Biomedical Signal Processing and Control 31 (2017) 366–376

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Microwave breast cancer detection via cost-sensitive ensemble classifiers: Phantom and patient investigation

Yunpeng Li*, Emily Porter, Adam Santorelli, Milica Popović, Mark Coates

Dept. of Electrical and Computer Engineering, McGill University, Montréal, Québec, Canada

ARTICLE INFO

Article history:
Received 18 February 2016
Received in revised form 23 July 2016
Accepted 5 September 2016

Keywords:
Microwave breast cancer detection
Sensor fusion
Biomedical monitoring
Clinical trial
Ensemble selection

ABSTRACT

Microwave breast screening has been proposed as a complementary modality to the current standard of X-ray mammography. In this work, we design three ensemble classification structures that fuse information from multiple sensors to detect abnormalities in the breast. A principled Neyman–Pearson approach is developed to allow control of the trade-off between false positive rate and the false negative rate. We evaluate performance using data derived from measurements of heterogeneous breast phantoms. We also use data collected in a clinical trial that monitored 12 healthy patients monthly over an eight-month period. In order to assess the efficacy of the proposed algorithms we model scans of breasts with malignant lesions by artificially adding simulated tumour responses to existing scans of healthy volunteers. Tumour responses are constructed based on measured properties of breast tissues and real breast measurements, thus the simulation model takes into account the heterogeneity of the breast tissue. The algorithms we present take advantage of breast scans from other patients or tissue-mimicking breast phantoms to learn about breast content and what constitutes a “tumour-free” and “tumour-bearing” set of measurements. We demonstrate that the ensemble selection-based algorithm, which constructs an ensemble of the most informative classifiers, significantly outperforms other detection techniques for the clinical trial data set.

1. Introduction

Early detection of breast cancer is vital for successful treatment [1]. Microwave imaging and detection methods have been intensely researched in recent years as a complementary modality for breast cancer screening. Such methods are based on the reported inherent contrast of dielectric properties of healthy and malignant breast tissues over the microwave frequency range [2,3]. Microwave techniques promise non-invasive screening with low-cost system fabrication and operation and have been applied to other fields including stroke detection [4]. Scans do not require breast compression and can be repeated frequently since no ionizing radiation is used. The aim is not to replace mammography, ultrasound, or MRI, but to develop an alternative approach that can act as an early warning system to flag the need for more comprehensive testing.

Most of the previous work on microwave breast cancer screening has concentrated on imaging. Algorithms are applied to measurement data to generate images that can be interpreted by a clinical expert. Microwave radar and microwave tomography are two common techniques in the microwave imaging field. Tomographic methods are used to reconstruct a dielectric profile of breast tissues [5] by solving an ill-conditioned inverse problem. Radar methods, on the other hand, generate a map of scattering regions within the breast. Tomography methods have been applied in experimental imaging of both phantoms [6,7] and patients [5,8]. Radar imaging approaches include beamforming algorithms [9–11] and hypothesis testing techniques [12]. Results have been reported for delay-and-sum and other beamforming algorithms on data collected in clinical trials [13,14,11].

Recently, some research has explored the application of machine learning techniques, in particular classifiers, to measurements collected from microwave breast cancer screening systems [15–19]. Classification techniques have been applied to characterize a tumour using microwave backscatter [15,16] with the assumption that the tumour has already been detected. In [15],

---

* Corresponding author at: Room 633, McConnell Engineering Building, 3480 University Street, Montreal, Québec, Canada H3A 0E9.

E-mail addresses: yunpeng.li@mail.mcgill.ca (Y. Li), emily.porter@mail.mcgill.ca (E. Porter), adam.santorelli@mail.mcgill.ca (A. Santorelli), milica.popovich@mcgill.ca (M. Popović), mark.coates@mcgill.ca (M. Coates).

http://dx.doi.org/10.1016/j.bspc.2016.09.003
1746-8094/

architectural tissue features such as shape and size are inferred from the backscatter by using linear classifiers with local discriminant bases and principal component analysis (PCA). Conceição et al. introduce a support vector machine (SVM)-based classifier that distinguishes between benign and malignant tumours according to their shape [16].

There has also been some recent work towards the detection task for microwave breast cancer screening systems. In [17], a discrete cosine transform (DCT) is applied to the received signal for feature extraction, and neural networks are used to detect the tumour existence, size and location. In [18], Byrne et al. apply SVM to features extracted from backscattered signals using PCA. A separate SVM classifier is applied to each measured signal; a tumour is detected for the breast if the majority of the classifiers decided that it was present. Building on this work, we described a strategy in [19] that fused data from random antenna pairs to improve the SVM classifier accuracy.

Our work focuses on the development of a microwave breast cancer screening system and the associated algorithms that can process measurements to make a decision as to whether a tumour is present in the breast. This system could offer women the option of self-screening at home on a regular (e.g., monthly) basis. The monthly (as opposed to annual) tests would be especially beneficial to those in the high-risk category, as frequent monitoring increases the chance of early-stage tumour diagnosis and, consequently, successful treatment. We envision that the system will track breast health by comparing the current breast scan to past scans of the same patient and to other patient scans, stored in a clinical database.

