Two-Dimensional ARMA Modeling for Breast Cancer Detection and Classification

Jerzy Zielinski
Department of Applied Science
University of Arkansas at Little Rock
Email: jszielinski@ualr.edu

Nidhal Bouaynaya
Department of Systems Engineering
University of Arkansas at Little Rock
Email: nxbouaynaya@ualr.edu

Dan Schonfeld
Department of Electrical
and Computer Engineering
University of Illinois at Chicago
Email: dans@uic.edu

***Abstract*—Computer aided diagnosis (CAD) paradigms have gained currency for discriminating malignant from benign lesions in ultrasound breast images. But even the most sophisticated investigators often rely on one-dimensional representations of the image in terms of its scanlines. Such vector representations are convenient because of the mathematical tractability of one-dimensional time-series. However, they fail to take into account the spatial correlations between the pixels, which is crucial in tumor detection and classification in breast images. In this paper, we propose a CAD system for tumor detection and classification (cancerous v.s. benign) in ultrasound breast images based on a two-dimensional Auto-Regressive-Moving-Average (ARMA) model of the breast image. First, we show, using the Wold decomposition theorem, that ultrasound breast images can be accurately modeled by two-dimensional ARMA random fields. As in the 1D case, the 2D ARMA parameter estimation problem is much more difficult than its 2D AR counterpart, due to the non-linearity in estimating the 2D moving average (MA) parameters. We propose to estimate the 2D ARMA parameters using a two-stage Yule-Walker Least-Squares algorithm. The estimated parameters are then used as the basis for statistical inference and biophysical interpretation of the breast image. We evaluate the performance of the 2D ARMA vector features in real ultrasound images using a k-means classifier. Our results suggest that the proposed CAD system based on a two-dimensional ARMA model leads to parameters that can accurately segment the ultrasound breast image into three regions: healthy tissue, benign tumor, and cancerous tumor. Moreover, the specificity and sensitivity of the proposed two-dimensional CAD system is superior to its one-dimensional homologue.**

***Index Terms*—Breast cancer, two-dimensional ARMA models, k-means algorithm.**

I. Introduction

Breast cancer continues to be a significant public health problem in the United States: It is the second leading cause of female mortality, and, disturbingly, one out of eight women in the United States will be diagnosed with breast cancer in her life time. Until the cause of this disease is fully understood, early detection remains the only hope to improve breast cancer prognosis and treatment. Breast cancer screening modalities are mainly based on clinical examination, mammography, ultrasound imaging, magnetic resonance imaging (MRI), and core biopsy. Mammography (breast x-ray imaging) is by far the fastest and cheapest screening test for breast cancer. Unfortunately, it is also among the most difficult of radiological images to interpret: mammograms are of low contrast, and features indicative of breast disease are often very small. Many studies have shown that ultrasound and MRI imaging techniques can help supplement mammography by detecting small breast cancers that may not be visible with mammography. However, these techniques often fail to determine if a detected tumor is cancerous or benign, and a biopsy may be recommended. Consequently, many unnecessary biopsies are often undertaken due to the high false positive rate.

Computer aided diagnosis (CAD) paradigms have recently received great attention for lesion detection and discrimination in X-ray and ultrasound breast mammograms [1]–[4]. The large amount of negative biopsies encountered in clinical practice could be reduced if a computer system was available to help the radiologists screen breast images. Broadly, the CAD systems proposed in the literature can be grouped into four major categories: geometrical [1], artificial intelligence [2], pyramidal (or multiresolution) [3], and model-based techniques [4], [5]. Geometrical methods employ morphological and other segmentation techniques to extract small specks of calcium known as microcalcifications from breast images [1]. However, this procedure usually requires a priori knowledge of the tumor pattern characteristics. Moreover, these techniques also tend to rely on many stages of heuristics attempting to eliminate false positives. Artificial intelligence techniques include neural networks and fuzzy logic methods. The performance of these systems is tied to the architecture of the network and the number of training data. Breast cancer is a heterogeneous disease which includes several subtypes with distinct prognosis. In particular, the variability associated with the appearances of the breast cancer, ranging from relative uniformity to complex patterns of bright streaks and blobs [2], makes the ANN require a large training data set to ensure a certain level of reliability. Pyramidal or multiresolution techniques refer mainly to the wavelet transform [3], which can be seen from a signal decomposition view point. Specifically, a signal is decomposed onto a set of the basis wavelet functions. A very appealing feature of the wavelet analysis is that it provides a uniform resolution for all the scales. However limited by the size of the basic wavelet function, the downside of the uniform resolution is uniformly poor resolution. Model-based methods include linear, non-linear and finite-element methods to build an accurate model of the breast [4], [5]. The model is subsequently used for image matching, detection, and classification [5]. The accuracy of the results are tied to the

accuracy of the considered model.

