

Copyright 2006 Society of Photo-Optical Instrumentation Engineers.

This paper was published in Proceedings of SPIE -- Volume 6146 Medical Imaging 2006: Image Perception, Observer Performance, and Technology Assessment, Yulei Jiang, Miguel P. Eckstein, Editors, 61460U (2006), and is made available as an electronic reprint with permission of SPIE. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

Mammographic texture synthesis using genetic programming and clustered lumpy background

Cyril Castella^a, Karen Kinkel^b, François Descombes^c, Miguel P. Eckstein^d, Pierre-Edouard Sottas^a,
Francis R. Verdun^a, François O. Bochud^{a*}

^aInstitut Universitaire de Radiophysique Appliquée, CH-1007 Lausanne, Switzerland;

^bClinique des Grangettes, Chemin des Grangettes 7, CH-1224 Chêne-Bougeries, Switzerland;

^cHaute Ecole Cantonale Vaudoise de la Santé, Bugnon 19, CH-1005 Lausanne, Switzerland ;

^dDept. of Psychology, University of California, Santa Barbara, CA USA 93106-9660

ABSTRACT

In this work we investigated the digital synthesis of images which mimic real textures observed in mammograms. Such images could be produced in an unlimited number with tunable statistical properties in order to study human performance and model observer performance in perception experiments.

We used the previously developed clustered lumpy background (CLB) technique and optimized its parameters with a genetic algorithm (GA). In order to maximize the realism of the textures, we combined the GA objective approach with psychophysical experiments involving the judgments of radiologists. Thirty-six statistical features were computed and averaged, over 1000 real mammograms regions of interest. The same features were measured for the synthetic textures, and the Mahalanobis distance was used to quantify the similarity of the features between the real and synthetic textures. The similarity, as measured by the Mahalanobis distance, was used as GA fitness function for evolving the free CLB parameters. In the psychophysical approach, experienced radiologists were asked to qualify the realism of synthetic images by considering typical structures that are expected to be found on real mammograms: glandular and fatty areas, and fiber crossings.

Results show that CLB images found via optimization with GA are significantly closer to real mammograms than previously published images. Moreover, the psychophysical experiments confirm that all the above mentioned structures are reproduced well on the generated images. This means that we can generate an arbitrary large database of textures mimicking mammograms with traceable statistical properties.

Keywords: image perception, texture analysis, image generation, mammography

1. INTRODUCTION

The problem of human perception and performance in radiology detection tasks has been studied in numerous frameworks in the past: detection of a tumor on computer tomographic images of the liver^{1,2}, stenosis in a blood vessel on fluoroscopic images³, nodules on pulmonary radiographs⁴, or microcalcifications on mammograms⁵. The aim of such studies is to determine the role on diagnostic decisions of the inherent parameters of the images like resolution or contrast, the imaging unit acquisition parameters or the anatomy in the detection process. Many of such studies are psychophysical experiments involving radiologists or trained naïve observers.

In particular, there has been a large interest in developing models that could predict human observer performance for detecting tasks as a function of the image characteristics and the observer properties⁶. These models aim at avoiding the subjective and time-consuming aspect of the psychophysical studies, as well as evaluating medical image quality^{7,8}. Models for objects superimposed on various types of real backgrounds or computer generated noises have been developed and applied to the detection of lesions in radiological images^{9,10,11}.

* francois.bochud@chuv.ch, phone +4121.623.34.40; fax +4121.623.34.35; <http://www.chuv.ch/public/instituts/ira/>

Both psychophysical and model observer approaches require a large number of images to obtain accurate results. Real images or regions of interest databases would be ideal, but in most cases the number of available clinical images is limited. In addition, the question arises about reproducibility of the results with other sets of images. An alternative to using real images is to use computer generated images which would allow for generation of unlimited number of samples with known and well-controlled statistical properties. Such images might have adjustable properties that would not depend on imaging device characteristics or digitization processes.

Two major methods have been explored for producing synthetic images mimicking mammograms. First, complete three-dimensional simulation of the breast components and properties, in conjunction with imaging device simulation, which is expected to produce very realistic images¹². However, the complexity and computational cost associated with such modeling and the difficulty of taking into account breast compression often takes from the quality of the resulting images. For that reason, 2D approaches have been investigated, using backgrounds constituted by the summation of elementary bright structures called blobs^{7,8,13,14}. These lumpy backgrounds, as named originally by Rolland and Barrett¹³, were designed to reproduce the textures observed in mammograms. Bochud *et al.*¹⁴ generalized the model to clustered lumpy backgrounds (CLB), and empirically optimized the parameters to obtain images which are as visually realistic as possible. CLB images have the advantage of having analytically computable statistical properties and are stationary within their boundaries. However, for this model as for most of 3D or 2D methods, thorough and objective assessment of visual and “statistical” realism has not been carried out. The main obstacle has been the difficulty of defining criteria for the assessment. They are nevertheless necessary for images intended to be used for experiments utilizing both human (visual realism) and model observers (statistical realism).

Our aim in this study was to extend and optimize the CLB model and to objectively assess the realism of the obtained images. For this purpose, we used a database of 1000 square regions of interest selected from real mammograms, and defined a metric based on the Mahalanobis distance to compute the statistical distance between real images and synthetic regions of interest produced by adjusting CLB parameters with a genetic algorithm. Psychophysical experiments were then designed in order to evaluate the visual realism of the synthetic images.

