AUTOMATIC SCALE-DEPENDENT ADAPTATION OF VARIABLE OBJECT MODELS

J. Heuwold, K. Pakzad

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

ABSTRACT:

A methodology for the automatic adaptation of models for the extraction of landscape objects with a variable number of similar object parts to a coarser image resolution is proposed in this paper. A given object model, consisting of parallel line-type object parts and represented as a semantic net, can with these methods be fully automatically adapted for the extraction of that same object in another coarser resolution. The methods are developed exemplarily for the landscape object road. Roads can be modelled with a beforehand unknown number of lanes, allowing thereby the description of different types of roads by a single model. This paper also introduces *dynamic relations* for realizing the variability in the models. The adaptation utilizes the linear scale-space for the prediction of the appearance of the road in the lower resolution and also considers the possible occurrence of scale-space events during scale change.

The developed automatic adaptation methods are demonstrated using an example for a variable road model. Intermediate steps and results of the adaptation process are presented and discussed.

1. INTRODUCTION

The spatial resolution of available image data plays a decisive role for the creation of suitable models for the extraction of landscape objects in aerial and satellite images. In images of high spatial resolution more and different object details might be distinguishable, compared to a lower resolution image of the same scene. Not only the type of objects, but also the attribute values change with increasing ground pixel size. Therefore, the models describing the objects for the extraction process have to be modified according to the different appearance of the same object in images of another resolution.

In this paper a concept and methods are presented allowing the adaptation of a model for the extraction of landscape objects in high resolution images to lower resolution images automatically. The landscape objects are described here by semantic nets, which allow a simple intuitive generation in an explicit representation. These models are called "object models" in the remainder of this paper, although the models not only describe the object parts and their hierarchical and spatial relation, but also their appearance in images. In relation to the definitions of [Förstner93], the model type to be adapted here, integrates both the object model and image model.

Algorithms for the adaptation of object models consisting of line-type object parts, represented as semantic nets, to a lower resolution were already presented in detail [Pakzad&Heller04], [Heller&Pakzad05]. The adaptation has so far only been possible for line-type objects with a well-defined number and distances of object parts (here called "constant object models"). The methods introduced here extend these algorithms regarding the variability of the adaptable object models (then called "variable object models"). By allowing for a variable number of similar line-type object parts, the adaptable object models are not as restrictive regarding the type of road as the constant models with static relations, which were automatically adaptable so far. The strategy and the methods for the

adaptation of constant object models are only summarized in this paper. For a detailed explanation cf. [Heller&Pakzad05].

In section 2 the strategy of the adaptation process is explained. An overview about the previously developed methodology for the adaptation of line-type object models including constraints for the automatic adaptation process is also given in this section. Section 3 includes the definition of variability in object models and a detailed description of the newly extended methodology concerning variable object models. The methods are demonstrated using an example for the adaptation of a variable road model to a lower resolution in section 4. The paper concludes with a discussion and future prospects in section 5.

2. ADAPTATION STRATEGY AND CONSTRAINTS

The scale-dependent adaptation of object models follows a process in three steps. The applied strategy is depicted in Fig.1.
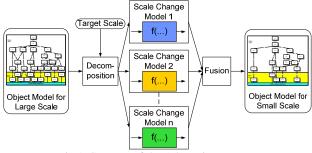

Fig.1: Strategy of the Adaptation Process

In order to facilitate the adaptation process, the given object model is at first decomposed into groups of object parts, which need to be examined together, as they influence each other during scale change due to its nearby location. This influence is dependent on the target scale of the adaptation.

For each of these groups, a scale change model then predicts the scale behaviour and the appearance of these object parts in the target scale. The adaptation process uses the linear scale-space [Florack94], [Lindeberg94] for the estimation of the scale behaviour of the object parts. Here, the attributes in the nodes of the semantic net describing the object parts in the target scale are predicted. The scale change models also consider possible scale-space events. For parallel line-type objects the events Annihilation and Merging are relevant.

