The Effect of Breast Composition on a No-reference Anisotropic Quality Index for Digital Mammography

Bruno Barufaldi[1,2(✉)], Lucas R. Borges[1,2], Marcelo A.C. Vieira[1], Salvador Gabarda[3], Andrew D.A. Maidment[2], Predrag R. Bakic[2], David D. Pokrajac[4], and Homero Schiabel[1]

[1] Department of Electrical and Computer Engineering, University of São Paulo, São Carlos, Brazil
Bruno.Barufaldi@uphs.upenn.edu
[2] Department of Radiology, University of Pennsylvania, Philadelphia, USA
[3] Spanish Council for Scientific Research, Institute of Optics, Madrid, Spain
[4] Department of Information and Computer Sciences, Delaware State University, Dover, USA

Abstract. There are several methods to evaluate objectively the quality of a digital image. For digital mammography, objective quality assessment must be performed without references. In a previous study, the authors investigated the use of a normalized anisotropic quality index (NAQI) to assess mammography images blindly in terms of noise and spatial resolution. Since the NAQI is used as a quality metric, it must not be highly dependent on the breast anatomy. Thus, in this work, we analyze the NAQI behavior with different breast anatomies. A computerized system was used to synthesize 2,880 anthropomorphic breast phantom images with a realistic range of anatomical variations. The results show that NAQI is only marginally dependent on breast anatomy when images are acquired without degradation (<12 %). However, for realizations that simulate the acquisition process in digital mammography, the NAQI is more sensitive (33 %) to variations arising from quantum noise. Thus, NAQI can be used in clinical practice to assess mammographic image quality.

Keywords: Breast anatomy · Image quality index · Anisotropy · Digital mammography

1 Introduction

The evaluation of mammography image quality is important in clinical practice to assure optimal performance of the radiologists. To execute this task, a large dataset of clinical mammograms must be assessed by a group of radiologists through standardized subjective methods [1]. These approaches are expensive, time-consuming and influenced by inter-observer subjectivity. An alternative is the use of objective image quality assessment, where the goal is to provide computational models that can automatically predict perceptual image quality [2].

A. Tingberg et al. (Eds.): IWDM 2016, LNCS 9699, pp. 226–233, 2016.
DOI: 10.1007/978-3-319-41546-8_30

There are several methods to evaluate the quality of a digital image objectively. Generally, such methods calculate the similarity between the degraded image and a reference image, which is assumed to have perfect quality (ground-truth) [2]. However, these methods cannot be applied to digital mammography because the ideal image without degradation is not available in clinical practice. Thus, for digital mammography, objective quality assessment must be performed without any reference. No-reference or "blind" image quality assessment is an extremely difficult task, as they have to predict the perceptual quality of distorted images with no information about the reference images [2].

In a previous paper [3], we investigated a no-reference image quality index to assess mammographic images acquired within a range of radiation doses (noise) and detector sizes (blur). The results reported by this index, the normalized anisotropic quality index (NAQI), followed the same behavior as other well-established full-reference indexes, such as the *peak signal-to-noise ratio* (PSNR) and *structural similarity index* (SSIM), when evaluating digital mammograms acquired with anthropomorphic breast phantoms.

However, the NAQI is normalized by the entropy of the image. The anatomical features of the breast exert some influence on this normalization. Since the anatomy varies among patients, it would be an undesirable feature if the NAQI were more sensitive to the variations on breast anatomy then to image quality. The purpose of the current study is to analyze the sensitivity of the NAQI to a range of breast anatomies and image qualities, using a large set of synthetic mammograms generated by an anthropomorphic breast software phantom [4, 5] with a realistic range of anatomical variations.

