GRAPH BASED METHODS FOR LOCALISATION BY A SKETCH

Matthias Kopczynski, Monika Sester

Institute of Cartography and Geoinformatics, University of Hannover
Email: Matthias.Kopczynski@ikg.uni-hannover.de and
Monika.Sester@ikg.uni-hannover.de Web: http://www.ikg.uni-hannover.de

ABSTRACT

This paper describes how a sketched location can be identified in a large collection of reference data. This is done by encoding the information of both datasets in a graph based structure, which preserves the topological properties of the sketched area. Finding the correspondence is a matching problem. In order to reduce complexity first a set of possible correspondences are identified using the following approach. From these graphs the frequency of relevant objects and relations is extracted and represented as a histogram. Distance measures for probability distributions can be used to judge if they are similar enough to represent the same location. Considerations on the information contents are leading to a way how to choose object types and relations used for the matching process. The distance measures are the fundamental tool for an efficient retrieval strategy. Since they declare a metric space, they can be used in a metric tree. Only promising locations are then considered in the matching process.

1 INTRODUCTION

In the past a considerable amount of the scientific work was dedicated to the problem of locating something or someone in space. The knowledge of the current location was a valuable information for war, trade and even for religion. Centuries ago it was a time consuming task to find a position, but getting easier and easier until today. Now it only requires to take a GPS-Receiver and just read your absolute position on earth with an accuracy of a few meters. Only in technical applications a greater accuracy is sometimes needed and specialized methods are available for most of them.
What you get from this techniques is a set of coordinates, being perfect for calculations but they lack some sort of descriptiveness if used in communication between humans. Indeed it is a rather mathematical view and most people need the help of a map to identify the set of coordinates with a place in reality. A well prepared additional medium is needed for this process. This overhead in the communication process leads to the question for the natural way of space description.
If two communicating partners know a common name for the place, they immediately understand what they are talking about. The name is directly connected with pictures, experiences and emotions regarding the place in mind. If one of the communicating partners does not know how to connect the name with a place, they have to describe the location in more detail. The reason for this may be that one of the partners does not know the name but the place or he knows the name but can't remember anything about the place. Another case occurs when no name for the place does exist at all because it is not important enough or it is defined by some complex constraints, which are too specific to have a name.
When the name is missing, people are using a description of the place which is a combination of known names and categories of topographic objects. Formally it may be a verbal description or a sketch drawn on a piece of paper. The sketched place is certainly easier to understand, especially when it is supported by a verbal description.
Today the computer is a common tool for everyone and many activities are planned and accomplished with it, including activities involving different locations. Considering the computer as a communication partner we can transfer the former considerations to this domain. Determining the current position with a GPS receiver is very easy, but what happens when another location is of special interest? If you know the name of the place and computer does know it as well, we don't have a problem. But in all other cases the computer cannot easily identify the location. Current available software solves this problem for example by presenting a map and asking the user for selecting a region [1][2].
As an alternative a sketch may be a more suitable input modality because it is very close to the behaviour of the user when he communicates with another person. It supports the knowledge about details of the place which are only visible in some scales and does not rely on the ability of the user to choose the right scale for the identification of the place.
The EU project SPIRIT [3] has developed a spatially aware search engine supporting a location to be one topic of the retrieved documents. The user interface can be accessed with different input modalities, including a sketching interface. Other useful applications are location retrieval in general GIS applications or PDA-based navigation systems.

XXII International Cartographic Conference (ICC2005) A Coruña, Spain, 11-16 July 2005
Hosted by:
ISBN: 0-958-46093-0 The International Cartographic Association (ICA-ACI)
Produced by Global Congresos

2 FUNDAMENTALS ABOUT SKETCHING

Definition of a sketch

The term "sketch" is not very strictly defined and it needs some clarification regarding the context of its use. In general a sketch is a rough description of something and contains only the relevant properties of the subject which it describes. In a more specific sense it is a quickly drawn picture, applied with a pen on some drawing media, containing only the outlines or prominent features of an object.
To make use of a sketch for localisation purposes this is still too general, because it may contain non spatially related topics. More useful is the definition of a sketch as a line drawing on a medium, especially a computer screen, which is following the style used for way finding descriptions. The place is sketched from the birds perspective and the objects are drawn in an abstract style [4], mainly using geometric primitives for topographic objects. This is also called a "map sketch".

