JUNCTION EXTRACTION BY ARTIFICIAL NEURAL NETWORK SYSTEM – JEANS

Arpad Barsi [a,b], Christian Heipke [b], Felicitas Willrich [b]

[a] Budapest University of Technology and Economics, Dept. of Photogrammetry and Geoinformatics, H-1111 Budapest, Muegyetem rkp. 3, Hungary – barsi@eik.bme.hu
[b] University of Hannover, Institute for Photogrammetry and GeoInformation, D-30167 Hannover, Nienburger Str. 1, Germany – {barsi,heipke,willrich}@ipi.uni-hannover.de

Commission III, Working Group III/4

KEY WORDS: Artificial Neural Networks, Automatic Object Extraction, Road Junctions

ABSTRACT: The paper presents a road junction operator, which was developed for medium resolution black-and-white orthoimages. The operator uses a feed-forward neural network applied for a running window to decide whether it contains a 3- or 4-arm road junction or not. The training set was created by a data analysis based feature selection. The best features took part in the training of 3-layer neural networks. The features are coming from the central kernel of the window (raster data) and from edge detection (vector data). The vectorized edges are only kept for training, if they are going through the central circle, which represents the junction central in a rotation invariant way. The edges fulfilling the circle criterion are applied to derive features, like edge length and direction measures. A set of identically structured networks with varied parameters was generated and trained by an efficient optimization algorithm. The evaluation of the networks was carried out in in-sample tests, where the main traditional methods are compared to the neural solution. The out-of-sample test was performed by real image chips with different rotations. The obtained results demonstrate the principal feasibility of the developed method.

1. INTRODUCTION

In the 20th and 21st century roads have more importance than ever. The traffic density is increasing, better navigation solutions are required. These systems rely on accurate and up-to-date road information. Because of the large demand there is much research concentrating on efficient road mapping.

In photogrammetry the general research focuses on automatic or automated solutions, also in the mapping of roads. Big efforts were undertaken on different levels of image processing: from the pixel based edge detection to the high level image understanding algorithms. Beside the general enhancement and edge detection techniques (Sobel, Laplace and similar convoluting operators, Canny and Deriche-algorithms) special road sensitive theories were developed. Steger worked out a line extraction algorithm, which was successfully applied for detecting road axes (Steger, 1998). Baumgartner et al. (1999) developed a very sophisticated road extraction algorithm for large resolutions (approx. 0.2 m per pixel), Mayer (1998) studied the suitability of the scale-space theory for extracting roads from images. Wiedemann (2002) enhanced the Steger-method with topological considerations: he studied the roads as elements of a network. The developed method works on satellite images and other medium resolution aerial imagery. Willrich (2002) refined the Wiedemann-strategy for verifying the roads of the German Official Topographic-Cartographic Information System (ATKIS). IKONOS imagery is also a suitable source for road mapping, as Dial et al. have proved by their texture based classification method (Dial et al., 2001).

Also artificial neural networks (ANN) have been used to extract objects from imagery. In the computer vision literature results can be found in airplane identification (Abdallah et al., 1995) and also for recognizing mechanical tools (Tang et al., 1996). Kepuska trained an ANN for detecting signalized points in an aerial image (Kepuska et al., 1991), Chiu used such kind of networks for single and multiple target processing in industrial measurements (Chiu et al. 1990).

In (Fiset et al., 1998) a neural strategy is implemented for matching the road elements of a geographic information system (GIS) database with images, the solution has potential for being applied in updating.

The mentioned references study mainly the road itself; crossings and junctions are treated mainly as special cases of road segments. Although e. g. in building extraction the vertices contain valuable information, such techniques were not tested for roads. The present paper summarizes the intermediate results of ongoing research. The goal of the work is to develop a technique based on artificial neural networks to detect road junctions in medium resolution aerial images of approximately 0.5 m per pixel. The network performs a kind of classification, which can be refined by further grouping and other high-level image understanding techniques. The trained neural network can find the objects without any additional data, and the method can be a part of a sophisticated updating procedure.

