Akademia Górniczo-Hutnicza

im. Stanisława Staszica w Krakowie

Wydział Geodezji Górniczej i Inżynierii Środowiska

Zakład Fotogrametrii i Informatyki Teledetekcyjnej

Praca dyplomowa

MAGISTERSKA

Imię i nazwisko: ***Piotr Tokarczyk***

Kierunek studiów: **Geodezja i kartografia**

Temat pracy dyplomowej: ***„Automatyzacja pozyskiwania danych i ich przetwarzania na potrzeby systemu nawigacji dla pieszych”***

Master’s thesis title: ***“Automatization of data acquisition and processing for pedestrian navigation system purposes”***

Promotor: *Recenzent:*

Ocena: ..

Kraków, rok akademicki:

Niniejsza praca dyplomowa opracowana została przeze mnie samodzielnie i zgodnie z Ustawa o prawie autorskim i prawach pokrewnych z dnia 4.02.1994 r. (Dz. U. 1994 nr 24 poz. 83) wraz z nowelizacja z dnia 25.03.2003 r. (Dz. U. 2003 nr 166 poz. 1610) oraz z dnia 1.04.2004 r. (Dz. U. 2004 nr 91 poz. 869).

....

(czytelny podpis studenta)

Program pracy i praktyki dyplomowej:

1.

Summary

Wayfinding is a vital part of our everyday life. Since it is our daily routine, hardly anybody realizes what a demanding task it is and that it requires certain spatial and cognitive abilities. To develop supporting tools for wayfinding, it is essential to know the mechanisms that control these processes. With a progress of technology, the size and weight of electronic devices have diminished significantly. The dropping of prices and wide-ranging availability of such devices increased interest in such systems. Factors like size and availability were the reason for calling them “ubiquitous systems”.

The market is full of perfectly working navigation systems for car drivers. They are widespread, have high user-acceptance level and their market is fast-growing. Can one modify such systems in a way that pedestrians can use it? The problem is more complex than one might imagine at first sight. Field tests show, that systems for car drivers do not meet the requirements of pedestrian users. Car and pedestrian navigation differ in: degree of freedom, velocity of movement and spatial resolution.

In this thesis I want to focus on data acquisition. The required data cannot be based on the same datasets used for car navigation systems. Automatization of the process of data acquisition is also required. Theoretical base of pedestrian navigation system was introduced in this thesis. The emphasis was put on the mechanisms responsible for surrounding environments’ perception. A conception of cognitive maps and image schemata, which are working in our minds, was presented. Moreover, this work explains what our needs and information categories are, while performing wayfinding task.

Elaborated wayfinding model, assumes construction of *decision scenes*, which will replace well known from car driver navigation systems decision points. The result of this research is an algorithm allowing automatic data processing for pedestrian navigation system. Through successive phases of preprocessing, binarization, skeletonization and application of Voronoi diagrams, the navigationable graph was obtained. It includes *decision scenes* and is ready for an implementation to the new system.

Streszczenie

Wyszukiwanie drogi (wayfinding) w nieznanym nam środowisku jest częścią naszego codziennego życia, dlatego też mało kto zdaje sobie sprawę z faktu, jak wymagające jest to zadanie. Wymaga ono od nas konkretnych przestrzennych i kognitywnych umiejętności. W celu stworzenia systemu nawigacji dla pieszych, istotne jest poznanie mechanizmów, które kontrolują procesy kognicji. Wraz z rozwojem technologii, znacznie się zmniejszyły rozmiar i waga elektronicznych urządzeń nawigacyjnych. Spadek cen i dostępność takich urządzeń spowodowały wzrost zainteresowania „wszechobecnymi" systemami nawigacji – takimi, które pomogą nam w prawie każdej sytuacji związanej z nawigacją.

Obecnie, rynek jest pełen doskonale działających systemów nawigacji dla kierowców. Są one dobrze dystrybuowane, posiadają wysoki stopień akceptacji użytkowników, a ich rynek z roku na rok wyraźnie wzrasta. Czy w takim razie można zmodyfikować te systemy w taki sposób, aby mogli ich używać piesi? Niestety, problem jest złożony. Badania dowiodły, iż takie systemy nie spełniają wymagań pieszych użytkowników. Zasadnicze różnice polegają na: stopniu swobody przestrzennej, prędkości poruszania i związanej z tym rozdzielczości przestrzeni. Trzeba również pozyskać na nowo dane potrzebne do takiego systemu.

W pracy przedstawiona została baza teoretyczna budowy systemów nawigacji dla pieszych. Nacisk został położony na mechanizmy, które są odpowiedzialne za to, jak każdy z nas postrzega otaczającą go przestrzeń i jakie są nasze umiejętności poznania świata. Przedstawia koncepcję map kognitywnych (cognitive maps) i schematów obrazowych (image schemata), które funkcjonują w naszym umyśle. Ponadto wskazuje, jakie są wymagania oraz kategorie informacji, które ludzie wykorzystują podczas zadania odnajdywania drogi.

Opracowany model wyszukiwania drogi, zakłada powstanie „scen decyzyjnych", które zastąpią punkty decyzyjne – dobrze znane z systemów nawigacji dla kierowców. Wynikiem przeprowadzonych badań jest algorytm pozwalający na automatyczne przetwarzanie danych dla systemu nawigacji dla pieszych. Poprzez kolejne etapy preprocessingu, binaryzacji, wykorzystania algorytmów szkieletyzacji i diagramów Voronoi, otrzymany został nawigowalny graf wraz ze scenami decyzyjnymi, gotowy na implementację do nowego systemu.

Spis treści

1. Introduction

The way people perceive surrounding space has fostered the curiosity of scientists ever since. To develop supporting tools for wayfinding, it is essential to know the mechanisms that control these processes. Cartographic maps were one of those first tools, together with development of satellite navigation. With a progress of technology, the size and weight of electronic devices have diminished significantly. The dropping of prices and wide-ranging availability of such devices increased interest in such systems. They were called “ubiquitous systems” (Rehrl et al. 2007), because of factors like size and availability.

At present, there are a number of different electronic tools available aiding wayfinding:

- Personal Navigation Devices (PND): nowadays they are widespread electronic devices. 95% of such devices are used in cars. The majority of currently working navigation systems for pedestrians is based on these devices. They are equipped with digital maps and provide turn-by-turn instructions via graphical and voice interface.
- Mobile Phone Navigation Systems: software installed on this new generation of mobile phones (also called “Smartphones”) allows to provide street maps and step-by-step instructions. They are more flexible devices, because they combine many services in one “box”.
- GPS maps: portable devices that compound cartographic maps with GPS receiver.
- Mobile guides: allow following a path predefined in advance. The focus is on giving extra information about a route.
- GPS watches: very simple devices in a form of a wrist watch, which allow us tracing a GPS path.
- Off-line systems, so-called “*route-planners*”: they are usually available on websites. They allow us to produce step-by-step instructions, most often as a printed document.

The greatest perspectives for development of electronic tools have Smartphones. Due to their mobility and flexibility they can be used also outside urban areas. They provide a significant possibility for people such as: tourists, bikers, skiers, etc.
The market is full of perfectly working navigation systems for car drivers. They are widespread, have high user-acceptance level and their market is fast-growing (Canalys 2006). Can one modify such systems in a way that pedestrians can use it? The problem is more complex than one might imagine at first sight. Field tests (Rehrl et al. 2007) of GPS maps and car navigation systems show that they do not meet the requirements of pedestrian users. GPS maps are not providing any instructions to the user in general. They only show the user's position on a map. The main problem of navigation systems for car drivers is providing confusing instructions for pedestrian. Car and pedestrian navigation differ in:

- Degree of freedom: car drivers are restricted to the street network (specifically to the lanes). Pedestrians are fairly free in movement. They do not have to consider road traffic principles. They can stop wherever they want to, turn back or move in whatever direction they like. Research has shown that instructions supporting drivers' navigation are mostly useless for pedestrians. For example, the command "*turn left in 200 meters*" has no meaning for pedestrians, as they mostly do not know how much 200 meters exactly is.
- Velocity of movement: pedestrians are moving with a speed that is few times slower than the speed of car. That results in different surrounding environment perception. While car drivers have to concentrate on car driving, the attention of pedestrians is drawn on the things that are enclosing them.
- Spatial resolution: because of our reduced velocity as pedestrians, we can observe surrounding environment in a more detailed way. That is the reason why tools that support our wayfinding have to be more specific. Instructions must be more detailed and semantically rich.

The differences between car drivers and pedestrians are classified when we elaborate a certain human wayfinding model, which could work for pedestrian users. The graph that is used as a simulation of pedestrian navigation is different from the one used in car navigation. In a paper by Gaisbauer and Frank (2008), the model of navigationable space was introduced.

I want to focus on data acquisition. Systems for drivers and for pedestrians differ significantly from each other, in consequence data supplying them will be also different. There are already working systems available, using the same databases as for car drivers. Many prototypes were created, but they need extensive geodata acquisition and preprocessing, and they are only useful for specific scenarios and area. The required data cannot be based on the same datasets used for car navigation systems (Elias 2007, Stark et al. 2007). Therefore, there is a need to acquire data from a source and to keep costs low. Automatization of the process of data acquisition is thus required.

My hypothesis states that there is a possibility of automatic production of navigationable graphs for pedestrian navigation systems based on topographic maps. Based on the wayfinding model provided by Gaisbauer and Frank (2008) I will elaborate an algorithm that allows automatization of this process in stages: preprocessing, binarization, skeletonization, and graph production.

2. State-of-the-art

This chapter introduces basis of navigation systems for car drivers, as well as numerous solutions for pedestrians that are available.

In subchapter 2.1 I will focus on navigation systems for pedestrians, which are currently working or are in prototype state and—moreover—are on some of the developing conceptions.

2.1 Conceptions and working systems

In this subchapter I will describe commercial and non-commercial systems that are widely available. I will refer to conceptions of data acquisition for this kind of system.

2.1.1 BMW Personal Navigator (BPN)

In a paper by Krüger et al. (2004) the authors describe conception and features of a system developed by BMW and DFKI (German Research Center for AI). BPN is a system that combines route-planner, car navigation system, and multi-modal in- and outdoor navigation system for car and pedestrians. Fig. 2.2 shows that the authors have introduced three types of situations, in which navigation systems can be useful.

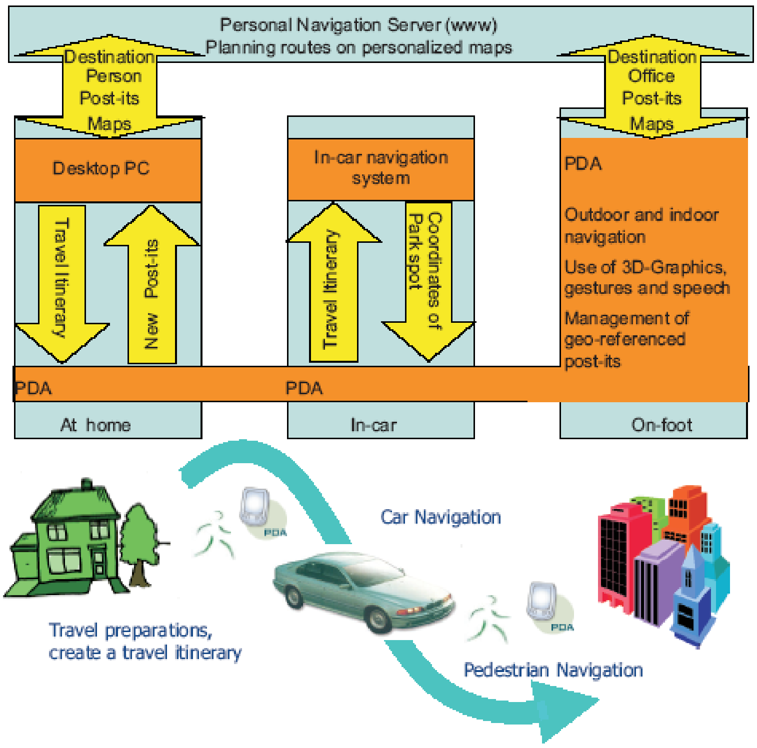


Fig. 2.1: Different situations during a navigational task and the role of PDA as the link between them.

Route finding and cartographic process is based on commercial maps provided by Navtech Navstreets Street Data Premium Germany (ArcView Format) and implemented in an open-source software GRASS. The data essential for navigation inside buildings was modeled manually in software described above. Distinction between navigational and non-navigational areas (line segments) was achieved by adding features (properties) to streets: name of the street and classification. GRASS software enriched with route-planning module was used in order to perform computations and their proper visualization. Designing path networks for the indoor environment is carried out manually, because there is a lack of commonly available databases. The data was acquired using scanning of architectural plans. An improved form of Dijkstra's algorithm is used for finding the shortest path in an indoor path network.

User interface is characterized by high level of multi-modal interaction with user. One can input the data (queries) using speech or speech combined with gestures. The output is expressed in a form of route description represented in a shape of 2-D/3-D visualization and audio segments. The interesting feature is the system's ability to recognize people's speech, used as user commands. Switching between in- and outdoor maps is achieved automatically when the user enters a certain area.

2.1.2 Deep Map

Conception of Deep Map system was elaborated by European Media Laboratory (EML)—Heidelberg, Germany. In a paper by Malaka & Zipf (2000), the fundamentals are described and an example of system application is presented with employment of working prototype.

At present, Deep Map is not a ready-made product, but more of a framework for a future research. In future, Deep Map will be a system able to create guided walks and help tourists to navigate through unfamiliar environments. Of particular interest is the fact that this system will be fully personalized. Despite the fact that Deep Map is a long term project, the existing prototype is successfully managing a major part of tasks. The first versions of such a system are developed to operate in the town of

Heidelberg, Germany. Geographic Information System is the core of Deep Map. User requests that refer to historical and temporal issues are of particular interest.
As a consequence of the system's complexity, object-orientation alone could be unsatisfactory, thus agent-oriented software paradigm was used. It is especially usable, when we have to deal with two levels of application: off-line web-based system for home users and the mobile system for tourists on site.
Deep Map architecture bases on three different layers on which a system is managing Human-Computer interaction. The first one is the interface layer that provides multiple types of input and output. The cognitive layer translates human concepts (input) into system requests. System responses are also translated into human-familiar representation (output). The knowledge layers consist of the GIS, databases, external services, and other systems that provide the knowledge of the contexts and solutions for problems concerning tour planning.
Malaka and Zipf described partly functional prototype and well-developed framework for pedestrian navigation system. They focused mainly on tourist market, which is a great application domain for IT research and interactive systems. Concepts of the high degree of personalization and user-friendliness of Deep Map can help system gain market acceptance for this kind of personal electronic-aid. It is another large step forward in ubiquitous navigation development.

2.1.3 Tailored geodatabase for routing

An interesting system is presented in a paper by Elias (2007). It stresses the fact that solutions already available on the market are still based on the same geodatabases as used in car navigation systems. There is information provided on footpaths and areas accessible for pedestrians. It emphasizes the influence of landmarks in pedestrian navigation, as with their assistance people keep orientating themselves and they are used to communicate route descriptions to other people.
In this case the goal was to develop a method to create such geodatabases described above using geodata sets that already exist and are area-wide. Obtaining data sources is only the first step. Afterwards one may need to analyze the content of provided databases regarding their role in pedestrian navigation. Areas accessible for pedestrians are of interest. The buildings and areas in a city, like streets and plazas, are extracted from a cadastral map.

A particular technique was developed to deal with elements included in indoor plans. The public shopping area and entrances onto train platforms were chosen from these plans, as well as all possible access points to those areas, such as entrances/exits. While different datasets are used, it is possible for some content of the data sets to partly overlap and the same real world objects are modeled repeatedly. For example, in cadastral maps, streets are represented as polygons in contrary to topographic map, where they are shown as vectors. Therefore, datasets had to be singled out, which suits better the needs of pedestrian navigation. It was also required to derive extra data from existing databases. Having such data available one is able to provide necessary navigationable data in form of connected linear graph network.

Elias (2007) shows that providing data sets for pedestrian navigation system may face various problems. The main reason of its imperfection is a lack of detailed data. In this case we were able to use only topographical database, which consisted of street network, which is not important for pedestrians. The issue of open spaces was not raised. This paper shows how many problems we need to face while providing data for pedestrian navigation system. It is also a large guidepost, showing the direction in which our attention should be drawn and which problems still have to be solved.

3. Human wayfinding

This chapter begins with the scientific background and research on spatial cognition, which underlies all processes of human wayfinding. Thus spatial cognition will be introduced first. Different aspects of human wayfinding: importance of landmarks, people's abilities and spatial knowledge, users' information needs and mental representations will also be reviewed.

3.1 Spatial cognition

Human spatial cognition is an active field of research and its role for navigation services is widely acknowledged (Gaisbauer and Frank 2008). It is a part of broader domain called "cognitive science". By its interdisciplinarity, cognitive science contributes from various fields including: psychology, neuroscience, linguistics, philosophy, computer science, anthropology, and physics. It investigates mind, reason, experience, and people's conceptualizations of the world they live in (Lakoff 1987). Cognitive science deals with the study of all facets of human intelligence, from perception and action to language and reasoning. But what does "human intelligence" mean? Pinker (1997, 62) defines it as a rational and humanlike thought, i.e., *"the ability to attain goals in the face of obstacles by means of decisions based on rational (or truth-obeying) rules"*. Other definition is by Wechsler (1944, 3): "*the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal efficiently with his environment"*.

The wayfinding model described in this thesis is based on findings of cognitive research, following a bottom to up approach. It is believed that including such concepts on each design level, will significantly improve the effectiveness and usability of navigation systems, and allow implementation of cognitively adequate user interfaces.

3.1.1 Cognitive views on space

In Geographic Information Systems the term "spatial" is used frequently. Webster's Unabridged Dictionary (1975, s.v. "spatial") defines spatial as "*relating to space; happening or existing in space"*. It further defines "space" as:

[…] distance extending without limit in all directions; that which is thought of as a boundless continuous expanse extending in all directions or in three dimensions, within which all material things are contained.

This definition misses the point of human conceptions of space, thus other sources of definition and terminology need to be considered. Gaile and Willmott (1989; quoted in Mark 1995) introduced this term from geographic perspective: *"place and its dimensions serve as the bases for geographic descriptions and explanations of events".* In this context, places are neither readily recognized as extending in three dimensions and containing all things, nor boundless. This definition describes conception of space shaped and influenced by factors that are not accounted in the previously quoted dictionary definition.

Two "views" of space: "physical" and "experiential" differ from each other and they are organized in a diverse way. The "physical" view of space perceives it as seamless and uniform. Each type of geometry (e.g., Euclidian) is believed to apply to all scales and all phenomena, although it may be recognized that some are valid approximations to geometries at some scales and not others.

The "experiential" view of space is organized differently. Experiential realism elaborated by Lakoff (1987; quoted after Mark 1995), based on the previous works of Rosh (1978; quoted after Mark 1995) and others, claims that cognitive concepts and categories come from the human interaction with surrounding environment. This interaction is straightforward and direct, restricted to size of human body, and at the scale of everyday, manipulable objects. Such objects, and spaces that contain them, are three-dimensional. Their relative locations are typically communicated through object-centered or viewer-centered frames of reference. The term "manipulable space" has been recently introduced by various authors and used to describe these spaces, such as "haptic space", "sensimotor space", "everyday-object space", etc. Downs and Stea (1977; quoted after Mark 1995) called them "small scale" or "perceptual" spaces.

