

Fourier irregularity index: A new approach to measure tumor mass irregularity in breast mammogram images

Gensheng Zhang · Wei Wang · Sung Shin ·
Carrie B. Hruska · Seong-Ho Son

© Springer Science+Business Media New York 2013

Abstract Shape descriptors have been identified as important features in distinguishing malignant masses from benign masses. Thus, an effective morphological irregularity measure could provide a helpful reference to indicate the likelihood of malignancy of breast masses. In this paper, a new Fourier-Transform-based measure of irregularity—Fourier Irregularity Index (F^2), is proposed to provide reliable malignant/benign tumor/mass classification. The proposed measure has been evaluated on 418 breast masses, including 190 malignant masses and 218 benign lesions identified by radiologists on film mammograms. The results show the proposed measure has better performance than other approaches, such as Compactness Index (CI),

This research was funded by the MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013.

G. Zhang

Department of Computer Science and Engineering, University of Texas at Arlington, Arlington, TX 76019, USA

e-mail: gensheng.zhang@mavs.uta.edu

G. Zhang

South Dakota State University, Brookings, SD 57007, USA

W. Wang (✉) · S. Shin

Department of Electrical Engineering and Computer Science, South Dakota State University, Brookings, SD 57007, USA

e-mail: wei.wang@ieee.org

S. Shin

e-mail: sung.shin@sdstate.edu

C. B. Hruska

Department of Radiology, Mayo Clinic, 200 1st St SW, Rochester, MN 55905, USA

e-mail: carrie_hruska@yahoo.com

S.-H. Son

Radio Technology Research Department, Electronics and Telecommunications Research Institute (ETRI), Daejeon 305-700, Korea

e-mail: shs@etri.re.kr

Fractal Dimension (FD) and the Fourier-descriptor-based shape Factor (FF). Furthermore, these mentioned measures are paired to investigate the possibility of performance improvement. The results showed the combination of F^2 and CI further enhances the performance in indicating the likelihood of malignancy of breast masses.

Keywords Breast masses classification · Contour analysis · Fourier Transform · Irregularity index

1 Introduction

Mammography is currently the most widely used breast screening technique and the modality shown to achieve the mortality reduction by early detection of breast cancer [2]. However, because of the various presentations of both malignant and benign lesions in the breast, mammographic interpretation is a complicated task requiring considerable domain expertise. It has been shown in previous studies that the rate of cases overlooked by radiologists is still not low enough [25].

The Breast Imaging Reporting and Data System (BIRADS), developed by breast imaging experts in the American College of Radiology, provides standardized terminology for interpretation of lesions identified by mammography [2]. The BIRADS lexicon classifies mammographic lesions into three types: masses, calcifications, and architectural distortion. “A mass is a space occupying lesion seen in two different projections [2].” For masses, the descriptors of shape, margins, and density can be used to further classify the lesion. Shape descriptors have been identified as important features in distinguishing malignant masses from benign masses. In general, a malignant tumor tends to have an irregular or lobular shape, while a typical benign mass tends to have a round or oval shape. The descriptors of “irregular shapes” and “spiculated margins” have previously been shown as two mass features with the highest likelihood of malignancy, with positive predictive values of 73 % and 81 %, respectively [12]. Irregular mass shape correlates pathologically to non-uniform growth and spiculated margins correlate with invasion of surrounding tissue, both common features of malignancy. Thus correlating shape descriptors to malignancy of masses could be an effective approach in early breast cancer detection.

In this paper, a measure to quantify the irregularity is proposed. The measure is based on the irregularity exhibited in contours of breast masses. We believe that the proposed measure can be used as a classifier to classify benign and malignant breast masses, as detailed in following sections. As the irregularity of shapes indicates malignancy of breast masses both in mammogram and in MT images, the results of this preliminary will be helpful to the future study of MT image analysis. Our goal is to build a breast cancer detection system, which is capable to aid medical personnel in interpreting Microwave Tomography (MT) [9] images by utilizing traditional breast imaging techniques and knowledge, such as mammography and Magnetic Resonance Imaging (MRI).

The paper is organized as follows. Section II reviews popular techniques that are widely used in this area. Section III presents the methodologies we use in this study, including the general concept of Fourier Irregularity Index (F^2) developed in this work, introduction of other three classic classifiers, and the method of pairwise combination of classifiers. The evaluation method with the description of the data sets and the evaluation protocol are also described in this section. Section IV compares the results obtained by both proposed measure and reported classifiers. This section also explores the possibility of performance improvement by pairing these classifiers. At the end, section V concludes this work.

2 Related works

Several studies have examined the use of computer algorithms to classify breast masses based on the irregularity exhibited in the contours. A measure of compactness has been used to characterize the diffused or compact nature of abnormalities, rather than roughness [3, 16, 21, 22]. A circumscribed/spiculated classification rate of 92.3 % was achieved by measuring the compactness of shapes [16]. A similar gross contour feature descriptor—tumor circularity, was also reported useful for classifying breast masses [8]. However, the utility of compactness is limited due to its sensitivity to noise along the contour. In this paper, we are going to investigate the performance of compactness instead of tumor circularity because the former is normalized.

The polygonal approximation based fractional concavity (f_{cc}) was used to investigate the morphological characteristics of malignant tumors, which usually have concave and convex segments as well as microlobulations and prominent spicules. Despite the high sensitivity (88.5 %) obtained in a previous investigation, f_{cc} resulted in a poor specificity of 60.7 % because it failed to look into the characteristics of the spicules in terms of their depth and narrowness [17].

