

SCALE-DEPENDENT ADAPTATION OF IMAGE ANALYSIS MODELS INCORPORATING LOCAL CONTEXT OBJECTS

Janet Heuwold, Kian Pakzad, Christian Heipke

Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover,
Nienburger Str. 1, D-30167 Hannover, Germany – (heuwold, pakzad, heipke)@ipi.uni-hannover.de

Commission III, WG III/4

KEY WORDS: Multiresolution, Scale Space, Modelling, Image Analysis, Interpretation, Urban, Method

ABSTRACT:

Local context objects have significant impact on object extraction from real aerial or satellite images, since occlusions or shadows can substantially hinder a successful extraction of a particular landscape object one is interested in. This paper presents an adaptation concept for image analysis object models considering local context objects to a lower image resolution. The scale-dependent adaptation of local context poses a severe problem for an unambiguous scale behaviour prediction, as the exact position of local context objects is generally unknown. However, by adapting the object models for the landscape object of interest and the local context objects separately, the influence of this contradiction is avoided while the loss of the exact combined scale behaviour prediction remains acceptable. An example for road extraction with vehicle as local context object demonstrates the capability of the adaptation approach. To conclude, the assumptions for the adaptation methods and its limitations are discussed.

1. INTRODUCTION

A main prerequisite of knowledge-based image analysis from aerial and satellite data is the availability of suitable models for the extraction of relevant landscape objects. In the last decade more and more sensors from space with very high spatial resolution capabilities became available. The range of available image resolution, thus, has been enlarged. An image scene changes with different resolution: area-type features change to lines or points, some details may get lost or small features disappear, while other features may become more dominant and important for the extraction of a particular object. As landscape objects change their appearance in images of different image resolutions, models for the extraction of landscape objects are in general not applicable to another image resolution. Automatic object extraction can require substantially different models for different image resolutions, because of the large amount of various details that are not visible or have merged with adjacent objects in coarser image resolution. Therefore, the need for different models for object extraction in different image resolutions is growing as well.



Figure 1. Road with vehicle as local context object

Furthermore, the extraction process of a landscape object may be severely influenced by local context objects, i.e. other nearby objects in the vicinity occluding the landscape object one is interested in. For example, roads are often partly obstructed by vehicles, trees, buildings or their shadows (cf. Figure 1). Then, the extraction of the object of interest can fail, since some object

parts cannot be completely extracted due to the influence of local context objects. Hence, modelling of a landscape object only by its inherent features is often not sufficient for reliable extraction results, but requires the incorporation of local context objects.

A methodology for the automatic adaptation of image analysis object models for 2D landscape objects represented in form of semantic nets was developed previously (Heuwold et al., 2007) using linear scale-space theory. This procedure is extended here to incorporate additional local context objects. The context objects are modelled as a semantic net as well with attributes describing their appearance and their specific feature extraction operators. While from the object models treated so far all parts of the landscape object of interest and their relative position was known, the presence and location of local context objects is usually not. The location of local context objects can be extremely variable and can mostly only be determined roughly, if at all, in the image analysis model. This variability is contradictory to a non-ambiguous scale behaviour prediction based on position and needs to be accounted for in the adaptation method. The extension of the existing method for the scale-dependent adaptation of object models towards image analysis considering additional local context objects is the topic of this paper. Due to the high relevance for database update in urban areas, our study focuses on road extraction with vehicles as local context objects. However, the presented adaptation concept enables the application to other landscape objects as well.

The remainder of this paper is organized as follows: The next section reviews related work concerning the automatic adaptation of image analysis object models and the incorporation of local context into object extraction. Section 3 derives the developed concept for the adaptation of object models considering local context, while section 4 gives an example for the adaptation of a road model incorporating vehicles as local context. At the end of this paper, the results of the example are discussed and conclusions are drawn.

2. RELATED WORK

The topic of this paper comprises two important aspects: first, the automatic scale-dependent adaptation of object models and secondly, the modelling of local context objects and their incorporation into the extraction model for the primary object, which in our case complies with roads.

