Breast Tumor Localization Using Simultaneous Perturbation Stochastic-Neural Algorithm

Ahmad Abdulsadda

Communication Department, Al Najaf Technical College, Foundation of Technical Education
*Corresponding Author: abdulsad@msu.edu

Copyright ©2013 Horizon Research Publishing All rights reserved.

Abstract Many impalpable or occult breast cancers cannot be localized using imaging techniques like mammography and ultrasound. An accurate localization of the tumor is, however, essential to guide the surgeon to the lesion, and ensure its correct and adequate removal with satisfactory excision margins. Current breast tumor localization techniques are invasive and often result in a cosmetic disfigurement. In this paper, we use the ultrawide band radar-based microwave breast imaging technique to non-invasively localize (impalpable) tumors in the breast. We consider four clinically important lesion features: location, size, depth and spatial orientation within the breast. A comparison of the energy of the received signal from healthy and cancerous breasts exhibits some remarkable differences in some frequency bands. We, therefore, use the energy spectrum of the receiving antenna signal decomposed by wavelet transform as the input to a Simultaneous Perturbation Neural Network (SPNN) classifier. Furthermore, we determine the optimum structure and gains of the SPNN, in terms of optimum initial weights and optimum number of nodes in the hidden layer. We use CST Microwave Studio to simulate a data set of 1024 cancer cases with various tumor location, size, depth and direction inside the breast. Our results show that the proposed algorithm gives accurate localization of the breast lesion, and possesses a high sensitivity to small tumor sizes. Additionally, it can accurately detect and classify multiple tumors with short learning and testing time.

Keywords Breast Cancer Localization, Radar-based Microwave Breast Imaging, Wavelet Decomposition, Simultaneous Perturbation Stochastic Analysis (SPSA), Neural Networks

1 Introduction

A complete understanding of the distribution of tumor before therapy is a key factor in managing any case of advanced breast cancer. This is particularly true if pre-operative chemotherapy is used in an attempt to allow breast preservation [1]. The key to surgical planning for breast conservation is tumor localization. Impalpable or occult breast cancers are being found with increasing frequency as microcalcifications, opacities or parenchymal distortions. If the tumor shrinks to 50% or more after chemotherapy or if the tumor shrinks to less than 2 cm, then the surgeon would need an accurate guide to the lesion in order to remove it from the breast with satisfactory excision margins, and be well tolerated by the patient. The goal in definite local therapy is to obtain tumor free margins of 1 cm in all dimensions around the residual tumor or calcifications [1].

Although a number of techniques have been proposed to localize impalpable lesions prior to surgery, needle or hook-wire localization is still the most common used technique to help guide definitive resection of a cancer [2]. For needle localization procedures, the breast is compressed and a needle is inserted through the skin and into the lesion with mammographic or ultrasonic guidance. After appropriate needle positioning is verified, the breast is taken out of compression. This procedure presents a number of disadvantages to the patient including discomfort and vasovagal syncope [3]. There are also clinical complications, which include pneumothoraces and wire migration, displacement or transection [3]. Also, the skin entry site of the needle is determined by the radiologist, and might not be the most optimal site for the surgeon. Radioguided occult lesion localization (ROLL) was proposed to eliminate some disadvantages of the needle biopsy [4]. In ROLL, tumor localization is carried out searching for the maximal radioactivity by gamma-probe designed for sentinel lymph node localization. It has been proven that ROLL reduces the excision volume, and enables better lesion centering within the specimen; thus, leading to better cosmetic results compared to wire localization. However, the procedure is still invasive, and there is the potential for radiation exposure to the patient and surgeon. Cash et al. [3] proposed a noninvasive method of breast tumor localization based on the coregistration of three-dimensional (3D) ultrasonographic (US) data with surface contour data obtained by using a 3D laser camera. However, as the authors point out, this technique fails in general with large breast patients because the image definition diminishes with depth, thus resulting in the need for compression. Compression will distort the breast con-
tour and result in a poor coregistration of US data with laser data. Moreover, breast compression is likely to cause a radial movement of the tumor, thus distorting the results of the coregistration.