Sketched objects and relations

The sketched objects usually have a clear manifestation in reality and cities, roads, railroad stations, buildings and street lights are some common examples. But the existence of the manifestation in reality is not a necessary. Examples for invisible objects are administrative boundaries or the inexact limits of areas like southern and northern Europe.
Many objects are aggregated from objects at a larger scale. Cities are aggregations of districts which are themselves aggregations of buildings and roads etc. Which of the representations is used depends on the importance of an object for the identification of the location and how detailed the location has to be determined.
In a map all the objects are presented in a consistent scale but a sketch usually is a mixture of different scales. This property has a big impact on the ability to produce a unique identification of a location.
Much of the information is encoded in the relations between the objects. Since the sketch is not a scaled representation of space, the topological relations are gaining importance. Objects which are neighboured in reality are likely to be drawn as neighbours in the sketch. Other relations are for example distance relations where a constraint is used to make the sketch more unique.

Graph Representation

The sketch itself is only a description of the location and it serves as a pattern that is compared to data about real locations. The reference data can be used to find a corresponding partial subset of objects which is representing the location of the sketch. If the reference data has coordinates attached, the sketched objects can be identified with coordinates from that data.
The reference data will in practice be some kind of GIS data, mainly collected for other purposes than locating sketches and thus it must be transformed into a suitable representation. In general it is stored as a collection of geometric primitives with implicit topology and this information must be made explicit because only this information is available from the sketch. A graph structure is the obvious choice when topological information is stored, because the edges represent the topological relations directly. The nodes are identical with the sketched objects and are connected by their relations. The example from Figure 1 shows how the general structure of alternating object and reference nodes [5].

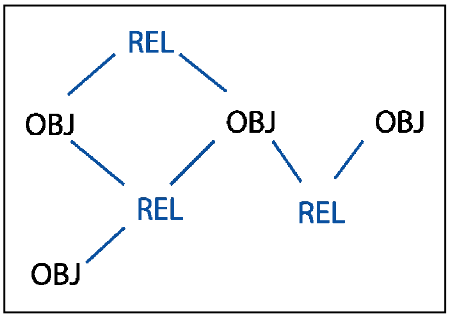

Figure 0 Here the objects and relations are both stored as nodes of the graph.

3 STATISTICAL DISTANCE MEASURES

For the pattern matching process a technique is needed which can determine if the graph from the sketch is fitting onto a sub graph of the reference data. This leads to the well known problem of sub graph isomorphism and inexact graph matching. Inexact matching is necessary because the sketch is very unlikely to be drawn as the reference data models the location. In most cases the user of an application that supports sketching would not draw all objects and relations but only a subset containing the important information. In other cases it is possible that the user simply does not remember the real place correctly and includes errors into the choice of objects and relations. Both cases are illustrated in Figure 2. Unfortunately inexact graph matching is a NP complete problem with exponential growth of comparison time depending on the graph size [6]. Even for small graphs compared to the average sketch sizes, calculation times can

