Detecting road junctions by artificial neural networks

Arpad Barsi, Christian Heipke

***Abstract*—Road junctions are important objects for all traffic related tasks, and are essential e. g. for vehicle navigation systems. They also play a major role in topographic mapping. For automatically capturing road junctions from images models are needed, which describe the main aspects. This paper presents an approach to road junction detection based on raster and vector information. The raster features are similar to the ones used in classification approaches. The vector features are derived from a road junction vector model containing edges as road borders. The whole feature serves as input to an artificial neural network. The neural classifier decides for a search window, whether its central pixel is a part of a road junction or not. The developed junction operator was tested on several black-and-white medium resolution orthoimages. The achieved results demonstrate that such junction models can successfully identify three- and four-arms road junctions.**

***Index Terms*—junction detection, object recognition, feature extraction, artificial neural networks**

I. Introduction

THE development of vehicle navigation systems and location based services has been very rapid over the last years. For these and other applications up-to-date, accurate and correct information of the road network is mandatory.

The current acquisition and updating of geospatial information is characterized by a large amount of manual work, and is thus rather costly and slow. Research and development have investigated new and effective solutions. Efforts were focused on different image based automatic solutions as well as the on data collection on the ground.

Automatic road extraction from images has a history of more than two decades. Most early approaches were not very useful, because they only used image processing algorithms such as edge and line detection. Exceptions to these attempts can be found in [1] considering a graph structure for the whole road network, and [2] combining different extractions cues. A very sophisticated method was developed by Wiedemann [3] based on the Steger line extractor [4]. The Wiedemann-method contains elements from graph theory (e.g. shortest path algorithms) and models the road segments extracted from the images as parts of a complete network. The method has been applied successfully to satellite and medium resolution aerial imagery, and was also implemented for quality control of existing road data, as Willrich has shown on black-and-white orthoimages combined with ATKIS (German Official Topographic and Cartographic Information System) road objects [5]. Other examples of automatic road extraction include e. g. the approaches of Baumgartner [6] for high resolution aerial images, and Dial et al. for satellite imagery such as those captured with Ikonos or Quickbird [7].

A relatively new technique in digital photogrammetry is the application of artificial neural networks (ANN) for object detection. These networks have proven their universality in several technical fields. In image processing, Chiu et al. [8] identified several objects with the help of ANN, Tang et al. [9] have reported similar work. Kepuska [10] used a neural network to recognize signalized control points in photogrammetric images, and Abdallah et al. [11] detected airplanes in images with the help of ANN.

The goal of our efforts was to develop a method for road junction detection from medium to high resolution aerial images based on ANN which should be able to deal equally well with three and four arm junctions. The motivation was twofold: firstly, there exists no good road junction detector in the literature, probably due to the fact that road junctions have a vastly differing appearance in images, and secondly, we wanted to learn more about ANN. Preliminary results of our work have already been reported in [12].

II. Modeling road junctions

The development of the junction operator started with a model-building step. In this phase image samples of different type of road junctions were collected. The obtained data set contained crossings with different road widths and orientations. In order somewhat simplify the task the type of the junctions was limited on three and four arm junctions.

The first junction model was very simple. Typical junctions were selected, binarized using gray value thresholding, and rotated into several positions in order to achieve rotation invariance; the generated data set was then used to train a feed-forward neural network. This junction model was based only on raster information and thus very strongly on the generalization ability of the ANN; just the image cut-off matrices with the intensity values were presented as input information. Not surprisingly, the results didn't meet our goals, and we therefore decided to enhance the model with other type of information.

In the current junction model we have kept the raster information, but extra vector type features was added. Within a window of pre-selected size we run an edge detection algorithm (we use the Lanser algorithm [13], which is a rotation invariant version of the Deriche edge finder and includes a hysteresis threshold, followed by edge smoothing using the Ramer algorithm). Results are depicted in Fig. 1.

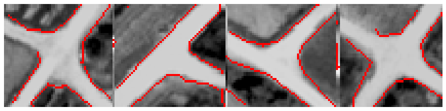

Fig. 1. The result of the edge detection

The road junctions have a common feature, which is visible in the derived vectors: most edges run in the direction of the junction. Due to the selected image resolution the junction is not a point-like (zero dimensional) object, and the edges form road borders; the road edges do not cross each other in a single point.

In order to model this finding we use a circle centered around the junction. This circle incorporates the edge evaluation criteria: an edge segment is further investigated if and only if the straight line through this edge segment (called the extended edge) intersects this circle, otherwise it is dropped. Fig. 2 shows three junction samples with an adequate sized junction circle.

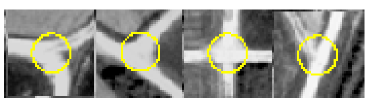

Fig. 2. The junction is represented by a centrally placed circle

The edge evaluation is based on the use of the central circle (Fig. 3). The circle belongs to a search window, which clips the found edge vectors. All the edge segments in the search window are evaluated by the circle test.

