

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

MATHEMATISCH-NATURWISSENSCHAFTLICHE FAKULTÄT  
INSTITUT FÜR NEUROBIOLOGIE  
LEHRSTUHL FÜR KOGNITIVE NEUROWISSENSCHAFTEN

BACHELORARBEIT DER KOGNITIONSWISSENSCHAFTEN

# Identifying hierarchical structures in feature graphs by using spectral clustering

vorgelegt von  
Frederic Becker

betreut von  
Prof. Dr. Hanspeter Mallot

August 13, 2019

# Erklärung

Hiermit erkläre ich,

- Dass ich diese Arbeit selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.
- Dass die eingereichte Arbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahren gewesen ist.

---

Ort, Datum

Unterschrift

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Models of spatial representation . . . . .	1
1.2	Hierarchy in spatial representations . . . . .	4
1.3	Hierarchy in human navigation . . . . .	5
1.4	Hierarchy in the feature graph . . . . .	6
<b>2</b>	<b>Methods</b>	<b>9</b>
2.1	Hierarchical graph structure . . . . .	9
2.2	Construction of the feature level . . . . .	10
2.3	Construction of higher-order levels . . . . .	12
2.3.1	Spectral clustering . . . . .	12
2.3.2	Similarity functions . . . . .	14
2.3.3	Clustering algorithm . . . . .	15
2.3.4	Connectivity . . . . .	16
2.3.5	Statistical program . . . . .	16
2.4	Experimental conditions . . . . .	16
<b>3</b>	<b>Results</b>	<b>19</b>
3.1	Graphical results . . . . .	19
3.2	Cohesion . . . . .	26
3.3	Expansion variance . . . . .	26
3.4	Separation . . . . .	27
3.5	Overlap . . . . .	27
3.6	Cluster similarity . . . . .	28
3.7	Border similarity . . . . .	28
<b>4</b>	<b>Discussion</b>	<b>29</b>
4.1	Metric properties of iterative levels . . . . .	29
4.2	Iterative vs. direct clustering . . . . .	30
4.3	Influence of abstraction factors . . . . .	31
4.4	Environmental properties of iterative levels . . . . .	32
<b>5</b>	<b>General discussion</b>	<b>34</b>
<b>6</b>	<b>Conclusion</b>	<b>36</b>
<b>7</b>	<b>References</b>	<b>37</b>

## Zusammenfassung

Eine Möglichkeit sich an eine kontinuierliche Umgebung zu erinnern, besteht darin Assoziationen von Bildmerkmalen in einem sogenannten „Feature Graphen“ zu speichern. Dabei entsteht eine räumliche Repräsentation von sehr hohem Detailgrad, die eine effiziente Wegplanung erschwert und abstrakte Raumkonzepte wie Orte oder Regionen vermissen lässt. Allerdings gibt es experimentelle Befunde die darauf hinweisen, dass der Mensch seine Umwelt hierarchisch kodiert und dass er verschiedene Detailgrade bei räumlichen Aufgaben nutzt, u.a bei der Wegplanung. (McNamara, 1986; Stevens & Coupe, 1978; Wiener & Mallot, 2003). Daher wurde mittels spektralem Clustering ein Feature Graph, der die Tübinger Innenstadt modelliert, auf die Existenz dicht vernetzter Sub-Netzwerke untersucht. Diese Netzwerke wurden dann als übergeordnete Einheiten in einen hierarchischen Graphen, bestehend aus mehreren Ebenen, eingebettet. Wichtig ist, dass diese Abstraktion iterativ mit abnehmendem Detailgrad angewendet wurde und ausschließlich auf der Vernetzung von Features basierte, ohne metrische Informationen zu berücksichtigen. Darüber hinaus wurden die übergeordneten Einheiten basierend auf der Sichtbarkeit der Features in der Umgebung lokalisiert. Die so entstandenen Regionalisierungen wurden hinsichtlich ihrer metrischen Eigenschaften analysiert. Mit zunehmender Abstraktionsstärke entstanden deutlichere Regionalisierungen, welche sich für hierarchische Wegeplanung qualifizierten. Außerdem zeigte sich, dass die Wahl der direkten und iterativen Abstraktionsfaktoren das Clustering beeinflusste. Clusterungen geringerer Granularität teilten jedoch eine gemeinsame Grundstruktur. Diese scheint eine dominante Grundstruktur des Feature Graphen widerzuspiegeln. Die Resultate legen nahe, dass ein Feature-Graph hierarchische Strukturen der Umgebung in seiner Konnektivität kodiert.

## Abstract

One approach to remember a continuous environment is by storing associations of image features in a graph. Hereby, a spatial representation of very high granularity arises which impedes efficient path-planning and lacks abstract concepts of space such as places or regions. However, there is profound evidence that humans encode their environment hierarchically and that they make use of this nested structure in path-planning and other tasks (McNamara, 1986; Stevens & Coupe, 1978; Wiener & Mallot, 2003). Therefore, an actual feature graph modeling the downtown of Tuebingen was searched for densely connected networks with spectral clustering. These networks were reduced to superordinate units embedded in a hierarchical graph consisting of multiple levels. Importantly, this abstraction was applied iteratively with decreasing granularity and it was solely based on the interconnectivity of features, disregarding any metric information. Further, superordinate units were localized in the environment based on the visibility area of features. The so created spatial divisions were analyzed with regard to metric properties. With increasing abstraction more explicit divisions of the environment developed which showed suitability for hierarchical path-planning. It was shown that the choice of direct and iterative abstraction factors influenced the clusterings. Clusterings of low granularity, however, shared a common composition of clusters which seemed to correspond to dominant structures within the feature graph. Conclusively, a feature graph seems to encode hierarchical structures of the environment in its connectivity.

# 1 Introduction

Humans are able to navigate efficiently in the city they live in or even describe the outline of the city to tourists. Additionally, we use different concepts of space such as neighborhoods, cities and states without difficulties. All these abilities require a comprehensive representation of our environment, stored somehow in our memory. This spatial representation is often referred to as cognitive map and its exact structure is a matter of research (Tolman, 1948). In this chapter different models of spatial representation will be discussed. All of them lack hierarchical-architectures even though there is strong evidence that humans use different levels of spatial detailing for instance in path-planning tasks (Wiener & Mallot, 2003). Therefore, in chapter two an approach based on spectral clustering is presented, that analyses a spatial representation of very high granularity for hierarchical structures. In chapters three to six results are presented and discussed.

## 1.1 Models of spatial representation

One possibility to think of the cognitive map is as a graph-like representation, for instance in form of a place-graph (G. Franz, Mallot, & Wiener, 2005). This representation evolves by associating a place with all its neighboring places and the corresponding actions to reach these. A process like this can be modeled by storing places as vertices and transitions between these as edges in a graph. Thereby, structures comparable to local transport maps develop as illustrated in Figure 1b). The model is mainly inspired by associative learning and the concepts of place-cells and head-direction-cells. O'Keefe and Dostrovsky (1971) discovered place-cells in the hippocampus of rodents. These specific cells adapt their firing rate depending on the rodents position in an environment. That way, place-cells are tuned to specific areas, whereas head-direction-cells are tuned to a specific head orientation of a rodent (Taube, Muller, & Ranck, 1990). These might encode allocentric directions, i.e directions referenced to a global orientation and not to the current body orientation. A place-graph can integrate local representations of places to a global representation of the environment by considering place connectivity and allocentric directions.

There is also evidence that our spatial representation is based on views. A view is considered a visual scene captured from a fixed position. Röhrich, Hardiess, and Mallot (2014) showed that the orientation of a recalled place is influenced by one's relative position to it. This was demonstrated by asking pedestrians in different locations to sketch the outline of a nearby square. The orientation of the drawings correlated with the approach direction of an imagined walk from the subjects posi-

tion to the target square. The orientation-variability of the recalled square suggests that its representation is composed of oriented parts, such as views (Röhrich et al., 2014). The finding of spatial view cells in humans and primates, which are tuned to specific views of an environment, also supports this idea (Ekstrom et al., 2003; Rolls, 1999).

In accordance with these findings, Schölkopf and Mallot (1995) proposed a view-graph model for spatial representation. This representation evolves by associating a view with all its possible subsequent views and the corresponding actions to reach these. In this way, a place is divided into the different views that are visible from it. For example, one can associate the view of a door with the view of the room lying behind that door. With multiple such association pairs, namely view sequences, one can traverse the environment. Specifically, the graph consists of nodes that represent views and labeled edges that represent transitions between these. A simple view-graph can be seen in Figure 1c). Since views are oriented transitions can be described not only with allocentric (i.e. world-centered, like "north") but also with egocentric (i.e. self-centered, like "left") directions. Further, the authors constructed a view-graph for a maze consisting of discrete views. It was shown that the network carried information about the structure of the maze and that it was suitable for path-planning.

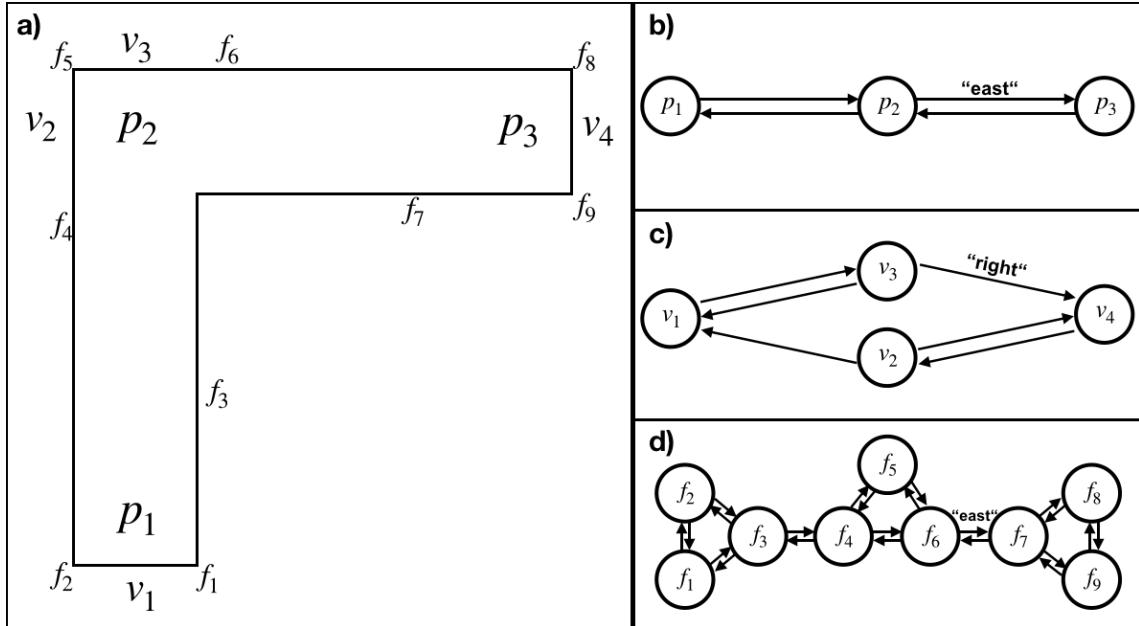
However, in real life we perceive visual input continuously while moving through our environment. The integration of a stream of views in such a discrete model is a major challenge. M. O. Franz, Schölkopf, Mallot, and Bülthoff (1998) proposed the idea of 'snapshots' for the view-graph. Snapshots are representative views with a certain divergence tolerance that discretize the space of views in sections. This concept shall extend the view-graph to open environments although the selection of representative views in an open environment is not trivial and might be influenced by various criteria such as salient landmarks and transitions.

Another approach to integrate stream of views in a discrete model is to disassemble views in components of high granularity. In general, different views cover different visual scenes containing different visual components. A dense structure of components is then able to encode views by the principle of population coding, i.e. each view can be described by a set of components. The component's granularity determines how sensitive the system is to view shifts. By using this approach, continuity of views can be approximated by a fine grid of component combinations.

Baumann (2019) implemented such a dense structure of view-components in form of a micro-snapshot graph. This graph consists of image features i.e. visual characteristics that are steadily detectable with regard to distance and viewing angle.