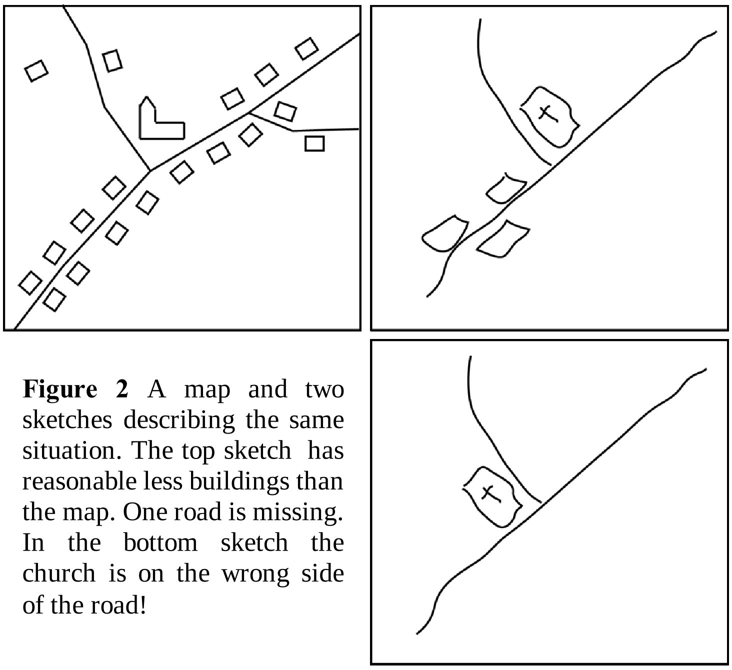

Figure 2 A map and two sketches describing the same situation. The top sketch has reasonable less buildings than the map. One road is missing. In the bottom sketch the church is on the wrong side of the road!

easily cumulate to years. Several approximation algorithms have been developed in the past [7][8]. Most of them are optimized for image matching with attributed relational graphs (ARGs) and the literature does report some good results in special applications. Own experiments have shown that their use for locating a sketched area is very limited, because they take far too much time for the calculations or don't produce any meaningful results.

All the primarily graph based methods have been designed for assigning individual nodes from one graph to another. Interestingly this is not necessary for the matching process because only one location for the whole sketch is sufficient. Only the existence of possibly matching objects is sufficient for an assignment. Counting and comparing the number of objects then provides a measure for the identity of two sub graphs.

Graph histogram

When it does not matter how the individual nodes are assigned it should be sufficient to count the number of objects to get a first impression whether it is consistent with the reference data. Doing this for every single object type and relation type is leading to the graph histogram. The idea is that the histogram of the sketch should have the same shape as the histogram of the matching location of the reference data. The order of counts for the object types and relations is not of any importance as long as the same order is used in the compared histograms. Because the user will discard the unimportant objects, the absolute counts will vary a lot. Taking the relative frequencies into account does not depend on the absolute number by preserving the shape of the histogram. The relative frequencies define a probability distribution over the feature space, that is formed by the single object types and relations types [9][10].

Comparing distributions

Two graphs are considered to describe the same location when their histogram has the same shape. The difficulty is now to find a measure for the discrepancy between two distributions. The solution is well known from statistical goodness of fit tests where the problem is to identify a certain distribution from collected data. Some of them do not require to know the details about the compared distribution. They can be used to decide whether two datasets stem from the same distribution. Two of them should be introduced here: the Chi Square Distance and the Kolmogorov-Smirnov Distance.

With two distributions f_1 and f_2 with $f_1(x)$ and $f_2(x)$ giving the relative frequency of feature $x=1..k$ the chi square distance is defined as

$$d_{\chi^2} = \sum_{i=1}^{k} \frac{(f_1(i) - f_2(i))^2}{f_2(i)} \tag{1}$$

The Kolmogorov-Smirnov distance can be computed as

$$d_{KS} = \max_{1<i\le k} \left| F_1(i) - F_2(i) \right| \tag{2}$$

Here F_1 and F_2 are the cumulated relative frequencies from the given distributions. The distances behave differently when the number of features is varied or the feature counts are taken from distinct magnitudes. An important property is that these measures are satisfying the axioms of metric spaces: non negativity, symmetry and triangle inequality. This is especially useful for the construction of retrieval strategies [9].

4 INFORMATION CONTENTS OF A SKETCH

A sketch always has a limited size and it tends to be small, this is due to the fact that it has to be created quickly and only contains the most important objects. This has an effect on the number of different sketch locations which can in principle be distinguished with the histogram method. If the number of possible places is small, it is not a good idea to search a large area with many locations, where most of them match to the same pattern. The information content [11] of the sketch can help to quantify this property.
Limiting the number of features in the histogram to one we can analyze the information content of a sketch covering the area A. The realizations shall be random distributed with density δ. For the average number n of realization of the feature, called points, it can be written

$$\delta = \frac{n}{A} \tag{3}$$

Because the position of the points is randomly generated, for example by an process like falling rain, the random variable X being the number of realized points on the area follows the Poisson distribution.

