Alternative options of using processing knowledge to populate ontologies for the recognition of urban concepts

Patrick Lüscher, Robert Weibel and Dirk Burghardt
Department of Geography, University of Zurich, CH-8057 Zurich, Switzerland
E-mail: {patrick.luescher, robert.weibel, dirk.burghardt}@geo.uzh.ch

Keywords: cartographic databases, ontologies, spatial data enrichment, pattern recognition, building types

Abstract

In this paper, we present an ontology-driven approach for cartographic pattern recognition in support of map generalisation. Spatial patterns are formalised by means of ontologies which are then used to deductively trigger appropriate low level pattern recognition techniques. Modelling ontologies suited for spatial pattern recognition is discussed by example of an ontology of terraced houses. The paper subsequently focuses on approaches for inferring the instances of higher level concepts. Three different approaches are employed to detect terraced houses in Ordnance Survey MasterMap® vector data: Weighted summation; Joint Bayes classifier; and Support Vector Machines. An evaluation by comparison to a manual classification reveals that weighted summation and the Joint Bayes classifier both have satisfactory prediction accuracy, but the Joint Bayes classifier has advantages when considering the calibration effort involved. In conclusion, we claim that the ontology-driven approach better captures the complex structure of spatial patterns and provides enhanced transparency and flexibility of the pattern recognition process in comparison to conventional, purely geometric and/or statistical techniques.

1. Introduction

Map generalisation aims to derive a model of the geographic reality that is appropriate for portrayal at a certain scale and purpose. It is important to note that this abstraction process is not just a matter of simplification of detailed situations in order to reduce spatial clutter and therefore guarantee legibility of a map; rather, different phenomena and patterns have to be portrayed at various scale levels. Bertin (1967, 1999:300) therefore distinguishes *conceptual generalisation* and *structural generalisation*. Conceptual generalisation happens when "a city emerges from a collection of houses and streets", or a "coal pan from a collection of coal mines". Structural generalisation simplifies geometry, but conserves conceptualisation. Latterly, this dichotomy has been termed *model (or model-oriented) generalisation* and *cartographic generalisation* (Grünreich, 1992).

It is our belief that conceptual abstraction can be achieved better when we understand the geography that maps model. In our research, we aim at understanding how complex higher level concepts can be derived out of simpler concepts. Rather than to deliver a specific set of algorithms to detect some of the concepts, our aim is to provide a framework that allows to formalise the concept definitions, and investigate how these formal definitions can be used to infer detection processes in order to enrich the database.

Of our particular interest are spatial patterns in the urban domain, since they provide the basis for a variety of tasks. Some concrete examples for the urban domain might be the derivation of the construction period of particular buildings to infer the typical copper concentration per building (Bergsdal et al., 2007), a more advanced application might be to connect urban land use patterns with urban evolution processes (Camacho-Hübner and Golay, 2007), or improved adaptation in mobile services such as navigation by considering spatial contexts specified in the cartographic database (Winter, 2002).

Our approach employs ontologies to better inform the pattern recognition process. Our paper in the previous ICA Workshop discusses the motivation of this approach (Lüscher et al., 2007). We explicitly model higher level semantic concepts and their (spatial) relationships to other, lower level concepts. While the lowest level concepts are extracted through traditional cartographic structure recognition processes, they can be used to infer the existence of higher level concepts.

The individual steps of this ontology-driven approach are illustrated in Figure 1: We depart from textual descriptions of urban spaces (step 1), then formalise these patterns, their context and hierarchical composition using methods from ontological engineering (Gómez-Pérez et al., 2003) (step 2). The ontological definitions of patterns are then used to deductively trigger appropriate 'low level' pattern recognition algorithms (step 3) in order to detect them in real spatial databases (step 4).

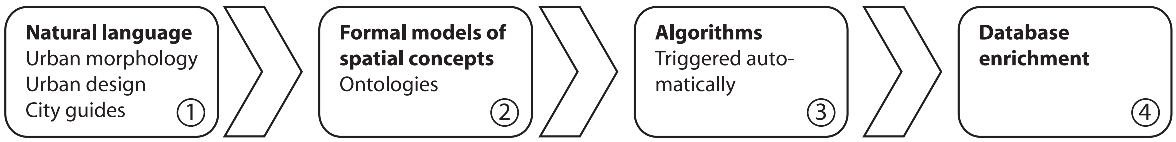


Fig. 1. Steps in the processing chain of ontology-driven pattern recognition.

A case study demonstrating the feasibility of the approach by a prototype implementation of the process will be presented at the Spatial Data Handling Conference (Lüscher et al., 2008). In this paper, we focus on the complexity of inferring instances of higher level concepts. After a brief review of related work (§ 2), alternative possibilities of capturing ontologies are explored. The example of a terraced house ontology is used to explain the formalisation of ontologies with a focus on spatial data enrichment (§ 3). We then turn to deriving computational models from ontologies. The theoretical background is introduced for three different methods for composition of higher level concepts (§ 4). In § 5 the three different classification methods are evaluated by comparing to a manually classified dataset. After a discussion of the approach in § 6, § 7 provides concluding remarks and an outlook on future research plans.

