FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

Automated Classification of Road-Surface Types Based on Crowd-Sourced Data

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieurin

im Rahmen des Studiums

Visual Computing

eingereicht von

B.Sc. Silvana Podaras
Matrikelnummer 0605836

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Associate Prof. Dipl.-Ing. Dipl.-Ing. Dr.techn. Michael Wimmer
Mitwirkung: Univ.Prof. Dipl.-Ing. Dr.techn. Norbert Pfeiffer
Senior Scientist Dipl.-Ing. Dr.techn. Markus Hollaus
Univ.Ass. Dipl.-Ing. Martin Wieser

Wien, 1. Jänner 2001

Silvana Podaras	Michael Wimmer

FAKULTÄT FÜR !NFORMATIK

Faculty of Informatics

Automated Classification of Road-Surface Types Based on Crowd-Sourced Data

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieurin

in

Visual Computing

by

B.Sc. Silvana Podaras
Registration Number 0605836

to the Faculty of Informatics

at the TU Wien

Advisor: Associate Prof. Dipl.-Ing. Dipl.-Ing. Dr.techn. Michael Wimmer
Assistance: Univ.Prof. Dipl.-Ing. Dr.techn. Norbert Pfeiffer
Senior Scientist Dipl.-Ing. Dr.techn. Markus Hollaus
Univ.Ass. Dipl.-Ing. Martin Wieser

Vienna, 1st January, 2001

Silvana Podaras	Michael Wimmer

Technische Universität Wien
A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.ac.at

Erklärung zur Verfassung der Arbeit

B.Sc. Silvana Podaras
Böckhgasse 6/3/16, 1120 Wien

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 1. Jänner 2001

Silvana Podaras

Kurzfassung

In dieser Diplomarbeit wird eine Methode zur automatisierten Klassifizierung von Straßen- und Wegoberflächen präsentiert, welche Radfahrern einen Überblick über die Wegbeschaffenheit entlang einer beliebigen Route geben soll.

Die automatische Klassifizierung von Landflächen ist ein aktives Forschungsfeld und befasst sich vor allem mit der Klassifizierung auf Basis von digitalen Luft- und Satellitenbildern. Die Klassifizierung von Straßenoberflächen anhand solcher Fotos birgt spezielle Herausforderungen, denn Straßen sind auf diesen Fotos nur wenige Pixel breit, wodurch die Anwendung gängiger Bildanalyse-methoden nur beschränkt möglich ist. Probleme bereiten unter anderem Mischpixel, die nicht einer einzelnen Oberflächenklasse zuordbar sind. Objekte wie Bäume oder Autos, die die Straßen überlagern, stellen ein weiteres Problem dar, da die eigentliche Straßenfläche oft nur unzureichend vom Rest des Bildes isoliert werden kann somit das Klassifizierungsergebnis negativ beieinflusst wird. Weiters sind Luftbildaufnahmen in hoher Auflösung nur mit einer geringen Anzahl an spektralen Bändern verfügbar.

Die hier präsentierte Methode verfolgt einen alternativen Weg basierend auf open-source Daten, wobei das Hauptaugenmerk auf Daten von OpenStreetMap (OSM) liegt. OSM ist ein online Projekt, bei dem geographische Informationen von Nutzern freiwillig gesammelt und in Form einer digitalen Weltkarte kostenfrei zur Verfügung gestellt werden. OSM bietet seinen Nutzern die Möglichkeit, beliebige Zusatzinformationen in Form von textlichen Annotationen zu Straßen und Wegen hinzuzufügen. Mit Hilfe dieser sogennaten "Tags" ist es möglich, Oberflächeneigenschaften für zahlreiche Straßen abzuleiten. Dies geschieht mittels bewährter Methoden aus dem Bereich der Mustererkennung. Das System ist so gestaltet, das zusätzliche Daten (wie zum Beispiel Höheninformationen) in die Klassifizierung mit eingebunden werden können. Weiters findet die Klassifizierung in zwei verschiedenen Feinheitsstufen statt, für die eine entsprechende Taxonomie ausgearbeitet wurde.

Zur Evaluierung wurden Testgebiete in Österrich und Liechtenstein herangezogen. Auf dem groben Level wurden rund 90% der Straßen korrekt klassifiziert, auf dem feinstufigen Level rund 60%. Der Vorteil der hier präsentierten Methode ist, dass sie schnell und für Gebiete weltweit einsetzbar ist, vorausgesetzt, dass ausreichend OSM Daten für die jeweilige Region verfügbar sind.

Abstract

This thesis presents a method to automatically estimate road-surface types based on crowd-sourced and open source data to give cyclists an overview of the road conditions along a cycle route.

Automatic classification of land-cover has been an active research field in recent years and mainly focuses on the classification of areas based on digital satellite and aerial imagery. Performing classification of road-surfaces based on such images bears some special challenges because roads have a width of only a few pixels on these photos, which makes it difficult to successfully apply classical image-analysis methods. Problems are caused by mixed pixels, which do not belong to a single surface class exclusively. Due to objects occluding the street, like for example trees and cars, it is difficult to isolate the street's actual surface from the rest of the image. This biases the classification procedure and may cause faulty results. Furthermore, aerial images of high spatial resolution are only available with a small range of spectral bands.

This thesis proposes an alternative approach for road-surface classification by utilizing open source data with a focus on data from the project OpenStreetMap (OSM). OSM is an online mapping project which collects geographical data and makes it available freely by providing a digital world map. Data is collected by users on a voluntary basis. OSM offers its users the possibility to add various properties to streets by making textual annotations. From these so-called "tags" it is possible to deduce road-surface properties for numerous roads by using methods from pattern recognition. The system is designed so it can be extended with additional data from other sources (e.g., height information) to improve classification results. Classification takes place at two levels, based on a coarse-to-fine-grained surface taxonomy.

The method was evaluated on different testing areas in Austria and Liechtenstein. At the coarse-grained level, up to 90% of streets were correctly classified. At the fine-grained level, up to 60% of streets were correctly classified. The advantage of the proposed method is that it is fast and applicable to regions worldwide at low cost, as long as sufficient OSM data for a certain region is available.

Contents

CHAPTER 1

Introduction

1.1 Motivation

This diploma thesis is part of a cooperation between the Institute of Computer Graphics and Algorithms and the Department of Geodesy and Geoinformation of TU Wien and the start-up company Bikemap. Bikemap offers a web service where users can create and share cycle routes around the world. Currently, Bikemap has a collection of around 3.1 million user-generated routes running in different terrains, which range from asphalted roads in cities to muddy tracks in the mountains [Bik].

For cyclists, it is an advantage to know which kind of surface they have to expect on those routes – for example to decide if their vehicle is suited for a certain route. Online map services like OpenStreetMap (OSM) [OSMn] provide information about the surface of roads, but it is not extensive because it is gathered by users manually. Instead of relying on the cumbersome process of manually annotating surface properties, an automated classification procedure for worldwide, extensive determination of surface types would be desirable.

Classifying the surface of roads is related to the problem of land-use and land-cover classification, which is often tackled with techniques from remote sensing. While aerial- and satellite imagery have already been used successfully for automated ground-cover classification, applying existing techniques to the classification of road-surfaces bears some special challenges. In aerial imagery, roads are long and thin elements, which are often occluded by vehicles or shadows cast on them. This makes the separation of the street from the rest of the image difficult.

Another challenge for classification is the mixed-pixel problem, which refers to the case where a single pixel represents more than one land-cover class. This happens when the spatial resolution of the image does not match the desired level of detail for classification. For such pixels, it is impracticable to apply a “hard” classifier, which assigns one distinctive class to a single pixel [MS05]. Increasing the spatial resolution may reduce the problem of mixed pixels, but also introduces other challenges: While aerial imagery has the advantage of being available in higher spatial resolution than satellite imagery, the spectral resolution is usually limited to the RGB and NIR bands.

But regardless of the chosen spectral and spatial resolution, image-analysis is of no use in forest areas, where roads tend to be mainly or completely occluded by trees.

Crowd-sourced data, on the other hand, may not be available in the same quantity as aerial images, but can provide surface information where image-analysis reaches its limits.

The strategy to gather land-cover information by crowd sourcing was already suggested in prior work. The collection and annotation of data could be presented in an entertaining way for individual users, for example by gamifying the process [SWPF13].

While a variety of crowd-sourcing projects for geographic data exist until today, the most important one with the biggest active community is OSM. OSM is not only the biggest online mapping project worldwide, but also tends to be more up-to date compared to aerial photography as it is updated several times a day, while aerial photographs are usually updated in yearly cycles. In addition to digital versions of the classical concept of a map, OSM provides detailed information on geographic objects thanks to a fine-grained tagging system. Those “tags”, which describe road types and usages in more detail, can be used to deduce road-surface properties automatically.

In addition to OSM data, other sources of useful open source data exist, such as height- and land-cover information from public databases. A system for road-surface classification should allow adding additional data in order to improve classification results.

1.2 Problem statement

The aim of this thesis is to combine the best of both worlds, and use a combination of crowd-sourced data and aerial imagery to develop methods for automatic classification of road surfaces.

While OSM data is up-to-date and available for free use, it tends to be highly heterogeneous regarding data quantity and quality among different regions. Data quantity is mostly dependent on how active the community is in a certain region, and some regions of the world tend to be more thoroughly mapped than others. While Europe (which is the core region of interest of the project partner, Bikemap) is one of the best mapped regions worldwide regarding roads, buildings and points of interest, this does not necessarily apply to the tags used to describe those geographical features. One challenge is to derive surface information for roads which are tagged with too little or no useful additional information.

Furthermore, the tagging system used in OSM is very extensive and rather inconsistent. Hundreds of possible tags exist, and though the users are advised to use a certain set of fixed tags for adding information, they always have the possibility to add custom tags where they think it is necessary. There are also no mandatory tags which have to be used, so the mere absence of a tag cannot be used to deduce certain assumption about the road's property – it is always possible that a user simply was not interested or too lazy to add additional information.

The problems described above make an extensive analysis of the available OSM data necessary: First, OSM data is crowd-sourced, so there is no guarantee hat the available data is actually correct. As evaluating the "trustworthiness" of OSM data is beyond the scope of this work, a review of related work on this topic is necessary. Then, tags which could be helpful for surface determination have to be identified. This is done both qualitatively by examining the OSM guidelines and quantitatively by evaluating the frequency of those tags for certain testing areas.

Regarding the analysis of aerial imagery, it is not guaranteed that images of sufficient resolution are available for all countries worldwide. Even when high-resolution images are available, it is not trivial to extract roads precisely due to occlusion by trees, cars and shadows cast by them. Furthermore, it is likely that images from different sources have different color properties, which is caused by different recording conditions and devices. Even when analyzing different streets from the same image, colors can vary due to shadows or different age of the paving. This has to be considered when using the spectral profiles for classification.

Finally, technical aspects have to be considered in order to combine the two heterogeneous data sources for classification. The crowd-sourced data has to be co-registered with aerial imagery in order to assign the results of the image classification to the ways represented in OSM.

1.3 Contributions

In this thesis, a method which estimates the surface type of a road by combining crowd-sourced OSM data with methods from remote sensing and image-analysis was developed and tested.

Classification is performed in two stages. First, OSM data is used to classify as many ways as possible by using a decision-tree classifier with the available tags. This phase is designed so it can be improved by additional data (also other sources than OSM) in the future. Image-analysis is performed as a second step for ways which couldn't be classified at all by the first phase, or for which the accuracy of the classification results is too low.

The developed procedure should be applicable worldwide, if OSM data and aerial images are available for this region.

The contributions of this work are summarized in the following list:

- An extensive qualitative and quantitative assessment of OSM data to find out which information can be used for determination of surface types. This includes a literature review to gain insights on trustworthiness of OSM data in the testing areas, a qualitative assessment of available tags and their usefulness for surface determination and a quantitative assessment of the occurrence of those tags in the testing areas.

- The design and evaluation of a classification method for road-surface types based on OSM data. This is done by utilizing a decision-tree classifier on preprocessed data based on the results of the OSM data evaluation.
- An analysis of the classified roads to verify the results of the classification step, which can also be used as a means of quality control. Furthermore, a graph-like neighborhood structure for the available OSM data is created, so that the surface properties of neighboring roads can be analyzed locally.
- The design and evaluation of a classification method for road-surface types based on aerial imagery, in order to supplement the results from the prior OSM analysis phase and to process roads which could not be classified in the OSM analysis phase due to a lack of relevant tags.

1.4 Structure

The work is structured as follows: In chapter 2, a general introduction to the theory of remote sensing, image classification and crowd-sourced geographical data is given. Chapter 3 presents prior work related to this thesis, with a focus on the domain of land-cover and road classification, because the prior chapter already covers current research on OSM data.

Chapter 4 starts with the presentation of the data analysis phase and gives a detailed taxonomy and description of the OSM data used. Then the developed algorithm is described, where important details of the implementation strategy are provided where considered necessary.

Finally, chapter 6 presents and discusses the results of the method, and also presents a list of possible future work and improvements.

CHAPTER 2

Fundamentals

As the contribution of this work is based on various core concepts, this chapter aims to give a brief introduction to them in order to facilitate the understanding and context of the following chapters. First, the topic volunteered geographical information, with a focus on OpenStreetMap, is covered, then remote sensing and spectral analysis are introduced, and finally an overview on relevant aspects in the field of pattern recognition is given.

2.1 Volunteered geographical information

For a long time, the creation of maps was in the hand of a small group of specialists such as cartographers. World-wide geographic data was not available for free to the general public, so acquiring this data usually involved high costs for individuals or companies.

This started to change with the beginning of the 2000s due to various reasons. First, selective availability of the GPS signal was abandoned, which made low-cost GPS receivers much more accurate. Before, the GPS signal for civilian use was intentionally downgraded due to security reasons. Second, GPS receivers became cheaper and thus more available for the mass market – nowadays, almost every smart phone is equipped with GPS. Finally, the introduction of the GPX-format (GPS eXchange format) facilitated sharing and exchanging GPS-tracked information via internet and computers [HW08].

Due to those developments, the internet-phenomenon of volunteered geographical information (VGI) emerged in the mid-2000s. The term was first introduced by Goodchild [Goo07], in order to highlight the spatial nature of this kind of crowd-sourced data. Users of a VGI platform voluntary collect, map and relate spatial data online on a virtual map and make it (sometimes freely) available for private persons or companies. This principle of user generated content is already known from other worldwide successful projects like Wikipedia or YouTube.

Contributing geographical data to a VGI platform can be done in various ways, the only condition being that the mapped data correlates to a real geographic position. One common method is to record GPS track points and upload them to the system. Another way which does not involve a GPS sensor is by tracing georeferenced aerial imagery [OSMk].

Nowadays, many different VGI-related projects exist, from bigger platforms like Wikimapia [Wik] to smaller ones like the German project Mundraub [Mun]. Each project normally has its own, distinctive mission statement or purpose. While Wikimapia aims to collect Links and Images and relate them to a geographical location, the purpose of Mundraub is to collect the location of fruit trees which can be accessed by the public.

One of the largest and oldest projects is OpenStreetMap (OSM), which primarily aims to provide free map-data to the general public. Because of its extensiveness and open-source nature, OSM was chosen as integral tool of this work and will be introduced in detail in the following subchapter [NZ14].

2.1.1 OpenStreetMap

OSM is not only one of the oldest, but also one of the most successful VGI projects with a very active community. Since it was founded in 2004, the self-proclaimed goal of the project is to provide a database of world-wide, license-free geographical information and to encourage people to make use of it. The data found on OSM can be used freely for various application needs, both commercially and non-commercially, as long as OSM is cited as source [OSMg], [OSMe]. A screenshot of the OSM website can be seen in figure 2.1. The website basically provides an interactive map interface, where users can search locations with textual input and explore the map by zooming or dragging. Detailed information about objects can be retrieved with the tools provided on the right side of the screen.

2.1.2 Data retrieval and sources

OSM has a very open approach when it comes to contributing data. First, a user has to register an account and provide a valid e-mail address. Immediately after the registration, he is allowed to add new or modify existing data and upload the changes to the global database. This open approach has the advantage that data can be kept up-to-date by updates happening in short intervals. Usually changes can be seen worldwide within less than an hour [OSMm]. The downside of this approach is that it also allows errors to spread quickly, and that it is easy for a malicious user to destroy data on purpose. To have some sort of control mechanism, the contributions of all users are logged in so-called changesets and OSM also keeps track of the last time an object was modified and which user edited it. The complete database is stored weekly as a backup [OSMo].

OSM users can collect data in many different ways and from numerous sources. In the early phase of the project, most data was mapped from scratch with GPS devices. A user simply went out with a GPS-capable device and tracked himself while wandering

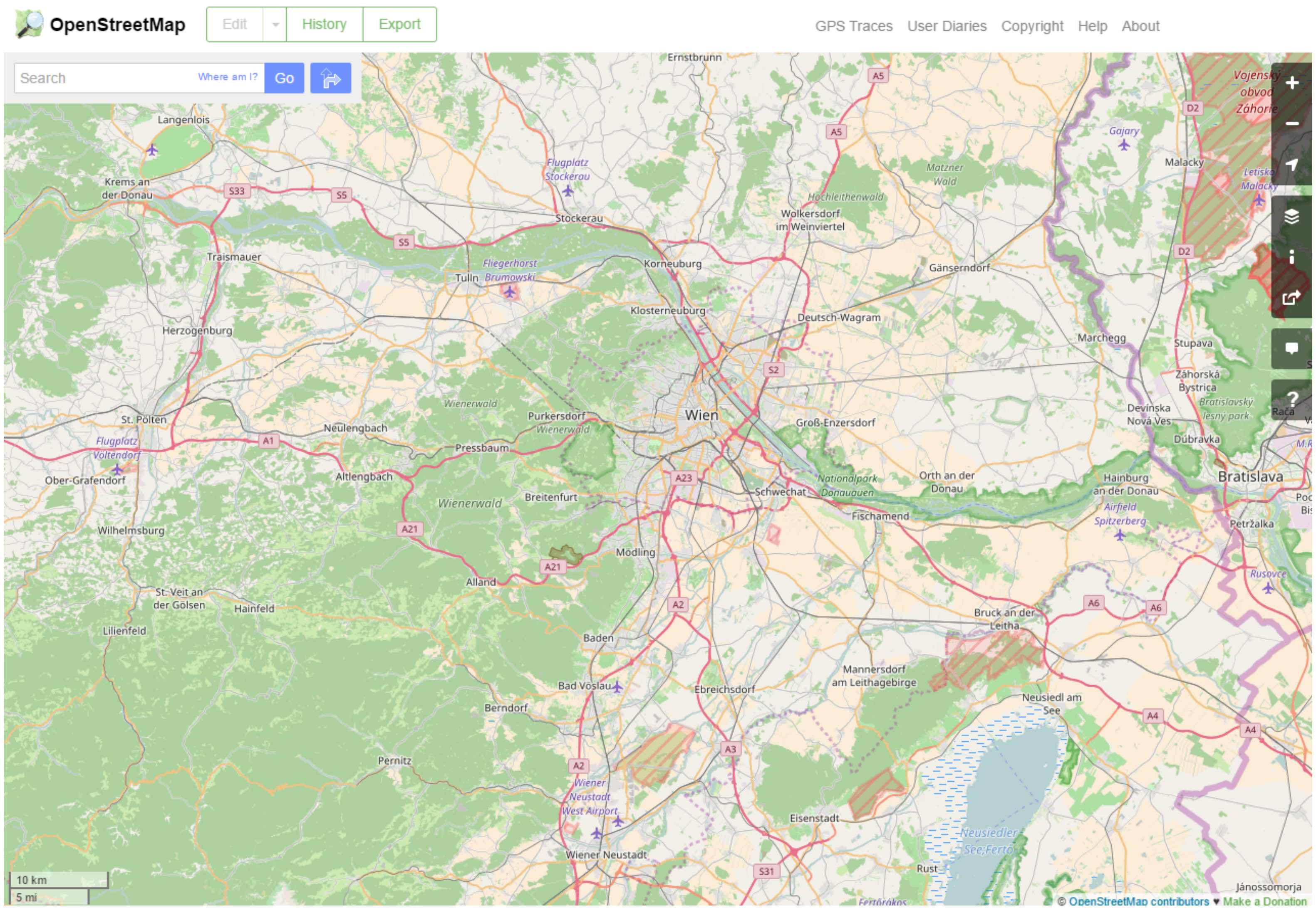


Figure 2.1: Screenshot of the OSM website, depicting the region around Vienna, Austria. Users can interactively explore the map by zooming, dragging or searching certain locations. Detailed information about objects can be retrieved with the tools provided on the right side of the screen. (Date: 15.3.2016)

around. When he finished his tour, he had to upload the track points to the OSM server and to connected them manually to form road networks, buildings and other points of interest out of these trackpoints. Later, also satellite imagery from “Yahoo!” [OSMq] and Airborne Imagery from various sources was made available for tracing. (Aerial imagery came mostly from Microsoft’s Service “Bing” [OSMa], but dependent on the country, eventually other sources were used.) Various editors, from online to offline solutions, exist to facilitate editing of OSM data [NZ14]. In some countries, public domain databases provided by companies or the state have been incorporated, as for example the TIGER Geodatabases from the U.S. Census Bureau, which provides information about many geographic features like streets, buildings and rivers for the United States.

2.1.3 Data quality

Because of the very open approach on collecting data, the question arises how accurate and reliable data found on OSM is. Various scientific studies about OSM focused on evaluating the quality of the data, and came to controversial conclusions.

Early studies focused on comparing the OSM database to data from commercial or governmental providers in regard of quantity and geometrical correctness.

Haklay [Hak10] did a comparison between OSM and the Ordnance Survey (OS) datasets for the regions London and England, where the OSM project was originally founded. The comparison method was to sum up the total amount of road length for both data sources. Those lengths were used as an indicator for completeness, so the set with more road length mapped was considered to be more complete. Haklay concluded that already 4 years after the beginning of the project, roughly 29% of the area of England were covered, and that around 80% of the motorways of the two datasets overlap. Just one year later, Haklay repeated the study and concluded that while the coverage in England has increased to 65% [Hak], there was still a lack of detailed information such as street names, and a quantitative heterogeneity between cities and rural areas.

Similar observations were made by Zielstra [ZZ10], who used the same comparison methods as Haklay, but focused on Germany and used data from a commercial provider, TeleAtlas, as counterpart to OSM. Again, OSM’s data was extensive in city areas and thus considered an alternative to products by commercial providers, whereas in rural areas the gathered information was still too sparse to be an alternative source of information. But one important observation made was the rapid growth rate of the OSM street network: Zielstra stated that within only 8 months, the difference between TeleAtlas and OSM shrunk down to 7% in total.

Other studies executed in the USA came to opposite conclusions regarding the differences between city and country side. They claimed that in rural areas, OSM data is much more accurate and up to date compared to commercial products [HW08]. Even though those observations seem contradictory, it is possible that both are correct – an explanation can be given when examining the OSM user community in further detail, which will be done in the following subchapter.

Figure 2.2: Comparison of the map for the Gaza Region found on Google Maps (left) to the one found on OSM (right). Even though comercially maintained, the map found on Google maps has far less information than the one found on OSM. (Date: 26.9.2016)

When comparing OSM to commercial products, it is also hard to make statements which apply globally, because it depends on both the region and the provider which information is more complete. A recent example for OSM data being much more complete compared to a commercial product is the Gaza region (see figure 2.2). Even though commercially maintained, the map found on Google maps has far less information than the one found on OSM.

2.1.4 OSM community worldwide

It is difficult to make statements regarding OSM data on a worldwide scale, because both quantity and quality of OSM data are highly dependent of the geographical location the data was retrieved for. When comparing the size and activity of the OSM community across different countries, it is far from being a homogenous group.

In general, bigger towns tend to be mapped more thoroughly than remote mountain areas. This is probably due to that technical requirements like hardware and internet access are more available in well-developed regions. There are exceptions from this rule, for example when some users were very active in a remote area and went so far as mapping hiking trails or paths in the mountains. For some remote regions with an exceptional amount of information, it is also possible that whole databases were provided by the state or private corporations, which explains a sudden “boost” of information in these areas.

In a worldwide comparison, Europe and the US tend to be mapped very thoroughly and with much detail, because the majority of OSM users lives on those continents [OSMb].

The biggest and most active OSM community is the German one, followed by the US, Russia, Italy and France [OSMb], [OSMi].

While OSM has over a million of users worldwide, a distinction has to be made between “registered” and “active” users. Not all users which are registered are actually contributing the the project – in fact, the majority has not mapped even a single node, while others are very active over a long period of time [Nei], [HW08], [NZ12].

A comparison between the overall distribution of active OpenStreetMap Contributors can be seen in figure 2.3, while a comparison regarding their activity in the last 12 months (from November 2015 to November 2016) can be seen in figure 2.4. The biggest and most active community is the German one, followed by the communities in the United Kingdom, Russia and France. Figure 2.5 shows the development of the OSM community worldwide from 2005 until 2016. While the number of registered users grew rapidly in the last 4 years of this period, only a fraction of the members is actively contributing to the project.