$$Po(X = k \mid \delta A) = \frac{(\delta A)^k}{k!} e^{-\delta A} \tag{4}$$

Information contents of a distribution

The information contents or entropy of a distribution in general is given by

$$I = -\sum_{j=1}^{\infty} p_j \cdot \log_2 p_j \tag{5}$$

where p_j is the probability of getting the realization count j of the random variable [12]. The unit of the Information I is measured in bits. Setting p_j as Poisson distributed $Po(X=j|\delta A)$ it provides a method to calculate the entropy for a random distributed feature. For example an area of 1 km² with a feature density of 10 points/km² has an information contents of 3.69 bits. This can be interpreted in a way that the feature has the potential to distinguish between $2^{3.69}=12.9$ different locations. As a consequence one feature with the given distribution is only capable to identify a location out of approximately 2^I times the area A. Any larger area will be dominated by ambiguous results.
In reality the distribution of the feature will not strictly follow the Poisson distribution. Cities are planned with some design pattern in mind and buildings are not spread randomly across a quarter. Nonetheless this features can be described with a distribution and the equation (5) can be applied.
When a feature is strictly following a grid pattern its distribution degenerates to a single point distribution. Only one realization of feature counts is possible and evaluation of the entropy expression gives a value of $I=0$ for the information content. Since $2^0=1$, such a feature can only describe one location as expected intuitively and is of no use for the localisation process.
As a consequence it is the best strategy to choose a feature which is as irregular distributed as possible [12]. This is the direct inversion of the entropy definition.

Information content with multiple features

Intuitively a place is not only described by only one feature and additional features increase the number of distinguishable locations. In this situation every feature adds an additional dimension to the random feature vector $\boldsymbol{X}$. With the assumption of uncorrelated features, the total entropy of the multidimensional variable X is described by

$$I(\boldsymbol{X}) = I(X_1) + I(X_2) \tag{6}$$

As the total entropy is the sum of all feature entropies, the number of possible locations described is roughly growing exponential with the number of features. In practice this is not strictly the case because the features are correlated. A road is vastly increasing the possibility of finding a traffic light. In this cases the equal sign in (6) must be exchanged against a less than sign [11].
When selecting features for matching a sketch to a reference data set, it is encouraged to choose as much independent features as possible. The density of the features is of minor importance while it should not form a regular structure.

However, the choice is restricted by the contents provided by GIS data sets and what can be expected to be in a sketch. A feature that is never sketched does not support identification and can be safely ignored.
When the reference data set is designed for navigation purposes, the following features may be used [13]:

- vertex degree at crossings
- road class
- acute, obtuse angles
- special aggregates, like T-crossings, Y-crossings, ...
- connections between different road classes
- orientation (coarse)
- distance classes
- names
- shape classes
- 9-intersection between objects
- points of interest

The term "point of interest" is an aggregation of many different object types, mainly buildings with a special function. The mentioned data set would contain types like

- petrol station
- police station
- hotel
- church
- airport
- museum
- shopping centre
- mountain peak
- library
- stadium
- lighthouse
- ...

Other data sets may set the priorities on topographic objects like the German ATKIS or on thematic objects like geo coded addresses or administrative boundaries like the European SABE data set. It is not expected to find data with a complete set of features appropriate for the sketch localisation. In practice a mix of certain data sources will be used. The information contents can be useful to decide if a feature helps to identify a place or not. Combined with an analysis of object types sketched by the users the result is an aggregated reference data set including all the necessary features. Dispensable features should be eliminated because they slow down the identification process and their use could lead to misinterpretations when the user introduces errors into the sketch while using them.

5 RETRIEVAL STRATEGIES

Now it is clear how to compare a sketch to a single location in the reference graph and how features can be selected to get as much information from the sketch as possible. But the reference data set contains a larger amount of locations and different strategies can be applied to find the best matching location. The limiting factor here is the observed retrieval time. Applications with sketch support are dominated by interaction with a user and they should give a response as soon as possible.

