Image Quality Comparison between Digital and Synthetic 2D Mammograms: A Qualitative and Quantitative Phantom Study

P. Barca¹², R. Lamastra¹, D. Caramella³, A. C. Traino⁴, R. M. Tucciariello¹ and M. E. Fantacci¹²
¹Department of Physics, University of Pisa, Pisa, Italy
²INFN, Pisa Section, Pisa, Italy
³Department of Translational Research on New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy
⁴Unit of Medical Physics, Pisa University Hospital “Azienda Ospedaliero-Universitaria Pisana”, Pisa, Italy

Keywords: Synthesized Mammography, Digital Mammography, Image Quality, Modulation Transfer Function, Noise Power Spectrum, Contrast to Noise Ratio, ACR Phantom.

Abstract: The recent introduction of digital breast tomosynthesis (DBT) have lead to improvements in sensitivity and specificity of breast cancer detection, especially in cases of tumors developed in dense breasts. Since DBT provides tomographic slices of an entire tissue volume, it reduces the inherent tissue overlapping limitation of digital mammography (DM). In addition, DBT combined with DM has been proven to decrease recall and increase invasive cancer detection rates in breast cancer screening. However, the employment of DBT+DM implies a not negligible increment of patients absorbed dose. Therefore, Synthesized mammograms (SMs) generated from the DBT data have been recently introduced to eliminate the need of an additional DM. However, several studies showed differences between DM and SM images and some studies found contrasting results in terms of image quality when DM and SM images were compared. In our phantom study, we objectively compare image quality of SM and DM images in terms of noise, spatial resolution and contrast properties. Additionally, a qualitative analysis of the ACR mammographic phantom was performed in both modalities to assess the detectability of different features. SM images were characterized by different texture with respect to DM images, showing lower overall performances in terms of contrast-to-noise ratio and modulation transfer function. However, the goal of SM images is to provide a useful two-dimensional guide complementary to the DBT dataset and the performances in terms of high-contrast features detectability were satisfactory in comparison to those obtained in DM.

1 INTRODUCTION

Digital mammography (DM) is a low radiation dose imaging technique and represents the reference modality for early detection and diagnosis of breast cancer (Lehman et al., 2015). However, DM is a projective imaging modality and its two-dimensional nature results in tissue overlapping which may lead to false positive lesion diagnosis. In this context, the recent introduction of digital breast tomosynthesis (DBT) is an important step to reduce the limitations of projectional DM (Vedantham et al., 2015; Sechopoulos, 2013; Sechopoulos et al., 2013).

DBT is a pseudo-three dimensional technique that allows to obtain a set of breast slices by acquiring a limited number of breast projections from a narrow angular range. In DBT, the x-ray tube rotates along a fixed direction; a projection of the compressed breast is acquired every few degrees and a set of fixed-thickness slices is reconstructed trough filtered back-projection (FBP) or iterative approach (Vedantham et al., 2015).

Since DBT provides tomographic slices of an entire tissue volume, the advent of this new modality may help to partially overcome the inherent tissue overlapping limitation of DM (Vedantham et al., 2015).

In spite of its strengths in terms of tissue localization, DBT produces 2D slices whose image quality is lower than DM (Rodriguez-Ruiz et al., 2016). In particular, a single 2D slice of the entire tomographic set exhibits lower spatial resolution
than a DM image and therefore the detection of microcalcifications could be more difficult if only DBT images were evaluated (Rodriguez-Ruiz et al., 2016; Rose et al., 2013). In addition, DBT combined with DM, has been proven to decrease the necessity to recall patients, avoiding unnecessary anxiety as well as possible disruptions of the busy workflow of Imaging Diagnostic Departments. Moreover, the use of DBT has proven to increase the sensitivity and specificity of invasive cancer detection rates in breast cancer screening (Houssami, 2018; Shin et al., 2014; Svahn et al., 2010). Therefore, DBT is frequently used in conjunction with DM. However, the employment of DBT plus DM leads to a relevant increase in the absorbed breast dose (Durand, 2018; Zuckerman et al., 2017; Alshafei et al., 2017).

Furthermore, since DBT and DM images are acquired in sequence, the breast compression time is much longer with respect a single DBT or DM procedure, amplifying patient’s discomfort and artifacts due to patient motion during the examination (Durand, 2018).