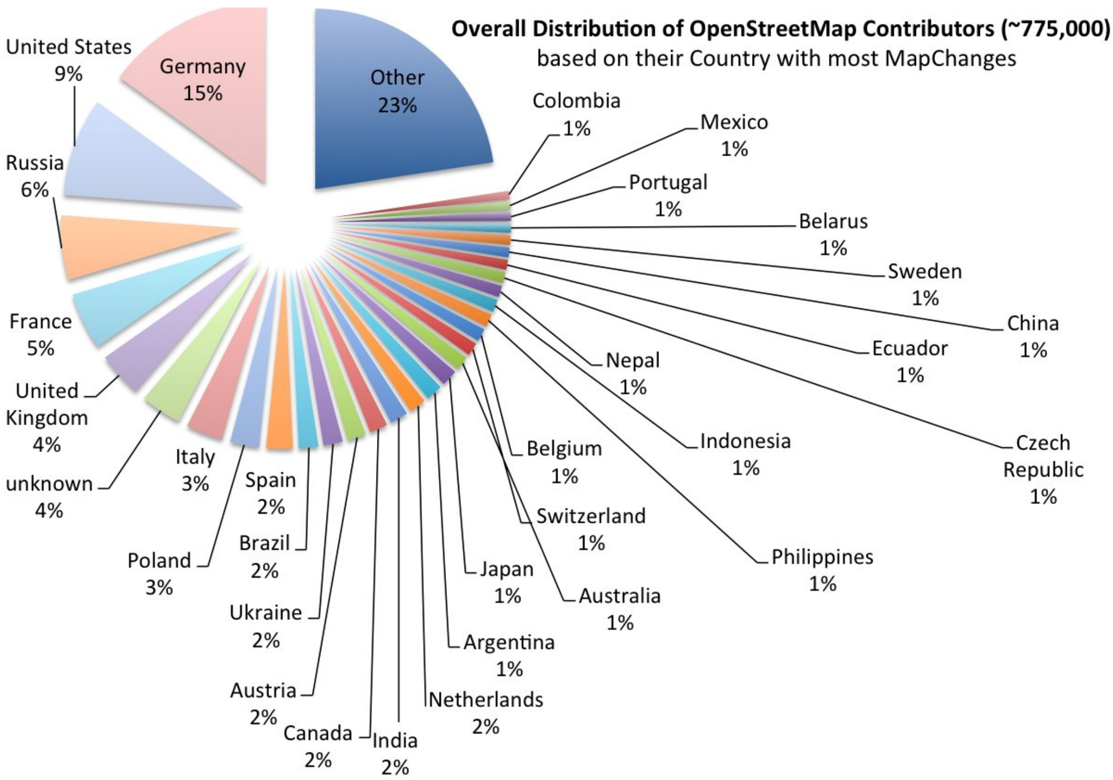


Figure 2.3: A comparison of OpenStreetMap contributors between countries since the project start. The German community is by far the biggest, followed by the US, Russia and France. Copyright by Pascal Neis [Twi].

Maron [Mar] tried to gain more insight on global differences by comparing the total sum of road lengths for each country in OSM with the total lengths listed in the “CIA

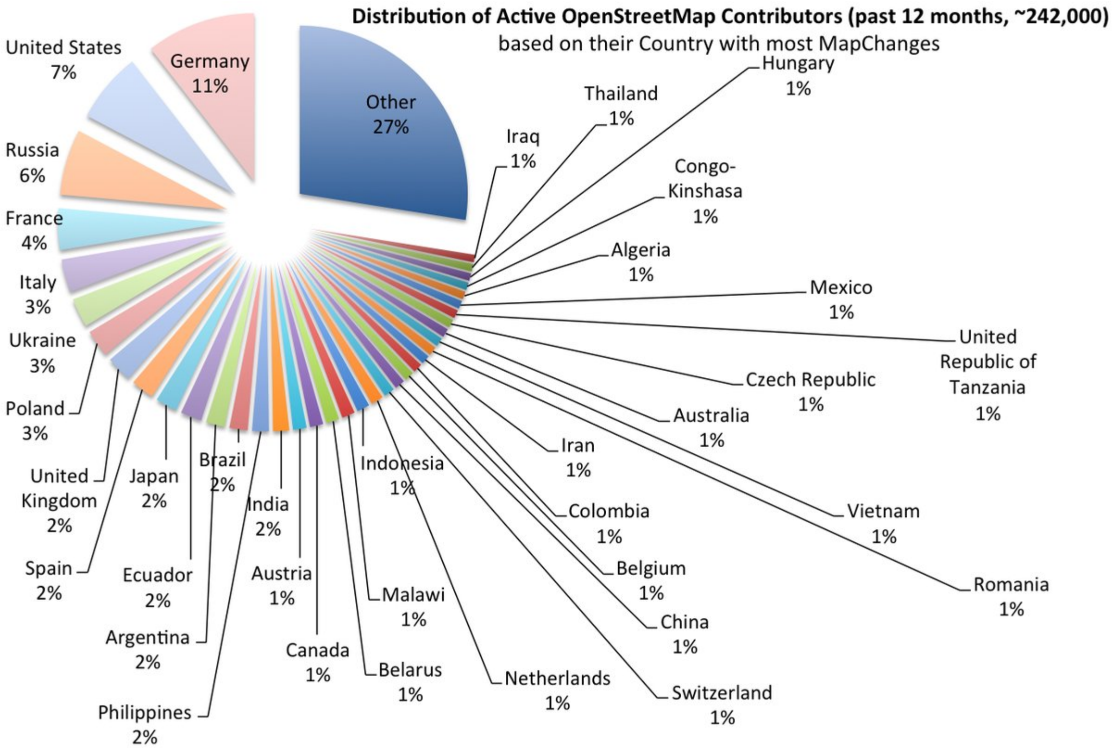


Figure 2.4: Comparison of active OpenStreetMap contributors by country, with “active” being defined as adding changes to the map in the recent period from November 2015 to 2016. Copyright by Pascal Neis [Twi].

World Factbook”, which is a freely available source of information published by the US government [CIA]. It provides a list of the total lengths of each countries road network, where both paved and unpaved roads are considered. The factbook does not define in detail what constitutes a road (e.g it is not clear if only roads suited for vehicles count or not), while Maron makes statistics both for roads suited explicitly for vehicles and all roads. He comes to the conclusion that OSM “performs well across the continents”, with some countries in Europe covering even more road length than listed in the factbook. (This can happen because in OSM, a broad mixture of roads including hiking trails or bicycle roads are mapped.) Just like prior studies, Maron also observes that data quantity and quality is dependent on the region, and mentions India and China as two examples which are still far away from reaching the kilometer count listed in the factbook.

So when utilizing OSM data, one should always keep this the regional dependency of the data in mind, and evaluate if OSM is the best choice for the region in which data is needed. Those regional differences should also be kept in mind before taking the results of scientific studies for granted. The temporal aspect is important, because even a relatively small time span of a few months can lead to a big growth of available information in the OSM database. The studies by Haklay and Zielstra are no exception, even though they

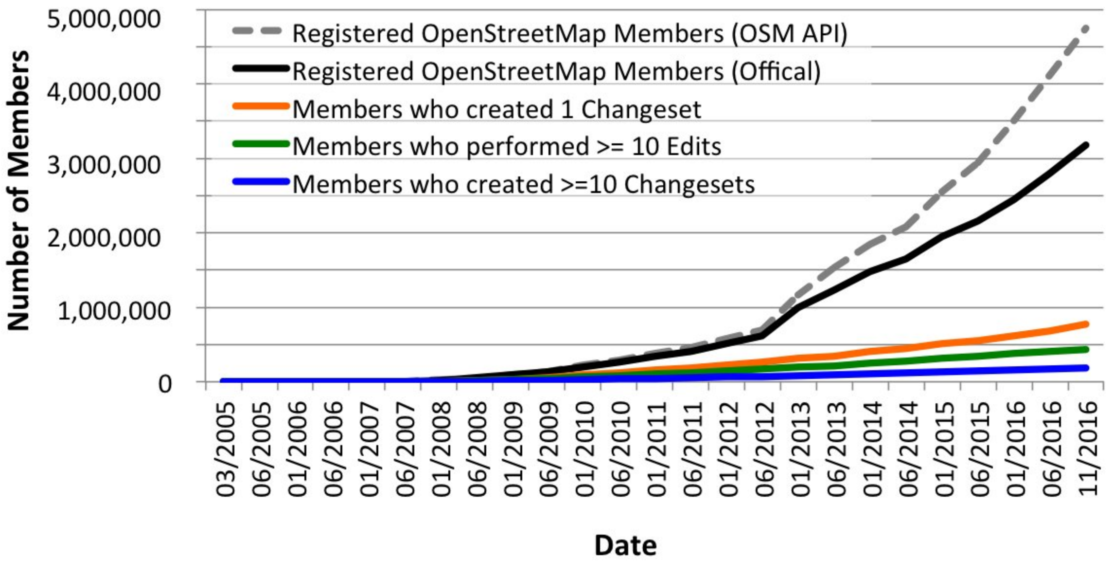


Figure 2.5: Development of the OSM community from 2005 until 2016, with a comparison between registered and active members. While the community grew rapidly in the last 4 years, only a fraction of the members is actively contributing to the project. Copyright by Pascal Neis [Twi].

are very extensive, one should keep in mind that they are over 6 years old. Furthermore, apart from mere road lengths, tag quantity is also a factor that should be considered when analyzing OSM's data. Tags are used to add various additional information to the map, and will be introduced together with OSM data representation in the following subchapter.

For this work, it is crucial that Europe is among the best mapped regions in the world, as Bikemap, the project partner, primarily operates in this region. Because there is a particular interest in analyzing tags which relate to surface properties, an evaluation of the testing areas regarding the quantity of such tags was performed and will be presented in later chapters.

2.1.5 Data Representation

In OSM, real-life geographical data is represented by three types of data elements, which are namely nodes, ways and relations. In the OSM database, all basic elements have some common attributes, like a unique ID, a reference to the user who created the element, a timestamp which indicates the last date when the object was modified and information to keep track of the version of the object. For understanding how the elements are used in OSM, a short overview based on the information found on the OSM Wiki is given [OSMd]:

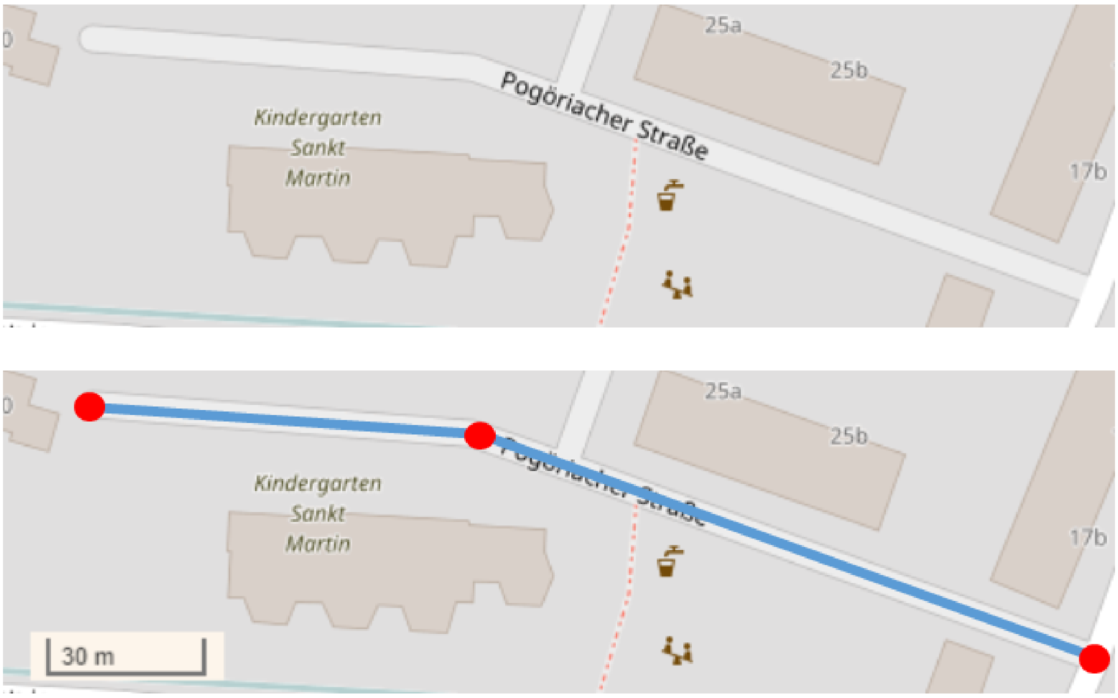


Figure 2.6: Top: A screenshot of an OSM way as it can be seen when viewing the map on www.openstreetmap.org. Bottom: Visualization of the data elements. The way itself is highlighted in blue, the red circles represent its nodes. (Date: 11.2.2016)

Nodes define points in space and are identified by their longitude and latitude coordinates. Nodes can either be part of a way or relation, or mark stand-alone points of interests in space. For example, a single node can be used to represent a tree or a telephone box.

Ways are a sequential (ordered) list of nodes which can also be part of a relation. Linear features such as roads, railway lines or rivers are represented by open ways. If the start- and endpoint of a way are equal, the way represents either a closed way (like a roundabout) or an area (like a building or a park). A distinction can only be made by considering the tags attached to a closed way.

Relations are used to model geographical relationships between multiple data elements. Thus, relationships have a list of nodes, ways or other relations as their members. For example, all ways of a national bicycle road network could be members of a relation, or all stops of a bus route could be members of a relation.

For visualization purposes, figure 2.6 shows a screenshot of a way with its nodes highlighted in red.

2.1.6 Tags

OSM uses a tagging system to describe features of ways, nodes and relationships in more detail. A tag consists of a key-value pair which describes one specific attribute of a data element. Each element can have one or more tags attached to it which describe attributes such as its name or usage. While there is no limit on how many tags can be assigned to a data element, each collection of tags for an element can contain a specific key only once.

Both key and value are text fields and separated by an equals sign, so for the human reader, a tag would be build up like this: “key=value”. For example, a cinema could be represented by a node which has the tag “amenity=cinema” attached to it. Because multiple tags per data element are allowed, the cinema can also have a name tag and a tag which denotes the operator.

Figure 2.7 shows how a way and its properties can be displayed in the browser. The panel on the left of the screen lists the tags and nodes of the currently selected way.

The tagging system in OSM is very extensive and flexible, because the set of available tags is not fixed, but instead each user can come up with his own tags if needed. To prevent users from creating masses of rarely-used tags and to introduce more consistency to the database, the OSM community has agreed on conventions on how specific tags are used for common purposes. Any user can propose and discuss the introduction of new tags [OSMf].

A list of the most common existing tags and a documentation of their intended usage can be found in the OSM wiki [OSMp] and on the website Taginfo [Tag]. Both websites are powered by the OSM community and basically work as a documentation for everything related to the OSM project. While the wiki provides general information for its users, e.g. how the data model works and how to edit data, Taginfo has the mere purpose to document and provide statistics about available tags.

One important task was a through analysis of the available OSM tags with respect to their usefulness for this work. Both quality and quantity were relevant aspects, as we not only need tags from which surface properties can be deduced, but also a relevant amount of them in order to classify ways with a high accuracy. A detailed overview on the chosen tags and how they were utilized to analyze different road-surface classes is given in chapter 4.

Figure 2.8 shows a screenshot of editing a way in the browser: on right part of the screen, the map is visible and the ways nodes can be manipulated. On the left of the screen, there is a panel which allows to add tags to the way.

2.1.7 Data structure

OSM data can be accessed in several ways by different APIs. For this project, the data was downloaded in the .osm file format, which is also known as OSM XML, where each OSM element is represented by an XML entity. For illustration purposes, a very simple .osm-file is depicted in listing 2.1.

Figure 2.7: Viewing an OSM way in the browser. When selecting a way, it is highlighted and tags are listed in the panel on the left side of the screen. (Date: 15.3.2017)

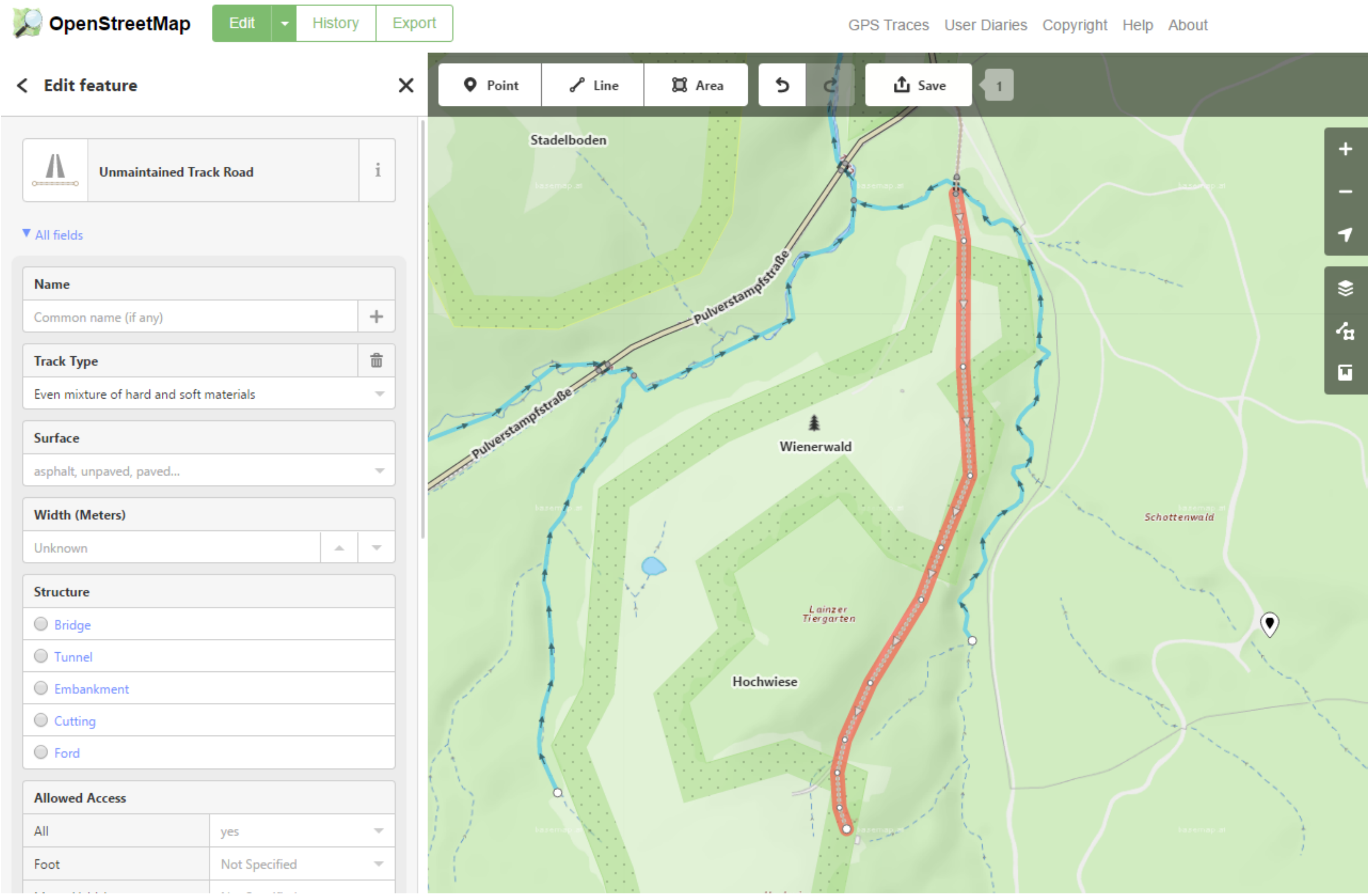


Figure 2.8: Editing an OSM way in the browser. The left panel on the screen allows editing tags of the currently active way, which is highlighted in light red on the map. The nodes, which are marked by little circles, a way consists of are also visible and can be edited. (Date: 15.3.2017)

The file provides a list of all OSM data primitives in an area described by the "bounds" entity. The data primitives are organized in blocks: first come the nodes, then the ways and finally relations.

Each primitive stores common attributes like its unique OSM ID, but also meta information on the last user who edited this element and a time stamp when the last change was done. If any tags are assigned to the element, they are nested as a list of child elements.

Listing 2.1: A simple example for an .osm-file.

```
<?xml version="1.0" encoding="UTF-8"?>
<osm version="0.6" generator="CGImap␣0.0.2">
 <bounds minlat="54.0889580" minlon="12.2487570" maxlat="54.0913900"
     maxlon="12.2524800"/>
 <node id="298884269" lat="54.0901746" lon="12.2482632" user="SvenHRO
     " uid="46882" visible="true" version="1" changeset="676636"
     timestamp="2008-09-21T21:37:45Z"/>
 <node id="261728686" lat="54.0906309" lon="12.2441924" user="
     PikoWinter" uid="36744" visible="true" version="1" changeset="
     323878" timestamp="2008-05-03T13:39:23Z"/>
 <node id="1831881213" version="1" changeset="12370172" lat="
     54.0900666" lon="12.2539381" user="lafkor" uid="75625" visible="
     true" timestamp="2012-07-20T09:43:19Z">
  <tag k="name" v="Neu␣Broderstorf"/>
  <tag k="traffic_sign" v="city_limit"/>
 </node>
 ...
 <node id="298884272" lat="54.0901447" lon="12.2516513" user="SvenHRO
     " uid="46882" visible="true" version="1" changeset="676636"
     timestamp="2008-09-21T21:37:45Z"/>
 <way id="26659127" user="Masch" uid="55988" visible="true" version="
     5" changeset="4142606" timestamp="2010-03-16T11:47:08Z">
  <nd ref="292403538"/>
  <nd ref="298884289"/>
  ...
  <nd ref="261728686"/>
  <tag k="highway" v="unclassified"/>
  <tag k="name" v="Pastower␣Strasse"/>
 </way>
 <relation id="56688" user="kmvar" uid="56190" visible="true" version
     ="28" changeset="6947637" timestamp="2011-01-12T14:23:49Z">
  <member type="node" ref="294942404" role=""/>
  ...
  <tag k="name" v="Bus␣Linie␣123"/>
  <tag k="operator" v="Regionalverkehr"/>
 </relation>
 ...
</osm>
```

2.2 Remote sensing

Remote sensing is the process of retrieving information about an object, area or phenomenon without getting into physical contact with it. There are numerous applications for remote sensing, of which many are related to the field of earth sciences. A distinction between two types of technologies can be made, namely passive and active sensors. Active sensors emit a signal and record its reflection, while passive sensors record an existing observable phenomenon. The latter one is the more common type, and digital photography is one example from everyday life which can be used to illustrate this technique:

Digital photography works analogue to the human visual system, which reacts to electromagnetic waves of a certain frequency. What a human observer "sees" are electromagnetic waves being reflected, absorbed or transmitted into the eye by his surroundings, and the human brain processes this information to form an "image". A digital camera mimics this process with a lens and a digital image sensor, which records and digitizes the portion of the electromagnetic spectrum visible for humans in order to represent it numerically by the Red, Green and Blue (RGB) channels of a digital image.

The remotely sensed data relevant for this work are spatial high-resolution images of the earths surface, which are usually retrieved by mounting sensors on satellites or air crafts. Those sensors record portions of the electromagnetic spectrum reflected by the earths surface.

Objects and materials reflect a broader spectrum of electromagnetic waves than those which are visible to the human eye, so sensors are not limited to record monochromatic images or the RGB color bands only. Multi- and hyper-spectral imaging sensors can record hundreds of different bands, including near infrared (NIR) or shortwave infrared bands (SWIR).

Because each material reflects the electromagnetic spectrum in a particular way, it is possible to characterize it by its so-called spectral signature. By examining those spectral signatures for individual or clusters of homogenous pixels which represent an object, it is possible to identify and classify the materials in a scene – this technique is also known as spectral analysis [LKC14], [CW11].

Practically, remote sensing and spectral analysis have many applications where an automated and non-invasive analysis is desired. Shaw [GAS03] distinguishes three generic application classes:

Anomaly detection The aim of anomaly detection is to search for rare or anormal spectral signatures among the pixels of a spectral image. One example application is the detection of man-made materials in a natural environment.

Target recognition Given a priori information about the spectral characteristics which could occur in a scene, target recognition aims to find pixels in the image which have a matching spectral signature. Applications like classifying ground-cover or determining the condition of asphalted streets in a road network fall into this category.

Background characterization Instead of detecting and identifying isolated regions in an image, the aim is to characterize a scene as a whole. One example application is the analysis of shallow water along coastlines.

With regard to Shaw's classification scheme, this work falls into the second category, target classification.

2.2.1 Airborne Imagery and Satellite Imagery

The first attempts to acquire images of the earths surface were made in the mid of the 19th century. Those photos were taken from balloons and kites until the airplane was invented in the beginning of the 20th century. The interest in airborne imagery for military purposes reached its peak during World War I and World War II [LKC14]. In the 1960s, the first satellites for remote imagery became available. In the beginning, only black-and-white images of rather coarse resolution were available, but the rapid technical development over the past 50 years led to major improvements. Nowadays, modern satellites like WorldView-3 can create images with a resolution of up to 31cm per pixel (Panchromatic), or 1.24m (Multispectral) [WV3], EROS-C can is capable of 40cm (Panchromatic) or 80cm (RGB + NIR) per pixel [ERO].

For the classification of road-surfaces, the currently available resolutions from satellite imagery are still not sufficient. Roads are narrow elements, where an average highway lane has a width around 3 meters, while other road types like cycle lanes or hiking trails tend to be much more narrow. In an image with a resolution of 1.24m per pixel, a street would cover around 2 pixels in width. This is not fine enough to extract spectral and textural information to classify the surface.