2. STRATEGY

The junction detection requires first of all a generic junction model. The model is created iteratitively: after setting up a coarse model, it is refined in order to obtain better description and behavior.

The model contains parameters, which are derived from the images by using image processing algorithms. Furthermore, a procedure belongs to the model that takes template window as input, performs several tasks (see below for details) and at the end it decides whether the window contains a junction or not. The result is related to the window center point. The model is based on a classifier, which is supported by adequate preprocessing steps (Fig. 1).

The classifier of the developed technology is a feed-forward neural network. Artificial neural networks are classifiers, which are trained by examples instead of "global" statistical measures, like maximum likelihood or minimum distance classifiers. The required samples were collected from black-and-white orthoimages with a ground resolution of 0.4 m. There were 60 junction (J) and 120 non-junction (NJ) samples, the windows had a size of 51×51 pixels (20.4×20.4 m). The junctions were 3- and 4-arm junctions with different road category.

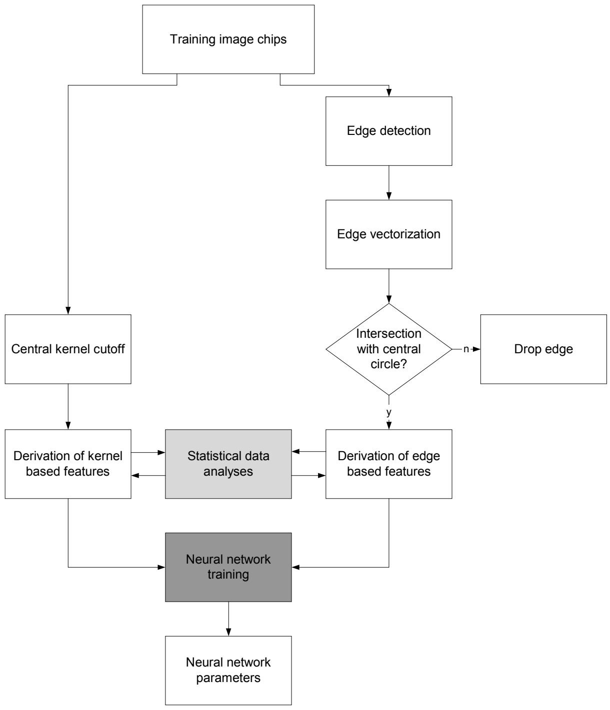

Figure 1. Strategy to create the junction detection neural network

The strategy uses both raster and vector data. There are two main processing lines: the first one is an application of a *central kernel* in the middle of the window. The size of the square kernel is a variable of the procedure. A number of parameters were derived for the kernel: minimum and maximum pixel intensities, average intensity, standard deviation and median value. Using these parameters an analysis was performed to determine the optimum kernel size. A combination of the following statistical methods was executed: discriminant analysis, principal component-, factor- and variance analysis. The largest differences (best separability) between junctions and non-junctions were found with a kernel size of 7×7, equivalent to about 3 x 3 m^2. This size is scale dependent, i.e. it depends on the image resolution. The kept features are the average pixel intensity value (*A*) and the standard deviation (*SD*) of the intensities within the kernel.
The second processing block of the strategy starts with an edge detection using the *Deriche algorithm*, which is a recursive approach coupled with smoothing and a hystheresis threshold. The edge detection was performed with non-maximum suppression. The results were edge amplitude (gradient magnitude) and edge direction images. For reasons of efficiency the edge detection was performed for the whole input orthoimage, not window-by-window. The all-in-one edge detection resulted also in a quality increase: the window-by-window edge detection finds many short "edge fractions", while the image-wise edge detection returns longer edge segments.
The amplitude image was filtered by an edge preserving 8-neighborhood method and vector objects were created followed by a smoothing with the Ramer algorithm.
The found edge vectors were evaluated by the *central circle* criterion. The meaning of the central circle for junctions (first row) and non-junctions (second row) is shown in Fig. 2.

