Report

Automated updating of road databases from aerial imagery

Author[s]:
Baltsavias, Emmanuel P.; Zhang, Chunsun

Publication Date:
2003

Permanent Link:
https://doi.org/10.3929/ethz-a-004657154

Rights / License:
In Copyright - Non-Commercial Use Permitted
Automated updating of road databases from aerial imagery

Emmanuel Baltsavias, Chunsun Zhang

Institute of Geodesy and Photogrammetry, Swiss Federal Institute of Technology Zürich
ETH-Hönggerberg, CH-8093 Zürich, Switzerland
{manos, chunsun}@geod.baug.ethz.ch

Abstract

This paper presents a practical system for automated 3-D road network reconstruction from aerial images using knowledge-based image analysis. The system integrates processing of color image data and information from digital spatial databases, extracts and fuses multiple object cues, takes into account context information, employs existing knowledge, rules and models, and treats each road subclass accordingly. The key of the system is the use of knowledge as much as possible to increase success rate and reliability of the results, working in 2-D images and 3-D object space, and use of 2-D and 3-D interaction when needed. Another advantage of the developed system is that it can correctly and reliably handle problematic areas caused by shadows and occlusions. This work is part of a project to improve and update the 1:25,000 vector maps of Switzerland. The system was originally developed to processed stereo images. Recently, it has been modified to work also with single orthoimages. The system has been implemented as a stand-alone software package, and has been tested on a large number of images with different landscape. In this paper, various parts of the developed system are discussed, and the results of our system in the tests conducted independently by our project partner in Switzerland, and the test results with orthoimages in a test site in the Netherlands are presented together with the system performance evaluation.

Keywords: roads; 3-D reconstruction; map updating; knowledge-based image analysis; color; context; edge matching; DSM/DTM; roadmarks; multiple cue integration; performance evaluation; semi-automated system

1. Introduction

The extraction of roads from digital images has drawn considerable attention in the last few decades, and is still one of the current challenges in digital photogrammetry and computer vision. Various existing and emerging applications require in particular up-to-date, accurate and sufficiently attributed road databases, including car navigation, tourism, traffic and fleet management and monitoring, intelligent transportation systems, internet-based map services, location-based services, generation of digital landscape models, etc. The fact that vendors of commercial photogrammetric, remote sensing and GI systems do not offer anything useful regarding automation of road extraction (not even practical semi-automatic methods) stresses the importance of this research topic.

The existing approaches for road extraction cover a wide variety of strategies, using different resolution aerial or satellite images. A quite extensive overview of such approaches is given in Zhang (2003a). Semi-automatic schemes require human interaction to provide interactively some information to control the extraction. Roads are then extracted by profile matching (Vosselman and Knecht, 1995; Airault et al., 1996), cooperative algorithms (McKeown et al., 1988), and dynamic programming or LSB-Snakes (Grün and Li, 1997). Automatic methods usually extract reliable hypotheses for road segments through edge and line detection and then establish connections between road segments to form road networks (Wiedemann et al., 1998). Data from different sources is often useful (Price, 1999; Hinz et al., 2001). Contextual information is taken into account to guide the extraction of roads (Ruskone, 1996; Heipke et al., 2000). Roads can be detected in multi resolution images (Baumgartner and Hinz, 2000). Several applications use map information (Willrich, 2002; Gerke et al., 2003). The map data is used either as approximation to start tracking or optimization process by Snakes (Bordes et al., 1997; Agouris et al.,...
2001), or to search for new roads (Vosselman and de Gunst, 1997). The road maps can be updated by map-image matching (Klang, 1998; Fiset et al., 1998). The existing approaches show individually that the use of road models and varying strategies for different types of scenes are promising. However, all the methods are based on relatively simplistic road models, and most of them make only insufficient use of a priori information, thus they are very sensitive to disturbances like cars, shadows or occlusions, and do not always provide good quality results. Furthermore, most approaches work in single 2-D images, thus neglecting valuable information inherent in 3-D processing.