The evaluation itself is carried out by computing the distance d from the center of the circle to the extended edge, where the edge runs from start point S to end point E. The extended edge intersects the circle, if d is smaller the radius R of the circle.

The evaluated and kept edge vectors are possibly roadside segments. Using these vector elements statistical features such as variation of the edge direction or edge length can be derived, which may be specific for road junctions and thus provide the possibility to differentiate junctions from other image features including straight road segments.

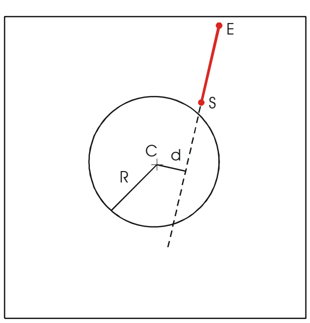

Fig. 3. The relation of a found edge to the junction circle

Beside this vector based features some raster features – e. g. average and standard deviation values of the gray values – can also be used. The raster data should be collected in a kernel of predefined size within the search window.

The differentiation (separation) of the junctions from the non-junctions is a decision procedure. The relevant features for this task are usually difficult to identify. A small number of the features is crucial for the neural network training: the complexity of the network and the training time requirement increases drastically with the number of features as input.

To ease this development step, various combinations of features were analyzed. The evaluation was mainly graphical: pair-wise scatter plots were created with junction–non-junction coloring. The goal of the analysis was to find a set, in which the point clouds of the classes are separated.

The influence of the input features was also investigated by principal component analysis. This tool can help to define the minimum number of adequate features.

The result of the statistical feature analysis was a collection of the most important input features. The set contains the standard deviation of the edge direction and edge length, and the coordinates of the center of gravity of the edge endpoint. As raster data the mean intensity value A and its standard deviation SD were kept in the feature set.

The obtained statistical features are used to train the neural networks by adding the information, whether the sample is a junction image or not. The reliability of the correct feature selection was increased by outlier elimination; ambiguous image samples – because of wrong light conditions, occlusions caused by trees, buildings, etc. – were removed. The artificial neural network was thus trained with clean data only.

The junction detection was realized by sliding a window of predefined size over the whole image, similar to a convolution operation. At each step the result of the decision is stored in the center pixel of the window. The result of the operator is a classified image, containing pixels, which are identified as road junctions and those which are considered non-road pixels.

III. The Neural Network Development

As mentioned in the introduction, ANN have already been used in a few instances in photogrammetry and image analysis. In our method they are applied as sophisticated *pattern classifiers*. The ANN inputs are derived using several image processing techniques, the desired output is thought to be a clear distinction between different patterns, in our case road junction and non-junction patterns.

We have chosen a feed-forward neural network which is one of the most frequently implemented network types. The network construction comprises an iterative back-propagation training method.

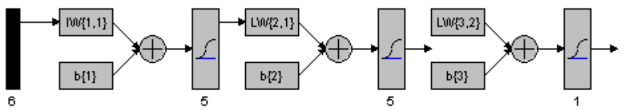

Fig. 4. The structure of the applied neural network

The network has three neuron layers with equally logistic sigmoid processing elements (Fig. 4). This neural transfer function has a simple derivative, which is a very valuable feature, since it is allows very fast training. The parameters of the network, which are to be defined during the training, are the weight matrices *IW* and *LW* and bias vectors *b* for all three layers.

The training of the neural networks is a time-consuming task. The Levenberg-Marquard algorithm was selected for our project, because of its efficiency. This algorithm uses a learning rule of second order gradient descent with momentum. The momentum helps to avoid the local minima on the error hyper surface. We used the mean square error measure, because it adjusts the network weights and it has moderate performance requirement. The implemented weight update was almost constant, which is achieved by adaptively modifying the optimization step length.

The training set and the network structure were fixed, and an acceptable decision accuracy was searched for changing the number of neurons per layer and using different feature sets. The whole training procedure was controlled by the difference between the required and obtained network results measured for the training samples. During training several hundreds of neural networks were generated, from which the most adequate one was selected in a follow-up evaluation step. During the network generation five repetitions with exactly the same structures were used to reduce the effect of the random initialization.

IV. Results

The developed junction operator was tested in on digital black-and-white orthoimages, which cover a region near Frankfurt am Main, Germany. The orthoimages have a ground resolution of 0.4 m.

The extracted training samples were image chips with 51 × 51 pixels, which equals a dimension of 20.4 × 20.4 m^2. Altogether, there were 180 image windows in the training, 1/3 with junctions and 2/3 with non-junctions. 38 image samples were considered to be “ambiguous” and were not further considered (see above).

The kernel size for the raster data acquisition was 7 × 7 pixels, which equals 2.8 × 2.8 m^2. The radius of the circle was empirically set for 11 pixels (4.4 m).

The trained junction operator was first tested on known images, i. e. with images fro m the training set. Fig. 5 shows the correct identification of a 4-arm junction.

