



Saudi Computer Society, King Saud University

Applied Computing and Informatics

(<http://computer.org.sa>)
www.ksu.edu.sa
www.sciencedirect.com



ORIGINAL ARTICLE

An inverse problem approach to pattern recognition in industry



Ali Sever *

University Of North Carolina at Charlotte, Charlotte, NC 28223, USA

Received 31 January 2013; revised 1 January 2014; accepted 27 February 2014
Available online 12 March 2014

KEYWORDS

Artificial neural networks;
Inverse problems;
Machine learning;
Pattern recognition;
Regularization

Abstract Many works have shown strong connections between learning and regularization techniques for ill-posed inverse problems. A careful analysis shows that a rigorous connection between learning and regularization for inverse problem is not straightforward. In this study, pattern recognition will be viewed as an ill-posed inverse problem and applications of methods from the theory of inverse problems to pattern recognition are studied. A new learning algorithm derived from a well-known regularization model is generated and applied to the task of reconstruction of an inhomogeneous object as pattern recognition. Particularly, it is demonstrated that pattern recognition can be reformulated in terms of inverse problems defined by a Riesz-type kernel. This reformulation can be employed to design a learning algorithm based on a numerical solution of a system of linear equations. Finally, numerical experiments have been carried out with synthetic experimental data considering a reasonable level of noise. Good recoveries have been achieved with this methodology, and the results of these simulations are compatible with the existing methods. The comparison results show that the Regularization-based learning algorithm (RBA) obtains a promising performance on the majority of the test problems. In prospects, this method can be used for the creation of automated systems for diagnostics, testing, and

* Tel.: +1 704 687 8580, +1 704 463 3118; fax: +1 704 687 8580.

E-mail address: ali.sever@uncc.edu

URL: <http://www.uncc.edu>

Peer review under responsibility of King Saud University.



Production and hosting by Elsevier

control in various fields of scientific and applied research, as well as in industry.
© 2014 Production and hosting by Elsevier B.V. on behalf of King Saud University.

1. Introduction

No patterns can be derived solely from empirical data (Yee and Haykin, 1993). Some hypotheses about patterns have to be chosen and, from among patterns satisfying these hypotheses, a pattern with a good fit to the data must be sought.

Neurocomputing brought a new terminology to data analysis: searching for parameters of their input/output functions is called learning, and samples of data training sets and a capability to satisfactorily process new data that have not been used for learning is called generalization.

The capability of generalization depends upon the choice of a hypothesis set of input/output functions, in which one searches for a pattern (a functional relationship) that matches the empirical data. So a restriction of the hypothesis set to only physically meaningful functions can improve generalization.

Inverse problems frequently arise in experimental situations when one is interested in the description of the internal structure of a system and is given indirect, noisy data. Estimating the response of a system given a complete specification of the internal structure, on the other hand, is the forward problem.

The modeling problem arises when one is given noisy data, observed over irregular intervals of space and time, and is asked to develop a reasonable model to fit those observed data (Vapnik, 1998).

With the advent of high-speed computers and artificial intelligence techniques, this modeling problem underwent a metamorphosis and emerged as a machine learning problem (Bauer et al., 2007; Gdawiec and Domanska, 2011). Tikhonov and Lanweber regularized that learning algorithms have recently received an increasing interest due to both theoretical and computational motivations (Abrukov et al., 2006; Kurkova, 2012; Tikhonov and Arsenin, 1977). Fractal, optimization, and a two-dimensional functional relational model have been used as a feature in several pattern recognition methods (Chang et al., 2010; Lo Gerfo et al., 2008; Nouredine, in press). Considerable attention is currently being devoted to new possibilities of using artificial neural networks (ANN) in view of their increasing importance for solving the problem of automated reconstruction of the inner structure of an object. Accompanying algorithms that effectively quantify uncertainties, deal with ill-posedness, and fully take the nonlinear model into account are needed. Therefore, it is necessary to both look for possible ways to improve the classical learning algorithms already existent in the literature, and to identify new methods which can compete with the traditional ones in speed, robustness, and quality of results.

