2009 Urban Remote Sensing Joint Event

Status and Further Prospects of Object Extraction from Image and Laser Data

(Invited Paper)

Franz Rottensteiner
Institute of Photogrammetry and GeoInformation
Leibniz Universität Hannover
Hannover, Germany
rottensteiner@ipi.uni-hannover.de

Abstract— **The automated extraction of topographic objects has been on the research agenda in the Photogrammetry and Computer Vision communities for more than two decades. Considerable progress has been achieved, though up to now there are hardly any commercial products that have been accepted by the market. Recent developments in the field of sensor technology, along with advanced techniques for data processing, have increased the potential of automated object extraction. This paper gives an overview on the status and further projects of automated object extraction, focusing on buildings and roads and on the application of high-resolution optical data.**

I. Introduction

The automated extraction of topographic objects such as buildings and roads from remotely sensed data has been an active field of research in photogrammetry and remote sensing for quite some time. Initially, the research focus was on the automatic classification of satellite images into relative broad classes such as ‘settlement’ and ‘farm land’ for applications such as topographic mapping at a regional scale. Since the late 1980s, progress achieved in sensor technology, especially the transition from analytical to digital photogrammetry and the advent of new sensors such as digital cameras or laser scanners, along with the application and adaptation of methodology originally developed in the computer vision community, have rendered possible the detection and 3-D reconstruction of individual objects at a relatively high level of detail. The demand for automated techniques for object extraction is driven by the urge to reduce the costs of primary data acquisition and continuous update of topographic data bases and 3-D city models, which have found wide-spread applications.

It is the goal of this paper to give an overview on the current state and future prospects for automatic extraction of buildings and roads from optical sensors. The paper will be restricted to methods that can detect and/or reconstruct individual objects, which implies that the sensor resolution has to be in the order of 1 m or better. The presentation will be structured by the type of object that is to be extracted. That is, Section II deals with building extraction techniques, whereas Section III is dedicated to the extraction of roads. A summary is given in Section IV.

II. Building Extraction

In this paper, the term *building extraction* is understood as the automated 3-D reconstruction of CAD models that represent buildings. In this context, it is important to consider the level of detail (LOD) that is to be achieved by the reconstructed CAD models. The CityGML standard, which was recently accepted by the Open Geospatial Consortium as a standard for the representation, storage and exchange of virtual 3-D city and landscape models [1], distinguishes five LODs:

1. LOD 0 – Regional model: A 2.5-D Digital Terrain Model (DTM) over which an image or a map is draped.
2. LOD 1 – Block model without roof structures: The buildings are modeled by vertical prisms with horizontal roof planes.
3. LOD 2 – Model with differentiated roof structures.
4. LOD 3 – Architectural model with detailed wall and roof structures, balconies, bays, etc.
5. LOD 4 – Architectural models including interior structure such as rooms and furniture.

Disregarding LOD 0 (which does not contain individual buildings), building extraction techniques based on aerial sensors can only deliver models corresponding to LOD 1 or LOD 2. If terrestrial sensors are used, LOD 3 can be achieved. LOD 4 requires primary data to be acquired in-door, which is not considered in this paper.

The 3-D reconstruction of buildings requires quite complex operations, so that it is reasonable to restrict the search space before these algorithms are applied. This is the reason why building extraction is usually carried out in two steps, e.g. [2]:

1. *Building Detection* requires the recognition of buildings in the sensor data. It is essentially a classification task and delivers regions of interest for building reconstruction. If building extraction aims at the generation of a 2-D building layer in a topographic data base or at 3-D models corresponding to LOD 1, building detection may include the precise location of the building

978-1-4244-3461-9/09/$25.00 ©2009 IEEE

outlines, from which prismatic models can be derived easily if information of the relative heights of the buildings relative to the ground is available.

2. *Building Reconstruction* is the generation of 3-D CAD models corresponding to LOD 2 or LOD 3 from sensor data. It is carried out based on previously detected regions of interest, which can be replaced by existing building outlines from a topographic data base [3] or by seed structures provided by a user [4].

In many industrialized countries, digital databases containing a building layer and / or 3-D city models do already exist. Keeping such a data base up-to-date has been estimated to require up to 40% of the costs of the initial data acquisition [5]. In such a situation it is not desirable having to acquire building data again from scratch. This is why there has been an additional focus on *change detection* for map updating [5], [6]. In the subsequent sections, the topics introduced here will be addressed in more detail.

A. Recognition of Buildings and Building Outline Detection

1) Data used for building detection: In the early work on building extraction, scanned aerial images were largely used for that purpose. The primary information content of such images is a color vector for each pixel. Scanned aerial color images were restricted to three components, either RGB or color infrared (CIR). CIR imagery was found to be better suited for classification due to its inherent potential for detecting vegetation [7]. The fact that color-based segmentation is influenced by illumination conditions and shadows [7], [8], [9], and because roofs may have a poor contrast with the background, alternatives to using only the color information were searched for at an early stage. Man-made objects are characterized by regular structures such as parallel or orthogonal straight lines. As these lines can be well extracted in digital images, extraction and grouping of straight lines according to some geometrical criteria has been used for detecting buildings [10]. It was also noted at an early stage that the 3-D information implicitly contained in aerial stereo images gives a very important cue for detecting the locations of buildings, because the building parts that are most consistently visible to aerial sensors (i.e., the roofs) are higher than the terrain surrounding a building. That is, Digital Surface Models (DSMs) generated from stereo imagery by image matching techniques soon became an important [7] if not the sole [11] cue for building detection. If a DTM is available, the height differences between the DSM and the DTM give direct access to the heights of objects above the terrain. A DSM also provides information about local surface properties via an analysis of the derivatives of the DSM: the first derivatives give access to the slope of potential roof surfaces and (via local maxima) to building outlines [12]. The second derivatives are more commonly used for building detection. As they are closely related to the local curvature of the DSM, they can be used to derive measures for surface roughness. Assuming that roofs mostly consist of planar or at least smooth surfaces, an analysis of surface roughness can help to separate buildings from trees. Various parameters have been used in the past to characterize surface roughness, e.g. the output of a Laplace filter applied to the DSM [12], [13], local curvature [13], or the local variance of the surface normal vectors [14].

The advent of Lidar brought about an improvement of the DSM quality compared to DSMs from matching techniques existing at that time. The main improvement is that by delivering 3-D points directly Lidar avoids the smoothing effects inherent to ‘traditional’ matching techniques. As a consequence, surface roughness can become more relevant for classification. Furthermore, Lidar can deliver points on the terrain in vegetated areas, which helps to generate better DTMs even in densely vegetated or built-up areas [15]. As an additional classification cue, the height differences between the first and the last pulse received by the Lidar sensor can be used to separate trees from buildings [16]. Finally, Lidar intensities, though relatively noisy compared to image data, can also be exploited [17]. Lidar promises a high degree of automation in building detection and has thus become a major data source for that purpose [2], [12], [13], [18], [19].