Cross correlation

The easiest strategy would be to perform the comparison for every location from the reference database. This is very similar to cross correlation in numerical applications, because neighboured locations share a lot of points and hence lead to nearly the same distances between the sketch and two locations. The sketch is moved as a pattern across the reference and the maximum of the correlation function indicates the best possible matching location. Similar techniques are well known from appearance based vision in image analysis [15]. They are for example used to identify the appearance of objects by giving a pattern with a fixed pixel size. Unfortunately the correlation algorithm can not be directly transferred to graph matching, because graphs have a significantly higher degree of freedom.
However, this is a slow strategy because it always compares the sketch to all the possible locations. This leads to impractical long retrieval times and should not be used. Despite of this limitations it is very useful to evaluate a new set of features and to explore how they influence the ability to identify a sketched place as unambiguous as possible. Figure 3 is the correlation picture with a part of the shown road network taken as a pattern (in red circle). The green areas are showing a good probability for matching the pattern against the sub network. Only the node degree at each crossing combined with the road class was used. Here the maximum of the correlation function is not found at the original pattern location because cutting out an area does introduce some nodes with degree one and thus the pattern histogram is distorted, increasing its distance to the original. More features should increase the red areas with a very low matching probability, leaving only the more or less correct areas as possible solutions.

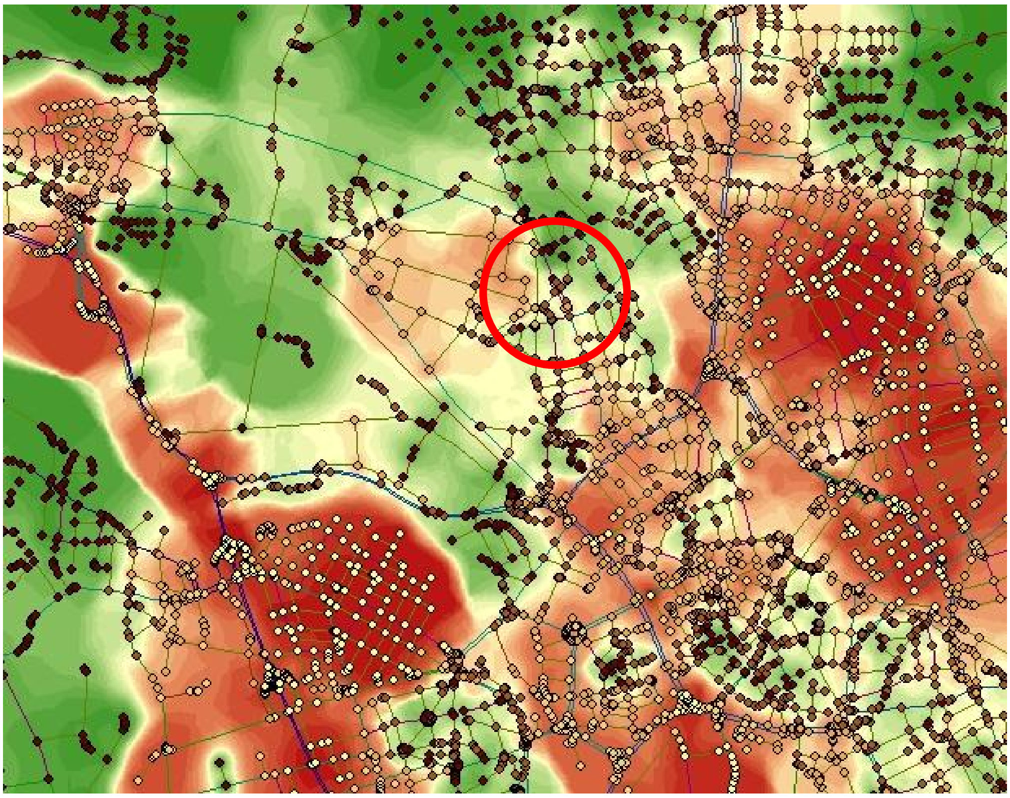

Figure 3 This is part of the Hannover road network matched against the area in the red circle. The green colour shows areas where a good matching probability is found while the red areas do not match very well.

Metric trees

In classical retrieval problems with one dimensional feature vectors, tree structures are very successful. The query time only grows logarithmical to the number of elements in the tree and only a small amount of elements is evaluated. With multidimensional feature vectors other trees like KD-trees, R-trees and Quad trees are the right choice under the precondition of Euclidian spaces. Unfortunately the distance measures for the comparison of distributions do not form an Euclidian space with the implication that this sort of trees can not be used directly.

