Clinical application of low-dose phase contrast breast CT: methods for the optimization of the reconstruction workflow

S. Pacilè,[1,2,*] F. Brun,[1,2] C. Dullin,[3] Y. I. Nesterest,[4] D. Dreossi,[1] S. Mohammadi,[1,5,6] M. Tonutti,[7] F. Stacul,[7] D. Lockie,[8] F. Zanconati,[9] A. Accardo,[2] G. Tromba,[1] and T. E. Gureyev[4,10,11,12]

[1]*Elettra - Sincrotrone Trieste S.C.p.A., Basovizza (Trieste), Italy*
[2]*Department of Engineering and Architecture, University of Trieste, Trieste, Italy*
[3]*Department of Diagnostic and Interventional Radiology, University Hospital Goettingen, Goettingen, Germany*
[4]*Commonwealth Scientific and Industrial Research Organisation, Melbourne, Australia*
[5]*The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy*
[6] *now at LAC+ USC Medical Center, Los Angeles, CA, USA*
[7]*AOU - Trieste Hospital, Department of Radiology, Trieste, Italy*
[8]*Maroondah BreastScreen, Melbourne, Australia*
[9]*Department of Medical Science-Unit of Pathology, University of Trieste, Trieste, Italy*
[10]*School of Physics and Astronomy, Monash University, Clayton, VIC, Australia*
[11]*School of Science and Engineering, University of New England, Armidale, NSW, Australia*
[12]*ARC Centre of Excellence in Advanced Molecular Imaging, The University of Melbourne, Parkville, Australia*

[*]serena.pacile@elettra.eu

Abstract: Results are presented of a feasibility study of three-dimensional X-ray tomographic mammography utilising in-line phase contrast. Experiments were performed at SYRMEP beamline of Elettra synchrotron. A specially designed plastic phantom and a mastectomy sample containing a malignant lesion were used to study the reconstructed image quality as a function of different image processing operations. Detailed evaluation and optimization of image reconstruction workflows have been carried out using combinations of several advanced computed tomography algorithms with different pre-processing and post-processing steps. Special attention was paid to the effect of phase retrieval on the diagnostic value of the reconstructed images. A number of objective image quality indices have been applied for quantitative evaluation of the results, and these were compared with subjective assessments of the same images by three experienced radiologists and one pathologist. The outcomes of this study provide practical guidelines for the optimization of image processing workflows in synchrotron-based phase-contrast mammo-tomography.

References and links

1. American Cancer Society, "Breast cancer facts & figures," Tech. rep., Atlanta: American Cancer Society, Inc. (2007-2008).
2. R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2015," CA-Cancer J. Clin. **65**, 5–29 (2015).
3. R. E. Hendrick, R. A. Smith, J. H. Rutledge 3rd., and C. R. Smart, "Benefit of screening mammography in women aged 40-49: a new meta-analysis of randomized controlled trials." J. Natl. Cancer I. pp. 87–92 (1997).
4. F. Kamangar, G. M. Dores, and W. F. Anderson, "Patterns of cancer incidence, mortality, and prevalence across five continents: Defining priorities to reduce cancer disparities in different geographic regions of the world," J. Clin. Oncol. **24**, 2137–2150 (2006).
5. E. D. Pisano, R. E. Hendrick, M. J. Yaffe, J. K. Baum, S. Acharyya, J. B. Cormack, L. A. Hanna, E. F. Conant, L. L. Fajardo, L. W. Bassett, C. J. D'Orsi, R. A. Jong, M. Rebner, A. N. A. Tosteson, and C. A. Gatsonis, "Diagnostic accuracy of digital versus film mammography: Exploratory analysis of selected population subgroups in DMIST," Radiology **246**, 376–383 (2008).
6. N. D. Prionas, K. K. Lindfors, S. Ray, S.-Y. Huang, L. A. Beckett, W. L. Monsky, and J. M. Boone, "Contrast-enhanced dedicated breast CT: initial clinical experience." Radiology **256**, 714–723 (2010).
7. A. O'Connell, D. L. Conover, Y. Zhang, P. Seifert, W. Logan-Young, C. . L. Lin, L. Sahler, and R. Ning, "Cone-beam CT for breast imaging: Radiation dose, breast coverage, and image quality," Am. J. Roentgenol. **195**, 496–509 (2010).
8. J. M. Boone and K. K. Lindfors, "Breast CT: Potential for breast cancer screening and diagnosis," Future Oncology **2**, 351–356 (2006).
9. B. Zhao, X. Zhang, W. Cai, D. Conover, and R. Ning, "Cone beam breast CT with multiplanar and three dimensional visualization in differentiating breast masses compared with mammography," Eur. J. Radiol. **84**, 48–53 (2015).
10. A. Mittone, A. Bravin, and P. Coan, "Radiation dose in breast CT imaging with monochromatic x-rays: Simulation study of the influence of energy, composition and thickness," Phys. Med. Biol. **59**, 2199–2217 (2014).
11. A. Olivo, S. Gkoumas, M. Endrizzi, C. K. Hagen, M. B. Szafraniec, P. C. Diemoz, P. R. T. Munro, K. Ignatyev, B. Johnson, J. A. Horrocks, S. J. Vinnicombe, J. L. Jones, and R. D. Speller, "Low-dose phase contrast mammography with conventional x-ray sources," Med. Phys. **40** (2013).
12. P. Coan, A. Bravin, and G. Tromba, "Phase-contrast x-ray imaging of the breast: Recent developments towards clinics," J. Phys. D Appl. Phys. **46**, 494007 (2013).
13. Y. Zhao, E. Brun, P. Coan, Z. Huang, A. Sztrkay, P. C. Diemoz, S. Liebhardt, A. Mittone, S. Gasilov, J. Miao, and A. Bravin, "High-resolution, low-dose phase contrast X-ray tomography for 3D diagnosis of human breast cancers," in Proc. Natl. Acad. Sci. USA **109**, 18290–18294 (2012).
14. S. Grandl, M. Willner, J. Herzen, A. Sztrkay-Gaul, D. Mayr, S. D. Auweter, A. Hipp, L. Birnbacher, M. Marschner, M. Chabior, M. Reiser, F. Pfeiffer, F. Bamberg, and K. Hellerhoff, "Visualizing typical features of breast fibroadenomas using phase-contrast CT: An ex-vivo study," PLoS ONE **9**, 97101 (2014).
15. A. Bravin, P. Coan, and P. Suortti, "X-ray phase-contrast imaging: From pre-clinical applications towards clinics," Phys. Med. Biol. **58**, R1–R35 (2013).
16. T. E. Gureyev, S. C. Mayo, Y. I. Nesterets, S. Mohammadi, D. Lockie, R. H. Menk, F. Arfelli, K. M. Pavlov, M. J. Kitchen, F. Zanconati, C. Dullin, and G. Tromba, "Investigation of the imaging quality of synchrotron-based phase-contrast mammographic tomography," J. Phys. D Appl. Phys. **47**, 365401 (2014).
17. E. Castelli, M. Tonutti, F. Arfelli, R. Longo, E. Quaia, L. Rigon, D. Sanabor, F. Zanconati, D. Dreossi, A. Abrami, E. Quai, P. Bregant, K. Casarin, V. Chenda, R. H. Menk, T. Rokvic, A. Vascotto, G. Tromba, and M. A. Cova, "Mammography with synchrotron radiation: First clinical experience with phase-detection technique," Radiology **259**, 684–694 (2011).
18. G. Tromba, M. A. Cova, and E. Castelli, "Phase-contrast mammography at the SYRMEP beamline of Elettra," Synchrotron Radiation News **24**, 3–7 (2011).
19. D. Dreossi, A. Abrami, F. Arfelli, P. Bregant, K. Casarin, V. Chenda, M. A. Cova, R. Longo, R. . Menk, E. Quai, E. Quaia, L. Rigon, T. Rokvic, D. Sanabor, M. Tonutti, G. Tromba, A. Vascotto, F. Zanconati, and E. Castelli, "The mammography project at the SYRMEP beamline," Eur. J. Radiol. **68**, 58–62 (2008).
20. Y. I. Nesterets, T. E. Gureyev, S. C. Mayo, A. W. Stevenson, D. Thompson, J. M. C. Brown, M. J. Kitchen, K. M. Pavlov, D. Lockie, F. Brun, and G. Tromba, "A feasibility study of X-ray phase-contrast mammographic tomography at the Imaging and Medical beamline of the Australian Synchrotron," Submitted to J. Sync. Rad. (2015).
21. G. Tromba, R. Longo, A. Abrami, F. Arfelli, A. Astolfo, P. Bregant, F. Brun, K. Casarin, V. Chenda, D. Dreossi, M. Hola, J. Kaiser, L. Mancini, R. H. Menk, E. Quai, E. Quaia, L. Rigon, T. Rokvic, N. Sodini, D. Sanabor, E. Schultke, M. Tonutti, A. Vascotto, F. Zanconati, M. Cova, and E. Castelli, "The SYRMEP beamline of elettra: Clinical mammography and bio-medical applications," in AIP Conference Proceedings **1266**, 18–23 (2010).
22. M. Boin and A. Haibel, "Compensation of ring artefacts in synchrotron tomographic images," Opt. Express **14**, 12071–12075 (2006).
23. F. Brun, A. Accardo, G. Kourousias, D. Dreossi, and R. Pugliese, "Effective implementation of ring artifacts removal filters for synchrotron radiation microtomographic images," International Symposium on Image and

