Paper presented at the 22nd Asian Conference on Remote Sensing, 5 - 9 November 2001, Singapore.

AUTOMATED CONSTRUCTION OF LEGEND FOR LAND COVER CLASSIFICATION OF ADEOS-II GLI IMAGE

Nguyen Dinh Duong, Nguyen Thanh Hoan, Le Kim Thoa
Department of Environmental Information Studies and Analysis
Institute of Geography
Hoang Quoc Viet Rd., Hanoi, Vietnam
Fax: 84-4-8361192, Phone: 84-4-7562417
Email: duong.nd@hn.vnn.vn,

KEY WORDS: Automated Classification, Land Cover, Digital Legend, ADEOS-II/GLI, Remote Sensing

ABSTRACT Automated classification of land cover by ADEOS-II GLI data is one of the research topics in the framework of NASDA ADEOS-II RA released in 1996. The authors carry out this research as pre-launch algorithm development for classification of Global Imager (GLI) data. One of the issues of automated classification is how to describe land cover objects so that the computer can understand and makes classification accordingly to it. The legend (a set of descriptors of land cover objects) for classification is constructed using image invariants in form of modulation of spectral reflectance curve and its various geometrical parameters. The construction can be made either manually or automatically based on some existing land cover maps. The manual construction is time consuming and requires knowledge on geographical and land cover conditions of the study area. To overcome this obstacle the authors have developed a method for automated construction of legend based on existing information on land cover categories. This method requires GLI dataset and land cover map, which was compiled by any of conventional, or remote sensing (visual interpretation, supervised or unsupervised classification) methods. By overlay of the given map on the GLI dataset and statistical computation a legend for following up classification can be constructed. This approach is useful especially for land cover mapping at regional, continental or global scale.

In the paper, the authors report on algorithm of automated construction of legend, result of usage of the legend for land cover classification using simulated GLI dataset. A discussion on advantages and disadvantages of the automatic versus manual construction of legend is also given in the paper. A comparison study between maximum likelihood and automated classification of land cover was reported.

1. INTRODUCTION

Automated classification of land cover seems to be a rational solution for global and regional environmental research due to its fast processing, objectiveness and high stability in accuracy that provide ability for use in operational basis, near real time processing and reliable change detection. It seems to be also an appropriate methodology for analysis of image data such as MODIS (launched in1999), GLI (to be launched in November 2001) and other which are in huge volume due to their high temporal resolution. The automated classification is carried out by so-called Graphical Analysis of Spectral reflectance Curve (GASC) developed in the framework of NASDA's ADEOS-II GLI research announcement. Instead of using various statistical measures such as mean vector, variance-covariance matrix and spectral distances like Mahalanobis or Euclidian that are commonly used in conventional classification methodology, the GASC algorithm proposes usage of graphical features of the spectral reflectance curve as principal classifiers. These graphical features, which are modulation of the spectral reflectance curve, total reflected radiance index (TRRI), normalized pixel vector and arithmetic combination of its components, are further called as image invariants due to their exclusiveness to particular land cover type. It means that every land cover category can be described by a set of these image invariants. A collection of sets of image invariants of different land cover types to be classified defines so-called digital land cover legend. Unlike conventional land cover legend that is a sets of statements describing land cover class and rooting from inter-disciplinary approach, which is suitable preferably for visual interpretation and expert analysis, the digital land cover legend is a sets of numeric values representing graphical feature of the spectral reflectance curve that are fully understandable by computer and therefore will be most suitable for automated classification.

Construction of legend for land cover category can be made either manually in interactive operation or automatically based on existing land cover map established by conventional methodology. In manual construction of legend the operator uses normalized pixel vector and his knowledge of spectral characteristics of remote sensing sensor combined with his experience and knowledge of local land cover conditions. This approach leads to very finely constructed legend, which can be short and comprehensive and provides the fastes t processing. However, the construction is time consuming, requiring very high expertise and subsequent long time tuning. When the remote sensing sensor is changed, the construction process needs to be done from the beginning. The automated construction, on the other hand, is more versatile and can overcome the mentioned obstacles. This paper will report on basis of the automated construction of land cover legend, comparison study to highlight its advantages versus manually constructed legend. Result of automated classification of land cover by simulated GLI data set and legend constructed by the proposed method is also compared with result of maximum likelihood classification. This comparison is subject for further discussion.

