

# TOROIDAL GAUSSIAN FILTERS FOR DETECTION AND EXTRACTION OF PROPERTIES OF SPICULATED MASSES

*Mehul P. Sampat<sup>1</sup>, Alan C. Bovik<sup>1</sup>, Mia K. Markey<sup>1</sup>, Gary J. Whitman<sup>2</sup>, Tanya W. Stephens<sup>2</sup>*

<sup>1</sup>The University of Texas at Austin, Austin, TX 78712, USA

<sup>2</sup>The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA

## ABSTRACT

We have invented a new class of linear filters for the detection of spiculated masses and architectural distortions in mammography. We call these Spiculation Filters. These filters are narrow band filters and form a new class of wavelet-type filter banks. In this paper, we show that unmodulated versions of these filters can be used to detect the central mass region of spiculated masses. We refer to these as toroidal gaussian filters. We also show that the physical properties of spiculated masses can be extracted from the responses of the toroidal gaussian filters without segmentation.

## 1. INTRODUCTION

The American Cancer Society estimates that 211,240 women will be diagnosed with breast cancer in the U.S. in 2005 [1] and 40,410 women will die of the disease. In the U.S., breast cancer is the most common form of cancer among women and is the second leading cause of cancer deaths, after lung cancer [1]. Early detection of breast cancer increases the treatment options for patients and also increases the survival rate. Screening mammography, which is x-ray imaging of the breast, is currently the most effective tool for early detection of breast cancer. Screening mammographic examinations are performed on asymptomatic woman to detect early, clinically unsuspected breast cancer. Two views of each breast are recorded; the craniocaudal (CC) view, which is a top to bottom view, and the mediolateral oblique (MLO) view, which is a side view. Radiologists visually search mammograms for specific abnormalities. Some of the important signs of breast cancer that radiologists look for are clusters of micro calcifications, masses, and architectural distortions. Calcifications are tiny deposits of calcium, which appear as small bright spots on the mammogram. A mass is defined as a space-occupying lesion seen in at least two different projections [2]. Masses are described by their shape and margin characteristics.

However, mammography is not perfect. Detection of suspicious abnormalities is a repetitive and fatiguing task. For every thousand cases analyzed by a radiologist, only 3 to 4 cases are malignant and thus an abnormality may be overlooked. As a result, radiologists fail to detect 10-30% of

cancers [3–5]. Computer-Aided Detection (CADe) systems have been developed to aid radiologists in detecting mammographic lesions that may indicate the presence of breast cancer [6–8]. These systems act as a second reader and the final decision is made by the radiologist. Most studies have shown that CADe systems, when used as an aid, have improved radiologists' accuracy in the detection of breast cancer [9]. Masses with spiculated margins carry a much higher risk of malignancy than other types of masses or calcifications. However, current CADe systems are dramatically better at detecting micro calcifications than masses.

We have been developing new methods for the detection of spiculated masses (SM) and architectural distortions (AD). For this task, we have developed a two step algorithm for the detection of spiculated lesions. In the first stage, we enhance spiculations on mammograms using a new enhancement algorithm. This entails filtering in the Radon domain. The goal of the second stage is to detect the spatial locations where the spicules converge. For this purpose, we have invented a new class of filters and filter banks that are specifically tuned to match the physical structures that define spiculated lesions. We call these as Spiculation filters (SFs) and Spiculation filter banks (SFBs). These filters are narrow band filters, which are designed to match the physical structures of these lesions and form a new class of wavelet-type filter banks. A key aspect of this work is that the parameters of the detection algorithm are based on measurements of physical properties of spiculated lesions. These measurements were made by two experienced radiologists (GJW, TWS). A detailed description of this algorithm can be found in [10, 11]. The goals of this paper are as follows. Previously, we have focused on the detection of the region where the spicules converge. In this paper we show that the gaussian toroidal filters can be used to detect the central mass region of SM. Secondly, we show that the physical properties of SM can be extracted from the responses of these filters. To verify the accuracy of these automatically extracted values, we compared them to the measurements made by the radiologists, using a statistical test of equivalence.

