Automatic Adaptation of Image Analysis Models for 2D Landscape Objects to a Coarser Image Resolution [1]

Janet Heuwold, Kian Pakzad & Christian Heipke, Hannover

Keywords: Multiresolution, Scale Space, Modelling, Image Analysis, Interpretation

Summary: This paper presents a new methodology for the automatic adaptation of image analysis object models for the extraction of 2D landscape objects to a lower image resolution. The object models are represented in form of semantic nets. The developed adaptation method includes a prediction of the object's behaviour in linear scale-space using analysis-by-synthesis. The scale behaviour prediction takes into account all scale events possibly occurring during 2D scale change of area-type object parts with arbitrary orientation. An example for the adaptation of an object model describing a road junction arm with road markings demonstrates the applicability of the methodology. Finally, conclusions point out potential enhancements of the method.

Zusammenfassung: *Automatische Anpassung von Bildanalysemodellen für 2D-Landschaftsobjekte an eine niedrigere Bildauflösung*. Dieser Beitrag beschreibt eine neue Methode für die automatische Anpassung von Bildanalyse-Objektmodellen zur Extraktion von 2D-Landschaftsobjekten an eine niedrigere Bildauflösung. Die Objektmodelle sind als semantische Netze repräsentiert. Die entwickelte Methode zur automatischen Anpassung nutzt Analyse-durch-Synthese zur Prädiktion des Verhaltens des Objektes im linearen Skalenraum. Die Prädiktion des Skalenverhaltens berücksichtigt sämtliche Skalenereignisse, die bei der 2D-Skalenänderung von flächenhaften Objektteilen mit beliebiger Orientierung auftreten können. Die Methode wird an einem Beispiel für die Anpassung eines Objektmodells eines Straßenarmes im Kreuzungsbereich demonstriert. Der Beitrag schließt mit Schlussfolgerungen für potenzielle Verbesserungen der Methode.

1 Introduction

The appearance of landscape objects varies in aerial or satellite images of different resolution. Hence, for knowledge-based object extraction in images of different resolution, different models describing the objects are usually required. The objective of this paper is to introduce a new methodology for the automatic adaptation of image analysis models for object extraction to a lower image resolution. The models for object extraction are represented as semantic nets (Tönjes et al. 1999). The previously developed method for the automatic adaptation of object models consisting of linear parallel object parts (Pakzad & Heller 2004) is only suitable for simple landscape objects such as roads, which can be described exclusively by parallel line-type objects[2]. As a result the pre-

[2] In our model *landscape objects* or *objects* are decomposed into *object parts*, which are the smallest entities in our model. In many cases, the extraction of an object is carried out through the extraction of its object parts. Note, however, that we extract the complete objects in cases where no decomposition is foreseen in the model. In the example described later the object is "Road junction arm" and the object parts are the road markings (e. g., "Arrow").

diction of scale behaviour is simplified to a 1D problem, since an analysis of the cross-section is in principal sufficient. However, for modelling of a road network, more complex roads (as present in junction and urban areas) or other lansdcape objects an analysis of the scale behaviour of 2D objects is necessary. In the 2D case, the scale behaviour of the objects is more complex. Particularly challenging is the prediction of scale events that may occur during scale change for 2D objects. This paper presents a new adaptation process for image analysis models consisting of a combination of 1D and 2D object parts.

Since the appearance of objects often severely changes between different image resolutions, the extraction model for objects in lower resolution is to be altered by the adaptation method. The central problem is the prediction of the scale behaviour of objects and object parts. The second core issue is the automatic processing of the given and adapted semantic nets.

We use linear scale-space theory for the prediction of the object's scale behaviour. The linear or Gaussian scale-space as defined first by WITKIN (1983) and KOEN-DERINK (1984) is created by convolution of an image $L(x, y)$ with the Gaussian $g(x, y; t)$ of varying width. Thereby, a family of signals is derived depending only on a single scale parameter t corresponding to the square of the Gaussian standard deviation σ, $t = \sigma^2$, while x, y denote the image coordinates. For details concerning the characteristics of linear scale-space see FLORACK et al. (1994). The analysis of image structure in different scales is also referred to as deep structure (KOENDERINK 1984). LINDEBERG (1993 and 1994) proposed a blob detection algorithm for the automatic scale behaviour prediction of 2D object models, which we use in our methodology.

