

FEATURE EXTRACTION IN HIGH-RESOLUTION RASTER IMAGES USING NEURAL NETWORKS

Adam IWANIAK¹, Tomasz KUBIK²,
Witold PALUSZYŃSKI², Przemysław TYMKÓW¹

¹Faculty of Environmental Engineering and Geodesy,
Agricultural University of Wrocław, Poland

²Institute of Computer Engineering, Control and Robotics,
Wrocław University of Technology, Poland

1 INTRODUCTION

Neural networks have already been applied to a range of problems in cartography and GIS, including: thematic (semantic) classification of the attributes of objects, classification of spatial objects, aggregation and generalization of objects, and others (Meng, 1993; Müller, 1992; Openshaw et al., 1991). Among the key properties of neural networks which make them useful in such applications are: their ability to process data with no well-defined data model, or even a good understanding of the problem nature, the ability to learn data processing algorithms without programming, to generalize, the ability to eliminate glitches and errors in data.

One specific area is the application of neural networks to pattern recognition and land use classification in raster images, particularly satellite images (Meng, 1993). The approaches here can be divided into: per-pixel classification methods, which concentrate on increasing the spatial accuracy limited by the available data resolution, and per-parcel, or per-field, classification methods, which combine the per-pixel techniques with further processing, often statistical classification filters (Berberoglu et al., 2000; German and Gahegan, 1996; Liu et al., 2002; Miller et al., 1995; Tatem et al., 2001; Venkatesh and S. Kumar-Raja, 2003).

The identification of textures was introduced to raster image processing (Haralick, 1979) and gradually developed into a mature technology (Haralick and Shapiro, 1992). Texture information has been recognized as a useful addition to spectral data obtained directly from images to help in the task of segmentation of images (Barr and Barnsley, 1999; Berberoglu et al., 2000; Carr, 1996; Clausi and Zhao, 2002; Haralick and Shapiro, 1992; Ryherd and Woodcock, 1996; Zhang, 1999). Various combinations of the above approaches have been also applied and described previously by some of the authors of this paper (Iwaniak et al., 2003, Iwaniak et al., 2005).

2 MOTIVATION

During the last decade the geographic information systems in Poland started utilizing aerial and satellite images on a large scale. Due to the implementation of the IACS project, high-resolution aerial photographs and, partially, satellite images of most of the country are being produced, and, for the first time, the whole country is covered with a current orthophoto map. Furthermore, within the scope of this project it is expected that around 20% of the country will have a new orthophoto map every year. Aside of

their utilization in the verification in the farmers subsidies system, these data can be used in other applications. In Poland the work is in progress on the National Geographical Information System and the Spatial Data Infrastructure. One of the components in the system's design is the Topographic Data Base (TDB), which make up the reference data. This component is responsible for doing several tasks, including topographic data collection on the regional level. It contains topographic objects related to the topographic maps in 1:10000 scale. By assumption these objects are retrieved from the updated orthophotomaps.

These efforts will create an opportunity and potential for creating and having the most actual TDB, but under the assumption that large amount of data will be effectively analyzed. As 10 thousand sheets of orthophotomaps are going to be processed every year, the need for supporting software tools is obvious. Such tools could be used for regions of special interest highlighting, like building, agriculture areas, etc., as well as in the process of database construction and verification. It could be also used for database updating, pointing outs parts of the map that should be updated by comparing the vector map created from the segmented image with the corresponding vector map stored in the TDB.

The need for new and effective techniques for the analysis and feature recognition in high-resolution images is therefore evident. The questions that need to be answered are: what algorithms can be applied, what kind of image characteristics should be used (eg.: RGB intensities, spectral intensities, various statistical features, textural features, etc.), and how to evaluate the processing results.

3 METHODS UNDER CONSIDERATION

3.1 Neural Networks

Feed-forward artificial neural networks were used in the method proposed. They were designed and trained as a single class classifiers. Their architecture was dictated by the kind of image processing applied. Thus the number of units in the networks input layer varied according to the method applied (size of the mask and the number of features considered, see Fig.1). All input values were normalized to the range of 0.0 to 1.0. There were two hidden layers in the networks, with 9 units for each layer. Networks were trained to recognize only one class, thus only one output unit was needed.

