Automatic Sensor Reconfiguration based on Adaptive Relevance Vector Machine for Uncertainty Reduction in Tomography Imaging

Daniel Ospina Acero, Shah M. Chowdhury, Fernando L. Teixeira,
ElectroScience Laboratory
The Ohio State University, Columbus, Ohio, USA
{ospinaacero.1, chowdhury.79, teixeira.5}@osu.edu

Qussai M. Marashdeh
Tech4Imaging LLC, Columbus, Ohio, USA
marashdeh@tech4imaging.com

***Abstract*—We apply the Adaptive Relevance Vector Machine to automatically select the measurement set in a tomographic setting, from all the arrangements or combinations of the measuring elements, that yield the lowest level of uncertainty about the estimated results, while maintaining good image reconstruction. To illustrate the proposed method, we present simulation results derived from Electrical Capacitance Tomography.**

I. Introduction

Usually, tomography involves the solution of an inverse problem which is ill-posed and ill-conditioned [1], [2]. This translates into a severely underdetermined system of equations. In the recent years there have been numerous advances in specific instances of tomographic applications that enable the acquisition of more data from the same physical sensor setup, helping to reduce the underdeterminedness [3]. However, in many of these cases, extra measurements do not always translate into improvements in image reconstruction and parameter estimation performance, and they often lead to a prohibitive computational burden [4], [5]. In this work we present the application of the Adaptive Relevance Vector Machine (Adaptive RVM) into tomographic problems as a way to reduce the data set and automatically select the subset of measurements/data points that reduce the most the uncertainty levels about the estimated values. We present simulations results from Electrical Capacitance Tomography (ECT) [6] to illustrate the Adaptive RVM.

II. Mathematical Model

In tomography, if the region of interest (RoI) is discretized into M pixels/voxels and there are N measurements/data points, the inverse problem that is usually present can be stated as follows [7], [8]:

$$\mathbf{t} = F(\mathbf{w}) \ , \tag{1}$$

where $\mathbf{t}$ represents the N (known) boundary measurements, $\mathbf{w}$ represents the M (unknown) values of the physical quantity involved in the particular instance of the problem, and F represents the functional relationship that describes the physical phenomenon connecting the values in $\mathbf{t}$ with the ones in $\mathbf{w}$.

Although the relation in (1) is almost always nonlinear, the problem can be linearized by considering the Born approximation, where a sensitivity map is determined between small perturbations in the RoI and measurement results [9]. With this idea, the model of (1) becomes

$$\mathbf{t} \approx \mathbf{\Phi w} \ , \tag{2}$$

where $\mathbf{\Phi}$ represents the sensitivity matrix, whose entries indicate how much change in the quantities of $\mathbf{t}$ results from a very small change in the quantities of $\mathbf{w}$ for their corresponding pixels/voxels.

III. Adaptive Relevance Vector Machine

A. Relevance Vector Machine

The Relevance Vector Machine (RVM) method is a Bayesian algorithm introduced by [10] to solve a regression problem, like the one in (2), by learning a probabilistic model from the data and by using a sparse set of basis vectors defined by the model matrix. The main features of the RVM method are: **(i)** it allows to incorporate some level of prior information (via Bayes rule) about the problem; **(ii)** it provides measures of certainty around the estimated values; and **(iii)** due to its sparse nature, it can achieve very good estimates with very few basis functions (i.e., columns of the model matrix) [11], [12]. The model that RVM solves takes the form of

$$\mathbf{t} = \mathbf{\Phi w} + \boldsymbol{\epsilon} \ , \tag{3}$$

where $\boldsymbol{\epsilon}$ is the noise/error vector, whose entries are assumed to be independent and identically distributed (iid) with Gaussian distribution of 0 mean and variance σ^2 (i.e., $\epsilon_n \sim \mathcal{N}(0, \sigma^2)$).