Signal Processing and Analysis, ISPA pp. 672–676 (2013).

24. F. Brun, S. Pacilè, A. Accardo, L. Mancini, D. Dreossi, G. Tromba, G. Kourousias, and R. Pugliese, "Enhanced and flexible software tools for X-ray tomography experiments at the Italian synchrotron radiation facility Elettra," Fundamenta Informaticae (in press).

25. A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging* (IEEE Press, 1988).

26. A. van der Sluis and H. van der Vorst, "SIRT- and CG-type methods for the iterative solution of sparse linear least-squares problems," Linear Algebra Appl. **130**, 257–303 (1990).

27. R. Gordon, R. Bender, and G. Herman, "Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography," J. Theor. Biol. **29**, 471–476 (1970).

28. J. Miao, F. Frster, and O. Levi, "Equally sloped tomography with oversampling reconstruction," Phys. Rev. B **72**, 052103 (2005).

29. J. Tang, B. E. Nett, and G. . Chen, "Performance comparison between total variation (TV)-based compressed sensing and statistical iterative reconstruction algorithms," Phys. Med. Biol. **54**, 5781–5804 (2009).

30. G. R. Myers, C. D. L. Thomas, D. M. Paganin, T. E. Gureyev, and J. G. Clement, "A general few-projection method for tomographic reconstruction of samples consisting of several distinct materials," Appl. Phys. Lett. **96**, 021105 (2010).

31. D. Paganin, S. C. Mayo, T. E. Gureyev, P. R. Miller, and S. W. Wilkins, "Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object," J. Microsc. **206**, 33–40 (2002).

32. T. E. Gureyev, Y. Nesterets, D. Ternovski, D. Thompson, S. W. Wilkins, A. W. Stevenson, A. Sakellariou, and J. A. Taylor, "Toolbox for advanced X-ray image processing," in Proceedings of SPIE - The International Society for Optical Engineering **8141**, 81410B (2011).

33. T. E. Gureyev, Y. I. Nesterets, F. De Hoog, G. Schmalz, S. C. Mayo, S. Mohammadi, and G. Tromba, "Duality between noise and spatial resolution in linear systems," Opt. Express **22**, 9087–9094 (2014).

34. S. M. Smith and J. M. Brady, "SUSAN - A new approach to low level image processing," Int. J. Comput. Vision **23**, 45–78 (1997).

35. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Trans. Image Process. **13**, 600–612 (2004).

36. Z. Wang and A. C. Bovik, "A universal image quality index," IEEE Signal Proc. Let. **9**, 81–84 (2002).

37. N. Damera-Venkata, T. D. Kite, W. S. Geisler, B. L. Evans, and A. C. Bovik, "Image quality assessment based on a degradation model," IEEE Trans. Image Process. **9**, 636–650 (2000).

38. W. S. Geisler and M. S. Banks, *Handbook of Optics*, vol. Visual performance (M.Bass, Ed. New York: McGraw-Hill, 1995).

39. Z. Wang and A. C. Bovik, "Embedded foveation image coding," IEEE Trans. Image Process. **10**, 1397–1410 (2001).

40. S. Mohammadi, E. Larsson, F. Alves, S. Dal Monego, S. Biffi, C. Garrovo, A. Lorenzon, G. Tromba, and C. Dullin, "Quantitative evaluation of a single-distance phase-retrieval method applied on in-line phase-contrast images of a mouse lung," J. Synchrotron Radiat. **21**, 784–789 (2014).

