

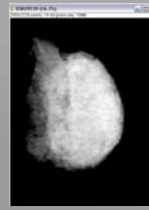


# A QUANTITATIVE METHOD FOR EVALUATING IMAGE PROCESSING ALGORITHMS IN DIGITAL MAMMOGRAPHY

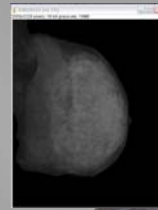
F. Zanca, C. Van Ongeval, J. Jacobs, H. Bosmans  
Leuven University Center of Medical Physics in  
Radiology, UZ Gasthuisberg, Leuven, Belgium



## DRAMATIC effect of Image Processing!!!!



For Processing



For Presentation

## Overview

1. Historical background
2. Method 3 steps: 1/Simulation 2/ Creation; 3/ Experiment
3. First Application: effect of anatomical background
4. Conclusions and Future : Comparison of different algorithms and how to get this into a standard?

## 1. Historical

- 2006: 4th edition of the European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis:
  - acceptance tests
  - routine inspections relevant for stability, image quality and dose for digital mammography
- Point 2b.3 Image processing :

“Image processing will not be considered in this version of the protocol. Special attention should be given to the visualization of microcalcifications and subtle structures.”

## PURPOSE

Develop a quantitative method to evaluate image processing algorithms for digital mammography

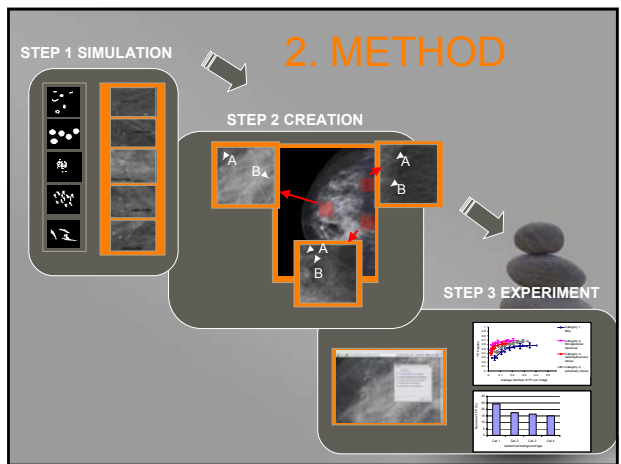
> generalization of the methodology to any processing

## 2 Method

- Part 1 We simulate microcalcifications on mammograms and use them as ground reference
- Part 2 We use this methodology for comparison of processing algorithms in a quantitative way

## Overview

1. Historical background
2. Method 3 steps: 1/Simulation 2/Creation; 3/ Experiment



## 2 Method

1. Simulation of Microcalcifications
2. Creation of Composite Images
3. Observer Performance Experiments



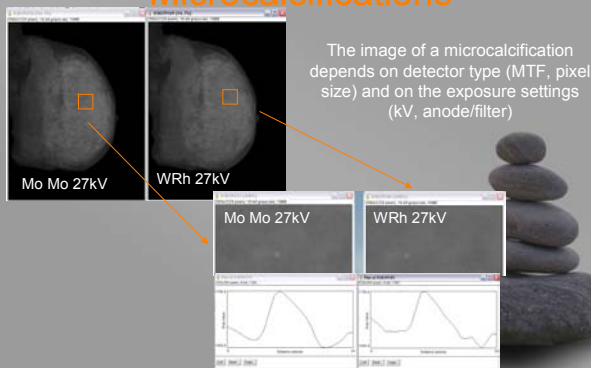
## 2 Method/Step 1 : Simulation of Microcalcifications\*

Annular round		Le Gal 1	
Pointy round		Le Gal 2	
Powder		Le Gal 3	
Pointy irregular		Le Gal 4	
Vermicular		Le Gal 5	

