ECT sensor design using machine learning techniques

Y. Oussar, C. Margo and J. Lucas
LPEM, UMR8213
ESPCI-ParisTech
10 Rue Vauquelin, 75005 Paris, France

S. Holé
LPEM, UMR8213
UPMC-Sorbonne Universités
10 Rue Vauquelin, 75005 Paris, France

Abstract—**Within the framework of image reconstruction by Electrical Capacitance Tomography (ECT) sensing, we investigate the relevance of the sensor structure embodied in both the number and the size of the electrodes. While most of the studies in the literature exhibit sensors with an assumed arbitrary structure, we consider that these two properties possess a significant impact on the sensor performances. In our study, the emphasis is about the detection of a single bubble with random size and position. We propose to determine the architecture of the sensor that leads to the most accurate image reconstruction. To achieve this objective, we propose to determine the image from a set of independent measurements using LS-SVM models selected with a sophisticated validation method. Indeed, this way to proceed ensures a faster image computation than inverting an underdetermined set of linear equations. By doing so, the computational burden is reduced since it leads to a direct calculation instead of an iterative optimization. Various numerical experiments are presented and discussed. They show the effectiveness of our assumption.**

I. Introduction

Electrical Capacitance Tomography (ECT) is a well-known and a widely used sensing technique. Generally, an ECT sensor is formed by a set of electrodes arranged circularly. The sensing technique is founded on measuring the electrical capacitances appearing between all the possible couples of electrodes. Typically, an ECT sensor can be used to inform about the spatial configuration (size, position) of a non-conducting material which permittivity is different from the one of the environment.

The image of a cross-section of the environment of interest can be obtained using the available measurements. Usually, the image pixels correspond to nodes from a spatial discretization scheme that involves a linear knowledge-based model. This model describes the relationship between the measurements and the permittivity at each node. The more the space discretization grid is fine, the better the image resolution. Constructing the image leads to invert an underdetermined set of linear equations [1].

The literature dedicated to ECT sensors presents various applications. While both the number and the size of electrodes have great importance regarding the sensor performances, they are usually set arbitrarily and seldom discussed [2]. Except for few works [3], the overall independent measurements are readily taken into account. One can wonder if all the measurements are relevant or just few of them are sufficient to achieve the best image reconstruction given the sensor structure. Moreover, one might ask: what should be the optimal structure of the sensor for a given application?

In the present study, we propose to pay attention to the design of ECT sensors. We assume that tuning the number and the size of electrodes allows to optimize the sensor performances for a given application. More precisely, we are interested in determining the sensor architecture that best predicts the size and the position of a bubble lying in the sensor environment. To achieve this goal, we propose to implement non linear black box models using LS-SVM to establish the relationship between the sensor measurements and the bubble size and position.

II. ECT sensor and problem statement

Figure 1 illustrates the structure of a typical ECT sensor considered in this study. It is formed by eight electrodes arranged circularly. The environment delimited by the sensor is characterized by a permittivity ε_1. In Figure 1, the gap between electrodes is relatively small. This may not be always the case since the size and the number of electrodes have a significant impact on the sensor performances.

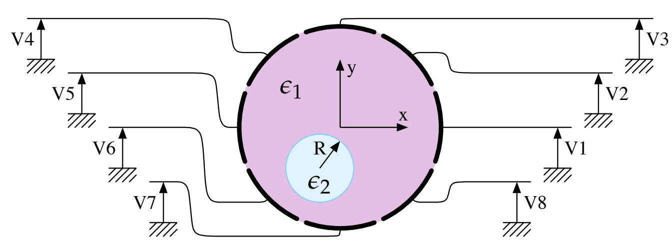

Figure 1. An 8-electrode ECT sensor

From the electrical point of view, each electrode can be polarized independently. When electrode i is polarized, the measurement of the charges on electrode j allows the capacitance C_{ij} between these two electrodes to be determined. Due to the circular symmetry of the ECT sensor, there are at most $S = \frac{1}{2}P(P-1)$ independent capacities for a given P-electrode sensor.

In the present study, we are interested in optimizing the structure of ECT sensors for image reconstruction by setting the number and the size of the electrodes to their optimal values for a given sensing task. We assume that an ECT sensor is fully defined by giving the number and the size of the electrodes. Indeed, this couple sets the complete sensor structure. The basic motivation for this study is to lead to the most accurate and computationally cost effective image reconstruction.

