DR, CNR, H, and Q presenting only some minor differences. This parameter weights the average value of intensity, resulting in a non-suitable ranking, which does not reflect the ability to μCa detection. As example, it can be seen from Figure 4 that central axis profiles for a 24 kV incident spectrum and 30 mm thickness of glandular breast tissue (green line) should not be better ranked than the image corresponding to red line. However, the obtained ranking predicts the opposite. The ranking using SNR values for these three examples was fifth, sixth and eight for the red, black and green lines, respectively.

Finally, the NCC parameter created a classification quite different from those presented earlier, but much more representative of the image quality. For the same examples used in Figure 4, the corresponding ranking of this parameter was fifth, sixth and tenth for the black, green and red lines, corresponding to 30, 70 and 30 mm thickness of glandular breast tissue, and 34, 40 and 24 kV incident spectrums, respectively. It can be observed that although the red line has greater contrast than the others, it fails to detect the μCa -9. This μCa is clearly detected by the other profiles. The image quality defined by these parameters represents the feasibility of μCa detection.

Conclusions

Different mammography image processing techniques have been proposed and investigated. A suitable simulation toolkit has been satisfactory implemented for this investigation. Image processing techniques reliability and suitability

for automatic detail detection have been carefully studied. Image segmentation and edge filtering approaches showed good performance for μCa detection. The template matching approach and its correlation values were excellent and helpful tools for μCa detection, and they could also be used as a quality parameter for different irradiation set-ups. These algorithms proved to be a valuable and promising tool for mammography images processing. The DR, CNR, H, and Q algorithms establish almost the same ranking. Such parameters measure the overall image contrast. But, SNR values weights average value of intensity, resulting in a nonsuitable ranking, which does not reflect the ability to μCa detection.

Finally, NCC was the only one capable of satisfactory reflecting image quality as well as providing feasible μCa detection. NCC makes few requirements on the image sequence and has no parameters to be searched by the user.

Acknowledgment

This paper has been partially supported by grants from research Projects PIP 11420090100398, PICT 2008-243 along SeCyT from *Consejo Nacional de Investigaciones Científicas y Técnicas* (CONICET), *Agencia Nacional*

de Promoción Científica y Tecnológica (ANPCyT) and *Universidad Nacional de Córdoba* (UNC) of Argentina.

References

1. Berns EA, Hendrick R, Solari M, Barke L, Reddy D, Wolfman J, et al. Digital and screen-film mammography: comparison of image acquisition and interpretation times. *Am. J. Roentgenol.* 2006;187(1):38-41.
2. Tiraio G, Quintana C, Malano F, Valente M. X-ray spectra by means of Monte Carlo simulations for imaging applications. *X Ray Spectrom.* 2010;39(6):376-83. DOI: 10.1002/xrs.1279.
3. Salvat F, Fernández-Varea J, Sempau J. PENELOPE – Version 2008, NEA: France; 2008.
4. Bushberg JT, Seibert JA, Leidholdt Jr EM, Boone JM. The essential physics of medical imaging. California: Lippincott Williams & Wilkins; 2001.
5. Pluim JPW, Maintz JBA, Viergever MA. Mutual-information-based registration of medical images: a survey. *IEEE Transactions on Medical Imaging.* 2003;22(8):986-1004.
6. Lewis JP. Fast Template Matching, *Vision Interface 95*, Canadian Image Processing and Pattern Recognition Society, Quebec, Canada, 1995; p. 120-3.
7. Wang Z, Bovik AC. A Universal Image Quality Index, *IEEE Signal processing Letters.* 2002;9:81-4.

