KNOWLEDGE-BASED IMAGE ANALYSIS FOR 3D EDGE EXTRACTION AND ROAD RECONSTRUCTION

Chunsun ZHANG, Emmanuel BALTSAVIAS
Institute of Geodesy and Photogrammetry,
ETH-Hoenggerberg, CH-8093 Zurich, Switzerland
Tel.: +41-1-6332931, Fax: +41-1-6331101
E-mail: {chunsun, manos}@geod.baug.ethz.ch

Working Group IC IV/III.2

KEY WORDS: Feature Extraction, Edge Matching, Model-based Processing, Reconstruction, Data fusion, Color.

ABSTRACT

Road network extraction from aerial images has received attention in photogrammetry and computer vision for decades. We present a concept for road network reconstruction from aerial images using knowledge-based image analysis. In contrast to other approaches, the proposed approach uses multiple cues about the object existence, employs existing knowledge, rules and models, and treats each road subclass differently to increase success rate and reliability of the results. Finding 3D edges on the road and especially the road borders is a crucial component of our procedure and is the focus of this paper. We developed an algorithm for automatically matching line segments across images. The algorithm exploits line geometrical and photometrical attributes and line geometrical structures. A framework to integrate these information sources using probability relaxation is developed and implemented to deliver locally consistent matches. Results of straight line matching are presented and future work is discussed.

1 INTRODUCTION

The extraction of the road network from 2D aerial images is one of the current challenges in digital photogrammetry. Due to the complexity of such images, their interpretation for mapping roads and buildings has been shown to be an extremely difficult task to automate (Gruen et al., 1995, 1997; Foerstner and Pluemer, 1997). As digital topographic databases have been created in many developed countries, and now need to be updated, several methods are explored to incorporate this existing knowledge for image interpretation. The effect of this incorporation is twofold: the existing information provides a rough model of the scene, that will help the automation process, while the old database gets revised and updated with the latest aerial images.

The here presented work is part of the project ATOMI. ATOMI is a cooperation between the Federal Office of Topography (L+T) and ETH Zurich. Its aim is to use aerial images and automated procedures to improve vector data (road centerlines, buildings) from digitised 1:25,000 topographic maps by fitting it to the real landscape, updating it, improving the planimetric accuracy to 1m and providing height information with 1-2m accuracy. In the current tests, we use 1:16,000 scale color imagery, with 30 cm focal length and 25% sidelap, scanned with 14 and 28 microns at a Zeiss SCAI. The input data are: the nationwide DTM (DHM25, 25 m grid spacing and accuracy of 2-3/5-7 m in lowland/Alps), the vectorised map data (VEC25), the raster map with its 6 different layers and the digital images. The VEC25 data have an RMS error of ca. 5-7.5 m and a maximum one of ca. 12.5 m, including generalisation. They are topologically correct, but due to their partly automated extraction, some errors might exist. The general overview of ATOMI is described in Eidenbenz et al. (2000). In this paper we deal with road reconstruction, while the building part can be found in Niederoest (2000). In a first step of road reconstruction, we aim at detecting existing roads, while roads that do not exist anymore or new ones will be treated later. We first concentrate on roads, ignoring other transportation network objects like railway lines, mountain paths etc.

In this paper, we present the ongoing work of the project on the interpretation of aerial images. The ultimate goal of this work is to build an automatic and robust system to reconstruct the road network from aerial images with the aid of an existing road database and possibly additional data. In order to increase the success rate and the reliability of the results the system will contain a set of image processing tools, and make full use of available information as much as possible. At the first stage, we will focus on the design of methodologies and techniques for image analysis tools, in particular we are interested in using information extracted from old database and applying rules and models to guide the process. Finding 3D edges on the road and especially the road borders is a crucial component of our procedure and is the focus

of this paper. Straight lines are extracted by fitting the detected edge pixels in each image and straight line matching across images is performed through exploiting rich line attributes and line geometrical structures.

