University of Zurich[UZH]

Department of Geography

Different levels of visual complexity in geographic visualizations - The effect on visual memory

GEO 620 Master’s Thesis

Dario Oertle
10870327

Supervised by
Dr. Arzu Çöltekin

Faculty representative
Prof. Dr. Sara I. Fabrikant

21. April 2017
Department of Geography, University of Zurich

Abstract

Realistic geographic visualizations such as satellite maps can affect the map readers' cognition and memory depending on what is depicted. This thesis investigates the effect of different levels of visual complexity in realistic geographic visualizations on visual short-term memory and whether spatial ability has an influence on pap memory performance. Previous research describes different ways to measure visual complexity of images and how these measurements can be connected with user studies involving visual search tasks.

In a first step, four existing image algorithms were used to measure visual complexity for four different map types. The map types (*blank*, *street*, *label* and *hybrid*) represent an increased complexity level of satellite maps. Based on these map types, the accuracy, response time and confidence of the participants were measured in four different visual search tasks. The user study took place in a controlled environment in order to minimize external effects and was combined with eye-tracking for further analysis. The results of the study indicate that different levels of visual complexity do not have a significant effect on the participants' accuracy, response time and confidence. Further investigations revealed that mainly the *street* map type affects the results in two tasks. In Task 1, requiring global visual search, it became obvious that the participants performed worse with the *street* map type compared to both with visually less and more complex map types. In Task 3, requiring local visual search, results revealed that participants feel more confident with the *street* map type than the *blank* and *hybrid* map type. Thus, the *street* map type does affect the participants' performance regarding accuracy and confidence in both positive and negative ways. The eye movement data from the user study supports the findings of results for Task 1 and 3 and shows that spatial ability does have an effect on the response time with different map types.

This study aimed to find a connection of measurements of existing image processing algorithms and experiment results with realistic geographic visualizations. The results indicate that there is a trend that visually more complex satellite maps seem to improve the participants' performance in visual search tasks. However, future research might improve the results and thereby come to another interesting conclusion.

Keywords

visual complexity, visual short-term memory, geographic visualization, eye-tracking,
spatial ability, image algorithm

Acknowledgments

I would like to thank Dr. Arzu Çöltekin from the Department of Geographic Information Visualization and Analysis (GIVA) form the Univeristy of Zurich for her constructive advice and support during this thesis.

I also want to thank Dr. Alžběta Brychtová for her help with the computational part of this thesis and feedback in the beginning of my study.

I am very grateful for Giuseppe Marbach and his help in the Linux environment.

Special thanks to all my participants, who took part in my experiment at the University of Zurich.

Last but not least, I would like to thank Paul Oertle, Urs Gruber, Haris Skenderovic and Cecilie Egholm for reading and correcting my thesis.

Table of Contents

List of Figures

List of Tables

List of Equations

1 Introduction

This section gives an overview about the motivation of this thesis and states three research questions which will be answered. At the end of this chapter the structure of this thesis is listed.

1.1 Motivation

Maps are used for different tasks by many people every day and can be designed with various levels of visual complexity. The term visual complexity has been used in many domains ranging from arts to design and science especially since the early 20th century.

Depending on the map design and the map-reader, visual complexity will have an influence on the process of map reading. For example, a situation where one looks on a digital city map on a screen of a smartphone/computer and has troubles finding an object of interest or linking a symbol to its correct physical entity. The question then arises how much it is possible for a human to retain from a previously seen map.

Thus, the motivation of this thesis lies in the study of the relationship between different levels of visual complexity and visuospatial memory with geographic visualizations. In particular, different levels of complexity within realistic visualizations e.g., satellite maps will be studied.

1.2 Research Gap

Despite the strong interest in studying visual complexity through computational methods (Fairbairn 2006; MacEachren 1982; Rosenholtz et al. 2007), there has been only little research done that validates the proposed metrics through user studies. Geographic visualizations have been studied for years, but experiments that relate map complexity to visual memory tasks (e.g., visual search) are rare (e.g., Beck et al. 2010) and therefore need to be further addressed. In this context, there may be an effect of different levels of visual complexity on the map reading performance. Thus it will be interesting to investigate whether computational methods can be applied to geographic visualizations and if the computational measurement can be validated through the results of a user study.

1.3 Research Questions

Based on the motivation and the research gap identified in the previous section, the following research questions (RQ) were formulated. For each RQ, a specific focus is outlined and a hypothesis is given followed by an explanation of the hypothesis with related literature.

<u>RQ 1:</u> *How is visual complexity currently measured, and do existing methods work well to assess the visual complexity of geographic imagery?*

- *Specifically, do (selected) visual complexity algorithms agree on the visual complexity of satellite maps of urban and rural areas and different map types?*

<u>Hypothesis</u>: *Existing image processing algorithms can be applied to geographic visualizations e.g., satellite maps and therefore determine visual complexity as a baseline for user studies.*

A number of studies are dealing with the theme of visual complexity. Bravo & Farid (2008) propose a measure of visual clutter for real images whereas Rosenholtz et al. (2007) studied visual complexity with three different image algorithms on maps and images. Additionally, there are studies that measure visual complexity through compressed file sizes and user rating (e.g., Tuch et al. 2009). Another study by Donderi & McFadden (2005) followed a similar idea, by measuring the ratio between an uncompressed and compressed file size of natural images on the user performance. So, these studies are mainly conducted with real images and thus it would be interesting to conduct a computational analysis with some of the image algorithms on satellite maps. As a result, an initial assessment of visual complexity will be obtained.

<u>RQ 2:</u> *Does the user performance decrease with visuospatial memory tasks as the level of visual complexity increases?*
- Specifically, how will participants perform in visuospatial memory tasks with visually more complex geographic visualizations, such as hybrid maps (cartographically-enhanced satellite maps), in comparison to regular satellite images?

<u>Hypothesis:</u> *Participants will perform worse with increased level of visual complexity, i.e., in this case, with the hybrid maps.*

Users will possibly struggle to remember most details with visually more complex maps due to an overload of information on the map. Each person has a limited storage capacity (Luck & Vogel 1997; Cowan 2001; Alvarez & Cavanagh 2004). Features such as colour and orientation play an important role in visual memory and according to Duncan (1984) and Cowan (2001) only four objects can be remembered. Therefore, more information to be processed can result in lower user performance. With the use of four map types in the user study, every participant will be confronted with maps containing different amount of details (i.e., different levels of visual complexity). The image algorithms by Rosenholtz et al. (2007) constitute the basis of how visually complex a display is. Thus, the user experiment is trying to investigate how visual complexity affects the user performance.

<u>RQ 3:</u> *What influence does the spatial ability have on visuospatial memory tasks with different level of visual complexity?*
- Specifically, do high spatial participants perform differently than low spatial participants with map-based visuospatial memory tasks?

<u>Hypothesis</u>*: Low spatial participants will have more difficulties with the map-based memory tasks in general, and this will be further amplified by more complex maps.*

Participants can be grouped into low and high spatial with results from spatial ability tests such as Mental Rotation Test (MRT) by Vandenberg & Kuse (1978) or a kit of factor-referenced cognitive tests by Ekstrom et al. (1976). According to literature by Hegarty & Waller (2005) spatial ability has an effect on the quality of spatial representations, namely do low spatial participants have more difficulties to maintain a complex image in their mind. A few studies indicate that high spatial participants perform better in a pointing task (Pazzaglia & De Beni 2006) or a map-based route learning tasks (Thoresen et al. 2016), but Richardson et al. (1999) indicated that psychometric measures can have weak predictive power. Therefore, it is interesting to see whether the spatial ability test is predicting that high spatial participants are performing better than low spatial participants.

1.4 Structure

This thesis has the following structure:

- Chapter 2 focuses on background and theories known in the field of visual complexity and the link to visual memory. The image processing algorithms that are used in the computational analysis are also presented.
- Chapter 3 is dedicated to the methods used in order to conduct a reasonable measure of visual complexity and explains the setup of the implemented user study.
- Chapter 4 describes the results of the computational analysis and user study.
- Chapter 5 discusses the results in a critical sense and explains some of the limitations in this thesis.
- Chapter 6 draws a conclusion of the thesis and comments on future research.

2 Background

This thesis is based on the theoretical framework of visual complexity in geographic visualizations. The term complexity is defined and grouped into visual and intellectual complexity followed by a detailed analysis into four different image processing algorithms. Additionally, three well-known image compression methods are presented in relation to visual complexity. To sum up the chapter, the term complexity is linked to the process of visuospatial memory.

2.1 Complexity

As the term of complexity is generally defined as the state of being complex, it is the term complex that has to be considered (MacEachren 1982). Definitions for the term *complex* are given, according to Banhart (1967): *"1. the mixture of interconnected parts; compound of multiple parts 2. characterized by a very complicated or involved combination of parts, units etc. 3. so complicated as to be difficult to understand."*

Complexity is defined as "the study of the behaviour of macroscopic collections of units that are endowed with the potential to evolve in time" (Highfield & Coveney 1995). The term complexity is wide and can therefore be connected to various research fields but the focus of this thesis lies on the complexity of geographic visualizations such as geographic maps. In the 1970s and 1980s a lot of research was done on cartographic complexity (Fairbairn 2006), concentrating on the visual and structural complexity of an image map (Montello 2002). Tufte (1989) indicated that there is a "level of complexity" best suitable for every human being. At this preferable point, the processing ability reaches a maximum level.

An important distinction suggested by MacEachren (1982) is that there are two types of map complexity: visual and intellectual. The next two subchapters describe both visual and intellectual complexity in more detail.

2.1.1 Visual Complexity

The visual or graphical complexity refers to the visual characteristics of a map e.g., feature density, number of colours or objects types. In other words, Brophy (1980) stated that visual map complexity is "a direct consequence of the spatial differentiation of the graphic content of the map." The quantity and range of objects as well as the variety of materials and styles do play a role for the perceived visual complexity. Discussed by Heylighen (1999), one could therefore argue that the more variety in a map or visualization exists, the higher the visual complexity seems to be.

Visual complexity entails the visual impact and the perceptual processes of viewing a map, which are often very different (Castner 1990). According to MacEachren (1982) visual complexity plays an important role for the process of map reading. In theory, every cartographer that produces a map and its design does have the greatest control over the visual complexity. Thus, it is of great importance for every cartographer to be aware of his or her influence on map-readers. In the studies by Oliva et al. (2004) and Heaps & Handel (1999) the term complexity is dependent on how difficult it is to provide a verbal description of an image. Patterns might occur in different areas on a map and therefore make it more or less complex to read. According to Oliva et al. (2004) visual patterns are complex whenever their parts such as features and symbols are difficult to detect.

For topographic maps the process of map generalization controls and reduces map complexity as for satellite maps map complexity cannot be controlled that well (e.g., Brassel & Weibel 1988; Buttenfield & McMaster 1991). The question of how to measure and quantify visual complexity in realistic geographic visualizations (i.e., satellite maps) will be described later in this chapter.

2.1.2 Intellectual Complexity

The counterpart of visual complexity is the intellectual complexity. This term can be described as the cognitive process that occurs as human beings work with maps. Intellectual complexity strongly influences the interpretation and analysis process of the map reader (MacEachren 1982). Brophy (1980) stated that intellectual complexity is radically different for each and every user and thus it will be difficult to measure it. Therefore, a lot of the studies from the past focused on the concept of visual complexity (Fairbairn 2006; Rosenholtz et al. 2007; Touya et al. 2015). The terms visual and intellectual complexity cannot be separated independently. Robinson (1952) claimed that all shapes within a map do have both visual and intellectual relationships to each other and whenever a map-reader observes a specific object within a map he or she tries to understand the symbol.

2.2 Measuring Complexity

Methods for measuring visual complexity have been proposed in recent decades. For example, one study attempted to identify possible relations between number of faces, edges, and vertices and the visual complexity of a map (MacEachren 1982). Fairbairn (2006) measured map complexity through data compression, whereas Harrie & Stigmar (2010) looked at the relation of number of objects, number of points and object line length to human judgement of complexity. There have been two relatively well-known measures suggested by Rosenholtz et al. (2007); namely feature congestion (FC) and subband entropy (SE). The former is based on the idea that it would be more difficult to place an object, which will catch the attention, on an already cluttered or disorganized work desk, whereas the latter is looking at the relation of clutter to the efficiency in which a display or visualization can be encoded (He et al. 1997; Reitsma 2001; Rosenholtz et al. 2007). In recent decades, Bjørke (e.g., 1996; 1997) has been one of the most active research-ers, who used entropy measures in cartographic displays. For example, Bjørke (1996) demonstrated how different entropy measures can be used as control variables for optimizing cartographic maps.

Additionally to FC and SE complexity measures, Mack & Oliva (2004) proposed a method that measures image complexity with edge density. The ratio of the number of edge pixels to the total number of pixels results in a measurement for visual complexity. While all these methods are designed to provide a 'holistic' measure of complexity, there are also methods (e.g., quad tree clutter) that are used to identify and measure locally more complex regions on a display (Jégou & Deblonde 2012). Jégou & Deblonde (2012) converted colour maps into greyscale images and then cut every cell into four smaller cells whenever the cell is heterogeneous. One major problem with this method is that it is limited to greyscale rates but it is interesting that this method is both global and local. Therefore, visual complexity could be measured for a whole image but also a specific area within an image. The method of image segmentation by Felzenszwalb & Huttenlocher (2004) is also considering a measure of visual clutter by counting the number of regions within an image. Bravo & Farid (2008) successfully used this algorithm for over 100 images in a visual search experiment. A few years ago, Ciolkosz-Styk & Styk (2013) did research in image processing methods, which could be suitable for dynamic maps, aerial images or 3D models. All those concepts illustrate how the visual complexity of an image or a generic geographic map can be measured.

Due to the large variability of algorithms measuring visual complexity, only the following algorithms will be used in this thesis: *FC, SE, edge density* and *image segmentation.* Those algorithms will now be studied in more detail in order to understand how results are being processed.

2.2.1 Feature Congestion

As mentioned in the previous section, Rosenholtz et al. (2007) proposed a few algorithms to conduct a measurable analysis of visual clutter. Two years before, Rosenholtz et al. (2005) already described clutter simply as the situation when something contains too many objects. To take a table as an example: there may be a lot of objects on the table and therefore, the owner cannot find specific objects anymore and somehow gets a notion of disorder. This chaos on the table wherein the object of interest is lost represents a clutter. Rosenholtz et al. (2005) gave a scientific definition for the term clutter as following: "*Clutter is the state in which excess items, or their representation or organization, lead to a degradation of performance at some task".*

Keeping the definition of clutter in mind, the model of statistical saliency plays an important role in order to understand the meaning of FC. According to Rosenholtz et al. (2005) the visual system of every human seeks to detect "unusual items". An item can be described as "unusual" whenever its features are outliers to the local distribution of the features in the actual display. The following equation [Equation 1] illustrates the model of statistical saliency; where T is the degree to which a vector is an outlier from the local distribution of feature vectors. The parameter μ_D is the mean of feature vectors, Σ_D the covariance, and where ()' indicates a vector transpose

$$\Delta = (T - \mu_D)' \, \Sigma_D^{-1} \, (T - \mu_D)$$

Equation 1. Statistical saliency, (Rosenholtz et al. 2005, p.764)

In Figure 1 the statistical saliency model is graphically shown. Two targets e.g., a difficult and an easy target marked as black dots can be seen. Generally, the further out a target is from the distractor distribution, the easier the visual search becomes. A study by Duncan & Humphreys (1989) did not only look at the distance of the target and the distractor distribution but also at the

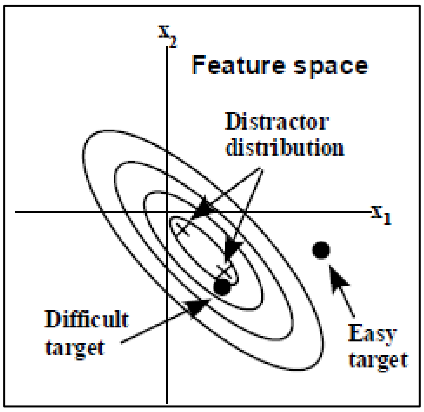


Figure 1. Statistical saliency model, (Rosenholtz et al. 2005, p.765)

role of target-distractor in comparison to distractor-distractor similarity in visual search. Increased similarity of targets to non-targets results in an increased level of difficulty as well. The more items exist in a display or specific area, the less room remains for new objects to be added. In congested areas it will be more difficult to locate a target object due to the unorganized and cluttered situation. In such situations, where too many colours, sizes and shapes of features are clamouring for attention FC measurement becomes relevant as suggested by Rosenholtz et al. (2005).
According to Rosenholtz et al. (2005), the implementation of FC includes the following four stages: a) computation of local feature (co)variance at multiple scales, b) combination across scale, c) combination of clutter across feature types, and d) consolidation of all measures in space to get a single measure of clutter for each input image. In the following part, a brief description explains what is done in every stage of the FC implementation.

a) Feature covariance

- Conversion of image into perceptually based CIE Lab (L*a*b) colour space
    - *CIE*[1] (Commission Internationale de l'Eclairage), *"L"* stands for lightness, "*a"* and "*b"* are chrominance indices [2]
- Procession of image at multiple scales and finding features at each scale
- Computation of local (co) variance for each feature
    - Linear filtering (to average over a local area)
    - Pointwise nonlinear operations (to compute the variance)
- Computation of the volume of the covariance ellipsoid

b) Combination across scales

- For each feature: combination of feature congestion across scale: maximum value at each pixel.

c) Combine across features

- By now: three clutter maps: colour congestion, orientation congestion and texture congestion
- Combination of colour clutter and contrast clutter at each point
- Measurement of how congested a feature space is, depending on how much of feature space is taken up by the covariance ellipsoid relative to how much feature space is totally available
- Scale the clutter value by the range of possible clutter values for that feature

d) Combine across space

- Take average clutter value of the entire image

[1] http://www.pcmag.com/encyclopedia/term/59755/cie-lab (accessed: 10.04.2017)
[2] https://www.konicaminolta.com/instruments/knowledge/color/part5/01.html (accessed: 20.2.2017)

2.2.2 Subband Entropy

The second measure applied by Rosenholtz et al. (2007)'s work is the SE clutter measure. It deals with the idea that clutter does have a relation to the number of bits that are required to encode an image while perceiving the same image quality. The image is first divided into several subbands with different orientations and spatial frequencies (Rosenholtz et al. 2007). So, the idea is to capture parts in an image that are redundant and therefore easier to encode. As an example, a space with similar objects grouped together appears as more organized and is therefore easier to encode. Additionally, other characteristics such as luminance, hue or size are having an effect on the SE clutter measure (Rosenholtz et al. 2007). Effective image codes such as JPEG or JPEG 2000 follow the concept of entropy encoding.

In order to give a proper view about how the implementation looks like, the following stages will give an overview of what is happening. First of all, the so-called Shannon entropy is computed within each subband. The equation is defined as illustrated in Equation 2.

The parameter p is the probability distribution of coefficients in each subband. The Shannon entropy includes the bits required to encode a subband and is dependent on the level of fidelity.

$$\sum_{i} -pi \log(pi)$$

Equation 2. Shannon entropy, (Rosenholtz et al. 2007, p.7)

The higher the fidelity, the more bits are needed to encode a specific image (Rosenholtz et al. 2007). In easier words the Shannon entropy equation estimates the minimum number of bits that are used to encode a line of symbols based on how often the symbols appear (Kaplan 2002).

After computing the Shannon entropy within each subband, all subband entropies are added and differentiated into luminance and chrominance, similarly to the FC measurement (CIE Lab) (Rosenholtz et al. 2007). The fact that image encoding leads to data compression and fewer bits than the original file used to have, three different compression techniques will be presented in section 2.3.

2.2.3 Edge Density

This method was first proposed by Mack & Oliva (2004) and then used again by Rosenholtz et al. (2007) as a measure of visual clutter. The method is basically focusing at the number of objects in an image and then calculating both the density of edges and a correlation of clutter with high-frequency content. Touya et al. (2015) do compare different image-based methods for quantifying visual clutter in images and do use the edge density method successfully used in the paper by Rosenholtz et al. (2007). The edge density is basically the relation of edge pixels within an image compared to the total pixel size of that same image. Matlab's Canny edge detector[3] was applied to measure the actual edge density of every image. With the function "edge" in Matlab a sensitivity threshold can be defined. Edges, which are not stronger than the given threshold, will be ignored by the function. So, a lower threshold results in a higher edge density measurement and respectively does a higher threshold detect stronger edges. A low and a high threshold, as Rosenholtz et al. (2007) applied in their project, was chosen for the computational part.

2.2.4 Image Segmentation

The segmentation algorithm is basically dividing an image into several segments and then counting the number of objects that are identified by the segmentation (Touya et al. 2015). The algorithm is very efficient, so that it was possible to test multiple large images (Bravo & Farid 2008). The original source code was developed by Felzenszwalb & Huttenlocher (2004) and then used by Bravo & Farid (2008) to measure visual clutter for real images, who give a brief and simplified explanation of the algorithm as following: *"The algorithm works by representing an image as an undirected graph with each vertex in the graph representing a pixel in the image. Neighboring vertices are connected by an edge, and the weight on the edge is proportional to the difference between the corresponding pixels. Segmenting the image into regions involves cutting edges in the graph to produce subcomponents, which are disjoint sets of interconnected vertices. The algorithm decides which edges to cut by comparing the minimum weight connecting two subcomponents with the maximum weight within the subcomponents".* The described algorithm was used for a search experiment with general images showing a bag filled with multiple objects. Nevertheless, it is the goal to adapt the algorithm and use it for geographic visualizations with different levels of visual complexity such as enhanced satellite maps.

2.2.5 Quad Tree Clutter

This method has first been proposed by Jégou & Deblonde (2012), who displayed visual clutter on geographic maps. Touya et al. (2015) worked on the effect of map generalization on visual clutter with Quad Tree Clutter as one of four methods. This algorithm generally computes a quad tree of a greyscale image. If the pixels inside a cell are heterogeneous the whole cell is cut in four. Each pixel greyscale value is compared to its neighbouring values trying to detect any heterogeneous areas. The maximum difference of neighbouring values is kept. If this maximum is higher than a manually set threshold, the area is considered as heterogeneous and cut into four. This algorithm is both working on measuring global and local clutter and therefore a very interesting method to use. In the beginning, it was meant to include this measurement in the computational analysis of the geographic visualizations but as the output from the code by Jégou & Deblonde

[3] https://ch.mathworks.com/help/images/ref/edge.html#buo5g4f-1 (accessed: 06.10.2016)

(2012) only generates a visual output, the method was dropped. Nevertheless, the algorithm was presented in this subsection as it might be interesting to work more on the output of the algorithm in order to get a useful numeric result and then compare it to other image processing algorithms.

2.2.6 Overview of Algorithms

In order to give a proper overview about the recently presented algorithms, Table 1 indicates what each algorithm actually measures. The Quad Tree Clutter algorithm is not listed above because it was not applied in the user study.

Table 1. Overview of algorithms and measures

Name of algorithm	***Function (what is measured)***
Feature congestion	colour, orientation, luminance
Subband entropy	colour and orientation (encoding)
Edge density	edge pixels
Image segmentation	number (n) of regions (with same characteristics)

2.3 Image compression

To measure the visual complexity of images multiple algorithms such as FC, SE, edge density, image segmentation and quad tree clutter were presented in the previous chapter. The measurement of SE is dealing with the concept of image encoding and therefore refers pretty well to image compression that can be used to quantify complexity of geographic visualizations, too. That is why the three compression algorithms: LZW, ZIP and JPEG will now be highlighted in the following subsection. All geographic visualizations where mainly designed and altered with Adobe Photoshop CS6 and Illustrator CS6. The fact that LZW, ZIP and JPEG are standardized algorithms in these Adobe tools and nowadays well known compression types turned the balance to use them for this analysis.

2.3.1 Lempel-Ziv-Welch

The Lempel-Ziv-Welch (LZW)[4] algorithm is a universal lossless data compression algorithm. It is named after the scientists Abraham Lempel, Jakob Ziv and Terry Welch and belongs to the group of dictionary lookup-based algorithms (Knieser et al. 2003). Dictionary based algorithms do have the characteristics that they are looking for sequences in the data that occur more than once. The dictionary stores all these sequences and then references are made to data that occurred several times in the compressed file.

[4] https://www.prepressure.com/library/compression-algorithm/lzw (accessed: 06.10.2016)

The LZW algorithm includes the following advantages and disadvantages. The algorithm is well suited for files that contain lots of repetitive data, so text or monochrome images are good examples to use the LZW with. Another advantage is that the algorithm is pretty fast but it is already quite an old technique and therefore current computer systems can use more efficient algorithms. Additionally, the popularity of LZW decreased due to the fact that users have to pay for using the algorithms. Whenever files, which do not contain any repetitive data, are compressed the file size can even grow bigger than it used to be.

2.3.2 ZIP

Compression method 8 (Deflate)[5] is the method used by default by most ZIP compatible tools like for example Photoshop. This is one possible compression method being used with ZIP compression but there are many other compression methods available as well. In a ZIP file every original file will be compressed and saved sequentially. That makes it easy to freely choose a single file within a file folder[6]. Both ZIP and LZW are lossless compression techniques that detect patterns and replace them with a single character. Therefore, the advantages and disadvantages of the LZW can be applied to the ZIP as well. The ZIP File format is nowadays very often used.[7].

2.3.3 Joint Photographic Experts Group

Joint Photographic Experts Group (JPEG)[8] is a very well-known compression algorithm. The compression is lossless so details or information will not be lost in comparison to the original file whenever you decompress a JPEG compressed file. For the use of JPEG compression, the ratio of the compression can be chosen. By lowering the compression ratio detail or information lost can be minimized and vice versa. This compression algorithm is going through following phases, explained in a more simplified way:

The JPEG compression is based on the different perception in colour and brightness. There is no analysis of RGB or CMYK values but instead, a luminance/chrominance based colour space called YCbCr is used. Luminance information is stored in Y and Cb/Cr contains data about colour. The image is cut up in separate blocks of 8 x 8 pixels and the compression algorithm is then calculated for each separate block. So, when too much compression is applied, those blocks may become visible.

In a next step, every 8 x 8 block consisting of 64 pixels is transformed into 8 x 8 Discrete Cosine Transformation (DCT) coefficients and no compression is happening.

---

[5]https://pkware.cachefly.net/webdocs/casestudies/APPNOTE.TXT (accessed: 06.10.2016)

[6] http://www.binaryessence.de/dct/de000219.htm (accessed: 06.10.2016)

[7] http://www.investintech.com/resources/articles/whatcompression/ (accessed: 06.10.2016)

[8] https://www.prepressure.com/library/compression-algorithm/jpeg (accessed: 07.10.2016)

After this step, the actual compression takes place by taking all values from the newly generated DCT coefficient block and dividing it by its corresponding constants in the 8 x 8 quantization table.

The advantage of using JPEG algorithm is the big compression ratio that can be achieved. It is possible to reduce a file to a fifth of its original file size. Nevertheless, the JPEG algorithm should not be used for images that have had a mask or shadow effect or if they do contain 256 or less colours. Images with sharp changes in tone should be avoided by the JPEG algorithm. At the moment, there are two known JPEG compression algorithms: JPEG and JPEG 2000. The JPEG 2000 algorithm adds three major things to the previous JPEG specifications. First, a wavelet-based algorithm that allows improved image quality at a very high compression ratio. Second, the algorithm is supposed to be 20% more efficient than the previous one. Third, there is an option of a lossless compression mode.

2.4 Visual Memory

There are various definitions of what makes memory a visual memory (VM). Luck & Hollingworth (2008) tried to give a proper definition without being too broad or narrow. They state that "visual memory encompasses memory representations that maintain information about the perceptual properties of viewed stimuli" (Luck & Hollingworth 2008). Nevertheless, it does not matter in which format that information is encoded. The information can either be encoded from low-level or high-level visual representations (Luck & Hollingworth 2008).

After looking at an image or a map, some features will remain in the memory and a person will be able to describe the content to a certain degree. The memory of a person is remarkably good even when loads of pictures are shown to a person in an experiment (Standing et al. 1970). It is known that people are focusing on objects rather than features, such as colour or orientation (Wolfe 1998). Nevertheless, it is interesting to focus on the effect of colour and orientation in more detail because the user experiment of this thesis contained colour, orientation and location effects.

The concept of visual memory can be divided into: visual sensory memory, visual short-term memory (VSTM) and long-term memory (LTM) or visual long-term memory (VLTM). Whenever researches are describing visual short-term memory the term visual working memory is often used to refer to the same system (Luck & Hollingworth 2008).

The focus in this thesis will remain with the concept of VSTM, as the user experiment is dealing with the process of object and colour recognition of visualizations for a very short time interval. Nevertheless, mainly detailed visualizations are retained over the short and long term (Luck & Hollingworth 2008).

2.4.1 Visual Short Term Memory

The following key properties define the system of VSTM. First, VSTM is created rapidly, only at rates of 20-50ms per item (Vogel et al. 2006, p.1449). Second, the storage capacity for VSTM is limited to a number of objects (Cowan 2001; Alvarez & Cavanagh 2004). There will be more discussion about storage capacity later on in this subchapter. Last but not least, only a limited amount of information will be stored for VSTM, which means that changes in an object or image might not be detected (Simons & Rensink 2005).

