DOCUMENTATION OF GEOGLYPHS ON NAZCA PLAIN, PERU USING REMOTELY SENSED DATA

Eva MATOUŠKOVÁ, Karolina HANZALOVÁ

Czech Technical University in Prague, Faculty of Civil Engineering
Thákurova 7, Prague 6, Czech Republic

eva.matouskova@fsv.cvut.cz, karolina.hanzalova@fsv.cvut.cz

Abstract: *Thanks to spatial resolution improvement of satellite images we now have the potential to document our cultural heritage using remote sensed data. This work deals with the possibilities of evidencing cultural heritage using GeoEye – 1 and QuickBird satellite data. The objects of interest are geoglyphs located around Palpa village in Peru. Geoglyphs are commonly known as large structures built on the Earth's surface. GeoEye – 1 scanner is distinguished by high spatial resolution, which is lower than 0.5m in panchromatic band. The second available image originates from BGIS2000 scanner located on QuickBird satellite and its spatial resolution in panchromatic band is higher – 0.61m. Both sensors acquire data also in multispectral mode and the spatial resolution is then 1.65m resp. 2.44m. Images are geocoded and then different classification methods are applied. In this paper different supervised and unsupervised classification methods are analyzed on both given datasets. The methods in question are then compared with the previously gained vectorized satellite data of the same area. In terms of the future analysis the best classification technique for identifying geoglyphs in the area of interest is found. Influences and demands of the normalized difference vegetation index (NDVI) for the given area are also specified. ENVI software was used for all our analyses.*

1. INTRODUCTION

Geoglyphs are commonly known as drawings produced on the ground. They are created either by accumulation of material (usually stone) or by dispatching the upper layer (10-20cm) of soil. Geoglyphs near Palpa, Peru were analyzed. The city of Palpa is located about 40 km northwest of the Nazca plain, where the best known geoglyphs were found. In the area of interest no figures like "colibri" or "monkey" can be found like on the Nazca plain, but many other geoglyphs are located there. They have the shape of geometric lines, triangles, trapezoids, circles or so-called cikcak lines and they are located in the desert all around the Palpa village.

Remote sensing is being used for a wide range of investigations. Their major advantage is that Earth's surface can be mapped and measured without any in-situ measurements. For easy and fast analysis of remote sensed data, different classification methods can be used. This paper's contribution is to find a suitable classification method for two different, very high resolution satellite images and also to explicate the use of NDVI (Normalized Difference Vegetation Index) in locating geoglyphs.

2. METHODOLOGY

2.1 Data

The data come from two very high resolution satellites – QuickBird-1 and GeoEye-1. Vector data were also used.

The **GeoEye-1** satellite was launched on the 6th of September 2008. It delivers multispectral images (blue, green, red, and near infrared) with a spatial resolution of 2 meters together with panchromatic images with a resolution of 0.5 meters. At the time of release, the GeoEye-1 was the world's highest resolution commercial earth-imaging satellite, and it assures the follow-up of Ikonos imagery. Table [1] shows the sensors specifications. The analyzed image was acquired on 14th of April 2010 and is located in central Peru near the village of Palpa.

The **QuickBird-1** was launched on 18th of October 2001 on Vandenberg Air Force Base, California, USA. Its spatial resolution is 2.62m in multispectral and 0.65m in panchromatic mode. It was designed to support applications ranging from map publishing to land and asset management to insurance risk management [2].

Our resolved data come from the 25th of April 2005 and they are also located near Palpa village, but there is no overlap between the two images. Figure 1 shows the location of both selected images and surrounding areas as seen on Google Earth. Table 1 shows the sensors specifications.

