

Contribution Separators Wide Margin Processing of Remote Sensing Data

A.Khireddine¹ and J.p.Salvestrini²

Abstract—One of the main goals of remote sensing is extracting meaningful information of the image. We are interested in one of the treatments used: segmentation. Thus, we tried to introduce the techniques of learning SVMs (separators Wide Margin) in the field of remote sensing. The SVMs, are tools designed to solve the problems of binary classification supervised, and then generalized approaches (one against one, one against all) classification multi-classes. According to the severability of data, SVMs are distinguished also by two models linear and nonlinear. So we made a classification segmentation of the contents of a satellite image representing the western region of Oran by the separator wide margin non-linear approach and one against all this by using a Gaussian kernel.

Index Terms—Remote sensing, segmentation, separators wide margin.

I. INTRODUCTION

The separators Wide Margin (SVMs) and Support Vector Machines in English, are supervised classification methods developed during the 90s by V. Vapnik. The SVMs aim to separate in a space dimension n appropriate, a set of vector data belonging to different classes of linear separators (right, plan or hyper plane). Among the multitude of linear separators that could be used, SVMs seek the most optimal separator, with a maximum distance between classes. According to the severability of data, SVMs are distinguished by two models: the linear SVM (where the data is linearly separable) and non-linear (in case the data are not linearly separable). We have therefore introduced the techniques of learning SVMs to the segmentation classification contained a satellite.

II. SEPARATE THEM WIDE MARGIN

The SVMs have been designed to solve the problems of binary classification supervised, and then extended through the strategies and approaches, the multi-class classification

A. The binary separators Wide Margin

Where the data is linearly separable and as part of a binary \in classification, it is a set of m vector data which is associated labels t_i (-1, 1) representative of their class. The linear separator of this data set is defined by

$$h(x) = w \cdot x + b = 0$$

Where $w = (w_1, \dots, w_n)$ is the normal vector,

And $x = (x_1, \dots, x_n)$ a vector splitter b linearly and the threshold. The function $h(x)$ will represent the border separating the two classes, the classes' t_i another example x is defined as:

$$t_i = \text{sign}(+ w \cdot x + b)$$

For Maximizing the distance between classes through maximizing the margin, we must close distance to vehicles known vectors of support.

As the margin is inversely proportional to the standard w [2], The search for optimal linear Separator:

$$\begin{cases} \text{Min} \frac{1}{2} \|w\|^2 \\ \forall i: t_i \cdot h(x_i) \geq 1 \end{cases} \quad (1)$$

The system obtained (1) represents the expression of primal optimization problem in SVMs. In order to simplify the constraints we solve the problem by its dual and using the method of Lagrange, we get $L(w, b, \alpha)$:

$$\begin{cases} \text{Min} \frac{1}{2} \|w\|^2 - \sum_{i=1}^m \alpha_i \cdot [(w^t \cdot x_i + b) \cdot t_i - 1] \\ \alpha_i \geq 0 \quad i = 1 \dots m \end{cases} \quad (2)$$

As the α_i represent the contribution of a x_i element in the design of linear separator $h(x)$, where only the α_i corresponding to support vectors are not zero. The Lagrangian term obtained will aim to minimize $L(w, b, \alpha)$ over $a: w$ and b , and maximize $L(w, b, \alpha)$ by providing $a: \alpha$. So the research on extremum of $L(w, b, \alpha)$:

$$\frac{\delta L(w, b, \alpha)}{\delta b} = 0 \Rightarrow \sum_{i=1}^m \alpha_i \cdot t_i = 0 \quad (3)$$

By replacing (3) and (4) in $L(w, b, \alpha)$ gives the following function:

$$F(\alpha) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m t_i \cdot t_j \cdot \alpha_i \cdot \alpha_j \cdot x_i^t \cdot x_j \quad (4)$$

The equation (4) allows us to recreate the primal optimization problem expressed in (5) by its dual:

¹Electronic Department, faculty of Technology, university of Bejaia(06000), Algeria(email : abdelkrim_khireddine@hotmail.com).

²LMOPS Laboratory, university of Metz, 56013, France(email : J.P. Salvestrini @univ-metz.fr).

$$\frac{\delta L(w, b, \alpha)}{\delta w} = 0 \Rightarrow w^* = \sum_{i=1}^m \alpha_i \cdot t_i \cdot x_i$$

$$\begin{cases} \text{Max} \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m t_i \cdot t_j \cdot \alpha_i \cdot \alpha_j \cdot x_i^t \cdot x_j \\ \sum_{i=1}^m \alpha_i \cdot t_i = 0 \\ \alpha_i \geq 0 \quad i=1..m \end{cases} \quad (5)$$

Thus, the search for optimal returns linear separator has a quadratic programming problem where and α_i are reckonable b and w can be deducted.

