TOOLS FOR MULTITEMPORAL ANALYSIS AND CLASSIFICATION OF MULTISOURCE SATELLITE IMAGERY

Antoine Masse[1], Danielle Ducrot[1], Philippe Marthon[2]

[1]Centre d'Etudes Spatiales de la Biosphère, UMR 5126, Toulouse, France

[2]Institut de Recherche Informatique de Toulouse, UMR 5505, Toulouse, France.

ABSTRACT

As acquisition technology progresses, remote sensing data contains an ever increasing amount of information. Future projects in remote sensing will give high repeatability of acquisition like Venµs (CNES[1]) which may provide data every 2 days with a resolution of 5.3 meters on 12 bands (420nm-900nm) and Sentinel-2 (ESA) 13 bands, 10-60m resolution and 5 days. With such data, process automation appears crucial. For that purpose, we develop several algorithms to automate image processing (classification, segmentation, interpretation, etc.). In this paper, we present an algorithm of automatic analysis which selects the best dataset of dates maximizing classification quality indices. We create two indices to evaluate jointly accuracy and precision. We present tests performed on Formosat-2 images which are similar to Venµs and Sentinel-2 for temporal repetitiveness. These tests allow validating the presented process for temporal discrimination improvement.

Index Terms-- land cover, multitemporal supervised classification, accuracy, precision

1. INTRODUCTION

A great number of tests showed that classification of a large amount of temporal images does not necessarily give the best results for every class [1]. Generally classification on a great number of dates gives the best overall accuracy whereas some classes present confusions with other classes. We know that too much information can impair meaning and spoil the results (figure 3), so we want to select automatically the dataset of images which maximize every class independently by quality indices [2]. For that purpose we create a new index which includes accuracy (correctness) and precision (fidelity) of classification.

Classification method based on genetic algorithm [3] is developed and applied to find the best data set of dates in term of classification accuracy. Confusion matrices (or error matrices) and indices extracted from them provide a classification evaluation. Only pixels of checking sample are classified, so the process is quicker than complete classification (on all image pixels). For a great number of dates (D>10), genetic algorithm and exhaustive research have similar accuracies but the first one is quicker.

The important novelties are automatic processes and their coordination: first, search of the best dataset (for overall accuracy and single class accuracy), second, classification of these data sets and third, fusion of these classifications (for more details and applications of fusion, see [4]).

These principles minimize the percentage of confusion between classes and improve the quality of classification. The sequence of these three processes is fully automatic. However, expert could modify parameters at every moment.

2. METHOD FOR CLASSIFICATION IMPROVEMENT

We introduce a new index that evaluates classification considering omission (deficit) and commission (excess) errors. Several indices exist (such as Overall Accuracy (OA) [5], Kappa index [6]) but they do not take into account all the characteristics which could be extracted from a confusion matrix (for example, absorbing or dissipating classes could have high accuracy but their precision is very low).

Users want to maximize accuracy and precision of the classification and minimize the number of images used (for data price reasons for example). Expert decision has to find compromise between these two optimizations.

Let us see the table 1 for notations and start with the first constraint, maximization of the classification evaluation. We develop an automatic algorithm which provides the best data sets of images depending on ground truth, thematic classes and the D dates used. Genetic algorithms (bases in [7] and new improvements in [8]) are applied to find the best data set because an exhaustive search is not time realizable for D > 20 (see equation 1 for number of possibilities with D dates, for example if D=20, we obtain about 10^6 possible combinations of dates).

$$\sum_{k=1}^{D} \binom{D}{k} = 2^D - 1 \text{ classifications} \quad (1)$$

[1] Centre National d’Etudes Spatiales

978-1-4577-1203-6/11/$26.00 ©2011 IEEE

Term	Definition
S	Dataset of images
I_d	Images of date d
N	Number of sampling pixels
n	Number of classes
D	Number of dates
m_{sij}	Number of pixels classed *i* in checking sample and *j* in classification of dataset *s*.
$m_{si.} and\ m_{s.i}$	Respectively sum of rows $\sum_{j=1}^{n} m_{sij}$ and sum of columns $\sum_{j=1}^{n} m_{sji}$
x_k	Pixel of classification image from dataset $S_k^*, k = 0..n$
$class(x_k)$	Class assigned to pixel x_k
S_k^*	Optimal dataset for class k=1..n, or overall if k=0

Table 1 : Terms and definition

Genetic algorithm is compound of three parts: mutation, crossing-over and selection. The important part is selection and we use an accuracy index based on confusion matrix to evaluate each phenotype. A phenotype is composed of D binary genes, each gene corresponding to utilization or not of a date in the classification.