We consider that a non conducting single bubble with permittivity ε_2 lies in the closed region circumscribed by the sensor electrodes. This bubble has unknown size and position respectively pointed by its radius R and its coordinates x and y. Hence, according to this formulation, the image reconstruction consists in estimating as accurately as possible both the bubble radius and coordinates.

Instead of dealing with an image reconstruction founded on inverting an undetermined set of equations, we assume that a black-box inverse model that predicts the size and the position of the bubble exists and can be trained using data generated by numerical experiments. This approach allows an almost instantaneous image reconstruction which makes it highly attractive for a real-time implementation. The black-box models are implemented using the LS-SVM technique.

III. Training and validating LS-SVM models

A. LS-SVM model description

In spite of several and efficient techniques for non linear static modeling, such as neural networks, Least Square Support Vector Machines (LS-SVM) are attractive candidates thanks to various interesting properties: they are linear-in-their-parameters models, their training algorithm consists in a quadratic minimization and they have a built-in regularization mechanism [4]. As a result, these properties confer to them the ability to build models with high generalization capabilities by avoiding overfitting and controlling model complexity. Given a set D formed by N measurements of S variables and a scalar z: $D = \left\{(\mathbf{t}_i, z_i) \in \mathbb{R}^S \times \mathbb{R},\ i = 1, ...N\right\}$, we are interested in nonlinear models of the form:

$$\hat{z} = \mathbf{w}^T \varphi(\mathbf{t}) + b \tag{1}$$

where $\mathbf{t}$ is the vector of input variables, $\hat{z}$ is the model output, $\mathbf{w}$ and b are unknown parameters and $\varphi(\cdot)$, usually introduced by a kernel function, is a nonlinear mapping from the original feature space to a high dimensional feature space. The LS-SVM training procedure consists in minimizing the cost function defined by:

$$J(\mathbf{w}, \mathbf{e}) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N} e_i \tag{2}$$

subject to the equality constraints $e_i = z_i - \hat{z}_i$ with e_i the prediction error taking into account the target output z_i and the predicted output $\hat{z}_i$. N is the size of the training set and C is the regularization parameter. This optimization problem can be cast into a dual form with unknown parameters $\boldsymbol{\alpha}$ and b, $\boldsymbol{\alpha}$ being the vector of the Lagrange multipliers [5]. The parameters can be computed by resolving the following system of linear equations:

$$\begin{bmatrix} \mathbf{K} + \frac{1}{2C}\mathbf{I_N} & \mathbf{1}_N \\ \mathbf{1}_N^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ b \end{bmatrix} = \begin{bmatrix} \mathbf{z} \\ 0 \end{bmatrix} \tag{3}$$

with $\mathbf{1}_N = [1, 1, ..., 1]^T$, $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, ..., \alpha_N]^T$ and $\mathbf{I_n}$ is the identity matrix. $\mathbf{K}$ is the kernel matrix: its component k_{ij} is defined as the dot product between the $\varphi(\mathbf{t}_i)$ and the $\varphi(\mathbf{t}_j)$ mappings. In our study, we use the gaussian kernel function. It introduces an additional parameter σ which is its standard deviation. Parameters σ and C are called the hyperparameters of the LS-SVM model. Their values can be optimized using a validation procedure. Hence, the expression of the model becomes:

$$\hat{z} = \sum_{i=1}^{N} \alpha_i \mathbf{K}(\mathbf{t}_i, \mathbf{t}) + b \tag{4}$$

where $\boldsymbol{\alpha}$ and b are the solutions of the system given by (3).

B. LS-SVM model selection procedure

Since LS-SVM models are linear in their parameters models, the solution of the training phase is unique and can be computed straightforwardly using the set of linear equations given by relation 3. This holds when the hyperparameters C and σ are known with fixed values. Usually, these hyperparameters are unknown and must be computed prior to the training phase. A suitable way to proceed consists in selecting the couple (C, σ) that best validates the LS-SVM model. In practice, the generalization capabilities of such black box model are estimated by computing the validation error. Various validation techniques exist in the literature [6]. The most popular techniques are probably the cross validation method and the Leave-One-Out (LOO) technique. Since a fine search is desirable to best optimize the model performance, the computational burden can rapidly become untractable when using either methods. In order to reduce substantially the computational time of the selection procedure without compromising its efficiency, we propose to estimate the validation error using the Virtual Leave-One-Out (VLOO) method. This method, first proposed for linear models [7] and later extended to nonlinear models [8], allows an estimation of the validation error to be computed by performing only one training involving the whole available data. This estimation is exact when dealing with linear-in-their-parameters models, such as LS-SVM models, while it remains an approximation for models which are nonlinear with respect to their parameters. More recently, a framework was described [9] to implement the VLOO method for LS-SVM models. For a given LS-SVM model, the VLOO error is computed as:

$$VLOO = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{\alpha_i}{(M^{-1})_{ii}}\right\}^2} \qquad (5)$$

where N is the size of the training set, and $(M^{-1})_{ii}$ is the i-th diagonal element of the inverse of matrix $\mathbf{M}$ that appears in relation 3. Thus, the VLOO permits a fast and exact estimation of the validation error which consists in a great benefit when optimizing the values of the hyperparameters according to a grid search.

