A NEURAL NETWORK BASED APPROACH TO SOLVE THE ELECTRIC CAPACITANCE TOMOGRAPHY INVERSE PROBLEM

Sara Carcangiu and Augusto Montisci

Department of Electrical and Electronic Engineering, University of Cagliari, Piazza d'Armi 09123 Cagliari (Italy)
1scarcangiu@diee.unica.it, 2amontisci@diee.unica.it

ABSTRACT

In this work a novel approach to solve the 2D Electric Capacitance Tomography (ECT) inverse problem is presented. The aim of this work is to improve the spatial resolution of the ECT method to test polymeric components for large-scale manufacturing and assembling industry, preserving the time resolution and the deterministic computational time of this non-destructive technique. The data obtained from Finite Elements techniques have been used to train an Artificial Neural Network (ANN) in order to solve the direct problem, namely calculating the capacitance for a given permittivity of the medium affected by defects in different positions and dimensions. In particular, the dimensions and the barycentre positions of the defect are the input of the ANN and the simulated capacitances represent the output. Once the ANN has been trained, in order to perform the test of the material, an inverse problem has to be solved, specifically determining the characteristics of the defect on the basis of the measured capacitances. To do this, the trained ANN can be exploited, by inverting the neural network instead of the problem itself. Results showed good accuracy in the identification of the position of the defects.

Key words: Electric Capacitance Tomography, Finite Element Method, Artificial Neural Network Inversion, Polymeric Material, Inverse Problems.

1. Introduction

In recent years, electrical capacitance and electrical resistance tomography systems have been developed for process tomography applications [1, 2]. Electrical capacitance tomography (ECT) is a non-invasive and non-destructive imaging technique that uses electrical capacitance measurement at the periphery of an object [3]. In particular, in the last years the ECT has been applied in the chemical control process engineering in order to detect the presence and the distribution of undesirable phases (solid or gaseous) in fluid flows and it has been appreciated mainly for its very high time resolution (up to 100 Samples/second) [4]. However the spatial resolution and the pattern recognition capability of this method at the moment are lower than other non-destructive techniques (NDT), in any case unserviceable for this kind of application (e.g. ultrasound NDT, hard-field X-ray tomography). The aim of this work is to improve the spatial resolution of the ECT method to test polymeric components for large-scale manufacturing
and assembling industry (i.e. automotive, aerospace, computer and mobile electronics), preserving the time resolution and the deterministic computational time of this NDT.

ECT involves the measurement of changes in capacitance from a multi-electrode sensor due to the change in permittivity of materials being imaged, and the reconstruction of cross-sectional images using the measured data and a suitable algorithm [5]. The problem of determining the unknown model parameters is usually identified as "inverse problem". Image reconstruction in ECT is a nonlinear and ill-posed inverse problem.

Different approaches have been used in literature in order to solve this problem. Generally, two types of reconstruction techniques can be used for ECT image reconstruction: non-iterative and iterative algorithms. Because of the nonlinear relationship between the measured capacitance and the permittivity distribution in ECT, iterative techniques are prevalently used [6]. Iterative algorithms for ECT are based on obtaining an estimate of the unknown permittivity distribution from the capacitance data (inverse problem) and calculating the capacitance based on the estimated permittivity distribution to update the image in the next scheme (direct problem). This process is repeated iteratively until the capacitance error is decreased to a satisfactory value [6].

With respect to the resolution of the inverse problem, in this work we propose the use of an innovative ANNs based methodology for solving the inverse problem of locating defects in a parallelepiped of polymeric material. In particular, a Multi Layer Perceptron (MLP) neural network is trained to solve the direct problem, namely associating the measured value of capacitance to an assigned defect, which univocally corresponds to a permittivity map. By means of a finite element method (FEM) software the physic problem is simulated. Numerical analysis of the polymeric material of rectangular cross section is performed to generate training and testing samples for neural network assessing task. In this case by means of the FEM simulations the capacitances for a given permittivity of the medium are calculated varying the position and the dimensions of the defects. Subsequently, in order to identify the spatial location (X, Y) and the dimensions (W, H) of a defect, an inverse problem has to be solved. The measured values of capacitance are imposed as output of the neural network, and the corresponding input is determined. Then the neural network is inverted instead of the problem itself.

