XIV International congress of the S.I.P. -Hambourg 1980-

AN AID FOR THE INTERPRETATION OF LANDSAT IMAGES.

Marie-France OUDIN, Denys CHAUME

Scientific Center I.B.M. France
36 Avenue Raymond Poincaré -PARIS 75016-

++++++++

INTRODUCTION

The use of LANDSAT data is getting more and more important in sciences refering to soil studies (agronomy, pedology, mining prospecting, etc...).
Taking advantage of these data requests some manipulations refering to mathematical methods used in statistical studies for multidimensional data processing. Principal Components Analysis and its application to LANDSAT data processing is developped in this article.

Because of the use of satellite data processing for getting a thematical map, most of the information coming from the four transformed channels has been displayed, trying to improve the representation technics.
We have studied different comparing criterions to supervise the trust of PCA for image display, and for extrapolation of its results in space. This shows the stability of the results for adjoining images.

PRINCIPAL COMPONENTS ANALYSIS

The purpose of Principal Components Analysis (cf. Lebart et F^{3}nelon 1975) is to provide synthetic representations of numerical values spacious entities. This allows to obtain a descriptive abstract (that can be represented as a graphic) for a table of n observations of p variables on which no particular hypothesis has been previously made.

From the set of points represented in the space Rp of variables, we look for the subspace Rk, with smaller dimensions, that maximizes its variance.

This new space is represented by the system of axes defined from the eigen vectors of the variance-covariance matrix of the set of points (corrrelation matrix when the data are reduced). Each eigen value of this matrix represents the projection dispersion of the set of points on the corresponding axe. That means that higher is the eigen value, and better is the representation on the observed axe.

APPLICATION TO THE LANDSAT IMAGES PROCESSING

The set of observations is there appointed with the radiometric values of every pixel of an image in the four original channels (variables) of the LANDSAT satellite.

The principal characteristic of the eigen values is that they decrease very quickly, and so, the two first eigen values generaly hold more than 95% of the amount of all variances.

That means that the statistical set will get flat very quickly in the direction of the axes corresponding to the decreasing eigen values. Therefore, if 95% of the whole variance is held in the two highest eigen values, this set of concentration has just two observable dimensions. They are represented by the two first eigen vectors of the inertia matrix, or factorial axes F1 and F2, which are therefore the only one to be taken into account.

In the original reference of observations, the whole realisations are strongly correlated. The transformation (Karhunen-Loeve transformation, Lowitz 1976) takes as new axes the principal axes of the set; the realisations projections on these new axes are uncorrelated.

These projections are numbered from 0 to 255, the mean value being on 128, to provide "artificial channels". A display of these channels gives two "principal images" which represent most of the information.

PRINCIPAL IMAGES DISPLAY

For photo-interpretation, we have reduced four images into two images, However, it is still difficult to analyse simultaneously two images.

Colored compositions usually used for simultaneous display of several spectral bands are not satisfying because they don't allow to separate easily the two components. It is indeed difficult, in additive colours for example, to find again the respective importance of F1 and F2 in a yellow made of green and red.

The decomposition of hues restorable with three filters (trichromy) in lightness and chrominance is not new; processes of colour televesion screen are based on it. It can also be found in the mathematical models of vision (Faugeras 1978).

Chrominance can itself be broken up into colour (argument) and saturation (module) (Chaume 1977).

If we display an image, assigning to each pixel a hue the light of which is proportional to its value on F1, and the colour of which is proportional to its value on F2, saturation being chosen as high as possible, we get a document on which we can simultaneously "read" the information contained in the two first principal images.

For the "thematician" who wishes to use LANDSAT images, this document offers several advantages. The most important are:

1) The document includes most of the information contained in the LANDSAT image (95% of the variance).
2) If the proportionality coefficients are correctly chosen, objects represented may appear with colours recalling their natural colours (forests and meadows in green, bare soils in very bright magenta, rocs and town areas in red, water in black, etc...).
3) An area of the principal plan F1/F2 corresponds to a hue of the image. This property may bring an important help for automatical classifications using the two-dimensional histogram (F1,F2) (Lowitz 1976,Chaume-Nguyen 1980).

