TEMPORAL SIGNATURE MATCHING FOR LAND COVER CLASSIFICATION

Sudhir Gupta [a] *, K. S. Rajan [b]

[a] The LNM Institute of Information Technology, Jaipur, India -sudhir@lnmiit.ac.in
[b] International Institute of Information Technology, Hyderabad, India -rajan@iiit.ac.in

Commission VIII

KEY WORDS: EVI, time series, land cover, classification

ABSTRACT:

Objects on the surface of earth describe unique patterns in time and this is exploited in the proposed land cover classification method enumerated in this paper. The applicability of various distance measures to reasonably find similarity or dissimilarity between time series vegetation index patterns for land cover classification is demonstrated. These distance metrics have inherent advantages for satellite image classification as land use practices such as field crops/vegetation exhibit temporal shifts from pixel to pixel based on the geographic location. Experimental results on time series of MODIS EVI (250m) and AWiFS derived NDVI (56m) demonstrate the applicability and validate the performance of the proposed method. The classification method is also compared with k nearest neighbor (kNN) and support vector machine (SVM) classifier to understand its utility.

1. INTRODUCTION

Land cover classification is an important remote sensing technique and the remote sensing community has been challenged to produce land cover data sets at various scales on a repetitive basis (DeFries and Belward, 2000). The range of clientele that use land cover datasets is expanding every day. Spectral and contextual supervised classification schemes with manual delineation of training samples remain the preferred choice of remote sensing scientists (Wilkinson, 2005). The most commonly used classification methodologies for satellite image classification are decision tree, ISODATA, k-means and maximum likelihood classification.

The multi-temporal data that is now available because of better revisit times of the satellites allows us to characterize objects based on their dynamic processes rather than static properties like color, shape, etc and this is being successfully demonstrated in this paper for a few classes. An important requirement for such a study is the definition or use of a variable whose time series can describe the process on the basis of which we can arrive at the classification. We have used time series of vegetation index to classify forest, crop and water. Although the value of satellite time series data for classification has been firmly established, only a limited number of methods for exploring such data series have been developed (Malingreau, 1986).

Till date, features derived from temporal and spectral response patterns of spatial locations have been used with conventional classification techniques for classifying satellite imagery (Jonnson et.al, 2002, 2004) but the time series itself has not been used as an inherent class identifier. We have viewed the classification of the pixel-time trajectories as a time series signature (curve) matching problem comparing the unknown vegetation index time series of a pixel to the temporal signatures available in the library derived from the training samples. An important part of such a curve matching problem is the definition of similarity measure. The conventional classification techniques typically measure closeness in Euclidean space. This leads to an anomaly that two similar classes of vegetation having an identical vegetation index profile with different growing practices (sowing, senescence, harvest, etc) appear as points separated by a large distance leading to different class labels. This paper demonstrates the application of a novel similarity measure to compare temporal signatures of various land cover classes.

2. DATA AND STUDY AREA

The MODIS 250-m Vegetation Index (VI) product (MOD13Q1), EVI data composited at 16-day intervals is one of the datasets used in this study. Time series of EVI images for cropping seasons 2003-04, 2004-05 and 2005-06 were constructed for Dharwad District of Karnataka state, India This district has an area of 4263 square kilometers. Each cropping year/time series has 23 points. A good classification algorithm has to demonstrate sensor independence. To prove this, the classification procedure was also applied on NDVI derived from AWiFS data which has a spatial resolution of 56 meter and revisit time of around 5 days. 24 images were available from December 2007 to April 2008 that were used to form NDVI time series. AWiFS data was available for Atmakur taluk of Kurnool district in Andhra Pradesh state, India. National land use and land cover dataset of 2005-06 produced by National Remote Sensing Centre (NRSC), Indian Space Research Organisation, Hyderabad, India with 16 classes and spatial resolution of 56 meters was used as a reference for evaluating classification results.

3. PROPOSED METHOD

Figure 1 shows the block diagram of the proposed classification procedure. The basic assumption on which the classification procedure is based is that the EVI of a target monitored over a growth period forms a temporal signature of that target which

* Corresponding author.

can be used to classify the target. Instances of such temporal signatures along with the corresponding labels are stored in a library which can act as templates for labelling unlabelled EVI time series. This is similar to classification of multispectral or hyperspectral data in which a spectral library comprising of unique spectral response curves is stored and classification of unlabelled samples is done based on the results of the matching between the unlabelled spectral curves and labelled spectral curves present in the spectral library. However, the same approach cannot be used to classify temporal signatures. One of the problems with temporal signature matching is that there can be variability in the growth pattern of similar vegetation based on the geographical area which conventional distance measures like Euclidean distance capture as belonging to different classes though the patterns belong to the same class.