Comparing the two data sources used most frequently for automatic building detection, i.e. Lidar and aerial imagery, it has been noted that they have complementary properties with respect to the problem to be solved, which suggests that much can be gained from a fusion of these data sources [20], [21], [22]. Lidar directly delivers 3-D points, so that the correspondence problem needs not be solved, and it gives a more direct access to surface properties. Due to the smaller opening angles of Lidar sensors, occlusion not as big a problem as with aerial imagery. There are no cast shadows, because Lidar is an active sensor technique. Finally, its potential for penetrating vegetation helps in distinguishing trees from buildings. On the other hand, the spectral content of Lidar data is very limited, and it gives only a poor representation of abrupt height changes, which leads to a poor definition of building outlines in Lidar data. Edges can be extracted relatively well from images, so that a better representation of building outlines is to be expected in image data. Finally, aerial imagery usually has a higher resolution, potentially in the order of a few centimeters, than Lidar, which is usually captured with an average point distance in the order of 0.5 m - 1 m. It has been shown that the fusion of Lidar data and multi-spectral information from aerial images can help to improve the classification accuracy for buildings smaller than 100 m^2 by 10%-15% [22]. Recent developments in sensor technology, namely the development of digital aerial cameras and of full-waveform Lidar systems, have improved the prospects of building detection from airborne sensor data considerably, which will be discussed below.

In addition to aerial images, high-resolution satellite imagery have become available as a source for mapping applications. However, up to now the resolution of these images has not been good enough for fully automatic building detection. Typically, semi-automatic approaches are favored [23]. Fully automatic methods seem to be capable of delivering large building structures [24], but suffer from problems related to the DSM quality [25]. This situation may be improved by new satellites having a resolution better than 0.5 m such as GeoEye-1.

978-1-4244-3461-9/09/$25.00 ©2009 IEEE

2) Classification techniques for building detection: A great variety of algorithms for building detection have been developed, making use of the data sources and classification cues described above. Apart from algorithms that are based on extracting and grouping image edges [10], building detection is usually carried out in two steps, namely an initial segmentation and a classification of the initial segments.

If single images are used as input data, the initial segmentation can be based on color [9] or on adaptive multi-scale segmenters [8]. Unsupervised classification techniques, e.g. the ISODATA [21] or the K-means algorithm [26] have been used for segmentation, too. From a DSM, segments can be generated in by detecting blobs [7], by applying a height threshold to the height difference between a DSM and an approximate DTM generated by morphologic filtering, [11] or by applying hydrological tools of a raster GIS [27]. Such approaches have problems separating buildings from trees that are very close to the buildings. In order to overcome these problems, many techniques apply classifiers using local features to obtain the initial segmentation. In this case, a great variety of classification techniques known from remote sensing and pattern recognition have been applied. These classification algorithms can be categorized into two main groups, namely model-based algorithms and probabilistic algorithms. Model-based algorithms often apply rules derived from knowledge about buildings such as minimum height, minimum/maximum roof slope and curvature in order to successively eliminate non-building areas by applying thresholds to the respective features [2], [11]. In order to obtain a solution that is not so directly dependent on the selection of thresholds, fuzzy rules have also been applied. In this case, the fuzzy membership functions are designed to reflect model knowledge about buildings [13], [18]. Among the probabilistic classification methods used for building detection are maximum likelihood classification [12], Bayesian classification [14], and classification based on the Dempster-Shafer theory of evidence [22], [26]. In this context, the statistical properties of the features can be obtained by training [12] or from model knowledge about buildings [14], [22]. Techniques from machine learning such as Support Vector Machines [28] also require training areas for classification. Training implies manual intervention, which restricts the degree of automation of an approach. On the other hand, taking decision based on training data may be more robust than selecting thresholds in an arbitrary way, especially for features related to surface roughness, where the actual feature values do not have an intuitive interpretation.

The results of the initial segmentation are usually post-processed in order to remove small errors, e.g. by morphological operators. After that, a classification of these segments is carried out, using similar features as for the initial segmentation. Additionally, shape features such as minimum size or ‘roundness’ can be applied. The techniques used for segment classification are essentially the same as those described previously in the context of initial classification. If single images are used, the lack of 3-D information makes it difficult to distinguish some roof planes from roads, because they may have similar radiometric properties. In this case, context can be used to support the classification. In aerial images, shadows reflect the third dimension of buildings and are, thus, often used for classification [8], [9]. Roads, which usually connect buildings or are aligned with them, can also be used as additional context objects [9].

The quality of the extraction results that can be achieved from Lidar and image data having a resolution as described above (i.e., DSMs with a resolution of 0.5 m – 1 m, aerial imagery with a resolution of about 0.1 m) has been well studied, e.g. [5], [13], [22]. Using these data, buildings having an area larger than about 100 m^2 can be extracted very reliably, with completeness and correctness rates [29] of 90% or better. The classification results deteriorate from that point; buildings smaller than 30 m^2 are hardly detectable from such data. A minimum resolution of 1.5 m has been found to be necessary for a successful detection of buildings [22].

3) Recent developments and future prospects: What has been said so far represents the current status of building detection from sensor data that are acquired using technology currently in use in many private companies and government agencies. However, the prospects for automated object extraction have been improved by recent developments in sensor technology, in combination with the development of new tools for data processing.

Firstly, aerial film cameras are being replaced by digital cameras [30], so that the bottleneck caused by the scanning of films that hampered the productivity of digital photogrammetry is about to disappear. These new cameras can deliver multi-spectral images at a dynamic range of 11 bit. The advent of digital aerial cameras has made the acquisition of multiple-overlap aerial imagery economically feasible, which has had a considerable impact on the prospects of automated object extraction from aerial image data in urban areas: the traditional problems of aerial imagery, namely occlusions and different object appearance due to perspective distortions, become less problematic as baselines become shorter and as there is also a 60% side lap of aerial photography. In addition, multiple views add redundancy to the object extraction task, which helps to make automated procedures more robust [28].

The impact of the new digital aerial cameras is further increased by the development of improved image matching techniques in the computer vision community, e.g. [31]. These methods apply global or semi-global matching techniques that can provide DSMs without smoothed object edges, at a resolution identical to the image resolution. If multiple overlap images are used, these DSMs can be generated nearly without occluded areas, at a height accuracy in the order of magnitude of the image resolution [32].