In the literature some other approaches dealing with scale behaviour analysis of 2D landscape objects from remote sensing data can be found, e. g., scale events for buildings were analysed in morphological scale-space by FORBERG & MAYER (2002); the scale-space primal sketch was used by HAY et al.

(2002) for the scale behaviour description of whole landscapes as complex systems. However, the prediction of the scale behaviour of complete 2D object models for image analysis and their adaptation to a lower image resolution is new.

This paper is organised as follows: The next section gives an overview of the adaptation process including strategy and adaptation method developed for object models consisting of line-type parallel (1D) and area-type (2D) object parts. An example for the adaptation of a model for a junction area with road symbol markings is outlined in section 3. The paper finishes with conclusions and an outlook for future work.

2   Adaptation Process

2.1 Adaptation strategy

The strategy for the automatic scale-dependent adaptation of object models comprising linear-parallel (1D) and area-type arbitrarily oriented (2D) objects follows a process in three main steps (cf. Fig. 1) – decomposition, scale change analysis, and fusion. Based on the type of object parts in the fine scale a decision is made whether the 1D or 2D scale change analysis method is to be applied. If there are only linear parallel object parts in the object model to be adapted, the 1D scale-dependent adaptation can be used; otherwise the more sophisticated adaptation for 2D objects is to be carried out.

The first stage of the automatic adaptation decomposes the given object model for the fine scale into object parts that can be analysed separately regarding their scale behaviour. The decomposition takes into account the mutual influence of nearby objects when scale changes – denoted as interaction. The possibility of interaction depends on the spatial distances of the object parts to each other. If the distances fall below a particular value, which depends on the quantity of scale change, interaction occurs. In this case, the respective object parts are analysed simultaneously in groups in the following scale change analysis phase.
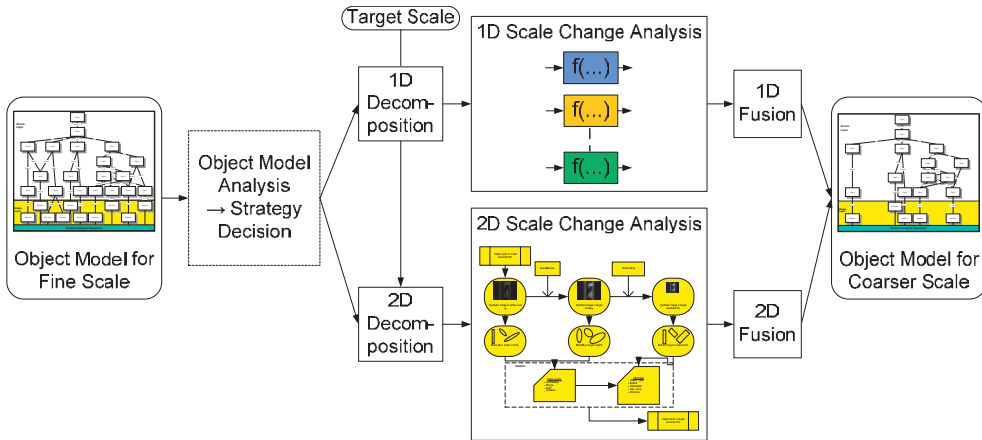
Fig. 1: Strategy for adaptation process (for details see text).

Scale change models predict the appearance and extractability in the target resolution for each interacting group or for single non-interacting object parts resulting from the decomposition. The prediction uses analysis-by-synthesis, simulating the appearance of the object parts in synthetic images of the target resolution.

In the last stage, fusion, all predicted object parts are recomposed back into a complete object model that is suitable for the extraction of the object in the target resolution.

2.2 Adaptation method

2.2.1 Decomposition

In order to facilitate the scale behaviour prediction of the object in the scale change models, a separate analysis of the individual parts of the modelled object is desirable. However, during scale change adjacent object parts may influence each other's appearance. This is the case, when they lie close to each other in the target resolution. This condition is checked by looking at their spatial distance in the object model. Object parts, which influence each other, need to be analysed together and form an *interaction group* in the successive scale change analysis stage. All other object parts that are not subject to interaction can be treated as single

object parts in the scale behaviour analysis. Depending on the size of the Gaussian kernel associated with the target scale σ_t an *interaction zone* is formed around the object parts. For more details concerning the interaction zone see HEUWOLD et al. (2007).