2. Related Work

Torres et al. (2005) use ontologies to describe the semantic content of topographic and thematic maps. Map features are described in terms of specialisations of point, linear and areal concepts and related to each other by means of cartographic and topological measures. One application of this semantic annotation is the discovery of geographic information on the web (Torres et al., 2007). The assumption is that cartographic databases are already semantically annotated (that is, they contain *populated ontologies*).

Arpinar et al. (2006) report on a framework for constructing and populating spatial ontologies and for carrying out spatial analytics based on these data. They emphasise the use of spatial reasoning techniques (such as proximity) in order to detect associations.

Hence, the need to annotate spatial databases with semantic information is perceived increasingly in the literature. In this context, Klien and Lutz (2005) discuss the automatic annotation of existing datasets with concepts defined in an ontology. Their approach emphasises spatial relations between features rather than individual feature properties.

Thomson and Béra (2008) are working on an approach to infer land use from spatial databases by use of ontological modelling. They present a methodology that creates increasingly complex spatial aggregates starting from atomic concepts like house, garden, or road.

3. Formalisation of urban space descriptions

3.1 Sources of urban space descriptions

We use 'ontology' in the sense of the engineering sciences, where it is usually defined as an explicit specification of a shared conceptualisation (Gruber, 1993). It is thus an attempt to capture the knowledge in a certain domain in a systematic way by breaking it down into the types of entities (*concepts*) that exist and the *relations* that hold between them. Ontologies can be classified regarding the degree of formalisation into informal (written in natural language), semi-formal (restricted language), and formal (artificial language) ontologies (Agarwal, 2005). Based on the level of formalisation we differentiate between four sources to find salient concepts:

1. Thesauri, such as WordNet (2008) and the Canadian Thesaurus of Construction Science and Technology (TC/CS, 2008), contain a controlled set of concepts that are important for the respective domain and provide cross-links between these concepts.
2. Literature of a specific domain, e.g. urban morphology (Conzen, 1969), contains a controlled set of concepts, since people involved in a certain field agree on a basic terminology. Descriptions of the concepts are richer than in thesauri, but the effort of extraction is also higher.
3. Controlled experiments with human subjects in order to find out how the urban space is perceived at different scales. For instance, Thomson and Béra (2007) report on a human subject study in the context of creating ontologies for the extraction of land uses from topographic maps.
4. Descriptions of the urban space can also be elicited from documents such as tourist guides. Mining such descriptions from the internet is also a very promising approach (e.g., Milne et al., 2006; Edwardes and Purves, 2007; for a formal approach, see Maedche, 2002).

The latter two approaches can also be termed a *naive geography* approach (Egenhofer and Mark, 1995) and are especially interesting since they capture how non-experts think about urban space.

3.2 Formalisation of extraction ontologies

It should be noted that according to the ontology definition by Gruber (1993), there can be multiple ontologies for the same concept depending on the purpose the ontology is modelled for. Our purpose is to model concepts for their detection in spatial databases. Such an ontology, in our case, has been built for the extraction of terraced houses as they are conceptualised in *urban morphology*. Relevant concepts of the domain were extracted from a thesaurus of urban morphology (Jones and Larkham, 1991); several case studies (e.g. Conzen, 1969) and a compendium about "The English Terraced House" (Muthesius, 1982) gave more insight in the understanding of the concepts. Figure 2 illustrates a typical terraced house settlement. Modelling of the terraced house ontology is explained in more detail in Lüscher et al. (2008).

Fig. 2. Terraced house settlement in Sheffield (Muthesius, 1982:110).

The ontology constructed in this way is shown in Figure 3. Such ontologies can be authored in standard ontology editors such as Protégé (2008). They are stored as OWL files (OWL, 2008) and imported into the pattern recognition engine. The engine starts with populating concepts at the lowest level, i.e. concepts that are atomic and not defined in terms of their relations to other concepts. Instantiation is then successively transferred to higher level concepts. The inference process is discussed in more detail in the next chapter.

We like to clarify the notions used in figure 3. Firstly, OWL differentiates between *datatype properties* (tagged with prefix DTP in figure 3) and *object properties* (tagged with prefix OP). Datatype properties relate a concept to an attribute, whereas object properties relate two (or more) concepts. This differentiation is also important in the orchestration of the pattern recognition process. Datatype properties denote restrictions on characteristics of concepts. Object properties constitute a composition hierarchy and therefore also a processing hierarchy: Before instances of *row of houses* are inferred, the corresponding instances of *connected set of houses* have to be known.

The basic object property commonly used in ontologies is the superclass-subclass relationship *is-a*. For general ontology editing tools, all other relations such as *hasArea*, *adjacent* etc. are identifiers that have to be interpreted by the application that uses the ontology. Therefore, a tool that compares two ontologies would simply check whether in both ontologies *terraced house* has a property *DTP:hasArea* to another concept called *small footprint*, but it would not interpret the properties in a spatial sense. The interpretation is carried out by the pattern recognition engine, which also has to provide a vocabulary of spatial primitives and properties the domain expert is confined to, constituting a *domain ontology* for spatial pattern recognition.

