Classification of croplands through integration of remote sensing, GIS, and historical database

M. J. ORTIZ

Imagem Sensoriamento Remoto, São José dos Campos (SP), Brazil

A. R. FORMAGGIO and J. C. N. EPIPHANIO

Instituto Nacional de Pesquisas Espaciais (INPE), Av. dos Astronautas, 1758, 12201-970 — São José dos Campos (SP), Brazil

(Received 29 June 1995; in final form 5 March 1996)

Abstract. This work presents a methodology to classify croplands using a multitemporal/historical dataset of images and ground ancillary data referring to three consecutive years. An image processing/geographic information system as well as a database management system (DBMS) were used to make the integration of these multisource data. In order to evaluate the usefulness of a database for crop classification, the area under study was digitally classified by two groups of interpreters, using two methodologies: (a) the proposed methodology using maximum likelihood classification assisted by an historical/multisource database, and (b) a conventional maximum likelihood classification only. Both results were compared using the Kappa statistics. The indices to both the proposed and the conventional digital classification methodologies were 0·669 (very good) and 0·472 (good), respectively. The use of the database rendered an improvement over the conventional digital classification. Furthermore, along with this study some problems related to multisource data integration are discussed.

1. Introduction

Crop forecasting is a task with the objective of anticipating the hectarage and crop yield in a defined region of interest (MacDonald and Hall 1980). This information is used by those institutions responsible for the infrastructure, transportation, sales, storing and distribution of crop production. A basic step for crop forecasting is the crop classification. As the satellite remote sensing data have been available since early 1970s and due to their characteristics, these data are being employed towards to the improvement of crop forecasting.

Historically crop forecasting has been made using statistical procedures with samples obtained in the field as input data. This kind of sampling procedure is both expensive and time consuming. In this framework, remote sensing techniques have been shown to be useful and this has been demonstrated by some major programs, like LACIE (Large Area Crop Inventory Experiment) and AgRISTARS (Agricultural Resources Inventories Through Aerospace Remote Sensing), developed in the United States (Cracknell and Hayes 1991). In addition, according to Allen and Hanuschak (1988), at the National Agricultural Statistics Service (NASS/USA), since the 1950s, remote sensing data has been used (initially using aerial photographs and more recently Landsat data) to build up sampling panels from regions of interest.

Similar to the experience developed at the United States Department of

Agriculture (USDA), the General Directorate for the Agriculture of the Commission of European Communities (CEC) and the Statistical Service of the European Communities (EUROSTAT) at the Joint Research Center (JRC), have developed the MARS (Monitoring Agriculture with Remote Sensing) program. According to Boissezon and Sharman (1993), the objective of this program is to provide up-to-date agricultural information to those governmental institutions responsible to direct the European agricultural policies. This program demonstrates the importance on the use of results from previous years as an useful information for classifying agricultural targets in remote sensing images.

The development of new processing techniques, and specially of Geographic Information Techniques (GIS), is opening new possibilities for the use of spectral, spatial as well as temporal information from orbital sensor systems, stored in geocoded databases. These databases can also store historical data from the areas of interest, thus allowing an integration with other available information of interest. As a consequence, and taking into account manipulations made in the system, it is possible to retrieve almost instantaneously more consistent and powerful information.

There is an increased interest on introducing GIS and multitemporal information into systems developed for agricultural evaluation purposes. Middelkoop and Janssen (1991) included temporal and expert relationships in their expert classification system and observed an improvement over the traditional classification, mainly in areas where there were high spectral overlaps among classes. The use of different kinds of ancillary data for improving land use classification is also a commonplace. Janssen *et al.* (1990) observed an improvement of more than 12 per cent in overall classification accuracy of agricultural fields when integrating topographical data from GIS into their Landsat TM classification.

In this direction, the basic premise of this investigation is that when one is classifying an actual image and there are historical remote sensing and ancillary data stored in a database, and provided that all this information is promptly available to the analyst, then the result of classification would be improved. Thus, the objective of this study is to present a methodology to identify crops by analysing historical Landsat TM data and a database (ancillary and historical ground data) in a GIS framework. Besides, a performance evaluation of the methodology is accomplished.