Inverse problems are often formulated by assuming that the underlying phenomenon is a dynamic system characterized by mathematical equations, although no such assumption is always essential. Often the goal is to build an algorithmic model of the underlying phenomena. In some contexts a model is only a means to an end. The ultimate goal in such cases is to test the validity of a hypothesis. In these cases, the model is used as a classifier (e.g., neural nets and decision trees), and it matters little whether the model is parametric or non-parametric; the classification accuracy becomes more important. From this point of view the entire field of Machine Learning can be treated as an exercise in solving inverse problems (Bauer et al., 2007; Prato et al., 2007). By their very nature, inverse problems are difficult to solve. Sometimes they are ill-posed. A *well-posed* mathematical problem must satisfy the following requirements: existence, uniqueness and stability. The existence problem is really a non-issue in many realistic situations because the physical reality must be a solution. However, due to noisy and/or insufficient measurement data, an accurate solution may not exist. More often, the major difficulty is to find a unique solution; this especially when solving a parameter identification problem. Different combinations of parameter values (including boundaries and boundary conditions) may lead to similar observations. One useful strategy to handle the non-uniqueness issue is to utilize *a priori* information as additional constraints. These constraints generally involve the imposition of requirements such as smoothness on the unknown solution or its derivatives, or positivity, or maximum entropy or some other very general mathematical property. A more aggressive approach would be the use of regularization. Given an observed data set, genetic algorithms and genetic programming can be used to search a hypothesis space.

In this paper, starting from a reformulation of the pattern recognition as an inverse problem, we introduce an alternative learning algorithm derived by a well-known regularization method. We use a Riesz-type kernel to solve classification tasks by transforming the geometry of input space by embedding them into higher dimensional, inner product spaces, and introducing a regularization method which adds to the derived integral equation a new term, called stabilizer, which penalizes undesired input/output functions. We split the problem into a simpler, ill-posed problem (an integral equation with a Riesz-type kernel) and a well-posed problem. In this way, we isolate and better control the propagation of errors due to the ill-posedness (Nouredine, in press). Then we show that this reformulation can be employed to design a learning algorithm based upon a numerical solution of a system of linear equations.

The rest of the paper is organized as follows: The next Section describes our model and justifies its use. In Section 3, we formulate the proposed regularized learning algorithm. Section IV presents main simulation results. We compare our Regularization-based Algorithm (RBA) with the Support Vector Machine (SVM) and Semanteme-based Support Vector Machine (SSVM) in Section 5. Finally, we conclude the paper with a summary of the work in Section VI.

2. Generalization model as Regularization

Let us formulate the generalized problem as regularization in the following way:

Find a function $\sigma_1 \in L_\infty(\Omega)$, $\Omega \in \mathfrak{R}^n$, given the function $B(x_k) = w(x_k)$, $x_k \in \Omega_k$, $\Omega_k \in \mathfrak{R}^n$. Therefore, we have the following integral equation of the first kind

$$A\sigma_1(x) = B(x), \quad x \in \Omega_k \quad (2.1)$$

where $A\sigma_1(x) = \int_\Omega k(x, y)\sigma_1(y)dy$ and $k(x, y) = (1/2\pi)^2|x - y|^{-2}$ and A is considered as an operator from $L_\infty(\Omega)$ into $L_\infty(\Omega_k)$. This integral equation is the Fredholm integral equation of the first kind with a Riesz-type kernel.

First we need to show that Eq. (2.1) represents a severely ill-posed problem. Then we have to prove that a solution $\sigma_1(x)$ to the Eq. (2.2) exists and is unique.

Theorem 1. Let us assume that Ω and Ω_k are nonintersecting domains in \mathfrak{R}^3 . Then the integral Eq. (2.1) with the Riesz-type kernel represents an ill-posed problem.