2. ALGORITHM DESCRIPTION

The algorithm for automated construction of legend is very simple. It is actually an overlay operation of given land cover map on image data and by statistical computation, pre-defined image invariants are computed. The concept of image invariant computation and automated classification is shown on figure 1.

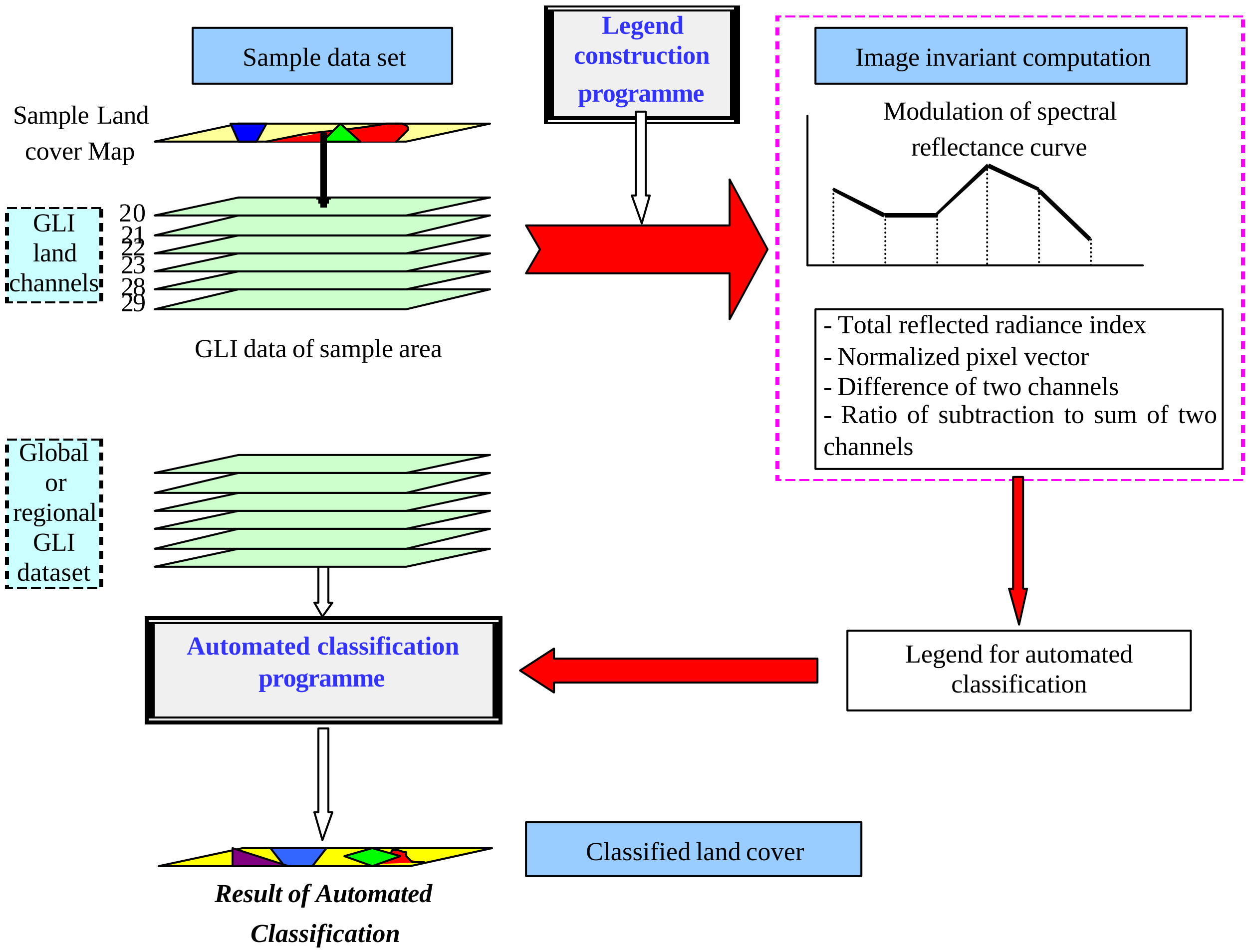


Fig.1 Concept of image invariant computation and automated classification

Because the computer cannot decide what invariant is relevant for a particular land cover class so all variants will be considered. For each land cover class, the following invariants are computed:

+ Modulation of spectral reflectance curve
+ Total reflected radiance index (TRRI)
+ Normalized pixel vector
+ Spectral differences
+ Ratio of subtraction to sum of a pair of spectral channels

The pixel vector should be converted from DN to radiance values before computation. Because the input land cover map was compiled by different techniques like supervised classification, visual interpretation combined with fieldwork etc. a filtering operation is required to exclude pixel which does not conform statistical model of the given land cover class. The filtering is done by using mean vector, standard deviation and pre-defined thresholds applied to the standard deviation of each spectral channel. Pixel vector that is in outside of spectral range specified by threshold values will be ignored and does not proceed to the next computation phase. After computation a digital legend is constructed. This legend, in contrast to training sampling in supervised classification that is applicable only for the image where it was generated, can be used for classification of many other image data. That is scene independent characteristic of the image invariants. The legend is constructed by syntax, which is obvious in example given on table 1.

Table 1.

Legend for automated classification of land cover	Explanation
1	Land cover class code
102-1	Short name of the class
102-1	Full name of the class
128 128 128	Color for visualization
M 011100	Modulation computation instruction
63.21 28.32 34.93 46.08 84.53 44.88	Characteristic vector of the class
T 38.07800 61.43700	Range of total reflected radiance index
P1 53.68000 69.91600	Range of normalized value of channel 20
P2 23.05500 33.14700	Range of normalized value of channel 21
P3 25.49300 43.15900	.
P4 34.24600 61.48600	.
P5 59.38200 112.1910	.
P6 30.96500 59.92500	Range of normalized value of channel 29
D12 29.66100 38.60300	Range of difference of channels 20 and 21
D13 22.04900 32.19100	Range of difference of channels 20 and 22
.	.
.	.
D56 25.68000 55.85900	Range of difference of channels 28 and 29
C1212 -+ 0.3420000 0.4197000	Range of subtraction and sum of channels 20 and 21
.	.
.	.
C5656 -+ 0.2382000 0.4060000	Range of subtraction and sum of channels 28 and 29
END	End of legend for class 1

On the table 1 is example of digital legend for one land cover class. In case of more classes, this structure is repeated for all of them.

3. CLASSIFICATION EXAMPLE

Study area has been chosen in south of Vietnam, which covers Lam Dong, Dong Nai, Binh Thuan provinces and a part of Hochiminh City. Land cover of this area is very diverse. It includes tropical evergreen broadleaf and needle leaf forest, various deciduous vegetation covers, savanna and grassland, mangrove and commercial plantation like rubber, coffee, tea and fruit trees. This area also includes different landscapes as semi-arid, high-plateau and mountainous zones and Mekong river delta. Simulated GLI data has been generated by visible and near infrared channels of two LANDSAT TM scenes 124/52 and 124/53 acquired on March 1, 1996. Spatial resolution was sampled from 30m to 250m to meet technical parameters of GLI sensor. Geographical location of study area and its false color composite are shown on figure 2.

Maximum likelihood classification has been carried out by WinASEAN software. Training data was generated for 39 land cover classes. The classification result was validated by field work and field photo database. The obtained classified image is used for overlay on GLI data set to generate digital legend for automated classification. On figure 3 is shown maximum likelihood classification result.