2. Study area and methodology

The study area is localized in the municipality of Guaíra, São Paulo State (Brazil), corresponding to a square of 15×15 km of the Guaíra topographic map (IBGE 1972), with geographical coordinates $20° 26' 07''$ S to $20° 17' 24''$ S and $48° 25' 02''$ W to $48° 14' 58''$ W (Figure 1).

According to Oliveira and Prado (1991) the predominant soil types in this region are two Oxisols derived from basic rocks. Agriculture is the main land use in the region, where the main crop season is in the rainy season (October to March). In the dry season, the focus of this study, the area is partially occupied by crops irrigated by central pivots and self-propelled systems. The main crops are beans, tomato, soybeans, sorghum, corn, onions and potatoes as well as other crops of secondary occurrence. Also, there are pastures, sugar cane, some patches of secondary forest, some small reforestation stands and some areas planted with *Hevea brasiliensis* (the rubber tree).

The Landsat TM images used were from the following dates: 22 June 1988, 25 August 1988, 24 May 1989, 12 August 1989, 12 June 1990, 30 July 1990,

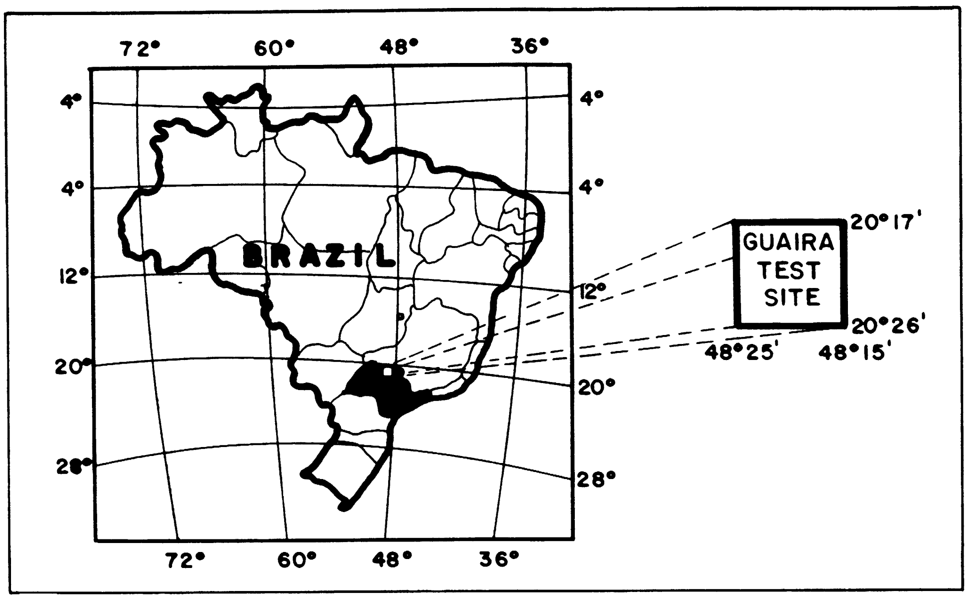

Figure 1. Guaíra test site at São Paulo State, Brazil.

16 September 1990, 17 July 1991 and 18 August 1991, mainly because these were the best available cloud-free scenes for the winter season.

The data manipulation and analysis were done using an image analysis system, a geographic information system both developed by Brazilian National Institute for Space Research (INPE) (Imagem 1993 a, 1993 b), and a database management system (DBMS). By integrating these three systems it is possible to display any stored image in a GIS framework, to inquire the database, and then to visualize spatially the output, as a support to the process of classifying agricultural areas.

Figure 2 depicts schematically the main procedures of this investigation and shows that three main methodological steps can be considered. The first one was the acquisition of current and historical ground information, during field campaigns, to obtain those data of the lots to be studied and to generate a tabular database. The second step was the data manipulation in the GIS/DBMS environments to create both a georeferenced database and a tabular database. In the last step two classifications were done: one using conventional digital classification and other using conventional digital classification assisted by the georeferenced and tabular databases in a GIS/DBMS environment. Afterwards, a statistical analysis was performed to obtain quantitative parameters for comparing the results obtained with the use of the georeferenced and tabular databases versus the results obtained without the use of these databases via Kappa statistics (table 1).