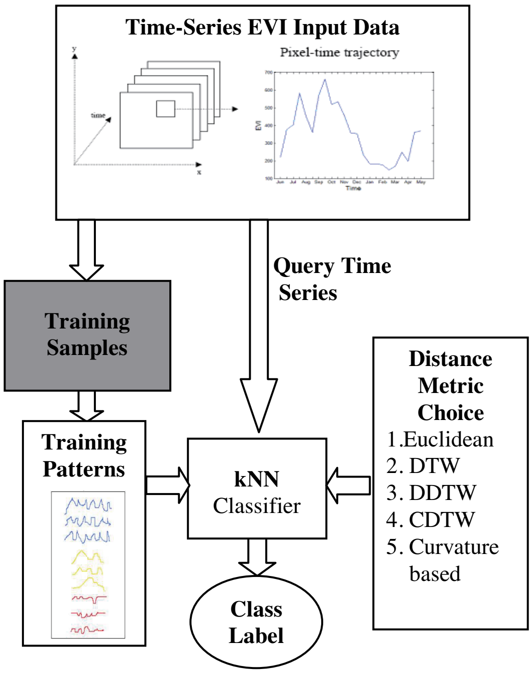

Figure 1. Proposed Time Series Classification Procedure

In other words, the conventional distance measures do not take temporal shifts and small variations in ordinate (EVI values in this case) into account. We have tried various distance metrics as discussed in section 3.2 and have proposed a distance metric that can be used for matching temporal signatures that can overcome misclassifications due to temporal shifts and small variations in EVI values.

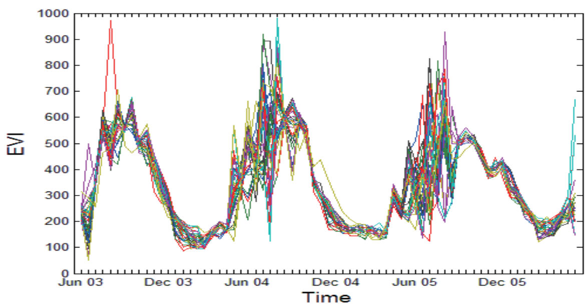

Figure 2. Crop EVI time series patterns

Figure 2 shows temporal signatures for fifty crop pixels. If all the fifty crops exhibited similar temporal responses, then they would have lined up on the time axis but this is not the case. Some variation in the responses can be seen that is not captured as intra class variation by conventional distance metrics like Euclidean distance because these distance measures compare the value of input pattern at time 't' to that of the template pattern at time 't'. The first step in the proposed classification procedure is to stack the vegetation index images to construct a time series. Knowledge of location of some representative samples for each class is used to construct a temporal library that consists of EVI patterns and the corresponding class labels. These are the training patterns or reference patterns. Thus the training phase of the proposed classification scheme consists of creating a temporal library unlike the conventional training phases that include determining certain classifier parameters. The classification of unlabelled EVI time series is a two step process. First, the distance between the unlabelled EVI pattern and all the reference patterns in the temporal library is found using a specified distance metric. In the second step, these distances are sorted in ascending order and 'k' most similar patterns are found. The label that occurs more frequently in these 'k' patterns is the label provided to the unlabelled EVI pattern and thus the name k nearest neighbour (kNN) classifier.

3.1 Formulation

The proposed classification scheme can be formulated as follows

$\{(y_1,\theta_1),(y_2,\theta_2),(y_3,\theta_3),..............,(y_n,\theta_n)\}$ be the training dataset

$\theta_i \in \{A,B,C,........\} \qquad \forall i$

A,B,C,........ are the class labels

Given an unclassified feature vector x, find its distances to the feature vectors y in the temporal library

Let $D_n = \{d(x,y_1),d(x,y_2),d(x,y_3),.......d(x,y_n)\}$ be these distances

Assuming, indices of the label feature vectors in D_n are permuted to satisfy

$$d(x,y_1) \le d(x,y_2) \le .... \le d(x,y_k)$$

$$d(x,y_j) \ge d(x,y_k) \qquad \text{for } j = k+1,.........,n$$

The k nearest neighbours are

$$\{Y_1,Y_2,Y_3,........Y_k\} \text{ where } Y_i = (y_i,\theta_i)$$

k nearest neighbor classification rule then becomes

$$\theta(x) = \max(n_A, n_B, n_C, ...)$$

Where

$n_A = |\{i|\theta_i = A\}|$; $n_B = |\{i|\theta_i = B\}|$; $n_C = |\{i|\theta_i = C\}|$ and so on

d is the distance metric or similarity measure. Experiments were performed with the various distance metrics described in the following section.