41. A. Mirone, E. Brun, and P. Coan, "A dictionary learning approach with overlap for the low dose computed tomography reconstruction and its vectorial application to differential phase tomography," PLoS ONE **9**, 114325 (2014).

42. B. Liu, H. Yu, S. S. Verbridge, L. Sun, and G. Wang, "Dictionary-learning-based reconstruction method for electron tomography," Scanning **36**, 377–383 (2014).

---

1. Introduction

Breast cancer is one of the two leading causes of cancer fatalities among women in most industrialised countries [1, 2]. Some types of breast cancer can be aggressive, with success of its treatment depending heavily on its early detection when it is small in size and less likely to have spread to other vital organs. Therefore, health authorities in many countries recommend regular mammographic screening of women over 40 years of age. Two-view 2D mammography is currently the technique used for this purpose.

The application of screening programs over large populations has contributed to an important reduction of breast cancer mortality [3, 4], however X-ray 2D conventional mammography still has two important limitations. The technique is intrinsically limited as it represents a 3D structure as a 2D projection image, leading to false negative cases where cancers are camouflaged by adjacent tissues, and false positive cases when "pseudo lesions" are created by superimposition of normal overlapping structures. The other major limitation is related to the minimal difference in X-ray attenuation between normal breast tissues, and also between normal and

abnormal breast tissues that implies low sensitivity imaging. These problems increase with increasing breast tissue density. The reduced sensitivity in patients with dense breast tissue has been only partially overcome with the introduction of new digital detectors [5]. Digital Breast Tomosynthesis (DBT) and Computed Tomography (CT) solve the overlap of anatomical structures/tissues affecting planar 2D mammography and prospectively give considerable improvement in the sensitivity and specificity of breast cancer diagnosis [6–9]. Despite that, the intrinsic limitation of standard radiology due to the small differences in X-ray attenuation between the tissues composing the breast (mainly the glandular and tumor tissue) remains unsolved. This has pushed the scientific research towards the development of new imaging modalities for breast cancer diagnosis. In this scenario, the use of synchrotron X-rays, as gold standard radiation, and phase contrast (PhC) techniques play an interesting role [10–14]. Among the PhC techniques [15], propagation-based imaging (PBI) is particularly attractive as its implementation requires simply that the detector is moved a certain distance from the sample under study, without the need for any additional experimental hardware.

The main concern for the application of CT in breast imaging with synchrotron radiation (SR) is the dose delivered to the patient. If the dose could be kept at a level comparable to that delivered in the present-day clinical 2D mammography or tomosynthesis, still ensuring sufficient image quality, mammographic CT is likely to outperform the other imaging techniques in terms of diagnostic value. To this aim, as a pre-requisite for successful translation of SR tomo-mammography into clinical practice, it is essential to evaluate, quantify and optimize the main parameters of the PBI-CT imaging technique, i.e. the choice of X-ray energy, sample-to-detector distance, detector characteristics, strategies for CT scans and the reconstruction workflow. This needs to be done in order to maximize the image quality and therefore the diagnostic information [16].

In this article, a comparison of different low-dose CT reconstruction workflows for tomo-mammography is proposed. The opinion of expert radiologists was considered to support the evaluation process. With the ultimate aim of an unsupervised comparison, several quantitative metrics were computed to assess the quality of the reconstructed images. In addition to the evaluation of common image characteristics (e.g., spatial resolution and signal-to-noise ratio - SNR), new image quality indexes are also presented, which are supposed to take into account the spatial resolution and noise sensitivity of the imaging setup simultaneously. The results presented here refer to a study performed at the SYnchrotron Radiation for MEdical Physics (SYRMEP) beamline of the ELETTRA synchrotron light source (Basovizza - Trieste, Italy), which is so far the only synchrotron beamline where trials on human patients have been already conducted in 2D projection mammography mode [17–19]. In this framework the SYRMA-CT (SYnchrotron Radiation for MAmmography) project aims to set-up a new clinical trial of PhC breast CT with synchrotron radiation in near future [20].

2. Materials and Methods

2.1. Imaging setup

The X-ray source at the SYRMEP beamline is a bending magnet and the monochromator is based on a double Si(1,1,1) crystal working in Bragg configuration, it can deliver an X-ray beam with an energy in the 8.5-40 keV range and an energy resolution of 10^{-3} [21]. The horizontal acceptance covered by the front-end light port is 7 mrad. Therefore, the beamline provides at the sample position located 23 m from the source, a laminar section X-ray beam with a maximum area of about $(160 \times 3)mm^2$ at 30 keV. A custom - built ionization chamber, placed upstream to the sample, is used to determine the entrance air kerma on the sample and hence to calculate the delivered dose. The detector used in this study is a DALSA Argus high resolution CCD TDI sensor with a pixel size of 27 μm and a maximum X-ray resolution up to

15 lp/mm. The phase contrast CT scans were taken in PBI modality with a sample to detector distance of 1 m.