With this long-term goal and motivation in mind, several essential milestones have been reached to date. We have developed a time-domain microwave radar system for breast screening. Time-domain measurements potentially offer advantages over frequency-domain, including faster scan times and more cost-effective equipment solutions, with the drawback of a slightly lower signal-to-noise ratio [6]. We have demonstrated successful imaging of tumours in realistic tissue phantoms [20]. Recently, we have conducted a clinical trial with 12 patient volunteers for breast health monitoring [21].

This paper presents a novel application of classification methods to clinical data collected from a microwave breast screening system to detect the presence of a tumour. Our main contributions compared to state-of-the-art work in this domain are the following: (i) we employ a principled Neyman–Pearson approach to select algorithmic parameters in order to control the false positive rate while minimizing the false negative rate (most past work in microwave breast cancer detection did not differentiate between these two types of errors); (ii) we design three ensemble classification architectures to fuse information from different antenna pairs; (iii) we demonstrate the performance of our classification techniques using data collected in a clinical trial that monitored patients monthly over an eight-month period. Preliminary results concerning this work and the efficacy of imaging-based algorithms with clinical trial data have been published in abbreviated forms [22,23]. This paper extends our previous work by proposing the ensemble selection-based classification method which significantly outperforms existing methods. It also provides a more detailed description of our algorithms, a new data-adaptive tumour response simulation procedure that factors in the heterogeneous propagation environment inside the breast, and a more complete performance evaluation involving both a breast phantom data set and a clinical trial data set.

The remainder of the paper is organized as follows: Section 2 introduces our system, data, and the ensemble classifier. We report and discuss experiment results in Section 3 and Section 4, and provide a summary in Section 5.

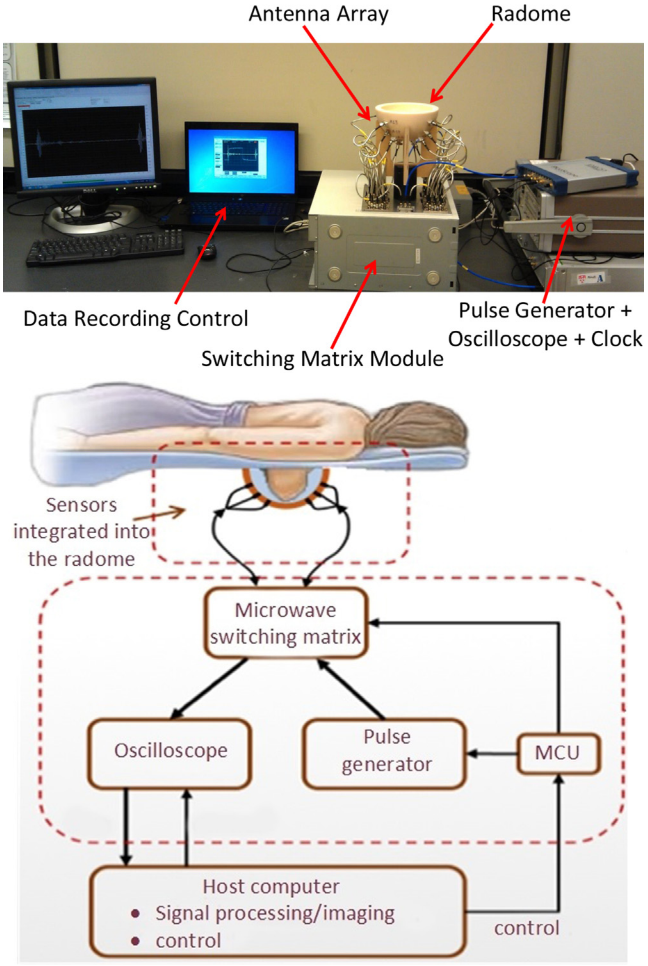

Fig. 1. Top: the experiment system we use to collect the data. Bottom: a graphical illustration of the system prototype for this experiment.

2. Materials and methods

2.1. System overview

The system uses multiple antenna sensors to collect the transmitted and reflected signals from the breast. The core of the system (Fig. 1) is a hollowed-out hemispherical dielectric radome, which houses both the breast under test and the 16-element antenna array. The radome is a ceramic dielectric made from alumina (with relative permittivity $\epsilon_r = 9.6$) [24]. The antennas are travelling-wave, resistively-loaded sensors that are designed for operation in the vicinity of breast tissues [25]. When a breast scan recording begins, a short-duration Gaussian-modulated pulse is generated and shaped, using a passive microwave filter, such that its frequency content is concentrated in the 2–4 GHz range [26]. The pulse is amplified and then input into an automated 16 × 2 switching matrix that selects each antenna as the transmitter in turn. The pulse is scattered off of the breast tissues, i.e., at all interfaces between tissue types, and is then collected by the selected receiving antenna. An equivalent-time sampling oscilloscope records the data. Then, a different transmit-receive antenna pair is selected until all possible combinations have been cycled through. With 16 antennas, a total of 240 signals are obtained per breast scan.

For performing breast scans on patients, the system is integrated in a way that it can be easily used to collect patient measurements in clinical trials. All equipment is placed under a table and the patient lies facing down on the table with their breast in the radome, which