MULTIPLE-MODEL BASED VERIFICATION OF JAPANESE ROAD DATA

M. Ziems[(a)], H. Fujimura[(b)], C. Heipke[(a)], F. Rottensteiner[(a)]

[(a)] IPI – Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover, Nienburger Str. 1, 30167 Hannover, Germany, {heipke, rottensteiner, ziems}@ipi.uni-hannover.de

[(b)] GSI – Geographical Survey Institute, Ministry of Land, Infrastructure, Transport and Tourism, Kitasato 1, Tsukuba, Ibaraki, 305-0811, Japan, hfu@gsi.go.jp

Commission IV, WG IV/2

KEY WORDS: Databases, Quality, Updating, Orthoimage, Classification, Modelling

ABSTRACT:

This paper describes a semi-automatic system for road verification from high resolution orthophotos in an urban context. The system combines several road detection and road verification approaches from current literature to form a more general solution. Each road detection / verification approach is realized as an independent module representing a unique road model and thus a unique strategy. The object-wise verification result of each module is formulated as a binary decision between the classes "correct road" and "incorrect road". These individual decisions are combined by Dempster-Shafer fusion, which provides tools for dealing with uncertain and incomplete knowledge about the statistical properties of the data. For each road detection / verification module a function for confidence is introduced that reflects degree of correspondence of an actual test situation with an optimal situation according to the underlying road model of that module. Experimental results achieved with four different test sites in Japan demonstrate the potential and confirm the reliability of the new system.

1. INTRODUCTION

The Geographical Survey Institute (GSI), the national mapping agency of Japan, is currently establishing an authoritative topographic cartographic database (Fundamental Geospatial Data, which then is introduced into Digital Japan Basic Map (Map Information)) corresponding to a mapping scale of 1:2500 for the whole urban area of Japan. The data structure is standardized, but the data are produced by different local governments and thus show heterogeneous quality. Therefore, quality control measures are required to achieve a consistent quality standard. In order to reduce the huge manual efforts for this quality check, this task has to be automated without compromising the data quality. We propose a semi-automatic strategy based on orthophotos for that purpose. The basic idea is that database objects found to exist at the location indicated in the database by an automated process need not be checked interactively. The database is not corrected fully automatically, but the automatic procedure highlights potential errors that have to be checked by a human operator. The main challenge is the general applicability required for the automatic system because the appearance of roads varies a lot in the images. Roads can be described as homogeneous areas with parallel edges, by their spectral or structural differences to their surroundings, by road markings, or by relations to context objects such as cars or buildings. As most algorithms for road extraction can only deal with a subset of these characteristics, their success is usually connected to a specific road type. Different road types can be defined by their appearance in the imagery, and thus can be extracted by different methods with different success. If each of these specific road extraction methods can deliver a measure for its own trust into its decision, the results of these algorithms can be combined by decision level fusion. This paper deals with the setup of such a fusion framework, with a focus on the definition of confidence values for different road extraction/verification methods.

2. RELATED WORK

In this review of existing approaches for road extraction, we focus on the aspects that are most important for the work presented in this paper, namely on the underlying road models and their consequences for road objects that do not conform to the models and on the robustness with respect to different appearance of road objects in different surroundings.

In (Baumgartner et al., 1999; Wiedemann, 2002) roads are modelled as linear objects in aerial or satellite imagery with a resolution of about 1-2 m. If the underlying line extraction algorithms are parameterized based on prior knowledge about the road width from an existing database, these algorithms can be applied to different road types. The line model can also be improved by incorporating parallel edge pairs (Baumgartner et al., 1999; Zhang, 2004). Nevertheless, these line based approaches only work in areas with homogeneous background, especially in rural areas. The EuroSDR test on automatic road extraction algorithms (Mayer et al., 2006) has shown the weakness of such approaches in an urban context. In contrast, the edge based approach by Youn et al. (2008) is specifically designed for dense settlement areas. The underlying model focuses on the discrimination of rows of buildings and roads and thus cannot deal with homogeneous background. The road model is also restricted to a grid-like road network.

Another group of approaches is based on aerial colour or texture classification in high resolution aerial imagery, e.g. (Mena and Malpica, 2005; Zhang & Couloigner, 2006). Roads are modelled as homogeneous image regions with certain radiometric properties. The geometrical aspects of the road model are considered in the post processing steps. Fujimura et al. (2008) use prior information from an existing database to define an appropriate shape of a window for a texture analysis that is carried out at all probable window positions around the position indicated by the database. For road tracking,

