

Methods for Flipping of Mammogram Images and Noise Removals

Prof. Samir Kumar Bandyopadhyay^{*,†}

Advisor to Chancellor, JIS University, India

DOI:<https://doi.org/10.15520/ijmhs.v9i4.2550>

Accepted 25 Mar 2019; Received 1 Mar 2019; Publish Online 28 Apr 2019

Reviewed By: Dr.
Daniel V.
Department: Medical

ABSTRACT

Cancer can be brain cancer, breast cancer, cervical cancer, etc. depending on the occurrence on the different parts of the body. to name a few. Cancer detected in later stage and it results into high mortality within few months. Early detection is the only hope for better prognosis and treatment. This paper proposed methods that digital mammogram image pair is initially pre-processed, the right oriented mammogram is flipped and any unwanted artifacts are removed from the image. Next noise removal is made using the Gaussian kernel.

Key words: Digital mammography–Computer Aided Diagnosis–Breast Region–Seed

1 INTRODUCTION

Digital mammography represents a technological advance in breast imaging. However, these images may contain artefacts. These artefacts include patient-related artefacts, hardware-related artefacts, detector-associated artefacts, collimator misalignment and underexposure and grid lines. Software processing artefacts like vertical processing bars, loss of edge and high-density artefacts can also be present. Although some of these artefacts are similar to those seen with screen-film mammography, many are unique to digital mammography. Such artefacts and noise in mammogram images is the major obstacle to develop fully automated Computer Aided Diagnosis (CAD) systems. This affects the result and accuracy of algorithms. Hence it is essential to perform preparation steps to suppress these artefacts and noise from background and enhance the breast region.

2 LITERATURE REVIEW

Prior to performing preparation process on digital mammogram it is essential to understand the different types of artefacts and noise which are present within the breast and non-breast region of digital mammogram. While the incidence of artefact on digital mammographic images are typically less than with film based mammography, artefacts can be produced on digital systems. Researchers [1, 2] have classified

the artefacts in “clinical digital mammogram”. Researchers also [4–5] reviewed all the artefacts in mammography encountered and classified the causes of these artefacts in a number of groups..

Researchers [3] described in detail what is known to improve image quality for digital mammography and make recommendations about how digital mammography should be performed to optimise the visualisation of breast cancers. Other researchers [6–10] have presented a study on “standardisation and comparison of image quality of Full-Field Digital Mammography (FFDM) versus screen-film mammography in a screening population”. However, they found that it was difficult to standardise and compare these.

Researchers also [11–18] classified the artefact especially on digital mammogram. According to them “some of these artefacts are similar to those seen with screen-film mammography, many are unique to digital mammography—specifically, those due to software processing errors or digital detector deficiencies. In addition, digital mammographic artefacts depend on detector technology (direct vs. indirect) and therefore can be vendor specific. All related personnel need familiar with the spectrum of digital mammographic artefacts and give careful attention to digital quality control procedures to ensure optimal image quality. [19]

Researcher [20] proposed one of the well-known artefact suppression algorithm based on “area morphology” to remove radiopaque artefacts from the background region of mammograms. Here, a comprehensive technique is proposed to suppress the unwanted artefact from the digital mammogram along with noise reduction to improve the quality of the image. At the same time images will be homogeneously oriented to meet the uniformity [21].

* Corresponding author.

† Email: 1954samir@gmail.com.

3 PROPOSED METHODS

Artefacts are common to digital mammograms. Recognition of these artefacts is critical for achieving optimal image quality. Digital mammography systems differ in the way they acquire, process, and display images, and artefacts can be a result of problems involving any one of these components. Many artefacts specific to screen-film mammography have been well documented and are recognisable on patient images. Some examples include dust artefacts, pickoff, processor roller artefacts, static artefacts, and fogging artefacts.

Artefacts may present at both screen-film mammography and digital mammography, such as patient-related artefacts and hardware-related artefacts. External artefacts are represented by tapes and other identification marks that are used on the mammogram to identify the patient and related data that may be required for identification of the mammogram. These markings provide high intensity regions on the mammogram and are inconsequential to the investigation of abnormalities within the mammogram. Presence of such artefacts also changes the intensity levels of the mammogram image significantly that may affect statistical analysis on the image. In this paper the algorithm proposed by me attempts to remove all such artefacts, markings on the non-breast region of the mammogram and replace them with the background colour. So image shows the breast region part only.

