ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

MASTER PROJECT

Classification of urban structural types (UST) using multiple data sources and spatial priors

Laboratory of Geographic Information Systems
(LASIG)

June 23, 2014

Author:
Arnaud PONCET-
MONTANGES

Supervisor:
Devis TUIA

Expert:
Gabriele MOSER

Abstract

Remote sensing and geographic information science offer many possibilities in terms of availability of diverse data. Some products like land cover layers or digital elevation models can be extracted from imagery and enable the realization of 3D city models. Starting from these morphological and geographical sources, an approach is proposed to extract information about urban structure types (UST), i.e. types of urban habitat at the neighborhoodscale. We propose an effective processing chain to describe UST : from the different data sources, we extract spectral and spatial indices and use them as features in a machine learning process to classify these urban structural types using support vector machine classification (SVM). Moreover, Markov Random Fields (MRF) are used to take into account the spatial distribution of the classe and increase the spatial consistency.

This study focuses on the city of Munich and uses as different data sources the land cover data, the 3D city model, spectral images from LandSat TM 8 and OpenStreetMap (OSM) vector data to characterize UST.

The main hypothesis is that we can discriminate among urban structural types by using land cover information, spectral properties and 3D structure: in other words, that an industrial area won't have the same structure nor the same properties as a residential or an agricultural area.

The proposed processing chain enables to predict with a precision of 70% the 11 UST. This opens possibilities to describe the urban footprint of the city, to detect the key areas for urban planification and to better understand the city dynamics.

Keywords : support vector machine (SVM), classification, urban structural types (UST), Markov random fields (MRF)

Abstract

   Les domaines de la télédétection et de l'information géographique
offrent aujourd'hui beaucoup de possibilités en termes de sources et
diversité de données. Des produits comme des couches d'occupation
du sol ou des modèles numériques de hauteurs peuvent être issus de
l'imagerie et permettent de réaliser des modèles 3D de villes. En
partant de ces sources de données morphologiques et géographiques,
une approche est proposée dans le but d'extraire les types structurels
urbains,c'est-à-dire les types d'habitat urbains à l'échelle du voisi-
nage. Une méthode efficace est proposée pour décrire ces différents
types structurels urbains: des différentes sources de données, une
série d'indices spectraux et spatiaux sont extraits et utilisés dans
un processus d'apprentissage automatique pour ensuite classifier les
différents types structurels urbains présents au niveau du voisinage à
l'aide de séparateurs à vaste marge (SVM). De plus l'usage des champs
aléatoires de Markov (MRF) permet de tenir compte de la distribution
spatiale des classes à priori et de promouvoir des solutions spatiale-
ment cohérentes.

   Cette étude se focalise sur la ville de Munich et exploite différentes
sources de données telles que la couverture du sol et le modèle 3D de
la ville, de même que les images Multispectrales LandSat TM 8 et les
données vectorielles OpenStreet Map (OSM).

   L'hypothèse prinicpale est donc que l'on peut déduire les types struc-
turels urbains sur la base d'informations propres à la géométrie, la
radiométrie et la couverture du sol observées, à savoir qu'une zone in-
dustrielle n'aura pas la même structure ni les même propriétés qu'une
zone résidentielle ou qu'une zone agricole.

   L'approche développée permet de prédire avec une précision de 70%
les 11 classes de types structurels urbains définies. Ceci ouvre des
possibilités pour décrire l'empreinte de la ville sur son environnement
et pour mieux comprendre la dynamique de la ville.

Mots Clefs : séparateurs à vaste marge (SVM), classification, types de
   structures urbaines (UST), champs aléatoires de markov (MRF)

Acknowledgments

For more informal Acknowledgement, please go to appendix D.

Contents

List of Tables and Figures

1 Introduction

Here are presented the motivation, the state of the art, goals of the study and the relevant previous work.