## 2. THEORY

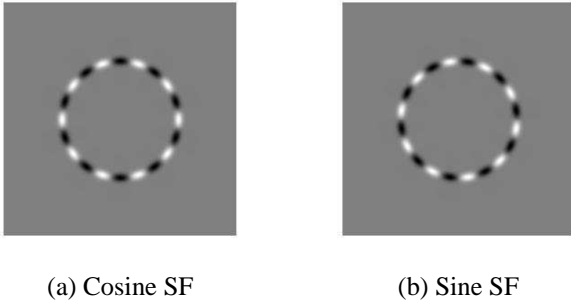
In this section we present a review of the SFs. A SF can be described as a Gaussian-modulated sine torus. The SF consists of quadrature components which are the cosine SF ( $f_{\cosine}(r, \theta; r_0, \sigma, \omega)$ ) and the sine SF ( $f_{\sine}(r, \theta; r_0, \sigma, \omega)$ ). These components are shown in Fig. 1 and are defined as follows:

$$f_{\cosine}(r, \theta; r_0, \sigma, \omega) = g(r; r_0, \sigma) * \cos(\omega\theta) \quad (1)$$

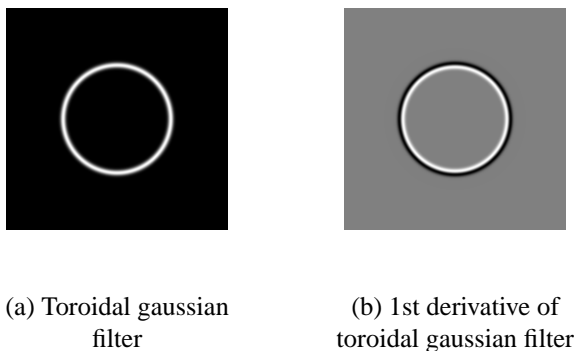
$$f_{\sine}(r, \theta; r_0, \sigma, \omega) = g(r; r_0, \sigma) * \sin(\omega\theta) \quad (2)$$

$$g(r; r_0, \sigma) = \exp[-(r - r_0)^2 / (2 * \sigma^2)] \quad (3)$$

where:  $r = \sqrt{x^2 + y^2}$ ,  $\theta = \arctan(y/x)$  and  $\sigma$  is the standard deviation in pixels. The parameter  $r_0$  is a size parameter measured in pixels, and  $\omega$  is the modulation frequency measured in cycles per circumference. The Gaussian torus is defined by equation (3). The complex SF form a new class of quadrature filters and have the advantage that the sum of the squared responses of the two quadrature components is phase independent. A detailed description of the SFs and filter banks can be obtained in [10, 11].



**Fig. 1.** Example of a Spiculation filter (SF): Figures (a) and (b) show the two quadrature components of the SF, namely, the cosine SF and the sine SF, respectively.



**Fig. 2.** Fig. (a) shows the toroidal gaussian filter and the first derivative of this filter is depicted in Fig. (b).

## 3. METHODS

In this section, we describe the methodology used to detect the central mass regions of SM on the complete mammogram. We also describe the methods to extract the properties of SM from regions of interest (ROIs) containing SM.

### 3.1. Data Description

The images for this study were obtained from the Digital Database for Screening Mammography (DDSM) [12]. The DDSM is the largest publicly available dataset of digitized mammograms and consists of 2,620 cases. Each case consists of four mammograms: a CC and MLO view of each breast. The outlines of the abnormalities are stored in “chain code” and are available in the database. From this “chain code”, borders of the abnormalities can be reconstructed. After reconstructing the border, the interior of the outline is filled to create a “mask” of the mass. A mass is said to be detected if the output of the detection algorithm lies within this mask.

To report the performance of the detection of the central mass region of SM, Free-response Receiver Operating Characteristic (FROC) curves were generated. A FROC curve is obtained by plotting sensitivity on the y-axis and the number of false positives per image on the x-axis. For this study, the different points on the FROC curve are obtained as follows. As mentioned in Section 3.3, we find the top  $N$  maxima for each image. The value of  $N$  was varied from 1 to 25. For each value of  $N$ , the sensitivity and the number of FPI were plotted to obtain the FROC curve.

To evaluate the performance of the central mass region detection algorithm, a set of images consisting of 45 images of SM was used. These images were scanned with a single digitizer, contained a single lesion and were randomly selected. The radiologists (GJW and TWS) had measured the properties on a different set of 12 SM. This set was used to extract the properties of the SM from the filter responses.

### 3.2. Measurements of properties by radiologists

The physical properties of SM measured by the radiologists (GJW and TWS) were: the major axis of the central mass region, the number of spicules and the length and width of spicules. These measurements were made on regions-of-interest (ROIs) containing a SM. We used the ROI Manager plugin of NIH ImageJ (<http://rsb.info.nih.gov/ij/>) to allow a user to make these measurements. A statistical test for equivalence was applied and it was shown that the measurements of the two radiologists were equivalent [13].