Therefore, aerial photographs are the current option for classification of such fine-grained features. Airborne images are created by mounting a camera on an aircraft instead of a satellite and thus provide a much higher image resolution of up to 10cm per pixel [Geoa]. Contrary to satellite images, airborne images usually provide not as many bands, with RGB and NIR being the most common ones.

Naturally, the region which a photo can cover is limited, so in order to record an extensive area by remote imagery, it is necessary to use several photos from one or more sources and tile them to gain the complete picture. Those "tiles" are usually recorded over a longer period of time and under different conditions, so one and the same ground cover type is likely to have different spectral characteristics in tiles from different sources. Furthermore, the creation of aerial images often lies in the countries' national competence, so the available image quality and resolution can differ significantly between countries. This variation in quality makes it necessary to extract spectral profiles of materials for each source individually, and thus adds additional complexity to image classification tasks.

Figure 2.9: Part of the orthofoto for the testing area of Liechtenstein with a resolution of 10cm per pixel. Right: Image at the zoomfactor of 10%. Right: Closeup of the region inside the red rectangle at full size.

Aerial imagery can be acquired from different providers, both commercial- and non-commercial services exist. Apart from the cost, quality (resolution) and update-cycles should be considered when choosing a provider.

In Austria, both the "Bundesamt fuer Eich- und Vermessungswesen" (BEV) which is maintained by the state and the private company "Geoimage" offer high-resolution aerial images. The BEV claims to provide a resolution of around 20cm per pixel and Geoimage claims a resolution between 12.5cm to 20cm per pixel on average, depending on the region. The images of both providers come from the same flights. A third of Austria's surface is updated a year, which means that the surface of the whole country is updated in a three-year cycle [Geob], [BEV]. An example for an orthophoto of Liechtenstein with a resolution of 10cm per pixel can be seen in figure 2.9.

Orthofotos

Photos from remote sensing are usually acquired by using a perspective projection, which causes several distortions, such as non-uniform scale and relief displacements. For GIS applications, those images have to be geometrically corrected so that they have uniform scale and show objects in their correct location. Those rectified images are called orthophotos and have the same characteristics as a map, so they can be used to measure "true" distances. An image of an aerial photo before and after rectification can be seen in figure 2.10.

For rectification, a digital elevation model (DEM) and the intrinsic and extrinsic camera parameters are needed. Then, mathematical mapping techniques can be applied to

Figure 2.10: Left: Orthofoto before rectification. Right: Orthofoto after rectification. Copyright by Nielsen [dN04].

transform the image and gain the correct, undistorted representation.

The images used for the practical part of this work were already orthorectified by their providers, so it was not necessary to consider this step explicitly in the implementation process, therefore the rectification procedure will not be described in more detail.

2.2.2 Limitations and challenges

While the concept of spectral analysis of images may seem straight-forward, there are various challenges which have to be considered when applying these techniques in practice. Those considerations include the spatial and spectral resolution of the sensor, recording conditions and spectral variability.

- Recording Resolution: Dependent on the application, a minimum image resolution is required. For this work, a resolution where streets would cover a single pixel in width would be useless, because it would be impossible to extract enough pixels and textural information for analysis of the surface. Another problem are pixels which represent multiple land-cover classes, so called “mixed pixels”. This problem occurs when the spatial image resolution is too low in relation to the desired level of detail for classification [MS05]. When using images of very high resolution, the challenge is to cluster homogenous pixels which represent an object in the image before performing classification.

- Recording conditions: Spectral information is heavily dependent on the external conditions during the recording process. Factors such as the recording angle, daytime, lighting conditions, atmospheric effects like scattering and absorption or motion of the sensor or objects in the scene can bias the recorded image [GAS03].

- Spectral variability: In the early years of spectral analysis, researchers assumed that each material has a clearly defined spectral signature, which is defined by the materials reflectance properties only and thus suited to distinguish materials from another. However, there is a lot of variability in the reflectance spectrum of most materials. Uncompensated errors and bias introduced in the recording process are one source of variability. But also factors like the age of an material or reflections made by its surroundings can be a cause [GAS03].

  In other words, it is not possible to determine a "universal" signature for a certain material. To illustrate this with an example, imagine that asphalted roads have to be classified in orthophotos from various sources. Because it is likely that the photos were taken under different lighting conditions with different sensors, the average RGB and NIR values for asphalt differ from image source to image source. So for each data source, an individual signature representing asphalt has to be determined.

  The spectral signature for a material is usually determined by taking multiple samples of regions which are known to have this particular surface type and averaging their spectral responses. When aiming to classify multiple material classes in an image, it is thus necessary to build a spectral library, which stores the spectral signature of each material, which is then used as reference for classifying the unknown regions of an image.

Using remote sensing to analyze the materials of a road network bears special challenges, because when photographed from above, roads are the "bottom elements" of our 3-dimensional world. They can be covered by trees, cars, houses or the shadows cast by them. Those influences from nearby geometry can cause a significant variability of the RGB values in one material within one single datasource.

Furthermore, roads are thin and narrow elements, which makes a precise extraction from the photographs difficult. When parts of the surrounding are also extracted and mixed with the pixels which represent the roads, a misleading spectral response is created and analyzed, which may result in a faulty or wrong classification [HGNR03].

A detailed description on the approach taken to evaluate the spectral signatures for materials of interest is given in chapter 4.

2.3 Pattern recognition and machine learning

Pattern recognition and machine learning are a subfields of computer science and focus on the development of methods which imitate the process of human learning and recognition. Contradictory to the classical concept of algorithms, where a strict set of operations is applied on some input data to solve a problem, methods from pattern recognition aim to find patterns in (or "learn" from) a specific set of input data to make predictions and classifications on a new dataset.

Pattern recognition plays a central role in the development of fully automatic classification routines and is applicable to various problems and type of data, including images, numerical and categorical data. In this work, pattern recognition methods are used on categrical and numerical data, and it is possible to add further information from other datasources in the future for gaining more accurate classification results.

The following pages give a brief introduction to aspects of pattern recognition which are relevant for this work, a detailed explanation on how these methods were used is given in chapter 4.

2.3.1 Learning strategies

Two main types of learning strategies can be distinguished, namely unsupervised and supervised methods, but also mixtures of both approaches exist.

The difference between the two strategies is the structure of the input data. Unsupervised learning routines take a set of unlabeled data samples as input and aim to find a certain structure in the data. Typical goals of unsupervised learning are clustering of similar data samples without prior knowledge of the classes, or discovering an underlying probability distribution between the samples.

The input sets for supervised learning consist of already labeled samples, where the label usually denotes the class a sample belongs to. Those labeled input sets are also referred to as “training sets”. The goal of the learning phase is to extract probabilities from the labeled samples in order to “learn” which information indicates the membership to a certain class. After the learning phase, the classifier can be used to classify a set of unlabeled samples. [Bis01], [DHS00]

2.3.2 Classification algorithms

For both supervised and unsupervised learning, different machine learning approaches exist. The choice for a certain classifier normally depends on the kind of data, but it is also common to test different classifiers in order to find the method suited best for a certain problem.

Because OSM provides us with a set of labeled data (which are those ways which already have a surface tag), the choice for a supervised method is straightforward. From the available supervised algorithms, the decision tree was chosen, as it is a “soft” classification method which does not assign a single label to a sample, but returns probabilities for each class. Furthermore, for humans it is rather easy to understand how the classification process is working compared to other methods, and it is also able to handle missing values.

Decision trees

Decision Tree Classifiers are one type of classification algorithms which follow a recursive divide-and-conquer approach to classify data samples. They have been used successfully in the field of land-cover and land-use classification (but they are of course not limited to this type of problem.)

The major advantage of this kind of classifier is that it can break down a complex classification process to a number of simpler decisions. Contrary to methods like neural networks, a decision tree is a white-box system, where every step of the decision process is transparent, thus yielding a solution which is easier to understand and interpret for humans. [SL91]

A decision tree consists of internal nodes, leaf nodes and edges which connect those nodes. When classifying a sample, the tree is traversed starting at its root, which is the topmost node. At each node, a decision has to be made based on a feature in the sample. Depending on the outcome of the decision, one of the outgoing edges of the current node has to be followed in order to reach the next node. This process is repeated until reaching a leaf node,which holds information about the samples (probable) class label. Leaf nodes have no outgoing edges, so reaching a leaf node terminates the classification process. For better understanding, figure 2.11 gives an example of a decision tree for deciding weather to play football or not. [Qui86]

Building a decision tree

Building a decision tree can be done automatically given a set of training data, where each sample in the data consists of several features and a class label. This building process is led by the aim to minimize entropy, which is done by subsequently separating the input samples into homogenous groups.

In the beginning of the building phase, the root node contains the set of all samples and all features. From all available features, the one which discriminates the samples most is chosen to divide them in homogenous subsets – this means that the resulting groups should consist of samples which are likely to have the same class. For each subgroup, a new child node containing those samples is appended to the root node. The chosen feature is used as decision rule for the root node. If a subgroup contains samples of only one class, it is turned into a leaf node which stores the probability for this class. For all the other nodes, the same procedure is applied recursively until either only leaf nodes remain or a stopping condition is met (e.g, a maximum tree depth is reached). A leaf node does not necessary have to store only one class label, it is also possible to store probabilities for each class. This makes it possible to perform "soft" classification with a DT.

The example in figure 4.4 is rather simplistic, and the features chosen to test on were taken more or less randomly until homogenous groups were gained. Any test could lead to a tree where on the final level, only homogenous groups exist, even if those groups consist of only one sample. In practice, the aim is to chose features which represent the structure of sample domain to create a tree that makes "meaningful" predictions. This means that the most discriminating feature for such a split has to be found in each step.

Outlook	Temp ($^\circ F$)	Humidity (%)	Windy?	Class
sunny	75	70	true	Play
sunny	80	90	true	Don't Play
sunny	85	85	false	Don't Play
sunny	72	95	false	Don't Play
sunny	69	70	false	Play
overcast	72	90	true	Play
overcast	83	78	false	Play
overcast	64	65	true	Play
overcast	81	75	false	Play
rain	71	80	true	Don't Play
rain	65	70	true	Don't Play
rain	75	80	false	Play
rain	68	80	false	Play
rain	70	96	false	Play

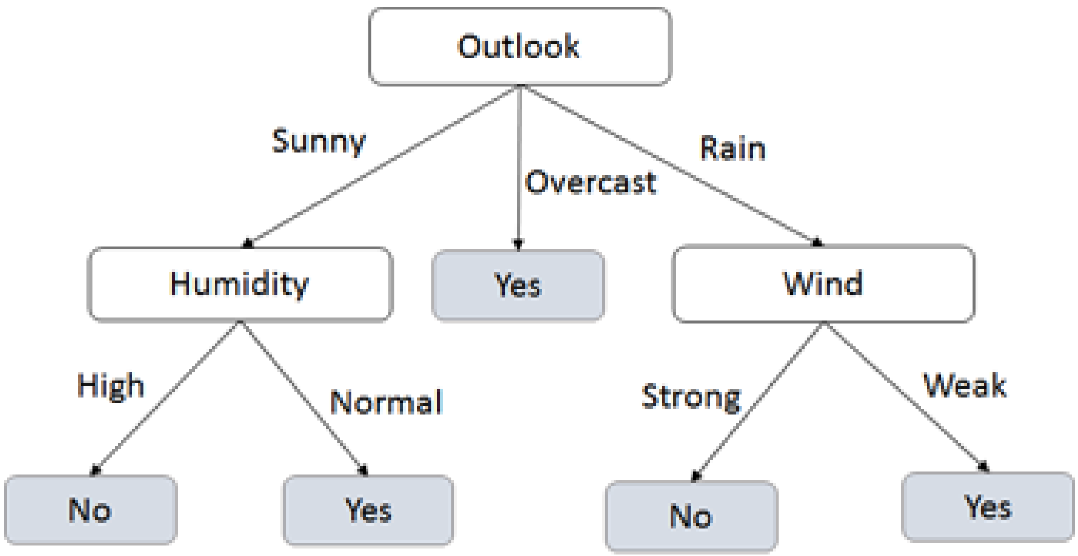


Figure 2.11: An example for a simple decision tree to decide weather to play football or not. [Qui86] On top, the sample data is depicted, which represents previous decisions based on weather conditions. There are two classes ("Play" and "Don't play"), and four features ("Outlook", "Temperature", "Humidity" and if it was raining or not). When choosing Outlook as the feature to test on, a group of samples belonging to the same class is gained – namely all the samples which have outlook = overcast. The other samples can be further divided when testing against humidity and windy, resulting in the decision tree depicted on the bottom.

There are different measures to determine the best suited feature for such a split, information gain or the Gini index are among the ones used in popular implementations. Because the former one is used in the implementation it will be explained it in more detail.

Information gain

Information gain is closely related to the concept of entropy used in information theory, which is used to measure the information in a system.

The mathematical definition of entropy H for a random variable X is described in equation 2.1, with px_i being the probability that a random variable X takes the value i. In terms of a decision tree classifier, the random variable is one of the available features, and px_i represents the probability that it corresponds to one of the i classes present in the current node.

$$H(X) = -\sum_{i=1}^{N} p(x_i) log_2 p(x_i) \tag{2.1}$$

Intuitively, the more values a variable can take, the more uncertain it is to make predictions about the outcome. If a variable can take one out of 8 possible classes with the same chance, the probability to make a right guess about the outcome is 1/8, whereas if a variable has the chance to take one out of 64 values, the probability to make a right guess about the outcome shrinks down to 1/64.

This is exactly what entropy, as defined in equation 2.1, measures. A high value for H(X) means high uncertainty, whereas a small value indicates that there are less possible values a random variable can take, and more certainty when making predictions about the outcome. Thus, reducing entropy is desirable for correct classification.

Based on this definition, information gain is defined as in equation 2.2, with T being the entropy of the parent node and a being the weighted sum (or conditional) entropy of the child nodes in case of decision trees. [WF05], [KMJ10] Information gain thus depends on the change of entropy between a parent node and its child nodes. A simple example of a parent and a child node is sufficient to illustrate this. If both parent and child have the same entropy, information gain will be 0. This is intuitive: no useful information is gained because the entropy stays the same, no matter if the path to the child node is followed or not. A high entropy on the parent node and a low entropy at the child node however, results in high information gain, which is desirable – thus, intuitively, information gain has to be maximized to build a good decision tree.

$$IG(T, a) = H(T) - H(T|a) \tag{2.2}$$

Missing values

When applying machine learning to real-life problems, missing values in the samples are a common problem. Values can be missing for several reasons, like for example a defect of the measuring device or sensor. OSM data tends to be incomplete not due to defects, but because users did not provide specific information (yet). This leads to a set of highly sparse data from which a decision tree should be built from.

There are different strategies to solve the problem of missing values. It is common to replace a missing value with a plausible one (for example by taking the most common value or by computing the mean of all available values).

Decision tree algorithms like ID3, C4.5 and CN2 can handle missing values for both training and testing sets. ID3 simply treats the missing value by creating an edge for the “unknown” in the tree. C4.5 is an extension which assigns weights to the edges. If an attribute is known, the probability to take that branch is 1, otherwise the probability to take the “unknown” edge is calculated based on the frequency of all the possible attributes which could occur. CN2 takes a simple approach which replaces missing values with the most frequent ones. [GLSGFV10]

Implementations

Different implementations of DT exist up to date, with ID3 and its successors C4.5 and C5.0 being the most popular ones. [Qui14] All of the three algorithms use entropy as splitting metric and aim to build rather simple decision trees, which are not necessarily optimal for classification. Contrary to ID3, C4.5 is able to handle both numerical and categorical data, and also to handle missing values discussed in the next section. C5.0 is a further development in regard of computational speed and memory usage, and available under a commercial license. [C50]

2.3.3 Image classification in remote sensing

The increasing availability of high-resolution digital imagery of the earth’s surface requires a practical application of the methods known from pattern recognition and image processing for creating efficient and automated classification techniques.

Up-to-date information about geographical features and land coverage in a country are required for practical applications such as change analysis, urban planning or monitoring the condition of road networks, woods and agricultural areas. [DLR03]

While related work and prior research in this field will be presented in chapter 3, the following section will give a brief introduction on aspects which are particularly relevant for classification of aerial and satellite imagery.

Pixel-based versus object-based approaches

In addition to the differentiation between supervised and non-supervised learning methods, methods for image-analysis can be further categorized into pixel-based and object-based approaches.

Pixel-based approaches are the older method and were used successfully with images of rather coarse resolution, in which objects like trees or cars covered less then a pixel. As the name suggests, those methods analyze the spectral properties of every image pixel individually, and also classification takes place on a per-pixel level. This means that spatial or contextual information of surrounding pixels is not considered when assigning a class to a single pixel.

When images of higher resolution became available, the counter intuitive observation was made that applying pixel-based approaches on those high-resolution images did not improve the classification results. Because an object in an high-resolution image consists of many pixels with big spectral variety, inspecting and classifying each pixel separately will lead to bad classification results. [WJRJ10] Misclassified isolated or groups of pixels (especially at object borders) are usually present in the final classification result, this phenomenon is also referred to as "salt-and-pepper-effect" in literature. [Bla10]

Therefore, high resolution imagery was used to be classified manually, which is a tedious and error prone process. In order to make automated classification of high resolution imagery possible, object-based approaches were developed.

They overcome the problems of traditional pixel based methods by first clustering the image into semantic "objects" (for example groups of homogenous pixels) before performing further classification on them. This has the advantage that in addition to spectral information, also spatial information such as texture and average brightness of a region can be taken into account for classification. Object-based methods add more complexity to the classification task and usually sophisticated image processing methods are required to cluster pixels and extract the objects (e.g. buildings, rivers, fields) from the given image data.

Even though pixel based approaches were the dominating method for years, object-based approaches recently grew popular the scientific community because they tend to outperform pixel-based approaches especially when used with high resolution images. [KSF05]

An object-based approach was also the choice for the practical part of this work. Thanks to the available OSM data, it is possible to extract streets without applying a preprocessing step of image clustering, which is usually necessary when photos are the only source for extracting features. For the analysis of an individual street, pixel-wise information is put into context in order to extract spectral profiles and textural information. For example, average profiles of RGB histograms are used to determine the spectral signature of asphalted roads. A detailed explanation of the process is given in chapter 4.

CHAPTER 3

Related Work

In this chapter, papers related to this thesis will be presented. Because the previous chapter already covered studies and scientific papers regarding OSM, this chapter will focus primarily on methods for extraction and classification of GIS features from imagery.

The final section will give an overview on current studies on OSM data with regard to land-cover and land-use classification, as this aspect was not covered in the previous chapters.

3.1 Automatic updating of GIS databases based on imagery

Keeping Geographic Information Systems (GIS) up to date manually is a cumbersome and error-prone process even for expert users, therefore different attempts were made to automate this process and make it more time- and cost efficient. With the availability of high-resolution, up-to-date satellite and air-borne imagery came the idea to automatically extract relevant information from images based on methods known from computer vision and machine learning.

Walter [WF00] presents a fully automatic method to detect changes in an already existing GIS database, ATKIS. The method combines CIR imagery from remote sensing with a geometrical resolution of 2 meters per pixel and laser scan data for change detection. Utilizing data with complementary characteristics increases the classification accuracy. For example, land-cover classes like forest and grass areas may have a similar spectral profile, but with the height information from laser scanned data, they can easily be distinguished. For that reason, the aim was to incorporate additional data from different sources to the developed method, which will be described in the next chapter.

The fully automatic approach consists of two steps. For the first step, remote sensing data is classified pixel-wise into different land-use classes by supervised maximum-likelihood classification. The main challenge when applying supervised classification is to automatically derive the necessary training areas for the learning phase. Normally, training areas are created by human expert users, but this is inapplicable when aiming to derive a fully automatic approach.

Based on the assumption that the majority of objects in the database is classified correctly and that all possible land-use classes are present in an image, training areas are derived by using already collected GIS objects in the ATKIS database. If too many objects in the database are erroneous or too few pixels of a certain class are present in an image, this approach fails.

For the classification, a maximum likelihood method was chosen which distinguishes between different land-use classes and streets. Only the location of streets is of interest, and not their surface coverage. Therefore, it is sufficient to extract only pixels which are located exactly at the streets centerline for the training areas. In cases where the the centerline is not aligned perfectly to the middle of a road, pixels from surrounding geometry are falsely selected and can bias the quality of the classification. Streets which are occluded by trees are excluded from the training areas for streets by using a vegetation index.

Finally, the classification results have to be matched and compared with existing GIS data to find inconsistencies and changes in the data. For this, different measurements (namely percentage, homogeneity and form) are introduced.

Walter's method updates an already existing, mostly complete database, from which training areas can be automatically derived. For this thesis, no such database is available. The existing OSM data is sparse and many streets have no surface information at all, therefore it is not possible to automatically derive training areas from surface tags only. Furthermore, Walter's method aims to extract streets from images, but not to classify their surface type.

3.2 Remote sensing of road infrastructures

When classifying surfaces for GIS, a distinction between techniques for area and line features has to be made, because they require different strategies due to their different geometric characteristics.

Streets are usually longer and thinner compared to area features, and thus the extraction of the pixels which represent a street in an image can be challenging. Roads form the "bottom layer" of an image, which means they can be occluded or shadowed by trees, buildings, cars and other objects. When naively extracting the road's pixels from an image, they are likely to be mixed up with pixels of those occluding or neighboring geometries, so that the spectral characteristics of the road-surface are biased and have an influence on the classification result. This is particularly a problem for road classification in urban areas, which are especially challenging due to the high occurrence of human made materials with similar spectral profiles and the density of street networks and buildings [HGNR03].

Naturally, when one aims to classify the surface properties of streets from images only, a high geometric resolution is needed compared to tasks where a mere detection of streets is the goal.

Herold et al. [HRGD04] created a spectral library for urban materials, which should help to determine the underground type of roads in urban areas. While surfaces like asphalt, gravel and concrete were classified successfully with a maximum likelihood classifier, problems occurred when distinguishing those roads from roofs in city areas. This was due to the spectral similarity of the materials which are typically used for roofs and streets.

Mohammadi [Moh12] introduced a method to classify road-surface materials and determine the pavement condition of asphalted roads in the city of Ludwigsberg by utilizing hyperspectral imagery with 4m resolution and 125 bands.

The three most common classes of pavement materials in the city were distinguished, namely asphalt, concrete and gravel. Different classification techniques were used, among which the application of mean and standard deviations on the brightness spectral feature over certain wavelengths proved to be most effective.

While asphalt has the lowest mean in a range between 445nm - 2448nm and is thus easily identified, gravel and concrete are distinguishable when examining the standard deviation in a range between 619.9nm-1323.7nm. Even though the study underlines the high accuracy of the method, one should mention that a big portion of the road network (52%) remains unclassified, which could be tackled by using imagery in higher spatial resolution.

The image material available for this thesis has different properties than the images used in the work described above. While the spatial resolution is significantly higher, the spectral resolution is limited to three bands, namely the RGB channels. Thus, hyper-spectral analysis like described in the methods above cannot be performed.

3.3 Remote sensing for land-use and land-cover classification

The papers presented so far use satellite imagery with rather coarse spatial resolution for classification, but compensate this by using a broader range of spectral bands or by adding data from other sources. The downside of those approaches is that they are relatively cost intensive and complex [NHRG02].

Methods based on high-resolution aerial imagery also exist, and are primarily focusing on detection and classification of urban land-cover. Land-cover mapping is an active research field due to the need for accurate Land-Use and Land-Cover (LULC) maps in land change science and landscape planning.

In order to create accurate maps, not only high-resolution (HR) imagery is necessary, but also sophisticated image processing techniques are required. Pixel-based classification approaches which are used with lower resolution imagery cannot be applied anymore, because of the high spectral ambiguity within objects in HR imagery. Instead, an object based approach has to be made to incorporate textural and spatial properties of objects.

Li and Shao [LS14] aimed to create highly accurate LULC maps with Object-Based Imagery Analysis (OBIA) by utilizing an 8-bit depth, 1m geometric resolution aerial orthophotography-mosaic of Tippecanoe County, Indiana. The images consist of four bands, including RGBB and NIR channels.

Using the spectral information directly did not prove to be practical for distinguishing land-cover types. Apart from “intricate” spectral characteristics within and among land-cover types, efficient processing of the high-resolution image material was a major concern.

In order to enhance useful information and reduce the spectral dimension of the images at the same time, three new bands were derived by performing principal component analysis (PCA) on the original images in a preprocessing step. Additionally, a binary Vegetation Index was used on a pixel-based level to split the images in vegetation and non-vegetation parts, which are processed independently in order to save computation time and gain more accurate classification results.

Objects are extracted from the images by performing image segmentation which divides each image into patches of homogenous regions. The classification process is modeled with a decision tree. First, vegetation- and non-vegetation objects are discriminated. Further distinction of the vegetation objects into the classes tree, grass and crop is made by analyzing the second and third bands gained from PCA. Non-vegetation objects can be either of the class building, road, water and open land/soil. In addition to brightness values in the first PCA band, characteristics such as shape, area and height gained from a Digital Elevation Model (DEM) are used to distinguish objects of those classes.