2 Method

The anisotropic quality index [6] is calculated using the generalized Rényi entropy (R_α) and the normalized pseudo-Wigner distribution (W_z)

$$R_\alpha = \frac{1}{1-\alpha}\log_2\left(\sum_n \sum_k P^\alpha[n,k]\right) \tag{1}$$

$$W_z[n',k'] = 2\sum_{m=-N/2}^{(N/2)-1} Z[n'+m]Z^*[n'-m]e^{-2i(2\pi m/N)k'} \tag{2}$$

where *P[n,k]* is the discrete space-frequency distribution, n and k are spatial and frequency components, and α is a constant for the space-frequency distributions; n' and k' represent the spatial and frequency discrete variables, m is a parameter used for shifting and z^* is the complex conjugate of z. To consider the entropy directions, *z[n']* is a 1D sequence of N gray values of pixels.

Based on Eqs. 1 and 2, the Rényi entropy for a pixel located at the position n, described at the pseudo-Wigner domain, is given by

$$R_3[n] = -\frac{1}{2}\log_2\left(\sum_{k=1}^{N} \check{P}_n^3[k]\right); \quad (3)$$

where the probability distribution P_n has been obtained from the normalization of coefficients W_z.

In considering the statistical dispersion of entropy, spatial resolution and noise have an inverse correlation with anisotropy. To overcome this problem, the proposed metric uses the anisotropy with directional dependency, since this measurement decreases in accordance with the amount of degradation.

The proposed metric, based on result of Eq. 3, is given by:

$$\bar{R}[\theta_s] = \sum\nolimits_n R_3[n, \theta_s] \Big/ M; \quad (4)$$

where M is the image size and $\theta_s \in [\theta_1, \theta_2, \ldots, \theta_S]$ represents the S different orientations that are considered to measure the entropy.

3 Materials

All images used in this work were generated using an anthropomorphic breast software phantom developed by the University of Pennsylvania [4, 5]. The tissue simulation is controlled by user-selected parameters, that can cover anatomical variations seen clinically. All phantoms had a voxel size of $(0.1\ \text{mm})^3$, a simulated breast volume of 1500 mL, and a compressed breast thickness of 6.4 cm. To test the robustness of the image quality measure, in this study we specifically varied:

- Thickness of Cooper's ligaments: 0.04 mm and 0.06 mm;
- Number of tissue compartments: 167, 333, 500, and 1,000;
- Percent of compartments labeled dense: 0 %, 5 %, and 10 %, corresponding to overall glandularity of 15 %, 21 %, and 26 %, respectively;
- Size and shape of compartments: three sets of parameters selected;
- Skin thickness: 1.2 mm and 1.5 mm.

The combination of the anatomical features mentioned above generated 144 different phantoms, which were then simulated 10 times using different position seeds, resulting in the equivalent of 1,440 patients.

Mammographic projections were synthesized based upon the simulation of a clinical mammography Selenia Dimensions system (Hologic Inc., Bedford, MA), with 70 μm pixel size. The x-ray acquisition parameters were chosen assuming automatic exposure control (AEC) and also half of the recommended mAs to simulate 50 % dose reduction on the radiation dose. Therefore, a total of 2,880 FFDM were generated. Figure 1 shows examples of simulated images.

A subset of the anatomical configurations was chosen, fixing the skin thickness to 1.2 mm and the size and shape of compartments to just one value. A total of 240 images were then generated considering an ideal acquisition system with no noise. This image set was used to study the impact of breast anatomy individually, without any influence of the image quality.