2.1 Automatic adaptation of object models to a coarser image resolution

The task of automatically adapting image analysis object models to a coarser image resolution was only rarely tackled previously. Heuwold et al. (2007) presented an algorithm for the adaptation of object models to a coarser target resolution. The adaptable models can consist of arbitrarily oriented line-type (1D) object parts and/or 2D object parts of area-type and are represented in form of semantic nets (Tönjes et al., 1997). At first, the high-resolution object model is decomposed into groups of object parts, which influence each other mutually during the reduction of image resolution (this influence is denoted as *interaction*). Based on linear scale-space theory (Witkin, 1983; Koenderink, 1984), the scale behaviour for each of these interaction groups is predicted using analysis-by-synthesis. The scale change analysis employs the blob detection algorithm by Lindeberg (1994). The analysis takes into account possibly occurring events in scale-space and predicts the shape and radiometric attributes of the individual object parts in the target resolution. These predicted low-resolution attributes serve as parameters for the respective feature extraction operators. In the last stage (fusion), all object part groups with their predicted appearance and adapted operators are composed back to a complete object model for extraction of this object in target resolution images.

While the adaptation of landscape objects with all types of object parts can be carried out, using the described approach the appearance and approximate relative location of the object parts must be specified in the object model to be adapted. Thus, an adaptation is up to now not possible with this algorithm for objects with unknown position. However, experiments with adapted object models for roads in image data of suburban areas (Heuwold, 2006) pointed out the need for incorporation of local context in order to achieve a high completeness for road extraction, as road extraction often suffers from the presence of nearby objects and their influence on the image data.

As far as we know only a single study concerning the scale behaviour analysis of local context objects exists (Mayer & Steger, 1998). A road lane with a vehicle is analysed in linear scale-space in order to choose a suitable image resolution for which an occluding vehicle does not disturb the extraction of the road. The scale for which the vehicle vanishes could be predicted analytically. However, the method of this study is only applicable to simple scenes, because the scale-space behaviour of more complex scenes cannot be predicted analytically.

2.2 Exploitation and modelling of local context

Local context objects can hinder the extraction of a landscape object by occlusion of or casting their shadow on the landscape object one is interested in, as context objects normally are situated very close to or on top of it. On the other hand, context objects can also support object extraction, if their pure existence is inferred. For instance, if a car is found in the image, with high probability a road must exist at this place. For road extraction the incorporation of local context proved to substantially

support extraction robustness in complex scenes (Baumgartner et al., 1997; Mayer, 2008).

As local context can greatly support a successful object extraction from remote sensing images, much work has been carried out studying the influence and exploitation of local context for image analysis. In general, reasonable context relations depend on the type of the individual context object and can be very diverse. For our application (roads) some of the proposed context relations are: *occludes, casts shadow on, is close to, is parallel to* for the most important context objects *vehicles, rows of trees and buildings* (Butenuth et al., 2003; Hinz & Baumgartner, 2003; Hinz, 2004; Gerke, 2006). The context objects to be used in the road extraction process can be extracted from images using object models, such as in (Hinz & Baumgartner, 2003) or derived from GIS, e.g. in (Zhang, 2004).

With growing demand for traffic monitoring in recent years many studies aim for vehicle extraction resulting in ample literature. Thanks to the availability of more visual remote sensors with high spatial resolution, explicit modelling with 2D or 3D models for such relatively small objects like vehicles became possible. Related work on vehicle extraction from visual image data can be divided into two groups based on the type of model used (Stilla, 2004): either appearance-based implicit models using training data or explicit models describing the vehicle in 2D or 3D. Most explicit approaches use edge extraction and subsequent grouping to a 2D rectangle applying methods such as the Generalized Hough Transform, or classification in a feature space using a Bayes net (Liu et al., 1999), or a 3D wire frame model for the vehicle (Zhao & Nevatia, 2003). Explicit object models with high amount of vehicle details for high-resolution aerial images were proposed by Hinz (2004): A car is geometrically modelled by a 3D wire-frame representation containing front, windshield, roof and hood as substructures. Colour constancy between hood and roof is included in the model as radiometric feature. Based on sun direction and image orientation parameters a shadow projection is computed from the wire-frame model in order to adapt the expected saliency of extracted edges.

3. ADAPTATION OF OBJECT MODELS CONSIDERING LOCAL CONTEXT OBJECTS

In this section a new concept for the automatic adaptation of image analysis object models considering local context to a lower image resolution is introduced and explained in detail. At first, several possibilities for the adaptation concept are discussed leading to the developed algorithm. Secondly, the adaptation method and its prerequisites regarding the modelling of local context and the incorporation of local context into the extraction strategy are described.