KARHUNEN-LOEVE TRANSFORMATION STABILITY.

Karhunen-loeve codification proves itself to be an interesting tool for multi-spectral images processing, as well to obtain automatical or semi-automatical classification than to obtain documents used for a manual analysis. However, it still remains a process suitable for the studied image, and one can ask the question whether the interpretation keys or the definied classes for a limited area are still valid when moving out of it.

In the purpose of studying the variation in space of factorial planes coming from different Principal Components Analysis, the image of Gale (Mali) requiring 480 lines on 480 points, representing about 1000 km2 surface, has been divided into sixteen sub-images (120 X 120) (table 1).

L11	L12	L13	L14
L21	L22	L23	L24
L31	L32	L33	L34
L41	L42	L43	L44

Tableau 1 : Division of the Gale image (L88) in 16 sub-images

The PCA were calculated for the whole image and for every sub-image, providing 1+16 systems of axes and for some of them, the principal caracteristics are in table 2.

The new radiometric values of every pixel are varying in the principal planes. Variation produces two aspects, according to the allowance the processing done before PCA may accept.

- Qualitative aspect (allowance for image display).
- Quantitative aspect (allowance for manual classification).

QUALITATIVE ASPECT.

The L23 sub-image was chosen as reference image and the principal images F1 and F2 were calculated and displayed for this area, into the different systems coming from the 17 PCA.

The radiometric values of this sub-image points were calculated with the transform matrix obtained for every PCA. It can be remarked that, provided contrast improvements adjusted to the resulting images (fitting of mean values and standard deviations), images diaplayed on eight levels of grey are all similar, as well for the first principal image than for the second one.

It means that differences between corresponding factors values are not high enough to induce in the radiometric values variations that would produce pixel class changes. Actually, only few pixels scattered at random on the whole image are affected by this class change, so that globaly all images look out the same.

QUANTITATIVE ASPECT.

Three comparaison tests have been held:

- Angles between homologous factors.
- Angle between principal planes.
- Correlation between principal images (Fi1 or Fi2).

1) Computation of the angle values.

Generaly, angles between the first axes are smaller than angles between the second axes; the first ones are less than two degrees, when the second ones may reach more than twelve degrees.

Though, these angles values are only significant regarding to their incidence on values assigned to the pixels when computing principal images.

If we project one point of the radiometric space on two axes F1 from analyses done on two different sub-images, the angle between these projections equals the angle between the axes and the radiometries relative variation on the axes varies as the sinus of this angle.

We see that variation is significant for image display (4 radiometries variation) when the angle between axes reaches 16° (cf. tableau 3).
This shows the very strong similarity observed between different image displays of the same area on eight levels of grey.

The variation is quantitatively significant (one radiometry) when the value of the angle reaches 4° (cf. tableau 3).
Therefore, the results that were obtained allow us to conclude that only a few angles between the axes F2 might lead some variations to take into account for an extension to a classification.

2) Computation of the correlation coefficient.

Correlation coeffecients between the principal images, born of the use of different PCA have been computed.

We verify that they are very close to 1, with nevertheless some divergences more important for the F2 than for the F1.

The correlation coefficient R, between the principal image i and the image obtained by projecting original channels on an axe that does with this one the angle x verifies the formule:

$$1 - R2 \leq (VP1 / VPi)\ tg2(x)$$

with: VP1 = first eigen value.
VPi = the others eigen values.

Looking at this formule, we see that the correlation coefficient is necessarily closer to 1 for the first principal images (VPi high) than for the other ones.

In the utmost case of a 12°39' angle, we have:

$$tg2(x) = 0{,}0452$$

That gives as absolute values of correlation coefficients, the following least values:

0,9771 for the first axe.