The artificial neural networks were trained with RProp (Resilient Propagation) learning algorithm. The training pattern, consisting of the set of provided input and expected output values, was created in the following way. At first an original training image was classified by an expert. The resulting image contained regions marked with labels corresponding to the classes assigned. Then a set of reference binary images was created by regions with the same label separation. The number of images was equal to the number of classes distinguished. Next for each considered position (for each pixel) in the original image the set of input values was computed according to the method applied. The set of output values consisted of a single binary value indicating whether the pixel of interest should be assigned to the particular class or not. This information was taken from a corresponding reference image.

After training phase artificial neural network was employed as a classifier. Testing pattern used then was prepared in the same way as the training pattern. The results of classification for the testing pattern (not used in the training process) and the training

pattern were stored in the resulting images. The resulting images and the accompanying reference images were used to evaluate classification results.

3.2 Texture features based on GLCM

Gray level co-occurrence matrix (GLCM) method allows estimation of an image features related to second-order statistics. The method starts with a counting of the number of occurrences of the pair of gray levels which are a distance apart in a given direction:

$$V_{l,\alpha}(i,j) = |\{(r,s),(t,v): I(r,s) = i, I(t,v) = j\}|$$

where $i, j = 0, \dots, N-1$; N – number of gray levels; l, α – distance and direction angle; $I(x,y)$ – image pixel at position (x,y) ; $(t,v) = (r + l \cos \alpha, s + l \sin \alpha)$. Then a symmetrical matrix $V'_{l,\alpha}$ is computed by adding $V_{l,\alpha}$ with its transpose divided by 2. After normalization of $V'_{l,\alpha}$ with a total number of pixels in the image:

$$P_{i,j} = \frac{V'(i,j)}{\sum_{i,j=0}^{N-1} V(i,j)}$$

resulting matrix P reveals certain properties about the spatial distribution of the gray levels in the image. On the base of it several features can be calculated. The features used in experiments were following:

Texture features based on GLCM						
ASM	energy	entropy	contrast	homogeneity	dissimilarity	max
$\sum_{i,j=0}^{N-1} P_{i,j}^2$	\sqrt{ASM}	$\sum_{i,j=0}^{N-1} P_{i,j} i-j $	$\sum_{i,j=0}^{N-1} P_{i,j} (i-j)^2$	$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2}$	$\sum_{i,j=0}^{N-1} P_{i,j} i-j $	$Max(P_{i,j})$

3.3 Measures of classification quality

For the confusion matrix $A = [a_{ij}]$, where a_{ij} is a number of sample pixels from the j th class that have been classified as belonging to the i th class, the following measures have been proposed (Cohen, 1960; Cohen, 1968; Hubert-Moy et al., 2001):

- user's accuracy of class i : $u_i = a_{ii} / a_{ri}$, where $a_{ri} = \sum_j a_{ij}$ (sum of i th row entries)
- producers accuracy of class i : $p_i = a_{ii} / a_{ci}$, where $a_{ci} = \sum_j a_{ji}$ (sum of i th column entries);
- overall accuracy of the method: $d = a_{ii} / a_t$, where a_t is the total number of pixels;
- Kappa coefficient: $\kappa = (P_o - P_e) / (1 - P_e)$, where $P_o = \sum_i a_{ii} / a_t$, $P_e = a_{ii} / a_t^2$;
- averaged error: $uDW = 1/K \sum_i \varepsilon_i$, where $\varepsilon_i = 1 - u_i$ (if $a_{ri} > 0$) or $\varepsilon_i = 0$ (if $a_{ri} = 0$) and K - number of classes.

4 IMAGE PROCESSING RESULTS

The methods proposed were tested for color aerial photographs, taken with 1m resolution. Training patterns were created on the base of original and reference images of size 2500x2500 pixels each. The images used for testing pattern creation were of size 1000x2500 pixels.

There were two kinds of processing performed (see Fig. 1):

- classifying on the base of the set of color pixels defined by the position of a mask moving over original photographs,

- classifying on the base of a single pixel color supported by texture features computed from GLCM in the region covered by the mask.

In both cases a position of the mask's center was a position of the resulting pixel.