1.1 Motivation

Land use is the characterisation by arrangements, activities and inputs people undertake in a certain land cover type to produce, change or maintain it (FAO/UNEP, 1999). It is therefore a relevant information for urban planning. It includes vital socio-economic and environmental information (Banzhaf & Hofer, 2008). Though the analysis of land use it is possible to derive factors such as energy demand, water supply, waste and traffic generation from it. Land use is a precious information for city planners to understand the city structure and the city dynamics and also on a temporal scale to get the evolution aspects of a city. It is also an economical interest since economical resources are not equally distributed. With the recent drastical size increase of some cities around the world (Taubenböck et al., 2012), land use changes are now in the center of attention. Urban planning and all the domains it considers (as energy resources management, environmental issues and protection of ecosystems, for example) requires that kind of geographic information.(Pauleit & Duhme, 2000)

On a city scale, land use information need to be aggregated. That's why we use urban structural types (UST) to describe the urban footprint of a city. We defined UST as specific spatial patterns of the urban structure at the neighborhood scale. UST are different from land cover (LC) since we do not map a tree or a building but a spatial arrangement of objects in a neighborhood of fixed size: in our case dense residential areas, ...

Most UST such as large commercial activities and industry include very similar types of urban structures and are therefore difficult to distinguish with classic image analysis. Moreover, there is no standardized definition of the urban structure types in the academic literature. (Banzhaf & Hofer, 2008)

Remote sensing and geographic information science offer many possibilities in terms of availability of diverse data. From high resolution satellite imagery and segmentation, it is now possible to obtain high resolution land cover data. (Taubenböck et al., 2006) By fusing building classes with digital elevation models, it is now possible to extract the building heights and vol-

umes, which means that 3D city models are easier to create.(Sirmacek et al., 2012) From these morphological and geographic information sources come possibilities to describe the UST of cities.

1.2 State of the art

Multiple studies have attempted using different data and different methodologies at multiple sites to determine the information of land use or the UST. We have chosen three of theses studies to show different appraches.

In their study, Walde et al. (2014) classify urban structural types. They define 5 classes (city center, residential block buildings, allotment, residential single family homes, industrial areas). They use an object based image analysis approach to extract land cover data. From these data, they extract some graph based measures : centrality indices, ajacency events, connectivity measures and building related measures. Using a random forest classifier, that they trained with training samples, they predict the urban structural class on the city of Rostock (North of Germany) with an overall accuracy of 87%. They found out that the most important feature in their classification algorithm was the building height with the highest node degree.

Another study from Puissant (2012) maps the Morphological Urban Areas and the urban fabrics from the urban areas of Strasbourg and Toulouse (France). The MUA is a delimitation of the city and is composed of Homogenous Urban Patches defined by Herold et al. (2003). Urban Fabrics are defined in a similar way as our UST. Their algorithm is based on a three level classification, first the urban blocks, then the urban materials and last the urban fabrics are classified. With this comes the disadvantage that if a classification error occures on the first level, the error propagates to the next levels. The overall accuracy is between 76 and 93 % depending on the method used.

The third study of Taubenböck et al. (2013) at delineating central business districts (CBDs) from their physical and morphological parameters. Their approch is based on three different steps. Firstly they formulate their hypothesis about how should be CBDs and test it with a statistical delineation, then they classify the entire Megacities and finally they make a cross-city spatial comparison of the CBDs using spatial metrics. They obtain an overall accuracy of 83 to 86% on the selection of city centers.

These studies are different but they all agree on the point that UST and mor-

phological and physical parameters are linked, which means that it should be possible to predict UST using the types of features considered in these papers. The data used in these studies are mainly imagery products produced from medium to very high resolution, LiDAR data, 3D city models, street networks data or urban cadastral data.

1.3 Goals of the study

The present study proposes a processing chain to predict urban structural classes from multiple data sources using support vector machine classification (SVM).