3.1 Adaptation concept

The exact position of local context objects in an individual scene is usually unknown beforehand. Roads, for instance, can be flanked by trees at their margin, but it is unknown where and if there are trees. Vehicles can even occur on all parts of a road, anywhere. While for the instance net of a particular scene the exact relative location of the objects is known, the concept net of the primary object needs to be very flexible with regard to the relative position of potential local context due to this location variability of local context. For our approach a scale-dependent adaptation of the concept net is desired. However, the behaviour of objects during scale reduction greatly differs for different relative object positions. For example: while for object *A* and *B*

lying relatively close together a scale event may occur for a certain resolution reduction ΔR , another object constellation of A and B with larger distance between each other may experience a different scale event or no scale event at all for the same change of resolution ΔR . Thus, the scale behaviour of a combination of objects strongly depends on the relative position of the objects, making an unambiguous prediction of the scale behaviour for local context objects with the primary object impossible. The result of the scale-dependent adaptation would be a number of different concept nets in the target resolution – one for each constellation. Hence, for a complete scale behaviour analysis an explosion in the number of target resolution object models can occur, making it unusable for real-world applications.

Nevertheless, several solutions regarding scale behaviour prediction for this variability problem are possible:

1. Highest probable constellations approach: The scale behaviour of the local context objects together with the primary object is predicted only for certain (a few) constellations with highest probability. The benefit is the exact prediction of the scale behaviour under consideration of the surrounding primary object, but the variability of the position of local context objects is not fully considered in the adaptation. This location constraint poses a contradiction to the inherent function of local context objects, as their occurrence and their resulting implications on the primary object are naturally extremely variable by their definition.
2. Statistics approach: The probability functions for the position of the individual local context objects are modelled from experience or derived by statistical investigations from training data. The scale behaviour for each constellation of objects is then predicted with the respective probability for their occurrence. The result is a set of concept nets in target resolution with statistical information.
3. Separation approach: The scale behaviour of the primary object and the local context objects is analysed separately. Thereby, the variability of the context object position is preserved on the cost of an exact combined scale behaviour prediction, because no embedding into the primary object takes place during the scale change prediction. The relation to the primary object is thereby partially neglected. However, most image analysis algorithms for context object extraction have to work in diverse surroundings due to the context object location variability making the consideration of all details of the primary object impossible. Thus, in most cases this approach will be sufficient for the adaptation of image analysis models for primary objects and context objects to a lower resolution.

Our adaptation concept uses the separation approach because of the advantage of preserving the flexibility of the context relation in the model, which we consider as a crucial aspect for realistic object extraction. Although the statistics approach represents a more exact solution, the analysis is not straightforward and its research effort would not be in accordance with our current resources.

3.2 Context Modelling

In our adaptation approach the landscape object one is interested in (in the following denoted as *primary object*) and the local context objects are regarded as separate objects. The local context objects are not incorporated into one single joint object model for the primary object and context objects, but

separated models for the primary object (road in our example) and one for each context object are assumed.

The connection between the primary object model and the context model(s) is defined by the relation *optional*, i.e. the local context object may be but does not have to be present. The context object can occlude or cast a shadow on object parts of the primary object, represented by the context relations *occludes* and *casts shadow on*.

Additionally, the incorporation of more than one context object is possible. The priority of several context objects for the extraction process can be defined by an ordering scheme for the individual context objects with their relations to the primary model. Thereby, a sequence for applying the context object models to the image for refining the extraction result is defined.

