

Crop type recognition based on Hidden Markov Models of plant phenology

P. B. C. Leite¹, R. Q. Feitosa¹, A. R. Formaggio², G. A. O. P. Costa¹, K. Pakzad³, I. D. A. Sanches²

¹Catholic University of Rio de Janeiro (PUC-Rio), ²National Institute for Space Research (INPE), ³Leibnitz University Hannover (IPI)

paula.leite@gmail.com, raul@ele.puc-rio.br, formag@dsr.inpe.br, gilson@ele.puc-rio.br, pakzad@ipi.uni-hannover.de, iedasanches@gmail.com

Abstract

This work introduces a Hidden Markov Model (HMM) based technique to classify agricultural crops. The method recognizes different crops by analyzing their spectral profiles over a sequence of satellite images. Different HMMs, one for each of the considered crop classes, are used to relate the varying spectral response along the crop cycles with plant phenology. The method assigns for a given image segment the crop class whose corresponding HMM presents the highest probability of emitting the observed sequence of spectral values. Experiments were conducted upon a sequence of 12 previously classified LANDSAT images. The performance of the proposed multitemporal classification method was compared to that of a monotemporal maximum likelihood classifier, and the results indicated a remarkable superiority of the HMM-based method, which achieved an average of no less than 93% accuracy in the identification of the correct crop, for sequences of data containing a single crop class.

1. Introduction

Given the importance of agriculture worldwide, socially and economically, the availability of precise and efficient information about agricultural activities in an appropriate time interval is highly relevant for a number of strategic decisions. Rural producers, export and import agents, companies in the food industry, suppliers, investors and the government are some of the players interested in this kind of information.

With accurate information about the status of different crops it is possible to develop commercial plans, to regulate agricultural products internal stocks, to make decisions on subsidies, and to draw strategies

for the negotiation of agricultural commodities in financial markets.

This work endeavours to combine two fields that have had a noticeable evolution in recent years, namely the research on multitemporal classification techniques using satellite imagery, and on plant phenology. In fact there are few reports on using phenological models to support the image classification process [1]. In this work Hidden Markov Models were used to relate the varying spectral response along crop cycles with plant phenology, for different crop classes.

Thus the general objective of this work was to evaluate the potential of Hidden Markov Models for crop classification from remote sensing temporal image sequences. Instead of relying on single date images, the methodology investigated in this work identifies different agricultural crops by analyzing the crop specific temporal profiles of spectral features over a sequence of medium resolution satellite images.

Section 2 shows the problem characterization, followed by a description of the Hidden Markov Model (section 3). The proposed methodology is presented in section 4 and a performance analysis is presented in section 5, followed by final comments and conclusions.

2. Problem characterization

2.1. Crops and their phenological cycles

The quantity of foliar area, phytomass volume and soil coverage temporal variations in a given area are determined by the planting and harvesting dates, and by the particular cycles of the different crops developed in it. The knowledge of these peculiarities provides the basis for understanding the spectral behaviors presented by different crop types in a certain period of the year.

2.1.1. Sugarcane. In São Paulo, Brazil, the sugarcane (SC) (*Saccharum* spp.) cultivation follows basically two cycles: one of 12 months (“one-year” sugarcane) and another of 18 months (“one-year-and-half” sugarcane). The one-year-and-half sugarcane is planted between January and March and the one-year sugarcane, between October and November. It is important to highlight that each sugarcane crop can be harvested during five or six consecutive agricultural cycles. For this reason the cycle is named “semi-perennial”, which is different from grain crops’ cycles, because of its duration, as well as of its phenological dynamics.

For areas where this crop is recently planted, a green mass of one-year-and-half sugarcane starts to completely cover the soil from October, when there is more heat and pluviometric precipitation; however, new areas of one-year sugarcane, should have full green coverage in April and May and then the green phytomass tends to increase its foliar area until the next harvesting period.