In a last step, the predicted object parts are fused back to a complete object model, which is suitable for the adaptation of the same object in the coarser resolution corresponding to the target scale. While the hierarchical relations remain unaffected for all object parts, the spatial relations have to be modified to account for the change of the object parts' extent.

3. ADAPTATION PROCESS OF VARIABLE OBJECT MODELS

3.1 Variability in object models

The term "variable object models" is used here for models for the extraction of landscape objects with a variable number of similar object parts in the semantic net (only in the concept net). For roads for instance, several lanes can be described in the same model, while knowledge about the exact number of lanes is not required. Hence, different types of roads can be described with only one model. One model is thus applicable to the extraction of several road types. The variable number can be realized in the semantic net with "dynamic relations". This kind of relation can determine hierarchical as well as spatial relations between parts in the semantic net. The number of the concerning object parts can be limited to a certain maximal and minimal value. An example for a possible dynamic relation for road lanes gives Fig.2. The lanes are here described by the parallel road markings as continuous lines.
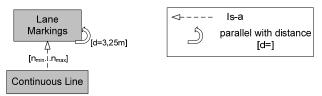

Fig.2: Dynamic Relation for Variable Lanes

3.2 Adaptation Process

According to the strategy of the adaptation (cf. section 2), the developed methods for the automatic adaptation of variable object models are explained in this section for each of the three strategy stages – decomposition, scale behaviour prediction and fusion.

3.2.1 Decomposition

The given object model for the extraction in high resolution images is in this first step of the adaptation process decomposed into single object parts or in groups of object parts. During reduction to the coarser target resolution some object parts might interact with each other, due to its spatial vicinity. Those interacting object parts have to be analysed together in the following scale behaviour prediction phase.

The coarse resolution determines the value of the scale change and hence, the target scale. The target scale represents the value of the scale parameter needed to transfer an image from the original scale to the coarser scale in linear scale-space. The linear scale-space is constructed by convolution of the original signal with the Gaussian function with the varying scale parameter as its standard deviation. The scale parameter therefore determines the width of the corresponding Gaussian kernel in the discrete space. With this background, the determination of the interaction of adjacent object parts is carried out by a comparison of the distance between the object parts and the width of the Gaussian kernel in the discrete space, corresponding to the target scale σ_t. The determination of interaction is done for each adjacent pair of object parts.

For the object parts in the dynamic relation, it is sufficient to determine the interaction only once, as the same distance d_{dyn} is separating all adjacent object parts. If there is no interaction between the lines in that group, all lines in the group are outlined as single lines. On the other hand, if the lines in the dynamic relation are subject to interaction, all lines form one group of interaction and need to be examined accordingly.

Fig. 3: Object Parts of the Dynamic Relation and their Distance

For neighbouring object parts of the dynamic group, a possible interaction with the most outlying lines in the dynamic group needs to be considered. If there is interaction, the interaction group has to be extended by those adjacent lines. Of course, the dynamic group might also have to split up, if there is no interaction in the dynamic group. In this case, only the lines at the border of the dynamic group will constitute an interaction group with the neighbouring outside object parts.

3.2.2 Scale Behaviour Prediction

The scale change models carry out the prediction of the behaviour of the object parts during scale change to the target scale. For each single line or line group, which were constructed in the decomposition phase, a scale change model examines possible scale events of the object parts and calculates their attribute values in the target scale. Hence, the scale change models predict the appearance of the object parts in the coarse resolution.

Scale Events

For dynamic relations with parallel line-type object parts possible scale events are Annihilation and Merging. While Annihilation means the disappearance of objects, Merging refers to the unification of two or more objects into a single object. If object parts in the dynamic relation form one interaction group without any neighbouring object parts from outside, scale events only occur for all dynamic relation parts simultaneously. This fact is due to the identical appearance of all object parts in the dynamic group concerning their attributes extent, grey value and distance, which are the critical values for the occurrence of scale events in general.