3.3 Extraction strategy

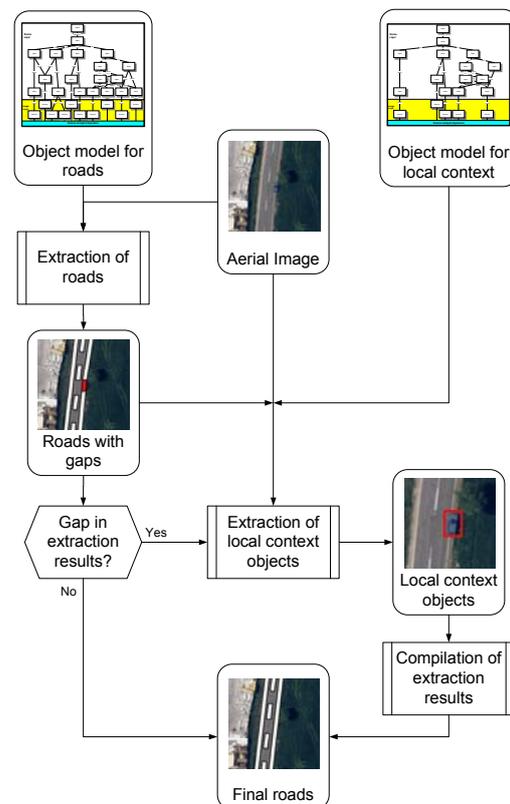


Figure 2. Extraction strategy with local context objects

It is assumed for our adaptation approach that the strategy for the extraction of a landscape object considering local context follows a two-step process. Figure 2 depicts the extraction strategy with local context objects. At first, the extraction of the primary object with its respective object model without considering local context is carried out. If this extraction partly fails, then we try to explain the emerged gap in the extraction result. The reason for the failure can be a local context object. Thus, an extraction of the local context using the corresponding object model(s) is carried out at the image area, where the gap occurred. If a local context object is found in the extraction gap area, the primary object is regarded as completely existent at this location, since the extraction failure can be explained by local context. If none of the context objects is verified in the gap, the extraction of the primary object remains disconnected in that area.

Local context objects are employed here as a means for gap filling in the extraction result for the primary object. The only allowed context relation is *optional*. Simple modelling for the context objects itself is in most cases sufficient, but a more detailed model, such as the one of Hinz (2004), for the context objects can also be incorporated into the extraction strategy, if necessary.

3.4 Adaptation method

The adaptation to the target resolution is carried out for the primary object model and the local context object model(s) independently applying the adaptation algorithm presented in (Heuwold et al., 2007), cf. section 2.1.

The scale behaviour prediction is carried out using analysis-by-synthesis in the scale change analysis stage of the adaptation algorithm. This prediction encompasses all four possibly occurring scale events during resolution change and the attributes for the object parts as defined in the initial object model. Scale events are predicted by comparing blob support regions and extrema in initial and target resolution. Occurring scale events can have a severe impact on the resulting net, as they dictate the number of remaining object parts in target resolution. The object part attributes for shape and radiometry are predicted from the characteristics of the blob support regions in the target resolution. The scale behaviour prediction of attributes is essential for the adaptation of feature extraction operators, since the predicted attributes can serve as adapted operator parameters. As pointed out in section 2, feature extraction operators for local context objects usually rely on shape properties as well as radiometric characteristics. The use of colour implies the scale behaviour analysis of more than one intensity channel for the image. However, in the analysis-by-synthesis a simulation of the scene with maximum contrast for a single channel is sufficient, if no restriction for colour is implied by the model. It is irrelevant in which channel the maximum blob contrast occurs, but the maximum contrast complies with maximum existence in scale space. The maximum contrast object might still exist in coarser scale, while an object with less contrast might not be extractable any more. Note however, that this approach is only permitted for objects with *optional* context relation.

As long as the context object is still predicted to be extractable in the target resolution, the type of context relation between primary object and context objects remains unchanged.

4. ADAPTATION EXAMPLE

This section gives an example for the adaptation of a high-resolution object model for a road (dual carriageway), which incorporates freight vehicles (trucks) as possible local context. The adaptation is demonstrated for an initial image resolution $R_0=0.03\text{m}$ to a target resolution $R_t=0.20\text{m}$. The object models are applied to a greyscale aerial image captured from Hanover region, Germany (see Figure 3).



Figure 3. Example image showing a road with local context

4.1 Model

The road model including the context relations for the initial high resolution $R_0=0.03\text{m}$ is given in Figure 4. The road model corresponds to the model in Heuwold (2006). The relation to local context objects is given by *optional occlusion* or *optional shadow*. As all road markings have a relation to local context, they all can be subject to occlusion or shadow by a vehicle.