Fractal Dimension (FD) is a popular shape descriptor of characterizing complex geometric form of regions-of-interests (ROIs) in mammograms [4, 6, 13, 15, 18]. However, one of the disadvantages of FD is that it is insensitive to large structure indentations and protrusions [10]. Rangayyan et al. showed that FD and f_{cc} can compensate for each other's weakness. The combined descriptor resulted in higher accuracy compared to other descriptors ($A_z=0.93$) [18]. Symmetry is also considered as one of those complex geometric features of ROIs. One popular symmetry measures is Symmetry Distance (SD) [26]. In order to get a symmetrical shape, the minimum distance required to move the points of the original shape is called the Symmetry Distance. However, a recent study [23] showed that it could not distinguish benign tumor from malignant ones effectively.

Even though Fourier Transform has been widely used in image processing, very few were applied to contour analysis. One of them is called Fourier-Descriptor-based shape factor (FF). The FF gives values proportional to energy distribution. For example, a small FF value would be obtained if many energies are distributed in the high-frequency part, indicating that the shape has a rough boundary. Previously reported results showed FF achieved a specificity of 84 % and a sensitivity of 84 % [20]. In our previous work [27], Fourier Transform was also utilized for classifying breast masses based upon their contours. Our preliminary irregularity index, F_{ff} , was used as a quantitative measure of energy distribution of the contour of breast masses. The energy distribution of benign tumors differs from malignant tumors, which makes it possible to distinguish them by using this Fourier-based irregularity index. However, some problems remain unsolved in the previous research, such as the definition of the most regular shape (irregularity index of 0). In this paper, we extend our previous research and propose an improved version of irregularity index— F^2 . The concept of F^2 and the computational methods are described in the following sections.

3 Methodologies

3.1 The proposed Fourier Irregularity Index (F^2)

The breast mass contours are represented by *origin distance*. The origin distance of a contour point is the squared Euclidean distance between the point and the origin (1). In this work, the

origin (x_0, y_0) of a contour is the left-bottom corner of the minimum bounding box of the contour, which has the minimum coordinates on x-y axes.

$$r(t) = (x_t - x_0)^2 + (y_t - y_0)^2 \quad (1)$$

where

$$(x_0, y_0) = \left(\min_t x_t, \min_t y_t \right) \quad t = 0, 1, 2, \dots, N-1$$

It is obvious that $r(t)$ is periodic with each traversal (clockwise or counter-clockwise) of the complete contour. By examining the plot of $r(t)$ carefully, more frequent and/or rapid changes can be observed if the shape has more roughness along the contour. In terms of frequency domain representation, the $r(t)$ sequence has a certain amount of high frequency components. In other words, some energy of the $r(t)$ sequence are distributed in high frequency components. To investigate this phenomenon, Fourier Transform was applied to analyze the energy distribution in details.

By using Fourier Transform, a Fourier series can be obtained by decomposing $r(t)$ into the sum of a set of sin and cosine functions [24]:

$$r(t) = \sum_{n=-N/2+1}^{N/2} c_n [\cos(nt) + \sin(nt)] \quad (2)$$

where

$$c_n = \sum_{n=-N/2+1}^{N/2} r(t) [\cos(nt) + \sin(nt)]$$

Fourier Irregularity Index (F^2) is defined as:

$$F^2 = 1 - \frac{\sum_{k=1}^{N/2} (\mathfrak{R}_k/k)}{\sum_{k=1}^{N/2} \mathfrak{R}_k} \quad (3)$$

Where

$$\mathfrak{R}_k = |c_k/c_1|, \quad k = 1, 2, \dots, N/2.$$

\mathfrak{R}_k is relative energy contribution to the shape descriptor, which is the modulus of the normalized Fourier coefficient at frequency k . The normalization is performed to alleviate the impact from the size of the shape by dividing the coefficient c_1 , which is one of significant Fourier coefficient reflecting the size of the shape (the radius of circle, as discussed below). That is also the reason why \mathfrak{R}_k is called “relative” energy contribution. Although the size of a breast mass may matter in diagnosis, the consideration of it is beyond the scope of this paper. As you may notice, F^2 only takes half of Fourier series. The reason of this is that Fourier series are symmetric on modulus at frequency 0. Thus we use the positive frequency components to represent the energy distributions of shapes.

By the definition of F^2 , a circle has an index value of 0: its $r(t)$ sequence could be expressed as a cosine function $r(t) = R^2(3-2\sqrt{2}\cos(t/2\pi R))$, where R is the radius of the circle; that means only magnitude of c_0 and c_1 could be larger than or equal to zero, and thus F^2 would be zero in this case. As a result, circles and other shapes with this property are considered as the most regular shapes, or the least irregular shapes. The $r(t)$ sequence of other irregular shapes will have more energy distributed in high frequency components (large n in (2)) because of variations along their contours, thus those shapes will have higher F^2 value. Furthermore, the impacts of noises along the contour are alleviated by the division of k , because the division *the frequencies of components* decreases the weight of the energy of high frequency components. As benign tumors tend to have round or oval shapes, while malignant tumors most likely have irregular shapes, it is reasonable to distinguish them by calculating F^2 values. Take two cases from the database [7] as examples: LEFT_CC view of case “A_1396_1” and RIGHT_MLO view of case “B_3016_1”. The first case (case A) was diagnosed as a benign breast mass with a nearly round shape, and the second case (case B) was diagnosed as malignant breast mass with an irregular shape (architectural distortion). As shown in Fig. 1, unlike case A (solid line), $r(t)$ sequence of case B (dashed line) has more energy distributed in high frequency components. It indicates that the contour of case B is more irregular, or rougher than the contour of case A, so do F^2 values (0.36 vs. 0.13).