For example, a feature can be a characteristic element of a shop-window or roof. Each feature is visible from a certain area in the environment, its so-called place field. Depending on the features visibility, place fields vary in size and outline but may also overlap. Hence, a grid of place field combinations emerge that divides the environment in discrete sections. Allocentric transitions between neighboring but not identical place fields are included in the graph as labeled edges between the corresponding features. A simple feature graph can be seen in Figure 1d).

It is important to note that the so far introduced representations do not correspond to a map in terms of its metric properties. The graphs include local relationships between their elements disregarding explicit metric information such as distances. In spite of that, path-planning still can be realized. For example, Baumann (2019) created actual feature graphs of virtual environments. A continuous path-planning algorithm then computed nearly optimal paths in the virtual environments based on the corresponding feature graphs.



*Figure 1:* Actual spatial representations for a simple maze. Own illustration based on Schölkopf and Mallot (1995). a) The outline of a maze consisting of two corridors.  $p$  marks different places within the maze,  $v$  marks different views perceivable in the maze and  $f$  marks different features located on the walls of the maze. b) An allocentric place-graph of the maze. c) An egocentric view-graph of the maze. d) An allocentric feature-graph of the maze.

## 1.2 Hierarchy in spatial representations

A feature graph illustrates a representation with very high granularity. In real life we do not use features to depict places or routes. Instead we use concepts of higher abstraction such as squares, cities or countries. There is profound evidence that our spatial representation includes such abstractions through a hierarchical architecture (Hirtle & Jonides, 1985; McNamara, 1986; Stevens & Coupe, 1978).

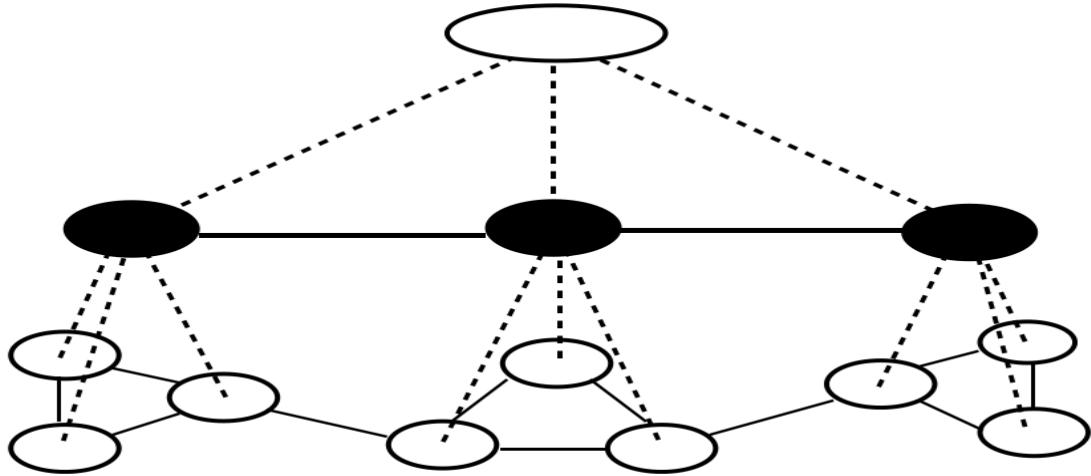
Stevens and Coupe (1978) asked subjects for the geographic relation of five location pairs. For instance the direction from Portland, Oregon to Toronto. The subjects estimated the direction to be north-east, even though Toronto lies east of Portland. According to the authors the distortion towards north was caused by the geographic relation of the countries in which the cities are located, namely Canada lies north of the USA. This effect was also observed when using artificial cities and regions. The results showed reliably that region relations influenced the estimation of city relations. Due to that, the authors concluded that our spatial representation includes superordinate units and relations in a hierarchical structure.

Moreover, McNamara (1986) conducted a spatial priming experiment. Priming is based on the idea that the recognition speed of a stimulus is influenced by a preceding stimulus called primer. The closer the primer and the questioned stimulus are associated in memory the faster the recognition (McNamara, 1986). Participants learned an environment consisting of objects either as a map or by navigating in it. This environment was segmented in four regions. After the learning phase, the participants were questioned in a reaction task whether certain locations (referenced via object names) were part of the environment or not. The response latencies showed the following pattern: In case the preceding location was within the same region as the questioned location, the response was quicker than in the case where the preceding location was part of a different region. This effect was especially meaningful in the case where both types of preceding locations had the same euclidean distance to the questioned location. Hence, locations primed locations of the same region more than locations of different regions.

In addition McNamara (1986) observed effects of regionalization in direction judgments, tasks and distance judgments tasks. In direction judgments, distortions corresponding to superordinate relations as described by Stevens and Coupe (1978) were registered. In distance judgments, subjects overestimated distances between locations of different regions whereas they underestimated distances between locations of same region. As artificial regions had a significant effect on subjects behavior the authors argued that spatial relations are encoded hierarchically.

Hierarchical representations of space are often proposed as graphs with multiple

levels of abstraction which is illustrated in Figure 2 (McNamara, 1986; Stevens & Coupe, 1978; Wiener & Mallot, 2003). The higher the level the more abstract the spatial concepts become. A higher-level node represents a superordinate spatial structure that maintains links with its components of lower abstraction. Spatial relations are represented by edges within the same layer, i.e. only concepts of the same abstraction level hold spatial relations.



*Figure 2:* A spatial graph with multiple levels. The higher the level the more abstract the spatial structures, for example place-district-city. Dotted lines illustrate which vertices are summarized to superordinate units. Solid lines illustrate connectivity within a level. Own illustration based on Wiener and Mallot (2003).

### 1.3 Hierarchy in human navigation

We do not only construct a model of our environment but also realize complex path-planning. There is evidence that humans make use of spatial hierarchies in path planning to increase planning efficiency (Huys et al., 2015; Wiener & Mallot, 2003).

Huys et al. (2015) conducted a path-planning experiment using a simple environment of six locations. Each location offered two transitions with earnings on it. The subjects were asked to plan a path of a given length which maximizes earnings. The results showed that subjects chose long paths suboptimal. However the paths often consisted of optimal solutions for shorter subpaths. Based on this, the authors suggested that the subjects planned paths hierarchically, i.e. long paths were fragmented in short sections which were then solved. Moreover, the fragmentation chosen by the subjects reduced computational effort almost optimally and mirrored

environmental properties such as salient rewards. The authors concluded that subjects used hierarchical fragmentation to plan paths more efficiently. However, the environment used in this study did not imply any regions, so it is unclear whether the fragmentation was task related or representation related.

Wiener and Mallot (2003) conducted a path-planning experiment to investigate the influence of explicit regions on path-planning. Participants learned a maze-like virtual environment populated with twelve landmarks. The environment was divided into three semantic regions of identical shape. Each region was defined by four neighboring landmarks belonging to the same object type, for instance animals. The participants' task was to pass certain landmarks using the fastest route possible. The result showed that in case of two shortest routes participants preferred the one that crossed fewer region boundaries. Moreover, the authors showed in a second experiment that participants entered target regions rather sooner than later. In summary, the study demonstrated the influence of explicit regions on subjects' path-planning.

According to Wiener and Mallot (2003) the participants' behavior implied that humans make explicitly use of spatial hierarchies in route planning by considering region connectivity. The authors proposed a path-planning scheme that implements this idea by considering different levels of spatial abstraction. This "fine-to-coarse" heuristic plans proximate steps in a fine-space level and more distant steps in a coarse-space level. Or more specifically, the sub-path from a start location to an adjacent region is planned as a place sequence, whereas the sub-path from the adjacent region to the target region is planned as a region sequence. This can best be understood in an example: Imagine to plan a route from Tuebingen to Athens. According to the heuristic, one would first plan which cities to pass in order to leave Germany and then plan which countries to pass on the way to Greece. The concept of hierarchical path-planning is crucial as it can explain both, efficient path-planning as observed in Huys et al. (2015) and region effects as observed in Wiener and Mallot (2003).

## 1.4 Hierarchy in the feature graph

As explained above, there is profound evidence that our spatial representation is structured hierarchically and that we make use of this nested structure in navigation tasks. However, proposed representations such as the view or feature graph do lack in hierarchical structures. In this bachelor thesis an actual feature graph of downtown Tuebingen is analyzed for implicit hierarchical structures. To do so, the graph is iteratively clustered to superordinate units. Further, it is assessed whether

the superordinate units show properties of spatial regions and how sensitive these are to the degree of abstraction.

The feature graph is a spatial representation of very high granularity. Therefore, it is especially in need of hierarchical structures. From a practical point of view, route planning in the feature graph is very inefficient due to a large number of nodes to be considered. With the help of superordinate structures and hierarchical path-planning this process can be drastically improved. From a theoretical point of view, the feature graph lacks explicit representations of spatial concepts. However, certain concepts can be found implicitly in the graph as a set of features. A particular view can be described by features visible in that view. A particular place can be described by features visible from that place. A particular region can be described by features visible in that region. As superordinate units depict sets of features, they are theoretically capable of representing spatial concepts.

In practice, features are projected into the environment based on the location where they were first detected. This means that a clustering of the feature graph corresponds to a certain degree to a division of the environment. Based on this idea the present study examines whether a hierarchical clustering of the feature graph mirrors a useful hierarchical division of the environment. Hierarchical clustering means that a feature graph is supplemented with higher-order levels, remodeling it into a multi-level graph as illustrated in Figure 2. Nodes of higher-order are constructed by clustering entities of the next-lower level. Importantly, all clusterings are solely based on feature interconnectedness explicitly disregarding metric or any other information. Due to that, it is even more of interest whether the superordinate units mirror useful spatial cluster in the environment.

The criterion 'useful' is hard to define in this context. The clustering method is designed to identify networks in the graph that show dense internal connectivity. It is expected that these networks mirror characteristic structures of the environment (e.g. squares). Still the obtained spatial divisions are not assessed primarily in terms of meaningfulness, i.e. whether the divisions assemble intuitively plausible divisions. This is because studies suggest that humans regionalize environments differently (Huys et al., 2015; Montello, Goodchild, Gottsegen, & Fohl, 2003) and after various criteria (Hirtle & Jonides, 1985; Schick, Halfmann, Hardiess, Hamm, & Mallot, 2019). Instead, the divisions are assessed on their usefulness for hierarchical path-planning.

More specifically, the divisions should consist of compact, non-overlapping spatial clusters of approximately equal size. Such a division divides the environment evenly in distinct regions like a grid does. This structure is beneficial for hierarchical

path-planning because a path segmentation exists, such that sub-paths exclusively traverse one region. This kind of segmentation describes an efficient abstraction of the path whereas in a more scattered regionalization overlapping regions hinder an efficient abstraction of the path. Further, it is evaluated whether the obtained spatial divisions are mirrored by dominant structures of the feature graph. To do so, four hierarchical feature graphs with different degrees of abstraction are constructed. The resulting spatial divisions are then compared.