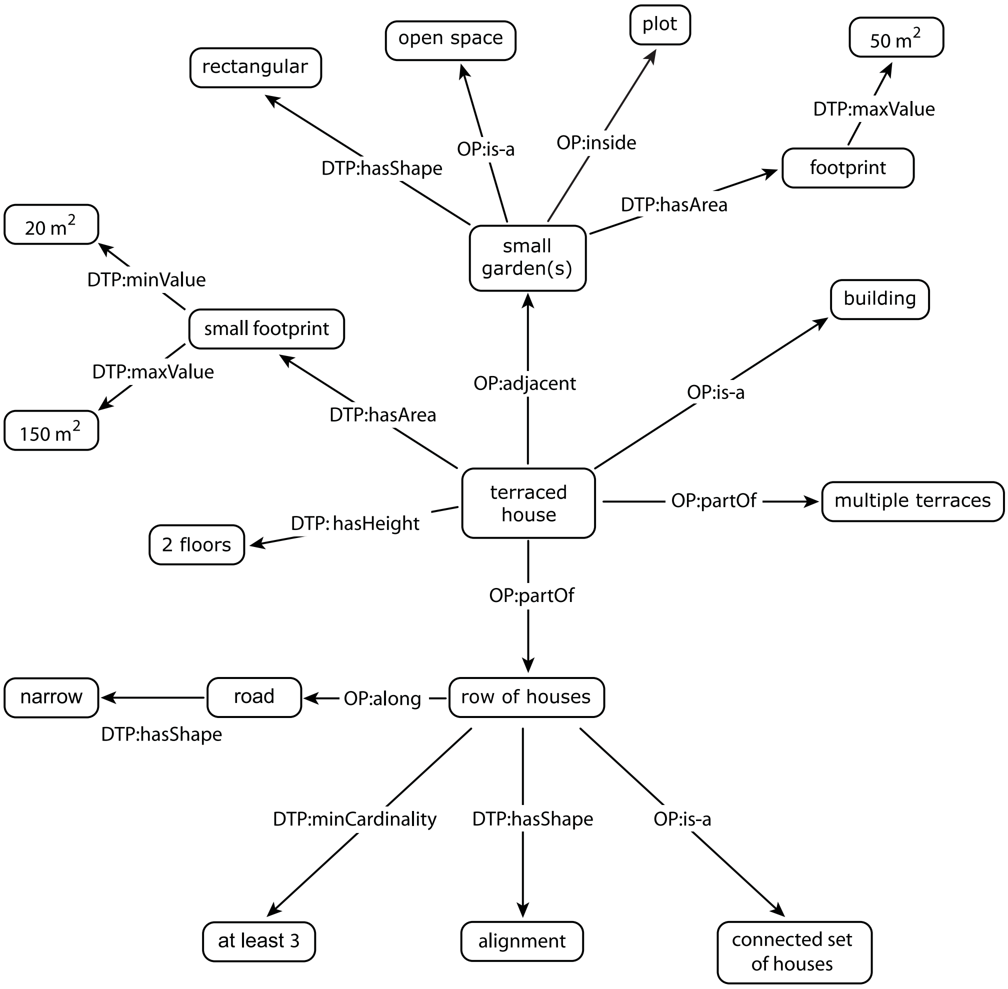


Fig. 3. An ontology of terraced houses suited for data enrichment. DTP: Datatype property. OP: Object property.

4. Derivation of a computational model

4.1 Inference of complex concepts

In order to carry out the data enrichment, concepts have to be related to measurable units, that is, the ontology has to be turned into a *computational model*. On the one hand a computational model consists of primitive algorithms and measures; on the other hand there must be a mechanism of inference that composes complex higher level concepts out of lower level concepts.

One of the standard applications of ontologies is to infer whether two concept descriptions are equal, or whether one concept description subsumes the other. To this end, Description Logics (DL), a family of logics for reasoning with categories, are employed. However, our aim is to populate ontologies with instances, and DL are weak when dealing with uncertainty. Therefore other types of 'inference' procedures have to be used in the presence of uncertainty.

The idea was to follow a modular approach which simulates forward reasoning known from Description Logics, but uses probabilistic inference methods, therefore combining advantages of the model-based approach with predictive power of probabilistic reasoning: We start populating the ontology with lowest level concepts such as *forming an alignment* or *narrow road*. Using cartographic measures or a cartographic pattern recognition algorithm, spatial objects (or a compound of objects) get assigned some *congruence value* with between 0 and 1, where 1 means the object is most likely an instance of that concept, and 0 means it is most likely not.

For higher level concepts (such as *row of houses*), the congruence value has to be inferred from congruence values of the composing lower level concepts. The question was which method would be most suitable for use in ontology-driven pattern recognition. Three alternative approaches for inference were explored, all of which are standard approaches in the pattern recognition and pattern matching domain: Weighted summation; Joint Bayes classifier; and Support Vector Machines (SVM). The experiments were carried out on a simplified part of the ontology, i.e. only the concepts *row of houses*, *multiple terraces* and *small footprint* were used and the concept *terraced house* was inferred thereof. Implementation details of necessary basic cartographic pattern recognition algorithms are discussed in Lüscher et al. (2008).

In the following sections of this chapter, we give some theoretical background for understanding the three approaches. Produced results are presented and compared in the following chapter.