All researchers who recognize more than one type of space acknowledge that geographic spaces are different from the spaces described above. Though we can only interact with them piece by piece, assemble and store them in our minds

through spatial reasoning. It is carried out predominantly by using drawn or printed maps.
Downs and Stea (1977) distinguish two basic types of space: “large scale space” or “geographic space”. They have called them “transperceptual spaces” to emphasize the fact, that they are known by direct integration of perceptual experiences. It may seem that “geographical space” is the best term for such spaces. After all geographic spaces are usually conceptualized as two-dimensional, with vertical height as an additional attribute, or omitted. Relative locations are often expressed by external reference frames, using cardinal directions or remote landmarks.
During last decades, numerous cognitive models of space were introduced. They have also raised a problem of two kinds of space:

- Ittleson’s (1973; quoted after Mark 1995) “object” and “large-scale” space,
- Downs and Stea’s (1977; quoted after Mark 1995) “small-“ and “large-scale” space,
- Mandler’s (1983; quoted after Mark 1995) “small-“ and “large-scale” space,
- Pixtens et. al’s (1983; quoted after Mark 1995) “physical” and “sociogeographic” space,
- Zubin’s (1989; quoted after Mark 1995) “A-“ and “D-“ space,
- Mark’s (1992; quoted after Mark 1995) “haptic” and “transperceptual” space and
- Montello’s (1993; quoted after Mark 1995) “figural” and “environmental” space.

Montello distinguished four major classes of psychological (cognitive) spaces defined in relation to their projective size with reference to the human body: figural, vista, environmental, and geographical. The vista and environmental space are most relevant for navigation tasks. The first one describes the surrounding of a person that can be visually apprehended from a position without locomotion. According to the wayfinding model described in this thesis, i.e., the space that is important for navigation when a pedestrian arrives at a decision point. The area conceived at a street intersection would be classified as a vista space. Environmental space is of large-scale type in which origin-to-destination navigation takes place. This kind of space is too big to be perceived from one position and therefore cannot be apprehended without locomotion.

3.2 Wayfinding

Wayfinding is one of our daily routine tasks. Finding a way from A to B is more complex and demanding task that one might think. It encompasses all the ways people orientate themselves in physical space and navigate from one place to another. Wayfinding requires extensive spatial knowledge and cognitive abilities. The term *"navigation"*, which is often connected with *"wayfinding"* is described as (Bowditch 2002, Wikipedia): *"the process of planning, recording and controlling the movement of a craft of vehicle from one place to another"*. Montello (2005, 258) refers to the term of navigation in a more particular way, which is of concern to presented research. He argues that it is *"coordinated and goal-directed movement of one' self (one's body) through the environment"*. He defines wayfinding as *"goal-directed and planned movement of one's body around an environment in an efficient way"* and introduces the definition of locomotion as the *"movement of one's body around an environment, coordinated specifically to the local or proximal surrounds"*. Referring to the research on human wayfinding it is essential to consider those two relevant cognitive processes.

There are some other definitions of the "wayfinding" term. Golledge (1999) explains wayfinding as the process of determining and following a path or route between an origin and a destination, which is a purposive, directed, and motivated activity. For the successful travel, it is important to identify start and endpoint, to determine turn angles, to identify segment lengths and directions of movement, to recognize route and distant landmarks, and to embed the route to be taken in some larger reference frame. If a destination is known but it is not directly connected by a path, road, or track to the origin, successful travel may involve search and exploration, use of landmarks, spatial updating of one's location, recognition of segment length and sequencing, identification of a frame of reference, and mental trigonometry (e.g., triangulation, dead reckoning).

The work of Lynch (1960) is one of the earliest and most influential contributions in the field of human wayfinding with regard to the cognitive structure of urban environments. The author gives an account of a research project, carried out in three American cities (Los Angeles, Boston, and Jersey City with comparisons to Florence and Venice). The project resulted in the evolution of the *"legibility"* concept depending on people's *"mental maps".*

Legibility is a term used to describe "*the ease with which people can understand the layout of a place*". By making questionnaire surveys, Lynch defined a method of analyzing legibility based on five elements. We define these as follows (Lynch 1960, 47):

1. *Paths*

Paths are the channels along which the observer customarily, occasionally, or potentially moves. They may be streets, walkways, transit lines, canals, railroads. For many people, these are predominant elements in their image. People observe the city while moving through it, and along these paths the other environmental elements are arranged and related. These are the major and minor routes of circulation that people use to move on. A city has a network of major routes and a neighborhood of minor routes.

2. *Edges*

Edges are the linear elements not used or considered as paths by the observer. They are the boundaries between two phases, linear breaks in continuity: shores, railroad cuts, edges of development, walls. They are lateral references rather than coordinate axes. Such edges may be barriers, more or less penetrable, which close one region off from another; or they may be seams, lines along which two regions are related and joined together. These edge elements, although probably not as dominant as paths, are for many people important organizing features, particularly in the role of holding together generalized areas, as in the outline of a city by water or wall.

3. *Districts*

Districts are the medium-to-large sections of the city, conceived of as having two-dimensional extent, which the observer mentally enters "inside of", and which are recognizable as having some common, identifying character. Always identifiable from the inside, they are also used for exterior reference if visible from the outside. Most people structure their city to some extent in this way, with individual differences as to whether paths or districts are the dominant elements. It seems to depend not only upon the individual but also upon the given city.

4. Nodes

Nodes are points, the strategic spots in a city into which an observer can enter, and which are the intensive foci to and from which he is traveling. They may be primarily junctions, places of a break in transportation, a crossing or convergence of paths, moments of shift from one structure to another. Or the nodes may be simply concentrations, which gain their importance from being the condensation of some use or physical character, as a street-corner hangout or an enclosed square. Some of these concentration nodes are the focus and epitome of a district, over which their influence radiates and of which they stand as a symbol. They may be called cores. Many nodes, of course, partake of the nature of both junctions and concentrations. The concept of node is related to the concept of path, since junctions are typically the intensive foci of districts, their polarizing center. In any event, some nodal points are to be found in almost every image, and in certain cases they may be the dominant feature.

5. Landmarks

Landmarks are another type of point-reference, but in this case the observer does not enter within them, they are external. They are usually a rather simply defined physical object: building, sign, store, or mountain. Their use involves the singling out of one element from a host of possibilities. Some landmarks are distant ones, typically seen from many angles and distances, over the tops of smaller elements, and used as radial references. They may be within the city or at such, a distance that for all practical purposes they symbolize a constant direction. Such are isolated towers, golden domes, and great hills. Even a mobile point, like the sun, whose motion is sufficiently slow and regular, may be employed. Other landmarks are primarily local, being visible only in restricted localities and from certain approaches. These are the innumerable signs, which fill in the image of most observers. They are frequently used clues of identity and even of structure, and seem to be increasingly relied upon as a journey becomes more and more familiar.

3.2.1 People's wayfinding abilities

A navigationable graph is used to simulate wayfinding. Car navigation uses different graph that simulates driver's behavior. They are expected to have certain spatial abilities. Before designing navigationable graph to simulate pedestrian wayfinding behavior we have to explain what the abilities that people incorporate during a task of wayfinding are.

Recent work of Allen (1999; quoted after Raubal 2001) suggested grouping people's spatial abilities according to their function, that is, to the tasks and situations in which they are applied. Such classifications are based on previous research in the psychometric information-processing, developmental and neuropsychology traditions. It consists of interaction between:

- A stationary observer and small manipulable objects,
- An observer and moving objects,
- A mobile observer and large stationary objects.

Allen claimed that people's spatial abilities depend mainly on the following four interactive resources: perceptual capabilities, fundamental information-processing capabilities, motor capabilities, and previously acquired knowledge, also called by Norman (1988; quoted after Raubal 2001) "the knowledge in the head". In Figure 3.1, the resources supporting different wayfinding means are shown. Because each person's wayfinding abilities are different (i.e., some people are better wayfinders than others), such a model could function as a framework for investigating individual wayfinding abilities.

Cognitive abilities, just like the spatial abilities also depend on the task and the context. Montello (2005, 261) described them as "*our abilities to perceive, remember and reason in space and place*". It is evident that wayfinding in a street network (Timpf et al. 1992, Car 1996; quoted after Raubal 2001) uses a different set of cognitive abilities than navigating from one room to another in a building, because these two environments differ significantly from one another (i.e., in a scale aspect). People are usually good in applying their individual skills to the task at hand. If their spatial skills are weak, they use verbal skills to navigate. When people get lost, they usually ask someone for help—and vice versa (Vanetti and Allen 1988; quoted after Raubal 2001).

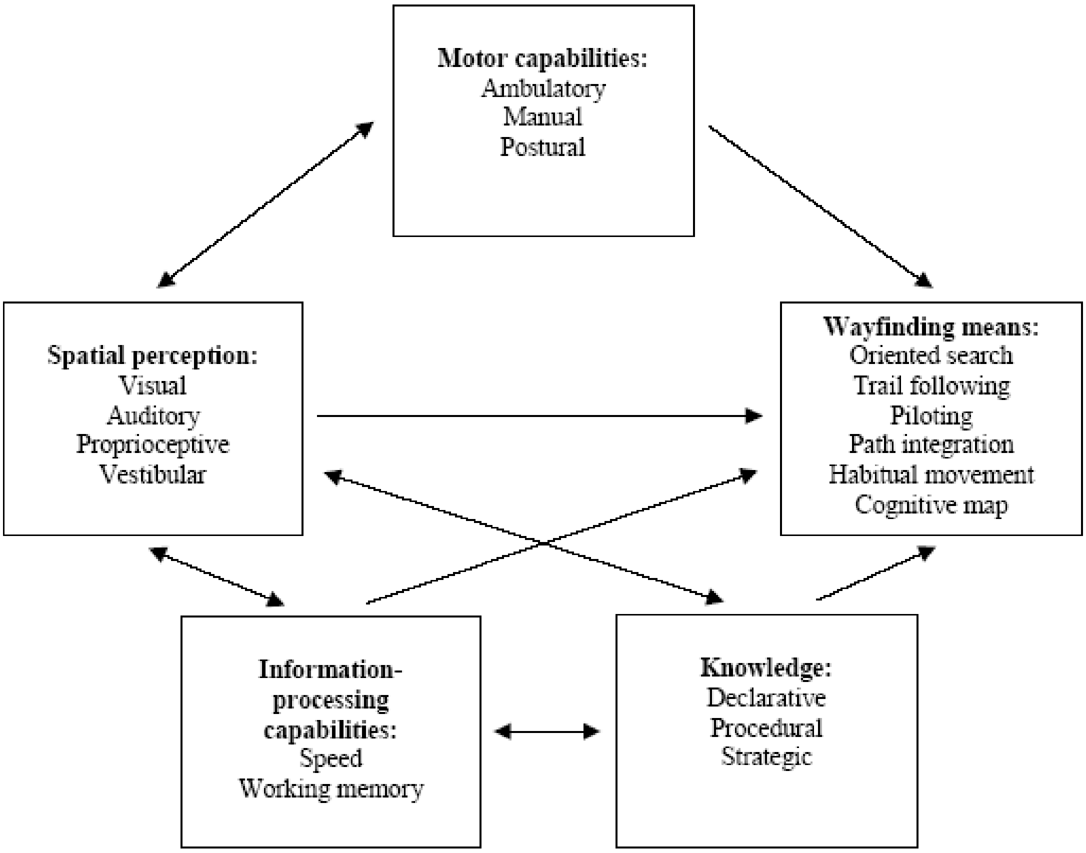


Fig. 3.1: Speculative model of relations among interactive resources and wayfinding means (Allen 1999, 79)

Human spatial knowledge, according to Siegel and White (1975; quoted after Raubal 2001) is assumed to develop in three successive stages:

- Landmark knowledge compromises salient points of reference in the environment,
- Route knowledge puts landmarks into a sequence,
- Survey of configurational knowledge allows people to locate landmarks and routes within a general frame of reference.

Because of its strict developmental sequence, this model has been recently criticized. Montello (1998; quoted after Raubal 2001) proposed a new framework for people's spatial knowledge acquisition in large scale environments. He believed that stages with pure landmark or route knowledge do not exist. He argues that this is true, as we never obtain 100% pure knowledge. There are no qualitative shifts from non-metric to metric forms of knowledge, because metric knowledge is obtained right from the beginning of the acquisition process, accumulated and then refined.

3.3 People's information needs

3.3.1 Personalization aspects

To design a successful pedestrian service there is a need of knowing what a future user of such system will demand. Such pedestrian aids need to be perfectly tailored for their demands. We have to be able to know people's information needs and understand the nature of the navigation task in such a way that will allow us to design a successful pedestrian service.

Guidance of pedestrians requires focusing on personalization aspect (Corona and Winter 2001). According to user tasks and information needs, they can be subdivided into groups, like: tourist, business traveler, resident, disabled person, etc. If the user links to a group, special preferences can be determined. For example, the service will create the best tailored route to the user's needs: the route with shortest distance, least time, fewest turns, or the most scenic one.

3.3.2 Identifying essential information

May et al. (2003) carried out empirical research, in which participants were given a series of complex pedestrian routes. They were asked to identify the information that they felt a pedestrian, unfamiliar with the area, would need in order to navigate these routes successfully. These navigation instructions were elicited from two separate participant groups: either from their memories (the cognitive map group), or based on the participants physically walking through the routes (the walkthrough group). In order to identify relevant navigation cues, the cognitive map group used route schematics, which were designed to provide just enough information to enable participants to recognize the routes without incorporating any potential navigation information, such as road names, buildings, other landmarks, or any indication of distance. On the contrary, the walkthrough group based their navigation instructions on being physically led through each of the chosen routes, i.e., the information acquired and identified by this group simply based on direct observation and experience, rather than the recall of these cues from the memory. Authors have distinguished these two groups, because they believed that the best environmental cues were those which were pertinent in participant's cognitive maps and reflected the importance of recognized mental representations of spatial environments, which was affirmed by Kitchen and Blades (2001; quoted after May et al. 2003), and Christon and Bulthoff (2000; quoted after May et al. 2003). The reason for authors to employ research on the walkthrough group was that apart from environmental cues lying in participant's cognitive maps, those which are usually prominent (i.e., which would depend highly on their perceptual-visual characteristics and thus would support an information processing perspective) are also important. These two different approaches had associated advantages and disadvantages, which are shown in Table 3.1.

Information based on:	Advantages:	Disadvantages:
Long-term memory (a cognitive map)	Based on repeated exposure to cues—information used by "expert" navigators	Individual's memory for navigation cues prone to subjective biases
Direct experience (walkthrough)	Based on direct observation: the view of an unfamiliar traveller	Potential inconsistencies between the experience of routes: limited by the specific views available, the time of day, etc.

Table 3.1: A comparison of different information sources for direction-giving studies (May et al., 2003, 332)

3.3.3 Information categories

To acquire information needs, it was essential to incorporate five top level categories of information required for pedestrian navigation (Table 3.2), i.e., which could be presented by mobile navigation devices:

1. distance
2. junction
3. road type
4. street name/number
5. landmarks

General category	Specific category	Example
Distance (qualitative or quantitative)	N/A	"Turn left in 100m" "Continue for some way"
Junction	Road junction Pedestrian "junction"	"Turn right at the T junction" "Go left where the path forks"
Landmark	40 sub-categories including: bank, bridge, car park, library, steps, tree	"Go past the NatWest bank"
Road type	Road/street Path (pedestrian) Pedestrianised area	"Follow the main road" "Go down the path..."
Street name/number	N/A	"Turn left down Market Place"

Table 3.2: Main information categories (May et al. 2003, 333)

This division was based on a pilot trial and cross-checked against other navigation information taxonomies, such as that used by Burnett (1998).

According to Table 3.2, results of this research are shown on Figure 3.2. It can be clearly seen that for research participants' landmarks are the most often related category.

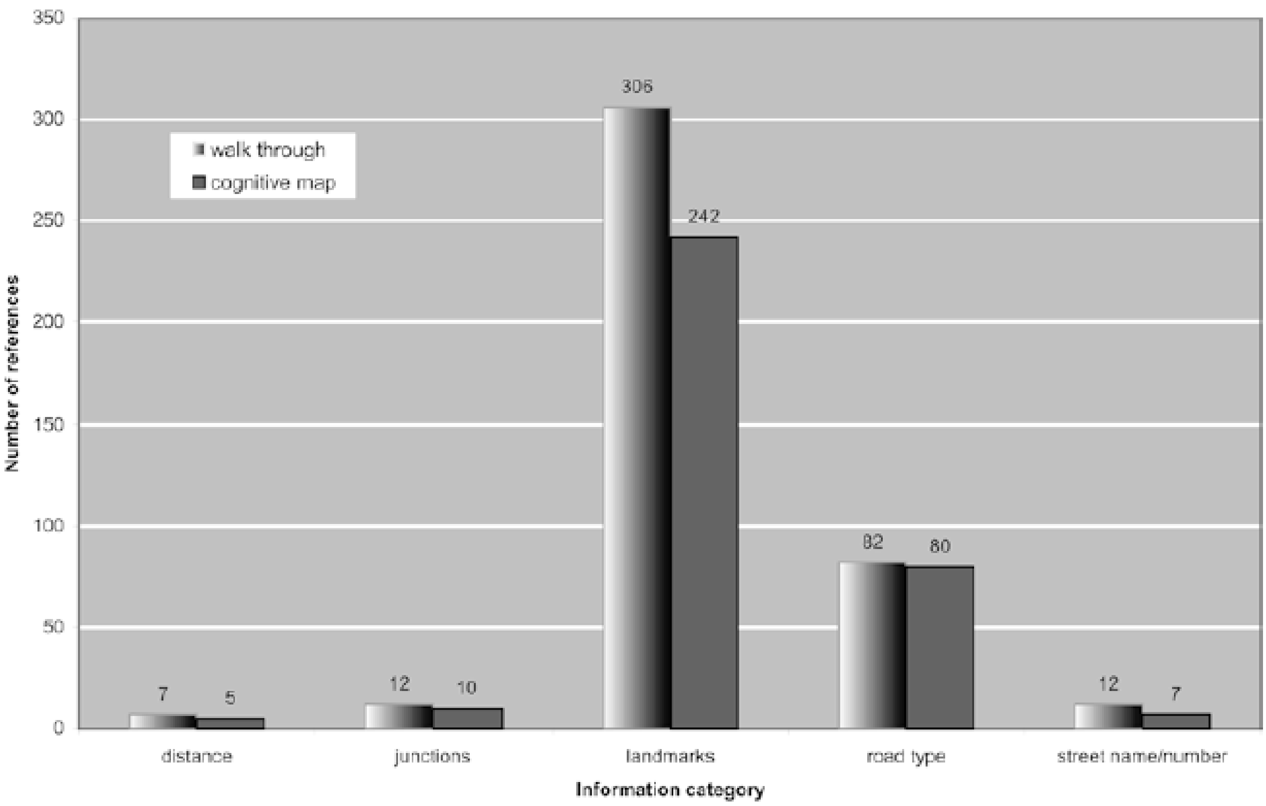


Fig. 3.2 : The total frequency of reference to general information categories by the walkthrough group and the map group participants (May et al. 2003, 335)

After landmarks, the next most frequently navigation cues used in this study were references to roads, pathways, streets or blacktop areas, i.e., information that describes the form or visual appearance of the pathway that the pedestrian was on, or was required to turn into (i.e., the "route" distinction of Hirtle and Jonides (1985; quoted after May et al. 2003)). The other categories of information were considerably seldom used as navigation cues. Another result that is intriguing showed that humans have difficulty in judging distances, and specifically the problems of mapping actual experience or visual representations, or routes onto distance judgments. Highlighting the potential danger of relying on distance information within pedestrian navigation devices was not only done in May's et al. (2003) research, but also proven before (Corona and Winter 2001).