Zhang et al. [ZQH+15] took a similar approach to analyze orthophotos of even higher resolution (9cm per pixel) in combination with a DEM. While such ultra-high-resolution (UHR) imagery is useful to analyze objects which are smaller than buildings, the high spectral ambiguity makes it necessary to incorporate spatial features in order to separate and classify objects in urban areas. Spatial information is incorporated by a so-called “dual morphological top-hat profile” (DMTHP), which is extracted both from the orthophoto and the DEM by utilizing techniques from mathematical morphology. Those DMTHP features have proven to be helpful for detecting urban features, especially buildings. Objects are extracted by height-assisted image segmentation with the DEM, and for the classification phase a random forest classifier is used on a combination of the spectral information and the DMTHP features. The combination of spectral and spatial information leads to improved classification results on UHR data sets compared to former methods.

Many methods developed for LULC-classification combine different data sources to increase classification accuracy. One method which uses 1m resolution, four-band imagery as only data source was developed by Li and Shao [LS13]. The aim was to identify different types of urban vegetation with an object-based approach by making use of both spectral and spatial features and by incorporating the hierarchical relationship between objects. The implementation is based on the software eCognition [eCo].

In a preprocessing step, Normalized Differenced Vegetation Index (NDVI) and PCA were applied on the original image to gain two more bands. For the classification, eCognition is used to build a hierarchical tree with four levels. First,vegetation and non-vegetation pixels are distinguished. On the second level, the vegetation is further discriminated into woody plants, grass and crops. On level three, woody plants are further categorized into trees and shrubs, whereas grass is divided into different lawn sizes. Level four distinguishes between single trees and forest areas.

To divide the image into meaningful objects and create hierarchical relationships between them, four different segmentation techniques were combined. Each technique contributes a distinctive kind of information to the objects, which is used to build the hierarchical classification tree in the next step. In this phase, each image object gains context information about its neighbors and its eventual super-and sub-level objects.

For the actual classification process, various spectral and spatial features of the objects are considered to determine their class. The features which are used in the decision tree are for example an objects area, its mean values in the red and green channels or the border to its neighbors or shadows.

The papers presented above aim to perform fine-grained classification by distinguishing on an object-based level, but no differentiation between the materials of these objects is made. Still, concepts and general ideas like combining different data sources and performing classification on different granularity levels can be adapted to be used for this thesis.

3.4 OpenStreetMap for surface classification

Chapter 2.1 already discussed the aspect of data quality in OSM and referenced various papers which focused on assessing OSM's data quality, including the completeness of street networks. Those studies, however, did not focus on the quality and quantity of OSM tags which annotate surface properties of streets.

In recent years, an interest in OSM data for LULC validation and classification has developed, as it could be an alternative to the cost-and time-intensive analysis of image data. Research in this field is quite recent and primarily focuses on evaluating the LU data quality of OSM for selected regions by comparing it to existing LU databases.

A study conducted for the region of continental Portugal evaluated the usefulness of OSM data for producing LULC databases by comparing available OSM data with CORINE level 1 land-cover data. (See 4.2.3 for more details on CORINE data.) With an accuracy of 76.7%, the authors conclude that OSM can be used as source for validating land-cover data [EP13].

Another study focused on comparing LU data in OSM with the GMESUA dataset for four metropolitan areas in Germany, in order to evaluate the completeness and reliability of OSM data. This study only considers the OSM tags “landuse” and “natural”, which are mapped to the corresponding LU classes of GMESUA for comparability. With a coverage between 40 and 60% in the study areas, and an accuracy between 63 and 77% (compared to “barely” 90% for GMESUA data), the authors conclude that OSM data holds a high potential and could be a promising input for updating and evaluating existing LU databases [AMZS15]. A similar study for the city of Vienna was also conducted, coming to the conclusion that LU classification based on OSM has several advantages, namely cost-efficiency and flexibility because it is independent of remotely sensed data. Missing data, inhomogeneity and inaccuracy of OSM data are listed as downsides, but the authors are positive that OSM and VGI is a growing field which will still develop [JAHB+13].

While the above studies all focus on LU classification, no studies for determining the surfaces of streets have yet been published. As the data quality for LULC classification seems promising, the expectation is that street surface classification based on OSM achieves similar results.

CHAPTER 4

Classification of road-surface types

The following chapter describes the overall strategy to implement the system and provides details of the implementation where necessary.

4.1 Algorithm outline

The algorithm consists of two classification phases, where the first focuses on the OSM data and the second on image-analysis of orthofotos. The order of the phases can be reversed, but the choice was made to perform the OSM phase first because it produces more reliable classification results and processes faster. In cases where the OSM classification cannot be applied (for example, because no tags are present) or where the classification results are too uncertain, the second phase of image-analysis can be applied directly.

Figure 4.1 illustrates the sequential steps of the algorithm: Given an OSM XML file as input, first the OSM analysis phase takes place. The data is preprocessed, which means that streets with surface tags are filtered out to build a decision tree for classifying streets which have no surface information. Quality analysis follows to determine if the results for the coarse and fine-grained classification are plausible. For ways which could not be classified by this procedure, image-analysis is performed. After extracting the ways from the orthofoto, k-means clustering and histogram analysis are performed to decide the surface type. Finally, the results are written out to a new OSM XML file. All steps of this algorithm will be presented in more detail in the following subchapters.

4.1.1 Study areas

Different areas were chosen to conduct this study on, namely the center and periphery of Vienna city, the city of Villach, Oberndorf, Spittal an der Drau and the state of

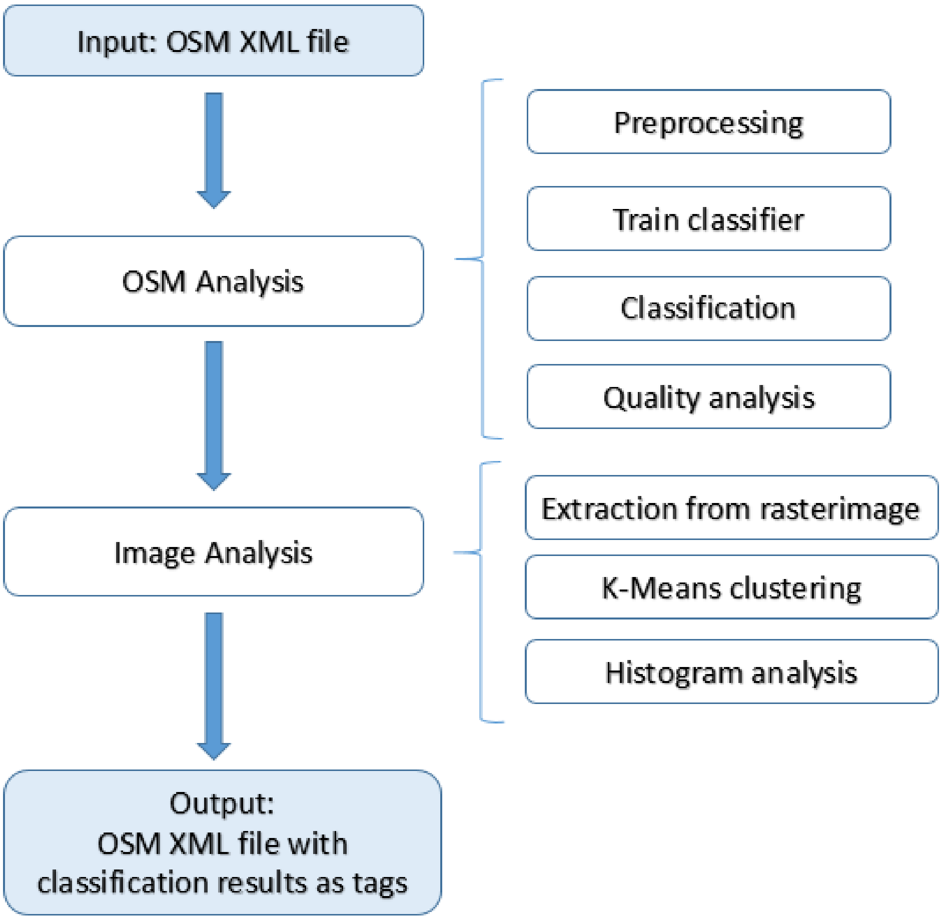


Figure 4.1: Overview of the algorithm and the sequential steps to classify ways. An OSM file with still unclassified ways is provided as input. In the first phase, classification is done based on OSM data only, with an option to add further textual data in a preprocessing step. For ways which could not be classified or which are likely to be misclassified, an image-analysis phase is applied. The final classification results are written out as additional tags to the input OSM file.

Table 4.1: Overview on the extent of the evaluation- and testing areas (Vienna's center and periphery are listed separately in this table, but are treated as one area in the testing phase). Vienna, Villach and Liechtenstein were used for the data analysis phase and also for classification, while Oberndorf and Spittal a.d. Drau were used exclusively for training an classification. This was done in order to have two testing areas which are "independent" of the data analysis phase. The coordinates describe the bottom left and top right corners of a rectangular bounding box. The data for all areas way downloaded on 25.6.2016, except Oberndorf and Liechtenstein, which were downloaded on 1.2.2017.

Testing Area	Latitude min	Longitude min	Latitude max	Longitude max
Vienna, city center	48.1996	16.3495	48.217	16.3897
Vienna, periphery	48.2115	16.2036	48.2624	16.3365
Villach	46.5564	13.7454	46.6444	13.9835
Liechtenstein	47.0311	9.4534	47.2847	9.6139
Oberndorf	47.5984	12.8425	48.2268	13.7334
Spittal an der Drau	46.5441	12.8063	47.0649	13.8951

Liechtenstein. The areas were used both for the evaluation of tags (e.g., for finding relevant tags and for generating tag-statistics, which is described in the following chapter in more detail) and for testing the results of our method.

The areas were selected with respect to their topology. Because OSM data differs both quality- and quantity-wise depending on the region, one aim was to test the method in regions with different geographical properties. The aim was to have a mix between rural and city areas, where also mountains and forest areas are present.

The extent of the study areas denoted in geographical coordinates can be seen in table 4.1

OSM Data for study areas

The OSM data for the study areas was downloaded online and consists of one .osm file for each area. A detailed description of the .osm fileformat was already given in chapter 2.

OSM data can be downloaded from the internet in various ways (an overview of sources can be found in the OSM Wiki [OSMc]), but the chosen method has no influence on the data itself. (Which means that each method will result in an xml file with nodes, ways and tags.) The download date on the other hand plays an important role because OSM data tends to change rapidly over time. Data can be added, removed or changed in short intervals, yielding a different input data set for one and the same region depending on the date of download.

4.2 OSM Data analysis

Because of the extensive nature of OSM's tagging system (as already described in chapter 2.1.6), it was necessary to perform thorough data analysis to define a subset of

Table 4.2: Potential tags by keys and how often they are used on the way-element worldwide.

Key	# Ways worldwide
highway	102 705 477 (26,06%)
surface	15508468 (3,94%)
tracktype	488094 (1,24%)

tags which are relevant for the classification of roads. This section will present how the analysis was performed and the resulting conclusions.

The selection process was driven by several criteria. First, all tags from which the surface properties of roads could be deduced were potentially considered. The most helpful tags are those which describe the surface directly, without the need to make any further assumptions.

Tags which describe how roads are used within a national road network were also of interest, because the assumption was made that the surface covering correlates with the road type. For example, a highway is likely to be paved with asphalt, whereas a trail in the mountains is more likely to consist of a muddy or earthy surface. Thus, from knowing the usage of the road, the surface properties can be deduced.

Furthermore, quantity was a factor which had to be considered. Because OSM's tagging system consists of a vast amount of possible keys and values, many tags exist which are used only for a very tiny fraction of ways (like 0.01% of all ways worldwide or less). Considering a tag which is only used on very few ways is eventually a waste of memory and processing power, and does not help to classify a bigger portion of ways than without using this tag.

Finally, because the aim is to classify the surfaces of roads, streets and paths, only tags which are used for ways are of interest, while tags which are used exclusively for buildings, areas and nodes were not considered.

4.2.1 Taginfo and OSM Wiki

To decide which tags are relevant for this thesis, the website "Taginfo" [Tag] was the first source of information. This website is maintained by the OSM community and provides detailed listings and statistics of all tags which can be used to describe nodes, ways and relations in more detail. Taginfo lists all existing keys for ways worldwide, and can display them ordered by their occurrence. From this sorted list, the keys "surface", "highway" and "tracktype" stood out as the most relevant ones, because their descriptions in the OSM wiki match all of the selection criteria described above. Statistics on how often those tags are used worldwide can be seen in table 4.2.

A short description for each key and the values considered relevant for this thesis follows:

Surface This key is used to describe the surface of roads and footpaths directly, thus it is the most relevant tag available for classification. The most basic distinction that can be made is between paved and unpaved surfaces, but the OSM Wiki encourages users to use more precise descriptions. [OSMj]

Highway This is the main key for defining the usage or type of a road, street or path. The values help to determine the importance of the road in a country's national road system. Roads which are primarily designed for motorized vehicles are assumed to be paved with asphalt- or concrete-like materials [OSMr].

Tracktype The key "tracktype" is used to describe the surface firmness of a minor road or track. It is intended to be used additionally on ways which are tagged as "highway: track", but can also be used for non-tracks, especially in places where many main roads are unpaved. [OSMl]

4.2.2 Analysis for the testing areas

Even though Taginfo gives a good first insight on OSM data and its distribution, it provides this information only on a world-wide level. To check if information provided by Taginfo also applies to the testing areas, it was necessary to create statistics specifically for those regions. A statistic of the 40 most common tags is presented in table 4.6.

In addition to verifying the information from Taginfo, the purpose of this statistic was to eventually find further common keys which could give a hint on surface properties for the testing regions.

The top-40 statistic verified that "surface", "tracktype" and "highway" are among the most common tags in the testing areas. Furthermore, "maxspeed" and "mtb:scale" were identified as common and potentially useful for surface determination. While the usefulness of "maxspeed" is rather intuitive (the higher the speedlimit, the more likely that the street is paved), the one of the tag "mtb:scale" may not seem obvious at first glance. This tag describes the hardness of a route for mountain biking, and could be used to identify unpaved ways. Another common tag was "lit", which describes if a way is lit or not. Further examination of tags was made by performing cross-correlation checks, which are described in section 4.2.2.

Additionally, statistics of how many ways have no tags at all and how many are tagged with tags of interest was made, which are presented in table 4.7.

Classification taxonomy

As the table with descriptions of possible surface values shows (see table 4.3), OSM allows a very fine-grained classification of surface types. This may seem like a welcome property because theoretically, detailed classification could be performed and many different surface types distinguished. But practically it is not, because there are not enough samples to effectively train a classifier for so many different surface types.

Table 4.3: Overview of common values for the surface tag, including a short description if considered necessary by the authors. The table is based on the OSM wiki, where more detailed information can be found [OSMj].

Value	Description
Paved surfaces	
paved	Road which is sealed with a material like asphalt or paving stones.
asphalt	Asphalted road.
concrete	Cement based concrete forming a large surface.
concrete:lanes	Two parallel lines of concrete plates, laid out so vehicles can drive along them with the wheels always hitting the concrete.
concrete:plates	Pre-fabricated concrete plates tiled closely together with narrow connections between them.
paving stones	Equally sized concrete stones with a flat top, tiled closely together.
cobblestone	Coarse, unshaped cobblestone.
cobblestone:flattened	Slightly flattened cobblestones, but still rough compared to paving stones.
sett	Sett paving made from regularly arranged, flat stones.
Unpaved surfaces	
unpaved	Used on ways which are unsealed and have a loose covering like compacted stone chippings.
compacted	A mixture of gravel and sand, which is compacted and thus more stable than loose gravel.
ground	General term for a natural, unpaved ground. Similar to “earth”.
dirt	Very similar to ground.
mud	Similar to ground, but wet most of the year.
earth	Very similar to ground.
grass	Grass covered ground.
grass_paver	Permeable paving with a cell structure made of a hard material like asphalt. The gaps are filled with grass so water can seep into the ground.
gravel	Broken and coarse rocks with sharp edges.
fine_gravel	Multilayer pavement based on stone or gravel with a topmost coat of firm, granular grit, basalt or quartz. Easy to walk, jog, cycle or ride on.
pebblestone	Lose stones rounded by waves or river flow.
sand	-
woodchips	-
wood	-
metal	-

Table 4.4: Overview of possible values for the highway tag, including a short description. The table is based on the OSM Wiki, where more detailed information can be found [OSMr].

Value	Description
General roads	
Motorway	Major highway with two or more lanes.
Trunk	According to the OSM wiki, "The most important roads in a country's road system which are not a highway."
Primary	Roads linking larger towns.
Secondary	Roads which link towns.
Tertiary	Roads which link smaller towns and villages
Unclassified	Artifact of the UK road system, does not mean that the road is actually unclassified. According to the OSM wiki, those are "Minor roads of a lower classification than tertiary, but which serve a purpose other than access to properties. Often link villages and hamlets."
Residential	Roads located in residential areas, mostly designed for accessing housing.
Service	Access roads which lead to or lie within industrial estates, camping sites, etc.
Link Roads	
Motorway Link	Link road between motorway and a lower class road
Trunk Link	Link road between a trunk road and a lower class road
Primary Link	Link road between a primary road and a lower class road
Secondary Link	Link road between a secondary road and a lower class road
Tertiary Link	Link road between a tertiary road and a lower class road
Special road types	
Living street	Residential roads where pedestrians have priority over motorized vehicles and the speed limit is kept very low.
Pedestrian	Street which is exclusively for pedestrians, often found in shopping and residential areas.
Track	Roads which are used in agricultural areas or forests, often unpaved with a rough surface. If available, the tag "tracktype" provides a more detailed description.
Road	(Yet) not classified road.

Table 4.5: Overview of possible values for the tracktype tag, including a short description. The table is based on the OSM Wiki, where more detailed information can be found [OSMl].

Value	Description
grade1	A solid, either paved or heavily compacted surface.
grade2	Mostly solid, unpaved surface consisting of a mix of gravel with sand, silt and clay.
grade3	Unpaved track with a mixture of "hard and soft materials".
grade4	Soft, unpaved track consisting of soil, sand or grass, but some hard or compressed materials can be mixed in between.
grade5	Soft, unpaved track consisting of soil, sand and grass.

Furthermore, some OSM surface tags describe very similar surface types, with barely any difference to the human observer – for example, both “surface=earth” and “surface=ground” refer to an earthy, unpaved surface.

Therefore the choice was made to group OSM surface tags that describe a similar ground type and create a coarse-to-fine-grained taxonomy for adaptive classification. Based on the information found in the OSM wiki, each surface tag was assigned to a group and the resulting groups were arranged hierarchically on two levels. On the first level (L1), “paved” and “unpaved” surfaces are distinguished. On the second level (L2), five different surface classes are distinguished. Paved surfaces can be either paved with asphalt or stones, while for unpaved surfaces a further distinction between gravel, earth and other materials is made. Figure 4.2 illustrates which surface-value pair contributes to which class.

The advantage of such a two-level approach is that if the classification accuracy is low on the fine-grained level, at least a coarse distinction between paved and unpaved surfaces can be made. For example, if a way is tagged with “highway=footway” and “tracktype=grade1”, it is likely that it is paved, but deciding if it is paved with asphalt or cobblestones may not be possible from those two features alone. So instead of having a very uncertain classification result, this approach allows a coarse distinction with higher accuracy.

Cross correlation

While the “surface” tags are the most useful ones as they directly describe surface properties, they only make up roughly 4% of tags worldwide. (See table 4.2.) Despite the rather low occurrence of the surface tag worldwide, it is possible to learn from those ways which have a surface tag together with other “useful” tags, in order to classify ways which have no surface tag but at least one other useful tag.

Several assumptions were made for the correlation between tag values and surface properties, for example that primary highways are very likely to be asphalted. Those

Table 4.6: The top 40 tags for the testing areas used in the analysis phase, sorted by frequency. "Highway" and "surface" are – as expected due to the information on Taginfo – ranked high in all areas. Vienna's city center and periphery are examined separately in this phase because of their different characteristics. "Tracktype" appears in all areas except the center of Vienna, which is due to the nature of tracks which usually don't appear in city ares. In addition to verifying that those tags are among the most common ones, other tags like "maxspeed" were identified as possible candidates for useful tags.

Liechtenstein (14292 ways total)	**Vienna Periphery** (8878 ways total)	**Vienna Center** (6752 ways total)	**Villach** (9741 ways total)
highway : 12220	highway : 6125	highway : 3942	highway : 8574
name : 4586	name : 2284	name : 1777	name : 3325
surface : 2472	maxspeed : 1844	maxspeed : 1374	is_in : 2932
tracktype : 1771	barrier : 1064	oneway : 1135	service : 1587
maxspeed : 1545	source:maxspeed : 860	lit : 1011	surface : 1578
service : 1508	oneway : 705	tunnel : 814	maxspeed : 877
bicycle : 1035	surface : 596	level : 804	tracktype : 848
foot : 963	access : 558	indoor : 614	ref : 772
oneway : 876	tracktype : 480	foot : 475	oneway : 712
layer : 781	foot : 446	sidewalk : 444	railway : 571
is_in : 746	lit : 412	bicycle : 406	gauge : 539
waterway : 650	railway : 365	railway : 334	foot : 523
bridge : 609	operator : 332	source:maxspeed : 329	operator : 459
lanes : 587	electrified : 329	surface : 325	layer : 436
ref : 574	voltage : 329	lanes : 303	electrified : 433
motor_vehicle : 564	frequency : 328	voltage : 302	voltage : 428
railway : 484	gauge : 327	electrified : 301	lanes : 427
gauge : 475	bicycle : 278	gauge : 301	frequency : 407
operator : 439	mtb:scale : 263	operator : 299	bicycle : 390
tracks : 372	vehicle : 236	incline : 298	access : 379
electrified : 360	service : 234	frequency : 294	bridge : 320
access : 358	waterway : 226	cycleway : 280	footway : 315
motorcar : 347	layer : 160	layer : 246	wheelchair : 312
motorcycle : 304	bridge : 127	barrier : 216	toll : 303
cycleway : 247	cycleway : 123	ref : 183	int_ref : 220
barrier : 244	admin_level : 120	service : 145	usage : 180
tunnel : 241	boundary : 120	step_count : 127	railway:radio : 175
lit : 232	tunnel : 107	conveying : 124	railway:track_class : 175
voltage : 225	oneway:bicycle : 101	wheelchair : 121	railway:traffic_mode : 174
sac_scale : 220	ref : 92	lcn : 115	structure_gauge : 170
frequency : 216	sidewalk : 80	access : 109	passenger_lines : 157
mtb:scale : 192	mtb:scale:uphill : 77	wikipedia : 105	tunnel : 141
width : 162	trail_visibility : 75	segregated : 104	sac_scale : 139
usage : 141	is_in : 69	is_in : 81	railway:pzb : 133
boundary : 139	boat : 60	turn:lanes : 80	tracks : 132
admin_level : 137	width : 54	motor_vehicle : 78	wikipedia : 130
passenger_lines : 136	lanes : 53	oneway:bicycle : 76	barrier : 113
toll : 132	wikipedia : 50	footway : 71	railway:bidirectional : 101
horse : 124	noexit : 48	man_made : 66	waterway : 93
railway:radio : 114	footway : 48	public_transport : 66	power : 90

Table 4.7: Selected tags by keys and how often they are used with the ways in the testing areas.

Ways - Testing areas	Vienna, center	Vienna, periphery	Villach	Liechtenstein	Oberndorf	Spital a.d. Drau	All together
in total	5517	7220	8906	12867	21431	10954	66895
with highway key	3346	6098	8519	12198	20005	9696	59862
with surface key	325	596	1578	2472	8620	772	14363
with tracktype key	1	480	848	1771	3468	1995	8563
no surface key	5192	6624	7328	10395	12811	10182	52532
no tag at all	1983	932	142	296	286	903	4542

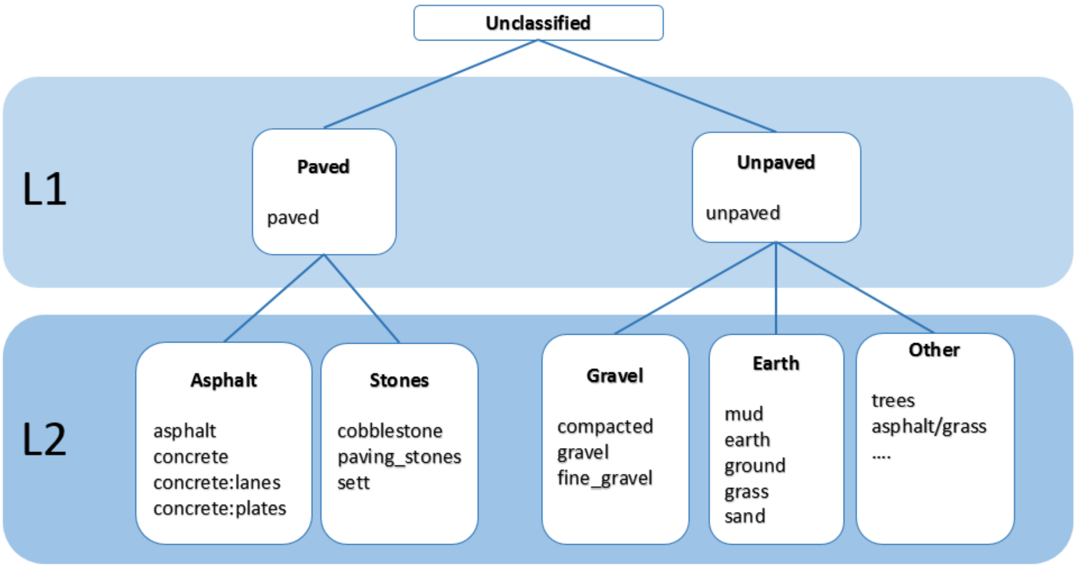


Figure 4.2: Illustration of the surface classes and which tag-values to contribute to each class. The classification is performed on two levels: Level 1 (L1) distinguishes between paved and unpaved ways, while Level 2 (L2) distinguishes between asphalt, stones, grass, gravel and remaining ("other") surfaces. Each class was created based on OSM-surface values which describe similar surfaces. For example, the OSM-surface tags "asphalt", "concrete", "concrete:lanes" and "concrete:stones" all contribute to the class asphalt for L2 classification. All those values and also the value "paved" contribute to the class "paved" for L1 classification.

assumptions are based on the tag usage description in the OSM wiki, but had to be statistically verified before being used practically.