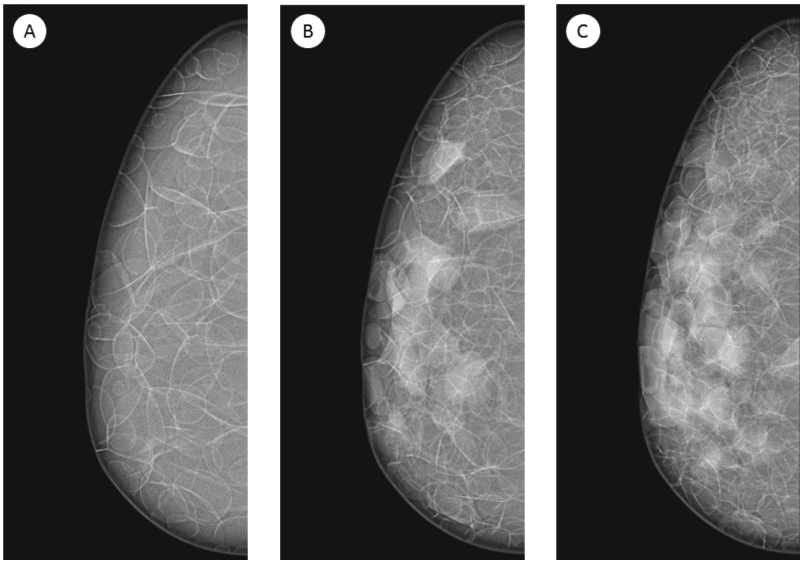

Fig. 1. Examples of three digital phantoms that simulate the skin thickness of 0.12 mm and 0.04 mm of Copper's ligaments. (A) Breast density overall of 15 % and number of compartments of 167; (B) 21 % and 500; and (C) 26 % and 1,000.

All experiments were performed using a ROI size of 2,791 × 709 pixels, containing as much of the inner breast tissue as possible to avoid background bias. The window size (w) and number of directions of the NAQI algorithm were set to 16 pixels and 5 angles ($\Theta = [0°, 45°, 90°, 135°, 180°]$), respectively.

4 Results

First, the NAQI was evaluated over a dataset of 240 ground-truth images, generated considering an ideal acquisition system with no noise. Second, all the 144 different breast anatomies, with 10 random realizations were evaluated at two different radiation dose levels: the one given by the AEC and 50 % of it. The total number of images considered in the second subsection is 2,880.

4.1 Impact of Breast Anatomy

Table 1 presents the NAQI reported for the phantom simulations, arranged by four classes of compartments (167, 333, 500 and 1000) and two groups of ligament thickness (0.04 mm and 0.06 mm).

Note that the NAQI is reduced as the number of compartments and the percent of compartment's density are increased. On the other hand, NAQI increases with the thickness of Cooper's ligaments.

Table 1. Proposed blind index (NAQI) calculated for 10 synthetic images of each type, without any degradation, categorized by number of compartments, Cooper's ligament thickness and breast skin thickness.

Average of NAQI				
No. Compartments	Overall percentage density (PD)			Total NAQI reduction
	15 %	21 %	26 %	
167	0.348	0.344	0.336	−4 %
333	0.333	0.326	0.318	−2 %
500	0.325	0.322	0.315	−3 %
1,000	0.318	0.315	0.310	−4 %
Total NAQI reduction	−9 %	−8 %	−8 %	
Ligament thickness				
0.06 (mm)	0.352	0.347	0.339	−4 %
0.04 (mm)	0.310	0.306	0.301	−4 %
Total NAQI reduction	−12 %	−12 %	−11 %	
Skin thickness				
1.2 (mm)	0.331	0.327	0.320	−3 %

4.2 Impact of Dose Reduction

The next set of experiments was performed using the set of 2,880 images simulated with 100 % and 50 % dose. In this experiment we investigated if variations on breast anatomy are more relevant then the changes on image quality.

Figure 2 shows the reported values of NAQI for different breast densities, number of compartments, and radiation dose. Figures 3 and 4 show the equivalent results for variations in the thickness of the Cooper's ligament and skin thickness, respectively.

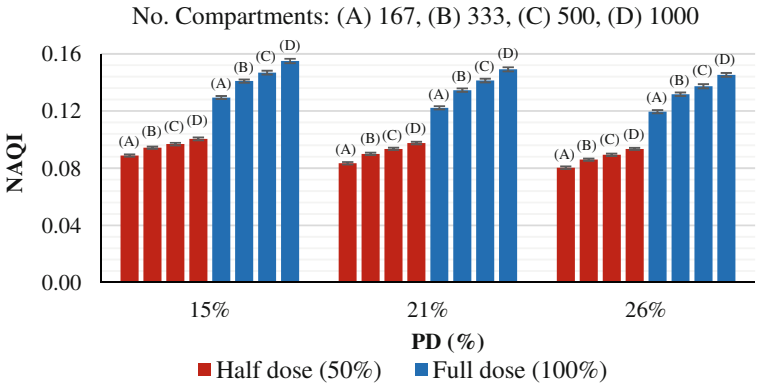

Fig. 2. Reported NAQI for each breast density and number of compartments. (Color figure online)

