Computer-aided detection of subtle signs of early breast cancer: Detection of architectural distortion in mammograms

Rangaraj M. Rangayyan

Department of Electrical and Computer Engineering, University of Calgary, Calgary, Alberta, CANADA
Breast cancer statistics

- Canada: lifetime probability of developing breast cancer is one in 8.8
- Canada: lifetime probability of death due to breast cancer is one in 27
- Prevalence: 1% of all women living with the disease
- Screening mammography has been shown to reduce mortality rates by 30% to 70%
Mammography

Signs of Breast Cancer:

- Masses
- Calcifications
- Bilateral asymmetry
- Architectural distortion
  (often missed)
Breast cancer causes a desmoplastic reaction in breast tissue

A mass is observed as a bright, hyper-dense object
Calcification

Deposits of calcium in breast tissue
Bilateral asymmetry

Differences in the overall appearance of one breast with reference to the other
Computer-aided diagnosis

- Increased number of cancers detected by 19.5%
- Increased early-stage malignancies detected from 73% to 78%
- Recall rate increased from 6.5% to 7.7%
- 50% of the cases of architectural distortion missed

1 (Freer and Ulissey, 2001)  2 (Baker et al., 2003)
Architectural distortion

- Third most common mammographic sign of nonpalpable breast cancer
- The normal architecture of the breast is distorted
- No definite mass visible
- Spiculations radiating from a point
- Focal retraction or distortion at the edge of the parenchyma
Architectural distortion

- Spiculated
- Focal retraction
- Incipient mass
Normal vs. architectural distortion
Normal vs. architectural distortion
Detection of architectural distortion

1. Extract the orientation field
2. Filter and downsample the orientation field
3. Analyze orientation field using phase portraits
4. Post-process the phase portrait maps
5. Detect sites of architectural distortion
Gabor filter

\[ g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left[ -\frac{1}{2} \left( \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right] \cos(2\pi fx) \]

**Design parameters**

- line thickness \( \tau \)
- elongation \( l \)
- orientation \( \theta \)

**Gabor parameters**

\[ f = \frac{1}{\tau}; \quad \sigma_x = \frac{\tau}{2\sqrt{2\ln 2}} \]

\[ \sigma_y = l\sigma_x; \quad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \]
Design of Gabor filters

\[ l = l_0 \]
\[ \tau = \tau_0 \]
\[ \theta = \theta_0 \]

\[ l > l_0 \]
\[ \tau = \tau_0 \]
\[ \theta = \theta_0 \]

\[ l = l_0 \]
\[ \tau > \tau_0 \]
\[ \theta = \theta_0 \]

\[ l = l_0 \]
\[ \tau = \tau_0 \]
\[ \theta > \theta_0 \]
Extracting the orientation field

Compute the texture orientation (angle) for each pixel

Gabor filtering (line detection)
Phase portraits

\[ \vec{v}(x, y) = \begin{pmatrix} v_x \\ v_y \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} + b \]

node    saddle    spiral
Texture analysis using phase portraits

Fit phase portrait model to the analysis window

\[
A = \begin{bmatrix}
1.1 & 0.3 \\
-0.2 & 1.7 \\
\end{bmatrix}
\]

\[
b = \begin{bmatrix}
-4.8 \\
-7.9 \\
\end{bmatrix}
\]
Texture analysis using phase portraits

Cast a vote at the fixed point in the corresponding phase portrait map

Orientation field

Node  Saddle  Spiral
Detection of architectural distortion
Initial results of detection (2004)

- Test dataset: 19 mammograms with architectural distortion (MIAS database)
- Sensitivity: 84%
- 18 false positives per image
Reduction of false positives
Rejection of confounding structures

- Confounding structures include
  - Edges of vessels
  - Intersections of vessels
  - Edge of the pectoral muscle
  - Edge of the fibro-glandular disk

“Curvilinear Structures”
Nonmaximal suppression

ROI with a vessel

Gabor magnitude output

Output of nonmaximal suppression (NMS)
Rejection of confounding CLS

Angle from the orientation field and direction perpendicular to the gradient vector differ by $< 30^\circ$
Improved detection of sites of architectural distortion

Node map (without CLS analysis)

Node map (with CLS analysis)
FROC analysis (2005)

With CLS analysis

Without CLS analysis

sensitivity (%)

False positives per image
# Effect of conditioning number of matrix $A$ on the orientation field