Haverkamp (2002) and Liang et al. (2008) assume a similar model. Here the image texture in a rectangular window is used to track the direction of a road from a starting point defined previously. Grote and Rottensteiner (2009) apply a normalized cuts segmentation and then classify the resulting segments. They are grouped in a process that considers the geometrical and topological properties of roads. All these area-based algorithms can in principle deal with complex scenes, but they have problems with objects that are radiometrically similar to roads, e.g. buildings and parking lots. In (Bacher & Mayer, 2005), a combined approach is introduced. A pixel-based classification is used to generate a "road class-image". The training data for the supervised classification is obtained from a very strict initial road extraction according to (Wiedemann, 2002), using also parallel edge information (Baumgartner et al., 1999). Various works, e.g. (Hinz & Baumgartner, 2003; Zhang 2004; Youn et al., 2008; Grote & Rottensteiner, 2009) introduce 3D information to make the road model more robust to confusion with buildings. The 3D information is used either as an additional feature in the classification step or for an internal evaluation of 2D road hypotheses. Context objects such as buildings or vehicles can also be used as an additional cue for road extraction. Hinz and Baumgartner (2003) use buildings and vehicles and connect them with the main road model by Fuzzy-Set theory. In (Zhang 2004), buildings and trees are integrated in a context model. In (Hu et al., 2004), vehicles are used to classify roads and parking lots.

It is the basic idea of this work to combine powerful approaches from the current research in the field of road extraction to create a more general solution for the verification of urban road data that is applicable for the whole road network of Japan. Rather than formulating a complex road model that is valid for all the roads of Japan, we exploit the specific strengths of existing algorithms for the situations that conform to their underlying road models. An alternative strategy could be the consequent use of machine learning algorithms, where all the low level features used in the existing approaches are collected in a large feature vector. In that case the different road types would correspond to different clusters in feature space. Such a strategy could deal with more types of roads, but it would require a high amount of training data, which is not available in our case.

3. THE MULTILPLE MODEL METHOD

Our method relies on a set of existing object extraction algorithms realized as so called *verification modules*. Every road object in the database is checked by every available verification module. Along with its decision about the correctness of the road object in the database, each module also delivers a confidence value C with $0 \le C \le 1$ that reflects the degree to which the situation encountered for the road object corresponds to the optimal situation according to the module's underlying object model. The decisions from all modules are combined in a decision level fusion process in which the confidence values control the impact of a single decision on the final result. It is possible to include algorithms for extracting other objects than roads into our framework. For instance, if a building detection algorithm is confident in detecting a building where the database indicates a road, this can be considered in the overall framework as a very confident vote against the correctness of the road object.

There are three advantages to our strategy. Firstly, the combination of an increasing number of approaches based on different models covers more road objects and thus improves the efficiency of quality control. Secondly, whereas road candidates not covered by any of the introduced road models still cannot be classified reliably, such a situation should be detectedable by analysing the confidence values, so that these objects can be passed on to the human operator for a final decision. Thirdly, the framework can be expanded easily by new verification modules.

The fusion of the results from the different verification modules is based on the theory of Dempster-Shafer, e.g. (Klein 1999). Our approach follows the thoughts of Gerke and Heipke (2008) and distinguishes the two classes *road* (R) and *non-road* (N). Consequently, the hypothesis space, which is called frame of discernment Θ in the terminology of Dempster-Shafer, contains only of two elements: $\Theta = \{R, N\}$. The power set of Θ, denoted by 2^{Θ}*-1,* is $2^{\Theta}\text{-}1 = \{R, N, R \cup N\}$. A probability mass m is assigned to each of the three classes by a "sensor" (verification module) such that $0 \le m(x) \le 1$ and $m(R)$, $m(N)$ and $m(R \cup N)$ sum up to 1. The sum of all probability masses assigned directly to a class $A \in 2^{\Theta}\text{-}1$ is called support $sp(A)$ of A. If p sensors are available, probability masses m_i have to be defined for all these sensors i with $1 \le i \le p$. The Dempster-Shafer theory allows the combination of the probability masses from several sensors to compute a combined probability mass for each class $A \in 2^{\Theta}\text{-}1$:

$$m(A) = \frac{\sum_{B_1 \cap B_2 \ldots \cap B_p = A} \left(\prod_{1 \le i \le p} m_i(B_j) \right)}{1 - \sum_{B_1 \cap B_2 \ldots \cap B_p = \emptyset} \left(\prod_{1 \le i \le p} m_i(B_j) \right)} \quad \text{with } B_j \in 2^{\Theta}\text{-}1 \qquad (1)$$

After combining the probability masses using equation 1, both $sp(R)$ and $sp(N)$ can be computed. The normal case is that the class obtaining maximum support will be accepted. However, we know that we will not find a solution for all objects. For our specific task it is most important to ensure that all errors in the database are detected. Road objects from the database that are detected by the automatic component are no longer inspected by the user. A false detection of a road thus directly leads to an error in the database. On the other hand, a road object that is not detected by the automatic component will be inspected by the user. Thus, this error type is less critical. As a consequence we only consider roads as verified for which the data reliably support that decision. This can be achieved by only accepting road hypotheses with a support $sp(R)$ larger than 0.75. This threshold of 0.75 was found by empirical analysis.

