AUTOMATIC EVALUATION OF QUALITY PARAMETERS IN DIGITAL MAMMOGRAPHY IMAGES USING THE PHANTOM CDMAM

M. A. Z. Sousa*, P. N. Siqueira*, R. B. Medeiros**, H. Schiabel*

*Department of Electrical Engineering – EESC/USP, S.Carlos/SP
**Department of Imaging Diagnosis – UNIFESP, S.Paulo/SP – Brasil

e-mail: angelicazucareli@usp.br

Abstract: Technical requirements of image quality in mammography established by several normative documents include quality parameters, which can be achieved by conducting periodic tests. It is recommended that some quality parameters be measured from images acquired by exposing specific phantoms, as the Contrast-Detail Mammography - CDMAM, in such systems. Accordingly, the purpose of this work is the development of a software to assist in the professional testing, allowing to ascertain image quality parameters: contrast-detail curve, correct observation ratio; image quality figure; and figure-of-merit. For this, we used 46 images obtained for five Computed Radiography (CR) systems and a method of detection based on circular correlators filters. The classification of the image discs between visible or not visible was made from a data-mining tool promoting the construction of decision tree models. The result was a system to aid the specialist, reinforcing the integrity of the assessment and reaching accuracies of up to 95%.
Keywords: Digital mammography, quality assurance in mammography, phantom CDMAM, quality parameters.

Introduction

An effective quality control system for digital mammography needs to evaluate the status of each stage of image formation – acquisition, display and processing, storage and archiving (Kanal et al, 2013).

According to Perez-Ponce (Perez-Ponce et al, 2013), image quality assessment approaches can be accomplished in two ways: based on the measurement of technical parameters related to the performance of digital detectors; and based on multiple-forced choice experiments on phantom images.

In the phantom tests, parameters that can determine the distinction between the signal of interest and the background need to be checked. These parameters include the high contrast details and the low contrast threshold, for instance (Jakubiak et al, 2013).

The European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis (Perry et al, 2006) recommends carrying out tests with CDMAM phantom (Bijker et al, 2000) which should be done annually by 3 observers who read two images verifying the discs location by means of a template in order to determine the contrast threshold.

The CDMAM phantom consists of an Aluminum plate inside acrylic, composing a matrix. Four other acrylic plates simulate the breast thickness. In each matrix cell, two identical gold discs are randomly disposed. They are between 0.03µm and 2.0µm thick and their diameters diverge from 0.06mm up to 2.0mm. Figure 1 illustrates an image of phantom CDMAM 3.4.

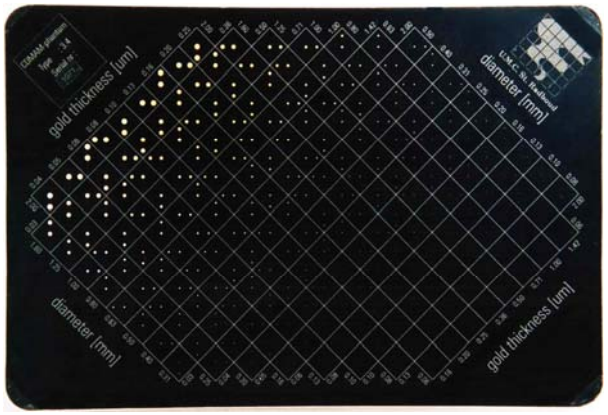

Figure 1. Frontal view of phantom CDMAM 3.4 – Artinis Contrast-Detail Phantom.

However, the phantom image reading is considered tiresome and time consuming. In addition, this reading is quite dependent on the reader subjectivity, which can cause errors, mainly among different observers. In order to minimize this subjectivity, computer systems have been developed to automate the process of reading the images (Thijssen & Karssemeijer, 1996; Figl et al, 2011; Monnin et al, 2011). Moreover, efforts have been applied in order to generate ideal detail-contrast curves based on comparison of image quality parameters (Thomas et al, 2005).

Therefore, this study has attempted to develop a computational tool capable of analyzing the phantom image, associating also the possibility of studying four parameters of image quality: contrast detail curve (CDC); correct observation ratio (COR); image quality figure (IQF) and figure-of-merit (FOM).

Methodology

The task of automating the reading of CDMAM phantom images involved the development of a

computational tool that detects the disks present in the image according to its diameters and thicknesses.