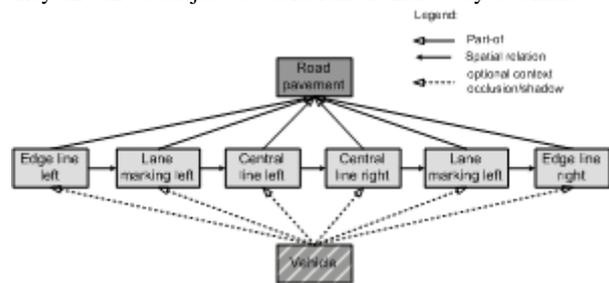


Figure 4. Road object model with context relations for 0.03m/pixel image resolution

The initial object model for the vehicle, depicted in Figure 5, is relatively simple and not as detailed as e.g. the one of Hinz (2004), but despite its simplicity it should suffice for demonstrating the capability of the adaptation concept. A large freight vehicle is modelled comprising two object parts as rectangles for front and trailer. The attribute values are given in pixels.

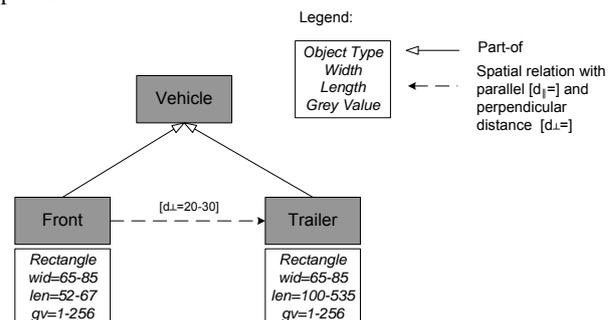


Figure 5. Freight vehicle object model for $R_0=0.03\text{m/pixel}$

The road model uses the operators for road markings and roadway, which are described in detail by Heuwold (2006). The operator for the vehicle parts uses region growing for segmentation. Suitable region candidates are selected by shape (convexity, compactness, excentricity, area) and radiometric properties (Meyer, 2006). The extraction results of the road model applied to the example image in high resolution yields an extraction gap for the road marking *lane marking left* at the context area. Figure 6 displays the results of the road extraction and of the vehicle extraction.

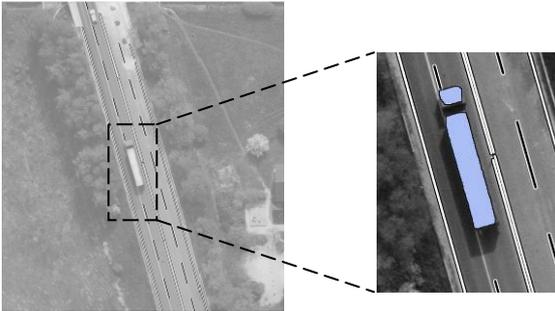


Figure 6. Road extraction result (black/white) superimposed on example image in 0.03m/pix and region of interest (ROI) for local context extraction

4.2 Adaptation Process

Both object models for road and vehicle are adapted separately. The adaptation of the road model was presented by Heuwold (2006). The adaptation process for the vehicle object model is explained in the following. The intermediate steps of the analysis-by-synthesis for the scale behaviour prediction of the vehicle are displayed in Figure 7. Note that the steps of the analysis-by-synthesis are depicted here in a single synthetic image for better illustration although there is no interaction between the vehicle parts.

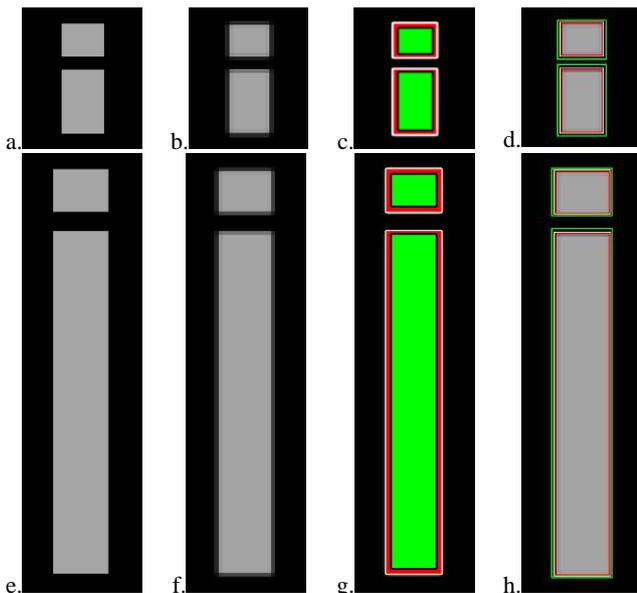


Figure 7. Intermediate steps of analysis-by-synthesis for minimum (upper row) and maximum constellation (lower row): synthetic image initial resolution L_0 (a.+e.), synthetic image target resolution L_{Rt} (b.+f.), blob extrema E_0 (red/white) and E_t (green/black) superimposed on L_{Rt} (c.+g.), blob support regions $Supp_0$ (red/white) and $Supp_i$ (green) on L_{Rt} (d.+h.)