In this paper, we present a practical system for automatic extraction of 3-D roads from aerial images that integrates processing of colour image data and existing digital spatial databases. The reported work has been carried out within the project ATOMI (Automated reconstruction of Topographic Objects from aerial images using vectorized Map Information), in cooperation with the Swiss Federal Office of Topography (L+T) with the aims to use aerial images and DTM data and automated procedures to improve vector data (road centerlines, buildings) from the digitised 1:25,000 topomaps by fitting them to the real landscape, updating them, improving the planimetric accuracy to 1 m and providing height information with 1-2 m accuracy. The extracted objects should have a high success rate, and more importantly should be reliable. After some initial work, the aims of ATOMI were restricted to improvement of the existing roads (i.e. no extraction of new roads) with first target the open rural areas. We currently use color aerial imagery of 0.22 m pixel size and 1:16,000 scale with 30 cm focal length. The existing geodatabase, called VEC25, is generated by digitization of the 1:25,000 topographic maps. The RMS error of VEC25 is ca. 2.5 m -7.5 m and the maximum one ca. 12.5 m. VEC25 roads are categorized into several classes, corresponding to the legend of the national map 1:25,000. In VEC25, the landcover map is also included, where the residential areas and the forests are distinguished. Other input data include a national DTM with 25 m grid spacing, and a DSM generated from the stereo images using MATCH-T of Inpho with 1 m or 2 m grid spacing. More details about ATOMI can be found in Eidenbenz et al. (2000).

This paper is organized as following. After this section we describe the general strategy in Section 2, while the details for the extraction of various cues are presented in Section 3. Section 4 explains the procedures of cue combination for road reconstruction, junction generation, as well as the measures for self-diagnosis and external evaluation. In Section 5, results are presented, together with the quality evaluation by comparing the results with manually measured data. The paper is concluded in Section 6 with a discussion and outlook.

2. General strategy

The developed system makes full use of available information about the scene and contains a set of image analysis tools. The management of different information and the selection of image analysis tools are controlled by a knowledge-based system (Fig. 1). The initial knowledge base is established by the information extracted from the existing spatial data and road design rules. This information is formed in object-oriented multiple object layers, i.e. roads are divided into various subclasses according to road type, land cover and terrain relief. It provides a global description of the road network topology, and the local geometry for each road. Therefore, we avoid developing a general road model; instead a specific model can be assigned to each road segment. This model provides the initial 2-D location of a road in the scene, as well as road attributes, such as road class, presence of roadmarks, and possible geometry (width, length, horizontal and vertical curvature, land cover and so on). A road is processed with an appropriate method corresponding to its model, certain features and cues are extracted from images, and roads are derived by a proper combination of cues. The knowledge base is then automatically updated and refined using information gained from previous extraction of roads. The processing proceeds from the easiest subclasses to the most difficult ones. Since neither 2-D nor 3-D procedures alone are sufficient to solve the problem of road extraction, we make the transition from 2-D image space to 3-D object space as early as possible, and extract the road network with the mutual interaction between features of these spaces. More details of the general strategy can be found in Zhang (2003a).
3. Feature and cue extraction

For each VEC25 road, our system focuses on the image regions around it. The regions are defined using the position of the road and the maximum error of the VEC25. Then, a set of data processing tools is activated to extract features and cues. 3-D straight edge generation is a crucial component of our procedure because the roadsides are among them, and they provide road edges with higher geometric accuracy than the other cues used. The 3-D information of straight edges is determined from the correspondences of edge segments between stereo images. An image classification method is implemented to find road regions. With the DSM and DTM data, the above-ground objects and ground objects are separated. We also exploit additional cues such as roadmarks and zebra crossings to support road extraction.

3.1. 3-D straight edge generation

Due to the complexity of aerial images, different view angles and occlusions, straight edge matching for 3-D edge generation is a difficult task in photogrammetry and computer vision. We developed a structural matching method that exploits rich edge attributes and edge geometrical structure information (Zhang and Baltsavias, 2000). The edge segments are extracted by the Canny operator in stereo images. We then compute the edge geometrical and photometrical attributes. The similarity measure for an edge pair is computed by comparing the edge attributes. The similarity measure is then used as a prior information in structural matching with probability relaxation, through which the locally consistent matching is achieved. The method seeks the probability that an edge in one image matches an edge in the other image, using the geometrical structure information and photometric information of the neighbouring image edges. The matching approach has a high success rate and most importantly is very reliable. Fig. 2 shows an example of straight edge extraction and matching in a suburban area. In this example, totally 81 and 75 straight edges are extracted in left and right image, respectively, 41 matches are found, of which only 1 is not correct.
Fig. 2. Straight edge extraction and matching. The extracted straight edge segments and the matched edge segments are shown in white and black lines respectively. (a) left image, (b) right image.