In the field of Lidar, the progress in sensor technology has rendered possible the acquisition of very dense point clouds along with the full waveform of the reflected signal [33]. Features extracted from the full waveform data can support the discrimination of buildings and trees [34], whereas using very high resolution data, e.g. acquired from helicopters [35], can help to improve the classification accuracy.

Thus, using new developments in the field of sensor technology, it should be feasible to detect small buildings reliably, which can help to make fully automatic building detection operational for areas such as the updating of cadastral

978-1-4244-3461-9/09/$25.00 ©2009 IEEE

maps. However, this cannot simply be achieved by just applying the 'old' algorithms to better data. With the new sensors, it becomes more and more important to consider the 3-D structure of the scene and the imaging process in the evaluation. For instance, in order to overcome the problem of occlusions not only in DSM generation, classification can be applied independently to each of the multiple-overlap images, and these individual classification results can be merged in object space together with additional cues derived from the DSM [28]. In the context of Lidar, very dense point clouds are often generated by flying over the area of interest several times. In this case, a 2.5-D perspective on data processing will no longer suffice. For instance, in multiple-overlap Lidar data, the sensor will collect points both on and underneath roof overhangs. In a 2.5-D processing environment this creates the impression of a high surface roughness and will affect the classification in a negative way. On the other hand, a strictly 3-D view on processing may even be able to detect walls both in image [36] and Lidar data [35]. In this way, an old problem of building detection using aerial sensors may become solvable: in cadastral data bases, building outlines are usually defined as the intersections of the walls with the ground, which has been unobservable in a systematic way up to now.

The fusion of data from different sensors still remains an important research topic, as it may especially increase the degree of automation of automated object detection. Whether this makes sense from an economical perspective still needs to be investigated: If the increase in productivity outweighs the additional costs for primary data acquisition, data fusion will be successful. The new developments in sensor technology have mitigated some of the problems of both aerial images and Lidar data, but very high resolution Lidar data remain more expensive than image data of the same resolution while still having advantages in scenes containing trees.

From the point of view of data processing, it may be worth while to find the transition from focusing on the detection of one object class only to more complex scene models that also consider the mutual interactions of objects of a different class. Even though context has been considered for a while in object detection [8], [9], new techniques such as Conditional Random Fields (CRF) [37] have become valuable tools for modeling context in a statistically sound way. In [37], buildings are detected in arbitrary terrestrial images with a relatively high success rate given the relatively 'complicated' appearance of the buildings in these images. Combining the domain-specific knowledge of the photogrammetric community with these new and statistically sound tools could be a good way to overcome the problems of existing object detection techniques.

Apart from using new data and developing new algorithms for building detection, research also has to investigate the quality of the results of such algorithms. There has not yet been a comparison of building extraction results using data from the new sensors, so that their full potential remains to be evaluated.

B. 3-D Reconstruction of Buildings

1) Data used for building reconstruction: Essentially, the same data sources that are used for building detection are also used for building reconstruction. In this context it has to be noted that algorithms that are based on images tend to favor processing based on the extraction, 3-D reconstruction, and grouping of edges, whereas algorithms based on Lidar data aim at a reconstruction of buildings by extracting and grouping planar patches corresponding to roof planes. In order to obtain models corresponding to LOD 3, terrestrial data sources have to be incorporated into the analysis, e.g. terrestrial laser scanner (TLS) data [38] or terrestrial imagery [39], [40].

2) Techniques for building reconstruction: Two different strategies have been applied for generating CAD models of buildings:

1. Top-down strategy: The ground plans of the building outlines are segmented into rectangles, and parametric models representing common building shapes such as saddleback roof or hip-roof buildings are fitted to the data, whereby the model achieving the best fit is accepted [12], [41].
2. Bottom-up strategy: The model is successively assembled from evidence found in the data, following some generic rules, e.g. assuming that the reconstructed building has to be a polyhedron [42], [43], [44], [45], [46].

Fitting primitives to data involves matching the wire-frame of the primitive to edges extracted from the digital images [47]. Since the results depend on the rectangular segmentation of the ground plans, such techniques are preferably applied if very good ground plans, e.g. from existing maps, are available [41], [47]. In [48], two building reconstruction techniques are compared, one using a bottom-up strategy and the other using primitives. The authors conclude that by using primitives, a more regular appearance of the models is achieved. However, it has to be noted that using primitives of a rectangular footprint restricts an algorithm to reconstructing rectangular structures, which may result in an over-regularization in areas that are characterized by more irregular building shapes [49]. It would seem that generic building models are more flexible with respect to modeling complex shapes.

Building reconstruction from images has been carried out by matching image edges and assembling polyhedral models from the resulting 3-D edges [4], [45]. In this context, it is important not only to have images of two-fold stereo overlap, because this would result in a poor geometric quality of 3-D edges that are nearly parallel to the baseline. In [4], the grouping of the 3-D edges is based on Bayesian model selection approach. Such probabilistic models are not frequently used for grouping due to the problems involved in learning the prior distributions. An elegant method to obtain planar hypotheses from 3-D edges is plane sweeping, where a plane is allowed to sweep around the 3-D edge and a correlation score is used to determine the actual tilt of the plane [45]. Another way of obtaining plane hypotheses from images is based on color segmentation in image space. The segmentation results are projected to a DSM, and image edges are matched with the original segment edges to obtain the delineating polygons of the roof planes [44].

Building detection from Lidar data usually starts with a planar segmentation of the point cloud or the DSM generated from the point cloud. The algorithms involved include region growing based on seed regions that are detected by an analysis of the local curvature of the DSM [2], clustering [3], techniques based on the random sample consensus (RANSAC) [41], and the application of the Hough transformation [48]. For grouping these planes, decisions about the local configuration of the planes have to be taken. For instance, a decision has to be taken whether two adjacent roof planes intersect, whether there is a height step between them, or whether they are actually co-planar. This problem has been solved by computing all possible intersections between planes and only keeping those that are actually possible according to the geometric configuration [2] or by a local analysis involving the comparison of geometric entities such as distances and angles to thresholds [43]. In order to make these decisions more robust, the decisions can be taken based on statistical tests, which requires the rigorous modeling of the uncertainty of the geometrical entities involved [46].

In [50], various techniques for building reconstruction from Lidar and image data were compared. The report, based on an international test carried out by EuroSDR (European Spatial Data Research – www.eurosdr.net), concluded that techniques based on Lidar data in general had a higher degree of automation and also received a better accuracy in height. On the other hand, the planimetric accuracy of the resulting 3-D models, especially at the building outlines, was found to be better for techniques based on aerial imagery. Again, due to the complementary properties of these data sources, the fusion of aerial imagery and Lidar data has been proposed, either to improve planar segmentation [51], or to improve the geometrical quality of the building outlines [52].