The main idea behind this method is that UST show geographical and morphological differences from type to type and that these differences can be expressed through multispectral image bands, spatial indices, morphological features extracted from the city model and thematic network indices. These indices can be quite numerous and that's why SVM are pretty well adapted for this purpose.

The second goal of this study is to compute Markov Random Fields (MRF) using the class Platt's probabilities extracted from the SVM to improve the classification using spatial priors. Using MRF help us in the classification accuracy, involving more spatial knowledge and encoding urban rules to achieve a better spatial consistency. It should enable to correct some classification mistakes which are not coherent with the prior knowledge.

1.4 Previous Work

During the previous semester, UST determination and a ground truth over the city of Munich (Germany) have been carried out. Since there is no standardized definition of the urban structural types in the academic literature(Banzhaf & Hofer, 2008), we used a mix of the values present in some studies. Eleven (plus one undefined classes) were retained, namely city center, residential high-medium-low density, industrial, commercial, leisure, agricultural, central business districts(CDBs), forest and water areas. (PoncetMontanges, 2014)

The defined UST are more detailed in the section 2.3.2.

2 Methodology

In section 2.1, we introduce the classification model used, SVM and the spatial post processing method used, MRF in section 2.2.

2.1 Support Vector Machine (SVM)

SVM is "one of the most successful modern machine learning algorithms".(Tuia, 2009)Thanks to the contributions of Boser et al. (1992), it is now considered as the "most studied" and the "most effective method for classification" and it is also easier to use than Neural Networks, which is another famous machine technique also use in many classification problems. (Chang & Lin, 2011)

2.1.1 What does a large margin classifier?

Support Vector Machine is a supervised classification algorithm whose goal is to find a linear decision function $f(\mathbf{x}) =< \mathbf{w}, \mathbf{x} > +b$ where $\mathbf{w}$ is a weight vector and $\mathbf{x}$ is a feature vector in $\mathbb{R}^d$, d is the number of bands and b is the bias. It is called large margin classifier because it maximizes the margins between the separated classes. This is equivalent to find the optimal hyperplane separating two classes.

2.1.2 SVM Formulation

Consider a training set composed of n samples $\{\mathbf{x}_i, y_i\}_{i=1}^n$, the margin maximization is equivalent to minimize the squared norm of the weight parameters $\mathbf{w}$. This can be formulated as

$$\min_{\mathbf{w}} ||\mathbf{w}||^2 \tag{1}$$
$$s.t. y_i(< \mathbf{w}, \mathbf{x}_i > +b) \geq 1, \forall i$$

which can be solved using Lagrange Mutlipliers and leads to the cost function:

$$L_D(\alpha) = \sum_{i=1}^n \alpha_i - 1/2 \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j < \mathbf{x}_i, \mathbf{x}_j > \tag{2}$$
$$s.t. \sum_{i=1}^h \alpha_i y_i$$
$$\alpha_i \geq 0$$

14

And the class prediction:

$$y_* = sign(\sum_{i=1}^{n} \alpha_i y_i < \mathbf{x}_i, \mathbf{z} > +b) \tag{3}$$

This is a cost function which can be solved by quadratic programming that provides a complex problem with a single minimum (i.e. SVM always converges to the same result). However, since data can be noisy, we might want to allow some flexibility to errors, if they allow a wider and more robust hyperplane: in this case we allow margin violations, that's why we need to include slack variables ξ_i, and the regularization term C to control the complexity of the model:

$$\min_{\mathbf{w},\xi} ||\mathbf{w}||^2 + C \sum_{i=1}^{n} \xi_i \tag{4}$$
$$s.t. y_i(< \mathbf{w}, \mathbf{x}_i > +b) \geq 1 - \xi_i, \forall i$$
$$\xi_i \geq 0$$

The term $||\mathbf{w}||^2$ wants to find the perfect separation to the training data while the right term allows some classification errors, with a penalisaton weighted by C. This is called the soft margin SVM formulation since we have a loss function and a regularizer term, and it is issued from the statistical learning theory.(Vapnik, 1998)

2.1.3   Kernel expansion

The extraordinary succes of SVM is due to the kernel expansion, also called kernel trick. Without it, SVM remains a robust linear classifier but cannot solve non linear problems. In reality we have rarely linear processes, that's where the kernel extension of the SVM comes into interest. The kernel trick enables to perform non linear classification through mapping the data into a higher dimensional feature space and to perform linear classification therein.