Confusion matrices are obtained on checking samples with a maximum likelihood classification. See equation 2 for notations of these matrices.

$$m_{sij},\ \ i,j \in [1,n], s \in [1, 2^D - 1]$$
$$S_s = \bigcup I_k \text{ , } k \text{ dates use by } s \tag{2}$$

The new index is named Precision and Accuracy Index (PAI) as it is a combination of accuracy and precision calculations. We defined PAI in equation 3 for a class i and a data set s:

$$\begin{aligned} PAI_i^s = exp(-|\ln(m_{si.}) - \ln(m_{s.i})| \\ + \ln(m_{sii}) \\ - \ln\,(m_{si.} + m_{s.i} - m_{sii})) \end{aligned} \tag{3}$$

More precisely, the first member of the exponential expression takes into account the precision of the class i compared to the others. Maximum precision is reached if sum of rows (deficit of class) and sum of columns (excess of class) are equals. The second member is the diagonal value of the confusion matrix (correspond to accuracy) and the third member gives weight to the evaluation of precision. For a class k, we maximize this index (equation 4).

$$S_k^* = \arg\max_s\ PAI_k^s \tag{4}$$

To find the best global dataset, we use an index derived from PAI: Overall Precision and Accuracy Index (OPAI), which is PAI weighted mean (equation 5)

$$OPAI_s = \frac{1}{N}\sum_{i=1}^{n} CAI_i^s * m_{si.} \tag{5}$$

OA and Kappa are the most used indices but they do not take into account all the information available from the confusion matrix (see equation 6 and 7). OA index only considers diagonal terms of the matrix and kappa index use sums of rows and columns but without an explicit calculation of the precision. The main advantage of OPAI is the combination of two characteristics: in one hand, precision calculation with comparison between columns and rows and in another hand, accuracy calculation with diagonal value. We obtain stricter index than OA or kappa because OPAI does not express the same nature, which explains different values. OA, Kappa and OPAI are complementary.

$$OA_s = \frac{1}{N}\sum_{i=1}^{n} m_{sii} \tag{6}$$

$$kappa_s = \frac{N\sum m_{sii} - \sum m_{si.} \times \sum m_{s.i}}{N^2 - \sum m_{si.} \times \sum m_{s.i}} \tag{7}$$

To obtain the dataset S_0^* which maximize the OPAI, we compute:

$$S_0^* = \arg\max_s OPAI_s \tag{8}$$

So, we have to solve the problem P_1 to maximize overall and class accuracies (respectively equation 8 and 4).

$$(P_1) \qquad \begin{cases} \arg\max_s OPAI_s \\ \arg\max_s\ PAI_k^s, k = 1..n \\ s = 1..2^D - 1 \end{cases}$$

For n classes, solving this problem will provide the n-best-dataset for each class $S_k^*, k = 1..n$ and the best-overall-dataset S_0^*.

To obtain classification image of each dataset, we apply a markovian algorithm of classification: Iterated Conditional Mode (ICM) [9] which class each pixels with constraints on its temporal and spatial context (segmentation image and Markovian cliques).

Then, we could apply a classification fusion process based on confusion matrix to optimize the combination of information of each dataset. With the k+1 classified images and confusion matrices, we apply the fusion formula (see equation 9) presented in [4].

$$x^* = fusion(x_0, x_k), \qquad k = class(x_0) \qquad (9)$$

Finally, we obtain an image with pixels x^* which combined best overall accuracy and best precision of class. If expert wants to minimize the number of dates used, he has to add a new constraint in the algorithm and extend the problem P_1 to problem P_2:

$$(P_2) \quad \begin{cases} \arg\max\limits_{s} OPAI_s \\ \arg\max\limits_{s} PAI_k^s, k = 1..n \\ \# \bigcup\limits_{k=0..n} S_k^* \leq number\ of\ wanted\ dates \\ s = 1..2^D - 1 \end{cases}$$

We can remark that problem P_1 is identical to P_2 with number of wanted dates equals to D. To conclude the theoretical part, the process is resumed in figure 1.