Therefore, cross-correlation tests between the tags were made. For ways which had both a surface tag and one of the other relevant tags, the correlation between the surface tag's value and the other tag's value was evaluated by counting those combinations. Those statistics were run for each of the testing areas in separate, and for all areas together, in order to see if differences between the areas exist.

Overall, the information found on the OSM wiki proved to be correct and also applies to the testing areas, and the results will be summed up in the following paragraphs, with tables 4.8, 4.9, 4.10 and 4.11 offering more detailed insight to the results.

- **surface/highway**

  The correlation between the surface classes and the values of "highway" shows that primary, secondary and tertiary highways tend to be either asphalted or paved, while tracks and paths have gravel or earthy surfaces.

surface	construction	cycleway	footway	living_street	motorway_link	path	pedestrian	primary	primary_link
Paved		73 (19.16%)	102 (13.44%)	2 (8.00%)		16 (2.70%)	19 (18.63%)	4 (6.45%)	
Asphalt	2 (100%)	289 (75.85%)	500 (65.88%)	4 (16.00%)	5 (100%)	24 (4.05%)	8 (7.84%)	58 (93.55%)	1 (100%)
Stones		3 (0.79%)	50 (6.59%)	17 (68.00%)			75 (73.53%)		
Unpaved		2 (0.52%)	28 (3.69%)			77 (13.01%)			
Earth		1 (0.26%)	22 (2.90%)			406 (68.58%)			
Gravel		12 (3.15%)	53 (6.98%)	2 (8.00%)		52 (8.78%)			
Other		1 (0.26%)	4 (0.53%)			17 (2.87%)			

surface	residential	secondary	secondary_link	service	steps	tertiary	track	unclassified
Paved	122 (16.16%)			76 (21.59%)	2 (5.00%)	8 (10.00%)	66 (4.07%)	29 (20.86%)
Asphalt	487 (64.50%)	32 (80.00%)	1 (100%)	213 (60.51%)	29 (72.50%)	50 (62.50%)	229 (14.12%)	102 (73.38%)
Stones	114 (15.10%)	8 (20.00%)		15 (4.26%)	4 (10.00%)	20 (25.00%)	2 (0.12%)	1 (0.72%)
Unpaved	7 (0.93%)			6 (1.70%)	3 (7.50%)		67 (4.13%)	2 (1.44%)
Earth	3 (0.40%)			5 (1.42%)	1 (2.50%)		392 (24.17%)	
Gravel	21 (2.78%)			37 (10.51%)	1 (2.50%)	2 (2.50%)	863 (53.21%)	5 (3.60%)
Other	1 (0.13%)						2 (0.12%)	

Table 4.8: The correlations between the surface classes and the values of the “highway” tag, accumulated for the testing areas Vienna periphery and center, Villach and Liechtenstein. Every column sums up to 100%, so for example from all “highway=cycleway” tags, 75.85% appear in combination with the surface class Asphalt, 19.16% with the class Paved, and so on.

Table 4.9: The correlations between the surface classes and the values of the "tracktype" tag, accumulated for the testing areas Vienna city and periphery, Villach and Liechtenstein. Every column sums up to 100%, so from all "tracktype = grade1" tags which occur together with a surface tag, almost 70% belong to the class Asphalt and 21.68% belong to the class Paved. Thus, more than 90% of ways which are tagged with "tracktype=grade1" are paved and probably asphalted.

surface	grade1	grade2	grade3	grade4	grade5
Paved	75 (21.68%)	4 (0.69%)			
Asphalt	242 (69.94%)	25 (4.33%)	2 (0.50%)		
Stones	1 (0.29%)				
Unpaved	3 (0.87%)	41 (7.11%)	5 (1.26%)	4 (1.89%)	1 (1.61%)
Earth	1 (0.29%)	23 (3.99%)	156 (39.29%)	159 (75.00%)	40 (64.52%)
Gravel	24 (6.94%)	481 (83.36%)	234 (58.94%)	48 (22.64%)	21 (33.87%)
Other		3 (0.52%)			

Footways, residential and service streets also have a strong tendency to be paved with asphalt or stones.

Thus, the highway tag is a suitable feature to deduce if a way is paved or not, and in many cases also an indicator for the type of paving.

- **surface/tracktype**

  The correlation between the surface classes and the values of "tracktype" for Vienna, Liechtenstein and Villach in sum can be seen in table 4.9.

  Because certain tracktype values appear only with certain asphalt values, it is possible to deduce surface properties from them.

  "Grade1" for example appears almost always together with the surface value "asphalt" or "paved", and thus is an indicator for a paved surface. "Grade3" to "grade5", on the other hand, never appear with "paved" at all. Thus, ways which are tagged with those values are likely to be unpaved.

  "Gravel" appears mostly in combination with "grade2" and "grade3", while "grade3" also appears in combination with "earth", and "grade4" and "grade5" tend to appear in combination with "earth" and "grass" mostly.

  This confirms the information found in the OSM Wiki, which states that tracks with a lower grade tend to be paved or covered artificially with a material like gravel, and tracks with a higher grade are more likely to be unpaved and have a earthy surface.

  When the areas are viewed separately, it is notable that none of the ways in the center of Vienna are tagged with "tracktype" – this was to expect, as tracks are not likely to appear in cities.

- **surface/maxspeed** and **surface/lit**

Table 4.10: The correlations between the surface classes and the values of the "maxspeed" tag, accumulated for the testing areas Vienna city and periphery, Liechtenstein and Villach. (Values are rounded and may not sum up to exactly 100% in each column.) It becomes apparent that not enough samples are available for building a reliable decision tree, because there are barely any maxspeed tags for a speedlimit of 5 or 10, but still the class Asphalt is the biggest one among all possible speedlimits- so the speedlimit would not be a distinctive feature. Therefore the decision was made to not use the tag for the decision tree. Nonethless, it is visible that for high speed limits, much more samples exist and as expected, they indicate a paved and asphalted road. Therefore, it eventually makes sense to use maxspeed as a "fallback tag", e.g if the correctness of classification results is rather uncertain.

surface	5 to 15	20	30	40 to 50	60	70 to 80
Paved		2 (8.7%)	19 (5%)	8 (3.4%)	1 (7.1%)	1 (9%)
Asphalt	3 (50%)	16 (69.6%)	232 (61.4%)	196 (84.1%)	12 (85.7%)	10 (91%)
Stones	1 (16.6%)	4 (17.4%)	108 (28.6%)	28 (12%)		
Unpaved			6 (1.6%)			
Earth			1 (0.3%)			
Gravel	1 (16.6%)	1 (4.3%)	12 (3.2%)	1 (0.4%)	1 (7.1%)	
Other	1 (16.6%)					

Compared to the other combinations, lit and maxspeed tags don't appear often enough in combination with surface tags to improve the results of the decision tree classification.

The value of the "maxspeed" tag indicates the speedlimit for a way, and as can be seen in the correlation statistics (see table 4.10), a way with a speed limit above 50 is most likely a paved, probably asphalted surface. Because there are some hundreds of ways which appear without a surface tag, but with a speedlimit tag, it is possible to use this tag for "fallback classification", which will be described in section 4.3.1.

The tag "mtb:scale" is another candidate for such a fallback, because it does not appear very often in combination with "surface", but generally describes unpaved tracks suited for mountain biking.

The "lit" tag was considered to be used as it may indicates a paved surface when appearing in combination with the value "yes". As the correlation tests (see table 4.11) show, it is also used with a value of "no" on many asphalted ways, therefore it is not encouraged to use it.

4.2.3 DEM and land-cover classes

Because OSM data is subject to continuous changes, the aim was to design a flexible system which allows to add further features easily in the future. Therefore the idea came

Table 4.11: The correlations between the surface classes and the values of the "lit" tag, accumulated for the testing areas Vienna city and periphery, Liechtenstein and Villach. Numerically, the majority of "lit=yes" tags may indicate asphalted streets, but on the other hand, not enough data is available to rely on that assumption. After all, procentually, "lit=no" also is used mostly on asphalted streets. Therefore the choice was made to not use the lit tag at all.

surface	no	yes
Paved	2 (4.17%)	7 (2.18%)
Asphalt	30 (62.50%)	157 (48.91%)
Unpaved	3 (6.25%)	2 (0.62%)
Earth	3 (6.25%)	3 (0.93%)
Gravel	3 (6.25%)	4 (1.25%)
Rest	1 (2.08%)	1 (0.31%)
Stones	6 (12.50%)	147 (45.79%)

up to also incorporate other, non-OSM data, to the samples to improve classification results.

In addition to the OSM tags, land-cover data from CORINE [CORa], as well as height and slope data from a DEM was used. All of this information is available freely in Austria and can be mapped to the corresponding OSM ways.

The land-cover information refers to the land-use class of a certain area, for example it determines if a region is an urban or agricultural area. Possible classes are listed in figure 4.3.

The original data is available as rasterimage, so it has to be co-registered with the OSM ways by utilizing GDAL (see section 5.1). If a way lies completely inside an area of one class, this class is added to the data sample. If a way crosses several areas, the pixelcount for each class is accumulated and the class with the majority of samples is chosen as input for the classifier.

The DEM is point-based height-information which also has to be mapped to the OSM ways. Each point of the DEM is sampled by using the node coordinates of a way and compute the maximum, minimum and average height of each way to use it with the classifier. The average slope is computed as the average of slopes between a ways nodes.

4.2.4 Preprocessing

Before being able to use OSM data for classification, some preprocessing has to be performed. As already stated, the choice was made to process data with GDAL, because it is suited not only for processing the OSM XML format, but also for relating OSM data to the orthofotos for image-analysis and provides handy functions for geoprocessing.

CORINE Landcover Nomenklatur (english)
44 classes (AUT: 28 classes)

1. Artificial surfaces	1.1. Urban fabric	1.1.1. Continuous urban fabric 1.1.2. Discontinuous urban fabric
	1.2. Industrial, commercial and transport units	1.2.1. Industrial or commercial units 1.2.2. Road and rail networks and associated land 1.2.3. Port areas 1.2.4. Airports
	1.3. Mine, dump and construction sites	1.3.1. Mineral extraction sites 1.3.2. Dump sites 1.3.3. Construction sites
	1.4. Artificial non-agricultural vegetated areas	1.4.1. Green urban areas 1.4.2. Sport and leisure facilities
2. Agricultural areas	2.1. Arable land	2.1.1. Non-irrigated arable land 2.1.2. Permanently irrigated land 2.1.3. Rice fields
	2.2. Permanent crops	2.2.1. Vineyards 2.2.2. Fruit trees and berry plantations 2.2.3. Olive groves
	2.3. Pastures	2.3.1. Pastures
	2.4. Heterogeneous agricultural areas	2.4.1. Annual crops associated with permanent crops 2.4.2. Complex cultivation patterns 2.4.3. Land principally occupied by agriculture, with significant areas of natural vegetation 2.4.4. Agro-forestry areas
3. Forests and semi-natural areas	3.1. Forests	3.1.1. Broad-leaved forest 3.1.2. Coniferous forest 3.1.3. Mixed forest
	3.2. Shrub and/or herbaceous vegetation association	3.2.1. Natural grassland 3.2.2. Moors and heathland 3.2.3. Sclerophyllous vegetation 3.2.4. Transitional woodland shrub
	3.3. Open spaces with little or no vegetation	3.3.1. Beaches, dunes, and sand plains 3.3.2. Bare rock 3.3.3. Sparsely vegetated areas 3.3.4. Burnt areas 3.3.5. Glaciers and perpetual snow
4. Wetlands	4.1. Inland wetlands	4.1.1. Inland marshes 4.1.2. Peatbogs
	4.2. Coastal wetlands	4.2.1. Salt marshes 4.2.2. Salines 4.2.3. Intertidal flats
5. Water bodies	5.1. Inland waters	5.1.1. Water courses 5.1.2. Water bodies
	5.2. Marine waters	5.2.1. Coastal lagoons 5.2.2. Estuaries 5.2.3. Sea and ocean

Table 1: nomenclature (classes in grey don't appear in Austria)

Figure 4.3: The CORINE land-use classes, which are available in three resolutions. Each column in the table corresponds to a resolution. This method uses the finest level, listed in the rightmost column [CORb].

With GDAL, it is also possible to access and manipulate OSM data primitives and their attributes as if working with an XML parser.

Removal of unwanted ways

Even though it is possible to configure GDAL to process ways only and ignore closed polygons, it is still necessary to filter out unwanted way primitives (like railways) and in some cases closed ways which turn out to be buildings before performing the classification phase. This is done by checking for the tags ‘barrier’, ‘railway’,‘tracks’, ‘waterway’, ‘indoor’, ‘building:part’ and ‘building’. The list is not necessarily complete for world-wide application, but proved to be sufficient for the testing areas. It can be easily extended or adapted if needed.

Because GDAL necessarily processes all nodes before being able to process ways, further filtering can be done simultaneously, such as elimination of nodes representing traffic lights or bus stations, to reduce the file size of the outputted OSM file after classification.

Adding additional data

For incorporating additional data from other sources, a flexible design which allows to add and remove data based on the .csv format was chosen. An arbitrary number of both numerical and categorical data can be added, as long as each value corresponds to one specific OSM way id.

Grouping surface classes

The OSM classes of the training data used for building the classification tree have to be mapped to the L1 and L2 classes which were described in section 4.2.2. This is simply done by utilizing a python dictionary as lookup table, which returns the class given an OSM surface tag as input. In this way, it is easy to detect spelling errors in an OSM file and detect surface-values which are not yet considered in the taxonomy.

Neighborhood structure

While it is possible to iterate nodes and tags of a way sequentially with GDAL, it is not possible to access information about its neighboring ways when reaching one of its end nodes. In order to process ways in a topologically meaningful order, it is necessary to explicitly create a neighborhood structure. This is possible because two neighboring or crossing ways always have one node in common. According to the OSM-Wiki, it is mandatory for two intersecting ways to share a node [OSMh], thus it is sufficient to examine its endpoints in order to find its neighbors. If other ways also share a reference to those endpoints, they are its neighbors. If an endnode belongs only to one way, it is a “dead end” and has no neighbor.

The neighborhood structure was realized with python dictionaries, which use the OSM ID of an element as key to which a list of all neighboring way's OSM IDs are mapped, and thus make it easy to access all neighbors of a way by having its ID.

Accessing information of those neighboring ways can be useful for estimating the surface of ways without any tags, which will be described in the following sections.

4.3 OSM Classification phase

OSM ways which have both a surface tag and at least one of the "useful" tags found in the analysis phase are used to build a decision tree for classifying ways which have no surface tag yet. Eventually, additional information (e.g., height information and land-cover data from external sources) can be added. The classifier used was the python package "DecisionTree 3.4.3" [Dtr]. This implementation was chosen because it is thoroughly documented and can handle both numerical and categorical data.

An arbitrary OSM XML file is used as input, from which the relevant tags are filtered out. After mapping the surface tag values to the L1 and L2 classes, a decision tree can be built. Files with additional data can be used if the data can be mapped to a specific OSM way. When classifying a way with this tree, the probabilities for all surface classes are computed. For the class with the highest probability, a new surface tag is added to the original OSM file after performing the accuracy check described in section 4.3.1. Ways that cannot be classified by the decision tree (e.g., because they are not tagged at all) can be classified with the image-analysis phase instead (see section 4.4).

For determining how accurate the classification is, training and testing sets were constructed which consist of ways which already have a surface tag in combination with either one or both highway and tracktype tags. Additionally, the land-cover classes and DEM and slope data were added to each way for the testing areas in Austria.

Regarding the treatment of missing values, there are so many missing values due to the incompleteness and inhomogeneity of OSM data that replacing missing values with the average or the most common value is out of question, as this would bias the classification results strongly. Instead, the choice was made to replace missing values with the value "Unknown", which is treated by the tree like a regular feature, so it appears as a node in the tree. This also makes it easier to understand how a classification result is made when values are missing. An example for a tree for L1 classification can be seen in figure 4.4. Because the input data varies from area to area, the resulting tree is likely to look different depending on the region. Also note that not all paths in the tree have to have the same depth, and that not only a single label is returned when reaching a leaf node, but probabilities for all possible classes.

Two different decision trees were built with the training samples. The first one performs the L1 classification to evaluate if a road is paved or not. The second one performs L2 classification on the same samples for finer-grained classification.

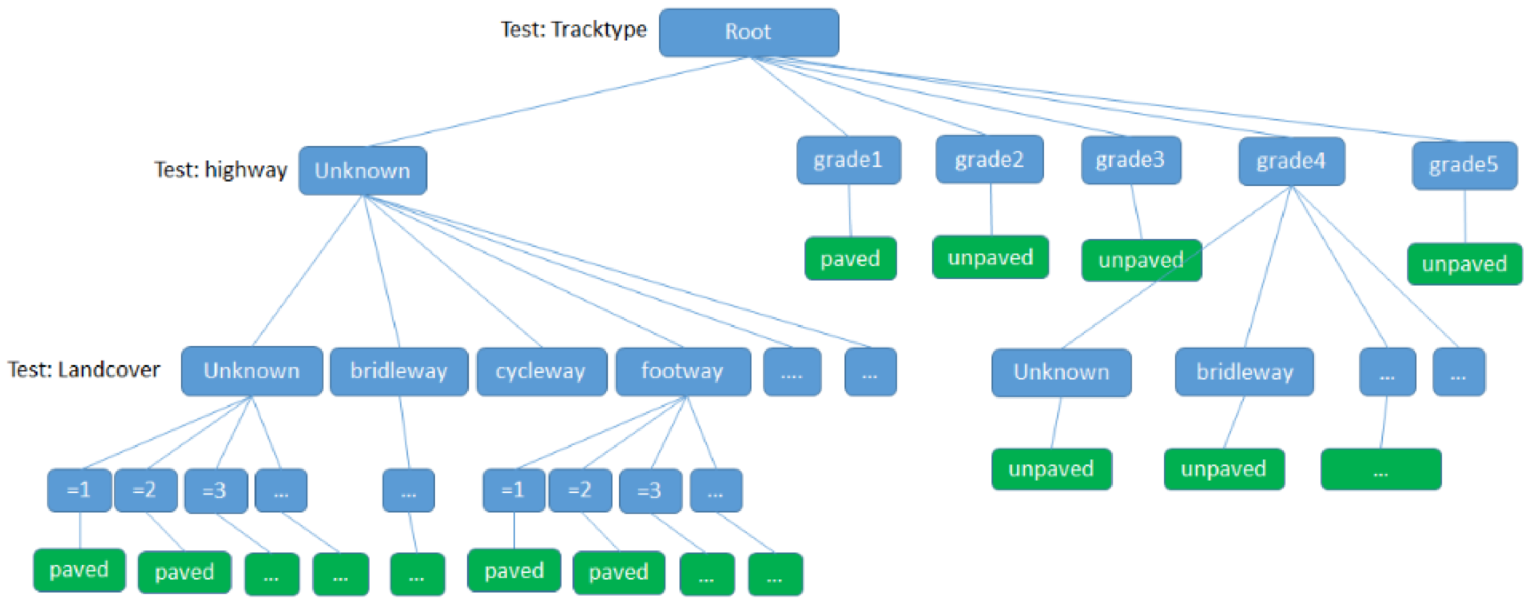


Figure 4.4: Example of one possible decision tree for L1 classification. In this example the feature "tracktype" is tested first. If its value is "Unknown", the "highway" feature is tested next. This procedure is repeated until a leaf node (green) is reached, which gives the classification label. Note that not all paths in the tree have to have the same depth: in this example, if the value of tracktype is "grade1", a leaf node would be reached immediately and the road would be classified as "paved". Not only a single label is returned when reaching a leaf node, but probabilities for all possible classes. For L1 classification, probabilities for "paved" and "unpaved" would be returned.

For testing, a sample from the testing set was classified with each tree. The correctness of the result was evaluated by comparing it with the actual surface value.

Training and testing was done both within and between the different testing areas. The features were also varied to test if land-cover and height information improve the classification result. When using different areas for training and testing, the whole area for both training and classification was used. When the same area was used for training and for testing, the data was randomly shuffled before splitting it into two sets to avoid that that patterns in the data might bias the classification results.

The results of these tests are described in more detail in section 5.2.

4.3.1 Fallback classification and accuracy check

As already described, the decision tree does not only classify a sample, but also provides a probability of how likely the classification is correct. One simple method to avoid misclassification of ways could be to discard or ignore results which are below a certain threshold, and let those ways remain unclassified. Thanks to the two-level classification approach, two simple mechanisms for testing the plausibility of the results can be proposed alternatively.

The classification taxonomy has a hierarchical structure, so if a sample is classified as “paved” in L1 classification, it should be either of the classes “asphalt” or “stone” in L2 classification – otherwise the classification would not be plausible, because the OSM surface tags contributing to the class “earth” don’t contribute to the class “paved” (compare figure 4.2).

After the classification is done, samples with high uncertainty in L2 can be tested against their L1 result. If the results contradict, the sample can either remain unclassified, or the result with the higher probability is taken, or further tags – if available – are checked:

“Maxspeed”, which is one of the more common tags in the testing areas, could be used for example to validate a surface based on the speed limit of a way. A way with a speed limit of 100km/h for example is almost certainly paved, so if this tag is available, it can be used as indicator or to determine if the classification result is likely to be correct. It would be desirable to already incorporate those tags when creating the classification tree, but unfortunately, they don’t appear often enough (yet) in combination with a surface tag to improve the resulting classification trees.

4.3.2 Neighborhood analysis

The classification only works on ways which are tagged at least with a “highway” or “tracktype” tag. Ways which have no tags at all or none of those relevant tags can thus not be classified. For those ways, a neighborhood analysis could help for determining the possible surface.

The idea is to examine the neighboring ways of an unclassified way and to deduce surface properties based on a majority vote. For this, the neighborhood structure created in the preprocessing step, described in chapter 4.2.4, has to be utilized.

The analysis is based on the assumption that the surfaces of roads appear to be steady, e.g., it is unlikely that between two asphalted ways, suddenly a way with the surface gravel appears. For each endpoint of an unclassified way, the surfaces of its neighboring ways can be checked. If the majority of the neighboring ways have the same surface class, it is possible that this way has the same surface as its neighbors.

It should noted that this method is of course rather speculative compared to actual classification, so in practical use, one has to determine if one prefers to have no surface information at all versus a suggestion. The neighborhood analysis should not be used for evaluating the classification results, but more as a fallback strategy if classification is not possible at all.

4.4 Image-analysis phase

The OSM analysis procedure described in the previous section performs well for ways which are tagged with at least one relevant tag or which have tagged neighboring ways, but unfortunately it cannot handle ways which have no relevant tags or tagged neighbors at all.

Because it is not possible to classify those ways based on OSM data, alternatively, image-analysis can be performed on orthophotos to gain information about their surface. Orthophotos for the testing area of Liechtenstein with a resolution of 10cm per pixel and RGB channels were used.

The following sections describe the individual steps of the image-analysis procedure.

4.4.1 Extraction of Ways

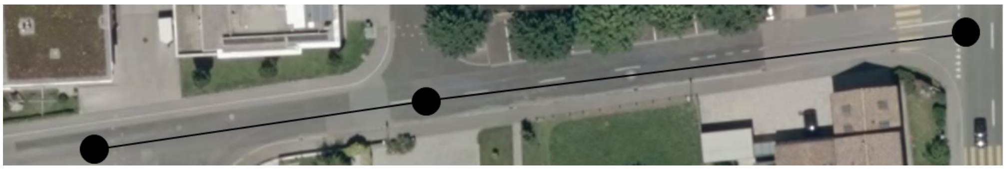
To perform image-analysis, the ways present in an OSM-File have to be extracted from the orthofotos, which can be done with GDAL.

In OSM, each way consists of several nodes, where each one provides its geographical coordinates by longitude and latitude. When connecting the successive nodes of a way linearly, they form a kind of "spine" for this way. A buffer distance can be defined to select a region perpendicular to this spine and thus extract the way. (See figure 4.5 for illustration).