The function associated decision becomes:

$$h(x) = \sum_{i=1}^m \alpha_i \cdot t_i \cdot x_i^t \cdot x + b \quad (6)$$

Where the data are not linearly separable, it is usually possible to find a linear separator performing an optimal projection of a space E to another area F larger and using functions (table.1):

TABLE.1

CORE :	generic form :	Parameter :
Laplacian	$K(x, y) = \exp(-\ x - y\)$	δ :Ecart type
Polynomial	$K(x, y) = (x \cdot y + 1)^p$	p :polynôme Order
Gaussian base radial		δ :Ecart type

This allows us to translate the optimization problem expressed (4) by:

$$\begin{cases} \text{Max} \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m t_i \cdot t_j \cdot \alpha_i \cdot \alpha_j \cdot x_i^t \cdot x_j \\ \sum_{i=1}^m \alpha_i \cdot t_i = 0 \\ 0 \leq \alpha_i \leq C \quad i=1..m \end{cases} \quad (7)$$

This transformation can be done implicitly by the kernel $K(x, y) = \Phi(x) \cdot t \cdot \Phi(y)$ including the following:

$$\forall i : t_i \cdot h(x_i) \geq 1 - \xi_i \quad (8)$$

The optimization problem becomes:

$$\text{Min} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i \quad (9)$$

Thus, the dual expression remains the same (for SVMs linear and nonlinear), the only difference is that all Lagrange multipliers α_i . The dual system optimization in the case of linear SVMs becomes:

$$\begin{cases} \text{Max} \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m t_i \cdot t_j \cdot \alpha_i \cdot \alpha_j \cdot K(x_i, x_j) \\ \sum_{i=1}^m \alpha_i \cdot t_i = 0 \\ 0 \leq \alpha_i \leq C \quad i=1..m \end{cases} \quad (10)$$

The dual system optimization in the case of non-linear SVMs becomes:

$$\begin{cases} \text{Max} \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m t_i \cdot t_j \cdot \alpha_i \cdot \alpha_j \cdot \phi(x_i)^t \cdot \phi(x_j) \\ \sum_{i=1}^m \alpha_i \cdot t_i = 0 \\ \alpha_i \geq 0 \quad i=1..m \end{cases} \quad (11)$$

III. SVMs MULTI-CLASSES

The application of SVMs to a classification with class k , with $k > 2$ requires a generalization of SVMs bi-level, an increase of binary classifiers. Among the different approaches that can be used for this purpose there is the approach one against one.

A. The approach one against One

This is to design every possible binary classifiers, and, for k classes will $k \cdot (k-1) / 2$ classifiers. To assign an element to a class e , e must be tested with all the classifiers designed, whenever e is assigned to a class, we increment a counter i associated with it (i is initially set to zero). e will be awarded to the class that has a meter at maximum value.

B. A approach against all approaches

In this approach and a classification with k classes, each class i is opposed to $k-1$ and other classes will develop k binary classifiers. To assign an element to a class e , e must be tested against all classifiers designed, then assigned to the class that presents.

IV. APPLICATION OF THE SEGMENTATION SVMs CONTENT OF A SATELLITE IMAGE

We have applied SVMs non-linear approach to one against all the segmentation of the content of satellite images type Landsat5 TM (Thematic Maps) from the region west Bejaia city dated 15 March 1993 at 9am 45mn. This study area was chosen for its varied landscape may be relevant to the application of SVMs to problems of classification. We started our application by loading three images corresponding to the three channels TM1 TM2 and TM3. To ease the use of these images, we first made a contrast enhancement and a composition combining colored blue filter channel TM1, the filter TM3 green channel and red channel TM4.(figure.1):

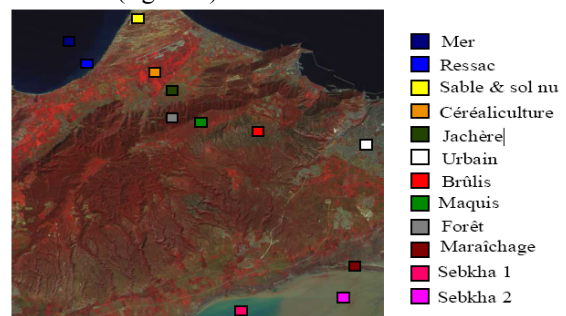


Figure 1: Identification of classes

Once the basic learning built we conducted our segmentation with the following entries: we set the compromise C to 500 and used a Gaussian kernel as well, we achieved the following results (figure.2):

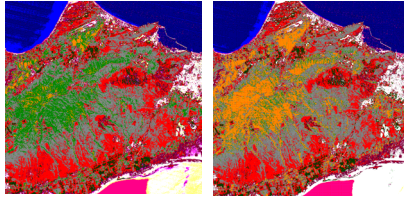


Figure 2: Images of the resulting test 1 and 3 Gaussian kernel (from left to right)

The first test with a standard deviation of a Gaussian 0.1 gave us the best recognition rate (TR) but also a large number of confusion mainly located at the classroom level Sebkh2, urban, sand and bare ground. The increase in the standard deviation gaussian level test 2 and test 3, has resulted in the decrease of recognition and the disappearance of classes Sebkh2, sand and bare ground. The Tests Settings percent is shown as:

Tests Settings	TR (%)
1 $\delta = 0.1$	74.51
2 $\Delta = 0.5$	64.98
3 $\delta = 1$	62.82

V. SIMULATIONS AND RESULTS

To ensure the recovery of the flow of data without learning sequence equalizer based on maximum likelihood is running a two-step procedure to identify the channel by the algorithm that we described earlier, and then we search the data using the results of this identification.