To achieve the desired goal a new algorithm has been introduced by the combination of modified seeded region growing with Thresholding. It has been observed that in MLO mammogram images, the breast portion is placed in the middle of the object irrespective of left and right breast. As per the characteristic feature of mammogram, breast region is represented by high intensity pixels, where as background consist of chest wall and skin to air are represented by ideally zero intensity or by very low intensity i.e. not more than 10 in grey scale images. The external artefacts and other irrelevant object are present in background, more specifically in the skin to air part of mammogram image.

The proposed algorithm is to search a seed from breast region of mammogram. It has been observed that the pixel that is located at height/2, width*3/4 may always lie in the breast region of mammogram with higher intensity value. Initially this pixel is considered as the seed of region growing algorithm. Otherwise, it will continue to search a new location given by height/2, width*5/8 and if necessary height/2, width/2. It is observed that in all the cases the seed will be obtained from these locations. It is shown in Figure 1.

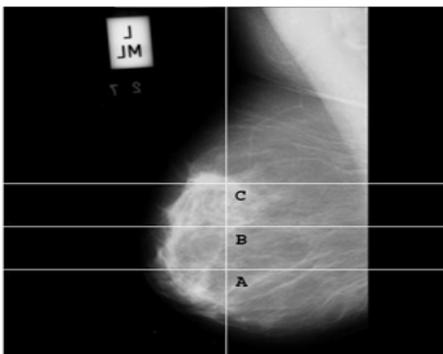


Figure 1. OriginalMammogram showing the preferred pixel location (A, B and C) to collect seed forregion growing algorithm

After finding the seed for a region, next objective is to extract only the breast region from the background, leaving the background with artefacts behind and copying the breast region on another blank image. The seed pixel is copied to the new image at the same location of the pixel in the original image. This seed pixel is being coloured black in

the original image. It is known that the breast regions of the mammogram have high intensity values whereas the background contains low or zero intensity. The algorithm starts searching for pixels that bounds the seed pixel. For each seed pixel the four boundary pixels located north, east, west and south of the pixel is also checked to find out whether they have high intensity value. Here algorithm uses the thresholding technique to divide the breast region pixel with background. If the pixels are with high intensity value, these will be used by the algorithm as seed for further searching. A stack is used to store the seeds to be investigated. The process continues by popping a seed from the stack and checking its intensity. If the seed is of high intensity value becomes the next seed. The value of intensity is copied to the new image at the same location of the pixel in the original image and the pixel is coloured black in the original image. This process continues till the stack is empty. The region grows from the single seed and stops when the entire breast region is blackened on the original image and the corresponding entire breast region is copied to the new image. The output image consists of only the breast region and remaining artefact are left behind in the original image. Now the external artefact free mammogram image will be the input for next processing.

Algorithm: Proposed Seeded Region Growing Artefact Removal Algorithm

SEEDED-REGION-GROWING (OrgImage, ImgWidth, ImgHeight)

$\Delta t \leftarrow 10$

If OrgImage[ImgHeight/2, ImgWidth*3/4]. Intensity > Δt

Then GROW-REGION (OrgImage, ImgHeight/2, ImgWidth*3/4, Δt)

Else If OrgImage[ImgHeight/2, ImgWidth*5/8]. Intensity > Δt

Then GROW-REGION (OrgImage, ImgHeight/2, ImgWidth*5/8, Δt)

Else If OrgImage[ImgHeight/2, ImgWidth/2]. Intensity > Δt

Then GROW-REGION (OrgImage, ImgHeight/2, ImgWidth/2, Δt)