2.2. Samples description and radiation dose assessment

A polycarbonate phantom with the diameter of 10 cm (Fig. 1(a)) containing 8 holes filled with different materials to simulate X-ray absorption and refraction characteristics of breast tissue was scanned at both high and low photon statistics (i.e. with different photon flux obtained by interposing filters between the source and the sample). A tissue specimen sampled during a surgical mastectomy obtained out of overleft material to be wasted was also scanned at high and low statistic (Fig. 1(b)). Before data acquisition, in addition to standard fixation, the specimen was sealed in a plastic bag under vacuum for a better conservation and handling. Prior to the surgical mastectomy the patient expressed the agreement about the possibility to use this material for scientific experiments. For both the considered samples the experimental parameters are: energy = 32 keV, sample-detector distance = 1 m, number of projections = 3600, detector binning = 2×2 with resulting voxel size = 54 μm (due to a small geometrical magnification of 24/23 =1.043 the effective pixel size of the detector, in the sample plane, is 51.75 μm). CT scans at low statistic were performed with 3600 projections at an Average Glandular Dose (AGD) calculated in the irradiated volume of 136 mSv. In order to obtain a value of AGD comparable to the clinical mammography, for each dataset only 1/10 of total projections with added simulated Poisson noise were considered. To be specific, noise was increased by measuring the standard deviation of a background area and multiplying it by a factor 3, thus the resulting AGD released to the sample is reduced to about 1.5 mSv. The evaluation of AGD has been performed from the values of air kerma measured by the ionisation chamber of the beamline and applying a Monte Carlo simulation code for the assessment of the delivered doses in the glandular tissue.
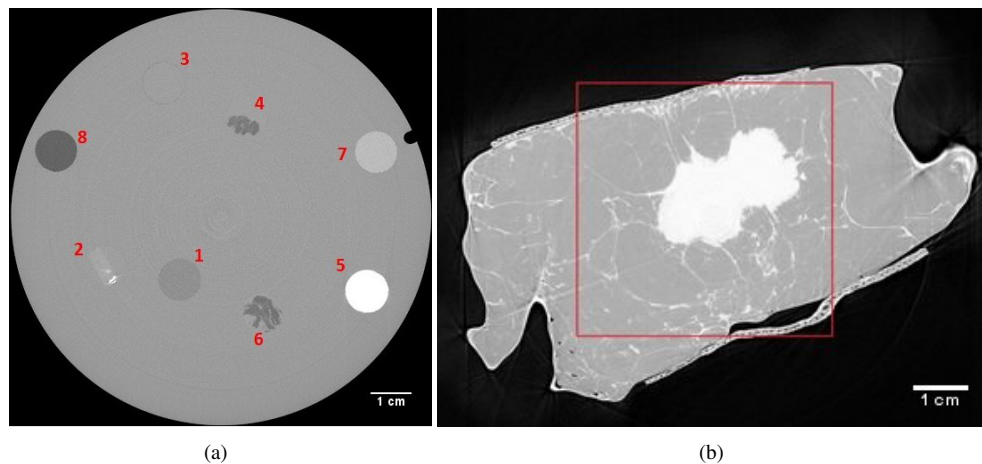

(a)   (b)

Fig. 1. (a) Polycarbonate phantom where 1 = Glycerol ($C_3H_8O_3$), 2 = Unknown tissue (Malignant), 3 = Water (H_2O), 4 = Fibrous tissue, 5 = Calcium Chloride ($CaCl_2$), 6 = Adipose tissue, 7 = Paraffin wax, 8 = Ethanol ($EtOH$); (b) Reference image for the mastectomy sample reconstructed with FBP algorithm and considering 3600 high statistic projections. The red square indicates the region-of-interest used for the image quality assessment (see Visualization 1).

2.3. Reconstruction workflows

In this manuscript reconstruction workflow refers to the entire process from the acquired projections to the final images, this means that it includes pre-processing steps, the reconstruction algorithm and post-processing steps. In all the considered workflows, before the actual reconstruction, the projections were flat-field corrected and normalized by considering the air windows on the left and right side of the image as a reference. Then a ring removal filter [22] was applied in order to better compensate the detector inhomogeneity [23, 24].

The considered CT reconstruction algorithms are: Filtered Back Projection (FBP) [25], Simultaneous Iterative Reconstruction Technique (SIRT) [26], Simultaneous Algebraic Reconstruction Technique (SART) [27], Conjugate Gradient Least Squares (CGLS) [26], Equally Sloped Tomography (EST) [28], Total Variation (TV) minimization [29] and an iterative FBP algorithm based on the image histogram updates [30]. In addition, the effect of in-line single distance phase retrieval was considered. The phase retrieval was performed prior to the actual reconstruction by processing each projection independently, in accordance with the Homogeneous Transport of Intensity (TIE-Hom) Algorithm [31], using $\delta/\beta = 1000$ (i.e. a value close to the theoretical value of the ratio of the imaginary part to the real decrement of the relative X-ray refractive index for glandular and adipose tissues) as implemented in X-TRACT software [32]. A post-processing of the reconstructed images was also considered based on simple spatial filtering over a neighborhood of $w \times w$ pixels, including the optimal Epanechnikov kernel [33] and the edge-preserving image denoising by the SUSAN method [34]. All these reconstruction algorithms and image filters are available in X-TRACT software which can be downloaded and verified on request (X-TRACT software, link to sign-up form available here: `http://www.ts-imaging.net/Services/SignUp.aspx`). A summary of the considered workflows is reported in Table 1 which also contains the abbreviations used throughout the manuscript.

Table 1. Considered reconstruction workflows.

Abbreviation	Phase retrieval	Reconstruction	Post-processing
FBP	no	FBP [25]	no
FBP-ITER	no	FBP-ITER [30]	no
SIRT	no	SIRT [26]	no
SART	no	SART [27]	no
CGLS	no	CGLS [26]	no
EST	no	EST [28]	no
phr FBP	yes	FBP [25]	no
phr FBP-ITER	yes	FBP-ITER [30]	no
phr FBP-ITER *Epan*17	yes	FBP-ITER [30]	Epanechikov [33] ($w = 17$)
phr FBP-ITER *Susan*5	yes	FBP-ITER [30]	Susan [34] ($w = 5$)
phr TV-MIN	yes	TV [29]	no
phr SIRT	yes	SIRT [26]	no
phr SART	yes	SART [27]	no
phr CGLS	yes	CGLS [26]	no
phr EST	yes	EST [28]	no

2.4. Quantitative evaluation of the images

The evaluation of the reconstructed images was carried out by means of *full-reference* and *no-reference* image quality assessment indexes [35]. In the full-reference approach, a complete reference image is assumed to be known. In this work, the reference image (Fig. 1) is the one

obtained by applying the FBP reconstruction algorithm and considering all the available projections acquired with high statistics. The underlying assumption is that, when a large number of high quality projections is available, the well-known and widely used FBP algorithm gives good results and therefore it can be assumed as the reference reconstruction workflow. The adopted indexes are: MSE - Mean Squared Error, SNR - Signal-to-Noise Ratio, UQI - Universal Quality Index [36], NQM - Noise Quality Measure [37] and SSIM - Structural Similarity Index [35]. These characteristics are typically evaluated within a representative sub-region of the image where the pixel intensity values can be considered approximately spatially stationary [38, 39]. While low values for the MSE index suggest a better image quality, higher values for SNR, UQI, NQM and SSIM are instead expected when a higher quality image is considered.

