Classification of breast tissue by electrical impedance spectroscopy

J. Estrela da Silva[1] J. P. Marques de Sá[1] J. Jossinet[2]

[1]INEB (Instituto de Engenharia Biomédica), Faculdade de Engenharia da Universidade do Porto, Porto, Portugal
[2]INSERM (Institut National de la Santé et de la Recherche Médicale), Lyon, France

Abstract—*Electrical impedance spectroscopy is a minimally invasive technique that has clear advantages for living tissue characterisation owing to its low cost and ease of use. The present paper describes how this technique can be applied to breast tissue classification and breast cancer detection. Statistical analysis is used to derive a set of rules based on features extracted from the graphical representation of electrical impedance spectra. These rules are used hierarchically to discriminate several classes of breast tissue. Results of statistical classification obtained from a data set of 106 cases representing six classes of excised breast tissue show an overall classification efficiency of ~92% with carcinoma discrimination >86%.*

Keywords—*Impedance spectroscopy, Tissue characterisation, Data classification, Discriminant analysis, Breast cancer*

Med. Biol. Eng. Comput., 2000, **38**, 26–30

1 Introduction

ELECTRICAL IMPEDANCE techniques have long been used for tissue characterisation and in monitoring applications, of which impedocardiography is the best known (KUBICEK *et al.*, 1970). These techniques have also enabled impedance mapping (TACHIBANA *et al.*, 1970; HENDERSON and WEBSTER, 1978) and, more recently, dynamic imaging (BROWN *et al.*, 1994).

Specific impedance (also termed 'impedivity') is the AC equivalent of resistivity for DC current. The specific impedance of a tissue is determined by its electric and dielectric properties, which depend, among other things, on the cell concentration, membrane capacitance, electric conductivity in interstitial space and the intracellular medium (SCHWAN, 1959; FOSTER and SCHWAN, 1989). The easiness, low cost and minimum invasiveness are praised features of the impedance techniques.

In recent decades, electric and dielectric measurements have been carried out in breast tissue under a range of experimental conditions including *in-vivo* or *ex-vivo* measurements and using various measurement techniques (SUROWIEC *et al.*, 1988; MORIMOTO *et al.*, 1990; CAMPBELL and LAND, 1992; HEINITZ and MINET, 1995). In the 488 Hz–1 MHz range, significant differences in the impedivity modulus and phase angle from among six groups of breast tissue were found by one of the authors (JOSSINET, 1998).

These findings suggest that electrical impedance spectroscopy (EIS) is potentially usable for the discrimination of breast tissue and especially for the detection of breast cancer. The present paper describes a method for classifying breast tissues based on EIS. These new results were attained using features derived from the Argand plot of the impedivity data collected in freshly excised tissue. The set of features used comprised those defined in a previous study (JOSSINET and LAVANDIER, 1998) as well as additional features selected for their discrimination ability. A statistical hierarchical approach was adopted using twelve-point and seven-point spectra.

Correspondence should be addressed to Dr J. P. Marques de Sá; email: jmsa@fe.up.pt

First received 5 July 1999 and in final form 11 October 1999

2 Materials and methods

2.1 *Data set*

The initial data set consisted of 120 spectra recorded in samples of breast tissue from 64 patients undergoing breast surgery. Each spectrum consisted of twelve impedance measurements taken at different frequencies ranging from 488 Hz to 1 MHz. Details concerning the data collection procedure as well as classification of the cases and frequencies used are given elsewhere (JOSSINET, 1996; JOSSINET, 1998). Prior to any classification attempt, 14 spectra were discarded since they exhibited manifestly abnormal features (erroneous I_0, low D_A) due to poor tissue collection care and/or data measurement. In the remaining 106 cases, the following six classes of tissue were represented:

Normal tissue classes:	connective tissue (**con**):	14 cases
	adipose tissue (**adi**):	22 cases
	glandular tissue (**gla**):	16 cases
Pathological tissue classes:	carcinoma (**car**):	21 cases
	fibro-adenoma (**fad**):	15 cases
	mastopathy (**mas**):	18 cases

2.2 *Bioelectrical features*

The classification features were extracted from EIS plots on the Argand plane the co-ordinates of which were the real impedance part (R_s) and the negative of the imaginary impedance part ($-X_s$). Ideally, those plots have the form of circular arcs, as shown in Fig. 1*a*, due to the underlying physical model (SCHWAN, 1959; FOSTER and SCHWAN, 1989). Such arcs are characterised by three parameters, classically the intercepts with the horizontal axis (low-frequency and high-frequency limit