Connection with FF [20] As described in the related works and following subsection (III.B), the proposed F^2 shares some commons with FF . For example, both methods use the energy distribution to capture the shape irregularity. One key difference between these two methods is the analysis object. FF analyzes the spatial information, i.e. coordinates, while F^2 analyzes the variations of distance from the contour points to the origin (x_0, y_0) . One advantage of using distances instead of coordinates is that the variations of irregular segments in the contour can be accumulated and amplified during the contour traversal, as shown in Fig. 2. This provides more chances to distinguish regular shapes from irregular shapes. Our experiments also confirm it with more accurate results by using F^2 than FF .

A few properties of F^2 can be also summarized according to the previous analysis and the definition:

- The range of F^2 index lies in $[0, 1]$;
- Circles are the least irregular shapes with the smallest F^2 index of zero.

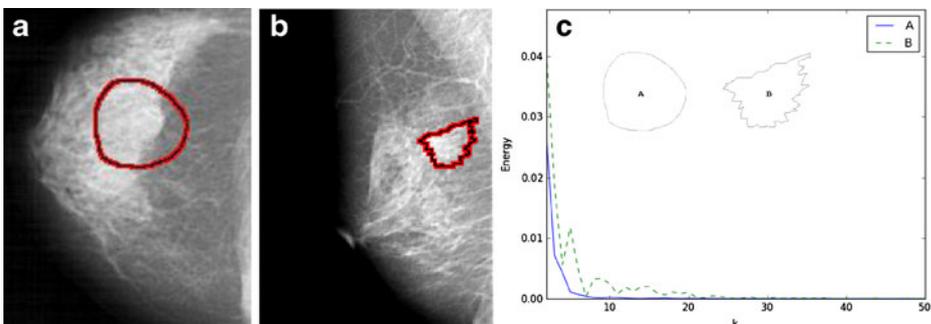


Fig. 1 A user case study. **a** Case A: LEFT_CC view of A_1396_1; **b** Case B: RIGHT_MLO view of B_3016_1; **c** energy distribution of case A and case B: B has more energy distributed in high frequency components than A does, which matches the fact that contour B is rougher than A

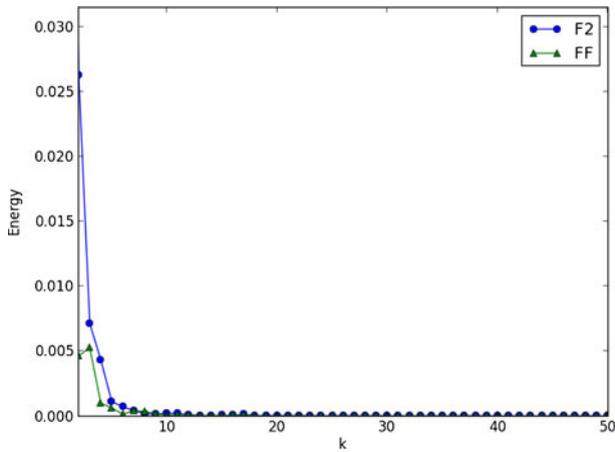


Fig. 2 Energy distribution of case A (LEFT_CC view of A_1396_1): comparing FF and F^2

- Irregular shapes would have bigger F^2 index than regular shapes due to the pattern of energy distribution of $r(t)$ sequences.
- F^2 is able to tolerate noises at some level because the high frequency components are suppressed by the division by *their corresponding frequencies*.

3.2 Other classifiers

Comparative analyses of F^2 are performed with three other previously reported classifiers, including Compactness Index (CI), Fractal Dimension (FD), and Fourier-descriptor-based shape Factor (FF). These classifiers are briefly described in the following subsections.

3.2.1 Compactness Index (CI)

Compactness Index (CI) measures the efficiency of a contour enclosing a given area, defined as $CI=1-4\pi A/P^2$, where A is the enclosed area and P is the perimeter of the contour [11]. Similar to F^2 , circles are defined as the most compact 2D shapes because they have the smallest index of zero. A higher CI value often implies the malignancy of a suspicious area due to concavities or spicules [1].

3.2.2 Fractal Dimension (FD)

Fractal Dimension (FD) was initially developed to characterize complex geometric forms. FD s can be derived from the self-similar dimension D , defined as (4) [14]. D could be estimated by the slope of a plot of $\log(n)$ vs. $\log(1/s)$.

$$n = \frac{1}{s^D} \quad (4)$$

where n is the number of self-similar pieces exhibited in a self-similar pattern and s is related to the measurement scale.

One of commonly used method for approximating FD is the ruler method, which uses different length of rulers to estimate the total length of a contour or pattern to different levels of accuracy [18]. Our implementation results in comparable output to the theoretical values when we tested in with Koch's snowflakes (less than 0.3 % error).