2 GENERAL STRATEGY OF ROAD NETWORK EXTRACTION

Much research work on road extraction was reported in the past (McKeown et al., 1988; Gruen and Li, 1997; Ruskone, 1996; Zhang and Baltsavias, 1999; Hinz et al., 1999; Trinder et al., 1999). The existing approaches cover a wide variety of strategies, using different resolution of digital aerial or satellite images. We develop a new approach for automatic extraction of 3D road network from aerial images which integrates knowledge processing of color image data and existing digital geographic databases. The information of existing road database provides a rough model of the scene. Color aerial images give the current situation of the scene, but are complex to analyze without the aid of other auxiliary data. Therefore, the information provided by the existing geographic database can help the understanding of the scene, while the images provide real data useful for improving the old road database and updating it. The system under development strongly relies on the following three aspects:

- Use and fusion of multiple cues about the object existence and of existing information sources. All cues have associated relevant attributes.
- Use of existing knowledge, "rules" and models. The road model includes geometric, radiometric, topological and contextual attributes.
- Object-oriented approach in multiple object layers (hierarchical division of classes in subclasses, division of a class according to terrain relief and landcover).

The initial database is established by the information extracted from existing geographic data and road design rules. This offers a geometric, topological and contextual description of road network in the scene. The database is automatically updated and refined using information gained from image analysis. Color cues, expressed in the form of color region attributes, are also used to support stereo matching and improve the performance of 2D and 3D grouping when combined with geometric cues. Since neither 2D nor 3D procedures alone are sufficient to solve the problem of road extraction, we propose to extract the road network with the mutual interaction of 2D and 3D procedures. Hence, the main steps of road extraction are: building up of the knowledge base for each road segment in VEC25, finding 3D straight lines in a search region defined by the VEC25 data, classification of image patches, extraction of other cues, combination of various cues guided by the knowledge database to find plausible groups of road edges for each VEC25 road segment and refinement and update of the knowledge database. The general strategy is shown in Fig. 1. Fig. 2 shows more details of the results of image processing and derivation of subclass vector attributes.