For the first type of processing the following masks were used: 7x7, 15x15, 25x25. In this approach the neural network was feed with a data which were the pixels color intensities defined by the placement of the moving window. Thus for a mask of the size 7x7 and three color channels (RGB) there were 3*7*7 input values.

For the second type of processing the masks used were of size 30x30. GLCM in this case was computed after transforming pixels colors to the gray scale with 32 levels. The co-occurrence was search in three directions: 45°, 90°, 135°, 180°, in the distance of 1 pixel. Such texture features as: AMS, energy, entropy, contrast, homogeneity, dissimilarity were computed and averaged. This resulted with 6 values which were treated as additional inputs to the neural networks. Together with 3 values corresponding to the colors of pixel from an original image there were 9 input values in total provided to the neural network.

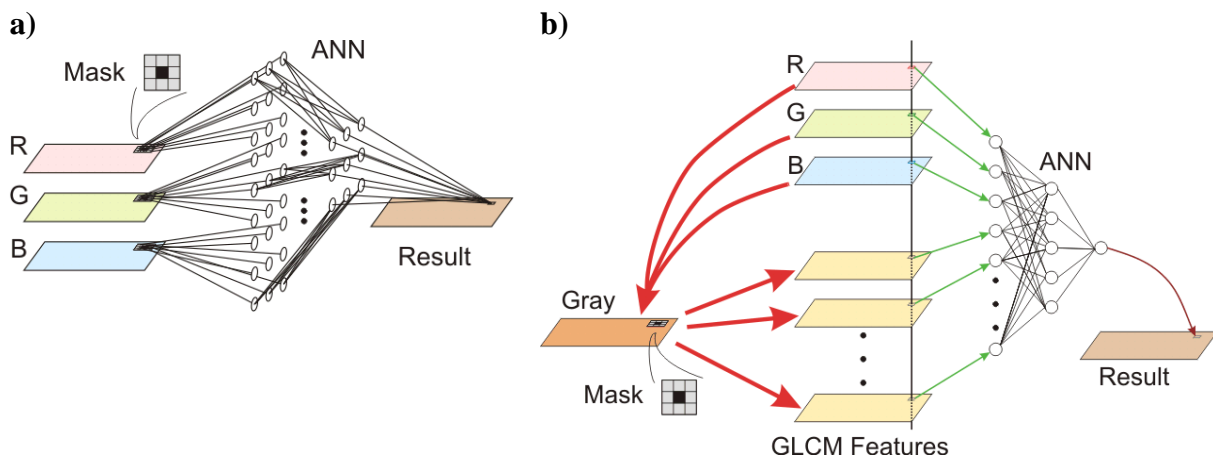


Fig. 1. The use of artificial neural networks for high-resolution raster image processing
 a) processing with a mask moving over color image (color channels are processed in parallel, thus in one position for a mask of size 7x7 there are 3*7*7 corresponding normalized values which are treated as an input to the neural network)
 b) processing using colors of a single pixel supported with a texture features computed on the base of the corresponding gray image (texture features are computed within an area covered by the mask moving over gray image)

The test with image processing were performed for the following classes: Trees, Buildings, Forests, Fields. The results of image processing are shown in the Fig.2 (for the training process) and Fig.3 (for the testing process). In these figures only parts of images processed are shown. Each part in Fig.2 is a square with 512x512 pixels in size and (what in the field was a square with a side width equals 512m). Each part in Fig.2 is a square with 500x500 pixels in size and (what in the field was a square with a side width equals 500m). To distinguish reference images from images representing results of classification former ones are painted in reverse.