3.2.3 Fourier-descriptor-based shape Factor (FF)

FF was designed to characterize the energy distribution of a contour's coordinate sequence [20]. The x - y coordinates of a contour are expressed as complex numbers in the form of $x+yi$, and then Fast Fourier Transform (FFT) is applied to the complex numbers to get the Fourier descriptor (FD) or the Fourier series ($A(t)$). The FF is defined as:

$$FF = \sum_{t=-N/2+1}^{N/2} \frac{\|E(t)\|}{|k|} / \sum_{t=-N/2+1}^{N/2} \|E(t)\| \quad (5)$$

where

$$E(t) = \begin{cases} 0; & k = 0 \\ A(t)/A(1); & k = 1, 2, \dots, N/2 + 1 \\ A(t+N)/A(1); & k = -1, -2, \dots, -N/2 + 1 \end{cases} \quad (6)$$

3.3 Combination of classifiers

Since each classifier captures different characteristics of shapes, a proper combination of classifiers may lead to better performance if they are complementary to each other. Each two classifiers are combined together to form a new classifier for investigating the possibilities of achieving better performance. In this work, we apply a Radial Basis Function (RBF) network¹ to automatically learn the combination model. Briefly speaking, given two functions f and g (inputs), and expected output, a RBF network is to learn a new function $\varphi(f, g)$ (output), such that the value of φ is close to the expected output. In our work, the inputs to the RBF network are the results of two classifiers, and the expected outputs are the diagnosis results of breast masses (benign/malignant). The output of the network φ is then used as the new (combined) classifier.

3.4 Evaluation methodology

3.4.1 Datasets of mass contours

The Fourier Irregularity Index concept developed in this study is applied to a set of mammograms, which are provided by the Digital Database for Screening Mammography (DDSM) [7]. This database includes digitized film mammograms from approximately 2,500 patients. Each study includes two images (craniocaudal and mediolateral oblique) of each breast. Associated pixel-level "ground truth" information are available including the locations and types of suspicious regions, as well as patient information and image acquisition information. An expert radiologist has provided an interpretation of each mammogram with identified abnormalities. Each abnormality has been described with BIRADS terminology including the lesion

¹ Our implementation is based on package PyRadbas: <http://cybercase.github.io/pyradbas/>

type, lesion description, and final assessment. Corresponding pathology findings are available. Outlines of suspicious regions are provided on the mammograms, derived from markings made on the film by an experienced radiologist. Each lesion boundary is encoded in Freeman Chain Code of Eight Directions [5]. For simplicity in this work, only three types of mass shapes were included in the testing data set: round, oval and irregular. In total, 418 mass were investigated in our study, including 190 malignant lesions and 228 benign lesions, with details shown in Table 1. In this study, no parameter derivation or classifier training was involved in the evaluation procedure; in other words, the dataset was not split into training and testing subsets.

3.4.2 Evaluation protocol

Each of the 418 masses obtained from DDSM is evaluated using the F^2 and three other classifiers: Compactness, Fractal Dimension, and Fourier Factor. To illustrate the characteristics and effectiveness of these classifiers, the sensitivity, specificity, accuracy, and area under the curve (A_z) for Receiver Operating Characteristic (ROC) are computed. An ROC curve is a plot of a test's sensitivity versus its false positive rate (FPR or 1-specificity), which are plotted on the vertical axis and the horizontal axis, respectively. The sensitivity and the specificity are given by (7) and (8). Each point on the curve is obtained with a different decision threshold. The A_z index is usually referred as a useful indicator of diagnostic performance [28]. The range of A_z lies in [0, 1] and larger A_z means better performance ($A_z = 1$ means perfect diagnostic accuracy achieved):

$$\text{sensitivity} = TP_s / (TP_s + FN_s) \quad (7)$$

$$\text{specificity} = TN_s / (TN_s + FP_s) \quad (8)$$

where TP is true positive, FN is false negative, TN stands for true negative, and FP stands for false positive.

In this study, the sensitivity, specificity, and accuracy of each classifier are obtained at the closest point to (0, 1) of the ROC curve. The area under the ROC curve (A_z) is computed by using trapezoidal rule.

3.4.3 Significance of improvement

Significance tests are performed to verify whether a classifier actually outperforms the other. As recommended in [19], a binomial test is used to compare two classifiers, which is based on

Table 1 Dataset of breast mass contours

| Pathology | Shape | Count |
|-----------|--|-------|
| Benign | ROUND | 58 |
| | OVAL | 168 |
| | ROUND-OVAL | 2 |
| Malignant | IRREGULAR | 143 |
| | IRREGULAR-ARCHITECTURAL_ DISTORTION | 47 |
| | Total | 418 |

disagreements on outputs. Specifically, let s be the number of instances that classifier A produces correct outputs, but classifier B produces incorrect outputs ($A > B$). Similarly, f be the number of instances that classifier A produces incorrect outputs, but B produces correct outputs ($B > A$). Among these different outputs, A has chance to perform better than B if the expected value $E(s) > E(f)$. In other words, if A and B perform equally well ($A = B$), then $E(s) = E(f)$. Therefore, our null hypothesis is: $H_0: A = B$. H_0 will be rejected if its likelihood (p) is smaller than the significant level, with which we can conclude classifier A performs better than classifier B. The value of p is computed using the binomial distribution (9).

$$p = \sum_{i=s}^{s+f} \frac{(s+f)!}{i!(s+f-i)!} 0.5^{s+f} \quad (9)$$

4 Results and discussion

Figure 1(a) and (b) show the typical contours of benign and malignant breast mass, respectively. The red-black lines are the outlines for the suspicious regions which are derived from markings made on the mammogram by an experienced radiologist [7]. The shape of benign mass (Fig. 1(a)) appears to be regular and smooth, whereas the malignant mass appears more irregular than the benign one.