<table>
<thead>
<tr>
<th>Example</th>
<th>Matrix $A$</th>
<th>Eigenvalues</th>
<th>Angle between principal axes</th>
<th>Conditioning number</th>
<th>Orientation field</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$\begin{bmatrix} 1 &amp; 0 \ 0 &amp; 3 \end{bmatrix}$</td>
<td>$\lambda_1 = 1$ \hspace{1cm} $\lambda_2 = 3$</td>
<td>90°</td>
<td>3</td>
<td><img src="image" alt="Orientation field for A" /></td>
</tr>
<tr>
<td>B</td>
<td>$\begin{bmatrix} 1 &amp; 7.46 \ 0 &amp; 3 \end{bmatrix}$</td>
<td>$\lambda_1 = 1$ \hspace{1cm} $\lambda_2 = 3$</td>
<td>15°</td>
<td>21.85</td>
<td><img src="image" alt="Orientation field for B" /></td>
</tr>
<tr>
<td>C</td>
<td>$\begin{bmatrix} 1 &amp; 0 \ 0 &amp; 20 \end{bmatrix}$</td>
<td>$\lambda_1 = 1$ \hspace{1cm} $\lambda_2 = 20$</td>
<td>90°</td>
<td>20</td>
<td><img src="image" alt="Orientation field for C" /></td>
</tr>
</tbody>
</table>
Results (2006)

• 19 cases of architectural distortion
• 41 normal control mammograms (MIAS)

• Symmetric matrix $A$: node and saddle only
• Conditioning number of $A > 3$: reject result

• Sensitivity: 84% at 4.5 false positives/image

• Sensitivity: 95% at 9.9 false positives/image
Prior mammograms

Detection mammogram 1997

Prior mammogram 1996
Prior mammograms

Detection mammogram 1997

Prior mammogram 1996
Prior mammograms

Detection mammogram 1997

Prior mammogram 1996
Interval cancer

- Indicates a case where breast cancer was detected outside the screening program in the interval between scheduled screening sessions.

- “Detection Mammograms” were not available.
Dataset

- 106 prior mammographic images of 56 individuals diagnosed with breast cancer (interval-cancer cases).
  - Time interval between prior and detection (33 cases)-average: 15 months, standard deviation: 7 months, minimum: 1 month, maximum: 24 months.
- 52 prior mammographic images of 13 normal individuals.
- Normal control cases selected represent the penultimate screening visits at the time of preparation of the database.
Interval cancer: site of architectural distortion

Mammogram

Gabor Magnitude
Interval cancer: site of architectural distortion

Orientation Field
Site of architectural distortion

Mammogram

Orientation field

Gabor magnitude

Node map
Interval cancer: potential sites of architectural distortion

Node Map  Automatically Detected ROIs
Examples of detected ROIs

True-positive

False-positive
# Automatically detected ROIs

<table>
<thead>
<tr>
<th>Data Set</th>
<th>No. of Images</th>
<th>No. of ROIs 128 x 128 pixels at 200 μm/pixel</th>
<th>No. of True-Positive ROIs</th>
<th>No. of False-Positive ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior mammograms of 56 interval-cancer cases</td>
<td>106</td>
<td>2821</td>
<td>301</td>
<td>2520</td>
</tr>
<tr>
<td>Prior mammograms of 13 normal cases</td>
<td>52</td>
<td>1403</td>
<td>0</td>
<td>1403</td>
</tr>
<tr>
<td>Total</td>
<td>158</td>
<td>4224</td>
<td>301</td>
<td>3923</td>
</tr>
</tbody>
</table>

*Notes:*
- ROI: Region of Interest
- Images are 128 x 128 pixels at 200 μm/pixel.
Feature extraction from ROIs

Potential Sites of Architectural Distortion

- Phase Portrait Analysis (Node value)
- Fractal Analysis (Fractal Dimension)
- Analysis of Angular Spread of Power
- Statistical Analysis of Texture (Haralick)
- Structural Analysis of Texture (Laws)

Feature Selection, Pattern Classification

Classification of ROIs
Fractal and spectral analysis

A TP ROI, $s(x, y)$

Fourier power spectrum, $S(u, v)$

Power spectrum in polar coords, $S(f, \theta)$

Angular spread of power, $S(\theta)$

Radial frequency spectrum, $S(f)$
Laws’ texture energy measures

- Operators of length five pixels may be generated by convolving the basic L3, E3, and S3 operators:
  
  $L_5 = L_3 \ast L_3 = [1, 4, 6, 4, 1]$  (local average)
  $E_5 = L_3 \ast E_3 = [-1, -2, 0, 2, 1]$  (edges)
  $S_5 = -E_3 \ast E_3 = [-1, 0, 2, 0, -1]$  (spots)
  $R_5 = -S_3 \ast S_3 = [1, -4, 6, -4, 1]$  (ripples)
  $W_5 = -E_3 \ast S_3 = [-1, 2, 0, -2, 1]$  (waves)