After having an idea of what VSTM means, the question arises of how to measure it. Scientists have used change detection or visual search tasks to study the effect of VSTM. Rensink (2002) highlights different types of tasks that try to detect a change. As an example, an array of objects with specific colours is shown for a given time. Then, participants will see a test array. Now they have to think whether they have detected a change or not in the test array. They can answer either by "yes/no" or choose from a fixed set of alternatives (Rensink 2002). Figure 2 illustrates two examples for a visual search and a detection change task by Eng et al. (2005). In the visual search

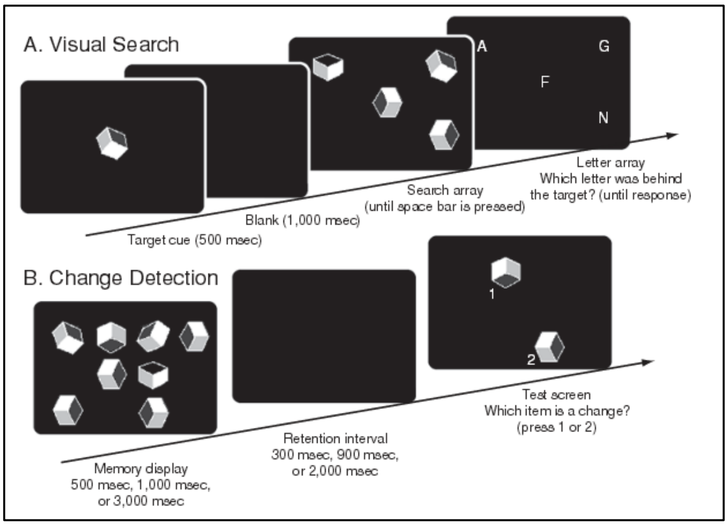


Figure 2. Visual search and change detection tasks, (Eng et al. 2005, p.1129)

task (A) participants have to study a target object (e.g., a cube) for a short time. After a blank space was shown for 1'000 milliseconds, a search array with four possible cubes appears. There, the participant has to find the cube matching with the target cue and press the space bar for continuing to the next page. On the last screen the location of the previous cubes is replaced by letters and the participant has to identify the letter behind the target cube (correct answer: letter F). In the task change detection (B) in Figure 2, multiple cubes are shown for either 500, 1'000 or 3'000 milliseconds. After the memory phase, a blank screen is shown for a given time. Then, the participant has to state which of the two cubes has changed (correct answer: cube number 2).

The presented task types in Figure 2 show how to possibly measure VSTM but by what gets the VSTM affected. As an example, does a blueish cube or a triangle affect the outcome of the study? At least how many objects can be memorized? In the following subsection, storage capacity and object features, such as colour, orientation or size are highlighted.
About 20 years ago Luck & Vogel (1997) conducted an experiment to analyse the storage capacity for simple features. Their findings indicate that participants were able to retain four colours or orientations in VSTM. Interestingly, participants were able to hold both colour and orientation of four objects. So, colour and orientation are integrated and stored in one object. Already Duncan (1984) came to the conclusion that retaining two features (e.g., colour and orientation) of an object is as easy as attending to a single feature. Cowan (2001) also reviewed storage capacity in VSTM and comes to the conclusion that a maximum of four objects can be retained. According to Wolfe & Horowitz (2004) the following attributes are undoubtedly attracting attention in images and therefore important for visual memory: Colour, motion, orientation and size.

2.4.2 Visual Complexity and Visual Memory

The goal of this section is to link the influence of visual complexity on recognition memory. The question arises whether different levels of visual complexity do have an influence on the amount of detail or information a person can memorize. As this thesis tries to find a connection between the level of complexity and memory, it is essential to elucidate research projects that link complexity and visual recognition or memory.
An article by Tuch et al. (2009) presented a research project that tries to find an effect of visual complexity of websites to the users' experience, performance and memory. In that user study participants were asked to look several websites with different levels of complexity. Different levels of complexity were assessed by compression rates of JPEG files and user rating. The experiment contained a passive viewing task (PVT), a visual search task (VST) and the actual memory test a week after the first two tasks. The memory assessment consisted of a test (seen/not seen) for each website. Tuch et al. (2009) found out that participants performed better on search and recognition tasks with websites of low visual complexity.
Another experiment by Donderi & McFadden (2005) analysed the effect of increasing JPEG file size on accuracy and search duration. Participants of the study had to solve a VST on colour photos of the earth and marine electronic displays. The goal was to find a specific target object as fast as possible. The results indicate that more complex displays (with increasing JPEG file size) result in lower accuracy and longer search time for the target object. Alvarez & Cavanagh (2004) analysed the effect of stimuli complexity on the storage capacity. Their results indicate that storage capacity for more complex stimuli (e.g., 3-D cubes) is lower than for simple objects (e.g., coloured square). There were not any geographic visualizations used in their study, but the results indicate that the more complex an object is, the less can be retained.

2.4.3 Eye Movement and Visual Memory

With the possibility to see something with your own eyes, information is being collected in our everyday life. Whenever you are walking into an area where you have never been before, you start screening it and your brain starts processing the information, which is being collected (Irwin & Zelinsky 2002).

Eye movements include two temporal phases: *fixations* and *saccades* (Henderson 2008). Fixation is the phase, in which the gaze positon is relatively still, whereas saccades are very fast eye movements (700-900 deg/sec) (Carpenter 1988). Henderson (2008) differentiates between three memory related eye movement types as following: *Transsaccadic memory, active online scene memory and long-term scene memory.*

Transsaccadic memory is within a time scale of tens of milliseconds and is determined over one saccade whereas *Active online scene memory* is happening in a time frame of seconds. According to Irwin & Andrews (1996) *both transsaccadic- and active online scene memory* can both be forms of short-term memory. *Long-term scene memory* can last minutes, hours and even years and an event goes beyond current perceptual episodes.

There are a few studies that linked eye movements to VSTM or VLTM (Nelson & Loftus 1980; Hollingworth & Henderson 2002; Irwin & Zelinsky 2002; Zelinsky & Loschky 2005). In the experiment of Irwin & Zelinsky (2002) participants were asked to view seven greyscale toys in a baby's crib. After a number of fixations (number of fixed objects), the display disappeared and participants were shown a specific location of the baby's crib without the toys being displayed. They were asked which object was placed at that location. The results indicate that participants did better in memorizing recently fixed objects.

3 years later, Zelinsky & Loschky (2005) investigated more in the recency effect related to eye fixations. They created a display with nine greyscale toys distributed in random order and asked participants to look at them. Compared to Irwin & Zelinsky (2002), in the study by Zelinsky & Loschky (2005) toys were shown only after a predetermined target and number of objects had been fixated. In Figure 3, the baby bottle represents the predetermined target object (E). In this example three fixations after the target object were set as a parameter. Consistent with the results from Irwin & Zelinsky (2002), an increasing number of fixating objects resulted in a poor memory accuracy. In a similar direction goes the study by Loftus et al. (1978), by indicating that participants may overwrite their old visual memory with present perceptual representations and therefore struggle to retain important memory.

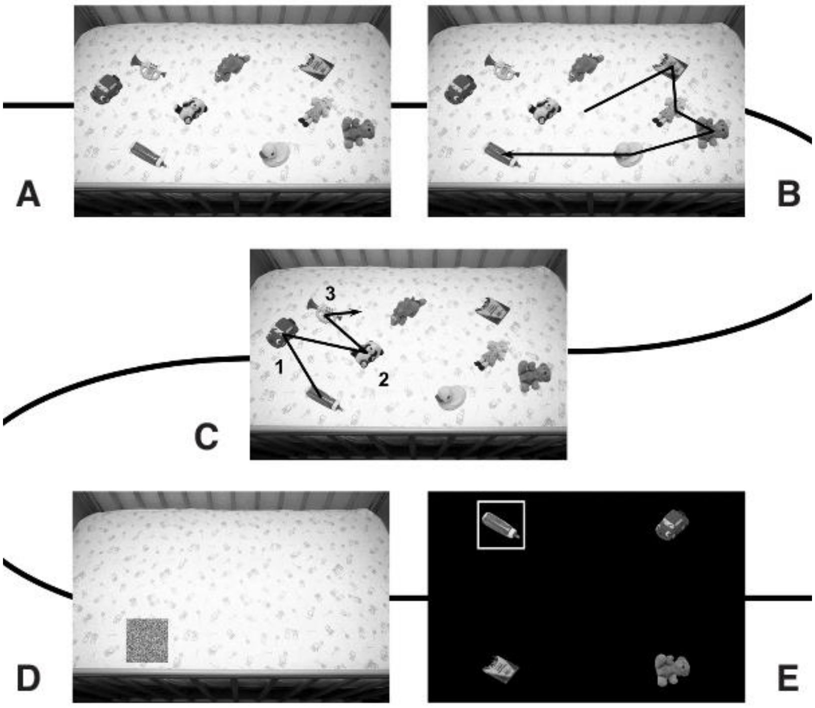


Figure 3. Memory test with predetermined target (baby bottle), three fixations after target object, (Zelinsky & Loschky 2005, p.679)

2.5 Spatial ability

Spatial ability is required in everyday tasks such as finding the way from a location A to location B or memorizing a path on a paper map (Hegarty et al. 2002). According to Hegarty & Waller (2005), spatial ability is basically the handling of spatial information. Cognitive research on spatial ability found an effect of low and high spatial individuals on the quality of the spatial representations (Hegarty & Waller 2005). Thus, low spatial individuals have more troubles to maintain a complex image after seeing it than high spatial individuals.
The two most important factors for spatial ability appear to be spatial visualization and spatial orientation. Spatial visualization is the ability of mentally manipulate a visually seen object. As an example, the mental rotation test suggested by (Shepard & Metzler 1971) measures the ability to manipulate an object in different dimensions. For such mental rotation tasks individuals need to store and process information at the same. Therefore, especially spatial visualization is linked with tests of spatial short-term memory (STM) (Miyake et al. 2001).

A lot of studies actually investigated in the effect of gender on spatial abilities (Voyer & Bryden 1990; Lawton 1994; Montello et al. 1999; Hegarty & Waller 2005). Nevertheless, this thematic is not described in more detail because gender differences are not included in the analysis of the results. It is interesting to know how spatial ability might affect the performance of spatial tasks even though Richardson et al. (1999) indicated that psychometric measures do have weak predictive power for large-scale (environmental) spatial ability. A study by Pazzaglia & De Beni (2006) differentiated between participants of low and high mental rotation (MR) and investigated the effect of MR in pointing tasks. The results indicate that participants with low MR (low spatial) tend to focus on salient landmarks or other visual features whereas high spatial participants use both landmarks and route perspectives. Another study by Thoresen et al. (2016) showed that participants with high mental rotation ability (MRT) perform better than low spatial participants in a map-based route learning task. Those studies are examples that try to link spatial ability to map-reading tasks and therefore it will be interesting to do so in this thesis as well.

3 Methods

The methodological chapter sets the focus on the computational analysis of the geographic visualizations regarding visual complexity and the experimental eye-tracking study conducted with multiple participants. The computational analysis sets the basis for the main study, because its algorithms quantify and therefore categorize the used satellite images into groups of low or high visual complexity. FC, SE, edge density and image segmentation are all well-known algorithms (see chapter 2.2) and have already been used in several studies so far. All code for the four algorithms runs in Matlab and was freely available online by the authors. Both 50 urban and rural satellite images from Swisstopo as well as four specific urban satellite maps were used as input for the following four algorithms.

3.1 Computational analysis

3.1.1 Feature Congestion / Subband Entropy

Both FC and SE codes were made available by Rosenholtz et al. (2007). The code was downloaded and then implemented in the Matlab environment. At first, local clutter for colour, contrast and orientation was computed. The input image (e.g., a RGB Tiff-File) is used to compute the colour, contrast and orientation clutter. In order to have an averaged clutter value, the three values for visual clutter were merged into a single metric. This metric was then used to compare input images.

The SE metric was also computed with a Matlab code by Rosenholtz et al. (2007). The code was available through the same package and works similarly to the FC algorithm. The final SE clutter value is computed through the number of scales ("wlevels") and the weight on chrominance ("wght_chrom"). For both algorithms, only default parameters were used, as Rosenholtz et al. (2007) did in their study a few years ago. The full code for FC and SE is listed in A 1. Feature Congestion and A 2. Subband Entropy as it was used for the computational process.

3.1.2 Edge Density

The edge density measurement was first applied by Mack & Oliva (2004) and then by Rosenholtz et al. (2007) combined with the FC and SE measurement. The algorithms were not computed with geographic maps that is why satellite maps are now used for those algorithms. Edges were detected with the canny edge function in Matlab and the density is basically the ratio between the number of edge pixels and the total number of pixels in an image. The *edge* function returns a binary image (black and white, BW) containing values (1) for places where the function detects edges and values (0) elsewhere. By default, the *Sobel*[9] edge detection is used as one possible method. As Rosenholtz et al. (2007) used the *Canny* method, the same method was used for this thesis. The Canny method requires the input image as intensity or binary image, a threshold and sigma. The parameters were set both to 0.11 and 0.27 as a low and a high threshold and 1 for sigma, as suggested by Rosenholtz et al. (2007). The full code is listed in A 3. Edge Density as it was used in Matlab.

3.1.3 Image Segmentation

The code of Felzenszwalb & Huttenlocher (2004) was freely available on their website. Due to technical problems in successfully implementing the code, an adapted code, which was created by Freytag (2014) on GitHub[10], was used.

The code includes a well-structured readme file that helps understanding what the files and the different parameters are used for. As default, a simple example image ("sheep image") is available to run the algorithm. Two different sets of parameters were used to calculate the number of regions in every file. The first set was default, the second one adapted in a reasonable way. The full code is listed in A 4. Image Segmentation with the default parametrization in Matlab.

3.2 Participants

A total of 40 participants (male and female) were recruited for the experiment, which took place in January 2017. Almost half of the participants (n=21) were or are still geography students at the University of Zurich whereas the others (n=19) do not have any experienced geography knowledge. All participants are aged 22 to 32. The experiment was held in the Eye Movement Recording Lab (Y25-L9) at the Institute of Geography University of Zurich (GIUZ) and took about 60 minutes per participant. All of them were fluent in German and completed all tasks on the same workstation. Based on technical issues with the software, eye-movement datasets of two participants had to be removed from the analysis. All participants received some candies after finishing the exhausting experiment.

[9] https://ch.mathworks.com/help/images/ref/edge.html (accessed: 27.10.2016)
[10] https://github.com/ (accessed: 27.10.2016)

3.3 Materials

In the beginning of the experimental design process *surveymonkey*[11] was operated but due to insufficient functions surveygizmo became the first choice. So, the user study was finally executed on *surveygizmo*[12]. The surveygizmo platform allows to combine several questionnaires within the same project and to include multiple images on every page. The most valuable function is a page timer for every single page in order to measure time as one of three dependent variables. When the time was up, the survey automatically skipped to the next page. Like the participants could conclude the experiment on their own.

3.3.1 Stimuli

The following map types are part of the user study and will now be described in more detail. All map types were created and adapted in Adobe Photoshop CS6, Adobe Illustrator CS6 and ArcMap 10.4.1.

Basic map *type*

Platforms such as USGS[13], Ordnance Survey[14] and Swisstopo[15] do provide several map layers in a free download. After testing the three platforms, Swisstopo was used because it proved to be very user friendly and has a very high quality standard regarding satellite and topographic images. The desired extent and scale can be manually chosen and the background map type can be switched between satellite and topographic mode. Once the extent is ready for download the file will be saved as a PDF and may later be imported into tools such as Adobe Photoshop in order to be saved as a TIFF image.

Both 50 urban and 50 rural maps were downloaded in order to conduct a first computational analysis. The scale was set to 1:5,000 as default and all locations are situated in Switzerland. Urban maps are characterized by object features such as buildings, streets, train lines and occasionally with some very small parks whereas rural maps are containing wide green areas, dense forest or hydrological features such as rivers or lakes. It was important to use maps representing heterogeneous objects and features. As an example, an image with a big river dividing a dense housing area into two or a large forest covering half of the image was avoided. The use of 100 images for an eye-tracking study would have blown up the time-frame, therefore, a subset of 4 urban images was chosen. The following four map types were used for the user study.

[11] https://de.surveymonkey.net/ (accessed: 17.10.2016)
[12] https://www.surveygizmo.eu/ (accessed: 11.01.2017)
[13] https://www.usgs.gov/ (accessed: 31.10.2016)
[14] https://www.ordnancesurvey.co.uk/ (accessed: 31.10.2016)
[15] https://map.geo.admin.ch/ (accessed: 31.10.2016)

Blank map type

This first type is basically a plain satellite image, which shows the real texture and colour from a top down view, Figure 4 illustrates a satellite image of the city of Berne. In the lower side of the image the source and the extent of the image are displayed for reasons of copyright and usability. This map type is meant to be the basis of both computational analysis and user study. From now on, this map type will be named *blank* map type.

Figure 4. *Blank* map type, (map.geo.admin.ch)

Street map type

The second map type consists of a plain satellite image with the street network as an overlay. In Figure 5 the street network is shown in a light greyish colour. Google maps stands as an example of how a street layer can be visually embedded into a satellite map view. The vector 25 street dataset of Swisstopo was used as street layer. In ArcMap 10.4.1 the street network was altered and clipped to the desirable extent. From now on, this map type will be named *street* map type.

Figure 5. *Street* map type, (map.geo.admin.ch)

Label map type

The third map type basically includes the same satellite image but also contains labels. In Figure 6 various labels added to this map type can be seen. In Adobe Illustrator CS6 the topographic map layer of the same extent was used to place the labels on the satellite map. The labels describe either public transportation stops or names of quarters. The colour and size of the labels was not altered in any way in order to keep it consistent. From now on, this map type will be named *label* map type.

Figure 6. *Label* map type, (map.geo.admin.ch)

Hybrid map type

The last map view combines both elements of the second and third map type. Both street layer and labels are displayed in Figure 7. According to Hoarau et al. (2013), the method of adding point, line or polygon features from topographic maps onto satellite images is called "adaptive symbolization". The overlay of the street layer highlights the street features and makes them visually more salient. According to Hoarau et al. (2013) roads are traditionally symbolized with a light or bright colour and a black outline to make them visible in every part of the satellite map. For this study, the method proposed by Hoarau et al. (2013) was applied as can be seen in Figure 7 with labels and the street layer in grey with a thin black outline. From now on, this map type will be named *hybrid* map type because it includes both features of the *street* and *label* map type.

Figure 7. *Hybrid* map type, (map.geo.admin.ch)

<u>Topographic map view</u>

In an earlier stage of the thesis it was planned to include two types of topographic map view into the computational analysis and the user study. The upper map in Figure 8 is altered with photorealistic texture from the buildings in the satellite map view whereas the lower map is basically a generic topographic map of Swisstopo. The merging of these two map types would have been an interesting issue but was abandoned due to lack of time. Nevertheless, both map types are illustrated and mentioned in this section because a big effort was put into the creation of the upper map in Figure 8.

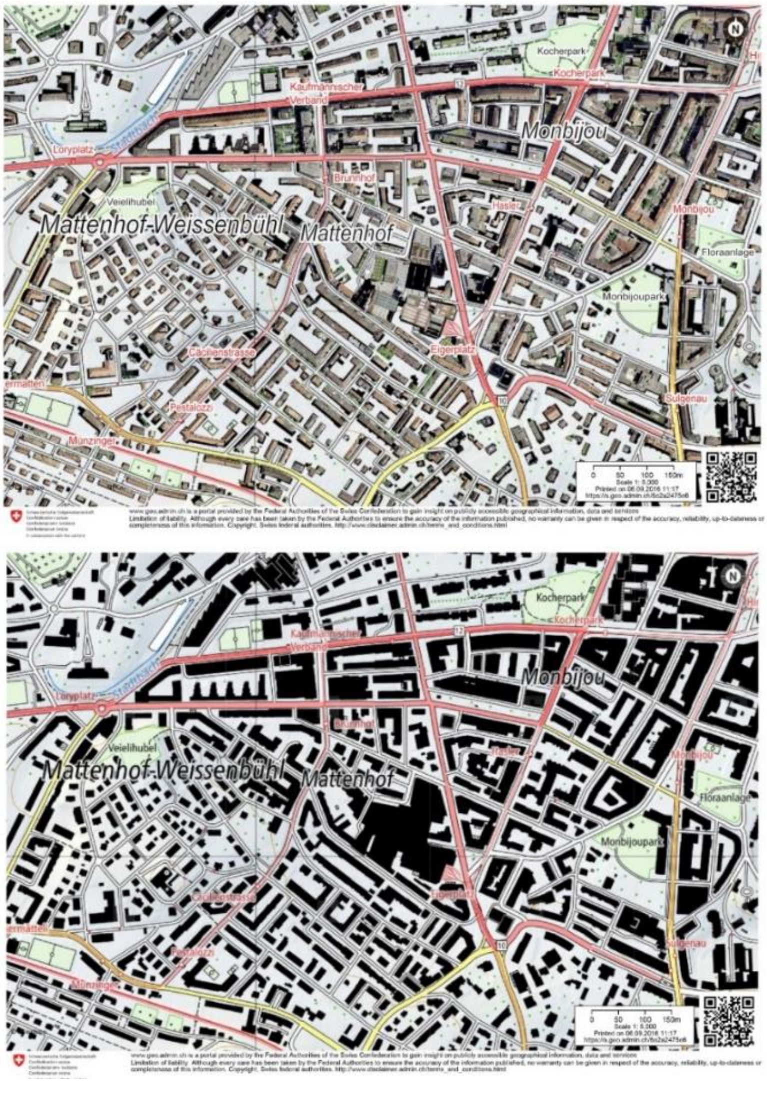


Figure 8. Above: Topotexture map type. Below: Topographic map type, (map.geo.admin.ch)

3.3.2 Apparatus

The experiment took place in the Eye Movement Recording Lab (EML) at the GIUZ. The room is windowless to ensure that participants will not be distracted by any environmental conditions. External factors were held constant as good as possible. All participants had both the same light conditions and the same room settings (e.g., same chair and monitor) and had to follow the same procedure.
The EML is equipped with a Tobii TX300 eye tracker, with a sampling rate of 300 Hz and a gaze accuracy of 0.4 degree[16]. The workstation was running on Windows 7 for the experiment. The 23'' display has a maximum resolution of 1920 x 1080. The data analysis itself was done at the EML, where the data was recorded because the dataset was very large.

3.3.3 Pre-Experiment Questionnaire

The questionnaire consisted of three pages grouped by different topics. On the first page, participants were asked about their age, gender and their degree of education. Then, questions had to be answered whether they use glasses or contact lenses or suffer from colour blindness. On the last page, participants were asked about their geographical knowledge and how often they use digital geographic maps such as Google maps and whether they prefer the satellite or map view when using digital maps.

3.3.4 Map Memory Task

Following the questionnaire, a map memory test was used to determine the spatial ability of the participants. The map memory test was used by Ekstrom et al. (1976) as one of multiple cognitive tests that tests the ability to remember part of a map and recognize it when it is shown again. The test consists of a learning- and a recognizing section. In the learning section, each participant has 3 minutes to study 12 maps in black and white such as illustrated in the left map of Figure 9. Thereafter, the recognizing section appears and 12 black and white maps are shown again, illustrated in the right map of Figure 9. The participant again has 3 minutes to choose whether he/she has seen ("Yes") or not seen ("No") every single map. Afterwards, in a second part, the same procedure is repeated with another set of 12 maps (learning- and recognizing section). Participants were allowed to skip to the next page before the time was up. Each participant has to give 24 (12 x 2) responses and can therefore achieve a total score of 24 points. Every wrong answer is subtracted from the total score. So, there was no advantage for the participants making a wild guess whenever they had no clue.

[16] https://www.geo.uzh.ch/en/units/giva/services/eye-movement-lab.html (accessed: 06.03.2017)

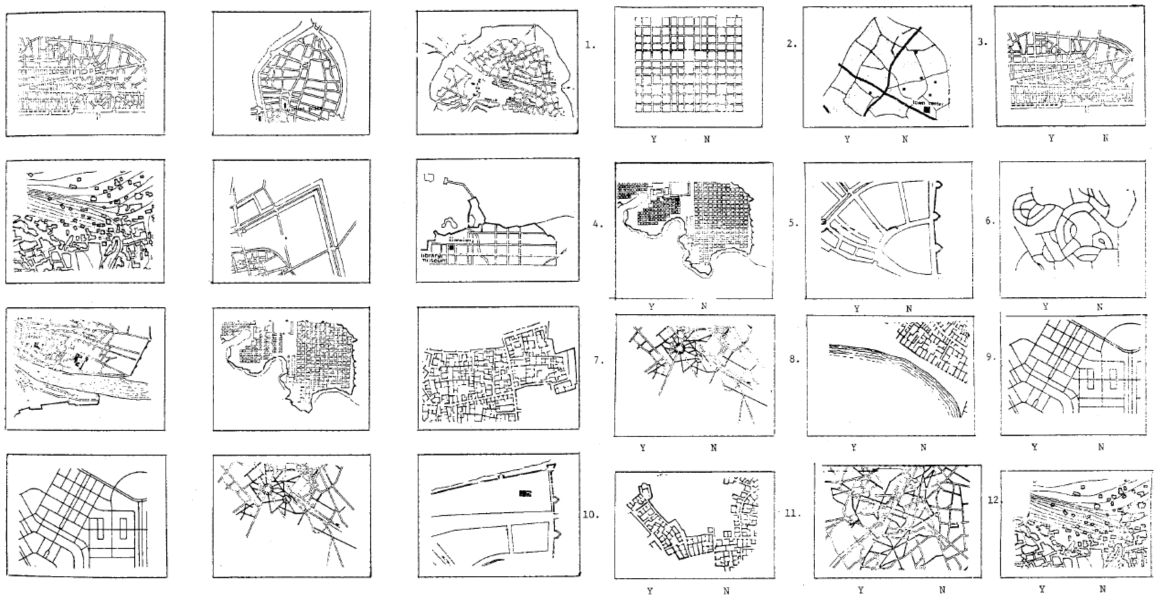


Figure 9. Left: Learning section of map memory test. Right: Recognizing section of map memory test, (Ekstrom et al. 1976, pp.155–157)

3.3.5 Main Experiment

In this section, all four tasks are presented in more detail. Each task follows the same experimental design: there is a learning section, followed by a recognizing section. The four tasks can be separated into a global and a local task group. The visual search and change detection task by Eng et al. (2005) was used as inspiration and adapted in a way that technical features remained reasonable. Four satellite maps with four different geographic locations (i.e., Basel, Bern, Lausanne and Geneva) were used for the learning section. Each satellite map comprises the four map types that were presented in section 3.3.1. Summarising, 16 initial satellite maps consisting of four different geographic locations and map types were shown for every map memory task.

It was taken into account that no more than four labels were placed on a map in the recognizing section as the storage capacity of VSTM is limited by the number of objects (Luck & Vogel 1997; Cowan 2001). Labels that were changed in the recognizing section were either from other satellite maps or the initial map itself. Thereby, participants were excluded, which were focusing at the labels only without even noticing the road network or background image. Principally, the intention was that all participants would recognize the map as a whole and see the big picture (e.g., Label XY is placed left of the green park).

Global task group

This group consists of two tasks (i.e., Task 1 & Task 2). The task group is called “global” because participants were asked to study the initial map as a whole.

In *Task 1* participants viewed a specific map type (*blank, label, street* and *hybrid,* see chapter 3.3.1) in full-size in preparation for a memory test. Participants had a maximum of 20 seconds to study the map as illustrated in Figure 10. The application automatically skipped to the next page when the time was up.

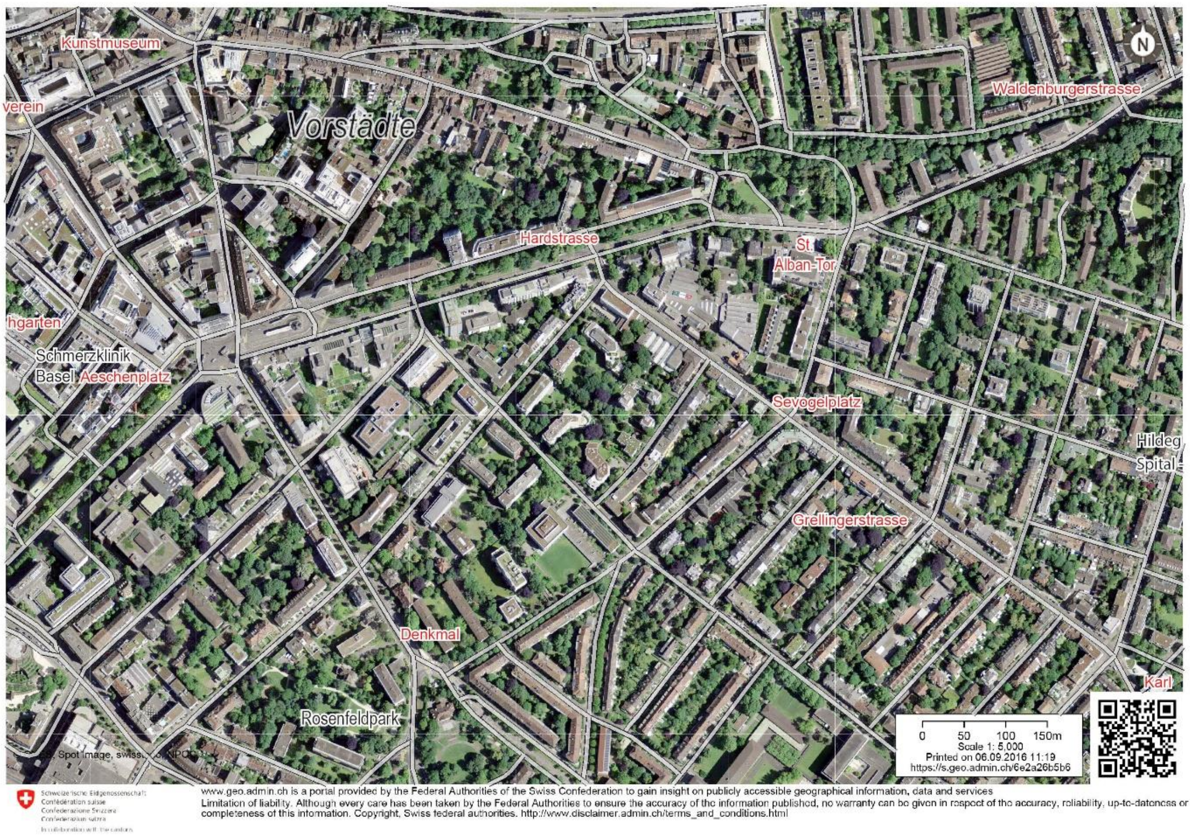


Figure 10. Learning section for Task 1 with *hybrid* map type, (map.geo.admin.ch)

Figure 11 illustrates the recognizing section for Task 1. Four maps are given as possible answers but only one is correct. Each participant had to answer which of the four maps he/she has already seen in the learning section. They had 15 seconds to answer the question and tick the right box (A, B, C or D) in this case is letter B correct.

