ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Segmentation of High Resolution Satellite Image Using S-Transform.

Meenakshisundaram N*

Department of Electronics & Telecommunication Engineering, Sathyabama University Chemmai-119

ABSTRACT

The resolution of remote sensing images increases every day, raising the level of detail and the heterogeneity of the scenes. Most of the existing geographic information systems classification tools (Stock well Transform) have used the same methods for years. With these new high resolution images basic classification methods do not provide satisfactory results. A region-based classification method segmentation is based on and a classification. In this paper, we have proposed an approach for the segmentation of very high resolution (VHR) satellite images using S-Transforms. Satellite images have many applications in meteorology, agriculture, geology, forestry, landscape, biodiversity conservation, regional planning, education, intelligence and warfare. The segmentation uses an S-Transform to divide the image into several homogenous regions. Then follows the region-based classification performed either with the method MCL (Maximum Likelihood classifier). The method was validated and a comparison between pixel-based and region-based classification was performed. This method provides better results comparing to the existing remote sensing classification tools, even if some work should be done to prove its robustness. We also proved that the prior segmentation significantly improves the results of classification, both from the quantitative and qualitative points of view.

Keywords: segmentation, S-Transform, Maximum Likelihood classifier, median filter.

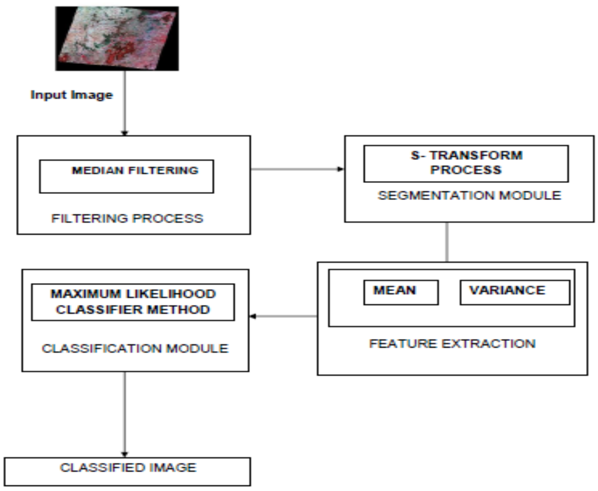
Corresponding author

INTRODUCTION

The resolution of images provided by the satellites increases every day. Some years ago, these images had a resolution of dozens of meters. The measured luminance of one pixel was representing the mean of luminance of several ground objects. Now the new satellites reach sixty centimeters of resolution, increasing the level of detail by a factor ten. With such images we can consider that each pixel is part of a single object. Thus, the heterogeneity of images has dramatically grown. Satellite images are mainly used in geographic information systems (GIS). Their classification is very useful for cartography. With low resolution satellite images, the intensity of the pixels is enough to individually classify each of them. On the contrary, high resolution image classification is more difficult. The increasing complexity of the scenes raises the level of details. For example a tree in a field or the shadows of the objects is visible, and the contextual information of the pixels becomes essential for a good classification. The existing GIS classification software's generally use the same methods for low and high resolution images. If satisfactory results can be obtained with low resolution shots, the effectiveness of this software's for high resolution images is questionable. To ensure a good accuracy, manual classification is sometimes preferred to automatic methods. The improvements of satellite imaging then require new classification methods. Some classifiers were recently developed for biomedical imagery or industry, but are still uncommon in remote sensing. Moreover in biomedical imagery a pre-processing step, the segmentation, is often added. Its aim is to divide the image into homogenous regions in order to extract contextual features. All these new methods are not fully exploited in remote sensing. Remote sensing provides a useful source of data from which updated land-cover information can be extracted for assessing and monitoring vegetation changes. In the past several decades, air photo interpretation has played an important role in detailed vegetation mapping, while applications of coarser spatial resolution satellite imagery such as Landsat Thematic Mapper (TM) and SPOT High Resolution Visible (HRV) alone have often proven insufficient or inadequate for differentiating species-level vegetation in detailed vegetation studies . Classification accuracy is reported to be only 40 percent or less for thematic information extraction at the species-level with these image types. However, high spatial resolution remote sensing is becoming increasingly available; airborne and space borne multispectral imagery can be obtained at spatial resolutions at or better than 1 m. The utility of high spatial resolution for automated vegetation composition classification needs to be evaluated. High spatial resolution imagery initially thrives on the application of urban-related feature extraction has been used, but there has not been as much work in detailed vegetation mapping using high spatial resolution imagery. This preference for urban areas is partly due to the proximity of the spectral signatures for different species and the difficulties in capturing texture features for vegetation. While high spatial resolution remote sensing provides more information than coarse resolution imagery for detailed observation on vegetation, increasingly smaller spatial resolution does not necessarily benefit classification performance and accuracy. With the increase in spatial resolution, single pixels no longer capture the characteristics of classification targets. The increase in intra-class spectral variability causes a reduction of statistical reparability between classes with traditional pixel-based classification approaches. Consequently, classification accuracy is reduced and the classification results show a salt-and-pepper effect, with individual pixels classified differently from their neighbors. To overcome this so-called H-resolution problem, some pixel-based methods have already been implemented, mainly consisting of three categories: (a) image pre-processing, such as low-pass filter and texture analysis, (b) contextual classification, and (c) post-classification processing, such as mode filtering, morphological filtering, rule-based processing, and probabilistic relaxation. A common aspect of these methods is that they incorporate spatial information to characterize each class using neighborhood relationships.