In this paper, we propose a new noninvasive tumor localization method based on the ultrawide band microwave backscatter technique. We use the wavelet decomposition to compress while enhancing the backscattered signal’s signature in the frequency domain, and an optimized simultaneous perturbation neural network for accurate tumor localization. Microwave imaging is an alternative imaging method for detecting breast cancer [5]. It is based on the electrical properties of the breast tissues. Cancerous tissues exhibit electrical characteristics that show significant and consistent contrast with healthy breast tissues. Cancer detection with microwave imaging is based on this contrast in electrical properties. Ultra wideband radar-based microwave imaging uses an antenna to highlight the breast with low-power ultra wideband pulses of microwave energy. The scattered energy is received by a receiving antenna. The early part of the received signal contains the skin-breast interface response, whereas the late-time response contains the backscattered signal from inside the breast, and might contain possible lesions [5]. The received signal, however, has a low resolution due to clutter or various healthy tissue backscattering. It is therefore difficult to reconstruct the fine characteristics, e.g., location, depth, size, and direction of a detected lesion in the breast.

So far, previous work in ultra wideband microwave imaging has focused on tumor detection and breast segmentation into healthy tissue and abnormal lesions. In particular, breast cancer detection using neural networks has been described by many researchers [5–8]. Al-Shenawee’s group [6–8] showed the applicability of artificial neural networks to breast cancer detection using radar-based microwave imaging. Specifically, they were able to detect tumors by using an estimate of the dielectric permittivity of the breast. However, they neither provided details about the neural network structure nor tried to optimize its design.

In this paper, we propose to localize breast lesions in terms of their (i) location, (ii) size, (iii) depth, and (iv) spatial orientation within the breast. Spatial orientation of the tumor with respect to the nipple may illustrate ductal extension of tumor, and this factor may have an influence on the shape and volume of the excised specimen. We found that there are remarkable differences in specific frequency bands between healthy and tumor tissues. We foster this difference in the frequency domain for tumor localization by using the wavelet transform of the received signal’s energy to obtain a discriminative feature vector for the neural network classifier. An important byproduct of the wavelet decomposition is the compression of the total number of input data to few wavelet coefficients. In order to avoid the common neural network problems of local minima, convergence, and overfitting of the data, we use the simultaneous perturbation stochastic approximation (SPSA) algorithm to select the optimum number of hidden nodes and the optimum initial weights for the network.

The rest of the paper is organized as follows: In Section 2, we present the breast model that we will use in our simulations to assess the performance of the proposed algorithm. The extraction of the wavelet feature vector index is presented in Section 3. Section 4 introduces the simultaneous perturbation stochastic-neural network and discusses the tumor classification and characterization algorithm and it addresses the extensive testing procedure and computes the performance of the proposed technique. Concluding remarks and future work are discussed in Section 5.

2 The Breast Model: Problem Formulation

Following the work by El-Shenawee et al [6], we model the breast as a homogenous region, which can possibly contain one or multiple lesions. In particular, our model ignores the skin-breast interface since it has been shown that its effect reduces to a dc level difference in the power of the received signal [8]. The transmitting antenna operates at a frequency of 6 GHz. This frequency has been determined by trial and error in order to determine the best radiation pattern [9]. To receive adequate signal at the pickup location without exceeding the specified radiation safety limits, an array of three patch antennas was designed to focus the main beam in the direction of the tumor. The distance between the breast and the antenna array is set to 2 cm with total feeding power of 300mW. The breast cancer model with the antenna position is described in Fig. 1. The received signal carries information about the tumor status inside the breast, and is related, through the Maxwell equations, to the dielectric permittivity of the breast [7].

In our breast model, we take into account the dependence of the relative dielectric permittivity of the breast, \( \epsilon_r \), on the frequency by using the following first-order Debye dispersion formula:

\[
\epsilon_r = \epsilon_\infty + \frac{\epsilon_s - \epsilon_\infty}{1 + \omega^2 \tau^2},
\]

Where \( \epsilon_s \) and \( \epsilon_\infty \) are the space permittivity, and the high frequency permittivity, respectively. \( \omega \) denotes the an-
gular frequency, and $\tau$ is the time constant. The values of the Debye parameters for normal and cancerous tissues have been determined experimentally in [6]. Figure 2 shows an example of the time domain measurement of the received signal for a healthy and cancerous breasts.