The dimensionless no-reference intrinsic quality characteristic Q_s which incorporates both the noise propagation and the spatial resolution properties of a linear system was computed according to the equation:

$$Q_s = \frac{SNR_{out}}{F_{in}^{1/2}\Delta x} \qquad (1)$$

where F_{in} is the mean value of the incident photon fluence (the number of incident photons per n-dimensional volume, in the current context, per unit area), Δx is the spatial resolution of the imaging system and $SNR_{out} = S_{out}/\sigma_{out}$ is the output signal-to-noise ratio (see details in [33], [16]). The Contrast-to-Noise ratio (CNR) and the Full Width Half Maximum (FWHM) no-reference indexes were also taken into account. The CNR was calculated by selecting two meaningful areas of the image containing two strongly different gray levels: a region inside the lesion and a region within the adipose tissue. Then, the following formula was applied:

$$CNR = A_{feature}^{1/2} \frac{|\langle\beta_{lesion}\rangle - \langle\beta_{adipose}\rangle|}{[(\sigma_{lesion}^2 + \sigma_{adipose}^2)/2]^{1/2}} \qquad (2)$$

where $A_{feature}^{1/2}$ is the area of the corresponding image feature measured in pixels, $\langle\beta_{lesion}\rangle$ and $\langle\beta_{adipose}\rangle$ are the average values of the imaginary part of the refractive index in the lesion and in the adipose tissue, respectively, and σ_{lesion}^2 and $\sigma_{adipose}^2$ are the variance of noise in the lesion and the adipose tissue, respectively [16]. Since the evaluation of FWHM requires a well defined image edge, this index was evaluated by considering the images of the polycarbonate phantom sample. The measure was carried out by taking ten line profiles (60 pixels length) at a polycarbonate - $CaCl_2$ detail interface. The profiles were then fitted with a non-linear sigmoid function. This function was then differentiated and the FWHM of the obtained Gaussian curve was evaluated (Fig. 2) [40]. The choice to evaluate edges quality on this particular detail was made because $CaCl_2$ is the only detail that presents sufficient differences in gray levels allowing the measure; further in some images reconstructed without phase retrieval it is the only visible detail. It has to be pointed out that the δ/β ratio, used for the phase-retrieved reconstructions, is not optimized on the $CaCl_2$ but on the glandular and adipose tissue, so the selected detail is heavily affected by the phase-retrieval blurring effect.

2.5. Radiological assessment

Three experienced radiologists and one pathologist with long experience in pathological correlation with radiological imaging expressed a blind opinion in terms of identification of the lesion borders and spiculations, visibility of small connectival residues included in the adipose tissue, perceived contrast and spatial resolution of each reconstructed image. They were asked to give a global score from 0, for the worst case, to 4, corresponding to the best image, i.e. the FBP reconstructed slice obtained from the full statistic dataset with 3600 projections (float
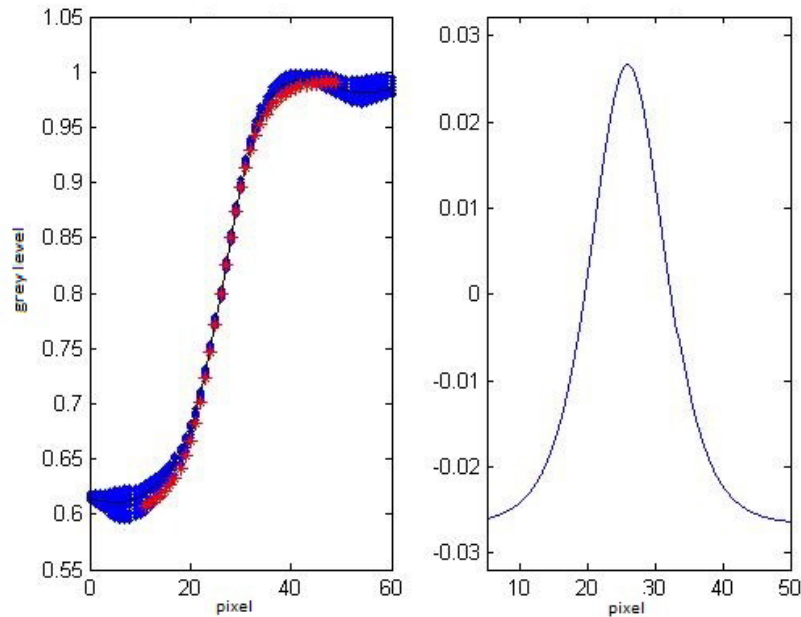
Fig. 2. Edges quality measurement method. On the left part, the figure shows ten individual profiles (blue dots) and the fitted sigmoid function (red dots) used for the evaluation; the right part displays the obtained Gaussian curve where FWHM was calculated.

values were allowed). After a first evaluation, they did a more detailed rating comparing groups of two or three images that previously received the same assessment.

Based on their opinion, the images were then classified into three categories: *no-diagnostic power* (mean radiological score from 0 to 2), i.e. it seems to be hard to diagnose the tumor; *poor diagnostic power* (mean radiological score from 2 to 3), i.e. it would be possible to diagnose the tumor but without an accurate evaluation of spiculations and/or connnectival residues existing in the tissue; *full diagnostic power* (mean radiological score higher than 3), i.e. all the relevant features are detectable and quantifiable.

3. Results and discussion

Figures 3, 4, 5 report some of the reconstructed images considered in this study. The images are grouped according to the three categories mentioned above.

Table 2 reports the computed values for the full-reference indexes when considering the image of the breast tissue sample after the application of all the considered reconstruction workflows. First, it has to be pointed out that the most widespread and well known reconstruction method based on the FBP algorithm is definitely outperformed by the other approaches. Indeed, the computed indexes suggest that the global quality of the images reconstructed with iterative algorithms is significantly higher than the one obtained with FBP. This is in line with the general consensus that when dealing with a limited number of noisy projections, algebraic techniques are better suited for the reconstruction of tomographic images. The benefit of the phase retrieval pre-reconstruction processing should be noted as well. When comparing the performance of the same reconstruction algorithm with and without the application of phase retrieval (e.g. SIRT vs phr SIRT), better MSE and SNR are revealed as well as a higher NQM
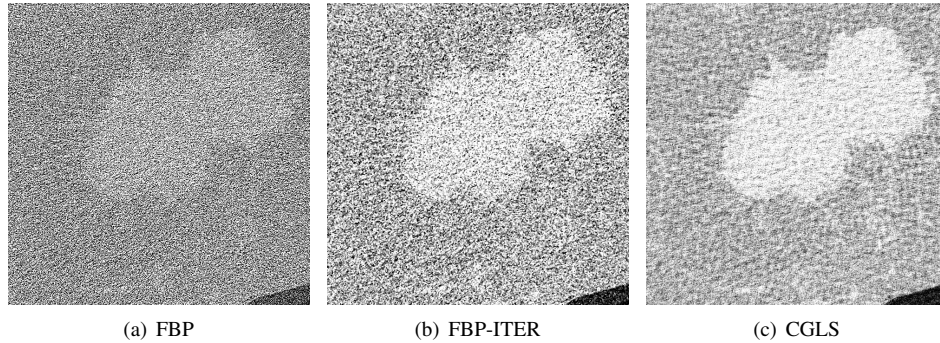
| (a) FBP | (b) FBP-ITER | (c) CGLS |