To investigate if the aforementioned shape descriptors are suitable for classifying breast masses, significance tests are performed on the statistics of the descriptors. The average F^2 for 192 malignant and 228 benign masses were 0.15 and 0.30, respectively. As it is designed, larger F^2 value indicates more irregular or severer malignancy. The *chi-squared* test between benign and malignant of F^2 produces a significance level of 5% ($p < 0.05$), indicating that F^2 is suitable to effectively distinguish between benign and malignant masses. The average compactness indexes of benign and malignant masses are 0.22 and 0.54, respectively ($p < 0.05$). It shows that there is more irregularity along the contours of malignant masses. The fractal dimension (FD) and the Fourier-descriptor-based shape Factor (FF) also exhibited that there are significant difference between benign and malignant masses, which makes them desirable to characterize breast masses. The average values of FD for benign and malignant masses are 1.06 and 1.13, respectively and that of FF were 0.82 and 0.68 (for both cases, $p < 0.05$).

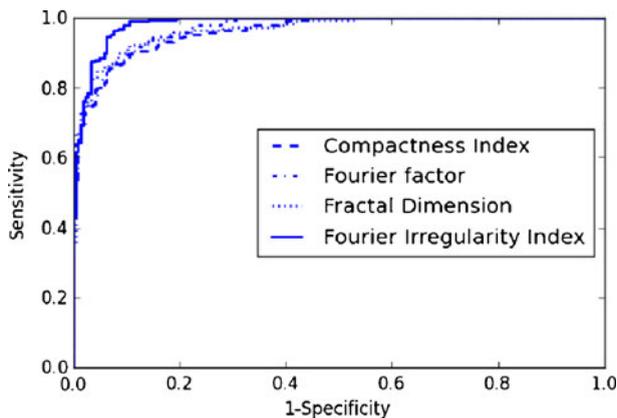


Fig. 3 ROC Curve when applying different classifiers to distinguish benign and malignant breast

Table 2 Performance of each classifier in terms of sensitivity, specificity, accuracy, and area under the ROC curve

| Classifier | Accuracy | Sensitivity | Specificity | A_z |
|------------|----------|-------------|-------------|-------|
| <i>CI</i> | 89.7 % | 0.91 | 0.89 | 0.96 |
| <i>FD</i> | 90.9 % | 0.90 | 0.92 | 0.97 |
| <i>FF</i> | 89.7 % | 0.91 | 0.89 | 0.97 |
| F^2 | 94.3 % | 0.95 | 0.94 | 0.98 |

Areas under the ROC curves (A_z) are computed to evaluate the performance of each classifier for benign/malignant classification. The results are shown in Fig. 3 and Table 2. As mentioned in Section III, the sensitivity, specificity and accuracy are determined by the closest point to (0, 1) on the ROC curve, and the A_z is calculated by using trapezoidal rule. A larger A_z indicates that the classifier achieves better performance in terms of sensitivity or specificity for characterizing benign and malignant breast masses.

As shown in Table 2, the results of the reported classifiers are similar to those in related literatures. Since the contours are tracked by an experienced radiologist, the negative impacts of noises to the compactness index are reduced. As a result, the compactness index gives a relatively high sensitivity (0.91) and a satisfactory specificity (0.89) with A_z index of 0.96. The overall accuracy of *CI* is 89.7 %. *FD* achieves a relatively accurate classification results with an accuracy of 90.9 %. One of the reasons that *FD* has the lowest sensitivity (0.90) is its insensitivity to significant structural irregularity. One of the advantages of *FF*—the tolerance of noise along contours—may be not perfectly presented in this study because the contours used in experiments are manually generated, as mentioned previously. Though the same results were achieved by *CI*, *FF* still yields a higher A_z value, which makes it more advantageous than *CI*. As discussed in Section III, the concepts of *FF* and F^2 are similar to each other: both methods measures irregularity based on energy distribution. One difference is that *FF* uses complex coordinates of contour points, whereas F^2 uses origin distances, which is attributable to the performance gains of F^2 . The variations along the contours are accumulated and amplified, which helps F^2 to recognize the existence of irregularity on these contours. Table 3 shows the significance tests of comparisons of classifiers. As indicated by the small differences in accuracy, *CI*, *FD*, *FF* fail to prove one performing better than another ($p > 0.19$). On the other hand, the results confirms that the proposed F^2 outperforms the other three ($p < 0.01$). In Sum, F^2 performed best among these classifiers with having the highest accuracy (94.3 %), sensitivity (0.95), specificity (0.94), and the largest area under the ROC curve (0.98).