For this, correlation filters were used (Gonzalez & Woods, 2002) composed by two concentric regions, one external and other internal. The internal region comprises the disc and the average of pixel values inside the circle is calculated. The external is a ring that provides the calculation of the average of pixel values in a region embracing the image background. Figure 2 shows an example of filters used.

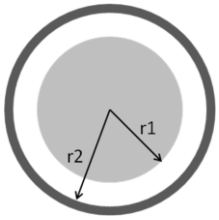

Figure 2. Filter model used for discs detection. The r1 and r2 radii vary with the radius of the disc analyzed, according to a reference image.

After discs detection, the classification step was started to define the discs as being visible or not according to the human vision.

Whereas there is much subjectivity in human visual analysis because the ability to identify the visibility of objects in an image is dependent on the lighting conditions and the visual system that is conducting the analysis, it is valid to select the image features that best represent it. Thus, it was decided to employ in this work an automatic technique for selection of the most relevant features for a classification method that uses decision trees to analyze the input data.

The software Waikato Environment for Knowledge Analysis - WEKA (Witten & Frank, 2000) was chosen. It has a free package of algorithms for mining and classification of data, widely used to pattern recognition (Cheng et al, 2010).

The characteristics used to create the classification models were predefined considering: the pixels values of the image $p_i(i,j)$; the pixel value of each disk on the image, $p_e(i,j)$, obtained from the internal region of the circular filter (Fig. 2); and the gray levels of the background image, $p_b(i,j)$, corresponding to the outer region of the filter. In addition, a *w*x*h* cropping around each disk was used to calculate the other attributes extracted for training the system. All this process was developed in MATLAB R2010b. Table 1 contains a description of each attribute used, where *(i, j)* is the position of the filter on the image.

A total of 2542 discs were extracted from 11 images selected exclusively for training. Half of them had been classified by specialists as visible and the other half as not visible in order to produce reliable models of learning.

Upon features extraction, files were created in standard format *.ARFF, allowing the insertion of the input data from each disk in the WEKA software to produce the classification models.

The training system involved the creation of decision trees for each disk diameter of low contrast using the J48 algorithm from the WEKA package (Salzberg, 1994).

Table 1. Attributes extracted from the image after detection of structures of interest (discs).

Attribute	Equation
Average pixel values of the disc	$\mu_e = \sum_{i=1}^{w}\sum_{j=1}^{h}\frac{p_e(i,j)}{(w\times h)}$
Average pixel values of the background	$\mu_b = \sum_{i=1}^{w}\sum_{j=1}^{h}\frac{p_b(i,j)}{(w\times h)}$
Difference of average pixel values of the disc and background	$\Delta\mu = \mu_e - \mu_b$
Weber ratio	$W = \frac{\Delta\mu}{\mu_e}$
Average pixel values of the image	$\mu_i = \sum_{i=1}^{w}\sum_{j=1}^{h}\frac{p_i(i,j)}{(w\times h)}$
Average pixel values of the equalized image	$\mu_{eq} = \sum_{i=1}^{w}\sum_{j=1}^{h}\frac{p_{ie}(i,j)}{(w\times h)}$
Variance of the image	$v = \sum_{i=1}^{w}\sum_{j=1}^{h}\frac{(p_i(i,j)-\mu_i)^2}{(w\times h)}$
Mode of the image	*Gray level more frequent in p(i,j)*

Then, from the results of classification four parameters of image quality (Thomas et al, 2005) were calculated by the developed program:

Contrast-detail curve (CDC): a graphic correlation between minimal correct reading diameter and disk thickness. Approximately linear in a two-log scale;

Correct observation ratio (COR): the ratio of the total number of correctly identified objects (N_i) to the total number of objects in the whole phantom (N_r) multiplied by 100, as in the equation:

$$COR = \frac{N_i}{N_r}.100; \quad (1)$$

Image quality figure (IQF): the sum of the products of the diameters for each of the smallest scored objects and their corresponding contrast, related by the equation:

$$IQF = \sum_{i=1}^{n^o cols} C_i.D_{i,mim} \quad (2)$$

where $D_{i,mim}$ is the minimum diameter of the disks for the contrast *i* and C_i is the thickness of the discs;

Figure-of-merit (FOM) (Borg et al, 2012): Diameter for which the contrast-detail curve crosses the threshold thickness axes. Obtained from a linear extrapolation of the contrast detail curve through the origin.