Each year the period of harvesting starts in April and ends in November, therefore, in a same satellite image it is possible to find: straw from harvested crop; recently planted sugarcane; as well as sugarcane in the growth phase and in the adult phase. It is also possible to find exposed soil, where the agricultural area is prepared for planting.

2.1.2. Short cycle crops (cereals). Soybean (*SB*) and corn (*CO*) are called “annual crops” or “short cycle crops”, once they can complete their phenological cycle in 110 to 140 days. They are planted, in general, in the end of October or in the beginning of November and they germinate about 10 days after being planted, beginning their vegetative growth and fully covering the soil surface around 60 days after the germination. In sequence, these crops reach the peak of green phytomass and then begin the grain filling process, when the quantity of green leaves starts to diminish, while the quantity of yellow leaves increases. They then dry out and fall, exposing again the soil background until the harvesting period.

2.1.3. Pasture. Pasture (PS) presents different phenological and spectral dynamics from the other crops mentioned above. These dynamics depend on the types of soil management used by cattlemen, however, in general, pastures are more dry and scarce between April and September. Their revigoration starts in the beginning of the rainy season along with, which increases the foliar area index and sustains the green vegetative vigor from November to March.

2.1.4. Other classes. Besides these crops, riparian forest (*RF*) was also considered in this work. Other classes of land cover are present in the study area: urban areas, roads, forest and water bodies. They appear as few, large segments that practically do not change through the image sequence, and for this reason, were not treated in this work.

3. Hidden Markov Models

A Hidden Markov Model (HMM) [2] represents a doubly embedded stochastic process. In an HMM, the observations (v_i) are regarded as symbols emitted by non observable states (S_i), following particular probabilistic functions, whereby the state sequence is a first order Markov Chain.

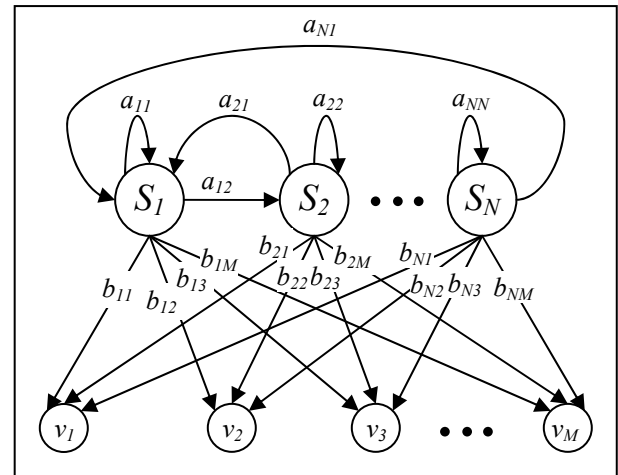


Figure 1. Example of a Hidden Markov Model ($S_i \rightarrow$ states, $v_k \rightarrow$ observation symbols, $a_{ij} \rightarrow$ state transition probability, $b_{ik} \rightarrow$ symbol emission probability).

An HMM is illustrated in Figure 1. N is the number of states in the model (the individual states are denoted as $\mathbf{S} = \{S_1, \dots, S_N\}$, and the state at time t as q_t) and M is the number of distinct observation symbols per state (the individual symbols are denoted as $\mathbf{V} = \{v_1, \dots, v_M\}$). A basic HMM consists of three sets of parameters:

- the symbol emission probabilities b_{jk} – the probability that symbol v_k is emitted by state S_j , i.e.

$$b_{jk} = P[v_k \text{ at } t | q_t = S_j], \quad 1 \leq j \leq N \text{ and } 1 \leq k \leq M$$
- the state transition probabilities a_{ij} – the probability of being in state S_j in the subsequent time instant given that the current state is S_i , i.e.

$$a_{ij} = P[q_{t+1} = S_j | q_t = S_i], \quad 1 \leq i, j \leq N$$

- c) the prior probability distribution π_i that the system is in a given state S_i at the initial time instant (not shown in the figure), i.e.

$$\pi_i = P[q_1 = S_i], \quad 1 \leq i \leq N.$$

If a state S_i can reach another state S_j , $a_{ij} > 0$, and if two states are not connected, $a_{ij} = 0$.

