DETERMINING THE OVERBURDEN DUMPS BY SUB-PIXEL METHOD

D. Borisova [a],*, H. Nikolov [a], S. Ivanova [b]

[a] STIL-BAS Acad.G.Bonchev str., bl.3, 1113 Sofia, Bulgaria – (hristo, dborisova)@stil.bas.bg
[b] UMG, Applied Geophyciscs Dept., Sofia, Bulgaria

KEY WORDS: Remote Sensing, Environment, Data Processing, Landsat Data, Classification Accuracy

ABSTRACT:

Delineation of overburden dumps by means of remotely sensed multispectral data with moderate spatial resolution (e.g. Landsat TM/ETM+ 30m) is a challenging task. The major difficulties arise from: 1) large period using the dump (introducing multitemporality); 2) the unknown proportions of vegetation and soil/rock samples in the marginal areas. A variety of methods have been proposed to overcome the problems with impure pixels, but a promising one is the soft classification which assign a pixel to several land cover classes in proportion to the area of the pixel that each class covers. In this scenario for every pixel of the data the correct proportion of the end-members should be found and then co-registered with the corresponding original pixel. As a result this sub-pixel classification procedure generates a number of fraction images equal to the number of land cover classes (end-members). All sub-pixel mapping algorithms have one property in common: accuracy assessment of sub-pixel mapping algorithms is impossible because of missing high resolution ground truth imagery. In this case one possible solution is to use laboratory and in-situ measured spectrometric data. This study presents a successful implementation of soft classification method with additional, precise spectrometric data for determination of dump area of the copper plant. The results were used for proving the in-situ gathered data and coincidence of 93.5% was achieved.

1. INTRODUCTION

Traditional classification techniques are "hard" in the sense that a single pixel is assigned to a single land cover class. At coarse resolutions pixels predictably become mixed. For these mixed pixels "soft" classifiers can be used, which assign a pixel to several land cover classes in proportion to the area of the pixel that each class covers.

Several researchers have addressed this sub-pixel mixture problem. Among the most popular techniques for sub-pixel classification are mixture modeling (Kerdiles and Grondona, 1996) and supervised fuzzy c-means classification (Foody,, 1994).

These techniques aim at estimating the proportions of specific classes that occur within each pixel. The result is a number of fraction images, one for each land cover class concerned. This information describes the class composition, it does not provide any indication as to how this is spatially distributed within the pixel. The result is thus quite different from the classic classification algorithms, where a single land cover map, containing all classes, is produced. One way to go from fraction images to a classic land cover map would be to use "hardeners", where rules can be formulated to determine which class dominates the pixel. The main disadvantage is that information is lost much in the same way as happens with classic hard classifiers. Atkinson (1997) has formulated an idea for an alternative approach. It consists of converting raster data to vector data by threading the vector boundaries through the original image pixels (instead of between pixels, as classic raster-to-vector conversion algorithms would do). This process is called sub-pixel mapping. The key problem is determining where the relative proportions of each class are most likely to occur.

2. METHODS

2.1 Sub-pixel Classification

The spectral unmixing technique that will be applied is based upon Linear Mixture Modeling. The basic hypothesis is that the image spectra are the result of mixtures of surface materials and that each of these components is linearly independent of the other. The mixtures in the image are expressed as linear combinations of their respective spectra in the image. Spectral response of each image pixel in every spectral band can be considered as a linear combination of the response of each component (end-member) present in the mixture. Therefore, every pixel contains information about the proportion (or fraction) and spectral response of each component.

Hence the spectral reflectance r, for every image pixel in any band, can be modeled as follows:

$$r_\Sigma(\lambda_i) = p_1 r_1(\lambda_i) + p_2 r_2(\lambda_i) \dots + p_m r_m(\lambda_i) + \varepsilon = \sum_{j=1}^{m} p_j r_j(\lambda_i) + \varepsilon \qquad (1)$$

where
p = fraction cover,
r for pure component reflectance
λ is wavelength
and ε is an error term.

This equation can become a linear system of n equations and k unknowns in two ways:
- applying the equation to every pixel in a single band (n pixels)
- applying the equation to every band of a single pixel (n bands)

---

* Corresponding author

In order to solve this set of equations, the number of unknowns cannot be larger than the number of equations. In case the linear mixture model is applied to every pixel in a single band, this condition is not likely to be limiting, as the number of pixels will normally largely exceed the number of end-members. In the other case, where the model is applied to every band for a given pixel, this condition is much more likely to be constraining. Thus the spectral dimensions of the image data set will limit the total number of end-members. In hyperspectral remote sensing this limitation generally is no restriction. In multispectral image analysis however this is a serious difficulty. The problem can be addressed by using a multitemporal data set, where the maximum number of end-members equals the number of spectral bands multiplied by the number of dates for which images are available. The first application has been widely used to calculate pure end-member reflectance, given the end-member fractions, while the second application is typically being used to derive end-member fractions, given the pure components reflectance.

In "soft" classifications the estimated variables (the fractions or proportions of each land cover class) are continuous, ranging from 0 to 100 percent coverage within a pixel. Consequently the above mentioned technique is not suitable for accuracy assessment. Nevertheless it is possible to apply statistical techniques that use a comparable approach and yield similar results.

Purposely dividing the known land cover proportions derived from the data with high spectral resolution into two data sets – the first one is used for unmixing model creation and the second one is for validation.

The validation could be achieved by calculating the sample correlation coefficient R, which is the estimator of the correlation coefficient , between the known land cover proportions from the validation set and the soft classification output. In the case of perfect agreement, the correlation coefficient would be 1 and the linear regression line would have a zero intercept and 45 degrees slope.