The more similar two data points are, the higher their dot product will be, it is then possible to build a kernel function that replaces the dot product and maps the points to a higher dimensional space where linear operations can be computed.

$$K : \chi \times \chi \rightarrow \mathbb{R}, < \mathbf{x}, \mathbf{z} > \rightarrow \mathbf{K}(\mathbf{x}, \mathbf{z}) \tag{5}$$

15

Where
$$\mathbf{K}(\mathbf{x}, \mathbf{z}) = <\phi(\mathbf{x}), \phi(\mathbf{z})> \tag{6}$$

is the kernel matrix.

As shown in Eq.(6), the results of a kernel evaluation is the dot product between the samples in the space spanned by $\phi()$. The beauty of it is that a kernel is computed using samples of the original spaces, i.e. without defining or computing explicitely the coordinates $\phi(\mathbf{x})$.

Here are a few examples of the most common kernels:

$$\text{Linear kernel}: K(\mathbf{x}, \mathbf{z}) = <\mathbf{x}, \mathbf{z}> \tag{7}$$
$$\text{Polynomial kernel}: K(\mathbf{x}, \mathbf{z}) = (<\mathbf{x}, \mathbf{z}> + 1)^p \tag{8}$$
$$\text{Gaussian kernel}: K(\mathbf{x}, \mathbf{z}) = exp(-\frac{||\mathbf{x}^2 - \mathbf{z}^2||}{2\sigma^2}) \tag{9}$$

As a requirement for the feature space where dot products are computed, all kernel functions must be symetric and positive definite.

We can then reformulate the SVM with kernel functions by substituing the kernel matrix to the dot product:

$$L_D(\alpha) = \sum_{i=1}^{n} \alpha_i - 1/2 \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_i) \tag{10}$$
$$s.t. \sum_{i=1}^{n} \alpha_i y_i$$
$$0 \le \alpha_i \le C$$

And the class prediction:

$$y_* = sign(\sum_{i=1}^{n} \alpha_i y_i K(\mathbf{x}_i, \mathbf{z}) + b) \tag{11}$$

2.1.4 Multi-class Classifier

SVM is a binary classifier, meaning that it separates only 2 classes. To have a solution for multiple classes, two types of SVM have been proposed. (Tuia, 2009) The first one is called one against all (o-a-a) which computes all the decision functions between one class and all the others. The second is called one against one (o-a-o) which computes $(k-1)k/2$ binary decision functions

for k classes. In the first case, the winning class is the one which is maximizing Eq.(11) on all the possible classes, in the second every binary classifier is evaluated and gives a vote to the winning class, the final decision is obtained by the class which gets the maximum number of votes. In this work, we use the one against one solution.

To fill this purposis, we used LibSVM which is the most common implementation of Support Vector Machine Classification and available on mutliple languages (C++, JAVA, Python, Matlab, Octave,...).[1]

2.1.5 Platt's probability estimates

Besides determining a class using SVM, it would also be interesting to know which classes could have been also possible, for example by assessing posterior probabilities $p(y_i = c|\mathbf{x}_i)$. Platt proposed a method to extract such probabilities from SVM. This method is mainly based on a Sigmoid fitting function.

2.2 Markov Random Fields (MRF)

Markov Random Fields modelling is widely used in many domains such as image processing because of its ability to describe and take into account the spatial information associated with an image, e.g. edge detection, classification, segmentation, etc.(Moser et al., 2013)It is especially interesting in the land cover classification domain where it can take into account the spatial distribution of classes and the patterns are not random, since a pixel belonging to a given class wil tend to be surrounded by others of the same class.