Proof. We should notice that there are no singularities in the Riesz-type kernel with the domains defined above. We then claim that whether $\sigma_1 \in L_2(\Omega)$ is continuous or not, $(A\sigma_1)(x)$ is continuous in the usual sense. In fact,

$$\begin{aligned} |A\sigma_1(x_1) - A\sigma_1(x_2)| &= \left| \int_\Omega |x_1 - y|^{-2}\sigma_1(y)dy - \int_\Omega |x_2 - y|^{-2}\sigma_1(y)dy \right| \\ &\leq |\sigma_1(y)| \left| |x_1 - y|^{-2} - |x_2 - y|^{-2} \right| dy \\ &\leq \|\sigma_1\|^2 \cdot \int_\Omega \left| |x_1 - y|^\sigma - |x_2 - y|^\alpha \right|^2 dy \end{aligned}$$

Since the integrand $\left| |x_1 - y|^\sigma - |x_2 - y|^\alpha \right|^2$ is uniformly continuous, we have if $|x_1 - x_2| < \delta$,

$$\left| |x_1 - y|^\sigma - |x_2 - y|^\alpha \right|^2 < \varepsilon \text{ for } \forall y \in \Omega.$$

Therefore,

$$|A\sigma_1(x_1) - A\sigma_1(x_2)| \leq \|\sigma_1\|^2 \cdot \varepsilon \cdot |\Omega|,$$

$|\Omega|$ stands for a certain measure of Ω for any given $\varepsilon > 0$.

Now it is clear that if we take any $B(x) \in L_2(\Omega_k)$ which is continuous, then there is no $\sigma_1 \in L_2(\Omega)$ such that $A\sigma_1 = B$. So the existence requirement of the well-posedness is violated. Therefore the Eq. (2.1) is ill-posed.

For the integral Eq. (2.1), with a Riesz-type kernel and non-intersecting domains Ω and Ω_k , there is uniqueness in $L_2(\Omega)$ [Prato and Zanni, 2008](#). Djatlov's work shows a logarithmic type of stability estimate ([Kress, 1989](#)).

We use the Tikhonov regularization method ([Kress, 1989](#); [Sever, 1999](#)) to solve the ill-posed problem in Eq. (2.1). In this method, instead of Eq. (2.1), we solve the following regularized equation:

$$(A^*A + \alpha I)\sigma_1(x) = A^*B(x) \quad (2.2)$$

where I is the identity operator, and α is a regularization parameter. Now we will show the solution $\sigma_1(\alpha)$ and its convergence to the solution B when $\alpha \rightarrow 0$, provided σ_1 exists and the uniqueness for the original Eq. (2.1). Now we need to prove the existence of the solution $\sigma_1(\alpha)$ and its convergence to the solution f when $\alpha \rightarrow 0$, provided σ_1 exists and is unique.

Theorem 2. A solution $\sigma_1(\alpha)$ to the Eq. (2.2) exists and is unique. Also $R_\alpha y$ defined as $\sigma_1(\alpha)$ is a regularizer to the Eq. (2.1) on $X = X_M$, provided that the equation $A\sigma_1 = 0$ has only zero solution.

Proof. First we assume that A is compact, so is A^*A , but A^*A is also self-adjoint. Thus A^*A has a complete eigenfunction system denoted by $\{e_k\}$ with the corresponding eigenvalues $\{\lambda_k\}$.

Then, we have $\sigma_1(\alpha) = \sum_k \sigma_{1k}(\alpha) e_k$, $A^*y = \sum_k B_{Ak} e_k$ using equation, we get

$$(\lambda_k + \alpha)\sigma_{1k}(\alpha) = B_{Ak}$$

which implies $\sigma_{1k}(\alpha) = B_{Ak}/(\lambda_k + \alpha)$ is uniquely determined. Uniqueness can be obtained without using the expression of the solution. In fact, since A^*A is positive, we have

$$((A^*A + \alpha)x, x) = (Ax, Ax) + \alpha(x, x) > 0 \text{ for } \forall x \neq 0.$$