Table 1 - GeoEye–1, QuickBird - Specifications

Spatial resolution	GeoEye - 1	QuckBird - 1
Panchromatic sensor	0.5 m	0.65 m
Miltispectral sensor	2 m	2.62 m
Spectral range	450–510 nm (blue)	43–545 nm (blue)
	510–580 nm (green)	466–620 nm (green)
	655–690 nm (red)	590–710 nm (red)
	780–920 nm (near IR)	715–918 nm (near IR)
Swath Width	15.2 km	18 km
Off-Nadis Imaging	up to 60 degrees	up to 60 degrees
Dynamic range	11 bits per pixel	11 bits per pixel
Revisit time	Less than 3 days	94,2 min
Orbital altitude	681	482 km
Equation crossing time	10:30 am	10:00 am

2.2 Processing

Before intrinsic classifications some pre-processing operation had to be performed. Geometric and radiometric corrections were made by the data producer. QUAC (Quick Atmospheric Correction) was used to get rid of atmosphere effects and for the reflectance – radiance computations. Pixel values of the analyzed data are calibrated radiance values in $W*sr^{-1}*m^{-2}$.

The images were classified into 6 classes – Water, Verdure, Communications, Desert, Buildings, and Geoglyphs. ITT's ENVI 4.8 software was utilized for computations and analysis.

Vector data are made by hand from the given satellite images. Water, communications and different types of geoglyphs were vectorized. For this purpose ESRI's software ArcGIS was used.
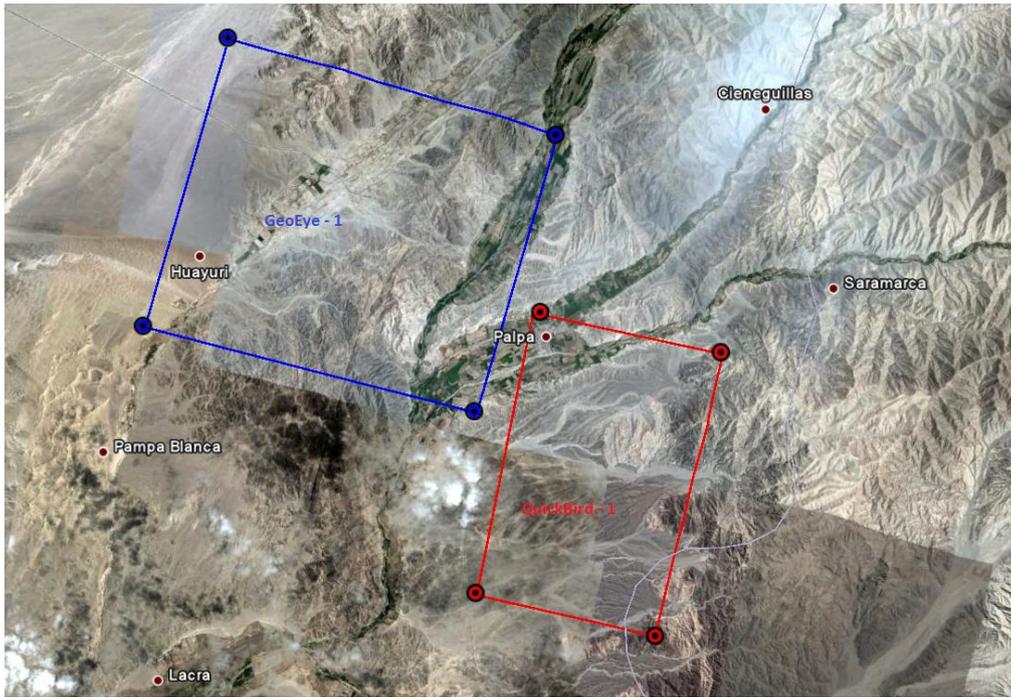
Figure 1: location of analyzed images as seen on Google Earth

2.2.1 Supervised Classification

Parallelpiped classification uses a simple decision rule based upon a standard deviation threshold from the mean of each selected class. If a pixel value lies above the low threshold and below the high threshold for all n bands being classified, it is assigned to that class. If the pixel value falls in multiple classes, ENVI assigns the pixel into the last class classified. Areas that do not fall into any class are designated as unclassified. [5]

Minimum Distance technique uses the mean vectors of each endmember and calculates the Euclidian distance from each unknown pixel to the mean vector for each class. The standard deviation value and distance threshold must be specified. [3]

Maximum Likehood classification assumes that the statistics for each class in each band are normally distributed and calculates the probability that a given pixel belongs to a specific class. Each pixel is assigned into the class that has the highest probability (i.e., Maximum Likehood). All pixels are classified until a probability threshold is individualized. [5]

