Title: Modeling realistic breast lesions using diffusion limited aggregation

Authors: Alaleh Rashidnasab, Premkumar Elangovan, David R. Dance, Kenneth C. Young, Oliver Diaz, Kevin Wells.


Published at: Proc. SPIE 8313, Medical Imaging 2012: Physics of Medical Imaging, 83134L (February 23, 2012);

doi: 10.1117/12.911512

Link: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1386471
Modeling Realistic Breast Lesions using Diffusion Limited Aggregation
Alaleh Rashidnasab, Premkumar Elangovan, David R Dance, Kenneth C Young, Oliver Diaz, Kevin Wells
Centre for Vision, Speech and Signal processing, University of Surrey, Guildford, GU2 7XH, UK
Department of Physics, University of Surrey, Guildford, GU2 7XH, UK
NCCPM, Royal Surrey County Hospital, Guildford, Surrey, GU2 7XX, UK

ABSTRACT
Synthesizing the appearance of malignant masses and inserting these into digital mammograms can be used as part of a wider framework for investigating the radiological detection task in X-ray mammography. However, the randomness associated with cell division within cancerous masses and the associated complex morphology challenges the realism of the modeling process. In this paper, Diffusion Limited Aggregation (DLA), a type of fractal growth process is proposed and utilized for modeling breast lesions. Masses of different sizes, shapes and densities were grown by controlling DLA growth parameters either prior to growth, or dynamically updating these during growth. A validation study was conducted by presenting 30 real and 30 simulated masses in a random order to a team of radiologists. The results from the validation study suggest that the observers found it difficult to differentiate between the real and simulated lesions.

Keywords: Diffusion limited aggregation, DLA, breast cancer, synthetic masses, imaging modality, fractal.

1. INTRODUCTION
Breast cancer can be described as the abnormal division of cells in breast tissues and accounts for approximately 17% of all cancers worldwide. Early diagnosis and treatment of breast cancer is known to be a key factor in reducing mortality rates and has motivated the introduction of national screening programmes in many western countries. Although X-ray film screen imaging has historically been the screening method of choice, there has been a shift over recent years towards using various digital imaging alternatives as these offer a number of advantages over film screen. However the relative merits of these competing technologies, with respect to the detection task is unclear. One approach to address this issue is to develop an image simulation chain in order to conduct virtual trials, presenting pathology to the radiologist that is representative in appearance to that produced on different imaging systems or with different imaging protocols.

Prior work in this area has used the process of random walks in 2D, shape descriptors and quantitative measurements from real masses to simulate the appearance of lesions in mammograms. However, most of these prior methods were designed for realistic appearance of simulated lesions in digitized films. In contrast to this approach, we introduce DLA (Diffusion Limited Aggregation), which is a fractal-like growth process to generate the appearance of lesions when inserted into digital mammograms. The fundamental advantage of using DLA is that the there is far greater control available in terms of the simulated lesion morphology.

2. MATERIALS AND METHODS
2.1 Mass generation using DLA
Diffusion Limited Aggregation is a type of fractal growth process in which particles undergoing random walk (Brownian motion) cluster together and form aggregates. The process may be explained, for simplicity in 2D, by starting with a

*Further author information: (Send correspondence to Alaleh Rashidnasab)
Email: a.rashidnasab@surrey.ac.uk
nominated growth center as the initial mass center. This is surrounded by concentric launch circles in 2D (spheres in 3D). Random walk processes are then launched from these circles, in some prescribed manner (see Fig. 1a). If the path of the random walk touches the growth center, then this adjacent site is labeled as part of the growth center or cluster. If a random walk fails to touch the growing mass center, then it is simply discarded and does not contribute further to the growth process. This process is repeated by launching many random walks from the launch circles in some prescribed manner, and thus the growth center steadily accumulates in size and grows until some terminal condition is reached. Figure 1 illustrates the growth of a DLA cluster alongside some examples of simulated masses grown using the DLA method. The structure of the DLA mass can be controlled by varying the parameters listed in Table 1 (also illustrated in Fig. 1) which provides flexible control of the microscopic appearance of lesion before it is digitally re-sampled and inserted into a mammogram. Each of the growth parameters has a different level of impact on the structure of the DLA cluster. A cluster formed by the DLA process follows well known fractal properties such as self similarity, irregular geometry and fine structure at arbitrarily small scales. DLA fractals can also be measured by the ‘Fractal Dimension’ metric, a statistical quantity that gives an indication of occupancy of a fractal at a given scale.