Fig. 3. *No-diagnostic power*. Due to the specific clinical context the main lesion is still noticeable but a global poor contrast to noise ratio can be perceived and connectival residues are not visible. For the high resolution images see the additional supplementary material available in Visualization 2, Visualization 3, Visualization 4, Visualization 5, Visualization 6, and Visualization 7.

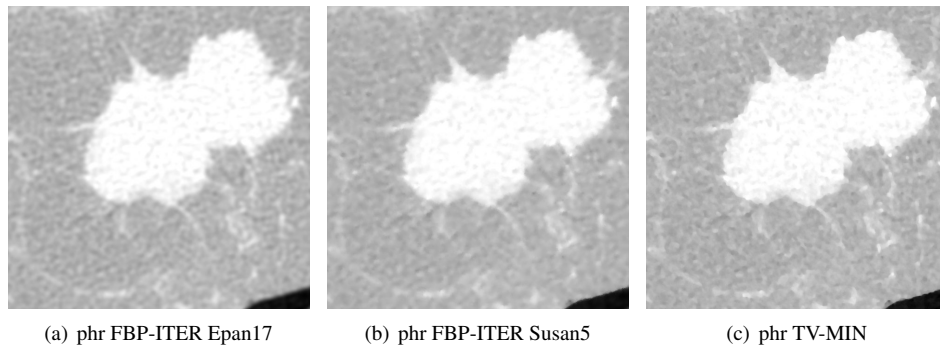| (a) phr FBP-ITER Epan17 | (b) phr FBP-ITER Susan5 | (c) phr TV-MIN |

Fig. 4. *Poor-diagnostic power*. Here the main lesion is well detectable thanks to the increased CNR, however the small connectival residues in the parenchimal area are yet not fully resolved, as well as the spiculations and the lesion borders. For the high resolution images see the additional supplementary material available in Visualization 8, Visualization 9, Visualization 10, Visualization 11, Visualization 12, and Visualization 13.

value. The structural similarity index suggests an increased image quality as well. On the other hand, the UQI index seems to be less sensitive to the benefits of phase retrieval. A lower UQI is recorded for e.g. the phr SART with respect to the SART reconstruction workflow while, instead, the MSE, SNR, NQM and SSIM suggest a better global image quality for the phr SART. Moreover, a similar situation is observed for the group of FBP-ITER methods. The MSE, SNR, NQM and SSIM indicate a better quality for this groups compared with the other reconstruction approaches when phase retrieval is considered. However, the UQI index does not seem to highlight the performances of this FBP-ITER based group. When adding post-processing filters to the FBP-ITER algorithm, the UQI index again seems unable to reveal differences that, on the other hand, MSE, SNR, NQM and SSIM seem to suggest.

The values of the no-reference quantitative metrics are reported in Table 3. Both the phantom and the tissue sample were considered to validate this approach. Since the the resultant trend was quite similar in both cases, for simplicity and experimental relevance, it was chosen to report only the numerical values computed for the tissue sample. For the sake of clarity, these
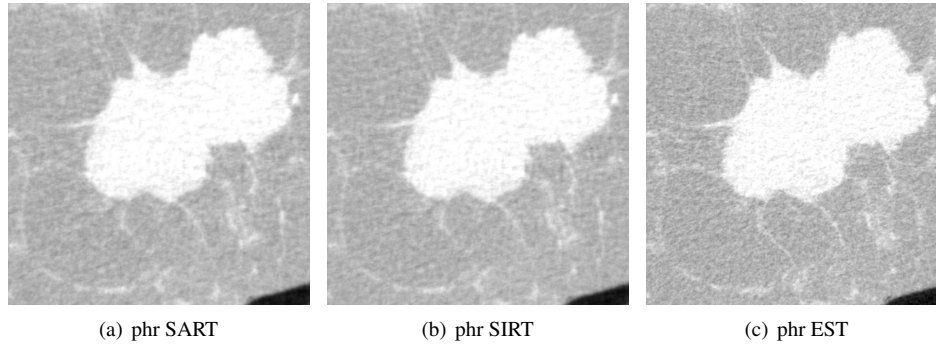
(a) phr SART   (b) phr SIRT   (c) phr EST

Fig. 5. *Full-diagnostic power*. The main lesion with all the relevant features is clearly visible and quantifiable.For the high resolution images see the additional supplementary material available in Visualization 14, Visualization 15, and Visualization 16.

Table 2. Numerical results of the full-reference quantitative image quality assessment.

	MSE	SNR [dB]	UQI	NQM	SSIM
FBP	$167.0E-07$	-2.91	0.0073	8.91	0.07
FBP-ITER	$21.1E-07$	6.09	0.0204	6.94	0.12
SIRT	$6.2E-07$	11.41	0.0403	10.00	0.28
SART	$4.7E-07$	12.54	0.0475	9.31	0.29
CGLS	$9.5E-07$	9.54	0.0338	9.19	0.25
EST	$20.1E-07$	6.29	0.0103	9.24	0.15
phr FBP-ITER	$2.1E-07$	16.17	0.0496	13.10	0.47
phr FBP-ITER *Epan*17	$1.6E-07$	17.25	0.0352	14.22	0.71
phr FBP-ITER *Susan*5	$1.5E-07$	17.57	0.0374	14.75	0.71
phr TV-MIN	$1.5E-07$	17.52	0.0419	14.29	0.69
phr FBP	$2.2E-07$	15.88	0.0403	12.91	0.51
phr SIRT	$1.7E-07$	16.92	0.0402	12.99	0.67
phr SART	$1.7E-07$	16.94	0.0408	12.82	0.67
phr CGLS	$1.8E-07$	16.88	0.0419	12.95	0.65
phr EST	$1.8E-07$	16.68	0.0361	12.98	0.61

metrics were also applied to the image reconstructed with the FBP algorithm and considering all the projections acquired at high statistics (reference image).