In order to investigate the possibilities of further improvement, each two classifiers are combined as a new classifier in this study. Figure 4 shows the results of classification based on an exact Radial Basis Function Network (RBFN), as described in Section III. Correctly classified benign and malignant masses are denoted by using symbols “■” and “○”,

Table 3 Significance test of performance improvement

| Classifier A | Classifier B | s | f | $p(A > B)$ |
|--------------|--------------|-----|-----|------------|
| <i>FF</i> | <i>CI</i> | 23 | 23 | 0.5585 |
| <i>FD</i> | <i>CI</i> | 22 | 17 | 0.2612 |
| F^2 | <i>CI</i> | 29 | 10 | 0.0017 |
| <i>FD</i> | <i>FF</i> | 13 | 8 | 0.1917 |
| F^2 | <i>FF</i> | 27 | 8 | 0.0009 |
| F^2 | <i>FD</i> | 23 | 9 | 0.0100 |

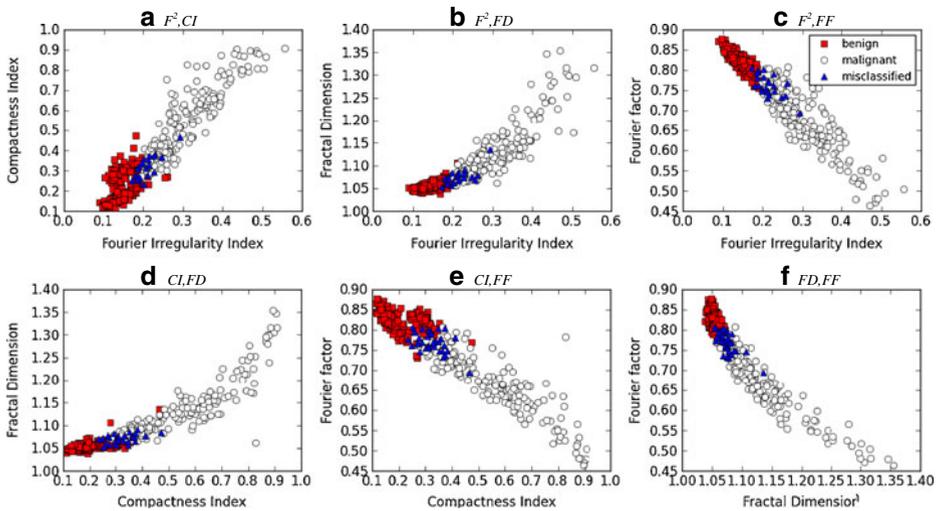


Fig. 4 Classification results of combined classifiers. Symbols “■”, “○”, and “▲” denote data points corresponding to benign, malignant and misclassified breast masses, respectively

respectively, whereas misclassified masses are denoted by symbols “▲”. The results are also presented in Table 4. The sensitivity, specificity, accuracy, and area under the ROC curve are calculated in the same way as described previously.

The performances of the combinations of classifiers are presented in Table 4. Bold data shows that it has the highest rate in performance. As shown in the table, by combining with F^2 , classifiers, CI , FF , and FD , performed better than using them individually. Specifically, the combination of CI and F^2 is improved about 0.7 % in accuracy, 1 % both in specificity, and area under the ROC curve, comparing to use F^2 alone FF provides the highest sensitivity of 0.96 when it is combined with F^2 . As shown in Table 5, the improvement of this combination is not so significant because they have similar physical characteristics (energy distribution), whereas the combination of CI and FF results in the highest improvement in accuracy (about 2.4 %) because they are supposed to capture different characteristics of contours (compactness nature vs. energy distribution). Similar observation can be made on the combination of CI and FD . The result of improvement significance tests confirm the possibility ($w/p < 0.05$) that the combinations of classifiers may perform better comparing to single classifiers. One exception is the comparison of F^2 . The most accurate classifier combination (F^2, CI) perform slightly better than F^2 , but only differed in nine cases ($s = 6, f = 3$), which was not significant

Table 4 Performance of combined classifier in terms of sensitivity, specificity, accuracy, and area under the ROC curve

| Combined classifier | Accuracy | Sensitivity | Specificity | A_z |
|---------------------|---------------|-------------|-------------|-------------|
| F^2, CI | 95.0 % | 0.94 | 0.96 | 0.99 |
| F^2, FD | 94.5 % | 0.95 | 0.94 | 0.99 |
| F^2, FF | 94.3 % | 0.96 | 0.93 | 0.99 |
| CI, FD | 92.8 % | 0.93 | 0.93 | 0.98 |
| CI, FF | 92.1 % | 0.94 | 0.90 | 0.98 |
| FD, FF | 90.9 % | 0.92 | 0.90 | 0.97 |