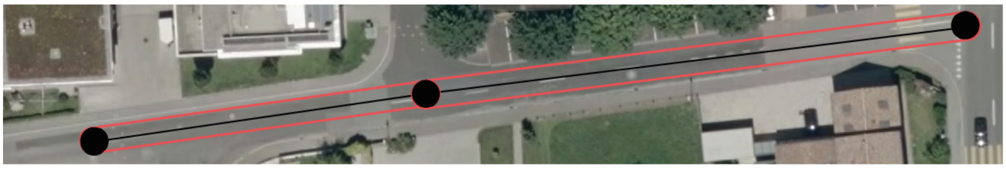
While this is a rather intuitive method, an automated exact extraction (so that only the actual way and its surface is selected) is unfortunately not possible due to several reasons:

Even though the geographical accuracy of OSM is remarkable, some nodes are "misplaced" and lie next to the actual way depicted on the ortophoto. This can be caused by a user

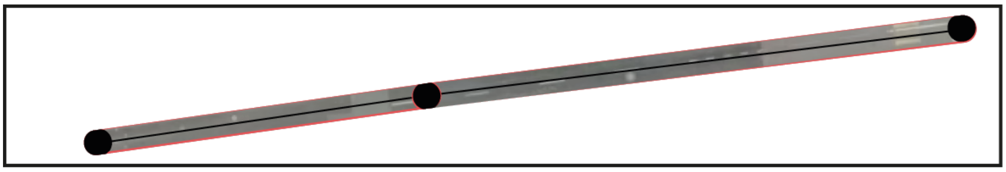
(a) An orthophoto and the corresponding OSM way, consisting of three OSM nodes (denoted with black circles)



(b) When connecting the nodes linearly, a “spine” is formed.



(c) Given a distance, a buffer can be defined by adding the distance left and right to the spine. The ends of this buffer are rounded off.



(d) In order to extract the way, the buffer is used as a mask. Only the pixels lying directly under that mask are processed further. The black rectangle indicates the extent of the original image.

Figure 4.5: Extracting an exemplary OSM way from an orthophoto.

who tracked a way with an inaccurate GPS device, or by a user who simply made a mistake when mapping the way in OSM (see figure 4.6).

Because the nodes are connected linearly to form the spine, a long, curved way which consists of rather few nodes, also tends to be cut out inaccurately.

Finally, because the extraction is done by creating a buffer in a certain distance perpendicular to the spine of a road, the most accurate buffer would be created if the nodes lie exactly on the centerline of the road and if the exact width of the road is known. But even if the mapped nodes lie on the way, it is not guaranteed that they are placed exactly on the centerline. Also, the actual road width cannot be estimated for completely untagged ways, because nothing about the road type is known to deduce width information from.

This leads to the discussion of another problem, the choice of an appropriate buffer distance. While the cases where nodes lie slightly off the centerline of a road can be treated

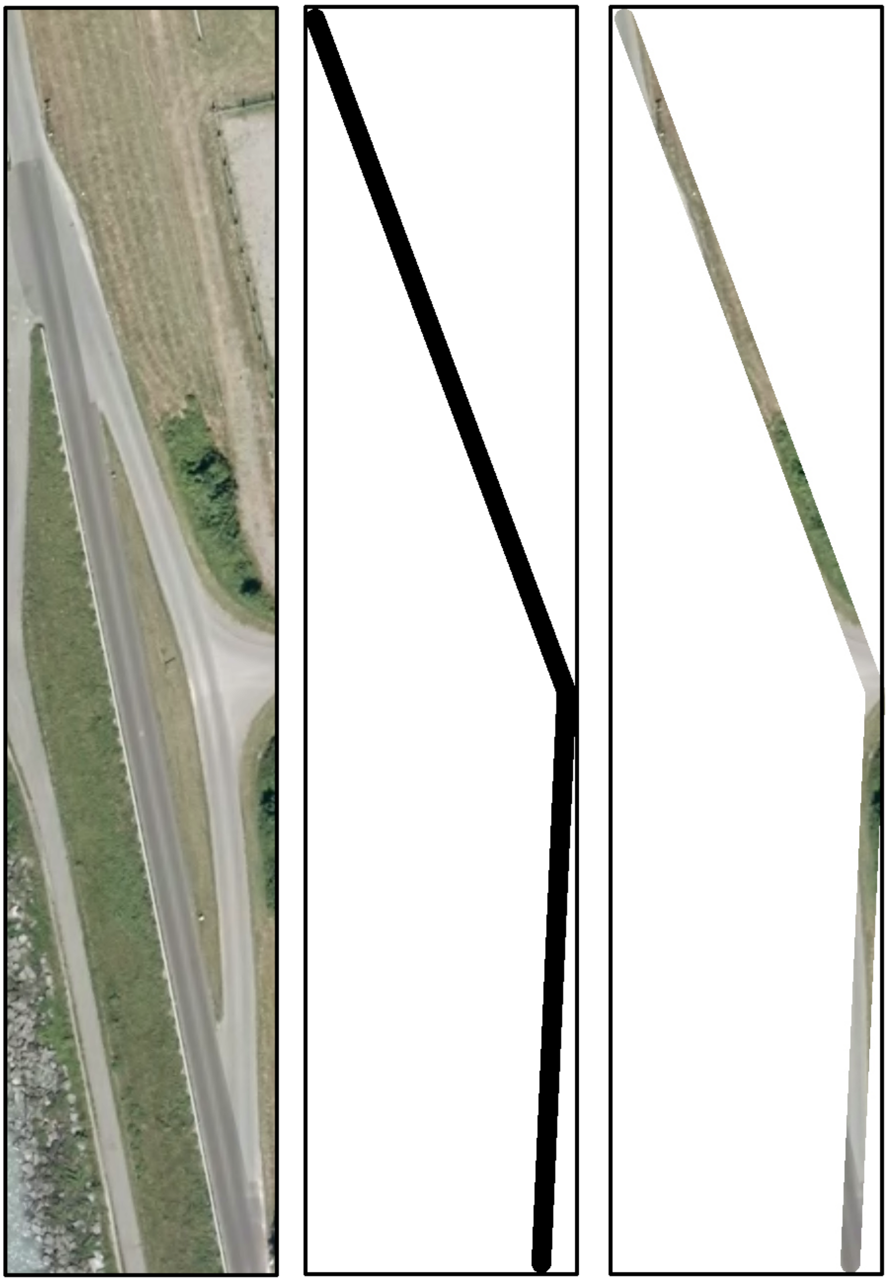


Figure 4.6: Example for extraction result of a way of which the nodes are mapped imprecisely in OSM. The original image on the orthofoto (left) is extracted based on the buffer mask formed by the nodes of the corresponding OSM way (middle). The extracted way (right) contains big portions of the surrounding.

Figure 4.7: An OSM-way occluded by tress before and after extracting it along the "spine" formed by its nodes. Before calculating the spectral signature of the way's surface, the occluding trees and the actual surface have to be separated, otherwise the result would be biased by the green pixels belonging to the trees. The asphalted areas are not connected because of the trees, which would make floodfill-like algorithms less suited than the k-means algorithm for clustering.

by defining a large enough buffer distance, this method also tends to add unnecessary portions of the surroundings to the extracted road, which can lead to faulty or wrong classification results when examining the RGB values. A too narrow buffer distance would not cut out enough pixels for a meaningful spectral analysis.

But even when being able to cut out a way perfectly, it is possible that it is occluded by cars or trees (see figure 4.7). So after extracting the way from the orthophoto with the buffer mask, it is necessary to separate the portions of the cut-out way which actually belong to the street from parts of the surrounding.

4.4.2 Clustering

When extracting a way with no tags from an image, there is absolutely no initial information about its surface, and eventually biasing pixels are present because the extraction is likely to include some pixels from the periphery. So the first task in image-analysis is to determine which pixels constitute the majority of the extracted way and then analyze them further.

This problem was tackled by a two-step clustering approach. The idea is to generate clusters of pixels with similar RGB values and examine their histograms to determine which surface classes are present. Two assumptions are made: If the way is extracted precisely and not occluded by too many objects, the cluster with the highest pixel count has to be the main surface class. If the way is paved, at least one cluster has to have high RGB values on average – this cluster represents a street with a paved surface. To find out which class is actually present, another step of clustering and analysis is performed.

If the road is not extracted precisely, image-analysis is not applicable.

For clustering, both the k-means and the flood-fill algorithm were evaluated, of which the former is suited better for this application. A naive implementation of the flood-fill algorithm proved to be impracticable because objects can occlude streets, forming unconnected surface patches as depicted in figure 4.7. Those isolated regions would end up as separate clusters when applying the flood fill algorithm, while the aim is to cluster similar pixels independent of their connectivity.

The k-means algorithm, on the other hand, can make clusters of topographically not connected areas. For the k-means algorithm to work it is necessary to have a numerical representation of the data which shall be clustered. The information available for each pixel is a triple if numeric values, which represents its RGB values. K-means starts with randomly choosing k seed points, which are assumed to be the cluster centers. Each pixel is assigned to the "closest" cluster centroid, according to a suited distance measure like for example the sum of squared errors between the RGB values. The cluster centers are recalculated after this assignment phase by taking the mean RGB values of all cluster members, and the procedure repeats until a stopping condition is reached. Hence, k-means works by minimizing the distance between the pixels and the centroids they are assigned to.

Crucial for the k-means algorithm to work as expected is the choice of a suited number of clusters ("k", hence the name), which has to be known or estimated beforehand. A too small k would assign pixels which should fall into different classes into the same class, while a too big value for k leads to a fragmentation of similar pixels [Bis01], [DHS00].

The choice for k was made after a series of empirical tests, and a number of 3 clusters for the first run proved to be optimal. Practically, when given a cut-out way, it seems plausible that around 3 classes are present: there are the pixels representing the way itself, eventually pixels which represent border areas such as grass to the left and right of the street and pixels of occluding objects and shadows cast by the surrounding. Figure 4.8 depicts the result of clustering a way with k=3. The dark red cluster represents the asphalted road-surface, the cyan and yellow clusters represent the dark green tree- and light green grass areas.

The k-means algorithm assigns a label to each pixel representing the class it belongs to. Given the labels, one has to find out which class actually represents the way's surface and to determine if the way is paved or unpaved.

For this, the histograms and statistical parameters such as the number of pixels, mean value and standard deviation are calculated for each cluster before eventually applying the second round of k-means clustering. The clusters are ranked from lowest to highest RGB mean values, with the cluster with the highest mean values being suspected as the one which represents the street. Furthermore, a threshold is computed by calculating the mean values of the RGB channels of the brightest clusters of all ways which are tagged with "asphalt", "cobblestone" or "gravel".

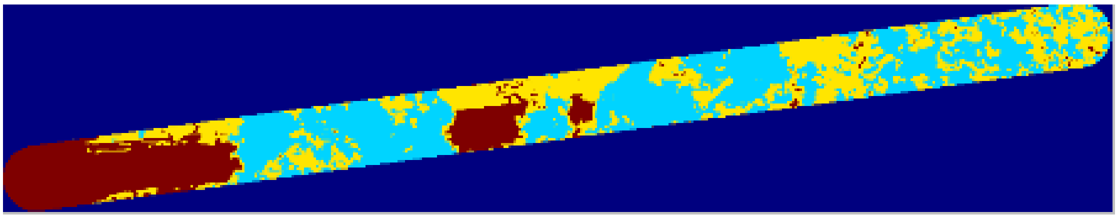


Figure 4.8: Result after clustering the way presented in figure 4.7 with a k=3. The dark blue region is the background and does not count as cluster.

If all clusters have high mean values for their RGB channels, all clusters represent a "bright" surface which is likely a paved street. No further clustering is applied in this case.

If the cluster with the highest RGB values has a mean value above a the computed threshold, this cluster is used as input for another round of k-means clustering with k=2, and the resulting clusters together with the cluster with the second highest mean-values are used for further analysis. Otherwise, the clusters with the highest and second highest mean values are used together for another round of k-means clustering with k=3. In both cases, the cluster with the lowest mean values from the first clustering is discarded.

The second round of k-means clustering aims to discard "dark" pixels from shadowing or occluding objects, and to extract a paved way as precisely as possible by assuming that a paved way consists mainly of "bright" pixels.

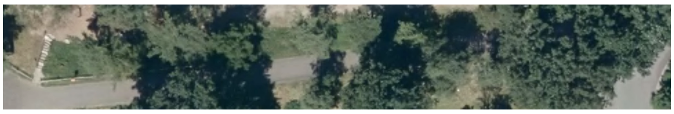
In all cases, three final clusters are obtained which can be examined with regard to the mean values of their RGB channels. If all three clusters have high mean values, it is likely that a paved street is detected. If the cluster with the highest mean values has a high amount of pixels compared to the total amount of pixels, and the other clusters have rather low mean values, it is likely that a paved way is present which is surrounded by trees and grass. An analysis of those clusters should yield information if the way's surface consists of asphalt, cobblestone or gravel.

If all clusters have rather low mean values, it is likely that the road lies in the woods or is unpaved.

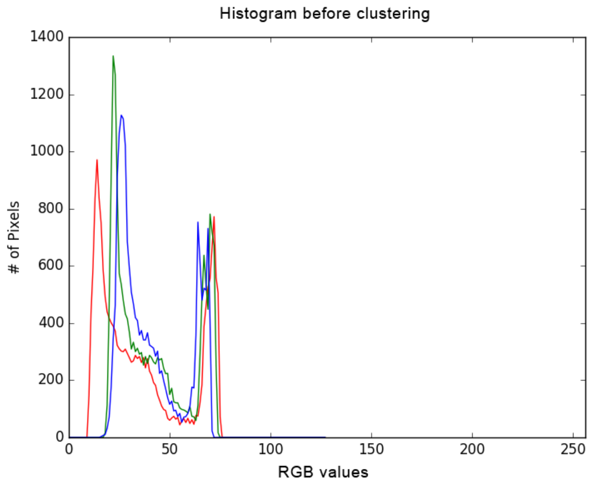
4.4.3 Analysis of centerline elements and street marks

A further idea was to extract centerline elements such as grass stripes in the middle of a road or street marks of a highway in order to deduce the surface type of a way.

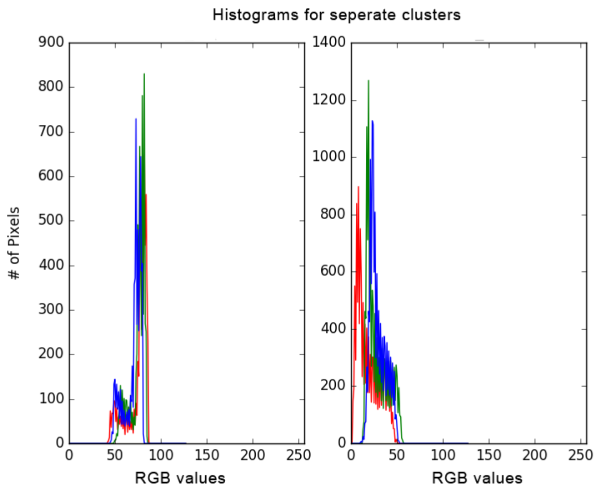
Centerline elements were extracted by utilizing the spine of a way as spatial reference. A Hough-transform was applied on those extracted regions to find linear elements like street marks or continuous stripes of grass (which appear often in the middle of a gravel way).



(a) Original way.



(b) Histogram of the complete way, without clustering.



(c) Histograms of the two biggest clusters after clustering the original way. The left histogram has much brighter and similar RGB values than the right histogram and corresponds to the asphalted parts of the way. The right histogram has darker values and a much higher pixelcount than the left histogram, which indicates that it corresponds to the trees occluding the way.

Figure 4.9: Comparison between the histogram for the whole way versus the histograms of separate clusters.

Unfortunately, the Hough-transform gave vague and ambiguous results due to the low resolution of the image material. Often, the OSM nodes were not placed precisely in the middle of the road, so linear elements were detected which did not lie in the centerline of the way. Furthermore, the image resolution turned out to be too low to detect fragile elements like street marks reliably. Therefore, the idea to use centerline elements was discarded, and instead the focus was set on distinguishing surface classes by k-means clustering and the examination of the resulting histograms.

CHAPTER 5

Results

In the following chapter, a short overview on the used programming language and software packages is given before presenting the testing results of the proposed method.

For testing, OSM ways which already have a surface tag were used. The classification accuracy could thus be evaluated by comparing the original surface tag with the results from the decision tree classifier and image-analysis phase.

5.1 Programming language and libraries

Programming was exclusively done in the Python programming language (version 3.4.2) [Pyt] with the help of suited open source libraries. This choice was made for several reasons: Due to the script-language nature of Python, programming can be done rapidly and it is also easy to run the code on different machines, without the need to consider platform-specific details. Furthermore, the project partner (Bikemap) has their whole system set up in Python and thus aims to reuse parts of the code commercially and without the need to port it to another language.

For geoprocessing, the Geospatial Data Abstraction Library (GDAL, version 2.1.0) [GDA] was used. It can handle both raster and vector data, and also provides methods to combine orthofotos with the corresponding OSM data. It also offers a driver for handling OSM data similar to an XML parser, but with additional geoprocessing methods. For image-analysis, methods and algorithms from both SciPy (version 0.18.2) [Scib] and OpenCV (version 3.1.0) [Ope] were used. While SciPy offers a set of various tools for scientific computation (ranging from mathematical methods for data processing to graph plotting tools), OpenCV provides fast implementations of methods from computer vision. For machine learning, both the decision tree from the data analysis package scikit-learn (version 0.18.1) [scia], and the DecisionTree package (version 3.4.3) [Kak] were tested.

5.2 Results of the OSM analysis phase

For testing the performance of the OSM analysis phase, OSM ways which already have a surface tag were used to construct a decision tree and evaluate its classification accuracy. In order to avoid biased results, different ways were used for the training and testing phase. When the same area was used for training and testing, the available data was split in two parts, so one half was was used only for training and the other half only for testing.

Different combinations of training- and testing areas were evaluated, in order to test if regions of similar structure perform better on each other than regions with different structure. The initial assumption was that the classifier performs better on regions which are similar. Intuitively, using an urban area as training set will hardly generate a tree which is capable of classifying hiking trails within a rural area and vice versa.

Separate tests were conducted for evaluating the method when using OSM data only, and OSM data in combination with additional data.

The results for the testing areas within Austria are presented in table 5.1. As expected, the classification results vary depending on which region is used as training region, and which one is used as testing region.

In all cases, L1 classification performs better than L2 classification, with around 80% to 90% of ways being classified correctly.

When using a training set that is similar to the testing set, better classification results can be achieved compared to using two sets with a different structure regarding surface and highway types.

But even when using urban areas for both training and testing, it is not guaranteed that an optimal classification tree will be built. When using Vienna as training area, it performs worse for Villach than when using Villach on Villach. Even though both areas are urban and from the same country, the type of ways occurring can be different.

In almost all cases, except for Spittal, the best classification results are achieved when using half of an area for training and the other half for testing. This is probably due to the distribution of the surface classes.

Figure 5.1 shows which surfaces occur how often in Villach, Oberndorf and Vienna. Villach has much less surfaces of the class “stone” (e.g. cobblestone, pavingstone) compared to Vienna. Vienna on the other hand has much less samples of the class “gravel” than Villach. This could be a reason why the two areas perform worse than one might expect when used for L2 classification on each other.

When comparing the structure of the surface classes for the cases where the datasets are split into halves in order to use the same area for training and testing, the distribution of surface classes is almost the same for both testing and training sets. (See figures 5.2, 5.3, 5.4.) This is one possible explanation why using the same area for training and classification yields the best results.

Table 5.1: Results of the classification tests. Overall L1 classification performs better than L2 classification, with 80% to 90% of the testing samples being classified correctly. In all cases the best classification results are achieved when using half of an area for training and the other half for testing. This is probably due to the distribution of the surface classes: When using a training set that is very similar to the testing set, better classification results can be achieved compared to when using dissimilar sets.

Train \Test	Villach	Oberndorf	Wien
Villach	L1: 686 / 789 (87%) L2: 540 / 789 (68.5%)	L1: 7065 / 8620 (82%) L2: 5030 / 8620 (58%)	L1: 842 / 921 (91.5%) L2: 501 / 921 (54%)
Oberndorf	L1: 1394 / 1578 (88%) L2: 996 / 1578 (63%)	L1: 3813/ 4310 (88.5%) L2: 3445 / 4310 (78%)	L1: 709 / 921 (78%) L2: 358 / 921 (39%)
Wien	L1: 1464/ 1578 (93%) L2: 835 / 1578 (53%)	L1: 7007 / 8620 (81%) L2: 2984 / 8620 (34.5%)	L1: 424 / 461 (92%) L2: 301 / 461 (65%)

Thus, in order to gain the best possible results, regions should either be used "on themselves" for training and classification, or one should make sure that the training and testing data has a similar structure regarding their surface class distribution.

Note that in figures 5.1, 5.2, 5.3 and 5.4, the classes "paved" and "unpaved" appear in the L2 training and testing sets. This was not intended when regarding the classification taxonomy, but numerous ways were tagged only as "paved" and "unpaved" in OSM. Those ways cannot be assigned to a finer grained class. The choice was made to use them in the L2 classification as they are instead of biasing the results by arbitrarily assigning them to one of the finer-grained classes.

Regarding the classification accuracy, figure 5.5 shows bar charts of the estimated accuracies for the classification results when using half of an area for training and half of it for classification. In all regions, L1 reaches higher probabilities, which matches the observation that L1 classification performs better than L2 classification.

The accuracy is measured by the probability that a result is correct, thus the scale on the x-Axis of all diagrams is ranging from 0 to 1. It becomes apparent that the estimated accuracies for L1 classification are higher than the ones for L2 classification, which also explains why more correct results were achieved for L1 than for L2.

In table 5.1, all classification results are considered regardless of the achieved classification accuracy. When considering only results with a certainty above a threshold (e.g., all the results which have a classification accuracy above 0.6), the relative amount of correct classified samples can be increased, but on the other hand, less samples will be classified. It lies in the hand of the user of the system to decide what is more important.

5.2.1 DTM and Slopes

For testing if DTM and slope data increases the classification accuracy, separate tests with and without these data were conducted. Tables 5.2 and 5.3 show that for the regions

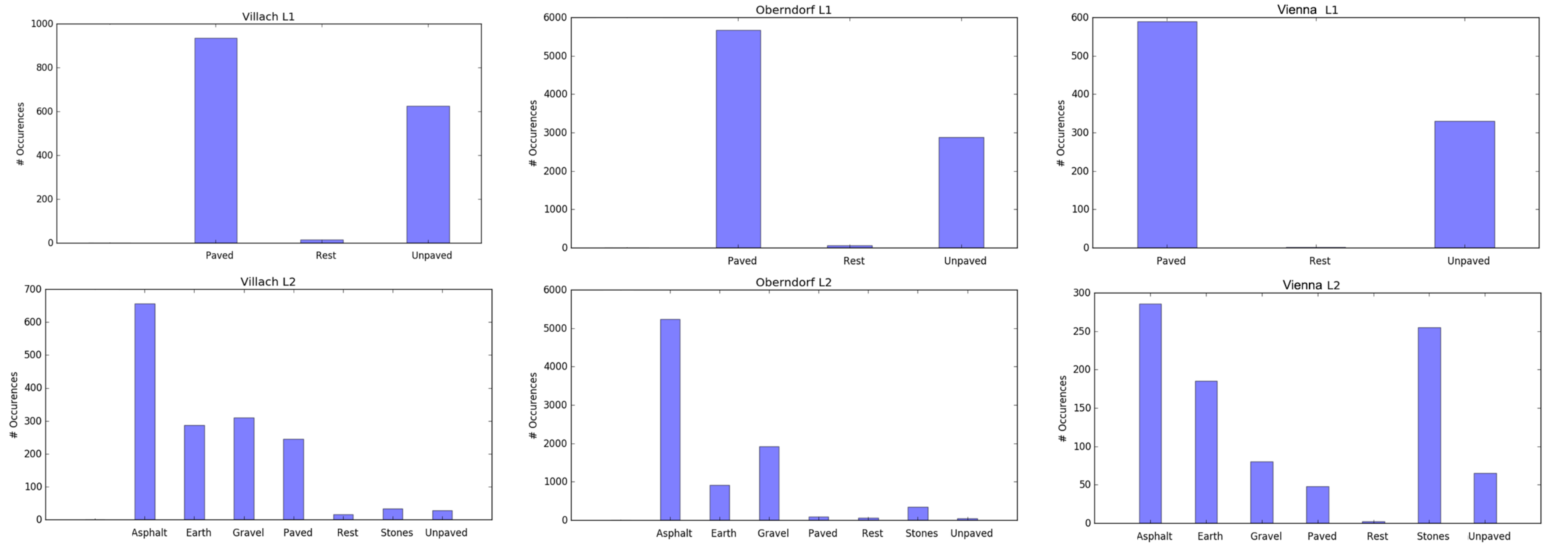


Figure 5.1: Structure of the surface classes for the testing areas Villach, Oberndorf and Vienna (city and periphery), for both L1 and L2 classification. While L1 classes have a similar structure in all areas, big differences occur for L2 classes. This is one possible explanation why L1 classification performs better than L2 classification across different testing areas.