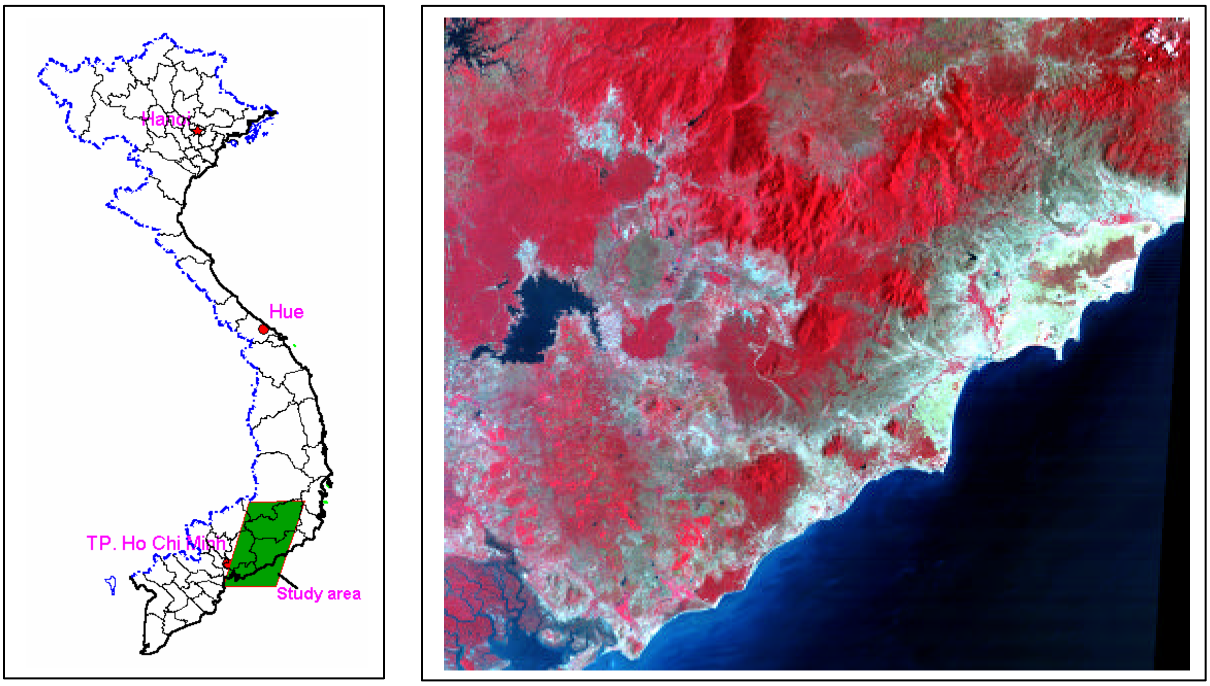
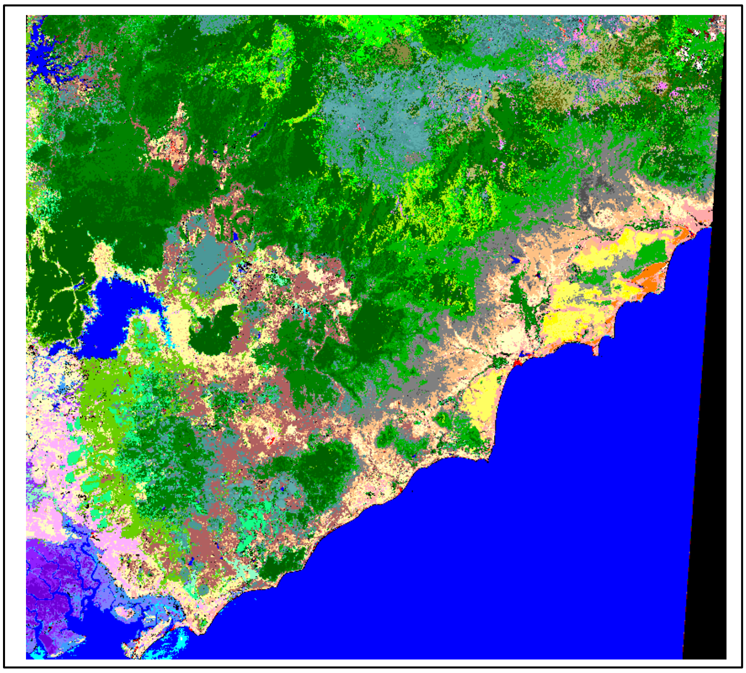


