World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:1, No:12, 2007

Detection of Ultrasonic Images in the Presence of a Random Number of Scatterers: A Statistical Learning Approach

J. P. Dubois, and O. M. Abdul-Latif

International Science Index, Computer and Information Engineering Vol:1, No:12, 2007 waset.org/Publication/3105

***Abstract*—**Support Vector Machine (SVM) is a statistical learning tool that was initially developed by Vapnik in 1979 and later developed to a more complex concept of structural risk minimization (SRM). SVM is playing an increasing role in applications to detection problems in various engineering problems, notably in statistical signal processing, pattern recognition, image analysis, and communication systems. In this paper, SVM was applied to the detection of medical ultrasound images in the presence of partially developed speckle noise. The simulation was done for single look and multi-look speckle models to give a complete overlook and insight to the new proposed model of the SVM-based detector. The structure of the SVM was derived and applied to clinical ultrasound images and its performance in terms of the mean square error (MSE) metric was calculated. We showed that the SVM-detected ultrasound images have a very low MSE and are of good quality. The quality of the processed speckled images improved for the multi-look model. Furthermore, the contrast of the SVM detected images was higher than that of the original non-noisy images, indicating that the SVM approach increased the distance between the pixel reflectivity levels (detection hypotheses) in the original images.

***Keywords*—**LS-SVM, Medical Ultrasound Imaging, Partially Developed Speckle, Multi-Look Model.

I. Introduction

RECENTLY, support vector machines (SVM's) have been introduced as a new method for solving classification and function estimation problems with many successful applications such as optical character recognition, object detection, face verification, and text categorization. In the medical imaging field, SVM has only been applied in the *post processing* phase for the classification of brain PET volumes [1], breast cancer diagnosis and prognosis based on binary classification (the tumor is malignant or benign) [2], and lesions segmentation in ultrasound images [3]. Our work is novel in the sense that we apply SVM in the *image acquisition* phase while considering advanced noise models such as the single look and multi-look *partially developed speckle*, which has been suggested and verified in [4]. Specifically, we apply least square-support vector machine (LS-SVM) in the detection stage to classify the received signal in order to construct the desired image with a relatively high precision.

Manuscript received July 15, 2005.
J. P. Dubois and O. Abdul-Latif are with the University of Balamand, Lebanon (phone: 961-3-841472; fax: 961-6-930250; e-mail: jeanpierre_dubois@ hotmail.com).

Biomedical ultrasound is an important modality for imaging body tissues in many physiological systems. Ultrasound is relatively inexpensive, safe, non-invasive, and portable. Despite these obvious advantages, ultrasound images suffers contrast and resolution degradation caused by contaminating speckle noise which is a form of biostructure or target-induced random noise [5].

II. Support Vector Machine

Traditional detection schemes generalize poorly on image detection due to the high dimensionality of the feature space. SVM generalizes well on difficult detection problems with a superior performance to traditional techniques.

In this section, we provide a succinct introduction to the SVM approach. The reader is referred to Vapnik [6] and Christianini [7] for in-depth treatment of the SVM theory.

As an intuitive approach to understanding SVM, consider a given set of points which belongs to either one of two classes. A linear SVM finds the hyperplane leaving the largest possible fraction of points of the same class on the same side, while maximizing the distance of either class from the hyperplane. This hyperplane minimizes the risk of misdetecting hypotheses of the test set. In SVM, the input vectors are nonlinearly mapped to a higher dimensional feature space. A linear discriminant function is then constructed in the new higher space, resulting in a non linear discriminant in the original input space.

We now present a brief theoretical approach to SVM. The relation between the capacity of a learning machine and its performance is ruled by a set of boundaries, which is referred to as the bound on the generalization performance. Statistical pattern recognition techniques face two problems: the identification problem and the parameters estimation problem. The identification problem determines the degree of freedom or complexity of the model and is generally the more complex problem [8]. The estimation problem obtains and optimal estimate of the model parameters regarding the training data set.