Mahalanobis Distance classification is a direction sensitive distance classifier that uses statistics for each class. It is similar to Maximum likelihood classification but assumes all class covariances are equal, and, therefore, it is a faster method. Pixels are classified to the closest region of interest (ROI) unless distance threshold is set. [3]

The Spectral Angle Mapper (SAM) is physically-based spectral classification that uses an n-dimensional angle to match pixels to reference spectra. The algorithm determines the spectral similarity between two spectra by calculating the angle between the spectra, treating them as vectors in a space with the number of dimensions equal to the number of bands. Endmember spectra were extracted directly from the image as a region of interest average spectra. [3]

2.2.2 Unsupervised Classification

K-means classification calculates initial class means evenly distributed in the data space then iteratively clusters pixels into the nearest class using a minimum distance technique. Each iteration recalculates class means and reclassifies pixels with respect to new means. All pixels are classified into the nearest class until a standard deviance or distance threshold is reached. This process continues as long as the

number of pixels in each class changes by less than the selected pixel change threshold or the maximum number of iterations is reached.

ISODATA classification method calculates class means evenly distributed in the data space then iteratively clusters the remaining pixels using minimum distance techniques. Each iteration recalculates means and reclassifies pixels with respect to new means. This process continues and ends similar to k-means classification. [3]

2.2.3 NDVI

Normalized Difference Vegetation Index (NDVI) is based on the assumption that chlorophyll (pigment in plant leaves) strongly absorbs visible light (wavelengths from 0.4 μm to 0.7 μm) for use in photosynthesis. On the other hand, the cell structure of the leaves heavily reflects near-infrared light (from 0.7 μm to 1.1 μm). The more leaves a plant has, the more affected these wavelengths are. Therefore, NDVI can quantify the photosynthetic capacity of the vegetation in a given pixel. [4]

Written mathematically, the formula is:

$$NDVI = (NIR - VIS) / (NIR + VIS) \tag{1}$$

2.2.4. Filtering

2.2.4.1. Texture filters

Images contain regions characterized by variation in brightness rather than any unique brightness value. Texture filters are based on the frequency of change and spatial arrangement of pixel values in an image, as a function of spatial scale. To be defined as a distinct textural area, the gray levels within that area must be more homogeneous as a unit than areas having different texture. [5]

Occurrence Measures applies five (data range, mean, variance, entropy and skewness) different texture filters that are based on occurrence measures. They use the number of occurrences of each grey level within the processing window for the texture calculations.

Co-occurrence measures apply texture filters (mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment and correlation) that are based on the co-occurrence matrix. It uses a gray-tone spatial dependence matrix to calculate texture values. [3]

2.2.4.2. Convolution filters

High Pass filter removes low frequency components of an image while retaining the high frequency. It enhances edges and sharpens an image. It uses a kernel with high central value, typically surrounded by negative values.

Laplacian filter is a second derivative edge enhancement filter that operates without regard to edge direction. The filter stresses maximum values within the image by using a high central value kernel surrounded by negative weights in north-south and east-west directions together with zero values at the kernels corners.

Directional filter is first derivative edges enhancement filter that selectively enhances image features having specific direction components (gradients). The sum of the directional filter kernel elements is zero.

Sobel filter is a non-linear edge enhancement, a special case filter that uses approximation of the true Sobel function. This function returns an approximation to the Sobel edge enhancement operator for images [6]. Sorbel's operator calculates the gradient of the image intensity at each point and gives the direction of the largest possible increase from light to dark and the rate of chance in that direction.

Roberts filter is a non-linear edge detector similar to Sobel. It is a special filter that uses a present 2x2 approximation of the true Roberts function. It is a simple 2D differencing method. The idea is to approximate the gradient of an image through discrete differentiation which is achieved by computing the sum of squares of the differences between diagonally adjacent pixels. [7]

3. RESULTS

3.1 Supervised Classification

Parallelpiped classification was the first one to analyze, and is also the fastest method used. With setting standard deviation threshold to 0.5, asphalt communications are very well classified, but geoglyphs are not so well visible. According to this knowledge, the standard deviation threshold was set to 1 for all classes. This configuration shows the best results for detecting geoglyphs even if there are some mistakes e.g. building were classified as dried up river-basins and highest peaks of local mountains. The biggest difference is in the desert. The threshold equal to 1 specifies these class parameters much better than if equal to 0.5. It is because the desert in the area of interest does not consist of the sand only, but also of different types of stones, so the threshold should not be so strict. On the other hand, if the threshold is set too robust, other classes like buildings are classified as desert. Figure 2B and 3B shows the resultant image of this classification on different geoglyph location.

