POLISH ASSOCIATION FOR SPATIAL INFORMATION
ANNALS OF GEOMATICS 2009 ❍ VOLUME VII ❍ NUMBER 2(32)

IMPROVING UNSUPERVISED CLASSIFICATION OF FOREST TYPES THROUGHT THE USE OF SPECTRAL INDICES

UDOSKONALENIE KLASYFIKACJI NIENADZOROWANEJ TYPÓW LASÓW PRZEZ WYKORZYSTANIE WSKAŹNIKÓW SPEKTRALNYCH

Piotr Krzysztof Stępniewski [1], Krzysztof Marek Będkowski[2]

[1] Techmex SA
[2] Katedra Urządzania Lasu, Geomatyki i Ekonomiki Leśnictwa, SGGW w Warszawie

Key words: forest, spectral indices, unsupervised classification
Słowa kluczowe: las, wskaźniki spektralne, klasyfikacja nienadzorowana

Introduction

The natural land cover class of Poland is almost entirely forest, but nowadays it is replaced by man-made ecosystems like arable, meadows, pastures and urban areas. The remaining forest forms islands, which are connected each other by corridors located mainly along rivers. Satellite images are an effective tool for the recognition of land cover structures which forms the first step for attaining a good understanding of landscape function.

One of the commonly used remote sensing methods is unsupervised land use classification derived from the ISODATA (Iterative Self-Organizing Data Analysis Technique) algorithm. The logic of this routine was described by (Ball, Hall, 1965). Although the ISODATA routine applied to spectral bands acquired from multispectral scanners usually leads to good classification results, new approaches continue to be developed. In this paper we describe how classification results can be improved through the use of spectral indices instead of the original spectral bands. This will be explained using an example of analysing the structure of a forested area. The main objective of this research is to analyse how the vertical and horizontal structure of mixed forests can be assessed from satellite images. In our study we've used the IDRISI32 ISOCLUST routine, which is a specific implementation of the ISODATA approach (Eastman, 1999).

Study area and data used

The study was carried out on a 100 km^2 subset of the Landsat scene p188r24 acquired on 7th May 2000 (http://www.landsat.org). This 7-band multispectral image with panchromatic band covers a region located in central Poland. The main land cover classes here are agricultural areas (arable fields, meadows, pastures) and forests. This landscape can be considered as representative of this part of Poland.

The entire study area was also recorded using aerial photographs taken with a DMC 2000 digital camera. The resolution (ground sample distance) of these images was about 0.15 m. For our study NC- (natural color) and CIR-composites were produced.

Methods

Spectral bands and spectral indices

The general overview of the data processing steps applied is shown in Figure 1. First a set of spectral indices was produced. Using 6 of the original bands (ETM1-ETM5, ETM7) 10 indices were calculated (Table 1). All indices were calculated using the Image Calculator module in IDRISI32. The resulting images were standardised by stretching them to 256 (0-255) DN-levels.

NDVI and *NRVI* indices are inversions – they contain the same information:

$$NRVI = \frac{\frac{RED}{NIR} - 1}{\frac{RED}{NIR} + 1} = \frac{RED - NIR}{RED + NIR} = -NDVI$$

These two indices weren't used in further classification procedures; their function was only to control whether the data processing is affecting (changing) or not the spectral information within the indices.

For further operations 6 indices were chosen: RATIO, TMMSI, TMMSI2, TMVI2, TMVI3, TMVI4. The key was the correlation level between indices. Low correlated indices were chosen, as it indicates their suitability to distinguish different land cover classes. The only exception was with RATIO, which is usually high correlated, but we've chosen it because of it's good usability in land cover classification.