The matched edge segments are then transformed to object space by finding the corresponding pixels within each matched edge pair (Zhang and Baltsavias, 2002). After the computation of 3-D edge pixels, 3-D straight edge segments are fitted to the 3-D edge pixels.

3.2. Image classification for road region detection

ISODATA (Jain and Dubes, 1988) is employed in our system to classify the color images and separate road regions from other objects. The algorithm automatically classifies the image data into desired clusters. It recursively generates a new partition by assigning each pattern to its closest cluster center and merges and splits existing clusters or removes small or outlier clusters. The success of image classification also depends on the quality of the input data. In our system, the original RGB image is transformed into HSI color space, and is also used to compute several artificial bands/indices to enhance features such as vegetation and shadows so that they are more isolated in feature space. The following 3 bands are selected for image classification: (1) the first component of the principal component transformed image, (2) a greenness band calculated with R and G bands in RGB space as (G-R)/(G+R); and (3) S band of HSI color space. With this classification, we determine five classes, corresponding to road regions, vegetation, shadow areas, dark roofs and red roofs. An example is given in Fig. 3b, showing the image classification result for the scene of Fig. 3a. The classification approach can be improved by modifying algorithms, input bands and number/type of classes.

3.3. DSM and DTM analysis

The DSM and DTM are used to support road extraction in our system. Subtracting the DTM from the DSM results in a so-called normalized DSM (nDSM), which enables the extraction of above-ground objects, including buildings and trees. The nDSM is employed in our 3-D road extraction system to verify if a 3-D straight edge is on the ground. The edges above the ground are thus discarded from further processing since roadsides are almost always on the ground. This information can be also used to reason if a region is on the ground. Since the DTM data in our project is not very accurate, subtracting the DSM from this DTM may lead to wrong results. Thus, we extract above-ground objects and the nDSM from the DSM data only, therefore avoiding introducing errors from the DTM. This is achieved by a Multiple Height Bin (MHB) method proposed by Baltsavias et al. (1995). The method is simple, fast and very effective. Fig. 3c shows the detected above-ground objects in Fig. 3a superimposed on the image in white. By combining the information of nDSM with image classification data, our system creates redundancy to confirm the existence of roads. Furthermore, it can partly compensate missing or wrong information in the classification data.

3.4. Roadmark and zebra crossing extraction

Roadmarks and zebra crossings are good indications of the existence of roads. They are generally found on main roads and roads in urban areas. Both of them have a distinct color (usually white or yellow). In
high-resolution images, such as the ones used in our project, roadmarks are white thin lines with a certain width. The zebra crossings are shown as yellow strips. The roadmarks give the road direction and even the road centerline, while the zebra crossings define the local road width and local road direction. Thus, they can be used to guide the road extraction process or verify the extraction results. In addition, in many cases the correct road centerlines can be even derived directly from the present roadmarks and/or zebra crossings. This is especially useful when the roadsides are occluded or not well-defined, such as in urban areas or city centers.

Since roadmarks are usually white, the image is first segmented by thresholding in R, G, B bands. The roadmarks are then extracted using an image line model and the geometrical description of each roadmark is obtained (Zhang, 2003a). The 3-D roadmarks are generated by our developed structural matching method. Zebra crossings are composed of several thin stripes with equal interval between. Using color information, the image is first segmented. We then obtain several clusters though morphological closing followed by connected labeling. Only the clusters with a certain size are kept, while the small ones are discarded. Then, the shape of the cluster is analyzed. The rectangle-like clusters are selected as zebra crossings. The center, the short and long axes of the detected zebra crossings are computed using spatial moments. Fig. 3d and Fig. 3e present the detected roadmarks and zebra crossing in the scene of Fig. 3a respectively. The detected roadmarks are shown in black lines superimposed on the image. In Fig. 3e, the center, the short and long axes of the detected zebra crossings are presented.

Fig. 3. Examples of image classification, above-ground object detection, roadmark detection and zebra crossing extraction. (a) Image. (b) Classification result. (c) The detected above-ground objects from DSM data are superimposed on the image in white. (d) The detected roadmarks shown as black lines. (e) The center, long and short axes of the detected zebra crossing.