The segmentation is a process which extracts the outline of the ground objects by defining homogenous regions. Most of the methods only use the intensity of each pixel to define the regions, but produce very noisy segmentations, particularly with the high resolution satellite images. Some algorithms now include contextual information in the process to reduce the heterogeneity of the segmentations. In some of them textural information extracted from the image is also used. The segmentation step is generally made using Stockwell Transform (S Transform). We first have the pre-processing, in which we select the features to use and eventually modify or re-scale the data. The second part is the initialization of the segmentation algorithm, if needed. Finally the third part is the segmentation itself.

FLOW GRAPH

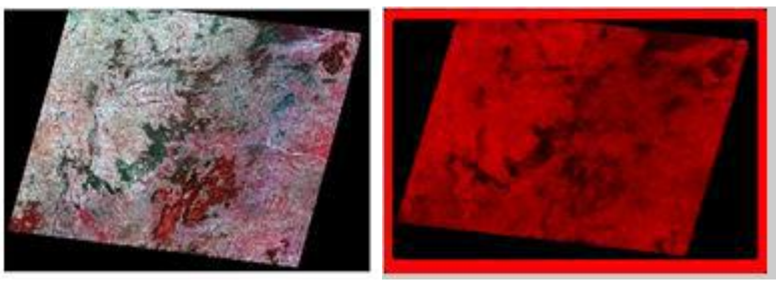


EXPERIMENTAL METHOD AND ALGORITHM USED

MEDIAN FILTER

The main idea of the median filter is to run through the signal entry by entry, replacing each entry with the median of neighboring entries. The pattern of neighbors is called the "window", which slides, entry by entry, over the entire signal. For 1D signal, the most obvious window is just the first few preceding and following entries, whereas for 2D (or higher-dimensional) signals such as images, more complex window patterns are possible (such as "box" or "cross" patterns). Note that if the window has an odd number of entries, then the median is simple to define it is just the middle value after all the entries in the window are sorted numerically. For an even number of entries, there is more than one possible median, see median for more details. Median filtering is one kind of smoothing technique, as is linear Gaussian filtering. All smoothing techniques are effective at removing noise in smooth patches or smooth regions of a signal, but adversely affect edges. Often though, at the same time as reducing the noise in a signal, it is important to preserve the edges. Edges are of critical importance to the visual appearance of images, for example. For small to moderate levels of (Gaussian) noise, the median filter is demonstrably better than Gaussian blur at removing noise whilst preserving edges for a given, fixed window size. However, its performance is not that much better than Gaussian blur for high levels of noise, whereas, for speckle noise and salt and pepper noise (impulsive noise), it is particularly effective. Because of this, median filtering is very widely used in digital image processing.