In our model for the original probability masses we assume that each verification module p delivers a binary decision for or against a road, i.e., either R_p or N_p, and a confidence value $C_p \in \{0, 1\}$ measuring the trust into this decision. Its negation $C_p^N = 1 - C_p$ corresponds to the degree to which no decision can be taken by the module given the data. This can be modelled by assigning a probability mass of $1 - C_p$ to Θ, thus $m_p(R_p \cup N_p) = 1 - C_p$. If the module's decision is R_p, we set $m_p(R_p) = C_p$ and $m_p(N_p) = 0$; otherwise, we set $m_p(R_p) = 0$ and $m_p(N_p) = C_p$. Thus, the decision is weighted by C_p in the Dempster-Shafer framework. If the confidence value C_p is low for all modules p, the support for R and N will be relatively low, too, so that these cases can be found by applying a threshold as described above.

4. THE VERIFICATION MODULES

Currently, our system includes seven verification modules. Three of them required considerable improvement compared to the way they were described in the literature and thus will be presented in more detail. This includes a short description of the underlying road model and an overview of the basic strategy for

its application to database verification with a focus on the way the information from the database is used to optimize the verification process. Finally, the derivation of the confidence value will be described. The further four modules will be described by a short summary.

All the modules are adjusted to deal with RGB orthophotos with 0.2 m ground resolution (Digital Japan Basic Map (Orthophoto Imagery), developed and maintained by GSI) as well as a Digital Terrain Model (DTM) and a Digital Surface Model (DSM) generated as by-products of the orthophoto generation process with 5 m grid spacing and 0.1 m resolution for the height component. From the DSM and the DTM a normalised DSM (nDSM) reflecting the height differences between the DSM and the DTM is generated. The road database corresponds to a cartographic map scale of about 1:2500. The geometrical accuracy standard for road objects in the database is stated as 2.5 m, but the real accuracy is significantly higher (approximately 1 m).

4.1 The SSH Module

SSH means "Sum of Similarity of Histograms" and was introduced as a road model by Fujimura et al. (2008) to remove a parallel shift of cartographically generalized road data.

4.1.1 Model: The basic idea is that the image region belonging to a road can be identified by the uniqueness of its intensity distribution compared to its surroundings. The model assumes that the intensity distributions in areas in the vicinity of a road are more similar to each other than to the distribution in the road. Since colour is used by another module, only the intensity channel is used for the calculations.

4.1.2 Strategy: Figure 1 shows the strategy of the SSH based verification in an exemplary way. The road geometry and width information from the database is used to define several image regions with identical shape and area. This is realized by shifting the road region in a direction orthogonal to the road axis. The region in the centre represents the road candidate from the database. All regions outside the tolerated GIS-specific error distance represent the neighbourhood and are thus expected to be non-road regions. To avoid a second road candidate in the neighbourhood the number of generated regions is parameterized by the expected city block size. This parameter depends on the character of a city district and has to be set by the user in advance. Subsequently, the intensity histograms for all regions are calculated. For each histogram s, a similarity value SSH is computed by comparing it to all other histograms t. The histogram similarity is expressed using the Bhattacharyya distance $BC(H_s,H_t)$ (Fukunaga 1990):

$$SSH(s) = \sum_{t=-S}^{S} (1-\delta_{s,t}) \cdot BC \;, \quad BC = \sum_{g \in G} \sqrt{H_s(g) \cdot H_t(g)} \qquad (2)$$

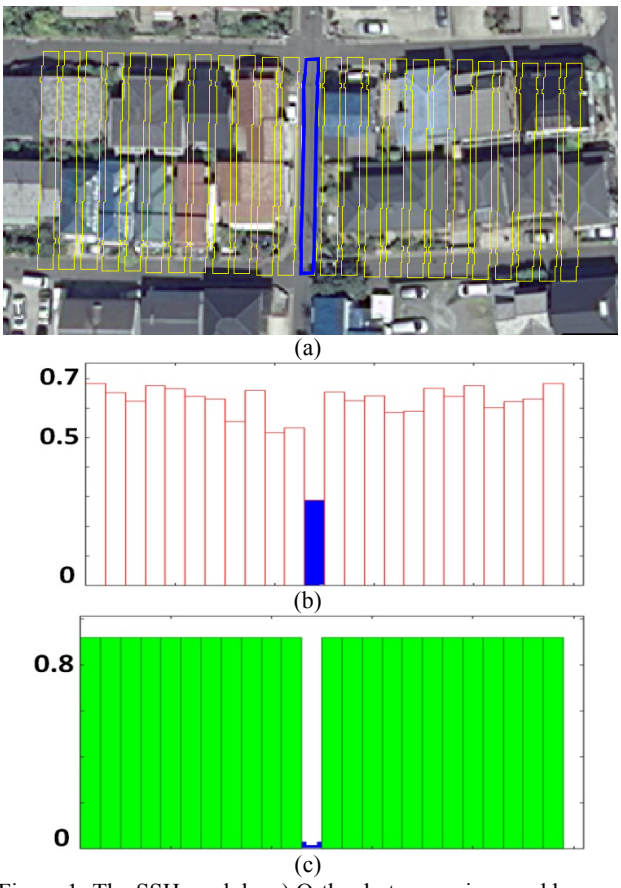
In equation 2, δ denotes the Dirac function and g denotes a grey value in the set of possible grey values G. Figure 1b shows an exemplary result of the SSH based approach in urban area. If the SSH of the histogram associated to the road in the database is significantly lower than the SSH scores of the non-road regions the road is classified as correct. For this significance test a Gaussian distribution of all SSHs from the off-road segments is assumed. This assumption has been evaluated empirically for relatively long roads with homogeneous background. However, it does not hold for all target objects. Therefore, the model should be enhanced by an adaptive distribution in the future.