No interaction of the vehicle parts takes place for the example target resolution $R_t=0.2m$, since they are situated too far apart for this resolution change. In order to predict the scale behaviour for the vehicles the minimum and maximum configuration with regard to the defined attribute ranges from the initial resolution is to be analysed. The scale behaviour simulation for the vehicle is carried out with the maximum contrast, since the initial object model does not contain any restriction with regard to grey value, i.e. all values are possible. Thereby, the most “optimistic” case is simulated, leading to the

maximum lifetime in scale-space. As explained in section 3.4 this procedure can be chosen, because the vehicle has an *optional* context relation. Neither for the minimum nor the maximum constellation a scale event takes place, but the attributes of the vehicle parts have to be changed for the target resolution. The attributes of the nodes and the spatial relation as well as the shape parameters for the feature extraction operators of the vehicle parts are automatically derived from the blob support regions in target resolution $Supp_t$.

4.3 Adaptation Result

The adapted road model is depicted in Figure 8. Details with attribute values of the nodes and spatial relations can be found in Heuwold (2006). The adapted vehicle object model is displayed in Figure 9 showing all attributes of the individual nodes and the spatial relations. Applied to the example image in target resolution the adapted road object model of Figure 8 also suffers an extraction failure in the area with the vehicle. This gap can again be filled with the adapted vehicle object model (Figure 9). Figure 10 shows both the road extraction gap and the vehicle extraction.

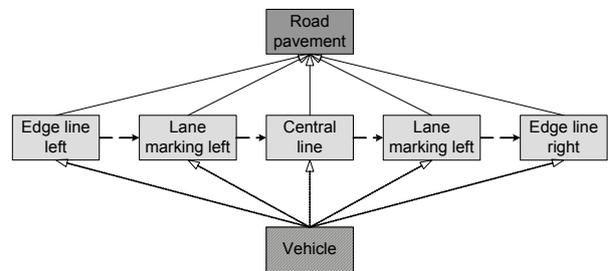


Figure 8. Adapted road object model for target resolution $R_t=0.2m/pix$

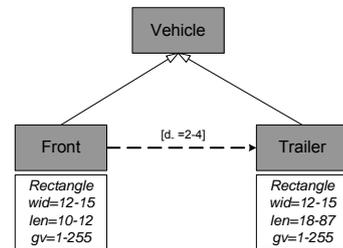


Figure 9. Adapted vehicle object model for target resolution $R_t=0.2m/pix$

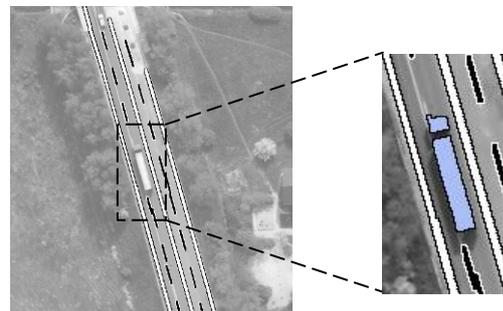


Figure 10. Road extraction failure for example image in $R_t=0.2m/pix$ and region of interest (ROI) for local context extraction

5. CONCLUSIONS

In this paper a method was presented for the automatic adaptation of image analysis object models, which consider local context objects, to a coarser image resolution. An example for the adaptation of a road model with vehicle as local context

object to a coarser resolution was given. Despite the simplicity of the context object model, the example indicates the capability of the adaptation method, as the adapted models successfully extract both the road and the vehicle in target resolution. Although the adaptation is demonstrated here only for a certain type of vehicles, other local context objects, such as trees or buildings can be incorporated in the same way by replacing the vehicle model with their respective model(s).