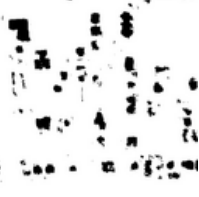



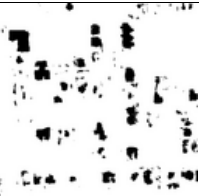



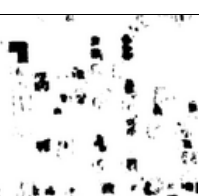



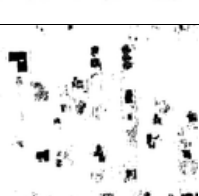

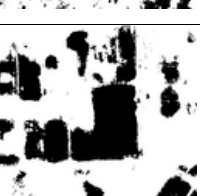
		Trees	Buildings	Forests	Fields
Original image					
Reference image					
Re su lts	25x25 RGB				
	15x15 RGB				
	7x7 RGB				
	GLCM RGB				

Fig. 2. Training data (raster images with resolution of 1m) and image processing results. Parts of images shown are of size 512x512 pixels















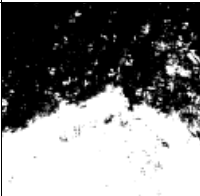

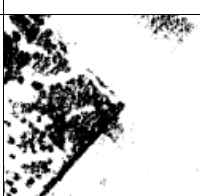



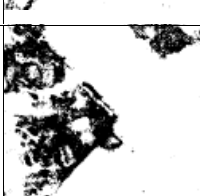



		Trees	Buildings	Forests	Fields
Original image					
Reference image					
Re su lts	25x25 RGB				
	15x15 RGB				
	7x7 RGB				
	GLCM RGB				

Fig. 3. Testing data (raster images with resolution of 1m) and image processing results. Parts of images shown are of size 500x500 pixels

Table 1: Numerical evaluations of results (numbers representing quality measures are computed for the whole images)

Type	Mask	Quality measures for the examples of <i>Testing Set</i>						
		κ	u_1	u_2	p_1	p_2	d	uDW
Forest	7x7	0.857	0.997	0.810	0.635	0.999	0.857	0.192
	15x15	0.926	0.995	0.894	0.815	0.997	0.926	0.111
	25x25	0.939	0.995	0.912	0.849	0.997	0.939	0.093
Fields	7x7	0.898	0.788	0.924	0.701	0.949	0.898	0.288
	15x15	0.888	0.759	0.919	0.690	0.941	0.888	0.322
	25x25	0.877	0.717	0.918	0.691	0.926	0.877	0.366
Trees	7x7	0.914	0.348	0.941	0.218	0.968	0.915	0.710
	15x15	0.926	0.441	0.937	0.139	0.986	0.926	0.622
	25x25	0.920	0.418	0.946	0.288	0.969	0.920	0.636
Buildings	7x7	0.975	0.806	0.977	0.325	0.997	0.975	0.217
	15x15	0.976	0.783	0.979	0.392	0.996	0.976	0.237
	25x25	0.977	0.753	0.982	0.461	0.995	0.977	0.265

5 CONCLUSIONS

Various architectures of feed-forward neural networks have been experimented with. Along with RGB intensities, including a pixel's surrounding neighborhood (mask), textural features have been used as inputs for classification. Visual evaluation was used to judge the results, as well as more objective statistical classification error measures. The results from these experiments were good or excellent for some types of spatial features and just satisfactory for some other features. The visual results can be examined in Figure 3, and the statistical quality measures are computed and collected in Table1. For example, the kappa coefficient comes out best for the Buildings class.

As can be expected, different neural network configurations and processing parameters (mask size), are required to achieve optimal results for various kinds of spatial features. However, for very large mask sizes computation was not reasonable due to the processing time.

The addition of texture parameters were expected to enhance processing and feature recognition. This enhancement was realized in some experiments for some features, but for some other features, there was degradation in their recognition. For example, buildings in figure 2 came out better without GLCM (texture) with most mask sizes. Trees and fields, however, benefit from the consideration of textures. On the other hand, it is unclear which mask size should be compared to GLCM results, which took into account only single-pixel RGB data, and a 32x32 texture field.

We consider the results presented highly promising, although these experiments should be treated as only preliminary, and more work is needed to further explore the problem. Some areas which require more work are:

- comparing our results to those obtained by traditional approaches, eg. manual, simple spectral analysis, filtering, etc.
- development of guidelines for (semi-)automatic neural network configuration and training for a specific processing task
- image preprocessing to account for: differences in image resolution, spectral/color balance characteristics of the imaging system, light quality, etc.

6 ACKNOWLEDGEMENTS

This work was supported in part by a grant from the Polish Ministry of Science and Higher Education number 4T12E 010 30

Computing-intensive calculations have been performed using the computers of Wrocław Networking and Supercomputer Center.

