Report

Edge matching and 3D road reconstruction using knowledge-based methods

Author(s):
Zhang, Chunsun; Baltsavias, Emmanuel P.

Publication Date:
2000

Permanent Link:
https://doi.org/10.3929/ethz-a-004657168

Rights / License:
In Copyright - Non-Commercial Use Permitted
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 edge segments across images. The algorithm exploits edge geometric and photometric attributes and edge geometric structures. A framework to integrate these information sources using probability relaxation is developed and implemented to deliver locally consistent matches. Results of straight edge matching are presented and future work is discussed.

Prologue
I first encountered the name Wrobel in some publications related to image correlation when I was doing my Master Thesis at OSU at the beginning of the 80s. After coming to ETHZ in 1984, and since my Ph.D. topic was on matching, a research field where Prof. Wrobel had many parallel activities, we had many opportunities to meet and discuss. 1985 the workshop on matching in Darmstadt, December 1985 watching together the winter swimmers in Cannes, 1987 jogging (not me) in Interlaken, and of course the many ISPRS events and some conventions of DGPF. Thus, it came as no surprise that Prof. Wrobel was finally the co-referent for my Ph.D. I remember that he read my lengthy text with great attention, making very concrete proposals on what to change, and being prompt in his reactions even under very difficult circumstances with his family. But what impressed me more, was the fact that in my overview of matching methods, I made several critical remarks on the object-based matching method that he co-developed, and later refined under the name FAST. He respected my opinion, we discussed and cleared-up several aspects, but at no stage did he try to „press“ me to make only positive comments on „his“ algorithm. In August 91, I was on holidays with my wife and one of my brothers in Sardinia. I had to make corrections to my Ph.D., deliver a new version, and have my examination (on a Friday the 13th!) two weeks after the holidays’ end. Thus, I took some work with me, including Prof. Wrobel’s comments to study them. I know it’s a rare thing in Italy, but it did happen to us. Our car was broken in the middle of the day at a beach full of people and everything was stolen. We could only contact the local police next morning. Finally, at around 11.30 the police office opened. An officer, il colonelo X, was typing very seriously and very slowly at an old typewriter his report. At the end he asked whether we had lost any things of particular value. I stated a few personal things, including Prof. Wrobel’s remarks, not hoping for anything. In the afternoon, we got a phone call to go to the police station. All things of exceptional value had been found! Presumably thrown in a field. Prof. Wrobel’s notes did not have a scratch. They were still placed neatly in a plastic enclosure as I had put them, all in the proper sequence. I was happy, and my brother was swearing because he had forgotten to name some exceptional value things, which he could have received back. This is the story of Prof. Wrobel’s corrections, which were respected even by the robbers.

Apart from our personal relations, Prof. Wrobel had always had close ties to our group at ETHZ. He was twice with us as guest professor giving lectures, and helping young researchers, was co-referent of Ph.D. dissertations and sent us a group of students for a common Praktikum. I always admired his serious and helpful comments, his consistent and persistent research work, digging deep into the problems, and not just scratching the surface or jumping here and there on „modern“ topics. His is definitely not the type of show-man, but his contributions in photogrammetry, for us and the coming generations, are significant not just because of his scientific output on important topics, like matching, but also attitude-wise. His paradigm was: you don’t have to do many things simultaneously, but those you do, do them well! We wish Prof. Wrobel and his family health, spiritual satisfaction and happiness in his „Ruhe“stand. As contribution to his
Festschrift, the Ph.D. student Chunsun Zhang and me would like to present some research on edge-matching, a topic that is related and complimentary to Prof. Wrobel’s work on his area-based matching approach, and 3D road reconstruction.

Introduction

The extraction of road networks 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; Förstner and Plümer, 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 co-operation 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 colour 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 Niederöst (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 edges are extracted by fitting the detected edge pixels in each image and straight edge matching across images is performed through exploiting rich edge attributes and edge geometrical structures.