4. Methodology

4.1. General Model Description

In this work each crop class has a specific HMM. Phenological stages correspond to states and the observable symbols are the vectors comprising the digital numbers registered by the orbital sensor in each spectral band, along with the NDVI computed from bands 3 and 4 according to:

$$NDVI = \frac{band4 - band3}{band4 + band3}. \quad (1)$$

The basic HMM shown in Figure 2 was chosen to model the temporal behaviour of sugarcane, soybean and corn. The arrows illustrate how the states are temporally related. According to plant phenology, states *PP*, *GR*, *AD* and *PH* correspond respectively to stages Post-Harvesting, Prepared Soil, Growth phase and Adult phase respectively.

For pasture and riparian forest there is no significant change in the radiometric features during the period covered by the images in the available data set. This can be explained by the fact that these cultures are neither planted nor harvested and are indeed found in the target areas for a long time. Thus a specific HMM is devised for these classes having a single state *AD*, which in these cases correspond to the adult phase (Figure 3).

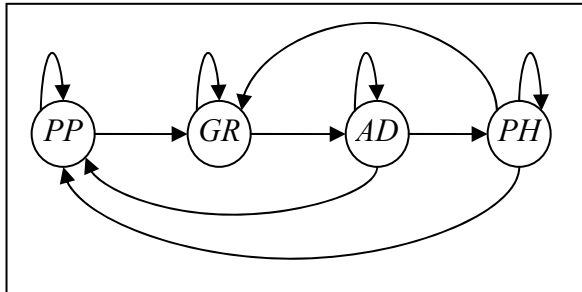


Figure 2. HMM used in this work for sugarcane, soybean and corn (*PP* = Prepared soil, *GR* = Growth, *AD* = Adult phase and *PH* = Post-harvesting).

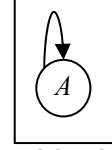


Figure 3. HMM used in this work for pasture and riparian forest (*AD* = Adult phase).

Even though pasture and riparian forest are actually not crop types, the term “crop” will be used hereafter to designate the set of all five classes to be recognized in our problem.

Tables 1 and 2 show examples of areas of each crop (and phenological stage) mentioned above in the satellite images (RGB composition). The sample crop areas occupy the center of the images on the tables.

4.2. Fitting the Model to the Application

The problem being considered in this work deviates in a number of ways from the basic HMM description presented in the preceding sections.

First, the symbol emission probabilities (b_{jk}) depend on seasonal effects that can not be fully compensated in the image pre-processing phase. Second, the prior probability distribution (π_i) is not constant along the year (see section 2). This happens because each crop has preferential months for planting conditioned mostly by the climate and by characteristics of the crop itself. In these months the crop is likely to be in the initial phenological states. As the time goes on this probability decreases, while the probability of being in the growth and adult stage increases. So, it is safe to affirm that the prior probabilities will vary according to these periods. Third, the basic model depicted in section 3 assumes that the symbols are emitted at a constant time rate. In most real applications not all images in a sequence acquired at a constant frequency are usable, mostly due to clouds over the target geographical area. It is also worth mentioning that the basic model shown in Figure 2 may also change for a larger interval between two consecutive images in the data set. For instance, a transition from *PP* to *AD* may become possible in these cases. Therefore, an HMM for our problem will have to consider distinct symbol emission probabilities, prior state probabilities, as well as state transition probability matrices for each pair of consecutive images in the available dataset.