- 2D 5×5 convolution operators:
  
  $L_5L_5 = L_5^T L_5$
  $W_5W_5 = W_5^T W_5$
  $R_5R_5 = R_5^T R_5$  etc.
Results of Laws’ operators
Laws’ texture energy

Sum of the absolute values in a 15×15 sliding window
Geometrical transformation for Laws’ feature extraction
Analysis of angular spread: TP ROI

- Frequency domain
- Gabor magnitude
- Gabor orientation
- Coherence
- Orientation strength
Analysis of angular spread: FP ROI

- Frequency domain
- Gabor magnitude
- Gabor orientation
- Coherence
- Orientation strength
Receiver operating characteristics with selected features

<table>
<thead>
<tr>
<th>Classifiers</th>
<th>AUC using the selected features with stepwise logistic regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLDA (Leave-one-ROI-out)</td>
<td>0.75</td>
</tr>
<tr>
<td>Bayesian (Leave-one-ROI-out)</td>
<td>0.76</td>
</tr>
<tr>
<td>SLFF-NN (Single-layer feed forward: tangent-sigmoid)</td>
<td>0.78</td>
</tr>
<tr>
<td>SLFF-NN* (Single-layer feed forward: tangent-sigmoid)</td>
<td><strong>0.78 ± 0.02</strong></td>
</tr>
</tbody>
</table>

* 2-fold random subsampling, repeated 100 times
Sensitivity =

80% at 5.8 FP/image
90% at 8.1 FP/image

with the selected features based on stepwise logistic regression and using the Bayesian classifier and the leave-one-image out method
Bayesian ranking of ROIs: unsuccessful case
Bayesian ranking of ROIs: successful detection
Characterization of Dispersion

The methods are based upon analysis of spicularity and angular dispersion caused by architectural distortion.

- Index of convergence of spicules (ICS)

\[
ICS = \sum_{i=1}^{P} \sum_{j=1}^{Q} M(i, j) | \cos[\theta(i, j) - \alpha(i, j)] |
\]

- \( P \times Q \): size of the ROI
- \( \theta(i, j) \): Gabor angle response within the range \([-89^\circ, 90^\circ]\)
- \( M(i, j) \): Gabor magnitude or coherence value
- \( \alpha(i, j) \): angle of a pixel with respect to the horizontal toward the center of ROI, in the range \([-89^\circ, 90^\circ]\)
ICS quantifies the degree of alignment of each pixel toward the center of the ROI weighted by the Gabor magnitude or coherence value.
Radially Weighted Difference

\[ RWD = \sum_{p=1}^{PQ} \sum_{q=1}^{PQ} |I_p - I_q||r_p - r_q| \]

- \( I \): attribute value (intensity or magnitude)
- \( r \): radial distance from the center of the ROI
- \( \alpha(i, j) \): angle of a pixel with respect to the horizontal toward the center of ROI, in the range \([0^\circ, 359^\circ]\)
Angle Weighted Difference

\[ AWD = \sum_{p=1}^{PQ} \sum_{q=1}^{PQ} |I_p - I_q| \sin(|\alpha_p - \alpha_q|) \]

- \( I \): attribute value (intensity or magnitude)
- \( r \): radial distance from the center of the ROI
- \( \alpha(i, j) \): angle of a pixel with respect to the horizontal toward the center of ROI, in the range \([0°, 359°]\)
Angle-weighted Difference in the Entropy of Spicules

\[ \text{AWDES} = \sum_{m=1}^{90} \sum_{n=1}^{90} |H_m - H_n| |\sin(|\alpha_m - \alpha_n|)| \]

\( \alpha \): angular bands or sectors with their angles with respect to the \( x \)-axis toward the center of ROI (with 90 bins over \([0^\circ, 359^\circ]\))

\( H \): entropy of the attributes (intensity, magnitude, or angle) in the angular bands
## ROC Performance of Features