3) Recent developments and future prospects: All the approaches described so far aim at a reconstruction of buildings at LOD 2. None of them has matured into a commercial product. In fact, the only commercial products that appeared on the market so far are semi-automatic [53]. Actually, one of only two such products mentioned in the survey paper [53] has already disappeared again. Even though some has certainly been achieved and despite the fact that it is certainly possible to reconstruct the main building structures automatically, there remain several reasons why there are no commercial systems for fully automatic building reconstruction:

1. Many approaches are not flexible enough to handle a large variety of building shapes. This especially applies to reconstruction techniques relying on parametric primitives.
2. It is very difficult to develop algorithms that are robust with respect to the selection of processing parameters. Many approaches apply ad-hoc methods for taking decisions in the reconstruction process.
3. Due to restrictions of the sensor resolution and/or problems such as lack of contrast, shadows, or occlusions, highly complex roof structures consisting of many small parts do not only result in over-generalized models, but these structures might actually cause algorithms to fail completely.
4. There is a lack of integration into digital workstations and, more importantly, a lack of efficient tools for post-processing in cases where the automatic procedure fails.

Again, what has been said so far applies to methods based on data from of sensors that are currently in practical use. The changes in sensor technology described above have also had a significant impact on the prospects of automated building reconstruction. As a matter of fact, a dense DSM of a building can already be viewed as a 3-D building model. By applying thinning algorithms to such a DSM, a 3-D representation without too much redundancy can be achieved, even though it does not correspond to the kind of model one would traditionally use in a CAD system. Microsoft claim to be able to produce 3-D building models fully automatically [32], though the required amount of manual intervention and the geometrical accuracy of the resulting models are not known. In this sense, the 3-D reconstruction of buildings corresponding to LOD 2 for visualization purposes can be seen as nearly solved, with future work concentrating on the improvement of the degree of automation and an improvement on the LOD that can be achieved. Whether or not this can be achieved in an economical way by the fusion of ALS and image data remains to be investigated.

In [54], a technique for building reconstruction based on existing maps and a dense DSM generated from imagery is described. Hypotheses for roof planes are generated based on the building outline, and a probabilistic approach is applied to select an appropriate model that fits the data well yet is not too complicated. A second technique is based on RANSAC for detecting planes. Using either technique, about 85%-89% of the buildings of a large test site in France could be reconstructed. The authors claim that by combining these two techniques in a semi-automatic environment, the methods could become operational, delivering 95% of correct buildings.

Using Lidar data of a higher resolution also has increased the prospects for building reconstruction. As noted above, a fully 3-D approach is required to do so. In [35], clustering techniques are used to detect planes in very high-density Lidar point clouds (20 points / m^2), which can be used to reconstruct very detailed building models, including a correct representation of the roof overhangs. Thus, structures such as large dormers that previously could cause an algorithm to fail entirely can now be correctly reconstructed. This is also the case for DSMs generated from images [55]. Based on a generalized model, rectangular regions of ‘outliers’, i.e. DSM points not fitting to the original model, are considered to be hypotheses for superstructures. The system provides parametric models for common types of superstructures, whose parameters are estimated from the DSM. Finally, in the case of competing hypotheses, the set of models achieving the overall best fit is accepted [55].

Thus, it is clear that by making use of the new developments in sensor technology and DSM generation, it is possible to increase the LOD of the resulting models. However, in order to really achieve LOD 4, that is, complete architectural models, the façades of the buildings must be refined. This requires the detection and 3-D reconstruction of structures such as windows, doors, and balconies, which has been investigated using TLS and/or image data [38], [39], [40].

In this context, new methodologies are investigated. In [56], grammars that describe rules for the combination of primitives to obtain complete building models are combined with minimum description length as evaluation function to control and guide the search process. Façades are often characterized by a very systematic structure, so that grammars are also used to reconstruct façades, not only in order to detect each individual façade object, but also to understand the underlying structure, e.g. symmetries or structures such as rows of windows [39], [57]. Techniques from pattern recognition and machine learning such as Markov Chain Monte Carlo [39] that can be used to obtain a sample for a distribution based on an evaluation function or Conditional Random Fields that can model context relations between different object classes [37] are used for façade interpretation. At the University of Bonn, a test data set is being made available that consists of a large set of fully labeled terrestrial images of façades. This test data set is to be used for benchmarking of techniques for structural learning [58].

For applications such as navigation, it may be helpful to obtain a 3-D model of the visual part of a street. There are efficient building reconstruction algorithms that generate coarse façade models from video cameras mounted on a car in real-time [57]. Since façades are often obstructed by cars in images taken at street level, the visual appearance of the models can be improved considerably if cars are automatically detected and removed in the resultant models [57]. This is yet another example showing how much can be gained by a more complex scene analysis taking into account multiple object types and their mutual interactions.

C. Change Detection

The automation of change detection requires a classification of existing objects in a data base according to whether they have remained unchanged or not, and the detection of new objects. This can be achieved separately, even using different techniques for the classification of existing objects and for the detection of new objects, or it can be done simultaneously, e.g. by first applying any of the building detection techniques outlined in Section II-A and then comparing the detection results to the existing data base. There has been a EuroSDR test project on automatic change detection for the updating of 2-D maps such as the cadastre [5]. The first data set consisted of Lidar data and a digital CIR orthophoto at a resolution of 1 m. The second data set consisted of a high-resolution multi-spectral orthophoto and a DSM generated by dense matching at a resolution of 0.2 m, and the third data set consisted of a multi-spectral orthophoto and a DSM generated from high-resolution satellite imagery (0.5 m). Four different techniques for change detection were compared in this test. The test shows that there are relatively small differences between the results from different algorithms, but that the success of change detection is mainly affected by the quality of the data used for that purpose. The results are in no way satisfactory for the satellite data set, due to the relatively poor quality of the DSM. The situation is better with the image-based data set, though all techniques produced too many false positives for changed and new buildings to be used in a fully automatic context [5]. In [59] it was shown that by integrating automatic procedures into a semi-automatic work flow where the human operator only needs to inspect buildings that are flagged as potential changes by the system, 40% of the work required for map updating can be changed if very small building structures are to be considered. This number increases to 60% if the data base only considers buildings larger than 50 m^2. These limitations arise from the problems related to the detection of very small buildings in Lidar data having a resolution in the order of 1 m. Currently, automatic tools for map updating are not operational, but by applying the strategies for improving the performance of building detection outlined in Section II-A, one can hope to build such systems in the future.

