A Software Module for Generating a Synthetic Model of the Breast from an X-Ray Image

Nikola Vasilev, Zheni Mincheva, Ventsislav Nikolov, Kristina Bliznakova

Abstract: *The main goal of this paper is to create a methodology for generation of synthetic 3D model from a digitalized X-Ray photo. Three-dimensional perspective gives more detailed and sophisticated in depth for analyzing different irregular lesions. In this research, the application is for detection and analysis in mammography realm.*

Keywords: *breast modeling; computer vision; image recognition; mammography, synthetic modeling.*

I. Introduction

Nowadays, the standard methods for detection and diagnosis of breast cancer have several disadvantages, especially not being sophisticated enough, which sometimes lead to false-positive backs or even death. The idea of the following software module is to generate a synthetic three-dimensional (3D) model of the breast, which is filled with geometrical primitives and is dedicated to early detection of breast malignancies. The computational model will be used to generate an X-Ray synthetic images which will be further subjected to various statistical analyses. In x-ray medical imaging, the formation of the x-ray image is based on the absorption of the body tissues which depends on both the density and the elemental composition of the tissues. The goal of the x-ray images is to screen and diagnose an injury or disease conditions. At the same time, however, the x-ray beam, which is used to produce the image, is harmful for the human body especially if the exposure to these rays is frequent. With the computer simulation and the use of synthetic models in virtual studies, unlimited tests and radiation could be generated which solves the problem with the need to use patients in these studies.

II. Materials and Methods

A. Algorithm overview

The fundamental part for performing diagnostic radiology and radiotherapy is the three-dimensional object modeling [1]. Every object is characterized by its mathematical attributes. The material properties include specific characteristics such as density, mass, weight, elemental composition, etc. The synthetic model is represented by simple geometrical primitives like sphere, cone, cube, cylinder, etc. Its shape is based on the boundaries of shapes by mathematical equations [2]. For the creation of the model, there are several tasks that should be accomplished before the automation process begin. The architecture of the module is shown in Fig1.

978-1-7281-2574-9/19/$31.00 ©2019 IEEE

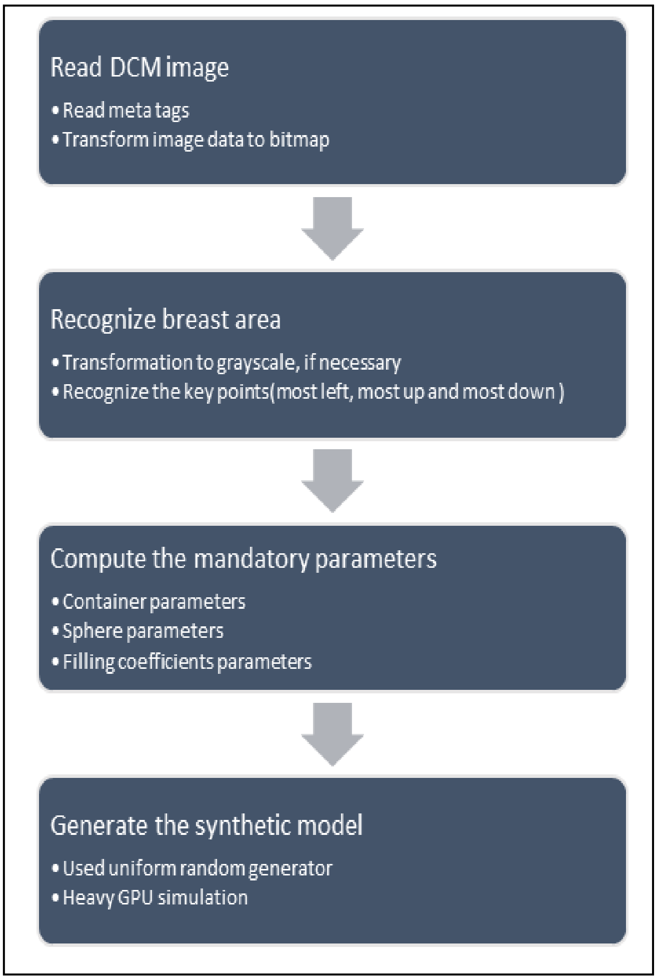

Fig. 1. Module workflow

B. Reading DCM image

The first step is to read a digital X-Ray image [3]. Nowadays most commonly used format is Digital Imaging and Communications in Medicine (DICOM). In addition, this format contains tags, which have important information as physical scaling, kilovoltage peak, tissue thickness, organ dose, etc.