Let us consider a mapping $\Phi : \mathbb{R}^d \mapsto H$, which maps the training data from $\mathbb{R}^d$ to a higher Euclidean space H, that may have an infinite dimension. In this high dimension space, the data is linearly separable, hence linear SVM formulation above can be applied for any type of data [9]. In the SVM formulations, the training data only appear in the form of dot products $\boldsymbol{x}.\boldsymbol{x}$. These can be replaced by dot products in the Euclidean space H, i.e., $\varphi(.).\varphi(.)$.

The dot product in the high dimension space can also be replaced by a kernel function. By computing the dot product directly using a kernel function, one avoids the mapping $\Phi(x)$. This is desirable because H has possibly infinite dimensions and $\Phi(x)$ can be tricky or impossible to compute. Using a kernel function, a SVM that operates in infinite dimensional space can be constructed [7].

Given a training set of N data points $\{y_k, x_k)^N$, the SVM aims at constructing a decision function

$$f(x) = sign\left[\mathbf{w}^T \varphi(x) + b\right] = sign\left[\sum_{k=1}^{N} \alpha_k y_k K(x, x_k) + b\right], \quad (1)$$

where $\mathbf{w}$ is the weight vector in the reproducing kernel Hilbert space (RKHS), α_k are support values (Lagrangian multipliers), b is the bias term, and the kernel function

$$K(x,x_k) = \varphi(x)\varphi(x_k). \quad (2)$$

For every new test data, the kernel functions for each SV (support vector) need to be recomputed.

For any kernel function suitable for SVM, there must exist at least one pair of $\{H, \Phi\}$, such that (2) is satisfied. The kernel that has these properties is said to obey the Mercer's condition, i.e., for any $g(x)$ with finite L_2 norm,

$$\int g^2(x)dx < \infty, \quad (3)$$

$$\iint K(x,y)g(x)g(y)dxdy \geq 0 \quad (4)$$

By choosing different kernel functions, the SVM can emulate some well known classifiers [10], as shown in Table I.

TABLE I
KERNEL FUNCTIONS' CLASSIFIERS

Kernel Function	Type of Classifier
$K(x,y) = xy$	Linear
$K(x,y) = \exp\left(\lVert x-y\rVert_2^2 / \sigma^2\right)$	Gaussian radial bias function (RBF)
$K(x,y) = (xy + \tau)^d$	Polynomial of degree d
$K(x,y) = \tanh(\kappa xy + \theta)$	Multi layer perceptron

While standard SVM solutions involve solving quadratic or linear programming problems, the least square version of SVM (LS-SVM), which has been adopted for this research, corresponds to solving a set of linear equations. In LS-SVM, the Mercer's condition is still applicable. Hence several types of kernels can be used, yet the RBF is the adopted one since it gives a Gaussian distribution for the errors in the feature space yielding an optimal estimate of the support values [11]. Many reasons could be stated for preferring LS-SVM over other models of SVM, yet the most important one is that LS-SVM is an iterative method that could be used to solve large scale problems with robustness in the sense of the choice of the regularization and smoothing parameters. Moreover, it offers a fast method for obtaining classifiers with good generalization performance in many real life applications.

So far, the formulation of SVM was based on a two-class problem (SVM is essentially a binary classifier). Various schemes can be applied to the basic SVM algorithm to handle the M-class pattern classification problem. Some of these schemes [9], [12], for solving the multi-class problem are:

- Using M one-to-rest classifiers.
- Using $M(M$-1)/2 pair-wise classifiers with one of the voting schemes: Majority voting; Pairwise coupling.
- Extending the formulation of SVM to support the M class problem: By considering all classes at once; By considering each class with only the training data points belonging to that particular class.

The derivation of the multi-class LS-SVM is based on the Lagrangian optimization formulation

$$\min_{\mathbf{w},b,\xi} \Im_{LS}(\mathbf{w},\xi) = \frac{1}{2}\mathbf{w}^T \mathbf{w} + \gamma\sum_{k=1}^{N}\xi_k, \quad (5)$$

subject to the equality constraint

$$y_k\left[\mathbf{w}^T \varphi(x_k) + b\right] = 1 - \xi_k, \ \xi_k \geq 0, \ k = 1, \cdots, N \quad (6)$$

where γ is the regularization factor and ξ_k is the difference between the output y_k and discriminant function $f(x_k)$. Using standard techniques, the Lagrangian for (5) and (6) is

$$\mathrm{L}(\mathbf{w}, b, \xi; \alpha) = \Im_{LS}(\mathbf{w},\xi) - \sum_{k=1}^{N}\alpha_k\left[y_k\left(\mathbf{w}^T \varphi(x_k) + b\right) - 1 + \xi_k\right], \quad (7)$$

where α_k are the Lagrangian multipliers corresponding to (6). The saddle point is obtained from

$$\max_{\alpha} \min_{\mathbf{w},b,\xi} \mathrm{L}(\mathbf{w}, b, \xi; \alpha). \quad (8)$$