General strategy of road network extraction

A large number of approaches for road extraction have been proposed and published. The strategies fall into two broad categories. Semi-automatic schemes require an operator to provide interactively some information to control the extraction. In McKeown and Denlinger (1988), Vosselman and de Gunst (1997), Airault and Jamet (1996), the extraction algorithms are based on road tracking starting from an initial point and direction, extraction of parallel edges by extrapolation and matching of road profiles. With a few points of a road segment provided by an operator, (Gruen and Li, 1995, 1997) developed a LSB-Snake method to extract roads simultaneously in multiple images. This is advantageous because the solution is more constrained and the result is more robust, particularly in areas which are occluded in only some of the images. These semi-automatic approaches can be extended to fully automatic operation by means of automatic seed point detection (Zlotnick and Carnine, 1993). The common automatic methods first extract reliable hypotheses for road segments through line and edge detection and then establish connections between road segments to form road networks (Bajcsy and Tavakoli, 1976; Fischler et al., 1981; Ton et al., 1989; Wang et al., 1992; Trinder et al., 1999). The contextual information was taken into account to guide the extraction of roads in Ruskone (1996). In Baumgartner et al. (1997) and Mayer et al. (1997), road detection is based on the extraction of lines in an image of reduced resolution through scale-space analysis (Steger, 1998). Baumgartner then extracted edges in the original image, where the final result was the combination of results at two resolutions based on a set of rules, while Mayer adopted ribbon snakes to verify roads and discriminate them from other line-type objects in the original images by means of width consistency. Knowledge-based methods involve the use of existing GIS databases or maps and rule-based systems. Stilla (1995) and Bordes et al. (1997) used maps and cartographic databases respectively as a guide for image interpretation. In Vosselman and de Gunst (1997), the old database is used not only to verify it but also to detect new road branches from the given data.
In ATOMI, we develop a new approach for automatic extraction of 3D road network from aerial images which integrates knowledge processing of colour image data and existing digital geographic databases. The information of existing road database provides a rough model of the scene. Colour aerial images give the current situation of the scene, but are complex to analyse 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. Colour cues, expressed in the form of colour 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 edges 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**

**Straight edge matching**

3D edge generation is a crucial component of our procedure. We are interested in 3D straight edges 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 localised more accurately. The 3D information of straight edges is determined from the correspondences of edge segments between two images.

Due to the complexity of aerial images, different view angles and occlusions, straight edge matching is a difficult task. Existing approaches to edge matching in the literature are generally categorised into two types. One is directly
comparing the attributes of a edge in one image with those of a set of edges 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 edge attributes, such as orientation, length, edge support region information etc. In another strategy, the edge 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).

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

The developed method in this paper exploits rich edge attributes and edge geometric structure information (Fig. 3). The rich edge attributes include the geometrical description of the edge and the photometric information in the regions right beside the edge (flanking regions). The epipolar constraint is applied to reduce the search space. The similarity measure for an edge pair is first computed by comparing the edge 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 edge matching

**Edge extraction**

The input images are first filtered with Wallis filter for contrast enhancement and radiometric equalisation (Baltsavias, 1991). The technique developed in a previous project (AMOBE) is used to extract straight edges (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 edge. For each straight edge 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 colour space conversion and include the median and the scatter matrix.
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 an edge segment in one image generate 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 edge included in this band (even partially) is a possible candidate, if it intersects the two epipolar lines (through the two edge endpoints in the left image) within this band. For example in Fig. 4, edges i, j, k are accepted and will be compared with edge \( pq \) in the left image for similarity measurement, while edge \( r \) is rejected because it intersects \( eq \) outside \( q_1, q_2 \).

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.

For each pair of edges 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 edges, divided by the maximum one. Thus, the similarity score is defined as

\[
V_{\text{len}} = \frac{\min(l_{\text{len}}, R_{\text{len}})}{\max(l_{\text{len}}, R_{\text{len}})}
\]  

(1)

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

\[
V_{\text{ang}} = w \cdot \frac{T_{\text{ang}} - |L_{\text{ang}} - R_{\text{ang}}|}{T_{\text{ang}}}
\]

(2)

where \( w \) is a weight related to edge length, and given by \( w = w_p w_r \), where \( w_p, w_r \) are computed for the edges in left and right image respectively. They are defined as a piecewise linear function as:

\[
\begin{align*}
& 0; \quad \text{len} < 5; \\
& \frac{1}{t - 4} (\text{len} - 4); \quad 5 \leq \text{len} < t \\
& 1.0; \quad \text{len} \geq t
\end{align*}
\]

(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 edge geometric properties, but also of photometric and chromatic properties of edge flanking regions. The photometric and chromatic edge 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 edges 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 edge pair is computed as:
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 edge pair. In our matching algorithm, we do not assume that the edge pair has same contrast; instead we only request that at least one side of the edge pair they demonstrate similar brightness. Thus, if both \( C_L \) and \( C_R \) are less than a predefined value, then the two edges are treated as different edges, 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.

\[
V_{\text{left}} = wC_L \\
V_{\text{right}} = wC_R
\]

The chromatic property of a region in the (a, b) data is represented by the scatter matrix \( C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \). 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:

\[
\text{Dir} = \frac{1}{2} \arctan \frac{2 \cdot C_{12}}{C_{11}^2 + C_{22}^2}, \quad Q = 4 \frac{\text{Det}(C)}{\text{tr}^2(C)}
\]

The similarity measurement of a chromatic property is performed by comparing ellipse orientation and roundness in the left and right regions for a edge 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 edges, while the very short ones (< 5 pixels) are ignored.