Results showed that almost 2/3 of information is given at nodes, and the rest of them are given along the route. They emphasize that a pedestrian is not simply directed from one navigation decision point to another, but he also requires information between those decision points in order to maintain his trust in the information source. His confidence and orientation throughout the route is underlining the continuous dynamic nature of navigation task (Kitchen and Blades 2001; quoted after May et al. 2003).

Figure 3.3 shows how different categories were used on a route. It can be clearly seen that the majority of information hints were landmarks being used to identify a point on a route and to confirm that correct navigation decision has been made. This is due to the progress of a pedestrian, that he does not need this information; in contrast to car drivers, who use this information to manage vehicle speed and correct lane positioning (May et al. 2001).

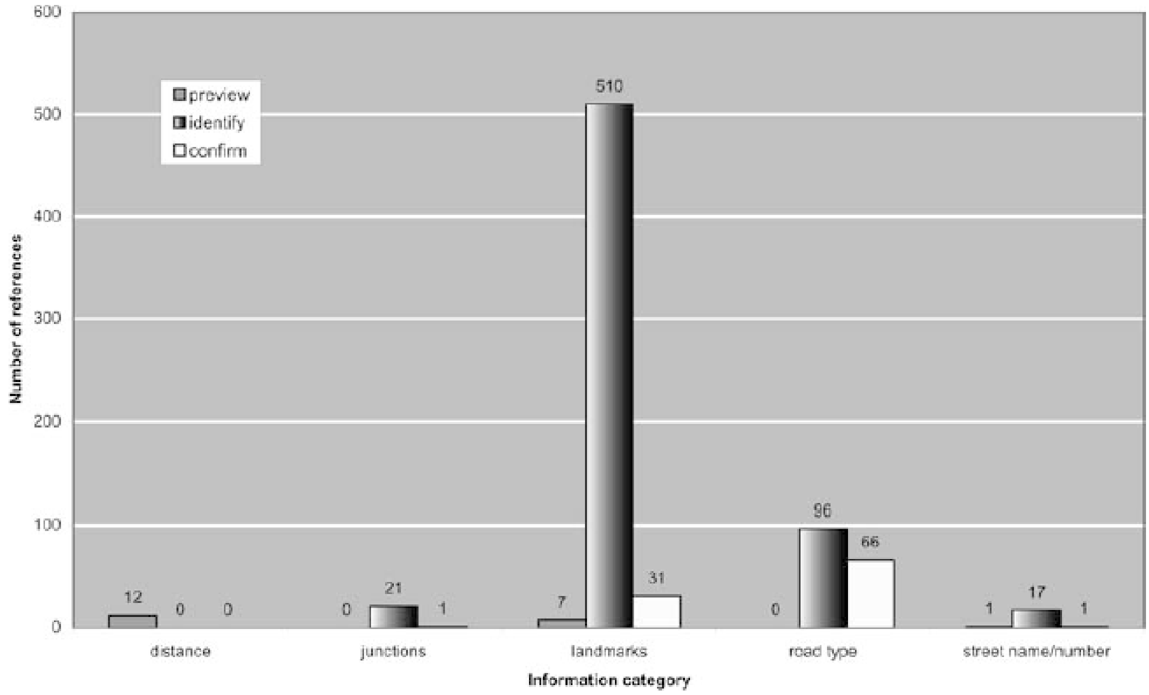


Fig 3.3.: The use of information for preview, identify or confirm purposes (May et al, 2003, 336)

3.3.4 Importance of landmarks

Landmarks are the significant element of numerous pedestrian navigation systems. The "landmark-based" navigation systems do not need positioning from a satellite or ground sensors. By providing people with step-by-step instructions they are able to guide users through unfamiliar environments. The instructions are mostly based on landmarks, which people can clearly notice, because they are the elements that are relatively easy to perceive. By their simple form they are easy to recognize and, i.e., we do not need "knowledge in the world" acquisition, as the patterns of majority of landmarks (buildings, towers, mountains, street lights, restaurants, etc.) are in our "knowledge in the head". Importance of landmarks in pedestrian navigation systems is shown by Goodman et al. (2005), May et al. (2003), Millonig and Schechtner (2007).

Landmarks are consistent with basic human navigational strategies (Burnett 2000). Studies carried out within the field of environmental psychology and human geography has indicated that landmarks are core components of peoples' mental representation of large-scale space, commonly termed as the "cognitive maps". In the cognitive science literature, there is a general agreement that landmarks are very important for acquisition and organization of our knowledge about surrounding space and play an important role in the environmental learning process (Evans et al. 1984; quoted after Burnett 2000). The model of spatial knowledge acquisition provided by Siegel and White (1975) sets landmarks to be the first kind of spatial knowledge acquired in a new environment. What is interesting, people recognize landmarks as places they have been before. Therefore, it is not surprising to find that they are widely used in traditional wayfinding strategies, such as a detail of giving direction that people provide to others (Alm 1990; quoted after Burnett 2000).

It was shown in numerous studies, that landmarks seriously increase wayfinding effectiveness. The directions (verbal or textual), that are given using references to landmarks are far more meaningful and useful. In addition to this, studies by Burnett (2000) showed that landmarks are a significant feature of directions that we request or value from others. It has been shown in a survey that landmarks are the second most important information type valued by drivers. The information that was most worthwhile for them were the “left-right” directions (Fig. 3.4).

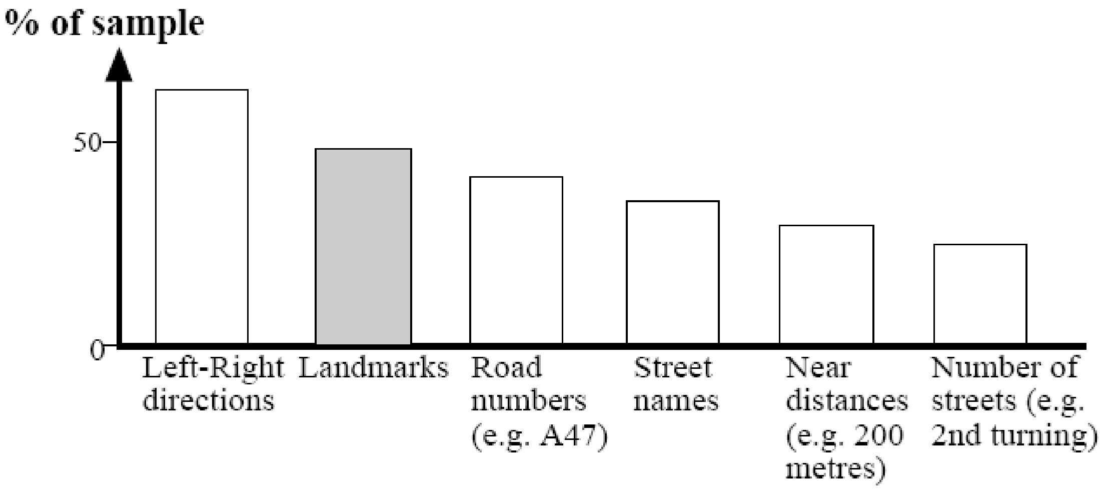


Fig. 3.4: Percentage of sample who requested particular information from the passenger to help them locate turnings (After Burns 1997)

The effectiveness of landmarks in pedestrian navigation systems was confirmed by Goodman (2005) and shown in a field study. By implementing landmarks in a pedestrian navigation system they increase efficiency in wayfinding and decrease mental and physical workload. Participants of Goodman’s research confirmed that landmarks gave visual identification or validation of location selection decision on the route. His studies showed that various modalities can be used effectively to provide information about landmarks, including photographs, text, and speech.

3.3.5 Ontology differences

Corona and Winter (2001) have chosen a different way for gathering knowledge about the user's information needs. The information requirements of car drivers are better known due to existence of mature car navigation systems and that ontology of guiding a car driver is derived from a comparative investigation of existing navigation systems, Corona and Winter suggested to acquire another ontology, one of guiding a pedestrian.

Considering these two ontologies, some commonalities and differences could occur. As it can be seen in Fig. 3.5, the elements of two ontologies belong to one of three sets: objects and actions relevant for car drivers only, objects and actions relevant for pedestrians only, and objects and actions in both ontologies. Regarding elements that are common, they represent the core of navigation systems in general. Concepts like route, map, and list of instructions are common for both of ontologies. Different kinds of destinations, landmarks and attributes can be represented in the overlap. The street network appears to be a common element, but it is not due to the different granularity. For car drivers the "world" is a linear network, but for pedestrians streets are sometimes two-dimensional spaces, such as places or parks. On the other hand, while considering differences, it appears that time indication does not interest car drivers, since they cannot measure distances in order of time. Other than pedestrians, car drivers do not travel with a constant speed, so it is not possible to estimate precisely their time of arrival. Inside the list of instructions a step for a car driver represents a segment of a route, whereas for pedestrian a unit instruction can also represent a reference indication or the position of a landmark (Gero 1999; quoted after Corona and Winter 2001). While driving the vehicle we usually care about the odometer and distance traveled so far. Hardly any car driver measures the distance traveled in a function of time.

These differences influence the various concepts of distance in pedestrian and car driver reasoning (Fig. 3.5). For a navigation system for pedestrians attention has to be paid to what users can cognize on their path or route: nameplates, stairs, and height of buildings. This information should also be added to pedestrian navigation systems (Dieberger 1999; Tedeschi 2000; quoted after Corona and Winter 2001).

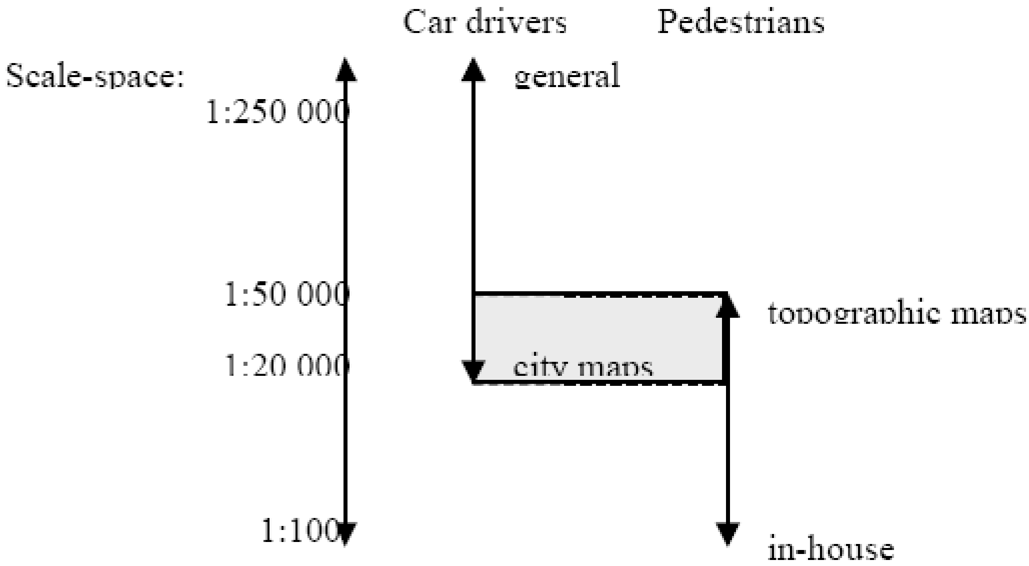


Fig. 3.5.: Scale-space representation both the ontologies and field of intersection (Corona and Winter 2001)

The following is a dilemma facing the future pedestrian service designer. Geographic distances are relatively easy to calculate, using location technologies. Street names are readily available via existing map databases and relatively stable. Thus the information quality associated with street names will have relatively slow decay rate over time, but it also depends on the urbanistic development of the area. Other types of information, such as: built features, pedestrian crosswalks and traffic lights present different difficulties due to the diversity of these data types and the current lack of standardized and centralized data sources.

3.4 Cognitive maps

Most of us travel everyday through the same, familiar environment, such as commuting from home to work. During this daily basis we use internal representations that persist in our memory. For some people navigating only one time through the same environment is enough to reach the same goal again. A useful metaphor made by Kuipers (1982) suggests that people have a cognitive map in their heads, which is some kind of mental representation that matches with people's perceptions of the real world. The term "cognitive map" was first introduced by Tolman (1948; quoted after Raubal 2001) in his paper concerning behavior of rats in a maze—learning task that required acquisition of spatial relations between start and goal.

Although this term should not be taken as anything more than a metaphor, there are still some other definitions, such as “cognitive collage” (Tversky 1993) or “cognitive atlas” (Hirtle 1998; quoted after Raubal 2001). Neisser (1976; quoted after Raubal 2001) uses the term “orienting schema” as a synonym for cognitive map. He stresses its information-seeking and active structure instead of defining it as a mental image. Downs and Stea (1973) referred to the cognitive map as “*an incomplete segmented, metrically distorted representation, with relational accuracy, both spatially and non-spatially varying within and across regions*”.

When one’s familiarity with an area increases, the use of the internal cognitive map gradually replaces the reliance on eternal sources of information (Golledge 1999). Traveling through an environment is commonly recognized as the most frequent way humans acquire spatial knowledge and store it as “the knowledge in the head”. Environments that are complex and difficult to navigate can lead to slower development of our cognitive maps, and may cause some representational errors or inaccuracies. Our cognitive maps are developing from a landmark map to a mental route map and should eventually result in a mental survey map. The last form of cognitive map is closest to the normal cartographic map although it is not metric and still contains inaccuracies and distortions. Davis (1990; quoted after Raubal 2001) pointed out main differences between cognitive and cartographic maps: a cognitive map may consist of different knowledge structures and it has to integrate incomplete, imprecise, and subjective knowledge. Cartographic representations are representing information only pictorially, and are reasonably accurate and complete. People construct and develop their cognitive maps based on their previous experiences: perception, natural language, and inferences.

Golledge partitioned the acquisition and utilization of geometric entities, like points, lines, areas, and surfaces, within a cognitive map into qualitative and quantitative terms. Qualitatively, information pertaining to, for instance, topological relations, inclusion/exclusion and order, can be extracted. When assessed quantitatively, these entities provide a basis for metrical geometric and trigonometric manipulation of the information. Golledge suggested that if cognitive maps were cartographically founded, map reading skills would be required prior to the use of cognitive maps (Golledge, 1999). However, studies have shown that skill of map reading has to be acquired well during child development, after the child exhibited use of a cognitive map (Downs and Liben 1985; quoted after Krafft 2001).

Researches from different fields of science have deeply investigated the role that cognitive maps play in spatial behavior, spatial problem solving, acquisition and learning (Kitchen 1994). What is remarkable is that much less has been found on about how people immediately understand different spatial situations while performing a wayfinding task. Gluck (1991; quoted after Raubal 2001) drew attention to this information deficiency by claiming that work done previously in the field of wayfinding focused on the description of cognitive maps and neglected affective and logistical concerns in most of the cases. He suggested exploration of our information needs and proposed the use of the "sense-making" investigation method, which he describes as a *"creative human process of understanding the world at a particular point in the time and space limited by our physiological capacities, our present, past and future"* (Gluck 1991, 117). Such understanding can be seen as a snapshot of common-sense knowledge of space and time. Therefore, it is important for the process of common-sense spatial reasoning. The idea behind the sense-making is to look at the wayfinding process itself instead of looking at the representation (i.e., the cognitive map) (Raubal 1997).

4. Image schemata

In this chapter I will focus on the concept of image schemata. I will introduce definitions, explain their meaning and show some examples and applications in the field of wayfinding. Generally, image schemata are concepts used by people to understand space (Johnson 1987; quoted after Raubal 1997). By using elements of human perception and cognition we are able to generate some spatial representation, which fits better with people's real-world spatial interactions. Those representations create the basis for spatial information and systems designed for simulation of real-world applications, such as wayfinding tasks.

4.1 Definitions

The term "image schemata" was first introduced by Johnson (1987) in his book titled "The Body in the Mind". He believed image schemata describe everyone's conceptualization of physical reality. They are a part of a philosophy believing that only when we take our bodily experiences into consideration, our mind could be understood. When speaking of bodily experiences we have multi-modal experiences on our mind; not visual experiences. Johnson defines an image schema as:

[…] a recurring, dynamic pattern of our perceptual interactions and motor programs that gives coherence and structure to our experience. In order for us to have meaningful, connected experiences that we can comprehend and reason about there must be a patterned order to our actions, perceptions and conceptualizations.

These recurring mental patterns are a direct consequence of our physical and cognitive abilities as well as our world's experience. They are more abstract than concrete images in the mind and they are believed to be independent of our cultural affiliation or other variances between individuals. It is believed that they allow us to establish connections between our experiences that have the same internal structure in common (Gaisbauer and Frank 2008). Johnson (1987) and Lakoff (1987) described an extensive list of image schemata (Table 4.1) and they identified seven schemata as being spatial in nature: container, surface, near-far, verticality, path, link and center-periphery.

CONTAINER	BALANCE	COMPULSION
BLOCKAGE	COUNTERFORCE	RESTRAINT REMOVAL
ENABLEMENT	ATTRACTION	MASS-COUNT
PATH	LINK	CENTER-PERIPHERY
CYCLE	NEAR-FAR	SCALE
PART-WHOLE	MERGING	SPLITTING
FULL-EMPTY	MATCHING	SUPERIMPOSITION
ITERATION	CONTACT	PROCESS
SURFACE	OBJECT	COLLECTION

Table 4.1: List of image schemata (Johnson 1987, 126)

4.2 Examples

Image schemata are dynamic embodied patterns, which take place *in* and *through* time. Those patterns of experience are not simply visual, but also multi-modal. As an example, let us think how the dynamic nature of the *containment* schema is reflected in the various spatial senses of the English word "out". The graphic description of this schema is shown in Figure 4.2 This word may be used in cases where a clearly defined trajectory (TR) leaves a spatially bounded landmark as in expressions (Wikipedia; Johnson 1987):

(1a) John went *out* of the room.
(1b) Mary got *out* of the car.
(1c) Spot jumped *out* of the pen.

In the most prototypical case, the landmark is a clearly defined container. However, *out* may also be used to indicate cases where the trajectory is a mass that spreads out, effectively expanding the area of the containing landmark:

(2a) She poured *out* the beans.
(2b) Roll *out* the carpet.
(2c) Send *out* the troops.

Finally, *out* is also often used to describe motion along a linear path where the containing landmark is implied and not defined at all:

(3) The train started *out* for Chicago.