## 2 Methods

### 2.1 Hierarchical graph structure

Hierarchical graphs include different levels of abstraction by multiple graph levels. Each level is a graph itself and is defined by a set of vertices and edges. The initial level  $L_0$  of the hierarchical graph constructed in this study equals a feature graph. Thus, level  $L_0$  consists of features and is from now on called the feature level. Each of the  $N_0$  features forms one vertex:

$$V(L_0) = \{f_i | i = 1, \dots, N_0\}. \quad (1)$$

A higher-order level  $L_k$  represents a partition of the feature level, i.e. all features are grouped in  $N_k$  subsets such that one feature belongs to one subset only. These subsets are called clusters. The elements of a partition, namely the clusters  $C_{ki}$ , are vertices in the higher-order level  $k$ :

$$V(L_k) = \{C_{ki} \subset V(L_0) | i = 1, \dots, N_k\} \quad (2)$$

$$\text{and } \forall i, j \in \{1, \dots, N_k\}, i \neq j : C_{ki} \cap C_{kj} = \emptyset$$

$$\text{and } \bigcup_{C \in V(L_k)} C = V(L_0).$$

Each level  $L_{k+1}$  is iteratively constructed on the basis of its predecessor level  $L_k$ . The vertices of the predecessor level  $L_k$  are clustered into groups. By uniting the corresponding sets of features in each group, the clusters of level  $L_{k+1}$  are created. This iterative clustering scheme can be seen best in an example as follows:

$$V(L_0) = \{f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8\}$$

$$V(L_1) = \{\{f_1, f_2\}, \{f_3, f_4\}, \{f_5, f_6, f_7\}, \{f_8\}\} = \{C_{21}, C_{22}, C_{23}, C_{24}\}$$

$$V(L_2) = \{\{f_1, f_2\}, \{f_3, f_4\}, \{f_5, f_6, f_7, f_8\}\} = \{C_{21}, C_{22}, (C_{23} \cup C_{24})\} = \{C_{31}, C_{32}, C_{33}\}$$

This clustering scheme leads to a number of desired properties. First, there is decreasing granularity. The number of clusters is reduced the higher the abstraction level. Thus, the average amount of features in a cluster increases per level which leads to broader spatial clusters with less granularity:

$$N_0 > N_1 \dots > N_K. \quad (3)$$

Second, there is a monotony of clusters across all levels. A higher-order level can not only be understood as a mathematical partition of the feature level but also as a conceptual partition of all levels beneath it. In detail, each cluster of a higher level can be seen as the unification of clusters of lower level. It has the effect that once a cluster of features is formed in a low level this grouping can be found in all higher levels. For example, the area of a city belongs to the same district and the same country and is not divided. This property is given by:

$$\forall l, h \in \{1, \dots, K\}, l < h : \forall C_{low} \in V(L_l) \exists C_{high} \in V(L_h) : C_{low} \subseteq C_{high}. \quad (4)$$

Third, there is monotony of borders across all levels. Borders of a higher level can be described by the edges of the feature level which connect features of different clusters. Monotony of borders means that the borders of a high level can be found as such in all levels beneath it. For example, the border between France and Germany can also be seen on the granularity level of districts:

$$\forall l, h \in \{1, \dots, K\}, l < h : Borders(L_h) \subset Borders(L_l) \quad (5)$$

$$Borders(L_k) := \{\{f_a, f_b\} \in E(L_0) | f_a \in C_{ki}, f_b \in C_{kj}, i \neq j\}. \quad (6)$$

In a multilevel graph there are two kinds of edges. On the one hand, there are inter-level edges which connect vertices of two successive levels. These depict which vertices of the predecessor level form the clusters of the next level. On the other hand, there are intra-level edges which connect vertices within one level. In higher-order levels two vertices are connected when there is an edge on feature level, that connects their corresponding feature sets:

$$E(L_k) := \{\{C_{ki}, C_{kj}\} | \exists \{f_a, f_b\} \in E(L_0) \text{ with } f_a \in C_{ki}, f_b \in C_{kj}, i \neq j\}. \quad (7)$$

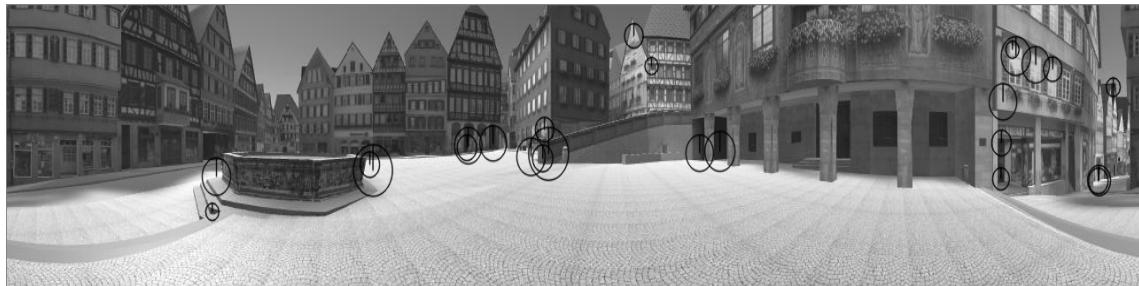
This way all intra-level edges are determined by the structure of the feature level. The intra-level connections are the focus area of this study, from now on these are referred to as edges.

## 2.2 Construction of the feature level

The feature level is constructed with an algorithm developed by Baumann (2019). This algorithm provides a virtual agent which explores a 3D model of the old town of Tuebingen, Germany by random walk. The walk consists of random steps of fixed size guiding the agent from one position to another. In each position the agent

receives a 360° image of its surrounding which it analyses for "speeded up robust features" (SURF) (Bay, Ess, Tuytelaars, & Gool, 2008). This process is illustrated in Figure 3. Up to 30 SURF features per position are encoded and stored as vertices in a complementary version of the feature graph. This version encodes feature co-occurrences rather than feature transitions. Two features are connected with an undirected edge in case they are visible at the same time, i.e in case their place-fields overlap. When the agent moves and not a single feature of the previous position is visible in the current position, the graph risks to become disconnected. In this situation features from the previous position are connected randomly with features of the current position. This variant of the feature graph is chosen because its connectivity constitutes place-field similarity rather than place-field transition. Because the clustering process is based on connectivity it will tend to extract dominant place-field structures. This is preferred, as it is expected that these place-field structures mirror environmental characteristics better than dominant place-field transitions.

To locate a feature's place-field in the environment, the position of the features detection is saved. In order to limit the size of the graph the agent explores just a part of the virtual environment including the central town square as well as a loop of alleys. The resulting feature level shows a very dense structure and can be seen in Figure 4. It includes 38,776 features detected in 4,511 different positions and 168,566 edges. A feature has 8.7 connections on average. In one position, 8.6 new features are detected on average.



*Figure 3:* Exemplary visual input of the agent in the virtual environment; a 360° image of the central town square. Detected SURF features are marked with black circles.



*Figure 4:* The feature level projected into the environment based on the location where a feature was detected for the first time. The environment consists of the central town square (Marktplatz) with its characteristic fountain (black circle) and a loop of alleys (Hirschgasse, Kornhausstraße, Marktgasse). Features are represented by red points. Note that usually multiple features overlay in one position. Co-occurrence edges between features are drawn in black.

## 2.3 Construction of higher-order levels

### 2.3.1 Spectral clustering

A superordinate level is constructed by clustering the predecessor level which represents an undirected and unweighted graph. Therefore, a graph clustering algorithm called spectral clustering is applied following the tutorial of Von Luxburg (2007). Spectral clustering groups the vertices of a graph in  $k$  groups based on the spectrum of a weight matrix. The number of clusters has to be set as parameter  $k$  in advance.

In a first step, a weight matrix  $W$  is constructed by weighting the edges of the graph by a similarity function  $s$ :

$$W : w_{i,j} = \begin{cases} s(i,j) & \text{if vertex } i \text{ is adjacent to vertex } j, \\ 0 & \text{otherwise.} \end{cases} . \quad (8)$$

The more similar two vertices, the higher the similarity score should be. Now, one would like to cluster the weighted graph such that the within-cluster-weights are high and the between-cluster-weights are low. This means that the vertices within a cluster should show high similarity and vertices of different clusters should show low similarity. Additionally, it is important to avoid correct but undesirable solutions, namely the separation of single vertices with low weights from the remaining graph. Mathematically, the minimum Ratio-Cut in the graph is searched:

$$\text{RatioCut}(C_1, \dots, C_k) = \frac{1}{2} \sum_{i=1}^k \frac{\sum_{a \in C_i, b \in \bar{C}_i} w_{a,b}}{|C_i|}. \quad (9)$$

This problem is NP-Hard, meaning that no deterministic algorithm exists yet that can solve this problem in practicable runtime for large graphs (Von Luxburg, 2007). At this point, spectral clustering offers an approximation to the problem using the unnormalized Laplacian  $L$  of the weight matrix  $W$ :

$$L : l_{i,j} = \begin{cases} \sum_{k=1}^N w_{i,k} & \text{if } i = j, \\ -w_{i,j} & \text{otherwise.} \end{cases} \quad (10)$$

One approach to think of the Laplacian is that it projects the weighted graph in a  $N$  dimensional space where  $N$  is the number of vertices. Each row of the Laplacian contains coordinates of a data point that corresponds to a vertex. In this space one would like to find  $k$  groups of data points such that the within-group-distances are low and the between-group-distances are high. For large dimensions it is difficult to identify such groups efficiently as the required information is spread over all dimensions. With the help of the Laplacians spectrum dimensions are reduced and information is concentrated. Therefore, the smallest  $k$  eigenvalues  $(\lambda_1, \dots, \lambda_k)$  with their corresponding eigenvectors  $(u_1, \dots, u_k)$  of the Laplacian  $L$  are calculated. The  $k$  eigenvectors are then combined column-wise to a matrix  $U$ , which has the dimensions  $N \times k$ . Now, matrix  $U$  projects the weighted graph in a  $k$  dimensional space:

$$U = \begin{bmatrix} \vdots & \vdots \\ u_1 & \dots & u_k \\ \vdots & \vdots \end{bmatrix}_{N \times k}. \quad (11)$$

The main benefit of this procedure is that the data points in the  $k$ -dimensional space shall form  $k$  well distinguishable clusters. Mathematically it can be shown that this clustering problem approximates the Ratio-Cut problem (Von Luxburg,

2007). A dataset cluster algorithm is then able to form  $k$  groups of data points. This corresponds to the final graph clustering as each data point represents one vertex.

To sum up, a superordinate level is constructed by weighting the connection of the predecessor level with a similarity function, then projecting the weighted graph in a  $k$ -dimensional space where it is clustered with a dataset cluster algorithm.

### 2.3.2 Similarity functions

Spectral clustering is mainly influenced by the choice of the similarity function. The similarity function  $s_1$  is used for the first abstraction, namely from  $L_0$  to  $L_1$ :

$$s_1(f_i, f_j) = |N(f_i) \cap N(f_j)|, \quad (12)$$

$$N(f_i) := \{f_j | \{f_i, f_j\} \in E(L_0)\}. \quad (13)$$

This function describes the similarity of feature  $i$  and  $j$  by the amount of shared neighbors. It resembles the idea that features that are visible from similar areas so have similar place-fields should hold a similar adjacency structure. However, this score highly depends on an informative and dense connectivity of the graph. In more abstract levels this is not always the case. For example, vertices with just one edge may exist which would then have a similarity score of zero.

Due to that, the similarity function  $s_2$  is utilized for higher abstractions namely, from  $L_1$  to  $L_2$  and from  $L_2$  to  $L_3$ :

$$s_2(C_i, C_j) = 5 * \frac{|N(C_i) \cap C_j| + |C_i \cap N(C_j)|}{|N(C_i)| + |N(C_j)|} + 5 * e^{-0.02 * \min(|N(C_i)|, |N(C_j)|)}, \quad (14)$$

$$N(C) := \left( \bigcup_{f \in C} N(f) \right) \setminus C. \quad (15)$$

This function describes the similarity of cluster  $i$  and  $j$  based on their degree of interweaving. The first term is the core term. It computes the proportional overlap between the neighborhood of cluster  $i$  and the features of cluster  $j$  and vice versa. Importantly, this comparison happens on feature level meaning that both sets  $N(C)$  and  $C$  consist of features. Further, the neighborhood of a cluster is defined by the adjacency structure on feature level and not by the adjacency structure of the current abstraction level. This means not all features in a neighboring cluster are considered as neighbors, but only the ones that have connections to the cluster on feature level. This way the similarity is computed based on the level with the most structural

information, namely the feature level.

The second term is a bias term. After the first abstraction clusters of varying sizes may exist. This is especially an issue when two clusters of very different sizes are compared. The overlap may be proportionally big for a small cluster but proportionally small for a big cluster. This conflict is often dominated by a large neighborhood of big clusters. In order to compensate, the bias term adds additional weight to similarity scores of small clusters. The factors 5 and -0.02 were chosen experimentally in the developing process.

### 2.3.3 Clustering algorithm

In order to cluster the data points in the  $k$ -dimensional space the partitioning around medoids (PAM) algorithm by Kaufman and Rousseeuw (1990b) is applied. PAM selects  $k$  data points which represent the cluster centers as optimally as possible. The representative data points are called medoids. The algorithm begins by randomly selecting  $k$  data points as medoids. After which,  $k$  clusters are created by assigning all data points to the nearest medoid. Then the average dissimilarity is calculated which averages the euclidean distance of the data points to their associated medoids. This objective function indicates the tightness of clusters and thus the goodness of the clustering. Next, a swap phase minimizes this function. To do so, a non-medoid data point is swapped with a medoid. In case the objective function increases the swap is undone and the next data pair is swapped. In case the objective function decreases the swap is maintained and the swap phase restarts. No decrement in the entire swap phase indicates that the set of optimal medoids has been found and the algorithm terminates.