For **Minimum Distance** classification, different standard deviation values and distance thresholds had to be tested. The best value for both thresholds was found and set to 1000. A geoglyph can be found using this method, but not all have been classified to the right class (due to their spatial differences). With this method desert at higher altitudes was classified as communications and vegetation was also classified incorrectly.

Processing **Maximum Likehood** classification was very tricky. Many tests were run in order to find the best parameters. It was found, that Maximum Likehood classification fits the best without any thresholds. If the threshold was set to minimum, i.e. to 0.001, only two classes were classified – the desert and geoglyphs. The geoglyph classes do not match the vectors, so the results are very poor. Without any threshold the asphalt communications were classified correctly, but high altitude peaks were wrongly put in the same class. Buildings were classified in the same class as dried up river-basins and rural roads. Geoglyphs were classified very incorrectly; therefore this method is not good for their detection. On the other hand verdure was classified very well.

In **Mahalanobis Distance** classification asphalt communications and verdure were classified correctly. The right threshold was set to 5 for all classes. Buildings were classified as dried up river-basins and rural roads. This method is a mismatch for detecting geoglyphs because they are not visible on the classified image.

The last supervised classification method was **Spectral Angle Mapper**. The setting of thresholds was very complicated. After many investigations 0.1 was set as the best one. Verdure and asphalt communications were very well classified, and only buildings were classified as dried up river-basins and rural roads. Generally the classification itself was good, but geoglyphs are hardly visible when using this method.

3.2 Unsupervised Classification

The outcome of both unsupervised classifications – **K-means** and **ISODATA** is similar. The classification was made for six classes that represent the same ones as those selected for supervised classifications, i.e. water, verdure, buildings, desert, communications and geoglyphs, and the number of iterations was set to 5. Change threshold value was equal to 5% in ISODATA classification and 1% for K-means. The maximum standard deviation from the mean and the maximum distance error were not set, because they did not lead to better results.

These methods excellently detect sharp lines and edges, so they are good for disclosing geoglyphs, because they consist of such precise arrows. Due to their differences, geoglyphs were not classified into the same class, but they are perfectly visible on the consequent image. Other classes were not distributed correctly, which is an obvious drawback when using unsupervised classifications. Figure 2C and 3C shows this classification's resultant image on two different geoglyphs.