**Structural matching with probability relaxation**

After performing similarity measurement computation, we construct a matching pool and attach a similarity score to each edge 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 edge is broken or occluded. In addition, matching using a very local comparison of edge 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 edge correspondence is realised through probability relaxation.

The edge structure in an image is described as a graph, where the nodes of the graph represent straight edges, and the links between edges the relations. To find a correspondence, both individual edges and the graphs should be matched. We represent the straight edges in the left image as a set \( L = \{ l_i \}, i = 1, 2, ..., n \), the straight edges in the right image as a set \( R = \{ r_j \}, j = 1, 2, ..., m \). 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 edge \( l_i \) matches \( r_j \), i.e. the matching problem becomes the computation of \( P \{ l_i = r_j | T \} \). Using Bayes formula, we can write:

\[
P\{l_i = r_j | T\} = \frac{P\{l_i = r_j, T\}}{P\{T\}}
\]

and applying the total probability theorem,
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). Factorising the numerator and denominator in (7) we obtain:

\[
P(k = n_j | T) = \frac{\sum_{r \in R} \sum_{r_i \in R} \sum_{r_l \in R} P(l = r_i \ldots l = n_j, T)}{\sum_{r \in R} \sum_{r_i \in R} \sum_{r_l \in R} P(l = r_i \ldots l = n_j, T)}
\]  

\(P(k = n_j | T)\)  

\( \sum_{k=1}^{m} P(k = n_j)Q(l = n_j)\)  

where

\[
Q(k = r_j) = \prod_{h=1, h \neq i}^{n} \sum_{j=1}^{m} P(T(l, r; l_h, r_h) | l_i = r_i, l_h = r_k)P(l_h = r) \]

The solution of the problem of edge 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 \(P^0(l = n_j)\) for a possible match pair \(l_i\) and \(r_j\). The constructed match pool greatly speeds up the probability relaxation process because only the edges in the match pool are involved.

\[P(T(l, r; l_h, r_h) | l_i = r_i, l_h = r_k)\] 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 \(l_h\) centres and the angle between edge \(l_i\) and the edge joining the centre 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 edge relations](image)

The final 3D position is computed from the original edge pixels and not the fitted straight edges. 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.

**Edge-matching results**

The proposed method for straight edge 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 analyse the major results. The dataset (Fig. 6) is extracted from a stereo pair at a test site of the ATOMI project. The edge extraction process resulted in 985 and 971 straight edges in the left and right images respectively (see Fig. 7). Only 789 and 809
edges 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.

3D road reconstruction

Road model
Several road models have been developed and can be found in literature. The basic model uses geometric and photometric attributes of roads, and relations between roads and background objects. In high resolution aerial images, roads are generally modelled as linear objects with parallel boundaries, bounded horizontal and vertical curvatures, constant width, and homogeneous area in between. These models are linked to the characteristics of the imagery (the resolution particularly) and to the characteristics of roads which have to be detected. Most often the road model chosen would not be suitable for another type of road or in another type of landcover. On the contrary, if the road model is too generic, other objects such as rivers or railways may be detected. In our knowledge-based road extraction system, a more specific road model can be derived from the constructed knowledge base, using the VEC25 data. Each road in VEC25 provides the rough location of a road in the scene, as well as road attributes, such as road class, type, presence of road marks, geometry with 3D information (width, length, horizontal and vertical curvature, landcover and so on). Fig. 9 shows some examples of roads in the Albis test dataset used in ATOMI. The road surface might include other objects (cars, shadows etc.), or be occluded by neighbouring objects.
Methodology for road centreline reconstruction

In our investigations, we focus first on roads in rural and suburban areas. In this case, often both roadsides are visible. The developed method extracts roads in such areas by finding 3D parallel lines that belong to a road and link them in sequence. In case of shadows, occlusions caused by trees, buildings etc., spatial reasoning is applied using the existing knowledge base. One of the important features of the knowledge base is that all information in it is spatially related, and relations between 2D edges and their corresponding 3D straight lines are kept. The method is shown in Fig. 10. The key of the method is the use of the knowledge as much as possible, working in 2D images and 3D space, use of 2D and 3D interaction when needed, and reasoning about possible problem areas. The details are reported below.

The method starts from finding 3D parallel lines. Only the lines on the ground located in the buffer defined by VEC25 and having similar orientation to VEC25 segments, and a slope less than a threshold are kept for further processing. Since roads are on the ground, lines above ground are removed by checking with the DHM25. By checking the image classification results, lines within a road class region (i.e. not close at its boundaries) are also removed. A line pair \((i, j)\) is considered as parallel, if the lines have similar orientation in 3D space. The measurement of the parallelity is based on two measures \(S_\alpha\) and \(S_\beta\) (which should ideally be equal to 1) as

\[
S_\alpha = \frac{T_\alpha - |\alpha_1 - \alpha_2|}{T_\alpha}, \quad S_\beta = \frac{T_\beta - |\beta_1 - \beta_2|}{T_\beta}
\]

\(\alpha_1, \beta_1\) and \(\alpha_2, \beta_2\) are directions in the XY plane and slopes respectively of the 3D lines \(i\) and \(j\) and \(T_\alpha\) and \(T_\beta\) thresholds. Lines \(i\) and \(j\) must overlap, and the distance between \(i\) and \(j\) must be within a certain range. Minimum and maximum distances depend on the road class as defined in VEC25.