2. MATERIAL AND METHODS

2.1. Clustered lumpy background (CLB) model

Lumpy backgrounds are digital images consisting of the superposition of elementary bright blobs. The number of blobs is randomly sampled according to a Poisson process and the blob centers are placed at random locations uniformly distributed in the image. Lumpy backgrounds were originally designed by Rolland and Barrett¹³ with circularly symmetric blobs $b(\mathbf{r})$, so that the image \mathbf{g} could be written as :

$$\mathbf{g}(\mathbf{r}) = \sum_{k=1}^K b(\mathbf{r} - \mathbf{r}_k), \quad (1)$$

where \mathbf{r}_k is the center position of the k^{th} blob, and K the total number of blobs in the image.

Later, Bochud *et al.*¹⁴ generalized this model to clusters of non-necessary circularly symmetric blobs. Clustered lumpy backgrounds (CLB) are produced by randomly choosing a number of clusters, K , following a Poisson process, and distributing them randomly on the image plane. For each cluster, a random number of blobs, N_k , are positioned randomly around the cluster center according to a probability density function (pdf) $\phi(\mathbf{r})$. Finally, all blobs belonging to the same k^{th} cluster are rotated by an angle θ_k before being summed to obtain the final image $\mathbf{g}(\mathbf{r})$:

$$\mathbf{g}(\mathbf{r}) = \sum_{k=1}^K \sum_{n=1}^{N_k} b(\mathbf{r} - \mathbf{r}_k - \mathbf{r}_{kn}, \mathbf{R}_{\theta_k}), \quad (2)$$

where all parameters and their distributions are summarized in Table 1.

Table 1: Definitions and distributions of the variables described in the CLB model.

Variable	Definition	Distribution
K	number of clusters	poisson with mean value K_0
\mathbf{r}_k	position of the k^{th} cluster	uniform across image
N_k	number of blobs within the k^{th} cluster	poisson with mean value N_0
θ_k	rotation angle of the blobs in the k^{th} cluster	uniform between 0 and 2π
\mathbf{R}_θ	rotation matrix of angle θ	N/A
\mathbf{r}_{kn}	position of the n^{th} blob within the k^{th} cluster	Gaussian pdf $\phi(\mathbf{r})$
$b(\mathbf{r}, \mathbf{R}_\theta)$	blob profile rotated at angle θ	N/A

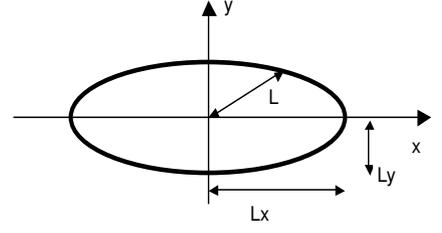


Fig. 1: Characteristic length of an ellipse having half-axes L_x and L_y

The general functional expression of the blob has been chosen as:

$$b(\mathbf{r}, \mathbf{R}_\theta) = \exp\left(-\alpha \frac{\|\mathbf{R}_\theta \mathbf{r}\|^\beta}{L(\mathbf{R}_\theta \mathbf{r})}\right), \quad (3)$$

where α and β are real parameters, and the characteristic length L is defined on Fig. 1.

One of the major advantages of CLB technique is that some statistical properties of $g(\mathbf{r})$ like its power spectrum can be analytically computed from the model parameters¹⁴.

The free parameters of the CLB model are thus $\{\alpha, \beta, L_x, L_y, \sigma_x, \sigma_y, K_0, N_0\}$, where σ_x and σ_y are the standard deviations of the Gaussian pdf, $\phi(\mathbf{r})$, in x and y directions respectively. These 8 parameters had been empirically optimized in the original study¹⁴, on the basis of visual inspections of the images and comparison of Wiener spectrum with that of real mammograms. These values were used as a starting point for our study.

In order to improve the realism of CLB images, we introduced two variations into the model. First, we superimposed another CLB onto the image computed from Eq. (2), with fixed parameters $\alpha = 2.0$, $\beta = 0.9$, $L_x = 50$ pixels, $L_y =$ pixels, $\sigma_x = 10$ pixels, $\sigma_y = 10$ pixels, and free parameters $K_0' < K_0$ and N_0' . The inclusion of a small amount of long and narrow blobs aims to better reproduce the fibrous structures of real mammograms.

The second variation was to favor oriented structures similar to those visible on real mammograms. At the whole breast scale, these structures arise from the projection of the ducts converging to the nipple, or from suspensory ligaments. For this purpose, the pdf of the rotation angle was changed from uniform to Gaussian with a mean equal to θ_0 and a standard deviation of $\pi/6$. With this change, the large scale oriented structures were constructed by the summation of clusters with similar orientation. The mean parameter, θ_0 , was changed randomly with uniform pdf between 0 and 2π for each realization.

2.2. Genetic algorithms and application to CLB model

Genetic algorithms are a family of computational models inspired by evolution¹⁵. The free parameters of a given optimization problem are encoded on a chromosome-like data structure, and selection and recombination operators are applied in order to allow a population of potential solutions to evolve towards the optimal solution of the problem. The initial population is usually chosen randomly in the search space, and the corresponding chromosomes are evaluated through a fitness function. The best chromosomes are given better reproduction and survival opportunities. Following, crossover and mutation operators are applied in order to generate a new population of equal cardinality. These processes of evaluation, crossover and mutation are repeated until a user-defined (sub-)optimal value of the fitness function is reached, or when the best chromosome of the population has not been improved for a given number of generations.