The main advantage of PAM is its robustness against outliers, which comes with high computational expenses for large data sets. Yet to realize the clustering of 38,776 data points in the first abstraction, the implementation of PAM for large datasets called CLARA is chosen (Kaufman & Rousseeuw, 1990a). This clustering algorithm does not apply PAM to the entire dataset but draws multiple samples to which PAM is applied. This procedure results in a set of medoids for each sample. The set of medoids with the lowest average dissimilarity for the entire data set is then selected as the final set. The idea of CLARA is to draw a 'good' sample that has a similar structure as the entire dataset. Hence, the optimal medoids of this 'good' data set should be similar to those of the whole data set.

To realize the first abstraction, namely from  $L_0$  to  $L_1$ , CLARA (25 samples of size 2500) is applied. In further abstractions, namely from  $L_1$  to  $L_2$  and from  $L_2$  to  $L_3$ , PAM is applied.

### 2.3.4 Connectivity

Navigating in a graph requires the graph to be connected, i.e. between each pair of vertices a path has to exist. A cluster of features represents a subgraph of the feature level that shall be used for navigation. Therefore, each cluster is tested on its connectivity after the clustering process. In case a cluster is not connected it is divided into its connected sub-clusters called components. The largest component remains as the original cluster. The other components are merged with existing clusters. More specific, a component is merged with the adjacent cluster it achieves the highest similarity score ( $s_2$ ) with. Since both are adjacent, the new cluster has to be connected again. This procedure maintains both, the number of clusters defined by the parameter  $k$  and the connectivity of clusters.

### 2.3.5 Statistical program

The implementation of the above described procedure is done with the statistical program R (R Core Team, 2017). The eigenvalues and eigenvectors are calculated with the package RSpectra (Qiu & Mei, 2018), the dataset is clustered with the CLARA/PAM implementation in the package cluster (Maechler, Rousseeuw, Struyf, Hubert, & Hornik, 2017) and the component analysis is done with the package igraph (Csardi & Nepusz, 2006). All calculations are done on a standard personal computer.

## 2.4 Experimental conditions

Overall four different hierarchical graphs based on the same feature level are created and assigned to two conditions.

In the direct reduction condition the feature level is directly clustered in six groups. The resulting graph consists of the feature level ( $L_0$ ) and a superordinate level ( $L_1$ ) including six superordinate units. This means that the feature level is abstracted by a factor of 6,462. The number of six is chosen after the eigengap heuristic. To identify the number of dominant clusters in a weighted graph, one can compare the development of the Laplacians smallest eigenvalues. The eigengap heuristic says that  $k$  should be chosen such that the first  $k$  eigenvalues are relatively small but the  $k + 1$  eigenvalue is noticeable larger (Von Luxburg, 2007). Graphically this is indicated by the first gap in the plot of the smallest eigenvalues of the Laplacian as displayed in Figure 5.

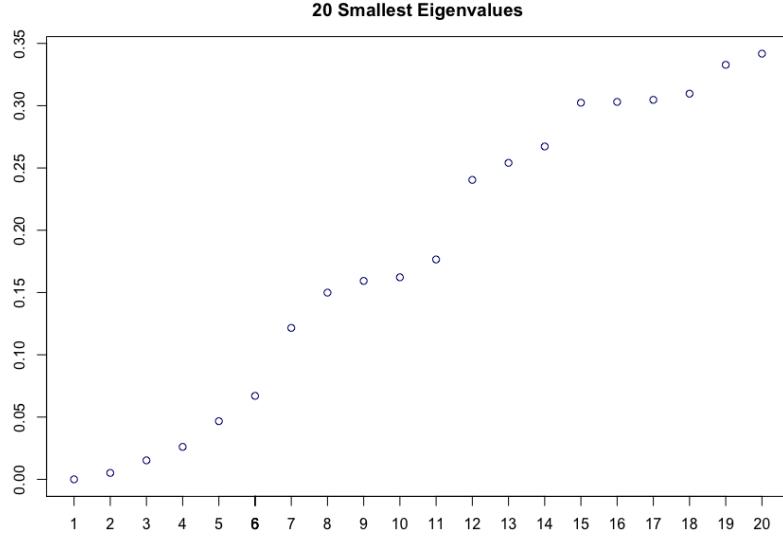


Figure 5: The 20 smallest eigenvalues of the Laplacian of the weighted feature level. The eigengap occurs after the sixth eigenvalue.

In the iterative reduction condition three different graphs are created. Each consisted of the feature level ( $L_0$ ) and three superordinate levels ( $L_1, L_2, L_3$ ). Since no method is known that can reasonably determine the number of levels or abstraction factors, they are chosen in a explorative manner. Hence, the number of clusters in levels one and two as well as the abstraction factors are varied between the graphs. The number of clusters in level three is set to six to maintain comparability between the graphs. The exact number of clusters and abstraction factors included in each graph is given in Table 1.

For each graph level a reference level including the same number of clusters is created. A reference level emerges by embedding all features in the two dimensional environment based on their position they were first detected. For each graph level, CLARA clusters the feature localizations in the corresponding number of clusters. This clustering is only based on the features metric localization and neglects any feature interconnectedness. In this case, CLARA tries to maximize metric proximity of features in one cluster. Thus, the resulting spatial divisions consists of compact, non-overlapping clusters of approximately equal size.

Using this constellation three questions can be assessed: First, whether the spatial divisions obtained in the iterative condition represent metric useful divisions. This can be achieved by contrasting the experimental divisions and the reference divisions. Second, whether the iterative scheme enhances the development of metric useful divisions in contrast to the direct reduction condition. This can be assessed by comparing the spatial divisions of the direct and iterative condition. Third, whether

the arbitrary chosen abstraction factors influence the shape of the spatial divisions. This can be assessed by contrasting the final six-cluster levels of the iterative condition.

Table 1

*The number of vertices/clusters per level per graph and the corresponding abstraction factors.*

	Direct	Iterative A	Iterative B	Iterative C
$N_3$		6	6	6
$N_2$		60	78	48
$N_1$	6	780	624	432
$N_0$	38776	38776	38776	38776

$\uparrow \frac{\div 10}{\div 13}$        $\uparrow \frac{\div 13}{\div 8}$        $\uparrow \frac{\div 8}{\div 9}$

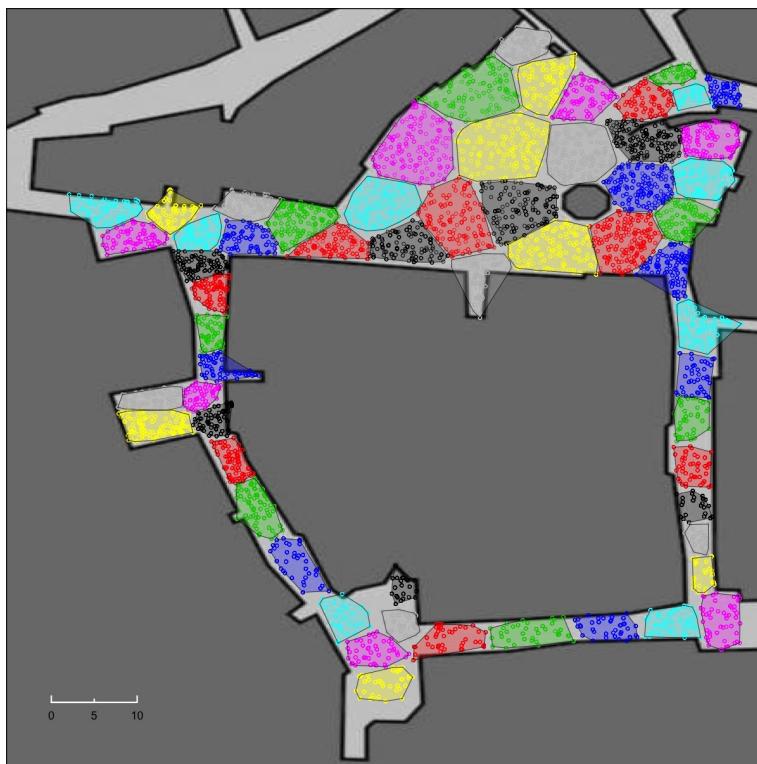
$\uparrow \frac{\div 50}{\div 62}$        $\uparrow \frac{\div 62}{\div 90}$

$\uparrow \frac{\div 6462}{\div 90}$

### 3 Results

Figure 6 to 18 illustrate the obtained spatial divisions indicated by each graph level. Features belonging to the same cluster have the same color. Additionally, each cluster's convex hull is plotted in the corresponding color. A convex hull is the smallest polygon that encloses all features of a cluster such that each straight connection of features lies fully inside the polygon. Note, different clusters may use the same color though they are distinguishable by distinct convex hulls. The spatial divisions are then analyzed in terms of cohesion and expansion variance as well as separation and overlap. Moreover, the different six-cluster divisions are compared in terms of cluster and border similarity.

#### 3.1 Graphical results



*Figure 6:* Reference division for Iterative Graph A - Level 2 - 60 clusters. This reference clustering was solely based on the features localizations. The clusters are compact, non-overlapping and approximately equal sized.

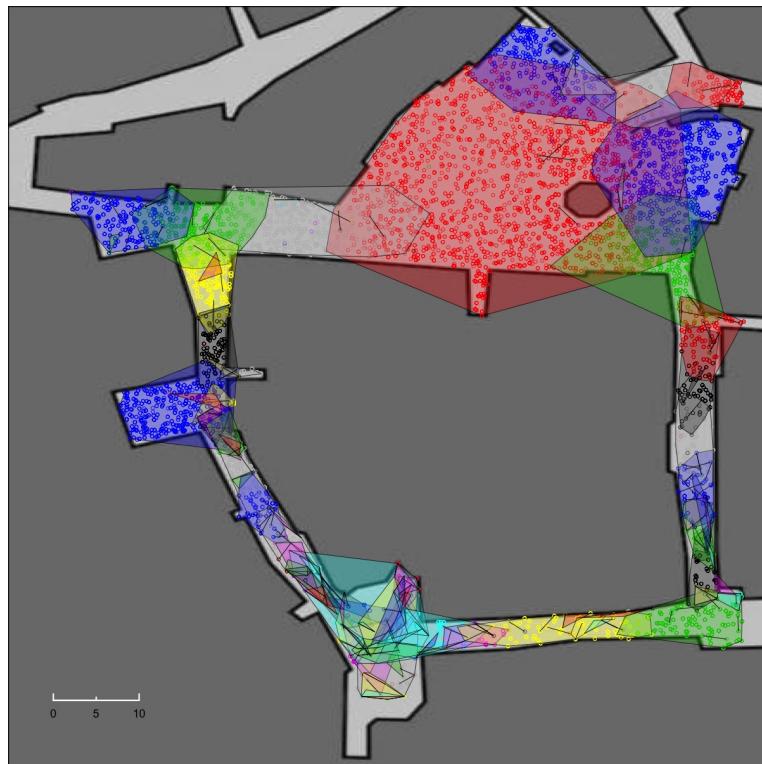
**Level 1**

Figure 7: Spatial division corresponding to: Iterative Graph A - Level 1 - 780 clusters.