In this work, we propose a new model-based CAD system for tumor detection and classification. We show that (x-ray, ultrasound, and MRI) breast images can be accurately modeled by two-dimensional autoregressive moving average (ARMA) random fields. The model parameters, being the fingerprints of the image, serve as the basis for statistical inference and biophysical interpretation of the breast image. ARMA models are parametric representations of wide-sense stationary (WSS) processes with rational spectra. The Wold decomposition theorem states that any WSS process can be decomposed as the sum of a regular process, which spectrum is continuous, and a predictable process, which spectrum consists of impulses. Since rational spectra form a dense set in the class of continuous spectra, the ARMA model renders accurately the regular part of the WSS process. It is, therefore, surprising that very few researchers have attempted to derive a general ARMA representation of the breast image, and use it for tumor detection and classification. In [5], the authors use a one-dimensional fractional differencing autoregressive moving average (FARMA) process to model the ultrasound RF echo reflected from the breast tissue. However, by considering separate scan lines, they do not take into account the two-dimensional spatial correlation between the pixels in the image. In [6], an autoregressive (AR) model is considered for improving the contrast of breast cancer lesions in ultrasound images. ARMA models, however, provide a more accurate model of a homogeneous random field than an AR model. As in the 1D case, the 2D ARMA parameter estimation problem is much more difficult than its 2D AR counterpart, due to the non-linearity in estimating the 2D moving average (MA) parameters.

This paper is organized as follows: In Section II, we define the 2D ARMA model of the breast image, and derive a Yule-Walker Least squares estimates of its parameters [7]. In Section III, we use the estimated ARMA coefficients as vector features for the k-means classifier. The simulation results, for ultrasound breast images showing cancerous and benign tumors, are shown in Section IV. Finally, in Section V, we summarize our contribution and provide concluding remarks.

II. 2D ARMA Modeling

We represent the breast image as a 2D random field $\{x[n,m], (n,m) \in \mathbb{Z}^2\}$. We define a total order on the discrete lattice as follows

$$(i,j) \leq (s,t) \Longleftrightarrow i \leq s \quad \text{and} \quad j \leq t. \tag{1}$$

The 2D ARMA(p_1, p_2, q_1, q_2) model is defined for the $N_1 \times N_2$ image $I = \{x[n,m] : 0 \leq n \leq N_1 - 1, 0 \leq m \leq N_2 - 1\}$ by the following difference equation

$$x[n,m] + \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{p_1} \sum_{j=0}^{p_2} a_{ij} x[n-i, m-j] = \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} b_{ij} w[n-i, m-j], \tag{2}$$

where $\{w[n,m]\}$ is a stationary white noise field with variance σ^2, and the coefficients $\{a_{ij}\}, \{b_{ij}\}$ are the parameters of the model. From Eq. (2), the image $\{x[n,m]\}$ can be viewed as the output of the linear time-invariant causal system $H(z_1, z_2)$ excited by a white noise input, where

$$H(z_1, z_2) = \frac{B(z_1, z_2)}{A(z_1, z_2)} = \frac{\sum_{i=0}^{q_1} \sum_{j=0}^{q_2} b_{ij}\ z_1^{-i} z_2^{-j}}{\sum_{i=0}^{p_1} \sum_{j=0}^{p_2} a_{ij}\ z_1^{-i} z_2^{-j}}, \tag{3}$$

with $a_{00} = 1$.

A. Yule-Walker Least-Squares Parameter Estimation

Assume first that the noise sequence $\{w[n,m]\}$ were known. Then the problem of estimating the parameters in the ARMA model in Eq. (2) would be a simple input-output system parameter estimation problem, which could be solved by several methods, the simplest of which is the least-squares (LS) method. In the LS method, we express Eq. (2) as

$$x[n,m] + \phi^t[n,m]\theta = w[n,m], \tag{4}$$

where

$$\begin{aligned} \phi^t[n,m] \quad = \quad & [x[n,m-1], \cdots, x[n-p_1, m-p_2], -w[n,m-1], \\ & \cdots, -w[n-q_1, m-q_2]], \end{aligned}$$