IV. Optimal sensor architecture for one-bubble detection

Whereas the number of electrodes is a discrete value beginning from 2 and going typically to few units, their size is a continuous value. As a result, investigating the influence of electrode size can lead to a large number of cases to study while evaluating the influence of the number of electrodes is more straightforward. However, various numerical experiments showed that it is convenient to express the electrodes size as a ratio of the sensor perimeter they cover and that three types of electrodes size must be distinguished: pinpoint, contiguous and intermediate size electrodes.

- Pinpoint electrodes: they cover a very small ratio of the sensor perimeter, typically 1%
- Contiguous electrodes: they cover the major sensor perimeter, typically 99%
- Intermediate size electrodes: they cover any other ratio of the sensor perimeter. We considered three values: 25%, 50% and 75%

Basically, there are three models: each of them predicts a component of the triplet (x, y, R). The search of the optimal sensor architecture leads to a combinatorial optimization. We proceed according to a grid search by considering all the sensors with a number of electrodes in {2, 3, 4, 5, 6} and electrodes size in {1%, 25%, 50%, 75%, 99%}. To generate the data, we intend to detect air bubbles with relative permittivity $\varepsilon_2 = 1$ in an oil flow having a relative permittivity $\varepsilon_1 = 3$. For each given sensor, a set of data dedicated to training and validation is generated. It is formed by 500 examples. Each example is computed by considering a bubble with both random radius and coordinates. We implement a knowledge-based model that allows to compute the capacitances between the sensor electrodes given a spatial configuration. Since the knowledge-based model is a set of differential equations, its implementation is performed using a finite elements modeling approach with the Gmsh software

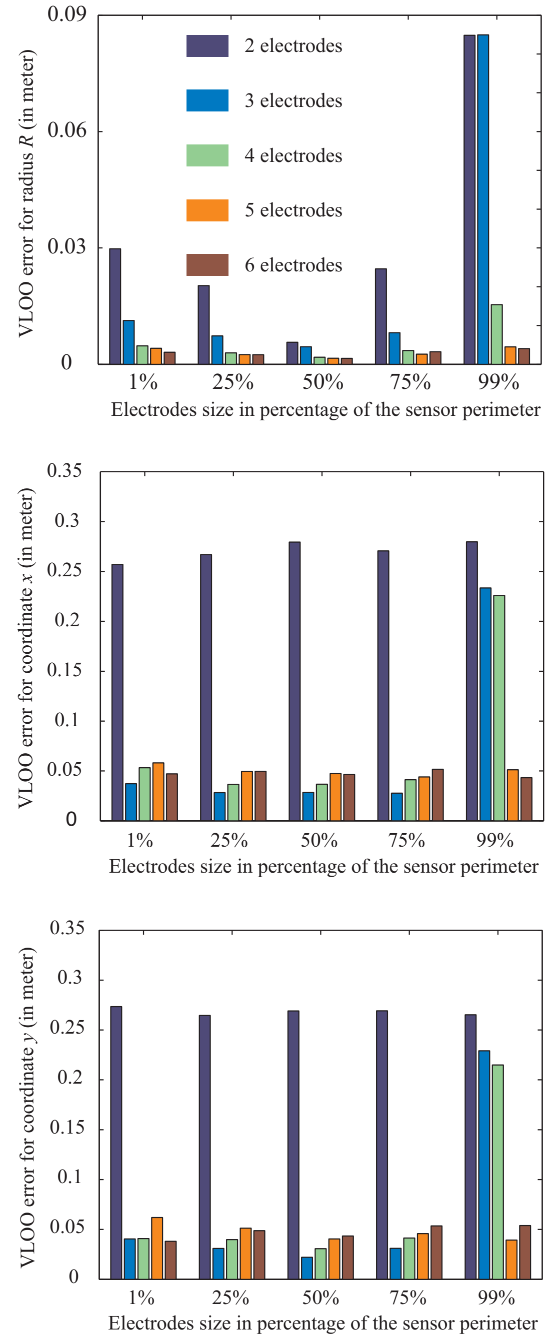

Figure 2. VLOO error for predicting radius R (top), coordinate x (middle) and coordinate y (bottom).