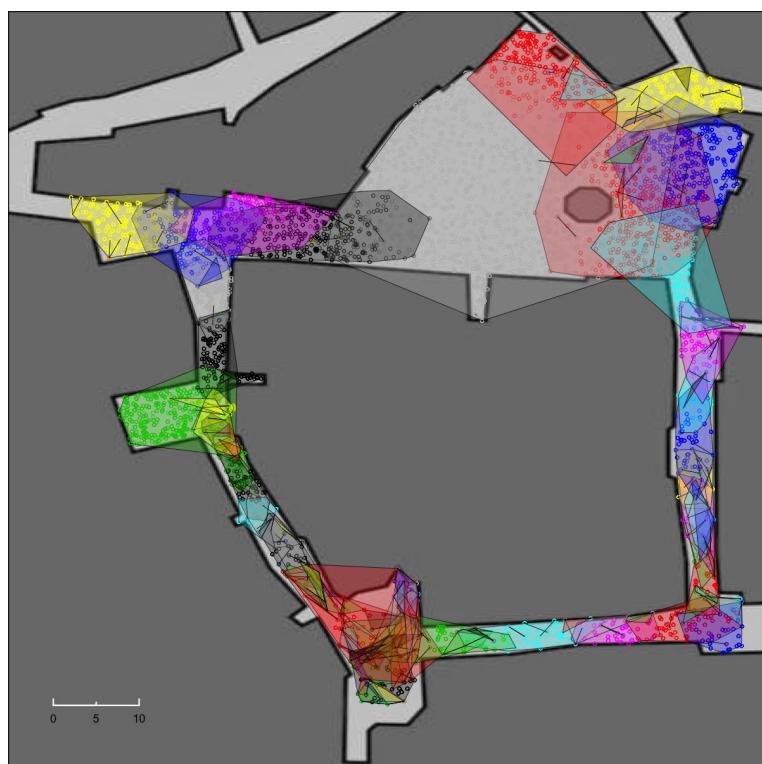


Figure 8: Spatial division corresponding to: Iterative Graph B - Level 1 - 624 clusters.

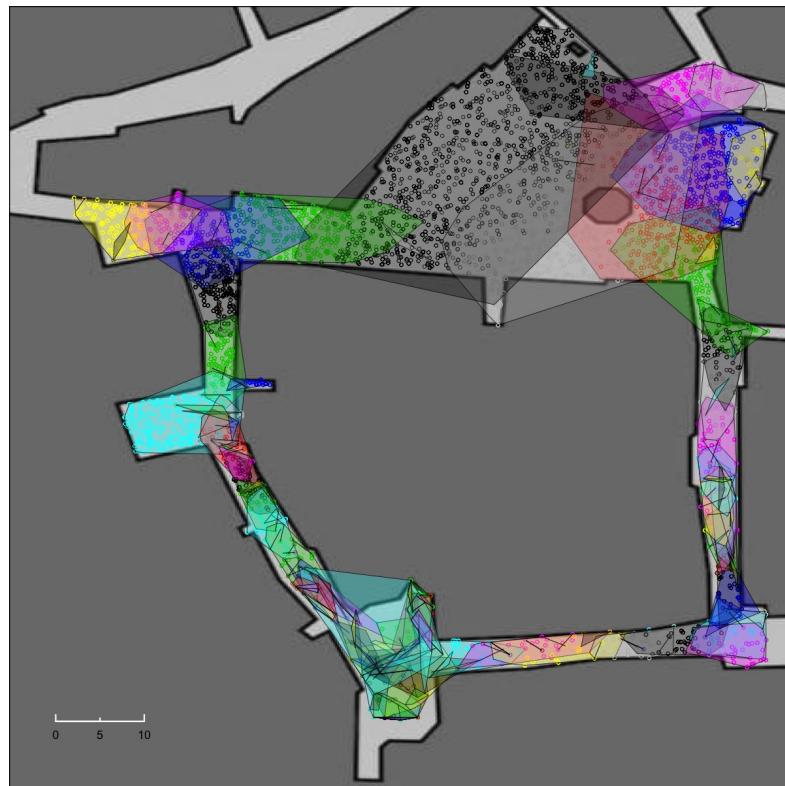


Figure 9: Spatial division corresponding to: Iterative Graph C - Level 1 - 432 clusters.

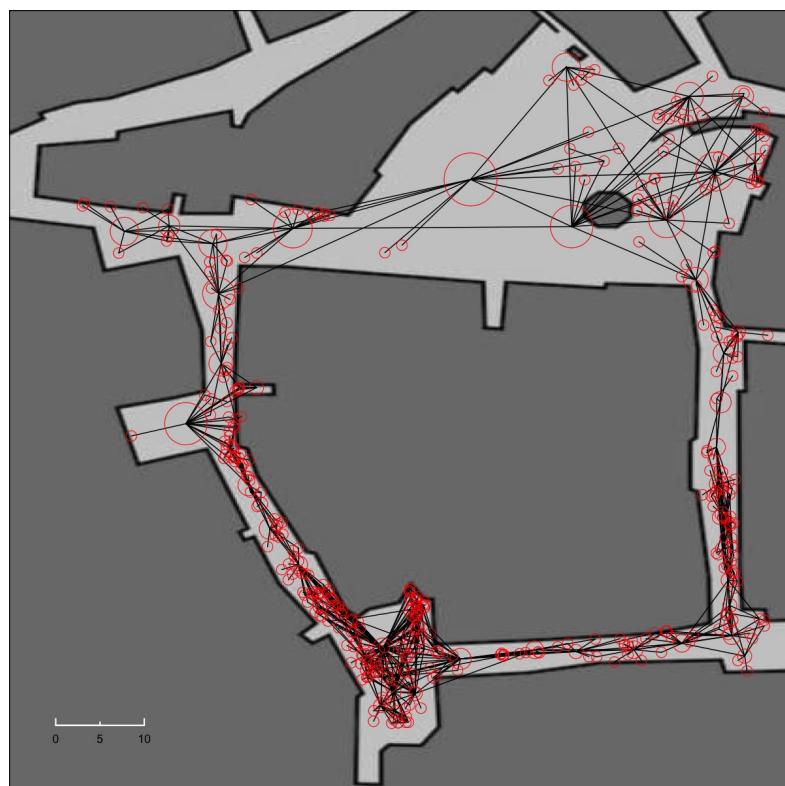


Figure 10: Network representation of: Iterative Graph C - Level 1 - 432 clusters. The number of features in a cluster is mirrored by the diameter of its node.

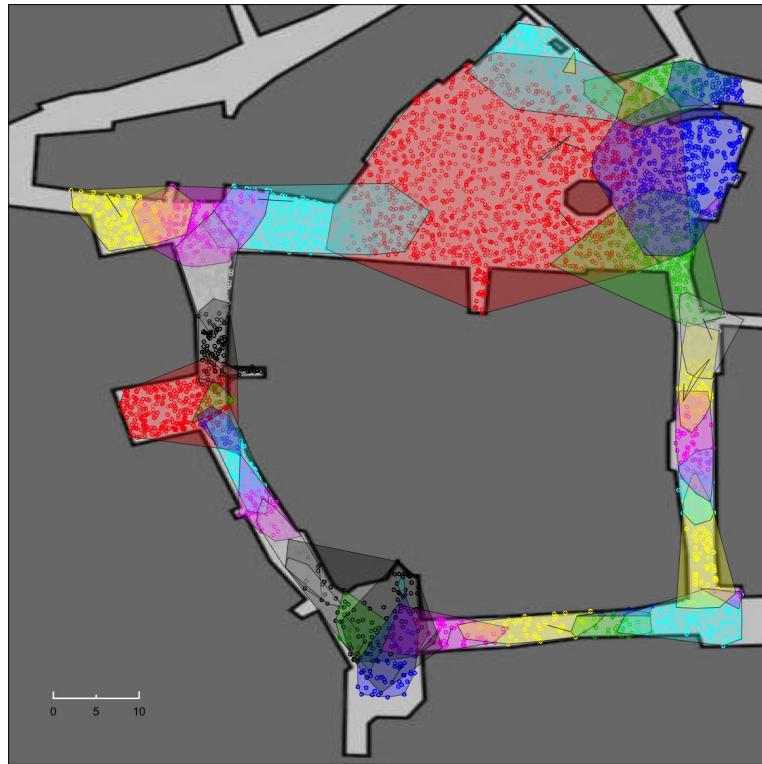
**Level 2**

Figure 11: Spatial division corresponding to: Iterative Graph A - Level 2 - 60 clusters.

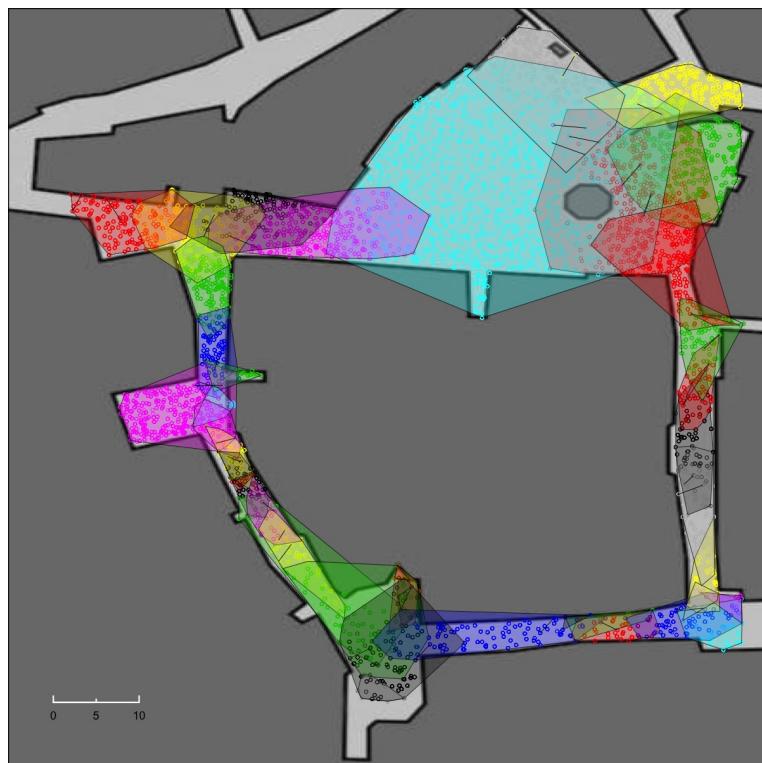


Figure 12: Spatial division corresponding to: Iterative Graph B - Level 2 - 78 clusters.

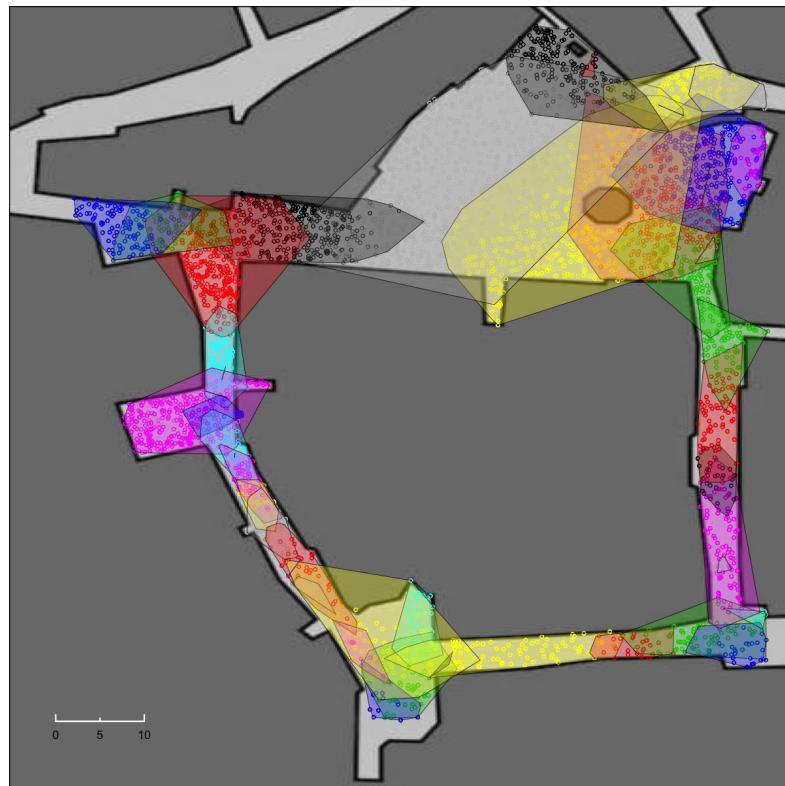


Figure 13: Spatial division corresponding to: Iterative Graph C - Level 2 - 48 clusters.