The Merging event only occurs when interaction is present and therefore only needs to be examined for line groups. Merging can be determined with only two adjacent lines. The result is

transferable to the whole dynamic group, i.e. when the two examined lines exhibit Merging, all lines in the dynamic group will merge into one single object.

which don't exhibit interaction, the determination of the attributes can be carried out again for only one of the objects parts as a sample for the whole dynamic group.

Fig.4: Resulting Grey Value *r* in Scale-Space under the Influence of Different Numbers of Object Parts in the Dynamic Group

If the decomposition of the dynamic group resulted in single lines, only Annihilation is possible. As there apply exactly the same circumstances for that scale event, the examination of only one line as a sample for the whole dynamic group is sufficient.

For an interaction group however, the scale event Annihilation is dependent on an additional parameter: the precise number *n* of the involved object parts. In general, Annihilation is occurring when the operator, which is assigned for the extraction of that object, fails to extract that object due to the low contrast of the object to the background. The contrast of the interaction group is influenced by the actual number of object parts, because of the interaction of the neighbouring lines. For large σ and small distances, the corresponding Gaussian filter can overlay more than one line. If the distance d_{dyn} is smaller than the width of the Gaussian filter w_f, the grey value in the resulting convolved image is then generated under the influence of several participating lines. Fig.4 sketches (for better recognisability only small quantisation steps are included) the relationship between the resulting grey value in the target scale and the actual number of involved object parts (here for three and four involved object parts of the dynamic group). Hence, the more object parts with equal characteristics are present, the more object parts might contribute to the resulting grey value, the larger the grey value in the interaction group can get and the later in scale-space the Annihilation will occur. Therefore, Annihilation is not unambiguously and reliably predictable for an unknown number of object parts in the dynamic group. The occurrence of a scale event affects the structure of the adapted object model in terms of the expected presence of one or more object parts. The result of the adaptation therefore encompasses more than one possible object model (as concept net) for the coarser resolution depending on how many object parts are present in a specific image.

However, for an object in the real world, it is normally possible to specify a minimal and maximal limit for the number of the reasonable object parts in the dynamic group. In the example for the lanes for instance, one could set the reasonable number of lanes as between 1 and 4 lanes. This allows us to restrict the possible resulting object models to a manageable amount. To enable the adaptation of variable object models automatically, the number of possible object parts should be limited in the given object model. This requirement can be seen as an additional constraint for the format of the given object model to be adapted. For other constraints on the object model in the high resolution please cf. [Pakzad&Heller05].

Attributes

The values of the attributes in the nodes of the adapted semantic net describe the appearance of the object parts in the images. For the case of single lines of the dynamic group,

On the other side, the grey value and width of an interaction group, which is subject to Merging in the target scale, is influenced by the actual number of the participating object parts, as explained above for Annihilation. Because of the variability of the number of object parts, the real attribute values can fall into a range, which can be stated accordingly in the concept net of the target scale. Here again, if the number of object parts is maximal and minimal limited, the grey value will be in a predictable range.

If despite interaction, the object parts do not exhibit Merging, grey value and width can be determined exactly in the same way as for single lines.

3.2.3   Fusion of the Adapted Object Model

The fusion assembles all the predicted nodes in the target scale to a complete semantic net, which describes the object in the coarser resolution. The structure of the adapted semantic net will only change compared to the original object model in case of the detection of scale events by the scale change models. In the case of scale events the number of object parts, and therefore of nodes, is to be reduced in the adapted semantic net. If Merging has appeared, the translation of the position of the merged object has to be considered in the fusion phase. This position translation affects the distance of the merged object to the neighbouring object parts. The distance of the merged dynamic group *L1* to an adjacent object part *L2* in the target scale σ_t can be calculated from the distances d_{dyn} of the dynamic group in the original scale σ_0 and the number *n* of the participating object parts in the dynamic group:

$$d_{\sigma_t L1,L2} = d_{\sigma_0 L1,L2} + \frac{1}{2}\sum_{i=1}^{n-1} d_{dyn} = d_{\sigma_0 L1,L2} + \frac{(n-1)\cdot d_{dyn}}{2} \qquad (1)$$

for *n>0*.