and

$$\theta = [a_{01}, \cdots, a_{p_1 p_2}, b_{01}, \cdots, b_{q_1 q_2}]^t.$$

Writing Eq. (4) in matrix form for $n = L+1, \cdots, N_1 - 1$, and $m = M+1, \cdots, N_2 - 1$, for some $L > \max(p_1, q_1)$, and $M > \max(p_2, q_2)$, gives

$$\mathbf{x} + \Phi\theta = \mathbf{w}, \tag{5}$$

where

$$\begin{aligned} \mathbf{x} &= [x[L+1, M+1], \cdots, x[N_1 - 1, N_2 - 1]]^t, \\ \mathbf{w} &= [w[L+1, M+1], \cdots, x[N_1 - 1, N_2 - 1]]^t, \end{aligned}$$

and Φ is displayed below. Assume we know Φ, then we can obtain a least-squares estimate of the parameter vector θ in Eq. (5) as

$$\hat{\theta} = -(\Phi^t \Phi)^{-1} \Phi^t \mathbf{x}. \tag{6}$$

Observe that the input model noise $\{w[n,m]\}$ in Φ is unknown. Nevertheless, it can be estimated by considering the noise process $w[n,m]$ as the output of the linear filter $\frac{1}{H(z_1,z_2)} = \frac{A(z_1,z_2)}{B(z_1,z_2)}$ with input $x[n,m]$. From Nirenberg's proof of the division theorem in multi-dimensional spaces [8], we can write the inverse ARMA filter $\frac{A(z_1,z_2)}{B(z_1,z_2)}$ as the infinite order AR filter $\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \alpha_{ij} z_1^{-i} z_2^{-j}$. In the time domain, we obtain

$$x[n,m] + \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{\infty} \sum_{j=0}^{\infty} \alpha_{ij}\ x[n-i, m-j] = w[n,m]. \tag{7}$$

Therefore, we can estimate $\{w[n,m]\}$ by first estimating the AR parameters $\{\alpha_{ij}\}$ and next obtaining $\{w[n,m]\}$ by filtering $\{x[n,m]\}$ as in Eq. (7). Since we cannot estimate

$$\Phi = \begin{pmatrix} x[L+1,M] & \cdots & x[L+1-p_1,M+1-p_2] & -w[L+1,M] & \cdots & -w[L+1-q_1,M+1-q_2] \\ x[L+2,M] & \cdots & x[L+2-p_1,M+1-p_2] & -w[L+2,M] & \cdots & -w[L+2-q_1,M+1-q_2] \\ \vdots & & \vdots & \vdots & & \vdots \\ x[N_1-1,N_2-2] & \cdots & x[N_1-1-p_1,N_2-1-p_2] & -w[N_1-1,N_2-2] & \cdots & -w[N_1-1-q_1,N_2-1-p_2] \end{pmatrix}.$$

an infinite number of (independent) parameters from a finite number of samples, we approximate the finite AR model by one of finite order, say (K_1, K_2). The parameters in the truncated AR model can be estimated by using a 2D extension of the Yule-Walker equations as follows

$$r[k,l] + \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{K_1} \sum_{j=0}^{K_2} \alpha_{ij} r[k-i,l-j] = \sigma^2 \delta[k,l], \qquad (8)$$

where $\{r[k,l]\}$ are the autocorrelation values of the field $\{x[n,m]\}$, computed as follows

$$\begin{aligned} r[k,l] &= \frac{1}{(N-k)(M-l)} \sum_{i=1}^{N-k} \sum_{j=1}^{M-l} x[i,j]x[i+k,j+l], \\ r[-k,-l] &= r[k,l], \text{ for } (k,l) \geq (0,0) \\ r[k,-l] &= r[-k,l], \text{ for } (k,l) \geq (1,1), \end{aligned} \qquad (9)$$

and $\delta[k,l]$ is the 2D Kronecker delta function. Equation (8) is a system of linear equations that can be written in matrix form and solved for the coefficients α_{ij}. Finally, the Yule-Walker Least-Squares algorithm is summarized below

1) Estimate the parameters $\{\alpha_{ij}\}$ in an AR(K_1, K_2) model of $x[n,m]$ by the Yule-Walker method in (8). Obtain an estimate of the noise field $\{w[n,m]\}$ as

$$\hat{w}[n,m] = x[n,m] + \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{K_1} \sum_{j=0}^{K_2} \hat{\alpha}_{ij} x[n-i,m-j],$$

for $n = K_1+1, \cdots, N_1$, and $m = K_2+1, \cdots, N_2$.

2) Replace the $w[n,m]$ by $\hat{w}[n,m]$ computed in Step 1. Obtain $\hat{\theta}$ in (6) with $L = K_1 + q_1$, and $M = K_2 + q_2$.

