A Comparison of Statistical and Fuzzy Approaches for Cascade Classification of Multitemporal Remote Sensing images

R. Q. Feitosa[a,c*], L.M.T. Alvarado[a], T. Klinger[b], G.A.O.P. Costa[a], C. Heipke[b]
[a] Department of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro, Brazil,
[b] Institute for Photogrammetry and GeoInformation, Leibniz Universität Hannover, Germany,
[c] Department of Computer Engineering, Rio de Janeiro State University, Brazil
Corresponding author: raul@ele.puc-rio.br, +55 21 35271212

Abstract: This work empirically compares a probabilistic and a fuzzy technique for cascade multitemporal classification based respectively on Hidden Markov Models and on Fuzzy Markov Chains. The analysis follows the OBIA paradigm and is conducted upon two bi-temporal data sets. The performance of variants of each approach relative to maximum likelihood mono-temporal counterparts is assessed. The experiments simulate distinct operational conditions. In particular, the robustness of each method against outliers in the training sets is investigated. Further, the study assesses the improvement resulting from incorporating the prior-knowledge concerning the admissible class-transitions in the target site, within the underlying time frame, into the classification model.
Keywords: Multitemporal Analysis, Hidden Markov Models, Fuzzy Markov Chains

1. Introduction

Differently from multitemporal methods designed to detect changes having occurred in an area over time, cascade-classification approaches explore the correlation contained in the temporal data sets in order to classify one or more images in the multitemporal sequence. Few cascade methods have been proposed so far, including statistical as well as fuzzy approaches. Whereas statistical methods benefit from a well-established theoretical foundation, fuzzy techniques can be more naturally incorporated into Object-Based Image Analysis (OBIA) interpretation models, which can rely on fuzzy rules for knowledge representation. To our knowledge no extensive comparison between these two approaches has been reported in the literature so far.

This work addresses this issue and aims at experimentally comparing a probabilistic and a fuzzy approach for cascade multitemporal classification under distinct operational conditions. The probabilistic method is based on Hidden Markov Models (HMM), whereas the fuzzy approach relies on Fuzzy Markov Chains (FMC).

2. Multitemporal Schemes

2.1. Probabilistic Approach – Hidden Markov Models

The probabilistic multitemporal model used in this analysis is based upon HMM as described in (Leite et al., 2011). Let $\mathbf{\Omega} = \{\omega_1, \omega_2, \dots, \omega_n\}$ be a set of n distinguishable LC/LU classes. For any given image object the hidden, not directly observable variables ${}^{A}\omega$ and ${}^{B}\omega$ denote the class labels, i.e., ${}^{A}\omega, {}^{B}\omega \in \mathbf{\Omega}$, at times t_A and t_B, whereas ${}^{A}\mathbf{x}$ and ${}^{B}\mathbf{x}$ indicate the observed data, actually the feature values of those image objects, also at t_A and t_B.

Published online May 2014

Without loss of generality we assume that the image acquired at t_B, hereafter referred as the *target image*, is the one to be classified, whereas the *supplementary image* acquired at time t_A only provides additional data from the same geographical object.

HMM provides a way to derive the expression for the likelihood of each possible sequence of classes given the observed data, from which the most likely class ${}^B\Omega_p$ at the *target date* can be inferred as:

$$ {}^B\Omega_p = \arg\max_{{}^B\omega} \left[p\left({}^B\mathbf{x} \mid {}^B\omega\right) \sum_{{}^A\omega} p\left({}^A\omega\right) p\left({}^A\mathbf{x} \mid {}^A\omega\right) p\left({}^B\omega \mid {}^A\omega\right) \right] \qquad (1) $$

where $p({}^A\mathbf{x}|{}^A\omega)$ and $p({}^B\mathbf{x}|{}^B\omega)$ represent the likelihood of observed features for classes ${}^A\omega$ and ${}^B\omega$ respectively, $p({}^A\omega)$ denotes the prior probability of class ${}^A\omega$ at t_A and $p({}^B\omega|{}^A\omega)$ the transition probability from class ${}^A\omega$ at t_A to class ${}^B\omega$ at t_B , for ${}^A\omega$, ${}^B\omega \in \mathbf{\Omega}$..

2.2. Fuzzy Approach – Fuzzy Markov Chains

The fuzzy multitemporal approach considered in this comparison is described in details in (Feitosa et al., 2011). Based on the features ${}^A\mathbf{x}$ and ${}^B\mathbf{x}$ measured at times t_A and t_B, two monotemporal classifiers deliver membership values $m({}^A\mathbf{x}|\,{}^A\omega)$ and $m({}^B\mathbf{x}|\,{}^B\omega)$ respectively for classes ${}^A\omega$ and ${}^B\omega$.