4.   EXAMPLE FOR THE ADAPTATION

In this section, the adaptation process of a given variable object model is demonstrated exemplarily for a target scale of $\sigma_t = 50$, corresponding to a spatial resolution of about 5m. The object model in this example describes a road with a variable number of lanes. The variability of the lanes is realized by the previously in section 3.1 introduced dynamic relation, here with a variable number of dashed lane markings. The object model (see Fig. 5) was created for the extraction of a road with 1-4 lanes, i.e. up to 3 dashed lane markings, in aerial images of about 3cm spatial resolution.
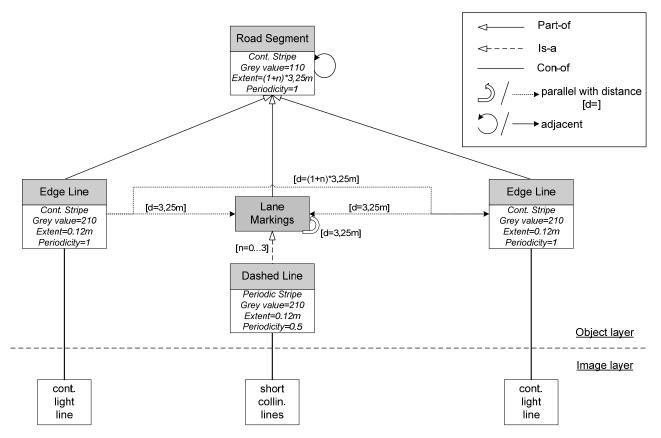
Fig. 5: Variable Road Model for ~3cm/pel Resolution

The road is modelled by its roadway and markings. In high resolution aerial images road markings are detectable as bright lines. This example model describes a road segment with two continuous lane markings at the edges of the road pavement. In between these edge lines, the model allows the existence of up to three more dashed lane markings, which are realized by the dynamic relation. Therefore, the extractable road must have got at least one lane, and might consist of two, three or four lanes (separated by the variable number of lane markings). The lane markings are parallel to each other in the same distance d_{dyn}, the width of one lane. Furthermore, the lane markings itself are dashed (periodic) lines in contrast to the edge lines (continuous lines), which is expressed as different object types. Other attributes in the nodes concerning the appearance of the lines in the images are given for the grey value, extent and periodicity. Except for the periodicity, which is different for continuous or dashed lines, all lane markings have the same attribute values and thus, will have a similar appearance in the image. The appearance of the pavement itself is also modelled in the upper node "Road Segment". The contrast of the markings to the road surface is given implicitly by the difference of their respective grey values. The modelling here is made under the assumptions that the road markings are detectable by a feature extraction operator.

From the decomposition of the given road model results one single group of interaction for all markings, as they all hold the same distance to the next lane marking, which is determined by the lane width (3.25m). This distance is smaller than the width of the Gaussian filter in the target scale σ_t. All object parts of the lowest hierarchical level in the semantic net, i.e. all lane markings, are interacting in the given target scale.

By limiting the number of the participating lanes to a realistic numer, the automatic prediction of scale-space events is feasible. For the determination of the attribute value limits, the minimal and the maximal configuration concerning the number of lanes in the dynamic relation have to be examined in the scale change models. Hence, in this example the scale change models have to predict the attribute values, namely the extent of the lines and the grey value for one lane, delimited by the two edge lines, and for four lanes with three lane markings between both edge lines, altogether five lane markings. In the scale behaviour prediction phase, the scene is analysed by a simulation of a synthetic image. This synthetic image is created using the information about the line's appearance in the nodes and edges of the semantic net. Fig. 6 and 7 display cut-outs of the synthetic images for the minimal and maximal lane configuration (here with three dashed lines). Because of the parallel nature of the lines, it is sufficient to carry out the simulation only with parts of the lines.