To assess the performance of the equalizer, we begin by presenting the accuracy of the identification algorithm for different SNR

A. Identification of Images of Gaussian kernel:

The identification of the Images of the resulting test 1 and 3 Gaussian kernel (figure.2) is used to detect information symbols.

We worked with a comment that contains ten information symbols ($N = 10$). The estimated coefficients are represented according to the index iteration, for reports RSB equal to 0, 5 and 10 dB.

Given the symmetry only the estimated values of half + 1 coefficients are presented.

The curves represent average estimated values of impulse response samples conducted on independent simulations. The coefficients of the initial impulse response are:

$$h((n-1)/2) = 1, h(i) = 0, i \neq (n-1)/2,$$

Where n is the number of coefficients.

After the tenth iteration, we note that RSB exact estimated coefficients are:

0 dB	[0407, 0.815, 0.407]	[0.5 0.65 0.5]
5 dB	[0407, 0.815, .407]	[0.25 0.58 0.25]
10 dB	[0407, 0.815, 0407]	[0.3 0.7 0.3]

For RSB = 10dB estimating coefficients, after the tenth iteration, we note that the identification algorithm converges to the RSB Exact estimated coefficients:

0 dB	[0,227 0.46 0,688 0,460 0,227]	[0.35 0.45 0.52 0.45 0.35]
5 dB	[0,227 0.46 0,688 0,460 0,227]	[0.1 0.34 0.5 0.34 0.1]
10 dB	[0,227 0.46 0,688 0,460 0,227]	[0.2 0.4 0.52 0.4 0.2]

A good estimate is obtained with a SNR = 10 dB.

We see that the noise influence on the estimated coefficients.

VI. CONCLUSION

The main difficulty relating to the use of separators wide margin lies in the choice of kernel. The latter may significantly influence the results of which type of kernel must be regarded as a parameter to take into account however, this method has the advantage of a clear mathematical formulation and elegant After applying separators Broad-margin non-linear approach to one against all the segmentation of satellite images, we noticed that the results are generally above 70%.

REFERENCES

- [1] A.Khireddine, K. Benmahammed, « Application of intelligent matching to automatic speech recognition »-ISH99 (international symposium on high voltage), IEE conference publication, vol. 4, pp.270-P6, code ISBN.0852967195, London UK, 1999.
- [2] R. Chellapa, A. Sawchuk, « Digital image processing and analysis ». Vol 1, IEEE Computer Society, 1985.
- [3] A.Khireddine, K. Madi, « Geometrical Correction of satellite images by analytical aproach » IMPAC'00 conference, vol.1, pp65-72., code ISSN 1112-3052, Boumerdes university, Algeria, 2000.
- [4] Zeng,B.,Neuvo,Y., Optimal Parallel Stack Filtering under the Mean Absolute Error Criterion,IP(3),No.3,May1994,pp.324-327.
- [5] Links,J.M.,Prince,J.L.,Gupta,S.N., A vector Wiener filter for dual-radionuclide imaging, MedImg(15),No.5,October1996,pp.700-709.
- [6] Concetti,A.,Jetto,L., Two-Dimensional Recursive Filtering Algorithm with Edge-Preserving Properties and Reduced Numerical Complexity, CirSysSignal(44), No. 7, July 1997, pp. 587-591.
- [7] Wang,X., Two-Dimensional Bayesian Estimator for Image Filtering, IP(8), No.7,July1999,pp.993-996.
- [8] Towghi, N.[Nasser], Javidi, B.[Bahram], LP-Norm optimum filters for image recognition. Part I. Algorithms, JOSAA(16), No. 8, August 1999, pp. 1928-1935.
- [9] Ghariieb,R.R.,Cumulant-based LP method for two-dimensional spectral estimation, VIP(146), No. 6, December 1999, pp. 307-315.
- [10] Knudsen,K.S.,Bruton,L.T., Mixed domain filtering of multidimensional signals, CirSysVideo(1), No. 3, September 1999, pp. 260-268.
- [11] Ghariieb,R.R., Cumulant-based LP method for two-dimensional spectral estimation, VISP(146), No. 6, December 1999, pp. 307.
- [12] Pokric, B., Allinson, N.M., Bergström, E.T., Goodall, D.M., Combining linear filtering and radial basis function networks for accurate profile recovery, VISP(146), No. 6, December 1999, pp. 297.
- [13] Kumar, B.V.K.V., Mahalanobis, A., Takessian, A., Optimal Tradeoff Circular Harmonic Function Correlation Filter Methods Providing Controlled In-Plane Rotation Response,IP(9), No. 6, June 2000, pp. 1025-1034.