This yields the Karush-Kuhn-Tucker optimality conditions:

$$\begin{cases} \dfrac{\partial \mathrm{L}}{\partial \mathbf{w}} = 0 \rightarrow \mathbf{w} = \sum_{k=1}^{N}\alpha_k y_k \varphi(x_k) \\ \dfrac{\partial \mathrm{L}}{\partial b} = 0 \rightarrow \sum_{k=1}^{N}\alpha_k y_k = 0 \\ \dfrac{\partial \mathrm{L}}{\partial \xi_k} = 0 \rightarrow \alpha_k = 2\gamma\xi_k \\ \dfrac{\partial \mathrm{L}}{\partial \alpha_k} = 0 \rightarrow y_k\left[\mathbf{w}^T \varphi(x_k) + b\right] = 1 - \xi_k \end{cases} \quad (9)$$

III. PARTIALLY DEVELOPED SPECKLE NOISE MODEL

In the single-look model, a resolution cell is assumed to contain a collection of N elemental point scatterers randomly distributed throughout the resolution cell, with each elementary scatterer's position distributed independently of the positions of other scatterers. The random spatial distribution of the scatterers is described by a point process with a set of points having associated complex marks E_k, $(k = 1, \dots, N)$, corresponding to the backscattered ultrasound field from the k-th scatterer. Each backscattered ultrasound component E_k has a constant amplitude A_k equal to the reflectance strength of the k-th scatterer and a random phase φ_k, uniformly distributed over the interval $[0,2\pi)$: $E_k = A_k e^{j\varphi_k}$. We assume that the number of scattering points within a resolution cell is Poisson distributed with parameter Γ.

When ultrasound waves are scattered from such a surface, the resultant scattered field is the superposition of the ultrasound fields scattered by each of the elemental scatterers. The resulting process is a marked Poisson point process and $\lambda = \Gamma/A$ is the intensity (or rate) of this process over a resolution cell with area A. When $\lambda(.)$ is a random process, the point process is referred to as doubly stochastic marked Poisson point process [13].

Ultrasound systems record intensity measurements of the mapped surface according to:

$$S_N = \left| \sum_{k=1}^{N} A_k e^{j\phi_k} \right|^2 \qquad (10)$$

In (10), speckle is referred to as partially developed.

In the multi-look model, L-statistically independent diversity measurements are taken over the resolution region. This technique involves the noncoherent sum of L statistically independent single realizations of the mark measurements S_{Nl} $(l = 1,2,..., L)$:

$$T_{NL} = \sum_{i=1}^{L} S_{NL} \qquad (11)$$

Multi-look measurements can be obtained by changing the frequency of the ultrasound transducer (between 500 kHZ and 30 MHz). Such a transducer is placed onto the patient's skin over the imaged region and transmits a pulse which travels along a beam to the tissue. The reflected ultrasound energy is converted by the transducer into a single-look echo signal. The various echos resulting from different frequency transmissions can be sent to signal processing circuits in the imaging hardware to form a multi-look 2-dimensional image. Although multi-look models are not widely used in ultrasound commercial imaging devices, it would be of interest to us to view the multi-look estimation results, especially that many coherent imaging techniques rely on mathematically similar phenomena, and consequently, there could be significant snipoffs to applications in these systems.

IV. Simulation Results And Discussions

The simulations are done according to the LS-SVM model using Matlab code downloaded from [14]. Without loss of generality and for simulation purposes, we assumed in (10) and (11) unit amplitude scatterers ($A_k = 1$) and unit parameter for the Poisson process ($\Gamma = 1$).