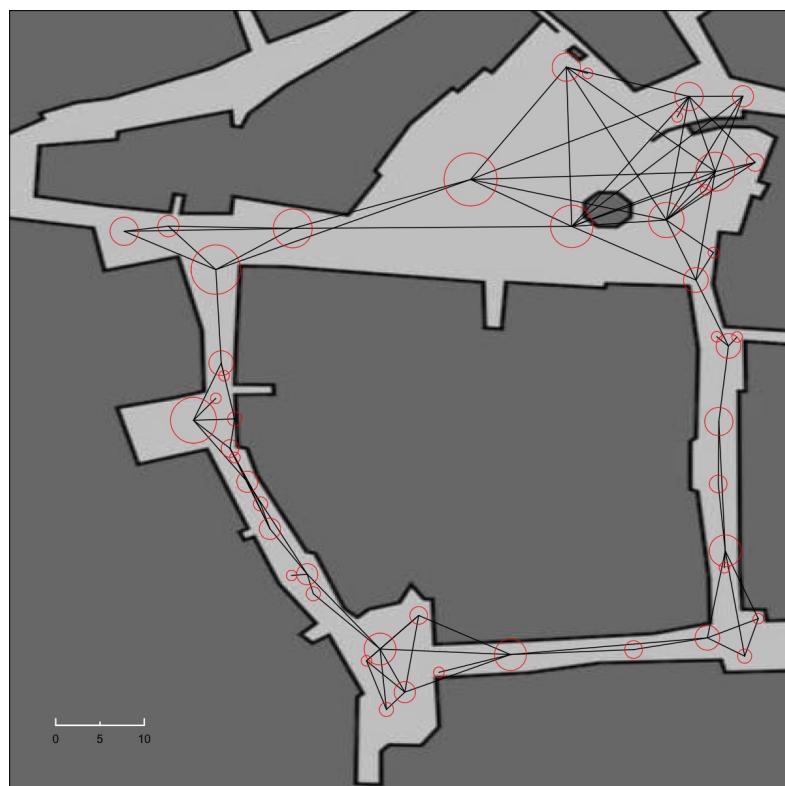


Figure 14: Network representation of: Iterative Graph C - Level 2 - 48 clusters. The number of features in a cluster is mirrored by the diameter of its node.

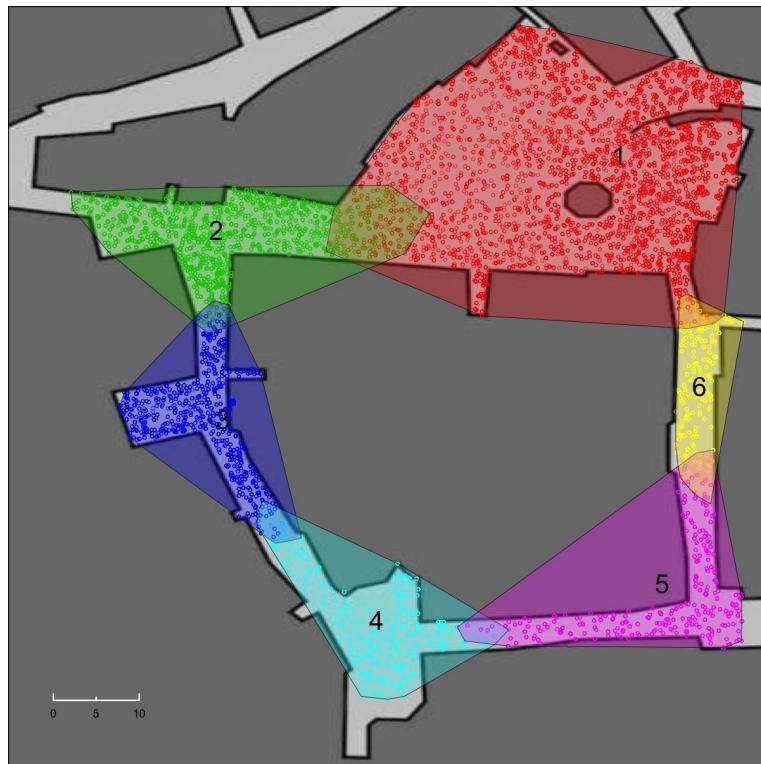
**Level 3**

Figure 15: Spatial division corresponding to: Iterative Graph A - Level 3 - 6 clusters.

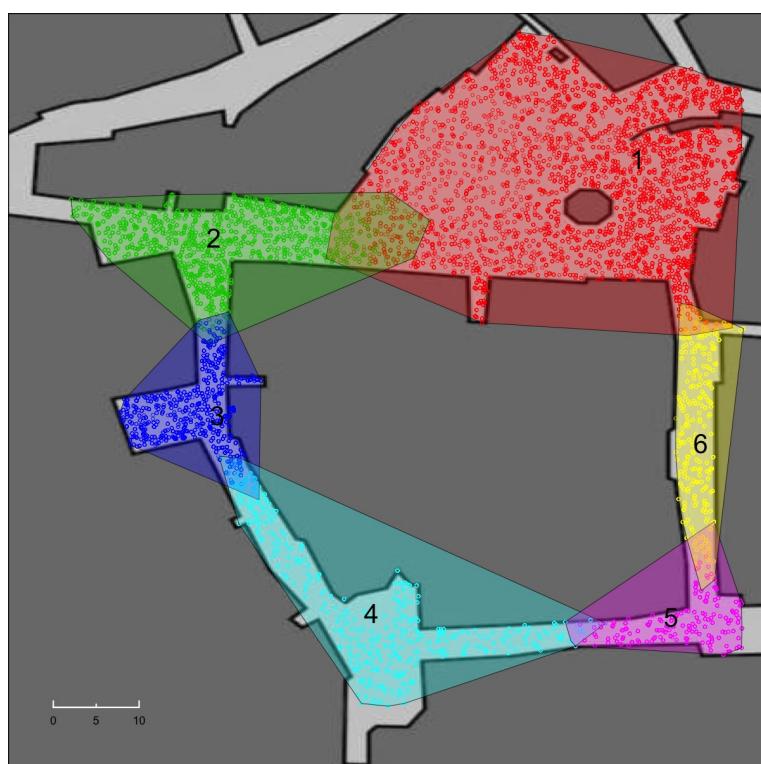


Figure 16: Spatial division corresponding to: Iterative Graph B - Level 3 - 6 clusters.

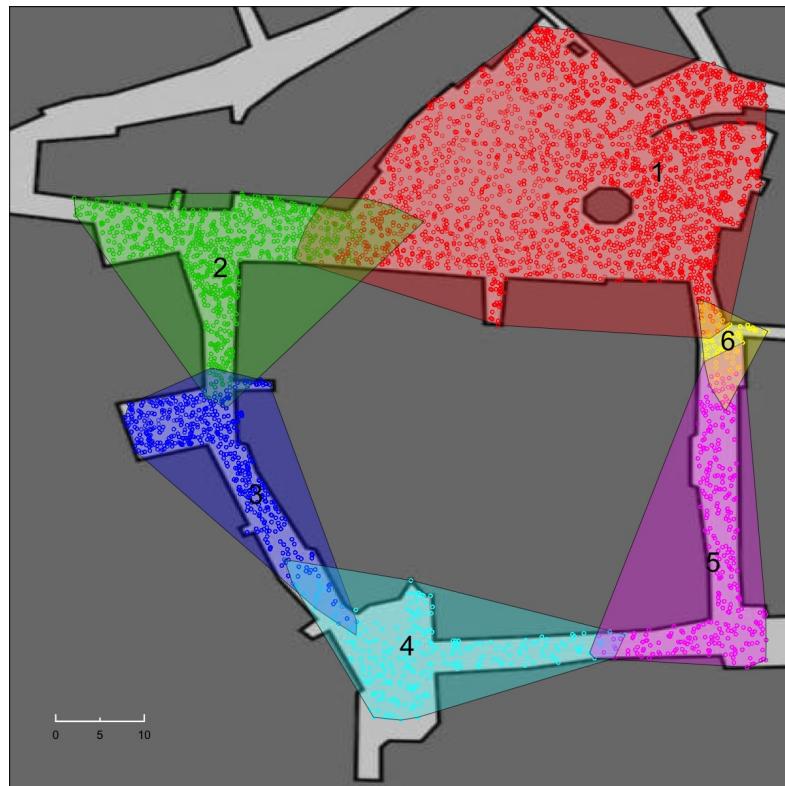


Figure 17: Spatial division corresponding to: Iterative Graph C - Level 3 - 6 clusters.

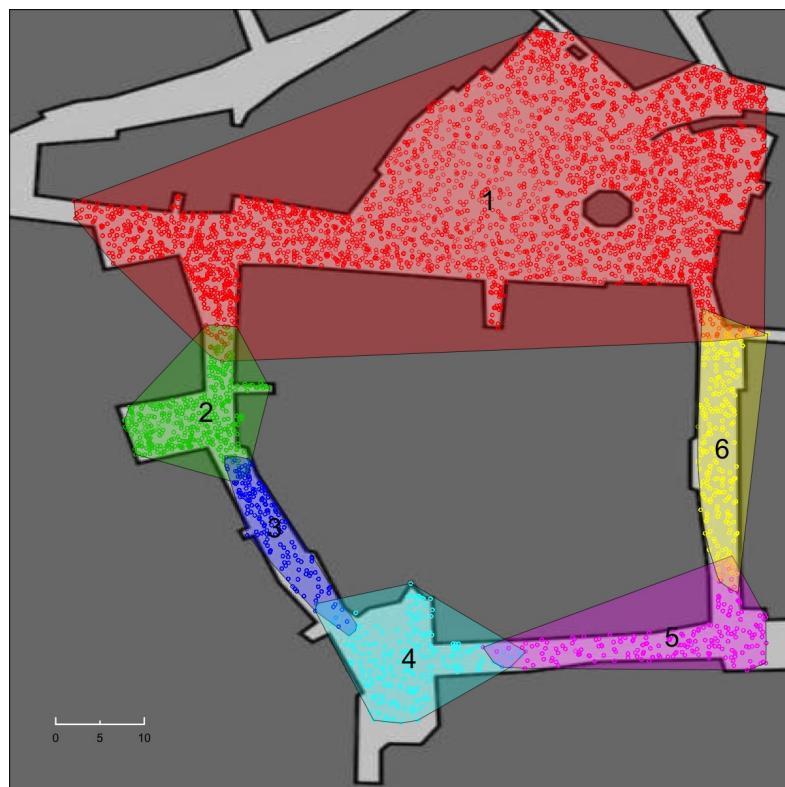


Figure 18: Spatial division corresponding to: Direct Graph - Level 1 - 6 clusters.

### 3.2 Cohesion

Cohesion describes the internal similarity of clusters. It is measured with the average distance of all features to their cluster centers.  $(x_f, y_f)$  denotes the location where a feature was first detected,  $(x_c, y_c)$  denotes the euclidean center of a cluster and function  $d$  computes the euclidean distance between both. Note, lower scores indicate greater cohesion than higher scores. It is given by:

$$\text{Cohesion}(L_k) = \frac{1}{N_0} \sum_{C \in V(L_k)} \sum_{f \in C} d((x_f, y_f), (x_c, y_c)), \quad (16)$$

$$(x_c, y_c) := \left( \frac{1}{|C|} \sum_{f \in C} x_f, \frac{1}{|C|} \sum_{f \in C} y_f \right). \quad (17)$$

Table 2

*The cohesion per level of the four different graphs. In brackets the reference value computed with the respective reference level.*

Graph Level	Direct	Iterative A	Iterative B	Iterative C
3		10.6 (10)	11.0 (10)	11.3 (10)
2		4.9 (2.4)	4.9 (2)	4.8 (2.7)
1	16.6 (10)	4.6 (0.5)	4.4 (0.6)	4.1 (0.7)

### 3.3 Expansion variance

This criterion describes the variance of spatial coverage in the clusters per level. For that, the standard deviation ( $sd$ ) of the clusters maximum expansion is calculated. Lower scores indicate less expansion variance than higher scores. It is given by:

$$\text{ExpansionVariance}(L_k) = sd(\max(\text{Expansion}(C_{k1}), \dots, \text{Expansion}(C_{kN_k}))), \quad (18)$$

$$\max(\text{Expansion}(C)) := \max_{f \in C} d((x_f, y_f), (x_c, y_c)). \quad (19)$$

Table 3

*The expansion variance per level of the four different graphs. In brackets the reference value computed with the respective reference level.*