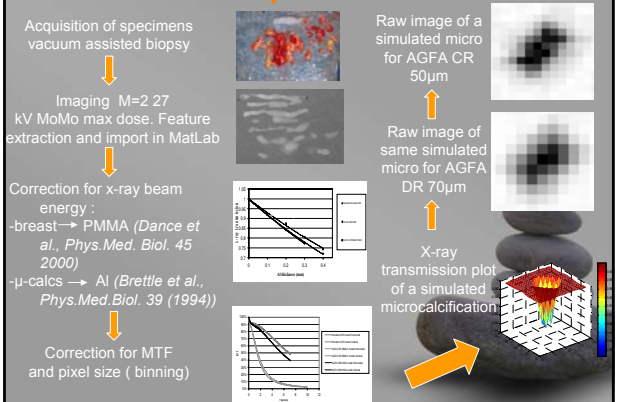
Validation of the methodology with ROC analysis:  $A_z = 0.52 \pm 0.4$

\*F. Zanca, C. Van Ongeval, Proceedings of SPIE Medical Imaging 2007

## 2 Method/Step 1 : Simulation of Microcalcifications

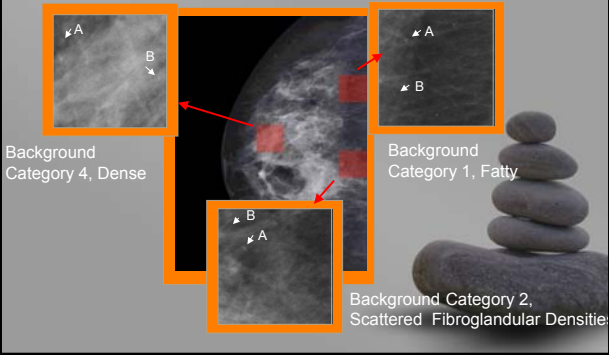


## 2 Method/Step 1 : Simulation



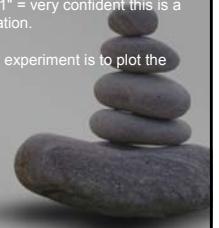
## 2 Method/Step 2 Composite images

In house developed software SimONA for the creation of composite images



## 2 Method/Step 3 Observer Performance Experiment

- Free response ROC: reader indicates (marks) and rates all locations on an image perceived as being suspicious.
- The confidence level used is an integer scale from "1" = very confident this is a microcalcification to "5" = probably not a microcalcification.
- One way to summarize the information in an FROC experiment is to plot the FROC curve
- Alternative to FROC analysis: JAFROC method



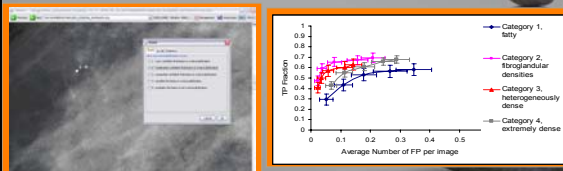
## 2 Method/Step 3 Observer Performance Experiment

### Analysis of the FROC data: 1/ FROC CURVES

- FROC curve: the plot of Lesion-Detected and Localized Fraction ( $\nu$ ) vs. mean number of False Positives per image ( $\lambda$ ).
- Error bars. Bootstrapping methodology

Graphical user interface OPerAA\*

Example of obtained FROC curves



\* F. Zanca, Proceedings of SPIE Medical Imaging 2006, vol. 6146

## 2 Method/Step 3 Observer Performance Experiment

### Analysis of the scoring data: 2 JACKKNIFE analysis \*

Two steps:

- scoring steps:** reduce data of a given observer and set of cases to a single number, a figure-of-merit  $\theta$  that quantify reader performance without any curve fitting
- statistical analysis steps:** estimate a confidence interval for the  $\theta$



\*Dev P. Chakraborty, Med. Phys. 31 .8., August 2004

## 2 Method/Step 3 Observer Performance Experiment

### Analysis of the FROC data: 2/JACKKNIFE analysis via software \*

Data for all readers for each processing are reduced to a Figure Of Merit  $\theta$ :  $\theta$  is the weighted-probability that a signal rating (scoring of a lesion) exceeds a noise rating (scoring of artefacts, noise) (ignore highest noise responses for abnormal images)

$$\theta = \frac{1}{N_N \sum_{i=1}^{N_A} \sum_{j=1}^{N_A} W_{ijk}} \phi(X_i, Y_{jk})$$