### 3.3. Detection of central mass region of SM

Previously, we have used the SFs for the detection of SM, which are marked by converging lines or spiculations. For the detection of the central mass region of SM and to extract

the information about the central mass region of a SM we use the first derivative of the gaussian torus. Let  $g'(r; r_0, \sigma)$  denote the first derivative of the toroidal gaussian (w.r.t  $r$ ) and be defined as:

$$g'(r; r_0, \sigma) = -(r-r_0)/(\sigma^2) * \exp[-(r-r_0)^2/(2*\sigma^2)] \quad (4)$$

Each image is filtered with the filter ( $g'(r; r_0, \sigma)$ ). If the filter “matches” a SM, a peak is obtained at the spatial location corresponding to the center of the SM. Thus, suspicious regions can be identified by detecting the local peaks in the overall output. The larger the output at a particular spatial location, the higher the likelihood that the spatial location corresponds to the center of a SM.

Since masses can be of various sizes, we first filter the image with a set of filters with different radii  $r_0$ . For each pixel, we find the maximum response across the set of filtered images. Thus, we obtain a single image where the intensity at each pixel denotes the maximum response obtained at that particular pixel. Finally, we find the top  $N$  maxima in this image. A pixel is only marked as a potential mass, if it was more than a fixed distance from previously marked points; the distance was set by the average size of a SM measured in our measurement study.

### 3.4. Extraction of physical properties of SM

We first describe our approach to extract the size of the central mass region of the SM. For this task, we use the first derivative of the gaussian torus. Each ROI is filtered with a set of filters with different radii  $r_0$ . If the filter is matched to the size of the mass, the response of the filter at the center of the image would be high. Thus, the size of the mass is the radius of the filter that generates the maximum response. Figure 3 (a) shows a ideal mass, modeled as a disc and the responses of the filters (at the center of the image) is show in Figure 3(b). For this task, we used a set of 18 filters with different radii  $r_0$ . The radius parameter  $r_0$  was varied from 80 to 188 pixels.

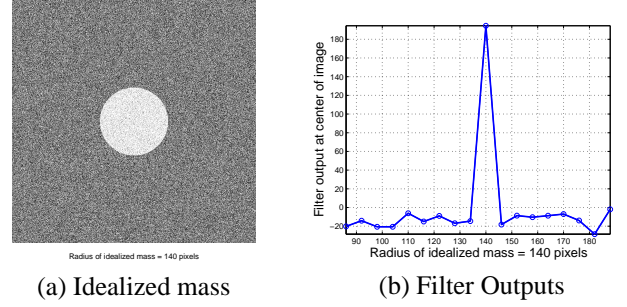
To determine the accuracy of these measurements, we compared them to the physical measurements made by our radiologists (GJW and TWS). This was done by applying a statistical test for equivalence. The null hypothesis ( $H_0$ ) was that the measurements made by the radiologist and the algorithm are not equivalent and the alternative hypothesis ( $H_1$ ) was that they are equivalent. The test statistic ( $t$ ) was computed as follows:

$$t = (\sqrt{n}(\bar{x} \pm \delta)/s) \quad (5)$$

where  $\bar{x}$  = the mean of the differences between the measurements of the radiologist and the algorithm.

$s$  = standard deviation of the differences between the measurement of the radiologist and the algorithm.

$\delta = 0.3$  (or  $0.25$ )  $\times$  mean of the radiologist’s measurement. The factor  $\delta$  accounts for the expected variability, based on our knowledge from other measurement studies in medical

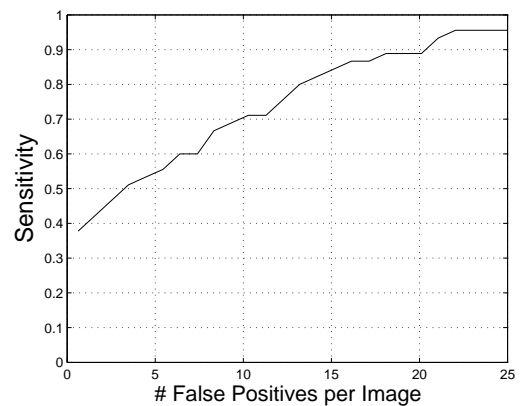


**Fig. 3.** Fig. (a) shows the an idealized mass (radius =140 pixels) with additive gaussian noise. The response of the filters (at the center of the image) of different radii is depicted in Fig. (b). The maximum response is obtained when the filter radius is matched to the size of the idealized mass.

imaging. Thus one can reject the null hypothesis ( $H_0$ ), if the p-value obtained for this statistical test is less than 0.05. Also, note that smaller values of  $\delta$  imply a stricter criteria for achieving equivalence.