Genetic algorithms have a great potential for non-linear function optimization in multi-dimensional spaces, since the intrinsic parallel structure of the optimization process is highly efficient for exploring multiple locations in the search space simultaneously, and avoiding local extrema. They can be used for binary or real coded problems, and many specific reproduction/mutation operators and techniques have been designed¹⁶ in order to design specific algorithms for handling a wide range of optimization problems.

According to Eq. (2) and (3), a classical CLB implementation requires a set of eight real parameters $\{\alpha, \beta, L_x, L_y, \sigma_x, \sigma_y, K_0, N_0\}$. For the 2-layer CLB, the addition of $\{K_0', N_0'\}$ increases the number of parameters to ten. The statistical properties of CLB images depend in a non-analytical way on the parameters. Their optimization is furthermore complicated by the stochastic nature of the realizations for a same set of parameters. The optimization of the eight parameters of the classical CLB model¹⁴ were limited to maximize the similarity of the gray level histogram properties and Wiener spectrum of the synthetic and real mammographic textures. No other consideration was taken into account in order to evaluate the mathematical realism of the obtained synthetic images. One key aspect of the present study was to introduce a metric based on Mahalanobis distance for quantifying similarity between synthetic and real images.

For this purpose, 36 statistical features based on the gray level histogram properties, the co-occurrence matrices^{17,18,19}, the primitives matrices¹⁷, the neighborhood gray tone difference matrix²⁰ and the fractal dimension²¹ were computed for 1000 square regions of interest within digital mammograms²². These 256 by 256 pixels square regions were manually selected from the central breast areas of digital mammograms. We used a database of 88 patients who underwent screening exams on a GE Senograph 2000D full-field digital detector^{23,24}, with one craniocaudal (CC) and one mediolateral oblique (MLO) view per breast for each of them. Once all 36 features of a given synthetic or real image were measured and grouped into a single vector \mathbf{v} , the Mahalanobis distance d was given by:

$$d^2 = (\mathbf{v} - \boldsymbol{\mu})^T \mathbf{K}^{-1} (\mathbf{v} - \boldsymbol{\mu}), \quad (4)$$

where $\boldsymbol{\mu}$ represents the mean vector over the real images and \mathbf{K} is the covariance matrix:

$$\boldsymbol{\mu} = \frac{1}{n} \sum_{i=1}^n \mathbf{v}_i \quad (5)$$

$$\mathbf{K} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{v}_i - \boldsymbol{\mu})^T (\mathbf{v}_i - \boldsymbol{\mu}), \quad (6)$$

with $n = 1000$ being the size of the reference database.

The chromosomes in our genetic algorithm implementation were sets of 8- or 10-dimensional real vectors representing CLB parameters values. The genetic algorithm used the average Mahalanobis distance computed over $m = 10$ successive CLB realizations as the fitness function for evaluating the chromosomes, and was designed to minimize it. This averaging was done in order to avoid erroneous evaluation caused by the random nature of the CLB algorithm. Rank-weighted selection of the parents, and elitist strategy were employed for the reproduction operators. The first parent chromosome was thus chosen with probability:

$$p(r) = \frac{1}{S} \left(\psi - (r-1) \frac{2\psi-2}{n-1} \right), \quad (7)$$

where r is the rank of the chromosome (rank 1 being attributed to the best chromosome), S is the population size, and ψ is a parameter influencing the relative weight of the best chromosomes compared to the worst. The second parent was chosen randomly with equal probability among the population, and the operation was repeated until $(S-1)/2$ couples of chromosomes had been chosen for crossover.

Crossover of two chromosomes \mathbf{c}^1 and \mathbf{c}^2 consisting in averaging half of the genes, keeping the others unchanged. The genes to be averaged were chosen randomly with equal probabilities. The crossover between \mathbf{c}^1 and \mathbf{c}^2 occurred with probability p_c , leaving both genes unchanged otherwise. The best chromosome remained unchanged from one generation to the next, which is the definition of elitist strategy.

After the crossover processes, all but the elite chromosome underwent individual gene mutation with probability p_m . The value of a mutated gene G' was given by:

$$G' = \begin{cases} G + (G_{\max} - G) \left(1 - r \left(\frac{1-t}{T}\right)^5\right) & \text{if } p = 1 \\ G - (G - G_{\min}) \left(1 - r \left(\frac{1-t}{T}\right)^5\right) & \text{if } p = -1 \end{cases}, \quad (8)$$

where G is the value of the original gene, G_{\min} and G_{\max} the user-defined bounds of the search space for gene G , T the number of generations for which mutations might occur, $t \leq T$ the current generation, and $r \in [0,1[$ and $p \in \{-1,1\}$ were chosen randomly each time a mutation occurred. Eq. (8) indicates that mutations are stronger at the beginning of the optimization, in order to extend the exploration of the search space, and to decrease monotonically in intensity. The mutation probability p_m itself evolved with t for the same reasons, as:

$$p_m(t) = \left(2 + \frac{L-2}{T}t\right), \quad (9)$$

where L is the number of genes in a chromosome.