Table 5 Significance test of performance improvement for combined classifiers

| Classifier A | Classifier B | <i>s</i> | <i>f</i> | <i>p(A>B)</i> |
|--------------|--------------|----------|----------|------------------|
| F^2, CI | CI, FF | 18 | 6 | 0.0113 |
| F^2, CI | CI, FD | 19 | 9 | 0.0436 |
| F^2, CI | F^2, FF | 7 | 4 | 0.2744 |
| F^2, CI | F^2, FD | 7 | 6 | 0.5 |
| F^2, CI | FD, FF | 54 | 9 | 3.1E-09 |
| CI, FD | FD, FF | 44 | 9 | 6.1E-07 |
| CI, FD | CI, FF | 11 | 9 | 0.4119 |
| F^2, FF | FD, FF | 49 | 7 | 3.7E-09 |
| F^2, FF | CI, FF | 18 | 9 | 0.0610 |
| F^2, FF | CI, FD | 17 | 10 | 0.1239 |
| CI, FF | FD, FF | 40 | 7 | 5.4E-07 |
| F^2, FD | FD, FF | 50 | 6 | 5.1E-10 |
| F^2, FD | CI, FF | 19 | 8 | 0.0261 |
| F^2, FD | CI, FD | 16 | 7 | 0.0466 |
| F^2, FD | F^2, FF | 3 | 1 | 0.3125 |
| FD, FF | CI | 19 | 42 | 0.9990 |
| FD, FF | FF | 10 | 33 | 0.9999 |
| FD, FF | FD | 12 | 40 | 1.0 |
| FD, FF | F^2 | 9 | 51 | 1.0 |
| CI, FF | CI | 21 | 11 | 0.0551 |
| CI, FF | FF | 13 | 3 | 0.0106 |
| CI, FF | FD | 15 | 10 | 0.2122 |
| CI, FF | F^2 | 10 | 19 | 0.9693 |
| CI, FD | CI | 22 | 10 | 0.0251 |
| CI, FD | FF | 21 | 9 | 0.0219 |
| CI, FD | FD | 10 | 3 | 0.0461 |
| CI, FD | F^2 | 10 | 17 | 0.9390 |
| F^2, FF | CI | 29 | 10 | 0.0017 |
| F^2, FF | FF | 26 | 7 | 0.0007 |
| F^2, FF | FD | 23 | 9 | 0.0100 |
| F^2, FF | F^2 | 2 | 2 | 0.6875 |
| F^2, FD | CI | 28 | 7 | 0.0003 |
| F^2, FD | FF | 27 | 6 | 0.0002 |
| F^2, FD | FD | 23 | 7 | 0.0026 |
| F^2, FD | F^2 | 3 | 1 | 0.3125 |
| F^2, CI | CI | 30 | 8 | 0.0002 |
| F^2, CI | FF | 28 | 6 | 0.0001 |
| F^2, CI | FD | 25 | 8 | 0.0023 |
| F^2, CI | F^2 | 6 | 3 | 0.2539 |

enough ($p=0.25$) to conclude that the combination (F^2, CI) would actually perform better. We believed that a larger test set would be able to draw a more confident conclusion. In general, we believe the results F^2 achieves make it possible to be applied to benign and malignant masses classification.

5 Conclusion

In this paper we have proposed a novel irregularity measurement (Fourier Irregularity Index: F^2) based on Fourier Transform of the *origin* distance. In our experiments, F^2 has exhibited high classification accuracy, making it possible to be integrated into a computer-aided diagnosis (CADx) system, as an effective factor for benign/malignant breast mass classification. Furthermore, we have also explored the possibilities of performance improvement of classifier pairs. The classifier pairs have been found that they are able to provide higher classification accuracy than the single classifiers. We believe that the outputs of this study and the proposed measurement may provide a valuable reference of malignancy diagnosis of breast masses to radiologists.

6 Future works

As mentioned earlier in this paper, the contours used in this study are tracked manually by a radiologist. Because of this, biases may be produced. Thus, the measurement still needs to be evaluated when dealing with automatic contour extraction algorithms. Furthermore, we still need a larger dataset to gain more confidence before we put it into practice.

References

1. Alto H, Rangayyan RM, Desautels JEL (2005) Content-based retrieval and analysis of mammographic masses. *J Electron Imaging* 14:023016–023016
2. D’Orsi C, Bassett LW, Berg W, Feig S, Jackson V, Kopans D (2003) Breast imaging reporting and data system: ACR BI-RADS-mammography. American College of Radiology (ACR), Reston
3. Davies D, Dance D (1990) Automatic computer detection of clustered calcifications in digital mammograms. *Phys Med Biol* 35:1111
4. Dey P, Mohanty SK (2003) Fractal dimensions of breast lesions on cytology smears. *Diagn Cytopathol* 29: 85–86
5. Freeman H (1961) On the encoding of arbitrary geometric configurations. *IRE Trans Electron Comput* EC-10:260–268
6. Guo Q, Ruiz V, Shao J, Guo F (2005) A novel approach to mass abnormality detection in mammographic images. In: *Proceedings of the IASTED International Conference on Biomedical Engineering*, pp 180–185
7. Heath M, Bowyer K, Kopans D, Moore R, Kegelmeyer WP (2001) The digital database for screening mammography. In: *Proceedings of the Fifth International Workshop on Digital Mammography*, pp 212–218
8. Kilday J, Palmieri F, Fox MD (1993) Classifying mammographic lesions using computerized image analysis. *IEEE Trans Med Imaging* 12:664–669
9. Kwon K-C, Lim Y-T, Kim C-H, Kim N, Park C, Yoo K-H, Son S-H, Jeon S-I (2012) Microwave tomography analysis system for breast tumor detection. *J Med Syst* 36:1757–1767
10. Lee TK, McLean DI, Atkins MS (2003) Irregularity index: a new border irregularity measure for cutaneous melanocytic lesions. *Med Image Anal* 7:47–64
11. Liang S, Rangaraj MR, Desautels J (1993) Detection and classification of mammographic calcifications. *Int J Pattern Recognit Artif Intell* 7:1403–1416
12. Liberman L, Abramson AF, Squires FB, Glassman J, Morris E, Dershaw D (1998) The breast imaging reporting and data system: positive predictive value of mammographic features and final assessment categories. *Am J Roentgenol* 171:35–40
13. Matsubara T, Fujita H, Kasai S, Goto M, Tani Y, Hara T, Endo T (1997) Development of new schemes for detection and analysis of mammographic masses, pp 63–66
14. Peitgen HO, Jürgens H, Saupe D (2004) *Chaos and fractals: new frontiers of science*. Springer Verlag