Else Error

Return

GROW-REGION (OrgImage, h, w, Δt) Stack \leftarrow New Empty Stack
NewImage \leftarrow New Blank Image Stack.Push (h)

```

Stack.Push (w)
While Stack ≠ Empty
Do x ← stack.Pop () y ← stack.Pop ()
GreyValue ← OrgImage[x,y].Intensity
If GreyValue > Δt
Then NewImage[x, y].Intensity ← GreyValue OrgImage[x, y].Intensity ← 0
If x-1 > 0 AND OrgImage[x-1, y].Intensity > Δt
Then Stack.Push (y)
Stack.Push (x-1)
If x+1 < OrgImage.height AND OrgImage[x+1, y].Intensity > Δt
Then Stack.Push (y)
Stack.Push (x+1)
If y-1 > 0 AND OrgImage[x, y-1].Intensity > Δt
Then Stack.Push (y-1)
Stack.Push (x)
If y+1 < OrgImage.width AND OrgImage[x, y+1].Intensity > Δt
Then Stack.Push (y+1) Stack.Push (x)
Return (NewImage)

```

Here the image size is $N \times N$. But there is no iteration in the algorithm to read the entire image. So, complexity will never exceed n^2 . The pixels are traversed directly using x and y value. The numbers of pixels accessed on the image depend on the area covered by the breast region. By observation, it is found that half of the mammogram image is covered by the high intensity breast region. So, it may be said that half of the pixels out of $N \times N$ pixel are traversed by the algorithm. Hence the average time complexity of the algorithm will be approximately $n^2/2$. Figure 2 shows original Mammogram without external artefact.

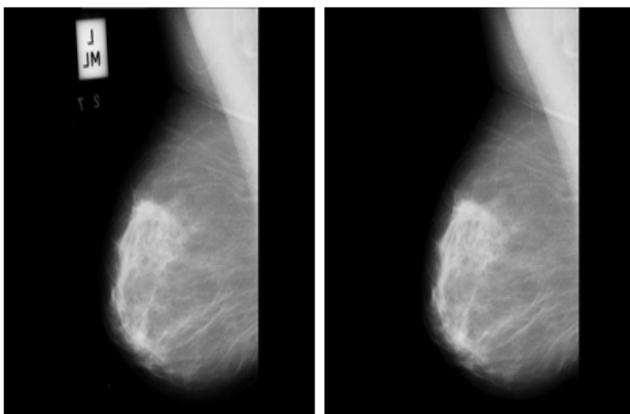


Figure 2. Original Mammogram and Mammogram without external artefact after SRGA

One of the major conditions of any system for detection of breast cancer is to standardise the input. The side of the image containing the pectoral muscle is on the upper left corner of the image in MLO view after transformation of image. So at this phase it is needed to identify the entire breast mammograms that have a left orientation, which is desirable from the point of view of execution of proposed methods. These left breast mammograms are represented by their chest wall on the left side and the pectoral muscles is on the top left corner of the mammogram. The breast boundary and nipple are on the right side. The right breast mammogram need to be turned horizontally at 180°. So it is an exact mirror reflection of the image. So the image now obtained after flipping is of similar orientation to the left breast mammogram. This process will allow the breast regions to be compared and analysed in a similar way by applying same automated algorithms for both the mammogram images. This is especially significant when two mammograms of the same pair need to be registered for determining the symmetry between a pair of mammograms.

Identification of the mammogram depends on image orientation whether it represents the left or right breast region. In case of left, algorithm will ignore the orientation process as it is already in the desired orientation. If the mammogram is the right breast, it is needed to be flipped horizontally. To identify the orientation, algorithm scans the pixel intensities from left to right within a row of pixel intensities, at fixed intervals of rows (Δd). The background of the mammogram is represented by very low intensity but the breast regions have higher intensities. As soon as, it finds a high intensity value, it stops scanning. On performing such scans on a number of rows at fixed interval the algorithm tries to find out the column value (K) that represents the breast region. If it falls on a straight line it can be safely concluded that it is the chest wall and the orientation is left. So, it does not require further checking. If the subsequent column values ($\neq K$) do not fall in a straight line, then algorithm can conclude that it represents the skin air interface of the breast region and it is right side up. Such mammograms will be required to be flipped.