The map type remained the same within one learning- and recognizing section. So whenever the map was displayed with labels in the learning section, the four possible options all included labels as well. Beforehand, a target box was chosen. The target box always had the size of 1000 x 600 pixels. Every extent (map location) and the four corresponding map types do have the same target boxes in order to compare the results within the different map types. The location of the target box was chosen depending on the visual saliency of objects in a map (Itti & Koch 2000; Itti &

Koch 2001). Thus, very salient objects were excluded. The centre of the map was not chosen as a possible location for a target box because people first focus on the centre as it is the shortest way from the eyes to the monitor.

Figure 11. Recognizing section for Task 1 with four possible answers, (map.geo.admin.ch)

Task 2 symbolizes a typical VST, as proposed in the study by Eng et al. (2005). Test persons were asked to study a small map with an extent of 1000 x 600 pixels for 10 seconds, illustrated in Figure 12. Afterwards they had to find the recently viewed map section in a bigger map again. One or two options out of four could be correct. They had 25 seconds to solve this VST and to confirm by clicking either A, B, C or D. As four images did not have enough room on a regular webpage, participants were asked to scroll up and down on the page in order to find all possible options.

Figure 12. Learning section with *hybrid* map type for Task 2, (map.geo.admin.ch)

Local task group

The local task group consists of Task 3 and Task 4. Both the tasks in this group required participants to study only a local area of a map. The task design remains the same as in the global task group, with one learning- and recognizing section.

In the learning section of *Task 3*, participants were asked to study the area around a visible blue dot. In the test example four possible target boxes (above/left/below/right) were illustrated as in Figure 13. All four boxes do border on the blue dot and do have a size of 1000 x 600 pixels. They indicate to each participant how far away out of the blue dot his or her range of vision should reach. After 15 seconds, the application switched to the recognizing section. There, participants saw four map extents and had another 15 seconds to answer the following question: "Which map is situated left of the blue dot?" The spatial components ("above/left/below/right") were chosen for each of the four different locations of the satellite map and remained constant. So, the question for the one satellite map could be like: "which map was left of the blue point" whereas the question for another satellite map was like: "which map was underneath the blue point."

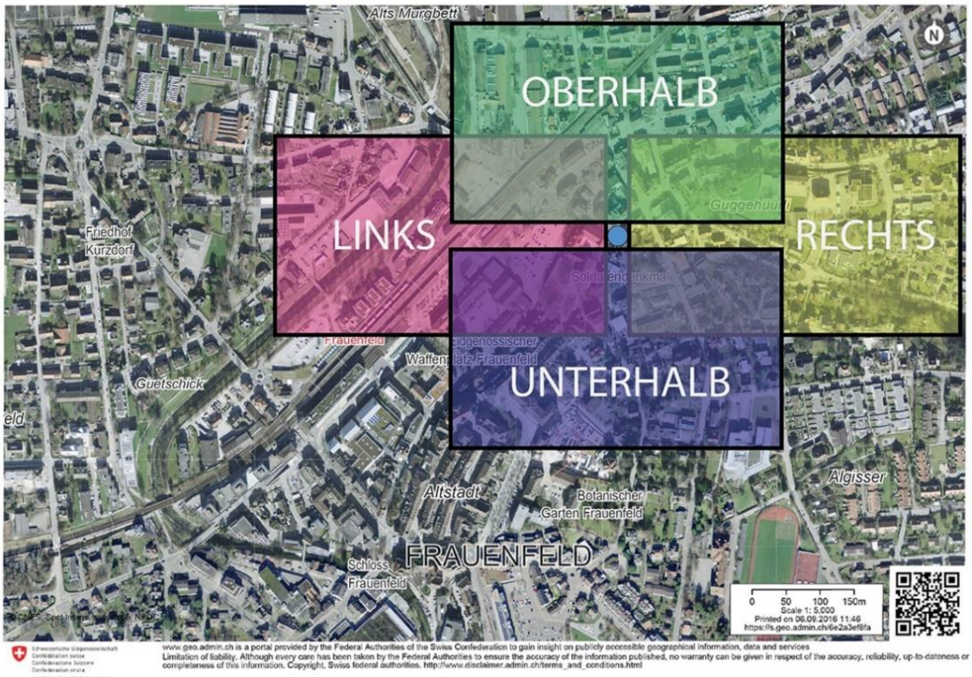


Figure 13. Blue point (centre) with target boxes (coloured) for Task 3, (map.geo.admin.ch)

In *Task 4*, the task design remains the same as in Task 3 but in the learning section two instead of one blue dot are visible. Participants were asked to focus on the surroundings of the blue dots for 25 seconds. On the next page, only one of the two dots was visible within an image frame for a very short time. Then, subjects had to guess the correct map extent to the given question within 15 seconds. As in Task 3 the question in Task 4 included a spatial component (e.g., left of…/underneath of…). As you might expect this is by far the most difficult task of all.

3.4 Experimental Design

3.4.1 Experimental Procedure

Potential participants received an E-mail containing a Doodle link[17] with slots to choose from. One slot stands for an hour and there was no buffer time between two slots. Participants were asked to read a consent form the day before they came to the experiment. In the experiment room each participant signed this form and received a signed copy. Then, they were asked to take a seat in front of the eye tracking system. All participants were instructed to take a comfortable sitting position and to be able to reach the crystal ball placed on the table on their left hand side. The crystal ball is supposed to stabilize the participants' upper body and consequently control the calibrated eye position.

Before the eye tracking calibration started, participants were also instructed to look at the four corners of the computer screen. This makes sure whether extreme positions are being monitored by the eye tracker. For the calibration of the eye tracking device 9 fixations points were used because it is the default setting in eye-tracking software.

Once the calibration process was successful, participants could start the experiment. At first, a short pre-questionnaire about participants' personal background had to be filled out, then followed by the map memory task, see section 3.3.4. Finally, the main experiment with four visual search tasks started (see section 3.3.5) with a given order, as illustrated in Table 2.

The experimenter was connected with the system through TeamViewer 12[18] in order to have full control over the experiment in case something unpredictable happened. Before every task, the experimenter typed in a unique ID number for each participant. Participants having questions of understanding were instructed to ask them after the example was finished and before the main experiment started. Additionally, a summary of task objectives was illustrated. During the experiment neither questions were answered by the experimenter nor could any verbal interaction take place as it would have influenced the concentration as well as the result.

The four tasks from the main experiment were counterbalanced according to the scheme proposed by Martin (2008). This scheme helps the experimenter to keep the learning effect as low as possible. Probably the most useful scheme is a balanced Latin Square (Martin 2008). There, every level (e.g., task) appears at every position equally often as one can see in Table 2. Additionally, each level precedes and follows each of the other levels equally often (Martin 2008). As an example, Task 1 was followed by Task 2, Task 3 and Task 4 equally often. The balanced Latin Square was continued to the last participant (e.g., see Table 2).

[17] http://doodle.com/ (accessed: 10.1.17)
[18] https://www.teamviewer.com/de/ (accessed: 10.1.17)

As the whole experiment in Tobii Studio was fully automated, a page timer in the bottom centre of the screen showed the remaining time. Participants were allowed to skip to the next page even before the time was up.

Table 2. Balanced Latin Square, (Martin, 2008)

	Order of presentation			
Participant 1	Task 1	Task 3	Task 2	Task 4
Participant 2	Task 3	Task 4	Task 1	Task 2
Participant 3	Task 4	Task 2	Task 3	Task 1
Participant 4	Task 2	Task 1	Task 4	Task 3
…	…	…	…	…
Participant 40	Task 2	Task 1	Task 4	Task 3

3.4.2 Independent Variables

Based on Martin (2008) a within-subject design was chosen. All participants were exposed to all map types and had to go through all four task types. An advantage to use a within-subject design is that fewer participants are needed in order to conduct a user study (Martin 2008). As participants did not get any financial compensation the recruiting process can sometimes be a slightly difficult. Thus, it was good to know that a total of 40 participants were enough for a study with a within-subject design. On the contrary, a disadvantage is that participants might learn through the order effect of the input images. The balanced Latin Square scheme tried to minimize the effect of learning for the participants (see section 3.4.1).

The design of the study is based on the two independent variables, namely *map type* and *task type.* The map type is considered as the main and the task type as a minor independent variable. Both independent variables do have four different levels, as presented in chapter 3.3.1, which results in a 4 x 4 factorial design (Martin 2008). Map types were randomly ordered by surveygizmo[19] in order to keep a confounding effect as low as possible and task types were ordered with a balanced Latin Square scheme (Martin 2008).

[19] https://www.surveygizmo.eu/, accessed: 11.1.17

3.4.3 Dependent Variables

The following three dependent variables were measured during the experiment:

Accuracy/Effectiveness. Participants had to find correct answers in all the tasks. The sum of all correct answers divided by the total number of tasks resulted in the accuracy score. By example, a participant that was able to answer half of the questions, of which all were correct, in the given time. Then, the accuracy score would be 50% because the other half of the questions remained blank. Wrong answers were not counted as minus.

Response Time (RT)/Efficiency. This variable was derived from the time participants needed to give each of the answers. A survey tool, namely "page timer" from surveygizmo saved the time for every page.

Confidence. After each map type was tested at last, participants were asked about their confidence. They were instructed to mark their confidence level at a five-point Likert scale (Likert 1932) between 1 (unsafe), 2, 3 (safe), 4 and 5 (very safe).

3.5 Statistics

All results were analysed with the IBM SPSS 21 Software[20]. Participants used all independent variables (e.g., map type) in the whole experiment. That is why repeated-measures analysis of variance (ANOVA) were performed (Field 2013). The independent-samples T test is used for the comparison of urban and rural satellite maps and pairwise comparisons were used to test any found connections between two map types. In order to use the repeated-measures ANOVA a few conditions, such as normally distributed data and variance homogeneity, have to be fulfilled. The results indicated that all data is normally distributed and the Mauchly's Test of Sphericity showed no significant difference for which reason the data has variance homogeneity[21].

Descriptive statistics are always declared with a mean (M), standard deviation (SD) and number of participants (N). Each result of the ANOVA test includes F-ratio (F), degree of freedom (df), p-values and the partial eta squared ($\eta 2p$). All diagrams were created and adapted in IBM SPSS 21.

[20] https://www.ibm.com/analytics/ch/de/technology/spss/spss-trials.html (accessed on 13.3.2017)
[21] http://www.methodenberatung.uzh.ch/de.html (accessed: 10.02.2017)

4 Results

This chapter shows separate results of the computational analysis and the user study. The computational analysis will differentiate between scene type (urban/rural) and the four different map types (*blank*, *street*, *label* and *hybrid*). The results of the experiment are shown for every task and map type including the effect of spatial ability. Additionally, eye-movement data is included to support the results. Significant results are marked with asterisk in diagrams. One asterisk ($p < .05$), two asterisks ($p < .01$) or three asterisks ($p < .001$) are used in the bar charts.

4.1 Computational Analysis

4.1.1 Image Compression

For the image compression only a subset of urban and rural satellite maps was used. The image file sizes for both 25 urban and rural satellite maps were calculated uncompressed and with the following compression methods: LZW, ZIP and JPEG. Figure 14 indicates that JPEG compression is most effective with a compression ratio of 36.82% for urban and 40.36% for rural areas. Compression ratios for both ZIP and LZW are between 17% and 26%. An independent-samples T test reveals that compression ratio for rural maps is higher than for rural maps for LZW *[t(48) = -6.638, p = .000]*, ZIP *[t(48) = -6.655, p = .000]* and JPEG *[t(48) = -6.644, p = .000]*. So, after the compression the file sizes for urban maps are higher than for rural maps regardless of the compression method.

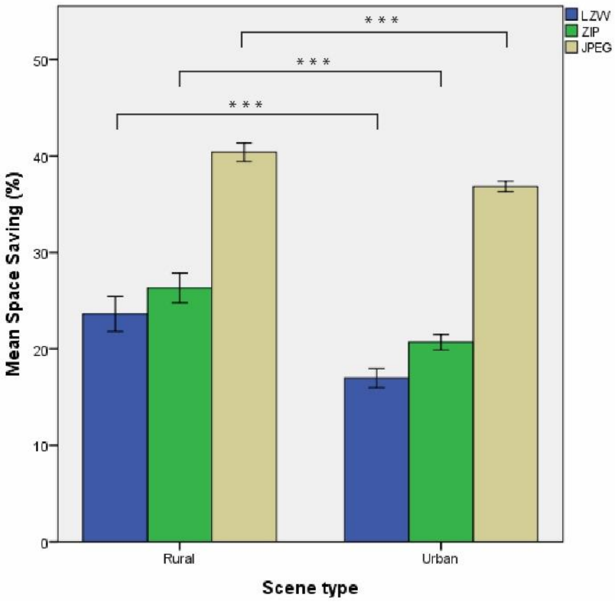


Figure 14. Compression ratios for urban and rural areas. Error bars: CI

The image compression rate for the subset of urban map types is only calculated for the JPEG compression method, because the images are used in the JPEG format only for the user study. Figure 15 shows the differences in mean JPEG file sizes within the four map types.

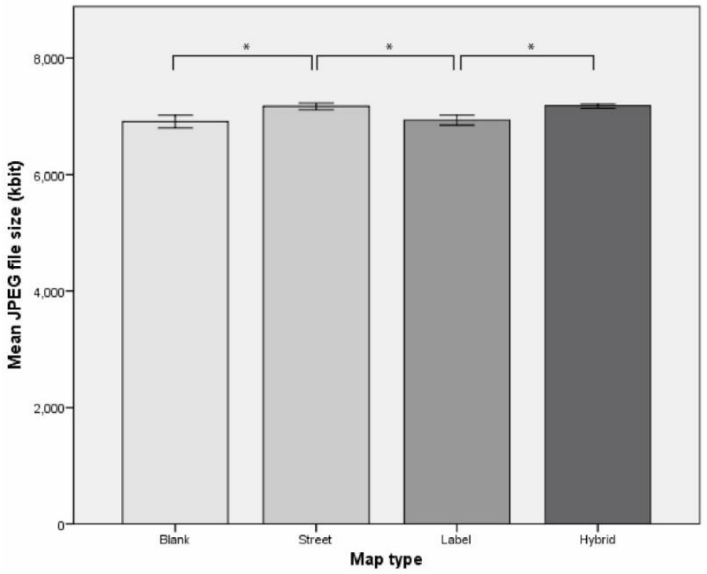


Figure 15. JPEG file size (kbit) for different map types. Error bars: CI (95%)

A repeated-measures ANOVA finds an effect of the map type on the JPEG file size *[F(3,9) = 41.726, p = .000,* η_p^{2} *= .933]*. Pairwise comparison shows significant differences for *blank-street (p < .05), street-label (p < .05)* and *label-hybrid* map types *(p < .05)*. The highest mean file size is for the *hybrid* map *(M = 7178.750, SD = 22.867)* and lowest for the *blank* map type *(M = 6910.750, SD = 67.663).* The JPEG file size for *blank* and *hybrid* map type are slightly not significant *(p > .05)* but indicate that the more information is within an image, the higher the JPEG file sizes are.

4.1.2 Scene Type

_FC/SE_

The FC/SE algorithm proposed by Rosenholtz et al. (2007) was applied to the sample of both 50 urban and rural satellite images. As shown in the left bar chart of Figure 16, the mean FC clutter score for satellite images from an urban area are higher (*M = 5.239, SD = 0.421*) than for rural satellite images (*M = 3.209, SD = 0.277*). That results in a difference of about 39% between the mean FC score for urban and rural satellite images.

An independent-samples T test was conducted in SPSS between the two variables and indicates that urban and rural FC measures are significantly different from each other *[t(98) = 28.451, p = .000]*. This is not a surprise as urban satellite images contain a lot of rooftops with different colours and dense infrastructure leads to high orientation contrast.

In the right bar chart of Figure 16, the mean SE score for both urban and rural satellite images is shown. Mean SE score are higher for the satellite images of urban areas (*M = 4.888, SD = 0.089*) than for satellite images of rural areas *(M = 4.251, SD = 0.212*). The results of the independent-samples T test indicate that there is a significant difference between rural and urban satellite images *[t(98) = 19.567, p = .000]* regarding mean SE values. Nevertheless, the mean between urban and rural for SE is less pronounced than for the FC algorithm. A possible reason for that will be addressed in more detail in section 5.

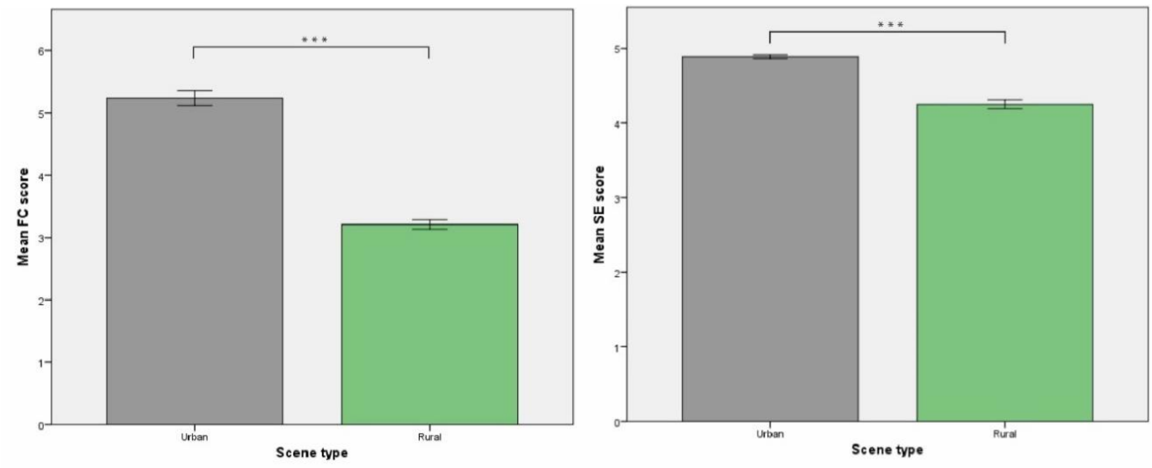


Figure 16. Left: Mean FC score. Right: Mean SE score. Error bars: CI (95%)

Edge density

The bar chart (left) in Figure 17 illustrates the mean edge density for both urban and rural images. Two different thresholds were set (0.11 and 0.27) for the computation of the mean scores. The first two bar charts represent measures for the low threshold for urban (grey) and rural (green) satellite maps. The last two bar charts represent the measures for the high threshold for urban and rural satellite maps. Around 15% of edge pixels were detected for urban satellite images for the lower threshold (e.g., 0.11). Consequently, a higher threshold (e.g., 0.27) results in a lower mean edge density of 8.3 % for urban satellite images. The same trend exists for rural satellite images but with lower mean values of 10% and 2.6 % for low respectively high thresholds. An independent-samples T test reveals significant results within the mean edge density for urban and rural measures for the low *[t(98) = 9.832, p = .000]* and high threshold *[t(98) = 20.749, p = .000]* parameter. So, dense areas with houses and street infrastructure in urban areas lead to a higher edge density than in rural satellite images and therefore represent higher visual clutter.

Segmentation

Image segmentation was calculated with default parameters according to Felzenszwalb & Huttenlocher (2004). The image segmentation processing algorithm computes the same pictures as in the previous algorithms. In the bar chart (right) of Figure 17 the default computation results in 9,000 and 4,600 regions for urban (grey) respectively rural (green) satellite images. The fact that the outcome for each of the four algorithms indicates urban satellite images to be visually more complex than rural satellite images, the experimental study was only conducted with urban images. An independent-samples T test reveals a significant effect of the scene type on the number of regions *[t(98) = 25.087, p = .000]* with the default parameters.

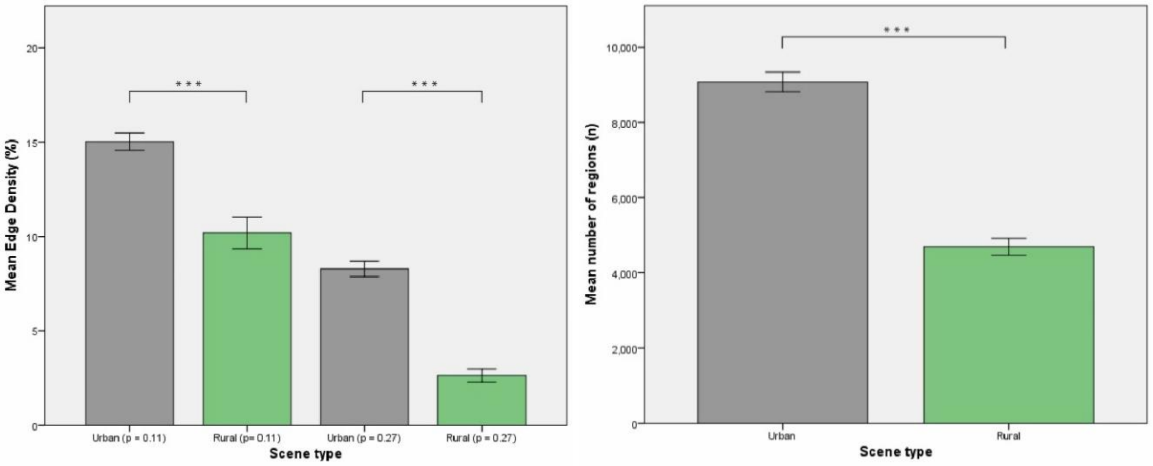


Figure 17. Left: Mean edge density (%). Right: Mean number of regions (n). Error bars: CI (95%)

4.1.3 Map Type

A subset of 10 urban satellite images was chosen for the computational analysis for each map type. In order to have a basis of how visually complex the four map types are their results are highlighted in this subsection.

<u>*FC/SE*</u>

Default parameters as in the work by Rosenholtz et al. (2007) were applied to the subset of satellite maps. In Figure 18 both bar charts show that scores (left/right) for FC and SE are lowest for the *blank* map type placed on the left-hand side in every chart. The *blank* map type is regarded as the basis for comparison between the other map types because it generally represents a raw satellite image without additional information overlaid.

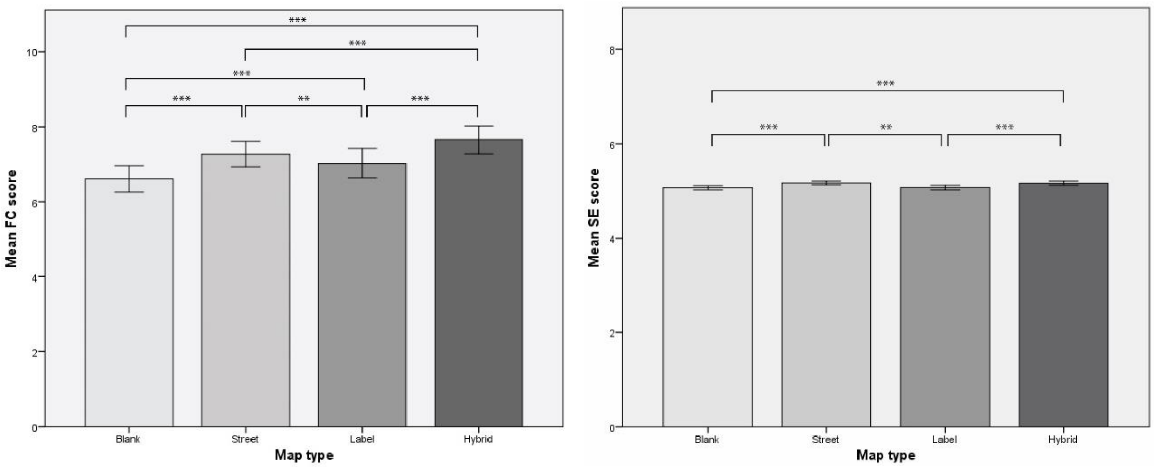


Figure 18. Left: Mean FC score and map types. Right: Mean SE score and map types. Error bars: CI (95%)

The other map types do all entail some additional information and scores are consequently different for each map type even if very little. In the left bar chart of Figure 18 the *hybrid* map type reaches the highest score (*M= 7.653, SD = 0.520, N=10*) and a repeated-measures ANOVA finds a significant effect of the map type on the FC measure *[F(3,27) = 116.651, p = .000,* η_p^2 *= .928]*. There is a significant difference (p < .05) between all map types for FC in Figure 18 (left) based on pairwise comparison.

The SE algorithm results in similar values between all four map types, as illustrated in the right bar chart of Figure 18. A repeated-measures ANOVA reveals that the map type has an effect on the SE complexity measure *[F(3,27) = 156.266, p = .000,* η_p^2 *= .946]*. There is no significant difference for map types *blank-label* and *street -hybrid (p > .05)*, as illustrated in Figure 18 (right).

Edge density

For the sample of 50 urban and rural satellite images a low and a high threshold was used in order to show an effect in the result. Now, only the low threshold of (p = 0.11) was used for the computational analysis for the subset of 10 images in order to be consistent with other algorithms that also use only one parametrization. The mean of edge pixels in the four map types ranges from *blank* satellite images ($M = 18.870$, $SD = 0.011$, $N = 10$) to hybrid satellite images ($M = 19.395$, $SD = 0.009$, $N = 10$), illustrated in Figure 19. A repeated-measures ANOVA shows that the map type has an effect on the edge density $[F(1,9) = 15.025, p = .004, \eta_p^2 = .625]$. There is a significant difference between the *blank* map type to the other three map types and from the *street* map type to the *hybrid* map type $(p < .05)$.

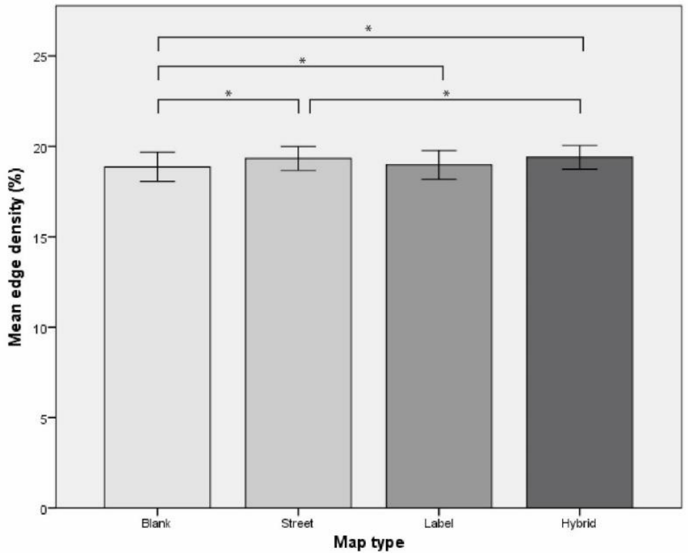


Figure 19. Mean edge density (%) for different map types. Error bars: CI (95%)

Segmentation

As for the edge density algorithm, the image segmentation algorithm only uses one default parametrization. The parameters from Felzenszwalb & Huttenlocher (2004) were used for the algorithm. A repeated-measures ANOVA shows that the map type has a significant effect on the image segmentation measure $[F(1,9) = 49.012, p = .000, \eta_p^2 = .845]$. The chart bar in Figure 20 indicates that the mean for number of regions (n) is increasing from the *blank* map type to the *hybrid* map type. Pairwise comparison shows that the *hybrid* map type has the highest mean $(M = 9265.600, SD = 661.560, N = 10)$ and is significantly different to the *blank*, *street* and *label* map type ($p < .05$). Additionally, the *blank* map type $(M = 8621.600, SD = 648.029, N = 10)$ is significantly different $(p < 0.05)$ to the *label* map type.