In order to preserve the flexibility of the position of local context objects, the scale behaviour for the landscape object of interest and the context object is predicted separately. Due to this separation, the scale behaviour prediction is not exact for the local context, since no scale events between the primary object and the local context objects or among the context objects can be detected. However, manually created object models for a certain image resolution and the corresponding feature extraction operators of local context objects also cannot consider possible scale events of local context due to their extreme variability of location. Against this background, the presented separate adaptation approach appears reasonable. Nevertheless, a holistic scale behaviour approach with statistical prediction including both the landscape object of interest and the local context could improve the presented more pragmatic solution.

ACKNOWLEDGEMENTS

This study has been funded by Deutsche Forschungsgemeinschaft under grant HE 1822/13. The project is part of the bundle "Abstraction of Geoinformation in Multi-Scale Data Acquisition, Administration, Analysis, and Visualisation".

REFERENCES

- Baumgartner A., Eckstein W., Mayer H., Heipke C., Ebner H., 1997. Context-supported road extraction. In: Grün A., Baltsavias E., Henricsson O. (Eds.), *Automatic Extraction of Man-made Objects from Aerial Images (II)*, Birkhäuser, Basel, pp. 299-308.
- Butenuth, M., Straub, B., Heipke, C., and Willrich, F., 2003. Tree Supported Road Extraction from Aerial Images Using Global and Local Context Knowledge. *LNCS*, 2626, pp. 162-171.
- Gerke, M., 2006. Automatic Quality Assessment of Road Databases Using Remotely Sensed Imagery. *DGK, Series C*, No. 599, Munich, 103 p.
- Heuwold, J., 2006. Verification of a methodology for the automatic scale-dependent adaptation of object models. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXVI, Part 3, pp. 173-178.
- Heuwold, J., Pakzad, K., Heipke, C. 2007. Scale behaviour prediction of image analysis models for 2D landscape objects. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXVI, Part 3/W49A, pp. 43-49.
- Hinz, S., and Baumgartner, A., 2003. Automatic extraction of urban road networks from multi-view aerial imagery. *ISPRS Journal of Photogrammetry & Remote Sensing*, 58, pp. 83-98.
- Hinz, S. 2004. Automatische Extraktion urbaner Straßennetze aus Luftbildern. *DGK, Series C*, No. 580, Munich, 131 p.
- Koenderink, J., 1984. The Structure of Images. *Biological Cybernetics*, 50(5), pp. 363-370.
- Lindeberg, T., 1994. *Scale-Space Theory in Computer Vision*. Dordrecht, The Netherlands, Kluwer Academic Publishers, 423p.
- Liu, G., Zhuge, F., Gong, L., and Haralick, R., 1999. Vehicle Detection in Aerial Imagery and Performance Evaluation. Technical Report, University of Washington, Seattle, WA.
- Mayer, H. and Steger, C., 1998. Scale-space events and their link to abstraction for road extraction. *ISPRS Journal of Photogrammetry & Remote Sensing*, 53(2), pp. 62-75.
- Mayer, H., 2008. Object extraction for photogrammetric image analysis. *ISPRS Journal of Photogrammetry & Remote Sensing*, 63(2), pp. 213-222.
- Meyer, H., 2006. Extraktion von Fahrzeugen als Kontextobjekte in optischen Bilddaten. Bachelor thesis, University of Hannover, Germany, 73p., unpublished.
- Stilla, U., Michaelsen, E., Soergel, U., Hinz, S., and Ender, J. 2004. Airborne Monitoring of Vehicle Detection in Urban Areas. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXV, Part B3, pp. 973-979.
- Tönjes, R., Growe, S., Bückner, J. and Liedtke, C.-E., 1999. Knowledge Based Interpretation of Remote Sensing Images Using Semantic Nets. *Photogrammetric Engineering & Remote Sensing*, 65(7), pp. 811-821.
- Witkin, A., 1983. Scale Space Filtering. In: *Proceedings of the 8th Int. Joint Conference on Artificial Intelligence*, Karlsruhe, pp. 1019-1022.
- Zhang, C., 2004. Towards an operational system for automated updating of roaddatabases by integration of imagery and geodata. *ISPRS Journal of Photogrammetry & Remote Sensing*, 58, pp. 166-186.
- Zhao, T., and Nevatia, R., 2003. Car detection in low resolution aerial images. *Image and Vision Computing*, 21, pp. 693-703.