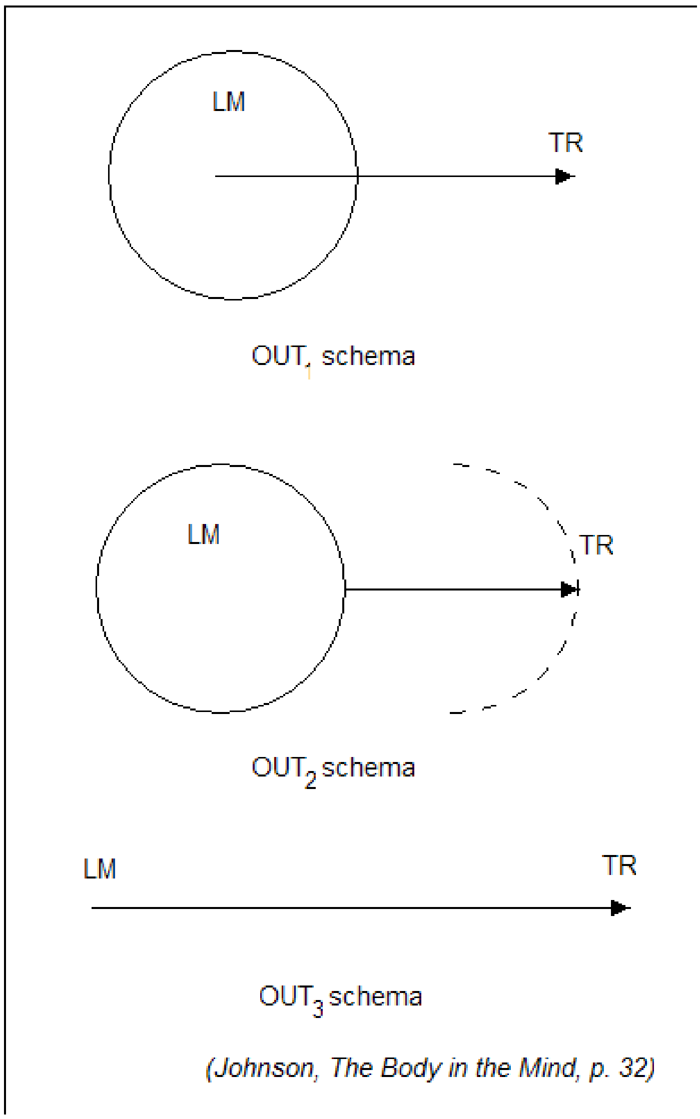


Fig. 4.2: The "out" schema

4.3 Applications

There are multiple applications of image schemata. Since many of them have some spatial characteristics, they have a prominent field of application in the spatial sciences. Image schemata have already been used to model the physical environment of humans (Gaisbauer and Frank 2008). For example, Raubal and Rüetschi and Timpf (2004, 2005) have used them to model the domains generally navigated by pedestrians, such as airports and train stations. That is the reason why we employ the image schemata to the wayfinding model of an outdoor navigation system, which will be described in detail in the next chapter. This wayfinding model currently relies on the *container* and the *link* schema to connect our wayfinding model to cognitive needs of users. They are considered to design our wayfinding model, because they structure the pedestrian environment on the vista-space level of granularity (Gaisbauer and Frank 2008).

- *Container schema*: The container schema represents containment. It consists of an inside, an outside, and a boundary separating them. This schema is regularly used when entering or leaving a building.
- *Link schema*: The link schema generally represents a connection between objects as well as non-physical linkage. Non-physical links can be the relation to siblings and parents or the direct visual connection from spectator to object.

5. Wayfinding model

In this chapter I will introduce the core of developed pedestrian navigation system: the wayfinding model. Assumptions included in this model are a basis for a design of data acquisition and automation of this task. In the first subchapter I will shortly call attention to differences between pedestrians and car driver information needs. Then I will define the wayfinding model. In the third subchapter I will explain how spatial cognition and image schemata can be employed to work in this model and introduce some graph basis that is believed to be useful. The last subchapter describes the details of wayfinding model. This chapter is based on and strongly influenced by Gaisbauer and Frank's (2008) paper, because their contribution was a foundation for this thesis.

5.1 Different information needs

In this subchapter I will briefly recall what was previously mentioned, namely, that environmental affordances and user requirements concerning pedestrians differ significantly compared for those of car drivers.

Stark et al. (2007) and Walter et al. (2006) focused their attention on the fact that pedestrians are not constrained to the street network as car drivers are. They face different problems and need different cognitive resources, because they are moving with a higher degree of freedom. As they move with significantly lower speed compared to car drivers, they have more cognitive resources available. They do not have to use them for paying attention to the surrounding traffic. Concerning pedestrian's degree of freedom, it is obvious that the normal, conventional graph of traffic lanes does not reflect their unconstrained movement in open space. For example, consider an open space area, such as Stephansdom (Fig. 5.1). If one wants to cross this area, the most obvious behavior would be to directly head towards an "exit" of this open area. This type of reaction cannot be supported by available navigation systems, where the whole area would be reduced to one single decision point. This solution can cause distortion of physical distance and allow the device to propose routes that are in fact detours (Gaisbauer and Frank 2008). We will sketch a solution that fits this obstacle in the last subchapter.

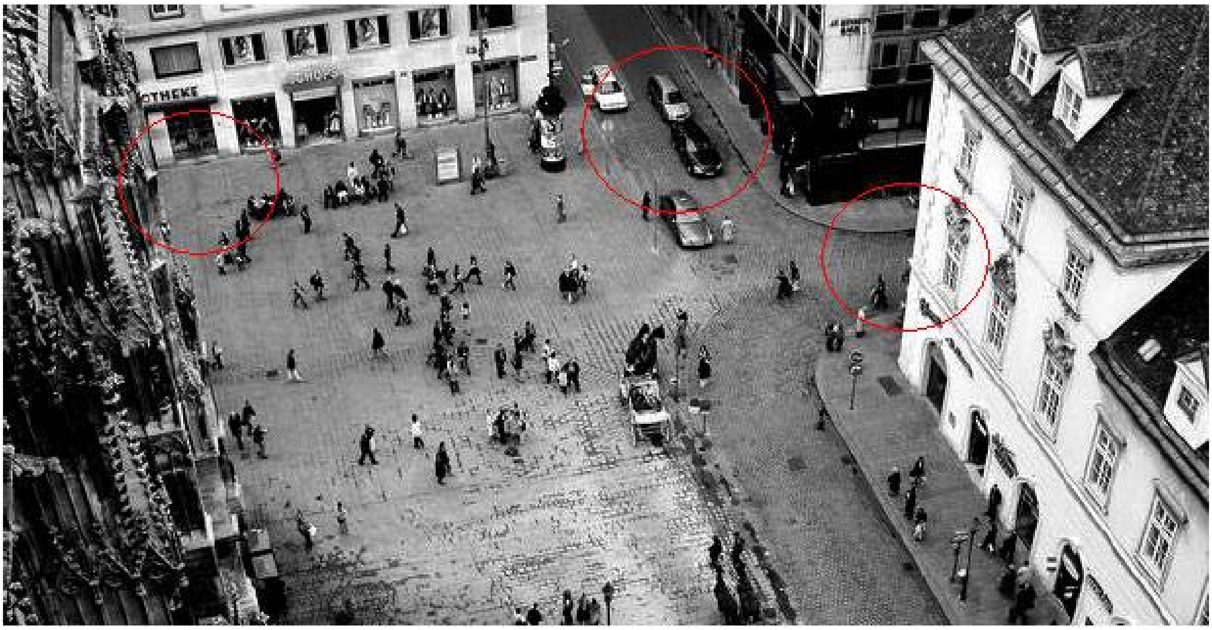


Fig. 5.1: A part of Stefansplatz in Vienna. "Exits" are marked in red. (Gaisbauer and Frank 2008, 2)

5.2 Defining wayfinding model

Going back to chapter 3, where the issue of wayfinding was introduced, we will shortly bring back Montello's (2005) definition of wayfinding as the "*coordinated and goal-directed movement of one's self (one's body) through the environment*". In this process, wayfinding represents the "*goal-oriented and planned movement of one's body around an environment in an efficient way*".

Fig. 5.2 explains the principles of wayfinding model operation. Pedestrians have certain capabilities that allow them to interact with the environment. We are able to perform some physical operations, we can also act upon this environment and to receive a sensory input in a recurring loop, which is shown as collection of the "knowledge in-the-world" and storing it as the "knowledge-in-the-head". This sort of interaction requires structuring the knowledge in a way that makes it easy to be mentally available, therefore we require a cognitive representation of this environment. Such mental representation allows us not only to define the walkable space, but also its internal semantics (Gaisbauer and Frank 2008).

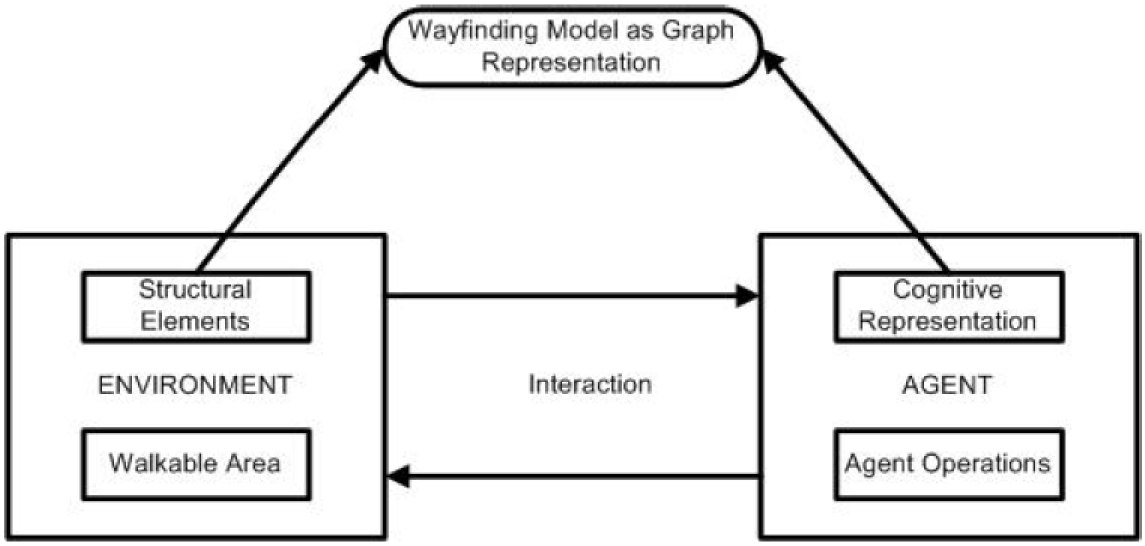


Fig. 5.2: The wayfinding model (Gaisbauer and Frank 2008, 3)

5.3 Role of spatial cognition and image schemata

<u>Spatial cognition</u>

Designing the wayfinding model, space that is surrounding us, and the way that people perceive and conceptualize it needs to be considered. Over the years, several cognitive models of space have been proposed. They were mentioned in chapter 3.1.1. In chapter 3.2 we introduced the one of the most influential contribution in the field of wayfinding – Lynch's book titled "The Image of the City". Lynch distinguished and put in order elements of the urban structure: landmarks, nodes, paths, and regions.

Montello (1993) distinguished four major classes of psychological spaces arranged by their projective size in the relation to the human body: figural, vista, environmental, and geographical. We claim that vista and environmental space are most relevant for the most common navigation tasks and therefore can be used in our wayfinding model (Gaisbauer and Frank 2008).

<u>Image schemata</u>

We devoted the whole chapter for explaining the issue of image schemata. We employed them for our wayfinding model of an outdoor navigation system and at the current state, the wayfinding model relies on two schemata: *container* and *link*. It connects cognitive needs of users to our model. These schemata structure the environment of pedestrians on the vista-space level of spatial resolution (Gaisbauer and Frank 2008).

5.4 Details of the wayfinding model

In the beginning, it is important to know which data is necessary for the envisioned navigation system. First, we have to define which space we can distinguish as the navigationable one for pedestrians. *Walkable space* is composed of sidewalks, pedestrian zones and park areas. We introduce passages such as crosswalks and underpasses, and blocking areas, such as traffic lanes and buildings.

5.4.1 Decision scenes

Lynch (1960) defines decision points as "*…strategic foci into which the observer can enter, typically either junctions of paths, or concentrations of some characteristics*". He pointed out that despite the fact that decision points are conceptualized as nodes, they may represent a large spatial area that is internally structural. It is believed that this internal structure of decision points, although negligible for car navigation, must not be neglected for pedestrian navigation systems (Gaisbauer and Frank 2008). Therefore, an aggregation of the vista space around the decision point to a simple node is an oversimplification of the environment and will not represent the choices of shortcuts that can be used by freely moving pedestrians. In the presented wayfinding model we introduce the *decision scenes*, after considering the immediate environment in conjunction with the local decision points. We can define *decision scenes* as the local vista space around a particular decision point. It can be entered and left and its physical boundaries are fixed by building or other solid obstacles that prevent movement. These constrains are imposed by, for example, traffic lanes that are not suited for passage (excluding zebra crosswalks). *Decision scenes* are adjacent to each other, forming a partition of the walkable space. As *decision scenes* are physically bounded and can be entered and left, they concur with the mental representation of the *container* schema (Gaisbauer and Frank 2008). The borders between two decision scenes are labeled *portals* and connect neighboring scenes. These portals are not constrained only to the border of adjacent *decision scenes*, but they may also encompass doorways of buildings allowing us to combine our model with one for indoor wayfinding. The concept of *portal* allows the transition from one scene (*container*) to another, which is the *link* schema. The option of walking directly from one portal to another is represented by the complete graph constructed between the portals. Using this connection, direct navigation between portals is modeled without inclusion of the local decision point, which results in the fact that these points are no longer vital for navigation unless they are the start or the destination. Figure 5.3 explains the concept of the *decision scenes*. It includes walkable space with decision points, the conventional navigation graph and the imposed complete graph within the decision scenes (Gaisbauer and Frank 2008). The next chapter shows how to implement this model into real life and construct decision scenes from the real data.

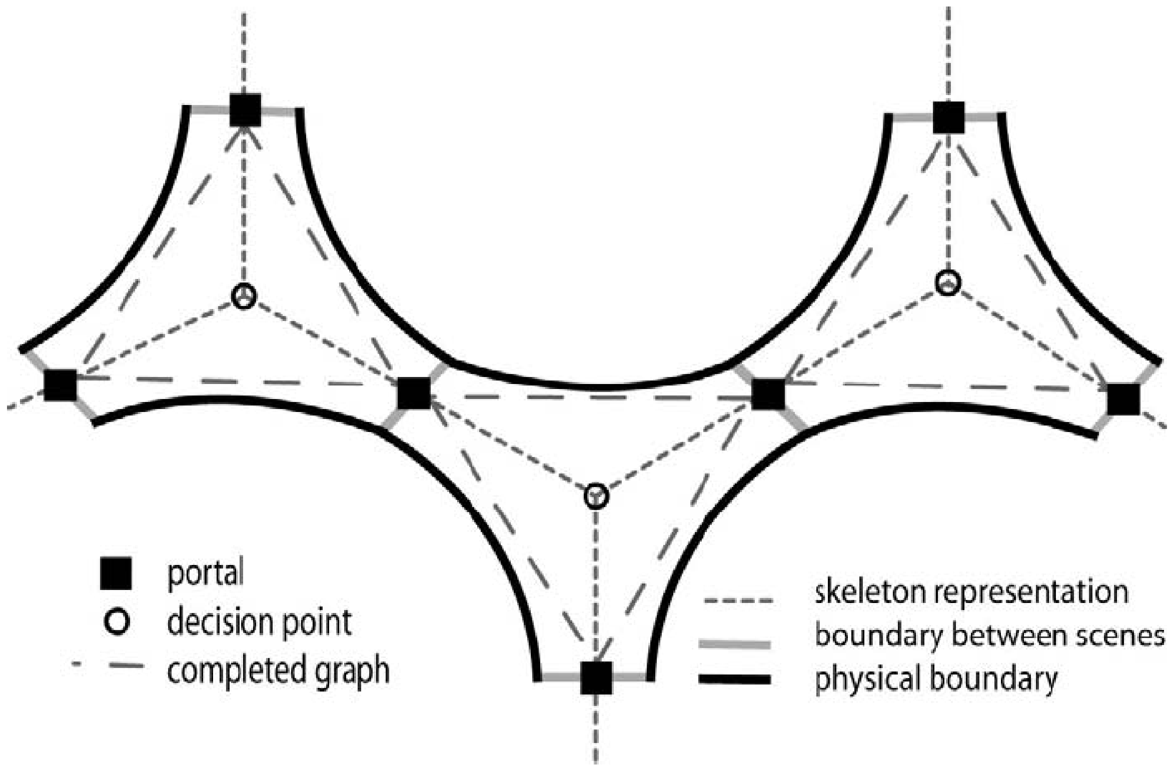


Fig. 5.3: The final graph representation of the wayfinding model (Gaisbauer and Frank 2008, 7)

6. Digital Image Processing

In this chapter I will describe a design of pedestrian navigation systems, which is focused on providing datasets for the designed system. The process of data acquisition is utterly connected with processing of the digital image data. In the first subchapter the short introduction and some basic definitions of Digital Image Processing (DIP) will be explained. Next subchapters will concentrate on the process of data processing itself: preprocessing and morphological operations. The main assumptions will be introduced in general, and then each stage will be described in detail. All the practical work done here was carried out using the Matlab programming language.

6.1 Digital image

In Digital Image Processing we are dealing with digital images. This term is not very broad and represents the definition of an image in a more specific way. A *digital image* is (Wikipedia, s.v. “digital image”):

[…] representation of a two-dimensional image using ones and zeros (binary). Depending on whether or not the image resolution is fixed, it may be of a vector or a raster type. Without qualifications, the term <<digital image>> usually refers to raster images.

In “Digital Image Processing” by Gonzalez and Woods, the authors extend the definition of image to a two-dimensional function, *f(x, y)*, where *x* and *y* are spatial coordinates. The terms *intensity* and *gray level* are also presented as to be the amplitude of *f* at any pair of coordinates (x, y). When *x, y,* and the amplitude values of *f* are finite, discrete quantities, then we can call image a *digital image* (Gonzalez and Woods 2002).

6.1.1Raster

Digital image can be either raster of vector type. *Raster* images have a finite set of digital values, called picture elements or pixels. The digital image contains a fixed number of rows and columns of pixels, which are the smallest individual element in an image, holding quantized values representing the brightness of pixels. The raster image is usually referred to as bitmap, because it contains information that is directly mapped to the display grid. Raster images can be created by a variety of input devices, such as digital cameras, scanners, etc. Each pixel of a raster image is associated to a position in some 2-D region, and has a value consisting of one or more quantities related to that position. Digital images can be classified, if based on those quantities –their number and nature—into some groups (types): binary, grayscale, color, false-color, multi-spectral, thematic, picture function (Wikipedia, s.v. “digital image”). For the image processing performed in this thesis, the binary type of raster data was of particular interest.

6.1.2 Vector

V*ector* graphic formats are complementary to raster type. In vector graphics the storing of lines, shapes, and colors that make up an image, is made by using mathematical formulae. A vector graphic program uses these formulae to construct the screen image by building the best quality image possible, given the screen resolution, from the mathematical data. Unlike the raster data, where with the growth of quality (resolution) of an image, the data size grows significantly. The mathematical formulae determine where the dots that make up the image should be placed, for the reason of obtaining best results when displaying the image. Because of the fact that these formulae can produce an image that is sizable to any size, the quality of the image is only determined by the resolution of the display. Figure 6.1 shows the effect of scaling vector data versus raster graphics. Raster images are based on pixels, thus they are dependent on a data granularity (resolution) and scaling those causes a loss of sharpness. On the contrary, vector-based images can be scaled without degradation.

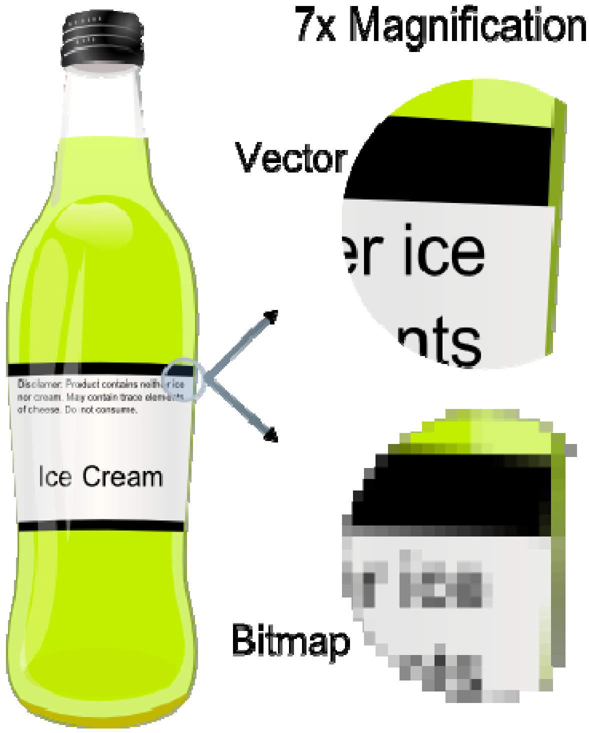


Fig. 6.1: Raster vs. vector image

In our case, the selection of the data type used to perform the task of image processing was constrained by availability of datasets. Fortunately, the City of Vienna provides test-data covering some parts of the city in raster and vector data.