4. 3-D road reconstruction

With the extracted features and cues, our next step is to combine them to extract the road. Firstly, irrelevant edges are removed. 3-D parallel overlapping edges are then searched for, and are evaluated to find possible roadsides. Due to occlusions or shadows, a road and its sides may be totally or partially invisible in images. The road segments with only one side visible in one or two images in the occluded/shadowed areas are inferred and reconstructed. In addition, gaps, where the roadsides are totally invisible, are bridged. Finally, the road is reconstructed by finding an optimal path among the road segment candidates that maximizes a merit function. Highways, first class roads and most second class roads are also extracted using the detected roadmarks and the zebra crossings. With the extracted roads, the road junctions are reconstructed and modelled. The main procedures are described below. More details on all about procedures are given in Zhang (2003a).
4.1. Finding 3-D parallel roadsides

The system checks extracted edges to find 3-D parallel edges. Only edges located in the buffer defined by the VEC25, having a similar orientation to the VEC25 segments and a certain slope are further processed. Edges above ground are removed by checking with the nDSM. Two edges are considered as parallel if they have similar orientation in 3-D space. The edges of a pair must overlap in the direction along the edges, and the distance between them must be within a certain range determined by the road class defined in the VEC25. In addition, the heights of the two edge segments should be similar. The found 3-D parallel edges are projected onto the images and evaluated using multiple knowledge. The region between the projected edges must belong to the class road as determined by the image classification. If roadmarks are presented on this road, the extracted roadmarks are used to confirm that the edge pair corresponds to correct roadsides. The found 3-D parallel edges have high probability of belonging to a road; they are called Possible Road Sides that are Parallel (PRSP). Thus, each PRSP is geometrically described by a pair of 3-D straight edges and its corresponding 2-D edges, and holds a set of attributes.

4.2. Reconstruction of missing roadsides

Not all the 3-D road segments can be obtained from the procedures described above. The absence of 3-D roadsides can be caused by shadows, occlusions, or roadsides do not actually exist, e.g. in the area where a parking lot is situated next to the road. Depending on the relations between the road segments and the neighbouring objects, sun angle, viewing direction, existence of moving cars on the road etc., there are various types of missing roadsides in images. We have made an investigation and classified them into 11 types. Each type of them is then treated, and the missing roadsides are inferred and validated by a specific procedure (Zhang, 2003a). As an example, Fig. 4 shows three cases of missing roadsides, together with the solution provided by the system. In the figure, the image patches of the occluded/shadowed road segments (where the missing roadsides occur) are shown together with the 2-D and 3-D edges extracted in the problematic areas. The 2-D and 3-D straight edge segments are shown in white and black lines respectively. The quadrilaterals with dark grey lines are the generated PRSPs. The dashed black lines in the figures are solutions for the missing roadsides provided by our system.

![Fig. 4. Examples of missing roadsides. See text for details.](image)
Based on the type of the missing roadsides, corresponding types of road segment candidates (RSCs) are obtained. They are then evaluated using the knowledge obtained from the cues. Common reliability measures for all RSCs are evaluated using the results from the image classification and the nDSM. This is carried out in image space and object space respectively. However, the evaluation process also uses the interactions between these two spaces. In addition, specific reliability measures for each type of RSC are also computed. The specific reliability measures mainly depend on the relations between the RSC and its adjacent PRSPs. We refer to Zhang (2003a) for the detailed procedures of RSC evaluation.

4.3. Gap bridging

As a special case of missing roadsides, a gap between neighbouring PRSPs that belongs to a road actually represents a road part where no roadside is visible in the images. In our system, a gap is bridged either by directly linking the neighbouring PRSPs or by adapting the shape of the VEC25 road corresponding to the gap area. That is, the vertices of the VEC25 road in the gap area are shifted to close the gap, based on the coordinate differences between the end points of the PRSPs and their corresponding points on the VEC25 road. Linking the adjacent PRSPs to close gaps is efficient on straight roads with short gap length, while using the shape of VEC25 road might be more useful for long and curved occlusion areas. We determine the solution for gap bridging in an evaluation process using various knowledge, e.g. the shape of the solution for the gap should approximately comply with that of the VEC25 road; it should be either a road region or a shadow or shadow mixed with road region; or roadmarks are extracted within the hypothesised road area. Based on the evaluation, we compute a measure for the gap hypothesis, \( s_{gap} \). The range of values for \( s_{gap} \) is [0,1], with decreasing value for long and inconsistent gaps. Fig. 5 shows the gap hypotheses generated by the above two methods in an occluded road area (between the two white dots). The VEC25 roads, the PRSPs and the road centerlines are presented in white, dark grey and black lines respectively. The light grey lines are the generated gap hypotheses. In Fig. 5a the hypothesis is created by directly linking the neighbouring PRSPs, while in Fig. 5b, it is created by adapting the shape of the VEC25 road in the gap area. The value of \( s_{gap} \) for the hypothesis in Fig. 5a is below 0.1. The value is 0.58 for the hypothesis in Fig. 5b. Thus, the hypothesis in Fig. 5b is used to bridge the gap.