3 Wavelet-based Feature Vector

Even though the response signal of the breast to an ultra wideband microwave frequency can be easily measured using patch antennas [9], the raw received signal cannot be used directly to fully identify the abnormal lesions within the breast due to the high amount of clutter from healthy tissues. An appropriate representative index vector has to be constructed from the received signal, and used for tumor characterization and classification. A comparison of the energy of the received signal from healthy and cancerous breasts exhibits some remarkable differences in specific frequency bands. This is because tumors will suppress or enhance certain frequency components of the received signal, and consequently cause energy increase or decrease of the received signal. Therefore, the energy of the received signal contains information on abnormal lesions. Additionally, the energy variation of one or several frequency components of the signal can indicate a special characteristic of the lesion. In order to extract the tumor information from the breast response signal, we first decompose the signal into multiple sub-signals in various frequency bands using the wavelet decomposition as follows

$$S_{0,0}[n] = \sum_{j=1}^{2^{k-1}} S_{k,j}[n], \quad n = 0, 1, \cdots, N - 1, \quad (2)$$

Where $S_{0,0}[n]$ denotes the sampled received signal of length $N$, and $S_{k,j}[n]$ is the sub-signal with orthogonal frequency band and $k$ indicates the layer number of the tree structure of the wavelet decomposition. The energy of the $j^\text{th}$ order sub-signals can be expressed as

$$U_{k,j} = \sum_{n=1}^{N} |S_{k,j}[n]|^2. \quad (3)$$

Assuming that the energy of the $j^\text{th}$-order sub-signals of the healthy and cancerous tissues are $U_{k,j}^h$ and $U_{k,j}^c$ respectively, the wavelet index vector can be constructed as follows:

$$V_d = [v_1, v_2, \ldots, v_{2^k-1}]^T = [1 - \frac{U_{k,1}^h}{U_{k,1}^c}, 1 - \frac{U_{k,2}^h}{U_{k,2}^c}, \ldots, 1 - \frac{U_{k,2^k-1}^h}{U_{k,2^k-1}^c}]^T \quad (4)$$

The elements of different index vectors not only indicate the differences between healthy and cancerous tissues, but also suggest the differences between various lesion types. We will use the energy vector as the input feature vector to the genetic neural network classifier [10].

4 Optimum design of the neural network using the Simultaneous perturbation algorithm

4.1 Simultaneous perturbation algorithm

Stochastic approximation (SA) represents an important class of search algorithms used in error function minimization. Many well-known techniques are applied to the problems of minimizing loss functions or root finding with noisy input information, including general network back propagation, perturbation analysis for discrete-event systems, recursive least squares and least mean squares, and some forms of simulated annealing cases of SA [11]. Therefore, any improvement in general SA methodology can have a potential bearing on a wide range of practical implementations [12]. As with all search algorithms, there are adjustable coefficients in SA that must be specified and can have a profound effect on the algorithm performance. It is known that choosing these coefficients according to an SA analogue of the deterministic Newton-Raphson algorithm provides an optimal or near optimal form of search. However, directly determining the required Hessian matrix (or the Jacobian matrix for root finding) to achieve its optimal form has often been difficult or impossible in practice. In particular, the problem of minimizing a (scalar) differentiable loss function $L(\theta)$ (in our case it is represented by the error to weight gradient) where $\theta \in \mathbb{R}^p$, $p = 1$ is considered [11]. A typical example would be some measure of mean-square error for the output of a process as a function of design parameters. For many cases of practical interest, this is equivalent to finding the unique minimum such that [12]:

$$g(\theta) \equiv \frac{\partial L}{\partial \theta} \bigg|_{\theta} = 0 \quad (5)$$

where $\theta^*$ in our case is represented by the optimal neural network weight. For the gradient-free setting, it is assumed that measurements $L(\theta)$ denoted as $y(\theta)$ are available at the various values of $\theta$. These measurements may or may not include random noise. No direct measurements (either with or without noise) of $g(\theta)$ are assumed available in this setting. A recursive procedure has been used to calculate $\theta$ in general SA form, given as [8]:

$$\hat{\theta}_{k+1} = \hat{\theta}_k - a_k \hat{g}_k(\hat{\theta}) \quad (6)$$

where $\hat{\theta}_k$ represents the estimate of $\theta$ at the $k^{\text{th}}$ iteration, $a_k > 0$ represents a scalar gain coefficient, $\hat{g}_k(\hat{\theta})$
represents an approximation of $g(\hat{\theta})$; typically, the $i^{th}$ component of $\hat{g}_k(\hat{\theta})(i = 1, 2, \ldots, p)$ is given by [7]:

$$\hat{g}_k(\hat{\theta}) = \frac{y(\hat{\theta} + c_k\Delta_k) - y(\hat{\theta} - c_k\Delta_k)}{2c_k\Delta_k}$$