Graph Level	Direct	Iterative A	Iterative B	Iterative C
3		5.7 (3)	7.5 (3)	9.5 (3)
2		4.8 (1.1)	5.0 (1)	5.8 (1.2)
1	12.5 (3)	2.3 (0.4)	2.6 (0.5)	3.2 (0.6)

### 3.4 Separation

Separation describes the dissimilarity between all clusters of a level. It is measured with the average distance between all clusters centers of a level. Note, higher scores indicate greater separation than lower scores. It is given by:

$$Separation(L_k) = \frac{1}{N_k * (N_k - 1)} \sum_{C_i \in V(L_k)} \sum_{C_j \in V(L_k)} d((x_{c_i}, y_{c_i}), (x_{c_j}, y_{c_j})). \quad (20)$$

Table 4

*The separation per level of the four different graphs. In brackets the reference value computed with the respective reference level.*

Graph Level	Direct	Iterative A	Iterative B	Iterative C
3		52.2 (53.1)	52.7 (53.1)	51.4 (53.1)
2		48.7 (47.5)	47.0 (47.0)	49.2 (47.3)
1	46.7 (53.1)	47.2 (45.9)	45.8 (46.9)	44.2 (46.7)

### 3.5 Overlap

Overlap is measured with the help of the clusters' convex hulls. Features located in areas overlapped by multiple convex hulls are counted and set in relation to the overall number of features. The score depicts the share of features lying in overlapping areas in percent. It is given by:

$$Overlap(L_k) = \frac{100}{N_0} * |\{f \in V(L_0) | (x_f, y_f) \text{ in OverlapArea}\}|. \quad (21)$$

Table 5

*The overlap per level of the four different graphs. The overlap of the reference levels were all zero.*

Graph Level	Direct	Iterative A	Iterative B	Iterative C
3		8.0	9.0	9.5
2		51.2	61.8	63.7
1	5.7	53.4	63.0	71.2

### 3.6 Cluster similarity

To depict the similarity of different clusterings the adjusted rand index (ARI) is calculated (Hubert & Arabie, 1985). This index is a common measure to compare clusterings. It is based on pairwise cluster comparison which is additionally adjusted for random grouping of elements. Thus, the index can range from -1 to 1, where negative values indicate dissimilarity above chance and positive values similarity above chance. The adjusted rand index is given by:

$$ARI(L_a, L_b) = \frac{\sum_{ij} \binom{|C_{ai} \cap C_{bj}|}{2} - t/\binom{N}{2}}{\frac{1}{2}t - t/\binom{N}{2}}, \quad (22)$$

$$t = [\sum_i \binom{|C_{ai}|}{2} \sum_j \binom{|C_{bj}|}{2}]. \quad (23)$$

Table 6

*Pairwise cluster similarity between the four different six-cluster levels.*

Levels	Direct ( $L_1$ )	Iterative A ( $L_3$ )	Iterative B ( $L_3$ )
Iterative A ( $L_3$ )	0.592	-	-
Iterative B ( $L_3$ )	0.6	0.897	-
Iterative C ( $L_3$ )	0.567	0.852	0.825

### 3.7 Border similarity

Further, border similarity of the different clusterings is described on a scale from zero to one. To do so, the overlap of border edges, as defined in Equation 6, between two clusterings is calculated and set in relation to the overall number of border edges:

$$BorderSimilarity(L_a, L_b) = \frac{2 * |Borders(L_a) \cap Borders(L_b)|}{|Borders(L_a)| + |Borders(L_b)|}. \quad (24)$$

Table 7

*Pairwise border similarity between the four different six-cluster levels.*

Levels	Direct ( $L_1$ )	Iterative A ( $L_3$ )	Iterative B ( $L_3$ )
Iterative A ( $L_3$ )	0.125	-	-
Iterative B ( $L_3$ )	0.133	0.658	-
Iterative C ( $L_3$ )	0.39	0.189	0.205

## 4 Discussion

Spatial representations based on features show a high spatial granularity and consist of a great number of elements. Hence, navigation is costly and more abstract concepts of space are missing as well. So far an actual feature graph was analyzed for implicit hierarchical structures which were then related to hierarchical divisions of the environment. Further, metric properties of these divisions were computed to asses their usefulness for path-planning. Multiple divisions were compared to asses the influence of two different clustering schemes and varying abstraction steps.

### 4.1 Metric properties of iterative levels

In spatial divisions corresponding to the first level of the iterative graphs relatively large clusters evolve around the central square whereas in the southern alleys smaller clusters can be found (see Fig. 7 to 9). This leads to a sparse node structure around the central square and a dense node structure in the southern alleys (see Fig. 10). Both the expansion variance as well as the cohesion value is higher in these levels compared to their reference levels (Table 3 and 2). This means that the clusters show higher variation in size and less compactness than the reference clusters. The cluster centers are separated nicely as illustrated in Table 4, although roughly 53 to 71 percent of the features are located in overlapping areas (Table 5). All in all, these levels represent rather suboptimal spatial divisions. With abstraction factors between 50 to 90, these levels are low abstraction levels. However, large clusters around the central square illustrate strong abstractions which do not reduce path-planning costs. Moreover, distorted and stretched clusters with a high degree of overlap exist especially in the southern alleys. These clusters do not divide the environment in distinct regions preventing a plausible abstraction of paths.

In the second level of the iterative graphs former smaller clusters are integrated in larger clusters resulting in more consistent divisions with homogeneous node structures (see Fig. 11 to 14). The expansion variances as well as the cohesion values are higher than the reference values (Table 3 and 2). Interestingly, the difference of cohesion between the experimental and reference levels becomes smaller compared to level one. This means that the clusters still vary in size but gain in compactness. Again the cluster centers are separated nicely (Table 4). The percentage of the features located in overlapping areas decreases slightly compared to level one and varies from 51 to 64 percent (Table 5). Still a clear division of the environment is missing. However, the patch-like clusters are arranged evenly and can be interpreted as a chain of overlapping spatial sections. In such structures distinct regions can be

either described by decisive clusters or by combinations of overlapping clusters. In principle, paths can be abstracted even though not very efficient.

In the third level of the iterative graphs six clusters evolve that clearly correspond to spatial regions (see Fig. 15 to 17). This is mainly illustrated by cohesion values close to the reference value and an overlap of under ten percent (Table 2 and 5). Variance of expansion is higher in these levels compared to the reference level which means that the clusters still vary in size (Table 3). Cluster centers are again nicely separated (Table 4). The clusters divide the environment into clear and distinct regions that only overlap in their transitions. Paths can easily be abstracted in cluster sequences making this level attractive for efficient path-planning.

Especially the first level of the iterative graphs depict more scattered spatial division. This might be related to the procedure used. Spectral clustering is foremost applied to identify  $k$  clusters, such that  $k$  corresponds to the first eigengap which was not the case here. Due to the large number of clusters and the related high-dimensional data points inaccuracies in the clustering process might have accumulated. In contrast, levels two and three of the iterative graphs represent spatial divisions that can be considered metrically useful. These two segmentations are linked in a hierarchical architecture, enabling them to be used by a hierarchical path-planning heuristic. This implies that implicit hierarchical structures of the feature level, which were crystallized based on feature interconnectedness only, also mirror hierarchical spatial divisions that are metrically plausible. Iterative graph A seems to be best suitable for hierarchical path-planning. Its levels consistently show the greatest cohesion (except for level one), the smallest expansion variance and the fewest overlap of clusters ( Table 2 to 5).

## 4.2 Iterative vs. direct clustering

The results show that the iterative scheme identifies other feature clusters than the direct scheme. When comparing the six-cluster divisions of the indirect and direct condition the major difference is in the distribution of the clusters, see Fig. 15 to 18. A major expansion of the red cluster in the direct condition leads to a shift in the arrangement of all clusters. This is also mirrored in the cluster comparison values. The clusters of the iterative graphs match the clusters of the direct condition just moderately (cluster similarities around 0.6). The border similarities vary from 0.13 to 0.39, see Table 7. The indirect condition shows consistently better cohesion, less variance in cluster sizes and a better separation of cluster centers but more overlap (Table 2 to 5). According to these comparisons, the indirect condition yields overall more evenly distributed clusters. Assuming that the direct scheme works

properly and identifies dominant cluster borders, the differences might be related to the similarity functions used. In detail, salient cluster transitions identified in the neighborhood structure of features ( $s_1$ ) might become less salient in the interweaving structure of clusters ( $s_2$ ). On the contrary, the iterative scheme seems to enhance the identification of feature clusters that cover the environment more evenly which is beneficial for path-planning.

### 4.3 Influence of abstraction factors

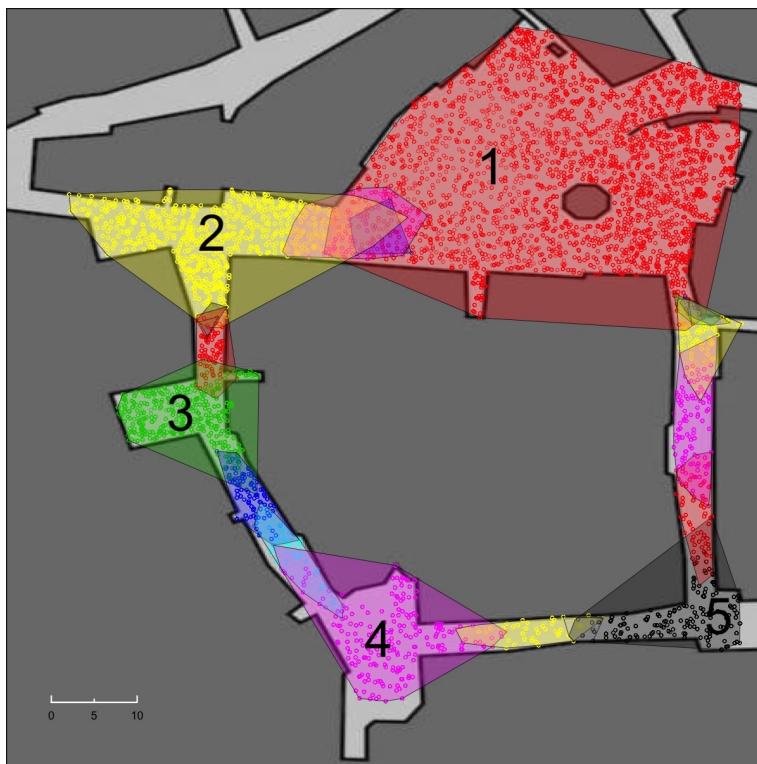
The abstraction factors of the three different iterative graphs were chosen more or less arbitrary. Whether the abstraction factors do have an impact on the resulting hierarchical structures can be evaluated by contrasting the third level of iterative graph A, B and C, see Figure 15 to 17. Level three of graph A and B consist of very similar clusters (cluster similarity of 0.9) that even share a majority of border edges (border similarity of 0.66). The graphical results show that they mainly differ in the transitions of cluster four, five and six. Clusters of graph C are overall similar to the clusters of graph A (cluster similarity of 0.85) and graph B (cluster similarity of 0.83). However, graph C only shares about a fifth of border edges with the clusterings of graph A and B. Again, the graphical results show that the three divisions mainly differ in the transitions of clusters. Additionally, cluster six is very small in the clustering of graph C in contrast to the other two clusterings.

In general, the shape of clusters one and two is very steady in the three clusterings whereas cluster three to six share a common core but vary vastly in their transitions. This can be seen very good in the Figure 19. This plot displays the characteristic clusters of the three clusterings. Characteristic clusters are groups of features that were grouped together independently in all three clusterings. Five major characteristic clusters can be found. Between these, multiple smaller clusters are located which illustrate varying cluster transitions in the three divisions.

Again, varying cluster transitions might be related to similarity function  $s_2$ . Consider the following case: Figure 19 indicates that the southern yellow cluster was grouped once with cluster four and once with cluster five. Since we have a very symmetric constellation the degree of interweaving between the three clusters seems to be indecisive, which leads to different groupings. This observation indicates that interweaving based clustering is partially responsible for the varying cluster transitions.

All in all, different abstraction factors led to a common cluster structure where mainly the cluster borders varied. This finding indicates that the obtained feature clusters do represent dominant superordinate structures within the feature level,

which become less explicit in their transitions. The choice of abstraction factors remains an unsolved problem. If the number of clusters does not suit the network, suboptimal and varying clusters are created as illustrated in the example above. Besides the eigengap heuristic no method provides advise how to choose abstraction factors especially not for intermediate abstraction levels. A clustering algorithm that allows overlapping clusters or an analytical approach such as the identification of characteristic clusters might help here.