Table 1. Spectral indices

Spectral index	Formula (numbers are indicating spectral bands)
RATIO	04/03
TMVI2	05/02
TMVI3	04/02
TMVI4	05/03
TMMSI	05/04
TMMSI2	07/04
GNDVI	(04-02)/(04+02)
SLAVI	04/(03+05)
NDVI	(04-03)/(04+03)
NRVI	[(03/04)-1]/[(03/04)+1]

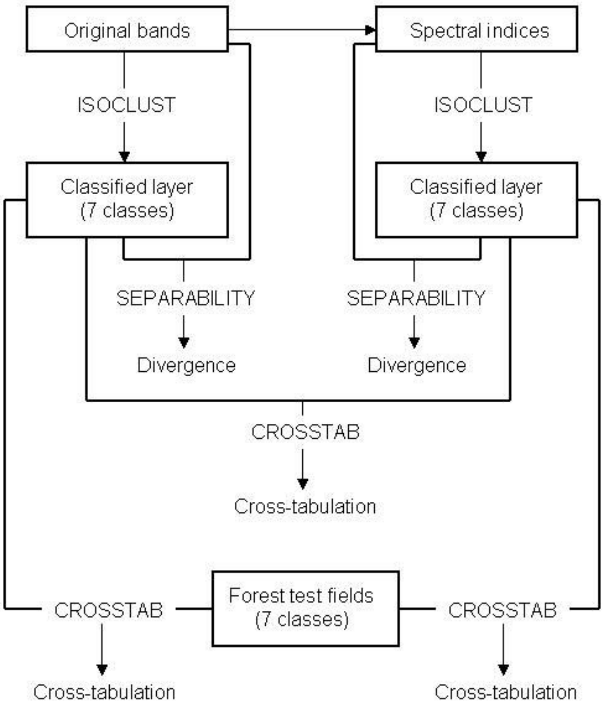


Figure 1. Work flow of operations

Unsupervised classification

The unsupervised ISOCLUST classification was performed using the original spectral bands and, separately, the spectral indices. The routine implemented in IDRISI32 requires a color composite image for use in the cluster seeding process. For this we have used bands 3, 4 and 5 from original data set and TMMSI2, TMVI2 and TMVI4 indices.

Both the spectral bands and spectral indices were classified into 5 clusters. The results obtained were compared with an existing topographic map. This classification produced unsatisfactory results, because some important classes, like water and coniferous forests were grouped together into one class. Thus we decided to expand the number of desired clusters to 7.

The land cover layers (Fig. 3, 4) produced with these two methods seem to be very similar to each other and it is impossible to determine their quality without detailed investigation. To assess the results we have compared the layers using cross-tabulation, in which the categories of one layer were compared with those of the second layer. The result of this operation is a table (Table 2) listing the number of cells in each combination of clusters between the layers.

Where there are only one dominant number in every row, then the layers could be assumed to be very similar. This is the case only for clusters 3 and 7. This indicates that the classification results differ – the pixels from one cluster in first layer are assigned to a number of different clusters in the second layer.

For further assessing of cluster layers we have applied an approach known from supervised classification. Both layers were now used as training fields to obtain spectral signatures based on the original images or the spectral indices. The resulting signatures were compared

Table 2. Cross-tabulation of classification results achieved with ISOCLUST routine performed using original spectral bands (columns) and spectral indices (rows) [numbers shown are pixels]

		Clusters formed based on spectral bands							
		1	2	3	4	5	6	7	Total
Clusters formed based on spectral indices	1	434	1316	442	3 341	5 490	66	4 356	15 385
	2	11 802	2718	4	110	0	6 347	2 093	23 074
	3	0	65	8 837	1	577	0	873	10 353
	4	10	159	3 048	27	8 302	0	3 797	15 343
	5	5 589	1654	0	0	0	15 594	70	22 907
	6	6 210	963	1 271	830	31	65	8 270	17 640
	7	37	1496	39	5 891	4	13	89	7 569
	Total	24 082	8371	13 641	10 200	14 404	22 025	19 548	112 271

using a signature comparison chart (Fig. 2a,b). In addition, a separability measure – divergence (refer to IDRISI32 help system) was calculated.

The overall separability (divergence) for clusters extracted from the original spectral bands was equal to 108.57 and the separability for clusters extracted from the spectral indices was 1186.97. Both the signature comparison charts (Fig. 2a,b) and divergence values proved better separability of clusters extracted using the spectral indices than those extracted from the spectral bands.