The outcome of the monotemporal classifier at t_A is updated to time t_B on the basis of a binary relation defined by a transition matrix $\boldsymbol{\tau}=\{\tau_{{}^A\omega{}^B\omega}\}$, where $\tau_{{}^A\omega{}^B\omega}$ stands for the possibility that an image object belongs to class ${}^A\omega$ at time t_A and to class ${}^B\omega$ at time t_B, for ${}^A\omega,{}^B\omega \in \mathbf{\Omega}$. The membership value $m({}^B\mathbf{x}|\,{}^B\omega)$ and the corresponding updated versions of $m({}^A\mathbf{x}|\,{}^A\omega)$ are combined by an aggregation function yielding the final consensus membership for the image object being classified at the *target date*. Following (Feitosa et al. 2011) the class label ${}^B\Omega_f$ assigned by the Fuzzy Markov classifier is given by

$$ {}^B\Omega_f = \arg\max_{{}^B\omega} \left\{ w\,{}^B m({}^B\mathbf{x} \mid {}^B\omega) + (1-w) \max_{{}^A\omega} \left[\tau_{{}^A\omega{}^B\omega} m({}^A\mathbf{x} \mid {}^A\omega) \right] \right\} \qquad (2) $$

This model involves up to n^2+1 parameters, namely the weight w and the transition possibilities $\tau_{{}^A\omega{}^B\omega}$, which must be estimated based on a set of training samples. $\tau_{{}^A\omega{}^B\omega}$ is set to 0 (zero) for class transitions not represented in the training set. This may reduce considerably the number of parameters to estimate. As in (Feitosa et al., 2011), a genetic algorithm has been used in our experiments for parameter estimation.

3. Experimental Analysis

3.1. Data sets

The first test site corresponds to a 14.4 km^2 area, situated in the city of Rio de Janeiro. Two pan-sharpened, orthorectified IKONOS II images were used in the experiments acquired in 2008 and 2009. The images were segmented using the multiresolution segmentation algorithm available in the Definiens Developer 7® software. All segments generated for the two images were duly, visually classified by specialists. The land-use classes considered in this work are listed in the leftmost column of Table 1(a). After classification, all segments from each class in 2008 were merged to the adjacent segments of the same class, generating

larger area segments. Then, only the segments from 2009 that fell completely inside the large 2008 segments (generated by the merging procedure) were selected. The features computed for each segment were: mean values of the four spectral bands and the textural entropy feature for all bands in all directions.
The second test site is situated in the Municipality of Alcinópolis in the State of Mato Grosso do Sul, Brazil, and encompasses an area of 950 km^2. Bands 5, 4 and 3 of two co-registered LANDSAT 7 images, acquired during the dry seasons of 1999 and 2001 were used in the experiments. The bands of both images were stacked forming an artificial six-band image. The resulting image was segmented using a variation of the watershed algorithm, which is described in detail in (Mota et al. 2007). The LC/LU classes present in the test area are listed in Table 1(b). The reference classification was produced visually with aid of a video image sequence produced in October 2001, a drainage map, and a digital elevation model.

Table 1. Class transitions in Rio (a) and Alcinópolis (b) test sites (units are segments)

	2009			
2008	Rock	Field	Urban	Trees
Rock	188	10		
Field	11	200	66	153
Urban		9	200	200
Trees		200	200	200

(a)

	2001				
1999	Soil	Riparian	Pasture	Water	Savannah
Soil	38		38		
Riparian		56			
Pasture	9		200		
Water				26	
Savannah					126

(b)

3.2. Experimental Procedure

For each class transition half of the samples were randomly selected for training and half for testing. The results in the next section are averages over 100 runs, each time with a different random selection of training and test sets. For HMM a Gaussian distribution was assumed for $p({}^{B}\mathbf{x}|\ {}^{B}\omega)$. It was further assumed that the ground truth in the *supplementary date* is known. Thus, for HMM $p({}^{A}\omega)\ p({}^{A}\mathbf{x}|{}^{A}\omega)$ was set to 1 (one), if ${}^{A}\omega$ was the true class, and to 0 (zero) otherwise. Similarly, for FMC $m({}^{A}\mathbf{x}|{}^{A}\omega)$ was set to 1 (one) or 0 (zero), if ${}^{A}\omega$ was either the true or the false class.