2.2.2 Scale change analysis for 2D

Scale change models predict the scale behaviour for single object parts and interaction groups. The prediction is carried out in an analysis-by-synthesis procedure, analysing the objects in synthetic images in original scale, and target resolution (the target resolution includes downsampling, the target scale does not). The result is a description of the appearance of the object parts in the given target resolution in terms of attributes for the nodes and edges of the semantic net. As some simplifications carried out for linear parallel object parts are not applicable to 2D object parts, the previously developed method for linear parallel object parts is no longer appropriate for 2D scale change analysis. Therefore, a new workflow for the analysis-by-synthesis process was developed and is depicted in Fig. 2.

First, from the specifications of the object parts' appearance in the nodes of the given semantic net for the high resolution a synthetic image is created for each object part or interaction group to be analysed. This in-
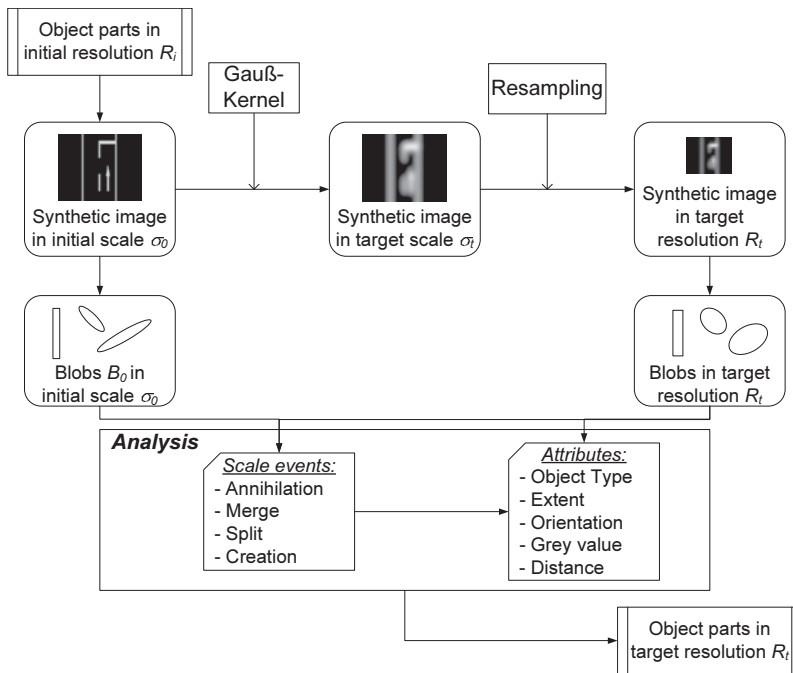
Fig. 2: Analysis-by-synthesis process for 2D adaptation.

itial image L_0 simulates the appearance of the object parts in the original scale σ_0. In a second step, the initial image is transferred to the target scale σ_t by convolution with the respective discrete Gaussian into the target scale image L_t. Since the object's appearance and extractability can vary between the target scale image and the image in the corresponding target resolution, the target scale image L_t is subsequently down-sampled to the corresponding lower resolution R_t. Both the analysis of the possibly occurring scale events and of the attributes describing the appearance of the object parts in the low resolution are carried out in this image. Although the resampling does not result in exactly the same image as a remote sensing sensor of a lower resolution produces, the process proved to be generally sufficient for a simulation of a remote sensing image (HEUWOLD 2006). In order to obtain an exact simulation, one would have to model the individual sensor characteristics and apply to the target scale image. Considering the amount of different imaging sensors

(aerial and satellite), this approach does not seem practicable.

Scale event prediction

During scale change so-called scale events may occur. There are four types to be distinguished: Annihilation, Merging, Split and Creation. In contrast to the one-dimensional case, where only the first two events need to be considered, all four different scale events may occur in 2D images. The prediction of scale events of interacting object parts is thus not as straight forward and requires a more sophisticated approach to scale behaviour prediction than the previously presented method for linear parallel object parts.