Regarding the symbol emission probabilities, it is assumed throughout this paper that they have a Gaussian distribution. Hence, the emission probability density of a symbol \mathbf{x} (a vector consisting of the spectral bands and NDVI) will be given by:

$$p = \frac{1}{(2\pi)^{d/2} |\Sigma_{cs}|^{1/2}} \exp \left[-\frac{(x - \mu_{cs})^T \Sigma_{cs}^{-1} (x - \mu_{cs})}{2} \right] \quad (2)$$

where μ_{cs} , and Σ_{cs} denote respectively the mean vector, the covariance matrix for culture c and state s , and d is the dimension of \mathbf{x} .

4.3. Estimating parameters values for each model

Parameter estimation is performed for each crop type separately in the following way. State transition possibilities a_{ij} are estimated for each pair of consecutive images in the data set. Only segment samples of the crop class being modeled are considered in the next steps. First, an accumulator matrix is created with rows and columns corresponding to states respectively in the earlier and in the later date. Next, for all rows and columns, the element at row i and column j is incremented for each sample segment in state S_i and S_j respectively in the earlier and in the later epochs. Finally the accumulator matrix is normalized by dividing each of its elements by the sum across the corresponding row. The result is the estimate of the transition probability matrix.

The procedure to estimate the prior probability distribution π_i is similar. Again, only sample segments of the crop class being modeled at that date are considered. An accumulator vector is created having one element per state. The i th element is incremented for each training sample at state S_i . Finally, the accumulator vector is divided by the sum of its elements. The result is taken as the estimate of the prior state probability.

As mentioned before, the problem of estimating emission probability turns into estimating mean vectors and the covariance matrices for each image. Instead of taking segments as samples, we consider each pixel within a training segment as a training sample. This increases the amount of samples and provides more accurate estimates that could be otherwise a major performance degradation factor. Mean vector estimates are simply computed as the average of the training vectors representing the crop type and the phonological stage at that date. Similarly we use the sampled covariance [3] to estimate the covariance matrix.

4.4. Classification

Once the HMM have been established and their parameters estimated, the classification of an image segment is done in the following way. The segment is

represented at each date by a symbol vector comprising its average spectral values and NDVI observed at that date. The classifier computes for each model, the probability that the corresponding crop class emits the observed sequence of symbol vectors. The segment is assigned to the class whose model delivers the highest emission probability. A detailed description about how emission probabilities are computed in an HMM can be found in [4]. The algorithm that solves this problem also permits to infer for each model the most probable state at each point in time.

5. Performance analysis

5.1. Data Set

5.1.1. Study area. The study area corresponds to three cities in the State of São Paulo, Brazil: Ipuã, Guar e So Joaquim da Barra (inside a rectangle defined by the following coordinates: 20°16'30"S to 20°40'00"S x 47°37'36"W to 48°13'50"W), covering an area of 124.100ha (Figure 4). Agriculture is the main activity in this area. The main crops found are: sugarcane, soybeans and corn. This region has a plane to slightly undulated relief, a tropical climate with dry winter, with annual mean temperature of 22,9°C and annual mean precipitation of 1480mm.

Table 1: Soybeans, corn and sugarcane areas in the satellite image.



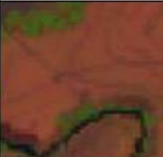
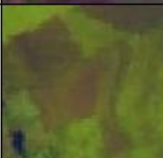






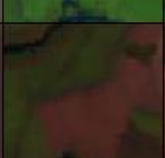

	Soybeans	Corn	Sugarcane
Prepared Soil			
Growth Phase			
Adult Phase			
Post-Harvesting			

Table 2: Pasture and riparian forest areas in the satellite image.



5.1.2. Image Sequence. The dataset contains a total of 12 images from the Landsat satellite, orbit/point WRS 220/74, from 2002 to 2004 (Table 3), from TM/Landsat-5, as well as from ETM+/Landsat-7 sensors [5]. Bands 1 to 5 and 7 were used in this work.