Comparision of classification results with true reference data

The best way to check out the results of unsupervised classification is to compare them with true data. As a reference, a layer with land use classes produced with other methods should be used. Generally, this scene can be divided into 7 classes of objects: fields with green cover, fields with bare soil, meadows, lakes, coniferous forest, mixed forest and deciduous forest.

Although using satellite data forests are usually classified into only three classes (coniferous, deciduous and mixed) we decided to form some subclasses, to reflect differences in tree species and the structure of the forest stands within the whole scene. Seven classes of forests were distinguished:

1. Coniferous forest
   1.1. Coniferous, one storey (scotch pine – *Pinus silvestris*).
   1.2. Coniferous, one storey (*norway spruce – Picea abies*).
2. Mixed forest
   2.1. Coniferous in first storey (Scotch pine) and deciduous in second storey (oak – *Quercus sp*., red beech – *Fagus silvatica*, hornbeam – *Carpinus betulus* and other).
   2.2. Mixed forest with approximately even rates of scotch pine and oak in first storey.
3. Deciduous forest
   3.1. Deciduous, one storey (oak).
   3.2. Deciduous, one storey (black alder – *Alnus glutinosa*).
   3.3. Deciduous, one storey (birch – *Betula verrucosa*).

Polygons indicating boundaries of these 7 forest types were drawn by stereo vectorization using a DVP 5.0 – Digital Video Plotter photogrammetric station and aerial images taken with a DMC 2000 digital camera. Following this, a mask of polygons representing forest types was produced. This mask was used for cross-tabulation with the layers shown on Figure 3, 4. The results are in Table 3 and 4. Note that there are no polygons representing non-forest classes, as they were not the subjects of our study.

Table 3 shows that one (5th) cluster dominates. This indicates, that it is not possible to distinguish between forest classes using clusters derived from the original spectral bands. There is a clear distribution of pixels shown in Table 4. The numbers of pixels seems to be connected with the share of coniferous and deciduous trees within each forest class: (a) scotch pine, (b) scotch pine in first storey, (c) deciduous trees in second storey, (d) scotch pine in mixed forest, (e) deciduous trees in mixed forest, (f) deciduous forest – oak, (g) deciduous forest – black alder. The samples for birch and norway spruce are too small and cannot be interpreted in this way.

Table 3. Cross-tabulation of clusters from layer created by ISOCLUST classification performed using original bands (columns) against forest classes (rows) [numbers shown are pixels]

		Clusters formed based on original image bands							
		1	2	3	4	5	6	7	Total
Forest classes	1.1	1	40	3	381	267	0	2	694
	1.2	0	5	2	3	12	0	0	22
	2.1	0	12	5	92	613	0	0	722
	2.2	0	15	23	11	743	0	0	792
	3.1	0	0	41	1	219	0	0	261
	3.2	0	1	22	3	125	0	2	153
	3.3	0	0	1	0	7	0	0	8
	Total	1	73	97	491	1986	0	4	2652

Table 4. Cross-tabulation of clusters from layer created by ISOCLUST classification performed using spectral indices (columns) against forest classes (rows) [numbers shown are pixels]

		Clusters formed based on spectral indices							
		1	2	3	4	5	6	7	Total
Forest classes	1.1	457[a]	6	1	62	2	25	141	694
	1.2	5	0	3	11	0	0	3	22
	2.1	486[b]	1	4	195[c]	0	12	24	722
	2.2	367[d]	0	44	377[e]	0	4	0	792
	3.1	4	0	55	201[f]	0	1	0	261
	3.2	19	0	17	110[g]	0	7	0	153
	3.3	3	0	1	4	0	0	0	8
	Total	1341	7	125	960	2	49	168	2652

Conclusions

We have compared unsupervised classification methods based on the original bands of a satellite image and using spectral indices derived from them. The results obtained with this research showed that the classification accuracy for coniferous and deciduous forests will increase with the use of spectral indices.