2. **ECT system**

An ECT system is generally composed of three different units (see Fig. 1).

![Fig. 1: Electrical Capacitance Tomography system.](image)

The components of the system are: 1) the capacitance sensor which comprises of an array of electrodes attached to periphery of the structure of polymeric material which has to be imaged, 2) the capacitance measuring unit to acquire and process the signals obtained from the
capacitance sensor, and 3) the control computer to reconstruct and display the permittivity distribution image using the data obtained.

The sensor, as depicted in Fig. 1, consists of a number of N electrodes placed symmetrically around the region of interest providing number of independent capacitance measurement used for image reconstruction. During the period of a scanning frame, in order to increase the scanning speed, multiple excitation signals, e.g., alternating current voltages at different frequencies, are applied simultaneously to multiple electrodes on the ECT sensor positioned in the same side of the structure as one of the electrodes in the opposite side is kept at the ground potential, acting as detector electrode. Subsequently, the voltage potential at each of the remaining electrodes is measured, one at a time, by the measurement electronics to determine the capacitance between the electrodes. The electric field distribution inside the region of interest is a function of permittivity distribution according to Poisson equation

\[ \nabla \cdot \left( \varepsilon(x, y) \nabla \varphi(x, y) \right) = -\rho(x, y) \]  

where \( \varepsilon(x, y) \) is permittivity distribution, \( \varphi(x, y) \) is electrical potential distribution, and \( \rho(x, y) \) is charge distribution. The Poisson equation is a linear partial differential equation in terms of \( \varphi(x, y) \). The nonlinearity of ECT results from the nonlinear dependence between the field distribution \( V(x, y) \) (and, hence, the measured capacitance data) and the (unknown) permittivity distribution (solution) \( \varepsilon(x, y) \). The mutual capacitance between two pair of electrodes, source, and detector is obtained through \( C_{ij} = \frac{Q_{ij}}{\Delta V_{ij}} \), where \( C_{ij} \) represents the mutual capacitance between electrodes i and j, \( \Delta V_{ij} \) the potential difference, and \( Q_j \) is the charge on the sensing electrode which is found by applying Gauss law

\[ Q_j = -\oint_{I_j} \varepsilon(x, y) \nabla V(x, y) \cdot \hat{n} dl \]  

where \( I_j \) is a closed path enclosing the detecting electrode and \( \hat{n} \) is a unit vector normal to \( I_j \).

Considering this nonlinear dependence, as already said in the introduction, the iterative procedure usually used to solve this problem has been replaced in this work by a new neural network based procedure. In order to solve the inverse problem it is not suitable developing the neural model of the inverse problem itself. This is due to the fact that the relationship between map of the permittivity and capacitance is not bi-univocal [7].

In the first step, an ANN is trained to solve the direct problem, by means of a set of examples, to associate the properties of the polymeric material affected by defects to a set of capacitance values. The input patterns are positions and dimensions of the defects inside the structure to be inspected. The output patterns are the capacitance values obtained considering a known configuration of electrodes. After the training, the ANN generalization capability can be exploited to estimate the capacitance values corresponding to a new defect position and dimension. In the second step, the trained ANN is inverted in order to solve the inverse problem. On the basis of values of measured capacitance values, the defect position and dimension can be identified.

3. Forward problem

The forward problem consists in determining the output response of an ECT system given the permittivity distribution in the region of interest with defect.

3.1 Data set

The capacitance values of a structure of polymeric material with defect have been simulated by means of the FEM, using electrostatic module of Comsol®. An ECT sensor of 10 electrodes in a
symmetrical arrangement as illustrated in Fig. 2 is used to collect the capacitance data set \( C_{ij} \) used for training. The numerical computations have been carried out simulating the layout in Fig. 2 where the input is a sinusoidal voltage of 1000 V and frequency of 10 MHz and the resistance \( R \) is equal to 100 \( \Omega \). A transient numerical analysis has been carried out by means of the finite element method. Data have been generated by simulating the electrostatic phenomenon in a polymeric structure of permittivity \( \varepsilon \) equal to 2.3 with voids of various dimensions in different positions. An analysis of the model without defects has been firstly performed in order to set the simulation parameters, i.e., the dimensions of the elements in the mesh of the FEM model, and the sampling time.