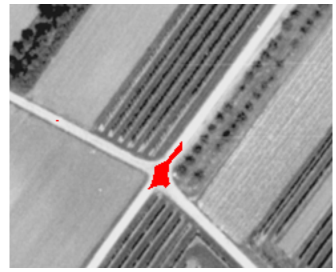

Fig. 5. Recognition accuracy of the junction operator on known sample

Fig. 6 shows the results of another image, the crossings in this image were not used in the training. The junction operator has detected several potential road crossings. All but one recognized positions are correct, they are marked with circles. There is an additional potential crossing, marked with a square, where the main road is widened. In addition, there are some pixels on road segments in the image, which are also identified as junctions (see the segment between the two circles on the right in Fig. 6).

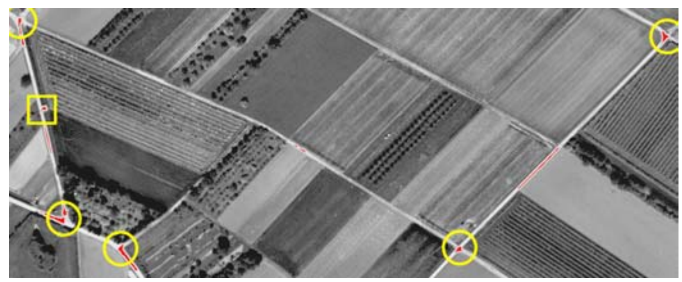

Fig. 6. Correctly recognized road junctions (marked with circles) and a position with a widening road (marked with square)

The current junction operator found another interesting place, namely a bridge (see Fig. 7). This is not surprising, since the bridge also fulfills the requirements of the model: there are road segments, which are crossing each other.

Fig. 7. A bridge as a road junction

The last example is shown in Fig. 8. All junctions were detected, except one in the left bottom corner. The enlargement (see Fig. 9) shows that there is a shadow of some trees falling into the junction. The shadows has the effect that no straight edge vector was detected, thus the crossing was not found. This last examples shows that our methods, while successfully being able to detect many road junctions, is still a little sensitive to some aspects of road junction detection.

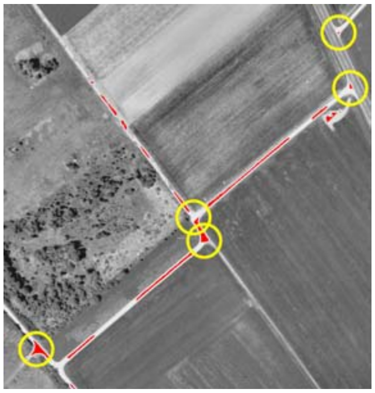

Fig. 8. Detected junctions in the other part of the study area

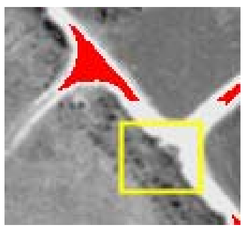

Fig. 9. Enlargement of Fig. 8 with tree shadows

V. Conclusion

The experiments have proven the general applicability of the new road junction detector based on integrating raster and vector information in an artificial neural network.

The used vectors are derived by edge detection techniques. Obviously, these techniques do not only find road edges, but also all similar gray level edges all over the image. In order to reduce this ambiguity we have focused on the evaluation of the extracted edges by introducing the central circle criterion. The selected edge vectors were used to derive features, which in conjunction with raster features were the inputs for the artificial neural network.

The presented junction operator currently finds mostly the pixels, which belong to a road junction. The robustness can be improved by introducing additional rules, like considering the fact that road edges outside the central circle should be parallel, and gray value constancy between the parallel edges. Further improvements are of course possible by increasing the number of training samples. We plan to incorporate these and similar enhancements in the near future.

References

[1] Fischler M, Tennenbaum J., Wolf H., 1981: Detection of roads and linear structures in low resolution aerial imagery using a multisource knowledge integration technique, Computer Graphics and Image processing, Vol. 15, pp. 201-223.

[2] McKeown D., Denlinger J, 1988: Cooperative methods for road tracking in aerial imagery, Computer Vision and Pattern Recognition, pp. 662-672.

[3] Wiedemann, C., 2002. Extraktion von Straßennetzen aus optischen Satellitenbilddaten. *Deutsche Geodätische Komission, Reihe C*, No. 551, München

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

[5] Willrich, F., 2002, Quality control and updating of road data by GIS-driven road extraction from images, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Science (34)4*, pp. 761-767.

[6] 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.

[7] 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

[8] 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

[9] 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

[10] 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

[11] 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

[12] Barsi, A., Heipke, C., Willrich, F., 2002, Junction extraction by artificial neural network system – JEANS, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Science (34)3B*, pp. 18-21

[13] Lanser, S., Eckstein, W., 1991, Eine Modifikation des Deriche-Verfahrens zur Kantendetektion, *13. DAGM Symposium, München*, pp. 151-158