The final tests were conducted with 46 CDMAM images from five CR units: Agfa 75, Agfa 85, Fuji 50, Fuji 100 and Kodak 975. Parameters as kVp and mAs were set according to the European Protocol.

Results

The results allowed gauging the images quality parameters. They were analyzed statistically according to the rate of right detections for each diameter examined.

Figure 3 illustrates the contrast-detail curves for two processed images compared with its respective reference curves obtained by the average of the reading from five expert observers.

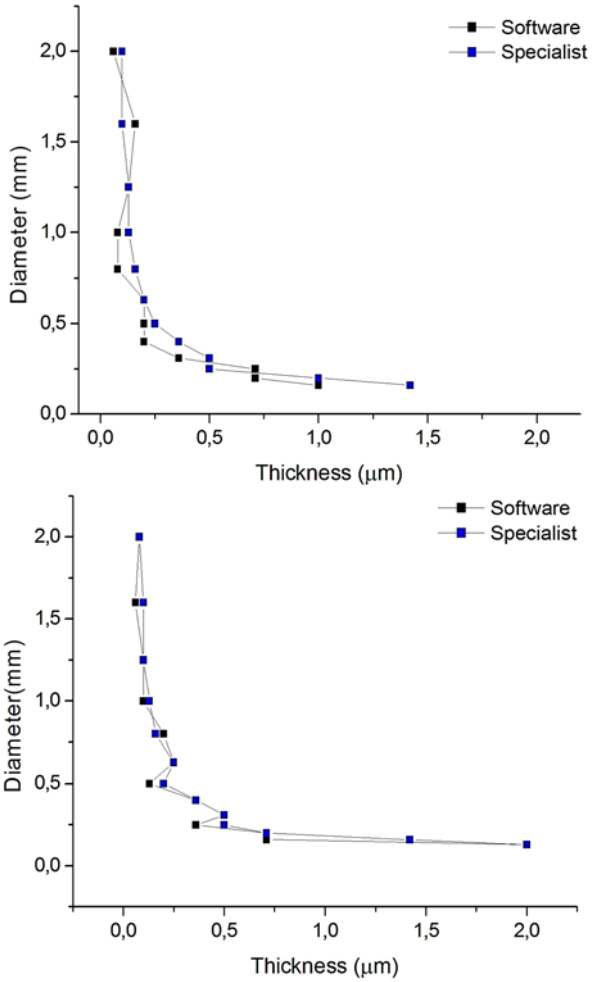

Figure 3. Examples of contrast-detail curves for two images obtained by the developed software in comparison with its respective reference curves (specialists' reports).

Knowing that the diameters of the discs decrease along the phantom lines while the contrast increases logarithmically in accordance with the thickness, the thresholds discs may be considered those with intermediaries' diameter and thickness, causing confusion even in human visual reading. This fact is reproduced by the system, since classifying accurately the disks located on the threshold of visibility is hard, which generates a superior percentual error relative to the specialists' reports, as shown in Table 2.

Table 2. Average of percentual errors of contrast-detail curves and the accuracy rate in relation to reports.

Diameter (mm)	Percentual error	Accuracy rate (%)
2.00	16.99	83.01
1.60	8.02	91.98
1.25	7.35	92.65
1.00	7.69	92.31
0.80	28.57	71.43
0.63	20.39	79.61
0.50	10.66	89.34
0.40	10.29	89.71
0.31	9.93	90.07
0.25	6.67	93.33
0.20	5.46	94.54
0.16	12.67	87.33
0.13	13.73	86.27

Considering the other parameters of image quality gauged by the developed program (COR, IQF and FOM), we could notice that the best results point out lower values for COR and higher for IQF, not allowing accurate conclusions about FOM. The data corresponding to Table 3 exemplify the values reached for five different images.

Table 3. Parameters of image quality collected by the developed software and accuracy rate for five tested images.

Image	COR	IQF	FOM	Accuracy rate (%)
1	282.86	0.16	14.69	0.91
2	260.53	0.09	66.35	0.83
3	319.35	0.16	1.12	0.86
4	257.14	0.26	-68.27	0.95
5	282.86	0.11	-135.78	0.91

These results confirm the nonlinear relation with parameters shown in the Thomas work (Thomas et al, 2005) as well as highlight the need of caution when analyzing the data from different images from the CDMAM phantom. We should also consider the relation between the COR and the accuracy rate that achieved identical values for most of the images.