With 3-D city models becoming available for many cities of the world, the issue of change detection for keeping such 3-D models up-to date will arise sooner rather than later. However, up to now there is hardly any work on the topic of change detection for 3-D city models. It remains to be investigated whether more efficient techniques can be applied to check whether a building has changed both in planimetry and height; the alternative would be to recreate the whole 3-D city model from scratch every time new primary data are collected, which would seem to be very unsatisfactory from an economic point of view.

III. Road Extraction

Roads have a very specific appearance in remotely sensed data, and this appearance is highly dependent on the resolution of the images that are used for road extraction, e.g. [60]:

1. In data having a resolution of about 1 m, roads appear as thin dark or bright lines, which can be extracted by line detection techniques such as the Steger operator [61].
2. In data having a higher resolution, e.g. 0.1 m, roads are areas of relatively homogeneous reflectance properties, which are, however, influenced by objects such as cars, shadows, and road markings [60], [62].

Accordingly, different algorithms are used depending on the sensor resolution. Line-based algorithms and area-based algorithms can be combined in a hierarchical process [60]. In any case, it shall already be noted here that model-based techniques for road extraction make use of spectral, geometric, topologic, and contextual properties of roads [63]. The road topology is especially important: it is the function of roads to connect human settlements, so that roads form a continuous network. These network characteristics can be used to improve a road network in case where the initial road detections are incomplete [64]. In many works on road extraction, context has been considered in terms of global context, i.e. a global information about whether a road is situated in a rural, forested, or settlement area [60], [64], and of local context, which explicitly models other object classes that might cause problems for road detection, e.g. cars, trees, buildings, and shadows [60]. The reason why context has been used more frequently in road extraction techniques than in building extraction may be that roads are usually among the lowest objects in a scene, so that their appearance in airborne sensor data is more frequently disturbed by higher objects.

As stated above, in images of a resolution of about 1 m, roads appear as thin bands. Reconstruction techniques based on lines deliver road centre lines and road edges. Disturbances such as cars or road markings have no or only limited influence on line detection algorithms such as the Steger operator [61] or the method described in [65]. It is a common approach to first extract lines and then to select promising candidates based on geometric criteria such as width, length and curvature [66] and / or spectral criteria such as the normalized difference vegetation index [65]. Finally, a graph-based evaluation is applied to complete the road network [64], [66].

Road extraction algorithms that consider roads as areas more frequently consider the radiometric appearance of these areas in the data. There are area-based approaches that work with images at a resolution of about 1 m, usually based on multi-spectral and / or textural classification. In [67], several such cues are combined by the Dempster-Shafer theory of evidence, and initial road segments are generated based on a skeleton. Finally, a graph-based analysis involving two different scales is used to obtain roads. In [68], K-means clustering is applied to segment the image. Fuzzy classification is used to determine which of the clusters in feature space corresponds to roads, and a second fuzzy classification is applied to remove large areas that belong to the 'road' cluster, but actually correspond to buildings and parking lots. Finally, road segments are found from the refined clusters via a localized iterative Radon transform

It is a common feature of all these methods that hypotheses for road crossings are generated when the road network is completed. It has to be noted that, line detection algorithms have problems in areas where the model of a thin line is not fulfilled and in areas where other objects interfere with the roads, which is both the case with road crossings. It is thus no surprise that the extraction of road crossings is a problem in its own right that has been tackled by various means, e.g. by a topological analysis that aims at removing cycles in the local network sub-graph [69] or by neural networks [70]. The network properties of roads are also considered in the final evaluation.

Road extraction techniques are often applied for the updating of existing data bases or for quality control of these data bases [62], [64]. In this case, the roads in the existing data base give approximate values for the road position, which improves the execution speed and avoids false positive detections in areas that are far away from these roads.

It is worth noting that in contrast to building extraction, high-resolution satellite images play an important role in automatic road extraction. In an international test, again conducted by EuroSDR, several methods for road extraction from Ikonos (1 m) and aerial images (0.5 m) were compared [71]. The results suggest that current road extraction techniques deliver good results for rural scenes of low to medium complexity. In settlement areas the results were relatively poor. This is caused by the fact that the model assumptions of these algorithms are often hurt by interfering objects such as buildings in the direct neighborhood of roads. Furthermore, in settlements there are many road crossings, which, as we have seen above, also cause problems for road extraction techniques. Ikonos images only deliver multi-spectral information at a resolution of 4 m. Some of these problems may be overcome by satellite imagery providing a resolution of about 1 m also for the multi-spectral bands. Similarly as with building extraction, the problems encountered in road extraction can be tackled by incorporating 3-D information from a DSM and by using more rigorous models of context objects than the ones used in common road detection approaches.

The benefits of using 3-D data have been shown by approaches making use of Lidar data for extracting roads [72], [73], sometimes in combination with image data [74]. Such techniques rely on high-quality DTMs that can be generated from Lidar data both in forested and in densely built-up areas [15]. Road extraction makes use of the fact that roads have to be situated on the terrain. Furthermore, roads have very characteristic reflectance properties in the wavelength of Lidar systems: they appear as dark bands. Using these two cues, along with some local density parameters, it is possible to generate a binary road image. Complex convolution with a phase-coded disk can be used to determine the road centerlines and the road widths. The results presented in [72], achieved for Lidar data of about 1 m resolution in a suburban area, show that incorporating the third dimension does help in road extraction. It can even be used to detect bridges [75], [76]. Problems also occur at road crossings, especially with roundabouts and road crossings involving a change of the road widths. Furthermore, parking lots could not be separated from roads in [72]. It remains to be investigated whether DSMs from image matching can be used in a similar way to improve the performance of road extraction techniques.

Road extraction algorithms working with images of a higher resolution (typically, 0.1-0.2 m) have to consider the fact that many additional objects are visible on a road surface, e.g. road markings. Thus, highly complex road models that consider the appearance of roads at different scales, objects on the road surface, and context objects have been elaborated. In [60], the model knowledge is represented by a semantic net. More recently there have been attempts to automatically predict the scale behavior of road objects depending on the resolution. Starting from a semantic net describing the road model at a fine resolution, the road model corresponding to a coarser resolution is automatically derived by an analysis of scale-space events [77].