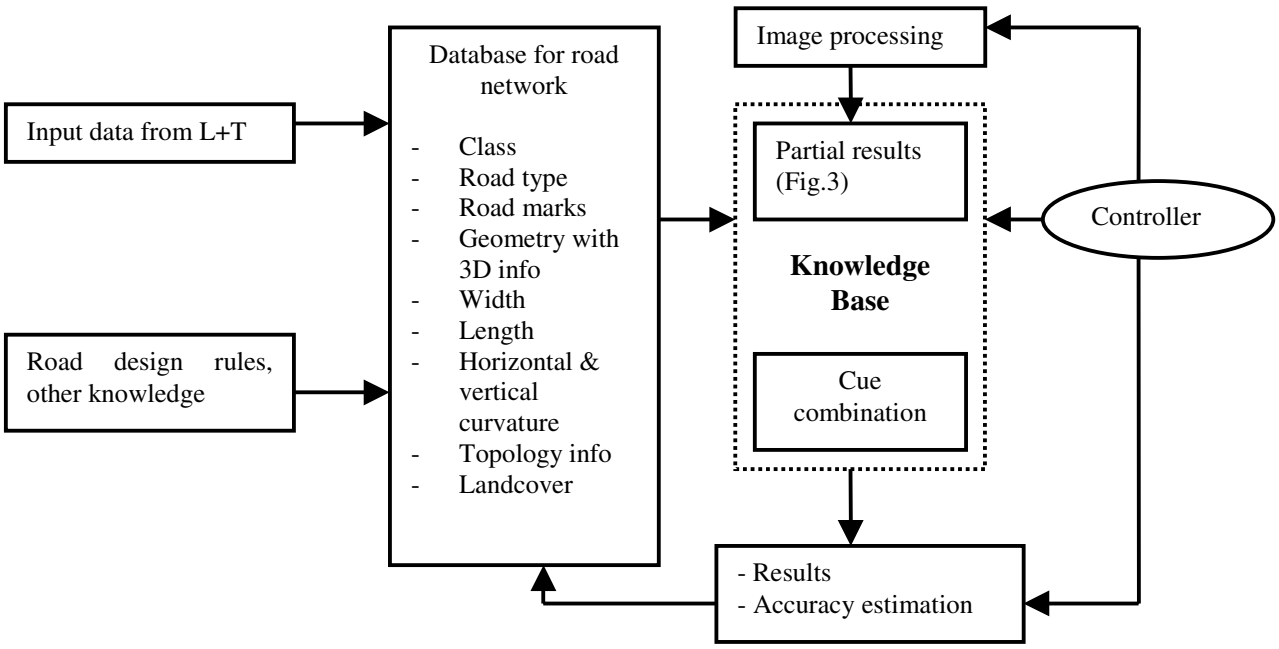

Figure 1. Strategy of road network extraction in ATOMI.

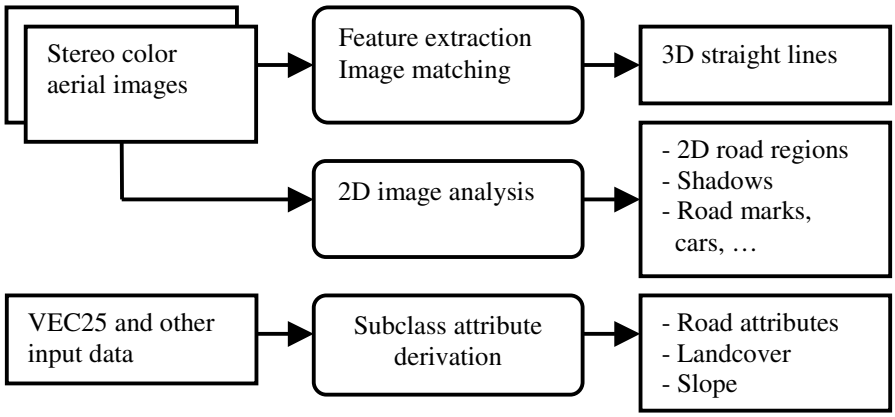

Figure 2. Details of image processing and derivation of subclass vector attributes.

3 STRAIGHT LINE MATCHING

3D edge generation is a crucial component of our procedure. We are interested in 3D straight lines because they are prominent in most man-made environments, and usually correspond to objects of interest in images, such as buildings and road segments. They can be detected and tracked relatively easily in image data, and they provide a great deal of information about the structure of the scene. Additionally, since edge features have more support than point features, they can be localized more accurately. The 3D information of straight lines is determined from the correspondences of line segments between two images.

Due to the complexity of aerial images, different view angles and occlusions, straight line matching is a difficult task. Existing approaches to line matching in the literature are generally categorised into two types. One is directly comparing the attributes of a line in one image with those of a set of lines in another image and selecting the best candidate based on a similarity measure (McIntosh and Mutch, 1988; Medioni and Nevatia, 1985; Greenfeld and Schenk, 1989; Zhang, 1994). The similarity measure is a comparison of line attributes, such as orientation, length, line support region information etc. In another strategy, the line correspondence is found by performing structural matching. Structural matching seeks to find the mapping between two structural descriptions. A structural description consists of not only features but also geometrical and topological information among features. A number of methods have been developed for structural matching (Vosselman, 1992; Haralick and Shapiro, 1993; Christmas et al., 1995; Cho, 1996; Wilson and Hancock, 1997).

The developed method in this paper exploits rich line attributes and line geometrical structure information (Fig. 3). The rich line attributes include the geometrical description of the line and the photometrical information in the regions right beside the line (flanking regions). The epipolar constraint is applied to reduce the search space. The similarity measure for a line pair is first computed by comparing the line attributes. The similarity measure is used as prior information in structural matching. The locally consistent matching is achieved through structural matching with probability relaxation. The details of the method are described below.