The time to execute the total processing is less than 1 minute by image using a conventional Windows environment without process control. This makes the system quite advantageous since the visual analysis can take up to 15 minutes to be performed, as we have already experienced.

Conclusions

The methodology employed allowed to create a computational tool to aid the specialist in preparing the final report on the quality of the image generated from known parameters.

The classification models allowed achieving accuracy rates of up to 95% in the implementation of the decision tree algorithm and reaching contrast-detail curves very close to those obtained by specialists' reports.

Regarding the other parameters of image quality gauged by the developed program (IQF, COR and FOM), the hypothesis already mentioned by other authors (as Thomas et al, 2005, for example) could be confirmed, that is, the use of different parameters of CDMAM image quality can potentially lead to different conclusions about the image quality. However, the remarkable correlation between the COR and the rate of right answers for each image denotes a convergent route for future studies.

Acknowledgements

To FAPESP and CAPES due to the financial support and to the team from Laboratorio de Qualificação de Imagens Médicas (QualIM) – UNIFESP by acquiring and providing the images used in our tests.

References

[1] Bijker KR, Thijssen MAO, Arnoldussen, ThJM, 2000. Manual of CDMAM-phantom type 3.4, University Medical Centre Nijmegen. St, Radboud Department of Diagnostic Radiology, Section of Physics and Computer Science.

[2] Borg M, Badr I, Royle GJ, 2012. The use of a figure-of-merit (FOM) for optimisation in digital mammography: a literature review. Radiation Protection Dosimetry. v. 151, n. 1, pp. 81–88.

[3] Cheng HD, Shan J, Ju W, Guo Y, Zhang L, 2010. Automated breast cancer detection and classification using ultrasound images: A survey, Pattern Recognition, v. 43, pp. 299-317.

[4] Figl M, Hoffmann R, Kaar M, Semturs F, Brasik N, Birkfellner W, Homolka P, Hummel J, 2011. Factors for conversion between human and automatic read-outs of CDMAM images. Medical Physics. v. 38, n. 9, pp. 5090-5093.

[5] Jakubiak RR, Gamba HR, Neves EB, Peixoto JE, 2013. Image quality, threshold contrast and mean glandular dose in CR mammography. Physics in Medicine and Biology. v. 58, pp. 6565–6583.

[6] Kanal KM, Krupinski E, Berns EA et al., 2013. ACR–AAPM–SIIM Practice Guideline for Determinants of Image Quality in Digital Mammography. Journal of Digital Imaging. v. 26(1), pp. 10-25.

[7] Karssemeijer N, Thijssen MAO, 1996. Determination of contrast-detail curves of mammography systems by automated image analysis. In DIGITAL MAMMOGRAPHY'96, Proceedings of the 3rd International Workshop on Digital Mammography, Chicago: 1996. pp. 155-160.

[8] Monnin P, Marshall NW, Bosmans H, Bochud FO, Verdun F.R, 2011. Image quality assessment in digital mammography: part II, NPWE as a validated alternative for contrast detail analysis. Physics in Medicine and Biology. v. 56, pp. 4221–4238.

[9] Perry N, Broeders M, Wolf CDe, Törnberg S, Holland R, Von Karsa L, 2006. European Guidelines for Quality Assurance in Mammography Screening and Diagnosis. 4th edition, Luxembourg: Office for Official Publications of the European Communities.

[10] Perez-Ponce H, Daul C, Wolf D, Noel A, 2013. Validation of a digital mammographic unit model for an objective and highly automated clinical image quality assessment. Medical Engineering & Physics. v. 35, pp. 1089-1096.

[11] Salzberg SL, 1994. Book Review: C4,5: Programs for Machine Learning by J, Ross Quinlan. Journal Machine Learning, Morgan Kaufmann Publishers, Inc.

[12] Thomas JA, Chakrabarti K, Kaczmarek R, Romanyukha A, 2005. Contrast-detail phantom scoring methodology. Medical Physics. American Association of Physicists in Medicine.

[13] Witten I and Frank E, 2000. Data Mining – Practical Machine Learning Tools. 3rd edition, Morgan Kaufmann.