Therefore it cannot happen that there is some $\sigma_1^* \neq 0$ such that

$$(A^*A + \alpha)\sigma_1^* = 0$$

which means

$$\text{kernel}(A^*A + \alpha I) = \{0\}.$$

Now we show that $\sigma_1(\alpha)$ is actually a regularizer. To this end, we may assume now $A\sigma_1 = B$ where $\sigma_1 \in X_M$. Noticing that $A^*A + \alpha I$ is also self-adjoint, we have

$$((A^*A + \alpha)\sigma_1(\alpha), e_k) = (A^*B, e_k)(A^*A\sigma_1, e_k)$$

or

$$(\sigma_1(\alpha), ((A^*A + \alpha)e_k)) = (\sigma_1, A^*Ae_k)$$

$$((\lambda_k + \alpha)(\sigma_1(\alpha), e_k) = \lambda_k(\sigma_1, e_k)$$

with $\sigma_1(\alpha) = \sum_k \sigma_{1k}(\alpha) e_k$, $\sigma_1 = \sum_k \sigma_k e_k$, therefore, we get

$$\sigma_{1k}(\alpha) = (\lambda_k/\lambda_k + \alpha)\sigma_1$$

and

$$\sigma_{1k}(\alpha) - \sigma_{1k} = (\lambda_k/\lambda_k + \alpha)\sigma_{1k} - \sigma_{1k} = -(\alpha/\lambda_k + \alpha)\sigma_{1k}.$$

Therefore

$$\begin{aligned} \|\sigma_{1k}(\alpha) - \sigma_{1k}\|^2 &= \sum_k |\sigma_{1k}(\alpha) - \sigma_{1k}|^2 = \sum_k (\alpha^2/(\lambda_k + \alpha)^2) \sigma_{1k}^2(\lambda_k > 0) \\ &< \alpha \sum_{k \leq K} K K (\lambda_k + \alpha)^{-2} |\sigma_{1k}|^2 + \sum_{k > K} |\sigma_{1k}|^2 \end{aligned}$$

Here we assume that $\lambda_k \geq \lambda_{k+1} \geq \dots$

For all $\varepsilon > 0$ since $x \in L_2$, $\sum_k |\sigma_{1k}|^2$ converges. We first choose K such that $\sum_k |\sigma_{1k}|^2 < \varepsilon^2/2$.

Then, for the fixed K , we may chose δ such that

$$\alpha \sum_{k \leq K} (\lambda_k + \alpha)^{-2} |\sigma_{1k}|^2 < \varepsilon^2/2.$$

Consequently,

$$\|\sigma_{1k}(\alpha) - \sigma_{1k}\| \leq \varepsilon$$

It remains to prove that R_{α} , i.e., σ_1 is continuous. By observing that

$$\sigma_{1k}(\alpha) - \sigma_{1k}(\alpha_0) = ((\alpha_0 - \alpha)\lambda_k)/(\lambda_k + \alpha)(\lambda_k + \alpha_0)\sigma_{1k}$$

And using a similar argument to what we had above, we get the continuity of σ_{1k} .

Spectral representation of self-adjoint operators in Hilbert space gives the general case. Assumption $A\sigma_1 = 0$ if and only if $\sigma_1 = 0$ guarantees that the limit of $\sigma_1(\alpha)$ is unique.

3. Regularized learning algorithm

In this section we formulate a regularized learning algorithm based upon the Tikhonov regularization algorithm. For computational reasons, let Ω and Ω_k be the domain in \mathfrak{R}^2 and we will regard the integral operator

$$A\sigma_1(x) = \int_{\Omega} \sigma_1(y)/|x - y|^2 dy$$

as defined from $L_2(\Omega)$ into $L_2(\Omega_k)$. By using the definition of an adjoint operator in L_2 , we have $A^*: L_2(\Omega_k) \rightarrow L_2(\Omega)$ defined by

$$A^*B(y) = \int_{\Omega_k} B(x)/|x - y|^2 dx \quad y \in \Omega \quad (2.3)$$

and $A^*A\sigma_1(y)$ becomes

$$A^*A\sigma_1(y) = \int_{\Omega_k} \int_{\Omega} \sigma_1(y')|x - y|^2|x - y'|^2 dy' dx \quad (2.4)$$