In order to understand the theory, we need to explain the notions *feature* and *feature space*. In pattern recognition terminology, characteristics of an object that is to be classified are expressed as *features*. A feature is a specific property of the object (such as a building footprint), and the set of all features used for classification comprise the *feature space*. For a higher level object, the feature space is constituted by congruence values of the composing features. For example, the feature space of *row of houses* has four dimensions which are formed by the concepts *connected set of houses*; at *least 3 houses*; *forming an alignment*; and *aligned with narrow road*.

4.2 Weighted summation

The simplest method is to calculate a weighted sum of the congruence values of composing concepts:

$$con(C_i, R_k) = \left(\sum w_j con(C_j, R_k)\right) \Big/ \sum w_j \quad (1)$$

Where $con(C_j, R_k)$ is the congruence value of a constituent concept of C_i and the weight w_j is an influence value of the subconcept.

For some of the low level concepts, congruence values are to be defined by setting thresholds by looking at known instances of the concept. Hence, their congruence value is either 0 or 1. For example, the congruence value for *small footprint* equals 1 if the building area is between 20m^2 and 150m^2, and 0 otherwise.

The calculation of congruence values starts with the patterns at the bottom and then propagates iteratively to higher order concepts. At the end of this process, all spatial objects have a congruence value that denotes their similarity to the concepts defined in the ontology.

This approach has some disadvantages: Firstly, results depend heavily on the weights in the summation equation. Secondly, the setting of strict thresholds (Boolean logic) for low level concepts is somewhat undesirable, since it demands some profound domain knowledge and does not reflect uncertainty in the data – there might be a terraced house with a footprint larger than 150m^2, it is just fairly unlikely. Because of these reasons, two supervised machine learning approaches were evaluated, which are both commonly used in pattern recognition in images (computer vision). Instead of defining classification thresholds, the operator has to provide pre-classified training data.

4.3 Joint Bayes classifier

Bayesian inference is profoundly covered by Duda et al. (2001), Rice (1988) and Russel and Norvig (2003). Assume that the value of a variable C is conditionally dependent on a set of feature variables $F_1,..., F_n$. The prediction for a given realisation $F_1 = f_1,\ldots, F_n = f_n$ is the value c which maximises the conditional probability:

$$\hat{C} = \arg\max_c P(C = c \mid F_1 = f_1 \wedge \ldots \wedge F_n = f_n) \tag{2}$$

Using Bayes' theorem, equation (2) can be reformulated such that probabilities can be learned from training data:

$$\hat{C} = \arg\max_c P(F_1 = f_1 \wedge \ldots \wedge F_n = f_n \mid C = c) P(C = c) \tag{3}$$

Given a test dataset with known classifications, the learning part consists of estimating the probability distributions for possible outcomes of C. Let's consider C as *terraced house* (having outcomes 'true' and 'false') and a two feature variables F_1 denoting *small footprint* and F_2 denoting *row of house* present (both with outcomes 'true' or 'false'). The probability distributions can be estimated by simple counting:

$$\begin{gathered} P(terraced) = \frac{\#terraced}{\#terraced + \#(\neg terraced)} \\ P(small\ footprint \wedge row\,of\ houses \mid terraced) = \frac{\#(small\ footprint \wedge row\,of\ houses \wedge terraced)}{\#terraced} \end{gathered} \tag{4}$$

In equation (4) *terraced* denotes *terraced = true* and ¬*terraced* denotes *terraced = false*. The hash (#) denotes 'number of occurrences in the training data'.

Until now, only discrete variables were considered. Rather than deriving instances of the concept *small footprint* by setting thresholds, we can also learn this concept from training data by introducing a continuous variable *footprint*. In the prediction function for continuous variables, the joint conditional probability is replaced by a joint probability density function $g_{C=c}$:

$$\hat{C} = \arg\max_c g_{C=c}(\vec{f}) = \arg\max_c g_{C=c}(F_1 = f_1,\ldots, F_n = f_n) \tag{5}$$

The joint probability density function is estimated by calculating a kernel density using sample points from the training data as input – thus, there is no division into learning part and classification part:

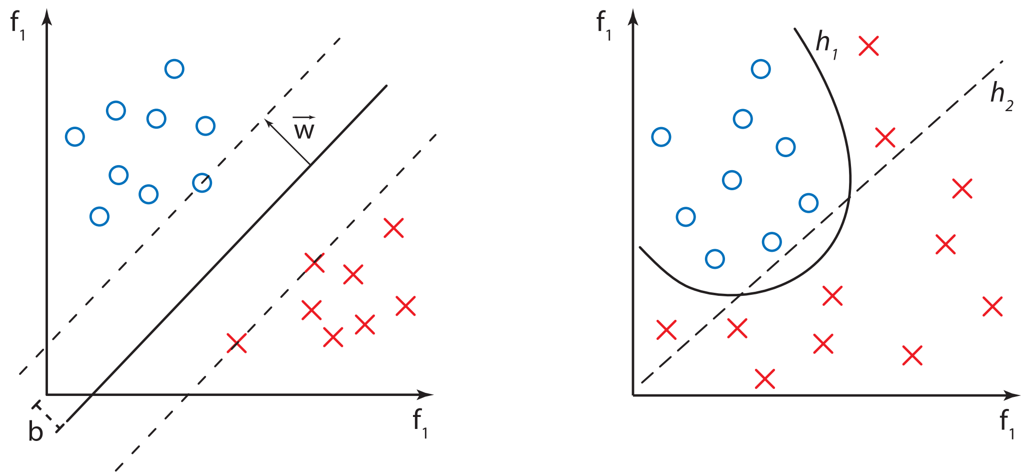
$$\hat{g}_C(\vec{f}) = \frac{1}{N\vec{h}} \sum_{i=1}^{N} K\left(\frac{\vec{f} - \vec{f}_i}{\vec{h}}\right), \text{ where } K(\vec{x}) = \frac{1}{(2\pi)^{N/2}} e^{-0.5 * \vec{x}^T \vec{x}} \tag{6}$$