6.2 What is Digital Image Processing?

Interest in digital image processing methods originates from two principal application areas: improvement of pictorial information for human interpretation and processing of image data for storage, transmission, and representation for autonomous machine perception (Gonzalez and Woods 2002).
Digital Image Processing as the usage of computer algorithms to perform image processing on digital images (Wikipedia, s.v. “digital image processing”). It allows a much wider range of algorithms to be applied to the input data. What is more, it can avoid problems such as the build-up of noise and signal distortion during the process. The history of digital image processing goes back to the 60s of the previous century, when it was developed at Jet Propulsion Laboratory, MIT, Bell Labs, and University of Maryland. In those years it was called *“digital picture processing”* and found its application in the field of satellite imagery, medical imaging, photo enhancement and

many more. But due to high costs of the processing resulting from the poor computing equipment, the technique was quite inefficient. Limited with fast computers and signal processors in 90s and 2000s, digital image processing has become the most common form of image processing, and it is generally used not only because it is the most versatile method, but also because it is least expensive one. The application of DIP can be found in many fields: such as medical science (X-rays), remote sensing, photogrammetry, geoinformation, computer engineering, climatology, sociology, and many more.

Gonzalez and Woods (2002) distinguished numerous fundamental subfields in digital image processing, such as:

- Image acquisition
- Image enhancement
- Image restoration
- Color image processing
- Wavelets and multiresolution processing
- Compression
- Morphological processing
- Segmentation
- Representation & description
- Object recognition

My work will be connected with few of the fields specified above. While developing an automatization of data acquisition I have used mainly morphological processing and segmentation.

6.2.1 Matlab as a key to Digital Image Processing

To perform all of the image processing tasks I decided to use the Matlab programming environment. It has a very important advantage: implementation of huge variety of ready-made functions and operators in DIP. In Matlab this set of additional functions specially designed for the image processing is called Image Processing Toolbox (IPT).

Matlab is a high-level language and interactive environment to perform computationally intensive tasks faster than with traditional programming languages (www.mathworks.com). Created by MathWorks, Matlab language allows matrix manipulation, plotting of functions and data, implementation of functions, creation of user interfaces, and interfacing with programs written in different languages. Matrix manipulation makes Matlab the one of the best programming languages to deal with image processing. Matlab is built around Matlab language, called also *M-code* or basically *M*. The built-up of the Matlab—Image Processing Toolbox provides a set of algorithms and graphical tools for image processing, analysis, visualization and algorithm development. Most of toolbox functions are written in the Matlab language, giving the possibility to inspect the algorithms, modify the source code, and create custom functions.

6.2.2 Morphological Image Processing

Introducing the topic of morphological image processing is essential for understanding the following chapters, as most of the functions are based on this domain.

In the context of *mathematical morphology* we use the same word here for extracting image components that are useful in the representation and description of region shape, such as boundaries, skeletons, convex hulls and many more. (Gonzalez and Woods 2004). Mathematical morphology (Serra 1982) is a theory and technique for the analysis and processing of the geometrical structures, based on set theory, lattice theory, topology, and random functions. In the following, I explain the operators, which were used the most during pre-processing.

<u>Structuring elements</u>

Morphological operations, just like filter operations, are one of the *neighbor* operators. Differentiating from filtration, in morphological operations the structuring element is used instead of mask and is defined as 1's in a logical matrix. In such matrix we can distinguish central point. Examples of structuring elements are shown in Fig. 6.2. Image Processing Toolbox is equipped with a function *strel* that constructs structuring elements with a variety of shapes and sizes (e.g., diamond, disk, line, square, octagon, etc.). During the pre-processing of the data, structuring elements are used very often.

1	0	1
0	1	0
1	0	1

1	0	1	1	1	0	1

1	1	1	1	1	1	1	1

1
1
1

1
0
1

Fig. 6.2: Examples of structuring elements

Dilation and erosion

The operations of *dilation* and *erosion* are fundamental to morphological image processing. In reference to binary images, *dilation* is an operation that "grows" or "thickens" objects. The specific manner and extent of these operations is controlled by structuring element. Mathematically, *dilation* is defined in terms of set operations. The dilatation of A by B is defined as:

$$A \oplus B = \{z \in E | (B^s)_z \cap A \neq \varnothing\}.$$

where $\varnothing$ is the empty set and *B* is the structuring element. Basically, the *dilation* of *A* by *B* is the set consisting of all the structuring element origin location where the reflected and translated *B* overlaps at least some portion of *A*.

Whereas *erosion* of the binary image *A* by the structuring element *B* is defined by:

$$A \ominus B = \{z \in E | B_z \subseteq A\}.$$

Like in *dilation*, when the structuring element *B* has a center, and this center is located on the origin of *E*, than the *erosion* of *A* by *B* can be understood as the locus of points reached by the center of *B* when *B* moves inside of *A* (Serra 1982).

Concerning grayscale images, the result of *dilation* operator in a certain point is the highest pixel value amongst image pixels covered by structuring element. Whereas discriminance of the lowest pixel value amongst these covered by structuring element is an effect of *erosion*. Fig. 6.3 illustrates the examples *dilation* and *erosion* operators.

A

1	1	1
1	1	1
1	1	1

B

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	1	1	1	1
0	0	1	2	1	1
0	0	1	1	1	3

C

0	0	0	0	0	0
1	1	1	1	1	1
1	1	1	1	1	1
1	1	2	2	2	2
1	1	2	2	3	3
1	1	2	2	3	3

D

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	1	1	1
0	0	0	1	1	1

Fig. 6.3: Example of the dilatation and erosion operators.
A-structuring element; B- original image; C-dilatation result; D-erosion result

Opening and closing

Opening and closing are the morphological operators defined by using dilation and erosion. The *opening* of *A* by *B* is obtained by the erosion of *A* by *B*, followed by dilation of the resulting image by *B*:

$$A \circ B = (A \ominus B) \oplus B$$

Whereas the *closing* of *A* by *B* is obtained by the dilation of *A* by *B*, followed by erosion of the resulting structure by *B*:

$$A \bullet B = (A \oplus B) \ominus B$$

During processing of a raster image containing the test data, the closing operator is thought to be useful concerning filling the gaps that connect two neighboring areas. The selection of the adequate structuring element, which should be in shape and size of the biggest gap we want to fill, is important.

6.3 Data sets

The data has to present the urbanized areas, containing not only streets, but also some open spaces and data has to be provided in a raster type. That is because of the fact that vector data for large areas are only available for streets and not for the pedestrian areas. There is a possibility of rasterizing such vector data, although there is no need to do it, since we are able to obtain raster data that is of a good quality. At last, it was useful that one should be able to verify data on-site.

Amongst many datasets that were considered, the one that fits the conditions best was the “Mehrzweckkarte” provided by “Stadtvermessung Wien – Magistratsabteilung 41” (Land map provided by Vienna State Land Survey). For the purposes of this thesis, we have used the test datasets. The data can be obtained in raster or vector format. But as the raster dataset consists of many different layers (each layer is described in a different color), there was no need to focus on vector data. The Matlab programming environment operates only on raster data. The advantage of vector data was that to some extent it was more detailed than raster data:

- File size: 720Kb,
- Image size: 1200 x 1200 pixels,
- File format: TIF,
- Resolution: 72 pixels per inch.

The map included the area in Vienna, constrained from:

- North: Kärtner Ring,
- South: Resselpark,
- West: Wiedner Hauptstraße,
- East: The building of Musikverein.

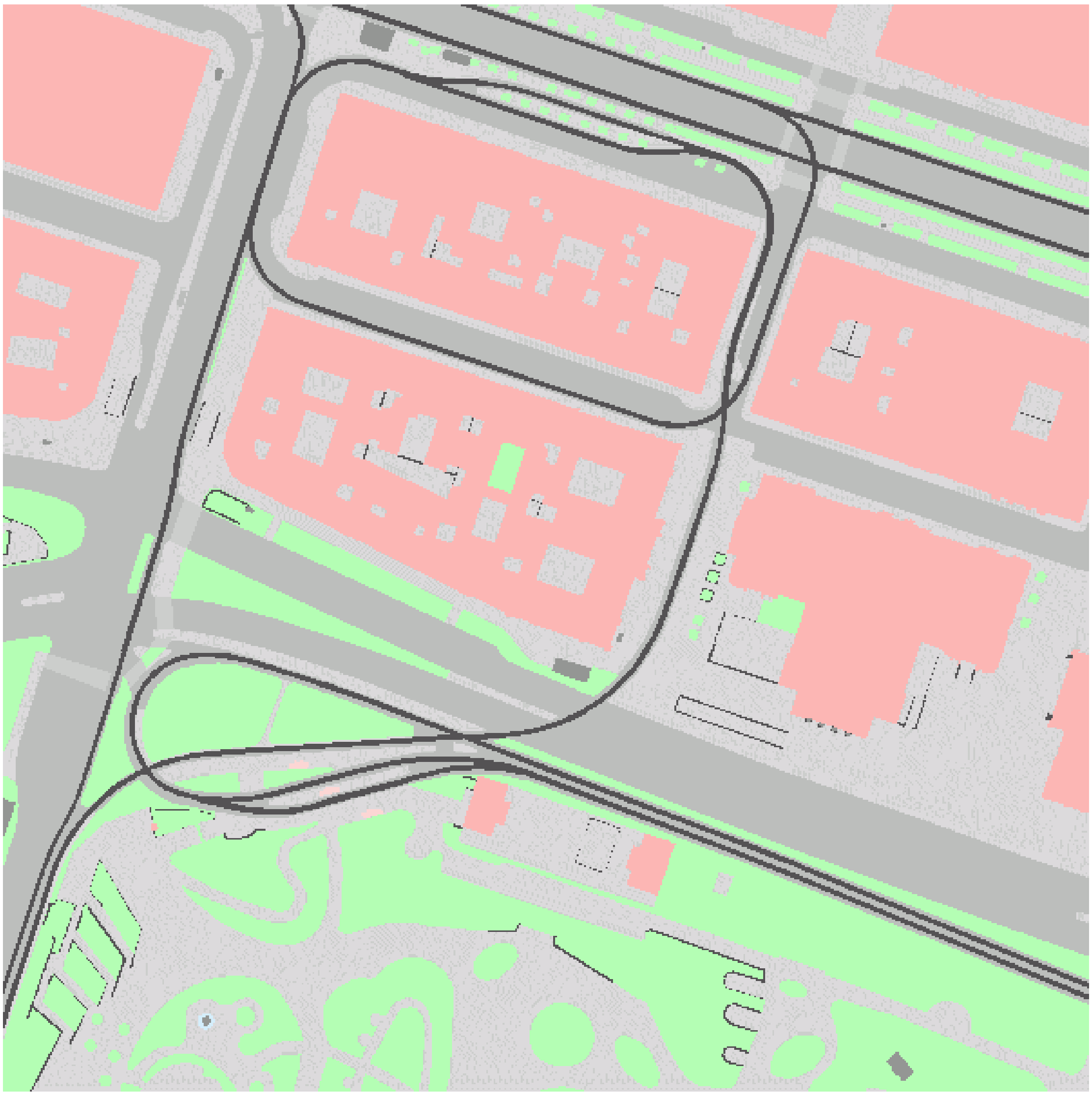


Fig. 6.5: Raster map

6.4 Processing steps overview

In this subchapter I will propose the procedures—steps that have to be performed in order to obtain the result: the decision scenes.

The step-by-step procedure of processing, further called forth the algorithm, consisted of different stages as described in the Figure 6.6. The idea of structuring an algorithm to such steps was also proposed by Walter et al. (2006).

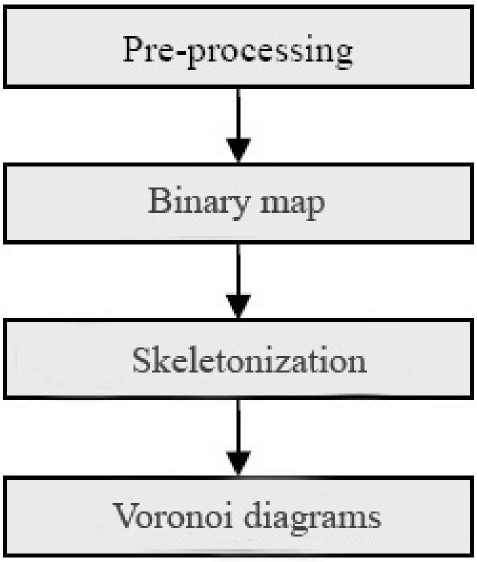


Fig. 6.6: Workflow

In the pre-processing (1) step using the raster data, the navigationable and non-navigationable areas for pedestrians are derived automatically. The fact that input data was raster type, with various separate features (segmented into colors), significantly facilitated this step. After that we have obtained the binary image (2) (map), which consisted of 0's for non-navigationable areas, and 1's for navigationable areas. When having a binary image of the test area obtained, it was possible to proceed with the skeletonization (3) method that provided us with navigationable graph retaining initial map topology. The next step was to clean the skeleton and eliminate the non-connected branches and nodes. Having the cleaned skeleton ready it was possible to distinguish the decision points affiliated to the graph. Based on the distinction of decision points—division of space by using Voronoi diagrams (4)—we were able to construct decision scenes.

This description shows the whole process in a very general way. Each step will be described in detail in the next subchapters. The algorithm is presented in the last subchapter.

6.5 Step I: Pre-processing

The first step before processing the data, or getting data ready for further processing, was to analyze it. As it was mentioned before, the raster data provided by the City of Vienna was in a form of TIF image with indexed color. While Matlab operates on matrices, it was more convenient to convert color coded data into grayscale, so that each color would get one grayscale value. Dealing with only one matrix (pixel grayscale values) is more useful than with three matrices (each for one color: RGB). After this operation we could perform some analysis and investigate which pixel values correspond to which land use features. Color coding was as follows:

- 83: Walls,
- 149: Station devices,
- 189: Roads,
- 203: Buildings,
- 205: Zebra crosswalks,
- 219: Sidewalks,
- 223: Green areas,
- 225: Würstelstand,
- 232: Monuments.

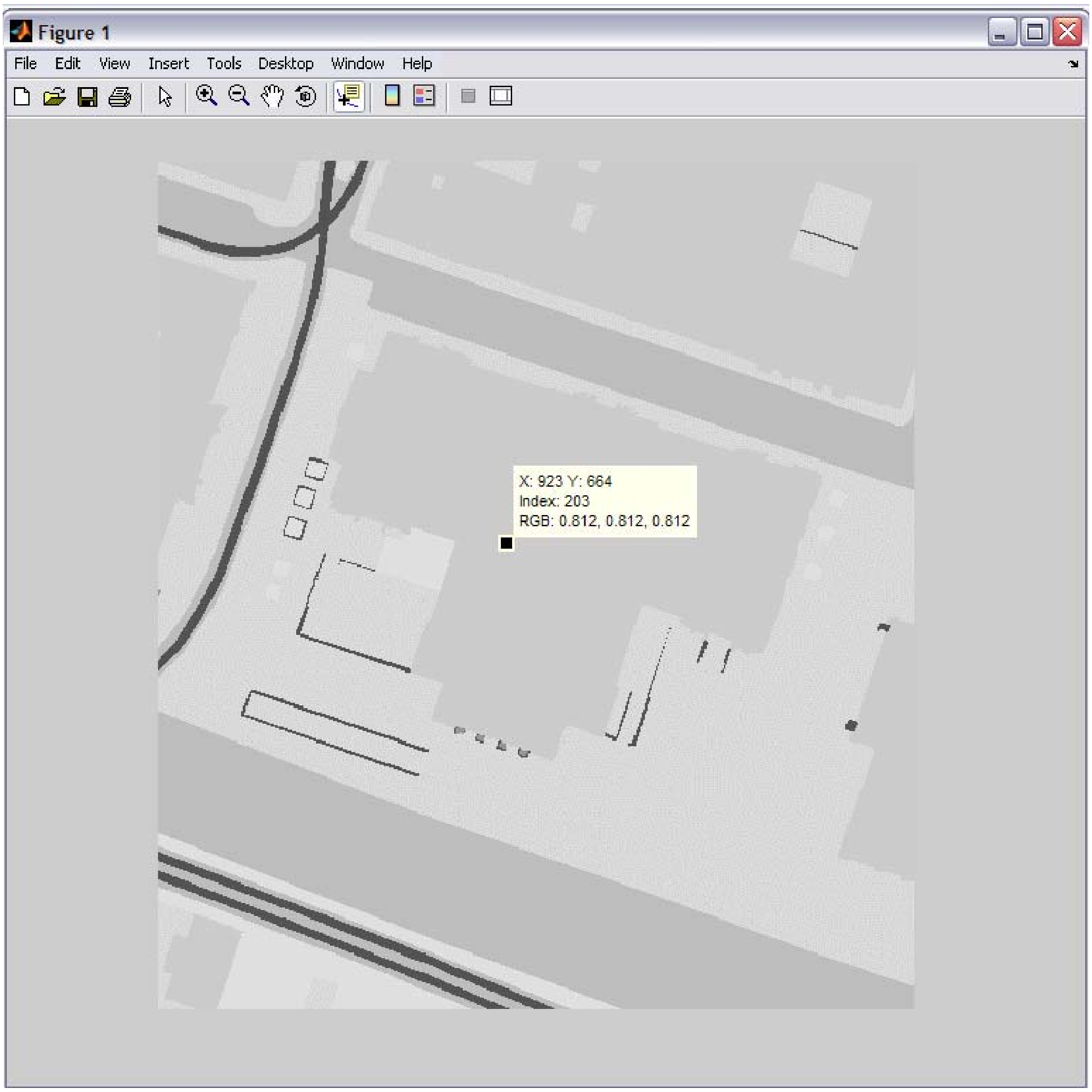


Fig. 6.7: Analyzing image

Having this information provided, the next step was to distinguish between areas that can be used by pedestrians and areas that cannot. It is obvious, that only "Sidewalks" (219) and "Zebra crosswalks" (205) can be used as navigationable areas. We excluded "Green areas" (223) from navigationable areas, because the map was detailed enough to provide precise information about park alleys.
During analysis we encountered some problems. To be certain, we compared the raster data with reality and it turned out that some objects in real world were wrongly transferred on the map. For example (Fig. 6.8) the railway tracks were classified as walls.

Fig. 6.8: Wrongly classified railway tracks

After analyzing raster data we binarized the map. Following the definition of CoderSource.net (2005), the binary image is:

[…] a digital image that has only two possible values for each pixel. Typically the two colors used for a binary image are black and white though any two colors can be used. The color used for the object(s) in the image is the foreground color, while the rest of the image is the background color.