4.1.3 Confidence: The optimal realization for the SSH model is characterized by a homogeneous neighbourhood in a direction orthogonal to the given road axis. But the model can also deal with inhomogeneous neighbourhoods if there is no single non-road region having a high SSH score. Thus, a continuous function for the confidence C_{SSH} of the SSH module is useful. C_{SSH} is modelled as the difference of the actual configuration (Figure 1b) and the optimal SSH model SSH_{opt} (Figure 1c):

$$C_{SSH} = \frac{1}{R} \cdot \sum_{r=1}^{R} \left(SSH(r) - SSH_{opt}(r) \right) \cdot C_{Length}(l) \qquad (3)$$

In equation 3, r is the region index and R is the number of regions considered. C_{SSH} mainly depends on the surroundings of the road. A confidence $C_{SSH} = 1.0$ is achieved if all non-road regions have the same intensity histogram and only the road region is different. A confidence $C_{SSH} = 0.0$ corresponds to a situation where every non-road region has a histogram that is totally different from the histograms of all the other non-road regions. The term $C_{Length}(l)$ depending on the road segment length l was introduced to model the fact that longer road segments are more reliable because they contain more pixels to compute the histograms and because they are less likely to be confused with building roofs. This term $C_{Length}(l)$ is assumed to be 1 for road objects longer than a predefined threshold L_{max}. The parameter L_{max} corresponds to the length of big buildings in a scene and has to be defined by the user. For $0 \le l \le L_{max}$, $C_{Length}(l)$ is described by a cubic parabola:

$$C_{Length}(l) = 3 \cdot \left(\frac{l}{L_{\max}} \right)^2 - 2 \cdot \left(\frac{l}{L_{\max}} \right)^3 \quad \forall \; l \mid 0 \le l \le L_{max} \qquad (4)$$



(a) (b) (c)

Figure 1. The SSH module. a) Orthophoto superimposed by one road region (blue) and 24 non-road regions (yellow); b) SSH; the x-axis shows the region index. The blue column denotes the true road candidate; c) Optimal SSH configuration.

4.2 The Colour Module

We use the approach described in (Fujimura et al., 2008) as a verification module evaluating colour information

4.2.1 Model: An image region belonging to a road has specific radiometric properties that can be defined in advance. More precisely, the RGB feature space is used to define probable clusters for roads and for non-roads.

4.2.2 Strategy: As the goal of this approach is the verification of a database and not the extraction of new roads, the problem can be relaxed compared to other approaches. No pixel-wise analysis is required, and the entire road candidate region is considered at once. Regions are defined in the neighbourhood of the road segment in a similar way as for the SSH module (see Figure 1), and the histogram of every region is calculated independently for all bands. A modal filter is applied to extract only the five most typical gray values for every region and every band and thus exclude small objects from the analysis. The verification problem is turned into a classification problem for the two classes road and non-road. We use the support vector machine (SVM) classifier (Vapnik, 1998) in the implementation of the open-source library LIBSVM (Chang & Lin, 2001). The necessary training data are selected manually and should include representative examples of road and non-road regions in a scene. The SVM classifier provides a decision for every tested road candidate region.

4.2.3 Confidence: It was shown by Zhang and Couloigner (2006) that a colour based classification is weak in areas with parking lots, houses and industrial buildings having similar radiometric properties as roads (Figure 2b). Hence, the colour model is only applicable if the surrounding area shows good contrast, e.g. for houses with coloured roof (see Figure 2a). Therefore, the SVM classification is applied to all the regions defined in a similar way to Figure 1a. If the region corresponding to the road in the database is classified as non-road, the road is decided to be incorrect. If there are multiple regions classified as roads, including areas that are supposed to be non-road regions by the model, no decision can be taken based on this module, which has to be reflected by a confidence value C_{col} of 0. Only if the region corresponding to the road in the database is classified as road and all surrounding regions are classified as non-road, the colour module provides a decision for the correctness of the road.

(a) (b)

Figure 2. Two orthophotos superimposed by road edges (yellow). a) Buildings with coloured roofs and good contrast to the road; b) Buildings with a low contrast.