Figure 1. (a)-(f) The DLA growth process. (g)-(j) A set of four examples of masses grown using DLA but with very different parameter prescriptions.

Table 1. DLA parameters and inference

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Details</th>
<th>Impact on the mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>ρ</td>
<td>Density of points on launching circle/sphere</td>
<td>Texture</td>
</tr>
<tr>
<td>ΔC</td>
<td>Spacing between launch circles/spheres</td>
<td>Size</td>
</tr>
<tr>
<td>ΔS</td>
<td>Step size of random walk</td>
<td>Density</td>
</tr>
<tr>
<td>N</td>
<td>Total number of launching points</td>
<td>Size</td>
</tr>
<tr>
<td>Ps</td>
<td>Sticking probability</td>
<td>Density</td>
</tr>
</tbody>
</table>

For the current experimentation, numerous DLA masses were grown by extending the aforementioned 2D growth process to a three-dimensional lattice (voxel size: 35μm). Parameter selection was guided using the results of a pilot study that provided expert feedback on mass simulation appearance. The most appropriate range of parameters involved simulated DLA masses with $10^5 \leq N \leq 10^7$, $16 \leq \Delta S \leq 64$ and $0.001 \leq P_s \leq 1$ for generating lesions with realistic appearance. While growing certain masses, the step size was dynamically varied in order to generate masses with dense centers. This was accomplished by initiating the growth process with a bigger step size and switching to a smaller step size after the mass had reached a certain size. In some cases, sticking probability was set to a lower value to achieve a dense center. Sticking probability is the probability of a particle undergoing random walk to stick to an occupied site when it comes in contact with it. Due to the fractal nature of DLA and the random walk process, repeated simulation of DLA masses with identical parameters yielded masses with broadly similar shape but different individual appearance. The masses were quantified by size (S) and fractal dimension (D). The size was estimated by constructing a bounding sphere that encompasses all the points of a DLA mass and using the resulting diameter of the fitted sphere. The fractal dimension...
was estimated from the radius of gyration plot, which shows how the average distance between the points of a DLA mass and the centre varies with increasing N (launch points). The relation follows a power law with a positive exponent (G(r) \sim r^{-\beta}). If \beta is the slope of the curve on a log scale, then the fractal dimension is given by (1/\beta). Masses were simulated with an apparent size range between 6mm and 15mm and the fractal dimension range between 2.7 and 2.9. The lesions were inserted into raw clinical mammograms at random locations using the approach explained in section 2.2.

2.2 Insertion

Each mass was grown within a pre-defined 3D array that was sufficient large to ensure that the resulting mass would not experience any boundary effects near the edge of the array. In order to insert the mass with the appropriate contrast, Volpara software was utilized to estimate the local breast composition from the raw digital mammogram prior to mass insertion. Volpara generates a density map of the raw mammogram with values representing the volume of the dense tissue above each pixel point. Using metadata available from the DICOM header (breast thickness, source-to-detector (SID) and spectrum information) the relative increase in attenuation due to the mass at the insertion site was estimated.

The resulting 2D array of transmission factors was then filtered employing the method described by Yip et al, using the MTF (Modulation Transfer Function) and pixel sampling representative of the image acquisition process of an Hologic mammography system (Bedford, Massachusetts, USA). In order to account for scatter at the insertion site, a look-up table of SPR (Scatter-to-Primary Ratio) values was generated from Monte Carlo simulation data for a range of breast thicknesses and glandularities. Scatter was accounted for by first estimating the scatter present at the insertion site, using the aforementioned metadata and then initially subtracted from this region in order to estimate an appropriate transmission factor of the primary beam. The transmission factor of the primary was then modified to account for the relative increase in attenuation due to insertion of the mass. Only then was the scatter added back into the image to ensure that both processes (transmission and scatter components) were appropriately accounted for, and thus realistic mass contrast was obtained. A detailed description of the insertion process, is given in Rashidnasab et al. The entire mammogram was then passed through the Hologic LORAD FFDM Selenia V5.0 image processing package to produce a digital mammogram suitable for viewing.

