VII Simposio Latinoamericano de Percepción Remota

Sexta Reunion Nacional SELPER-Mexico

Latinoamérica Evaluada desde el Espacio

Puerto Vallarta, México

Memorias

Noviembre, 1995

SAR TEXTURE DISCRIMINATION USING AR-2D MODELS FOR AMAZONIAN LAND USE CLASSIFICATION

Sidnei J. S. Sant'Anna
Luciano V. Dutra
INPE — Instituto Nacional de Pesquisas Espaciais
DPI — Divisão de Processamento de Imagens
Avenida dos Astronautas, 1758
12227-010 São José dos Campos
SP - Brasil
{sidnei;dutra}@dpi.inpe.br

ABSTRACT

In tropical ecology studies, forest classification is a key issue. Although there is no widely accepted forest classification criterion, it is well known that texture is an important factor to discriminate forest types and other land cover, particularly when using radar images. The two-dimensional autocorrelation function (ACF-2D) seems to characterize better the SAR textures, but its direct estimation is not precise in general. Autoregressive moving average (ARMA) modelling for time series is a method used to obtain better estimates of the spectral density function (SDF), with less training data. In this paper is proposed and tested a method for SAR texture sample classification based on two-dimensional ARMA modelling. Fifteen samples of primary forest and five samples of non-forest (pasture and agricultural crops) were collect from SAREX data (C-band, HH polarization, $6m$ resolution, 7 looks) in the Tapajós National Forest (Flona) region in Pará state, Brazil. Two-dimensional AR models (a subset of ARMA models) were estimated for each one of the samples. Euclidian distances were computed between the model coefficient vector of the samples and the average coefficient vector for the two classes (defined here as the class vectors). The coefficient vectors formed two perfect clusters, corresponding to each one of the classes, in a non-linear mapping of the coefficient space. This result demonstrated that these classes can be discriminated by using this method. Similar procedure was tested using the angle between sample vectors and the class vectors, and clusters were also observed.

1 INTRODUCTION

Texture is an important factor to discriminate forest types and other land cover, particularly when using radar images. However there is no widely accepted mathematical definition for texture which comprises all types of texture that can be found in nature. Traditional texture measures, like concurrence matrix derived features and structural approaches have failed to characterize texture in synthetic aperture radar (SAR) images adequately because of the strong influence of the speckle noise. Other commonly used features, like the coefficient of variation, although well fitted for the SAR theory, can not gather all the texture information, being considered a simple roughness measure.

The two-dimensional autocorrelation function (ACF-2D) seems to characterize better the SAR textures, but its direct estimation is not precise in general. Distance of ACF-2D has been successfully proposed as a tool to classify SAR texture samples (Dutra, 1990), but it needs a large amount of training data to have a good estimation of the involved ACFs. Autoregressive-moving average (ARMA) modelling is a method normally used to obtain better estimates of the spectral density function (SDF), with less training data. Due to the possibility of choosing different types of models and the method of order estimation, ARMA modelling provides more flexible assumptions than the implicit one that unobserved data outside the window is zero, used in the traditional ACF estimates.

The use of ARMA models does also have other advantages, like the use of fewer coefficients to represent a particular ACF structure and the direct use of the model for simulation. In this paper is proposed and tested methods for SAR texture sample classification based on two-dimensional ARMA models (ARMA-2D). The theory and methods for estimating ARMA-2D models are not still fully developed, being available only methods for estimating two-dimensional Autoregressive (AR-2D). An approximation to two-dimensional ARMA modelling is achieved by considering the texture samples as seasonal one-dimensional ARMA process by concatenating rows of the image data.

2 ARMA MODELS FOR UNIVARIATE TIME SERIES

A time series can be described by a sequence of random variables $y_1, y_2, \cdots, y_N$. These series are modelled as being generated by a sequence of independent *shocks* w_i, sample values of a white noise process with zero mean and σ_w^2 variance, which is the input of a linear filter that characterizes the process. This filter is defined by the equation:

$$y_i = \sum_{k=0}^{q} \alpha_k w_{i-k} + \sum_{j=1}^{p} \beta_j y_{i-j} + \mu \tag{1}$$

where μ is the expected value of y_i.

This is called an autoregressive-moving average model of order p and q (ARMA(p,q)); α_0 is normally set to one. Specialized models are derived from eq. (1); for $q = 0$ an autoregressive model of order p (AR(p)) is obtained and for $p = 0$ (no regressive terms), a moving average model of order q (MA(q)) is defined. Note that these models are *causal*, because the output, given certain initial conditions, depend only on the past values of the random process and past *shocks*.

Preliminary estimation of model parameters are obtained, from existing training data, by examing the plots of the autocorrelation function (FAC) and the partial autocorrelation function(FAP) to help decide the model orders. A non-interative method is used (Box and Jenkins, 1970) to determine a first set of model coefficients which is used as initial guess to an iterative maximum likelihood approach.