The initial population was chosen randomly between the bounds G_{\max} and G_{\min} defined individually for each gene. We made the assumption that the empirical optimization proposed in the original study¹⁴ ($G = G_{\text{Oex}}$) could be used as starting point, and we computed G_{\min} and G_{\max} as:

$$\begin{cases} G_{\min} = 0.8G_{\text{Oex}} \\ G_{\max} = 1.2G_{\text{Oex}} \end{cases}, \quad (10)$$

thus allowing $\pm 20\%$ deviations from these values. Values for the various G_{\min} and G_{\max} are summarized in Table 2. All parameters of the genetic algorithm and their meaning are given in Table 3. Four variations of the CLB model were successively optimized: 1-layer classical CLB with isotropic orientation of the clusters (this will be referred further in text as *simpiso* type), 2-layer CLB with isotropic orientation of the clusters (*doubiso*), 2-layer CLB with favored orientation of the clusters (*doubori*), and 1-layer CLB with favored orientation of the clusters (*simpori*).

Table 2: Lower and upper bounds of the search space for each gene.

Variable	$[G_{\min}, G_{\max}]$
α	[1.6,2.4]
β	[0.4,0.6]
L_x	[4,6]
L_y	[1.68,2.52]
σ_x	[9.6,14.4]
σ_y	[9.6,14.4]
K_0	[480,720]
N_0	[16,24]
K_0'	[50,500]
N_0'	[5,50]

Table 3: Genetic algorithm parameters used for optimizing CLB variables

Parameter	Meaning	Value
L	Number of genes in a chromosome	8 or 10
S	Size of the chromosomes population	51
m	Number of realizations of a chromosome for evaluating its fitness function	10
T	Time constant of the mutation operator	200
ψ	Relative weight of the best chromosomes	1.5
p_c	Crossover probability	0.8
p_m	Mutation probability	Eq. (9)

2.3. Evaluation of the visual realism of synthetic images

The role of the genetic algorithm is to ensure that the synthetic CLB images would have statistical properties similar to real images. Although this point is necessary for future model observer experiments for example, it is certainly not a sufficient condition for using them in psychophysical detection experiments. Human perception is highly dependent on properties of the background as well as those of the neural processing and coding of visual information. Thus, similar statistical properties for a pair of images does not necessarily imply their visual resemblance. To evaluate the visual realism of the four optimized CLB types, a study was conducted with three radiologists and two radiographers.

The three main structures types that are likely to be found in real mammograms were evaluated: glandular areas, fatty areas, and fibers. The observers were first presented a series of 20 real images representative of each structure type. The selection of these reference images was based on the choices of one of the radiographers, and then confirmed by the opinion of a radiologist. The presentation of the reference images also allowed the radiologists to get acquainted to the laptop screen, light conditions, and definitions used for the three structure types. After this training phase, 50 realizations of each CLB model variation were presented in random order. The four variations developed with the GA, and the original CLB¹⁴ were displayed in 10 blocks of 25 images. The order of presentation for each CLB type was randomized within each block.

For each image, the observers were asked to tell whether or not they observed a given structure (glandular areas, fatty areas, fibers). For each affirmative answer, they were asked to grade the realism of the structure, indicating whether it could be expected to be observed on real mammograms or not, based on a 10-grade scale evaluation. Grades 7 to 10 were considered to be realistic enough, and 1 to 6 to be insufficiently realistic. In the latter case, the observers were given the possibility to further evaluate which features looked unrealistic by using one or more checkboxes representing possible defaults: too disorganized, too rectilinear, too much contrast, too fuzzy, or appearance of 3D-like artifacts. Additionally, the radiologists were asked to mention if some structure resembled a tumor (mass). This latter question was aimed at determining whether unwanted pathological (tumor-like) patterns arose from the CLB superimposition algorithm.

The 12-bits CLB images were converted to 256 gray levels before being displayed on the laptop screen. Their mean gray level value and standard deviation were adjusted to 110 and 35 respectively, in order to obtain images lying in the central dynamic range region of the display. The observers had the possibility to adjust the display brightness and contrast by observing a mammography test pattern at the beginning of the experiment. All radiologists and radiographers reported satisfactory conditions to confidently assess the realism of the three structure types, since they were to be compared to real digital mammograms regions of interest displayed on the same screen at the beginning of the test. The 256 by 256 pixels synthetic images display size was 9 by 9 cm. According to preliminary discussions with the radiologists, the size of the synthetic images structures at this scale corresponded to the scale obtained when zooming on the GE Senograph 2000D display unit.

3. RESULTS

3.1. CLB parameters optimizations

The genetic algorithm designed to optimize the different variations of CLB model ran on a Pentium 4 (3GHz processor, 512 MB RAM) processor. Typical total time for evaluations, crossovers and mutations was around 150 minutes per generation. More than 95% of the computation time was spent for computing the m realizations of each chromosome in the population, whereas applying evolution operators was a faster process. Although genetic algorithms with elitist strategy usually have the property to be monotonically converging towards extrema of the fitness function, the example fitness function history on Fig. 2 shows that it decreased relatively regularly during 20-30 generations, and then had a more chaotic behavior. This was observed for all model variations, and can be explained by the random nature of the m realizations per chromosome that were computed for evaluating its fitness function. Same CLB parameters lead to images with similar overall statistical properties, but the 36 features we used in this study were able to evaluate their variations much more precisely. The fitness function of a given chromosome could thus vary from a generation to another, and the best chromosome of generation T' could be rejected to a higher rank at $T'+1$, even by chromosomes that had worse performance at generation T' . The upper series in Fig. 2 shows that the median fitness function of the population was less sensitive to this phenomenon. The evolution process was conducted during 100 generations for each

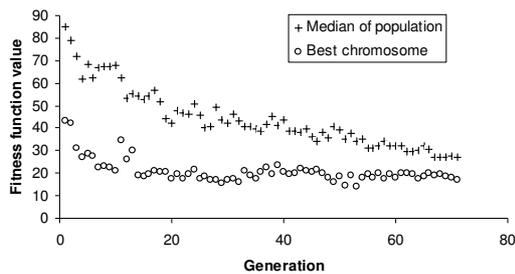


Fig. 2: Example of fitness function history (70 first generations). The upper series represents the median value of the fitness function evaluated on the population at generation t , and the lower series indicates the value for the best chromosome.