Reading of the image data is done by third party java library (ImageJ) and it is not part of the current research. The major tasks for this stage are reading the image and converting it in suitable for algorithms format (e.g. bitmap) and obtaining the key data tags.

C. Recognizing the breast area

Transformation from image to 3D breast model is done by image analysis and processing. The image is considered as a picture taken by an appropriate tool (for example x-ray emitting tubes), observed by a camera and it is the intensity of the rays passed by the breast organ. Formally, the image is represented as a two-dimensional function:

$$I = f(x, y) \quad (1)$$

where x and y are the coordinates of a given point and I is the brightness of this point.

Thus, the character of the changes of the points of f(x, y) determines the details of the environment under interest. If the image is colorful, then the first step is transformation to a gray scale image by calculating the mean squared value of the color components: red, green and blue and assigning this value to these color components:

$$m = \sqrt{\frac{r^2 + g^2 + b^2}{3}} \quad (2)$$

$$r = m; g = m; b = m;$$

where m is the mean squared value and r, g and b are respectively the values of each Red, Green and Blue components of the pixel. If components values are equal to the mean squared the image is already gray scaled.

When every pixel of the image is transformed into the grayscale, an automatic search is done of a threshold t. A new image is then created with the same width and height as the original one. All pixels with value m (bigger than the threshold value) are colored in white in the new image and all other are colored in black. In the new image then the set of the white pixels must be as connected to one island as possible. This means that there must be as minimum islands of black pixels surrounded by white pixels. Then the white pixels are analyzed, and the most left white pixel coordinates are taken as well as the longest distance from top to bottom within the breast. That parameters are used to build the mathematical model (the center of the sphere, its radius, etc.).

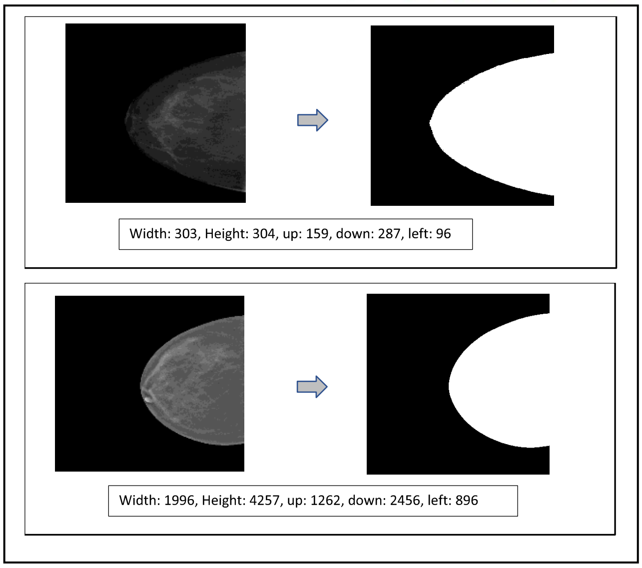

Fig. 2. Example of transformation of input images to white islands images

Examples of x-ray images and their transformation into images with black and white pixels only are shown in Fig 2. There is also information about the size of the picture and the positions of the very left, up and down white pixels. By convention in images processing the coordinate start point is in the upper left corner. In this case the X axis values are increasing from left to right and Y axis values are increasing from up to down.

D. Computing the mandatory parameters

Parameters that are taken from the DICOM image tags are the *Body Part Thickness* and the *Pixel Spacing*. The thickness is given in millimeters corresponds to height of the model. The next parameter pixel spacing describes the ratio of physical to logical distance. It is later used for the calculation of logical value of other parameters. Expression of the physical value of any argument consists of corresponding logical value and physical to logical distance ratio. In most cases the big logical value (e,g, 3000) is multiplied by the small ratio (e.g. 0.01) to get eligible physical value:

$$V_{Ph} = V_L.C \quad (3)$$

where V is physical and logical values and C is the physical to logical distance ratio.

The radius of the model is taken from the X-ray image after its transformation into new image that has only black and white pixels. In the new transformed image, white pixels are parts of the tissue. The area of the image that belongs to the breast is estimated using CUDA cores and Monte Carlo algorithm (the environment used for the implementation is Visual Studio 2017 Community). The total count of pixels in the image is the width multiplied by the height. Random numbers are generated and used for the indexing. Indexed pixels are checked if they are white (in the breast area). With other words, the estimation of the breast area is based on the ratio of the count of the white pixels(breast area) and count of all pixels(image resolution - width multiplied by height) :

$$S = S_1 . \frac{q}{p} \quad (4)$$

where S is the area of the breast, S_1 is the total count of pixels, p is the count of random generated pixels and q is the count of white pixels from p.