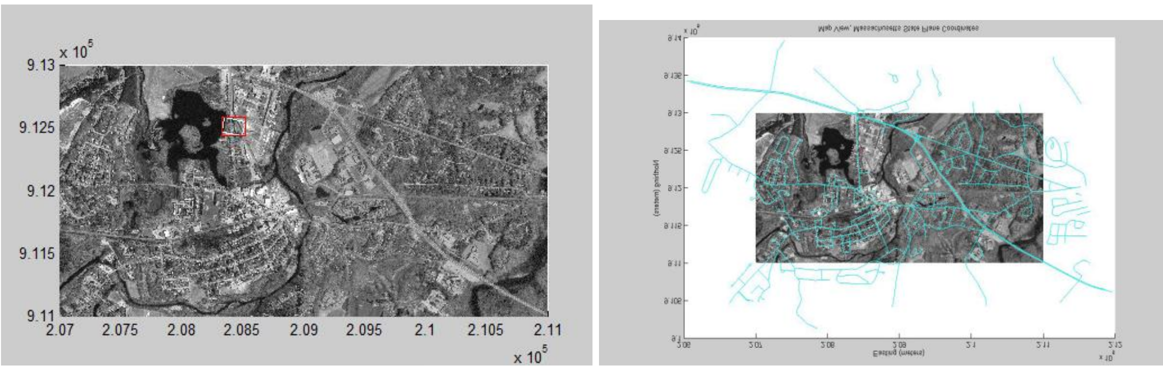
a. Input Image **b. Median filter output image**



STOCK WELL TRANSFORMS (S TRANSFORM)

The S-transform (ST) is a time-frequency representation known for its local spectral phase properties. A key feature of the S-transform is that it uniquely combines a frequency dependent resolution of the time-frequency space with absolutely referenced local phase information. This allows one to define the meaning of phase in a local spectrum setting, and results in many advantageous characteristics. It also exhibits a frequency invariant amplitude response, in contrast to the wavelet transform The S-transform is shown to have absolutely referenced phase information, a quality that the continuous wavelet transform is lacking. The S-transform is shown to have a frequency invariant amplitude response in contrast to the continuous wavelet transform which attenuates high frequency signals relative to the low frequency signals. The S-transform (ST) is similar to a continuous wavelet transform in having progressive resolution but unlike the wavelet transform, the S-transform retains absolutely referenced phase information and has a frequency invariant amplitude response. Absolutely referenced phase information means that the phase information given by the S-transform refers to the argument of the co sinusoid at zero time (which is the same meaning of phase given by the Fourier transform). The S-transform defines what local phase means in an intuitive way, at a peak in local spectral amplitude (indicating a quasimonochromatic signal), as well as off peak, where the rate of change of the phase leads to a channel Instantaneous Frequency analysis. The S-transform not only estimates the local power spectrum, but also the local phase spectrum. It is also applicable to the general complex valued time series. It is often useful to think of the time series as a single vector in an N-dimensional vector space. The basis vectors of this time series in the time domain are the vectors (1,0,0,...,0), (0,1,0,...,0) and so on. The action of the Fourier transform is simply a change of basis on the time series, from these delta function basis vectors (time domain), to sinusoidal basis vectors (frequency domain). The time series itself, which is a defined single vector in this N-dimensional vector space, remains unchanged. One of the reasons that the Fourier transform is ubiquitous in the analysis of geophysical phenomena is that the sinusoidal basis functions are the solution to the mathematical equations describing a small perturbation of a physical system about a stable equilibrium, and thus provides a suitable framework for studying such phenomena. Also a number of theoretical predictions concerning the evolution of such systems are easily couched in terms of Fourier theory. Thus changing the representation of the time series may present the information contained in the time series in a more easily assimilated form. The S-transform produces a time-frequency representation of a time series. It uniquely combines a frequency dependent resolution with simultaneously localizing the real and imaginary spectra.