<table>
<thead>
<tr>
<th>Feature symbol</th>
<th>Feature name</th>
<th>$A_z$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node</td>
<td>Node value</td>
<td>0.61</td>
</tr>
<tr>
<td>$ICS_m$</td>
<td>ICS of magnitude</td>
<td>0.65</td>
</tr>
<tr>
<td>$ICS_c$</td>
<td>ICS of coherence</td>
<td>0.64</td>
</tr>
<tr>
<td>$RWD_i$</td>
<td>RWD of intensity</td>
<td>0.53</td>
</tr>
<tr>
<td>$RWD_m$</td>
<td>RWD of magnitude</td>
<td>0.62</td>
</tr>
<tr>
<td>$RWD_a$</td>
<td>RWD of angle</td>
<td>0.64</td>
</tr>
<tr>
<td>$AWD_i$</td>
<td>AWD of intensity</td>
<td>0.53</td>
</tr>
<tr>
<td>$AWD_m$</td>
<td>AWD of magnitude</td>
<td>0.62</td>
</tr>
<tr>
<td>$AWD_a$</td>
<td>AWD of angle</td>
<td>0.64</td>
</tr>
<tr>
<td>$AWDES_i$</td>
<td>AWDES of intensity</td>
<td>0.63</td>
</tr>
<tr>
<td>$AWDES_m$</td>
<td>AWDES of magnitude</td>
<td>0.64</td>
</tr>
<tr>
<td>$AWDES_a$</td>
<td>AWDES of angle</td>
<td>0.53</td>
</tr>
</tbody>
</table>
## Performance of Combinations of Features

<table>
<thead>
<tr>
<th>Feature Set</th>
<th>ROC Analysis ($A_z$)</th>
<th>FROC Analysis: Bayesian (FP/patient at sensitivities shown)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>80%</td>
</tr>
<tr>
<td>Node</td>
<td>0.61</td>
<td>8.2</td>
</tr>
<tr>
<td>All (ANN-RBF)</td>
<td>0.73</td>
<td>5.7</td>
</tr>
<tr>
<td>Selected set:</td>
<td>0.76 (ANN-RBF)</td>
<td>5.3</td>
</tr>
<tr>
<td>RWDi, RWDm, RWDa, AWDi, AWDm, AWDa, AWDESm</td>
<td>0.76 (ANN-RBF)</td>
<td>5.3</td>
</tr>
</tbody>
</table>

**ANN-RBF:** Artificial neural network based on radial basis functions
Sensitivity = 80% at 5.3 FP/patient

Sensitivity = 90% at 6.3 FP/patient
Other Approaches to Detect Architectural Distortion

Karssemeijer and te Brake, IEEE TMI 1996: multiscale-based method using the output of three-directional, second-order, Gaussian derivative operators

Sampat et al., IEEE SW Symp. Im. An. Int. 2006: linear filtering of the Radon transform of the given image for the enhancement of spicules; the enhanced image was filtered with radial spiculation filters
Other Approaches to Detect Architectural Distortion

Matsubara et al., CARS 2003, 2004: detection of architectural distortion near the skin line

Nemoto et al., IJ CARS 2009: lines corresponding to spiculation of architectural distortion differ in characteristics from lines in the normal mammary gland; modified point convergence index weighted by the likelihood of spiculation calculated to enhance architectural distortion
Fig. 7 Extracted lines with likelihood of spiculation. Lines with high likelihood are displayed in red and those with low likelihood are drawn in white.

Nemoto et al.
IJCARS 2009
Fractal Analysis

Guo et al. IJ CARS 2009: fractional Brownian motion model; regions with masses and architectural distortion have lower fractal dimension and higher lacunarity than normal regions.


Rangayyan et al. IJ CARS 2007: fractal analysis and texture analysis of ROIs detected in prior mammograms of cases of screen-detected cancer.
Expected Loci of Breast Tissue

CBMS 2012, IJ CARS 2012
Landmarking of Mammograms: Breast Boundary, Pectoral Muscle, Nipple

Second- and fifth-order polynomials fitted to parts of breast boundary
Derivation of Expected Loci of Breast Tissue: Interpolation
Distance between curves decreases with equal steps from AB to O

Number of curves = $N_1$; $L_{1\text{ max}} = N_1 - 1$

Distance between curves = 1 at AB

All curves contain $M_1$ points

Decrement along y-axis = $1/M_1$

$i$-th point of $n$-th curve:

\[
x_i(n) = x_i(1) - \left( \frac{n-1}{M_1-1} \right)[i-1]
\]

\[
y_i(n) = y_i(1)
\]
Number of points in curve = $M$

$L_i = \perp$ length between two curves at the $i$-th point

$L_{\text{max}} = \max(L_i)$

Number of curves = $N = L_{\text{max}}+1$

Distance at $i$-th point = $L_i / L_{\text{max}}$

= $L_i / (N-1)$

$i$-th point of $n$-th curve:

$$x_i(n) = x_i(1) - [x_i(1) - x_i(N_2)] \left( \frac{n - 1}{N_2 - 1} \right)$$

$$y_i(n) = y_i(1) - [y_i(1) - y_i(N_2)] \left( \frac{n - 1}{N_2 - 1} \right)$$
Number of points in curve = $M$