The *scale-space primal sketch* was introduced by Lindeberg as an explicit representation of features in scale-space and their relations at different levels of scale (LINDEBERG 1993). The sketch was designed as a basis for the extraction of significant image features at stable scales for later processing

towards object extraction. Blobs serve as primitives of the scale-space primal sketch. *Grey-level blobs* are smooth image regions that are brighter or darker than the background and thereby stand out from their surrounding. By definition a grey-level blob $B(E)$ is a region of a scale-space image $L(x, y; t)$ associated with a pair of critical points (or regions in discrete scale-space) consisting of one local extremum E and one delimiting saddle S. The grey-level blob is a three-dimensional object with extent both in the spatial and the grey-level domain. The spatial extent of the blob is given by its *support region Supp(B)*.

At first, for the prediction of scale events blob detection is carried out in both the initial image L_0 and the target resolution image L_{Rt} using the sequential blob detection algorithm of LINDEBERG (1994). Blob linking between the initial image and the target resolution image is then carried out by matching blobs with intersecting support regions in original and target resolution. We assume that most blobs are not subject to a scale event during the scale change given by the specified target resolution. In the blob linking process we thus first try to establish a so-called *plain link* between the initial and target resolution for all blobs. Based on the set of support regions in initial resolution $Supp_0$ and the set of support regions in target resolution $Supp_{Rt}$ a plain link must fulfil the following condition (with m and n being the number of blobs in initial and target resolution):

Plain Link: One particular blob in initial resolution has one and only one direct correspondence in target resolution. The support region of a blob B_i in initial resolution $Supp_0(B_i)$ intersects the support region of a blob B_q in target resolution $Supp_{RT}(B_q)$. All other blob support regions in target resolution must not intersect $Supp_0(B_i)$.

$$\exists^{=1} q(Supp_0(B_i) \cap Supp_{Rt}(B_q) \neq 0),$$
$$i \in \{1 \dots m\}, q \in \{1 \dots n\} \qquad (1)$$

If a plain link cannot be established for all blobs, scale events must have occurred. The types of scale events must then be resolved automatically. We set up the following postulates for the occurrence of blob events during scale change:

Annihilation: One particular blob in initial resolution has no correspondence in target resolution. The set of support regions in target resolution $Supp_{Rt}$ is empty at the position of a blob support region in initial resolution $Supp_0(B_i)$.

$$\exists i(Supp_0(B_i) \cap Supp_{Rt} = 0), i \in \{1 \dots m\} \quad (2)$$

Merging: Two (or more) initial blobs have one and the same blob as correspondence in target resolution. The support regions of at least two initial blobs $Supp_0(B_i)$ and $Supp_0(B_j)$ intersect the support region of one and the same blob in target resolution $Supp_{Rt}(B_q)$.

$$\exists q((Supp_0(B_i) \cap Supp_{Rt}(B_q) \neq 0)$$
$$\wedge (Supp_0(B_j) \cap Supp_{Rt}(B_q) \neq 0)),$$
$$i, j \in \{1 \dots m\}, q \in \{1 \dots n\}, i \neq j \qquad (3)$$

Split: One initial blob has two (or more) blobs as correspondence in target resolution. The support region of an initial blob $Supp_0(B_i)$ intersects the support regions $Supp_{Rt}(B_q)$ and $Supp_{Rt}(B_s)$ of at least two blobs in target resolution.

$$\exists q \exists s((Supp_0(B_i) \cap Supp_{Rt}(B_q) \neq 0)$$
$$\wedge (Supp_0(B_i) \cap Supp_{Rt}(B_s) \neq 0),$$
$$i \in \{1 \dots m\}, q, s \in \{1 \dots n\}, q \neq s \qquad (4)$$

Creation: One particular blob in target resolution has no correspondence in initial resolution. The set of support regions in initial resolution $Supp_0$ is empty at the position of a blob support region in target resolution $Supp_{Rt}(B_q)$.

$$\exists q(Supp_0 \cap Supp_{Rt}(B_q) = 0), q \in \{1 \dots n\} \quad (5)$$

Based on these conditions the scale events that have occurred during scale change are inferred. The procedure is described more detailed in HEUWOLD et al. (2007).