![Fig. 2: Logical layout of the set-up.](image)

In a second phase, the same model has been considered, which presents different defects. Varying the position of the cavity into the inspected section of polymeric material, several simulations have been performed by varying the height \( H \) and the width \( W \) of voids from 1 cm to 3 cm with a step of 1 cm and from 2 cm to 4 cm with a step of 1 cm respectively. The resulting data set is composed by 266 defect cases plus the fault free case. For each defect case a matrix of dimension 5x5 of capacitance values is obtained as output. In Fig. 3 the plot surface of the electrical potential and the streamlines of the electrical field obtained after a numerical simulation of a defect case with a particular electrodes configuration are shown.

![Fig. 3: Numerical simulation of a defect case. The black streamlines represent the electric field.](image)

### 3.2 Neural model construction

An ANN is a network of elementary processing elements (neurons) which are interconnected by means of weighted connections. By means of said weights, the ANNs are able to store different kind of information, such as patterns, input-output relationships, dynamics and so on. The most appealing feature of ANNs is that they can store the desired information simply on the basis of a collection of examples, rather than on a mathematical model of the system at hand. The process of storing information into an ANN is called training, and usually it is performed by applying an
algorithm whose purpose is minimizing some error function. One of the most popular layouts of ANN is the MLP (see Fig. 3) [8]. The MLP is arranged in two or more layers of neurons, where each layer is connected only to the previous layer and to the successive one, whereas there aren’t connections between neurons belonging to the same layer. The neurons of the first layer (input) receive the signal to be processed, while the last layer (output) returns the result of the elaboration. Between the input and the output layer the MLP can have one or more intermediate (hidden) layers, which do not communicate with the outside.

The training of the NN has to be performed carefully, in order to guarantee the possibility to invert it in the successive phase. To this end, the trained ANN has to fulfill three fundamental proprieties: precision, generalization capability, invertibility. These properties are in conflict each other, because overtraining the NN on the training set in general we reduce the generalization capability, namely the attitude to furnish precise output in correspondence of examples not included in the training set. On the other hand, we can improve the precision of the network by increasing the number of intermediate neurons and in general there is a range of such number that does not deteriorate the generalization capability, but in order to univocally associate an input to a given output of the NN, the intermediate neurons has to be less than or equal to the output ones, so as the same constraint holds for the input neurons with respect to the intermediate ones. The reason of that is explained in the following.

The first decision to make concerns the structure of the ANN. As demonstrated in [9] an MLP ANN with only one hidden layer of sigmoidal neurons is an universal approximator, therefore we will assume such an elementary structure for the ANN. Due to the constraint on the number of neurons of each layer, it is preferable to have high dimensional vectors to represent the output and low dimensional vectors to describe the inputs. For this reason in this work the capacitance values, which represent the output of the NN are used without applying any data compression technique, whereas as input four coordinates are used which correspond to the dimensions and the position of the defect. As said before, the number of output neurons represents an upper bound for the number of hidden neurons. In order to maximize the degrees of freedom of the ANN, the number of hidden neurons is assumed equal to that of output neurons.

As described in detail below, in order to make univocal the inversion of the ANN, it is necessary that both the two matrices of weights (between input and hidden layers and between hidden and output layers) are full rank. Furthermore, it is important to avoid that the hidden neurons saturate for some example. In fact, a saturating function, such as the sigmoid, in correspondence of the flat zone is impossible to be inverted. Usually the training algorithm do not care of the rank of weights matrices or the saturation of the neurons, therefore some supplementary check have been included during the training phase. Such control is made any few training epochs, and if the check result is positive a small correction is performed on the weights in order to fix the problem. In particular, if the full rank condition is not verified, a diagonal matrix with small values is added to the matrix, while if the saturation condition is not verified, the whole matrix is multiplied by the minimum factor which allow to avoid the saturation. After that the adjustment has been performed, the training resumes, until the goal is met or a new constraint violation occurs. It is important that the training does not terminate with an adjustment, because it unavoidably introduces an error in the NN.
Another aspect of training which affects the generalization capability concerns the construction of the training set. Indeed, if the ANN is trained on a non representative set of examples the generalization capability will be unsatisfactory. In order to face this problem, a growing method has been used to select the training set. It consists in a first stage in training the ANN on a limited subset of available examples and in using the remaining ones as validation set. The performances of the trained ANN are then evaluated on the validation set and the examples which corresponds the worst performance is added to the training set. After that, the NN is trained on the new set and so on, until the error on the validation set is below a prefixed threshold or a prefixed number of examples have been included in the training set.