Segmentation output



FEATURE SELECTION

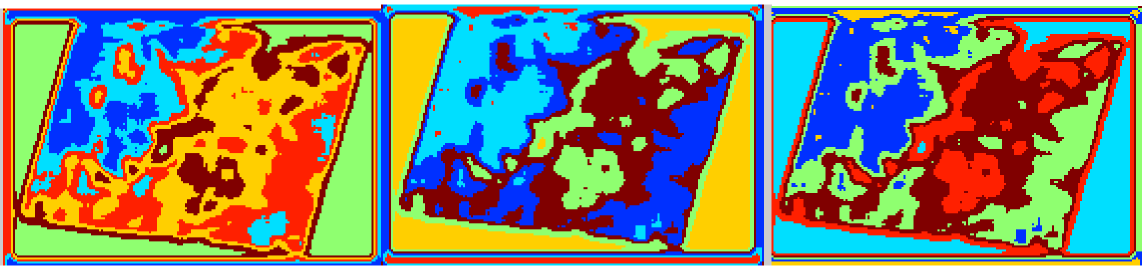
After the segmentation by the modified graph cut theory, features are extracted from the segmented regions. Feature Extraction is of vital importance because on the basis of extracted features the maximum likelihood classifier will be trained and the final classification is carried out. The features extracted include Standard deviation. The mean value is calculated by the formula:

$$M = \sum_{j=1}^{D} \frac{I_j}{D}$$

Where, j I is intensity of the pixels of the image
D is the number of pixels in the segment.

Image intensity is the most used feature in the segmentation methods. Satellite images are usually multi-spectral, which means that the image is made up of several bands. For example a Quick bird image consists of four bands: blue, green, red and near infrared. The intensity of each band can be extracted, we can then use four intensity features. Some ground objects are better defined with their texture than with their intensity. For example, with an intensity-based segmentation a forest will be divided into several homogenous parts, which is not the case of texture based segmentation. Although the texture is not directly available, it can be extracted from image intensity.

Feature Extraction output



MAXIMUM LIKELIHOOD CLASSIFICATION

Maximum Likelihood classification (ML) is a supervised classification method which is based on the Bayes theorem. It makes use of a discriminated function to assign pixel to the class with the highest likelihood. Class mean vector and covariance matrix are the key inputs to the function and can be estimated from the training pixels of a particular class. In this study, we used ML to classify a diverse tropical land covers recorded from Landsat 5 TM satellite. The classification is carefully examined using visual analysis, classification accuracy, and band correlation and decision boundary. The results show that the separation between mean of the classes in the decision space is to be the main factor that leads to the high classification accuracy of ML. Each pixel is assigned to the class with the highest likelihood or labeled as unclassified if the probability values are all below a threshold set by the user.

The general procedures in ML are as follows:

- The number of land cover types within the study area is determined.
- The training pixels for each of the desired classes are chosen using land cover information for the study area. For this purpose, the Jeffries-Matusita (JM) distance can be used to measure class reparability of the chosen training pixel.
- The training pixels are then used to estimate the mean vector and covariance matrix of each class.
- Finally, every pixel in the image is classified into one of the desired land cover types or labeled as unknown.
- In ML classification, each class is enclosed in a region in multispectral space where its discriminated function is larger than that of all other classes.

These class regions are separated by decision boundaries, where, the decision boundary between class i and j occurs when:

Training areas were established by choosing one or more polygons for each class. Pixels fall within the training area were taken to be the training pixels for a particular class. In order to select a good training area for a class, the important properties taken into consideration are its uniformity and how well they represent the same class throughout the whole image. Class reparability of the chosen training pixels was determined by means of the JM distance. Fifty pairs have JM distance between 1.9 and 2.0 indicating good reparability, four from 1.0 to 1.9 indicating moderate reparability and one less than 1.0 indicating poor reparability. The worst reparability, possessed by the urban – industry pair (0.947), was expected since both have quite similar spectral characteristics. For each class, these training pixels provide values from which to estimate the mean and covariances of the spectral bands used. This information is to be used by the ML classifier to assign pixels to a particular class. The outcome of ML classification after assigning the classes with suitable colors, coastal swamp forest (green), dry land forest (blue), oil palm (yellow), rubber (cyan), cleared land (purple), coconut (maroon), bare land (orange), urban (red), industry (grey), sediment plumes (sea green) and water (white). Clouds and their shadows are masked black. The areas in terms of percentage and square kilometers were also computed; the classes with the largest area are oil palm, cleared land and industry. Although being similar, coastal swamp forest and dry land forest can be clearly seen in the south-west and northeast of the classified image, as indicated by the reference map. Coastal swamp forest covers most of the Island and coastal regions in the south-west of the scene. Most of the dry land forest can be recognized as a large straight-edged region in the north-east. Oil palm and urban dominate the northern and southern parts respectively. Rubber appears as scattered patches that mostly are surrounded by oil palms. Industry can be recognized as patches near the urban areas, especially in the south-west and north-east. Coconut can be seen in the coastal area in the north-west of the image. A quite large area of bare land can be seen in the east, while cleared land can be seen mostly in the north, south and south-east of the image. The most critical setting for the classification step is probably the choice of the input data, and more precisely the choice of the relevant features. Actually, for each class there are some relevant and some useless features. If the user chooses to perform the classification step with all the presented features, the results will probably not be satisfying. The problem is that the useless features reduce the importance of the relevant ones, resulting in a ineffective classification.

The feature selection is an essential step to ensure the best results. One of the possible methods is to let the user choose which features he wants to use. However, it is very difficult to know which features are relevant or not, because it is generally not observable. In the following section we will propose a feature selection method combining two techniques, the cross-validation and the sequential generation.

The maximum likelihood classifier is one of the most popular methods of classification in remote sensing, in which a pixel with the maximum likelihood is classified into the corresponding class. The likelihood Lk is defined as the posterior probability of a pixel belonging to class k.

$$Lk = P(k/\mathbf{X}) = P(k)*P(X/k) / P(i)*P(\mathbf{X}/i)$$

where P(k) : prior probability of class k

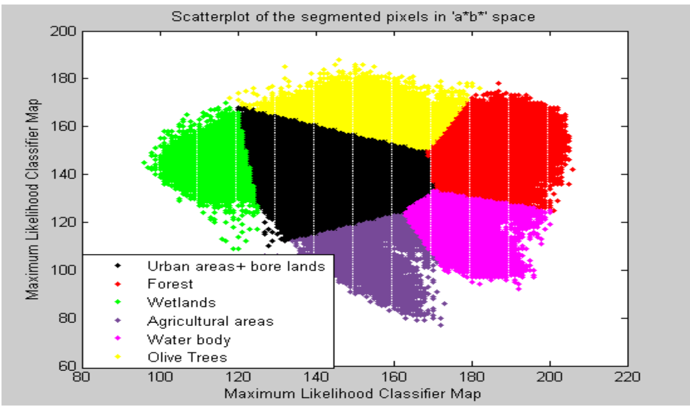
P(X/k) : conditional probability to observe X from class k, or probability density function.

Usually P(k) are assumed to be equal to each other and P(i)*P(**X**/i) is also common to all classes. Therefore Lk depends on P(**X**/k) or the probability density function. For mathematical reasons, a multivariate normal distribution is applied as the probability density function. In the case of normal distributions, the likelihood can be expressed as follows.

$$L_k(\mathbf{X}) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma_k|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(\mathbf{X}-\mu_k)\Sigma_k^{-1}(\mathbf{X}-\mu_k)^t\right]$$

where n: number of bands
X: image data of n bands
Lk(X) : likelihood of X belonging to class k
k : mean vector of class k
k : variance-covariance matrix of class k