Figure 2. Location of study area and false color composite



ID	Color	Description
1		Closed evergreen forest
2		Medium evergreen forest
3		Opened evergreen forest
4		Conifer forest
5		Deciduous forest
6		Closed and medium mangrove forest
7		Opened mangrove forest
8		Shrub land
9		Grassland
10		Dry cropland
11		Wet cropland
12		Bare ground
13		Red sand
14		White sand
15		Basalt ground
16		Urban
17		Clear water
18		Turbid water
19		Cloud

Figure 3. Maximum likelihood classification

After overlay operation and statistical computation, a digital legend was constructed and used for automated classification. On figure 4 is result of automated classification using automatically constructed legend (left) and manually constructed legend (right).

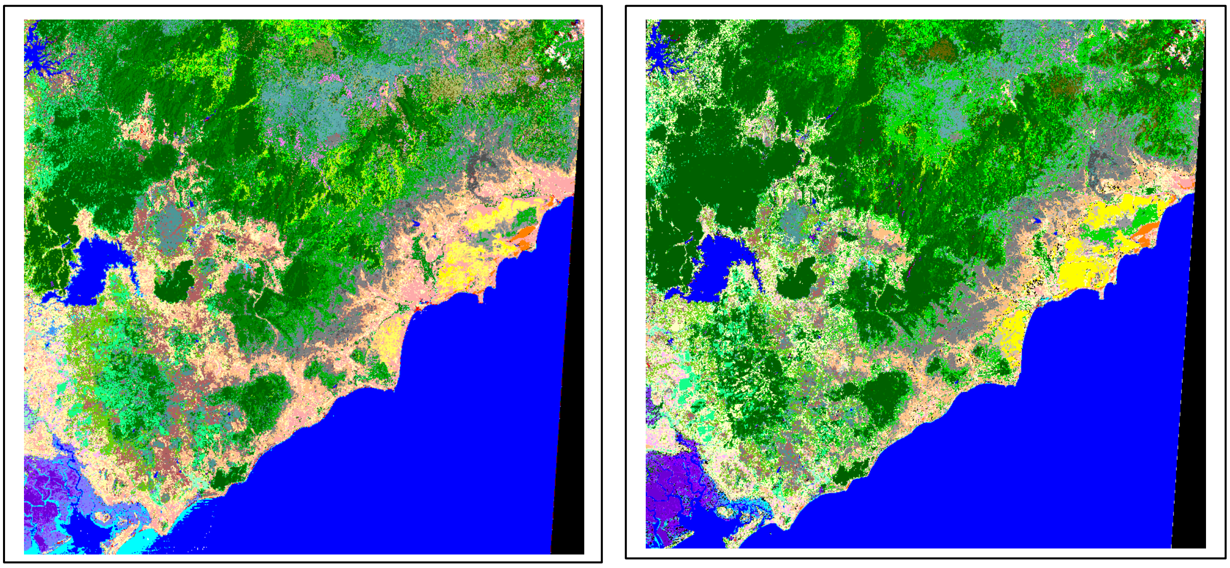


Figure 4. Result of automated classification using automatically (right) and manually (left) constructed legend

On Table 3 are demonstrated legends constructed by training selection, automatic and manual methods for class: Closed forest.

Table 3.

Interactive training sampling	Automatically constructed legend	Manually constructed legend
Class: C_Forest	161	161
	161	Closed forest
Mean vector	0 96 0	C_Forest
	M 111111	0,96,0
73.3 27.4 24.4 70.6 47.1 13.0	50.07 20.30 17.46 65.11 41.17 11.96	M 011100
	T 23.70850 44.76710	73,27,24,70,47,13
Variance - Covariance matrix	P1 40.15000 64.50400	R34 0 0.3
.8 .4 .3 .5 .2 .1	.	D16 38 60
.4 .4 .3 .1 .3 .2	.	T 36 45
.3 .3 .4 -.2 .6 .3	P6 7.435000 21.01000	H 180 360
.5 .1 -.2 4.0 -.8 -.8	D12 23.51750 38.7120	END
.2 .3 .6 -.8 2.5 1.0	.	
.1 .2 .3 -.8 1.0 .6	.	
	D56 15.5275 37.4620	
Standard deviation	C1212 -+ 0.38257 0.48158	
	.	
.9 .6 .6 2.0 1.6 .7	.	
	C5656 -+ 0.45003 0.58720	
	END	