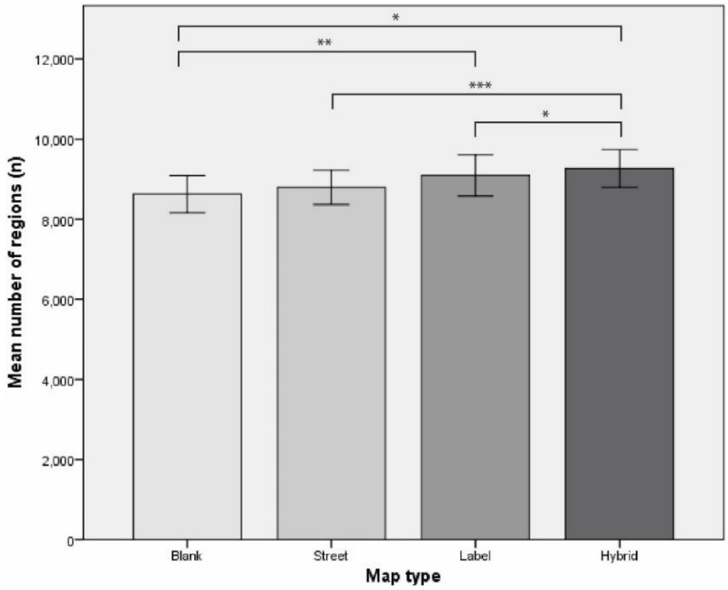


Figure 20. Mean number of regions (n) for four different map types. Error bars: CI (95%)

Summary

The four map types were tested with the four proposed algorithms: FC, SE, edge density and image segmentation. The default parameters proposed by Rosenholtz et al. (2007) and Felzenszwalb & Huttenlocher (2004) were used. The results show that the *hybrid* map type has the highest mean in three of the four algorithms (FC, edge density and image segmentation). The map types *street* and *label* are mostly between the range of the lowest (i.e., *blank*) and highest mean (i.e., *hybrid*). This indicates that the more information (i.e., *hybrid*) there is, the more cluttered a visualization becomes. To test whether more cluttered map types also affect participants' performance the results of the experiment are analysed in 4.2.3.

4.2 Experiment Results

4.2.1 Participants

A total of 40 participants took part in the experiment. There were 29 (*72.5%*) male and 11 (*27.5%)* female persons aged from 22 and 32 *(M = 25.75, SD = 2.250)*. The majority (60%) of the participants were aged between 25 and 27. Almost half of the participants (52.5%) were or are still studying geography at the University of Zurich. The rest (47.5%) are either students from other subjects or people with general education (e.g., apprenticeship).

Preferences

First, participants were asked about how often they use digital maps (e.g., google maps) on their smartphone, tablet and/or computer. They had to choose on a Likert-scale between 1 (never), 2, 3 (weekly), 4 and 5 (daily). Participants most often use digital maps weekly *(M = 3.65, SD = 0.893)*, with 70% of all participants marking option 3 (weekly) and 4.

Second, participants were asked whether they prefer the satellite or map view when using digital maps. Again, a Likert-scale (Likert, 1932) between 1 (map view), 2, 3 (no preference), 4 and 5 (satellite view) was designed. Most (70%) participants preferred the map view *(M = 2.3, SD = 1.454)*. Only 17.5% of all participants marked that they prefer the satellite view.

4.2.2 Map Memory Task

Accuracy/time

The results in the map memory task indicate that participants achieve mainly good results *(M = 0.864, SD = 0.160)* with 35% of all participants achieving a task accuracy of 100%. The task accuracy ranges from a minimum of 41.7% to a maximum of 100%. Participants could move to the next page before the time limit was up, in order to save time. The total task time was 6 minutes (360 seconds). They accomplished the task in a range between 54 seconds (100% accuracy) and another one with 289 seconds (41.7% accuracy). The mean response time (RT) was at 154.35 seconds *(SD = 60.237)*.

4.2.3 User Study

In this subsection, the results of the user study will be presented. First, overall results are presented to show the bigger picture. Then, the performance is analyzed in more detail. All results are shown for the three dependent variables: accuracy, RT and confidence. RT for wrong answers were not included in the results, as suggested by Çöltekin et al. (2010).

Main effect - Map type

To get a sense of how participants generally performed in the experiment the results are displayed for each map type, summarized over the four tasks. Figure 21 illustrates the mean accuracy for all four map types. The mean accuracy for all is around 50%. The *hybrid* map type results in the highest mean accuracy *($M = 0.563$, $SD = 0.178$)* whereas participants perform at a lower mean accuracy *($M = 0.511$, $SD = 0.165$)* for the *street* map type. Those map types represent the lowest respectively highest overall accuracy in this experiment.

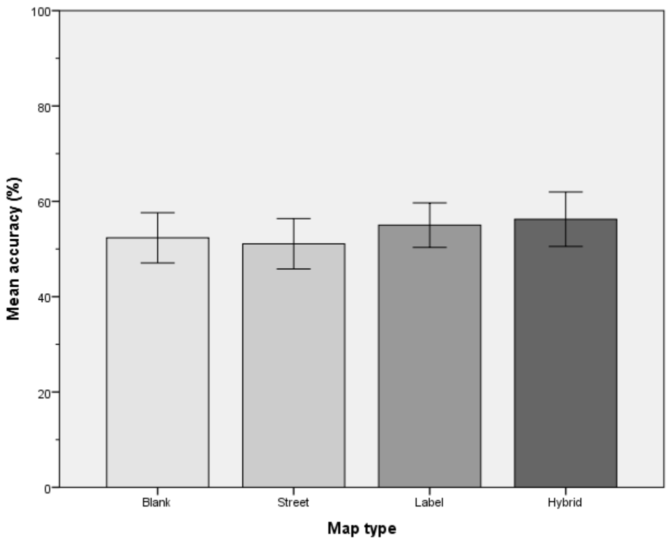


Figure 21. Mean accuracy (%) and map type. Error bars: CI (95%)

A repeated-measures ANOVA reveals that the map type does not have a significant effect on the overall accuracy *[$F(3,117) = 1.540$, $p = .208$, $\eta_p^2 = .038$]* and so pairwise comparison shows no significant effect either *($p > .05$)*. Even though there is not a significant effect, both *blank* and *street* map type reach similar mean accuracy results *($M = 0.523$; $M = 0.511$, $N = 40$)*. The *label* and *hybrid* map type also result in a similar mean accuracy *($M = 0.550$; $M = 0.563$, $N = 40$)*.

The results for RT do not indicate a grouping of two map types. A repeated-measures ANOVA indicates that the map type does not have a significant effect on RT *[$F(3,93) = 1.075$, $p = .364$, $\eta_p^2 = .034$]* and pairwise comparison shows no significant difference. Participants perform similarly fast with all four map types.

A repeated-measures ANOVA shows that the map type does not have an effect on the confidence *[$F(3,93) = 1.509$, $p = .217$, $\eta_p^2 = .046$]*. Mean confidence ranges from a minimum for the *blank* map type *($M = 2.229$, $SD = 0.680$, $N = 32$)* and a maximum for the *street* map type *($M = 2.463$, $SD = 0.769$, $N = 32$)* meaning that participants generally do not feel very confident with any of the map types.

<u>Task type – effect of map type</u>

As all four tasks are different from each other it therefore makes more sense to review results for each map type within a task. A repeated-measures ANOVA was performed for all tasks reporting all three dependent variables. Results are reported based on alpha level (p-value) minimum at .05 and, as effect size, partial eta squared. (η_p^2) is reported in order to calculate Cohen's f (Ellis 2010). In order to remind the reader of what the objective in every task was, a short description is given in the beginning of the reporting. Only measures of significant results are illustrated in the diagrams.

In *Task 1* participants were asked to view a satellite map as a whole for 20 seconds. Then on the next page, participants were asked which map extent they had seen in the previously shown map. A repeated-measures ANOVA shows that the map type does have a significant effect on the accuracy [*$F(3,93) = 6.083$, $p = .001$, $\eta_p^2 = .164$)*. According to Cohen (1992), Cohen's f is at 0.443 which corresponds to a strong effect. The bar chart in Figure 22 illustrates the results of the four map types.

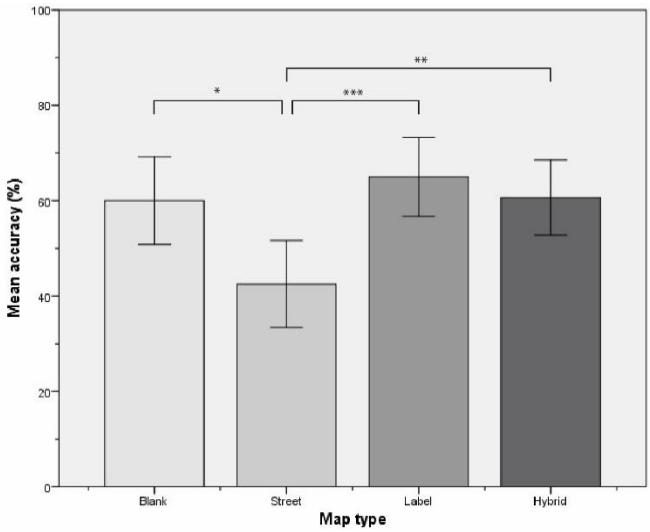


Figure 22. Mean accuracy (%) for Task 1. Error bars: CI (95%)

A pairwise comparison shows that participants perform significantly worse *($p < .05$)* with the *street* map type *($M = 0.425$, $SD = 0.284$)* in comparison to the *blank ($M = 0.6$, $SD = 0.288$), hy-*

brid (M = 0.606, SD = 0.246) and *label* map type *(M = 0.65, SD = 0.258)*. Both charts (left/right) of Figure 23 indicate that the map type does not have an effect on RT *[F(3,93) = 1.704, p = .172,* η_p^2 *= .052]* and confidence *[F(3,93) = 1.724, p = .167,* η_p^2 *= .053]* and therefore, pairwise comparison shows no significant effect either.

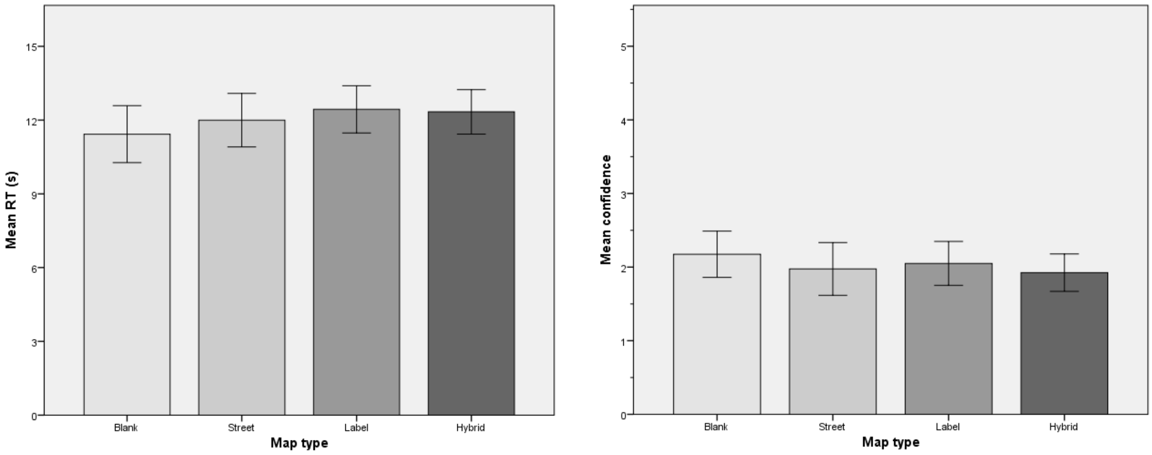


Figure 23. Left: Mean RT (s) for Task 1. Right: Mean confidence for Task 1. Error bars: CI (95%)

In *Task 2* participants viewed a small map extent for 10 seconds and then had to find this extent in four satellite maps again. For this task, map type shows no significant effect on the accuracy *[F(3,117) = 0.231, p = .875,* η_p^2 *= .006]* and confidence *[F(3,117) = 0.897, p = .445,* η_p^2 *= .022]*. However for the RT, the map type shows an effect *[F(3,63) = 3.841, p = .014,* η_p^2 *= .155*]. According to Cohen (1992), an effect size of 0.423 corresponds to a strong effect. Pairwise comparison of all map types for the three variables shows no significant effect.

In *Task 3* participants were asked to focus on the surrounding area of a blue point, which was visible on the satellite map. After 15 seconds, participants had to choose the correct map depending on the stated question. So in some cases, participants were asked what was e.g., left of the blue point. Results in this task show that the map type does not have an effect on the accuracy *[F(3,117) = 1.794, p = .152,* η_p^2 *= .044]* and RT *[F(3,45) = 1.702, p = .180,* η_p^2 *= .102]*, as illustrated in both charts (upper left/upper right) of Figure 24. As expected, pairwise comparison does not result in a significant effect.

For confidence, repeated-measures ANOVA revealed a significant effect for the different map types *[F(3,117) = 8.674, p = .000,* η_p^2 *= .182]*. That results in an effect size of 0.472, which corresponds to a strong effect according to Cohen (1992). Pairwise comparison between the map types shows that participants feel more confident *(p < .05)* with the *street* map type *(M = 2.8, SD = 1.344)* than with the *blank (M = 1.875, SD = 1.067)* and *hybrid* map type *(M = 1.975, SD = 1.025)*, illustrated in Figure 24 (lower centre).

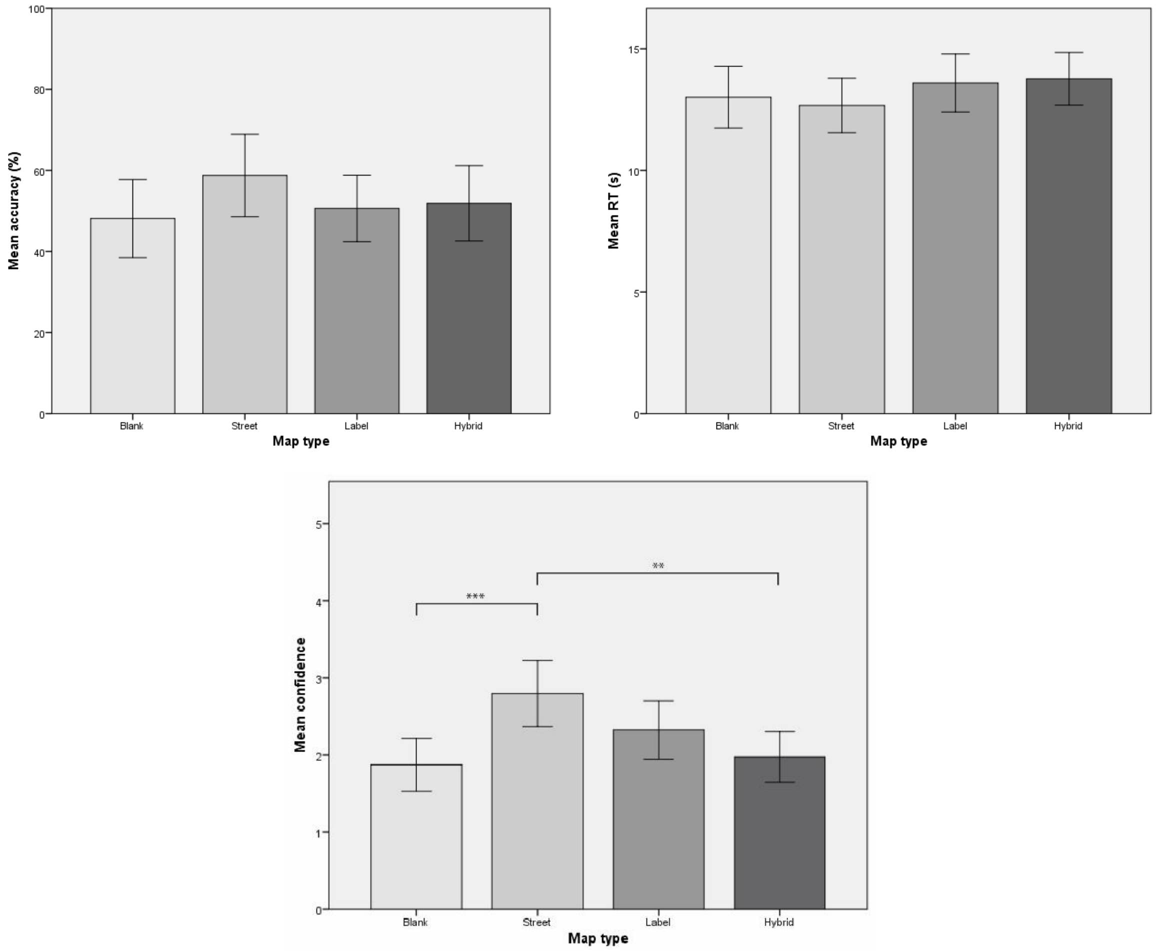


Figure 24. Upper left: Mean accuracy (%) and map type for Task 3. Upper right: Mean RT (s) and map type for Task 3. Lower centre: Mean confidence and map type for Task 3. Error bars: CI (95%)

In *Task 4* participants were asked to focus on the surrounding areas of two blue points for 25 seconds. After that, one of the blue points was chosen from the system and participants were then required to choose the correct map type depending on the stated question in the recognizing section. In this task no significant effect is detected for accuracy *[F(3,117) = 1.088, p = .357,* η_p^2 *= .27]*, RT *[F(3,24) = 1.143, p = .352,* η_p^2 *= .125]* and confidence *[F(3,111) = 0.713, p = .546,* η_p^2 *= .019]*. The mean accuracy for *blank* is the lowest *(M = 0.375, SD = 0.277)* and highest for *hybrid* map types *(M = 0.469, SD = 0.261)*. The mean values for the two remaining map types are between *(M = 0.4313, SD = 0.300)* for *street* and *(M = 0.425, SD = 0.261) label* conditions.

<u>*Effect of spatial ability – Map & task type*</u>

In order to find out whether results from the map memory task have a connection with the main experimental results a median split was conducted. The participants' data was divided into low *(N = 15)* and a high spatial *(N = 17)*. A few datasets were excluded because their score represents the median *(N = 8)*. A bivariate correlation analysis was also conducted. Like that no participants had to be excluded from the analysis. Results from both median split and bivariate correlation analysis are included because a bivariate correlation has the advantage that no datapoints are excluded and therefore the statistical power is higher.

Main effect – overall

In Figure 25 the main effect of spatial ability to the total mean accuracy (upper left), RT (upper right) and confidence (lower centre) is shown. The total mean accuracy for the low spatial participants *(M = 0.512, SD = 0.116)* is slightly lower than for the high spatial participants *(M = 0.554, SD = 0.130)*. The total mean RT for the low spatial participants *(M = 14.179, SD = 2.136)* is slighty lower than for the high spatial participants *(M = 15.278, SD = 1.615)*. The total mean confidence is following the same trend indicating that the total mean confidence is slighty lower for the low spatial participants *(M = 2.186, SD = 0.547)* than for the high spatial participants *(M = 2.396, SD = 0.661)*. An independent-samples T test shows that the spatial ability does not have a main effect on the accuracy *[t(30) = -0.917, p = .340]* RT *[t(30) = -1.624, p = .115]* and confidence *[t(30) = -0.981, p = .335]*.

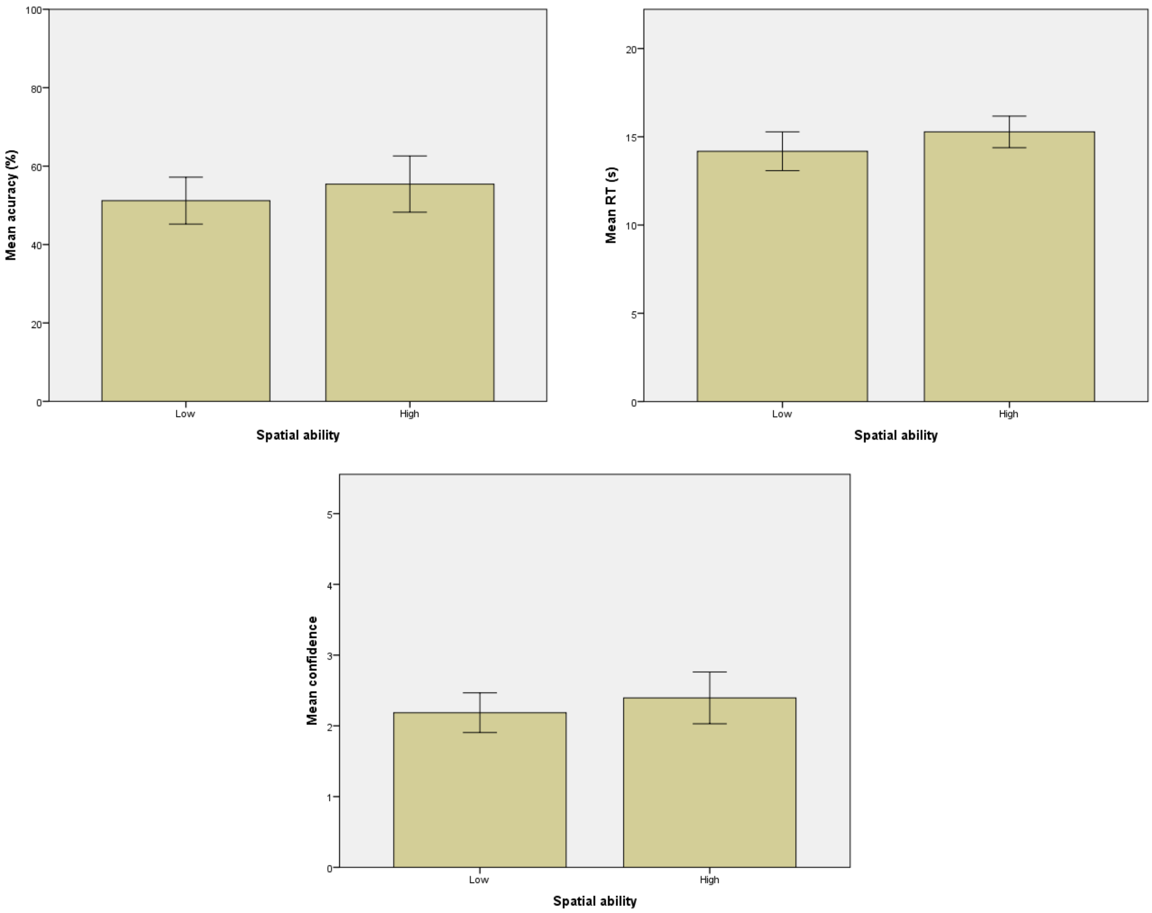


Figure 25. Upper left: Mean total accuracy (%) and spatial ability. Upper right: Mean RT (s) and spatial ability. Lower centre: Mean confidence and spatial ability. Error bars: CI (95%)

Main effect – map type

To assess whether one specific map type improves the results of low and high spatial participants, the total mean accuracy, RT and confidence is analysed for each map type. In Figure 26 (left) the mean accuracy for map types *blank*, *street* and *label* (pillar: left to right) is lower for low spatial than for high spatial participants. Only the mean accuracy for the *hybrid* map (furthest right) is higher for low spatial participants than for high spatial participants. Nevertheless a bivariate correlation analysis showed no effect of spatial ability and the *hybrid* map type. Therefore the result in Figure 26 is not representative.

The two remaining variables: RT (centre) and confidence (right) follow the same pattern, as displayed in Figure 26. That basically means that the mean RT and confidence is higher for high spatial than for low spatial participants. Again, a bivariate correlation analysis for the map types and the total mean accuracy, RT, and confidence was conducted. For the total mean accuracy, no map type indicates any correlation with the spatial ability.

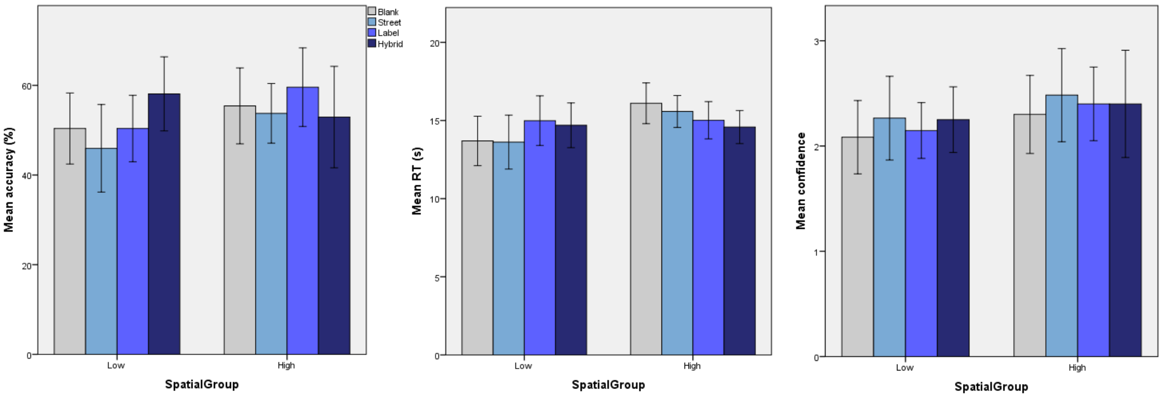


Figure 26. Left: Mean accuracy (%) and spatial ability per map type. Centre: Mean RT (s) and spatial ability per map type. Right: Mean confidence and spatial ability per map type. Error bars: (95%)

Only for RT a correlation of spatial ability to a map type was revealed. The scatter plot in Figure 27 shows that spatial ability correlates to the mean RT for *blank* map types *($r = .338$, $p = .038$, $N = 38$).* According to Cohen (1992), a correlation coefficient of $r = .338$ indicates that there is a medium effect between the two variables. So, the high spatial participants generally use more time to solve tasks with the *blank* map type. Beside that, the mean RT of the three remaining map types (*street*, *label* and *hybrid*) does not correlate with the spatial ability of the participants.

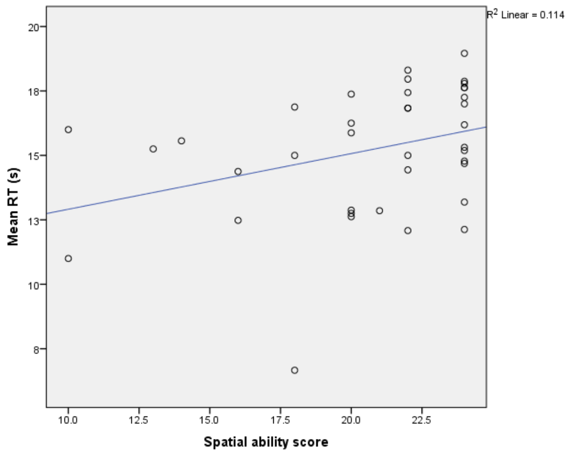


Figure 27. Mean RT (s) and spatial ability for *blank* map type. Linear fit (blue line)

<u>*Main effect – map and task type*</u>

In this section, the effect of spatial ability on the map type for each task type is discussed. For *Task 1*, no significant effect on accuracy for the map types is revealed between the low and high spatial participants. In Table 3 the mean accuracy for all four map types and both spatial groups (low, high) is illustrated. The independent-samples T test reveals no significant effect for accuracy *($p > .05$)* between the two spatial groups. The lowest mean accuracy is registered for the *street* map type for the low *($M = 0.485$, $SD = 0.241$)* and high spatial participants *($M = 0.4$, $SD = 0.325$)*. That reflects the findings in Figure 22 with the global mean accuracy for Task 1 that participants generally perform worse with the *street* map type than the three remaining map types. However, it is interesting that the mean accuracy for the *street* map type is lower for the high than the low spatial group (see Table 3). Nevertheless, a bivariate corrrelation analysis shows no correlation between spatial ability and mean accuracy for every map type *($p > .05$)*. That is why the use of a median split is problematic and therefore an additional measure is useful.

Table 3. Mean accuracy (%) and spatial ability for Task 1.

Task 1

	SpatialGroup	N	Mean	Std. Deviation	Std. Error Mean
Blank	Low	17	.5294	.26343	.06389
	High	15	.6833	.31997	.08262
Street	Low	17	.4853	.24159	.05859
	High	15	.4000	.32459	.08381
Label	Low	17	.6324	.25183	.06108
	High	15	.7333	.25820	.06667
Hybrid	Low	17	.6029	.26603	.06452
	High	15	.5500	.21547	.05563

When comparing the results for RT of the two spatial groups, a significant effect for the *street* map type is detected *($t(27) = -2.403$, $p = .023$)*. The low spatial participants need less *($p < .05$)* time *($M = 10.938$, $SD = 3.079$, $N = 16$)* for solving tasks with the *street* map type than the high spatial participants *(M = 13.583, SD = 2.778, N = 13)*. The bivariate correlation analysis shows that there is a medium correlation effect *(r = .405, p = .016, N = 35)* between the RT and the *street* map type (Cohen, 1992). The higher the spatial ability, the more time the participants need with the *street* map type in Task 1 [Figure 28]. Regarding confidence, a bivariate correlation analysis indicates that there is a medium effect *(r = .362, p = .022, N = 40)* between spatial abil-

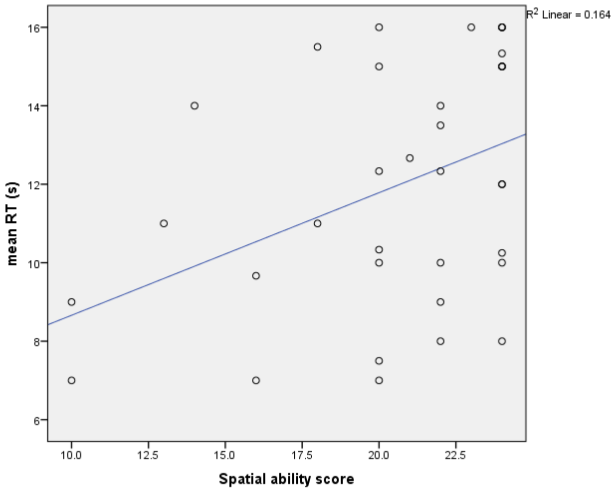


Figure 28. Mean RT (s) and spatial ability for *street* map type for Task 1. Linear fit (blue line)

ity and confidence for the *label* map type (Cohen 1992). That means that high spatial participants use more time to answer questions with the label map type in Task 1 than the low spatial participants. *Task 2* and *Task 3* results for mean accuracy, RT and confidence show no significant effect (*p > .05)* either for low and high spatial participants and will therefore not be addressed in more detail.