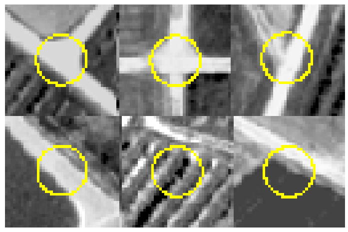

Figure 2. Central circle on junction (first row) and non-junction (second row) samples

The central circle criterion expresses, that the road side edges must run in the junction direction. The circle represents the junction in the model in a rotation invariant way.

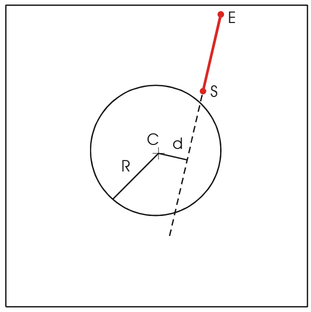

Figure 3. The running window containing the central circle and a found edge vector (*ES*)

The square in Fig. 3 represents the template window; its center point is marked (*C*). The found edge vector is shown as a thick line, having *S* as startpoint and *E* as endpoints. The intersection problem of the *ES* line and the central circle can be formulated as a test of $d < R$, where d is the calculated distance of the circle's origin (*C*) and the extension of *ES* line. The optimal value of the radius was found as 11 pixels (4.4 m).
Depending on the test result, edges were dropped (no junction) or they were applied to derive several features: center of gravity of the edges in X and Y direction (*COGX* and *COGY*), the average edge length (*AVELEN*) and average direction angle (*AVEDIR*), as well as the standard deviation of the edge lengths (*SDLEN*) and of their directions (*SDDIR*).
The goal of the analysis carried out so far was to select the right features for good separability between junctions and non-junctions. As an example Fig. 4 shows the relation of the average pixel intensity value of the kernel (*A*) to the average of the accepted edge vector directions (*AVEDIR*). Junction samples are marked with circles, while the non-junction cases are the crosses.
The outlier, ambiguous junction samples (e.g. junctions with occlusion caused by trees) were removed from the training set manually, because they represent mixtures between the two groups. After the removal 44 samples remained.
The data analysis was used for iteratively refining the junction model. The best features were used as input for the neural network training. The applied artificial neural networks were 3-layer feed-forward (back-propagation) networks.

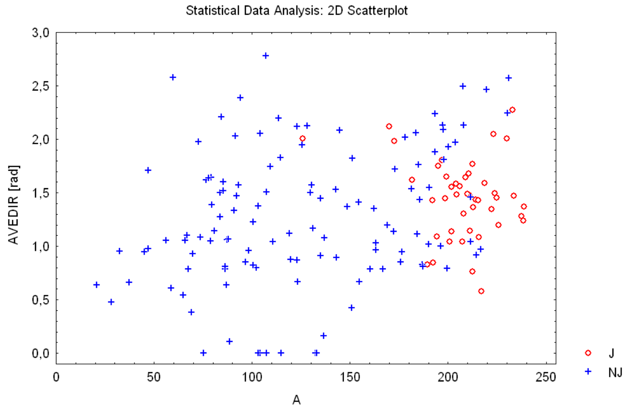

Figure 4. Scatterplot for features *A* and *AVEDIR*

The output layer contains only a single neuron, which gives a "neural possibility" between *0* (non-junction, NJ) and *255* (junction, J). The neurons have logistic sigmoid transfer function. Because of the known efficiency of this method, the networks were trained by the Levenberg-Marquard optimization technique. The network structure was varied during the training to obtain the highest network recognition accuracy. The number of neurons in the first and second layers, and the desired network error rate were varied. These variations created several hundreds of neural networks for every feature set. The resulting networks were finally evaluated by *in-sample* and *out-of-sample* tests.

3. RESULTS

In this chapter we present results obtained with the method described before. The used test data represent black-and-white orthoimages of a resolution of 0.4 m from Frankfurt am Main, Germany.