The probabilistic view that is involved in (3) yields $\mathbf{w}|\mathbf{\Phi}, \mathbf{t}, \boldsymbol{\alpha}, \sigma^2 \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$, where $\boldsymbol{\alpha}$ represents the vector of M independent hyperparameters that help to express a preference for smoother models (avoiding overfitting), and

$$\boldsymbol{\mu} = \sigma^{-2}\mathbf{\Sigma}\mathbf{\Phi}^T\mathbf{t} \ , \quad \mathbf{\Sigma} = (\sigma^{-2}\mathbf{\Phi}^T\mathbf{\Phi} + \mathrm{diag}(\boldsymbol{\alpha}))^{-1} \ . \tag{4}$$

Under this setting, the estimate of $\mathbf{w}$ is simply given by $\boldsymbol{\mu}$. Additionally, $\mathbf{\Sigma}$ represents the level of uncertainty about the estimated values.

B. Adding data to the original model: Adaptive RVM

One of the main strengths of RVM is its ability to provide, through $\boldsymbol{\Sigma}$, a measure of "certainty" around the set of estimated values. By exploiting this fact, the authors of [13] proposed an adaptive method that allows to incorporate new information into the model, which in turn reduces the uncertainty level directly in each iteration of the algorithm. The approach of [13] is based on computing the (differential) entropy of $\mathbf{w}$ (which is a measure of its level of uncertainty, its variability) and the change that adding a new row to the matrix $\boldsymbol{\Phi}$ induces over it. This is as follows:

$$h_{\text{new}}(\mathbf{w}) = h(\mathbf{w}) - \frac{1}{2}\log\left(1 + \frac{1}{\sigma^2}\boldsymbol{\varphi}_{N+1}\boldsymbol{\Sigma}\boldsymbol{\varphi}_{N+1}^T\right) , \quad (5)$$

where $h(\mathbf{w})$ represents the entropy of $\mathbf{w}$ from the original N measurements, and $\boldsymbol{\varphi}_{N+1}$ represents the new row that is added to the matrix $\boldsymbol{\Phi}$ (part of a new measurement).

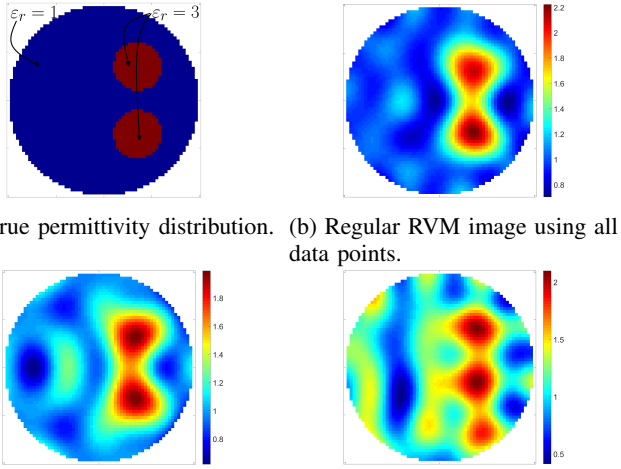
In order to minimize (5), the expression $\boldsymbol{\varphi}_{N+1}\boldsymbol{\Sigma}\boldsymbol{\varphi}_{N+1}^T$ must be maximized. For a fixed $\boldsymbol{\Sigma}$, this is achieved by its leading eigenvector (appropriately normalized). This solution, however, cannot be used in practical tomographic problems, as the leading eigenvector of the covariance matrix $\boldsymbol{\Sigma}$ does not have a direct connection with physical quantities. A more practical way of adaptively adding new data can be obtained by establishing two subprocedures: one *offline* (i.e., with empty RoI) and another one *online* (i.e., during the actual measurement). In the former, a dictionary $\mathcal{L}$ with all the desired sensitivity vectors is constructed. These sensitivity vectors correspond to different measurement configurations across all that are available in the physical setup.

The online subprocedure, on the other hand, is as follows: (*i*) select the sensitivity vector from $\mathcal{L}$ that maximizes $\boldsymbol{\varphi}_{N+1}\boldsymbol{\Sigma}\boldsymbol{\varphi}_{N+1}^T$ and append it at the end of $\boldsymbol{\Phi}$, (*ii*) obtain the new entry for $\mathbf{t}$ by performing the measurement dictated by the selected $\boldsymbol{\varphi}_{N+1}$, and (*iii*) perform a new reconstruction through RVM using the updated $\mathbf{t}$ and $\boldsymbol{\Phi}$. These steps must be repeated until a desired level of certainty has been reached.