where $x \in \Omega_k$, and $y, y' \in \Omega$. By discretizing Eq. (2.4), we have

$$\begin{aligned}
A^* A \sigma_1(y_i) &= \sum_{k,j=1}^n w_j w_k \sigma_1(y'_j) / |x_k - y_i|^2 |x_k - y'_j|^2, y_i, y'_j \in \Omega \\
&\cong \sum_{j=1}^n \Theta_{I,j} \sigma_1(y'_j)
\end{aligned} \tag{2.5}$$

where

$$\Theta_{I,j} \cong \sum_{k,j=1}^n w_j w_k / |x_k - y_i|^2 |x_k - y'_j|^2$$

and w_j , and w_k are the weight functions. By discretizing (2.3), we have

$$A^* B(y_i) \cong \sum_{k=1}^n w_k B(x_k) / |x_k - y_i|^2 \tag{2.6}$$

From (2.3)–(2.6) we have the discretized matrix equation in the form of

$$(\alpha I + \Theta) \sigma_1(y) = A^* B \tag{2.7}$$

for some regularization parameters, $\alpha > 0$. Now the problem is reduced to solving systems of linear equations.

4. Simulation results

In this section we want to investigate the effectiveness of the regularized learning algorithm introduced in Section 3.

Let us denote ‘real object’ by σ_t , and ‘computed object’ by σ_c . To test the method numerically, it is necessary to generate ‘ $B(x)$ ’ in the integral Eq. (2.1). We do this by specifying σ and evaluating the integral numerically. Once we have the numerical values of $B(x)$, we use these as our data and recover the pattern inside the required region. The steps of test calculation are:

- (1) specify σ_t ,
- (2) calculate the integral (2.1),
- (3) use (2.7) to find σ_c ,
- (4) compare σ_t with σ_c .

Our test calculation used

- (i) smooth surface (Fig. 1)
- (ii) $\sigma_t = \sigma_0 + x_1 \sigma_1 + x_2 \sigma_2$ (two objects)

where x_i ’s are characteristic functions of unknown objects, and $\sigma_0 = 1/2$ and $\sigma_t = 1$ (Fig. 3). In the case of smooth surface (i), the numerical calculations have shown that when σ_t is a smooth polynomial, the reconstruction is a very good approximation of σ_c (Fig. 2).

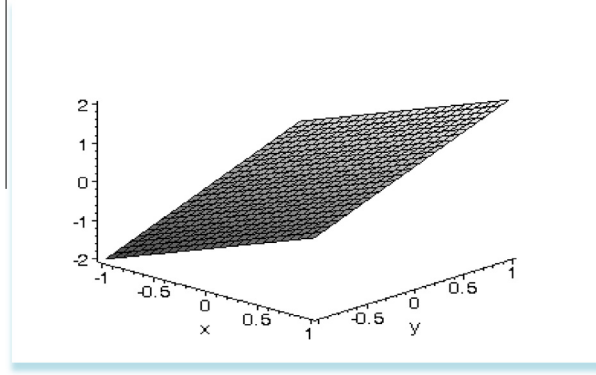


Figure 1 Smooth surface $\sigma_t = x + y$.

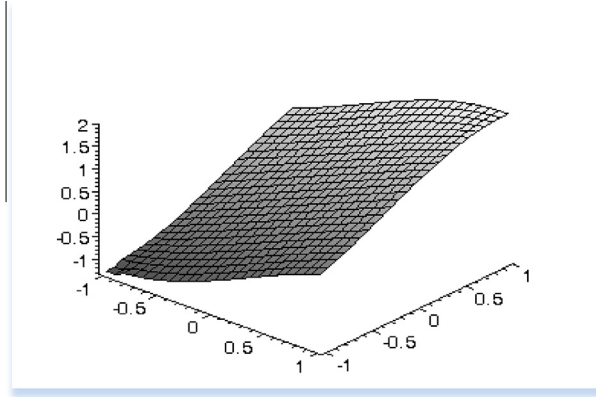


Figure 2 Reconstruction of surface in [Fig. 1](#).