For the calculation of the physical radius of the semi-cylinder is used the formula for circle area ($A = \pi r^2$), where A is the area of the circle. For semi-circle this formula is the following:

$2S = \pi r^2$

where S is the area of the semi-circle and r is the radius.

The last equation is used for the evaluation of the logical size of the radius. To get the physical size of the radius it should be multiplied by the ratio of the physical to the logical distance. The final expression of the physical size of the radius is given by:

$$R = C\sqrt{2\frac{S}{\pi}} \quad (5)$$

where R is the radius of the semi-cylinder in physical measure, S is the count of the white pixels and C is the physical to logical distance ratio [4].

Representing tissue with spheres is a common practice in phantom synthesis [5]. Most of the phantoms consist of spheres made of material that has properties similar as breast tissue. For instance, the physical phantom of Cockmartin *et al* [6] is based on Gang et al work's, which proposed that same volumes of different sized acrylic spheres provide an abstract representation of real biological tissues. For the generation of model are used six types of spheres with different radii. In this paper, the spheres radii are 7.94, 6.35, 4.76, 3.17, 1.59 and 0.79 in physical measure.

Fig. 3. A physical phantom dedicated for 2D and 3D breast imaging [7]

III. Results

Development of 3D computer model of the human breast includes modelling of two semi-cylinders, consecutively filled with 6 different types of spheres[8]. They are of different size and color, analogous to individual mammary glands. The generation of the 3D model is a heavy filling algorithm and thus is used Graphical Processor Unit (GPU) for it. Computing power which GPU provides is cheaper and faster, especially for tasks, which are suitable for massive parallelism. The used technology for the generation is CUDA[9], because it provides efficiency and development tools, which helps the overall progress of building and maintaining the module. Furthermore, with more detailed and larger model, the conventional solution with CPU are not satisfactory enough because the lack of efficient random number generator as well as the computation time for the generation.

The generation process is based on following constrains:

1. Six types of spheres should be placed randomly within a half cylinder, where the total volume for every sphere type should be the same:

$$\sum_{i=1}^{n} V_{1i} = \sum_{j=1}^{m} V_{2j} = \sum_{k=1}^{o} V_{3k} = = \sum_{r=1}^{p} V_{4p} = \sum_{s=1}^{t} V_{5s} = \sum_{u=1}^{h} V_{6u} \quad (6)$$

,where:

- n – number of spheres with radius r1;
- m – number of spheres with radius r2;
- o – number of spheres with radius r3;
- p – number of spheres with radius r4;
- t – number of spheres with radius r5;
- h – number of spheres with radius r6;

2. Every sphere should not be included either intersected with other spheres;
3. The random sphere packing is giving a 64% filling quote, as it described in [10].

The described algorithm becomes more complex and need more computation power, because along the filling process, the free space in which will be placed newly generated spheres becomes smaller and smaller and several random trials will be discarded. The spheres are generated from these with largest diameter to the smallest. An example of the filling phases is shown in figure 4.

The architecture of the generation is slightly different from the conventional solution[11] and it is using the following approach:

1. The filling of the container is sequentially per type and it is performed using several stages;
2. Within every stage, the processor thread generates a random sphere;
3. The generated sphere is firstly checked against all spheres in the container from previous stage, which stays unchangeable, and then a try to place among container constrains. Steps 2 and 3 are performed in parallel by all the threads;
4. All successfully placed spheres are then checked against each other and then placed into the model. This is a sequential task, but the average number of successfully placed spheres is small;
5. Each stage ends when the desired number of generated spheres is satisfied. The process is repeated from point 2 for all stages.

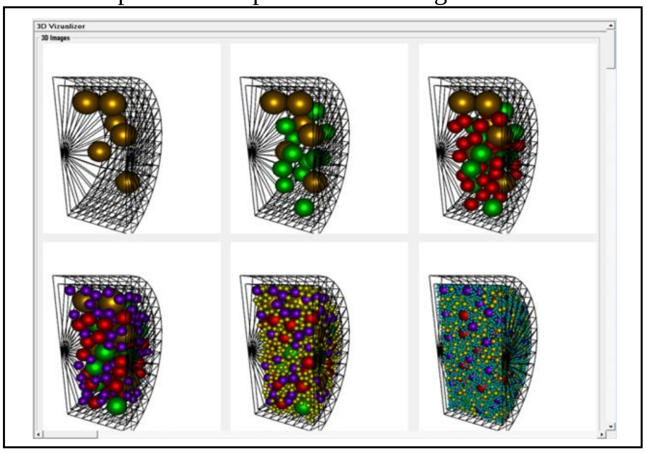

Fig. 4. Filling algorithm result

The random sphere packing task was run on 3 hardware platforms: CPU with 8 cores, GPU with 160 cores and GPU with 512 cores. The comparison is performed against a CPU whit 1 core where the calculation time is 30 hours. The sphere packing time for the container above enclosing 6 000 spheres was 20 sec. on the third platform. Test results are shown on table 1.