$\vec{f}$ is the feature vector of a spatial object that needs to be classified. $\vec{f}_i$ are the feature vectors of sample points, N is the number of samples and $\vec{h}_i$ the bandwidth, a smoothing factor for the density function. In our experiments, we used a bandwidth of 5m for the *footprint*, and a bandwidth of 0.05 for *row of houses* and *multiple terraces*.

The composition of concepts to higher order concepts is similar to weighted summation, except that congruence values of composed concepts are conditional probabilities $g_{C\,=\,true}$ given the congruence values of their constituting concepts. Binary decisions are made by comparing whether the probability $g_{C\,=\,true}$ or $g_{C\,=\,false}$ is larger.

4.4 Support Vector Machines

In order to identify the predictive power of Bayesian networks and separate effects of inappropriate training data, we implemented a second machine learning approach that uses the same training data. Support Vector Machines (SVM, cf. Duda et al., 2001; Cristianini and Shawe-Taylor, 2000) determine a hyperplane that separates classes in feature space such that the margin between the plane and sample points is maximised (figure 4a). Often, linear boundaries do not optimally separate classes. Therefore, SVM provide a method to simulate non-linear transformation into a higher dimensional space, where linear separation may be possible. It consists of using non-linear kernel functions. Figure 4b illustrates a case where a non-linear boundary h_1 classifies better than a linear boundary h_2.



(a) (b)

Fig. 4. (a) Hyperplane $\vec{w}-\vec{b}=0$ that maximises the margins between classes. (b) A case where a non-linear boundary classifies better than a linear boundary.

We used a Radial Basis Function (RBF) kernel. SVM classification with a RBF kernel has two parameters: C is a penalty function for misclassified sample points of training data, and $\gamma > 0$ is an exponent factor in the RBF function. They were determined by systematically trying out different parameter combinations and evaluating the classification accuracy by means of four-fold cross-validation. The parameter combination $C = 2^{16}$, $\gamma = 2^{5.8}$ yielded the highest accuracy. A second combination $C = 2^{16}$, $\gamma = 2^{3}$ was selected to determine overfitting effects of high γ values.

Other than the two previous approaches, SVM return only binary classifications and not congruence or probability values.

5. Evaluation of inference approaches

The ontology was transferred into a prototype for ontology-driven pattern recognition implemented in the Java programming language. JTS (Vivid Solutions, 2008) and OpenJUMP (2008) were used as a basis for spatial analysis, data display and data manipulation. SVM experiments were carried out using the open source software package libsvm (Chang and Lin, 2001).

We like to evaluate the three different inference methods in terms of predictive power (comparison of classification results with 'ground truth'), and in terms of their integration into the framework of ontology-driven pattern recognition.

5.1 Predictive power

Buildings were extracted from Ordnance Survey MasterMap® for two different areas. The authors manually attributed buildings in both datasets with 'terraced' / 'not terraced' by visual inspection. Besides MasterMap®, aerial photographs provided by Google Earth were used for the manual classification.

Dataset A was used as training dataset. 764 of the houses in the dataset were tagged with 'terraced' and 1 039 with 'not terraced'. Dataset B served as test dataset and was exhaustively classified with 'terraced' or 'not terraced'. The class 'not terraced' was not split up further, i.e. whether they represented some other type of residential housing or industrial / commercial housing. Of 22 505 buildings, 4 913 (21.83%) were classified as terraced house and 17 592 (78.17%) as some other type of housing.

Table 1 presents the confusion matrices for comparison of automated prediction with manual classification. In the following interpretation it is important not to mistake the manual classification for ground truth: It approximates ground truth generally well, but some border cases were hard to classify even by a human inspector. Therefore, an uncertainty factor stemming from human interpretation has to be kept in mind. Since the weighted summation provides a congruence value but not a classification, a threshold of $con(\textit{terraced house}, R_k) \geq 0.7$ was arbitrarily chosen to denote an instance of *terraced house*.