It should be noted that blobs can be composed of several individual object parts that

are adjacent to each other. Hence, the number of blobs in initial or target resolution does not necessarily equal the number of nodes in the semantic net for the respective resolution. However, as we know the composition of blobs in initial scale from the synthesis process, we can separate the individual object parts that compose the same blob in target resolution, if no scale events have occurred.

Attribute prediction

The attributes in the nodes specify the appearance of an object part in the image. The values of the attributes in the nodes of the adapted semantic net for the lower resolution are therefore analysed in the synthetic target resolution image L_{Rt}. Blobs are assumed to correspond to the individual object parts. Hence, the number of blobs in the target resolution equals the number of nodes in the semantic net for the target resolution. For each blob in the target resolution the following attributes are derived: *object type* (e. g., line or pattern), *spatial extent* (width and length), *grey value*, and *orientation*.

2.2.3 Fusion

The fusion is the last stage of the automatic adaptation process. All nodes remaining after scale change including their attributes representing the object parts in the given target resolution are compiled to a complete semantic net.

The hierarchical relations between the nodes remain unchanged as long as no scale event occurred. In case of Annihilation, the respective node is simply omitted. For merged blobs only a single *part-of* relation remains. For the Split and Creation events new *part-of* relations are introduced into the respective hierarchy level.

The type of the spatial relation (parallel or perpendicular distance) stays unaffected. However, the distances between the object parts are to be adapted. The adapted distance values are derived from the position of the borders of the blob support regions $Supp_{Rt}$ in target resolution.

3 Adaptation Example

In order to demonstrate the applicability of the adaptation methodology, this section gives an example for the automatic adaptation of a high-resolution object model to a coarser image resolution. A semantic net for the extraction of a road junction arm with 2D symbols (arrows and stop lines) serves as given model for high image resolution (3–5 cm per pixel). We chose $\sigma_t = 8$ as target scale – corresponding to an approximate target resolution of $R_t \approx 0.8$ m. This junction model can be seen as a part of an object model for a road network consisting of roads and junction arms (cf. Fig. 3). The node for an adjacent *Road*, for instance, can be represented by the model given in HEU-WOLD (2006), which consists of linear parallel objects and can be adapted with the 1D adaptation strategy.

The road junction in our example consists of a number of road junction arms with a dashed central line. These arms meet in the junction area. They contain lane markings and additional traffic markings (in our example direction arrows and stop lines).
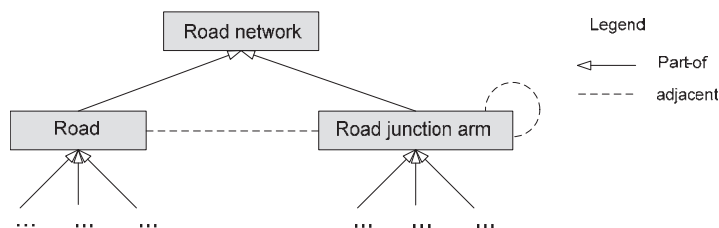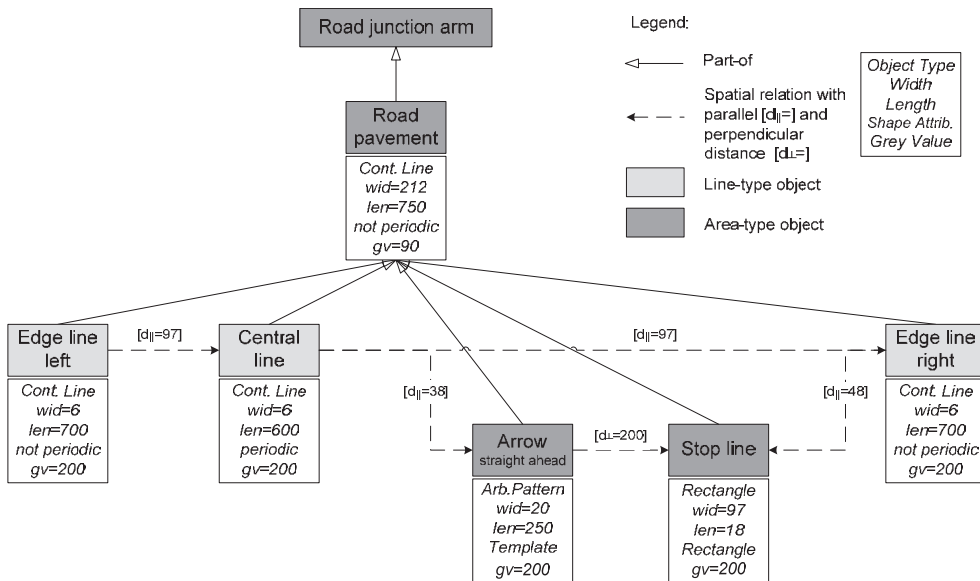