3.3. Results

The results of the experiments are expressed hereafter in terms of reduction of classification errors obtained by each multitemporal method relative to the corresponding monotemporal maximum likelihood (ML) classifier, whereby Gaussian class probability distributions were assumed. The first results, shown in Figure 1.a, correspond to ideal operational conditions, specifically, that (i) the training samples have no outliers, and (ii) each class transition occurs with the same relative frequency in the training and in the test sets. Figure 1.b shows the results of experiments in which none of these conditions holds. 20% randomly selected training samples were assigned to some incorrect label. In consequence, the overall accuracy of all classifier decreased; in particular, the ability of the multitemporal methods to reduce misclassification deteriorated, noticeably for the fuzzy approach. The impact of outliers can be reduced by exploiting prior knowledge concerning the admissible class transitions, as shown in Figure 1.c. Finally, Figure 1.d presents the results of experiments having no

Issue
Vo3, No2S, 2014

outliers, i.e., condition (i) holds, for training sets containing the same number of samples for each admissible class transition, i.e., condition (ii) does not hold. Clearly, in these experiments the estimation of class transition probabilities/possibilities becomes less accurate. The results shown in Figure 1.d suggest that, under such circumstances, the probabilistic approach is more robust than the fuzzy counterpart.

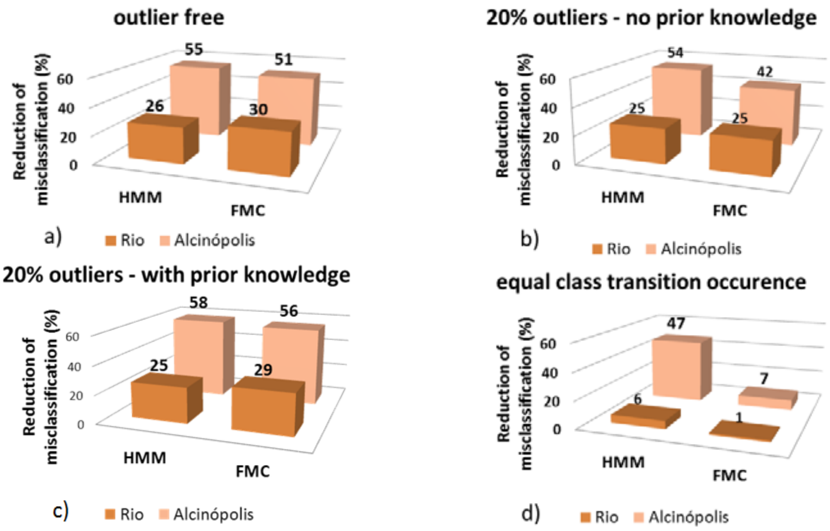

Figure 1. Reduction of classification errors for a. outlier free and consistent frequency of class transitions in the training and test sets, b. 20% outliers, c. 20% outlier with the incorporation of prior knowledge regarding possible class transitions, d. outlier free, but inconsistent occurrence of class transitions in the training set.

4. Conclusions

The results presented in this work indicate that the probabilistic as well the fuzzy multitemporal cascade classification methods may substantially reduce the classification errors of the monotemporal counterpart. The fuzzy approach was more sensitive to outliers than the probabilistic method, although the exploitation of prior knowledge w.r.t. admissible class transitions may counterbalance this weakness. Moreover, the experiments revealed that inaccuracies regarding the relative occurrence of class transitions may degrade substantially the performance of both multitemporal methods, specially the one based on Fuzzy Markov Chains.

Acknowledgments

The authors thank the Marie Curie International Research Staff Exchange Scheme of the European Commission for the support to this work as part of the TOLOMEO project.

References

Costa G.A.O.P., Feitosa R.Q., 2014, A generalized fuzzy Markov chain-based model for classification of remote-sensing multitemporal images. International Journal of Remote Sensing, Volume 35, Issue 1: 341-364.

Leite P.B.C., Feitosa R.Q., Formaggio A.R., Costa G.A.O.P., Pakzad K., 2011, Hidden Markov Models for Crop Recognition in Remote Sensing Image Sequences. Pattern Recognition Letters, Volume 32, Issue 1: 19-26.

Feitosa R.Q., Costa G.A.O.P., Mota G.L.A., Feijó B., 2011, Modeling alternatives for Fuzzy Markov chain-based classification of multitemporal remote sensingdata. Pattern Recognition Letters, Volume 32, Issue 7: 927–940.

Mota G.L.A., Feitosa R.Q., Coutinho H.L.C., Liedtke C.E., Müller S., Pakzad K., Meirelles M.S.P., 2007, Multitemporal Fuzzy Classification Model Based on Class Transition Possibilities. ISPRS Journal of Photogrammetry & Remote Sensing, Volume 62: 186-200.