To overcome these issues, recent advances in DBT technology were focused on the reconstruction of 2D “synthesized mammograms” (SMs) from the DBT acquisition dataset (Durand, 2018; Zuckerman et al., 2017; Smith, 2015). The advent of SMs is expected to be decisive to avoid an additional DM, particularly in screening programs in which is essential to achieve the lowest dose compatible with the clinical task.

The SMs are generated from the DBT dataset by a vendor-specific algorithm that combines the projections acquired from different angles into a single image. In the reconstruction process, filters can be applied and high-contrast features, glandular tissue and calcifications can be emphasized on SMs (Durand, 2018; Smith, 2018; Ratanaprasatporn et al., 2017). The goal of SMs is to provide a useful 2D image complementary to DBT slices, highlighting important features that could be less evident in DBT images. Moreover, SMs allow to profit from the well honed capabilities of expert breast radiologists who are used to analyze projection images.

However, since SMs are reconstructed from a DBT dataset, different image quality properties are exhibited with respect to DM images (Baldelli et al., 2018; Nelson et al., 2016). Therefore, an objective and comprehensive assessment of image quality of SMs in comparison to DM images is required.

Previous objective studies have shown some differences in terms of image quality between SMs and DM. Specifically, a phantom study conducted by Ikejimba et al. showed that the detectability of low-contrast signal was significantly lower for SMs than for DM (Ikejimba et al., 2016). Additionally, Nelson et al. in their work found that SMs offered a better depiction of objects of certain size and contrast with respect to DM. However, SMs provided poorer overall resolution and noise properties (Nelson et al., 2016). In another phantom study, Baldelli et al. concluded that the image quality of SMs does not exhibit significant differences with respect to DM (Baldelli et al., 2018). Furthermore, some clinical studies reported similar performance between SMs and DM (Zuckerman et al., 2017; Mariscotti et al., 2017).

Given these previous results, more insights into the comparison between SM and DM could be of practical interest. Therefore, we aimed at conducting a qualitative and quantitative comparison between image quality in SMs and DM. Noise, contrast and spatial resolution properties were objectively evaluated through the assessment of Noise Power Spectrum (NPS), Contrast-to-Noise Ratio (CNR) and Modulation Transfer Function (MTF) both for SMs and DM. Additionally, the ACR phantom was analyzed to compare the detectability of many high and low contrast features in automatic exposure conditions.

This work is part of the “RADIOMA” project (RADiazioni IOntizzanti in MAmmografia, ionising radiations in mammography) whose aim is to evaluate absorbed dose and image quality in DM and DBT (Traino et al., 2018; Sottocornola et al., 2018).

2 MATERIALS AND METHODS

Images of three different phantoms were acquired in DM and DBT modality on a Selenia Dimensions device (Hologic, Bedford, Mass, USA) by using the automatic exposure settings. This machine implements both DM and DBT modalities (angle range of +/- 7.5°). SMs (“C-view” images, Hologic) were reconstructed from the DBT dataset (Selenia Dimensions User Manual, 2010). Processed (“for presentation”) DM images were considered in this study.

Image quality was quantitatively evaluated through the assessment of noise, spatial resolution and contrast-to-noise ratio properties of the images. In addition, a qualitative analysis on the detectability of a number of features was performed.

A 30x24x4 cm³ PMMA plate was imaged to study the noise spectral properties. The NPS was computed by following the Siewerdsen approach.
A radial average of the 2D NPS was calculated to better appreciate differences between SM and DM images.

To evaluate the spatial resolution an home-made phantom with a 12.5 \(\mu\)m tungsten wire (Fig. 1) was used to measured the Line Spread Function (LSF) of the system. The wire was placed in a thin PMMA support (2 mm of thickness) with an angle of 3 degrees, as suggested by the EUREF DBT protocol (EUREF, 2016). The PMMA support was placed under a PMMA plate of 4 cm of thickness to approximate a 4.5 cm thick breast.

The MTF was obtained by computing the module of the LSF Fourier transform (EUREF, 2016).

For the CNR calculation and for a qualitative detectability analysis the ACR phantom (Fluke Biomedical, Everett, WA, USA; Fig. 2) was employed.

The 4.4 cm thick ACR phantom is made up of a 7 mm wax block insert containing 16 sets of test objects, a 3.4 cm thick acrylic base, and a 3 mm thick cover. It approximates a 4.5 cm compressed breast of average glandular/adipose composition. Included in the wax insert are aluminium oxide, \(\text{(Al}_2\text{O}_3)\) specks to simulate micro-califications. Six different size nylon fibers simulate fibrous structures and five different size lens shaped masses simulate tumors (Fluke Biomedical, ACR mammographic phantom - Operators Manual, 2005).