Pixels representing two-storey stands (2.1 in tab. 4) have been classified, using spectral indices, into cluster 1 (apparently "coniferous") or into cluster 4 – "deciduous". Similar results were obtained for one-storey mixed forest stands (2.2) with equal share of scotch pine and oak. Because of this it was not possible to create separate clusters for two-storey stands and for mixed forest. This problem could probably be solved by increasing the number of desired clusters, as this is one of the important factors controlling the ISOCLUST routine. Further research is needed to recognise seasonal changes in spectral reflectance of two-storey and mixed forests. The use of multitemporal images should improve the classification results.

References

Ball G.H., Hall D.J., 1965: A novel method of data analysis and pattern classification. Stanford Research Institute, Menlo Park, USA.

Eastman J.R., 1999: IDRISI32. Guide to GIS and Image Processing. Vol. 1 and 2. Clark Labs, Clark University, USA.

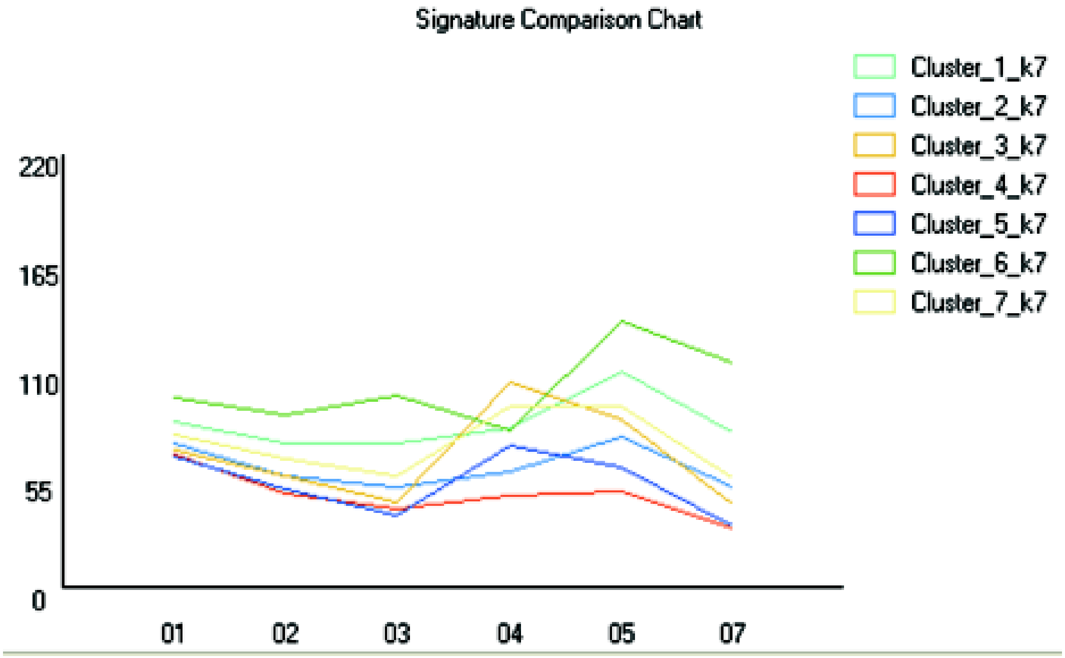
http://www.landsat.org/. FREE Global Orthorectified Landsat Data via FTP. Landsat.org.