Maximum likelihood classifier map



ISSN: 0975-8585

RESULTS, DISCUSSIONS AND PERFORMANCE ANALYSIS

The choice of the classifier is not obvious and strongly depends on the data. Among the supervised classifiers, a study assessed that the support vector machine is generally more effective than neural network and maximum likelihood classifiers. According to this study, support vector machine is often considered to be the best compromise between results and complexity. However we have to keep in mind that for neural network and support vector machine classifiers the choice of the internal parameters is essential to ensure good results and is sometimes very hard, while the Mahalanobis distance classifier for example does not need any parameter adjustment. One of the main goals of this study is to evaluate the need of a prior segmentation for the classification. We found much remote sensing classification software, but those which include a prior segmentation are uncommon. Most of them only perform pixel-based classification. The library of classifiers used in these software's is almost always the same: unsupervised classifiers (K-means) and supervised classifiers (maximum likelihood, K-nearest neighbors, Mahalanobis distance, and sometimes neural network). For supervised classification, testing areas are selected by the user, or the specifications of the classes are chosen from a library of pre-defined values. It is also important to note that some of this software allows doing different levels of classification. For example we can create a classifier with two levels, the first one separating water and land, and a second step only for land areas separating urban and forest regions

DISCUSSION

The results presented above allow us to do some conclusions on the segmentation step. The S-Transform algorithm used for the segmentations seems to be an effective method. The qualitative evaluation shows that the results of segmentation are good and provide very few noises. Moreover the basic principles of this method allow to segment simultaneously small and large areas, thus improving the conservation of the geometry. The quantitative evaluation proves that the error caused by the segmentation is very small, which confirms the visual observations. Although quantitative evaluations were not performed on other type of images, user can suppose with the help of the visual observations that this method also works for Aster shots, and probably for other types too. However the effectiveness of the segmentation algorithms depends on the quality and the quantity of information of the image. The comparison of our segmentation results with an E-cognition segmentation demonstrated that the accuracy of both methods is quite equivalent. However, when performing segmentation with E-cognition, the user must select the size of the segments. This method then provides segments with equal size, which is not representative of the reality. From a quantitative point of view, user can use more number of components to reduce the error. However, we must keep in mind that the more the number of components, the more the segments, which can lead to an over-segmentation. In other terms, the real regions of the image will be divided into several smaller parts, decreasing the information contained in the features extracted from these parts, and reducing the performance of the classification. It is then very important to find a compromise between geometry and accuracy to ensure the best classification results. Finally, note that the S-Transform algorithm would not provide satisfactory results if the pre-processing and the initialization were not properly performed. The two prior steps presented for the segmentation are as important as the segmentation itself. By using the suggested method for pre-processing and initialization, the probability of finding good segmentation results is high. When we look at the manually classified image, it seems obvious that the classes do not have the same spectral and geometrical characteristics. For example, we can say that the compactness is relevant for urban areas, since these regions are very elongated. On the contrary, farming areas do not present singularities in shape features. We can then say that each class has its own set of relevant features. The principle of the multi-level method is to divide the classification into several levels. Each level discriminates one class c from the remaining ones, and excludes the regions belonging to this class for the next level. Then the same method is applied to the remaining regions, until all the classes are determined. For each level, only the relevant features are used, so that each class is constructed with its best features. By its definition the support vector machine is the most appropriate for one against the rest classification. We evaluated the potential of this method for the training set. We can see that shape features are essentially used for urban areas, as expected. The individual scores are very good, and by combining these classification levels in the order "Urban, Farming 1, Scarce vegetation, Wooded, Open land; Farming 2", we obtained a global classifier accuracy of 84.6 percent. The improvement comparing to the classical support vector machine classifier (79.5 percent) is not really significant. Moreover this method requires more time and there is a lot of parameters to estimate. Therefore, we require executing the satellite image classification using maximum likelihood classifier.

CONCLUSION

In this paper, a parametric supervised classification algorithm based on Maximum Likelihood applied to LIDAR data using data fusion is presented. Several bands such as first, last echo and intensity LIDAR data and co-registered line scanner bands such as aerial and near infra-red photos are employed to build up a feature space. Four classes are classified and their individual accuracy is assessed. The results show that detached objects (buildings, vegetation) and bare earth are correctly classified up to 88.17%. The performance of class car, however, shows potential for further improvement due to its ambiguous appearance within the different features and the limited number of samples which can be collected compared to other classes. The aim of this paper is to provide some classification tools for remote sensing images, and we can affirm that this aim was successfully achieved. For the segmentation step we demonstrated the effectiveness of the S-Transform algorithm, both from the qualitative and quantitative points of view. Some improvements could also be added to this method. A post-processing could be inserted to suppress the noise in the homogenous regions. The segmentation step could also be improved in order to increase the potential of the classification step. However, due to the limits imposed by remote sensing acquisition systems, we do not think that exceptional progress could be provided to this classification process.