The CNR was evaluated in the four larger masses of the phantom through the following relationship:

\[
\text{CNR} = \frac{PV_{\text{insert}} - PV_{\text{background}}}{\sigma_{\text{background}}} \quad (1)
\]

where \(PV_{\text{insert}}\) and \(PV_{\text{background}}\) are the mean pixel values in a region of interest (ROI) placed within the insert and in the background region respectively; \(\sigma_{\text{background}}\) is the standard deviation computed in the background ROI.

Additionally, the ACR phantom was employed for a qualitative analysis performed by a senior reader (more than 25 years of radiological experience) and two junior ones, who evaluated the detectability of the specks, fibers and masses. As recommended by manufacturers, a good imaging system should be able to see the 3rd speck group, 4th fiber, and the 3rd mass (Fluke Biomedical, ACR mammographic phantom - Operators Manual, 2005; EFOMP, 2015). This criterion was used for a qualitative comparison between SM and DM images by employing typical clinical settings.

Image analysis was performed by using ImageJ (Wayne Rasband, National Institute of Health, USA) and Origin (Origin-Lab Corporation, MA, USA) software packages.
3 RESULTS

Figure 3 shows the radial NPS obtained from SM and DM images acquired with automatic exposure parameters.

From Fig. 3 clear differences can be appreciated in NPS shape. More in detail, the NPS of DM images appeared almost flat within the entire spatial frequency range, while an evident peak around 1 mm\(^{-1}\) is expressed by SM images. Therefore, the image texture of SM images appeared more coarse grained than the texture of DM images, as shown in Fig. 4.

Results about spatial resolution are presented in Fig. 5. A comparison between MTFs of DM and SM images was carried out in both x (i.e. along the tube-travel direction) and y directions (i.e. orthogonal to the tube-travel direction). The MTF for DM images shown in Fig. 5 is representative of both x and y direction (i.e. no differences were found in those cases). Evident discrepancies are exhibited between the MTF computed in the y direction from SM images and the MTF of DM images. Additionally, great differences can be observed between the x- and y-MTFs of SM images as well as between the x-MTF of SM images and the MTF of DM images. The spatial frequency corresponding to MTF\(_{50}\)\% and MTF\(_{10}\)\% are presented in Tab. 1.

In Tab. 2 the CNRs obtained from the four largest masses of the ACR phantom are presented. The 0.25 mm mass was not visible in the ACR phantom images, for this reason it was not considered in this analysis.

The CNR values obtained in DM images are slightly higher than those obtained in SM (C-view) images for all analyzed inserts, especially for 2 mm and 1 mm insert sizes.

---

**Table 1:** Spatial frequency corresponding to MTF\(_{50}\)\% and MTF\(_{10}\)\% for DM and SM (C-view) images acquired with automatic exposure parameters. The presented values were extracted from MTF curves shown in Fig. 5.

<table>
<thead>
<tr>
<th>Insert size (mm)</th>
<th>DM</th>
<th>C-view</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(f_x) (mm(^{-1}))</td>
<td>(f_y) (mm(^{-1}))</td>
</tr>
<tr>
<td>2</td>
<td>7.1</td>
<td>7.1</td>
</tr>
<tr>
<td>1</td>
<td>6.7</td>
<td>6.7</td>
</tr>
<tr>
<td>0.75</td>
<td>12.3</td>
<td>12.3</td>
</tr>
</tbody>
</table>

**Table 2:** CNRs calculated for four masses of the ACR phantom from images acquired with automatic exposure parameters.

<table>
<thead>
<tr>
<th>Insert size (mm)</th>
<th>CNR (DM)</th>
<th>CNR (C-view)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2.23±0.04</td>
<td>1.51±0.15</td>
</tr>
<tr>
<td>1</td>
<td>1.25±0.03</td>
<td>1.14±0.12</td>
</tr>
<tr>
<td>0.75</td>
<td>0.88±0.03</td>
<td>0.84±0.11</td>
</tr>
<tr>
<td>0.5</td>
<td>0.54±0.03</td>
<td>0.42±0.10</td>
</tr>
</tbody>
</table>
Some examples of the ACR phantom inserts images employed for a qualitative evaluation of detectability are shown in figures 6-8.

