

SEGMENTATION OF SPICULATED MASSES IN MAMMOGRAPHIC IMAGES

Sègbédji R. T. J. Goubalan ^{*,†}

Yves Goussard ^{*}

Hichem Maaref [†]

^{*} Ecole Polytechnique de Montréal, Université de Montréal, Montréal, Canada

[†] Laboratoire IBISC, Université d'Evry Val d'Essonne, Evry, France

email: {segbedji.goubalan, yves.goussard}@polymtl.ca, Hichem.Maaref@ibisc.univ-evry.fr

1. INTRODUCTION AND PROBLEM STATEMENT

Breast cancer is among the leading causes of cancer deaths for women. In 2012, 522 000 deaths have been recorded worldwide, representing a 14% increase compared to 2008. Generally, the benign masses class is associated with the presence of circular or oval shapes, while spiculated masses are more likely to belong to the malignant masses class. Thus, spicule is a leading discriminant factor in the classification of various masses. Its extraction is a complex task because of their low contrast, variable widths and the overlapping of blood vessels, fibers and ducts.

With the increase of images obtained during the screening, mammography interpretation by radiologists is becoming more difficult, time-consuming, and leads sometimes the increase the ratio of false positives due to tissue superimposition. Hence, to help radiologists improve detection and diagnosis accuracy the design of computer aided detection systems (CADE) known as a great leap forward in recent years due to their ability to provide an objective and reproducible second opinion. The CADE are structured in three steps: segmentation of the region of interest which contains the mass and description of the segmented mass.

Based on these steps, the proposed method for automatic breast mass detection can be described as follows: the segmentation is based on MRF using the Pickard random field (PRF) which is much faster, more robust and nearly unsupervised compare to most of MRF-based methods, which require complex and time-consuming computations [1, 2]. The description step presents the main contribution of this paper, since, to the best of our knowledge, we present the first attempt to extract the spicules with the mixture of a Markovian framework and an *a contrario* model.

2. MASS SEGMENTATION BASED ON PRF

Segmentation is a crucial step in all CADE systems used in mammography, as the segmented image is used as a base for subsequent processing such as feature extraction and eventually detection and classification of suspicious breast masses.

Several segmentation techniques suited for mammogram processing have been proposed. In this paper, we focus on MRF-based methods due to their ability to take into account the local interactions between neighboring pixels that define different parts of the image through the specification of local energies. As a consequence, these techniques generally perform well; for over two decades, several studies using various types of MRFs for segmentation and mass detection have been presented.

Clearly, a critical point in the development of MRF-based segmentation methods is the determination of an appropriate trade-off between sophistication of the model and the practicality of the resulting algorithm; classical choices include the multilevel logistic model used by Szekely *et al.* [1], or the Potts model which is used by Suliga *et al.* [2] for example. Then, in accordance with the probabilistic nature of MRFs, a maximum *a posteriori* (MAP) approach is generally used to perform segmentation. Unfortunately, maximization of the posterior likelihood is made difficult by its nonlinear structure and by the discrete nature of some unknown variables. To perform the actual optimization, one may turn to simulated annealing which can be shown to converge to a global optimum, but generally involves large and complex computations. This is why many authors [1, 2] revert to the simpler iterated conditional mode (ICM) algorithm, which is suboptimal and may not converge to a global optimum. In either case, automatic determination of the model parameters, also referred to as *hyperparameters*, is either a difficult and time-consuming task, or just cannot be solved numerically.

These elements indicate that most segmentation techniques mentioned above earlier are far from suitable for use in a clinical environment, because they lack simplicity, robustness and efficiency and because they are generally supervised. So, while MRF-based techniques are potentially very powerful, they involve complex computations and are time-consuming. In this work, we attempt to overcome some of these limitations by using a particular MRF known as PRF. In addition, the PRF model is suitable to derivation of almost unsupervised segmentation techniques, as only the number of labels has to be specified beforehand; moreover, in contrast to other methods, initialization of the procedure is straightforward, and the overall amount of computation is greatly

This work was supported by Mitacs and Campus France (Research Grant IT5501).

reduced with respect to standard MRF-based approaches. The presentation of the main characteristics of the proposed method (denoted *log-PRF*) and all technical and algorithmic details can be found in [3, 4].