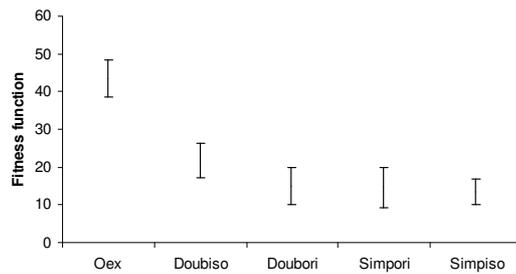


Fig. 3: Fitness function computed from 200 realizations with the optimized set of CLB parameters for model variation. The error bars represent standard deviation of the realizations' fitness function.

of the variation of the CLB model, and the best chromosome of the evolution history was selected for computing the fitness functions presented on Fig. 3, on the basis of 200 realizations per model.

Fig. 3 shows that the gain obtained by tuning the CLB parameters with the genetic algorithm is at least a factor of 2 for average Mahalanobis distance, compared to the original values¹⁴ (*Oex* series), depending on the model used. One-sided t-test (significance level of 5%, power of 0.8, under the null hypothesis that the means are equal) showed that the means of the last three presented series were significantly lower than for the series *doubiso*, and that *Oex* series was significantly higher than the other four variations.

Fig. 4 presents representative examples of images created with the different CLB parameters. The real mammogram region of interest (ROI) was selected from a medium-density breast. Superimposition effects for 2-layer backgrounds *doubiso*, *doubori* and oriented structures for *doubori* and *simpiso* series are visible. The optimized CLB parameters for generating these 256 by 256 pixels images are detailed in Appendix A. Typical computation time needed for computing the 200 realizations and their associated Mahalanobis distance was 40 minutes, which represents 12 seconds/realization.

3.2. Evaluating the realism of synthetic textures

Table 4 summarizes the results for visual realism evaluation experiments performed by the radiologists (KK, ES, NH) and the radiographers (FD, PS). About 2% of the data was classified as outliers according to Chauvenet criterion²⁵. The outliers arose from the fact that the observers sometimes used grades which were too extreme for qualifying subjectively excellent or poor images, compared to their usual answers. The corresponding data were removed before results analysis.

For each synthetic image, the radiologists evaluated the realism of the three structures types (glandular areas, fatty areas, and fibers), whereas the two radiographers chose not to give their opinion in some cases, when they judged that a given structure covered a too small part of the ROI to be evaluated. This mainly happened for the evaluation of fibers, which are less visible in *Oex* and *simpiso* series. This latter series (*simpiso*) was only evaluated by the first two observers which took part in the study. T-tests of the first two observers' data (with significance level equal to 5% and power equal to 0.8) showed that the *simpiso* series was significantly lower from the apriori selected threshold for realism (6.5) and thus visually unrealistic.

Bold values in Table 4 indicate that nearly all structures were considered significantly realistic (above a grade of 6.5) for the first four CLB models. *Oex*, *doubori* and *simpori* series obtain comparable overall performances for all structures types, while *doubiso* series outperforms them for all structures types when the results are averaged over the 5 observer (Fig. 5). This series was evaluated by the three radiologists as significantly better than *Oex* for all structure types, and

better than all other series by ES and NH. The radiographers, however, did not significantly prefer a unique series for all structure types. For these two observers, *Oex* and *doubiso* series obtained the best evaluations for glandular and fatty areas, while observer PS preferred *doubiso* and *simpori* series for the fibers. No significant differences between the four series were observed for the evaluation of fibers by observer FD.