Weighted summation	Manual classification		
		¬terraced	terraced
	¬terraced	74.77%	3.78%
	terraced	3.39%	18.05%

(a) Weighted summation

Joint Bayes	Manual classification		
		¬terraced	terraced
	¬terraced	70.53%	3.93%
	terraced	7.64%	17.90%

(b) Joint Bayes Classifier

SVM	Manual classification		
		¬terraced	terraced
	¬terraced	74.38%	12.46%
	terraced	3.79%	9.38%

(c) SVM with $C = 2^{16}$, $\gamma = 2^{5.8}$

SVM	Manual classification		
		¬terraced	terraced
	¬terraced	70.41%	5.36%
	terraced	7.76%	16.47%

(d) SVM with $C = 2^{16}$, $\gamma = 2^{3}$

Table 1. Comparison of results produced by pattern recognition algorithms and manual classification.

Classification accuracy can be measured by means of *precision* and *recall*. Precision indicates the probability that a terraced house found by a classification algorithm is indeed a terraced house in reality. Recall indicates the probability that a real terraced house is found by the classification algorithm:

$$precision = \frac{true\ positives}{true\ positives + false\ positives}$$
$$recall = \frac{true\ positives}{true\ positives + false\ negatives} \qquad (7)$$

A comparison of precision and recall shown in table 2 reveals that the weighted summation produces the highest accuracy. Most striking is that a relatively good classification can be achieved with a quite simple model. Machine learning approaches have a slightly lower precision of around 70%. The SVM run with $C = 2^{16}$, $\gamma = 2^{5.8}$ produced a very low recall of only 43%. It seems indeed that this choice of parameters led to overfitting to the training data. The Joint Bayes classifier and the SVM run with $C = 2^{16}$, $\gamma = 2^3$ produced similar results. Our interpretation is that both methods are equally applicable to the purpose, and the choice of training data is more important than the choice of the classification algorithm.

Classification approach	Precision	Recall
Weighted summation	84.17%	82.70%
Joint Bayes classifier	70.08%	81.99%
SVM with $C = 2^{16}$, $\gamma = 2^{5.8}$	71.24%	42.95%
SVM with $C = 2^{16}$, $\gamma = 2^3$	67.97%	75.45%

Table 2. Comparison of classification accuracy.

It is also important not only to evaluate how well (or poorly) a classification approach performed statistically, but also *where* misclassifications happened. This was done by visual inspection of false positives and false negatives. We could discern between three different types of *false positives*:

- In some cases, the building structure indicated a terraced house, but it was unlikely when considering the spatial context. For example, a single row of terraced houses in an otherwise industrial area seems strange. Considering the garden structure, as it is modelled in the ontology, could avoid many of these misclassifications. All three approaches produced errors of this type.
- Looking at the confusion matrices of the machine learning approaches, it stands out that their precision is lower because they produced many false positives. Most of these were rows of garages that were classified as terraced houses. Rows of garages are often perfectly aligned and homogeneous, so they could be terraced houses, but they have a footprint of only 10 – 15m^2. In the training dataset, there were no garages at all, and therefore there was no evidence for the machine learning approaches that perfectly aligned, homogeneous rows of buildings of this size are *not* terraced houses. Extending the current set of training data could therefore have avoided this type of misclassification.
- A third but less frequent type of misclassification occurred when houses were connected by garages or sheds and were mistaken as row of houses.

False negatives occurred mostly where rows of terraced were less regularly arranged at the boundary of large terraced areas or in more recent residential areas, because regular arrangement of terraces was somewhat out of fashion in the second half of the last century.

5.2 Adequacy of inference approaches for ontology-driven pattern recognition

Besides good classification results, two criteria are essential for the adequacy for ontology-driven pattern recognition. Firstly, the orchestration process should be carried out with little user interaction. Secondly, classifications (or congruence values) should be transparent, allow a user to click on a concept instance to obtain a plausible explanation for its classification.

One main difference between the classification approaches is that weighted summation doesn't use training data, whereas the other two methods do. Therefore, the domain expert has to define thresholds for some sub-concepts having non-binary congruence values, building footprint being one example. In order to overcome the problem, fuzzy logic could be used. In this case, the domain expert needs to define membership values instead of strict thresholds. Also, weights in equation (1) that reflect the importance of sub-concepts have to be defined by the domain expert. The setting of thresholds / membership functions and weights requires considerable expertise, but it reflects the idea of ontological modelling which is purely declarative. The summation equation and resulting congruence values are easily comprehensible by domain experts.

Joint Bayes classification is a simple statistical approach that doesn't require any thresholds or weights. A further advantage of this approach is that it delivers a probability for possible instantiations of concepts analogous to the congruence values of the weighted summation approach. However, it can be difficult to interpret values for joint probabilities if there are more than three features. Therefore, this approach is not as transparent as the weighted summation approach.

When using Support Vector Machines, parameters have to be trained for every composed concept in the ontology. There are algorithms such as grid search (Chang and Lin, 2001) to obtain useful parameters, but they are computationally expensive and only suitable if training data are not noisy (otherwise there is a tendency for overfitting parameters). Classification results can only be understood with a good deal of knowledge of the SVM method, and by interactively investigating plots such as the one presented in figure 4. A further disadvantage is that there is no classification certainty for individual objects.

We conclude this chapter with a table that summarises the arguments discussed.

Classification approach	Predictive power	Computational complexity	Calibration effort	Transparency
Weighted summation	++	++	-	+
Joint Bayes classifier	+(+)[1]	+	++	-
Support Vector Machines	+(+)[1,2]	--	--	-

[1] Depending on training data
[2] Depending on parameter selection

Table 3. Comparison of inference approaches.