The process of Image flipping can be done with any general purpose image processing software but in this research it has been done automatically since the objective is to develop a fully automated Computer Aided Detection (CAD) system. The process involves the scanning of the image and copying the mirror reflection of the image pixels on another image. The process starts by scanning the pixel intensities of the image and then copying the pixel intensity to the resulting image exactly at position that is obtained by subtracting the position from the width of the image and it continues for all the pixels of each row and the subsequent rows till the last pixel of the last row are copied.

Algorithm: Homogeneous Orientation using Image Orientation Algorithm

```

HOMO-ORIENTATION (OrgImage, Height, Width)
ChestWall ← 0
Δd ← Constant
Loop i ← 0 to Height

```

```

Do Loop j ← 0 to Width
Do If OrgImage[i, j] > 0
Then If i = 0
Then ChestWall ← j
If ChestWall ≠ j
Then HORZ-FLIP (OrgImage, Height, Width)
Return
Break
i ← i + Δd
Return

HORZ-FLIP (OrgImage, Width, Height)
NewImage ← New Blank Image
Loop i ← 0 to Height
Do Loop j ← 0 to Width
Do NewImage[i, j].Intensity ← OrgImage[i, Width - (j
+ 1)].Intensity
Return (NewImage)
    
```

Here the image size is $N \times N$. The proposed method is consisting of two parts; initial part is to detect orientation and latter part is only applicable for right breast. By observation, it may be said that the chest wall part approximately occupies $\frac{1}{4}$ part of the width of mammogram image. Initial process will start scanning from $OrgImage[0, j]$ and terminated to Chest Wall i.e. $OrgImage[N/4, j]$. So, $N/4$ numbers of pixels are scanned. Further the algorithm scans the pixel intensities at fixed intervals of rows (Δd). So, it $N/\Delta d$ number of row are scanned. Finally, it may be concluded that time complexity of the initial part of the method is $(N/4) \times (N/\Delta d)$ i.e. $n^2/(4 \times \Delta d)$. The later optional part, will require $N \times N$ processing and generate time complexity of n^2 . If the image is consisting of left breast, then it will be much faster. Figure 3 shows Mammogram after Flipping.

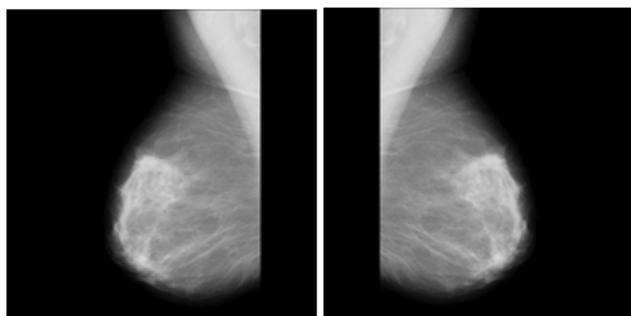


Figure 3. Right Mammogram and Mammogram after Flipping 180°

There are different types of noises, which appear in digital mammogram images. High intensity noises are embedded to the breast region of the mammogram thus resulting in loss of information from the breast region. These noises also make detection process for an automated CAD process to yield false results or negative detection. Such noise must be removed from the image to provide accurate results in the detection processes. In this research the well-known Gaussian filter is used to remove such noise by blurring these

noises before performing edge detection or other processing on the mammogram images. For 2-D, an isotropic (i.e. circularly symmetric) Gaussian form is used for the proposed method. Once a suitable kernel is obtained then the Gaussian smoothing can be performed using standard convolution methods. Figure 4 shows the kernel size.