2.2 Sub-pixel Mapping

The major challenge in sub-pixel mapping is determining the most possible locations of the fractions of each land cover class within the pixel. Creating unmixing model from data with high spectral resolution can solve this. A simple representation of the problem and a possible solution is given in Figure 1 (adapted from Atkinson, 1997).
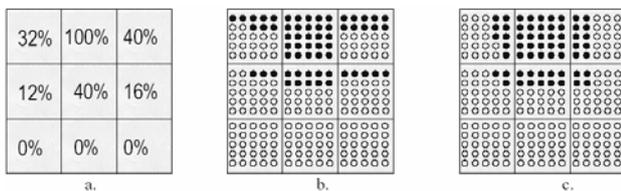

Figure 1. A raster grid of 3 by 3 coarse pixels, each discretised into 5 by 5 sub-pixels

3. STUDY AREA AND DATA

The geological exploration of the copper-bearing rocks in the Sredna gora region, located in the middle of Bulgaria, started in the late 50-ies of 20-th century. As a result the mining plant "Elacite" was built who started its production 1986. The main activity of this plant is production of copper concentrate from chalcopyrite. Two dump sites from the flotation process were established near village Benkovski. They are considered as a main source of contamination of neighboring agricultural areas. As a source for airborne data with moderate spectral resolution for the spectral reflectance of the land cover was used data sets from Landsat TM/ETM+ instrument (GLCF, 2005). The acquisition data is in first decade of June which guarantees good illumination conditions.

From this dataset the two target areas (two open dumps) were extracted by their visual discrimination from the surroundings. This way the dataset necessary for sub-pixel procedure was created.

Laboratory measurements of ore minerals were performed with laboratory spectroradiometer. The instrument provides data in the spectral range 0,5 – 2,55 um with bandwidth less than 0.01 um. The statistic reliability is guaranteed by integrating 10 spectra per sample per measurement. The obtained spectral reflectance of the samples from the dump sites were compared with similar ones taken from spectral library (Clark, 2003)

4. RESULTS

Since in our case the number of mixed pixels (which is the case for pixels from the borders of the areas) is not very big but it is crucial to find every correct pixel. This was achieved by implementing the methodology of the unmixing theory on the data from the dump slag. We also found a correct proportion for the boundary pixels between the slag and the water and between the slag and the surrounding agricultural areas. This helped us to map correctly (determining the surfaces of the dumps) the dump areas thus assessing their environmental impact. The surfaces determined this way allowed us to monitor the rate of reclamation activities for the older dump. These activities are summarized in the Table 1.

Activities	Reclamation as planned	Reclamation as identified
Pilot reclamation of dump boundary (m^2*1000)	50	30
Reclamation of dump boundary and neighboring agricultural areas (m^2*1000)	180	150
Biological purification of the water (NTUs)	80	140

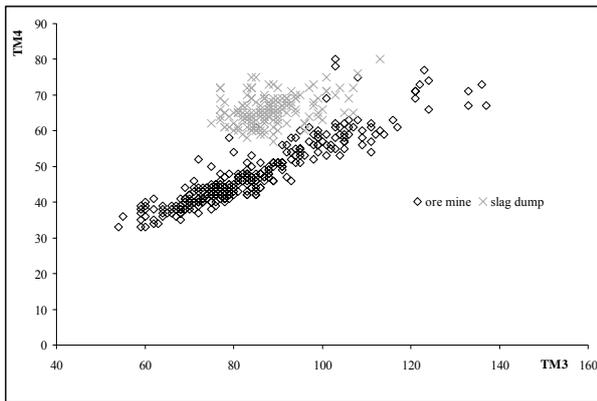Table 1. Reclamation activities in Elacite mine plant dump sites

Figure 2. DN corresponding to red and NIR for open mine and slag dump areas

In Figure 2 the minor overlapping of two areas (open pit mine and slag dump) show that two end-members are easily recognized. The values for dump slag were used in process of unmixing before the sub-pixel mapping was done.

If the ground resolution for the Landsat data is considered to be 28.5 m instead of 30 m and above sub-pixel unmixing is applied on the data the discrepancy between the calculated (see Table 1) and in-situ measured surfaces is less than 2%. This proves that the applied methodology for surface determination of the slag dumps is suitable for similar type of applications.

5. CONCLUSIONS

The above results prove that methodology for sub-pixel classification is capable of estimating the end-member fractions quite accurately.

The main advantage of the presented technique is that mixed pixels are used during the training phase. Compared to these other techniques, the present one is simple, cheap and objective. The results of the sub-pixel mapping exercise indicate that the technique can be useful to increase the resolution while keeping the classification accuracy high.

REFERENCES

Atkinson P. M., M. E. J. Cutler, H. Lewis, 1997. Mapping sub-pixel proportional land cover with AVHRR imagery. *International Journal of Remote Sensing*, vol. 18, no.4, pp. 917-935.

Clark R., G.Swayze, A.Gallagher, T.King, W.Calvin. 2003. The U. S. G. S., Digital Spectral Library Ver. 05: Version 1: 0.2 to 3.0 um. U. S. Geological Survey, http://speclab.cr.usgs.gov (accessed 03 February 2005)

Foody G. M., D. P. Cox, 1994. Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. *International Journal of Remote Sensing*, vol. 15, no. 3, pp. 619-631.

Global Land Cover Facility - University of Maryland, http://glcfapp.umiacs.umd.edu:8080/esdi/index.jsp (accessed 06 April 2006)

Kerdiles H., M. O. Grondona, 1996. NOAA-AVHRR NDVI decomposition and sub-pixel classification using linear mixing in the Argentinean Pampa. *International Journal of Remote Sensing*, vol. 16, no. 7, pp. 1303-1325.