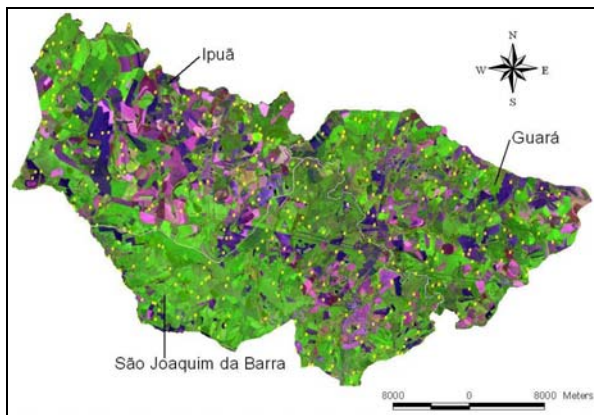


Figure 4. Study area in state of São Paulo, Brazil.

Table 3. Images available

	2002	2003	2004
January		(ETM+) 08/01/03	(TM) 19/01/04
February		(ETM+) 09/02/03 - 25/02/03	
April		(ETM+) 14/04/03 - 30/04/03	
May		(ETM+) 16/05/03	
July		(TM) 27/07/03	
August		(TM) 12/08/03	
September	(ETM+) 02/09/02		
October	(ETM+) 20/10/02	(TM) 15/10/03	

5.1.3. Image pre-processing. The Landsat images were in Geotiff format and for the geometric corrections, 13 control points gathered by GPS were used. The nearest neighbor resampling method was applied, because it well preserves the original image's radiometry [6].

A correction technique was applied to the multitemporal images to diminish atmospheric effects, once that the atmosphere, by its spread-spectrum, absorption and refraction phenomena, affects the radiance measured by the orbital sensors. The Dark-object subtraction technique, developed by Chavez [7], was applied.

As the same object may present distinct digital values in different acquisition dates' images, due to difference in the solar angles and to spread-spectrum effect, multitemporal groups of images must be radiometrically normalized. In this work, this process was done according to the methodology proposed by Gürtler et al.[8].

Classification algorithms are based on the spectral appearance of the objects being classified in images from different dates, so the grayscale values were converted to reflectance values, which have a physical meaning, in order to correctly represent the different objects and their conditions at the images' acquisition moments. This conversion was based on the methodology proposed by Luiz et al. [9].

5.1.4. Image Segmentation. Each image in the data set was individually segmented using the watershed based method detailed in [10]. The values of the segmentation parameters were selected after a series of trials-and-errors, till the outcome was visually consistent considering all images of the data set. It is worth mentioning that the segment contours generally do not match through the image sequence, although the images have been previously co-registered.

All segmentation algorithms involve parameters that must be adjusted so as to obtain the meaningful image objects in the particular application. This task is usually done (as in our case) by a photo-interpreter, who in this way, ultimately adapts the selected segmentation algorithm to the application needs. For this reason, in our target application, the segmentation method ended up not influencing significantly the final result. A region growing segmentation algorithm, for instance, would probably be also appropriate in this case.

5.1.5. Reference Data. A total of 316 reference image locations were selected in the study area and two experts classified them visually in each image indicating the crop class and corresponding phenological stage. First, the two experts worked individually, and each one of them interpreted all 316 reference segments. Then, both experts worked together, in order to conceive a consensus classification result. To accomplish this task, both results were compared, and when they differed, the

experts considered again the multitemporal image sequences, and decided on a final classification. Apart from that, there was also information gathered in two field works, taken place in March/2003 and August 2003, to aid the experts.

The segments enclosing each of the 316 image locations in each date were used to build the training and testing sets. Each segment was split in two parts with roughly the same amount of pixels. One of these parts was used for training and the other part for testing.

5.2. Experiment Results

For all the experiments there was a total of 7 attributes (6 spectral bands and NDVI).

The first experiment aimed at determining the set of n attributes which resulted in the best classification, for n varying from 1 to 7.