When comparing legends constructed by these three methods we could see that interactive training sampling produces a legend, which is composed of statistical values as mean vector, variance-covariance matrix and standard deviations of the mean vector with accuracy and reliability depending very much on skill and experience of an operator. These values are unusable for multi-scene analysis. It means that we need to generate a separate legend for each scene so there is almost no way to automate the classification. On the other hand, the legend manually constructed is very comprehensive and contains unique information derived from spectral characteristics of the land cover. It means that only image invariant relevant for the given land cover category is selected. The legend for class: Closed forest as shown on table 2 is composed of 6 invariants while automatically constructed legend has 249 invariants. The more invariants presented in the legend, the slower is the classification. To achieve good land cover map, the legend should be constructed by expert who understands well response of land cover to spectral characteristics of the sensor and also geographical condition of the study area. It is always time consuming process. However, when construction is done, the obtained legend can be used for classification of many other scenes. The automatically constructed legend, in contrast to manually constructed one, includes all image invariants. If we consider invariants as modulation of spectral reflectance curve, TRRI, normalized pixel vector, spectral differences of channels and ratio of difference to sum of a pair of spectral channels then for each land cover category we need to compute 249 invariants, which cover all of channel pair combinations of GLI dataset.

Using these legends for land cover classification of the study area we obtained land cover maps as shown on figures 3 and 4. The corresponding land cover categories on these maps are visualized by the same color so that they are comparable each to the others. Due to complexity of spectral reflectance and absorption of the land cover, the manually constructed legend is limited to some major land cover classes. Expansion of this legend to more detail land cover types and refining its parameters are possible by consecutive upgrading. The automatically constructed legend allows to achieve land cover map very closed to the one produced by maximum likelihood classification. All classes that were classified by maximum likelihood method are successfully extracted correspondently by GASC algorithm with automatically constructed legend. However, slight differences in boundaries of some classes are observed. These differences are caused principally by different classification mechanism of pixels, which are in the overlap of neighboring classes.

4. CONCLUSION

The proposed method of legend construction allows to shorten legend construction phase that is usually time consuming in case of manual construction. This approach provides ability to inherit and continue up-to-date land cover information gathered by different methodologies. It also enables to adapt quickly the GASC algorithm to any new remote sensing sensor to build up legend for land cover classification. The usage of image invariants like modulation of spectral reflectance curve, TRRI, normalized pixel vector, spectral differences of channels and ratio of difference to sum of a pair of spectral channels to describe land cover is proved by research result, however, such a mechanically constructed legend does not provide optimal classification from execution time point of view. Refining to exclude invariants that are not relevant for particular land cover category and to make the legend more comprehensive is subject of further research.

ACKNOWLEDGEMENT

The authors would like to thank National Space Development Agency of Japan (NASDA) for financial support through ADEOS-II/GLI RA. Sincere thanks are also given to the Basic Research Program of Vietnam for co-sponsorship to this research.

REFERENCE

Nguyen Dinh Duong. Graphical Analysis of Spectral Reflectance Curve. Proceedings of the 18^{th} Asian Conference on Remote Sensing. 20 – 24 October 1997, Kuala Lumpur Malaysia.

Nguyen Dinh Duong. Total Reflected Radiance Index - An Index to Support Land Cover Classification. Proceedings of the 19^{th} Asian Conference on Remote Sensing. 16 – 20 November 1998, Manila, Philippines

Nguyen Dinh Duong. Land Cover Category Definition by Image Invariants for Automated Classification. International Archives of Photogrammetry and Remote Sensing. Vol. XXXIII, Part B7/3, Commission VII. ISPRS 2000 Amsterdam, the Netherlands.