The two general trends that have been identified with the use of full-reference image quality indexes (i.e. the improvement of the quality as the consequence of using iterative CT reconstruction algorithms, compared to FBP, and also as a result of including the TIE-Hom phase retrieval into the reconstruction workflow), are repeated in the measured values of the no-reference quality index Q_s (Table 3). The negative effect of lower spatial resolution on the Q_s values can be seen for "phr FBP-ITER Epan17" method, where the use of a broad (17-pixel wide) Epanechnikov filter led to the deterioration of the spatial resolution and, as a result, to a lower Q_s score, even though the CNR was considerably higher than in the "phr FBP-ITER" method, i.e. without the post-processing with the Epanechnikov filter. On the other hand, the use of edge-preserving filters in the methods like "phr FBP-ITER Susan5" or "phr TV-MIN" did not lower the spatial resolution to the same degree and thus resulted in higher Q_s values.

As regards to the CNR and the FWHM values it has first to be reminded that they require a joint discussion since CNR solely can lead to misleading conclusions when computed after

e.g. post-processing the reconstructed image with a low pass filter. In fact, in this case, the the smoothing process produces an enhancement in contrast and noise but it negatively affects the spatial resolution. To better understand the duality of these two no-reference indexes, Fig. 6 reports in graphical form the relation between CNR' (contrast to noise ratio normalized by the selected area) and FWHM. The values for the reference image are also plotted to better show that an optimal reconstruction workflow should be as close as possible to the reference, i.e. it should present a low FWHM value and high CNR. The results for the FBP workflow are not represented in Fig. 6 because it was impossible to identify an edge in the reconstructed image to be used for the computation of the FWHM. This confirms what already underlined after the application of the full-reference indexes, i.e. the iterative approaches overcome the FBP when dealing with a reduced number of noisy projections. From both Table 3 and Fig. 6 it could be noticed that the already underlined benefit of the phase retrieval pre-processing is due mainly to an increased CNR but, as expected, a degradation of the edges quality, i.e. a lower value for the FWHM, is also noticeable in the images where phase retrieval has been applied. In order to better appreciate the effects of this pre-reconstruction processing, phase retrieval was applied also to the reference image and the no-reference indexes were assessed. The computed values are: $CNR' = 30.53$, FWHM = 5.82, $Q_s cancer = 62$ and $Q_s adipose = 38$. Again, a substantial improvement in the contrast with a slight reduction of the edges quality is noticeable. This confirms that the quality of the reconstructed image is altered mainly by the phase retrieval step and not by the specific reconstruction algorithm.

Table 3. Numerical results for the no-reference quantitative image quality assessment.

	CNR'	FWHM [pixel]	$Q_s cancer$	$Q_s adipose$
Reference image	8.88	2.03	18	12
FBP	0.20	-	1	0.5
FBP-ITER	0.67	11.47	2	1
SIRT	1.53	12.17	3	2
SART	1.89	13.56	4	3
CGLS	1.14	10.52	3	2
EST	0.73	8.48	2	1
phr FBP-ITER	4.62	11.01	8	5
phr FBP-ITER *Epan*17	15.18	12.88	5	4
phr FBP-ITER *Susan*5	16.47	12.67	17	13
phr TV-MIN	14.06	6.03	22	13
phr FBP	5.37	17.16	8	6
phr SIRT	12.58	16.35	13	10
phr SART	12.60	16.32	11	8
phr CGLS	10.58	16.43	14	10
phr EST	8.55	15.97	11	8

Table 4 reports the results of the subjective evaluation of the images performed by the expert radiologists and the pathologist. This subjective evaluation is fundamental to better understand if the benefits in terms of contrast resolution with the simultaneous reduction of edges quality induced by the phase retrieval are tolerated during the radiological evaluation process. The analysis of Table 4 with respect to Fig. 6 suggests that, at first, the observers focus the attention on the contrast resolution of the images. If an insufficient contrast is detected, a radiologist judges an image as meaningless for the diagnosis process. Beyond a level of contrast considered acceptable, the radiologists prefer the images having the highest edges quality (i.e. the lowest value of FWHM). Therefore, once this acceptable contrast level is identified, it seems
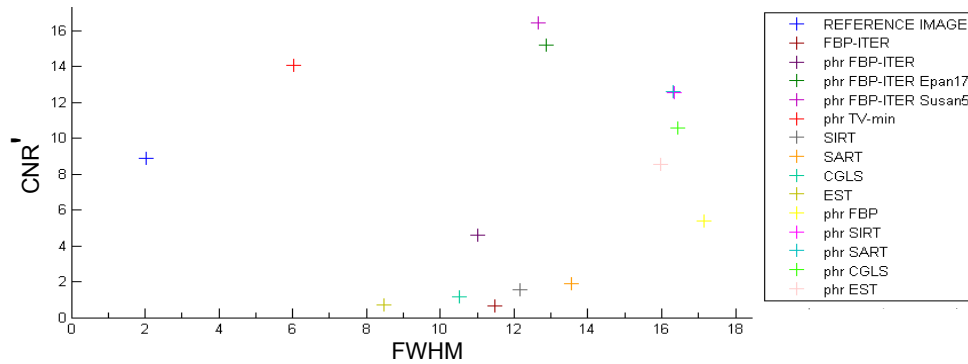
Fig. 6. Full Width Half Maximum with Contrast to Noise Ratio measured on polycarbonate phantom. CNR' measure is CNR normalized over the selected area.