## 4. RESULTS

The results of the detection of the central mass region are presented in the form of a FROC curve. This is shown in Fig. 4. A sensitivity of 90% was achieved at 20 FPI. For the extraction of the parameters of the central mass region, a statistical test of equivalence was used to compare the results of the algorithm and the measurements made by the radiologists. Table 1 shows these results. We show that the measurements are equivalent for reasonable choices of the parameter  $\delta$ .



**Fig. 4.** FROC results for the detection of the central mass regions of spiculated masses.

$\delta$	p-value
0.20 * GJW mean	0.024
0.25 * GJW mean	0.005
0.30 * GJW mean	0.001

**Table 1.** Results of the statistical test for equivalence for the measurement of the major axis. The null hypothesis  $H_0$  was that the estimates obtained from the algorithm and the measurements made by the radiologist are not equivalent.

## 5. DISCUSSION

Some researchers have used the difference of gaussian (DoG) filters for the detection masses [14]. The DoG filters and the toroidal filters are band-pass filters but the toroidal gaussian is a more specific narrowband filter. Thus these filters are more specific in their response and respond strongly to masses of a particular size. Also, the DoG filters are not well suited for the extraction of properties of masses because of their lack of specificity.

In conclusion, we have shown that the gaussian toroidal filter can be used for the detection of the central mass region of SM. We also show that these filters can be used to extract physical properties of SM accurately, without segmentation.

## 6. REFERENCES

- [1] "Cancer facts & figures 2005," *American Cancer Society*, 2005.
- [2] American College of Radiology, *ACR BI-RADS - Mammography, Ultrasound & Magnetic Resonance Imaging*, American College of Radiology, Reston, VA, fourth edition, 2003.
- [3] K. Kerlikowske, P. A. Carney, B. Geller, M. T. Mandelson, S.H. Taplin, K. Malvin, V. Ernster, N. Urban, G. Cutter, R. Rosenberg, and R. Ballard-Barbash, "Performance of screening mammography among women with and without a first-degree relative with breast cancer," *Ann Intern Med*, vol. 133, no. 11, pp. 855–63, Dec 2000.
- [4] T. M. Kolb, J. Lichy, and J. H. Newhouse, "Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations," *Radiology*, vol. 225, no. 1, pp. 165–75, Oct 2002.
- [5] R. E. Bird, T. W. Wallace, and B. C. Yankaskas, "Analysis of cancers missed at screening mammography," *Radiology*, vol. 184, no. 3, pp. 613–7, Sep 1992.
- [6] M. L. Giger, "Computer-aided diagnosis of breast lesions in medical images," *Computing in Science & Engineering*, vol. 2, no. 5, pp. 39–45, Sept. 2000.
- [7] K. Doi, H. MacMahon, S. Katsuragawa, R. M. Nishikawa, and Y. Jiang, "Computer-aided diagnosis in radiology: potential and pitfalls," *Eur J Radiol*, vol. 31, no. 2, pp. 97–109, Aug 1999.
- [8] M. P. Sampat, M. K. Markey, and A. C. Bovik, *Handbook of Image and Video Processing*, chapter Computer-Aided Detection and Diagnosis in Mammography, pp. 1195–1217, Academic Press, 2nd edition, 2005.
- [9] T. W. Freer and M. J. Ulissey, "Screening mammography with computer-aided detection: Prospective study of 12,860 patients in a community breast center," *Radiology*, vol. 220, pp. 781–786, 2001.
- [10] M. P. Sampat and A. C. Bovik, "Detection of spiculated lesions in mammograms," in *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, 2003, vol. 1, pp. 810–813 Vol.1.
- [11] M. P. Sampat, G. J. Whitman, M. K. Markey, and A. C. Bovik, "Evidence based detection of spiculated lesions and architectural distortions," in *SPIE Medical Imaging, Image Processing*, Apr 2005, vol. 5747, pp. 26–37.
- [12] M. Heath, K. W. Bowyer, and D. Kopans, "Current status of the digital database for screening mammography," *Digital Mammography*, pp. 457–460, 1998.
- [13] M. P. Sampat, G. J. Whitman, L. D. Broemeling, A. C. Bovik, and M. K. Markey, "Inter- and intra-observer variability in measuring properties of spiculated lesions on mammography," in *Medical Imaging Perception Conference*, 2005.
- [14] W. E. Polakowski, D. A. Cournoyer, S. K. Rogers, M. P. DeSimio, D. W. Ruck, J. W. Hoffmeister, and R. A. Raines, "Computer-aided breast cancer detection and diagnosis of masses using difference of gaussians and derivative-based feature saliency," *Medical Imaging, IEEE Transactions on*, vol. 16, no. 6, pp. 811–819, 1997.