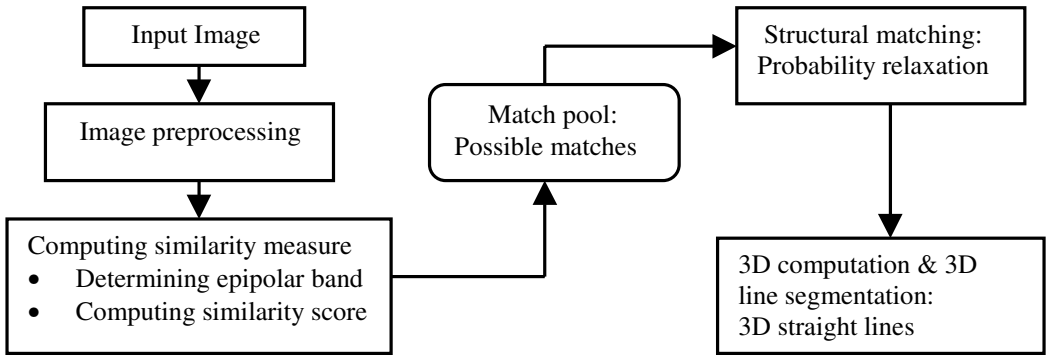

Figure 3. Flow-chart of straight line matching.

3.1 Line Extraction

The input images are first filtered with Wallis filter for contrast enhancement and radiometric equalization (Baltsavias, 1991). The technique developed in a previous project (AMOBE) is used to extract straight lines (Henricsson, 1996). The edge pixels are detected with the Canny operator. An edgel aggregation method is applied to generate contours with small gaps bridged based on the criteria of proximity and collinearity. All segments are checked using their direction along their length and split at points where the change in direction exceeds a given value. A test is conducted to see whether the consecutive segments can be merged into a single straight line. For each straight line segment we compute the position, length, orientation, and photometric and chromatic robust statistics in the left and right flanking regions. The photometric and chromatic properties are estimated from the “L”, "a" and "b" channels after an RGB to Lab color space conversion and include the median and the scatter matrix.

3.2 Computation of the Similarity Score

With known orientation parameters, the epipolar constraint can be employed to reduce the search space. The two end points of a line segment in one image generates two epipolar lines in the other image. With the approximated height information derived from DHM25 or DSM data, an epipolar band is defined (Baltsavias, 1991). Fig. 4 illustrates this idea. Therefore, a search region is determined in the right image for each segment in the left image. Any line included in this band (even partially) is a possible candidate, if it intersects the two epipolar lines (through the two line endpoints in the left image) within this band. For example in Fig. 4, lines *i, j, k* are accepted and will be compared with line *pq* in the left image for similarity measurement, while line *r* is rejected because it intersects *eq* outside *q1, q2*. The size/height of this search band is decreasing with edge length and orientation difference to the epipolar lines. The comparison with each candidate edge is then made only in the common overlap length, i.e. ignoring length differences and shifts between edge segments.

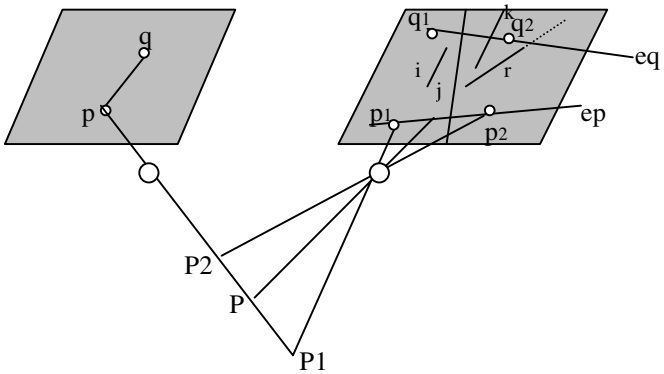

Figure 4. The epipolar band $p_1p_2q_2q_1$ defines the search space for line pq

For each pair of lines which satisfy the epipolar constraints above, their rich attributes are used to compute a similarity score. Therefore, the similarity score is a weighted combination of various criteria. A similarity measurement for length is defined as the ratio of the minimum length of the two lines, divided by the maximum one. Thus, the similarity score is defined as