that spatial resolution has to be privileged. This is because the diagnosis process for this specific application is based mainly on the morphology of the lesions in the tissue rather than contrast. According with this, phr FBP-ITER *Susan*5 and phr FBP-ITER *Epan*17 reconstructions work-flow should have been judged as *full diagnostic power* images. Nevertheless additional remarks have to be made about the texture of these reconstructed images. One of the involved radiologists commented about this and the subjective evaluation for these reconstruction workflows resulted strongly penalized by unrealistic image texture. Figure 7 reports the images reconstructed with phr TV-MIN (affected as well by the same problem) and phr EST, respectively. A different texture, and a consequently lower visibility of spicules, is observed after the application of the considered implementation of a TV minimization algorithm. Interestingly, none of the considered quantitative metrics seemed able to reveal this aspect. Further optimization of this work should include also a quantitative evaluation of textural features in the reconstructed images. Moreover, it seems that the radiologists are not concerned by a strong amount of noise
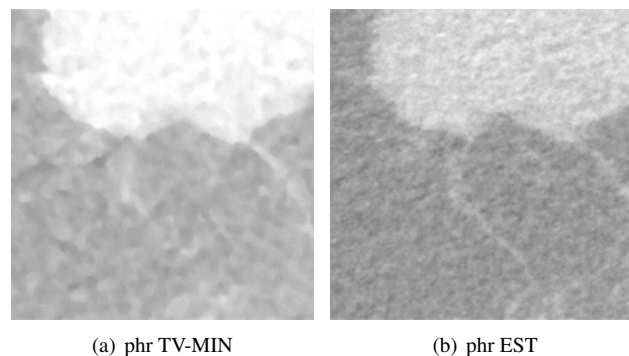

(a) phr TV-MIN   (b) phr EST

Fig. 7. Close-up of tissue sample slices pre-processed with phase retrieval and reconstructed with TV-Minimization (a) and EST (b) algorithm. In this small image portion the different textures are better noticeable.

in the reconstructed image. This could help the optimization of the reconstruction workflow since, in order not to sacrifice the spatial resolution, usually more noise can be tolerated.

With the ultimate aim of an unsupervised evaluation of breast CT images, an additional correlation between the subjective score proposed by the medical personnel and some of the proposed indexes is reported in Fig. 8. There is a general agreement between most of the eval-

Table 4. Qualitative assessment of the considered images performed by expert supervisors.

	Radiol. 1	Radiol. 2	Radiol. 3	Pathologist	Mean Score
FBP	1	1	1	1	1
FBP-ITER	1	1	1	0.5	0.87
SIRT	1.5	1	1.5	1.75	1.43
SART	1.5	1.25	1.5	1.75	1.5
CGLS	1	1.25	1.5	1.5	1.31
EST	1	1.25	1.5	1.5	1.31
phr FBP-ITER	3	3	2.5	2	2.62
phr FBP-ITER $Epan17$	2.5	3.25	2.5	2.5	2.68
phr FBP-ITER $Susan5$	2.5	3.25	3	2.5	2.81
phr TV-MIN	1.5	2.75	2.75	2.75	2.43
phr FBP	2.5	2.5	2.75	3.5	2.81
phr SIRT	3.25	3.25	3.5	3	3.25
phr SART	3.25	3.25	3	3	3.12
phr CGLS	2.5	3.25	3	3	2.93
phr EST	3.5	3.25	3.5	3.5	3.43

uated image quality scores and the radiological assessments, at least as far as the general trends are concerned. Overall, among all the evaluated objective image quality measures, the SSIM score (multiplied by the factor of 6) tracked the radiological assessments the closest.
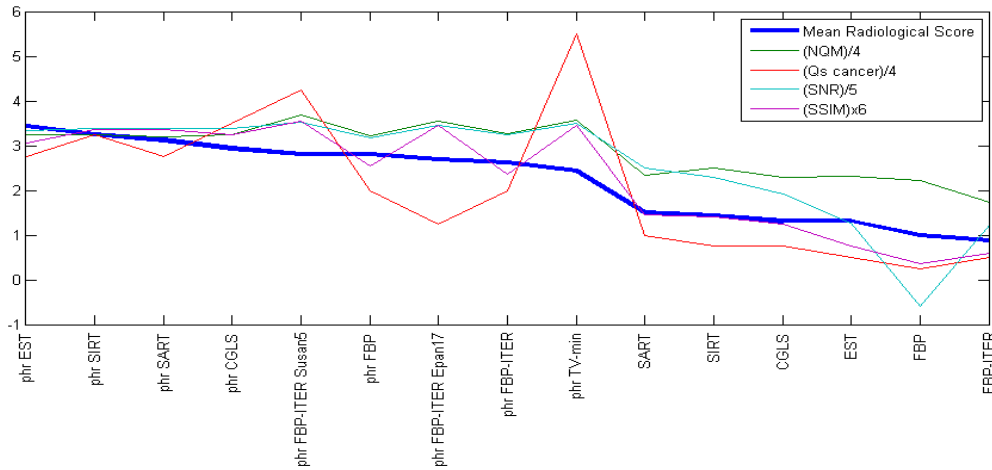

Fig. 8. Correlation between the radiological score and various image quality indexes.

It is worthy to underline that the considered mastectomy sample includes an opacity compatible with a malignant lesion of about 3 cm where irregular edges, blurred margins and spicules are strongly visible. Moreover the lesion is part of a convoluted breast where thin connectival trabeculae are clearly noticeable. When considering this specific diagnostic context, the radiologists confirm that a diagnosis is easy to perform. With the support of the anatomo-pathology unit, further activities will be focused in the search of a more challenging mastectomy sample where e.g. the excised tissue presents small satellite lesions and therefore a diagnosis could be less easy to perform. Together with the evaluation of additional reconstruction methods (e.g

dictionary learning [41, 42]), this will help a future refinement of the reconstruction workflows identified in this article as those producing images with full diagnostic power. This should be done in order to better identify the reconstruction workflow most suitable for an *in vivo* application of low-dose phase contrast breast CT.

Finally, in the framework of a clinical trial, further work will be dedicated to the application of the proposed methodology to other experimental setups, such as e.g. different sample-to-detector distances and X-ray energies. An optimization of the detector to be used is also necessary. The one adopted in the experiment presented in this article was chosen for its high spatial (an effective pixel size of 54 μm) and contrast resolution, nevertheless for CT *in-vivo* studies on patients, an optimized version with a faster readout will be required.

4. Conclusions

With the help from expert radiologists, it has been shown in the present paper that phase retrieval techniques and refined CT reconstruction methods significantly improve the image quality, thus allowing low dose phase-contrast CT to be effectively applied to mammography. It was also demonstrated that, at a dose comparable to conventional 2D planar radiography (1.5 mGy), it was possible to produce CT images with a high diagnostic value. This can become a crucial point in view of the introduction of breast tomography into large-scale screening programs. Further, the adoption of the phase retrieval increases the diagnostic power of the phase-contrast images, even if there is an intrinsic blurring effect affecting the spatial resolution. This paves the way to a new diagnostic tool for the clinical practice that takes advantage of 3D tomographic images of the breast.

Acknowledgments