At the image analysis system the following preliminary procedures were done with the nine multitemporal Landsat TM scenes available: atmospheric correction, gray level transformation to reflectance values (Markham and Barker 1986), registration of image versus topographic map and registration of image versus image. The resultant images were transferred to the GIS environment at UTM (Universal Transverse Mercator) projection and at 1 : 50 000 scale.

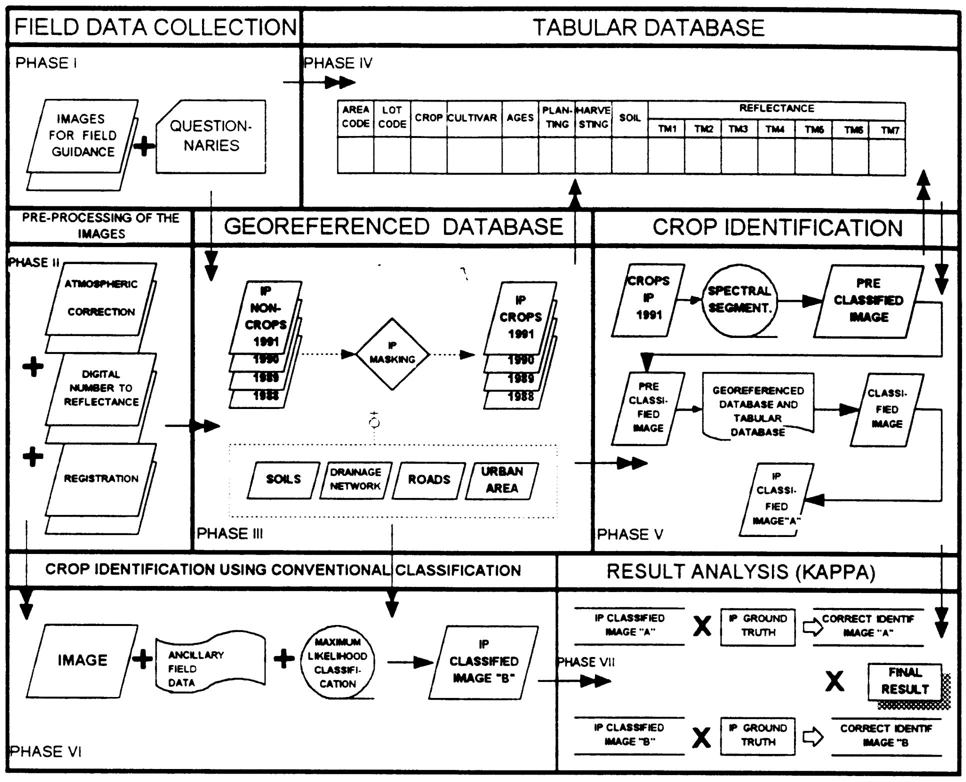

Figure 2. Main activities developed in the investigation.

The georeferenced database was elaborated using two categories of data: 'polygons' and 'spectral images'. The category 'polygons' included the information planes (IPs) whose information was originated from the topographic map, the soil map, and the results of digital image classifications. Table 2 shows the main IPs of the category polygons and their descriptions. The category spectral images included the nine Landsat scenes. A set of these images was generated including only those lots of interest and that were selected for the further classification while the remaining areas were cut off to avoid misclassification.

The tabular database was generated using all the information obtained in the field campaigns (e.g., crop types and cultivars, crop age in each used scene date, planting date, harvesting date, soil type, etc.). Furthermore this database was fed

Table 1. Classification quality associated to a 'Kappa' statistics value.

Kappa	Quality
<0·00	Worst
0·00–0·20	Poor
0·20–0·40	Reasonable
0·40–0·60	Good
0·60–0·80	Very good
0·80–1·00	Excellent

Table 2. Information planes of the category polygons that were included in the georeferenced database.