The procedure defined above has two basic benefits. First, if more information of the same setup can be obtained, then the Adaptive RVM can reduce the level of uncertainty for the reconstructed images. Second, if no additional data can be obtained from the setup, then the Adaptive RVM can provide "good enough" results with fewer measurements and with comparable amount of confidence.

IV. Adaptive RVM in ECT

We illustrate the approach in a ECT process tomographic setting [6], [14] where the permittivity distribution in the RoI is to be determined from a set of capacitance measurements taken from boundary electrodes. To illustrate the Adaptive RVM method (in particular, the second benefit mentioned above), we show reconstruction results from an ECT simulated setup in Fig. 1. In these plots we can observe that the Adaptive RVM is capable of producing good results with fewer measurements as compared to simply adding measurements sequentially one by one .



(a) True permittivity distribution. (b) Regular RVM image using all data points.

(c) Adaptive RVM image with 54% of measurements (d) Regular RVM image by adding 54% of measurements sequentially.

Fig. 1: Reconstruction results from an ECT simulated setup.

Acknowledgment

This work was supported in part by NASA under grant NNX16CC10C and by the DOE under grants DE-SC0011936 and DE-SC0010228.

References

[1] R. A. Williams and M. S. Beck, *Process Tomography: Principles, Techniques, and Applications*, ser. Measurement Science and Technology. Butterworth-Heinemann, 1995.

[2] Q. Marashdeh, W. Warsito, L. Fan, and F. Teixeira, "A nonlinear image reconstruction technique for ect using a combined neural network approach," *Meas. Sci. Tech.*, vol. 17, no. 8, p. 2097, 2006.

[3] M. Wang, *Industrial Tomography: Systems and Applications*. Woodhead Publishing, 2015.

[4] M. Wang, Q. Wang, and B. Karki, "Arts of electrical impedance tomographic sensing," *Phil. Trans. R. Soc. A*, vol. 374, no. 2070, p. 20150329, 2016.

[5] L. W. Escalona Galvis, P. Diaz-Montiel, and S. Venkataraman, "Optimal electrode selection for electrical resistance tomography in carbon fiber reinforced polymer composites," *Materials*, vol. 10, no. 2, p. 125, 2017.

[6] Q. Marashdeh, F. Teixeira, and L.-S. Fan, "Electrical capacitance tomography," in *Industrial Tomography*, ser. Woodhead Publishing Series in Electronic and Optical Materials, M. Wang, Ed. Woodhead Publishing, 2015, pp. 3 – 21.

[7] R. C. Aster, B. Borchers, and C. H. Thurber, *Parameter Estimation and Inverse Problems*. Academic Press, 2011, vol. 90.

[8] M. Soleimani, "Computational aspects of low frequency electrical and electromagnetic tomography: A review study," *Int. J. Num. Analysis Model.*, vol. 5, no. 3, pp. 407–440, 2008.

[9] W. Yang and L. Peng, "Image reconstruction algorithms for electrical capacitance tomography," *Meas. Sci. Tech.*, vol. 14, no. 1, p. R1, 2002.

[10] M. E. Tipping, "Sparse bayesian learning and the relevance vector machine," *J. Machine Learn. Res.*, vol. 1, no. Jun., pp. 211–244, 2001.

[11] A. E. Fouda and F. L. Teixeira, "Bayesian compressive sensing for ultrawideband inverse scattering in random media," *Inv. Probl.*, vol. 30, no. 11, p. 114017, 2014.

[12] A. Fouda and F. Teixeira, "Ultra-wideband microwave imaging of breast cancer tumors via bayesian inverse scattering," *J. Appl. Phys.*, vol. 115, no. 6, p. 064701, 2014.

[13] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Proc.*, vol. 56, no. 6, pp. 2346–2356, Jun. 2008.

[14] Q. M. Marashdeh, F. L. Teixeira, and L.-S. Fan, "Adaptive electrical capacitance volume tomography," *IEEE Sensors J.*, vol. 14, no. 4, pp. 1253–1259, 2014.