The found 3D parallel lines are projected to the 2D images for photometric evaluation. Firstly, the region between the projected lines must belong to the class “road” as determined by image classification. Image processing tools will be activated to extract road marks and detect cars within this region (for additional cues about the road existence and its approximate position), however this part has not been implemented yet.

The parallel lines passing the above check are considered as Possible Road Sides that are Parallel (PRSP). They compose a graph. The nodes of the graph are PRSPs, the arcs of the graph are the relations between PRSPs. In occlusion areas, the arcs also represent the missing parts of the road between a pair of PRSPs. The width of two PRSPs should be similar. If there is no gap between two PRSPs with similar width, i.e. one PRSP shares points with the other, and the linking angles between them both in XY plane and 3D space comply with the VEC25 data, they are connected directly. The width similarity measurement is defined as

\[
Sw = \frac{\text{min (widths of PRSPs)}}{\text{max (widths of PRSPs)}}
\]

In case of gaps between PRSPs, the gap area is checked. This is called spatial reasoning in our work. If the gap is not too long, based on a threshold, and

1) Within the gap there is road region, e.g. a parking lot right beside the road, or
2) Within the gap there is shadow, or shadow mixed with road region, or
3) The gap is caused by tree occlusion, or
4) Within the gap there is terrain as determined by the DSM, or
5) There are extracted road marks within the gap (not implemented yet),

and the connecting angles between PRSPs and gap comply with the VEC25 (see Fig. 11), then a link is made for the two PRSPs.

Once two PRSPs are considered as connectable, the relation between them is evaluated by

\[
S_\alpha = w \left(1.0 - \frac{\alpha}{180.00}\right), \quad S_\beta = w \left(1.0 - \frac{\beta}{180.00}\right)
\]

where \(w\) is a weight defined using the gap length, as

\[
\begin{cases} 
0; & L_{\text{gap}} \geq T; \\
1.0 - \frac{L_{\text{gap}}}{T}; & L_{\text{gap}} < T
\end{cases}
\]
A road hypothesis is found by searching the graph using a depth-first method. Each hypothesis is associated with a score that is the summation of the relation measurement of the PRSPs it contains. Currently, the hypothesis with the highest score is selected as road.

Figure 10. Flow chart of developed method
Test results

The developed method was implemented, and several tests were conducted in rural and suburban areas. Fig. 12a shows the matched straight edges of Fig. 9a after removal of irrelevant edges. In Fig. 12b the found PRSPs are displayed as thick white lines, and the extracted road centreline is presented in Fig. 12c.

Another test is shown in Fig. 13, where a dirt field road is extracted. The found PRSPs and extracted road centrelines are shown in one image. Fig. 14 shows an extracted rural road with building occlusions and Fig. 15 a case with tree occlusions.
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 matching of straight edges. The matching approach has high success rate and most importantly is very reliable. It makes use of rich attributes for matching, including edge geometrical properties and edge flanking regions photometric and chromatic properties. This is an advantage over other approaches that only use edge geometry and edge grey 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 prior 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 edge 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 centrelines of VEC25 and their estimated maximum error). By combining the 2D edges 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. We developed (currently incomplete) methods and presented first results for road extraction in rural and suburban areas. The key of the developed method is the use of knowledge, working in 2D image and 3D space, and use of 2D and 3D interaction when needed. The road hypotheses are generated in 3D. This not only enables us to apply more geometric criteria to generate hypotheses, but also largely reduces search space. The hypotheses are evaluated in the images, using photometric information. Whenever 3D features are incomplete or entirely missing, 2D information is used to infer the
missing features. By incorporating multiple knowledge sources, the problem areas caused by shadows and occlusions can be handled. This is realised in our development through spatial reasoning.

The developed method still needs improvement so that it can work with various road classes in rural and suburban areas. As reliability is the most important figure in our system, the system will use as much as possible information to make use of redundancy. Cues like cars and road marks will be extracted to confirm the reconstruction results. A metric for quality estimation will be developed, and the results will be evaluated with manually extracted ground truth reference data. Based on the gained experiences, a more general method will be developed to handle the problems of road extraction in more complex areas, i.e. urban ones and city centres.

Acknowledgements
We acknowledge the financial support of this work and of the project ATOMI by the Swiss Federal Office of Topography.

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