Results for *Task 4* show that there is a significant difference for mean RT for the *blank* map type (*t (22)* = -*2.188, p < .05*). For the *blank* map type [Figure 29] the high spatial participants need more *(p < .05)* time (*M = 12.271, SD = 2.942, N = 12*) than the low spatial participants *(M = 14.375, SD = 1.565, N = 12)*. For the three remaining map types (*street*, *label* and *hybrid*) show no significant effect on the dependent variables.

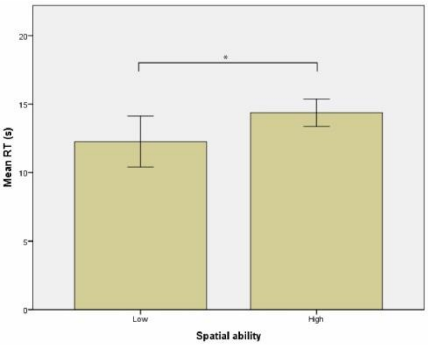


Figure 29. Mean RT (s) and spatial ability for *blank* map type for Task 4. Error bars: CI (95%)

<u>*Summary of experiment results*</u>

The overall results indicate that the four map types do not have an effect on accuracy, RT and confidence. Though, a detailed analysis of every task type reveals that in Task 1, participants perform worse with the *street* map type than with the three remaining map types: *blank*, *label* and *hybrid* [Figure 22]. Regarding RT and confidence, no significant effect is found in Task 1, where participants were asked to study the whole part of a satellite map.

In Task 3, participants feel more confident with the *street* map type than with the other three map types [Figure 24]. Those two results represent the only significant findings on every map type. In Task 3, participants were asked to study the surrounding area of a blue point.

In Table 4 the mean accuracy of all map types is rated for all tasks. The map types on top represent the highest mean accuracy and mean accuracy is decreasing from top to bottom.

The table gives a sense of how participants performed with the four different map types. As one can see, participants score the highest mean accuracy for the *hybrid* map type in three out of four tasks.

Regarding spatial ability, an overall analysis reveals that high spatial participants need more time to solve tasks with the *blank* map type than the low spatial participants. So, spatial ability has a correlation with the mean RT for *blank* map types.

For all tasks, no correlation of spatial ability with accuracy is revealed. But the results indicate that for Task 1 low spatial participants need significantly less time to solve tasks which included *street* map types [Figure 28]. That is interesting though because for the *street* map type the mean accuracy of high spatial participants is lower than for low spatial participants. Additionally, in Task 1 high spatial participants feel more confident with the *label* map type. In Task 4, results indicate that low spatial participants need less time for solving tasks with *blank* map types [Figure 29]. The mean accuracy for this map types is slightly higher for high spatial participants. In Task 4, participants were asked to study the surrounding area of two blue points and the system then chose one of the points for visual memory question.

Table 4. Overview of map type results (mean accuracy)

Accuracy (%)	*Task 1*	*Task 2*	*Task 3*	*Task 4*
High	Hybrid	Hybrid	Street	Hybrid
	Label	Blank	Hybrid	Street
	Blank	Label	Label	Label
Low	Street	Street	Blank	Blank

4.2.4 Eye Movement Data

During the experiment, all gaze data was recorded. This data will now be viewed in more detail. The idea is to give an overview of the gaze data for the different map types. As there are loads of data available from the eye-tracker, the focus is mainly set on the results of Task 1 and Task 3 because in these, significant effects are revealed. The goal is that results of the gaze data analysis will support the findings about accuracy, RT and confidence. The gaze density maps are all of the same design and the range is between low (blue colour) and high values (white colour).

__Task 1 – gaze density__

In Task 1, participants perform worse (mean accuracy) with street map types ($p < .05$) than with the three remaining map types [see 4.2.3]. In Figure 30 the gaze density of one specific image of the *street* map type is illustrated. The image is chosen because participants had the most difficulties with this specific image when comparing it with the other map types. The gaze density for the *street* map type indicates that there are three main fixation points where participants looked at the most. The white rectangular box is the area which represents the solution for this specific map memory task. This area is called Area of Interest (AOI). The three main fixation points (white colour) are situated mainly on the border of the AOI.

Figure 30. Gaze density map for *street* map type for Task 1. Highlighted AOI (white)

The gaze density map for the *hybrid* map type in Figure 31 shows that participants mainly focused on all visible labels. This results in a very point-based gaze density map. The visual comparison of the two maps indicates that for the *street* map type participants fix mainly three points whereas for the *hybrid* map type many other points are fixed.

Figure 31. Gaze density map for *hybrid* map type for Task 1. Highlighted AOI (white)

In Figure 32 the mean total AOI fixation duration is shown. A paired-samples T test reveals that participants' mean fixation duration within the AOI is higher for the *hybrid* ($t(36) = -2.034$, $p = .049$) than for the *street* map type.

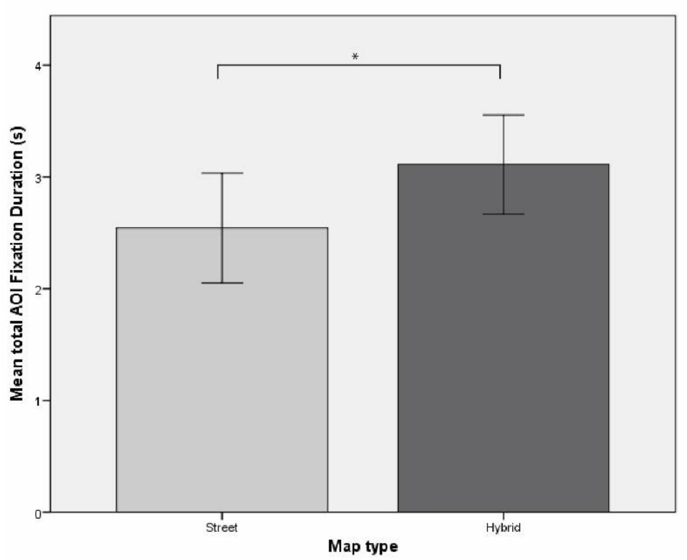


Figure 32. Mean total AOI fixation duration for *street/hybrid* map type for Task 1. Error bars: CI (95%)

The gaze density maps of two the map types, *blank* (left) and *label* (right), are illustrated in Figure 33. The *blank* map type on the left shows that participants basically screen the whole map. There are no clear point fixations visible. In the left map the maximum fixation count is at 20.29 whereas for the map on the right it is at 31.50. That means that for the *blank* map type participants fixated less in absolute time. For the *label* map type participants fixated strongly on some points with a higher absolute count, which can be seen right in Figure 33. This gaze density map is similar to the one of the *hybrid* map type in Figure 31.

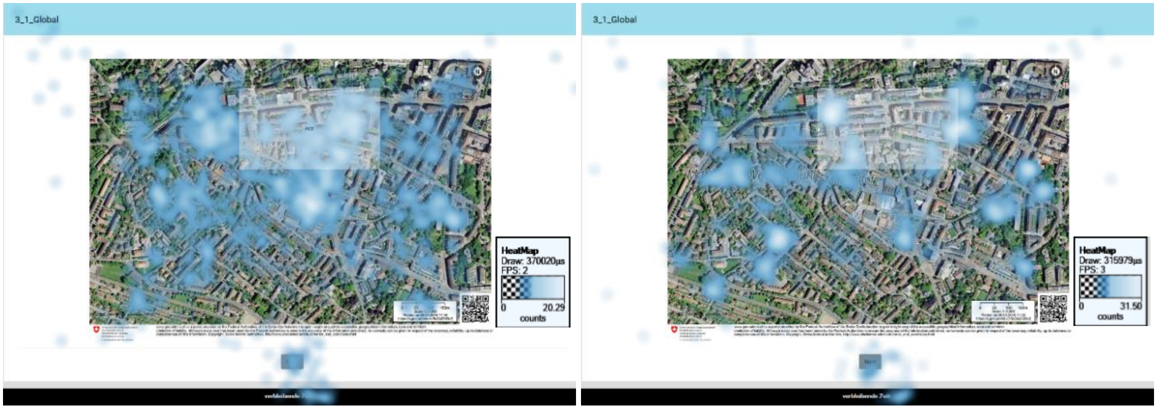


Figure 33. Left: Gaze density map for *blank* map type for Task 1. Highlighted AOI (white) Right: Gaze density map for *label* map type for Task 1. Highlighted AOI (white)

Task 3 – gaze density

The results in section 4.2.3 shows that in Task 3 participants feel more confident with the *street* map type than with the *blank* and *hybrid* map type *($p < .05$)*. For that reason, the gaze density maps are highlithed in this subsection to emphasize the results of the user experiment. As a reminder, participants were asked in this task to focus on one point and its surrounding area. For the chosen image the mean accuracy for the *street* map type was higher than for the *blank* and *hybrid* map type. In Figure 34 the gaze density for the *street* map type is illustrated. The red rectangle represents the AOI for this task.

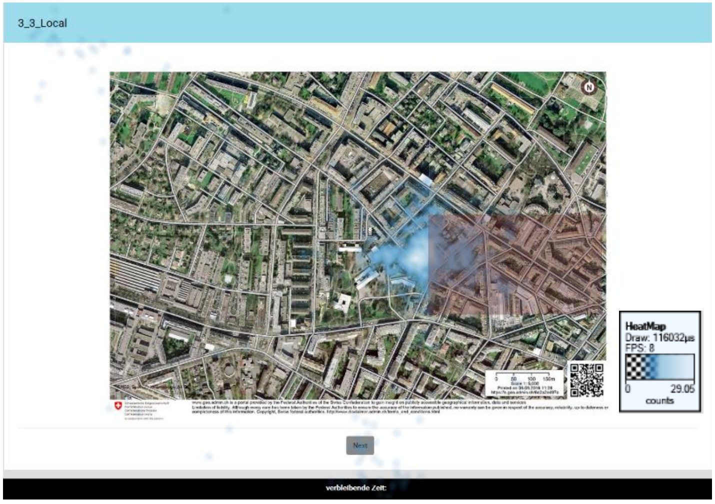


Figure 34. Gaze density map for *street* map type for Task 3. Highlighted AOI (red)

The comparison of the gaze density map in Figure 34 and of the *blank* (left) and *hybrid* (right) map type in Figure 35 does not give a clear indication why participants feel more confident with the *street* map type. The visual comparison only shows that the gaze density map for the *hybrid* map follows the labels that are visible on the map. This can be seen best in the right map of Figure 35 were main fixation points are measured above the point (fixation line above the blue point).

Figure 35. Left: Gaze density map for *blank* map type for Task 3. Right: Gaze density map for *hybrid* map type for Task 3. Highlighted AOI (red)

A deeper analysis between the three map types though, indicates and supports different levels of confidence within this task. In Figure 36 the bar chart on the left shows the mean total AOI fixation duration whereas the bar chart on the right illustrates the mean total AOI visit duration. The

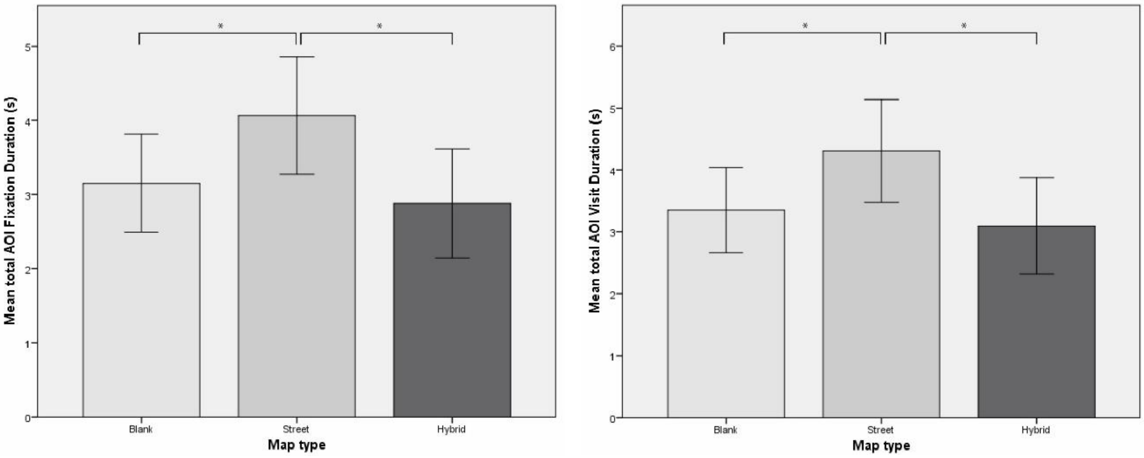


Figure 36. Left: Mean total AOI Fixation Duration (s) for Task 3. Right: Mean total AOI Visit Duration (s) for Task 3. Error bars: CI (95%)

AOI fixation duration is basically the total time that participants fixated on objects whereas the AOI visit duration is the time from when participants start viewing the AOI until they leave it. When a participant' eye movements are very fast they will not be recorded as fixations. Therefore, the mean AOI visit duration is slightly higher than the mean AOI fixation duration. In the result of the left diagram, a pairwise comparison shows reveals that the mean total AOI fixation duration for the *street* map type *(M = 4.065, SD = 2.374)* is higher than for the *blank (M = 3.153, SD = 1.987)* and *hybrid* map type (*M = 2.882, SD = 2.205)* A repeated-measures ANOVA reveals that the map type does have an effect on the mean total AOI fixation duration $[F(2,72) = 7.975, p = .001, \eta_p^2 = .181]$.

In the results of the right diagram, the mean total AOI visit duration indicates the same trend. A repeated-measures ANOVA results in significant effects between the three map types $[F(2,72) = 7.471, p = .001, \eta_p^2 = .172]$. The mean AOI visit duration is longer for the *street* map type *(M = 4.308, SD = 2.491)* than for the *blank (M = 3.351, SD = 2.073, p < .05)* and *hybrid* map type *(M = 3.094, SD = 2.335, p < .05)*.

<u>*Summary of eye movement data results*</u>

The eye movement data was meant to support the findings in the previous section of the experimental results (see 4.2.3). Therefore, the eye movement data was only analysed for two main findings from the user experiment.

First, the gaze density maps for a specific image in Task 1 highlight the different fixating behaviours depending on the map type. Map types that include labels result in a point fixed gaze density map. Participants basically look at all the labels in more detail. On the contrary, the blank map type results in a widespread gaze density map. The gaze density map of the *street* map type makes participants fix points but less than on the *label* and *hybrid* map type. For the *street* map type participants view less time in the AOI than for the *hybrid* map type ($p < .05$).

Second, the analysis of Task 3 indicates that participants fixate and visit the AOI in the *street* map type longer than in the *blank* and *hybrid* map type ($p < .05$) in Task 1. The gaze density maps highlight that labels lead the participants' eye fixations as it is the case with the results in Task 1.

5 Discussion

In this chapter, the main findings will be discussed in more detail and compared to relevant literature. Additionally, three stated research questions at the beginning of the thesis will be answered and finally, the limitations and future research will be addressed.

5.1 Research Question 1: Measuring Visual Complexity

A comparative investigation of the visual complexity of satellite maps with different image processing algorithms is one of the core contributions of this thesis. The computational analysis conducted in the scope of this thesis gives an impression of how a computer classifies images into different levels of visual complexity, and whether different algorithms agree on what is visually complex. For this reason, the following research question was defined:

<u>RQ 1:</u> *How is visual complexity currently measured, and do existing methods work well to assess the visual complexity of geographic imagery?*

- *Specifically, do (selected) visual complexity algorithms agree on the visual complexity of satellite maps of urban and rural areas and different map types?*

Visual complexity of natural scenes (images and photography as an input) has been studied extensively in vision research (Mack & Oliva 2004; Bravo & Farid 2008). On the other hand, only few studies have examined the measurement of visual complexity for geographic imagery and generalized geographical maps (e.g., Fairbairn 2006; Rosenholtz et al. 2007; Touya et al. 2015). Based on these previous studies, we hypothesized that existing image processing algorithms can be applied to realistic geographic visualizations as well, and thus would predict their visual complexity successfully; and that these algorithms would rank the complexity of examined images similarly.

The computational analysis performed in this thesis with the four image processing algorithms on urban and rural maps reveals that urban maps are visually more complex than rural maps (see chapter 4.1), and all four algorithms agree on this point. Based on Rosenholtz et al. (2007)'s visual clutter algorithm, the difference between urban and rural maps for FC is at 38.7%, whereas it is 12.9% for SE. However, note for rural maps, the SE score is higher than the FC score. One might have expected the score for FC to be higher than for SE, as FC measures more features within an image (Rosenholtz et al. 2007). An explanation might be that in terms of encoding and redundancy (measured by SE), urban and rural maps do not differ that much, as much as, for example, colour variability and luminance (measured by FC). This might be a reason for the increase of SE values for rural maps. Thus, the relative difference as measured by SE between rural and urban maps is smaller than differences measured by FC.

Results for image segmentation and edge density (Felzenszwalb & Huttenlocher 2004; Mack & Oliva 2004) follow the same trend as for FC and SE. Image segmentation is fairly high for urban maps compared to rural maps. Urban maps with dense housing areas, street infrastructure and small parks result in a very heterogeneous and thus much segmented space. Therefore, measurements for both edge density and image segmentation are high for urban maps. On the other hand, rural maps generally consist of large homogeneous areas such as water, grass or rocks. These features do have much finer-scaled and low coarse-scale structures (Bravo & Farid 2008). Since image segmentation detects and segments along similar characteristics, a homogeneous space results in a lower number of regions. As a consequence of this, the image segmentation measurement for rural maps is lower.

The analysis of the four different map types for urban scenes indicates that more enhanced maps are visually more complex. For FC and SE, all map types are significantly different from each other (Figure 18). However, as there are pronounced differences within the map types for FC, the results for all map types for SE and edge density follow a flat line trend. Rosenholtz et al. (2007) state that FC explicitly measures colour variability, orientation and luminance whereas SE measures colour variability (implicitly) and orientation. The visually enhanced map types do mainly contain coloured labels and a street network layer. Therefore, the SE measurement could remain on a stable level due to the fact that colour variability and luminance are not taken into account.

The edge density measurement detects white and black pixels in a black and white image. The major changes between map types such as coloured labels and street network layer are vector features and do consequently not affect the black and white images. Thus, the edge density measure remains on a relatively stable level. The results for image segmentation algorithm follow a similar tendency as in the measurements for FC, SE and edge density. All three enhanced maps (*label*, *street* and *hybrid*) result in higher numbers of regions. The more information there is, the more regions are detected by the algorithm. The overlay of information causes a more segmented space and therefore an increased number of regions, which symbolizes a high visual clutter.
Mack & Oliva (2004) found indications that complexity is dependent on following features: quantity and variety of objects, detail, colour and symmetry. A few years later Rosenholtz et al. (2007) discovered a correlation of visual clutter and the quantity and variety of objects. Consequently, measurements of visual clutter can be linked to visual complexity. Thus, the use of satellite maps as input data for image processing algorithms is not a problem at all. The results indicate that

enhanced map types (*street*, *label* and *hybrid*) do increase visual clutter for all map algorithms. Therefore, and as an answer to the first question, image processing algorithms can be used in order to predict visual complexity for satellite maps.

5.2 Research Question 2: Visual Complexity and Memory

The second research question builds on the results of the first research question. Without a useful measurement of visual complexity, the following research question cannot be answered:

<u>RQ 2:</u> *Does the user performance decrease with visuospatial memory tasks as the level of visual complexity increases?*
- Specifically, how will participants perform in visuospatial memory tasks with visually more complex geographic visualizations, such as hybrid maps (cartographically-enhanced satellite images), in comparison to regular satellite images?

Rosenholtz et al. (2007)'s findings indicate that measures of visual clutter do correlate with visual search performance. The more visually complex a display is, the longer it will take participants to solve a VST. Bravo & Farid (2008)'s research has also shown that high visual clutter results in higher RT in visual search tasks. Based on these findings, we predicted that visually more complex displays do affect task and search performance, such as accuracy, RT and confidence.

Regarding accuracy, the results from this study only show a significant effect in Task 1, where participants perform significantly worse with the *street* map type than with the *blank*, *label* and *hybrid* map types. In other words, the visually most complex map (*hybrid*), as identified by the visual complexity measures, is not the intellectually most complex in a visuospatial memory task. *Hybrid* map types contain labels, which might have simplified the process of understanding and interpreting a map and consequently result in a lower intellectual complexity measure (MacEachren 1982). Therefore, the visual complexity might not have affected the participants' visuospatial memory that much. Interestingly, the examination of eye movement data shows that the mean AOI fixation duration is longer with the visually most complex map (*hybrid* map) than the *street* map type (Figure 36), as well as others. Our observation regarding eye movements confirm Hegarty et al. (2012)'s findings in which they show that complex maps result in longer eye fixations and greater number of fixations. According to the computational measurements the *hybrid* map type is visually more complex than the *street* map type. While the accuracy and RT in Task 1 differ from the visual complexity assessed by the image processing algorithms, the eye movement data supports the previous findings.

Furthermore, for Task 1, the comparison between *street-label* and *street-blank* map types does not show a significant effect in fixation duration. In order words, the fixation duration is not different when comparing two visually different complex map types (*street-label* and *street-blank*). Task 1 required participants to scan the whole map. However, the gaze density map for the *street* map type shows that participants mainly fixate to a limited number of points. When it comes to the *blank* map type, participants basically scan locations on the whole map equally often. This can be the reason why the participants in Task 1 have a lower accuracy with the *street* map than the *blank* map. Participants do focus on fewer points and consequently the chance of holding important features is lower. Since there are no significant results from the gaze data, this remains a proposition that further research could verify. The fact that the *street* map type is visually more cluttered than the *blank* map type could be a reason for the decreased performance and accuracy for the *street* map type (Rosenholtz et al. 2007; Bravo & Farid 2008).

When comparing *blank-street* map types, the question addresses why the *street* map type makes participants focus on a restricted number of points or on features. Probably, different attributes guide the participants' attention towards specific features on a map. Wolfe & Horowitz (2004) list colour, motion, orientation and size as examples of attributes that undoubtedly guide visual attention. Therefore, the greyish colour and changes in shape/orientation of the street network might have attracted or at least influenced the viewers' attention in this task. This analysis is based on gaze density data of one specific image in Task 1. Therefore, the finding has to be carefully addressed because it does not represent all stimuli data. As an example, the gaze density map for a specific street map location (Figure 30) in section 4.2.4 indicates that participants focused on a small green park amongst others. The path in the park is circular and highlighted by the street network layer whereas the rest of the street network in the map is mostly angular. According to Wolfe & Horowitz (2004) the search is easier when the target/distractor (TD) difference is larger. This finding supports Wolfe et al. (1992)'s claim that the search for a target is easier with homogeneous than with heterogeneous distractors. Thus, the reason why participants focused on the path in the park (target) might be because its shape clearly differs from the rest (distractor) of the street network but as the park is not a part of the AOI, participants mainly focused on parts of the map that are not of great importance. The actual AOI is located in the centre of the heterogeneous street network and therefore makes it quite difficult for the participants to perform well with the *street* map type.

Another interesting result from Task 1 is the difference in accuracy for the map pairs, *street-label* and *street-hybrid*. We observe higher recall accuracy scores for both *label* and *hybrid* map types than for the *street* map. *Label* and *hybrid* maps both have labels in them, and as text (i.e., labels) contains semantic information, they will be processed differently than visual features. This effect can be seen in the gaze density maps, and thereby indicate that participants mainly focus on all visible labels. According to Carroll et al. (1992), it is known that people tend to view textual and non-textual information in an isolated way by focusing on textual information first, before actually viewing and interpreting the background picture (Hegarty 1992a; Hegarty 1992b). Additionally, a study by Rayner et al. (2001) confirms that people focus more on text than pictorial information. Hence, text and visual information are processed differently, and this influences the results in this study by helping people remember better when there is verbal information as well as visuospatial (Mayer & Moreno 2010).

Edler et al. (2014) recently investigated the effect of complexity for object location memory in topographic maps. The results show that road maps cause weaker recall performance than maps with additional point and area features. Labels in the *label* and *hybrid* map type can be regarded as additional features. This might support the findings that the performance is lower for the street map type (road map) than the *label* and *hybrid* map type with additional features. However, since the study was conducted with topographic maps the finding has to be carefully interpreted and eventually further verified.

In another study, Wolfe & Horowitz (2004) show that colours are undoubtedly attributes that guide attention. Consequently, the attention on the coloured labels in the user experiment of this thesis is even more pronounced. In addition to this, people are able to store both colour and orientation characteristics of four objects according to Luck & Vogel (1997) but as there were more than four labels shown on a map in the experiment of this thesis, participants were not able to retain all of them. However, a trend is visible in Table 4 that semantic information seems to have helped and therefore influenced the participants' performance with the different map types.

There is especially one very interesting point from the results regarding the confidence level of the participants. In Task 3, participants feel more confident with the *street* than the *blank* and *hybrid* map type. This was rather surprising because participants actually perform worst with the *street* map type in Task 1 but as the task objective was completely different in Task 1 and 3 it would be difficult to investigate in these findings. Nevertheless, it is interesting to consider why and how the task objective might have affected the participants' performance and confidence.

In Task 3, the task-relevant area is limited because participants are asked to focus on the surrounding area of a blue point. Thus, participants already know where the AOI might be located. This might also be the reason why the street network layer has not guided the participants' visual attention as much as in Task 1 where the AOI could have been anywhere. It might even have helped some participants to visualize and memorize the surrounding area with the help of memorizing nodes and lines of the existing network. As an example, participants might have retained four anchor points and their spatial relation to the point (e.g., blue point on the map) (Luck & Vogel 1997).

For the *blank* map type, such anchor points are more difficult to locate because there is no highlighting feature such as the street network available. This could be a possible explanation for the lower confidence with *blank* map types compared to *street* map types in Task 3. The comparison of the *street* and *hybrid* map type indicates that the inclusion of labels on the street network harms the confidence level. Labels and their attributes clearly attract the participants' attention, which is visible in the gaze density maps for the *hybrid* map type (Carroll et al. 1992; Hegarty 1992a; Hegarty 1992b; Rayner et al. 2001; Wolfe & Horowitz 2004). Participants actually visited and fixated the AOI less for the *hybrid*/*blank* map than for the *street* map type. The fixation of labels is not a negative influence per se, but as the location of labels changes in the recognizing section, the participants might have gotten insecure.

As people are able to hold about four objects only in their working memory at a time, the inclusion of labels on top of the street network potentially results in an increased cognitive load (Duncan 1984; Luck & Vogel 1997; Cowan 2001). Thus, participants might have felt overstrained and insecure in their decision-making process. On the other hand, the mean confidence in Task 3 for the *label* map type is slightly higher than for the *hybrid* map type. This does not contradict the suggestion that the amount of information leads to cognitive load, i.e., there is more information (as measured by the visual complexity algorithms) in the *hybrid* map type (including the labels) than the *label* map type. Hence the participants' cognitive load increases to retain more detail with the *hybrid* map.

The existence of the street layer network affects the participants' performance and behaviour in both negative and positive ways. In Task 1, the street network layer clearly guides the participants' attention by highlighting some features more than others. This makes participants look on a restricted number of features, which might not be part of the AOI. In this case, the participants' eye movements could be misled and guided towards non-important features. In Task 3, the street network layer alone might have helped the participants because the location or the names of the labels changed in the recognition section. So, the focus on a fixed street network itself might in this case have increased the confidence and somehow affected the accuracy for Task 3 and Task 4 where the street and hybrid map type reached the highest accuracy (see Table 4).

Some of the results indicate a slight tendency in the direction that the more visually cluttered a map is, the higher the mean accuracy is. However, according to Hegarty et al. (2009) map design-ers should be careful not to add too much information, which might be extraneous or task-irrelevant. The initial expectation was that the more visually cluttered a map is, the lower the mean accuracy is, but results now show a tendency of the opposite. Labels and their semantic information might have helped a lot even though the location and name were changed in the tasks. Unfortunately, the research question cannot be clearly answered because there are no statistically significant differences between the different map types in this case. These initial observations, nonetheless, pave the way for future research.

5.3 Research Question 3: Spatial Ability and Memory

The distinction between low and high spatial participants is based on the results of the map memory task published by Ekstrom et al. (1976). The following research question connects the results of the map memory task with the user experiment:

RQ 3: *What influence does the spatial ability have on visuospatial memory tasks with different level of visual complexity?*
- Specifically, do high spatial participants perform differently than low spatial participants with map-based visuospatial memory tasks?

Overall, the mean accuracy for high spatial participants is about 4-5 % higher but not significantly different from low spatial participants. According to Miyake et al. (2001) tasks containing working memory have been linked to spatial ability. A significant effect of spatial ability to the overall accuracy in this experiment was however not found, as the range of accuracy of the results is too large across the whole experiment.
According to Miyake et al. (2001), low spatial participants generally take more time solving working memory tasks but the overall results for RT indicate that for the *blank* map type low spatial participants take less time than high spatial participants. A reason for this could be that low spatial participants only focus on landmarks whereas high spatial participants focus on landmarks and other characteristics within the image and therefore need more time (Pazzaglia & De Beni 2006).