Algorithms for road extraction at a high resolution also use line detection algorithms, but in order to obtain road markings such as the solid and broken lines that separate individual lanes. If no such evidence is found, grey level edges that may correspond to road edges are used. From these data, hypotheses for road lanes are extracted in multiple images [60], and these hypotheses are merged in object space. Problems due to occlusions or poor contrast in one image may be overcome by information contained in other images. 3-D information in the form of a DSM generated by image matching is used to

978-1-4244-3461-9/09/$25.00 ©2009 IEEE

exclude building areas from processing both in [60] and in [62]. The radiometric properties can be used for the evaluation of road lane hypotheses [60] or directly for the extraction of regions of interest [62]. Apart from linear road marks, zebra crossings also give cues for roads [62]. Finally, context is used to complete the road network. For instance, shadow areas are extracted. They can support the hypothesis of a road by giving an explanation for the failure of the initial extraction algorithm [60], [62], [78]. Trees, cars [78], or rows of cars [60] can also support road extraction. However, it has to be noted that context objects are often used in an ad-hoc manner. Statistical models for context objects could help to improve the situation.

The algorithms aiming at a reconstruction of roads at a very high resolution have been successful in improving the geometric quality of roads from an existing data base [62]. In an urban context, the main roads could be successfully extracted in [60], with problems remaining with small suburban roads where hardly any road markings can be found, and, both in [60] and [62], with road crossings, which are not explicitly modeled in either case.

The problem of extracting suburban roads has recently been tackled in [78]. Normalized cuts are used for a segmentation of high-resolution CIR orthophotos, initial segments are grouped, and road candidates are selected based on geometric criteria. Context objects are used to complete the suburban road network. A new technique for generating highly detailed models of road crossings based on snakes, including the detection of traffic islands and the reconstruction of roundabouts, is described in [79]. The approach does not make use of context, which would be necessary especially in scenes with many cars waiting in front of a road crossing.

It would seem that three strategies can be applied to improve the prospects of automated road extraction beyond what has been achieved so far. They are similar to what has been suggested in the context of building extraction. Firstly, the integration of 3-D information into the classification must be achieved in a more systematic way. Promising results have been shown for Lidar, but it may also be supported by high-resolution DSMs generated by the new image matching techniques. Secondly, the role of context has to be emphasized, and context needs to be treated in a statistically rigorous way. Combining the model-based techniques developed in the photogrammetric community with techniques from pattern recognition may be a way to go. Thirdly, more emphasis has to be laid on specific problem areas, especially on detailed models of road crossings. The two strategies previously mentioned play an important role here, too, because on the one hand, traffic congestion in urban areas might especially affect road crossings, and on the other hand, some situations of crossing roads, e.g. crossing motor ways, can only be resolved if 3-D information is considered.

IV. Summary

This paper has given an overview on the current state and some future prospects of object extraction, with a focus on buildings and roads. It was shown that recent developments in the field of sensor technology have considerably improved the prospects of automated object extraction, and so have algorithmic developments in the computer vision and pattern recognition communities. The main trends are:

1. Usage of more and / or better data: This adds redundancy to the object extraction task. With new digital cameras, it is economically feasible to generate multi-view images of the same object. New matching techniques have rendered possible the generation of high-resolution DSMs in densely built-up areas. Lidar data are available at higher resolutions, and the full waveform of the reflected pulse provides valuable information on the structure of the underlying object. Data fusion is another option to be considered.
2. Systematic integration of 3-D information: The rich information content of new data sources supports the transition from 2-D or 2.5-D approaches, still common in building detection, to a really 3-D interpretation of a scene.
3. Adoption of recent developments in the fields of computer vision and pattern recognition, especially with respect to statistical modeling of the data and/or of context objects.
4. Development of more sophisticated scene models that do not only consider one, but multiple classes of interest along with their mutual interactions.
5. Increasing the level of detail of the models that are to be generated. Too coarse models might not simply be over-generalized, but small structures that are not modeled may actually hinder the success of object reconstruction.
6. Provision of data sets for benchmarking: benchmarking is important to make different object extraction techniques more comparable.

Considering these trends may be the key to success in automated object extraction from remotely sensed data.

References

[1] OpenGIS® City Geography Markup Language (CityGML) encoding standard. http://www.opengeospatial.org/standards/citygml (accessed 10 March 2009).

[2] F. Rottensteiner, "Automatic generation of high-quality building models from LIDAR data," *IEEE Computer Graphics & Applications*, vol. 23 (6), pp. 42–51, June 2003.

[3] M. Peternell and T. Steiner, "Reconstruction of piecewise planar objects from point clouds," *Computer-Aided Design* vol. 36, pp. 333–342, 2004.

[4] S. Scholze, T. Moons, and L. Van Gool, "A probabilistic approach to roof extraction and reconstruction," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXIV–3B, pp. 231–236, 2002.

[5] N. Champion, L. Matikainen, X. Liang, J. Hyyppä, and F. Rottensteiner, "A test of 2D building change detection methods: comparison, evaluation and perspectives," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVII–B4, pp. 297–302, 2008.

[6] G. Vosselman, B. Gorte, and G. Sithole, "Change detection for updating medium scale maps using laser altimetry," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXV-B3, pp. 207–212, 2004.

[7] O. Henricsson, F. Bignone, W. Willuhn, F. Ade, O. Kübler, E. Baltsavias, S. Mason, and A. Grün, "Project AMOBE: Strategies,

current status and future work," *Int. Arch. Photogrammetry, Remote Sens.*, vol. XXXI-B3, pp. 321–330, 1996.

[8] A. Stassopoulou, T. Caelli, and R. Ramirez, "Automatic extraction of building statistics from digital orthophotos," *International J. Geographical Information Sc.*, vol. 14 (8), pp. 795–814, 2000.

[9] S. Müller, and D. W. Zaum, "Robust building detection in aerial images," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVI-3/W24, pp. 143-148, 2005.

[10] C. Lin and R. Nevatia, "Building Detection from a monocular image," in *RADIUS: Image Understanding for Imagery Intelligence*, O. Firschein and T. M. Strat, Eds, California: Morgan Kaufmann Publishers, 1997, pp. 153–170.

[11] U. Weidner and W. Förstner, "Towards automatic building reconstruction from high resolution digital elevation models," *ISPRS J. Photogrammetry Remote Sens.*, vol. 50 (4), pp. 38–49, 1995.

[12] H.-G. Maas, "Fast determination of parametric house models from dense airborne laserscanner data," *Int. Arch. Photogrammetry, Remote Sens.*, vol. XXXII-2W1, pp. 1–6, 1999.

[13] T. Vögtle and E. Steinle, "On the quality of object classification and automated building modeling based on laserscanning data," *Int. Arch. Photogrammetry, Remote Sens.*, vol. XXXIV-3W13 pp. 149–155, 2003.

[14] A. Brunn and U. Weidner, "Extracting buildings from digital surface models," *Int. Arch. Photogrammetry, Remote Sens.*, vol. XXXII/3-4W2, pp. 27–34, 1997.

[15] K. Kraus, and N. Pfeifer, "Determination of terrain models in wooded areas with aerial laser scanner data," *ISPRS J. Photogrammetry Remote Sens.*, vol. 53 (4), pp. 193–203, 1998.