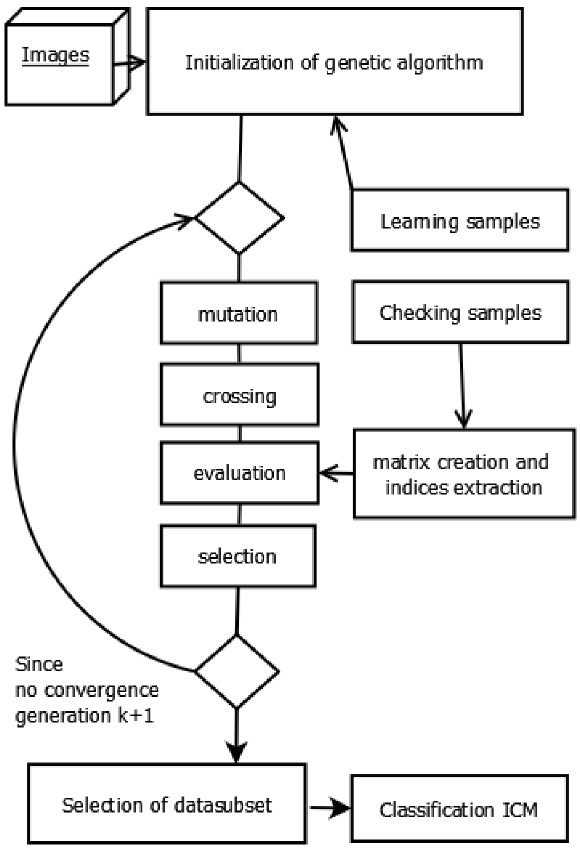

Figure 1 : Genetic algorithm for one best dataset research: for overall dataset, use OPAI for selection and for single class, use PAI.

3. APPLICATIONS AND DISCUSSION

To validate the presented automatic process and new accuracy indices PAI and OPAI, we test them on Formosat-2 (NSPO, Taiwan) images which provide high temporal revisit (2 days) and high spatial resolution (8m) imagery to improve land-cover mapping. These images have been accurately co-registered and corrected from atmospheric effects, making possible their temporal comparison. These scenes were acquired with constant viewing angle such as the future VENµS mission, minimizing temporal noise due to bidirectional effects.

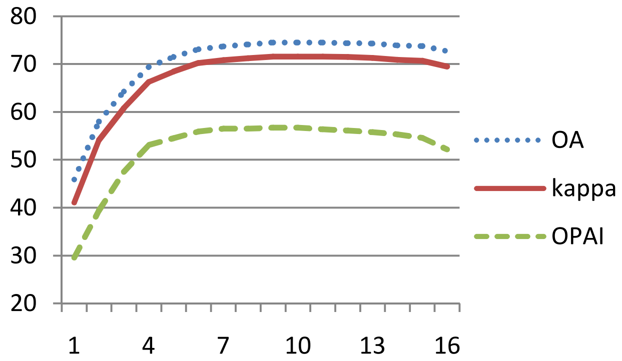

Figure 2 : Evolution of the best datasets found by overall indices depending on number of dates and for all classes.

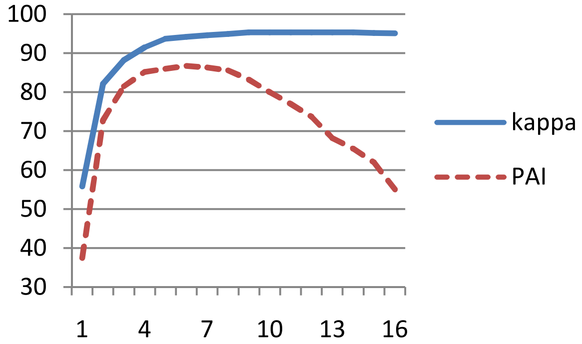

Figure 3 : Evolution of the best datasets found by indices on sunflower class depending on number of dates.

This work is achieved over an agricultural site located in the South-West of France. This site is particularly well instrumented with continuous field observations, allowing accurate training and validation of land cover mapping system. From a thematic point of view, high temporal resolution is crucial to observe phenological stage which discriminates classes (for example, summer crops like corn and soybean have similar stages in some periods, so we must use specific dates to discriminate these classes). So experts need to identify the dates which correspond to their thematic analysis.

We use images of year 2009 (16 dates: 02/15, 03/17, 03/21, 03/30, 05/03, 06/23, 07/01, 07/12, 07/26, 08/05, 08/14, 08/22, 08/30, 09/06, 09/24 and 10/16) and compute a genetic algorithm with those parameters: 1/100 probability of mutation, 6/10 probability of crossing and OPAI index as evaluation function. We use 500 phenotypes and a maximum of 50 generations. Figures 2 and 3 show the difference between datasets found by several indices, PAI and OPAI.