![Fig. 5. Comparison of gap hypotheses generated by directly linking the neighbouring PRSPs and by adapting the shape of the VEC25 road. The VEC25 roads, the PRSPs and the road centerlines are shown in white, dark grey and black lines respectively. The light grey lines are the gap hypotheses. (a) gap hypothesis created by directly linking the neighbouring PRSPs, (b) gap hypothesis created by adapting the shape of the VEC25 road in the gap area.](image)

4.4. 3-D road segment linking for 3-D road reconstruction

With the extracted PRSPs, the road is reconstructed by linking the PRSPs belonging to a road. The goal of linking road segments is twofold. First of all, this implies that PRSPs belonging to a road should be selected and connected with the gaps bridged by the linking algorithm. Secondly, this also implies that PRSPs not belonging to a road should be rejected. Therefore, the algorithm must be very selective in which PRSP it adds to a road. The linking function in our system is defined as
\[
\sum l_i \cdot s_i + l_{gap} \cdot s_{gap} + l_j \cdot s_j
\]  

where, i and j are adjacent PRSPs, \(l_i\) and \(l_j\) are their lengths, \(s_i\) and \(s_j\) are their reliability measures. \(l_{gap}\) is the gap length between i and j, and \(s_{gap}\) is the gap evaluation measure.

The function takes high values for long curves with a shape similar to the VEC25 road. From the definition of the linking function, the goal now is to find a subset among all PRSPs that maximizes the linking function. The subset that provides the maximum value is the extracted road. This can be achieved using dynamic programming (Gruen and Li, 1997).

Fig. 6 demonstrates the performance of the linking algorithm. In the figure, the VEC25 roads are shown in white, the extracted straight edge segments and the PRSPs are shown in light grey and solid black lines respectively. The dashed black lines are the RSCs. The extracted road centerline is presented in dark grey line. The result shows the algorithm works very well. The PRSPs are correctly connected, and the roads are reconstructed.

Fig. 6. 3-D road segment linking for road reconstruction. The VEC25 roads, the extracted straight edge segments and the PRSPs are shown in white, light grey and solid black lines respectively. The dashed black lines are the RSCs. The extracted road centerline is presented in dark grey line.

For main road classes, on which the system knows that roadmarks are present, the system also extracts them using the detected roadmarks and zebra crossings. The roadmarks are linked using a similar method as described in the previous paragraph. This procedure increases the effectiveness and reliability of our system. In complex areas, such as in city centers, the roadsides are generally occluded very much, while sometimes they are not defined. However, the road centerlines are successfully extracted by the system using roadmarks. In rural and suburban areas, the extracted road using roadmarks is used by the system to verify the extraction results using 3-D parallel edges.

4.5. Road junction generation and modelling

Road junctions are important features of the road network. However, it is even more difficult to model and extract road junctions from images than road segments. This might be one of the reasons that this issue has been rarely touched in past research. In our system, we reconstruct junctions through intersecting the extracted roads guided by the topology of the VEC25 data, and further model the junctions with road class information and the shape of the VEC25. During the process, each road is assigned a weight corresponding to the evaluation value of the road segment in the junction area. In addition, an angular constraint is applied. That is, the angle of the two reconstructed roads at the reconstructed junction point should be similar to that defined by the corresponding VEC25 roads at the VEC25 junction point. Furthermore, in the angular constraint higher priority is given to higher class roads or continuous roads. In such formulation, the junction point is correctly located while the shape of the roads at the junction point is well formed: complying with the reality and following the road design
rules. Figs. 9 and 12 show several examples of road junctions generated by our system. Taking Fig. 9(d) as an example, it can be seen that the results of the higher class roads continue smoothly through the reconstructed junction points allowing fast, smooth and safe traffic, while the results of the lower class roads meet the higher class roads at the junction points with similar angles to the VEC25 data, while following the image roads.