Fig.6: Synthetic Image (cut-out) of the Minimal Configuration in the Original Scale σ_0
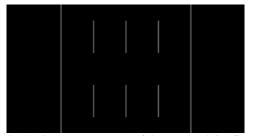
Fig. 7: Synthetic Image (cut-out) of the Maximal Configuration in the Original Scale σ_0

The scale change models examine scale-space events for the interaction group with all possible lane markings. For the minimal configuration with only the two edge lines and no lane markings, the interaction group exhibits Merging (cf. Fig.8). In the target scale, there is only a single object existent, which is emerged from all object parts in the original semantic net.
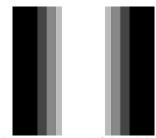

Fig. 8: Synthetic Image (cut-out) of the Minimal Configuration in the Target Scale $\sigma_t=50$

In the scale-space image of the target scale σ_t there is no definite Merging for all lines in the maximal configuration with three dashed lane marking, as in the cross section of the altogether five lines there is more than one extreme value (maximum) of the resulting grey values present (cf. Fig.9). However, there is Merging for all dashed lane markings. The adapted semantic net for the maximal configuration will therefore consist of three object parts.
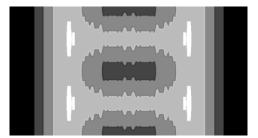

Fig. 9: Synthetic Image (cut-out) of the Maximal Configuration in the Target Scale $\sigma_t=50$

For this scale change, different scale events might occur depending on the actual number of participating lines in the dynamic group. Thus, different object models are the result of the adaptation of the variable object model for a scale change of $\sigma_t = 50$ in this example.

Annihilation, on the opposite side, is determined by the resulting contrast in the centre of the line group, which varies with the number of the involved lane markings. In a line group with many lines, the grey value in the target scale might be larger as for a line group with fewer lines. In the presented example, the resulting grey value of the Merging group is for the minimal configuration with 2 lines as large as the largest grey value of the maximal configuration with 5 lines. The contrast is in the coarse target resolution of ~5m still large enough to extract the lines with the assigned Steger operator [Steger98]. Therefore, no Annihilation has occurred during scale change.

The appearance of the remnant object parts is predicted by the attribute values in the scale change models as well. The grey value is determined in the synthetic image in the target scale σ_t. The extent (width for parallel objects) of the merged line is calculated by the gradient image in the target scale σ_t. The width of the line varies according to the actual number of lines in the dynamic group. In this case, the resulting extent is a range of values, delimited by the minimal and maximal configuration. The gradient images in the target scale σ_t of both configurations are displayed in Fig. 10 & 11.
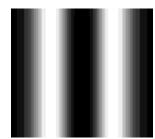

Fig. 10: Gradient Image of the Synthetic image (cut-out) of the Minimal Configuration in the Target Scale $\sigma_t=50$
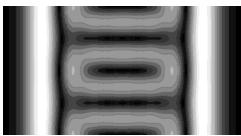

Fig. 11: Gradient Image of the Synthetic image (cut-out) of the Maximal Configuration in the Target Scale $\sigma_t=50$

The resulting object type of the merged lines in the minimal configuration without lane markings is a continuous line. The periodicity of this type is always 1. In the maximal configuration, however, the Merging occurs for the dashed lane markings. Here the resulting object type for the merged object parts is a periodic line. Fig. 12 & 13 depict the corresponding synthetic images in the coarse resolution of about 5m/pel. These images were created by resampling the smoothed synthetic images of the target scale to the appropriate spatial resolution. Please note that the images displayed here are visually enhanced by grey value stretching because of its originally low contrast.
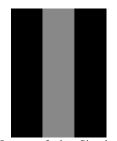

Fig.12: Synthetic Image of the Simulation with no Lane Markings in the Target Resolution ~5m/pel
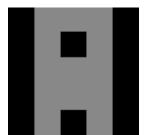
Fig.13: Synthetic Image of the Simulation with 3 Dashed Lane Markings in the Target Resolution ~5m/pel

The result of the automatic adaptation is no uniform structure of the adapted semantic net due to different possible scale events depending on the actual number of the involved object parts. For the minimal configuration with two continuous edge lines only one object part is left in the lowest hierarchy level of the semantic net. All spatial relations and their corresponding distances are then omitted. The adapted semantic net for the minimal configuration to the coarse resolution of ~5m is presented in Fig.14.