3.2 Similarity Measures

We have experimented various similarity measures with the proposed classifier. These are described here.

3.2.1 Euclidean distance

The simplest time-domain algorithm for computing a similarity metric between time series is the Euclidean distance between two discrete time series ***x[n]*** and ***y[n]*** where the distance between the two series is defined as:

$$D(x,y)=\sqrt{\sum_{n=0}^{M}(x[n]-y[n])^2}$$

where M is the length of the time series.

3.2.2 Dynamic Time Warping (DTW)

Suppose we have two time series X and Y of length n and m respectively, where

$$X = x_1, x_2, x_3, ..., x_n \text{ and } Y = y_1, y_2, y_3, ..., y_m$$

To align two sequences using DTW, an n-by-m matrix is constructed whose element at (i,j) position contains the distance $d(x_i, y_j)$ between the two points x_i and y_j (Typically the Euclidean distance is used). This is illustrated in Figure 3. A warping path W, is a contiguous set of matrix elements that defines a mapping between X and Y. The k^{th} element of W is defined as $w_k = (i, j)_k$, so we have: $W = w_1, w_2, w_{3,.......}, w_K$ where $\max(m,n) \le K < m+n-1$

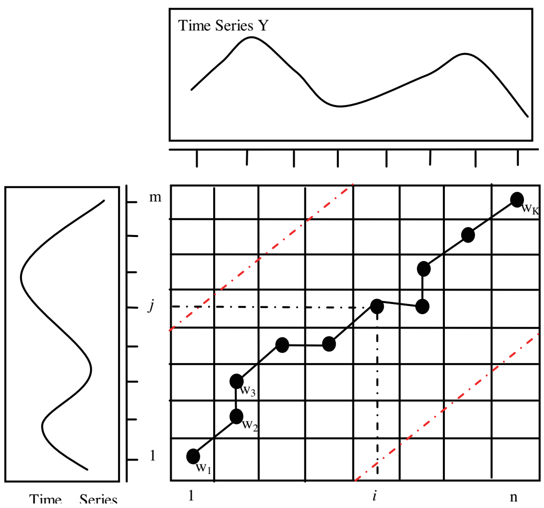

Figure 3. Illustration of Dynamic Time Warping

The warping path is typically subject to several constraints like those of boundary conditions, continuity and monotonicity.

3.2.3 Derivative Dynamic Time Warping (DDTW)

An additional problem with DTW is that the algorithm may fail to find obvious, natural alignments in two sequences simply because a feature (i.e peak, valley, inflection point, plateau etc.) in one sequence is slightly higher or lower than its corresponding feature in the other sequence.

With DDTW, the distance measure $d(x_i, y_j)$ is not Euclidean but rather the square of the difference of the estimated derivatives of x_i and y_j.

$$D_x[x] = [(x_i - x_{i-1}) + ((x_{i+1} - x_{i-1})/2)]/2 .$$

Similarly, the other sequence is also processed and then DTW of these sequences is calculated.

3.2.4 Constrained Dynamic Time Warping

Although DTW has been successfully used in many domains, it can produce pathological results. The crucial observation is that the algorithm may try to explain variability in the Y-axis by warping the X-axis. This can lead to unintuitive alignments where a single point on one time series maps onto a large subsection of another time series. Such examples are called singularities. A variety of ad-hoc measures have been proposed to deal with singularities. All of these approaches essentially constrain the possible warpings allowed. However they suffer from the drawback that they may prevent the "correct" warping from being found. Of the many solutions proposed to overcome the singularities, we have used the windowing method (Sakoe and Chiba, 1978). Allowable elements of the matrix can be restricted to those that fall into a warping window, $|i-(n/(m/j))| < R$, where R is a positive integer window width. This effectively means that the corners of the matrix are pruned from consideration, as shown by the red dashed lines in figure 3. This is also known as Sakoe Chiba band. Euclidean distance can be treated as a limiting case of CDTW with R = 0. Detailed explanation of the above metrics is provided by Keogh and Pazzani.