For Task 1, low spatial participants need less time to solve tasks with *street* map types than high spatial participants, which is not consistent with the findings in research of Miyake et al. (2001), who claim that low spatial participants generally require more time. Here, the mean accuracy for low spatial participants was 8 % higher than for high spatial participants. Overall, low spatial participants take less time to solve the task and perform slightly better than the high spatial participants in this experiment. Due to the small sample size of 10 participants it is rather difficult to come to any conclusions as the statistical power is simply too low. In addition to this, it is also relevant to ask whether it is meaningful to consider a difference of 2 seconds when it comes to memorizing maps.

Additionally, a correlation with a medium effect between spatial ability and confidence for *label* map types is revealed in Task 1 (see 4.2.3). So the higher the spatial ability, the more confident a participant gets with the *label* map type. The fact that labels changed their position in the recognizing section might have been more confusing for the low spatial participants than for their counterpart. As there is no street network layer visible on this map type, low spatial participants might have struggled to locate labels as landmarks, due to the changed location of the labels. On the contrary, high spatial participants might have benefited from not only looking at landmarks but also other visual objects and anchor points, as this is also mentioned by Pazzaglia & De Beni (2006).

For Task 4, the low and high spatial participants require a different amount of time to solve the task with *blank* map types. Again, the RT for low spatial participants is less than for high spatial participants. This result follows the overall trend that low spatial participants need less time with *blank* map types than the high spatial participants averaged over the whole experiment. However, a bivariate correlation analysis for Task 4 does not reveal any correlation between spatial ability and RT for *blank* map types. Therefore, the significant result from the median split for the low and high spatial group has to be carefully interpreted because it represents a small sample size.

Spatial ability does influence the participants’ performance regarding RT and confidence on a smaller scale. However, no overall effects regarding spatial ability and user performance were discovered. As there are only minor significant results with spatial ability and user performance, the research question cannot be clearly answered but a trend shows that in some cases high spatial participant need longer to give an answer than low spatial participants.

5.4 Limitations

In this user study no overall effects for visual complexity and task performance were found. This is not very surprising because the design of every task is slightly different. Nevertheless, a separation into two local and global tasks was interesting and revealed remarkable findings regarding the street network layer. The range of answers in this study was very large. For all tasks and map types accuracy results ranged from 10 to 100 %. Therefore, it was extremely difficult to obtain any significant results. A higher number of both geographers and non-geographers might have made space to investigate more in the effect of geographic knowledge in map reading tasks, but as there were only about 20 geographers the statistical power was too low.

The main insight from this study was that the tasks were very difficult which is reflected in an overall mean accuracy between 50 and 60% and in some of the participants' comments at the end of the experiment. One possibility to improve this would be to increase the time limit for every task or to simplify the tasks. For that, the number of tasks would have to be reduced in order to keep a maximum study duration of 60 minutes or less. Fewer tasks might result in more consistent results because participants have more time to answer and they will be presented to fewer different tasks in total, which might help them to stay focused. However, it is not clear whether more time would automatically result in higher accuracy, because participants might get more disturbed e.g., by the change of location of the labels. The adjustment of the time limitation could also result in more pronounced results of RT because the time range will be higher within the participants.

In order to better compare the overall results the same task design could be used in the whole experiment. This would require more maps to be generated in order to get enough experimental results from different map types. Still, two or three of the same tasks would increase participants' exhaustion and simultaneously decrease their motivation during the experiment. Therefore, an improved balance between similar and different task objectives is of great importance. In terms of map design, it could be worth a try to reduce the visual saliency of the labels and street network layer and report the results. The characteristics of guiding attributes such as colour and luminance (Wolfe & Horowitz 2004) could be slightly changed in order to make the visual attention less pronounced and thus prevent participants from focusing too much on labels.
Overall, the results of the map memory test by Ekstrom et al. (1976) did not have a great effect on the experimental results of this experiment. The initial map memory test and the actual experiment seemed to be too similar. Therefore, it might be relevant to consider another test like e.g. Vandenberg Mental Rotation Test (MRT) suggested by Vandenberg & Kuse (1978) as a baseline for spatial ability.

6 Conclusion and Future Research

This thesis investigated the effect of visual complexity in realistic geographic visualizations such as satellite maps on visual map memory performance. Four different working memory tasks (local and global) were designed and a restricted time limit was set for every single task.

The results indicate that current measures of visual clutter (Mack & Oliva 2004; Rosenholtz et al. 2007; Bravo & Farid 2008) are good indicators for visual complexity in geographic displays e.g., satellite maps. First, a computational analysis with four different image algorithms shows that an urban scene is visually more complex than a rural scene. Consequently, only the visually more complex urban scene type is used in order to design and measure the four following map types: *blank*, *street*, *label* and *hybrid*. Again, the image processing algorithms indicate that visually enhanced satellite maps are visually more cluttered and can therefore be used a baseline for map memory tasks. So generally, geographic visualizations can be used in large sums in order to predict visual complexity. The image algorithms are very efficient what accelerates the computational measurement.
Overall results of the user experiment do not show significant results between the four different map types. So the main research question about whether higher visual complexity decreases the participants' performance could unfortunately not be answered. However, there is a tendency towards that visually more enhanced satellite maps do result in higher map accuracy, RT and confidence. The fact that visually more complex displays contained labels amongst others supports the findings in literature, suggests that participants read text and visual features isolated and that the focus is more on textual than pictorial information (Carroll et al. 1992; Rayner et al. 2001). The labels are visually noticeable because of the red and black colour and the white outline (Wolfe & Horowitz 2004). Depending on the task design labels can also harm the map performance, as the results of Task 3 show, where the mean accuracy of map types with labels is slightly lower than with the *street* map type. It is very important to carefully analyse the results and come up with final awareness because some of the map types contain labels and others not.

Spatial ability does not have a great effect on the results of the user experiment because of various reasons. The map memory test by Ekstrom et al. (1976) seems to be too similar to the main experiment design. Another test such as MRT by Vandenberg & Kuse (1978) could be interesting to use as a baseline for further map memory tasks. The use of eye-movement data in displays with different visual complexity levels reveals that in some cases more complex displays result in longer eye fixations also suggested by Hegarty et al. (2012) Therefore, different levels of visual complexity seem to have an effect on the eye movement behaviour.

6. Conclusion and Future Research

Realistic geographic visualizations such as satellite maps with different levels of visual complexity might have an effect on the participants' performance. This thesis highlights that depending on the map task and the participants' ability more information on a map can be either helpful or harmful. Therefore, the design of enhanced satellite maps is very important and influential on map-readers' performance.

Future work should include more research in image algorithms in order to retrieve not only measures of global visual complexity but also local visual complexity within an image. Thus, local areas of high visual complexity could be revealed and used in user experiments for further analysis. The quad tree clutter measure by Jégou & Deblonde (2012) is one example of a measurement of local clutter, but the current output should be refined. Additionally, more studies that relate visual complexity in satellite maps to visual working memory should be addressed and verified with more user experiments.

Bibliography

Alvarez, G.A. & Cavanagh, P., 2004. The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, 15(2), pp.106–111.

Banhart, C., 1967. *The american college dictionary* 21st ed., Random House.

Beck, M.R., Lohrenz, M.C. & Trafton, J.G., 2010. Measuring search efficiency in complex visual search tasks: global and local clutter. *Journal of experimental psychology: Applied*, 16(3), pp.238–250.

Bjørke, J.T., 1996. Framework for entropy-based map evaluation. *Cartography and Geographic Information Science*, 23(2), pp.78–95.

Bjørke, J.T., 1997. Map generalisation: an information theoretic approach to feature elimination. In *Proceedings, International Cartographic Conference, (ICA)*. Stockholm, pp. 480–486.

Brassel, K.E. & Weibel, R., 1988. A review and conceptual framework of automated map generalization. *International Journal of Geographical Information Science*, 2(3), pp.229–244.

Bravo, M.J. & Farid, H., 2008. A scale invariant measure of clutter. *Journal of vision*, 8(1)(23), pp.1–9.

Brophy, D.M., 1980. Some reflections on the complexity of maps. In *Technical Papers of ACSM 40th Annual Meeting*. pp. 343–352.

Buttenfield, B.P. & McMaster, R.B., 1991. Map generalization: making rules for knowledge representation. In *Map generalization: making rules for knowledge representation*. Longman, Harlow/Wiley, New York, pp. 150–239.

Carpenter, R.H., 1988. *Movements of the eyes*, London: Pion Limited Movements of the eyes.

Carroll, P.J., Young, J.R. & Guertin, M.S., 1992. Visual analysis of cartoons: A view from the far side. In *Eye movements and visual cognition*. Springer New York, pp. 444–461.

Castner, H., 1990. *Seeking new horizons: a perceptual approach to geographic education*, Montreal: McGill–Queens University Press.

Ciolkosz-Styk, A. & Styk, A., 2013. *Advanced image processing for maps graphical complexity estimation*, Warsaw.

Cohen, J., 1992. A power primer. *Psychological Bulletin*, 112(1), pp.155–159.

Çöltekin, A., Fabrikant, S.I. & Lacayo, M., 2010. Exploring the efficiency of users' visual analytics strategies based on sequence analysis of eye movement recordings. *International Journal of Geographical Information Science*, 24(10), pp.1559–1575.

Cowan, N., 2001. The magical number 4 in short term memory. A reconsideration of storage capacity. *Behavioral and Brain Sciences*, 24(4), pp.87–186.

Donderi, D.C. & McFadden, S., 2005. Compressed file length predicts search time and errors on visual displays. *Displays*, 26(2), pp.71–78.

Duncan, J., 1984. Selective attention and the organization of visual information. *Quarterly Journal of Experimental Psychology*, 113(4), pp.501–517.

Duncan, J. & Humphreys, G.W., 1989. Visual search and stimulus similarity. *Psychological review*, 96(3), pp.433–458.

Edler, D. et al., 2014. Grids in topographic maps reduce distortions in the recall of learned object locations. *PLoS ONE*, 9(5), p.e98148.

Ekstrom, R.B., French, J.W. & Harman, H.H., 1976. Kit of factor-referenced cognitive tests.

Ellis, P.D., 2010. *The essential guide to effect sizes: statistical power, meta-analysis, and the interpretation of research results*, Cambridge University Press.

Eng, H.Y., Chen, D. & Jiang, Y., 2005. Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review*, 12(6), pp.1127–1133.

Fairbairn, D., 2006. Measuring map complexity. *The Cartographic Journal*, 43(3), pp.224–238.

Felzenszwalb, P.F. & Huttenlocher, D.P., 2004. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2), pp.167–181.

Field, A., 2013. *Discovering statistics using IBM SPSS Statistics* 4th ed., SAGE Publications.

Freytag, A., 2014. Felzenszwalb-Segmentation. Available at: https://github.com/cvjena/Felzenszwalb-Segmentation [Accessed October 27, 2016].

Harrie, L. & Stigmar, H., 2010. An evaluation of measures for quantifying map information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(3), pp.266–274.

He, Z., Zhu, G. & Pang, X., 1997. A study of cartographic information theory used in map-making. In *Proceedings of the International Cartographic Conference (ICA)*. pp. 2249–2261.

Heaps, C. & Handel, S., 1999. Similarity and features of natural textures. *Journal of Experimental Psychology: Human Perception and Performance*, 25(2), pp.299–320.

Hegarty, M., 1992a. Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(5), pp.1084–1102.

Hegarty, M., 1992b. The mechanics of comprehension and comprehension of mechanics. In *Eye movements and visual cognition*. Springer New York, pp. 428–443.

Hegarty, M. et al., 2002. Development of a self-report measure of environmental spatial ability. *Intelligence*, 30(5), pp.425–447.

Hegarty, M., Smallman, H.S. & Stull, A.T., 2012. Choosing and using geospatial displays: Effects of design on performance and metacognition. *Journal of Experimental Psychology: Applied*, 18(1), pp.1–17.

Hegarty, M. et al., 2009. Naïve cartography: how intuitions about display configuration can hurt performance. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 44(3), pp.171–186.

Hegarty, M. & Waller, D.A., 2005. Individual differences in spatial ability. In *The Cambride Handbook of Visuospatial Thinking*. pp. 121–169.

Henderson, J.M., 2008. Eye movements and scene memory. In *Visual Memory*. pp. 87–121.

Heylighen, F., 1999. The growth of structural and functional complexity during evolution. In F. Heylighen, J. Bollen, & A. Riegler, eds. *The evolution of complexity*. Dordrecht: Kluwer Academic, pp. 17–44.

Highfield, R. & Coveney, P., 1995. *Frontiers of complexity: the search for order in a chaotic world*, Ballantine.

Hoarau, C., Christophe, S. & Mustière, S., 2013. Mixing, blending, merging or scrambling topographic maps and orthoimagery in geovisualization? In *International Cartographic Conference*. pp. 1–17.

Hollingworth, A. & Henderson, J.M., 2002. Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception & Performance*, 28(1), pp.113–136.

Irwin, D.E. & Andrews, R.W., 1996. Integration and accumulation of information across saccadic eye movements. In *Attention and performance XVI: Information integration in perception and communication*. pp. 125–155.

Irwin, D.E. & Zelinsky, G.J., 2002. Eye movements and scene perception: Memory for things observed. *Perception & psychophysics*, 64(6), pp.882–895.

Itti, L. & Koch, C., 2000. A saliency-based search mechanism for overt and covert hifts of visual attention. *Vision Research*, 40(10), pp.1489–1506.

Itti, L. & Koch, C., 2001. Computational modelling of visual attention. *Nature reviews. Neuroscience*, 2(3), pp.194–203.

Jégou, L. & Deblonde, J.P., 2012. Vers une visualisation de la complexité de l'image cartographique. *Cybergeo: European Journal of Geography*.

Kaplan, I., 2002. Shannon Entropy. Available at: http://www.bearcave.com/misl/misl_tech/wavelets/compression/shannon.html [Accessed February 20, 2017].

Knieser, M.J. et al., 2003. A technique for high ratio LZW compression. In *Proceedings of the conference on Design, Automation and Test in Europe*. p. 10116.

Lawton, C.A., 1994. Gender differences in way-finding strategies: Relationship to spatial ability and spatial anxiety. In *Sex Roles*. pp. 765–779.

Likert, R., 1932. A technique for the measurement of attitudes. *Archives of Psychologgy*, 22(140), p.55.

Loftus, E., Miller, D. & Burns, H., 1978. Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology; Human Learning and Memory*, 4(1), pp.19–31.

Luck, S.J. & Hollingworth, A., 2008. *Visual memory*, New York: Oxford University Press.

Luck, S.J. & Vogel, E.K., 1997. The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), pp.279–281.

MacEachren, A.M., 1982. Map complexity: Comparison and measurement. *Cartography and Geographic Information Science*, 9(1), pp.31–46.

Mack, M.L. & Oliva, A., 2004. Computational estimation of visual complexity. In *12th Annual Object, Perception, Attention, and Memory Conference*.

Martin, D.W., 2008. *Doing psychology experiments* 7th ed., Cengage Learning.

Mayer, R. & Moreno, R., 2010. Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist*, 38(1), pp.43–52.

Miyake, A. et al., 2001. How are visuospatial working memory, executive functioning, and spatial abilities related? *Journal of Experimental Psychology: General*, 130(4), pp.621–640.

Montello, D.R., 2002. Cognitive map-design research in the twentieth century: theoretical and empirical approaches. *Cartography and Geographic Information Science*, 29(3), pp.283–304.

Montello, D.R. et al., 1999. Sex-related differences and similarities in geographic and environmental spatial abilities. *Annals of the Association of American geographers*, 89(3), pp.515–534.

Nelson, W.W. & Loftus, G.R., 1980. The functional visual field during picture viewing. *Journal of experimental psychology. Human learning and memory*, 6(4), pp.391–399.

Oliva, A. et al., 2004. Identifying the perceptual dimensions of visual complexity of scenes. In K. Forbus, D. Gentner, & T. Regier, eds. *Proceedings of the 26th Annual Cognitive Science Society*. Austin, TX: Cognitive Science Society, pp. 1041–1046.

Pazzaglia, F. & De Beni, R., 2006. Are people with high and low mental rotation abilities differently susceptible to the alignment effect? *Perception*, 35(3), pp.369–383.

Rayner, K. et al., 2001. Integrating text and pictorial information: Eye movements when looking at print advertisements. *Journal of experimental psychology. Applied*, 7(3), pp.219–226.

Reitsma, F., 2001. *Spatial complexity*. University of Auckland.

Rensink, R.A., 2002. Change detection. *Annual review of psychology*, 53(1), pp.245–277.

Richardson, A.E., Montello, D.R. & Hegarty, M., 1999. Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory & Cognition*, 27(4), pp.741–750.

Robinson, A.H., 1952. *The look of maps: an examination of cartographic design*, Madison: University of Wisconsin Press.

Rosenholtz, R. et al., 2005. Feature congestion: a measure of display clutter. *Proceedings of the Special Interest Group for Computer-Human Interaction*, pp.1–11.

Rosenholtz, R., Li, Y. & Nakano, L., 2007. Measuring visual clutter. *Journal of Vision*, 7(2)(17), pp.1–22.

Shepard, R.N. & Metzler, J., 1971. Mental rotation of three-dimensional objects. *Science*, 171(3972), pp.701–703.

Simons, D.J. & Rensink, R.A., 2005. Change blindness : Past , present , and future. *Trends in Cognitive Sciences*, 9(1), pp.16–20.

Standing, L., Conezio, J. & Haber, R.N., 1970. Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science*, 19(2), pp.73–74.

Thoresen, J.C. et al., 2016. Not all anxious individuals get lost: Trait anxiety and mental rotation ability interact to explain performance in map-based route learning in men. *Neurobiology of Learning and Memory*, 132, pp.1–8.

Touya, G. et al., 2015. Comparing image-based methods for assessing visual clutter in generalized maps. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-3/W5, pp.227–233.

Tuch, A.N. et al., 2009. Visual complexity of websites: Effects on users' experience, physiology, performance, and memory. *International Journal of Human Computer Studies*, 67(9), pp.703–715.

Tufte, E.R., 1989. *Envisioning information*, Cheshire, Connecticut: Graphics Press.

Vandenberg, S.G. & Kuse, A.R., 1978. Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and motor skills*, 47(2), pp.599–604.

Vogel, E., Woodman, G. & Luck, S., 2006. The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 32(6), pp.1436–1451.

Voyer, D. & Bryden, M.P., 1990. Gender, level of spatial ability, and lateralization of mental rotation. *Brain and Cognition*, 13(1), pp.18–29.

Wolfe, J.M. et al., 1992. The role of categorization in visual search for orientation. *Journal of Experimental Psychology Human Perception and Performance*, 18(1), pp.34–49.

Wolfe, J.M., 1998. Visual memory: what do you know about what you saw? *Current Biology*, 8(9), pp.R303-304.

Wolfe, J.M. & Horowitz, T.S., 2004. What attributes guide the deployment of visual attention and how do they do it? *Nature reviews. Neuroscience*, 5(6), pp.495–501.

Zelinsky, G.J. & Loschky, L.C., 2005. Eye movements serialize memory for objects in scenes. *Perception & Psychophysics*, 67(4), pp.676–690.

Appendix A. Matlab Code

A 1. Feature Congestion

```
function [clutter_scalar_fc, clutter_map_fc] = getClutter_FC(p);
%
% [clutter_scalar_fc, clutter_map_fc] = getClutter_FC(filename, [p]);
% computes Feature Congestion measure of visual clutter.
% Outputs:
%   "clutter_scalar_fc" is a scalar, which gives the Feature Congestion
%     clutter of the whole image.
%   "clutter_map_fc" is a clutter map (same size as the input image),
%     which gives local clutter information.
% Inputs:
%    "filename": the file name of an image
%    "p" (optional, default 1): a parameter when combining local clutter over
%      space; the combination can be considered Minkowski distance of order p

% This measure (Feature Congestion) of visual clutter is related to the
% local variability in certain key features, e.g., color, contrast, and
% orientation.
%
% Reference:
% Ruth Rosenholtz, Yuanzhen Li, and Lisa Nakano. "Measuring Visual Clutter".
% Journal of Vision, 7(2), 2007. http://www.journalofvision.com/7/2/

% Ruth Rosenholtz, Yuanzhen Li, and Lisa Nakano, March 2007.


if ~exist('p')
    p = 1;
end
S = dir %
[m,n] = size(S) %
jmena = cell(m,1)
for i=1:1:m
    x = S(i);
    j = x.name;
    jmena(i)=cellstr(j);
end
jmena

for i=3:m
filename=jmena{i}
% computelocal clutter in color, contrast, and orientation
% (please see "computeClutter.m" for info about the paramters and the outputs)
[color_clutter, contrast_clutter, orient_clutter] = computeClutter(filename, 3,
1, 3, 3, 0, 0, 0);

% combine color, contrast, and orientation:
```

```
%clutter_map_fc(i-2) = color_clutter{2}/0.2088 + contrast_clutter{2}/0.0660 +
orient_clutter{2}/0.0269;
clutter_map_fc = color_clutter{2}/0.2088 + contrast_clutter{2}/0.0660 + ori-
ent_clutter{2}/0.0269;
% shows me the image:
%imagesc(clutter_map_fc)


% combine over space using a Minkowski mean of order p, then take the average
% clutter_scalar_fc(i-2) = mean(clutter_map_fc(:).^p).^(1/p)
clutter_scalar_fc(i) = mean(clutter_map_fc(:).^p).^(1/p);

end
jmena(3:i)
clutter_scalar_fc(3:i)
```

A 2. Subband Entropy

```
function [clutter_se_final] = getClutter_SE(wlevels, wght_chrom)
% [clutter_se] = getClutter_SE(map_name, [wlevels], [wght_chrom])
% Subband Entropy measure of visual clutter.
% Outputs:
%     "clutter_se": the subband entropy clutter of the image.
% Inputs:
%     "map_name": input image. It can be a string (file name of the image),
%       or an array (the image itself).
%     "wlevels":  the number of scales (optional, default 3)
%     "wght_chrom": the weight on chrominance (optional, default 0.0625)

% This measure (Subband Entropy) of visual clutter is based on the notion
% that clutter is related to the number of bits required for subband
% (wavelet) image coding.
%
% Reference:
% Ruth Rosenholtz, Yuanzhen Li, and Lisa Nakano. "Measuring Visual Clutter".
% Journal of Vision, 7(2), 2007. http://www.journalofvision.com/7/2/
% Ruth Rosenholtz, Yuanzhen Li, and Lisa Nakano, March 2007.

if ~exist('wlevels')
    wlevels = 3;
end
if ~exist('wght_chrom')
    wght_chrom = 0.0625;
end
S = dir
[m,n] = size(S)
jmena = cell(m,1)
for i=1:1:m
    x = S(i);
    j = x.name;
    jmena(i)=cellstr(j);
end
jmena

for i=3:m
map_name=jmena{i}
% map_name='images/8d_map.jpg'

% load input image
if ischar(map_name)
    try
        [tmp, map] = imread(map_name);
    catch
        error(sprintf('Unable to open %s image file.', map_name))
    end
    map = double(tmp);
else
    if isnumeric(map_name)
        map = double(map_name);
```

```
    end
end

[ht, wth, dchrom] = size(map);
% if the image is a color rgb image, convert it to luminance and
% chrominance (from RGB to CIE Lab):
if dchrom == 3
    map_lab = RGB2Lab(map);
else
    map_lab = map;
end

wor = 4;
% luminance channel:
map = map_lab(:,:,1);
en_band = band_entropy(map, wlevels, wor);
clutter_se = mean(en_band);

if dchrom == 1
    return;
end
% chrominance channels:



for jj = 2:3
    map = map_lab(:,:,jj);
    if max(map(:))-min(map(:)) < 0.008
        map = zeros(size(map));
    end
    en_band = band_entropy(map, wlevels, wor);
    clutter_se = clutter_se + wght_chrom*mean(en_band);
end

clutter_se_final(i) = clutter_se/(1+2*wght_chrom);

end
jmena(3:i)
clutter_se_final(3:i)
```

A 3. Edge Density

```
% read in data, convert it to black and white, and then do the edge detection
with Canny edge detector
% https://ch.mathworks.com/matlabcentral/answers/1172-loop-that-creates-arrays
% (Array to store results)

srcFiles =
dir('\\service.geo.uzh.ch\private\doertle\data\Desktop\MsC\Experiment\Urban\Task
1\ClutterImage\*.jpg');  % the folder in which ur images exists
EdgeDensity_Rural = zeros (length(srcFiles), 1);
for i = 1 : length(srcFiles);
filename =
strcat('\\service.geo.uzh.ch\private\doertle\data\Desktop\MsC\Experiment\Urban\T
ask1\ClutterImage\',srcFiles(i).name);
I = imread(filename);

BW = rgb2gray(I);

% Rosenholtz et al. (2007),
% edge (I, canny, threshold, sigma), according to Rosenholtz et al. for
% thresholds 0.11 and 0.27 were used, sigma = 1

BW1 = edge(BW, 'Canny', 0.27, 1);

% blackcount (background) and whitecount (edges)
https://ch.mathworks.com/matlabcentral/newsreader/view_thread/279293

blackcount = sum(sum(BW1==0))
whitecount = sum(sum(BW1==1))

total = sum(blackcount + whitecount)

edge_density = (whitecount / total)
rest_density = (blackcount / total)

EdgeDensity_Rural(i) = edge_density;
end
```

A 4. Image Segmentation

```
% demoFelzenszwalbSegmentation
%
% author: Alexander Freytag
% date:   28-03-2014 (dd-mm-yyyy)
%
% brief:  small demo showing how to use the felzenszwalb segmentation code
%         via Matlab and how to adjust parameter specifications.
```

```
srcFiles =
dir('\\service.geo.uzh.ch\private\doertle\data\Desktop\MsC\Experiment\Urban\10_S
electedJPEG\Cropped\*.jpg');  % the folder in which ur images exists
Segmentation_rural = zeros (length(srcFiles), 1);
for i = 1 : length(srcFiles);

filename =
strcat('\\service.geo.uzh.ch\private\doertle\data\Desktop\MsC\Experiment\Urban\1
0_SelectedJPEG\Cropped\',srcFiles(i).name);
img = imread(filename);

%read image ...

% show input image in figure
%figOrig = figure;
%set ( figOrig, 'name', 'Input image');
%imshow ( img );

% call felzenszwalb segmentation using our mex-interface
%segResult = segmentFelzenszwalb(img);

% show segmentation result in separate figure
%figSegResult = figure;
%set ( figSegResult, 'name', 'Segmentation result');

%imshow ( segResult );
% make region colors visually distinguishable
%colormap ( 'lines' );

d_sigma               = 0.5;   % default: 0.5
i_k                   = 500;   % default: 500, segmentation threshold, influences
the threshold accept a boundary between regions, larger k -> larger regions
i_minSize             = 50;    % default: 50, minimum component size enforced by
post-processing stage
b_computeColorOutput = true;   % default: false
s_destination         = ''; % default: '', filename where the image will be
stored
b_verbose             = true; % default: false, print additional information

[srManSpec,noRegions]  = segmentFelzenszwalb( img, d_sigma, i_k, i_minSize, ...
                                 b_computeColorOutput, ...
```

```
                                 s_destination, ...
                                 b_verbose...
                               );

% show segmentation result in separate figure
%figSRManSpec = figure;
%set ( figSRManSpec, 'name', 'Segmentation result, man. specified');

%imshow ( srManSpec );

% actually, another linespec is not needed, since the output is already a
% nice RGB image. However, for easier visual comparison, we again apply the
% matlab color mapping here.
%colormap ( 'lines' );

Segmentation_rural (i) = noRegions;
% wait for user input
%pause

% close images
%close ( figOrig );
%close ( figSegResult );
%close ( figSRManSpec );
```

Appendix B. Experiment

B 1. Questionnaire

2. Geschlecht

- ○ weiblich
- ○ männlich

3. Alter

4. Welches ist Ihre höchste abgeschlossene Schulbildung?

- ○ Sekundarstufe 1
- ○ Sekundarstufe 2: Gymnasium
- ○ Sekundarstufe 2: Berufslehre
- ○ Pädagogische Hochschule, Fachhochschule, Höhere Fachschule
- ○ Bachelor (Universität / ETH)
- ○ Master (Universität / ETH)
- ○ Doktorat (Unviersität / ETH)
- ○ Anderes

Next

5. Benutzen Sie eine Brille oder Kontaktlinsen?

- ○ Ja
- ○ Nein

6. Nutzen Sie diese jetzt?

- ○ Ja, Brille
- ○ Ja, Kontaktlinsen
- ○ Nein

7. Wurde Ihnen von einem Arzt o.ä eine Farbfehlsichtigkeit diagnostiziert?

- ○ Ja
- ○ Nein

8. Wie viele Stunden haben Sie letzte Nacht ungefähr geschlafen (in h)?

Next

9. Bewerten Sie ihre Erfahrung in den folgenden Gebieten

	1 Keine	2	3 Durchschnittlich	4	5 Hoch
Kartographie	○	○	○	○	○
Raumplanung / Stadtplanung	○	○	○	○	○
GIScience	○	○	○	○	○
andere Gebiete der Geographie	○	○	○	○	○
Computer Grafik	○	○	○	○	○
Bildanalyse / Bildinterpretation	○	○	○	○	○
Fotografie	○	○	○	○	○

10. Wie oft benutzen Sie digitale Karten auf ihrem Handy / Tablet / Computer (z.B. Google maps)?

1 Nie	2	3 Wöchentlich	4	5 Täglich
○	○	○	○	○

11. Welche Ansicht bevorzugen Sie, wenn Sie digitale Karten benützen?

1 Kartenansicht	2	3 Keine Bevorzugung	4	5 Satellitenansicht
○	○	○	○	○

B 2. Map Memory Task

Page 1

1. Teilnehmendennummer (wird von der Studienleitung ausgefüllt) *

Page 2

2_1_MapMemory

Dies ist ein Test, um ihre Fähigkeit, Teile einer Karte zu merken, zu testen. Der gesamte Test dauert ungefähr 12 Minuten.