For the maximal configuration with two continuous edge lines and three dashed lane markings there are still the two edge lines existent, but only one periodic line for the merged lane markings left. Fig.15 represents the resulting semantic net for the maximal configuration.
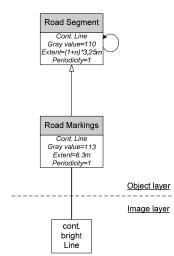

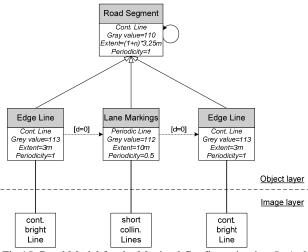Fig.14: Road Model for the Minimal Configuration in ~5m/pel

Fig.15: Road Model for the Maximal Configuration in ~5m/pel

5. CONCLUSION

In this paper new methods for the automatic scale-dependent adaptation of variable object models (represented as semantic nets) are presented here. The methods aim at the adaptation of object models with a variable number of certain object parts and are a further development of the already existing adaptation methods for object models with a constant well-known number of object parts. By introducing variability into the object models, one variable object model is capable to describe more than one type of roads. Thereby, the models become more flexible regarding the extraction process. An example for the adaptation of a variable road model demonstrates the adaptation process of a variable object model to a coarser spatial resolution.

Despite the newly introduced flexibility of the adaptable object models due to the variability, the presented example does still not meet the requirements of a successful and reliable road extraction procedure in real aerial images. In the future, the methods for the automatic adaptation are thus to be augmented to the adaptation of more realistic object models with non-parallel and area-type object parts. In this case, not only the profile of the parallel object parts, but the scale behaviour of the 2-dimensional object needs to be examined in the scale change models. Then different events in scale-space can occur, which the methods should be able to predict.

Moreover, the example object models are currently implemented in the knowledge-based interpretation system GeoAIDA [Bückner04]. The extraction results of the original object models and the automatically adapted object models in the corresponding resolution can then be used for the verification of the developed methods.

6. ACKNOWLEDGEMENTS

7. REFERENCES

Bückner, J., Pahl, M., Stahlhut, O., Liedtke, C.-E., 2002: A Knowledge-Based System for Context Dependent Evaluation of Remote Sensing Data. In: *LNCS* 2449, DAGM2002, Springer Verlag, Berlin, pp. 58-65.

Florack, L., Haar Romeny, B., Koenderink, J. and Viergever, M., 1994. Linear Scale-Space**.** *Journal of Mathematical Imaging and Vision*, 4(4), pp. 325-351.

Förstner, W., 1993. A future of photogrammetric research. *NGT Geodesia*, 93(8), pp. 372-383.

Heller, J., Pakzad, K., 2005. Scale-Dependent Adaptation of Object Models for Road Extraction. In: *IntArchPRS*, Vol. XXXVI, Part B3/W24, Vienna, pp. 23-28.

Lindeberg, T., 1994. *Scale-Space Theory in Computer Vision.* Kluwer Academic Publishers, The Netherlands, 423 p.

Pakzad, K., Heller, J., 2004. Automatic Scale Adaptation of Semantic Nets. In: *IntArchPRS*, Vol. XXXV, Part B3, Istanbul, pp. 325-330.

Steger, C., 1998. "An Unbiased Detector of Curvilinear Structures." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(2), pp. 113-125.