The Markov Random Fields modelling helps us to take into account the prior knowledge of urban patterns and land cover class distributions, which is not available in the SVM, since SVM is a non-contextual classifiera and considers regions as independant between each other. There is usually a strong spatial correlation between the classified regions (class distributions are not random), that's why it is interesting to use MRF modelling to favor a logical, i.e. smooth or driven by urban knowledge, class distribution pattern.(Moser & Serpico, 2010)

The main idea is to use the joint probability estimates of a sample (Platt's

---

[1]The LibSVM documentation and code are available freely at http://www.csie.ntu.edu.tw/ cjlin/libsvm/

probability estimates from SVM) with its prior knowledge from the class distribution and the spatial patterns to influence the labelling using the maximum a posteriori (MAP) estimation.(mar, 2009) The method considered is Iterated conditional Modes.

2.2.1 Iterated Conditional Modes (ICM)

To maximize the joint probability of a Markov Random Fields, we need to iterate for each class y_i and over each pixel $\mathbf{x}_i$

$$U(Y|X) = -\sum_i \ln(\frac{P(y_i|\mathbf{x}_i)}{p(y_i)}) + \beta \sum_i \sum_j 1 - \delta(y_i, y_j) \qquad (12)$$

- With $U(Y|X)$ which is the total MRF energy.

- The first term $-\sum_i \ln(\frac{P(y_i|\mathbf{x}_i)}{p(y_i)})$ is the class distribution related term where :

  - $P(\mathbf{x}_i|y_i)$ is the Platt probability estimate for pixel $\mathbf{x}_i$ estimated from the SVM to be classified a posteriori as y,
  - and $P(y_i)$ is the prior class occurence probability.

  If a class has more chances to occur than the others, $p(y_i)$ will be greater for this class. If there is not, the probability is usually equal for each class which means $p(y_i) = 1/k$ with k classes.

- The second term $\beta \sum_i \sum_j \delta(y_i, y_j)$ includes the neighborhood information for each pixel, i.e. the spatial context. It penalizes the non homogeneity of pixel classes.

  - β is a positive parameter
  - The $\delta(y_i, y_j)$ function is a dirac function called the Kronecker function, outputting 1 if the neighboring class y_j is the same class as the one considered y_i, and can be customized to influence more or less surrounding pixels and can also favors directional shapes.

3 Proposed Processing Chain

This section contains the proposed processing chain for building a classifier on multiple data sources and predict urban structural types.

3.1 General

Here is a plan of the main steps contained in the proposed processing chain :
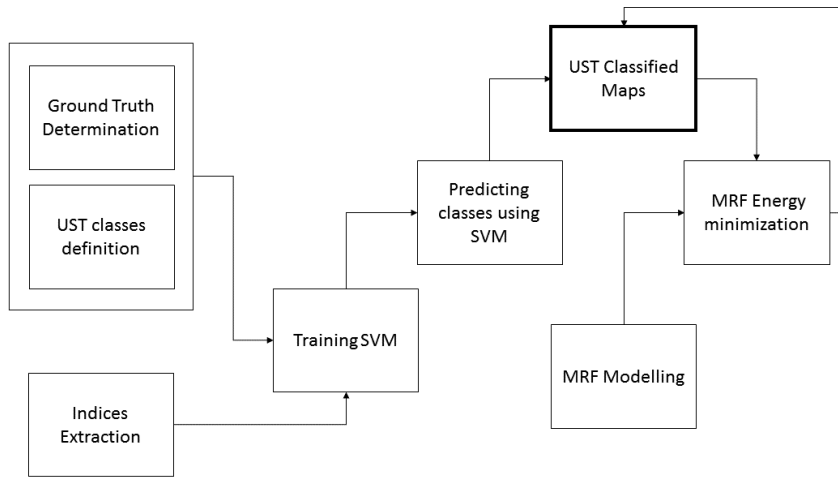

Figure 1:   Flow chart of the processing chain

3.2 Scale selection

The scale is a very important factor in the detection of urban structures, which is why it was necessary to choose a suitable mesh size, consistent with the size of buildings and blocks, without being too high to avoid being too general. In the case of the city of Munich, a mesh of 200x200m was chosen after a series of tests including the examples in Fig. 2.