In case (ii), we used two different regularization parameters for our reconstruction: simply, $\alpha = 10^{-7}$ and $\alpha = 10^{-10}$. We provided the cross sections of σ_t and σ_c . The proposed model was able to distinguish the objects. We observed that the computed σ_c was always smoothed. The location of the objects was well produced by σ_c , and also the shape of the σ_c was a fair indication of the objects ([Figs. 4 and 5](#) and [Table 1](#)).

Equation (2.7) involves main parameters that must be adjusted for greatest efficiency: the regularization parameter and the number of grid points. Looking at the reconstructions, the numerical experiments described above have shown that the reconstructed surface is smooth and close to the true surface. The reconstruction was usually a fair representation of the shape of the σ .

Summarizing, the simplicity and the reconstruction accuracy make the proposed regularized learning model well suited for the considered application.

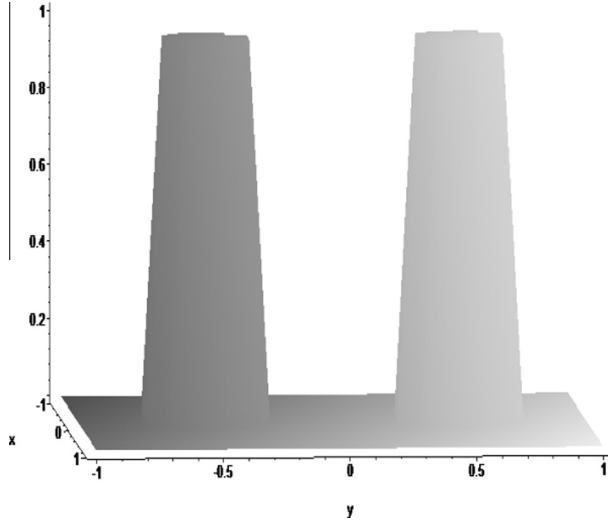


Figure 3 Test domain with two objects.

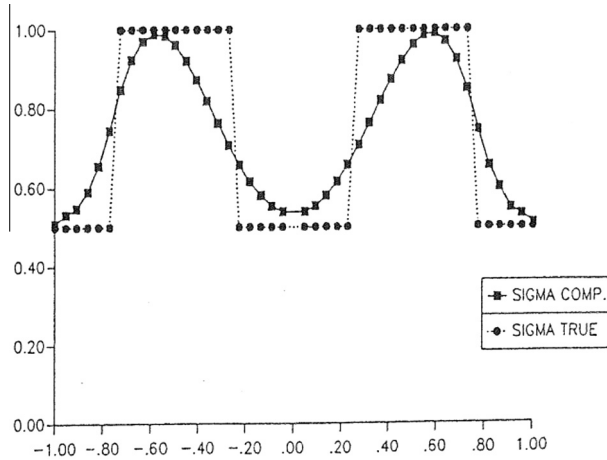


Figure 4 Cross-section from the reconstruction of $\sigma_t = \sigma_0 + x_1\sigma_1 + x_2\sigma_2$ where x_i 's are characteristic functions of unknown objects, and $\sigma_0 = 1/2$ and $\sigma_1 = 1$, and $\alpha = 10^{-7}$.

5. Results and discussion

In this section, misclassification rate (Li and Wang, 2009) is used to evaluate the efficiency of our algorithm. Misclassification rate refers to the ratio of the number of misclassified exemplars to the total number of exemplars in the dataset. The ratio is computed using the Formulation (5.3). Correspondingly, the classification accuracy is determined by the Formulation (5.2).