6. Discussion

In order to carry out spatial data enrichment, we employed ontologies to model domain knowledge about urban structures. One of the main benefits of the approach is enhanced transparency: Assumptions about the structure (and relationships) of concepts are explicitly stated. Ontologies can be modelled and validated in collaboration with domain experts (making sure they fit their view), and different conceptualisations of the same terms can be compared, for example in order to reveal culturally different conceptions. Also, a modular and flexible pattern recognition approach is made possible.

Furthermore, we explored how such ontologies could be transferred into a pattern recognition process, concept composition being one part of the problem. Besides prediction accuracy, the effort for model calibration is critical since the inference process has to be carried out for every higher level concept. In this respect, the Joint Bayes classifier is preferred, since it requires no calibration parameters at all, but further testing is needed to check whether the prediction accuracy could be improved with more and better training data. Finally it should be noted that the chaining of Bayes classifiers in the pattern recognition process corresponds to Bayesian networks (Russel and Norvig, 2003) and has therefore its theoretical foundation.

One of the long-term goals of our research is to transfer the ontology automatically into a pattern recognition process. The automatic orchestration of the inference process is complicated by dependencies of concepts in the ontology. For example, the concept *multiple terraces* in the ontology shown in figure 3 presumes on the one hand that some *terraced houses* have already been detected, on the other hand having detected multiple terraces also supports the assumption of having found *terraced houses*. We are currently investigating whether a rule-based inference mechanism based on a categorisation of object properties would solve such problems. For instance, *is-a* properties have to be evaluated first since superclasses are inferred before subclasses.

Another difficulty lies in the choice of the implementation of low level concepts. For example, there exist multiple algorithms that calculate alignments and hence could be used for deriving rows of houses. In the current implementation, the arrangement of roads of houses respective to roads has to be derived in a second step. On the other hand, there is an algorithm by Burghardt and Steiniger (2005) which directly derives houses aligned along a road and which might be better suited for this specific application – but it is not as modular as our approach. The influence of the choice of extraction algorithms for low level concepts on the final classification will have to be evaluated further in future work.

7. Conclusions

We firmly believe that ontological models of the geographical reality are important because they allow abstraction of cartographically relevant patterns over large scale changes and for different usages. The automated semantic annotation of cartographic databases is a key success factor in this respect. Therefore, in this paper, we discussed a novel approach to cartographic pattern recognition, focusing on the complexity of deriving a sound ontological model suitable for pattern recognition, and on alternative methods of inferring composite concepts. The approach renders cartographic pattern recognition more transparent by explicitly stating assumptions of pattern structure, and more flexible through its modularity. One of the problems that yet have to be addressed is the automatic orchestration of the pattern recognition process.

Our future work will focus on the implementation of more concepts and a further formalisation of the pattern recognition vocabulary in order to find out whether the orchestration of the extraction process could be automated better; on the evaluation of the choices of algorithms for low level concepts and their influence on extraction results; and on experiments using people to study where and how they visually detect concepts such as terraces.

Acknowledgment

The research reported in this paper is part of the PhD project of the first author. Funding by the Swiss State Secretariat for Education and Research (SER) through COST Action C21 (project ORUS, grant no. C05.0081) is gratefully acknowledged.

References

Agarwal, P. (2005). "Ontological considerations in GIScience", *International Journal of Geographical Information Science*, **19** (5), 501–536.

Arpinar, I.B. et al. (2006). "Geospatial Ontology Development and Semantic Analytics", *Transactions in GIS*, **10** (4), 551–575.

Bergsdal, H., Brattebø, H., Bohne, R.A. and Müller, D.B. (2007). "Dynamic Material Flow Analysis of Norway's Dwelling Stock", *Building Research & Information*, **35** (5):557–570.

Bertin, J. (1967, 1999). *Sémiologie graphique: Les diagrammes – Les réseaux – Les cartes.* Paris: Les réimpressions des Éditions de l'École des Hautes Études en Sciences Sociales, 431 pages. (Original edition published 1967 by Gauthier-Villars, Paris)

Burghardt, D. and Steiniger, S. (2005). "Usage of Principal Component Analysis in the Process of Automated Generalisation", *Proceedings of 22nd International Cartographic Conference*, July 9–16, 2005, La Coruña, Spain.

Camacho-Hübner, E. and Golay, F. (2007). "Preliminary insights on continuity and evolution of concepts for the development of an urban morphological process ontology", in Teller, J., Lee, J.R. and Roussey, C. (eds.), *Ontologies for Urban Development.* Studies in Computational Intelligence, Vol. 61, 95–108.

Chang, C.-C. and Lin, C.-J. (2001). *LIBSVM : a library for support vector machines*. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm

Conzen, M.R.G. (1969). *Alnwick, Northumberland: A Study in Town-plan Analysis*. London: Institute of British Geographers, 131 pages.

Cristianini, N. and Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge: Cambridge University Press, 189 pages.

Duda, R.O., Hart, P.E., Stork, D.G. (2001). *Pattern Classification*. New York et al.: John Wiley & Sons, Inc., 654 pages.