3.3 Normalized Difference Vegetation Index (NDVI)

NDVI was calculated for the GeoEye-1 image. Light areas show higher NDVI, which represents more vegetation. Because geoglyphs do not have any chlorophyll, they cannot be seen on this image and therefore they cannot be found using NDVI.
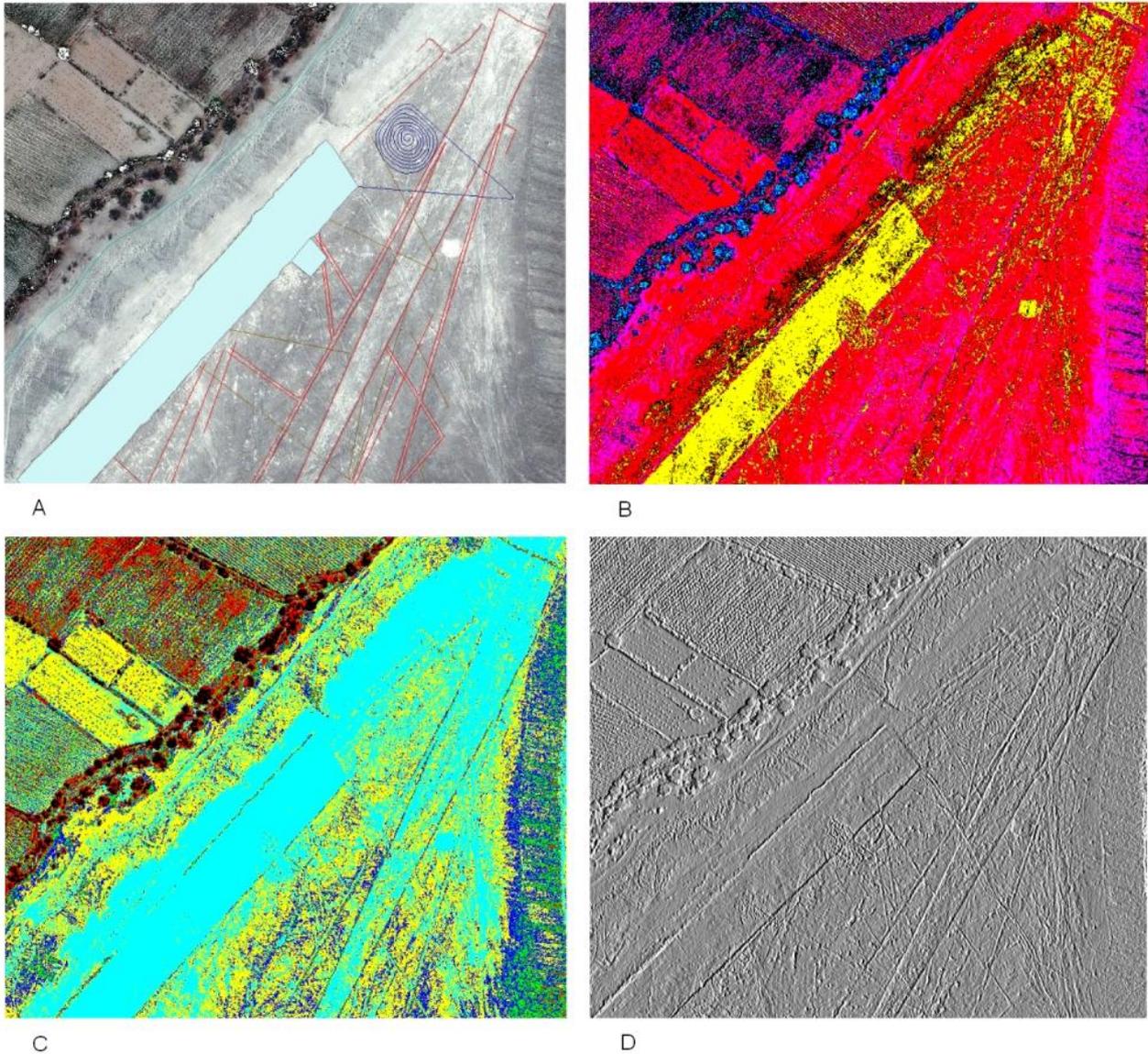

Figure 2: A: vectors illustration of geoglyph; B: supervised classification - Parallelpiped; C: unsupervised classification - ISODATA; D: Filtering – directional filter

3.4 Filtering

3.4.1. Texture filters

Occurrence and co-occurrence filters can detect areas very well, but they fall behind in retrieval of lines. Occurrence filters compute data range, the mean, variance, entropy, and skrewness. Co-occurrence filters outcome is composed of the mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation. Variance, homogeneity, contrast and dissimilarity images are the best ones for locating geoglyphs.