The first complete quality check was the *in-sample* test. The confusion matrix and the main quality measures (total accuracy, accuracy of the junctions and non-junctions) were calculated on the training samples. For reasons of comparison the in-sample test was performed with traditional methods, like linear discriminant functions (*LF*), minimum distance (*MD*) and maximum likelihood (*ML*) methods, as well as by neural networks (*NN*).

Table 1 shows the results obtained with the different methods with their recognition accuracies.

	method name	total accuracy %	Junction accuracy %	non-junction accuracy %
traditional methods	*LF*	91.5	86.4	93.3
	MD	79.3	97.7	72.5
	ML	84.1	100.0	78.3
neural methods	*NN1*	97.6	95.5	98.3
	NN2	97.6	90.9	100.0
	NN3	97.6	97.7	97.5

Table 1. In-sample recognition accuracy of different methods

There were many neural networks created, trained and evaluated during the development. We present only a subset of interesting versions. *NN1* was a neural network with a 3-3-1 neuron structure, the final (inner) network error was 0.0241. It was a rather simple network with few neurons. *NN2* was a 3-7-1 network with a final error of 0.0227. There were more neurons in the middle layer to obtain better flexibility, but the difference to NN1 was not significant during the tests. The last network (*NN3*) had a structure 9-9-1 and final network error of 0.0232. This network was a complex one, with most neurons in the free layers. The highest recognition accuracy was expected from the last version, but the reached network errors are almost the same, independently of the structure.

As one can notice the traditional methods have less total recognition accuracy than the selected neural networks. The *LF* method could detect the non-junction pixels rather well, the total accuracy is unexpectedly high. The minimum distance method is ranked as the worst traditional solution because of the low non-junction recognition. The reason is the high variety of the non-junction window content. The best traditional method (*ML*) can detect all junctions, but unfortunately had poor *NJ*-recognition accuracy – again because of the mentioned facts.

The selected neural networks have overall accuracy measures above 95 %. The simplest network had average accuracy measures in all categories. More neurons in the second layer resulted in better *NJ*-recognition, but was coupled with lower junction detection rate. The most complex network is adjusted both in junction and non-junction cases.

Because of the increased number of neurons within the same layer structure, there is more flexibility given to the network, so even complicated decision functions can be realized. This feature is advantageous in many practical applications, but it also has a drawback, namely the network may learn the training set perfectly and loose the desired generalization ability. This is why the out-of-sample tests are of extreme importance.

After the in-sample tests the junction operator was tested on images (*out-of-sample* test). There were 4 open field images in the test series, rotated by 0, 30 and 90 degrees respectively. Using these images the rotation invariance can also be proven. Fig. 5 shows the result of the smallest network.

The orthoimage subset is rotated by 30 degree in order to study how the network behaves. The road junctions are detected perfectly, but there are wrongly classified pixels inside of a parcel and on a road segment. This is caused by the regular texture combined with detectable edges. Future developments will add more stabilizing features, which can be helpful to reduce such errors and increases the recognition accuracy.

Fig. 6 shows another image (with no rotation) studied for the out-of-sample test. The test shows again the generalization feature of the networks, because the intensity value range is differing from the training set. The detection accuracy is quite good: all the visible junctions were marked with no misclassifications.

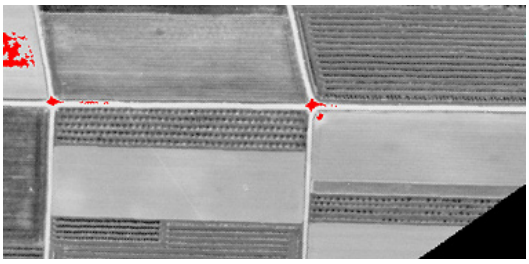

Figure 5. Out-of-sample test with 30 degree rotated image (the training was executed with no rotation)