REFERENCES

[1] L. A. Ruiz, A. Fdez-Sarria, and J. A. Recio, “Texture feature extraction for classification of remote sensing data using wavelet decomposition: A comparative study,” IAPRS, vol. 35, 2002.

[2] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural features for image classification,” IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-3, pp. 610–621, November 1973.

[3] N. Sebe and M. S. Lew, “Wavelet based texture classification,” International Conference on Pattern Recognition, vol. 3, pp. 947–950, 2000.

[4] S. Liapis, N. Alvertos, and G. Tziritas, “Maximum likelihood texture classification and bayesian texture segmentation using discrete wavelet frames,” International Conference in Digital Signal Processing DSP97,1997.

[5] V. Lakshmanan, V. DeBrunner, and R. Rabin, “Texture-based segmentation of satellite weather imagery,” International Conference on Image Processing, vol. 2, pp. 732–735, 2000.

[6] B. R. Corner, R. M. Narayanan, and S. E. Reichenbach, “Application of principal component analysis to multi sensor classification,” Proc. 27th Applied Imagery Pattern Recognition Workshop, pp. 202–210, 1998.

[7] J. A. Bilmes, “A gentle tutorial of the em algorithm and its application to parameter estimation for gaussian mixtures and hidden markov models,” Univeristy of Berkeley, ICSI-TR-97-021, pp. 1–7, 1998.

[8] L. Amouraandw.Pieczynski, “Hierarchical markov fields and fuzzy image segmentation,” Proceedings Second IEEE International Conference on Intelligent Processing Systems, pp. 154–158, 1998.

[9] P. Lanchantin and W. Pieczynski, “Unsupervised non stationary image segmentation using triplet markov chains,” Advanced Concepts for Intelligent Vision Systems ACVIS04, 2004.

[10] J. Zhang and W. Modestino, “Maximum-likelihood parameter estimation for unsupervised stochastic model-based image segmentation,” IEEE Transactions on Image Processing, vol. 3, pp. 404–420, July 1994.

[11] M. Neubert and G.Meinel, “Evaluation of segmentation programs for high resolution remote sensing applications,” Proceedings of the ISPRS Joint Workshop High Resolution Mapping from Space, 2003.

[12] X. Gigandet, “Segmentation d’images satellites multispectrales par classification statistique,” tech. rep., Swiss Federal Institute of Technology (Lausanne), 2004.

[13] W. Pieczynski, “Mod`eles de markov en traitementd’images,” Traitement du Signal, Presses Universitaires de Grenoble, vol. 20, no. 3, pp. 255–277, 2003.

[14] M. B. Cuadra, Atlas-based segmentation and classification of magnetic resonance brain images. PhD thesis, Swiss Federal Institute of Technology (Lausanne), November 2003.

[15] C. J. L. C. W. Hsu, C. C. Chang, “A practical guide to support vector classification,” 2003.

[16] H. Liu and H. Motoda, Feature Selection for Knowledge Discovery and Data Mining. Kluwer, 1998.

[17] C. M. Bishop, Neural Networks for Pattern Recognition. Oxford Clarendon Press, 1995.

[18] M. Pal and P. M. Mather, “Support vector classifiers for land cover classification,” Map India, Image Processing & Interpretation, 2003.

[19] P. Watanachaturaporn, P. K. Varshney, and M. K. Arora, "Evaluation of factors affectiong support vector machines for hyperspectral classification," The American Society for Photogrammetry & Remote Sensing ASPRS, 2004.
[20] A. Karlsson, "Classification of high resolution satellite images," tech. rep., Swiss Federal Institute of Technology (Lausanne), August 2003.
[21] R. G. Congalton and K. Green, Assessing the Accuracy of Remotely Sensed Principles and Practices, ch. 5, pp. 45–50. Lewis, 1999.