Fig. 3: Object model for road network.

Fig. 4: Object model for junction arm in a resolution of 0.03−0.05 m/pixel.

This means the road junction arms represent the part close to the roads before the junction centre, where the roads contain additional road markings. The width of the road junction arm is modelled to be constant. The 2D position of the object parts is given by distances (in number of pixel in the respective resolution) between them in two perpendicular directions. The model also contains information concerning the extent of the object parts. The model for road junction arms for the high resolution is depicted in Fig. 4.

An essential part of an object model for image analysis is also a set of image analysis operators. They represent the procedural knowledge needed to extract particular object parts from images. The image analysis operators search for the particular parts in the image, which are represented as nodes in the object model. Therefore, the nodes of the semantic net are connected to the corresponding image analysis operators. Different image analysis operators are assigned to different types of object parts. Two object types were defined for area-type objects: *Geometric Shape* (e. g., *Rectangle*) for simple object parts and *Arbitrary Pattern* for more complex ones. The shape of the latter kind

of object part is defined by templates. The operators for *Arbitrary Pattern* use cross-correlation matching with provided templates, whereas the line-type objects and the *Rectangle* use the road marking operators based on differential geometry developed in HEUWOLD (2006).

As mentioned before, the scale behaviour prediction of the junction example is carried out in the scale change analysis stage by analysis-by-synthesis. Due to the width of the discrete Gaussian used for the creation of the target scale image L_t, all object parts of the bottom level except *edge line left* form an interaction group. Hence, the scale behaviour of these object parts is analysed jointly. Fig. 5a) depicts all simulated object parts in a synthetic image, while Fig. 5b) shows the down-sampled filtered image in target resolution L_{Rt}. The number of blobs in initial resolution (5) differs from the number of blobs in target resolution R_t (4), suggesting the occurrence of a scale event: a Merging event is detected for a part of the central line and the direction arrow. Here, the corresponding blobs have merged. The results of the blob detection in the synthetic images of the initial and target resolu-
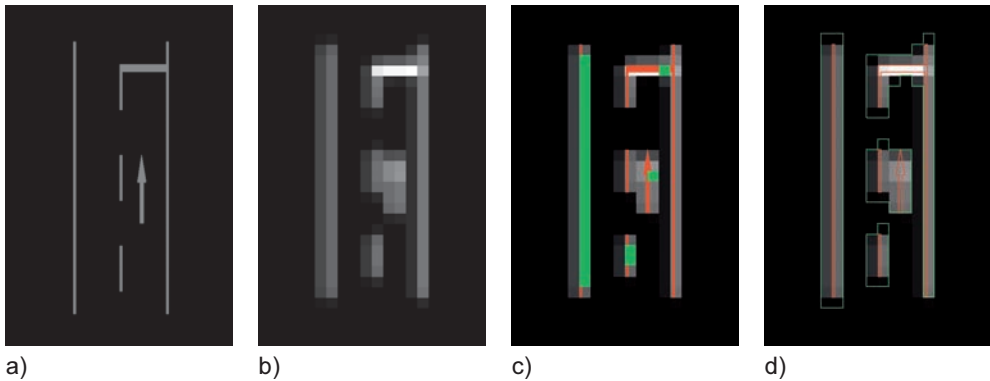
a)　　　　　　　b)　　　　　　　c)　　　　　　　d)

Fig. 5: Blob detection results: initial blob features and target blob features superimposed on synthetic images; a) initial image L_0, b) target resolution image L_{Rt} (grey-value stretched and enlarged), c) extrema E_0 (red), E_{Rt} (green) on target resolution image L_{Rt}, d) support regions $Supp_0$ (red), $Supp_{Rt}$ (green) on L_{Rt}.

tion, illustrated in Fig. 5 d), reflect the postulated condition for Merging events – the support regions $Supp_{RT}(B_q)$ in target resolution R_t intersect the two support regions $Supp_0(B_1)$ and $Supp_0(B_2)$ of the merged blobs from initial resolution. Fig. 5c) depicts the position of the blob extrema in initial and target resolution.