Binary images often arise in DIP as masks or as the result of certain operations such as segmentation, thresholding and dithering. A binary image is usually stored in memory as a bitmap (Wikipedia, s.v. "binary image"). In our case we used binary image to distinguish between areas that are navigationable and non-navigationable by pedestrians. In order to accomplish with that, we had to replace pixel values responsible for navigationable areas with 1's and non-navigationable areas with 0's.
After the first test another problem showed up. The areas that are courtyards (classified on map as navigationable) were not connected to the street (or sidewalk) network and therefore existed as separate areas. I decided to treat the whole building area (between the streets) as one area, and therefore—fill those gaps. It could be done by using *imfill* function that fills regions in an image.

It resulted in having building areas being closed and having pixel values of 1 (Fig. 6.9). In order to combine it with the rest of the data, two images had to be added up (Fig. 6.10).

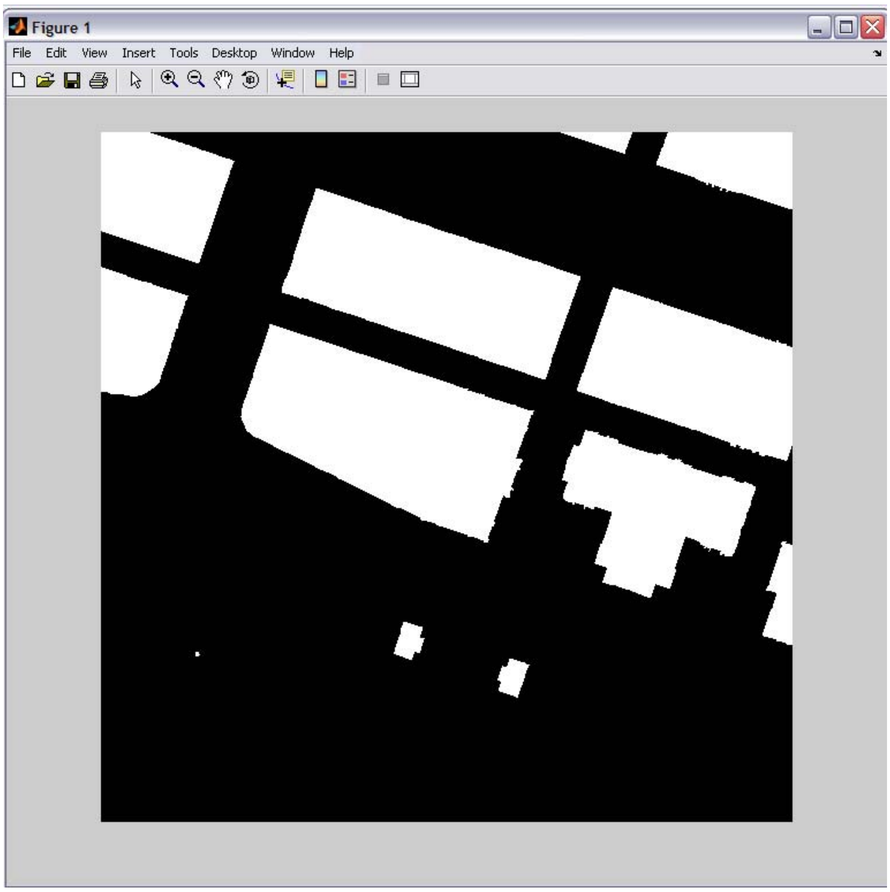


Fig. 6.9: Result of closing building areas

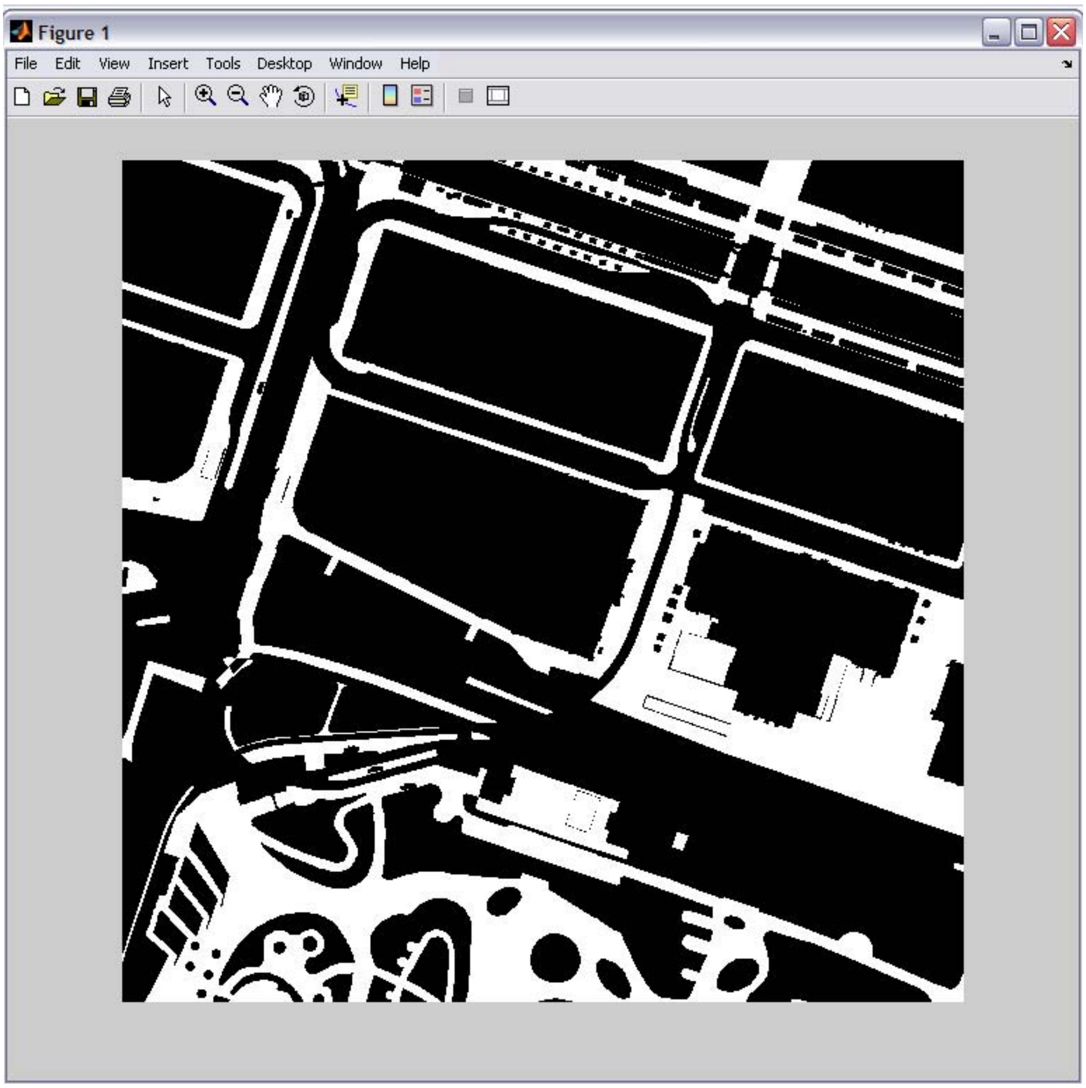


Fig. 6.10: Result of image addition

6.6 Step II: Binary map

The main aim of this step of processing was to prepare data for the skeletonization procedure. We had to analyze the data obtained so far after pre-processing stage, from the data consistency and correctness point. In the Fig. 6.10, one can see that there are some gaps between navigationable areas that should not exist, and therefore need to be connected. The example of such erroneous data is shown in Fig. 6.11.

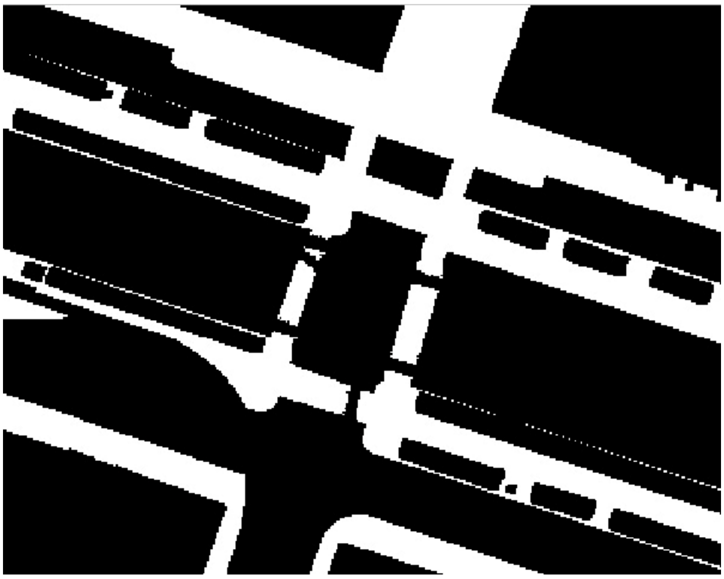


Fig. 6.11: Erroneous data example

The reason for crosswalks to be disjoint is due to the fact, that railway tracks were crossing zebra crosswalks. We had to find a good way to cope with such cases. It turned out that such gaps can be filled automatically with the help of image *closing* function. Details of this operator were described before. Therefore we have chosen the structuring element that will fit the best, and help to connect these disconnected areas. An *ones* structuring element with a size of 18 pixels was chosen. We also needed to solve another problem. The railway tracks were disjoining some navigationable areas. This could cause a problem in the next processing steps. We had to perform operation that could allow to join two neighboring navigationable areas disjoined by railway tracks as in this case pedestrians are fairly free to cross the railway tracks. Result is show on Fig 6.12.

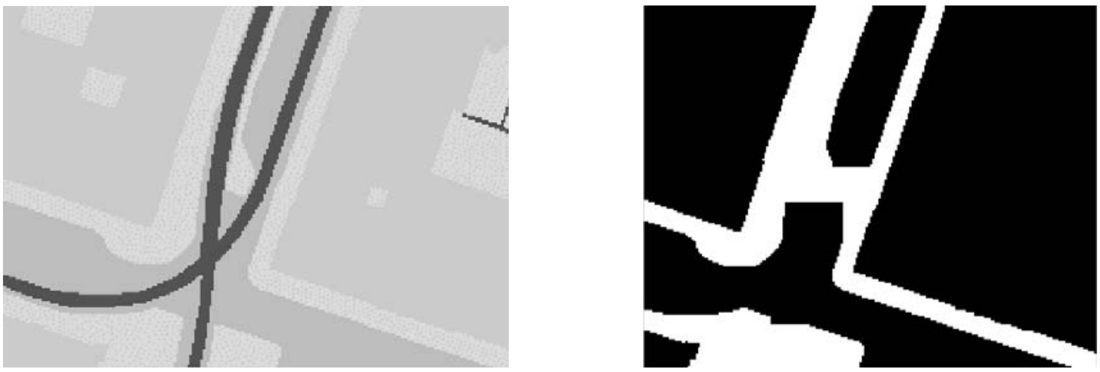


Fig. 6.12: Before and after joining two navigationable areas affected by railway tracks

Having in mind that this automatization is not 100% correct, and in some cases, there are navigationable areas that have to be connected. We made a module that allows to manually fill the gaps between disconnected areas (Fig 6.13).

Fig. 6.13: Filling the gaps

6.7 Step III: Skeletonization

Before explaining the skeletonization function, there is exceptional need of introducing details concerning skeletonization.

6.7.1Definition and examples

In his work, Serra (1982) explains the notion of a skeleton of a set X:

(1) if x is a point of skeleton and B_x the largest disk centered at x and contained in X, one cannot find a larger disk (not necessarily centered at x) containing B_x and included in X (one says that B_x is a maximum disk),

(2) the disk B_x hits the boundary ∂X at two or more different places.

The skeleton of A can be expressed in terms of erosions and openings. That is, it can be shown that:

$$S(A) = \bigcup_{k=0}^{K} S_k(A),$$

with

$$S_k(A) = \bigcup_{k=0}^{K} \{(A \ominus kB) - [(A \ominus kB) \circ B],$$

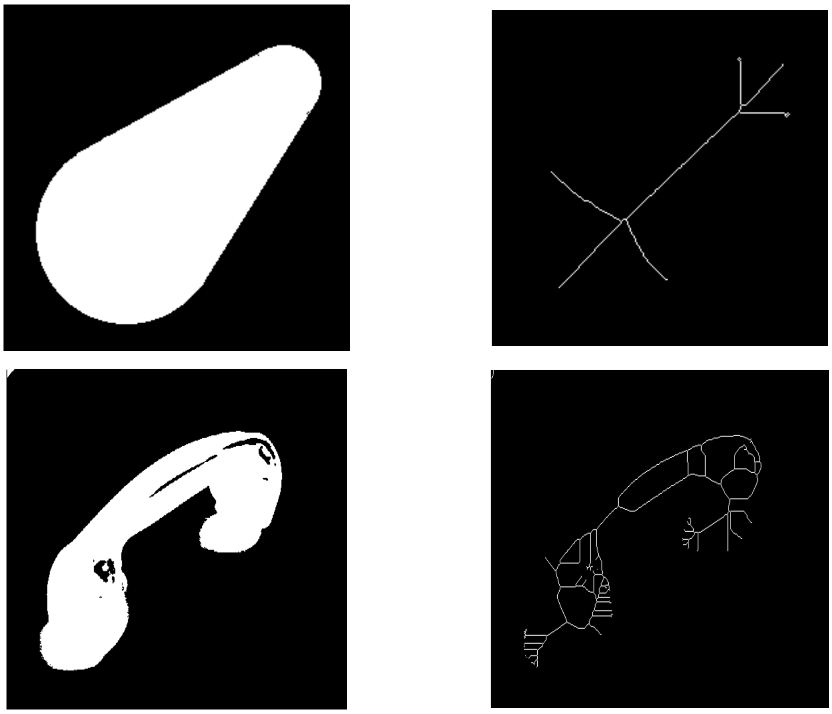
.

Implementing this definition, we are able to introduce the term of *straight skeleton*, which is a method of representing a polygon by a topological skeleton in geometry. It is similar to the medial axis.

One should bear in mind to represent the structural shape of a plane region is reducing it to a graph. This reduction is an important approach and may be accomplished exactly by thinning a region, which is also called *skeletonizing*. The skeleton of a region may be defined using the medial axis transformation. The medial axis transformation of a region *R* with border *B* is as follows: for each point *p* in *R*, we find its closest neighbor in *B*. If *p* has more than one such neighbor, it is said to belong to the medial axis of *R* (Blum 1967).

Thus the medial axis transformation of a region is an intuitive concept. Discreet implementation of this definition is computationally very expensive, because it involves calculating the distance from every interior point to every point on the boundary of a region. Numerous algorithms have been proposed for improving computational efficiency. In Matlab, the Image Processing Toolbox function *bwmorph* allows us to generate the skeleton of all regions contained in a binary image.

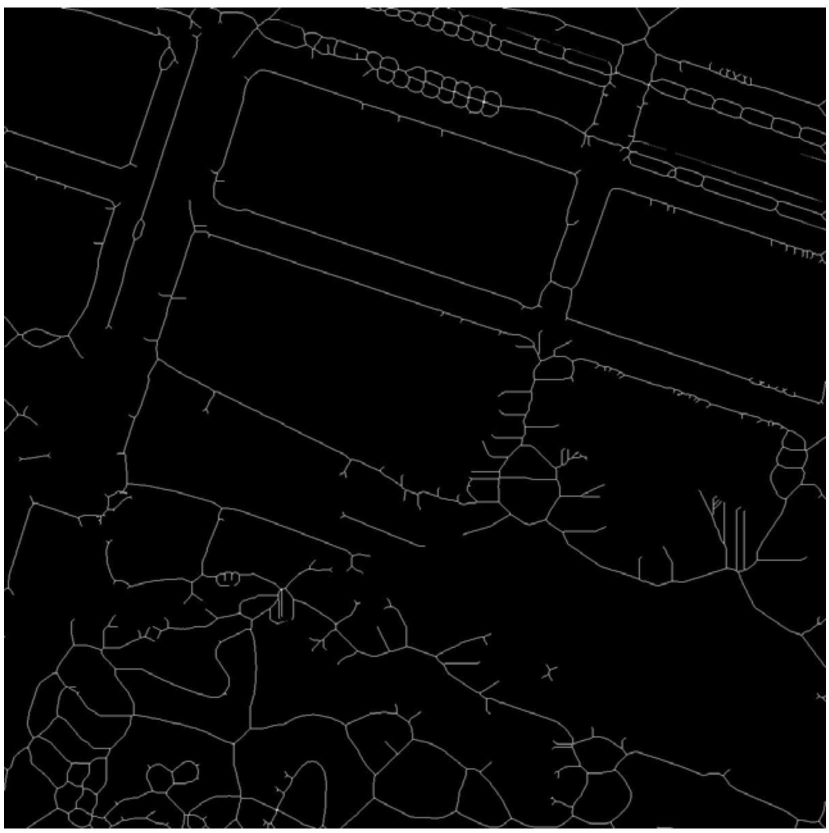
Below are some examples of how skeletonization function works in practice:



6.14: Examples of skeletonization

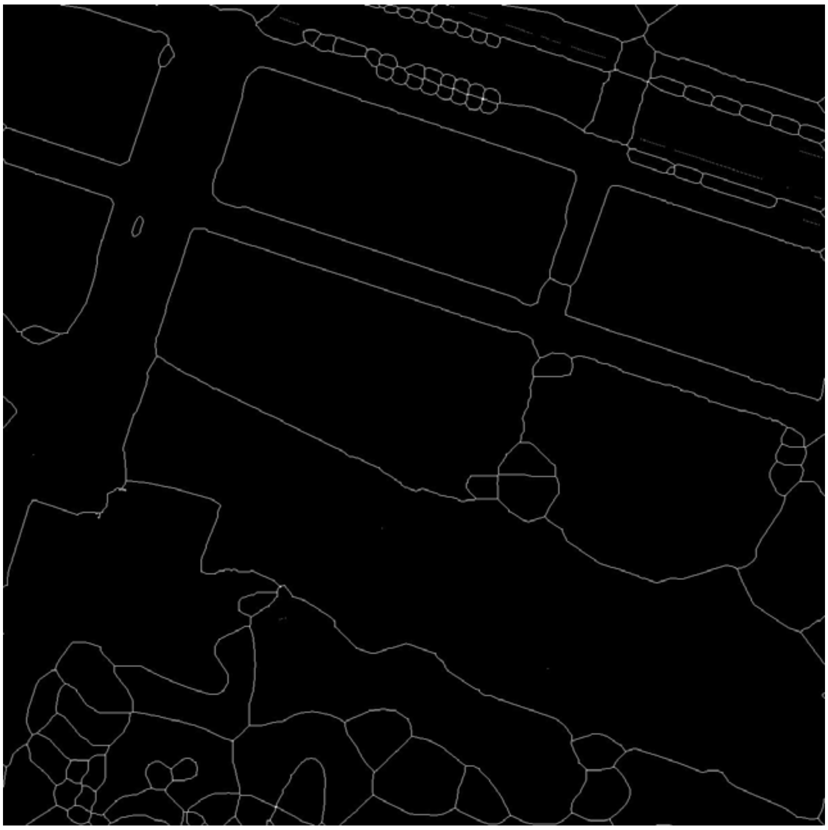
6.7.2 Application

The result of the skeletonization function applied on the image is as follows (Fig. 6.15).



6.15: Result of skeletonization

The skeleton has a lot of dead-end branches that are not needed for further processing. Therefore, they have to be cut. This operation in DIP is called “skeleton spurring”. The result of spurring is shown below (Fig. 6.16). Spurred skeleton image is the navigationable graph.



6.16:Spurred skeleton

6.8 Step IV: Voronoi diagrams

In this chapter we will discuss and explain the creation of “decision scenes” using Voronoi tessellation, also called Voronoi diagrams.