3.3 Training Samples

Approximately, 1% of the total pixels to be classified are deemed sufficient for training. Since the proposed classification procedure matches the query EVI patterns with the template patterns in the temporal library, it is necessary that the temporal library and thus the training samples be selected carefully so that the patterns in the temporal library are truly representative of the classes to be classified. We have also developed a method for identifying training samples automatically, the description of which is beyond the scope of this paper. However it is not necessary to use the automatic method with this classification. Training sample locations obtained through any other method can also be used with the proposed classification procedure.

4. RESULTS

Classification results were compared with 56-m resolution land use/cover map obtained from National Remote Sensing Centre. This image had 16 classes which were merged to obtain the required three classes. After appropriate scaling, it was compared with the derived map. Accuracy results are reported using several definitions of agreement between the map and primary or alternate reference land cover labels. Pixel-to-pixel comparison is the most restrictive protocol for defining agreement. It reflects a 'conservative bias' (Verbyla and Hammond, 1995) due to the confounding of true classification error with errors attributable to misregistration or inability to confidently photo-interpret a sample unit. The second definition of agreement allows a match between the photo-interpreted label of a sample pixel and the most common class within a 3 by 3 pixel block centered on the sample pixel. This comparison takes into consideration that, for many applications, a certain level of spatial generalization from the original full resolution land cover data is appropriate. Yet another set of accuracy estimates are derived using a subset of the original samples, i.e., the sample pixel is located within a homogeneous area in which only one land cover type exists within the 3x3 pixel block. The estimates based on this comparison are likely have an 'optimistic bias' (Hammond and Verbyla, 1996) because of the restriction to areas where land cover is homogeneous and generally is easily identified. The latter method is used for validation of the land use map in this study as the reference map used has a resolution of 56-m.

4.1 MODIS Dataset Results

MODIS time series data for the cropping years 2003-04, 2004-05 and 2005-06 were constructed and used as input with the proposed classifier. A value of k=5 for kNN classifier in case of MODIS dataset was found optimal.

		Reference		
Our Method		Forest	Crop	Water
	Forest	3540	3628	3
	Crop	408	36350	117
	Water	1	2055	163
Producer's Accuracy:			User's Accuracy:	
Forest	89.64295		Forest	49.3655
Crop	86.47967		Crop	98.57627
Water	57.59717		Water	7.345651
Overall % correct = 86.573003				

Table 1. Confusion matrix for classification using Euclidean distance

		Reference		
Our Method		Forest	Crop	Water
	Forest	3208	2258	4
	Crop	740	37123	118
	Water	1	2652	161
Producer's Accuracy:			User's Accuracy:	
Forest	81.23576		Forest	58.64717
Crop	88.3187		Crop	97.74098
Water	56.89046		Water	5.721393
Overall % correct = 87.521885				

Table 2. Confusion matrix for classification using DTW

		Reference		
Our Method		Forest	Crop	Water
	Forest	2260	4785	30
	Crop	1101	17567	74
	Water	588	19681	179
Producer's Accuracy:			User's Accuracy:	
Forest	57.22968		Forest	31.94346
Crop	41.79335		Crop	93.73066
Water	63.25088		Water	0.875391
Overall % correct = 43.242192				

Table 3. Confusion matrix for classification using DDTW

		Reference		
Our Method		Forest	Crop	Water
	Forest	3434	2638	0
	Crop	514	38470	135
	Water	1	925	148
Producer's Accuracy:			User's Accuracy:	
Forest	86.95872		Forest	56.55468
Crop	91.52333		Crop	98.34096
Water	52.29682		Water	13.78026
Overall % correct = 90.893764				

Table 4. Confusion matrix for classification using CDTW

Tables 1, 2, 3 and 4 show the confusion matrix, producer's accuracy and user's accuracy for proposed classifier with Euclidean, DTW, DDTW and CDTW similarity measure respectively.

Figure 4(a) shows the ground truth for the area under study and Figure 4(b) shows the classified map with CDTW as the distance measure used with proposed classification scheme and Sakoe Chiba radius of 4.