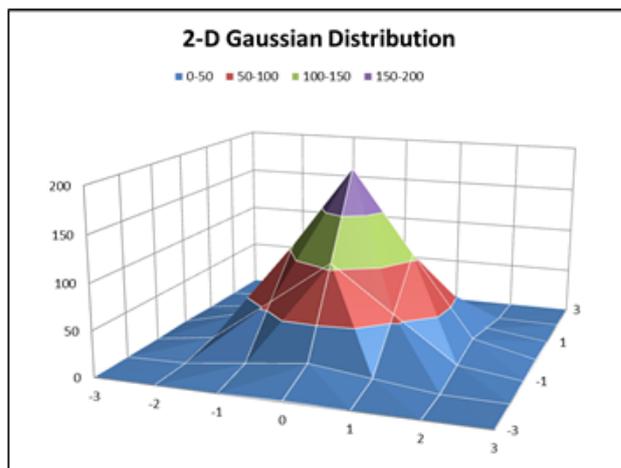


Figure 4. Proposed Method uses 7×7 kernel as a Convolution Filter

In this proposed method, 7×7 kernel has been taken as a convolution filter. The mammogram images are broadly categorized into 3 categories namely Fatty, Fatty- Glandular and Dense-Glandular depending on the density of fatty tissues and abundance of glands in the breast. Each category of mammogram displayed a varied intensity value which is distinct for each category. This property of the mammograms has helped the choice of value of deviation (Ω) for each category, thus able to adjust the level of smoothing for each category. Figure 5 shows Mammogram before and after Gaussian Smoothing.

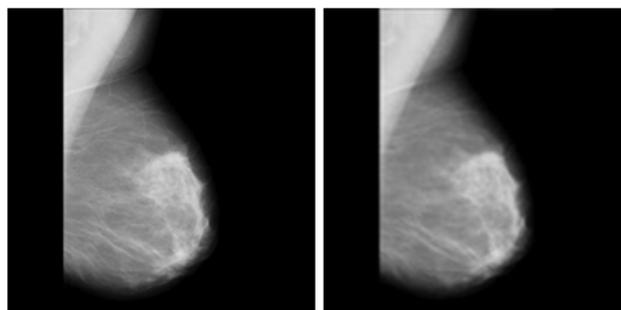


Figure 5. Mammogram before and after Gaussian Smoothing

Experimental Results

The algorithms have been tested with several mammographic images including on all mammograms from MIAS mammogram database and other available databases containing normal and abnormal cases. Almost all cases output is as expectation. Some of the appropriate test results are

depicted here with the mammograms taken from the MIAS database to prove the accuracy of the algorithm. Figure 6 to Figures 7, 8 and 9 shows the result.

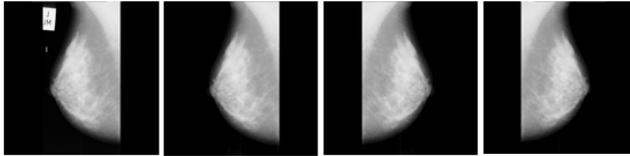


Figure 6. Images showing Original Mammogram followed by image after Artefacts Removable, Flipping and Noise Removable (Images from left to right)

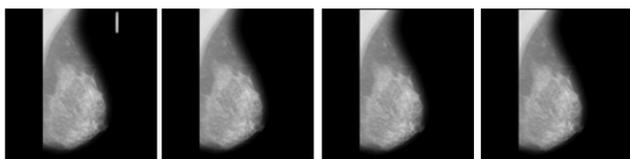


Figure 7. Images showing Original Mammogram followed by image after Artefacts Removable, Flipping (no change) and Noise Removable (Images from left to right)

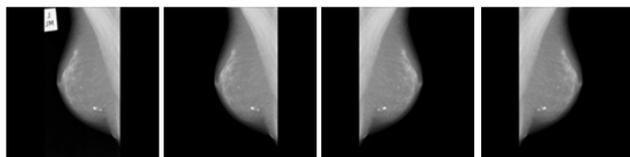


Figure 8. Images showing Original Mammogram followed by image after Artefacts Removable, Flipping and Noise Removable (Images from left to right)

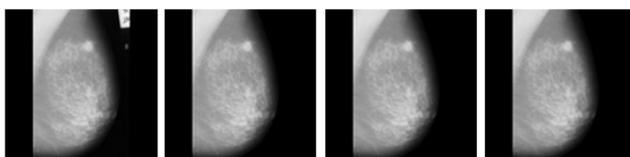


Figure 9. Images showing Original Mammogram followed by image after Artefacts Removable, Flipping (no change) and Noise Removable (Images from left to right)

Quantitative Analysis

The preparation includes three broad areas namely, artefact removal, flipping of right sided breast and noise elimination. All the three process is done automatically by the system without any user intervention which is a prerequisite for any real-time system. Very few authors included this vital step of preparation in their dissertations. Noise elimination has been done with standard Gaussian kernel. Flipping method has been successfully done by the proposed method except for 5 images in MIAS database where there is operator induced errors (refer to Figure 10) are present. The Accuracy of the proposed method is 100%. It can be noted that flipping may not be relevant to other methods

proposed by different authors but for the newly proposed method, it is of vital importance as registration of mammogram pairs depends on the success of flipping.