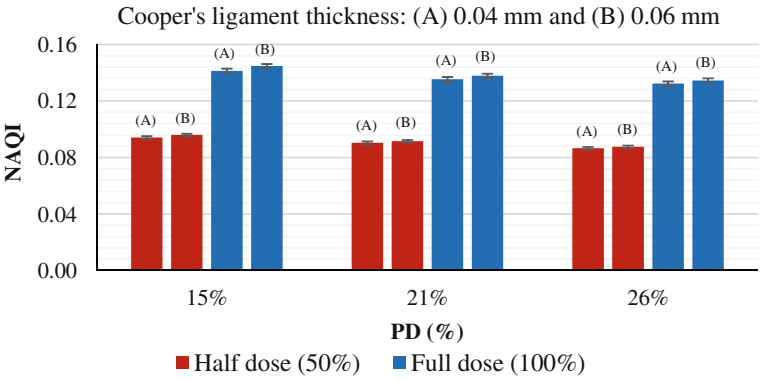

Fig. 3. Reported NAQI for each breast density and thickness of Cooper's ligament. (Color figure online)

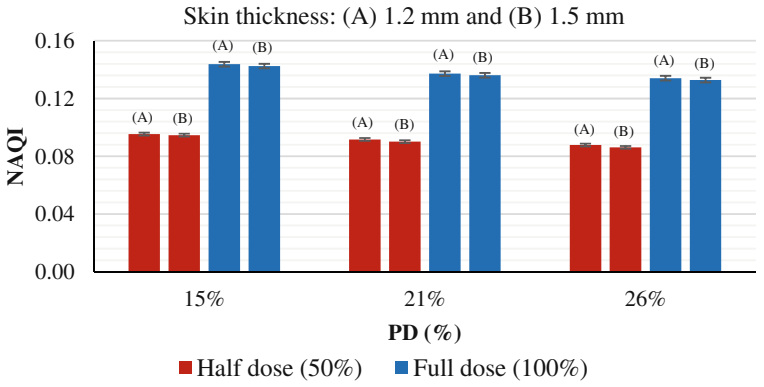

Fig. 4. Reported NAQI for each breast density and skin thickness. (Color figure online)

As a final experiment, we compared the distributions reported by the NAQI values at both radiation levels, as shown by Fig. 5. Using both distributions, we calculated the ROC curve that illustrates the performance of a binary classifier that would apply the NAQI to discriminate images at different doses. The reported AUC was 0.994.

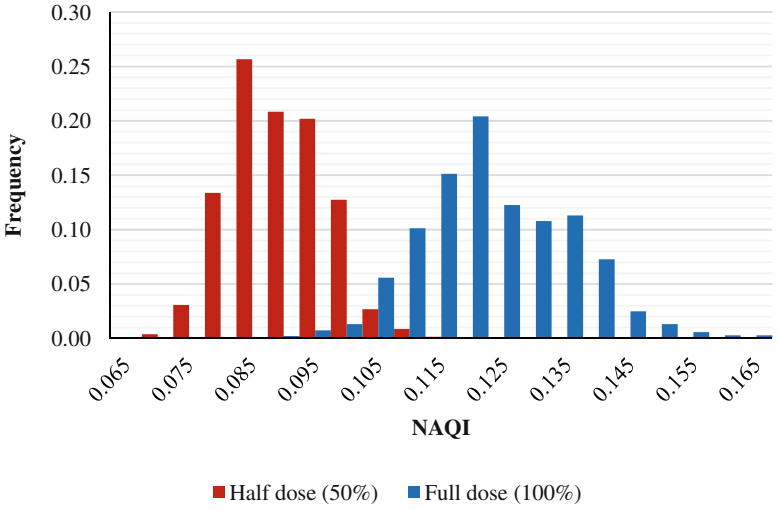

Fig. 5. Histogram distribution of NAQI showing the variation in accordance with the breast anatomy and radiation dose (half and full dose). (Color figure online)

5 Discussion and Conclusions

In this work, we evaluated the dependency of the NAQI to breast anatomy and radiation dose. To assess the image quality properly, the proposed index must report a stronger dependency to image quality than to breast anatomy.