$$V_{len} = \frac{min(L_{len}, R_{len})}{max(L_{len}, R_{len})} \tag{1}$$

We use the absolute difference between two line angles and an expected maximum difference in a ratio form similar to (1). The maximum value T_{ang} is a predefined threshold value. Since in aerial image the Kappa angle has a large effect on the line angles, we rotate the lines with Kappa in left and right image respectively before computing the line similarity score by

$$V_{ang} = w \frac{T_{ang} - |L_{ang} - R_{ang}|}{T_{ang}} \tag{2}$$

w is a weight related to line length, and given by $w = w_l w_r$, where W_l, W_r are computed for the lines in left and right image respectively. They are defined as a piecewise linear function as:

$$\begin{cases} 0; & len < 5; \\ \frac{1}{t-4}(len-4); & 5 \le len < t \\ 1.0; & len \ge t \end{cases} \tag{3}$$

The same weight is applied for the computation of the following flanking region similarity scores.

We do not only compute a similarity score of line geometric properties, but also of photometric and chromatic properties of line flanking regions. The photometric and chromatic line region attributes include the median and standard deviation of the L band, and median and scatter matrix for the chromatic components, i.e. (a, b) data. Note that when these properties are computed, the outliers in the flanking regions of each band are removed; for details see (Henricsson, 1996). We first compare the medians of L, a, and b band in left and right flanking regions for a pair of lines in two images. The similarity measurement is defined as the ratio of minimum median divided by the maximum. For example, the left region median similarity measurement of the L band for a line pair is computed as:

$$C_{Ll} = \frac{\min[\ median\ \ (left\ \ image\ \ ), median\ \ (right\ \ image\ \ )]}{\max[\ median\ \ (left\ \ image\ \ ), median\ \ (right\ \ image\ \ )]} \tag{4}$$

This computation is also applied to the right regions of L band. Similarly, the median similarity measurements of a and b bands are obtained. Then, we average these scores of L, a, b bands as one region similarity measurement, i.e. we obtain the left and right region similarity measurement C_l and C_r for a line pair. In our matching algorithm, we do not assume that the line pair has same contrast; instead we only request that at least one side of the line pair they demonstrate similar brightness. Thus, if both C_l and C_r are less than a predefined value, then the two lines are treated as different lines, and we stop computing similarity scores for them. Finally, the similarity scores for the median C_l and C_r are multiplied with the weight, where the weight is defined as in (3), i.e.

$V_{ml} = wC_l$ For left region
$V_{mr} = wC_r$ For right region

The chromatic property of a region in the (a, b) data is represented by the scatter matrix $c = \begin{pmatrix} c11 & c12 \\ c21 & c22 \end{pmatrix}$. We further describe this property as an ellipse. The shape of the ellipse is determined by its axes and orientation, derived from the scatter matrix. Thereby, the similarity of a region chromatic property can be achieved through comparing the shapes of the respective ellipses. We compute the orientation and roundness of ellipse as below:

$$Dir = \frac{1}{2}\arctan\frac{2 * C12}{C11^2 + C22^2}, \quad Q = 4\frac{Det(C)}{tr^2(C)} \tag{5}$$

The similarity measurement of a chromatic property is performed by comparing ellipse orientation and roundness in the left and right regions for a line pair. The similarity score of ellipse orientation is computed with a form similar to (2). For ellipse roundness, we use a form similar to (1).

All the scores are from 0 to 1, and the total similarity score is the average of all scores. The similarity score computation starts from the longer lines, while the very short ones (< 5 pixels) are ignored.

3.3 Structural Matching with Probability Relaxation

After performing similarity measurement computation, we construct a matching pool and attach a similarity score to each line pair. However, one still has problems to determine the best matches. The difficulty comes first how to decide on a threshold and how to treat the case when a line is broken or occluded. In addition, matching using a very local comparison of line attributes does not necessarily give results consistent in a local neighbourhood. For this reason, structural matching receives more and more attention in computer vision and photogrammetry.

Structural matching establishes a correspondence from the primitives of one structural description to the primitives of a second structural description (Haralick and Shapiro, 1993). Several methods for structural matching were developed in the past (Vosselman, 1992). In this paper, the structural matching for line correspondence is realized through probability relaxation.