Genetic algorithms computed on these tests (see table 2 and 3) converge very quickly to their optimum (between 10 and 14 generations). For example, PAI finds 03/21, 07/01, 08/14 and 08/30 as best dataset for sunflower class and 4 dates. In the same condition, Kappa gives the following dates: 02/15, 05/03, 08/14 and 09/06.

ALL CLASSES	Genetic algorithm	Exhaustive research		
Problem	OPAI	OPAI	OA	kappa
P_1	56.7 %	56.7 %	74.5 %	71.6 %
$P_2$with wanted dates_number=4	53 %	53.1 %	69.4 %	66.3 %

Table 2: OPAI for 28 classes obtains on best dataset found by the two methods with problem 1 and 2.

SUNFLOWER	Genetic algorithm	Exhaustive research		
Problem	PAI	PAI	Accuracy	kappa
P_1	86.1 %	86.7 %	96.2 %	95.3 %
$P_2$with wanted dates_number=4	85.2 %	85.2 %	92.9 %	91.5 %

Table 3: PAI of class "Sunflower" computed on the best dataset found by the two methods with problem 1 and 2.

The first dataset presents little confusion with other classes (sunflower does not have more than 2% confusion with other classes) whereas the second one provides high accuracy even if there is a high confusion between sunflower and another class (12% confusion between sunflower and corn).

These tests allow concluding that the genetic algorithm used to find the best datasets is powerful: the calculation time is two weeks for the exhaustive research and two days for genetic algorithm in same condition and the found datasets of the two methods have similar values. We can also note that the new index is well adapted to evaluate a classification, PAI or OPAI superior to 80 % is a guarantee of classification quality.

4. CONCLUSION AND PROSPECTIVE

The difference between the presented process and other methods like exhaustive research will be more important with launch of new satellite with high temporal repetitiveness and large amount of data. A goal of this method is to give information of discrimination to users, to help them to choose the dataset which correspond to their thematic research.

The OPAI and PAI indices are well adapted to select dataset because best overall accuracies do not necessarily mean good classifications. These two indices are stricter than OA or Kappa but they clear up all the available information that we could extract from a confusion matrix. A research on other indices extracted from different tools than confusion matrix, will be a future work.

Another current and future work is utilization of the temporal discrimination to interpret automatically unsupervised classification. The most important problem is multiannual difference of phenological stage; a best dataset for a thematic is not the same year after year and spectral discrimination is also very important. Characteristics of new satellite will be better to discriminate classes with multi temporal and multi spectral information.

ACKNOWLEDGEMENTS

We thank the Cesbio members for their contribution: G. Dedieu and O. Hagolle for Formosat 2 correction, J.F. Dejoux, C. Marais-Sicre for ground truth samples and E. Ceshia ("Sud Ouest" project manager).

REFERENCES

[1] Ducrot, D., Sery, F., Sassier, H., Goze, S., And Planes, J.G., Classification and fusion of optical and radar satellite data for land use extraction, Fusion of Earth Data 98, 15, pp. 5-20, 1998.

[2] Hughes, G.F., "On the mean accuracy of statistical pattern recognizers", IEEE Trans Inform Theory, vol. IT-14, pp. 55-63, 1968.

[3] Holland, J., "Outline for a logical theory of adaptive systems", Journal of the Association of Computing Machinery, 3, 1962.

[4] Ducrot, D., Masse, A., Marais-Sicre, C. and Dejoux, J.F., "Multisensor and multitemporal image fusion methods to improve remote sensing image classification", RAQRS'III, 2010.

[5] Congalton, R.G., A review of assessing the accuracy of classification of remotely sensed data, Remote Sensing of Environment, 37, pp. 35-46, 1991.

[6] Cohen, J.A., "A coefficient of agreement for nominal scales", Educational and Psychological Measurement, 20, pp. 37-40, 1960.

[7] Golberg, D.E., "Genetic Algorithms in Search, Optimization and Machine Learning", Reading MA Addison Wesley, 1989.

[8] Alliot, J.M. and Durand, N., "Algorithmes génétiques", Centre d'Etudes de la Navigation Aérienne, 2005.

[9] Idbraim, S., and Ducrot, D., An unsupervised classification using a novel ICM method with constraints for land cover mapping from remote sensing imagery, International Review on Computers and Software (IRECOS), 4, 2009.