Zu Beginn werden Ihnen mehrere Kartenausschnitte gezeigt. Sie haben dann 3 Minuten Zeit, diese Kartenausschnitte möglichst gut einzuprägen. Sie werden nach der Zeitlimite automatisch zur Aufgabe weitergeleitet.

Danach haben Sie wiederum 3 Minuten Zeit die Kartenausschnitte auszuwählen, welche Sie vorhin schon einmal gesehen haben oder nicht. Sie beantworten immer mit Ja oder Nein.

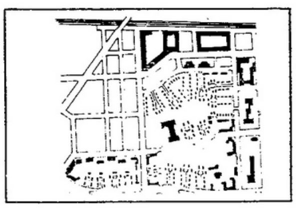
Auf der nächsten Seite folgt ein Beispiel. Sie haben dabei 30 Sekunden Zeit die Karten einzuprägen und werden danach zur Aufgabe weitergeleitet.

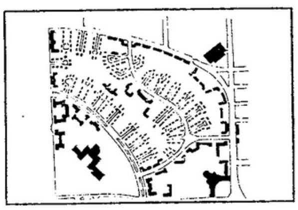
Drücken Sie auf "Next" um das Beispiel zu starten.

Next

Page 3

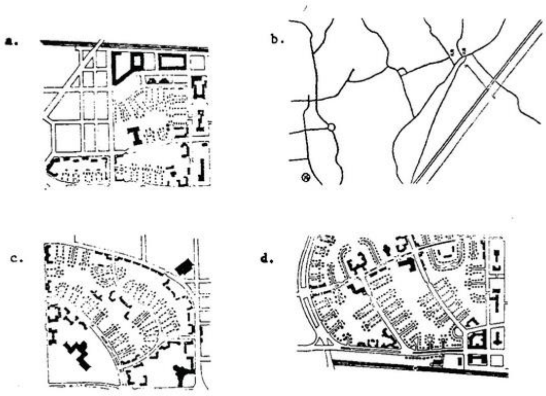
Sie haben nun 30 Sekunden Zeit die Kartenausschnitte einzuprägen.





Next

Page 4



2. Haben Sie diesen Kartenausschnitt gesehen?

	Ja	Nein
a	○	○
b	○	○
c	○	○
d	○	○

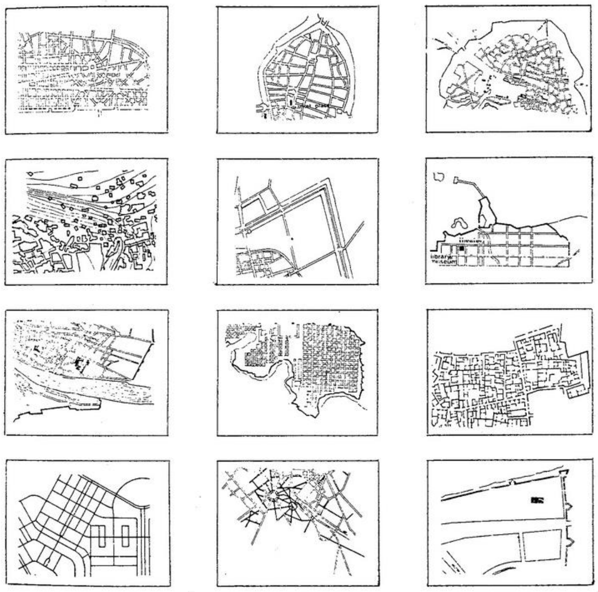
Page 5

Der Test hat zwei Teile, welche jeweils aus zwei Phasen bestehen. In der ersten Phase haben Sie 3 Minuten Zeit, die Kartenausschnitte einzuprägen. Danach werden Sie automatisch zur Aufgabe weitergeleitet. Sie haben dann wiederum 3 Minuten Zeit für die Lösung der Aufgabe. Sie können jederzeit auch früher "Next" drücken.

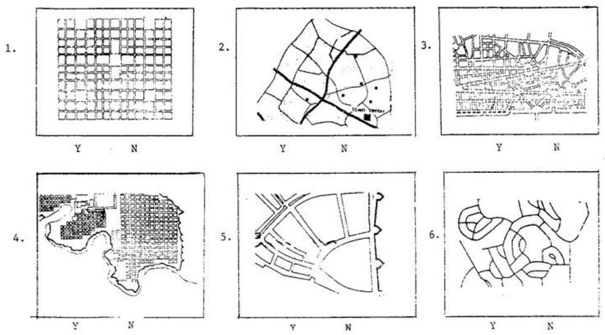
Ihr Testergebnis wird berechnet, indem die Anzahl der richtigen Antworten aufsummiert und davon die Anzahl falscher Antworten abgezogen wird.
Aus diesem Grund ist es *nicht vorteilhaft*, eine Lösung zu erraten, ausser Sie können einige Kartenausschnitte als definitiv falsch ausschliessen.

Drücken Sie auf "Next", um den Test zu starten.

Page 6



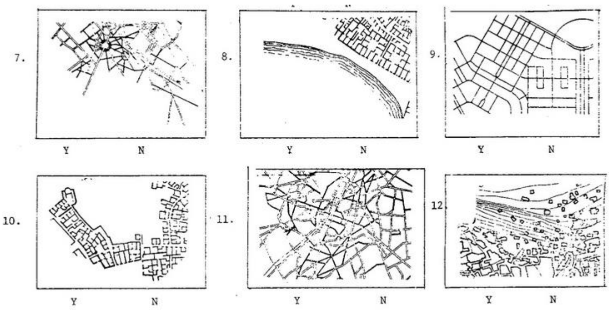
Page 7



3. Haben Sie diesen Kartenausschnitt gesehen?

	Ja	Nein
1.	○	○
2.	○	○
3.	○	○
4.	○	○
5.	○	○
6.	○	○

Page 8



4. Haben Sie diesen Kartenausschnitt gesehen?

	Ja	Nein
7.	○	○
8.	○	○
9.	○	○
10.	○	○
11.	○	○
12.	○	○

Page 9

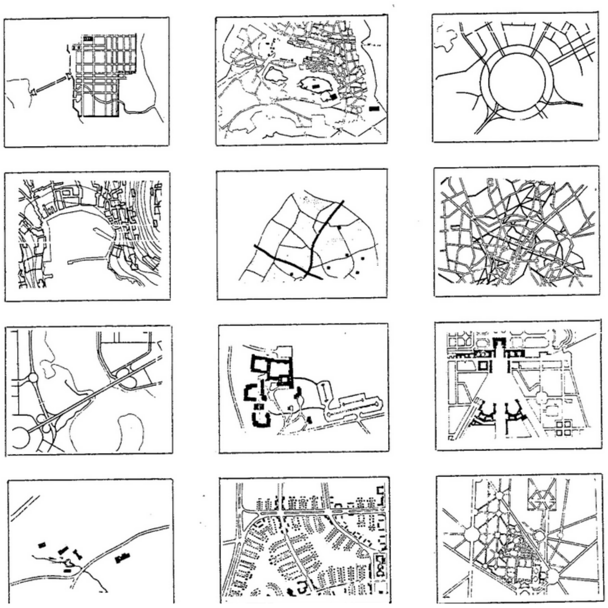
Nun kommen Sie zum zweiten Teil. Sie haben wiederum zuerst 3 Minuten Zeit für die Lernphase und werden dann getestet.

Drücken Sie auf "Next" um zu starten.

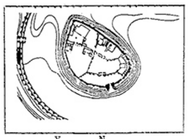
Page 10

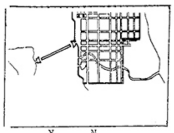
Betrachten Sie für die nächsten 3 Minuten die folgenden Kartenausschnitte.

Page 11



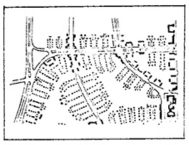
Page 12

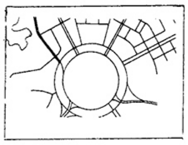
13.

Y N

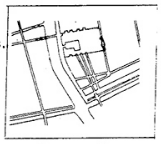
14.

Y N

15.

Y N

16.

Y N

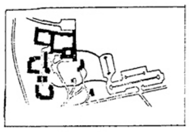
17.

Y N

18.

Y N

5. Haben Sie diesen Kartenausschnitt gesehen?

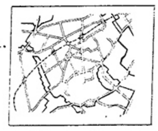
	Ja	Nein
13.	○	○
14.	○	○
15.	○	○
16.	○	○
17.	○	○
18.	○	○

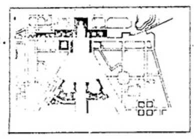
Page 13

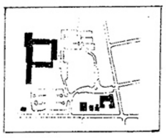
19.

Y N

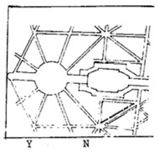
20

Y N

21.

Y N

22.

Y N

23.

Y N

24.

Y N

6. Haben Sie diesen Kartenausschnitt gesehen?

	Ja	Nein
19.	○	○
20.	○	○
21.	○	○
22.	○	○
23.	○	○
24.	○	○

B 3. Main Experiment

Only a few examples per task are shown due to large amount of data. The full survey can be downloaded on the CD.

Filenames:

Task 1: Userstudy_Task1.pdf

Task 2: Userstudy_Task2.pdf

Task 3: Userstudy_Task3.pdf

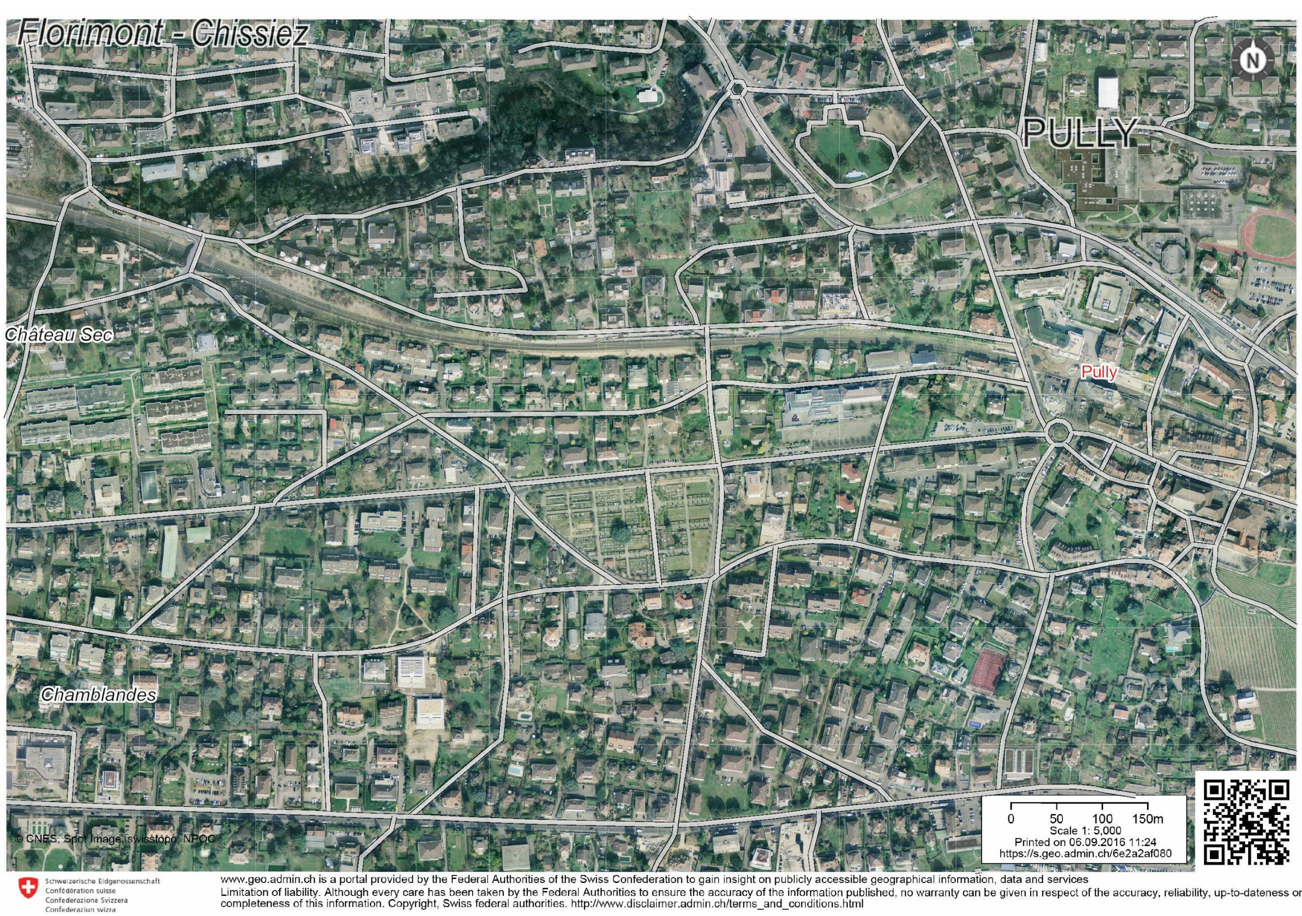
Task 4: Userstudy_Task4.pdf

The tasks are illustrated in the following order: Task 1, Task 2, Task 3 and Task 4 labelled in each header.

Anleitung

Zu Beginn der Aufgabe gibt es immer eine Lernphase, in welcher Sie 20 Sekunden ein Satellitenbild zu sehen bekommen. Das Ziel ist es, in dieser kurzen Zeit möglichst viele Details dieses Bildes einzuprägen.
Auf der nächsten Seite (Testphase) sollen Sie dann erkennen könnnen, welchen Ausschnitt Sie schon gesehen haben.

Sie haben nun 20 Sekunden Zeit das Satellitenbild anzuschauen. Keine Panik, es handelt sich vorerst um ein Beispiel.



Task 1

Beispiel Testphase

Nun sollen Sie in der Testphase den passenden Ausschnitt zum Satellitenbild wieder erkennen und in der Auswahl A,B, C oder D auswählen.

Sie haben 15 Sekunden Zeit dafür. Sobald die Zeit abgelaufen ist, werden sie automatisch zur nächsten Seite weitergeleitet.

A

B

C

D

2. Welchen Ausschnitt haben Sie vorhin gesehen?

- ◯ A
- ◯ B
- ◯ C
- ◯ D

Die richtige Antwort wäre hier "C".

Task 1

Start

Nicht vergessen, Sie haben jeweils 20 Sekunden Zeit für die Lernphase und 15 Sekunden für die Testphase. Sobald die Zeit abgelaufen ist, werden Sie automatisch zur nächsten Seite weitergeleitet. Sie können aber auch jederzeit schon vor der Zeit weiter springen, indem sie auf "Next" klicken.

Das Satellitenbild wird jeweils mit verschiedenen Informationen (z.B. Orstnamen, Strassennetz) überlagert, kann aber auch ohne erscheinen. Alle Satellitenbilder haben immer den gleichen Massstab (1:5'000).

Drücken Sie auf "Next", um die Studie zu starten.

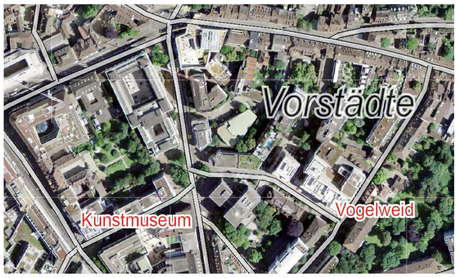
Task 1

Einprägungsphase (20 Sekunden Zeit)

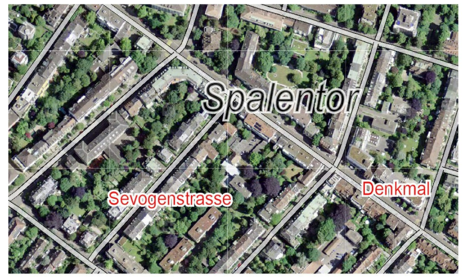
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
In collaboration with the cantons

Task 1

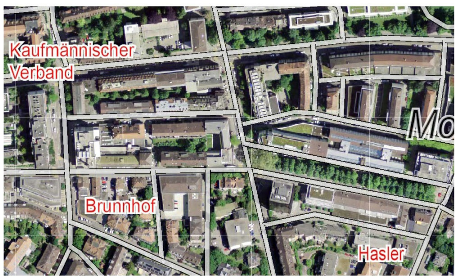
Testphase (15 Sekunden Zeit)



A



B



C

D

3. Welchen Ausschnitt haben Sie vorhin gesehen?

○ A ○ B ○ C ○ D

Task 1

Einprägungsphase

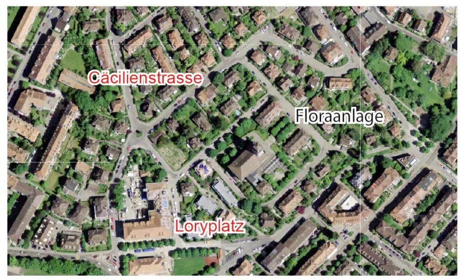
Task 1

Testphase

A

B

C



D

4. Welchen Ausschnitt haben Sie vorhin gesehen?

○ A ○ B ○ C ○ D

Beispiel: Einprägungsphase

Zu Beginn der Aufgabe gibt es immer eine Einprägungsphase. Sie sehen dabei für 10 Sekunden einen kleinen Ausschnitt von einem Satellitenbild. Sie sollen sich dann diesen Ausschnitt einprägen, um ihn kurze Zeit später in einem grossen Satellitenbild wiederzuerkennen.

Sie haben nun 10 Sekunden Zeit den Ausschnitt einzuprägen. Keine Panik, es handelt sich vorerst um ein Beispiel.

Beispiel - Einprägungsphase

Beispiel: Testphase1

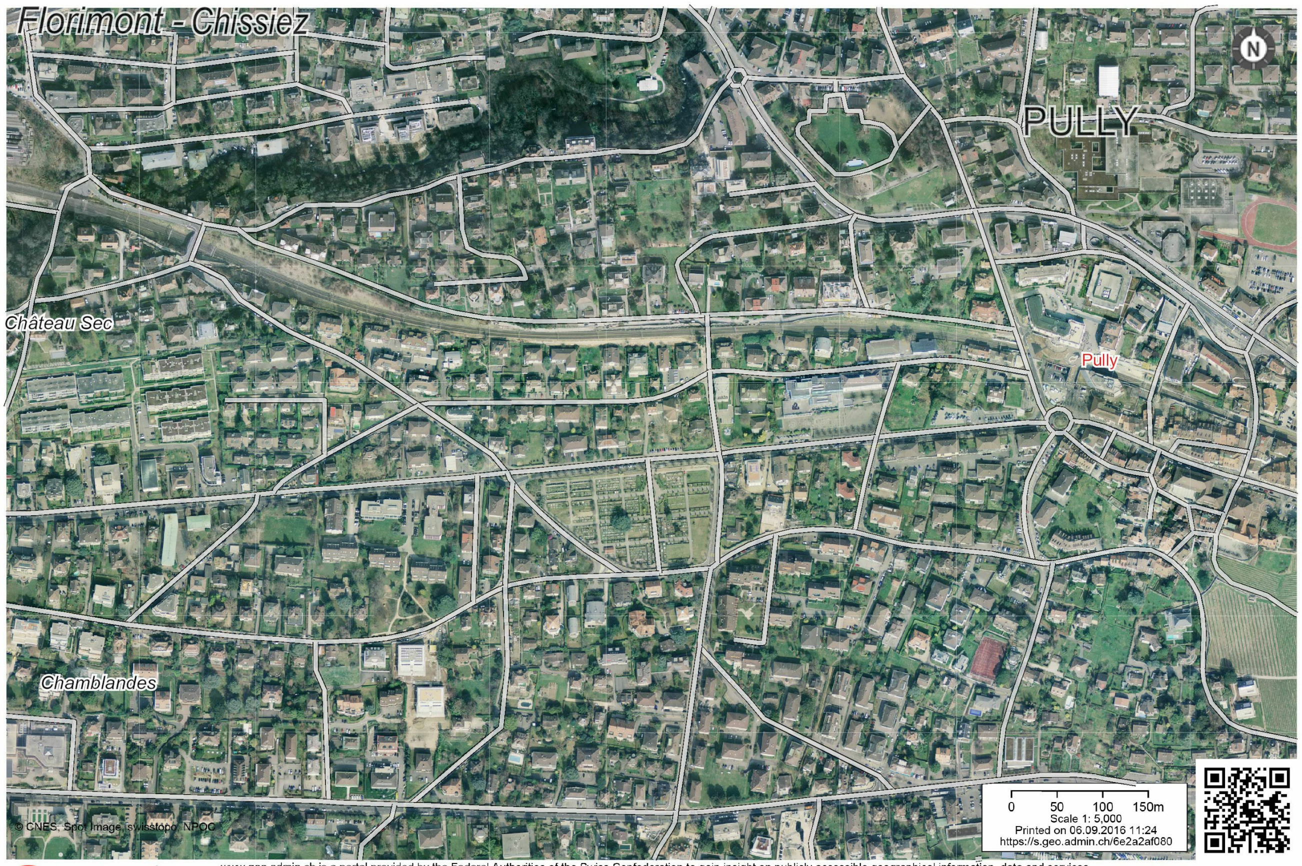
Nun sollen Sie in der Testphase den passenden Ausschnitt zum Satellitenbild wieder erkennen und in der Auswahl A,B, C oder D auswählen. Es können auch mehrere Optionen richtig sein.

Sie haben 25 Sekunden Zeit dafür. Scrollen Sie nach unten, um alle Bilder zu sehen.

A

Task 2

B



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
In collaboration with the cantons

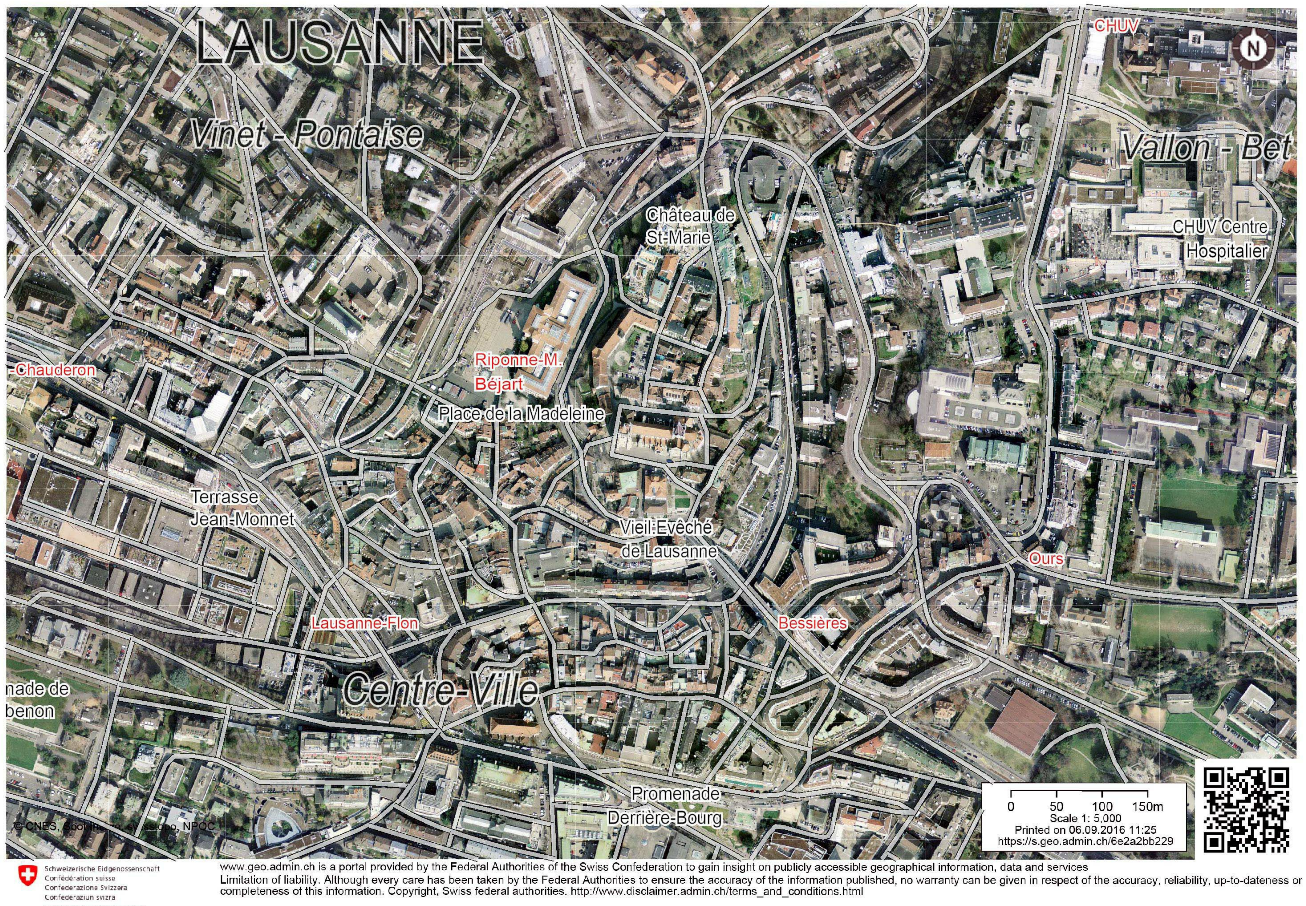
www.geo.admin.ch is a portal provided by the Federal Authorities of the Swiss Confederation to gain insight on publicly accessible geographical information, data and services

Task 2

C


Strassburgerallee
Universitätsspital Basel Augenklinik
Universitäts-Kinderspital beider Basel (UKBB)
N
Universitätsspital Basel Frauenklin
Psychiatrische Poliklinik (PUPM)
Burgfelderplatz
Psychiatrische Klinik Sonnenhalde
Universitätsspital Basel
Hegenheimerstrasse
Pilgerstrasse
Universitätsspital Basel Institut für Pathologie
Adullam-Stiftung Basel
Botanischer Garten der Universität Basel
Spalentor
Allschwilerplatz
Birmannsgasse
Universität
Merian Iselin Spital Basel
© CNES, Spot Image, swisstopo, NPOC
0 50 100 150m
Scale 1: 5,000
Printed on 06.09.2016 11:20
https://s.geo.admin.ch/6e2a27b8d1

D



2. Welchen Ausschnitt haben Sie vorhin gesehen?

- [ ] A
- [ ] B
- [ ] C
- [ ] D

Task 2

Start

Es gibt pro Ausschnitt jeweils vier 4 Satellitenbilder zur Auswahl. Es können auch mehrere Satellitenbilder richtig sein.

Nicht vergessen, Sie haben jeweils 10 Sekunden Zeit für die Lernphase und 25 Sekunden für die Testphase. Sobald die Zeit abgelaufen ist, werden Sie automatisch zur nächsten Seite weitergeleitet. Sie können aber auch jederzeit schon vor der Zeit weiter springen, indem sie auf "Next" klicken.

Das Satellitenbild wird jeweils mit verschiedenen Informationen (z.B. Orstnamen, Strassennetz) überlagert, kann aber auch ohne erscheinen. Alle Satellitenbilder haben immer den gleichen Massstab (1:5'000).

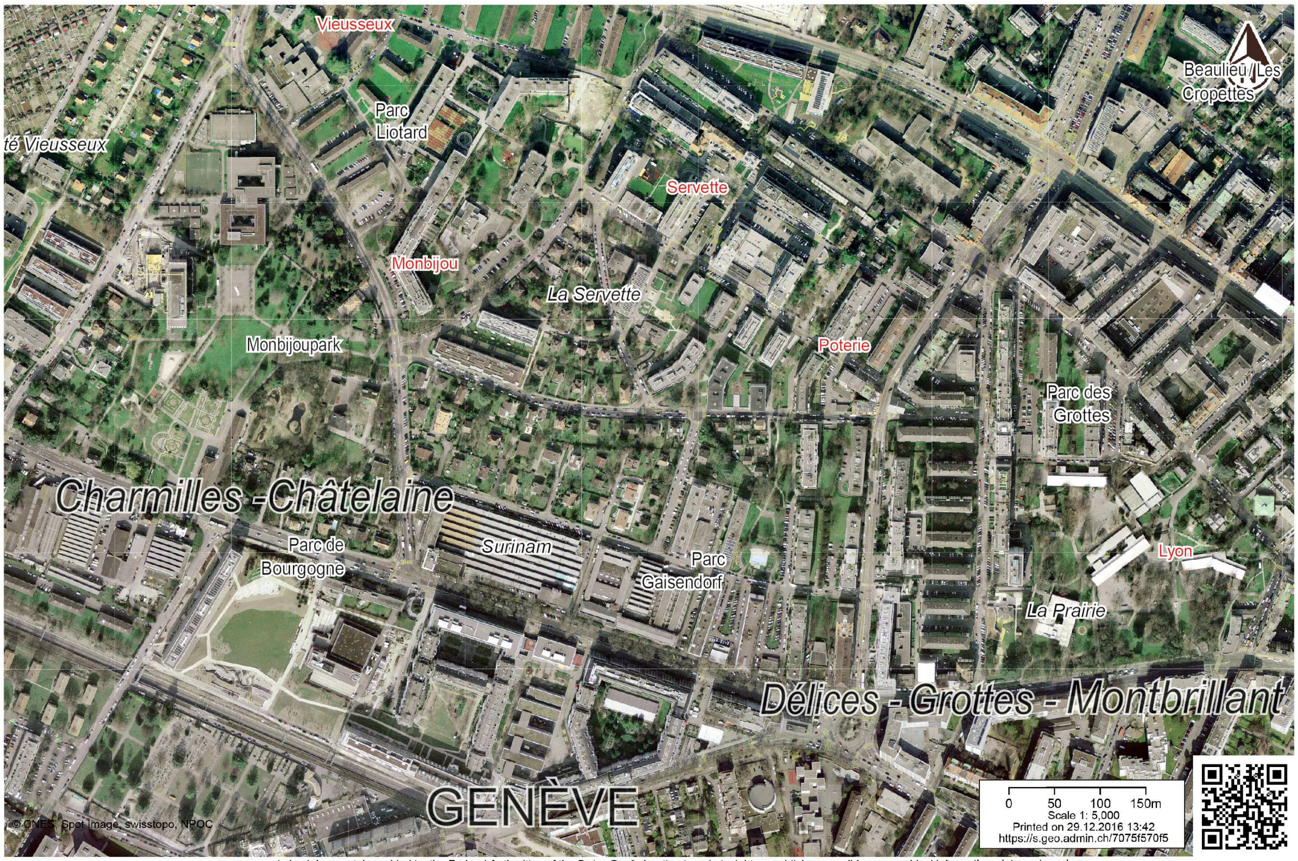
Drücken Sie auf "Next", um die Studie zu starten.