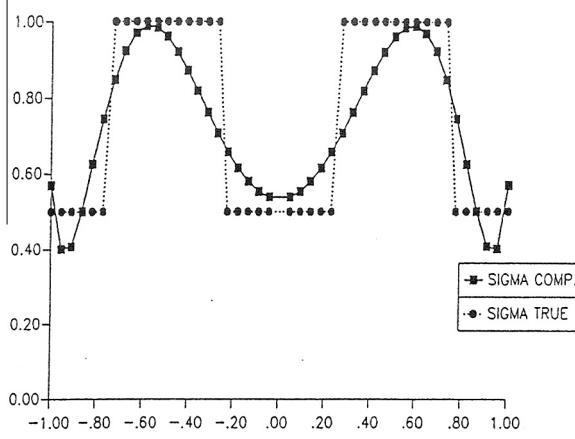


Figure 5 Cross-section from the reconstruction of $\sigma_t = \sigma_0 + x_1\sigma_1 + x_2\sigma_2$ where x_i 's are characteristic functions of unknown objects, and $\sigma_0 = 1/2$ and $\sigma_1 = 1$, and $\alpha = 10^{-10}$.

Table 1 Experimental datasets.

Cross-sectional points	$\sigma_0 = 1/2$ and $\sigma_1 = 1$				
	-1	-0.5	0	0.5	1
σ_t - True value	0.5	1	0.5	1	0.5
$\alpha = 10^{-7}$	0.507	0.982	0.53	0.982	0.507
$\alpha = 10^{-10}$	0.582	0.973	0.532	0.973	0.591
Absolute error for $\alpha = 10^{-7}$ (%)	1	2	6	2	1
Absolute error for $\alpha = 10^{-10}$ (%)	16	3	6	3	18

$$\gamma_{\text{error}} = \sum_{i=1, -1} \text{number}_i / n \quad (5.1)$$

$$\gamma_{\text{error}} = 1 - \gamma(\text{error}) \quad (5.2)$$

where number_i refers to the number of misclassified exemplars in the positive and negative classes, and n is the total number of X . A smaller γ_{error} value indicates higher classification accuracy and better classification efficiency. Conversely, a bigger γ_{error} value indicates worse classification efficiency.

Experiments are performed on the purely syntactic datasets from the UCI machine learning repository (UCI, 1998). We compare our Regularization-based Algorithm (RBA) with the Support Vector Machine (SVM) and Semanteme-based Support Vector Machine (SSVM) in the Table 2. In each test, all patterns with missing attribute values are initially removed. Continuous Dataset (CDS), Discontinuous Dataset I (DDS-I), and Discontinuous Dataset II (DDS-II) have their fixed real and computed pattern sets. The 9% outliers existing in DDS-II's training patterns are identified, and the dataset is classified as unbalanced. The continuous

Table 2 Misclassification ratio of different algorithms.

Dataset	SVM (%)	SSVM (%)	RBA (%)
CDS	2.9	3.1	5.32
DDS-1	16.3	9.1	4.95
DDS-2	15.5	9.42	9.45
Average	11.5	7.2	6.57

attributes in datasets are preprocessed using the Formulation (5.3) Li and Wang, 2009.

$$x_{ij} = (x_{ij} - \min_i \{x_{ij}\})(\max_i \{x_{ij}\} - \min_i \{x_{ij}\}) \quad (5.3)$$

A number of classification algorithms depend on the similarity or dissimilarity of exemplars, such as Euclidean distance and inner production, among others. However, majority of these algorithms only process continuous-attributed data, not discontinuous data. Discontinuous data (surface in our examples) are extreme, having disordered and unbalanced distribution.

For parameter selection and optimization, regularized learning algorithm faces the problem of selecting parameter α . In many algorithms, the standardization of the selection method for α is rarely performed (Han and Zhao, 2009). The heuristic method or some optimization algorithm is used to select α . In this study, the heuristic method is used in each experiment to standardize datasets. Table 1 shows the relationship between classification accuracy and the value α in the RBA algorithm.

6. Concluding remarks and future work

In this section, we briefly discuss several results obtained and issues related to the proposed RBA learning and recognition technique. Some of these issues may be viewed as merits while others as limitations leading to open research problems for the future.

The implementation of the proposed algorithm shows that the method is reasonably accurate for the reconstruction of two objects, using artificially generated data whose distributions are known. We have seen, both theoretically and experimentally, that pattern classification can be viewed as an ill-posed, inverse problem to which a method of regularization may be applied. As shown in Table 2, our proposed regularized learning algorithm has already shown promising performance in comparison with the state-of-the-art approaches, such as Support Vector Machines (SVMs), on benchmark datasets and real-life test problems.