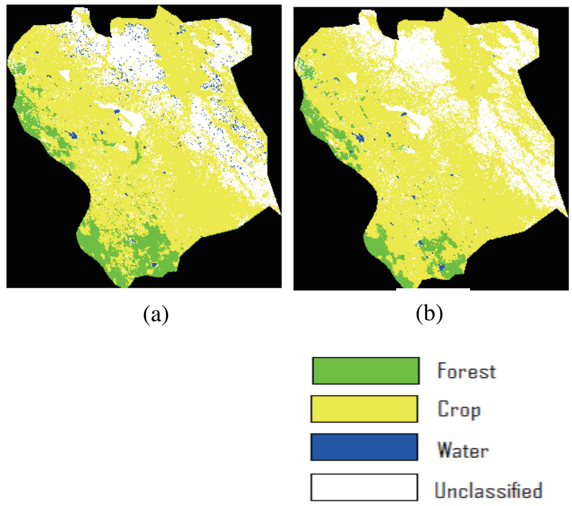

Figure 4. a) Ground truth b) Classified map of Dharwad district

4.2 AWiFS Dataset Results

To demonstrate sensor independence, the proposed classification procedure was also run on NDVI time series constructed from AWiFS data which has a spatial resolution of 56m and temporal resolution of 5 days. Classification results using the proposed classification method and CDTW as similarity measure for AWiFS data is quantified by the confusion matrix in table 5. Since the ground truth was also at a resolution of 56 meters, a pixel to pixel comparison was done to obtain this confusion matrix.

		Reference		
Our Method		Forest	Crop	Water
	Forest	224949	33232	183
	Crop	7110	43764	2199
	Water	1485	2728	21338
Producer's Accuracy:			User's Accuracy:	
Forest	96.32		Forest	87.07
Crop	54.89		Crop	82.46
Water	89.96		Water	83.51
Overall % correct = 86.0716				

Table 5. Confusion matrix for classification using CDTW distance for AWiFS data

5. OBSERVATIONS

It can be observed that the classifier using CDTW as distance measure performs better than classifier using Euclidean distance as similarity measure. Thus, the basic assumption that was made about Euclidean distance not being able to capture temporal shifts in growth patterns as intra class variation holds true. Also, derivate dynamic time warping distance measure based classifier does not perform well as this works on the derivates and not the absolute values of EVI which are important for characterizing the vegetation classes. From figure 4(b), it can be observed that a lot of misclassifications happen with water. This is largely due to the inaccurate training samples. Also vegetation index is not a proper descriptor for characterizing water. The error may also be due to the way time series is constructed using compositing in which the highest value during the compositing period is used in the time series. The performance of classifier using CDTW as distance metric is around 4% higher as compared to the classifier using Euclidean distance metric which is not phenomenal but the gap between these accuracies will increase as large spatial areas are used as input for classification. Since the input used for the results shown in the previous section was a district, the variation in the vegetation growth patterns may not have been significant. As the area used for classification increases, the probability of variations in vegetation growth patterns increases and thus the ability of the Euclidean distance to capture these variations decreases. This is where the proposed classification procedure using CDTW as the distance metric is the best way of classifying temporal signatures.

To evaluate the performance of the proposed algorithm in comparison with some commonly used classification techniques, we used the same training set to train a k-NN classifier with k=5 using LNKNET and support vector machine classifier using SVM multiclass whose results are shown in table 6 and table 7 respectively. The low classification accuracy of these classifiers for time series satellite image classification can be attributed to the use of Euclidean distance in the feature space.

		Reference		
		Forest	Crop	Water
Our Method	Forest	3126	8308	24
	Crop	820	33641	127
	Water	3	84	132
Producer's Accuracy:			User's Accuracy:	
Forest	79.159		Forest	27.28
Crop	80.035		Crop	97.26
Water	46.643		Water	60.27
Overall % correct = 79.75				

Table 6. Confusion matrix for classification results of kNN classifier (LNKNET)

		Reference		
		Forest	Crop	Water
Our Method	Forest	3196	6254	16
	Crop	647	35477	127
	Water	106	302	140
Producer's Accuracy:			User's Accuracy:	
Forest	80.932		Forest	33.76
Crop	84.403		Crop	97.86
Water	49.47		Water	25.55
Overall % correct = 83.89				

Table 7. Confusion matrix for classification results of SVM classifier (SVM multiclass)

5.1 Effect of constraining the time warp

The classifier using CDTW as distance measure was further tested with different values of Sakoe-Chiba radius and the classification accuracies noted. Figure 5 shows the result of this experiment. It is observed that the maximum classification accuracy is obtained at radius equal to 4(equivalent to 60 days) for MODIS dataset and at radius equal to 5 (equivalent to 25 days) for AWiFS dataset. Further observations that can be made from this plot are that CDTW performed better than Euclidean distance (radius = 0) and DTW (radius = 69) because CDTW is able to capture temporal shifts from pixel to pixel based on the geographic location as intra class variation. The classification accuracies saturate to a constant value after radius equal to 23 for MODIS dataset. Since there are 23 images per year, this also shows that most of the vegetation phenological patterns repeat during the subsequent years. The reduction in classification accuracy from maximum to this constant value can be attributed to the patterns that do not repeat during the subsequent years.