The line structure in image is described as a graph, where the nodes of the graph represent straight lines, and the links between lines the relations. To find a correspondence, both individual lines and the graphs should be matched. We represent the straight lines in the left image as a set *L, L={li}, i=1, 2, … n*, the straight lines in the right image as a set *R, R={rj}, j=1,2,… m*. The mapping from the left description to the right one is represented as *T*. Assuming the right type of mapping *T*, we seek the probability that line l_i matches r_j, i.e. the matching problem becomes the computation of $P\{ l_i = r_j \mid T\}$. Using Bayes formula, we can write:

$$P\{l_i = r_j \mid T\} = \frac{P\{l_i = r_j, T\}}{P\{T\}} \tag{6}$$

and applying the total probability theorem,

$$P\{l_i = r_j \mid T\} = \frac{\sum_{r1\in R} \dots \sum_{r_{i-1}\in R} \sum_{r_{i+1}\in R} \dots \sum_{r_m\in R} P\{l_1 = r_{1,\dots}, l_i = r_j, T\}}{\sum_{r1\in R} \dots \sum_{r_i\in R} \dots \sum_{r_m\in R} P\{l_1 = r_{1,\dots}, l_i = r_j, T\}} \tag{7}$$

We assume that the relationship between (l_i, r_j) and $(l_h = r_k)$ is independent of the relations of other pairs, and that (l_i, r_j) does not by itself provide any information on $(l_h = r_k)$ (Christmas et al., 1995). Factorizing the numerator and denominator in (7) we obtain:

$$P\{l_i = r_j \mid T\} = \frac{P\{l_i = r_j\} Q(l_i = r_j)}{\sum_{h=1}^{m} P\{l_i = r_h\} Q(l_i = r_h)} \tag{8}$$

where

$$Q(l_i = r_j) = \prod_{h=1, h\neq i}^{n} \sum_{j=1}^{m} P\{T(l_i, r_j; l_h, r_k) \mid l_i = r_j; l_h = r_k\} P\{l_h = r_k\} \tag{9}$$

The solution of the problem of line matching defined in (6) can be obtained by combining (8) and (9) in an iterative scheme (see Rosenfeld et al., 1976; Hummel and Zucker, 1983; Christmas et al., 1995). The previously computed similarity scores are taken here as the initial probabilities $P^0\{l_i = r_j\}$ for a possible match pair l_i and r_j. The constructed match pool greatly speeds up the probability relaxation process because only the lines in the match pool are involved.

$P\{T(l_i, r_j; l_h, r_k) \mid l_i = r_j, l_h = r_k\}$ is called compatibility function. Indeed, the compatibility function plays an important role in the process of structural matching. In our implementation, the evaluation is derived from the geometrical relation measurement between (l_i, l_h) and $(r_j = r_k)$. Three relations constitute the relational description between l_i and l_h (Fig. 5). They are the angle between l_i and l_h, the distance between l_i and l_h centers and the angle between line l_i and the line joining the center points of l_i and l_h.

Structural matching is conducted bidirectionally from left to right and from right to left. The next step is a combined 2D and 3D grouping of straight segments. Thereby, information in one space helps bridging gaps and combining segments in the other space. Thus, small gaps are bridged, edges broken in multiple straight segments are combined, matched segments of different length are extended.

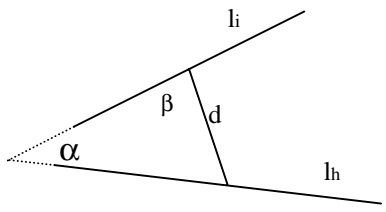

Figure 5. Three neighbouring line relations.