*Figure 19:* Characteristic clusters of the six-cluster divisions of iterative graph A,B and C. Characteristic clusters are groups of features that were grouped together independently in all three iterative clusterings.

#### 4.4 Environmental properties of iterative levels

So far, the divisions have been reflected upon their spatial properties independent of the given environment. Obviously a spatial representation should reflect the environment, this also accounts to hierarchical representations. When analyzing the characteristic clusters in Figure 19 one can notice five major clusters. Cluster one covers the area of the central square surprisingly accurate. Clusters two to five cover areas where two alleys intersect. Such an abstraction of the environment is very attractive for path-planning as intersections and squares mark locations where direction decisions are made.

It is assumed that the specific abstraction of spacious areas such as intersections and squares is mainly influenced by the size of the features place-fields. In spacious areas features have larger place-fields due to extensive visibility. Large place-fields tend to overlap with other place-fields, leading to dense feature connectivity. In more narrow alleys the visibility of features is limited and the resulting place-fields are smaller with less overlap. In these areas dense networks consist of fewer features leading to small but numerous clusters. This pattern can be especially seen in levels of low abstraction (see Fig. 7 to 9).

Furthermore, the granularity of a level determines what size of dense networks, called communities, dominate the cluster structures. In levels of low abstraction foremost large communities dominate as seen in level three of the iterative graphs. These networks are mainly centered in the most spacious areas as explained above. With increasing granularity smaller communities dominate as seen in level two of the iterative graphs. These networks are also centered in less spacious areas such as alleys. Figure 19 indicates that alleys are represented by a chain of medium sized communities within the feature graph, whereas more spacious areas are represented by single larger communities in the feature graph. The results indicate that an environment can be best represented by a mix of different granularity levels. That way dominant structures of different size can be crystallized, which represent the environment better when combined. Moreover, superordinate units are able to explicitly represent spatial concepts. For instance, the central square can be reliably found as a superordinate node in the iterative clusterings.

## 5 General discussion

It is important to consider that the localization of place-fields might have influenced the results. A feature’s place-field was embedded into the environment based on the location where the feature was first detected. However, place-fields can extend over large areas and thus can not be easily localized in one position. Assuming that most place-fields are shaped circularly it is unlikely that the localizations denote the center of place-fields but rather its margins. This means that the localization of place-fields was certainly very noisy and with that the results. This might explain distorted and overlapping spatial clusters especially in levels of high granularity. When using place-field centers or even better full place-field outlines, one would expect better and more conclusive results. However, to compute all features place-fields is quite costly and therefore has not been realized in this work.

Interestingly, levels of low abstraction might help the exploration algorithm which created the feature graph to explore the environment more evenly. The density of clusters in low abstraction levels may also indicate the degree of exploration. For instance, the feature graph used in this study seems to be under-explored in southern areas of the environment, where also a high density of clusters can be found in levels of low abstraction. On the one hand, redundant features could be identified in sparse cluster areas and removed from the feature graph to reduce its size without significant information loss. On the other hand, dense cluster areas might indicate under-explored regions. That is why single walks are designed to leave a trace of small communities which are then connected to larger communities by further exploration. Based on this, under-explored areas can be identified, approached by navigation and then further explored by random-walk. This would also mean that the clustering results were partially influenced by the distribution of features. This relation gives indication for further investigations and might help to improve feature based exploration behavior.

Further, the results indicate that the method offers approaches for improvements. First, the similarity function used for higher abstractions ( $s_2$ ) is mainly based on interweaving and can be improved. It seems that the interweaving of clusters do not reliably resemble salient community borders that were detected in the direct scheme, which led to a high variance of cluster transitions. An alternative is to sum the  $s_1$  scores of inter-cluster edges to depict cluster similarity. Second, spatial divisons with plausible metric properties emerged foremost in high abstraction levels. Abstraction factors are recommended such that clusters contain at least 500 features on average. Third, the localization of feature place-fields can be improved either by considering place-field centers or full outlines. Fourth, a procedure that is specified to

identify dominant clusters is desirable. For instance, replace PAM with a data-point clustering algorithm that allows overlapping clusters. The so obtained clusters might differ in dominant structures, whereas their overlap might cover weaker components.

As described in the introduction humans seem to include spatial abstractions in a hierarchical representation of space and use it for efficient path-planning, distance and direction judgments (McNamara, 1986; Stevens & Coupe, 1978; Wiener & Malilot, 2003). A hierarchical feature graph can use its superordinate structures to realize similar tasks. Proximate steps can be planned in levels of high granularity, whereas distant steps can be planned in levels of low granularity. Distance judgments can be approximated by a simple heuristic: Places already grouped together in levels of low abstraction can be considered closer, whereas places that are only grouped together in levels of high abstraction can be considered more distant. Direction judgments can be approximated by superordinate relations of regions, which were so far neglected but may be based on allocentric labels of inter-cluster edges. That way the created hierarchical structures can be used to accomplish spatial tasks.

Obviously, humans are likely to consider various information to create regions. For instance, we might take floor coverings and architectural designs into account. Still, a feature graph might encode environmental properties that might be considered by humans as well. It remains an interesting question whether humans identify similar regions in the environment as spectral clustering has done. It might be the case that the variability of clusters is also mirrored in human regionalizations, i.e. invariable region centers and variable region transitions. Moreover, higher-order levels show a certain similarity to place-graphs although it is very difficult to identify a clear assignment of places in an open environment such as the old town of Tuebingen. Certainly, superordinate units are capable of representing spatial concepts of the environment like the central town square. All in all, the explanatory power of this thesis is limited to a practical assessment of the identified structures. To asses the identified structures on their environmental meaningfulnesses further research is necessary.

## 6 Conclusion

With the help of spectral clustering hierarchical structures of a feature graph can be extracted and integrated in a hierarchical spatial representation. It was argued that these structures reflect hierarchical divisions of the environment which show metric properties suitable for hierarchical path-planning when combined with a high degree of abstraction. The results also indicate deficiencies in the clustering scheme which were discussed along with corresponding improvements. It was shown that the choice of direct and iterative abstraction factors influence the resulting divisions. Clusterings of low granularity, however, share a common composition of clusters which seem to correspond to dominant structures of the feature graph. A dominant superordinate unit represents a dense network within the feature graph and a characteristic structure of the environment. Additionally, superordinate units corresponding to abstract spatial concepts such as squares were observed which can enrich the feature graph with explicit representations of these. Conclusively, the feature graph seems to encode hierarchical structures of the environment in its connectivity.

## 7 References

- Baumann, T. (2019). *Personal consultation*. (Unpublished Master thesis at Cognitive Neuroscience group, Faculty of Science, Eberhard Karls Universität Tuebingen)
- Bay, H., Ess, A., Tuytelaars, T., & Gool, L. V. (2008). Speeded-up robust features (surf). *Computer Vision and Image Understanding*, 110(3), 346 - 359. doi: <https://doi.org/10.1016/j.cviu.2007.09.014>
- Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, 1695. Retrieved from <http://igraph.org>
- Ekstrom, A. D., Kahana, M. J., Caplan, J. B., Fields, T. A., Isham, E. A., Newman, E. L., & Fried, I. (2003). Cellular networks underlying human spatial navigation. *Nature*, 425(6954), 184. doi: <https://doi.org/10.1038/nature01964>
- Franz, G., Mallot, H., & Wiener, J. (2005). Graph-based models of space in architecture and cognitive science-a comparative analysis. In *Proceedings of the 17th international conference on systems research, informatics and cybernetics* (Vol. 3038).
- Franz, M. O., Schölkopf, B., Mallot, H. A., & Bülthoff, H. H. (1998). Learning view graphs for robot navigation. *Autonomous robots*, 5(1), 111–125. doi: <https://doi.org/10.1023/A:1008821210922>
- Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & cognition*, 13(3), 208–217. doi: <https://doi.org/10.3758/BF03197683>
- Hubert, L., & Arabie, P. (1985). Comparing partitions. *Journal of classification*, 2(1), 193–218. doi: <https://doi.org/10.1007/BF01908075>
- Huys, Q. J., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J., ... Roiser, J. P. (2015). Interplay of approximate planning strategies. *Proceedings of the National Academy of Sciences*, 112(10), 3098–3103. doi: <https://doi.org/10.1073/pnas.1414219112>
- Kaufman, L., & Rousseeuw, P. J. (1990a). Clustering large applications (program clara). In *Finding groups in data: an introduction to cluster analysis* (Vol. 344, pp. 126–163). Hoboken, US: John Wiley & Sons.
- Kaufman, L., & Rousseeuw, P. J. (1990b). Partitioning around medoids (program pam). In *Finding groups in data: an introduction to cluster analysis* (Vol. 344, pp. 68–125). Hoboken, US: John Wiley & Sons.
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2017). cluster: Cluster analysis basics and extensions [Computer software manual]. Retrieved from <https://cran.r-project.org/package=cluster> (R package version

- 2.0.6)
- McNamara, T. P. (1986). Mental representations of spatial relations. *Cognitive psychology*, 18(1), 87–121. doi: [https://doi.org/10.1016/0010-0285\(86\)90016-2](https://doi.org/10.1016/0010-0285(86)90016-2)
- Montello, D. R., Goodchild, M. F., Gottsegen, J., & Fohl, P. (2003). Where's downtown?: Behavioral methods for determining referents of vague spatial queries. *Spatial Cognition & Computation*, 3, 185–204. Retrieved from <http://geog.ucsbs.edu/~montello/pubs/downtown.pdf>
- O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat. *Brain research*. doi: [http://dx.doi.org/10.1016/0006-8993\(71\)90358-1](http://dx.doi.org/10.1016/0006-8993(71)90358-1)
- Qiu, Y., & Mei, J. (2018). Rspectra: Solvers for large-scale eigenvalue and svd problems [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=Rspectra> (R package version 0.13-1)
- R Core Team. (2017). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Röhrich, W. G., Hardiess, G., & Mallot, H. A. (2014). View-based organization and interplay of spatial working and long-term memories. *PLoS one*, 9(11), e112793. doi: <https://doi.org/10.1371/journal.pone.0112793>
- Rolls, E. T. (1999). Spatial view cells and the representation of place in the primate hippocampus. *Hippocampus*, 9(4), 467–480. doi: [https://doi.org/10.1002/\(SICI\)1098-1063\(1999\)9:4<467::AID-HIPO13>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1098-1063(1999)9:4<467::AID-HIPO13>3.0.CO;2-F)
- Schick, W., Halfmann, M., Hardiess, G., Hamm, F., & Mallot, H. A. (2019). Language cues in the formation of hierarchical representations of space. *Spatial Cognition & Computation*, 1–30. doi: <https://doi.org/10.1080/13875868.2019.1576692>
- Schölkopf, B., & Mallot, H. A. (1995). View-based cognitive mapping and path planning. *Adaptive Behavior*, 3(3), 311–348. doi: <https://doi.org/10.1177/105971239500300303>
- Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive psychology*, 10(4), 422–437. doi: [https://doi.org/10.1016/0010-0285\(78\)90006-3](https://doi.org/10.1016/0010-0285(78)90006-3)
- Taube, J. S., Muller, R. U., & Ranck, J. B. (1990). Head-direction cells recorded from the postsubiculum in freely moving rats. i. description and quantitative analysis. *Journal of Neuroscience*, 10(2), 420–435. doi: [10.1523/JNEUROSCI.10-02-00420.1990](https://doi.org/10.1523/JNEUROSCI.10-02-00420.1990)

- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological review*, 55(4), 189. doi: <http://dx.doi.org/10.1037/h0061626>
- Von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and computing*, 17(4), 395–416. doi: <https://doi.org/10.1007/s11222-007-9033-z>
- Wiener, J. M., & Mallot, H. A. (2003). 'fine-to-coarse' route planning and navigation in regionalized environments. *Spatial cognition and computation*, 3(4), 331–358. doi: [https://doi.org/10.1207/s15427633scc0304\\_5](https://doi.org/10.1207/s15427633scc0304_5)