Name of plane	Auxiliary planes
Soil	IP with soils of module
EDRE	IP with drainage of module
Roads	IP with roads of module
City	IP with city of Guaíra
	Execution planes
IP88	IP with fields planted during winter 1988
IP89	IP with fields planted during winter 1989
IP90	IP with fields planted during winter 1990
IPPIVOT	IP with fields planted during all years analysed
SOILR	IP presenting soils of selected fields
IP91	IP with fields planted during winter 1991 – 'ground data'

with the six reflective TM bands reflectance values corresponding to each field and each date. The tabular database was connected to the GIS by a function that associated a record of the tabular database to a given polygon, through a label. This identifier must be unique. So it is not possible to have one record associated to two or more polygons and vice versa.

These databases (georeferenced plus tabular databases) contained all information available for the years 1988, 1989 and 1990 and will be named GTD. GTD had a structure of 120 fields and 70 records for each field. Each field was the intersection of those three years—the smallest area element with an unique information, a necessary condition to perform the integration of databases from different years. During the classification process the interpreter could access GTD through search expressions (queries). Queries were useful to provide evidence in order to decide if a spectral behaviour could be associated to a given crop or to another in a given lot. Once defined what was the crop in some specific lot it could be used as a training area in the digital classification.

In order to test the methodology two groups of analysts were selected: one using GTB (historical images and ancillary data) as a help for the digital image classification (*WI*th databases group), and a second group that did not use the databases (*W*ith*O*ut databases group). The image to be classified was the winter 18 August 1991 image, and the field work was close to this date. Three experienced image interpreters were invited for the first group (named here as WI_A, WI_B and WI_C). The methodology and the databases were presented to these interpreters and they were informed on all the possibilities offered by the databases during the classification. As a result, three 1991 classified images were generated. For the second group three different interpreters were selected (named here as WO_D, WO_E, and WO_F). The only information provided to this last group was some training areas from the previous year and which kind of crops were planted in 1991. They did a conventional digital image classification and, thus, three other 1991 classified images were generated.

The next step was to cross the results of the classifications of the WI and WO groups with the 1991 ground data in order to evaluate quantitatively and validate the proposed methodology. By using the results of these crossings, it was possible to evaluate the performance of each interpreter and each methodology of classification. To perform this analysis both the error matrices and the 'Kappa' statistics were applied. The objective of Kappa statistics is to evaluate the degree of similarity between maps, such as between a classified thematic image and a reference map. The values resulting from the use of this method may vary from below 'zero' to 'one'. Values 'lower than zero' indicate no similarity (total independence) and 'one' indicates equality (total dependency). Landis and Koch (1977) grouped the 'Kappa' values as they are shown in table 1, thus allowing the evaluation of the classification quality. Congalton and Mead (1983) and Congalton *et al.* (1983) used the Kappa statistics to improve the evaluation of the results of a thematic classification.

3. Results and discussion

3.1. *Tabular database versus 'farmer cropping tradition' concept*

After the elaboration and manipulation of the tabular database, the first result was that it was possible to confirm the concept that most of the farmers presented the characteristic of cropping the same crops year after year. This concept was named here as 'farmer cropping tradition', and it strengthened our hypothesis that historical information may be used to improve the procedures of identification of agricultural targets in remote sensing products. In this case we are coupling the multi-years or historical information concept with digital image analysis for classifying current images. Another useful concept was the 'regional cropping tradition'. This concept means that in the function of soil types, climate, available infrastructure, and social-economic factors of a given region a tendency is shown to have the same suite of crops that are cultivated year after year, at least for some years.

Both concepts were useful for identifying the agricultural targets of interest. The crop identification in a specific field was done by inspecting GTD for that field of interest in different years. Areas planted with the same crop in successive years create some interpretation patterns for the interpreter. These spectral patterns were helpful to identify agricultural targets during the interpretation process in the digital image analysis. However, there were a few cases whose patterns were not followed.