Streszczenie

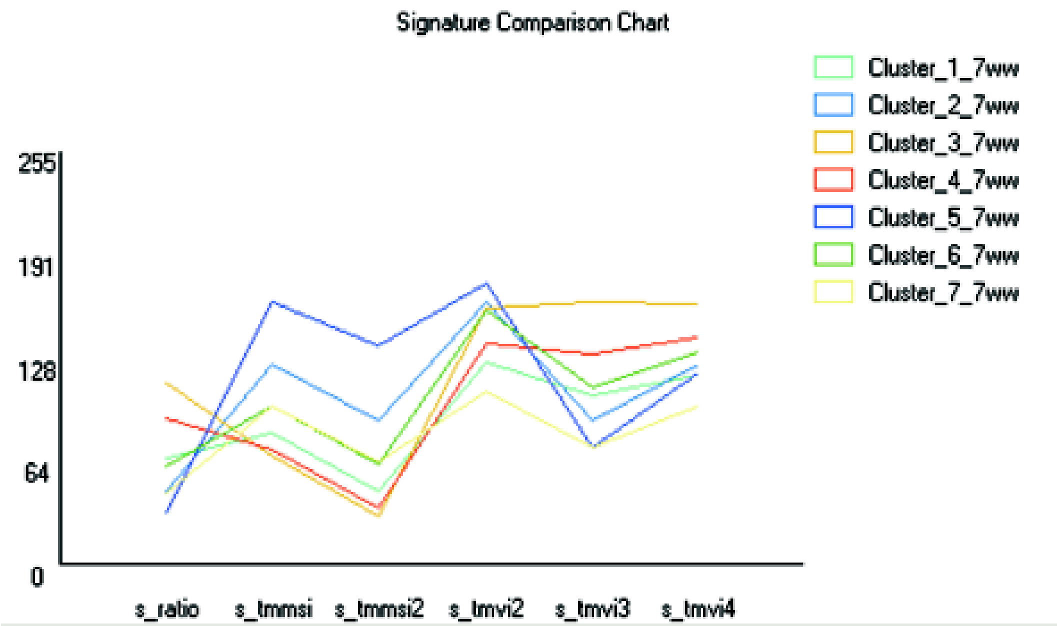
Naturalną formą pokrycia terenu w Polsce są lasy, które zostały jednakże zastąpione w znacznym stopniu przez obszary użytkowane rolniczo – pola uprawne, łąki i pastwiska. Pozostałości lasów tworzą w krajobrazie wyspy, które są połączone siecią korytarzy. Rolę korytarzy pełnią najczęściej lasy i zarośla położone wzdłuż cieków wodnych. Obrazy satelitarne są uważane za efektywne narzędzie służące do rozpoznawania struktury i funkcjonowania krajobrazu na dużych obszarach. Jedną z popularnych metod analizy jest procedura klasyfikacji nienadzorowanej ISODATA. W niniejszej pracy przedstawiono wyniki klasyfikacji sceny Landsat p188r24 za pomocą algorytmu ISOCLUST, który jest implementacją procedury ISODATA w programie IDRISI32. Głównym celem badań było sprawdzenie, czy można poprawić wyniki klasyfikacji nienadzorowanej przez zastosowanie indeksów spektralnych, zamiast oryginalnych kanałów obrazu satelitarnego. Specjalną uwagę zwrócono na odróżnianie drzewostanów mieszanych o złożonej strukturze poziomej i pionowej w sytuacji, gdy scena jest klasyfikowana do niedużej liczby klas. Wykazano, że zastosowanie wskaźników spektralnych poprawia wynik klasyfikacji. Rozpoznano drzewostany jednogatunkowe, jednak nie udało się utworzyć oddzielnej klasy drzewostanów mieszanych – piksele reprezentujące te obiekty były klasyfikowane do klasy drzewostanów iglastych bądź liściastych. Wynik taki otrzymano zarówno w przypadku drzewostanów mieszanych jednopiętrowych, jak i dwupiętrowych – z sosną w piętrze górnym i dębem lub innymi gatunkami liściastymi w piętrze dolnym. Autorzy sugerują, że wynik klasyfikacji można poprawić przez zwiększenie liczby tworzonych klastrów-skupień (parametr procedury ISOCLUST), a także przez uwzględnienie sezonowej zmienności lasów, czyli prowadzenie analizy na obrazach wieloczasowych.

mgr inż. Piotr Krzysztof Stępniewski
kb.ur.35@gmail.com

dr hab. inż. Krzysztof Marek Będkowski
krzysztof.bedkowski@wl.sggw.pl



a



b

Figure 2. a – signatures for forest classes calculated based on original spectral bands, b – spectral indices. The graph shows mean values of pixels in clusters being achieved with ISOCLUST routine