$L_i = \perp$ length between two curves at the $i$-th point

$L_{\text{max}} = \max(L_i)$

Number of curves = $N = L_{\text{max}} + 1$

Distance at $i$-th point = $L_i / L_{\text{max}}$

= $L_i / (N-1)$

$i$-th point of $n$-th curve:

$x_i(n) = x_i(1) - [x_i(1) - x_i(N_2)] \left( \frac{n - 1}{N_2 - 1} \right)$

$y_i(n) = y_i(1) - [y_i(1) - y_i(N_2)] \left( \frac{n - 1}{N_2 - 1} \right)$
Divergence with Respect to the Expected Loci of Breast Tissue

\[
\gamma(i, j) = \frac{\sum_{m=1}^{L} \sum_{n=1}^{L} |M(m, n) \cos[\theta(m, n) - \phi(i, j)]|}{\sum_{m=1}^{L} \sum_{n=1}^{L} M(m, n)}
\]

\(M\): Gabor filter magnitude response

\(\theta\): Gabor filter angle response

\(\phi\): expected orientation of breast tissue

\(L\): 25 pixels at 200 \(\mu\)m/pixel

180 Gabor filters used over \([-90, 90]\) degrees

\[
D(i, j) = 1 - \gamma(i, j)
\]
Orientation Field of Breast Tissue Obtained Using Gabor Filters

Original image  Gabor magnitude  Gabor angle
Divergence with Respect to the Expected Loci of Breast Tissue

Original image

Divergence map

Thresholded map
Automatically Detected Regions of Interest

ROC:
AUC = 0.61

FROC:
Sensitivity = 80%
at 9.1 FP/patient
### Performance of Selected Features

<table>
<thead>
<tr>
<th>Initial number of ROIs selected</th>
<th>Feature selection using stepwise logistic regression</th>
<th>Using all of the 12 features proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ROC analysis (AUC) FROC analysis (FP/patient at the sensitivities shown)</td>
<td>ROC analysis (AUC) FROC analysis (FP/patient at the sensitivities shown)</td>
</tr>
<tr>
<td>30</td>
<td>Bayesian 0.74 ANN 0.75 80% 5.3 90% 6.6</td>
<td>Bayesian 0.69 ANN 0.71 80% 5.9 90% 7.7</td>
</tr>
<tr>
<td>25</td>
<td>Bayesian 0.73 ANN 0.73 80% 6.0 90% 7.5</td>
<td>Bayesian 0.68 ANN 0.68 80% 5.6 90% 7.1</td>
</tr>
<tr>
<td>20</td>
<td>Bayesian 0.70 ANN 0.70 80% 6.7 90% 8.0</td>
<td>Bayesian 0.67 ANN 0.69 80% 5.8 90% 7.0</td>
</tr>
<tr>
<td>15</td>
<td>Bayesian 0.71 ANN 0.73 80% 5.8 90% 7.3</td>
<td>Bayesian 0.67 ANN 0.70 80% 5.7 90% 6.5</td>
</tr>
</tbody>
</table>
FROC Performance of Features
Combination of 86 Features

- Spiculation features IDS, RWD, AWD, AWDES: 12
- Haralick’s and Laws’ texture features, fractal dimension: 25
- Angular spread, entropy: 15
- Haralick’s measures with angle cooccurrence matrices: 28
- Statistical measures of angular dispersion and correlation: 6
- Feature selection with stepwise logistic regression
- Bayesian classifier with leave-one-patient-out validation: 80% sensitivity at 3.7 FP/patient (IJ CARS 2012)
Reduction of False Positives
Reduction of False Positives
“Our methods can detect early signs of breast cancer 15 months ahead of the time of clinical diagnosis with a sensitivity of 80% with fewer than 4 false positives per patient”

Future work:

- Detection of sites of architectural distortion at higher sensitivity and lower false-positive rates
- Application to direct digital mammograms and breast tomosynthesis images
Thank You!

- Natural Sciences and Engineering Research Council (NSERC) of Canada
- Indian Institute of Technology Kharagpur
- Shastri Indo-Canadian Institute
- University of Calgary International Grants Committee
- Department of Information Technology, Government of India
- My collaborators and students:
  Dr. J.E.L. Desautels, Dr. S. Mukhopadhyay, N. Mudigonda, H. Alto, F.J. Ayres, S. Banik, S. Prajna, J. Chakraborty

www.enel.ucalgary.ca/ People/ Ranga/