The basic precondition for the SVM classifier is the availability of representative training data. For practical use it is hard to fulfil such a rigid demand. Especially for the non-road class this seems to be impossible with a realistic amount of training data. As the SVM classifier always delivers a solution independently of the distance of a point in feature space to the nearest training cluster, the resulting decisions could be based on insufficient data without any indications for the user. In order to overcome this problem, the confidence C_{col} of the classification result is described by the degree of correspondence between training data and a test object. The concept of Support Vector Domain Description (SVDD) (Tax, 2001), also known as SVM-based outlier detection, is used for this purpose. This approach constructs a hypersphere around the training data in the transformed feature space (based on the Gaussian kernel in our SVM approach) so that it encloses most of the training data with minimal volume. As the hypersphere is constructed in the transformed feature space, it corresponds to a region of arbitrary shape in input feature space. As mentioned before the reliability of the SVM results decreases for high distances to the clusters representing the used training dataset. The sequential combination of the conventional two-class SVM and the SVDD allows a very sharp separation of two classes through the SVM, whereas vague decisions based on unexpected data can be prevented through the SVDD. However the use of SVDD leads to a decreasing efficiency because one important advantage of two-class SVM, the generalization effect, is significantly weakened at the periphery of the trained clusters. Therefore, for points outside the hypersphere, their distances from the hypersphere are used to define a softer criterion so that objects close to but outside of the hypersphere will still receive a high confidence value compared to objects that are far away. We use the method introduced by Guo et al. (2009) to compute the distance d of a test point to the hypersphere. The average distance D_{avg} of test points enclosed by the hypersphere to their nearest hypersphere point is calculated to define an appropriate scale. In our experience the objects outside the hypersphere that are further away than $½D_{avg}$ show significant differences to the training dataset and thus are assigned a zero confidence. Objects inside the hypersphere obtain full confidence ($C_{col} = 1$). For objects outside the hypersphere, but having a distance smaller than $½D_{avg}$ to the hypersphere, the confidence C_{col} is modelled by a cubic parabola function.

$$C_{col} = 3 \cdot \left(\frac{d}{1/2\,\mathrm{D_{avg}}} \right)^2 - 2 \cdot \left(\frac{d}{1/2\,\mathrm{D_{avg}}} \right)^3 \cdot C_{Length}(l) \qquad (5)$$

$$\forall\ 0 \le d \le 1/2\,\mathrm{D_{avg}}$$

In equation 5, again a length dependent factor (see equation 4) is included to consider the weakness of the model for short objects. The colour module provides a decision for the correctness (R_{Col}) or against the correctness (N_{Col}) of a road object and a value for the confidence C_{col}. Note that a high percentage of small confidence values for a scene is taken as a hint for the operator to define a new training dataset. This occurs usually if the radiometric conditions or the style of buildings change dramatically.

4.3 The Intersection Module

This module is based on the method developed by Youn et al. (2008) for the task of road extraction in dense urban areas.

4.3.1 Model: The model takes into account the structural differences between a road and a row of buildings. In an edge image the direction of the majority of detectable edges within a road region conforms to the direction of the road. In the area next to a road, objects such as houses lead to multiple edge directions. This difference is used for classification.

4.3.2 Strategy: The original method by Youn et al. (2008) was adapted to be used for verification. This module defines a set of lines that are parallel to the road segment from the database and counts the number of intersections of these lines with edges extracted from the image (see Figure 3). According to the model, there should be no edges across the road, so that the intersection count should be close to zero for lines inside the

road (cf. Figure 3c). Figure 4a shows the distribution of this count for the lines in Figure 3. The small values in the centre indicate the true position of the road, whereas greater values indicate buildings. The number N of shifted elements is restricted by the expected city block size L_{Blo} that is one of the parameters to be set by the user. To check the positional accuracy the histogram of intersections is smoothed by a Gaussian filter of width σ (Figure 4b). The attribute *road width* available from the input data is used to define a suitable value for σ. The minimum of the filtered histogram is assumed to correspond to the road centreline. If the distance between the position of this minimum and the position of the centreline indicated by the database is below the maximum allowable error according to the specifications of the database and if the number of intersections is smaller than a threshold the road is classified to be correct and thus R_{Int} is true. Whereas the model assumes that there are no intersections at all for the road, a value larger than 0 allows for small disturbances on the road surface as they might be caused by single road markings or shadows. The threshold was set to *2* in all experiments.

4.3.3 Confidence: For dense urban areas containing a lot of small houses, the model is robust. However, more homogeneous context such as grassland, paddy fields or huge industry halls is not covered by the model. Therefore, the confidence value is modelled as a function of the surrounding structure elements. The actual histogram is compared with a histogram based on the optimal situation for that model. This optimal situation is an absolute free passage through the expected road and a number of intersections on each side of the road, which occur if a row of buildings of standard size is situated next to the road. The standard size of a building varies with the area and is another parameter to be set by the user. Figure 4c shows such an optimal histogram for the example depicted in Figure 3. The confidence C_{int} for the intersection model is calculated as the area ratio between the optimal histogram H_{Model} and the actual histogram H. As the neighbouring areas may be fairly different on both road sides, the area to the left and the area to the right of the road are considered by different terms.

$$C_{Int} = \frac{\sum_{n=1}^{A-1} \Delta(n) \cdot \sum_{n=A+1}^{N} \Delta(n)}{\left(\sum_{n=1}^{N} H_{Model}(n) \right)^2} \cdot C_{Length}(l) \qquad (6)$$

In equation 6, n is the profile index of A is the index of the centreline profile. $\Delta(n) = H_{Model}(n) - H(n)$ if $H_{Model}(n) \geq H(n)$ and $\Delta(n) = 0$ otherwise. Again a length dependent term $C_{Length}(l)$ (see equation 4) is included to prevent instabilities of the model with respect to very short roads.