6.8.1Definitions and examples

The concept of Voronoi diagrams is a simple but intuitively appealing one. Given a set of two or more but a finite number of distinct points in the Euclidean plane, we associate all locations in that space with the closest member(s) of the point set with respect to the Euclidean distance. The result is a tessellation of the plane into a set of regions associated with members of the point set. We call this tessellation the *planar ordinary Voronoi diagram* generated by the point set, and the regions constituting the Voronoi diagram *ordinary Voronoi polygons* (Okabe et al. 2000).

Mathematically, a Voronoi Diagram for a set *S* of *N* planar points is a partition of the plane into *N* polygonal regions, each of which is associated with some point p_1 and S_1 and is the locus of points closer to any point than to any other points (Preparata and Shamos 1985).

Let

$$P = \{p_1, \ldots, p_n\}, \text{ where } 2 \leq n \leq \infty \text{ and } x_i \neq x_j \text{ for } i \neq j,\ i, j \in I_n.$$

The region given by

$$V(p_i) = \{x: \| x - x_i \| \leq \| x - x_j \| \text{ for } j \neq i,\ i \in I_n\}$$

is called ordinary Voronoi polygon associated with p_i or the Voronoi polygon of p_i and the set given by

$$V = \{ V(p_i), \ldots, V(p_n)\}$$

is called the planar ordinary Voronoi diagram generated by *P* or simply Voronoi diagram of *P*.

There are many fields where Voronoi diagrams may be useful. For example, in polymer physics they can represent free volume of the polymer. It is also used in derivations of the capacity of wireless networks. In climatology, Voronoi diagrams are used to calculate the rainfall of an area, based on a series of point measurements. In this usage, they are generally referred to as Thiessen polygons. Voronoi diagrams are also used in computer graphics to procedurally generate some kinds of organic looking textures. The examples of Voronoi tessellation are shown below:

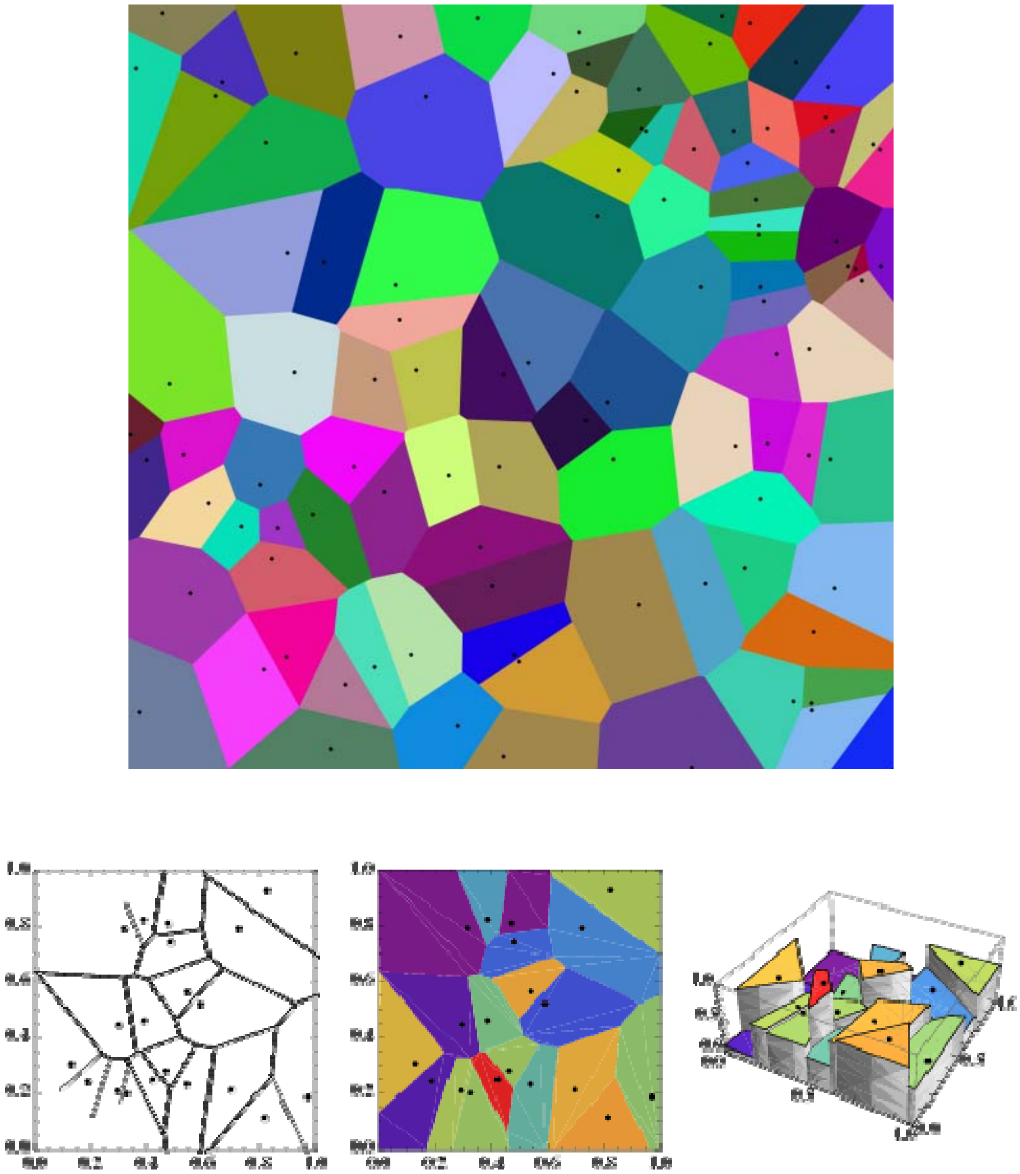
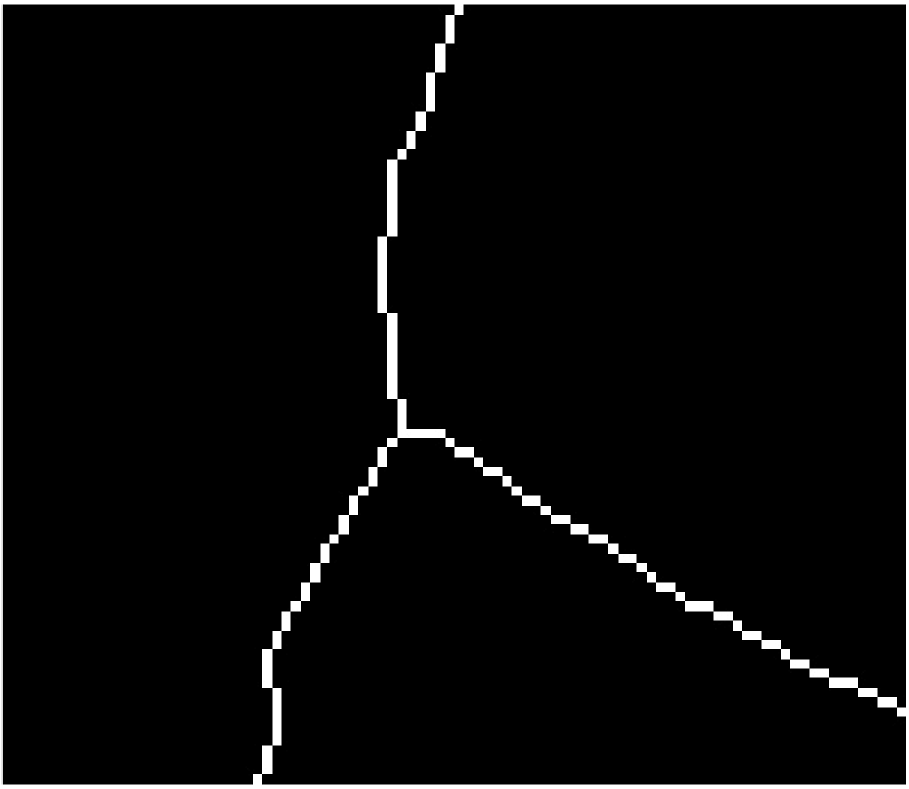


Fig. 6.17: Examples of Voronoi tessellation

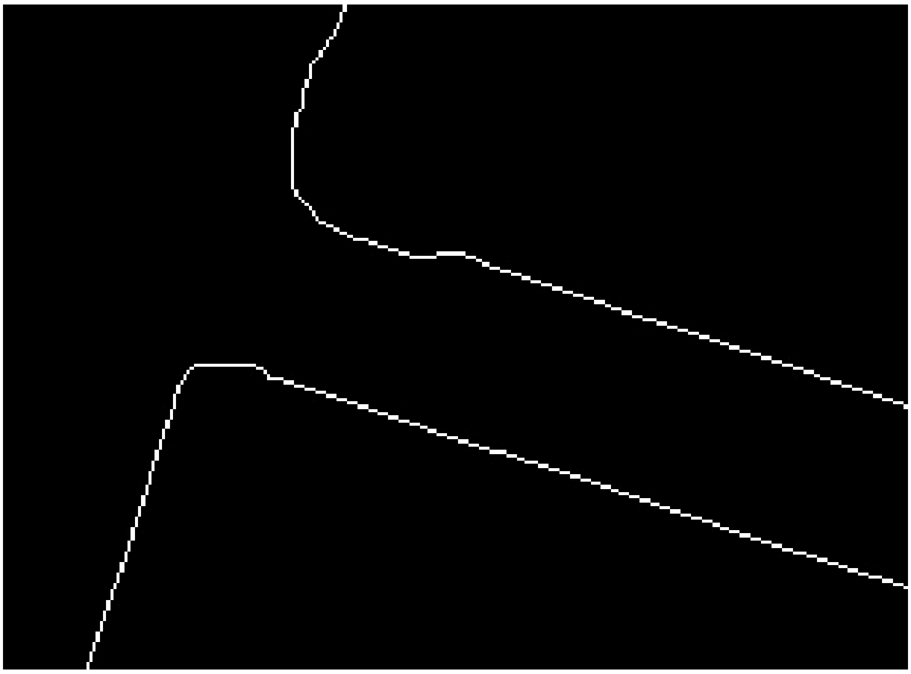
6.8.2 Application

The last processing step yielded in the spurred skeleton image. This is the navigationable graph that will be a base for creating decision scenes using Voronoi diagrams. First, we need to decide which points will be the basis for tessellation. These points are also acknowledged as the decision points. Considering assumptions of the wayfinding model, the decision points need to be present where the decisions of a user take place—at nodes. Therefore it is needed to find where the navigation graph branches off and to separate nodes as single pixels (Fig. 6.18).



6.18: Example of node

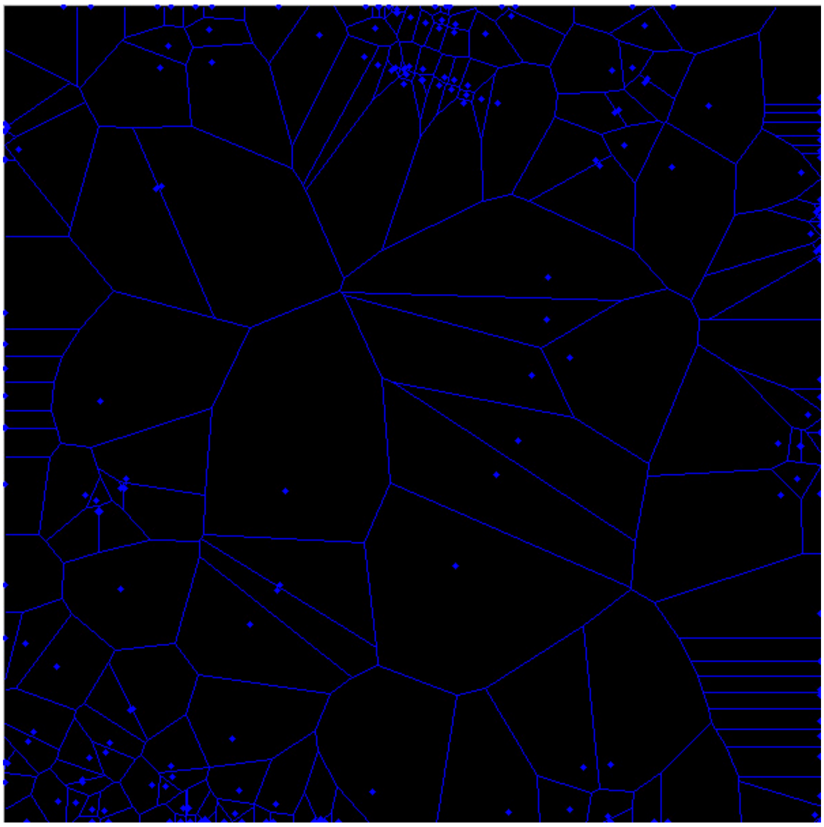
After separating nodes, certain data incompleteness was uncovered. It emerged that there are many zebra crossings missing. For example, around building block in the middle of an image (Fig. 6.19), the graph is not connected to any of paths opposite the street. It should be, as there are zebra crossings at the crossroads. This problem can be solved in two ways. We can either manually fill the gaps between opposite sidewalks, which will yield in connecting a graph, or we can use street graph (car navigation graph) to identify the points where streets intersect and create decision scenes basing on these points. In this case, the decision point will be in the middle of the street intersection, not at the end of zebra crossing as it would be possible, if one connected graph manually. Making one decision scene—at the street intersection is more efficient than dividing it into numerous small scenes.



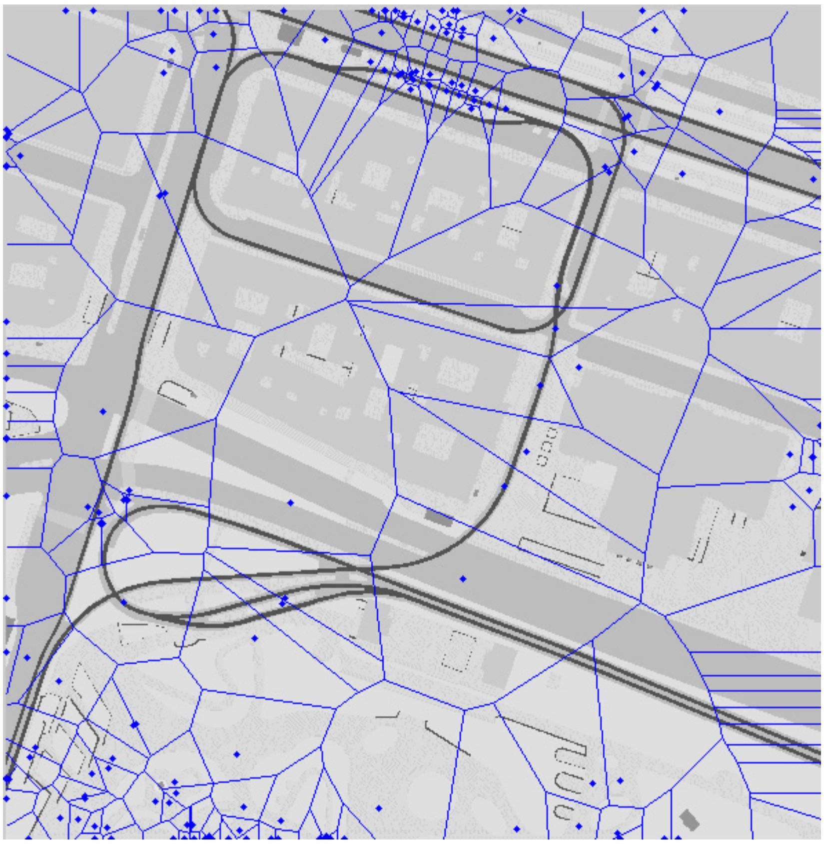
6.19: Example of graph around building block

In the first stage of selecting points for Voronoi diagrams construction, we separated only nodes derived from navigationable graph. The second step was to process the image from the beginning to separate only streets. Then the same actions of skeletonization and skeleton pruning were performed, resulting in the street network graph. To make Voronoi diagrams construction more reasonable we had to reject nodes that are too close to each other.

The last step was to tessellate the space using Voronoi diagrams (Fig. 6.20).



6.20: Space tesselation using Voronoi diagrams



6.21: Space tesselation using Voronoi diagrams overlaid on base map

7. Conclusions and future research

This chapter starts with the summary of research and problems raised in this thesis including current state-of-the-art, focuses on theoretical assumptions concerning pedestrian navigation systems, and introduces the way of carrying out the automatization of data acquisition and processing for such system purposes. In the second part of this chapter we will present results of my work, and in the last chapter, directions for the future development.

7.1 Summary

The main objective of this thesis was to elaborate an automatic method of acquisition and processing data for pedestrian navigation system, as well as to describe theoretical base for such a system. By reasons that were mentioned in previous chapters, available navigation systems for car drivers significantly differ from systems for pedestrians. In the second chapter we introduced prototypes or ready-made systems for pedestrian guidance aids. However, they still cannot be acknowledged as pedestrian systems, because of imperfections and unsolved problems. They are not fulfilling user information needs.

Most of all, datasets used for building a base for these systems do not meet expectations of future users. Research has proven that data structure in system for pedestrians, has to be different from previous one. Then it is impossible to transform datasets used in systems for car drivers, what has been already done unsuccessfully and described in chapter 2. Thus it was crucial to elaborate conception of data structure, and its method of acquisition. In a paper by Gaisbauer and Frank (2008), the data model was proposed. My contribution to their work was to automate the data acquisition for creating navigationable graph of street network and building the decision scenes. For testing I have used raster maps of the City of Vienna.

The second chapter describes prototypes of navigation systems for pedestrians, phases of adaptation datasets for such system purposes and explains basis of navigations systems for car drivers.

The next two chapters specify theoretical problems connected with designed conception of a system: chapter three describes current state of knowledge on spatial cognition and human wayfinding. The fourth chapter raises some important issues concerning conceptualization of space that surrounds us. It focuses on the concepts of human mental images of environment they perceive. In chapter five the wayfinding model introduced by Gaisbauer and Frank (2008) is presented and some future directions and improvements are also described. The last chapter of this thesis contains the most important part—the specification of data acquisition and processing automatization process. This is the most relevant part of this thesis, as it shows a practical result and solution to problems that were stated at the beginning. It makes a base for implementing this system, fulfilling all assumptions concerning its functionality simultaneously.

7.2 Results

The most relevant outcome of this thesis is an elaborated consideration of automatic acquisition and processing of data provided for navigation systems for pedestrians. I have proven my hypothesis – it is possible to automatically produce navigationable graph for pedestrian navigation system. In addition, this data is compatible with a wayfinding model proposed by Gaisbauer and Frank (2008), which differs from datasets used in recent research. Using Matlab programming environment we developed an algorithm that processes step-by-step raster map of Vienna to a navigationable graph consisting of nodes, segments and decision scenes.

The first stage of this work was to use binarization algorithm to distinguish navigationable areas—which pedestrians have the opportunity to walk on without obstacles (e.g., squares, parks, sidewalks, off-traffic roads); and non-navigationable areas, such as streets, buildings, water areas, etc. While having binary map obtained we were able to build a navigationable graph using skeletonization. Decision scenes posed another very important element of wayfinding model. The problem connected to creation of decision scenes was solved with the help of Voronoi diagrams, which was characterized in the same chapter as skeletonization.

Data quality was most of all problematic, researching for this thesis. The map that was used for automatization was lacking of connections between certain areas, such as sidewalks. Numerous zebra crossings were missing. This caused an explicit need for elaborating module which allows to correct such errors manually, as the algorithm cannot know where the zebra crossing should be. I have also proposed a solution for this problem that bypasses manual actions. It is possible to create nodes that are base for decision scenes from the points of street intersections.