Figure 2:   Grid scale selection

3.3   Urban structural classes definition

To establish the classes, I based myself on Walde et al. (2014), Taubenböck et al. (2013) and Banzhaf & Hofer (2008) and also taking into account a good illustration of Wickop (1998) who proposes to characterize the types of urban structures with their imperviousness, the degree of homogeneity or mixture in their structure, the building height. Some ratios were estimated, especially when they were not based on existing literature or not detailed enough.

Figures 3 and 4 summarize selected urban structural types, along with an example issued from the city of Munich. A more complete definition can be seen in Appendix A.

Miniature	Classe	Miniature	Classe
	1. City Center Multi-floors buildings contiguous in large blocks adjacent to roads and highly impervious area (possibly courtyard)		**5. Central Business District** Very large buildings close communication channels, often original forms
	2. Residential High Density Multi-floors buildings in blocks but with common areas (gardens / playgrounds)		**6. Industry** Large elongated buildings near highways or railway lines with substantial parking buildings
	3. Residential Medium Density Multi-floors buildings separated by gardens and vegetation.		**7. Commercial Area** Tall buildings of several floors adjacent to large parking areas and close to main streets and highways
	4. Residential Low Density Simple two-floors villas and less, spaced by gardens and vegetation zones		**8. Leisure Area** Public parks and gardens stages. Include only a few low-rise buildings except sports complexes.

Figure 3: Example of UST classes
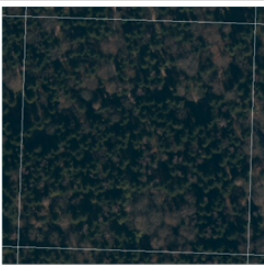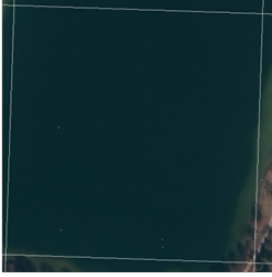
	9. Agricultural Land Very low presence of buildings, fields very present.		**10. Forest** Little or no buildings, strong presence of trees
	11. Water Little or no buildings, strong presence of water		**0. Not Attributed**

Figure 4: Example of UST classes (continued)

3.4 Ground truth determination

To map the UST on the city of Munich, we chose a large grid (70 by 50 regions of 200x200m each). On this large grid, we chose three squared patches of 400 regions each where we determined visually the UST and which serve as ground truth (which makes 1200 sampled regions) as shown in fig 5.
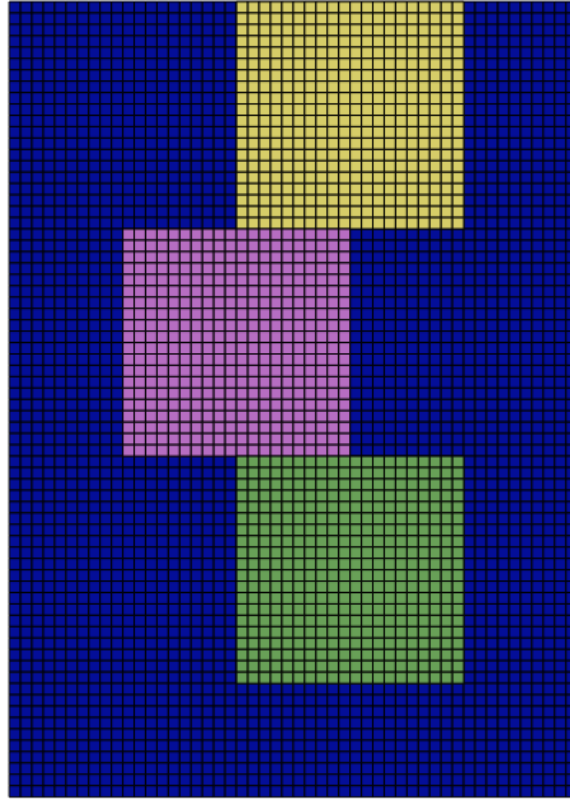
Figure 5:   General grid and sample patches