The information obtained from a preliminary analysis is by no means exhaustive of the method discussed here and suggests several areas of additional investigation. We recognize the clear connection between regularization theory for inverse problems, and pattern recognition as learning, and this allow us to

introduce a new learning algorithm. On one front, improvements have to be done both on the algorithm (different regularizer properties must be investigated) and the applications (non-homogeneous 3-D object recognition). More detailed work is needed to improve the effectiveness of the numerics in general. In addition, the answer to exactly how sensitive the method is to moderate amounts of noise is an open question.

References

- Abrukov, V.S., et al., 2006. Artificial Neural Networks and Optical Diagnostics. In: Sixth International Conference on Intelligence Systems Design and Applications.
- Bauer, F., Preverzev, S., Rosasco, L., 2007. On regularization algorithms in learning theory. *J. Complexity* 35, 52–72.
- Chang, Y.F., Lee, J.C., Mohd Rijal, O., Syed Abu Bakar, S.A.R., 2010. Efficient online handwritten Chinese character recognition system using a two-dimensional functional relationship model. *Int. J. Appl. Math. Comput. Sci.* 20 (4), 727–738. <http://dx.doi.org/10.2478/v10006-010-0055-x>.
- Gdawiec, K., Domanska, D., 2011. Partitioned iterated function systems with division and a fractal dependence graph in recognition of 2D shapes. *Int. J. Appl. Math. Comput. Sci.* 21 (4), 757–767. <http://dx.doi.org/10.2478/v10006-011-0060-8>, ISSN (Print) 1641-876X.
- Han, Xian-Pei, Zhao, Jun, 2009. The creation of semantic metadata based on Wikipedia. *J. Chin. Inform. Process.* 23 (2), 108–114.
- Kress, R., 1989. Linear integral equation. In: *Applied Math. Sciences*. Springer-Verlag, Berlin, New York.
- Kurkova, V., 2012. Complexity estimates based on integral transforms induced by computational units. *Neural Netw.* 33, 160–167.
- Li, Zhi-Hua, Wang, Shi-Tong, 2009. Clustering with outliers-based anomalous intrusion detection[J]. *Syst. Eng. Electron.* 31 (5), 1227–1230.
- Lo Gerfo, L., Rosasco, L., Odone, F., De Vito, E., Verri, A., 2008. Spectral algorithms for supervised learning. *Neural Comput.* 20 (7), 1873–1897.
- Nouredine, S., in press. An optimization approach for the satisfiability problem, *Appl. Comput. Inform.* <http://dx.doi.org/10.1016/j.aci.2011.11.002>.
- Prato, M., Zanni, L., 2008. Inverse problems in machine learning: an application to brain activity interpretation. *J. Phys.* 135.
- Prato, M., Zanni, L., Zanghirati, G., 2007. On recent machine learning algorithms for brain activity interpretation. In: 23rd Annual Review of Progress in Applied Computational Electromagnetics, March 19–23, 2007, Verona, Italy, pp. 1939–1946.
- Sever, A., 1999. On uniqueness in the inverse conductivity problem. *Math. Methods Appl. Sci.* 22 (12), 953–966.
- Tiknonov, A.N., Arsenin, V., 1977. *Solutions of Ill-posed Problems Transl from Russian*. John Wiley Sons, New York, Toronto.
- UCI repository of machine learning database [EB/OL], 1998. Available from: <<http://www.ics.uci.edu/~mllearn/MLRepository.html>> .
- Vapnik, V.N., 1998. *Statistical Learning Theory*. John Wiley and Sons, New York.
- Yee, P., Haykin, S., 1993. Pattern classification as an ill-posed, inverse problem: a regularization approach. In: ICASSP'93 Proceedings of the 1993 IEEE International Conference on Acoustics, Speech, and Signal Processing: Plenary, Special, Audio, Underwater Acoustics, VLSI, Neural Networks, vol. I.