3. BREAST MASS SPICULES EXTRACTION

The second part of the proposed method is based on several assumptions on the structure of spicules. The four more important ones are the following: (a) In 2-D mammography, the spicules, which are almost straight lines at a small scale in 3-D, have a curvilinear aspect due to tissue superimposition and partial occlusions caused by projection of a 3D object onto a 2-D plane; (b) in a whole mammogram, the spicules are not the most elongated structures due to the presence of blood vessels, mammary and lymphatic ducts, fibers and other structures; (c) all segments present in the image do not necessarily belong to the same structure, the discriminating factors being constraints on the orientations of neighboring segments and distance between their end-points; (d) spicules are structures which intersect the mass boundary and converge toward the mass center.

In order to make use of the above assumptions, the proposed method proceeds along the following steps: first, the mammogram is separated into patches onto which the curvilinear structures are discretized into segments. Then, Markov modeling and contextual information are used to refine the segment positions and associate segments into curvilinear structures. Finally, spicules are detected based on geometric assumptions (b) and (d). The first part of our algorithm is similar to some research contributions in remote sensing. In order to perform detection of the spicules among all modeled curvilinear structures, an *a contrario* modeling was developed to the spicules detection. All technical and algorithmic details about the spicules segmentation method can be found [5].

4. RESULTS AND DISCUSSION

In order to provide a visual evaluation of the efficiency of log-PRF method, we compare it to a recent and efficient segmentation approach using a deep learning method [6]. Result example is presented in Figure 1. It can be observed that, despite the density in which the mass is embedded, our method is able to extract the spiculated margins of malignant masses fairly accurately, so that the result is very close to the manual segmentation. Figure 2 shows the result of the proposed spicules extraction approach. More details about the log-PRF and spicule segmentation method can be found respectively in [3, 5].

5. REFERENCES

[1] N. Szekely, N. Toth, and B. Pataki, "A hybrid system for detecting masses in mammographic images," *IEEE TIM*, vol. 55, no. 3, pp. 944–952, 2006.

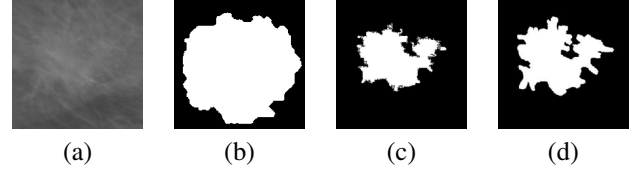


Fig. 1. Malign breast masses segmentation : (a) ROIs, (b) automatic segmentation result obtained in [6], (c) log-PRF automatic segmentation and (d) ground truth.

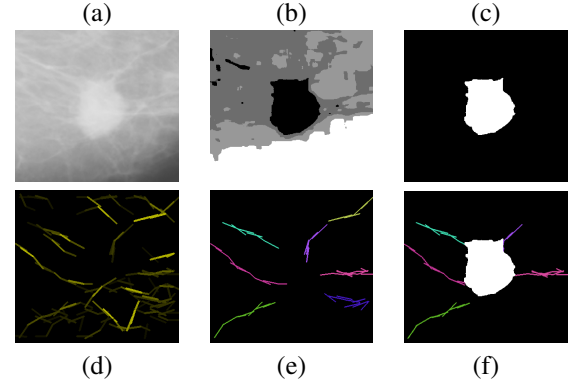


Fig. 2. Annotation of spicules : (a) mass ROI, (b) mass segmentation via log-PRF algorithm, (c) binary version of (b), (d) curvilinear structure discretization, (e) annotated spicules and (f) result of fps reduction.

[2] M. Suliga, R. Deklerck, and E. Nyssen, "Markov random field-based clustering applied to the segmentation of masses in digital mammograms," *Computerized Medical Imaging and Graphics*, vol. 32, no. 6, pp. 502–512, 2008.

[3] S.R.T.J. Goubalan, Y. Goussard, and H. Maaref, "Unsupervised malignant mammographic breast mass segmentation algorithm based on pickard markov random field," in *ICIP*. IEEE, 2016, pp. 2653–2657.

[4] Y. Goussard, J. Idier, and A. De Cesare, "Unsupervised image segmentation using a telegraph parameterization of Pickard random fields," in *ICASSP*, Munich, Germany, 1997, IEEE, vol. 4, pp. 2777–2780.

[5] S.R.T.J. Goubalan, Y. Goussard, and H. Maaref, "Markovian approach to automatic annotation of breast mass spicules using an a contrario model," in *IWDM*, 2016, pp. 461–468.

[6] N. Dhungel, G. Carneiro, and A. P. Bradley, "Tree re-weighted belief propagation using deep learning potentials for mass segmentation from mammograms," in *ISBI*. IEEE, 2015, pp. 760–763.