(7)

where the elements of the user specified random perturbation vector $\Delta_k = (\Delta_{k1}, \ldots, \Delta_{kp})$ are assumed to be independent and symmetrically distributed around zero, and $c_k$ is a positive scalar. SPSA makes use of decaying gain sequences $a_k$, and $c_k$ in the form [8]:

$$a_k = \frac{a}{(A + k + 1)^\gamma},$$

(8)

$$c_k = \frac{c}{(k + 1)^\gamma},$$

(9)

where $A$, $\gamma$, $a$, and $c$ are non-negative coefficients. We use the simultaneous perturbation algorithm to determine the optimal initial weights and the optimum number of nodes in the hidden layer of the neural network. The optimal initial weights will guarantee a faster convergence to the global minimum of the objective function, and the optimum number of nodes in the hidden layer will avoid the bias-variance dilemma of neural networks. Given a data set, it is known that a small number of nodes in the hidden layer captures the data trend only in a small region of the pattern, whereas a large number of nodes in the hidden layer results in overfitting the data [13]. In the tumor localization problem, it is important to obtain a high accuracy in order to design a computer-aided system, which will ultimately replace preoperative invasive techniques for tumor localization.

### 4.2 Simultaneous perturbation neural network

We use a three-layer artificial neural network (ANN) with a hyperbolic activation function, which has been shown to approximate any smooth mapping.: (1) An input layer that receives the index vector of the received signal, which has 32 nodes; (2) A hidden layer, which processes the data; and (3) An output layer that indicates the tumor status we are interested in classifying.

In this paper, we consider the following breast cancer statuses: In this paper, we consider the following breast cancer statuses:

1. Tumor spatial orientation of $0^\circ$ and $90^\circ$ corresponding to a longitudinal or a transverse lesion, and expressed using one bipolar nodal output of -1 or 1.

2. Three possible tumor depth: $5\%H$, $10\%H$, and $15\%H$ denoted using the nodal outputs with two digits $(-1,1), (-1,1)$, and $(11)$, respectively.

3. 15 possible tumor cell lengths: $1\%L$, $2\%L$, $\ldots$, $15\%L$ represented using the nodal outputs with four digits $(1, -1, -1, -1), (-1, 1, -1, -1), \ldots, (1, 1, 1, 1)$, respectively.

4. 15 possible ranges of x-coordinates, with respect to the breast length L, and 15 possible ranges of y-coordinates, with respect to the breast width B, of the tumor location at $5\%, 15\%, 25\%, 30\%, 35\%, 40\%$, $45\%, 50\%, 55\%, 60\%, 65\%, 70\%, 75\%, 85\%$, and $95\%$. The $x, y$ coordinates are described using two sets of nodal outputs with four digits $(1, -1, -1, -1), (-1, 1, -1, -1), \ldots, (1, 1, 1, 1)$, respectively.

Figure 3 shows 60 examples of cancer statuses among the 32,767 possible cancer statuses that can be described by the NN. Observe that we need 15 nodes in the output layer of this NN. The healthy case (no tumor) will be represented by the value 0 at all 15 nodes in the output layer.

Our ultimate goal would be to assess the performance of the proposed tumor localization system in breast cancer patients. However, due to the legal and technical difficulties involved in obtaining ultrawideband microwave backscatter signals from patients, we opted to use the CST Microwave Studio to simulate the backscattered microwave signal from cancerous breasts. We model the breast as a homogeneous cylindrical shape attached to a half-sphere, and containing a tumor, modeled as a small sphere inside the breast. In this reported set of simulations, we use a tumor of radius 8 mm. Observe that it is very difficult to observe such a small size tumor using the standard imaging methods like mammography and ultrasound imaging. The Debye parameters for healthy breast tissues are given by $\epsilon_s = 10, \epsilon_\infty = 7, \tau = 6.4$ ps, and for cancerous tissues are $\epsilon_s = 40, \epsilon_\infty = 5.573, \tau = 9.149$ ps at 6 GHz [8]. Therefore, using Eq. (1), the relative dielectric permittivity of a healthy (resp., cancerous) breast is $\epsilon_r = 9.89$ (resp., $\epsilon_r = 38.20$).