Figure 6. Test image from an orthoimage, where no samples were taken

4. CONCLUSION

Our research concentrates on the detection of road junctions using medium resolution black-and-white orthoimages. A junction model has been created, which was refined using the tools of statistical data analysis. The model is built on raster as well as vector based information. The applied features were derived from a central kernel and from detected edges, which intersect the center of the search window. The adequate edges were selected by the central circle criterion, then several edge parameters were calculated to define the parameters of the junction operator. The developed neural technique was compared with the main traditional methods.
Future work will focus on extending the training set by further samples, not only from the current training area. Because artificial neural networks are trained by samples, a larger training set increases the accuracy. The size of the sample window can be increased, so the occluded junctions may be also taken into consideration and the unwanted effects of disturbing texture (e.g. rows in field parcels) can be better eliminated.
The increase of the recognition accuracy can be achieved by the use of existing data, e.g. using vector road data. The up-to-date road maps and suitably scaled topographic maps can therefore expand the training sets.
We plan a more comprehensive test for the developed method. The number of the applied parameters must be expanded to achieve higher recognition stability and accuracy. The junction operator is to be checked in other independent test areas.

ACKNOWLEDGEMENTS

The authors are grateful for the financial support to the Alexander von Humboldt Foundation, for the data to the Federal Agency for Cartography and Geodesy and the Hungarian Higher Education Research and Development Project.

REFERENCES

Abdallah, M.A., Samu, T.I., Grisson, W.A., 1995. Automatic target identification using neural networks. SPIE Proceeding, Intelligent Robots and Computer Vision XIV, Philadelphia, Vol. 2588, pp. 556-565

Baumgartner, A., Eckstein, W., Heipke, C., Hinz, S., Mayer, H., Radig, B., Steger, C., Wiedemann, C., 1999. T-REX: TUM research on road extraction, In: Festschrift für Prof. Dr.-Ing. Heinrich Ebner zum 60. Geburtstag, pp. 43-64

Chiu, W.C., Hines, E.L., Forno, C., Hunt, R., Oldfield, S., 1990. Artificial neural networks for photogrammetric target processing, SPIE Proceedings, Close-Range Photogrammetry Meets Machine Vision, Vol. 1395, No. 2, pp. 794-801

Demuth, H., Beale, M., 1998. Neural Network Toolbox – For Use with Matlab. User's Guide Version 3.0, The MathWorks Inc., Natick

Dial, G., Gibson, L., Poulsen, R., 2001. IKONOS satellite imagery and its use in automated road extraction. In: Baltsavias et al. (eds): Automatic extraction of man-made objects from aerial and space images, Swets & Zeitlinger, Lisse, pp. 357-367

Fiset, R., Cavayas, F., Mouchot, M.C., Solaiman, B., Desjardins, R., 1998. Map-image matching using a multi-layer perceptron: the case of the road network, ISPRS Journal of Photogrammetry and Remote Sensing, 53(2), pp. 76-84

Halcon 2001. Halcon Version 6.0.1 C++ Reference Manual, MVTec Software GmbH, München

Kepuska, V.Z., Mason, S., 1991. Automatic signalized point recognition with feed-forward neural network, IEE Conference Publication, 2nd International Conference on Artificial Neural Networks, London, No. 349, pp. 359-363

Mayer, H., 1998. Automatische Objektextraktion aus digitalen Luftbildern. Deutsche Geodätische Komission, Reihe C, No. 494, München

Steger, C., 1998. An unbiased extractor of curvilinear structures, IEEE Transactions of Pattern Analysis and Machine Intelligence(20)2, pp. 113-125.

Tang, H.W., Srinivasan, V., Ong, S.H., 1996. Invariant object recognition using a neural template classifier. Image and Vision Computing, 14(7), pp. 473-483

Wiedemann, Ch., 2002. Extraktion von Straßennetzen aus optischen Satellitenbilddaten. Deutsche Geodätische Komission, Reihe C, No. 551, München

Willrich, F., 2002, Quality control and udating of road data by GIS-driven road extraction from images, IntArchPhRS (34)4, pp. 761-767.