Although there are only four blobs in the target resolution image L_{Rt}, the resulting object in the target resolution R_t consists of

six individual object parts. Because the upper part of the dashed *central line*, the *stop line* and the *edge line right* touch each other, they form one single blob – not only in target resolution but already in initial resolution. For all resulting six object parts in the target resolution the node attributes are derived from the target resolution image L_{Rt}, as illustrated in Fig. 5d): there are four *Continuous Lines* (*edge line left*, *edge line right* and two parts of the dashed *central line*), one
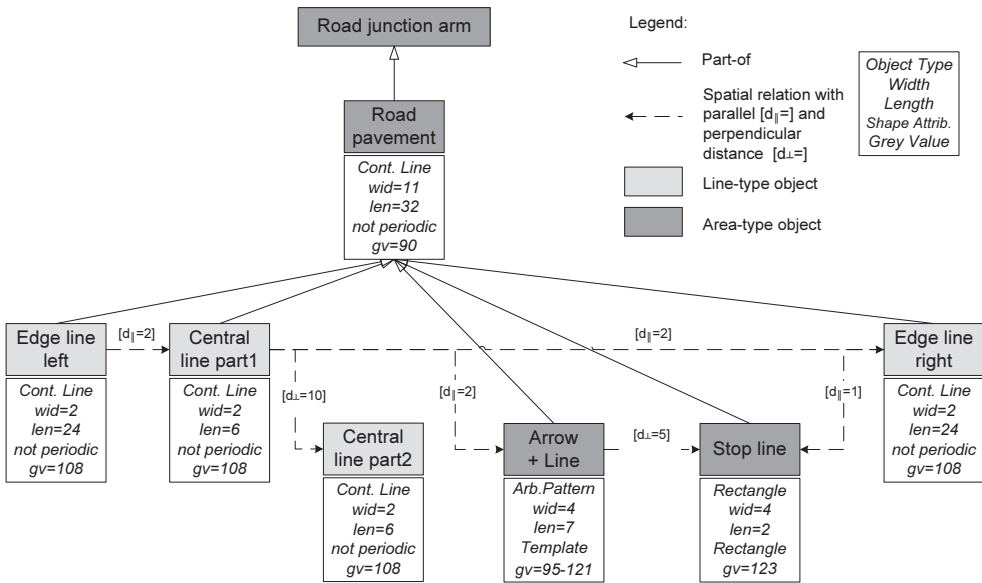

Fig. 6: Adapted example object model in target resolution $R_t \approx$ 0.8 m.

Rectangle (*Stop Line*) and one *Arbitrary Pattern* (*Arrow + Line Part*) for the merged part of the central line and the direction arrow; their extents are given by the width and length of a fitted rectangle of their blob support regions $Supp_{Rt}$; the grey values are determined from the blob brightness inside its predicted extent (mean value for area-type object parts, maximum value in cross-section for line-type object parts). The distances being attributes of the spatial relations (edges) are derived from the shortest distance of the blob support regions for area-type object parts.

The adapted object model of the road junction arm for the target resolution $R_t \approx 0.8\,\mathrm{m}$ is illustrated in Fig. 6. Most line-type object parts changed only their attributes (width and contrast); whereas one part of the previous dashed central line was subject to a scale event and was merged with the adjacent area-type arrow symbol during scale change. The resulting object is another arbitrary pattern, which is to be extracted by pattern matching. The operator for this new object part can use a template that is derived from the support region $Supp_{Rt}$ of the merged blob in the target resolution image L_{Rt} for cross-correlation matching.

4 Conclusions and Outlook

In this paper a new methodology was presented for an automatic adaptation of image analysis object models, created for a specific high resolution, to a lower resolution image. The modelled landscape objects can consist of both arbitrarily oriented line-type and area-type object parts. In order to adapt the representation of the objects, the scale behaviour of the object parts is analysed taking into account 2D scale events and changes in the object's appearance. Using an object model for a road junction area, the adaptation method is exemplarily described.