4.6. Performance evaluation

With the extracted roads and road junctions, the road network is obtained. The results inherit other attributes from the VEC25 data with the road lengths updated, and road widths appended. The number of lanes can be inferred from the known road width and also possibly roadmarks. The 3-D information permits the derivation of other useful attributes like horizontal and vertical curvatures and road slope.

We introduce two types of measure for self-diagnosis of the extraction results: an overall quality measure for the whole road, and measures for the road segments. If a result does not pass the overall quality test, a further test is conducted to find in which segments the errors occur. The overall quality of the extraction result can be obtained from the following criteria:

- The lengths of the extraction result and the VEC25 road should be similar
- The shape difference between the extraction and the VEC25 road should be small
- Total length of PRSPs should cover a large part of the extraction result

We also define an internal quality measure for each road segment, especially gap segments, using the shape similarity measure between the segment and the corresponding VEC25 road, and the information in the classification result and the nDSM in the segment area. The detailed computation and examples are given in Zhang (2003a) and Zhang (2003b). For first and second class roads, the assessment of the reliability of the extraction results is also done through comparison of the results by using parallel edges and roadmarks. If the two extraction results are similar, the extracted road centerlines are considered reliable. Otherwise, if they differ, the system will give a low reliability to the result, and ask the operator to check it. Fig. 7 presents several examples of the internal quality assessment of the results. In the figure, the white lines are the VEC25 roads. The accepted results and the rejected results are shown in black and light grey lines respectively, the dark grey lines represent the results that have to be checked/verified by an operator.

Fig. 7. Internal quality assessment of the results. The accepted results, the rejected results and the results to be checked/verified are shown in black, light grey and dark grey lines respectively. (a)
accepted result, (b) result that has to be checked/verified, (c) accepted gap solution, (d) rejected gap solution.

The external evaluation is conducted by comparing the extracted results with precise reference data. The quality measures used in this work aim at assessing completeness and correctness as well as geometrical accuracy. Completeness measures the percentage of the reference data that lies within the buffer of the extracted roads, while correctness is the percentage of the extracted roads within the buffer of the reference data (Heipke et al., 1998). The buffer distance is defined using the required accuracy of the project ATOMI. The geometrical quality is assessed by the mean and RMS of the distances between the extracted roads and the reference data. Additional measures have been defined, e.g. to evaluate the shape quality, but have not been used in the test results presented here.

5. Results

The described system has been implemented as a stand-alone software package with a graphic user interface running on SGI platforms, while porting on Windows XP should be completed at the beginning of 2004. The system imports color stereo imagery, the existing road database and other input data, and outputs the extracted roads in 3-D Shapefile format that is readily imported by existing GIS software. Other data formats can be easily accommodated. The system has been tested using more than 20 models in various landscapes. Some reports of the system performance can be found in Zhang and Baltsavias (2002), Zhang (2003a) and Zhang (2003b). A benchmark test has been conducted independently by our project partner using new flight imagery, in the test site Thun, Switzerland. The terrain height ranges from 550 m to 2200 m. Almost all road types in Switzerland can be found in this area. The images were acquired in October 2001, and the image data have the same specifications as described in Section 1. During the test, our system is only applied to extract roads in rural areas, while roads in urban and forest areas are not processed. Fig. 8 presents a portion of 3-D road extraction and road network generation (only the left image is shown). The landscape of Fig. 8 includes open rural, forest areas and small settlements.

![Fig. 8. Extracted 3-D roads and road network (black lines) in the test site Thun, Switzerland.](image)

The details of automatic 3-D road extraction and junction generation in rural areas with varying complexity are presented in Fig. 9, where the VEC25 roads are shown in white lines and the extracted roads in black lines. Note that the road junctions are also well extracted and modeled. Recently, the system has been under extensive test by our project partner. Around 2,500 km roads have been processed. Almost all roads in rural areas were correctly extracted using visual check.
Table 1 summarises the external evaluation of the extraction results in Fig. 8 using the reference data measured by L+T at an analytical plotter. It can be seen that our system achieves very good results. The completeness and correctness are very high. The accuracy of the extracted road network is about 0.5 m both in planimetry and height, fulfilling the accuracy requirements of the project ATOMI.