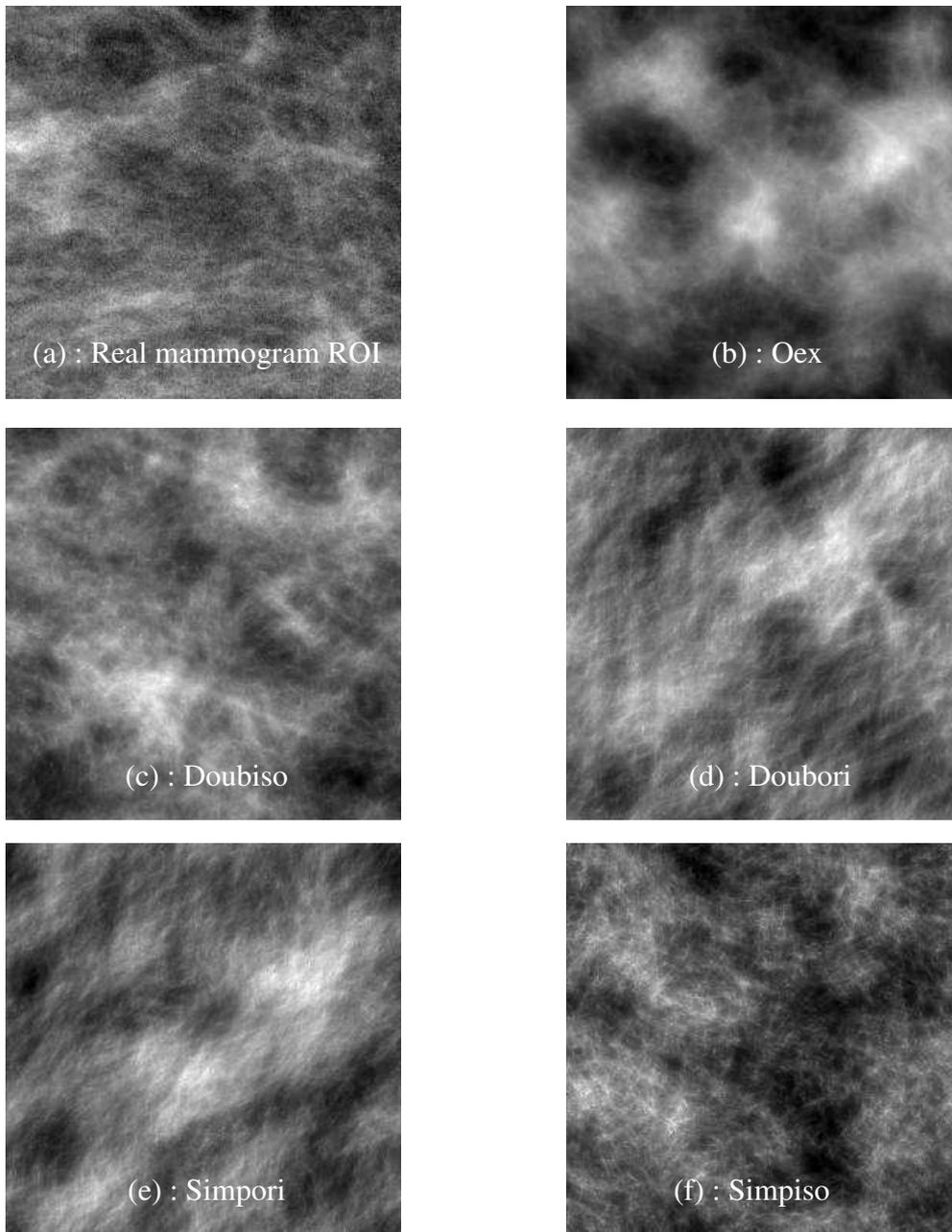


Fig. 4 Examples of realizations for the different types of CLB variations. (a) ROI selected from a real mammograms; (b) 1-layer CLB, *Oex* parameters (referred in text as *Oex*); (c) 2-layer CLB, isotropic orientation of the clusters (*doubiso*); (d) 2-layer CLB, favored orientation of the clusters (*doubori*); (e) 1-layer CLB, favored orientation of the clusters (*simpori*); (f) 1-layer CLB, optimized version of (b) (*simpiso*)

Table 4: Realism evaluation by the five observers. Bold values indicate substantially realistic evaluations. The statistical test performed was one-sided t-test, with significance level = 5%, and power = 0.8, under the null hypothesis H_0 that the average is equal to 6.5.

	Oex	7.38		Doubiso	7.99		Doubori	7.37		Simpori	7.32		Simpiso	5.02	
	Gland.	Fatty	Fibers												
KK	6.80	7.18	7.85	7.18	7.96	7.94	7.40	7.96	8.78	7.22	7.76	7.92	5.00	6.47	4.93
ES	6.65	6.87	6.87	7.68	7.70	7.73	6.68	6.81	6.81	6.98	7.05	7.11			
NH	7.68	7.68	7.68	9.00	9.00	9.00	7.68	7.68	7.68	6.96	6.96	6.96			(b)
FD	7.47	8.16	7.00	7.68	8.04	6.87	7.54	7.93	7.03	7.44	7.71	7.08			
PS	7.65	7.77	(a)	8.40	8.14	7.58	7.23	7.15	6.18	7.55	7.46	7.67	4.67	4.95	4.09
Mean	7.3	7.5	7.4	8	8.2	7.8	7.3	7.5	7.3	7.2	7.4	7.3	4.8	5.7	4.5
	± 0.2	± 0.2	± 0.3	± 0.3	± 0.2	± 0.4	± 0.2	± 0.2	± 0.4	± 0.1	± 0.2	± 0.2	± 0.2	± 0.8	± 0.4

(a) This observer did not report any observation of fibers in that series

(b) This series was only evaluated by two observers.

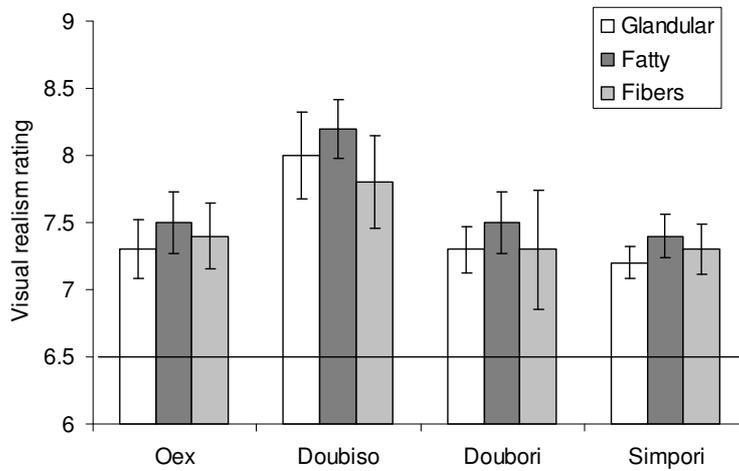


Fig. 5: Average values obtained by the four best series for the realism of glandular areas, fatty areas, and fibers. Error bars represent standard deviation of the mean