The GTD allowed an individual analysis for each farmer and the outline of the agricultural history of each area. This history also allowed the study of the agricultural evolution of each lot, improving the performance of the identification of cultures. This agrees with the findings of Boissezon and Sharman (1993) who emphasized the utility of a database to identify crops. This can be associated to a growing experience of the photo-interpreters. If the photo-interpreters are always the same, the 'human visual interpretation memory' is kept as a further and powerful element of analysis. This way we optimized the multitemporal and multi-year dimensions of the satellite remote sensing in a GIS to function as a ready available memory for the photo-interpreter.

Even though the use of the databases had been efficient, it became evident that a more complete database with a longer time record and with more information (e.g., statistical data of the previous harvests, pluviometric data, changes of markets due to prices) would be useful. Another aspect to be considered is that the construction of the database should reflect as closely as possible the instantaneous picture of the field during each satellite overpass. Raafat *et al.* (1991) also stressed these aspects, showing the importance of the use of databases that have a good capability to store, to manipulate and to retrieve the largest amount of data of multitemporal origin.

Despite some problems for the integration of multisource data (image analysis system, DBMS, and GIS) the benefits obtained with this integration should be emphasized, as reported by Wang (1991).

3.2. *Classification assisted by databases*

Each interpreter of the WI group (using GTD) generated a classified image. These images were moved to the GIS and transformed into IPs that were afterwards compared with ground data. The three WI interpreters used all the facilities available in the system, performing initially a spectral separation (nonsupervised classification algorithm) of the 1991 scene. This was important so that the interpreters could have a feeling for some spectral differences that cannot be perceived by the human eye. We used TM3, TM4, TM5 and TM7 spectral bands, which are representatives of the main reflective spectral portions for crop discrimination. This procedure often allowed a rough spectral discrimination among different crops. Next, several queries could be applied to the databases in order to label the classes of the nonsupervised classification. This allowed improvement for the acquisition of training areas to classify thematically the 1991 image using supervised classification algorithms.

Table 3 presents the crops and their respective areas identified by the interpreters who were invited to test the methodology. One can observe that the use of the databases uniformed the classification performed by the interpreters for various targets. This is clear for areas of beans, bare soils and residuals of crops.

As the databases contained historical information, they could help the interpretation of some targets even if these targets were not very clear for classification in the current image. For instance, the databases helped the interpretation of the 'soybean' class in the agricultural 1991 season (even though soybean was not included in the file for the training of the interpreters). Soybean is generally planted very late in the season and it does not appear very clearly in the image. However, two interpreters, while analysing the history of the soybean in the test site using GTD, concluded that some areas that were fallow or had residuals of vegetation should be classified as soybean. This could be done mainly due to the 'farmer cropping tradition' concept.

Table 3. Result of the digital classification performed by the interpreters using georeferenced plus tabular databases.

	Lots and areas identified by interpreter					
	WI_A		WI_B		WI_C	
Classes	N	*Area	N	Area	N	Area
Bare soil	06	138·14	06	136·90	06	151·36
Residuals	09	315·45	09	293·42	09	295·95
Beans	13	337·30	11	345·66	13	311·75
Maize	07	165·69	09	268·03	07	119·93
Tomato	05	93·86	03	64·90	06	124·82
Sorghum	03	87·65	02	46·64	00	–
Squash	00	–	02	7·36	00	–
Soybeans	01	17·41	00	–	05	160·22
Non-class	00	–	00	–	00	–
Total	44	1163·80	42	1163·90	46	1164·00

Note: N = Number of lots; *Area in hectares.

The use of the GTD allowed the class 'non-class' (non-classified areas) to be reduced. This class had an area of zero hectare after the final classification. However, the use of GTD seemed to have worsened the classification of the 'tomato' class (mainly for the interpreter WI_B). The previous years were indicating a decrease in area for tomato, but in 1991 there was an increase, which might have caused some confusion for the interpreters.