III. Tumor Detection and Classification

The estimated ARMA parameters, $\{a_{ij}\}, \{b_{ij}\}$, are used as a basis for inference about the presence of a tumor and its nature: benign or cancerous. We use the k-means algorithm to segment the breast image into 3 classes: healthy tissue, benign tumor and cancerous tumor. Our method consists of representing each pixel in the image by an ARMA model whose parameters are estimated by using an appropriate neighborhood for the pixel. We make the assumption that all pixels in the considered neighborhood belong to the same class, and hence, for computational efficiency, we replace the entire neighborhood by the vector value of the estimated ARMA parameters. This procedure is repeated for the entire image, creating a new block by block vector-valued image, which will be the input to the k-means classifier.

IV. Simulations

Although the proposed algorithm is independent of the imaging modality of the breast, we perform our simulations on ultrasound images, collected from the Radiology department, College of Medicine at the University of Illinois at Chicago. Our database of cancerous images show intraductal carcinoma, which is the most common type of breast cancer in women. Intraductal carcinoma is usually discovered through a mammogram or an ultrasound as microcalcifications. Our benign tumor images show the Fibroadenoma of the breast, which is a benign fibroepithelial tumor characterized by proliferation of both glandular and stromal elements.

Our extensive simulations indicate that ARMA$[2,2,2,2]$ is a sufficient model order, in terms of mean square error, to accurately represent ultrasound breast images. Figure 1 shows two ultrasound images, one with a cancerous tumor and one with a benign tumor, and their respective 2-D ARMA$[2,2,2,2]$ and 1-D ARMA$[2,2]$ models. It is visually clear that the 2D-ARMA model accurately represents both ultrasound images, whereas the 1-D model fails to capture any image feature. We estimate the 2D-ARMA parameters using a window of size 16×16. The choice of the window size presents an inherent trade-off between the accuracy of the representation and the accuracy of the classification. A large window size would lead a better representation of the 2D-ARMA model, but might include pixels from different classes. We found that for 256×256 images, a 16×16 window size leads to a good segmentation performance. Each image is therefore represented by a number of 1×8 2D-ARMA feature vectors, which contain the 8 parameters $a_{00}, a_{01}, a_{10}, a_{11}, b_{00}, b_{01}, b_{10}, b_{11}$ for each 16×16 sub-block image. Without loss of generality, we chose $a_{00} = b_{00} = 1$. Therefore, the size of the feature vectors reduces to 6 instead of 8. We decide that an image has a cancerous (resp., benign) tumor if at least one of the sub-block images is classified as a cancerous (resp., benign) tumor. Otherwise, we conclude that the image is healthy and contains no tumors.

We conducted our simulations using 573 ARMA feature vectors of healthy, benign and cancerous ultrasound breast images. The ARMA feature vectors were used as the input to a k-means classifier. Figures 1(c) and 1(f) show the segmentation outputs of the cancerous and benign tumor images, respectively. We can observe clear delineations of the tumors from the healthy tissues in both cases. The accuracy, sensitivity and specificity of the 2D-ARMA and 1D-ARMA k-means classifiers are shown in Table I. It is clear that the 2D-ARMA feature vectors are more selective than their one-dimensional homologue.