In the past the problem was discussed in the context of retrieval in large multimedia databases and a generalized approach was found. The solutions only require the distance to form a general metric space. Under this conditions it is possible to find tree structures being able to approximate the behaviour of trees in Euclidian space. One of the structures is taken here as an example, the M-tree [14]. The M-tree is designed to work like any balanced tree. The internal nodes are specified as centres of a hyper sphere in the metric space and give a hint about the maximum distance that is expected between two elements.

The construction is starting with a root node and an associated radius that is equal to the largest distance to any of the remaining elements. A number of children is selected from these elements. They should have the greatest possible distance and their associated radius should be as small as possible. The rest of the elements is allocated to the children if they have a distance to a child that is smaller than the distance to all the other children. The associated radius must be larger than the

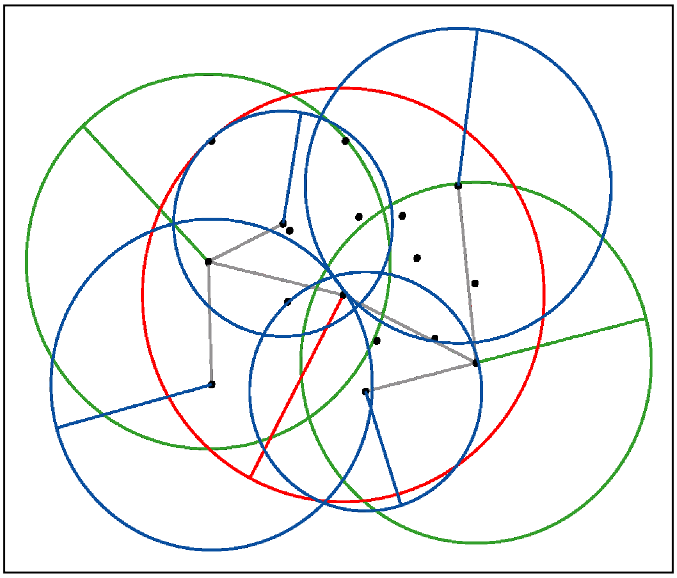

Figure 4 M-tree with euclidian distance. The red hyper sphere is the root node, then the green and the blue hyper spheres continue to partition space

distance from the child to all the allocated elements. This process is repeated until the number of allocated elements to a child is smaller than a chosen maximum (Figure 4).
To perform a nearest neighbourhood search (NN-search) for a given element, it must first be compared to all children of the root element. If the distance is smaller than the associated radius, the children are stored in a sorted list of elements with a distance to the hyper sphere of zero. If the distance is larger, it is included into the list but only if this distance is smaller than the distance of the last element in this list. The elements of the list are taken as root nodes and the comparison of hyper sphere distances is repeated.
The effect is that some of the branches are pruned and their leaf elements don't need to be considered. Unlike the trees in Euclidian space, more than one branch is followed in every iteration. This way the runtime behaviour is expected to be somewhere between the cross correlation approach and the trees in Euclidian space. The exact complexity depends on the used metric, the space division strategies and the data distribution.
The construction of the tree has to be performed before the first NN-search and is a limiting factor because it clearly needs more distance evaluations than one cross correlation run.

6 CONCLUSION

Sketch based localisation is a difficult task and to achieve user friendly retrieval times only approximations to the stricter purely graph based methods can be used. Using histogram based distance measures has the advantage of simplicity and high speed. The runtime complexity is linear depending on the size of the sketched area. In combination with metric trees a large database can be searched in a few seconds. The price paid for the speed is the loss of the information originally encoded in the sketch and the reference graph. The complicate pattern of connections in that graph is discarded and only the frequency of typical objects is taken into account. However, in order to refine an approximate solution more exact techniques can be used on this very limited number of best solutions.
The next step is to identify object types and relations that are both used in the GIS based reference and in a sketch drawn by humans and additionally really allow to distinguish between locations. Data sets for navigation purposes do provide a lot of information about roads, which are indeed found in the discussed kind of sketches, but many sketched features like buildings and vegetation types are missing. Additional data sources must be accessed, e.g. cadastral data sets.

ACKNOWLEDGEMENTS

This work is funded by the EU in the IST-5th framework project SPIRIT.