In Fig. 6 images of the fibers set are presented both for DM and SM modality.

According to Fig. 6, five fibers are detectable from the DM image, while the fifth fiber (diameter of 0.54 mm) was not visible on the SM. The sixth nylon fiber (diameter of 0.4 mm) was not identifiable both on DM and SM images.

DM and SM images of the five specks patterns of the ACR phantom are shown in Fig. 7.

As shown in Fig. 7, the two specks groups corresponding to 0.24 mm and 0.16 mm diameter sizes were not detectable in both modalities. The third specks group (diameters of 0.32 mm) appeared slightly clearer in SM images.

Figure 8 shows the results in terms of ACR phantom masses detectability.

The mass corresponding to 0.25 mm of thickness was detectable neither in DM nor in SM images (Fig. 8), while the other masses were clearly visible in both modalities.

4 DISCUSSION

The advent of SMs represents a useful diagnostic support to DBT images, offering a 2D guide complementary to 3D DBT information (Durand, 2018; Zuckerman et al., 2017; Smith, 2015). Therefore, SMs have been proposed to substituting DMs when DBT procedures requires additional DM exams, to avoid additional absorbed dose. (Durand, 2018; Zuckerman et al., 2017; Ratanaprasatporn et al., 2017).

Given the potential of SMs, a careful evaluation of SM image quality in comparison to DM is necessary.
In this phantom study a quantitative and qualitative assessment of SM and DM image quality is presented. In order to perform a more clinically relevant analysis, automatic exposure parameters where employed in images acquisitions and only processed images where analysed. It should be noted that the automatic exposure settings were different for DM and DBT modalities. In particular, to optimise the DBT acquisition and reconstruction, a dedicated W-Al anode-filter combination is employed in the Holocig system. The different anode-filter combination expresses different X-ray spectra than those employed in DM (W-Rh combination) and therefore different tube voltage and tube load values are employed in the two modalities for the same phantom/breast thickness.

Previous studies have shown some differences in terms of image quality between SM and DM images.

More in detail, Ikejimba et al. conducted a contrast-detail study by assessing detectability reader performances in DM, DBT and SM images (Ikejimba et al., 2016). They concluded that the observer performances were significantly lower for SM than for DM or DBT, for all imaging conditions.

In their phantom study Nelson et al. objectively analysed CNR and noise properties of DM and SM images (Nelson et al., 2016). The ACR phantom was used to obtain CNR-based object visibility thresholds, while a 3D printed anthropomorphic breast phantom was employed to perform a visual inspection of simulated calcifications and evaluation of the high-contrast resolution features. Image texture was also studied. They found that the SM image quality differs from DM, offering a better depiction of objects of certain size and contrast, but providing poorer overall resolution and noise properties.

Another work, presented by Baldelli et al., studied sizes and contrasts of several patterns of microcalcification specks, shapes and contrasts of circular masses, as well as NPS of background regions (Baldelli et al., 2018) by employing two different phantoms. No significant differences in image quality were found between SM and DM images.

Moreover, in spite of differences in the appearance of SM and DM images, recent clinical studies showed similar interpretive performance between SM and DM images, confirming the potential role of SMs as alternative to DM in addition to DBT (Zuckerman et al., 2017; Mariscotti et al., 2017).

As expected, our findings confirm the different image appearance of SMs compared to DM images. In particular, the noise texture of SM images was more coarse grained than the one observed in DM, confirming the results found by Nelson et al. (Nelson et al., 2016). This aspect was quantified by NPS curves showed in Fig. 3. The NPS of DM was typical of uncorrelated noise, showing an almost constant trend within the entire spatial frequency range. Conversely, SMs expressed a different NPS shape with an evident peak centered at 1 mm⁻¹. This behaviour is strictly related to the reconstruction process involved in SMs generation, which introduces spatial correlation in image noise. The coarse grained texture of SMs could impact on the detectability and/or boundaries of low contrast structures as can be visually appreciated in Fig. 4.