7.3 Future work

Automatization of data acquisition and processing is just a part of project aiming at revolutionary approach to human wayfinding and creating a modern navigation system for pedestrian purposes. The main directions of future development are stressed principally on implementing this solution into mobile device that will be able to cope with the new system. What is also of significant importance, there are some improvements on the algorithm that has to be made. Moreover, trying to rewrite and re-implement this algorithm to some different programming environment would also be very useful. The language that can meet the needs best could be a functional programming language—Haskell.

There are still numerous issues that have to be solved in the future. For example, we still do not know where we should put decision points along straight route. Creating decision scene in long sidewalk corridor is not always efficient and convenient.

As it has been shown before, we have explicitly focused on processing particular raster data. There is a need for testing it on various types of data and different maps containing various areas. What is also significant is the interoperability of an algorithm. The future work should also focus on making it more flexible to various dataset types and making it context-aware.

8. Acknowledgements

This research has been supported by the SemWay project within the research programme FIT-IT: Semantic Systems by the Austrian Ministry for Transport, Innovation and Technology (BMVIT).

9. Bibliography

Allen, G. (1999): *Spatial Abilities, Cognitive Maps, and Wayfinding - Bases for Individual Differences in Spatial Cognition and Behavior*. In: R. Golledge (Ed.), Wayfinding Behavior - Cognitive Mapping and Other Spatial Processes. pp. 46-80, Johns Hopkins University Press, Baltimore.

Alm, H. (1990): *Drivers' cognitive models of routes*. In: W. van Winsum, H. Alm, J.M. Schraggen & J.A. Rothengatter (Eds.), Laboratory and field studies on route representation and drivers' cognitive models of routes (DRIVE II V1041 GIDS, Deliverable GIDS/NAV2, pp. 35-48). Groningen, The Netherlands: University of Groningen, Traffic Research Centre.

Arthur, P., and Passini, R. (1992): *Wayfinding: people, signs, and architecture*. New York: McGraw Hill.

Blum, H. (1962): *An Assiociative machine for dealing with the visual fields and some of its biological implications*. Biol. Prot. And Synth. Syst. 1, pp. 244-260.

Bowditch, N. (2002): *The American Practical Navigator*. Bethesda, MD: National Imagery and Mapping Agency. ISBN 0939837544.

Burnett, G.E. (2000): *Turn right at the traffic lights*. The requirement for landmarks in vehicle navigation systems, The Journal of Navigation, 53(3), pp.499-510.

Burnett, G.E. (1998): *Turn right at the King's Head: Drivers' requirements for route guidance information*. Unpublished PhD dissertation, Loughborough University, UK. S. Fontaine and M. Denis, "The production of route instructions in underground and urban environments," in *Proc. COSIT*, C. Freksa and D. M. Mark, Eds., pp. 83–94.

Canalys (2006): *Mobile GPS navigation market doubles year-on-year in EMEA*, Canalys research, UK, 2006. Available online as: http://www.canalys.com/pr/2006/r2006081.htm.

Car, A. (1996): *Hierarchical Spatial Reasoning: Theoretical Consideration and its Application to Modeling Wayfinding*. Department of Geoinformation, Technical University Vienna, Vienna.

Christon, C., Bulthoff, H.H. (2000): *Perception, representation and recognition: a holistic view of recognition*. Spat Vis 13(2–3), pp. 265–275.

CoderSource.net (2005): *Conversion of a Color Image to a Binary Image.*

Corona, B. and Winter, S. (2001): *Guidance of car drivers and pedestrians*. Department of Geoinformation, Technical University Vienna, Austria, 2001.

Davis, E. (1990): *Representations of Commonsense Knowledge*. Morgan Kaufmann Publishers, Palo Alto.

Dieberger, A. (1999): *A City Metaphor to Support Navigation in Complex Information Spaces*. In: Spatial Information Theory - A Theoretical Basis for GIS, vol. 1329, Lecture Notes in Computer Science, S. C. Hirtle and A. U. Frank, Eds. Berlin-Heidelberg: Springer-Verlag, pp. 53-67.

Downs, R.M., and Stea, D. (1973): *Image and Environment: Cognitive Mapping and Spatial Behavior.* Transaction Publishers.

Downs, R. M. and Stea, D. (1977): *Maps in Minds*: Reflections on Cognitive Mapping. New York: Harper and Row.

Downs, R.M. and Liben, L.S. (1985): *Children's understanding of maps*. In: P. Ellen and C. Thinus-Blanc, editors, Cognitive Processes and Spatial Orientation in Animal and Man, volume 2, pages 202-19. Martinus Nijhoff, Dordrecht, Nederlands.

Elias, B. (2007): *Pedestrian Navigation - Creating a tailored geodatabase for routing*. 4th Workshop on Positioning, Navigation and Communication 2007 (WPNC'07), Hannover, Germany, 1-4244-0871-7/07 IEEE, pp.41-47.

Evans, G.W., Skorpanich, A.A., Gärling, T., Bryant, K.J., & Bresolin, B. (1984): T*he effects of pathway configuration, landmarks and stress on environmental cognition*. Journal of Environmental Psychology 4 (4), 323-335.

Gaile, G. L. and Willmott, C. J. (1989): *Geography in America*. Columbus, Ohio: Merrill Publishing Company.

Gaisbauer, C. and Frank, A.U. (2008): *Wayfinding Model For Pedestrian Navigation*. Institute of Geoinformation and Cartography, Vienna University of Technology.

Gero, J.S. (1999): *Representation and Reasoning about Shapes: Cognitive and Computational Studies in Visual Reasoning in Design*. In: Spatial Information Theory - Cognitive and Computational Foundations of Geographic Information Science, vol. 1661, Lecture Notes in Computer Science, C. Freska and D. M. Mark, (Eds). Heidelberg -Berlin: Springer-Verlag, pp. 315-330.

Gluck, M. (1991): *Making Sense of Human Wayfinding: Review of Cognitive and Linguistic Knowledge for Personal Navigation with a New Research Direction*. In: D. Mark and A. Frank (Eds.), Cognitive and Linguistic Aspects of Geographic Space. Series D: Behavioural and Social Sciences 63, pp. 117-135, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Golledge, R. (1999): *Human Wayfinding and Cognitive Maps*. In: R. Golledge (Ed.), Wayfinding Behavior - Cognitive Mapping and Other Spatial Processes. pp. 5-45, Johns Hopkins University Press, Baltimore.

Gonzalez, R.C., Woods, R.E. (2002): *Digital Image Processing (Second Edition)*. Prentice-Hall.

Gonzalez, R.C., Woods, R.E., Eddins, S.L. (2004): *Digital Image Processing using MATLAB*. Prentice-Hall.

Goodman, J., Brewster, S.A., Gray, P. (2005): *How we can best use landmarks to support older people in navigation.* In: Behaviour & Information Technology, Vol. 24, No. 1, Taylor and Francis Ltd, pp. 3-20.

Hirtle, S.C, Jonides, J. (1985): *Evidence of hierarchies in cognitive maps*. Mem Cogn 13(3), pp. 208–217.

Hirtle, S. (1998): *The Cognitive Atlas: Using GIS as a Metaphor for Memory*. In: M. Egenhofer and R. Golledge (Eds.), Spatial and Temporal Reasoning in Geographic Information Systems. pp. 267-276, Oxford University Press.

Ittelson, W. H. (1973): *Environment perception and contemporary perceptual theory*. In: W. H. Ittelson (Ed.), Environment and Cognition. New York: Seminar, pp. 1-19.

Johnson, M. (1987): T*he Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason.* The University of Chicago Press, Chicago.

Kitchin, R. (1994): *Cognitive Maps: What are they and why study them?* Journal of Environmental Psycholog*y* 14, pp.1-19.

Kitchin, R., Blades, M. (2001): *The cognition of geographic space*. I. B. Tauris, London

Krafft, M.F. (2001): *A Neural Optimal Controller Architecture for Wayfinding Behavior*. Departament of Computer Science, Swarthmore College, PA, USA.

Krüger, A., Butz, A., Müller, C., Wasinger, R., Steinberg, K., Dirschl, A. (2004): *The Connected User Interface: Realizing a Personal Situated Navigation Service*. In: Proceedings of the International Conference on Intelligent User Interfaces, pp. 161-168. ACM Press.

Kuipers, B. (1982): *The 'Map in the Head' Metaphor*. Environment and Behaviour 14(2), pp. 202-220.

Lakoff, G. (1987): *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. The University of Chicago Press.

Lynch, K. (1960): *Image of The City*. Cambridge, MA: MIT Press.

Malaka, R. and Zipf, A. (2000): *DEEP MAP – Challenging IT research in the framework of a tourist information system*. In: Fesenmaier, D.R., Klein, S., Buhalis, D., (Eds): Information and communication technologies in tourism 2000. Springer, Berlin, Heidelberg, New York, pp. 15-27.

Mandler, J. M. (1983): *Representation*. In: P. Mussen (Ed.): Handbook of Child Psychology, Vol. III (4th ed.). New York: John Wiley and Sons, pp. 420-494.

Mark, D. M. (1992): *Spatial metaphors for human-computer interaction*. Proceedings, Fifth International Symposium on Spatial Data Handling, v. 1, 104-112.

Mark, D. M. and Freundschuh, S. M. (1995): *Spatial Concepts and Cognitive Models for Geographic Information Use*. In: Nyerges, T. L., Mark, D. M., Laurini, R., and Egenhofer, M., (Eds.): Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems. Dordrecht: Kluwer Academic Publishers, pp. 21-28.

May, A., Ross, T., Bayer, S.H., Burnett, G.E. (2001): *Using landmarks to enhance navigation systems: driver requirements and industrial constraints*. In: Proceedings of the 8th World Congress on Intelligent Transport Systems, Sydney, Australia.

May, A.J., Ross, T., Bayer, S.H., Tarkainen, M.J. (2003): *Pedestrian navigation aids: information requirements and design implications*. In: Pers Ubiquit Comput (2003) 7, Springer-Verlag London Limited, pp. 331–338.

Millonig, A., Schechtner, K. (2007): *Developing Landmark Based Pedestrian-Navigation Systems*. In: IEEE Transactions on Intelligent Transportation Systems, Vol. 8, No. 1, pp. 43-49.

Montello, D. (1993): *Scale and Multiple Psychologies of Space*. In: A. U. Frank and I. Campari (Eds.), Spatial Information Theory: A Theoretical Basis for GIS. Berlin: Springer-Verlag, pp. 312-321.

Montello, D. (1998): *A New Framework for Understanding the Acquisition of Spatial Knowledge in Large-Scale Environments*. In: M. Egenhofer and R. Golledge (Eds.), Spatial and Temporal Reasoning in Geographic Information Systems. Spatial Information Systems pp. 143-154, Oxford University Press, New York.

Montello, D.R. (2005): *Navigation*. In: P. Shah & A. Miyake (Eds.), The Cambridge handbook of visuospatial thinking, pp. 257-294. Cambridge University Press.

Neisser, U. (1976): *Cognition and Reality - Principles and Implications of Cognitive Psychology*. Freeman, New York.

Norman, D. (1988): *The Design of Everyday Things*. Doubleday, New York.

Okabe, A., Boots, B., Sugihara, K., Chiu, N.S. (2000): *Spatial Tesselations, Concepts and Applications of Voronoi Diagrams (Second Edition)*. John Wiley & Sons.

Pinker, S. (1997): *How the Mind Works*. Penguin Books, London.

Preparata, F., Shamos, M.I. (1985): *Computational Geometry*. Springer-Verlag Berlin and Heidelberg.

Raubal, M. (1997): *Structuring Wayfinding Tasks With Image Schemata*. M.Sc. Thesis, University of Maine.

Raubal, M. (2001): *Agent-based Simulation of Human Wayfinding: A Perceptual Model for Unfamiliar Buildings*. Ph.D. Thesis, Vienna University of Technology.

Rehrl, K., Leitinger, S., Gartner G. (2007): *The SemWay Project - Towards Semantic Navigation Systems*. In Proceedings of the 4th International Symposium on LBS & TeleCartography, Hong Kong.

Rüetschi, U.J. and Timpf, S. (2005): *Modelling wayfinding in public transport: Network space and scene space*. In: Freksa, C., Knauff, M., Krieg-Brückner, B., Nebel, B., Barkowsky, T. (Eds.) Spatial Cognition IV: Reasoning, Action, Interaction; International Conference Frauenchiemsee. LNCS (LNAI), vol. 3343, pp. 24–41. Springer, Heidelberg.

Serra, J. (1982): *Image Analysis and Mathematical Morphology*. Academic Press, London.

Siegel, A. and White, S. (1975): *The development of spatial representations of large-scale environments*. In: H. Reese (Ed.), Advances in child development and behavior. 10, pp. 9-55, Academic Press, New York.

Stark A., Riebeck M., Kawalek J. (2007): *How to Design an Advanced Pedestrian Navigation System: Field Trial Results*. IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications. Dortmund, Germany.

Tedeschi, B. (2000): *Cyberscout: Turn Left at the Shoe Store...*. In NYTimes on the web.

Timpf, S., Volta, G., Pollock, D. and Egenhofer, M. (1992): *A Conceptual Model of Wayfinding Using Multiple Levels of Abstraction*. In: A. Frank, I. Campari, and U. Formentini (Eds.), GIS - From Space to Territory: Theories and Methods of Spatio-Temporal Reasoning, Pisa, Italy, pp. 348-367.

Tolman, E. (1948): *Cognitive maps in rats and men*. Psychological Review 55, pp.189-208.

Tversky, B. (1993): *Cognitive Maps, Cognitive Collages, and Spatial Mental Model*. In: A. Frank and I. Campari (Eds.), Spatial Information Theory: Theoretical Basis for GIS. Lecture Notes in Computer Science 716, pp. 14-24, Springer Verlag, Heidelberg-Berlin.

Vanetti, E., and Allen, G. (1988): *Communicating Environmental Knowledge: The Spatial Impact of Verbal and Spatial Abilities on the Production and Comprehension of Route Directions.* Environment and Behavior 20(6), pp.667-682.

Walter, V., Kada, M., Chen, H. (2006): *Shortest path analyses in raster maps for pedestrian navigation in location based systems*. International Symposium on “Geospatial Databases for Sustainable Development”, Goa, India, ISPRS Technical Commission IV (on CDROM).

Webster’s Unabridged Dictionary (1975).

Wechsler, D. (1944): *The measurement of adult intelligence* (3rd ed.). Baltimore: Williams & Wilkins.

Wikipedia: The Free Encyclopedia. Available online as: http://www.wikipedia.org.

Zubin, D. (1989): Oral presentation, NCGIA Initiative 2 Specialist Meeting, Santa Barbara, CA. Reported in Mark, D. (Ed.), Languages of Spatial Relations: Researchable Questions & NCGIA Research Agenda, NCGIA Report 89-2A, NCGIA.

Attachment I: Algorithm

```
%File pt_final.m
%Main algorithm

close all
[FILE,map] = imread('Raster-FMZK.tif');
I = ind2gray(FILE,map);

f = I;
f(f==203) = 1;
f(f~=1) = 0;
f = imfill(f, 'holes');
budynki = f;

% Highlight navigationable areas

g = I;
g(g==205) = 1;
g(g==219) = 1;
g(g~=1) = 0;

%Intersect Innenhofs with Navigationable areas
% var f - buildings image; g - navigationable areas image

f(f==1) = 10;
BW = f + g;
BW(BW==11) = 0;
BW(BW==10) = 0;
%imshow(BW,[0;1]);

Ltr=I==83;

%figure; imshow(BW,[])
%figure; imshow(Ltr,[])

Lobszary_dop=xor(imclose(BW,ones(18)),BW);
figure; imshow(BW,[])

Lobszary_linie_tr=Lobszary_dop&Ltr;
figure; imshow(Lobszary_linie_tr,[])



LolrBW=bwlabel(Lobszary_dop);
Lt=BW;
h = waitbar(0,'Please wait...');
for i=1:max(LolrBW(:))
   Li=LolrBW==i;
   Lii=Li&Ltr;
   if sum(sum(Lii))>0
    Lt=or(Lt,Li);
   end
waitbar(i/max(LolrBW(:)),h)
end
close(h)
figure; imshow(Lt,[])
```

```
skelet= bwmorph(Lt, 'skel', Inf);
figure;imshow(skelet,[])
spurred=bwmorph(skelet,'spur',Inf);
figure;imshow(spurred,[])



SE1 = [0 0 0; 1 1 1; 0 1 0];
SE2 = [0 1 0; 0 1 1; 0 1 0];
SE3 = [0 1 0; 1 1 0; 0 1 0];
SE4 = [0 1 0; 1 1 1; 0 0 0];
SE5 = [0 1 0; 0 1 0; 1 0 1];
SE6 = [1 0 1; 0 1 0; 0 1 0];
SE7 = [1 0 0; 0 1 0; 1 0 1];
SE8 = [1 0 1; 0 1 0; 0 0 1];
SE9 = [1 0 1; 0 1 0; 1 0 0];
SE10= [0 0 1; 0 1 0; 1 0 1];


L1 = bwhitmiss(spurred,SE1);
L2 = bwhitmiss(spurred,SE2);
L3 = bwhitmiss(spurred,SE3);
L4 = bwhitmiss(spurred,SE4);
L5 = bwhitmiss(spurred,SE5);
L6 = bwhitmiss(spurred,SE6);
L7 = bwhitmiss(spurred,SE7);
L8 = bwhitmiss(spurred,SE8);
L9 = bwhitmiss(spurred,SE9);
L10 = bwhitmiss(spurred,SE10);

LF = L1+L2+L3+L4+L5+L6+L7+L8+L9+L10;
LF(LF>1)=1;
imshow(LF,[]);
LLF=bwmorph(LF, 'shrink', Inf);
figure; imshow(LLF,[]);

% Highlight streets

str_ = I;
str_(str_==189) = 1;
str_(str_~=1) = 0;
strclose=imclose(str_,ones(18));
skelet_str=bwmorph(strclose, 'skel', Inf);
spurred_str=bwmorph(skelet_str,'spur',Inf);
figure; imshow(spurred_str,[]);

L1_str = bwhitmiss(spurred_str,SE1);
L2_str = bwhitmiss(spurred_str,SE2);
L3_str = bwhitmiss(spurred_str,SE3);
L4_str = bwhitmiss(spurred_str,SE4);
L5_str = bwhitmiss(spurred_str,SE5);
L6_str = bwhitmiss(spurred_str,SE6);
L7_str = bwhitmiss(spurred_str,SE7);
L8_str = bwhitmiss(spurred_str,SE8);
L9_str = bwhitmiss(spurred_str,SE9);
L10_str = bwhitmiss(spurred_str,SE10);
```

```
LF_str =
L1_str+L2_str+L3_str+L4_str+L5_str+L6_str+L7_str+L8_str+L9_str+L10_str;
LF_str(LF_str>1)=1;
figure; imshow(LF_str,[]);
LLF_str=bwmorph(LF_str, 'shrink', Inf);
figure; imshow(LLF_str,[]);

LLF_final = LLF_str + LLF;
figure; imshow(LLF_final,[]);
[c,r,v] = find(LLF_final == 1);
hold on
voronoi(r,c)
```

```
%File fill_gaps_2.m
%Filling gaps v.2

figsize = get(0, 'DefaultFigurePosition');
[BWcrop, rect] = imcrop(BW);
figure('Position',[360, 278, 500, 500]);
imshow(BWcrop);
h = fspecial('gaussian',25,15);
BWroi = roipoly;
BWcrop = roifilt2(h,BWcrop,BWroi);
```