15. Pohlman S, Powell KA, Obuchowski NA, Chilcote WA, Grundfest-Broniatowski S (1996) Quantitative classification of breast tumors in digitized mammograms. *Med Phys* 23:1337
16. Rangayyan RM, El-Faramawy N, Desautels JEL, Alim OA (1997) Measures of acutance and shape for classification of breast tumors. *IEEE Trans Med Imaging* 16:799–810
17. Rangayyan RM, Mudigonda NR, Desautels JEL (2000) Boundary modelling and shape analysis methods for classification of mammographic masses. *Med Biol Eng Comput* 38:487–496
18. Rangayyan RM, Nguyen TM (2007) Fractal analysis of contours of breast masses in mammograms. *J Digit Imaging* 20:223–237
19. Salzberg S (1997) On comparing classifiers: pitfalls to avoid and a recommended approach. *Data Min Knowl Disc* 1:317–328
20. Shen L, Rangayyan RM, Desautels JEL (1994) Application of shape-analysis to mammographic calcifications. *IEEE Trans Med Imaging* 13:263–274
21. Spiesberger W (1979) Mammogram inspection by computer. *IEEE Trans Biomed Eng BME-26(4)*:213–219
22. Wee WG, Moskowitz M, Chang NC, Ting YC, Pemmeraju S (1975) Evaluation of mammographic calcifications using a computer program. *Radiology* 116:717–720
23. Wei Y, Su Z, Yazhu C, Yaqing C, Wenying L, Hongtao L (2008) Effective shape measures in malignant risk assessment for breast tumor on sonography. In: *International Multisymposiums on Computer and Computational Sciences. IMSCCS '08*, pp 51–56
24. Wikipedia (2012) Fourier series. Available: http://en.wikipedia.org/wiki/Fourier_series
25. Xiangyang X, ShengZhou X, Lianghai J, Shenyi Z (2010) Using PSO to improve dynamic programming based algorithm for breast mass segmentation. In: *2010 I.E. Fifth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA)*, pp 485–488
26. Zabrodsky H, Peleg S, Avnir D (1995) Symmetry as a continuous feature. *IEEE Trans Pattern Anal Mach Intell* 17:1154–1166
27. Zhang G, Shin S, Wang W, Li Z, Choi H. A new Fourier-based approach to measure irregularity of breast masses in mammograms. In *Proceedings of the 2012 ACM Research in Applied Computation Symposium (RACS '12)*. ACM, New York, NY, USA, pp 153–157
28. Zhang G, Wang W, Moon J, Pack JK, Jeon SI. A review of breast tissue classification in mammograms. In: *Proceedings of the 2011 ACM Symposium on Research in Applied Computation (RACS '11)*. ACM, New York, NY, USA, pp 232–237



Gensheng Zhang received his MS degree in Computer Science from South Dakota State University in 2012, USA. His research interests include breast cancer image processing, computer vision and medical informatics. He is currently working on Ph.D. degree in Computer Science in University of Texas Arlington.



Dr. Wei Wang is an Assistant Professor with the department of Electrical Engineering and Computer Science, South Dakota State University, Brookings, SD, USA. He received his Ph.D. degree in Computer Engineering from University of Nebraska—Lincoln, USA, 2009. His major research interests include wireless sensor networks, multimedia computing, information security, and educational robotics. He won two Best Paper Awards of IEEE WCNC 2008 and ANSS 2011. He serves as an Associate Editor of Wiley Security in Communication Networks Journal, the Guest Editor of three Special Issues for Hindawi IJDSN, the workshop co-chair of ICST BodyNets 2013, the program vice chair of ACM RACS 2013, the chair of IEEE CIT-MMC track 2012, the vice-chair of IEEE ICCT-NGN track 2011 and the program chair of the ICST IWMMN 2010.



Dr. Sung Shin is a Full Professor and the graduate coordinator chair with department of Electrical Engineering and Computer Science, South Dakota State University, SD, USA. He also served as an Invited Professor at Seoul National University of Korea in 2007, and a National Aeronautics and Space Administration (NASA) Visiting Scientist at Johnson Space Center at Houston of Texas in 1999. His research interests include software engineering, software fault tolerance, internet data security, distributed real time systems, and telemedicine. He has published over 120 research papers in these research areas. He currently serves as the chair of ACM SigAPP. He has served as a chair for many international conferences and symposiums, such as Association for Computing Machinery (ACM) Symposium on Applied Computing, and in the editorial board for many international journals.



Dr. Carrie Hruska is an Assistant Professor of Radiologic Physics at Mayo Clinic, Rochester, MN, USA. Her major research interests include Nuclear medicine and breast imaging.



Dr. Seong-Ho Son received the B.S. degree in control and mechanical engineering (summa cum laude) from Pusan National University, Pusan, Korea, in 1997, and the M.S. and Ph.D. degrees in mechanical engineering from Pohang University of Science and Technology (POSTECH), Pohang, Korea, in 1999 and 2009, respectively. From 1999 to 2001, he was with the technical research center of the Daewoo Motors (now the GM Korea), Incheon, Korea, as a member of research staff. Since March 2001, he has been with the Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea, as a senior member of research staff. He was involved in the development of the mobile active phased-array antenna systems for satellite communications. He is now involved to develop the microwave tomography system for breast cancer detection. His research interests include active phased-array design, RF-related mechatronic system, and medical microwave imaging and computer-aided diagnosis for cancerous tumor. Dr. Son was the recipient of the 2008 Best Poster Paper Award of the International Conference on Mechanics of Composite Materials, and he also received the 2010 ETRI Best Paper Award for his paper published on the IEEE Transactions on Antennas and Propagation.