Flipping failed in these few cases are due to the presence of some shadow or vertical high intensity line, as noise on one of the sides of the mammogram. This line mimics the chest walls so the algorithm fails to distinguish a right breast mammogram, falsely interprets it to be a left mammogram and it does not perform flipping.

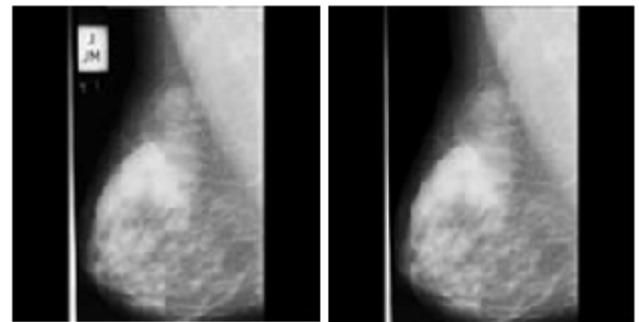


Figure 10. Original Mammogram and Image showing failure of Flipping Algorithm due to presence of vertical strip which is also overlapped with breast region

Artefact removal method is implemented on 322 mammograms but showed failure in 8 cases and falsely classified two mammograms of artefact where there was none present. A detailed statistical analysis has been performed using Receiver operating characteristic (ROC) analysis. ROC methodology is a popular method for comparing the performances of two or more imaging modalities. ROC is a binary concept where the object is present or not present and the result is in binary. The resulting 2 x 2 truth-response table defines positive decisions (true positives, true negatives) and negative decisions (false positives and false negatives). Using the obtained ROC quantities one can define True Positive Fraction (TPF), False Positive Fraction (FPF) and resultant ROC curve.

The findings of ROC curve in total number of cases 322, number of correct cases is 312 with Accuracy of 96.9%, Sensitivity of 99.1% and Specificity value of 91.9%. Total positive cases missed is 2 and negative cases missed is 8. The Empiric ROC curve area enclosed is 0.955. The Empirical ROC curve is shown in Figure 11.

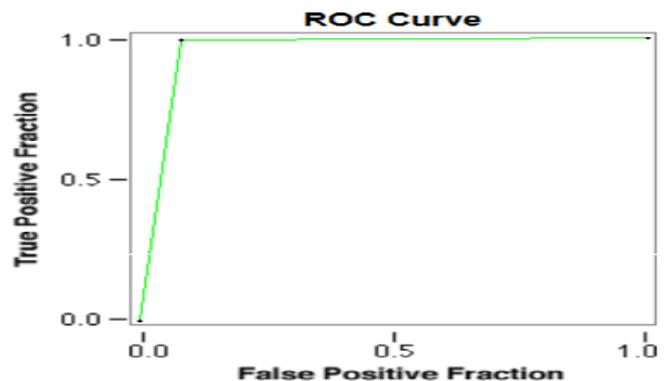


Figure 11. Empirical ROC Curve for Artefact Removal

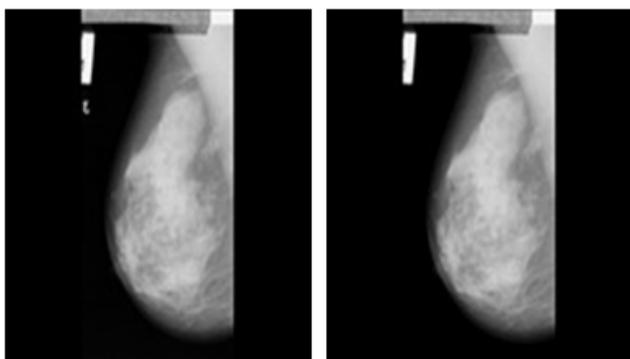


Figure 12. Original Mammogram and Image showing failure of Artefact Removal Algorithm due to artefact is overlapped with breast region

The main reason for failures (as shown in Figure 12) are mostly due to the artefact being embedded or linked to the breast region of the mammogram. The algorithms fail to distinguish between the breast and non-breast artefact. Such cases are few in number but their occurrence is due to the failure of the technicians or defective equipment.