The final 3D position is computed from the original edge pixels and not the fitted straight lines. This is done for each pixel in the overlap length of the corresponding edges. A 3D straight line is then fitted. This approach can also handle the problem case when a road edge goes up and downhill but appears straight in 2D. This processing is very useful as we suppose to reconstruct roads using straight lines in both 2D and 3D spaces. The developed method for 3D straight line fit recursively subdivides a segment at the points with a certain deviation from a line connecting its endpoints, this

process is repeated and produces a tree of possible subdivisions. Then, unwinding the recursion back up the tree, a decision is made at each junction point as to whether to replace the current low level description with a single higher level segment by checking the connecting angle.

3.4 Experimental Results

The proposed method for straight line matching is implemented and experiments have been performed on a number of areas extracted from aerial images. The test areas cover different terrain and landcover, including rural areas, suburban, urban, and hills. Because of a limited space, the authors describe one of the dataset used, and report and analyze the major results. The dataset (Fig. 6) is extracted from a stereo pair at a test site of the ATOMI project. The line extraction process resulted in 985 and 971 straight lines in the left and right images respectively (see Fig. 7). Only 789 and 809 lines in the left and right images are longer than 5 pixels. 460 matches are found, of which only 5 are wrong matches. The matching result is shown in Fig. 8.

Figure 6. The test image pair.

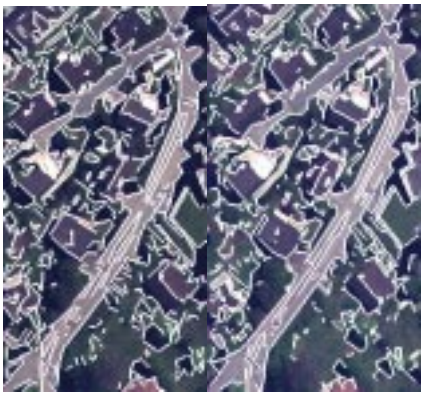

Figure 7. Extracted straight lines.

Figure 8. The matching results.

4 DISCUSSION AND CONCLUSIONS

We presented a new scheme for road reconstruction from aerial images. The proposed idea uses as much information as possible to increase success rate and reliability of the results. As one of the key components of the system, we presented a method for 3D line generation through stereo matching. The matching approach has high success rate and most importantly is very reliable. It makes use of rich attributes for matching, including line geometrical properties and line flanking regions photometrical and chromatic properties. This is an advantage over other approaches that only use line geometry and line gray scale information. The developed structural matching method achieves locally consistent results, allows matching in case of partially occluded edges, broken edges etc. The use of similarity scores priori to structural matching greatly speeds up the process. Although used here for straight edges, this method can be easily extended to arbitrary edges, or even points, if some of the matching criteria (feature attributes) are excluded or adapted.

Besides the work on straight line matching and 3D line generation, we completed a multispectral image classification method to find road regions. Guided by the initial knowledge base, we excluded the lines outside the road buffer area (this area is defined using the road centerlines of VEC25 and their estimated maximum error). By combining the 2D lines with the classification result, a relation with the road region (in, outside, at the border) is attached to each line. Lines with a slope difference to the slope from the known local DHM25 larger than a certain value are excluded. Our future work will focus on the extraction of roadsides by fusion of various 2D and 3D information extracted from images and knowledge from a road database, and modeling roads of different classes in various terrain relief and landcover. As reliability is the most important figure in our system, the system will use as much as possible information to create redundancy. Cues like cars on road and road marks will be extracted to confirm the reconstruction results. Finally, a metric for quality estimation will be developed, and the results will be evaluated with manually extracted ground truth reference data.

ACKNOWLEDGEMENTS

We acknowledge the financial support of this work and of the project ATOMI by the Swiss Federal Office of Topography. The discussions with IGP members and members of the projects ATOMI and AMOBE are gratefully appreciated. The first author would also like to thank Prof. A. Gruen for his guidance and comments.

REFERENCES

Baltsavias E.P., 1991. Multiphoto Geometrically Constrained Matching. Ph.D. Thesis, Report No. 49, Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland.

Cho W., 1996. Relational Matching for Automatic Orientation. IAPRS, Vol. XXXI, PartB3, Vienna, pp.111-119.

Christmas W. J., Kittler J., Petrou M., 1995. Structural Matching in Computer Vision Using Probabilistic Relaxation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(8), pp.749-764.