4. DISCUSSION

The genetic algorithm, although made complex by the inherent random nature of CLB model, produced images which statistical properties were significantly closer to real mammographic images. As the chromosomes' evolution continued for more than 50 generations after the optimal parameters presented in Table 5 had been found, these values can thus be confidently considered as optimal for the developed model variations. It is difficult to intuitively interpret the absolute Mahalanobis distances in Fig. 3, since several statistical parameters, among the 36 used for defining that metric, are correlated in a complex way. However, a benchmark can be given by the distance computed for the 1000 real mammograms regions of interest, which is equal to 5.6 ± 2.1 . This indicates that from the statistical point of view, the synthetic images obtained by the models tuned by the genetic algorithm are much closer to real images than the original *Oex* series, but also that they cannot be considered indistinguishable from real images. Allowing enlarged bounds for G_{\min} and G_{\max} would lead to optimized chromosomes with better fitness function, but preliminary tests had shown that when given more freedom, the blobs dimensions evolved to points as small as $G_{\min,Lx}$ by $G_{\min,Ly}$, lowering the average Mahalanobis distance down to about 10-15 depending on the model, but losing all visual realism. For that reason, we limited the search space of the genetic algorithm to reasonably close values to the original CLB model.

Concerning the model variations and their effect on the visual evaluation by the radiologists and radiographers, the favored orientation of the structures in *simpori* and *doubori* series was generally recognized as such by the observers, and their main drawback was that in some cases this orientation was too obvious and artificial, giving them the impression of seeing three-dimensional structures instead of flat projections. This defect was particularly mentioned in *simpori* series, while the few large scale structures of *doubori* seemed to hide or mask the principal layer composed of the smaller blobs. On the other hand, the observers found that some of the isotropic images were too disorganized to correctly represent real mammograms. This was the main reason for discarding the display of *simpiso* series for the last three observers in the psychophysical study. The presence of the second layer CLB in *doubori* did not improve or deteriorate significantly the visual aspect of *simpori* series, but the difference was clearly shown by the observers' evaluations for the isotropic series: they reported unorganized images with too much contrast for the 1-layer series, and selected the 2-layer *doubiso* images as best overall series. The only limitation mentioned by the radiologists for that series was that for some images (about 10% of the set), bright points caused by blobs superimposition might be interpreted as clusters of microcalcifications. However, for visual experiments of mass detection, they confirmed that this downside would not be critical, since they are not affected by the presence of mm-scale microcalcifications when looking for cm-scale structures like masses.

A still open question is the visual variability of the synthetic textures. As for most of other models, it is much smaller than that of real images. This has been partly solved by converting the CLB output float images to 12-bit image using randomly chosen values of mean gray level and standard deviation following real image corresponding distributions. One could have imagined using "floating" values for CLB parameters as well, but this possibility was not applied in our study.

5. CONCLUSION

Using a genetic algorithm and variations of the original CLB model, we were able to synthesize images having significantly closer visual and statistical properties than the original model. These models and parameters allows for generating an arbitrary number of such images while guaranteeing their realism. The synthetic images may find direct applications in detection experiments involving human or model observers. We would recommend *doubiso* series in particular, since they have excellent visual characteristics, even if their statistical properties are more distant from real images than for the other model variations *simpori* and *doubori*.

Compared to other image synthesis techniques, our models are limited to the generation of square regions of interest. However, they have the advantage of being able to quickly generate a large number of images, with traceable statistical properties, and visually representing all major structures types (glandular areas, fatty areas, fibers) that are visible on real mammograms. The methodology presented in this study is not limited to mammography and may be easily generalized to other medical or non-medical images: the only need in such cases would be a sufficiently large database of reference textures for defining the Mahalanobis distance used as fitness function by the genetic algorithm for tuning the CLB parameters.

ACKNOWLEDGMENTS

The authors are grateful to Elsabe Scott, MD, Nigel Howart, MD, Christel Elandoy, and Philippe Spring for their participation to psychophysical experiments, and to Cormac Mac Gearailt for his careful revision of the manuscript. This work was supported by Swiss National Science Foundation FN 3252BO-104273.

APPENDIX

A. Optimal CLB parameters for each model variation

The CLB parameters mentioned used for generating the ROIs of Fig. 4 are given in Table 5.

Table 5: Optimized CLB parameters for the various CLB models

Series	α	β	Lx	Ly	σ_x	σ_y	K_0	N_0	K_0'	N_0'
Oex	2.1	0.5	5	2	12	12	600	20	N/A	N/A
Doubiso	2.31	0.57	4.09	1.76	13.27	13.92	643.81	20.21	61.47	5.60
Doubori	2.49	0.53	4.14	1.61	11.17	14.29	709.96	20.37	78.00	5.01
Simpori	2.51	0.54	4.53	1.66	10.67	13.33	714.44	23.14	N/A	N/A
Simpiso	2.47	0.59	4.19	1.63	11.02	12.27	674.97	16.53	N/A	N/A