4.4 Further Modules

A frequently used characteristic of roads are parallel edge pairs to represent road borders in an edge image. In our system, this information is used by a module based on the approach by Baumgartner et al. (1999). Another module is based on (Hinz and Baumgartner, 2003) where roads are detected as valleys in nDSM data. To define an appropriate confidence for these modules the extraction results are evaluated by a statistical approach presented in (Gerke & Heipke, 2008).

Finally, two modules extracting buildings and vegetation, respectively, are included. Building detection is based on nDSM information and follows the work of Hinz and Baumgartner (2003), whereas vegetation detection is realized similar to (Zhang, 2004). With regard to our task the nDSM based building model turns out as very robust in any situation. The only weak point is the low spatial resolution of the nDSM. Therefore, the confidence function describes its proportion to the assumed road width. Since the vegetation module uses radiometric properties, the confidence function is based on the colour saturation of the detected vegetation areas. In contrast to the other modules the building and the vegetation modules can only vote for the incorrectness of a road object. Possible contradictions are resolved by the Dempster-Shafer approach.

(a) (b) (c)

Figure 3. The intersection module. a) Image detail. b) Extracted edges. c) Edges (black) superimposed by expected road centreline (cyan, dotted) and several parallel profiles (blue).

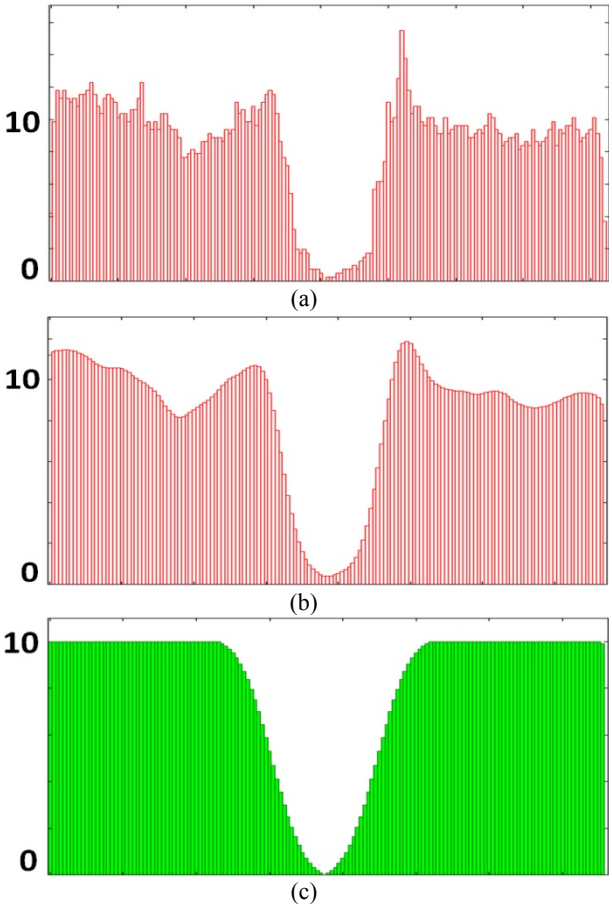


Figure 4. Histogram of the intersection counts over the profile index. a) Raw results of the example in Figure 3. b) Results after Gaussian filtering with σ = road width. c) H_{Model}: Optimal situation according to the model.

5. EXPERIMENTS

In order to evaluate the proposed verification strategy, a test based on four urban test sites with each 0.76 km x 0.92 km was carried out (Figure 5). The Uraga test site is situated in a hilly sub-urban area with many small residential buildings. The Okayama test site shows mainly big modern houses in flat terrain. There are hardly any shadows. This is the only scene with nDSM data. The test site Kyoto1 is situated in the historic centre of Kyoto and is characterised by grey roofs and narrow roads. There are significant shadow effects. The second test site from Kyoto (Kyoto2) shows some big residential and community buildings. The roads are wide, partly covered by trees, and there are significant shadow effects.