2.3 Validation Study

The raw mammograms with inserted DLA masses were processed using manufacturer image processing software and compared with mammograms containing real masses. The processed mammograms with inserted DLA masses were presented to 7 expert radiologists for independent observation wherein the lesions were a-priori identified by specifying their approximate location in the mammogram. A total of 60 mammograms (30 pathology proven malignant cases and 30 normal mammograms with simulated masses) were presented in random order to the readers. The dataset was individually randomized for each reader to make sure every mammogram was equally likely to appear at the start, middle or end of the dataset. This was to avoid bias due to observer fatigue. The average clinical screening experience of the observers who participated in the validation study was around 6 years. Observer assessment required rating the realism of the mass on a 5 point scale categorized as: 1.definitely simulated, 2.probably simulated, 3.indeterminate, 4.probably real and 5.definitely real, and the level of suspicion as: 1.normal, 2.benign, 3.indeterminate, 4.suspicious and 5.malignant. A screenshot of the GUI utilized for the study is shown in Figure 3, which was developed in-house as a Java plug-in to ImageJ.
A selection of inserted DLA masses is shown in Figure 4. Figure 5 shows the histogram of observers’ ratings on the realism of the masses. It is apparent from the nature of the histogram that observers had difficulty in differentiating real masses from those simulated using DLA. Subsequently, ROC analysis was performed to confirm this observation. Table
2 and Figure 6 show the results of ROC analysis in which the False Positive Fraction (FPF) was plotted against the True Positive Fraction (TPF) for each observer, and AUC (Area Under Curve) estimated using the trapezoidal method. The Confidence Interval (CI) was calculated using the method presented in Hanley et al.\(^{13}\) This suggests that for all observers used here, no statistically significant difference was found between ratings of real and DLA simulated masses. The overall mean of the AUC was 0.50±.06, which implies that the likelihood of an observer to correctly identify a simulated lesion is essentially random. Figure 7 shows the histogram of observers’ ratings of the masses in terms of suspicion. It was not surprising to see most of simulated lesions being rated as a ’recall’ as simulated lesions were designed to appear malignant. On the contrary, it was interesting to see that some of the real masses which were originally recalled where not rated as a ’recall’ by the observers; this demonstrates a degree of inter-observer variations which might be expected.

![Figure 5. Global histogram of ratings on the realism of the mass for all the observers.](image1)

![Figure 6. ROC curves of the ratings of the masses per observer.](image2)
Table 2. ROC analysis results

<table>
<thead>
<tr>
<th>Observer</th>
<th>Area Under Curve</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer 1</td>
<td>0.49</td>
<td>(0.34, 0.63)</td>
</tr>
<tr>
<td>Observer 2</td>
<td>0.57</td>
<td>(0.42, 0.71)</td>
</tr>
<tr>
<td>Observer 3</td>
<td>0.43</td>
<td>(0.29, 0.58)</td>
</tr>
<tr>
<td>Observer 4</td>
<td>0.40</td>
<td>(0.26, 0.54)</td>
</tr>
<tr>
<td>Observer 5</td>
<td>0.60</td>
<td>(0.46, 0.75)</td>
</tr>
<tr>
<td>Observer 6</td>
<td>0.53</td>
<td>(0.38, 0.67)</td>
</tr>
<tr>
<td>Observer 7</td>
<td>0.49</td>
<td>(0.34, 0.64)</td>
</tr>
<tr>
<td>Average</td>
<td>0.50</td>
<td>(0.44, 0.56)</td>
</tr>
</tbody>
</table>

Figure 7. Global histogram of ratings on the level of suspicion for all the observers.

4. CONCLUSION

In this paper, we have introduced a new method for generating synthetic masses, Diffusion Limited Aggregation. From the pilot study it was observed that the realism was critically dependant on a combination of parameter selection and nature of the insertion site. Thus, in the current study we used local estimates of the glandularity at the insertion site to control the apparent contrast of the mass. The results of an ROC study suggest that masses generated using DLA may be indistinguishable in appearance from genuine pathologically-confirmed masses.

ACKNOWLEDGEMENTS

This work is part of the OPTIMAM project and is supported by the CR-UK & EPSRC Cancer Imaging Programme in Surrey, in association with the MRC and Department of Health (England). The authors thanks Ralph Highnam for granting access to the Volpara breast volume measurement tool and providing technical assistance. The authors would also like to thank colleagues at NCCPM, the clinical observers at Jarvis Breast Screening Centre, Guildford, St. Georges Hospital, London and Addenbrookes Hospital, Cambridge and specially Dr. Julie Cooke for their invaluable assistance.

REFERENCES