- $N_N$  total number of normal cases
- $N_A$  number of abnormal cases
- $X_i$  highest noise rate for a normal case  $i$
- $Y_{jk}$  signal rating for the  $k$ th target on case  $j$ .
- $W_{ijk}$  is the relative importance of detecting the  $k$ th signal on abnormal case  $j$

$$\phi(X, Y) = \begin{cases} 1.0 & \text{if } Y > X \\ 0.5 & \text{if } Y = X \\ 0.0 & \text{if } Y < X \end{cases}$$

$$\sum_{i=1}^{N_A} W_{ijk} = 1.$$

\* 95% confidence interval is calculated each calculated figure of merit

\*Dev P. Chakraborty, Med. Phys. 31 .8., August 2004

## 2 Method/Step 3 Observer Performance Experiment

Example:

Result of JAFROC analysis	MODALITY	PROCESSING
1 F-statistic and p-value		
F-statistic : 0.622423		whether or not the test revealed one or more significant differences between the modalities.
p-value : 0.447953		
NB: If the p-value is less than 0.05 there is a significant difference between at least one pair of modalities		
2 FOM = New JAFROC Figure of Merit, to appear in Academic Radiology		
(a) Reader Averaged FOMs and confidence intervals		
Modality 1 : 0.708920, and 95% CI = [ 0.653753 , 0.759592 ]		
Modality 2 : 0.736350, and 95% CI = [ 0.687199 , 0.781189 ]		
(b) FOMs for modalities (columns) and readers (rows)		
Modality :	1	2
Reader ID 1 :	0.686950	0.673800
Reader ID 2 :	0.717200	0.728150
Reader ID 3 :	0.762100	0.724700
Reader ID 4 :	0.729350	0.791050
Reader ID 5 :	0.649000	0.763950
(b) The FOMs for all combinations of readers and modalities.		
3 Inter-modality differences and 95% confidence intervals		
FOM(Modality 1) - FOM(Modality 2) = -0.0274100		For all pairings of modalities, the difference in reader averaged FOM, and the 95% confidence interval.
and 95% CI = [ -0.104828 , 0.050056 ] (NOT Significant Difference)		

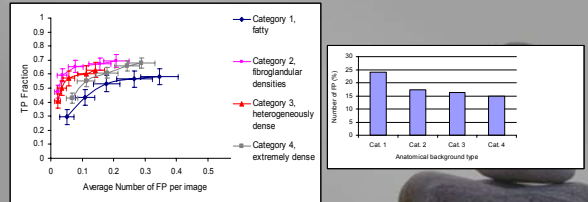
## Overview

1. Historical background
2. Method 3 steps: 1/Simulation 2/Creation; 3/ Experiment
3. First Application: effect of anatomical background



## 3/ Application

- Effect of anatomical background on detectability of microcalcifications



## Overview

1. Historical background
2. Method 3 steps: 1/Simulation 2/Creation; 3/ Experiment
3. First Application: effect of anatomical background
4. Conclusions : Comparison of different algorithms and how to get this into a standard?



## 4 Conclusions and Future

- A methodology for evaluation of processing algorithms for digital mammography has been established
- Next steps are:
  - To fully understand the effect of processing on the detectability and characterization of microcalcifications
  - To compare in a quantitative way different processing algorithms used in clinical practice.
  - To get the methodology into a standard usable by other people, more research is needed



## Acknowledgment

- We sincerely thank the 9 radiologists who scored the images.
- Reading session was organized in the frame of a multi center workshop within the Sentinel project. This is an EC supported project (FP6-012909) on Safety and Efficacy for New Techniques and Imaging using New Equipment to support European Legislation.
- We would like to thank dr. Thomas Mertelmeier and dr. Daniel Fisher from Siemens (Erlangen, Germany) for all the practical help with the reprocessing of the Novation DR images.
- Herman Pauwels is greatly acknowledged for his help with several practical problems.