TABLE 1. PERFORMANCE BENCHMARKS

Processor unit	Cores@Clock rate	Time	Faster than single core
CPU – intel i7 6700K	8 cores@4.0GHz	300 sec.	x 8
GPU – nVidia GTX960M	5 blocks x 32 cores@1097MHz	80 sec.	X 50
GPU - nVidia GTX1070	16 blocks x 32 cores@1506MHz	20 sec.	X 200

The most important task for the further development is the improvement of the model. The phantom contains six types of spheres that represent adipose tissue, so the next step is generating authentic mammary gland represented by simple stereometric shapes. Once a 3D model is created it can be irradiated using different techniques giving the opportunity to find more precise, clear, detailed and nonradioactive way of breast scanning. It also can be used as an authentic container of synthetic model of breast tumor for research with new treatments that is impossible in real world.

IV. CONCLUSION

By using complex mathematical analysis, the targeted synthetic models are used for generation of abnormalities and they could be watched depending on different circumstances (medicaments, treatments, different tissues, etc.). Using the GPU computation power, we create possible aspect to make a précised analysis for the breast cancer with ad-hoc time.

ACKNOWLEDGMENT

The module prototype is part of the Phantom Explorer, which is performed by Eurorisk Systems Ltd. and it is based on advanced IT technologies: rule based, neural network prediction, multi-core CPU/GPU, 3D images and visualization. The object modeling is supported and powered by biomedical engineering community in Varna represented by Dr. Kristina Bliznakova and her team of researchers at the Laboratory of Computer Modelling and Simulations in Medicine at the Technical University of Varna. The involvement of Kristina Bliznakova was also supported by the Bulgarian National Science Fund under grant agreement DN17/2.

REFERENCES

[1] Bliznakova K et al. (2003) A 3D breast software phantom for mammography simulation. Phys. Med. Biol. 48:3699-3721.

[2] Ruschin M. (2005) Using simple mathematical functions to simulate pathological structures – input for digital

mammography clinical trial. Radiation Protection Dosimetry
[3] Yaffe MJ, Digital mammography in Physics and Pshophyssics, Handbook of medical imaging, SPIE Bellingham, WA, 2000
[4] Monte Carlo Estimation of Area (2014) http://demonstrations.wolfram.com/MonteCarloEstimationOfArea/
[5] Kiarashi, N., A. C. Nolte, et al. (2015). "Development of realistic physical breast phantoms matched to virtual breast phantoms based on human subject data." Med Phys 42(7): 4116-4126.
[6] Cockmartin, L., N. Marshall, et al. (2012). Design and evaluation of a phantom with structured background for digital mammography and breast tomosynthesis. Philadelphia, PA. 7361 LNCS: 642-649
[7] Bliznakova K., Buliev I., Bliznakov Z. (2018) Anthropomorphic Phantoms in Image Quality and Patient Dose Optimization, Institute of Physics Publishing, 0750313242, 9780750313247
[8] Carton, A. K., P. Bakic, et al. (2011). "Development of a physical 3D anthropomorphic breast phantom." Med Phys 38(2): 891-896.
[9] NVIDIA CUDA Runtime API (2019) https://docs.nvidia.com/cuda/cuda-runtime-api/
[10] MathWorld Sphere Packing (2019) http://mathworld.wolfram.com/SpherePacking.html/
[11] Lazos D, Bliznakova K, Kolitsi Z and Pallikarakis (2000), A software data generator for radiographic imaging investigations, Comput. Mehods Prog. Biomed

About the Authors

Nikola Vasilev, Zheni Mincheva, Dr. Ventsislav Nikolov

31, Eurorisk Systems Ltd, General Kiselov Street, 9002
Varna, Bulgaria

E-mail: nvasilev at euroriskssystems dot com

Dr. Kristina Bliznakova

Department of Computer Science and Engineering,
Technical University of Varna, 1 Studentska str, 9010
Varna, Bulgaria

Laboratory of Computer Simulations in Medicine,
Technical University of Varna, 1 Studentska str, 9010
Varna, Bulgaria