Eidenbenz Ch., Kaeser Ch., Baltsavias E., 2000. ATOMI – Automated Reconstruction of Topographic Objects from aerial images using vectorized Map Information. IAPRS, Amsterdam, Netherlands, Vol. XXXIII.

Foerstner, W., Pluemer L. (Eds.), 1997. Semantic Modeling for the Acquisition of Topographic Information from Images and Maps. Birkhaeuser Verlag, Basel.

Greenfeld J. S., Schenk A. F., 1989. Experiments with Edge-Based Stereo Matching. Photogrammetric Engineering and Remote Sensing, 55(12), pp. 1771-1777.

Gruen A., Li H., 1997. Semi-Automatic Linear Feature Extraction by Dynamic Programming and LSB-Snakes. Photogrammetric Engineering and Remote Sensing, 63(8), pp. 985-995.

Gruen A., Kuebler O., Agouris P. (Eds.), 1995. Automatic Extraction of Man-Made Objects from Aerial and Space Images. Birkhaeuser Verlag, Basel.

Gruen A., Baltsavias E.P., Henricsson O. (Eds.), 1997. Automatic Extraction of Man-Made Objects from Aerial and Space Images (II). Birkhaeuser Verlag, Basel.

Haralick R. M., Shapiro L G., 1993. Computer and Robot Vision, Volume II. Addison-Wesley Publishing Company.

Henricsson O., 1996. Analysis of Image Structure using Color Attributes and Similarity Relations. Ph.D. Thesis, Report No. 59, Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland.

Hinz S., Baumgartner A., Steger C., Mayer H., Eckstein W., Ebner H., Radig B., 1999. Road Extraction in Rural and Urban Aeras. SAMTI'99, Muenchen, pp. 133-153.

Hummel R.A., Zucker S.W., 1983. On the foundation of relaxation label process. IEEE Transaction on Pattern Analysis and Machine Intelligence, 5(3), pp.267-286.

McIntosh J. H., Mutch K. M., 1988. Matching Straight Lines. Computer Vision, Graphics, and Image Processing, 43, pp. 386-408.

Medioni G., Nevatia R., 1985. Segment-based stereo matching. Computer Vision, Graphics and Image Processing, 31, pp. 2-18.

McKeown Jr. D. M., Denlinger J. L., 1988. Cooperative methods for road tracking in aerial imagery. Proc. Computer Vision and Pattern Recognition, Ann Arbor, MI, June 5-9, pp. 662-672.

Niederoest, M., 2000. Reliable Reconstruction of Buildings for Digital Map Revision. IAPRS, Amsterdam, Netherlands, Vol. XXXIII.

Rosenfeld A., Hummel R. A., Zucker S. W., 1976. Scene labeling by relaxation operations. IEEE Transaction on System, Man, Cybernetics, 6(6), pp. 420-433.

Ruskone R., 1996. Road Network Automatic Extraction by Local Context Interpretation: Application to the Production of Cartographic Data. Ph.D. Thesis, Marne-La-Vallee University, France.

Trinder J., Nachimuthu A., Wang Y., Sowmya A., Singh S., 1999. Artificial Intelligence Techniques for Road Extraction from Aerial Images. IAPRS, Vol.32, Part 3-2W5, "Automatic Extraction of GIS Objects from Digital Imagery", Muenchen, Sep. 8-10.

Vosselman G., 1992. Relational Matching. Lecture Notes in Computer Science 628, Springer-Verlag, Berlin.

Wilson R. C., Hancock E. R., 1997. Structural Matching by Discrete Relaxation. IEEE Transaction on Pattern Analysis and Machine Intelligence, 19(6), pp. 634-648.

Zhang C., Baltsavias E.P., 1999. Road Network Detection by Mathematical Morphology. Proc. ISPRS Workshop "3D Geospatial Data Production: Meeting Application Requirements", Paris, France, pp. 185-200.

Zhang Z., 1994. Token Tracking in a Cluttered Scene. Image and Vision Computing, 12(2), pp. 110-120.