Task 2

Einprägungsphase (10 Sekunden Zeit)

Testphase

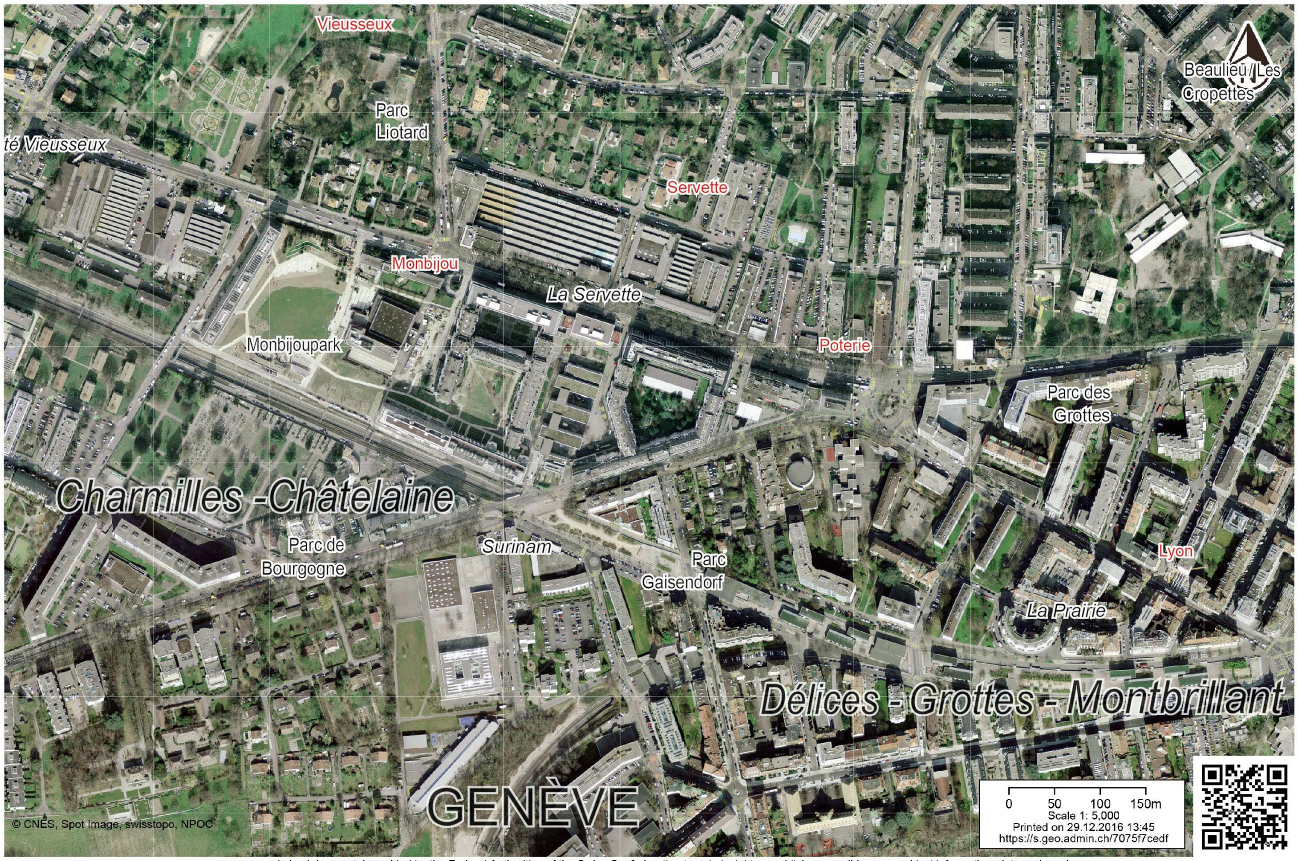
A



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
In collaboration with the cantons

Testphase (10 Sekunden Zeit)

B



Testphase

C


Vieusseux
Parc
Liotard
té Vieusseux
Servette
Beaulieu / Les
Cropettes
Monbijou
La Servette
Monbijoupark
Poterie
Parc des
Grottes
Charmilles - Châtelaine
Parc de
Bourgogne
Surinam
Parc
Gaisendorf
Lyon
La Prairie
Délices - Grottes - Montbrillant
GENÈVE
© CNES, Spot Image, swisstopo, NPOC
0 50 100 150m
Scale 1: 5,000
Printed on 29.12.2016 13:46
https://s.geo.admin.ch/7075f8c5aa

Testphase

D

3. Welchen Ausschnitt haben Sie vorhin gesehen?

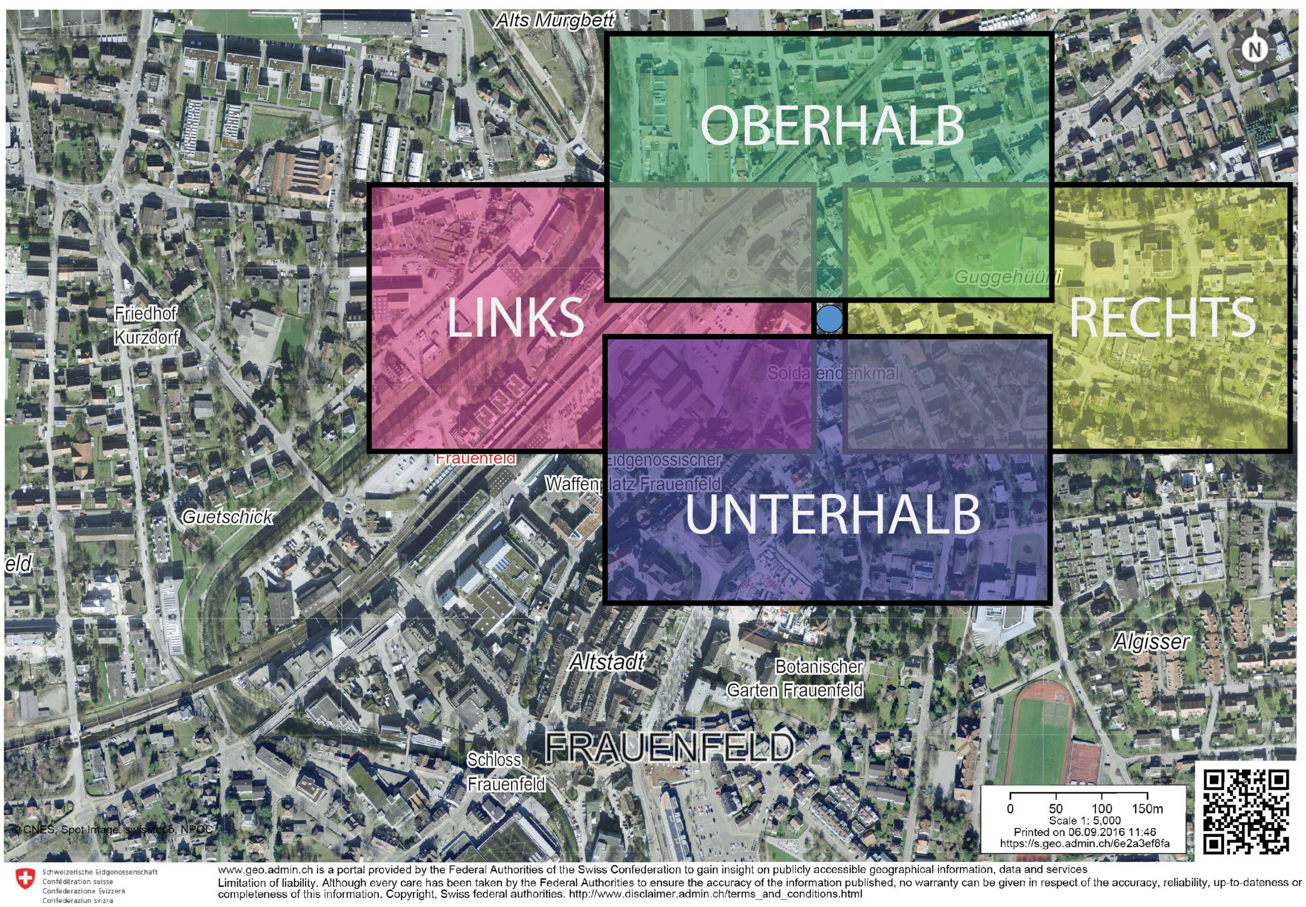
- ☐ A
- ☐ B
- ☐ C
- ☐ D

Anleitung

Zu Beginn der Aufgabe gibt es immer eine Lernphase. Sie sehen dabei für 15 Sekunden ein Satellitenbild mit einem blauen Punkt. Im Beispiel unten sehen sie vier Rechtecke (links, rechts, oberhalb und unterhalb) vom blauen Punkt. Die Rechtecke grenzen jeweils an den blauen Punkt. Die Lösung für die Aufgabe ist immer eines dieser Rechtecke.

In der Studie werden die Rechtecke nicht eingeblendet.

Sie haben nun 15 Sekunden Zeit den Ausschnitt einzuprägen. Keine Panik, es handelt sich vorerst um ein Beispiel.



Task 3

Beispiel Testphase

Nun sollen Sie in der Testphase den passenden Ausschnitt zum Satellitenbild wieder erkennen und in der Auswahl A,B, C oder D auswählen. Sie haben 15 Sekunden Zeit.

Wichtig: In der Studie wird nicht immer nach dem Ausschnitt links vom Punkt gefragt. Es kann auch nach dem Ausschnitt oberhalb, unterhalb oder rechts gefragt werden.

2. Welcher Ausschnitt liegt direkt links vom blauen Punkt?

◯ A ◯ B ◯ C ◯ D

A

B

C

D

Antwort "C" wäre hier richtig.

Task 3

Start

Nicht vergessen, Sie haben jeweils 15 Sekunden Zeit für die Lernphase und 15 Sekunden für die Testphase. Sobald die Zeit abgelaufen ist, werden Sie automatisch zur nächsten Seite weitergeleitet. Sie können aber auch jederzeit schon vor der Zeit weiter springen, indem sie auf "Next" klicken.

Das Satellitenbild wird jeweils mit verschiedenen Informationen (z.B. Orstnamen, Strassennetz) überlagert, kann aber auch ohne erscheinen. Alle Satellitenbilder haben immer den gleichen Massstab (1:5'000).

Drücken Sie auf "Next", um die Studie zu starten.

Task 3

Einprägungsphase (10 Sekunden Zeit)

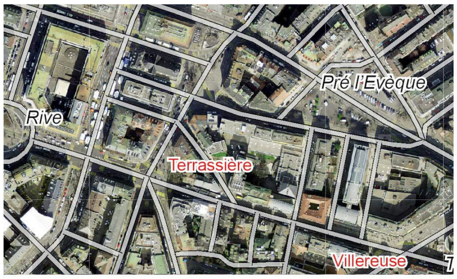
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
In collaboration with the cantons

Task 3

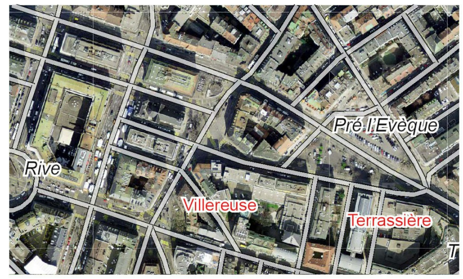
Testphase (15 Sekunden Zeit)

3. Welcher Ausschnitt liegt direkt <u>oberhalb</u> vom blauen Punkt?

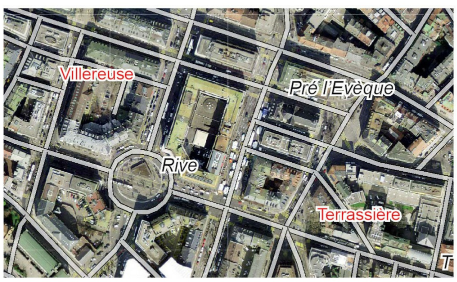
◯ A ◯ B ◯ C ◯ D



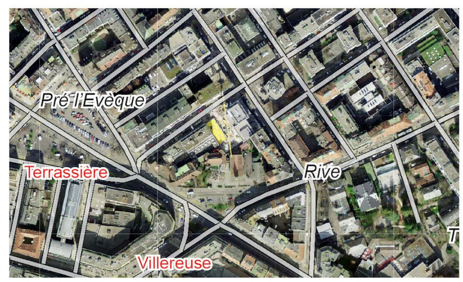
A



B



C



D

Task 4

Anleitung

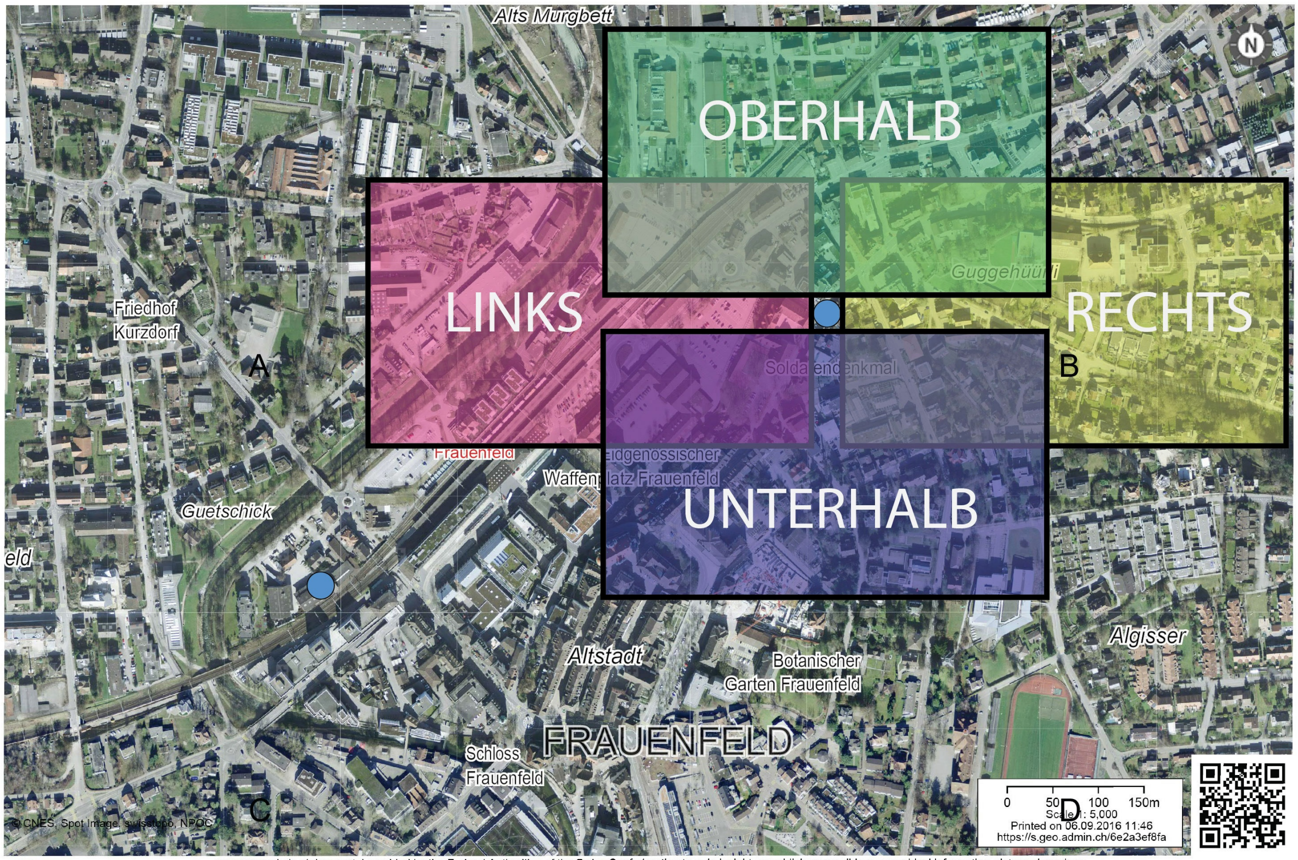
Zu Beginn der Aufgabe gibt es immer eine Lernphase. Sie sehen dabei für 25 Sekunden Zeit ein Satellitenbild mit zwei blauen Punkten. Im Beispiel unten sehen sie vier Rechtecke (links, rechts, oberhalb und unterhalb) von einem der zwei blauen Punkte. Die Rechtecke grenzen jeweils an den blauen Punkt.

In dieser Aufgabe müssen sie die Umgebung beider blauen Punkte einprägen. Auf der nächsten Seite werden Sie für eine kurze Zeit einen dieser 2 Punkte in einem Bilderrahmen wiedersehen. Das Satellitenbild wird jedoch nicht sichtbar sein. Der Punkt befindet sich an exakt gleicher Stelle wie im Originalbild.

In der Testphase werden Sie dann gefragt, welcher der vier Ausschnitte links/rechts/unterhalb/oberhalb des zuvor gesehenen blauen Punktes liegt.

Sie haben nun 25 Sekunden Zeit den Ausschnitt einzuprägen. Keine Panik, es handelt sich vorerst um ein Beispiel.

Task 4

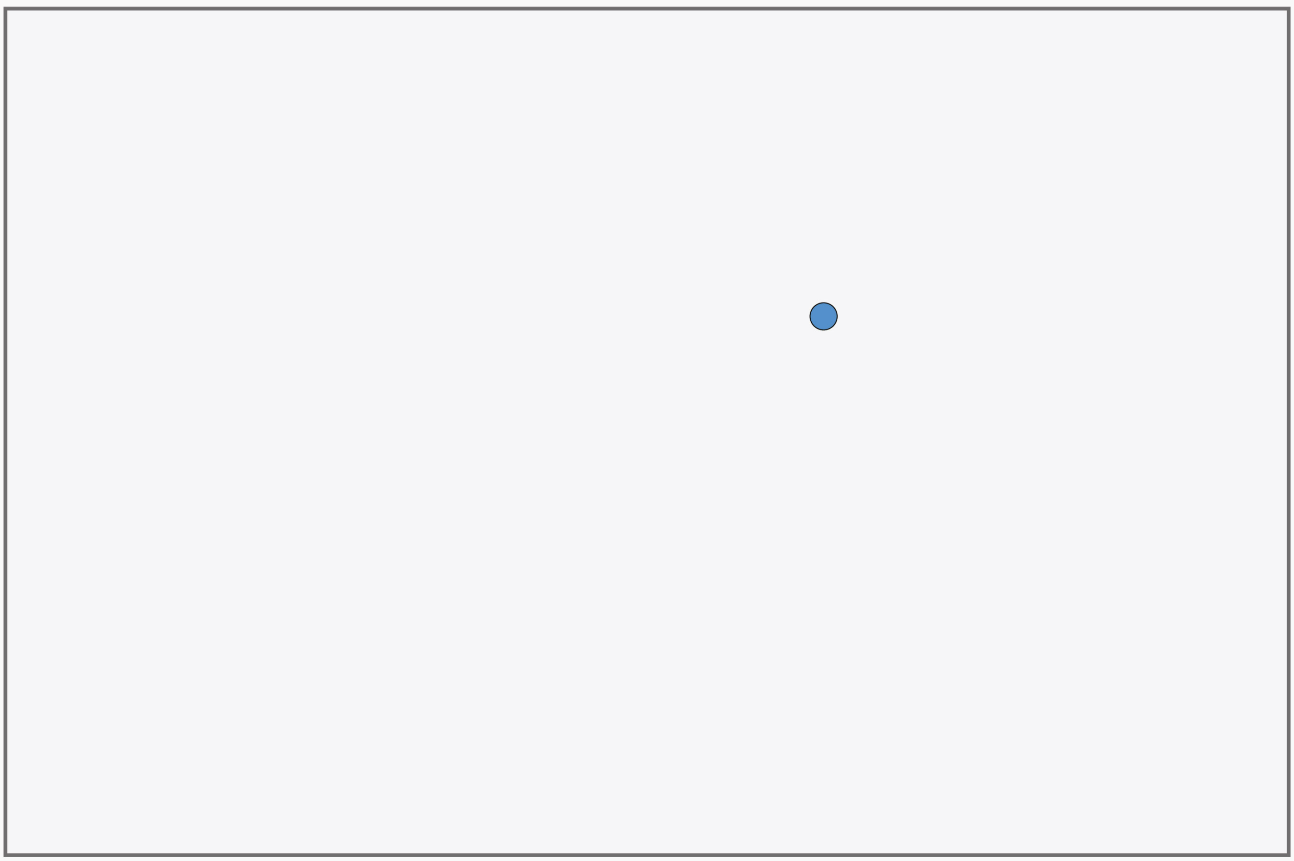

Alts Murgbett
OBERHALB
LINKS
RECHTS
UNTERHALB
Guggehüürli
Soldatendenkmal
Friedhof
Kurzdorf
A
B
C
D
Frauenfeld
Eidgenössischer
Waffenplatz Frauenfeld
Guetschick
Altstadt
Botanischer
Garten Frauenfeld
Algisser
FRAUENFELD
Schloss
Frauenfeld
0 50 100 150m
Scale 1: 5,000
Printed on 06.09.2016 11:46
https://s.geo.admin.ch/6e2a3ef8fa
© CNES, Spot Image, swisstopo, NPOC

Task 4

Zuweisung Beispiel

Der unten abgebildete Punkte wurde vom System ausgewählt. Dieser befindet sich an exakt genau gleicher Stelle wie zuvor im Originalbild. Merken Sie sich, um welchen der beiden Punkte es sich dabei handelt.
Auf der nächsten Seite sollen Sie dann zur Umgebung dieses Punktes eine Frage beantworten.

In der Studie wird diese Seite jeweils nur 2 Sekunden angezeigt.



Task 4

Beispiel Testphase

Nun sollen Sie in der Testphase den passenden Ausschnitt zum Satellitenbild wieder erkennen und in der Auswahl A,B, C oder D auswählen. Sie haben 15 Sekunden Zeit.

Wichtig: In der Studie wird nicht immer nach dem Ausschnitt rechts vom Punkt gefragt. Es kann auch nach dem Ausschnitt oberhalb, unterhalb oder links gefragt werden.

2. Welcher Ausschnitt liegt direkt rechts vom blauen Punkt, welcher Ihnen zuvor gezeigt wurde?

◯ A ◯ B ◯ C ◯ D

A

B

C

D

Antwort "C" wäre hier richtig.

Task 4

Start

Nicht vergessen, Sie haben jeweils 25 Sekunden Zeit für die Lernphase und 15 Sekunden für die Testphase. Der gesuchte Punkte wird jeweils nur für 2 Sekunden angezeigt.

Sobald die Zeit abgelaufen ist, werden Sie automatisch zur nächsten Seite weitergeleitet. Sie können aber auch jederzeit schon vor der Zeit weiter springen, indem sie auf "Next" klicken.

Das Satellitenbild wird jeweils mit verschiedenen Informationen (z.B. Orstnamen, Strassennetz) überlagert, kann aber auch ohne erscheinen. Alle Satellitenbilder haben immer den gleichen Massstab (1:5'000).

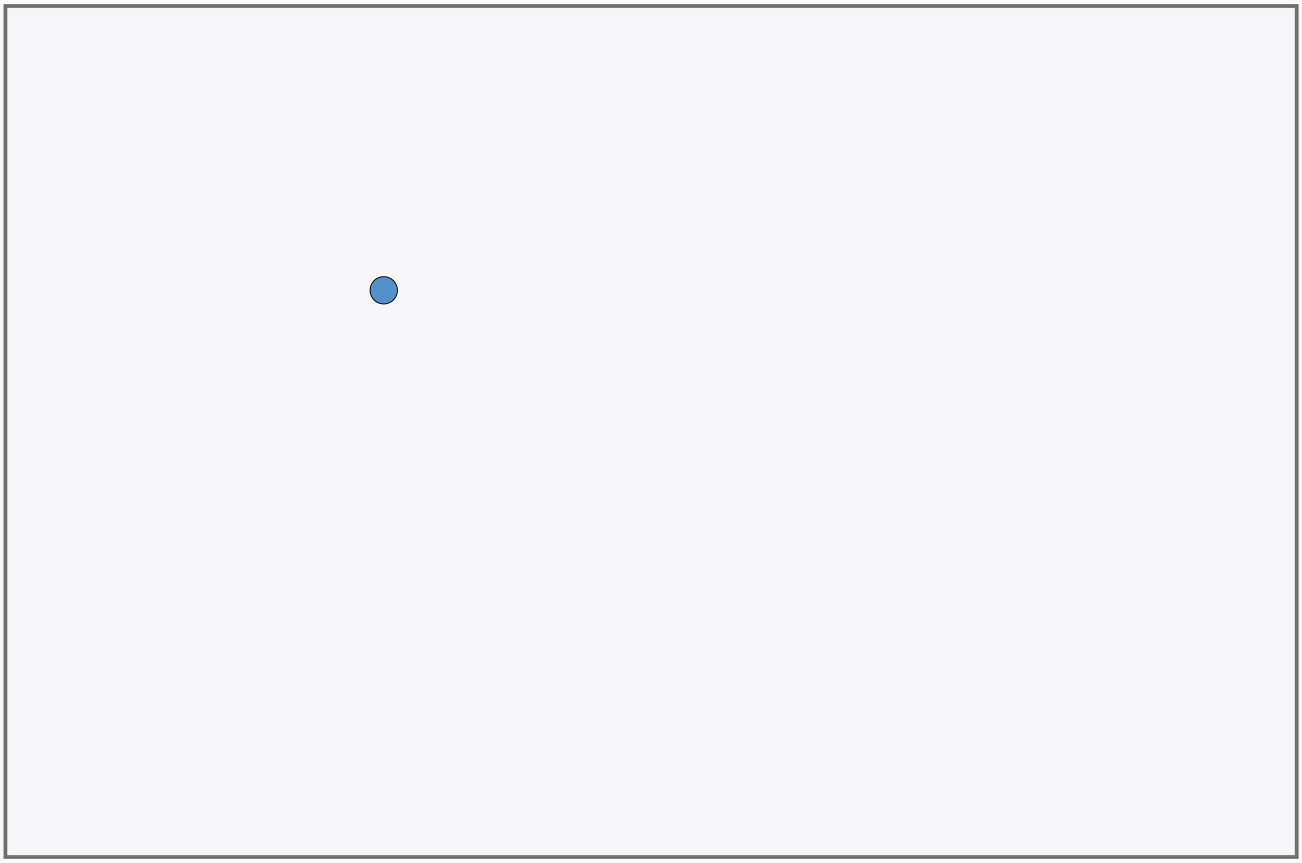
Drücken Sie auf "Next", um die Studie zu starten.

Task 4

Einprägungsphase (20 Sekunden Zeit)

Task 4

Zuweisung (3 Sekunden Zeit)



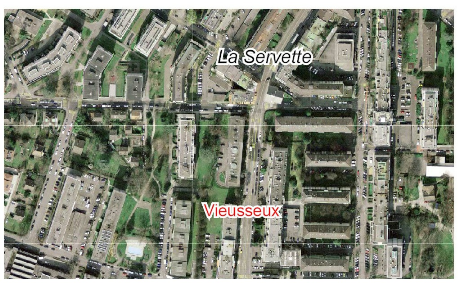
Testphase (15 Sekunden Zeit)

3. Welcher Ausschnitt liegt direkt <u>oberhalb</u> vom blauen Punkt, welcher Ihnen zuvor gezeigt wurde?

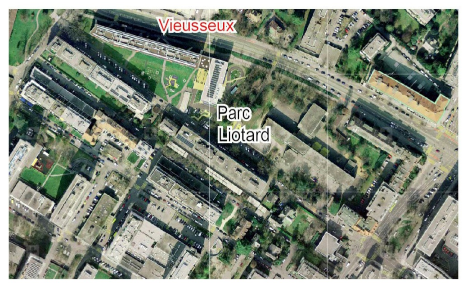
- ◯ A
- ◯ B
- ◯ C
- ◯ D

A

B



C



D

Personal declaration

I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

Dario Oertle
Zürich, April 21th 2017