3.4.2. Convolution filters

Directional 3x3 filter was found optimal to detect geoglyphs. The outcome images are very rounded and the areas of interest are clearly visible at first sight. Highpass filters are better in detecting narrow lines, but larger areas can be found using these filters as well. Sobel and Roberts filters give similar results, but the return cannot match directional or highpass filters. The Laplacian filter is an average among convolution filters, but it gives sufficient results. Lines are not so sharp as in other cases, but they are still suitable for geoglyph detection. Resultant image for two locations can be seen on Figure 2D and 3D.
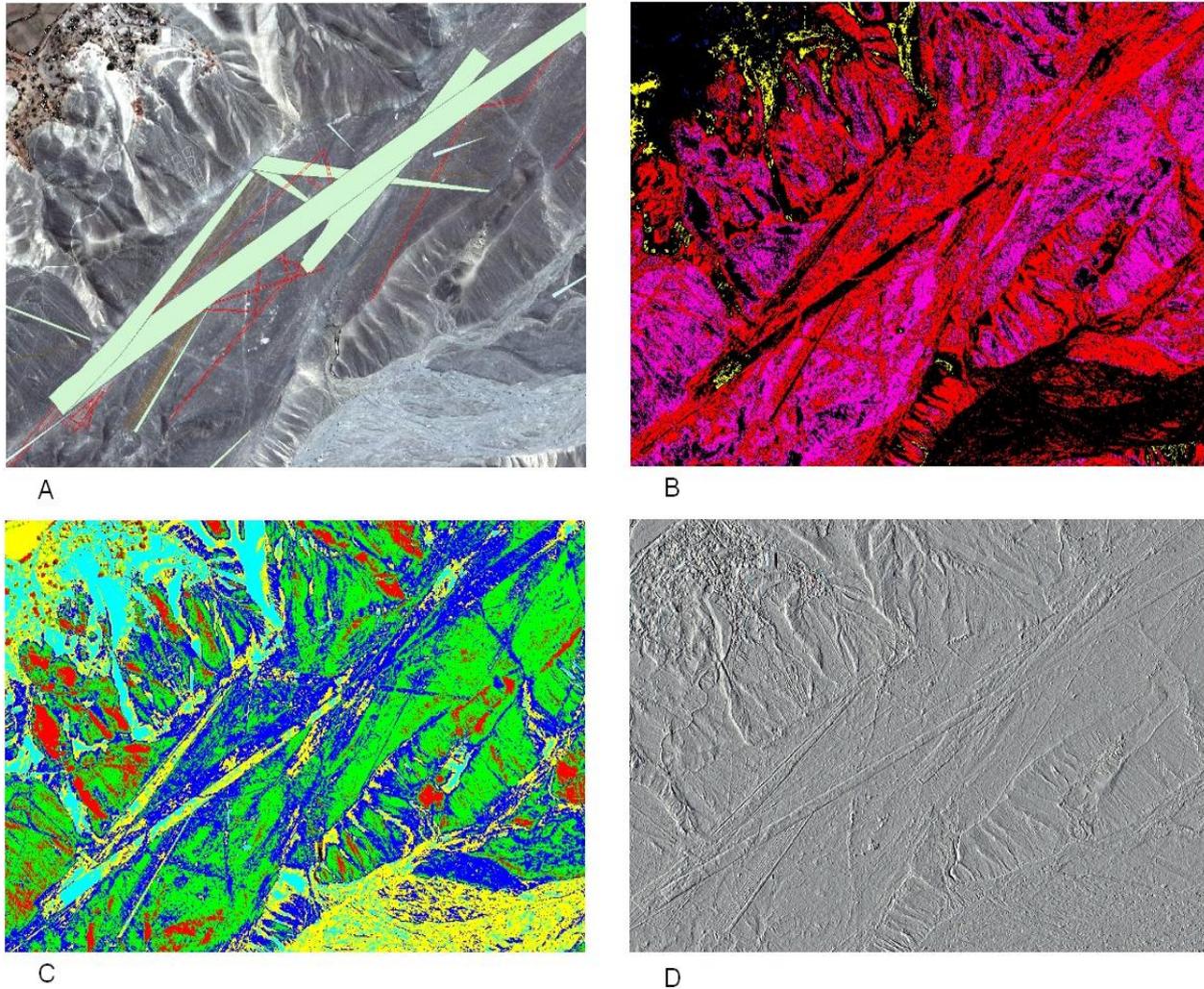

Figure 3: A: vectors illustration of geoglyph; B: supervised classification - Parallelpiped; C: unsupervised classification - ISODATA; D: Filtering - directional filter

4. CONCLUSIONS

It was seen that classification methods can be used for detecting geoglyphs only with great care, because their results can be misguided. Classification outcomes must always be compared with aerial or satellite imagery to check if the geoglyph locations match. The reason is that reflectance for geoglyphs and the desert is not so diverse and the desert is then classified as a geoglyph and vice versa. Classification was used for extraction of partial primitives from multiimage. The best results of classification methods for determining geoglyphs are supervised Parallelpiped and Minimum Distance classification and both unsupervised classifications (ISODATA and K-means).