[10] as a mesh generator and the GetDP software [11] as a solver.

A first attempt to build inverse models to predict bubbles configuration with a linear LS-SVM leads to poor generalization capabilities. Therefore, the use of nonlinear models is necessary.

Figure 2 illustrates the main result of our study. It shows the validation error for the prediction of radius R (top graph), coordinate x (middle graph) and coordinate y (bottom graph). Obviously, a 2-electrode sensor is not satisfactory to predict the bubble configuration: it provides only 1 measurement while there are 3 unknown parameters. Beginning from 3 electrodes, all the sensors provide enough information to estimate accurately the bubble size and position with some differences depending on electrodes size.

Regarding the coordinates x and y, a deeper analysis of the results shows that the most accurate model is obtained with 3 electrodes of intermediate size 50%. The VLOO error computed using relation 5 is less than 3% of the sensor radius. Increasing the number of electrodes leads systematically to worse models. For intermediate size electrodes, the error increases almost monotonically. However, the behavior of sensors with contiguous electrodes is quite different. At least 5 electrodes are required to achieve a satisfactory prediction accuracy. A physical explanation can be found in the penetration of the electric field in the sensor. Contrarily to intermediate size electrodes, and with contiguous electrodes, the electric field is indeed concentrated at the junction of electrodes. Therefore, the impact of the bubble on measurements is limited.

Predicting the bubble size is clearly less critical than the bubble coordinates. A 2-electrode sensor of intermediate size 50% allows to predict radius R with an error less than 1% of the sensor radius. Similarly to the coordinates, the radius prediction with a contiguous electrodes sensor requires at least 5 electrodes.

V. Conclusion

This study was dedicated to investigate the influence of both the number and the size of an ECT sensor electrodes on its performances. Our goal was to determine the optimal sensor structure for a given application. We have implemented LS-SVM models validated by a sophisticated procedure to build nonlinear inverse models that predict the size and the position of a single bubble using independent measurements.

Regarding the number of electrodes, our results show that increasing this parameter do not lead systematically to an improvement of the sensor performances. The best accuracy is achieved with a sensor formed by 3 electrodes. This result is counterintuitive.

Albeit to a lesser extent, the electrodes size also has an impact on the sensor performances. The most accurate sensor we obtained uses intermediate size electrodes that occupy 50% of the sensor perimeter. The overall results show that the sensor performances are roughly stable with pinpoint or intermediate size electrodes (between 25% and 75% of the sensor perimeter). However, the use of contiguous electrodes necessitates the involvement of more electrodes to achieve satisfactory performances.

The main result of our study: the use of a 3-electrode sensor, associated to an image reconstruction via a direct calculation for a single bubble localization and sizing, leads to a computational cost effective sensing approach. This makes it highly attractive for real-time implementations.

References

[1] W. Q. Yang and L. Peng, "Image reconstruction algorithms for electrical capacitance tomography," *Measurement Science and Technology*, vol. 14, pp. R1–R13, 2003.

[2] X. Chen, H. Hu, F. Liu, and X. Xuang Gao, "Image reconstruction for an electrical capacitance tomography system based on a least-squares support vector machine and a self-adaptive particle swarm optimization algorithm," *Measurement Science and Technology*, vol. 22, pp. 1–9, 2011.

[3] H. X. Wang and L. F. Zhang, "Identification of two-phase flow regimes based on support vector machine and electrical capacitance tomography," *Measurement Science and Technology*, vol. 20, pp. 1–8, 2009.

[4] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer, New York, 1995.

[5] D. P. Bertsekas, *Nonlinear Programming (Second ed.)*. Cambridge, MA.: Athena Scientific., 1999.

[6] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. Springer, 2009.

[7] D. Belsley, E. Kuh, and R. Welsh, *Regression Diagnostics: Identifying Influential Data And Sources Of Collinearity*. John Wiley and Sons, New York, 1980.

[8] Laurent and R. Cook, "Leverage, local influence and curvature in nonlinear regression," *Biometrika*, vol. 80, 1993.

[9] G. C. Cawley and N. L. C. Talbot, "Preventing over-fitting during model selection via bayesian regularization of the hyper-parameters," *JMLR*, vol. 8, pp. 841–861, 2007.

[10] Available online: http://geuz.org/gmsh/.

[11] Available online: http://www.geuz.org/getdp/.