Table 4. Optimum set of n spectral attribute and corresponding classification performance.

Number of spectral features	Optimum set of spectral features	Average Class Accuracy (%)
1	NDVI	84,47
2	4 NDVI	86,51
3	4 5 7	89,92
4	2 3 5 NDVI	89,95
5	1 2 4 5 7	91,89
6	1 2 3 4 5 7	92,72
7	all	91,22

In this experiment we determined by exhaustive search the optimum set of n spectral features, for n varying between 1 and 7. Table 4 summarizes the results. With the NDVI alone the HMM method achieves approximately 84%. By adding the remaining 6 spectral features to the NDVI the performance increases about modest 7%. Interestingly, the highest performance reported in Table 4 was achieved with 6 features. These features (bands 1, 2, 3, 4, 5 and 7) were the ones used in all the following experiments.

A second experiment aimed at identifying crop types, as well as the phenological stages during the dates in the test sequences.

For this experiment, there was only one crop type per sequence.

Table 5 and 6 show the accuracies and the confusion matrix for crop class classification respectively. These tables show that the method

performed well for all crop classes, i.e. with 93% average class accuracy.

Table 5. Crop classification accuracy.

Class Accuracy (crops)	
Crops	Rates (%)
Soybeans (SB)	95
Corn (CO)	90
Sugarcane (SC)	96
Pasture (PS)	92
Riparian forest (RF)	91
Overall accuracy:	95
Average class accuracy:	93

Table 6. Crop classification confusion matrix.

Confusion matrix (crops)					
	SB	CO	SC	PS	RF
SB	95	0	4	1	0
CO	2	27	1	0	0
SC	2	1	191	4	0
PS	0	0	1	23	1
RF	0	0	2	1	29

Table 7, 8 and 9 refer to stage identification, considering all the sequences available and only sequences correctly identified by the HMM classification model separately.

Even if recognizing the phenological stage is not the main interest, looking at Table 7 helps understanding how the model behaves. The phenological stages were also well identified, with the exception of the Growth phase.

This can be explained by the temporal evolution of the crops throughout the phenological cycle.

During the prepared soil, adult and post-harvesting phases, there is no significant changes in the crop's spectral response. However, the spectral response of the growth phase is continuously changing from prepared soil to adult-phase. So its spectral response could be close to the response of these other two stages, or something in between, which can lead to misclassification.

Tables 8 and 9 confirm this interpretation, as the confusion matrices shows that the growth phase was often misclassified as adult phase and prepared soil.

This varying behaviour along the time can not be properly modeled by a single Gaussian distribution of the spectral response vectors. This reasoning suggests that the growth phase could be better recognized by looking at the temporal derivative instead of the absolute values of the spectral response.

When comparing both columns in Table 7, it is clear that the results for all sequences available and for

only the sequences correctly identified do not diverge too much. This was expected, once the crop recognition rates were high, 93% (as shown in table 5).

Table 7. State classification accuracy.

Class Accuracy (states) (%)		
States	Sequences correctly identified	All sequences available
Post-harvesting (PH)	97	77
Prepared soil (PP)	80	55
Growth phase (GR)	57	95
Adult phase (AD)	96	96
Overall accuracy:	88	87
Average class accuracy:	82	81

Table 8. State classification confusion matrix considering only the sequences correctly identified.

Confusion matrix (states)				
	PP	GR	AD	PH
PP	437	37	31	39
GR	54	229	111	7
AD	15	64	1847	4
PH	2	1	5	235

Table 9. State classification confusion matrix considering only the sequences correctly identified.

Confusion matrix (states)				
	PP	GR	AD	PH
PP	452	39	52	41
GR	56	231	129	7
AD	20	78	1914	6
PH	2	1	8	239

A final experiment was carried out for comparison purposes. The aim was to assess how a single-date approach would perform under similar conditions. It can be demonstrated that for single-date sequences the HMM devised in this work turns into a bank of maximum likelihood classifiers, where each crop type is represented by as many classes as the number of phenological stages.