<table>
<thead>
<tr>
<th>Quality measures</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Completeness</td>
<td>94.2%</td>
</tr>
<tr>
<td>Correctness</td>
<td>96.9%</td>
</tr>
<tr>
<td>Length of reference (km)</td>
<td>12.4</td>
</tr>
<tr>
<td>Length of extraction (km)</td>
<td>11.69</td>
</tr>
<tr>
<td>RMS (m)</td>
<td></td>
</tr>
<tr>
<td>dx</td>
<td>0.40</td>
</tr>
<tr>
<td>dy</td>
<td>0.33</td>
</tr>
<tr>
<td>dz</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Table 1. Quality measures for the benchmark test dataset in Fig. 8.

A higher completeness has been achieved by our partner when spring photography is used. This is because of less tree occlusions. Even some roads along forest borders can be extracted in spring images (see Fig. 10). In addition, some roads in the field, invisible in summer images, are visible in spring images and are extracted, thus further contributing to a higher completeness of the extracted road network.
Fig. 10. Roads along the forest border, invisible in the summer image, are visible in the spring image and are extracted. (a) summer image, (b) spring image with the extracted roads in black lines and outdated roads in white lines.

We also tested our system on unknown data provided by the National Geographic Institute, Belgium. Although the images are black and white with quite poor radiometric quality, and no DSM is available, the performance of our system is also quite good in flat open rural areas. By comparing the results with the manually measured reference data for ca. 13 km roads, the achieved completeness and correctness are 97.6% and 98.1% respectively (Zhang and Baltsavias, 2002).

Recently, the system has been modified to work even with single orthoimages, whereby the 3-D information is not extracted by image matching, but by overlaying the 2-D information on the DSM and DTM. Extensive tests conducted by our Swiss project partner using various resolution orthoimages (0.20 m ~ 0.50 m) have shown that the results are quite similar to that from the stereo imagery, and that an increase of the pixel size leads to a much smaller deterioration of the geometric accuracy of the extracted roads. With this development, the Dutch Ministry of Transport, Public Works and Water Management (MTPWWWM) awarded to us, after an evaluation of various research systems, a project for a feasibility study of semi-automated updating of the Dutch road database, using color orthoimages of 0.5 m pixel size from aerial images of 1:25,000 scale with 15 cm focal length. The study site is situated near the city of Weert in the province Limburg (in the south of the Netherlands), covering an area of 12 * 12 sq. km. The landuse changes gradually from open rural to urban, with the complexity of the scenes increasing correspondingly. The images were taken in June, 2000. The images do not have good quality; they are too green and noisy. In many cases the roads show very poor contrast with surroundings. The image edges are poorly defined; also color shifts between bands are observed. In addition, trees at roadsides usually occlude roads very much in these summer images; some roads are even totally occluded. We also observed that the roadmarks on roads are very weakly represented in such images. The old road databases are created by digitizing 1:10,000 topographic maps, with an RMS error of about 10 m. The database allows distinguishing national roads and a small part of the provincial roads in the Netherlands, and provides the number of lanes for them. The other roads are in a single class. There are no clues that can be used to infer the approximate road width. Raw height data are from laser scanning. Also the filtered height data are provided. Both datasets have points regularly distributed with a 5m x 5m spacing. During the test, our system is only applied in open rural areas. Fig. 11 shows a portion of the test results. The roads in rural areas are correctly and reliably extracted by our system. In Fig. 12, the details of road extraction and junction generation for this dataset are presented in several examples.
Fig. 11. Extracted 3-D roads and road network in the test site in the Netherlands superimposed on image as black lines.

Fig. 12. Details of road extraction and junction generation in the Netherlands dataset. The extracted roads are shown in black lines and the outdated roads in white lines.

Reference data for the Dutch dataset is not available at moment. The accuracy of the extraction result cannot be accessed. In each test image we computed the ratio of the length of the extracted roads to the length of rural roads in the existing database (the total length of the rural roads in the old database is ca. 500 km). The ratio values range from 80% to 92%, depending on the complexity of the scene. Generally, the performance is worse compared to the performance on the Swiss dataset. This is mainly caused by: (1) the poor image quality, (2) insufficiently road information in the existing road database,
especially the lack of the road classes or road width, (3) the images are taken in summer, many roads are occluded by trees, (4) the worse spatial resolution of 0.5m compared to 0.22m of the Swiss data.