The first experiment used noiseless images with a range of anatomies to assess how the breast composition individually influences the NAQI. The results of this experiment, presented in Table 1, showed that the NAQI decreases at most 12 % when only the anatomy changes.

In the second experiment, the images were simulated with quantum noise in accordance with the AEC in digital mammography, corresponding to the standard radiation dose (full dose) and 50 % dose of a clinical mammographic projection. The variation on image quality due to dose reduction yielded a decrease of 33 % in the NAQI. Figures 2 and 3 show how each anatomic feature affects the metric on image quality. It is evident that changes on image quality (between blue and red columns) cause greater variation in the metric than do variations in breast anatomy (within red and blue columns).

Moreover, in a detailed analysis, the NAQI trend is inverted in terms of the number of compartments, i.e., the anisotropic index tends to be higher because of the increased complexity arising from the larger number of compartments (from 167 to 1,000). This inverse relationship proves that the index is highly sensitive to quantum noise. Thus, although the AEC adjusts the exposure factors to different breast profiles, denser phantoms tend to achieve lower NAQI.

The discrete histogram distribution based upon the NAQI values differentiates with high accuracy (AUC = 0.994) good quality images (full dose) and low quality (half dose).

For future work, the authors will analyze a large institutional clinical data set to provide a more complete statistic.

Acknowledgments. The authors would like to thank São Paulo Research Foundation (FAPESP grant #2013/18915-5) and the Brazilian Foundation for the Coordination of Improvement of Higher Education Personnel (CAPES grant #99999.014175/2013-04 and grant #88881.030443/2013-01) for the financial support given to this project. The authors would also like to acknowledge the support of the National Institutes of Health/National Cancer Institute grant 1R01-CA154444 and the U.S. National Institute of General Medical Sciences (P20 GM103446) from the National Institutes of Health. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies. We thank Real Time Tomography (RTT) for providing assistance with image processing. We also acknowledge Abdullah-Al-Zubaer Imran, Niara Medley, Rick Emory, and Vernita Adkins for their assistance generating software breast phantoms and projections. ADAM is a member of the scientific advisory board and shareholder of RTT.

References

1. Li, Y., Poulos, A., Mclean, D., Rickard, M.: A review of methods of clinical image quality evaluation in mammography. Eur. J. Radiol. **74**, e122–e131 (2010). doi:10.1016/j.ejrad.2009.04.069
2. Wang, Z., Bovik, A.C.: Modern image quality assessment. Synth. Lect. Image Video Multimed. Process. (2006). doi:10.2200/S00010ED1V01Y200508IVM003
3. Oliveira, H.C.R., Barufaldi, B., Borges, L.R., Gabarda, S., Bakic, P.R., Maidment, A.D.A., Schiabel, H., Vieira, M.A.C.: Validation of no-reference image quality index for the assessment of digital mammographic images. In: SPIE Med Imaging, San Diego, CA, p. 978713-2 (2016). doi:10.1117/12.2217229
4. Bakic, P.R., Pokrajac, D.D., De Caro, R., Maidment, A.D.: Realistic simulation of breast tissue microstructure in software anthropomorphic phantoms. In: Fujita, H., Hara, T., Muramatsu, C. (eds.) IWDM 2014. LNCS, vol. 8539, pp. 348–355. Springer, Heidelberg (2014). doi:10.1007/978-3-319-07887-8_49
5. Bakic, P.R., Zhang, C., Maidment, A.D.A.: Development and characterization of an anthropomorphic breast software phantom based upon region-growing algorithm. Med. Phys. **38**, 3165–3176 (2011). doi:10.1118/1.3590357
6. Gabarda, S., Cristóbal, G.: Blind image quality assessment through anisotropy. J. Opt. Soc. Am. Opt. Image Sci. Vis. **24**, B42–B51 (2007). doi:10.1364/JOSAA.24.000B42