4 CONCLUSION

In this proposed method, three distinct algorithms are used and the combinations of these have provided elimination of different artefacts, making homogeneous orientation and excellent removal of noise from mammogram images. SRGA algorithm removes the external artefacts in most cases. Next is changing the orientation of mammogram image to get uniform mammogram image for both left and right pairs of mammogram. Finally, Gaussian smoothening is used to remove noise that is internal to the breast region. The output of this preparation processing are the mammograms that are free from most of the artefacts and noise; can be used for other medical image processing applications and further studies on mammogram for detection of abnormalities.

REFERENCES

- [1] Abdel-Mottaleb et al, "Locating the Boundary between the Breast Skin Edge and the Background; 1996. p. 467–470.
- [2] About.com: Breast Cancer Open Surgical Breast Biopsy by Pam Stephan, Health's Disease and Condition content is reviewed by the Medical Review Board, About.com Guide Updated 13; 2008.
- [3] Albregtsen et al, "Adaptive grey level run length features from class distance matrices". vol. 3; 2000.
- [4] American Cancer Society, "Breast Cancer Facts & Figures". Atlanta: American Cancer Society; 2009.
- [5] American Cancer Society, "Cancer Facts & Figures". Atlanta: American Cancer Society; 2012.
- [6] of Radiology, "Illustrated Breast Imaging Reporting and Data System (BI-RADS)". American College. 1998;.
- [7] Antoniou et al, "Segmentation of the pectoral muscle edge on mammograms by tunable parametric edge detection. Scientific and Engineering Society WSES Press. 2001;p. 55–60.
- [8] Welfare, "Breast Screen Australia Achievement Report 1997-1998". Australian Institute of Health and. 2000;.
- [9] Aylward et al, "Mixture modeling for digital mammogram display and analysis", Digital Mammography. Computational Imaging and Vision. 1998;13:305–312.
- [10] Baheerathan et al, "New texture features based on complexity curve". Pattern Recognition. 1999;32:605–618.
- [11] Bahlmann. Automated Detection of Diagnostically Relevant Regions in H&E Stained Digital Pathology Slides; 2012. .
- [12] Bakic et al, "Effect of breast compression on registration of successive mammograms". Springer; 2004.
- [13] Balachandran et al, "Cancer - an ayurvedic perspective". Pharmacological Research, Vol. 2005;51:19–30.
- [14] Ball. Digital Mammogram Spiculated Mass Detection and Spicule Segmentation using Level Sets; 2007. p. 4979–4984.
- [15] Ball. Digital Mammographic Computer Aided Diagnosis (CAD) using Adaptive Level Set Segmentation; 2007. p. 4973–4978.
- [16] Bamford et al, "A water immersion algorithm for cytological image segmentation"; 1996. p. 75–79.
- [17] Bankman et al, "Segmentation Algorithms for Detecting Microcalcifications in Mammograms". Institute of Electrical and Electronics Engineers Transactions on Information Technology in Biomedicine. 1997;1:2–141.
- [18] Bassett et al, "Breast sonography: technique equipment and normal anatomy". vol. 10; 1989.
- [19] Bawa, "Edge Based Region Growing". Department of Electronics and communication Engineering, Thapar Institute of Engineering & Technology (Deemed. 2006 6;.
- [20] Beaulieu et al, "A hierarchy research article in picture segmentation: a stepwise optimisation approach". vol. 11; 1989.
- [21] Belhomme et al, "Towards a computer aided diagnosis system dedicated to virtual microscopy based on stereology sampling and diffusion maps". Diagnostic Pathology. 2011;6(3):1–4.

AUTHOR BIOGRAPHY

Prof. Samir Kumar Bandyopadhyay Advisor to Chancellor, JIS University, India