A general constraint of the described method is that interaction with other objects can always influence the objects' appearance in the target resolution; also other non-modelled objects in the vicinity of the modelled objects can influence the appearance. Therefore, strictly speaking the described method is only valid for a modelled object, if the neighbouring object is a homogeneous area and no other objects influence the adaptation. This means that the spatial distance to neighbouring object must be larger than the computed threshold for interaction (which depends on the amount of scale change). This is also true for the case that a modelled object can influence itself, because it is strongly curved (e. g., serpentine roads may cause adaptation problems). This condition forms a current limitation with regard to automatically adaptable object models using our methodology.

Our future work will therefore focus on the incorporation of relevant local context objects such as trees or buildings into the adaptation process in order to be able to deal with more realistic scenes. This will allow us to account for occlusions and shadows in the images. Due to the high complexity of the scale behaviour simulation for local context objects with unknown position, however, we will not process the road objects simultaneously with the local context objects, as the number of possibly resulting low resolution object models can easily become very high. Instead, we intend to consider the local context object as a separate object model, for which a scale change is derived independently. In cases where in the low resolution the road extraction fails, we will then try to explain this failure through the introduction of a building, a tree or their shadows.

The 2D concept presented in this paper is until now essentially limited to theoretical work. However, once extended this approach to automatic object model generation could prove useful in a number of applications: e. g., for multi-sensor image interpretation without the need to create more than one object model manually, as the models for sensors with lower resolution can be automatically derived; generally, the developed method for blob linking can also support many-to-many matching of spatial entities in different representations.

References

FLORACK, L., HAAR ROMENY, B., KOENDERINK, J. & VIERGEVER, M., 1994: Linear Scale-Space. – Journal of Mathematical Imaging and Vision **4** (4): 325–351.

FORBERG, A. & MAYER, H., 2002: Generalization of 3D Building Data Based on Scale-Spaces. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **34** (4): 225–230.

HAY, G., DUBÉ, P., BOUCHARD, A. & MARCEAU, D., 2002: A scale-space primer for exploring and quantifying complex landscapes. – Ecological Modelling **153** (1–2): 27–49.

HEUWOLD, J., 2006: Verification of a methodology for the automatic scale-dependent adaptation of object models. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **36** (3): 173–178.

HEUWOLD, J., PAKZAD, K. & HEIPKE, C., 2007: Scale Behaviour Prediction of image analysis models for 2D landscape objects. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **36** (3/W49A): 43–49.

KOENDERINK, J., 1984: The Structure of Images. – Biological Cybernetics **50** (5): 363–370.

LINDEBERG, T., 1993: Detecting Salient Blob-Like Image Structures and Their Scales with a Scale-Space Primal Sketch: A Method for Focus-of-Attention. – International Journal of Computer Vision **11** (3): 283–319.

LINDEBERG, T., 1994. Scale-Space Theory in Computer Vision. – Kluwer Academic Publishers, Dordrecht, The Netherlands, 423 p.

PAKZAD, K. & HELLER, J., 2004: Automatic Scale Adaptation of Semantic Nets. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **35** (B3): 23–28.

TÖNJES, R., GROWE, S., BÜCKNER, J. & LIEDTKE, C.-E., 1999: Knowledge Based Interpretation of Remote Sensing Images Using Semantic Nets. – Photogrammetric Engineering & Remote Sensing **65** (7): 811–821.

WITKIN, A., 1983: Scale Space Filtering. – 8th Int. Joint Conference on Artificial Intelligence: 1019–1022.

Address of the Authors:

Dipl.-Ing. JANET HEUWOLD, Dr.-Ing. KIAN PAKZAD, Prof. Dr.-Ing. CHRISTIAN HEIPKE, Leibniz Universität Hannover, Institut für Photogrammetrie und GeoInformation, D-30167 Hannover, Tel.: +49-511-762-19953 (Heuwold) 3893 (Pakzad), 2482 (Heipke), Fax: +49-511-762-2483, e-mail: heuwold | pakzad | heipke@ipi.uni-hannover.de