[16] K. Kraus, "Principles of airborne laser scanning," *J. Swedish Soc. Photogrammetry & Remote Sens.*, vol. 1, pp. 53–56, 2002.

[17] G. Vosselman, "On the estimation of planimetric offsets in laser altimetry data," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXIV-3A, pp. 375–380, 2002.

[18] L. Matikainen, J. Hyyppä, and H. Hyyppä, "Automatic detection of buildings from laser scanner data for map updating," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXIV-3W13, pp. 218–224, 2003.

[19] A. Sampath and J. Shan, "Building boundary tracing and regularization from airborne lidar point clouds," *Photogrammetric Eng. Remote Sens.*, vol. 73 (7), pp. 805-811, 2007.

[20] T. Schenk and B. Csatho, "Fusion of lidar data and aerial imagery for a more complete surface description," *Int. Arch. Photogrammetry, Remote Sens. Spatial Information Sc.*, vol. XXXIV-3A, pp. 310–317, 2002.

[21] N. Haala and C. Brenner, "Extraction of buildings and trees in urban environments," *ISPRS J. Photogrammetry Remote Sens.*, vol 54, pp. 130–137, 1999.

[22] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik, "Building detection by fusion of airborne laserscanner data and multi-spectral images: Performance evaluation and sensitivity analysis," *ISPRS J. Photogrammetry Remote Sens.*, vol. 62 (2), pp. 135–149, 2007.

[23] S. D. Mayunga, Y. Zhang, and D.J. Coleman, "Semi-automatic building extraction utilizing Quickbird imagery," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVI-3/W24, pp. 131-136, 2005.

[24] T. Krauß, P. Reinartz, U. Stilla, "Extracting orthogonal building objects in urban areas from high resolution stereo satellite image pairs," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVI-3/W49B, pp. 77-82, 2007.

[25] D. S. Lee, J. Shan, and J. S. Bethel, "Class-guided building extraction from Ikonos imagery," *Photogrammetric Eng. Remote Sens.*, vol. 69 (2), pp.143-150, 2003.

[26] Y. Lu, J. Trinder, and K. Kubik, "Automatic building detection using the Dempster-Shafer algorithm." *Photogrammetric Eng. Remote Sens.*, vol. 72 (4), pp. 395–404, 2006.

[27] M. Rutzinger, B. Höfle, T. Geist, and J. Stötter, "Object-based building detection based on airborne laser scanning data within GRASS GIS environment," *Proc. UDMS 2006 25th Urban Data Management Symposium*,Aalborg, Denmark, 12 p. (on CD-ROM), 2006.

[28] B. Gruber-Geymayer, A. Klaus, and K. Karner, "Data fusion for classification and object extraction", *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVI-3/W24, pp. 125-131, 2005.

[29] C. Heipke, H. Mayer, C.Wiedemann, and O. Jamet, "Evaluation of automatic road extraction," *Int. Arch. Photogrammetry, Remote Sens.*, vol. XXXII- 3/4W2, pp. 151-160, 1997.

[30] R. Passini and K. Jacobsen, "Accuracy analysis of large size digital aerial cameras," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVII–B1, pp. 507–514, 2008.

[31] Hirschmüller, H., "Stereo processing by semiglobal matching and mutual information", *IEEE T-PAMI*, vol. 30 (2), pp. 328-341, Feb. 2008.

[32] M. Gruber and S. Schneider, "Digital surface models from Ultracam-X images," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVI-3/W49B, pp. 47-52, 2007.

[33] W. Wagner, A. Ullrich, V. Ducic, T. Melzer, and N. Studnicka, "Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner," *ISPRS J. Photogrammetry Remote Sens.*, vol. 60 (2), pp. 100-112, 2006.

[34] C. Mallet, F. Bretar, and U. Sörgel, "Analysis of full-waveform lidar data for classification of urban areas," *Photogrammetrie Fernerkundung Geoinformation (PFG)* vol. 5/2008, pp. 337-349, 2008.

[35] P. Dorninger and C. Nothegger, "3D segmentation of unstructured point clouds for building modelling," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVI-3/W49A, pp. 191-196, 2007.

[36] L. Zebedin, A. Klaus, B. Gruber, and K. Karner, "Facade reconstruction from aerial images by multi-view plane sweeping," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVI-3, pp. 31-36, 2006.

[37] S. Kumar and M. Hebert, "Discriminative random fields", *Int. J. Computer Vision*, vol. 68 (2), pp. 179–201, 2006.

[38] N. Haala, S. Becker, and M. Kada, "Cell decomposition for the generation of building models at multiple scales," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVI-3, pp. 19-24, 2006.

[39] N. Ripperda and C. Brenner, "Data driven rule proposal for grammar based facade reconstruction," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVI- 3/W49A, pp. 1-6, 2007.

[40] H. Mayer, S. Reznik, "MCMC linked with implicit shape models and plane sweeping for 3D building facade interpretation in image sequences," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVI-3, pp. 130-135, 2006.

[41] C. Brenner, "Dreidimensionale Gebäuderekonstruktion aus digitalen Oberflächenmodellen und Grundrissen," *Deutsche Geodätische Kommission*, vol. C-530, 2000.

[42] G. Vosselman, "Building reconstruction using planar faces in very high density height data," *Int. Arch. Photogrammetry, Remote Sens.*, vol. XXXII-3/2W5, pp. 87–92, 1999.

[43] A. Alharty and J. Bethel, "Detailed building reconstruction from airborne laser data using a moving surface method," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXV-B3, pp. 213–218, 2004.

[44] B. Ameri, "Automatic recognition and 3D reconstruction of buildings from digital imagery," *Deutsche Geodätische Kommission*, vol. C-526, 2000.

[45] C. Baillard, C. Schmid, A. Zisserman, and A. Fitzgibbon, "Automatic line matching and 3D reconstruction of buildings from multiple views," *Int. Arch. Photogrammetry, Remote Sens.*, vol. XXXII-3/2W5, pp. 69–80, 1999.

[46] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik, "Automated delineation of roof planes in LIDAR data," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXIV-3/W19, pp. 221–226, 2005.

[47] I. Suveg and G. Vosselman, "Reconstruction of 3D building models from aerial images and maps," *ISPRS J. Photogrammetry Remote Sens.*, vol. 58 (3-4), pp. 202-224, 2004.

[48] G. Vosselman and S. Dijkman, "3D building model reconstruction from point clouds and ground plans," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXIV-3W4, pp. 37–43, 2001.