The resulting images from the simulation are presented in Fig. 1. We observe that the SVM-reconstructed ultrasound images are of very good quality for $L = 8$ looks. The quality of the reconstructed speckled images improved as the number of looks increased. The performance metric used to determine the quality of the images is the mean square error (MSE). The MSE results are listed in Table I.

TABLE II
The SVMPerformance in Terms of the Number of Looks

No. of Looks (L)	MSE
1	3.72
2	2.30
4	1.15
8	0.65

Despite the fact that SVM is essentially a binary classifier, the approaches mentioned in section II to extend the SVM to multi-class detection worked well, as is evident from the results of the detected images, which are clearly not binary.

Furthermore, the contrast of the SVM detected images was higher than that of the original non-noisy images, indicating that the SVM approach increased the distance between the pixel reflectivity levels (the detection hypotheses) in the original images.

The reconstructed ultrasound images are observed to contain spiky noise. This is due to simulation artifacts, which are inevitable in sample training based algorithms. However, the effect of spiky noise is diminished with larger L, and the 8-look detected image is of very good quality.

The artifacts in lower-look images is also attributed to the rather simplified noise model assumed in section III. In reality, ultrasound speckle patterns are correlated with ACF

$$R_\Psi(\Delta x) = P_{dif}^2 \left[1 + 4\left(\mathrm{sinc}^2\,\Delta x \otimes \mathrm{sinc}^2\,\Delta x \right)^2 \right], \qquad (12)$$

where Δx is the spatial lag and P_{dif} is the mean diffuse power [5]. Any more advanced speckle model should thus incorporate into it speckle pattern interference.

V. Conclusion

In this paper, we applied LS-SVM to the detection of ultrasound images corrupted with partially developed speckle noise using single look and multi-look models. SVM proved to be a learning machine suitable for detection problems of non binary speckled images and produced very good results for 8-look images.

The results of SVM also showed an enhanced contrast. As perspective to this work, this phenomenon could be mapped to the performance of wireless communication systems, where SVM detection will increase the distance between the signals in the sense that it will make signals representing different levels almost perpendicular. This will lead, in theory, to higher signal-to-noise ratio (which is what multi-look processing effectively does). This issue will be subjected to more analysis in order to fully understand the underlying phenomena behind such behaviour.

As future work, we also propose to examine post-detected image processing filters which are suited to the speckle noise statistical characteristics.

REFERENCES

[1] M. Bonneville, J. Meunier, Y. Bengio, and J. Soucy, "*Support Vector Machines for Improving the Classification of Brain Pet Images*," Proc. SPIE Medical Imaging, 1998, pp. 264–273.

[2] O. Mangasarian, W. Street, and W. Wolberg, "*Breast Cancer Diagnosis and Prognosis via Linear Programming*," Operations Research, vol. 43, 1995, pp. 570–577.

[3] C. Kotropoulos and L. Pitas, "*Segmentation of Ultrasonic Images Using Support Vector Machines*," Pattern Recognition Letters, vol. 24, 2003, pp. 715-727.

[4] J. Daba and M. Bell, "*Statistics of the Scattering Cross-Section of a Small Number of Random Scatterers*," IEEE Transaction on Antennas and Propagation, 1994, pp. 773-783.

[5] J. Goodman, *Laser Speckle and Related Phenomena*. Berlin: Springer-Verlag, 1984.

[6] V. Vapnik, "*Estimation of Dependences Based on Empirical Data*," Nauka English Translation, Springer Verlag, 1982.

[7] N. Christianini and J. Taylor, "*Support Vector Machine and Other Kernel Learning Methods*". London: Cambridge University Press, 2003.

[8] J. Christopher and C. Burges, "*A Tutorial on Support Vector Machines for Pattern Recognition*", Kulwer Publishers, vol. 12, no. 2, 1998, pp. 121-167.

[9] J. Weston and C. Watkins, "*Support Vector Machines from Multi-Class Pattern Recognition*," University Of London, unpublished.

[10] T. Joachims, "*Support Vectors and Kernel Methods*", Cornell University, unpublished.

[11] J. Suykers and J. Vardewalle, "*Multi-Class Least Square-Support Vector Machine*," Universite Catholique de Louvain, Belgium, unpublished.