In total 1124 road objects were automatically processed. For the Okayama test site we could use all verification modules, whereas for the others we could not use the building detection and the 3D line detection modules because they require a nDSM. Only road objects that show a support over 0.75 are accepted as correct by the automatic system. All the other objects are highlighted for manual post processing. For the Uraga test site (Figure 5a) where there is a significant difference between the database and the image due to a huge redevelopment zone, the automatic results could be compared to a manual reference of correct / incorrect database objects. The confusion matrix is shown in Table 1. The system finds enough support for 46.3% of the road objects (plus one false positive). This value marks the advantage of the semi-automatic approach against a completely manual procedure and can be called "efficiency". The false alarm rate, i.e. the rate of roads that correspond poorly with one of the road models used by the verification modules, is 36.3%. The low rate of undetected errors (only 1, corresponding to 0.2%) shows the high reliability of the system for the test scene.

	accepted (system)	rejected (system)
correct (reference)	46.3% (185)	36.3% (145)
incorrect (reference)	**0.2% (1)**	17.2% (69)

Table 1. Confusion matrix for the Uraga test site.

Efficiency analysis ∑= 1054 road objects		
	accepted (system)	rejected (system)
correct dataset	**57.8% (609)**	42.2% (445)
Sensitivity analysis ∑= 791 road objects		
	accepted (system)	rejected (system)
incorrect dataset	**0.5% (4)**	99.5% (787)

Table 2. Efficiency and stability analysis for the Okayama, Kyoto1 and Kyoto2 test sites.

For the other three scenes the number of errors in the vector data was too small to determine a meaningful confusion matrix. Therefore, the system was tested by two different setups: the first contains only correct data; the second contains only incorrect data. The first dataset corresponds to the input data without a few wrong objects. In this case, 100% of the objects should be accepted, whereas 0% should be rejected. The second dataset was generated by rotating the first dataset by 180°. In this case, the system should reject 100% of the objects and accept none. The results shown in Table 2 confirm the results for the Uraga test site, with an efficiency of 57.8% and less than 1% of undetected errors. The best results were achieved for the Okayama test site because the modules requiring a nDSM could be used there. Unmodelled shadow effects are a weak point of our system. In order to improve efficiency, a module considering shadows should be included.

6. CONCLUSION AND OUTLOOK

The results presented in this paper show that the manual efforts for road verification can be reduced by nearly 60% on the basis of the ideas presented in this paper. The introduced modules use comparably simple models, but more complex models that describe the relations between shadow, cars and buildings can be added without changing the overall framework. Thus, more and more road objects as well as background objects could be described in future, which can raise the efficiency of the automatic system step by step.

ACKNOWLEDGEMENTS

A part of the work was funded by the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) by supporting a one-year stay of H. Fujimura at IPI. A four-month stay of M. Ziems at GSI was financed by the Japan Society for the Promotion of Science (JSPS). We gratefully acknowledge this support.

REFERENCES

Bacher, U., Mayer, H., 2005. Automatic road extraction from multispectral high resolution satellite images. In: IAPRSIS XXXVI (B3/W24), pp. 29–34.

Baumgartner, A., Steger, C., Mayer, H., Eckstein, W., Ebner, H., 1999. Automatic road extraction based on multi-scale, grouping, and context. *PE & RS* 65(7):777–785.

Gerke, M., Heipke, C., 2008. Image based quality assessment of road databases. *International Journal of Geoinformation Science* 22 (8): 871-894

Grote, A., Rottensteiner, F., 2009. Assessing the impact of digital surface models on road extraction in suburban areas by region-based road subgraph extraction. In: IAPRSIS XXXVII(3/W4), pp. 27-33.

Fujimura, H., Ziems, M., Heipke, C., 2008. De-generalization of Japanese road data using satellite imagery. *PFG* 5 (2008): 363-373.

Fukunaga, K., 1990. Introduction to Statistical Pattern Recognition. Academic Press, 2nd edition.

Guo, S. M, Chen, L. C., Tsai S. H., 2009. A boundary method for outlier detection based on support vector domain description. *Pattern Recognition* 42(1):77-83.

Haverkamp, D., 2002. Extracting straight road structure in urban environments using IKONOS satellite imagery. *Optical Engineering* 41(9): 2107-2110.

Hinz, S., Baumgartner, A., 2003. Automatic extraction of urban road networks from multi-view aerial imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 58(1-2): 83 - 98.

Hu, X., Tao, C.V., Hu, Y., 2004. Automatic road extraction from dense urban area by integrated processing of high resolution imagery and LIDAR data. In: IAPRS XXXV(B3), pp. 288-292.

Klein, L., 1999. Sensor and Data Fusion, Concepts and Applications. SPIE Optical Engineering Press, 2nd edition.

Liang, Y., Shen, J., Lin, X., Bi, J., Li, Y., 2008. Road tracking by parallel angular texture signature. Proc. Earth Observation and Remote Sensing Applications (EORSA 2008), pp.1-6.

Mena J., Malpica J., 2005. An automatic method for road extraction in rural and semi-urban areas starting from high resolution satellite imagery. *Pattern Recognition Letters* 26(9):1201–1220.

Mayer H., Baltsavias E., Bacher U., 2006. Automated extraction, refinement, and update of road databases from imagery and other data. *Report Commission 2 on Image Analysis and Information Extraction*, European Spatial Data Research - EuroSDR, Official Publication 50, pp. 217-280.