Spatial resolution was evaluated through LSF measurements by computing the MTF. DM images showed higher MTFs than SMs over all spatial frequencies. No appreciable differences were found between x-MTF and y-MTF of DM images. On the other hand, the y-MTF of SM images was considerably lower than those obtained from DM images (Fig. 5). Moreover, for SM images, the x-MTF (i.e. the MTF along the tube travel direction) was significantly lower than the y-MTF (Fig. 5). This feature is a direct consequence of the DBT data acquisition process in which the x-ray tube rotates along the x-direction. The spread artefact along the x-direction is then expressed in reconstructed SM images. Table 1 summarises the spatial frequency corresponding to 50% and 10% of MTF curves of Fig. 5. These results about spatial resolution are in agreement to those reported in international reports (IAEA Human Health Series No. 17 - Table 17, 2011; CEP Report 05084 - Table 11, 2006) and observed in previous studies (Nelson et al., 2016; Jesneck et al., 2005).

Moreover, CNRs were computed from the masses inserts of the ACR phantom (Table 2). Higher CNRs were found in DM images for all considered masses sizes. Differences in CNRs between the two modalities depended on the masses size and the maximum discrepancy was found for the largest insert (2 mm of thickness). However, the four analysed masses were clearly detectable both in SM and DM images, as shown in Fig. 8. Uncertainties in CNR were estimated by computing the coefficient of variation from a set of images acquired under the same conditions.
From a visual inspection of DM and SM images, the acceptable detectability thresholds of the ACR phantom were respected for all groups of inserts. In fact, five fibers, four specks patterns and four masses were identifiable in DM images, while four fibers, four specks patterns and four masses were detectable in SMs (Fig. 6-8). Notice that a good imaging system should be able to see the 3rd speck group, 4th fiber, and the 3rd mass (Fluke Biomedical, ACR mammographic phantom - Operators Manual, 2005; EFOMP, 2015). In spite of these encouraging results, SMs seem to exhibit lower performances than DM images for low-contrast inserts (Fig. 6 and 8). Conversely, SMs exalted the visibility of specks with respect to DM images (Fig. 7). This aspects are confirmed by radiologists and can be related to the reconstruction process of SMs. In fact, filters applied in SMs reconstruction have the potential to highlight high-contrast features (Durand, 2018; Smith, 2018; Ratanaprasatporn et al., 2017). This characteristic of SMs may be advantageous especially in dense breasts, providing essential information complementary to DBT.

Our analysis was carried out by considering standard phantoms thicknesses to approximate a 4.5 cm thick breast (the ACR, the PMMA and the homemade phantom). It should be reminded that the automatic exposure parameters change when different thicknesses are involved. However, this is expected to have weak impact on the image quality comparison between SM and DM images. In fact, the automatic exposure is developed to guarantee an acceptable noise/contrast level on the final image in all modalities for each breast thickness.

Based on our results, even though SMs provide lower performances in terms of noise texture, spatial resolution and CNR, the ACR analysis showed that SMs could permit adequate visualization of low-contrast inserts and emphasized the appearance of microcalcifications. It should be reminded that SMs are not designed to be identical copies of DM images, instead, SM images should provide a useful additional information in supplement to DBT diagnostic capabilities. Therefore, SMs can not be employed as a stand-alone diagnostic tool but shall accompany the full DBT dataset (Murphy et al., 2018).

Although this preliminary study confirms some results of other previous studies, further studies are required to better assess the SMs performances in terms of image quality and objects detectability. Specifically, a contrast-detail analysis is crucial to assess the detectability performances of a mammographic imaging system. Therefore we are planning to extend this study with the employment of other dedicated phantoms and to perform additional analysis of clinical images.

5 CONCLUSIONS

In this study, a comprehensive comparison between image quality of SM and DM images was performed. Noise, spatial resolution, contrast and detectability properties of phantom images were assessed.

Even though SMs expressed overall lower noise, contrast and spatial resolution properties than DM images, their detectability performances were still satisfactory. In addition, high-contrast features were exalted in SM images. Therefore, our results confirm the potential role of SMs as useful 2D images complementary to the 3D DBT dataset.

Further studies should investigate contrast-detail properties of SM images, even in comparison to DM images.

Even though clinical images were not yet analysed in this study, our findings provide additional insights into the characterisation of SMs image quality performance and could be of practical interest toward a more appropriate employment of this new technology.

ACKNOWLEDGMENTS

The presented work is part of the RADIOMA project which is partially funded by "Fondazione Pisa", Technological and Scientific Research Sector, Via Pietro Toselli 29, Pisa.

REFERENCES


Sechopoulos I. 2013. A review of breast tomosynthesis. Part II. Image reconstruction, processing, analysis, and