REFERENCES

1. Barr, S. and Barnsley, M.J., 1999. A syntactic pattern recognition paradigm for the derivation of second-order thematic information from remotely sensed images. In: P.M. Atkinson and N.J. Tate (Editors), *Advances in Remote Sensing and GIS Analysis*. Chichester: Wiley, pp. 167-184.
2. Berberoglu, S., Lloyd, C.D., Atkinson, P.M. and Curran, P.J., 2000. The integration of spectral and textural information using neural networks for land cover mapping in the Mediterranean. *Computers and Geosciences*, 26(4): 385-396.
3. Carr, J.R., 1996. Spectral and textural classification of single and multiple band digital images. *Computers and Geosciences*, 22(8):849-865.
4. Clausi, D.A. and Zhao, Y., 2002. Rapid extraction of image texture by co-occurrence using a hybrid data structure. *Computers and Geosciences*, 28(6): 763-774.
5. Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20: 37-46.
6. Cohen, J., 1968. Weighted kappa: nominal scale agreement with provision for scale disagreement or partial credit. *Psychol. Bull.*, 70: 213-220.
7. German, G.W.H. and Gahegan, M.N., 1996. Neural network architectures for the classification of temporal image sequences. *Computers and Geosciences*, 22(9): 969-979.
8. Haralick, R.M., 1979. Statistical and structural approaches to texture. *Proceedings of IEEE*, 67(5): 786-803.
9. Haralick, R.M. and Shapiro, L.G., 1992. *Computer and Robot Vision*, vol. 1., 1. Addison-Wesley Publishing, Reading, MA.
10. Hubert-Moy, L., Cotonnec, A., Du, L.L., Chardin, A. and Perez, P., 2001. A Comparison of Parametric Classification Procedures of Remotely Sensed Data Applied on Different Landscape Units. *Remote Sensing of Environment*, 75(2): 174-187.
11. Iwaniak A., Krówczyńska M., Paluszyński W., Using Neural Networks for Urban Area Classification in Satellite Images, EARSEL, Ghent, Belgium 2003.
12. Iwaniak A., Kubik T., Paluszyński W., Tymków P., Classification of Features in High-resolution Aerial Photographs Using Neural Networks, w: XXII International Cartographic Conference (ICC'2005) , ISBN:0-958-46093-0, A Coruña 2005.
13. Liu, X.-H., Skidmore, A.K. and H. Van Oosten, 2002. Integration of classification methods for improvement of land-cover map accuracy. *ISPRS Journal of Photogrammetry and Remote Sensing*, 56(4):257-268.
14. Meng, L., 1993. Application of neural network in cartographic pattern recognition, *Proceedings 16th International Cartographic Conference*, Cologne, pp. 192-202.
15. Miller, D.M., Kaminsky, E.J. and Rana, S., 1995. Neural network classification of remote-sensing data. *Computers and Geosciences*, 21(3): 377-386.

16. Müller, J.-C., 1992. Parallel distributed processing: An application to geographic feature selection, Proceedings Fifth International Symposium on Spatial Data Handling, Charleston, SC, pp. 230-240.
17. Openshaw, S., Wymer, C. and Cross, A., 1991. Using neural nets to solve some hard analysis problems in GIS, Proceedings EGIS '91, Brussels, pp. 797-807.
18. Ryherd, S. and Woodcock, C., 1996. Combining spectral and texture data in the segmentation of remotely sensed images. Photogramm. Eng. Remote Sens., 62: 181-194.
19. Tatem, A.J., Lewis, H.G., Atkinson, P.M. and Nixon, M.S., 2001. Super-Resolution Target Identification from Remotely Sensed Images using a Hopfield Neural Network. IEEE Transactions on Geoscience and Remote Sensing(39): 781-796.
20. Venkatesh, Y.V. and S. Kumar Raja, 2003. On the classification of multispectral satellite images using the multilayer perceptron. Pattern Recognition, 36(9): 2161-2175.
21. Zhang, Y., 1999. Optimisation of building detection in satellite images by combining multispectral classification and texture filtering. ISPRS Journal of Photogrammetry and Remote Sensing, 54(1):50-60.