Tax, D., 2001. One-class classification. Unpublished doctoral dissertation, Delft University of Technology.

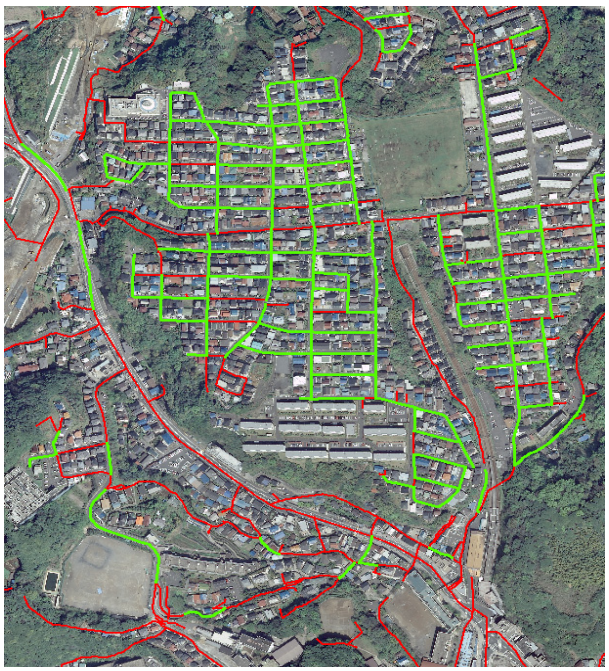
Vapnik, V. N., 1998. Statistical Learning Theory. Wiley, New York.

Wiedemann, C. 2003. External evaluation of road networks. In: IAPRSIS XXXIV, (3/W8), pp. 93-98.

Youn, J., Bethel, J. S., Mikhail, E. M., Lee, C., 2008. Extracting urban road networks from high-resolution true orthoimage and Lidar, *PE&RS* 74(2): 227-238.

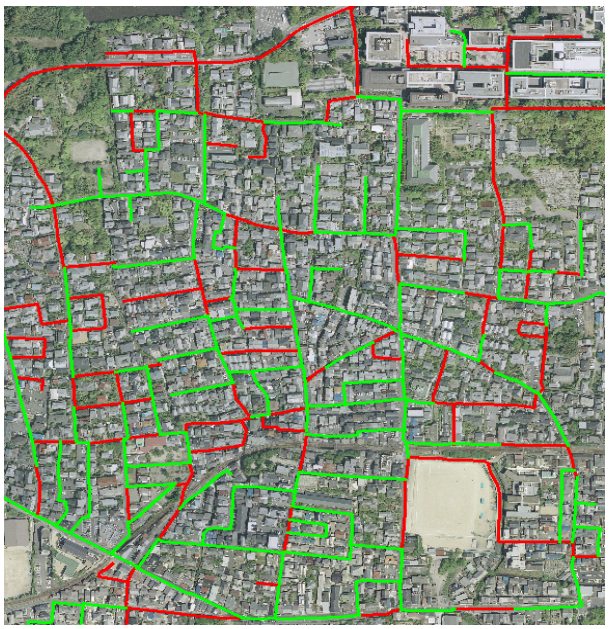
Zhang, C., 2004. Towards an operational system for automated updating of road databases by integration of imagery and geodata. *ISPRS Journal of Photogrammetry and Remote Sensing* 58 (3/4): 166–186.

Zhang Q., Couloigner I., 2006. Automated road network extraction from high resolution multi-spectral imagery. In: ASPRS 2006 Annual Conference, Reno, Nevada (on CD).

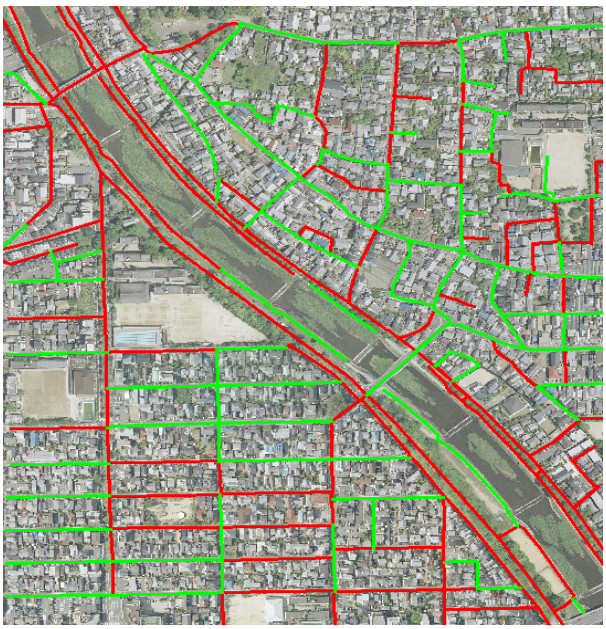


Uraga

Okayama



Kyoto1



Kyoto2

Figure 5: Results of road verification for the four test sites (red: rejected by the system, green: accepted by system.