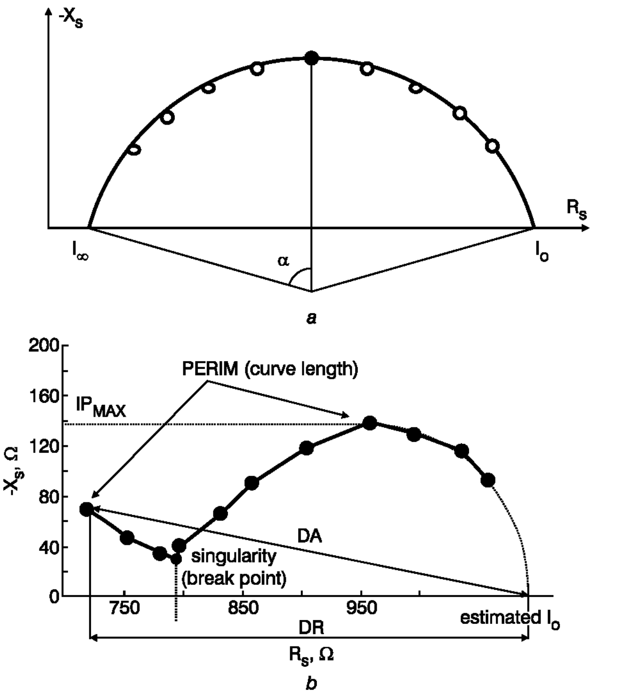

Fig. 1 *Impedance locus in Argand plane: a) ideal locus; b) locus example with computed features*

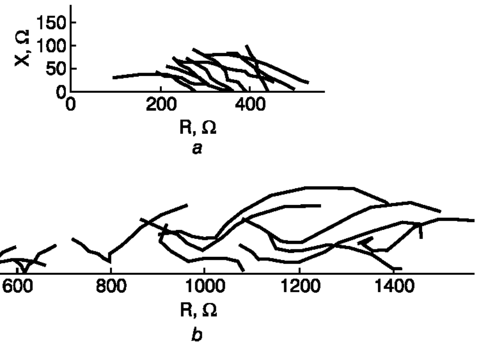

Fig. 2 *Argand plots: a) carcinoma tissue samples; b) connective tissue samples*

resistances, denoted I_0 and I_∞, respectively) and the fractional power, α, controlling the shift of the centre from the axis. The frequency at which the imaginary part passes by a maximum (IP_{MAX}) characterises the distribution of the frequency points on the arc.

However, due to the various time constants present in the tissues and the limited frequency range, superimposed circular arcs and, generally, incomplete arcs such as the ones shown in Fig. 2 are in practice obtained. In the present study this made it impossible to calculate some of the features derived from the circular arcs (mainly I_∞ and α). Non-conventional features were therefore used.

2.3 *Data pre-processing*

The objective of the data pre-processing stage was to obtain a group of features capable of characterising of the impedivity loci in the Argand plane. Initially the plots of the various cases in the Argand plane were visually compared. It was expected that obvious distinctions would be found. For instance, the plots of the carcinoma tissues (Fig. 2*a*) are visually distinct from the plots of the connective tissues (Fig. 2*b*). Previous reports on excised tissue discrimination (JOSSINET and LAVANDIER, 1998) used the following impedance features, which are shown in Fig. 1*b*:

I_0	impedivity at zero frequency (low frequency limit resistance)
PA_{500}	phase angle at 500 kHz
S_{HF}	high-frequency slope of phase angle (at 250, 500 and 1000 kHz points)
D_A	impedance distance between spectral ends
BREAK	spectral break point
NOTCH	spectral notch point

In the present study, we also used the following new non-conventional features with a view to achieving a better classification performance than reported previously:

AREA	Area under spectrum
AREA_D_A	Area normalised by D_A
IP_{MAX}	Maximum of the spectrum
D_R	Distance between I_0 and real part of the maximum frequency point
PERIM	Length of spectral curve
SLOPE	Slope of spectral curve obtained by linear regression

All the above features were calculated in Excel 7.0 using programs in Excel's Visual Basic For Applications (VBA) programming language for each of the 106 impedance spectra. PA_{500}, S_{HF} and D_A were directly calculated. The calculation of the remaining features was aided graphically.

2.4 *Classification*

Given the reduced number of cases representing each class in the available data set, we used a linear discriminant analysis approach to classify the spectra. This seemed to be a sensible choice since the design of linear classifiers is much less demanding in the number of parameters needed to estimate compared to other approaches such as neural nets or K-nearest neighbour classifiers, which are quite demanding in terms of the number of parameters needed to estimate and therefore on the size of the training sets.

Statistical characterisation of the aforementioned set of features was performed in order to assess the features' distribution and discrimination ability. The tool used was STATISTICA 4.5 for Windows (StatSoft™). This characterisation involved normality tests (Kolmogorov-Smirnov, Shapiro-Wilk), discrimination tests (Kruskal-Wallis), inspection of box-and-whisker plots and correlation computations. The Shapiro-Wilk test was used because of its high performance, compared to alternative tests, in the detection of small deviations from normal distribution (SHAPIRO *et al.*, 1968).

The inspection of the box-and-whisker plots and the Kruskal-Wallis test results suggested that a sensible approach to this difficult six-class classification problem was to perform it in two hierarchical stages. The straightforward one-step six-class discrimination (based on the design of six linear discriminant functions) performed quite poorly, as reported in Section 3.2. Hierarchical approaches may be helpful in difficult multiple class separation problems (SWAIN, 1977), which commonly characterise biomedical data.

Based on the discriminative power of the features and the inspection of other statistical results (e.g. scatter diagrams and box-and-whisker plots), a two-stage hierarchical approach was devised as follows.

In the first stage, two groups of class were considered: fatty tissues (adipose and connective tissues) and the other four classes taken together. The second stage involved classifying this four-class group by attempting to split it into the following