REFERENCES

[1] H. Tappe, C. Habel, Verbalisation of Dynamic Sketch Maps: Layers of Representation and their Interaction, 1998
[2] C. Freksa, Spatial Aspects of Task-specific Wayfinding Maps, In Visual and Spatial Reasoning in Design, J.S. Gero, B. Tversky, p. 15-32, Key Centre of Design Computing and Cognition, University of Sydney, 1999
[3] C. B. Jones et al., Spatial Information Retrieval and Geographical Ontologies: An Overview of the SPIRIT project, SIGIR 2002: Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval August 11-15, 2002, Tampere, Finland, ACM Press, pp.387 - 388. (2002)
[4] A. D. Blaser, Sketching Spatial Queries, The Graduate School, Maine, University of Maine (2000).
[5] M. Kopczynski, M. Sester, Representation of Sketch Data for Localisation in Large Data Sets, XXth Congress of the International Society for Photogrammetry and Remote Sensing (ISPRS), Istanbul, Turkey (2004)
[6] J. R. Ullmann, An Algorithm for Subgraph Isomorpism, Journal of the ACM (JACM), Volume 23, Usue 1, p. 31-42, ACM Press, New York (1976)
[7] R. C. Wilson, E. R. Hancock, Structural Matching by Discrete Relaxation, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 6, p. 634-648 (1997)
[8] E. Bengoetxea, Inexact Graph Matching Using Estimation of Distribution Algorithms, Ph. D. Thesis, Ecole Nationale Superieure des Telecommunications, Paris (2002)
[9] A. N. Papadopoulos, Y. Manolopoulos, Structure-Based Similarity Search with Graph Histograms, Proceedings of the DEXA/IWOSS International Workshop on Similarity Search, p. 174-178, IEEE Computer Society (1999)
[10] A. M. Ferman, A. M. Tekalp, R. Mehrotra, Robust Color Histogram Descriptors for Video Segment Retrieval and Identification, IEEE Transactions on Image Processing, Val. 11, No. 5, May 2002
[11] C. E. Shannon, W. Weaver: The Mathematical Theory of Communication, University of Illinois Press (1963)
[12] H. Rinne, Taschenbuch der Statistik, 3. Auflage, Verlag Harri Deutsch (2003)
[13] TeleAtlas Multinet User Guide, TeleAtlas 2003

[14] M. Patella, Similarity Search in Multimedia Databases, Ph. D. Thesis, University of Bologna (1999)
[15] B. Leibe, B. Schiele, Scale Invariant Object Categorization Using a Scale-Adaptive Mean-Shift Search, Springer LNCS, Vol. 3175, p. 145-153, 2004

GRAPH BASED METHODS FOR LOCALISATION BY A SKETCH

Matthias Kopczynski, Monika Sester

Institute of Cartography and Geoinformatics, University of Hannover
Email: Matthias.Kopczynski@ikg.uni-hannover.de and
Monika.Sester@ikg.uni-hannover.de Web: http://www.ikg.uni-hannover.de

BIOGRAPHIES

Matthias Kopczynski (Dipl. Ing.), born in 1976, studied Geodesy at the Technical University of Hannover and obtained the Master's degree (Dipl.-Ing.) in 2002. His master thesis was about .Implementation of a 3D visualisation module for the presentation of geo information.. Since 2002 he is a scientific assistant at the Institute of Cartography and Geoinformatics, University of Hannover. He is working in the EC-project SPIRIT and his primary interest lies in spatial data interpretation and geometric algorithms concerning sketching.

Tel: +49511/762-5422

Monika Sester (Professor Dr.-Ing.), born in 1961, studied Geodesy at the Technical University of Karlsruhe and obtained the Master's degree (Dipl.-Ing.) in 1986. Until October 2000 she was a staff member of the Institute for Photogrammetry of the University of Stuttgart, Germany, where she obtained her PhD in 1995. Since November 2000 she is Professor and head of the Institute of Cartography and Geoinformatics at the University of Hanover. Her primary research interests lie in multi-scale approaches in GIS and image analysis, data generalisation and aggregation, data interpretation, and integration of data of heterogeneous origin and type. She is chair of the Working Group on Multiple Representation of Image and Vector Data Commission II of the International Society for Photogrammetry and Remote Sensing (ISPRS). She is also active in the Commission on Generalisation of the ICA.

Tel: +49511/762-3588
Fax: +49511/762-2780