Normalized Difference Vegetation Index cannot be used for determining geoglyphs, because they do not contain any chlorophyll particles; therefore they are not visible on the image, but it is very easy to see

areas with vegetation. This is very important in our area of interest, because where there is vegetation there is often water.

Filtering was found as the best way to detect geoglyphs in a desert area. Lines detections are more unique and are very well seen by the naked eye. Using this technique is more time-consuming but gives better results. The bigger the filtration kernel the narrower and small lines are lost during the process. Directional filter with 3x3 kernel size was the best for geoglyph location. Textural filter did not satisfy the requirements; there was much noise and the geoglyphs were not clearly visible.

5. REFERENCES

[1] *GeoEye* [online]. 2011 [cit. 2011-06-13]. Imagery sources. Available on <http://www.geoeye.com/CorpSite/products-and-services/imagery-sources/>.

[2] *DigitalGlobe* [online]. 2011 [cit. 2011-06-13]. QuickBird. Available on <http://www.digitalglobe.com/index.php/85/QuickBird>.

[3] *ENVI Tutorial* [online]. 2011 [cit. 2011-06-13]. Classification methods. Available on <http://www.ittvis.com/portals/0/tutorials/envi/Classification_Methods.pdf>.

[4] NASA, Gov. *Earth Observatory* [online]. 2011 [cit. 2011-06-13]. Measuring vegetation. Obtainable on: <http://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_2.php>.

[5] COLWELL, Robert, N; AMERICAN SOCIETY OF PHOTOGRAMMETRY *Manual of Remote sensing : Volume I.*. 2nd edition. Falls Church, Virginia, USA : American society of photogrammetry, 1983. 1230 s. ISBN 0-937294-41-1.

[6] *The IDL Astronomy User's Library* [online]. 2007 [cit. 2011-06-15]. SOBEL. Available on WWW: <http://idlastro.gsfc.nasa.gov/idl_html_help/SOBEL.html>.

[7] COSTELLA, John, P. *Assassination science* [online]. 26.2.2011 [cit. 2011-06-15]. A superior edge detection gradient operator. Available on WWW: <http://assassinationscience.com/johncostella/edgedetect/>.

[8] Pavelka, K. - Svatušková, J. - Bukovinský, M.: Using of VHR satellite data for potential digs localisation and their verification using geophysical methods. In Proceedings of 1st EARSeL International Workshop on "Advances in Remote Sensing for Archaeology and Cultural Heritage Management. Paris: Earsel, 2008, vol. 1, p. 169-175. ISBN 978-88-548-2030-2

[9] Pavelka, K.: Rectification of Petroglyphs with photogrammetrical methods. Dresdener Kartographische Schriften. 2007, vol. 25, no. 12, p. 103-114. ISSN 1436-0004

[10] Pavelka, K.: Using of Remote Sensing and Photogrammetry for Documentation of Geoglyphs in Nazca/Peru Region. In Proceedings of 29th ACRS. Tokyo: URSI, 2008, vol. 1, p. 253-256

[11] Pavelka, K. - Klokočník, J.: Nasca,Peru: Der Astronaut ist ein Fischer. Amerindian research. 2010, Jahr. 5, nr. 2, S. 112-114. ISSN 1862-3867

[12] Pavelka, K.: Topographic and Thematic Mapping from Multi-Resolution Satellite Image. In Proceedings of 24th ICA Conference. Santiago: International Cartographic Association, 2009, vol. 1, p. 956-962

[13] Pavelka, K. - Bukovinský, M. - Svatušková, J. - Hanzalová, K.: Documentation of the Abandoned Town La Ciudad Perdida In Peru Combining of VHR Satellite and Terrestrial Mesurement. In Proceedings The 30th EARSeL Symposium [CD-ROM]. Paris: Earsel, 2010, p. 126-137. ISBN 978-3-00-033435-1