3 THE ESTIMATION OF TWO-DIMENSIONAL ARMA MODELS USING CONCATENATION

Univariate ARMA models can be generalized to generate two-dimensional random fields, by considering proper support regions on the plane. The notion of past is ill defined for the plane and it is replaced by the ***recursive computability*** notion (Dudgeon and Mersereau, 1984). A two dimensional model is recursively computable if one uses a support region which permits direct filtering or synthesis (obtaining y_i from w_i) and inverse filtering or analysis (obtaining w_i from y_i) in only one pass. The most used support regions are the so called non symmetrical half plane support (NSHP) and the quarter plane (QP) support (Kay, 1988). The estimation of two-dimensional ARMA models, either using NSHP support, also known as unilateral ARMA models (UARMA), or using QP support (QARMA) are not trivial, being available (Marple, 1987) only least squares or the so called Yule-Walker estimates solution for the QAR models (Therrien, 1989).

Trying to overcome the problems found for estimating true two-dimensional models, one assumes images as a two-dimensional ***separable*** process or an image is linearized through the concatenation of rows or columns. In these cases unidimensional methods can be readily applied, in spite of the inaccuracies implied by the given hypothesis. In this paper, SAR images are linearized by concatenating stacked portions of image rows. A non-zero coefficient at lag multiple of the size of the row (or column) would correspond to a pixel contiguous to the pixel being generated (Dutra, 1990).

The methodology can be briefly described by the following steps:

1. Obtain data from the training areas by concatenating segments of rows or columns.
2. Remove the average value of data and plot the autocorrelation function (ACF) and partial autocorrelation function (PAF). Keep the lag whose ACF value falls into confidence interval as initial maximum value for q. Keep the lag whose PAF value falls into confidence interval as a intial maximum value for p.
3. Obtain preliminary estimates of ARMA(p,q), AR(p), MA(q). Eliminate models that result non-causal or non invertible by sistematically reducing p and/or q for the three types of models.
4. By using confidence intervals for the coefficients eliminate those which are non-significant, which will not be considered for the next steps.
5. Submit the models to maximum likelihood estimation. Keep the model with smaller number of coefficients. This is the parsimony principle which considers that a better model is a simpler one. Smaller models also result in smaller filters which have better performance crossing region boundaries.

Note that the original two-dimensional nature of data will quickly reveal the models support regions.

4 ESTIMATION AND CLASSIFICATION OF FOREST TEXTURE MODELS

To test the methodology fifteen test sites were selected from Tapajós National Forest (Flona) in Brazilian Amazônia and five test sites were selected from regrowth areas beside Flona. Tapajós National Forest is a forest reserve under the administration of the Brazilian Institute of the Environment (IBAMA). Its geographic coordinates are: S 02°40′ to S 04°10′ and W 54°45′ to W 55°00′. The forest localization is shown in figure 1.

The estimations were initially restricted to UAR models that were estimated for all twenty training sites, using the methodology given in the previous section. Figure 2 presents the average models for the forest and