4. Inverse problem

The inverse problem consists in determining the dimensions and position of a defect inside a polymeric structure given the capacitances values obtained by an ECT system.

4.1 Neural Network inversion

Given an MLP trained as described above, the inversion procedure has to be applied to deduce the unknown permittivity map of the object under test on the basis of a series of measurements of capacitance. Such measurements constitute the output pattern of the trained ANN, and the inversion consists in determining the corresponding input.

The MLP (Fig. 3) realizes a relationship between input and output patterns described by the following algebraic equations system:

\[
\begin{align*}
\text{Input layer} & : W_1 \cdot x + b_1 = y \\
\text{Hidden layer} & : h = \sigma(y) \\
\text{Output layer} & : W_2 \cdot h + b_2 = u
\end{align*}
\]  

(3)

where: \(x\) is the input of the network, \(W_1\) is the weights matrix of the input layer, \(b_1\) is the bias vector of the input layer, \(y\) is the input of the hidden layer, \(h\) is the output of the hidden layer, \(\sigma(\cdot)\) is the hidden neurons logistic activation function, \(u\) is the output of the network, \(W_2\) is the weights matrix of the output layer, \(b_2\) is the bias vector of the output layer.

On the basis of the known output of the system, which derives from the measurements of capacitance, the corresponding input can be calculated exploiting the method described in [10].

On the basis of the third equation in (3), starting from the output \(u\), the vector \(h\) can be determined. Provided that the matrix \(W_2\) is full rank, and taking into account that in the present case such matrix is squared, the solution corresponding to the minimum sum squared error is equal to:

\[
h = W_2^{-1} \cdot (u - b_2)
\]

(4)

More in general, if the matrix \(W_2\) is rectangular it cannot be directly inverted. In order to guarantee the uniqueness of the solution, the rows (number of output neurons) must be more than the columns (number of hidden neurons) and the matrix has to be full rank. If this is the case, the equations system is overdetermined and the uniqueness is ensured by assuming the solution which corresponds to the minimum mean squared error.

Such solution can be found by solving the following modified equations system, whose coefficients matrix is squared.
where the mark \( T \) denotes the transposition operator. The second equation in (3) states a biunivocal relationship between \( y \) and \( h \), therefore the vector \( y = \sigma^{-1}(h) \).

Finally, provided that also the matrix \( W_1 \) is full rank, the input pattern \( x \) can be calculated as:
\[
    x = \left( W_1^T \cdot W_1 \right)^{-1} \cdot W_1^T (y - b_1)
\]
and it represents the size and the position of the defect.

### 5. Results

In the studied case, to solve the direct problem, a 4 inputs-25 hidden-25 outputs structured MLP is trained with 235 examples by means of the Levenberg-Marquardt (LM) algorithm \([11]\). The following results refer to the classification errors evaluated on an independent test set consisting of 22 examples, not used during the training procedure.

Figure 4 shows the actual values of the position of the barycentre and the corresponding values predicted by the neural classifier for 8 test examples. Figure 5 shows as the actual as the predicted tomography map of the studied polymeric structure showing the dimension of the defect in the case of 2 test examples.

As can be noted, the method is able to detect both the defect x and y coordinates and to estimate the dimensions with good accuracy in the tested configurations.

**Fig. 4:** Position of the barycentre in the XY plane: actual (○) and predicted (*) values.

**Fig. 5:** Dimensions of 2 defect examples: actual (black) and predicted (red).
6. Conclusions

In the present paper the use of an innovative ANNs based methodology for solving the inverse problem of locating defects in a parallelepiped of polymeric material has been proposed. The results show that the proposed technique can be efficient for the diagnosis purpose. Future research will investigate the method applicability using a more realistic model of plastic objects used in industrial environment. Moreover, tests on a real world products made of plastic will be performed.

7. Acknowledgment

This work has been performed within the scope of the project “Non destructive ultrasound tests based on pseudo orthogonal sequences for imaging and automatic classification of industrial products”, which is funded by MIUR within the framework of the programme PRIN 2009.

8. References