[49] F. Rottensteiner, "Consistent estimation of building parameters considering geometric regularities by soft constraints," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXIV-3, pp. 13–18, 2006.

[50] H. Kaartinen, J. Hyyppä, E. Gülch, G. Vosselman, H. Hyyppä, L. Matikainen, A. D. Hofmann, U. Mäder, Å. Persson, U. Söderman, M. Elmqvist, A. Ruiz, M. Dragoja, D. Flamanc, G. Maillet, T. Kersten, J. Carl, R. Hau, E. Wild, L. Frederiksen, J. Holmgaard, and K. Vester, "Accuracy of 3D city models: EuroSDR comparison," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXIV-3/W19 pp. 227-232, 2005.

[51] K. Khoshelham, "Region refinement and parametric reconstruction of building roofs by integration of image and height data," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVI-3/W24, pp. 3-8, 2005.

[52] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik, "Fusing airborne laser scanner data and aerial imagery for the automatic extraction of buildings in densely built-up area," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXV-B3, pp. 512–517, 2004.

[53] H. Mayer, "Object extraction in photogrammetric computer vision," *ISPRS J. Photogrammetry Remote Sens.*, vol. 63, pp. 213-222, 2008.

[54] M. Durupt and F. Taillandier, "Automatic building reconstruction from a digital elevation model and cadastral data: an operational approach," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXIV-3, pp. 142 – 147, 2006.

[55] F. Dornaika and M. Brédif, "An efficient approach to building superstructure reconstruction using digital elevation maps", *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVII-B3A, pp. 179-184, 2008.

[56] D. Dörschlag, G. Gröger, and L. Plümer, "Sementically enhanced prototypes for building reconstruction," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVI-3/W49A, pp. 111-115, 2007.

[57] L. Van Gool, G. Zeng, F. Van den Borre, and P. Müller, "Towards mass-produced building models," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVI-3/W49A, pp. 209-221, 2007.

[58] eTRIMS image data base. Institute of Photogrammetry, University of Bonn, http://www.ipb.uni-bonn.de/projects/etrims_db/ (accessed 13 March 2009).

[59] F. Rottensteiner, "Automated updating of building data bases from digital surface models and multi-spectral images: potential and limitations", *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVII-B3A, pp. 265-270, 2008.

[60] S. Hinz and A. Baumgartner, "Automatic extraction of urban road networks from multi-view aerial imagery," *ISPRS J. Photogrammetry Remote Sens.*, vol. 58 (1–2), pp. 83–98, 2003.

[61] C. Steger,"Extracting curvilinear structures: a differential geometric approach," *Proc. 4th European Conf. Computer Vision*, pp. 630–641, 1996.

[62] C. Zhang, "Towards an operational system for automated updating of road databases by integration of imagery and geodata," *ISPRS J. Photogrammetry Remote Sens.*, vol. 58, pp. 166–186, 2004.

[63] M.-F. Auclair-Fortier, D. Ziou, C. Armenakis, and S. Wang, "Survey of work on road extraction in aerial and satellite images", *Canadian J. Remote Sens.*, vol. 27 (1), pp. 76–89, 2001.

[64] M. Gerke, M. Butenuth, C. Heipke, F. Willrich, "Graph-supported verificationof road databases", *ISPRS J. Photogrammetry Remote Sens.*, vol. 58, pp. 152–165, 2004.

[65] V. Lacroix and M. Acheroy, "Feature-extraction using the constrained gradient", *ISPRS J. Photogrammetry Remote Sens.*, vol. 53, pp. 85-94, 1998.

[66] C. Wiedemann and S. Hinz, „Automatic extraction and evaluation of road networks from satellite imagery," *Int. Arch. Photogrammetry, Remote Sens.*, vol. XXXII, pp. 95–100, 1999.

[67] J. Mena and J. Malpica, "An automatic method for road extraction in rural and semi-rural areas starting from high-resolution satellite imagery," *Pattern recognition letters*, vol. 26, pp. 1201-1220, 2005.

[68] Q. Zhang and I. Couloigner, "Benefit of the angular texture signature for the separation of parking lots and roads on high resolution multi-spectral imagery," *Pattern recognition letters*, vol. 27(9), pp. 937-946, 2006.

[69] C. Wiedemann, "Improvement of road crossing extraction and external evaluation of the extraction results," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXIV-3B, pp. 297–300, 2002.

[70] A. Barsi, C. Heipke, and F. Willrich, "Junction extraction by artificial neural network system – JEANS," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXIV-3B, pp. 18–21, 2002.

[71] H. Mayer, S. Hinz, U. Bacher, and E. Baltsavias, "A test of automatic road extraction approaches", *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXIV-3, pp. 209-214, 2006.

[72] S. Clode, F. Rottensteiner, E. Zelniker, and P. Kootsookos, "Detection and vectorisation of roads from Lidar data," *Photogrammetric Eng. Remote Sens.*, vol. 73 (4), pp. 517-536, 2007.

[73] Hu, X., C. V. Tao, and Y. Hu, "Automatic road extraction from dense urban area by integrated processing of high resolution imagery and LIDAR data," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXV-B3, pp. 288–292, 2004.

[74] P. Zhu, Z. Lu, X. Chen, K. Honda, and A. Eiumnoh, "Extraction of city roads through shadow path reconstruction using laser data," *Photogrammetric Eng. Remote Sens.*, vol. 70 (12), pp. 1433–1440, 2004.

[75] S. Clode, F. Rottensteiner, and P. Kootsookos, "Improving city model determination by using road detection from LIDAR data," *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol, XXXVI–3/W24, pp. 159-164, 2005.

[76] G. Sithole and G. Vosselman, "Bridge detection in airborne laser scanner data," *ISPRS J. Photogrammetry Remote Sens.*, vol. 61, pp. 33-46, 2006.

[77] J. Heuwold, K. Pakzad, and C. Heipke, "Scale behaviour prediction of image analysis models for 2D landscape objects", *Int. Arch. Photogrammetry, Remote Sens., Spatial Information Sc.*, vol. XXXVI-3/W49A, pp. 43 -50, 2007.

[78] A. Grote, C. Heipke, F. Rottensteiner, and H. Meyer, "Road extraction in suburban areas by region-based road subgraph extraction and evaluation," *Proc. Joint IEEE-GRSS/ISPRS Workshop Remote Sens. Data Fusion over Urban Areas*, Shanghai, China, 6 p. (on CD-ROM), 2009.

[79] M. Ravanbakhsh, K. Pakzad, and C. Heipke, "Automatic extraction of traffic islands from aerial images", *Photogrammetrie Fernerkundung Geoinformation (PFG)* vol. 5/2008, pp. 375-385, 2008.