Edwardes, A.J. and Purves, R.S. (2007). "A theoretical grounding for semantic descriptions of place", *7th International Symposium on Web and Wireless Geographical Information Systems (W2GIS 2007)*, November 28–29, 2007, Cardiff, UK.

Egenhofer, M.J. and Mark, D.M. (1995). „Naive Geography", in Frank, A.U. and Kuhn, W. (eds.), *Spatial Information Theory: A Theoretical Basis for GIS. International Conference, COSIT '95*, September 21–23, Semmering, Austria, Lecture Notes in Computer Science 988, 1–15.

Gómez-Pérez, A., Fernández-López, M. and Corcho, O. (2003). *Ontological Engineering: With Examples from the Areas of Knowledge Management, E-commerce and the Semantic Web.* Berlin: Springer, 415 pages.

Gruber, T.R. (1993). "A translation approach to portable ontology specifications", *Knowledge Acquisition*, **5** (2), 199–220.

Grünreich, D. (1992). "ATKIS – a topographic information system as a basis for GIS and digital cartography in Germany", in Vinken, R. (ed.), *From digital map series to geo-information systems*, Geologisches Jahrbuch, Federal Institute of Geosciences and Resources, Hannover, A(122), 207–216.

Jones, A.N. and Larkham, P.J. (1991). *Glossary of Urban Form.* Historical Geography Research Series no.26, London: Institute of British Geographers.

Klien, E. and Lutz, M. (2005). "The Role of Spatial Relations in Automating the Semantic Annotation of Geodata", in Cohn, A. G. and Mark, D. M. (eds.), *Spatial Information Theory, International Conference, COSIT 2005*, Ellicottville, NY, USA, September 14–18, 2005, LNCS 3693, 133–148.

Lüscher, P., Burghardt, D., Weibel, R. (2007). „Ontology-driven Enrichment of Spatial Databases", *10th ICA Workshop on Generalisation and Multiple Representation*, August 2–3, 2007, Moscow, Russia.

Lüscher, P., Weibel, R., Mackaness, W. (2008). „Where is the Terraced House? On The Use of Ontologies for Recognition of Urban Concepts in Cartographic Databases", accepted for presentation at *SDH 2008 – The 13th International Symposium on Spatial Data Handling*, June 23–25, 2008, Montpellier, France.

Maedche, A.D. (2002). *Ontology Learning for the Semantic Web.* Dordrecht: Kluwer Academic Publishers Group, 272 pages.

Milne, D., Medelyan, O. and Witten, I.H. (2006). "Mining Domain-Specific Thesauri from Wikipedia: A case study", *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence*, December 18–22, 2006, Hong Kong, China.

Muthesius, S. (1982). *The English Terraced House*. New Haven and London: Yale University Press, 278 pages.

OpenJump (2008). http://openjump.org/. Accessed 18.05.2008.

OWL (2008). http://www.w3.org/2004/OWL/. Accessed 19.05.2008.

Protégé (2008). http://protege.stanford.edu/. Accessed 19.05.2008.

Rice, J.A. (1988). *Mathematical Statistics and Data Analysis.* Pacific Grove: Wadsworth and Brooks, 594 pages.

Russel, S. and Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*. Upper Saddle River: Prentice Hall, 1080 pages.

TC/CS (2008). Website of the Canadian Thesaurus of Construction Science and Technology. http://irc.nrc-cnrc.gc.ca/thesaurus/. Accessed 27.02.2008.

Thomson, M.K. and Béra, R. (2007). "Relating Land Use to the Landscape Character: Toward an Ontological Inference Tool", *Proceedings GIS Research UK 15th Annual Conference (GISRUK-2007)*, April 11–13, 2007, Maynooth, Ireland.

Thomson, M.K. and Béra, R. (2008). "A Methodology for Inferring Higher Level Semantic Information from Spatial Databases". *Proceedings of the GIS Research UK 16th Annual Conference (GISRUK 2008)*, April 2–4, 2008, Manchester, UK.

Torres, M., et al. (2005). "Ontology-Driven Description of Spatial Data for Their Semantic Processing", in Rodríguez, M.A. et al. (eds.), *GeoSpatial Semantics: First International*

Conference, GeoS 2005, November 29–30, 2005, Mexico City, Mexico, LNCS 3799, 242–249.

Torres, M., et al. (2007). “Retrieving Geospatial Information into a Web-Mapping Application Using Geospatial Ontologies”, in Nguyen, N.T. et al. (eds.), *Agent and Multi-Agent Systems: Technologies and Applications.* First KES International Symposium, KES-AMSTA 2007, May 31–June 1, 2007, Wroclaw, Poland, 267–277.

Vivid Solutions (2008). http://www.vividsolutions.com/. Accessed 18.05.2008.

Winter, S. (2002). „Ontologisches Modellieren von Routen für mobile Navigationsdienste“, in Kelnhofer, F. and Lechthaler, M. (eds.), *TeleKartographie und Location-Based Services*. Schriftenreihe der Studienrichtung Vermessungswesen und Geoinformation, Technische Universität Wien, 111–124.

WordNet (2008). http://wordnet.princeton.edu/. Accessed 27.02.2008.