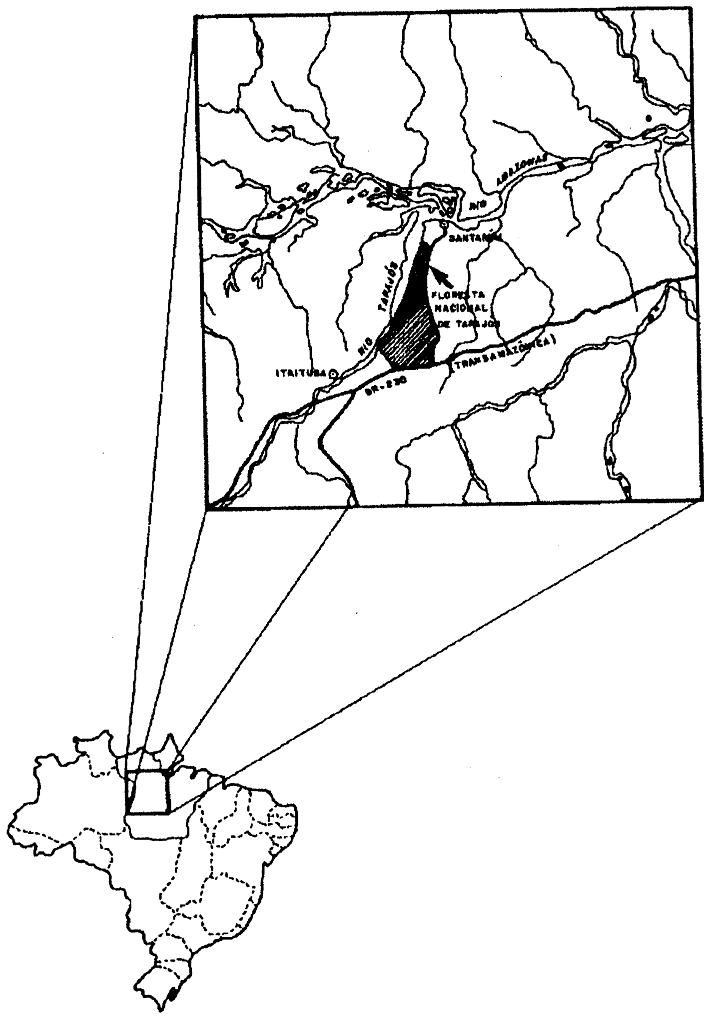

Figure 1: Tapajós National Forest localization.

non-forest classes. If ARMA models are good descriptors for texture, all models for the same class should cluster in the coefficient space. Due to high dimensionality of the coefficient space, a non-linear mapping using distances between each training pattern and the average model (Therrien, 1989) was used to visualize the distribution of model patterns in the feature space.

Table 1 presents the calculated distances between the model vectors and the average model vectors. Φ_F represents the average model vector for forest, Φ_i are the model vectors for the forest training areas, Ψ_{NF} is the average model vector for non-forest, being Ψ_i vectors for non-forest training areas. Table 2 presents the calculated angles between training vectors and the average vectors. Angle in this context is also a measure of how model vectors are similar.

Figure 3 plots the distances of all training vectors to non-forest average vector on the vertical axis and the distances to the average forest model vector on the horizontal axis. Figure 4 plots the angles of training vectors

(a)

–	•	0.556	-0.156
0.035	0.833	-0.344	0.060
-0.006	-0.276	0.097	–

(b)

–	–	–	–	•	0.491	-0.158	0.031
–	0.022	–	0.017	0.606	-0.231	0.052	–
-0.019	–	–	–	-0.164	0.060	–	–
0.018	–	–	–	–	–	–	–

Figure 2: Average UAR models for (a) Forest and (b) Non-Forest.

to the same average vectors. It is clear from the plots that distances between vectors have better performance.

From figure 3 is seen that vector distances to forest model vector would be enough to separate forest from non-forest, while with angles, the two features are needed to separate regions, although one non-forest vector can not be linearly separated.

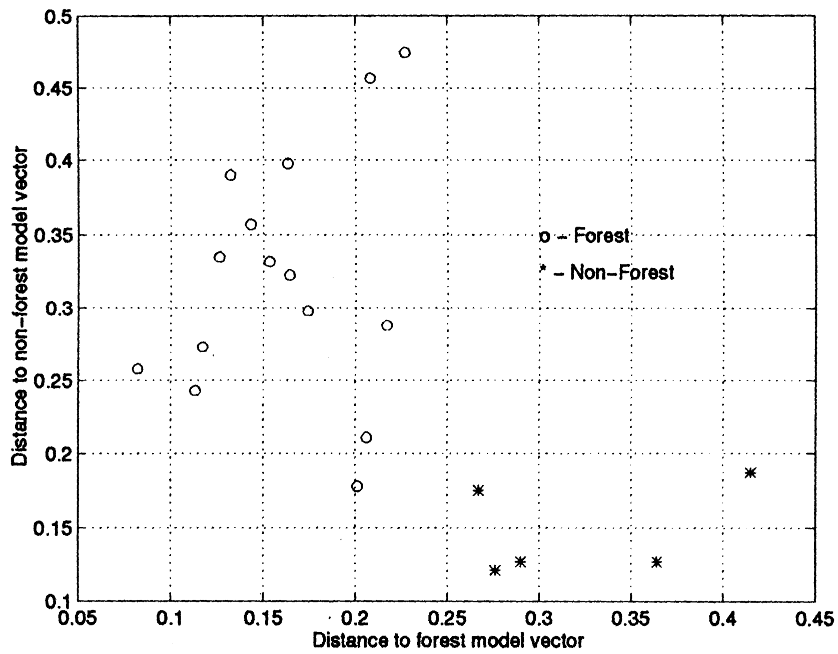

Figure 3: Distances between the model vectors and average model vectors.

	Φ_F	Ψ_{NF}
Φ_1	0.132	0.390
Φ_2	0.201	0.178
Φ_3	0.164	0.323
Φ_4	0.174	0.298
Φ_5	0.206	0.211
Φ_6	0.217	0.288
Φ_7	0.163	0.398
Φ_8	0.113	0.243
Φ_9	0.153	0.332
Φ_{10}	0.227	0.475
Φ_{11}	0.082	0.258
Φ_{12}	0.208	0.457
Φ_{13}	0.143	0.357
Φ_{14}	0.117	0.273
Φ_{15}	0.126	0.335
Ψ_1	0.290	0.127
Ψ_2	0.364	0.127
Ψ_3	0.415	0.187
Ψ_4	0.276	0.121
Ψ_5	0.267	0.175

Table 1: Distances between the model vectors and average model vectors.