0,8203 for the second axe.

These values are very much lower than those that were found (cf.tableau 4).

The three tests used to compare the two first axes of the PCA done on the sub-images of a same image gives importance to the very high stability of the Karhunen-Loeve transformation.

CONCLUSION

Karhunen-Loeve codification derived from PCA permitted us to build a document representing most of the Landsat image information, liable to be brought in ground mission. It is the link needed between the thematician's work and automatical classification.

The stability study of this codification allows to measure up the extent (that is to say the limits) of the keys interpretation validity, and of the classifications proceeding from it.

It remains to carry on by studying the spacial variation and also by looking at the codification stability between images taken at different times.

ACKNOWLEDGEMENT

The work here presented was done by the soil scientists of the "Institut de Recherches Agronomiques Tropicales des Cultures Vivrières" (IRAT), conjointly with the research workers of the the "image processing" group of the Scientific Center I.B.M-France. Let all be thanked.

BIBLIOGRAPHY

CHAUME D. : Image visualisation on TV screen. NATO-ASI Digital Image Processing and Analysis, Simon and Rosenfeld ed., Noordhorf 1977.

CHAUME D. et NGUYEN P.T. : Interactive processing of Landsat image for Morphopedological studies. Machine Processing of Remotely Sensed Data Symposium (june 1980), Purdue University, W-Lafayette, Indiana.

FAUGERAS O. : Application d'un modèle homomorphique de vision au traitement des images couleur. Congrès AFCET-IRIA, Châtenay-Malabry, (février 1978).

LEBART L. et FENELON J.P. : Statistique et informatique appliquées. Dunod éd. (1975).

LOWITZ G.E. : Etude de Morphographie Dynamique. Matra Espace (1976).

image	MSS4	MSS5	MSS6	MSS7
L88	22.66	30.06	35.46	14.90
L12	22.64	29.90	35.70	15.30
L23	22.05	28.93	34.71	14.72
L31	22.58	30.04	35.66	15.05
L33	22.25	29.58	34.63	14.37
L34	21.48	28.29	33.62	13.99
L44	22.31	29.43	34.10	13.91

a) Means values of the radiometries in the original channels.

	image	F1	S.D.	F2	S.D.	v.p.1	.p.2
	L88	128.05	6.91	127.99	2.63	3.33 0	46
	L12	126.30	6.36	128.05	2.89	3.35 0	47
	L23	128.01	6.72	128.00	3.00	3.14 0	62
	L31	126.81	4.86	128.37	2.29	3.26 0	49
	L33	127.82	6.67	128.78	2.97	3.20 0	56
	L34	129.65	6.21	128.11	2.74	3.22 0	52
	L44	128.32	6.13	126.82	2.81	2.85 0	78

b) Factors F1 and F2 means and standard deviations, first and second eighen values.

Table 2 : Statistical properties of the original channels and of the axes F1, F2.

angular err	radiometric err
2°23'	1
3°52'	1
12°39'	3
15°40'	4

<u>Table 3</u> : Relation between angular variation of an axe and the variation of the radiometric value computed by projection on this axe.

image	F1	F2	P. P.	corr F1	corr F2
L23/L88	0°48'	3°28'	0°08	0.9978	0.9841
L23/L12	0°43'	0°04'	0°00'	0.9979	0.9911
L23/L31	0°44'	12°39'	0°54'	0.9963	0.9263
L23/L33	0°09'	1°37'	0°00'	0.9983	0.9891
L23/L34	0°52'	7°17'	0°11'	0.9979	0.9847
L23/L44	2°07'	-5°51'	0°08'	0.9977	-0.9525

<u>Table 4</u> : Variations of the axes F1 and F2 computed on different sub-images comparing to the same axes computed on the L23 sub-image. (F1: angles between axes F1, F2: angles between axes F2, P.P.: angles between the F1-F2planes, corr F1: coefficient of correlation between the first principal images, corr F2: coefficient of correlation between the second principal images).