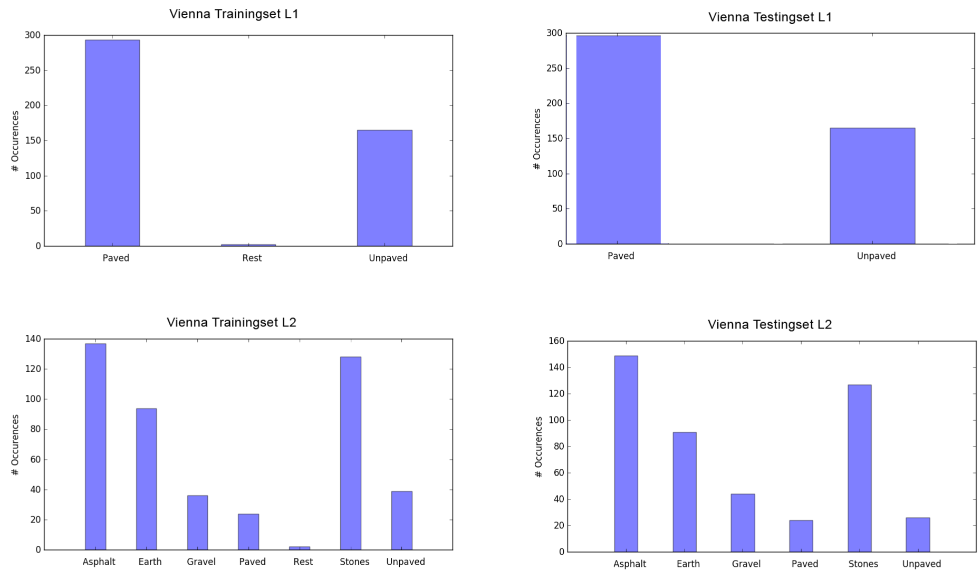


Figure 5.2: Surface classes for Vienna when splitting the data in two parts. Both L1 and L2 classes have a similar structure.

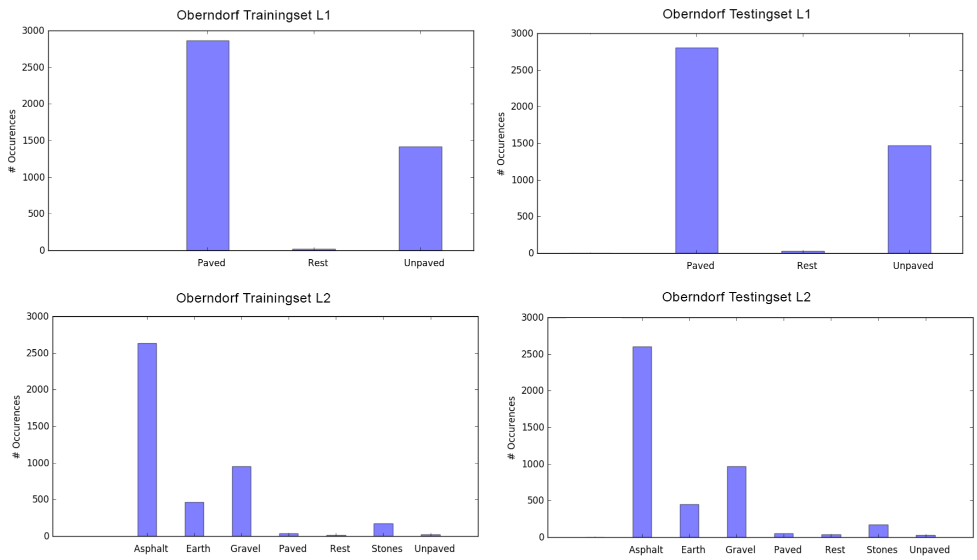


Figure 5.3: Surface classes for Oberndorf when splitting the data in two parts. Notice the similar structure between L1 and L2 for both halves of the data.

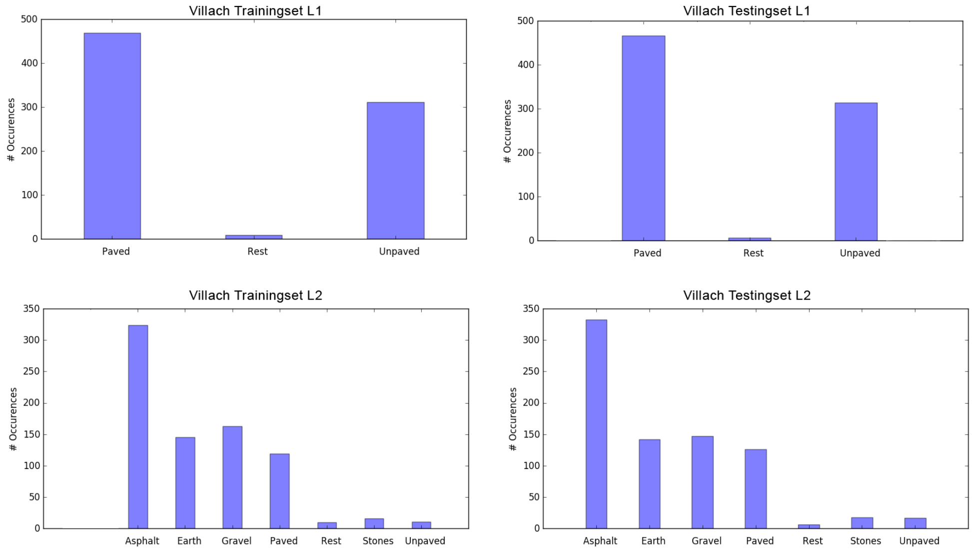


Figure 5.4: Surface classes for Villach when splitting the data in two parts.

Villach, Oberndorf and Spittal an der Drau, classification results improved significantly only in one case, while for all the other cases, results stay the same or even worsen slightly. This could be due to the urban character of the testing areas, which have a high amount of asphalted ways independent of the height and slopes. It is possible that for example in mountain areas in high altitude, DTM and slope data is beneficial for classification, but the evaluation would be a topic for further research.

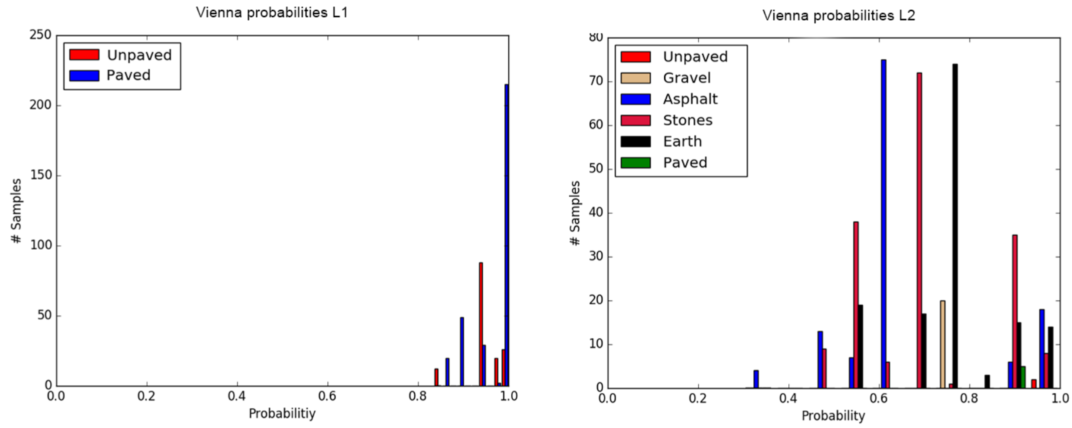
5.2.2 Manual analysis

As the accuracy of the classifier was tested on ways which already have a surface tag so far, the question how accurate the developed method works on ways which have no surface tag remains open. Because it is not possible to perform automated tests for those cases, manual tests on 200 ways were conducted by comparing the classifier's result with orthofotos.

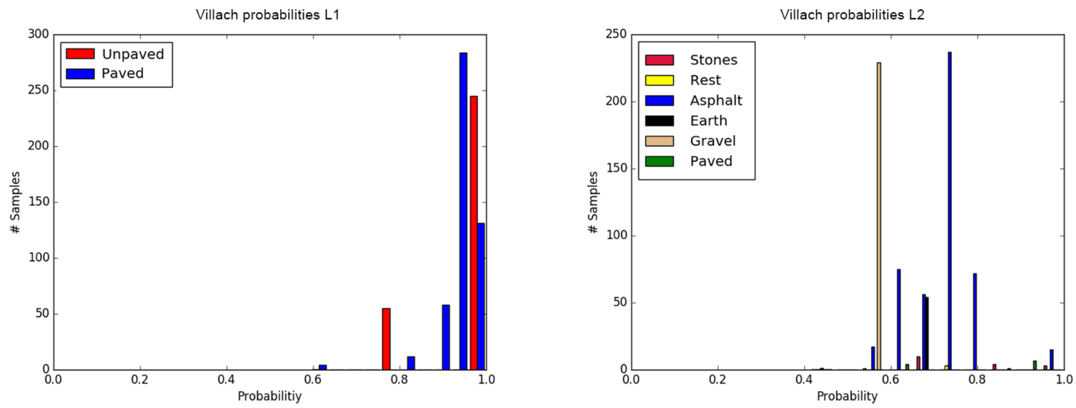
For this evaluation, Oberndorf was used as training area and Villach as testing area. 200 ways of those which had not had an OSM tag before the analysis were randomly selected for testing.

Out of those 200 Ways, 168 were correctly classified, 20 were wrongly classified, and the rest was correctly classified for L1, but not necessarily for L2.

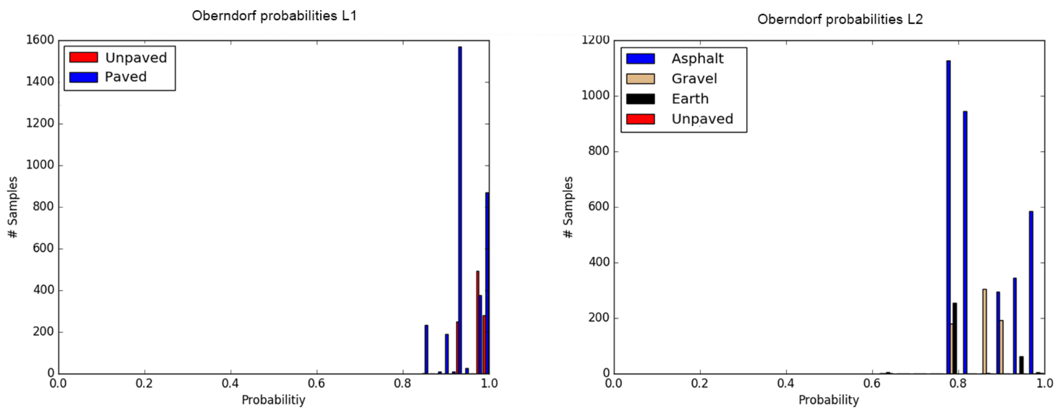
The incorrectly classified ways were all tagged with "highway=residential", "highway=track" or "highway=living-street". Eventually, there was a "tracktype=grade 1" tag. All of these ways were classified as paved and asphalted, but were actually covered with gravel or an earthy surface. This is because all the input samples for "tracktype=grade 1" in the area of Oberndorf are asphalted, so the decision tree learned that "grade1" always



(a) Vienna city center and periphery together

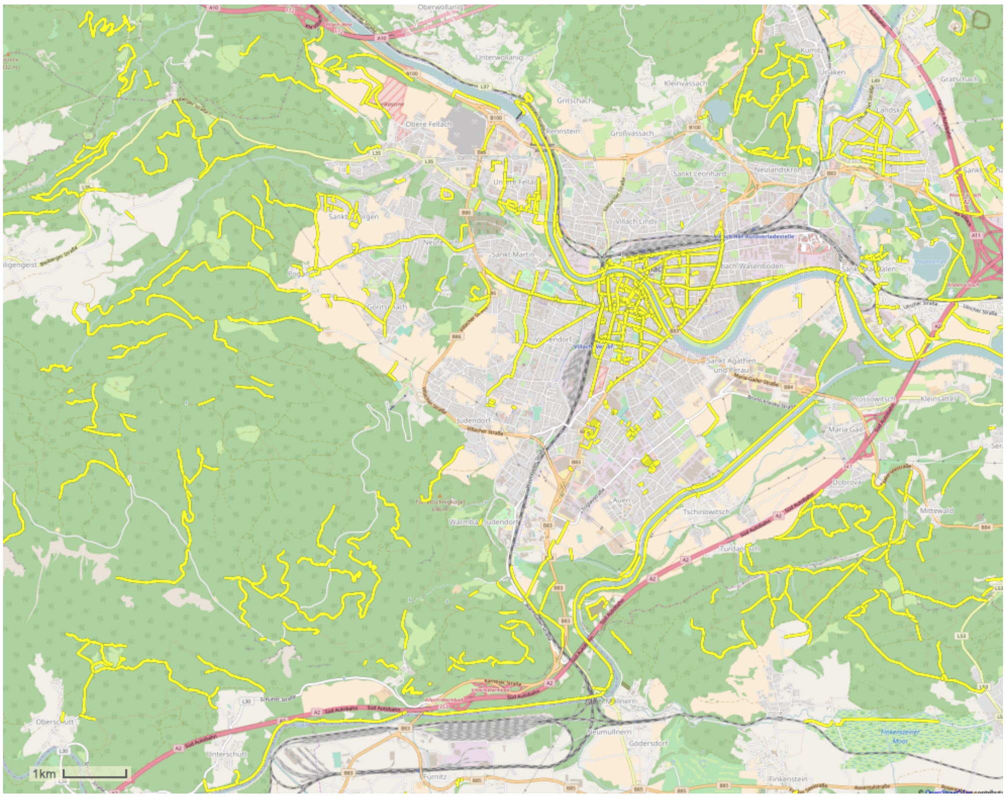


(b) Villach

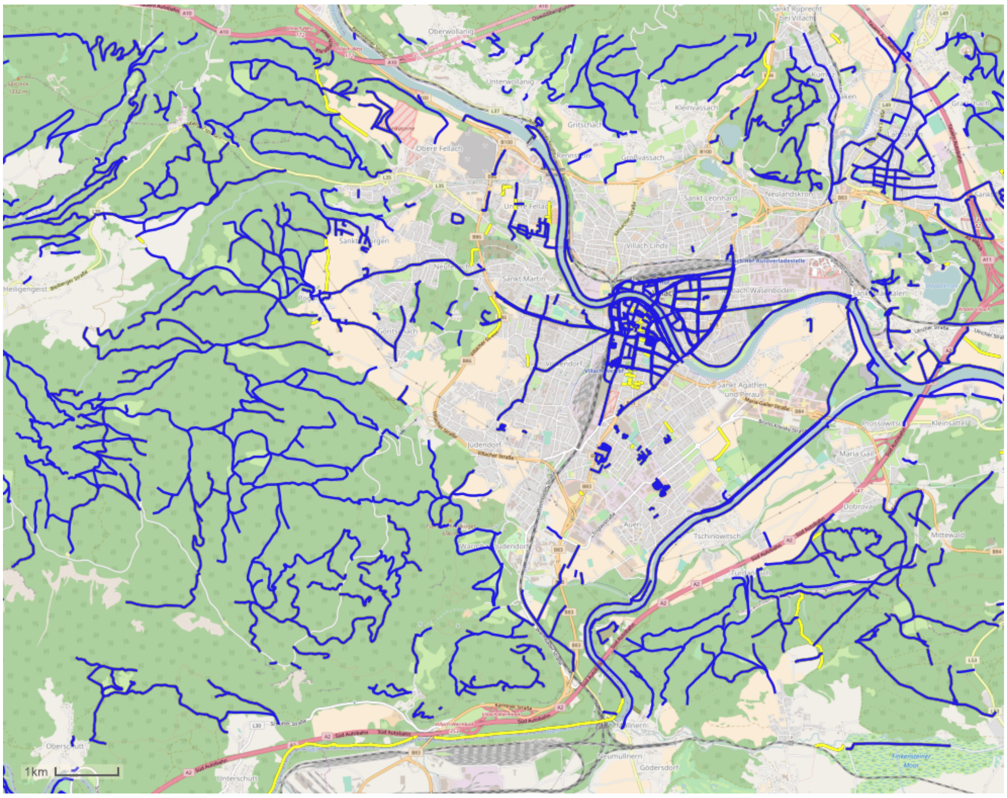


(c) Oberndorf. From all three areas presented, this one has the highest probabilities and thus estimated accuracy for classification. This is expected, becasue from all areas, Oberndorf performed best both in L1 and L2 classifiaction.

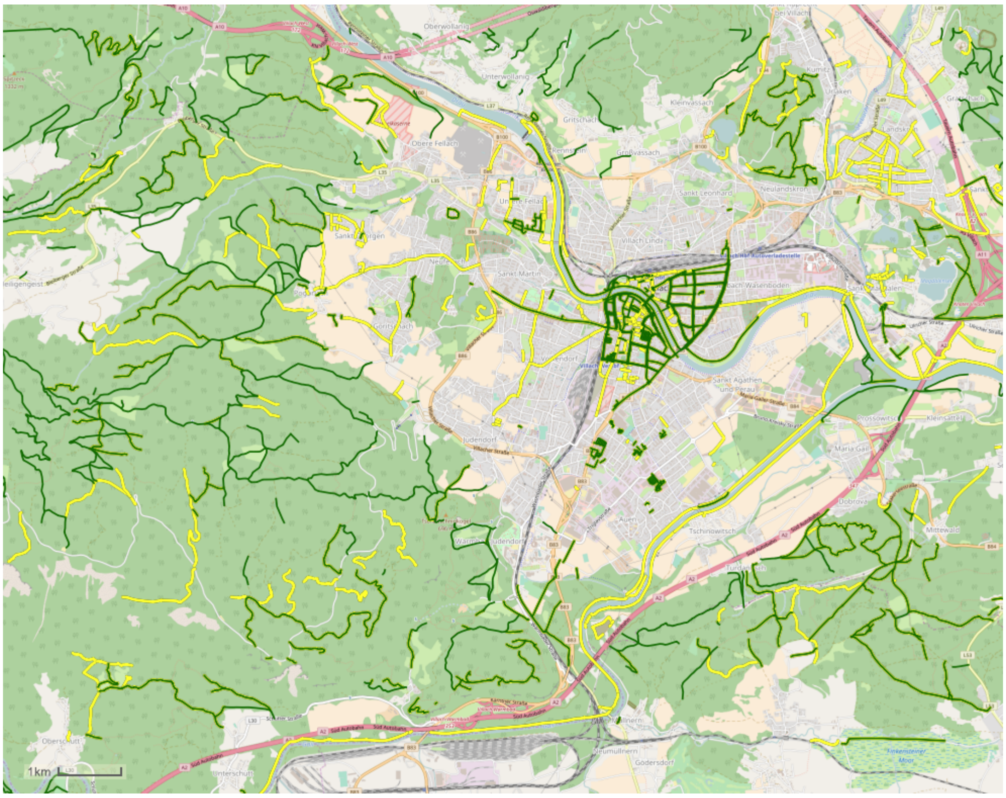
Figure 5.5: Probabilities of classification accuracy when using half of an area for testing and the other half for training. In all regions, L1 reaches higher probabilities, which matches the observation that L1 classification performs better than L2 classification.



(a) Yellow: All ways in the testing area of Villach which have a surface tag.



(b) Blue: All correctly L1-classified ways. Yellow: Incorrectly classified ways.



(c) Green: All correctly L2-classified ways. Yellow: Incorrectly classified ways.

Figure 5.6: Classification of Villach when using Vienna as training set. Wrongly classified ways tended to be paved with gravel, which is a class which does not appear very often in the area of Vienna.

denotes an asphalted way. For ways with no tracktype tag at all, it is likely that the land-cover value biased the results, because it often indicates an urban area in the testing region of Villach, which is most likely to be asphalted.

The ways which were correctly L1-classified but not necessarily for L2 were ways which had either gravel or earth as surface. Even when examining those ways visually, it was not possible to tell if the paving consists of earth or gravel for each case. What was sure is that those ways were unpaved based on the width, location and structure of the road.

The conclusion of the manual analysis is that the method performs also good on ways which have no surface tag, but that the training set should be balanced and suited for a certain testing area. It also suggests that more fine-grained land-cover data would be useful.

Figure 5.6 shows a visualization of the correctly classified L1 and L2 ways. All reference ways which already have a surface tag are highlighted in yellow. The correctly classified L1 ways are highlighted in blue, the correctly classified L2 ways are highlighted in green.

Table 5.2: Classification results for Villach, Oberndorf and Spittal a.d. Drau without using DTM and slope data.

Train\Test	Villach	Oberndorf	Spittal a.d. Drau
Villach	L1: 479 / 526 (91.06%) L2: 332 / 526 (63.11%)	L1: 7065 / 8620 (81.96%) L2: 4309 / 8620 (49.98%)	L1: 612 / 772 (79.27%) L2: 273 / 772 (35.36%)
Oberndorf	L1: 1394 / 1578 (88.33%) L2: 996 / 1578 (63.11%)	L1: 3821 / 4310 (88.65%) L2: 3440 / 4310 (79.81%)	L1: 510 / 772 (66.06%) L2: 367 / 772 (47.53%)
Spittal a.d. Drau	L1: 1051 / 1578 (66.60%) L2: 597 / 1578 (37.83%)	L1: 6787 / 8620 (78.73%) L2: 5471 / 8620 (63.46%)	L1: 288 / 386 (74.61%) L2: 183 / 386 (47.40%)

Table 5.3: Classification results for Villach, Oberndorf and Spittal a.d. Drau with DTM and slope data. Only the L2 classification accuracy of Oberndorf/Villach increased significantly, the other areas show no improvement or even worse performance to when using no DTM.

Train\Test	Villach	Oberndorf	Spittal a.d. Drau
Villach	L1: 477 / 526 (90.68%) L2: 332 / 526 (63.11%)	L1: 7065 / 8620 (81.96%) L2: 4337 / 8620 (50.31%)	L1: 614 / 772 (79.53%) L2: 290 / 772 (37.56%)
Oberndorf	L1: 1394 / 1578 (88.33%) L2: 891/ 1578 (56.46%)	L1: 3821 / 4310 (88.65%) L2: 3440 / 4310 (79.81%)	L1: 510 / 772 (66.06%) L2: 307/ 772 (39.76%)
Spittal a.d. Drau	L1: 1055 / 1578 (66.85%) L2: 599 / 1578 (37.95%)	L1: 6788 / 8620 (78.74%) L2: 5471/ 8620 (63.46%)	L1: 288 / 386 (74.61%) L2: 155 / 386 (40.15%)

5.3 Results of the image-analysis phase

The goal to create an image-analysis phase which classifies ways based on the analysis of RGB values was only partially met. Distinguishing paved and unpaved surfaces, analogue to L1 classification, is partially possible by examining the histograms of clusters of similar pixels.

A working example is depicted in figure 5.7, which shows one way covered with gravel and two other ones covered with grass (according to OSM data). After the clustering procedure, 3 clusters are gained, which are ordered according to their mean values of the RGB channels (see table 5.4). The cluster with the highest mean values is the cluster which contains the brightest values for human vision. Thus in the given table, cluster C is most likely to represent a paved way.

For the way covered with gravel, cluster B and cluster C contain the majority of pixels, while cluster A, the one with the darkest pixels, has a relatively low percentage of the total pixel count. Both B and C have relatively high mean values, and C has a relatively low standard deviation, which is an indicator that only few mixed pixels are present. Thus, it is likely that the road is paved.

The other two roads in the image have a low percentage of pixels in cluster C, which indicates that only a few bright pixels are present. Clusters A and B have rather low

Table 5.4: Comparison of spectral properties between ways covered with gravel and grass. For the way covered with gravel, cluster B and cluster C contain the majority of pixels, while cluster A (which represents darker pixels) has a relatively low pixel count. Both B and C have relatively high mean values, thus it is likely that the road is paved.

	Portion of total pixels			Spectral mean value				Spectral standard deviation		
Way	A	B	C		A	B	C	A	B	C
(a) Gravel	12	50	38	R	84.5	108.0	157.8	8.1	9.3	11.2
				G	98.3	115.4	157.6	7.8	7.2	11.3
				B	72.4	91.1	145.7	6.0	8.9	13.1
(b) Grass	31	63	6	R	80.2	92.6	159.9	5.4	7.4	18.6
				G	95.5	106.2	162.1	5.3	5.6	17.3
				B	70.1	80.3	146.8	4.2	5.8	23.2
(c) Grass	21	70	9	R	82.6	97.9	149.3	4.4	7.6	18.2
				G	97.5	109.9	151.4	4.1	5.9	15.3
				B	71.7	83.9	130.9	3.7	6.2	17.4

Table 5.5: Comparison of the spectral properties of asphalt and gravel. The classes are barely distinguishable when examining the mean RGB values, pixel count or standard deviation for the separate clusters.

	Portion of total pixels			Spectral mean value				Spectral standard deviation		
Way	A	B	C		A	B	C	A	B	C
(a) Asphalt	18	32	50	R	100.7	122.2	150.9	7.5	7.5	13.2
				G	115.5	124.8	147.8	6.9	6.3	14.5
				B	85.8	101.1	129.4	6.8	7.8	13.7
(b) Gravel	21	49	30	R	101.3	101.3	183.4	7.4	8.1	14.1
				G	118.0	118.0	185.5	6.1	7.9	13.4
				B	92.8	92.8	175.4	6.2	7.1	13.8

mean values and a small standard deviation, which is a strong indicator for an unpaved way in a grassy area.

While the analysis of clusters gives a first hint if a way is paved or not, further distinction of the paving classes was not possible, due to the high spectral similarity of the RGB channels of asphalt, stones and gravel.

As can be seen in figure 5.9, those two classes can barely be distinguished even by a human observer. When examining the mean values and standard deviations of this example (see table 5.5), there is no significant difference between the values for asphalt or gravel. The given example is representative for other paved ways. The evaluation of 100 ways which were either asphalted or covered with gravel showed that there is no significant statistical difference to tell those two types automatically apart with the proposed method.