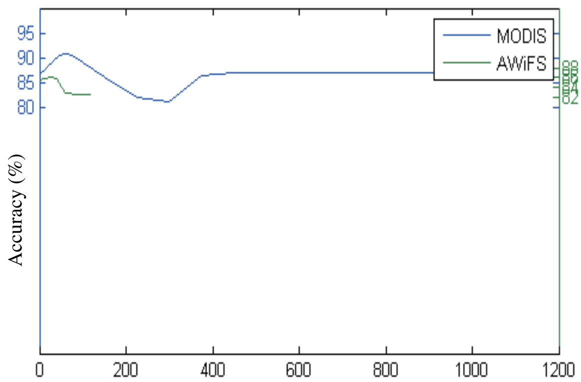

Figure 5. Plot of Classification accuracies Vs Sakoe Chiba radius

6. CONCLUSIONS

In this paper, we have successfully presented a classification procedure for time series satellite image classification to produce a land cover dataset. Temporal measurements and their analysis become important to understand the process (For e.g. vegetation growth) and thus the object which is being reinforced by this study. This is one of the first attempts at an investigation looking at these patterns as curves and applying curve matching in the time-space domain instead of clustering in the feature-space. A value of k=4(equivalent to 60 days) for the Sakoe-Chiba Band in case of MODIS data and k= 5 (equivalent to 25 days) for AWiFS data for constraining the warping path producing the highest classification accuracy indicates that there can be variations in sowing practices which can differ regionally by a maximum of one to two months. Further, this parameter itself can be used to find spatial variability in the cropping practices, a need for crop modeling and adaptation studies.

Acknowledgements – *The authors would like to thank Land Processes Distributed Active Archive Centre (LPDAAC), NASA for providing MODIS data and National Remote Sensing Centre(NRSC), India for providing AWiFS data and land cover dataset used for validation.*

REFERENCES

DeFries, R.S. and Belward, 2000. A.S. Global and regional land cover characterization from satellite data: an introduction to the special issue, *International Journal of Remote Sensing*, 21(6-7):1083-1092.

Hammond, T.O., Verbyla, D. L., 1996. Optimistic Bias in Classification Accuracy Assessment. *International Journal of Remote Sensing*. 17, 1261-1266.

Jonsson, P., and Eklundh, L., 2002. Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*, 40 (8), 1824-1832.

Jonnson, P., Eklundh, L., 2004. TIMESAT—A Program for Analyzing Time-Series of Satellite Sensor Data, *Computers & Geosciences*. 30, 833–845,

Keogh, E.J., Pazzani, M. J., 2001. Derivative Dynamic Time Warping, *Proceedings of First SIAM International Conference on Data Mining*.

LNKNET software package, MIT Open source, Available http://www.ll.mit.edu/mission/communicaTions/ist/lnknet/index.html (accessed November 2008).

Malingreau, J. P., 1986. Global Vegetation Dynamics: Satellite Observations over Asia, *International Journal of Remote Sensing*. 7, 1121–1146.

Rousseeuw, P.J., Leroy., 1987. A. M. Robust Regression and Outlier Detection, *Wiley Series in Probability and Mathematical Statistics*.

Sakoe, H., Chiba, S., 1978. Dynamic Programming Algorithm Optimization for Spoken Word Recognition, *IEEE Transactions on Acoustics, Speech, and Signal Proceedings*. Vol. -26, 43-49.

SVM light software for multiclass,Available: http://svmlight.joachims.org/svm_multiclass.html (accessed October 2008).

VERBYLA, D.L., HAMMOND T. O., 1995. Conservative Bias in Classification Accuracy Assessment due to pixel-by-pixel of 369 Classified Images with Reference Grids. *International Journal of Remote Sensing*.16, 581-587.

Wilkinson, G.G., 2005. Results and Implications of a Study of Fifteen Years of Satellite Image Classification Experiments, *IEEE Transactions on Geoscience and Remote Sensing*. 43 (3), 43.