3.3. *Conventional classification*

The results obtained by the three interpreters that used the method of digital classification without access to GTD are presented in table 4. In this step each interpreter generated an image containing the identification of the crops observed at the lots under study, for the year of 1991. These images were also transformed into IPs that were afterwards compared with the ground data within the GIS. In short, the interpreters inspected the 1990 images to acquire a spectral knowledge of the module under study and then they applied a maximum likelihood algorithm to classify the 1991 image.

The procedure for the classification of crops, in this case, was easier in comparison to the interpretation assisted by GTD. The WO interpreters did not have to access the databases, thus making the classification process faster. However, as GTD was not used, the conventional classification became less precise not only from the point of view of the correct identification of the targets but also with respect to the spatial form of the lots. Because the interpreter did not know the history of the agricultural practices adopted by the farmers in the study area, some lots were divided in nontypical shapes.

The performance of the interpreters who used the conventional classification was not as good as that of those interpreters who used the multitemporal/historical databases. The results for the conventional classification are presented in table 4 and show an heterogeneity among the interpreters. Two interpreters could not classify some fields, leaving some areas as 'non-class' (non-classified areas). Furthermore, the class 'tomato' was not classified, even though it did occur in the module.

Table 4. Result of the digital classification performed by the interpreters using conventional classification only.

	Lots and areas identified by interpreter					
	W0_D		WO_E		WO_F	
Classes	N	*Area	N	Area	N	Area
Bare soil	06	143·87	05	152·77	09	181·27
Residuals	10	303·81	09	318·46	17	520·67
Beans	10	232·10	06	139·32	11	263·83
Maize	12	311·43	12	409·00	05	96·99
Tomato	00	–	00	–	00	–
Sorghum	06	117·73	02	32·62	04	100·83
Squash	00	–	00	–	00	–
Soybeans	00	–	04	100·81	00	–
Non-class	04	54·86	01	10·84	00	–
Total	48	1163·80	39	1163·82	46	1163·59

Note: N = Number of lots; *Area in hectares.

The spectral knowledge alone is not enough to make a good classification of agricultural crops in a specific region. It is necessary to have further information on the area. This can be obtained from the use of databases or from the interpreters' experience accumulated over the same area year after year. As a good example, Project MARS (Boissezon and Sharman 1993) has taken advantage of this kind of experience.

3.4. *Statistical analysis*

Table 5 shows the global performance for each interpreter from both the group that was helped by the GTD as well as from the group that made the conventional classification only. These groups can be clearly distinguished from each other. The WI interpreters (who used the multitemporal/historical databases) achieved an average global performance of 73·65 per cent with a standard deviation of 1·57. The second group (WO interpreters), using just the conventional classification, achieved an average global performance of 57·12 per cent with a standard deviation of 5·0. The improvement of the performance of the former group can be attributed to the use of the GTD, since this was the sole factor to differentiate between both groups of interpreters.

The error matrix was derived from the statistical comparison between each interpretation and the ground data. The Kappa statistics were derived from this matrix and are presented in Table 6. In order to improve the evaluation of the Kappa statistics, the scale of values was used as proposed by Landis and Koch (1977). The classifications performed by all the interpreters of the WI group obtained the quality index 'very good' and the classifications performed by the WO group obtained the index 'good'.

Some comments about specific interpretations done by the interpreters must be addressed. The classes 'bare soil' and 'crop residuals' had a quite different spectral

Table 5. Performance of the two groups of interpreters.

Interpreters using databases	Global performance (%)	Average	Standard deviation
WI_A	75·69		
WI_B	73·43	73·65	1·57
WI_C	71·84		
Interpreters using no databases			
WO_D	63·81		
WO_E	51·73	57·12	5·00
WO_F	55·82		

Table 6. Results of 'Kappa' statistics.

Matrix	Lower limit	Kappa	Upper limit	Landis and Koch (1977)
WI_A	0·690	0·693	0·696	Very good
WI_B	0·661	0·664	0·667	Very good
WI_C	0·646	0·649	0·652	Very good
WO_D	0·557	0·560	0·563	Good
WO_E	0·419	0·422	0·425	Good
WO_F	0·431	0·434	0·437	Good