5 CONCLUSIONS AND FUTURE WORK

ARMA modelling is shown to be a potencially good tool to characterize and discriminate SAR textures. Although no maximum likelihood estimators are available to estimate true two-dimensional ARMA models, simplifications, in this case, transforming two-dimensional process into one dimensional by concatenation, can be used to determine aproximate two-dimensional ARMA models. Succesfull test were made with UAR models to discriminate different SAR textures. Future work will be focused into testing other SAR textures and types of models for modelling and discrimination. Separable ARMA estimation will be used to access its performance as texture descriptor and discriminatory properties. Also ARMA modelling provides a good theorectical framework to develop statistical texture descriptors.

6 ACKNOWLEDGMENTS

This work was partially supported by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo, Brazil), Project N° 91/3532 – 2 and by CNPq, Project Geotec N° 680.061 – 94 – 0 (ProTem/CC).

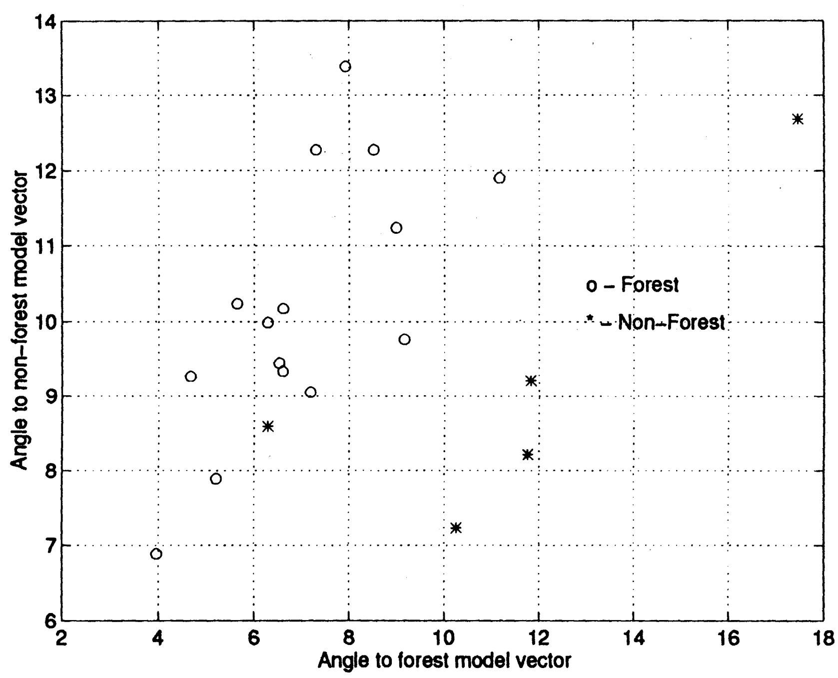

Figure 4: Angles between the model vectors and average model vectors.

	Φ_F	Ψ_{NF}
Φ_1	4.659	9.260
Φ_2	7.181	9.052
Φ_3	8.509	12.274
Φ_4	8.980	11.242
Φ_5	9.150	9.761
Φ_6	11.161	11.902
Φ_7	7.294	12.272
Φ_8	5.208	7.890
Φ_9	7.922	13.389
Φ_{10}	6.530	9.441
Φ_{11}	3.959	6.884
Φ_{12}	6.603	10.175
Φ_{13}	6.601	9.329
Φ_{14}	5.640	10.236
Φ_{15}	6.280	9.990
Ψ_1	6.289	8.590
Ψ_2	11.765	8.215
Ψ_3	17.451	12.680
Ψ_4	10.266	7.224
Ψ_5	11.827	9.202

Table 2: Angles between the model vectors and average model vectors.

7 REFERENCES

Box, G.E.P.; G.M. Jenkins, 1970, *Times series analysis: forecasting and control.* Holden Day, San Francisco.

Dudgeon, D.E.; R.M. Mersereau, 1984, *Multidimensional digital signal processing.* Englewood Cliffs, NJ, Prentice Hall.

Dutra, L.V., 1990, *Classificação de texturas usando modelos ARMA e distâncias da função de autocorrelação.* (Tese de Doutorado em Computação Aplicada) – INPE, São José dos Campos, SP, Brazil.

Kay, S.M., 1988, *Modern spectral estimation: theory and application.* Englewood Cliffs, NJ, Prentice Hall.

Marple, S.L., 1987, *Digital spectral analysis.* Englewood Cliffs, NJ, Prentice Hall.

Therrien, C.W., 1989, *Decision estimation and classification: An introduction to pattern recognition and related topics.* John Wiley and Sons.