[12] C. Hsu and C. Lin, "*A Comparison of Methods for Multi-Class Support Vector Machines*," IEEE Trans. Neural Net., vol. 13, 2002, pp. 415-425.

[13] D. Snyder and M. Miller, *Random Point Processes in Time and Space*. New York: Springer-Verlag, 1991.

[14] K. Pelckmans, J. Suykens, T. Gestel, J. De Brabanter, L. Lukas, B. Hamers, B. De Moor, and J. Vandewalle, "*LS-SVMlab Toolbox User's Guide version 1.5*", Katholiede Univeristeit Leuven, Belgium, unpublished. Available http://www.esat.kuleuven.ac.be/sista/lssvmlab/

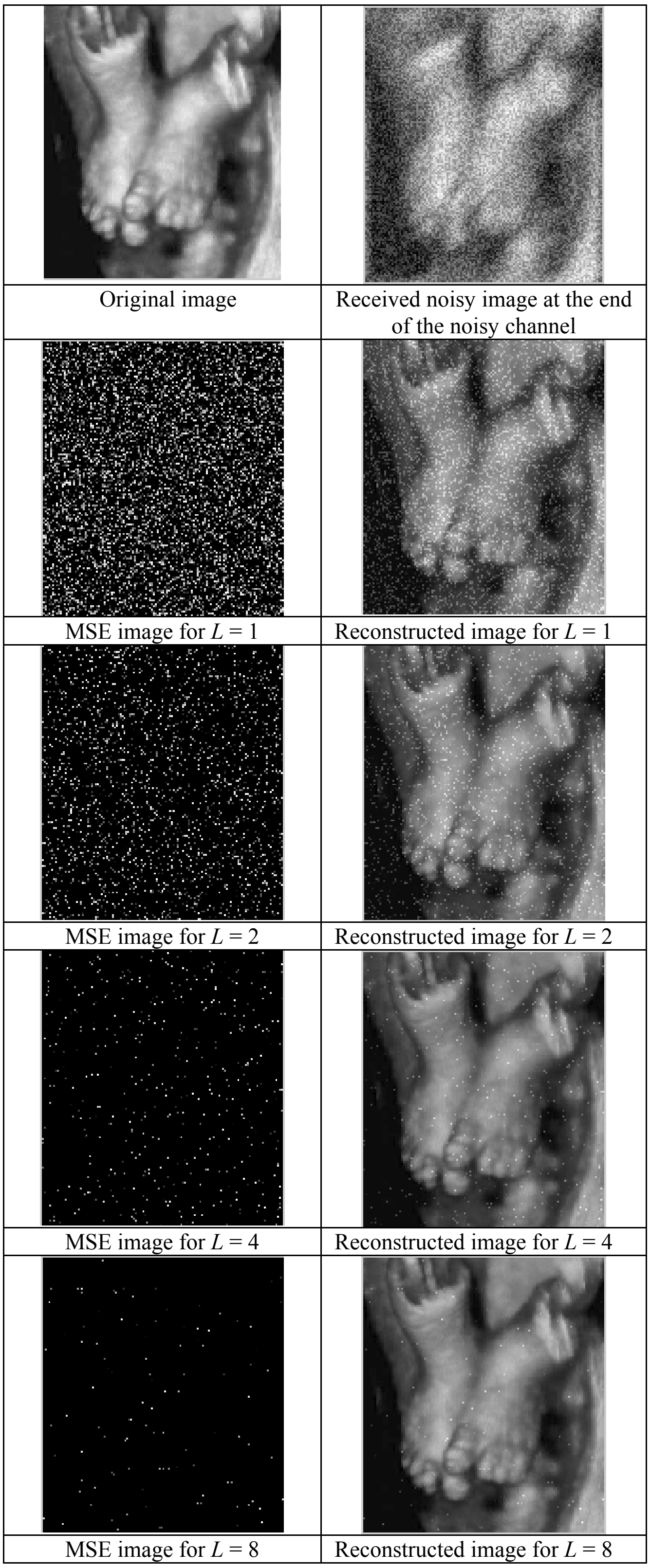

Fig. 1 LS-SVM detector applied to real phantom data