6. Discussion and outlook

In this paper, we have presented a practical automated system for road extraction from stereo and ortho images focusing on rural areas. The roads should have a minimum width of about 3 pixels in order that edges on both road sides are extracted. The system has several advantages over other approaches. It uses existing knowledge, image context, rules and models to restrict the search space, treats each road subclass differently, checks the plausibility of multiple possible hypotheses, therefore provides reliable results. The system contains a set of data processing tools to extract various cues about road existence, and fuses multiple cues and existing information sources. This fusion provides not only complementary information, but also redundant one to account for errors and incomplete partial results. Working on stereo images, the system makes an early transition from 2-D image space to 3-D object space. Road hypotheses are generated directly in 3-D object space. This not only enables us to apply more geometric criteria to generate hypotheses, but also largely reduces the search space, and speeds up the process. The hypotheses are evaluated in images using accumulated knowledge information. Whenever 3-D features are incomplete or entirely missing, 2-D information from stereo images is used to infer the missing features. By incorporating multiple knowledge, the problematic areas caused by shadows, occlusions etc. can be often handled. Based on the extracted roads, the road junctions are generated and modeled, thus the system provides an up-to-date and complete road network for practical uses. We also present in this paper the results of road extraction in benchmark tests conducted independently by our project partner. The quantitative analysis using accurate reference data is also presented. The comparison of the reconstructed roads with such data shows that more than 94% of the roads in rural areas are correctly and reliably extracted by the developed system, and the achieved accuracy of the road centerlines is about 0.5 m both in planimetry and height, fulfilling the requirements of the project ATOMI. The system can also work with black and white images and no use of DSM without significant performance reduction (as the tests in Belgium have shown).

We also showed that a higher completeness can be achieved if spring photography is used. An even better performance can be expected with the increasingly possible use of better quality DTMs and DSMs (e.g. from airborne laser scanning) and a near-infrared channel, which is available in many new digital photogrammetric cameras and high-resolution satellites.

The system can also process orthoimages. The extensive tests at the L+T in rural areas have shown that the performance is quite similar. We have presented the results in the study site of the Netherlands, with the dataset provided by MTPWWM, the Netherlands. Using a visual check, most of the roads in rural areas are extracted. The extraction results are quite good and reliable. There are only very few errors in the extraction results. Better results can be expected by using good quality images, which can be achieved by acquiring better film imagery and by better scanning. Use of new digital cameras, e.g. ADS40 developed by Leica Geosystems, providing near infrared information would be an extra plus.

First tests with ADS40 imagery at the L+T have shown that the road extraction results are similar, however with larger pixel size and no use of the infrared channel. Use of digital photogrammetric cameras may prove to be advantageous for road extraction, on top of other advantages like faster data acquisition, avoidance of errors during scanning etc.

The roadmarks in the Dutch images are only very weakly represented. The poor image quality makes the situation even worse. This caused some short roadmarks like the road division marks to remain unextracted. Roads in the images are usually occluded very much by trees. Also roads along forest borders are invisible. This drastically decreases the completeness of the extraction results. We suggest using early spring or autumn instead of summer images for road extraction in future production. The day time of the flight should be also carefully selected such that shadows are minimised.

Regarding 3-D reconstruction of new roads not included in the database, this can be achieved by manual on-screen digitisation of characteristic road seed points and continuation with the same procedure described in this paper with minor modifications. This approach can be also followed when no road database exists, for 3-D road reconstruction from scratch.
Due to the good system performance, the project ATOMI has been continued with the project ATOMIR (R standing for roads) with aim the development of an operational system for the L+T until spring of 2004. Apart from operational aspects (porting to Windows XP, batch processing of large datasets, speed increase, RAM optimisation etc.), several methodological aspects will be investigated or refined, like improvement of self-diagnosis and reliability criteria, integration of manual intervention in difficult cases (e.g. complex junctions) and performance improvement for operational use with roads along forest borders and in urban areas with low buildings.

Acknowledgements

We acknowledge the financial support for this work and for the project ATOMI by the Swiss Federal Office of Topography, Bern. We also thank the Ministry of Transport, Public Works and Water Management, the Netherlands for letting us use their data for this publication. The kind cooperation with Mr. Col Offermans at the MTPWWM, the Netherlands is very much acknowledged.

References