Table 10 shows the results. When considering samples in any phenological stage, the average class accuracy was about 70%.

The experiment was repeated 3 times taking away samples respectively in stages PP, PH, and GR, which are in this order less characteristic of the crop classes. The best result obtained in these experiments was 74%.

Table 10. Performance for single-date sequences.

Phenological states present in the sample	Average class accuracy (%)
PP, GR, AD, PH	70
GR, AD, PH	69
GR, AD	71
AD	74

Recalling the performance reported in Table 5, it becomes clear the superiority of multivariate approach in relation to the single-date counterpart.

It should be expected that the multitemporal would be better than the monotemporal approach. In fact the purpose of this experiment was merely to evaluate numerically this superiority.

6. Conclusion

This work evaluated the potential of Hidden Markov Models for crop classification. The experimental evaluation, based on sequence of 12 Landsat images for 5 crop types, indicated a remarkable superiority of the HMM-based method over a monotemporal maximum likelihood classification approach.

An analysis of the experimental results revealed that the performance of HMM-based classifier was severely impacted by the scarcity of training samples of some crop types. Hence, even better results could have been achieved if a more representative training set were available.

The HMM approach also performed well to recognize phenological stages. The exception was the growth-phase, which was frequently confused with prepared-soil and adult-phase. This observation suggests that symbol vectors used to characterize the growth-phase should take into account not only the absolute spectral values but also their variation along the time.

For this work, only sequences of data associated to one crop type were considered. It is planned for future an analysis of the behaviour of the method considering sequences containing samples of more than one crop type.

References

- [1] L. Aurdal, R. B. Huseby, L. Eikvil, R. Solberg, "Use of hidden Markov models and phenology for multitemporal satellite image classification: applications to mountain vegetation classification", *International Workshop on the Analysis of Multi-Temporal Remote Sensing Images*, 2005, pp.220-224.

- [2] H. Bunkle, T. Caelli, Hidden Markov Models – applications in computer vision, *World Scientific*, 2001.
- [3] R. A. Johnson, D. W. Wichern, *Applied Multivariate Statistical Analysis*, Prentice Hall, New Jersey, 1998.
- [4] L. R. Rabiner, “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition”, *Proceedings of the IEEE*, Vol. 77, n.2, 1989, pp. 257-286.
- [5] I. D. Sanches, “Sensoriamento remoto para o levantamento espectro-temporal e estimativa de áreas de culturas agrícolas”, Dissertação de mestrado, INPE, São José dos Campos, 2004, 172p.
- [6] P. M. Mather, *Computer processing of remotely-sensed images: an introduction*, Chichester: John Wiley & Sons, 1993, 352 p.
- [7] P. S. Chavez Jr., “An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data”. *Remote Sensing of Environment*, Vol.24, n.9, 1988, pp. 459-479.
- [8] S. Gürtler, “Estimativa da área agrícola a partir de sensoriamento remoto e banco de pixels amostrais”. Dissertação de Mestrado, 2003, 179 p. sid.inpe.br/jeferson/2003/06.02.07.29 (accessed june 2003.)
- [9] A. J. B. Luiz, S. Gürtler, J. M. Gleriani, J. C. N. Epiphânio, R. C. Campos, “Reflectância a partir do número digital de imagens ETM”. *Simpósio Brasileiro de Sensoriamento Remoto*, 11., Belo Horizonte. Anais. São José dos Campos, Brasil: INPE, 2003, pp. 2071-2078.
- [10] G. L. Mota, R. Q. Feitosa, H. L. C. Coutinho, C. E. Liedtke, S. Müller, K. Pakzad, M. S. P. Meirelles, “Multitemporal fuzzy classification model based on class transition possibilities”. *ISPRS Journal of Photogrammetry and Remote Sensing*. Vol.1, 2007, pp.1-2.