Applying the simultaneous perturbation to find the optimal neural network parameters, we found that the optimal number of nodes in the hidden layer is 13. We test the NN on 1024 cancer statuses and one healthy (no tumor) case. Table1 shows the desired (or target) outputs and the obtained outputs of the Simultaneous perturbation neural network for five selected samples. The proposed algorithm was able to discriminate between healthy and cancerous breasts with an accuracy of 100%. Therefore, the false positive and false negative rates for the tumor detection problem are at a minimum of 0%. However, the classification error of the cancer cases is evaluated at about 4%, thus leading to an accuracy of 96%.
5 Conclusion

Breast tumor localization is a very serious issue cancer type for women throughout the world. The commonly used diagnostic techniques, like mammography, are reported to lack high diagnostic capability. Therefore, there is an absolute necessity to develop better diagnostic techniques. Breast cancer signal can be conveniently acquired using the microwave imaging modality with transmitting and receiving patch antennas. In this paper, we showed that we can detect and classify breast tumor cases using the antenna spectrum of the breast responses decomposed by wavelet transform and fed to an optimized genetic neural network. Our simulations show that the required number of inputs to the neural network is greatly reduced, thus reducing the time and effort needed for the genetic neural network training. The simultaneous perturbation algorithm gave us the optimal number of nodes in the hidden layer with the best initial weights as well. The main contribution of this paper goes beyond tumor detection to tumor localization by determining important clinical properties of any detected tumor, such as location, size and depth. If known at the diagnosis time, these properties can provide the clinician with valuable head-start clues; for instance the type of the abnormality in the breast, namely benign or malignant, and if an immediate surgery is needed or if an appropriate treatment should be taken first in order to shrink the tumor before surgery. Furthermore, a computer-aided diagnosis system based on the proposed algorithm for tumor localization will save money and pain to the patients by avoiding expensive (invasive or non-invasive) techniques to localize the exact position of the tumor before surgery. Finally, the proposed algorithm can be applied to any tumor localization within the body as long as the tumor and healthy tissue properties exhibit significant differences in the wavelet or frequency domain.

REFERENCES

[1] S. E. Singletary, G. L. Robb, and G. N. Hortobagyi, Advanced therapy of breast disease, PMPH-USA, 2004.

[2] A. Haid, M. Knauer, S. Dunzinger, and et al. Z. Jasarevic, “Intra-operative sonography: a valuable aid during breast-conserving surgery for occult breast cancer,” Annals of surgical oncology, vol. 14, no. 11, pp. 3090–30101, November 2007.

[3] C. J. C. Cash, G. M. Treece, A. D. Purushotham, and et al. P. Britton, “Noninvasive method of preoperative localization with three-dimensional us and surface contour mapping,” Radiology: Volume 245: Number 2 2007, vol. 245, no. 2, pp. 556–566, November 2007.

[4] C. De Cicco, M. Pizzamiglio, G. Trifiro, and et al. A. Luini, “Radioguided occult lesion localisation (roll) and surgical biopsy in breast cancer. technical aspects.,” The quarterly journal of nuclear medicine, vol. 46, no. 2, pp. 145151, June 2002.

[5] E. Fear and M. Stuchly, “Microwave detection of breast cancer,” IEEE Transaction on Microwave theory and techniques, vol. 48, no. 11, pp. 1854–1863, 2000.

[6] D. Woten, J. Lusth, and M. Al-Shenawee, “Interpreting artificial network output for the microwave detection of breast cancer,” IEEE Microwave and Wireless Components Letter, vol. 17, no. 12, pp. 825 – 827, December 2007.

[7] D. Woten and M. Al-Shenawee, “Improvement of artificial neural network detection of breast cancer using broadband dual polarized antenna,” in IEEE Antennas and Propagation Internation Symposium, June 2007, pp. 261–264.

[8] D. Woten and M. Al-Shenawee, “Error analysis of breast tumor signature versus skin thickness at microwave frequencies,” in IEEE Antennas and Propagation Internation Symposium, July 2008, pp. 1–4.

[9] S. Suh, A comprehensive investigation of new planar wideband antennas, Ph.D. thesis, Virginia Polytechnic Institute, Blacksburg, VA, USA, 2002.

[10] L. Yam, Y. Yan, L. Cheng, and J. Jiang, “Identification of complex crack damage for honeycomb sandwich plate using wavelet analysis and normal networks,” Smart Materials and Structures, vol. 12, pp. 661671, 2003.

[11] J. C. Spall, “Multivariate stochastic approximation using a simultaneous perturbation gradient approximation,” IEEE Transaction Automation Control, vol. 37, pp. 332–331, 1992.

[12] J. C. Spall, “Developments in stochastic optimization algorithms with gradient approximation based on function measurement,” Winter Simulation Conference, pp. 207–214, 1994.

[13] S. Samarasinghe, Neural Networks for Applied Science and Engineering, Taylor & Francis Group, LLC, 2007.