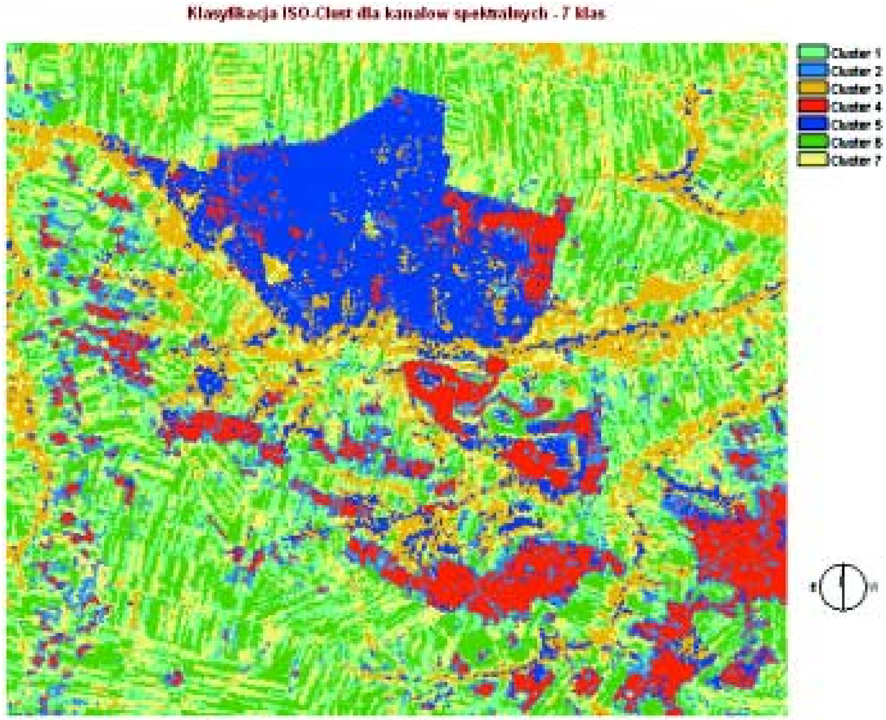


Figure 3. Results of unsupervised classification achieved with ISOCLUST routine using original spectral bands

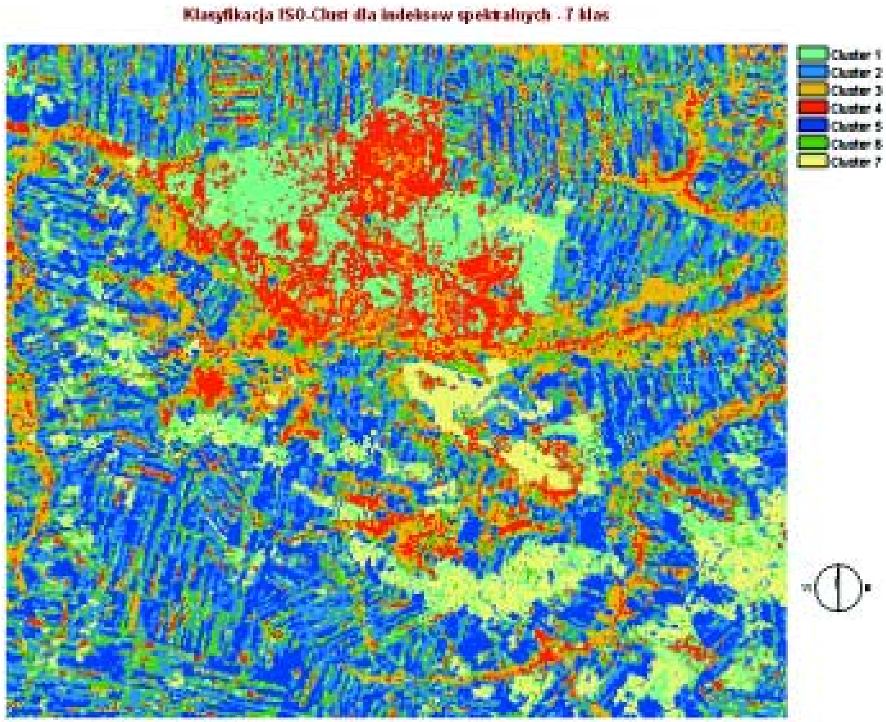


Figure 4. Results of unsupervised classification achieved with ISOCLUST routine using spectral indices