This means that for L2 classification, the presented method in combination with the available orthophotos proved to be impractical, because of the low spectral resolution of

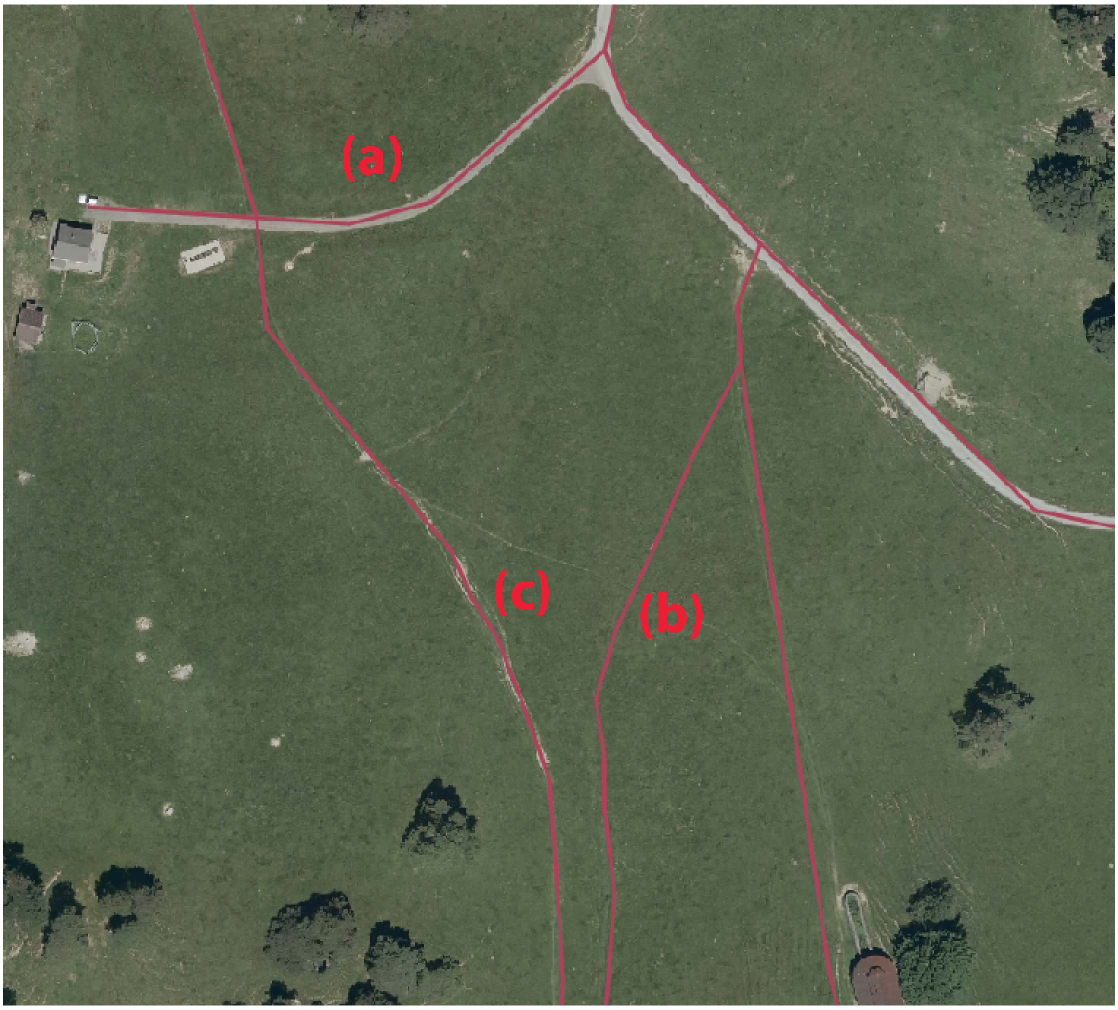


Figure 5.7: Comparison of ways covered with gravel and grass. The way with the gravel surface (a) is visually distinguishable from the ways with the grass surface (b, c), because it consists of much more bright pixels with relatively high RGB values.

Figure 5.8: Example of an asphalted way which can be clearly distinguished from unpaved surfaces by the image-analysis procedure. The cluster with the brightest RGB values has mean values of 146.8, 149.0 and 131.4, and contains around 40% of the total pixels. The second highest cluster has similar properties and the cluster with the lowest RGB values has mean values of 39.8, 57.6 and 51.4, and a low pixel count. These properties are an indicator for a bright, paved street.

the image material. Furthermore, the low spatial resolution and occlusions caused by trees and made it impossible to detect narrow objects like street marks automatically. With a higher spatial resolution, more sophisticated texture analysis could have been performed. A resolution which is sufficient to detect structures which are typical for certain types of paving would be desirable. For example, street marks are typical for asphalt, grid patterns are typical for paving stones, or a high amount of noise and grain would indicate a gravel surface. Eventually, the field of computer vision will develop more sophisticated methods in the future which can handle recognition of small and fine-grained elements in orthophotos.

Utilizing more spectral channels than only RGB could improve the classification results, as for example greenlands and trees could be identified more easily with a NIR channel, but this would increase costs drastically and thus oppose the idea to design a cost-efficient method.

Alternatively to an automated approach, a semi-automated approach could be taken where humans have to annotate a training set of roads which could be used with a suited classifier. A graphical user interface, which works based on a prior k-means clustering, could facilitate this process. A screenshot of a prototype developed for this purpose can

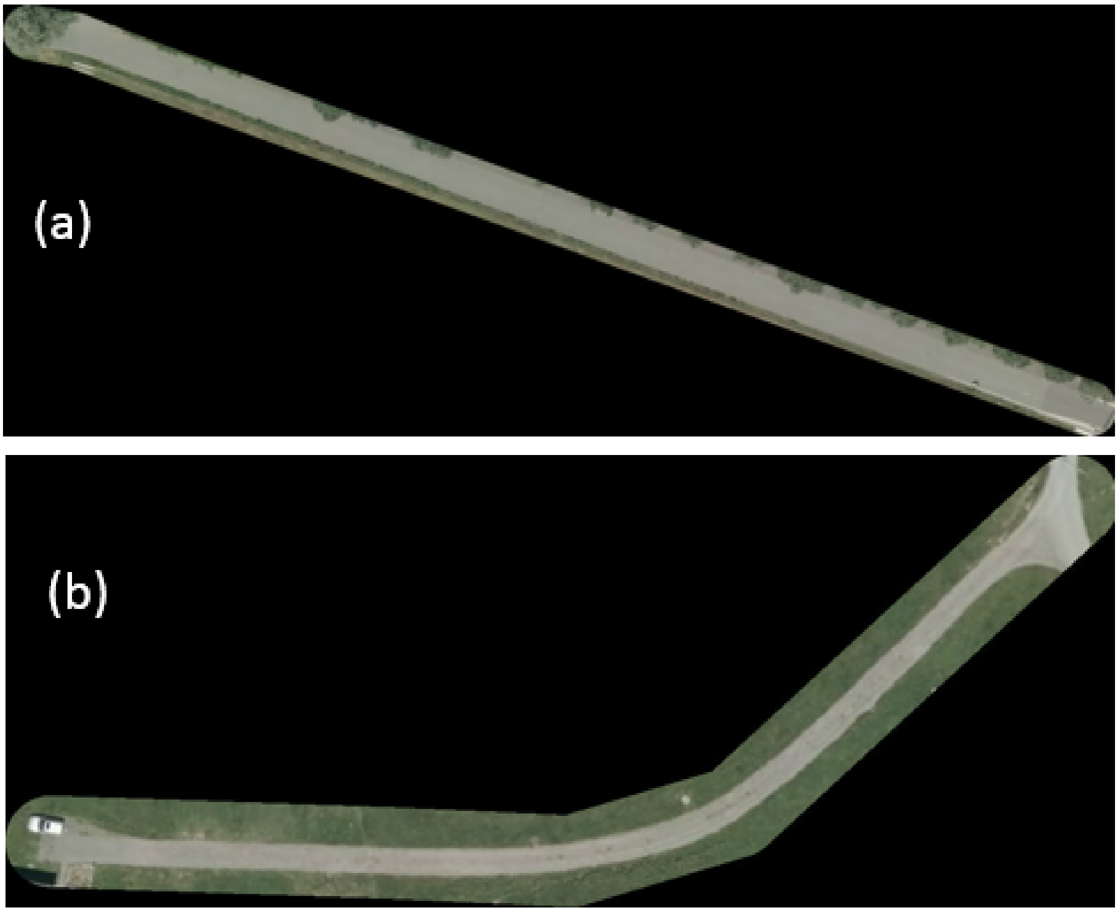


Figure 5.9: Comparison of two ways which are paved with asphalt (a) and gravel (b). Even for a human observer, the surfaces are hardly distinguishable due to the similar RGB values.

be seen in figure 5.10. With such a tool, a human user could select and annotate surface pixels easily.

The downsides of such a semi-automatic approach are that it is a cumbersome process, which would have to be repeated for every orthofoto source. Furthermore, even a human has a hard time distinguishing surfaces like gravel and asphalt only from a photo.

Figure 5.10: For a semi-automatic approach, a graphical interface based on prior k-means clustering could help a user to select and annotate surface regions. When a user selects a region by clicking on it, all pixels belonging to that cluster are selected. Visual feedback is given by highlighting the region semi-transparently. (Red in this example.) The k-means clusters used for this example are the ones depicted in figure 4.7.

CHAPTER 6

Conclusion and future work

In this chapter, a summary of the developed method is given, including a recap of the achieved results. Problems that occurred are discussed and an outlook for future improvements is given.

6.1 Conclusion

This thesis presented a fully automatic approach for classifying road-surfaces based mainly on crowd-sourced data gathered from OSM. In OSM, information about road properties is stored in so-called "tags", which are textual key-value pairs. From those tags, information regarding surface properties can be deduced.

This work includes a thorough analysis of the OSM tagging system, focusing on the quality and quantity of available tags and their distribution on a world-wide and regional scale.

While several tags proved to be useful for deducing road-surface properties, one challenge was the rather sparse availability of tags with useful information – missing data was the rule and not the exception.

Furthermore, DTM, slope and land-cover data was added for improving classification results. The system can be extended with further data, given it can be represented in textural or numerical form.

Classification is performed by a decision tree, which is automatically built from a set of input training data. A focus was laid on the decision tree classifier because it already proved successful in the field of land-use and land-cover (LULC) classification in the past, and because the classification process itself is easily understandable for humans.

The system is able to perform well when classifying if a road is paved or not. Up to 90% of roads in the testing scenarios were correctly classified. Performance for a finer-grained classification ranged between 60% to 75% in the testing scenarios. Adding further geospatial data may improve those results. Carefully selecting and preprocessing additional data so that it complements the testing areas in a meaningful way is recommended.

The advantage of the presented method is that it is fast and can be applied to regions world-wide under the condition that enough OSM data is available for a certain region.

Image-analysis of high-resolution orthophotos was performed on roads where no OSM data was available. The image-analysis phase was based on RGB channels only and help to distinguish between paved and unpaved ways for which no OSM data was available at all, but a more detailed classification was not possible due to the spectral similarity of the RGB values of asphalt, cobblestone and gravel. Extracting and classifying ways from orthophotos in a fully automatic way craves for higher resolution spatial and spectral images. A NIR channel for example would help to separate vegetation from non-vegetation and facilitate the extraction of streets. A higher spatial resolution could make more sophisticated methods from computer vision applicable. For example, textural analysis could be performed to differentiate asphalt from gravel based on the noise and grain of the surface.

6.2 Future work

As this work is one of the first attempts to classify road-surfaces based on crowd-sourced open source data, there is a lot of potential for future studies in this field.

Future research could focus on creating optimal and distinctive training data sets to improve the results of the OSM analysis phase. Generating a training set with selected samples that have a similar distribution of surfaces as the target area could be one way to improve classification results.

Another option would be testing other classification algorithms and see if they perform better than the chosen model. Random Forests, SVM classifiers and Naive Bayes classifiers are just some possible classifiers which could be evaluated.

There is also potential to improve classification accuracy by adding further features to the data samples. Useful data could be derived from Open Source GIS databases, like the already incorporated land-cover data and slope- and height information. Furthermore, more fine-grained data with regard to land-cover information could also improve classification accuracy.

Also, it is likely that OSM becomes more complete regarding tags over time, as the whole project is always in an “under construction” status. This means that not only the number of “surface”, “highway” and “tracktype” tags is likely to increase, but also that other tags will eventually become common enough to be used for the classification phase.

Another aspect for future work would be the improvement of the image-analysis phase. The most straightforward option would be to utilize orthophotos of higher resolution and with more spectral channels, but this would not resolve the problems when isolating ways from the photos (e.g., not-centered centerlines, shadow artifacts, occlusion by trees, etc.).

Another approach would be to evaluate how suited panorama photos like the ones available from Google Street View are to determine street surface covering.

83

Bibliography

[AMZS15] Jamal Jokar Arsanjani, Peter Mooney, Alexander Zipf, and Anne Schauss. Quality assessment of the contributed land use information from openstreetmap versus authoritative datasets. In *OpenStreetMap in GIScience*, pages 37–58. Springer, 2015.

[BEV] Bundesamt für eich- und vermessungswesen - updates. http://www.bev.gv.at/portal/page?_pageid=713,1573999&_dad=portal&_schema=PORTAL1. Accessed: 2016-07-18.

[Bik] Bikemap. http://www.bikemap.net. Accessed: 2016-07-3.

[Bis01] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, New York, 2001.

[Bla10] Thomas Blaschke. Object based image analysis for remote sensing. *ISPRS journal of photogrammetry and remote sensing*, 65(1):2–16, 2010.

[C50] C5.0. http://www.rulequest.com/see5-info.html. Accessed: 2017-02-10.

[CIA] The cia factbook. https://www.cia.gov/library/publications/resources/the-world-factbook/rankorder/2085rank.html. Accessed: 2016-09-05.

[CORa] Corine. http://www.eea.europa.eu/publications/COR0-landcover. Accessed: 2017-02-09.

[CORb] Corineclasses. http://www.umweltbundesamt.at/umweltsituation/raumordnung/rp_projekte/rp_corine. Accessed: 2017-02-09.

[CW11] James B Campbell and Randolph H Wynne. *Introduction to remote sensing*. Guilford Press, 2011.

[DHS00] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification (2Nd Edition)*. Wiley-Interscience, 2000.

[DLR03] Ahmed Darwish, Kristin Leukert, and Wolfgang Reinhardt. Image segmentation for the purpose of object-based classification. In *International Geoscience and Remote Sensing Symposium*, volume 3, pages III–2039, 2003.

[dN04] Morten Ødegaard Nielsen. True orthofoto genration. Master's thesis, DTU, Kgs. Lyngby, DK, 2004.

[Dtr] Decisiontree 3.4.3 - decision tree package for python by avinash kak. `https://engineering.purdue.edu/kak/distDT/DecisionTree-3.4.3.html`. Accessed: 2016-20-12.

[eCo] ecognition. `http://www.ecognition.com/`. Accessed: 2017-02-01.

[EP13] Jacinto Estima and Marco Painho. Exploratory analysis of openstreetmap for land use classification. In *Proceedings of the second ACM SIGSPATIAL international workshop on crowdsourced and volunteered geographic information*, pages 39–46. ACM, 2013.

[ERO] Eros-c satellite. `http://www.imagesatintl.com/eros-sat`. Accessed: 2016-09-15.

[GAS03] hsio-hua K. Burke Gary A. Shaw. Spectral imaging for remote sensing. *Lincoln laboratory Journal*, 14(1):3–28, 2003.

[GDA] Gdal - geospatial abstraction library official website. `http://www.gdal.org/l`. Accessed: 2017-04-1.

[Geoa] Geoimage website - resolution. `http://www.geoimage.at/geodatenprodukte.html`. Accessed: 2016-07-10.

[Geob] Geoimage website - update cycles. `http://www.geoimage.at/metadaten.html`. Accessed: 2016-07-10.

[GLSGFV10] Pedro J García-Laencina, José-Luis Sancho-Gómez, and Aníbal R Figueiras-Vidal. Pattern classification with missing data: a review. *Neural Computing and Applications*, 19(2):263–282, 2010.

[Goo07] Michael F Goodchild. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4):211–221, 2007.

[Hak] Completeness in volunteered geographical information — the evolution of openstreetmap coverage (2008–2009). `https://povesham.wordpress.com/2010/08/13/completeness-in-volunteered-geographical-information-%E2%80%93-the-evolution-of-openstreetmap-coverage-2008-2009`. Accessed: 2016-09-24.

[Hak10] Mordechai Haklay. How good is volunteered geographical information? a comparative study of openstreetmap and ordnance survey datasets. *Environment and planning B: Planning and design*, 37(4):682–703, 2010.

[HGNR03] Martin Herold, M Gardner, Val Noronha, and D Roberts. Spectrometry and hyperspectral remote sensing of urban road infrastructure. *Online Journal of Space Communications*, 3, 2003.

[HRGD04] Martin Herold, Dar A Roberts, Margaret E Gardner, and Philip E Dennison. Spectrometry for urban area remote sensing—development and analysis of a spectral library from 350 to 2400 nm. *Remote Sensing of Environment*, 91(3):304–319, 2004.

[HW08] Mordechai (Muki) Haklay and Patrick Weber. Openstreetmap: User-generated street maps. *IEEE Pervasive Computing*, 7(4):12–18, October 2008.

[JAHB+13] Jamal Jokar Arsanjani, Marco Helbich, Mohamed Bakillah, Julian Hagenauer, and Alexander Zipf. Toward mapping land-use patterns from volunteered geographic information. *International Journal of Geographical Information Science*, 27(12):2264–2278, 2013.

[Kak] Avinash Kak. Decision tree for python package. `https://engineering.purdue.edu/kak/distDT/DecisionTree-3.4.3.html`. Accessed: 2017-01-2.

[KMJ10] Asha Gowda Karegowda, AS Manjunath, and MA Jayaram. Comparative study of attribute selection using gain ratio and correlation based feature selection. *International Journal of Information Technology and Knowledge Management*, 2(2):271–277, 2010.

[KSF05] Florian P Kressler, Klaus Steinnocher, and Michael Franzen. Object-oriented classification of orthophotos to support update of spatial databases. In *Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS'05.*, volume 1, pages 4–pp. IEEE, 2005.

[LKC14] Thomas Lillesand, Ralph W Kiefer, and Jonathan Chipman. *Remote sensing and image interpretation*. John Wiley & Sons, 2014.

[LS13] Xiaoxiao Li and Guofan Shao. Object-based urban vegetation mapping with high-resolution aerial photography as a single data source. *International journal of remote sensing*, 34(3):771–789, 2013.

[LS14] Xiaoxiao Li and Guofan Shao. Object-based land-cover mapping with high resolution aerial photography at a county scale in midwestern usa. *Remote Sensing*, 6(11):11372–11390, 2014.

[Mar] How complete is openstreetmap? https://www.mapbox.com/blog/how-complete-is-openstreetmap. Accessed: 2016-07-20.

[Moh12] M Mohammadi. Road classification and condition determination using hyperspectral imagery. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*, 39:B7, 2012.

[MS05] Yasuyo Makido and Ashton Shortridge. Land cover mapping at sub-pixel scales: Unraveling the mixed pixel. In *Proceedings of the 8th International Conference on GeoComputation*, pages 1–3. Citeseer, 2005.

[Mun] Mundraub website. mundraub.org. Accessed: 2016-09-24.

[Nei] A comparative study between different openstreetmap contributor groups – outline 2016. http://neis-one.org/2016/11/osm-contributor-groups/. Accessed: 2016-11-18.

[NHRG02] Val Noronha, Martin Herold, Dar Roberts, and Meg Gardner. Spectrometry and hyperspectral remote sensing for road centerline extraction and evaluation of pavement condition. In *Proceedings of the Pecora Conference*, 2002.

[NZ12] Pascal Neis and Alexander Zipf. Analyzing the contributor activity of a volunteered geographic information project — the case of openstreetmap. *ISPRS International Journal of Geo-Information*, 1(3):146–165, Jul 2012.

[NZ14] Pascal Neis and Dennis Zielstra. Recent developments and future trends in volunteered geographic information research: The case of openstreetmap. *Future Internet*, 6(1):76–106, 2014.

[Ope] Opencv - open source computer vision. http://www.opencv.org. Accessed: 2017-04-1.

[OSMa] Osm wiki entry for bing aerial imagery. http://wiki.openstreetmap.org/wiki/Bing_Maps. Accessed: 2017-03-14.

[OSMb] Osm - coverage. http://wiki.openstreetmap.org/wiki/Press_Kit#Coverage. Accessed: 2016-09-26.

[OSMc] Overview on download possibilities of osm data. http://wiki.openstreetmap.org/wiki/Downloading_data. Accessed: 2016-11-14.

[OSMd] Osm wiki - osm elements. https://wiki.openstreetmap.org/wiki/Elements. Accessed: 2016-09-14.

[OSMe] Osm - license and copyright. http://www.openstreetmap.org/copyright. Accessed: 2016-07-20.

[OSMf] Osm map features: about tags and their usage. http://wiki.openstreetmap.org/wiki/Map_Features. Accessed: 2016-07-20.

[OSMg] Osm - mission statement. https://wiki.osmfoundation.org/wiki/Mission_Statement. Accessed: 2016-07-20.

[OSMh] Osm wiki - node element. http://wiki.openstreetmap.org/wiki/Node. Accessed: 2016-09-14.

[OSMi] Osm statistics by pascal neis. http://osmstats.neis-one.org/. Accessed: 2016-09-26.

[OSMj] Osm wiki - surface tag. http://wiki.openstreetmap.org/wiki/Key:surface. Accessed: 2016-06-17.

[OSMk] Osm wiki - mapping techniques. http://wiki.openstreetmap.org/wiki/Mapping_techniques. Accessed: 2016-09-09.

[OSMl] Osm wiki - tracktype tag. http://wiki.openstreetmap.org/wiki/Key:tracktype. Accessed: 2016-06-17.

[OSMm] Osm - update cycle, questions and answers. https://help.openstreetmap.org/questions/178/how-often-does-the-main-mapnik-map-get-updated. Accessed: 2016-07-20.

[OSMn] Open street map. https://www.openstreetmap.org. Accessed: 2016-08-12.

[OSMo] planet.osm - the osm weekly dump. http://wiki.openstreetmap.org/wiki/Planet.osm. Accessed: 2016-07-20.

[OSMp] Osm wiki. https://wiki.openstreetmap.org/wiki/Main_Page. Accessed: 2016-07-20.

[OSMq] Osm wiki entry for yahoo aerial imegery. http://wiki.openstreetmap.org/wiki/Yahoo!_Aerial_Imagery. Accessed: 2017-03-14.

[OSMr] Osm wiki - highway tag. http://wiki.openstreetmap.org/wiki/Key:highway. Accessed: 2016-06-17.

[Pyt] Python programming language. http://www.python.org/. Accessed: 2017-02-25.

[Qui86] J. R. Quinlan. Induction of decision trees. *Mach. Learn.*, 1(1):81–106, March 1986.

[Qui14] J Ross Quinlan. *C4. 5: programs for machine learning*. Elsevier, 2014.

[scia] scikitlearn - machine learning in python official website. http://www.scikit-learn.org. Accessed: 2017-04-1.

[Scib] Scipy - scientific computing tools for python official website. http://www.scipy.org. Accessed: 2017-04-1.

[SL91] S Rasoul Safavian and David Landgrebe. A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, 21(3):660–674, 1991.

[SWPF13] Tobias Sturn, Michael Wimmer, Peter Purgathofer, and Steffen Fritz. Landspotting - games for improving global land cover. In *Proceedings of Foundations of Digital Games Conference 2013 (FDG 2013)*, pages 117–125, 2013.

[Tag] Taginfo. https://taginfo.openstreetmap.org/. Accessed: 2016-07-20.

[Twi] Distribution of (active) open street map contributors. https://twitter.com/pascal_n/status/799957976480149505. Accessed: 2016-11-19.

[WF00] Volker Walter and Dieter Fritsch. Automated revsion of gis databases. In *Proceedings of the 8th ACM International Symposium on Advances in Geographic Information Systems*, GIS '00, pages 129–134, 2000.

[WF05] Ian H Witten and Eibe Frank. *Data Mining: Practical machine learning tools and techniques.* Morgan Kaufmann, 2005.

[Wik] Wikimapia. wikimapia.org. Accessed: 2016-09-24.

[WJRJ10] Robert C Weih Jr and Norman D Riggan Jr. Object-based classification vs. pixel-based classification: comparative importance of multi-resolution imagery. *Proceedings of GEOBIA 2010: Geographic Object-Based Image Analysis*, 38:6, 2010.

[WV3] Worldviev-3. http://www.satimagingcorp.com/satellite-sensors/worldview-3. Accessed: 2016-09-15.

[ZQH+15] Qian Zhang, Rongjun Qin, Xin Huang, Yong Fang, and Liang Liu. Classification of ultra-high resolution orthophotos combined with dsm using a dual morphological top hat profile. *Remote Sensing*, 7(12):16422–16440, 2015.

[ZZ10] Dennis Zielstra and Alexander Zipf. A comparative study of proprietary geodata and volunteered geographic information for germany. In *13th AGILE international conference on geographic information science*, volume 2010, 2010.