REFERENCES

- ¹ P. F. Judy, R. G. Swensson, R. D. Nawfel and K. H. Chan, "Contrast detail curves for liver CT", *Medical Physics* vol. 19, 1167-1174, 1992
- ² S. E. Seltzer, P. F. Judy, R. G. Swensson, K. H. Chan and R. D. Nawfel, "Flattening of the contrast-detail curve for large lesions on liver CT images", *Medical Physics*, vol. 21, 1547-1555, 1994
- ³ M.P. Eckstein and J.S. Whiting, "Visual signal detection in structured backgrounds. I. Effect of number of possible spatial locations and signal contrast," *Journal of Optical Society of America A*, vol. 13, 1777-1787, 1996
- ⁴ C. Herrmann., E. Buhr and D. Hoeschen, "Bildrauschen und Diagnose von Rundherden in Thoraxaufnahme", *Zeitschrift für Medizinische Physik* vol.6, 80-86, 1996
- ⁵ F.O. Bochud, J.-F. Valley, F.R. Verdun, C. Hessler and P. Schnyder, "Estimation of the noisy component of anatomical backgrounds", *Medical Physics*, vol. 26, 1365-1370, 1999
- ⁶ R.F. Wagner and D.G. Brown., "Unified SNR analysis of medical imaging systems", *Physics in Medicine and Biology* vol. 30, 489-518, 1985
- ⁷ L. Chen and H.H. Barrett, "Task-based lens design with application to digital mammography", *Journal of Optical Society of America A*, vol. 22, 148-167, 2005
- ⁸ M.A. Kupinski, E. Clarkson, J.H. Hoppin, L. Chen, H.H. Barrett, "Experimental determination of object statistics from noisy images", *Journal of Optical Society of America A*, vol. 20, 421-429, 2003
- ⁹ K. J. Myers, H.H. Barrett, M.C. Borgstrom, D. D. Patton and G. W. Seeley, "Effect of noise correlation on detectability of disk signals in medical imaging", *Journal of Optical Society of America A*, vol. 2, 1752-1759, 1985
- ¹⁰ F.O. Bochud., C.K. Abbey and M.P. Eckstein, "Visual signal detection in structured backgrounds. III. Calculation of figures of merit for model observers in statistically nonstationary backgrounds", *Journal of the Optical Society of America A*, vol. 17, 193-205, 2000
- ¹¹ F.O. Bochud, C.K. Abbey, M.P. Eckstein "Search for lesions in mammograms: Non-Gaussian observer response" *Medical Physics*, vol.31, 24-36, 2004
- ¹² B. Bliznakova, Z. Bliznakov, V. Bravou, Z. Kolitsi, and N. Pallikarakis, "A three-dimensional breast software phantom for mammography simulation", *Physics in Medicine and Biology*, vol. 48, 3699-3719, 2003
- ¹³ J.P. Rolland and H.H. Barrett, "Effect of random background inhomogeneity on observer detection performance", *J. Opt. Soc. Am. A* vol. 9, 649-658, 1992
- ¹⁴ F.O. Bochud, C.K. Abbey, and M.P. Eckstein, "Statistical texture synthesis of mammographic images with clustered lumpy backgrounds", *Optics Express* vol. 4, no.1, 33-43, 1999
- ¹⁵ D. Whitley, "A Genetic Algorithm Tutorial", *Statistics and Computing* vol. 4, 65-85, 1994
- ¹⁶ A.E. Eiden, R. Hinterding, and Z. Michalewicz, "Parameter Control in Evolutionary Algorithms", *IEEE Trans. On Evol. Comput.* vol. 3, 124-141, 1999

- ¹⁷ M. Sonka, V. Hlavak, R. Boyle. *Image processing, Analysis and Machine Vision*, 2nd ed. Pacific Grove: Brooks/Cole, 1999
- ¹⁸ M. Tuceryan and A. K. Jain. "Texture Analysis", in *The Handbook of Pattern Recognition and Computer Vision*, 2nd ed., C.H. Chen, L.F. Pau, P. Wang, Eds., River Edge, NJ: World Scientific Publishing Co, 207-248, 1998
- ¹⁹ R.M. Haralick, K. Shanmugam, and I. Dinstein. "Textural Features for Image Classification", *IEEE Trans. Syst., Man Cybern.*, vol.3, no. 6, 610-62, 1973
- ²⁰ M. Amadasun and R. King, "Textural Features Corresponding to Textural Properties", *IEEE Trans. Syst., Man, Cybern.* vol. 19, no. 15, 1264-1274, 1989
- ²¹ C.B. Caldwell, S.J. Stappleton, D.W. Holdsworth, R.A. Jong, W.J. Weiser, G. Cooke, and M.J. Yaffe, "Characterisation of mammographic parenchymal pattern by fractal dimension", *Phys. Med. Biol.* vol. 35, no. 2, 235-247, 1990
- ²² C. Castella, K. Kinkel, M.P. Eckstein, P.-E. Sottas, F.R. Verdun, and F.O. Bochud, "Automatic mammographic parenchymal patterns classification using multiple statistical features", *IEEE Trans. Med. Imaging*, submitted
- ²³ S. Vedantham, A Karellas, S. Suryanarayanan, D. Albagli, S. Han, E.J. Tkaczyk, C.E. Landberg, B. Opsahl-Ong, P.R. Granfors, I. Levis, C.J. D'Orsi, R.E. Hendrick, "Full breast digital mammography with an amorphous silicon-based flat panel detector: Physical characteristics of a clinical prototype", *Med. Phys.* vol. 27, no. 3, 558-567, 2000
- ²⁴ S. Muller, "Full-field digital mammography designed as a complete system", *European Journal of Radiology*, vol. 31, no.1, 25-34, 1999
- ²⁵ J. R. Taylor, *An Introduction to Error Analysis*, University Science Books, Mill Valley, 1982