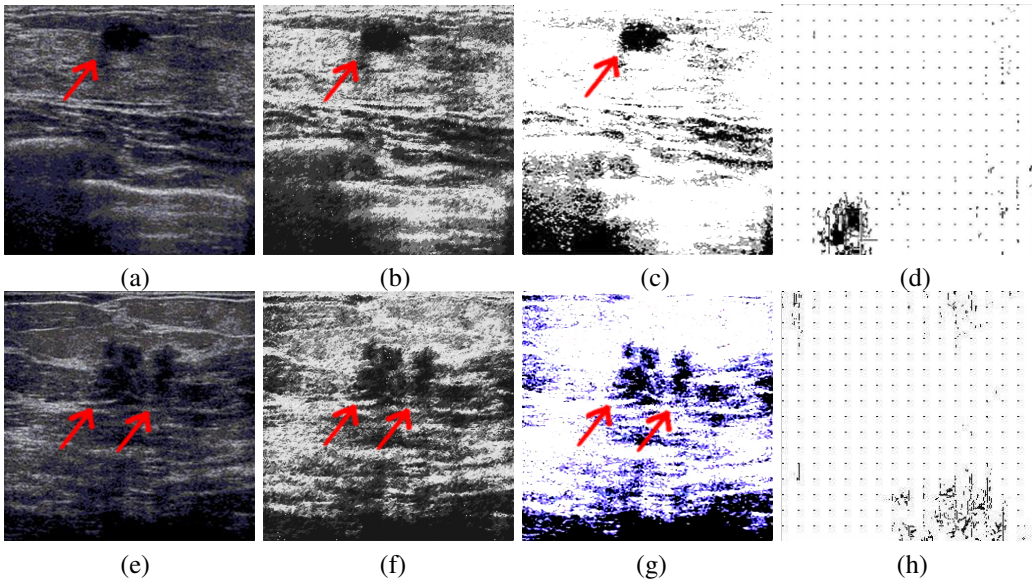

Fig. 1. ARMA modeling and segmentation of ultrasound breast images: (a) cancerous ultrasound image; (b) 2D-ARMA[2,2,2,2] representation of (a); (c) Segmentation of (b) using an appropriate threshold; (d) 1D-ARMA[2,2] representation of (a); (e) benign tumor ultrasound image; (f) 2D-ARMA[2,2,2,2] representation of (e); (g) Segmentation of (f) using an appropriate threshold; (h) 1D-ARMA[2,2] representation of (e).

TABLE I
CLASSIFICATION ACCURACY OF CANCEROUS AND BENIGN TUMORS

	Accuracy	Sensitivity	Specificity
2D-ARMA	93.87%	92.03%	94.14%
1D-ARMA	78.51%	59.54%	79.76%

V. CONCLUSIONS

We propose to exploit the high spatial correlation inherent in neighboring pixels to improve tumor detection and classification in ultrasound breast images. We achieve this goal by using a two-dimensional autoregressive moving average (ARMA) field model of the image. Current techniques often rely on one-dimensional representations of the image in terms of its scan lines in order to process it as a one-dimensional time-series [5], [6]. Such one-dimensional projections are advocated solely on the basis of the simplicity of their mathematical formulations. The analysis of two-dimensional fields is more involved mathematically and computationally than the study of one-dimensional time-series. In this work, we derive an efficient two-stage algorithm to estimate the parameters of the two-dimensional ARMA field model of the breast image. The estimated ARMA parameters are excellent discriminative features, which are used as the basis for statistical detection and classification of tumors in the breast image. Our simulation results on a library of benign and cancerous ultrasound breast images show the superiority, in terms of accuracy, sensitivity and specificity, of the two-dimensional ARMA model to its one-dimensional homologue. The proposed algorithm can be efficiently incorporated within a computer-aided diagnosis (CAD) system that clinically portends an accurate prognosis of breast cancer.

VI. ACKNOWLEDGEMENTS

The authors want to thank Dr. Karen Lin Xie from the Radiology department at the University of Illinois at Chicago for providing the ultrasound breast images.

REFERENCES

[1] N.H. Eltonsy, G.D. Tourassi, and A.S. Elmaghraby, "A concentric morphology model for the detection of masses in mammography," *IEEE Transactions on Medical Imaging*, vol. 26, no. 6, pp. 880–889, 2007.

[2] L. Wei, Y. Yang, R.M. Nishikawa, and Y. Jiang;, "A study on several machine-learning methods for classification of malignant and benign clustered microcalcifications," *IEEE Transactions on Medical Imaging*, vol. 24, no. 3, pp. 371–380, March 2005.

[3] A. Mencattini, M. Salmeri, R. Lojacono, M. Frigerio, and F. Caselli, "Mammographic images enhancement and denoising for breast cancer detection using dyadic wavelet processing," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 7, pp. 1422–1430, July 2008.

[4] P. Pathmanathan, D.J. Gavaghan, J.P. Whiteley, S.J. Chapman, and J.M. Brady, "Predicting tumor location by modeling the deformation of the breast," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 10, pp. 2471–2480, October 2008.

[5] B. Alacam, B. Yazici, N. Bilgutay, F. Forsberg, and C. Piccoli, "Breast tissue characterization using FARMA modeling of ultrasonic RF echo," *Ultrasound in Medicine and Biology*, vol. 30, no. 10, pp. 1397–1407, October 2004.

[6] E. von Lavante and J. A. Noble, *Improving the Contrast of Breast Cancer Masses in Ultrasound Using an Autoregressive Model*, vol. 4791 of *Lecture Notes in Computer Science*, Springer, Berlin/Heidelberg, 2007.

[7] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*, Prentice Hall, 2005.

[8] L. Nirenberg, "A proof of the malgrange preparation theorem," *Springer Lecture Notes in Math*, vol. 192, pp. 97–105, 1971.