The UST defined above have been determined visually using the previously defined grids and superimposed on a satellite/aerial image of the city (see figure 6).
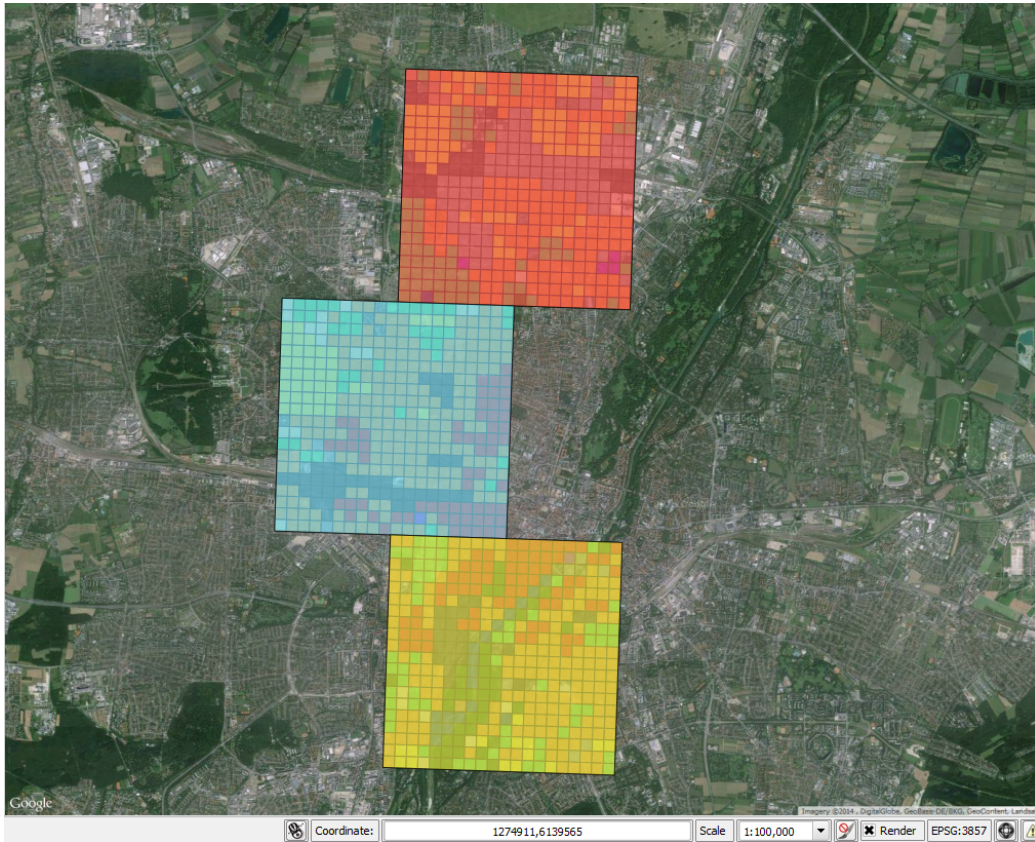
Figure 6:   Ground truth determination

To ease identification, we massively used tools such as Google Earth and Google Street View that allowed to disentangle areas where a vertical view only would have been analogous. (see figure 7)

Figure 7: Google earth disentangle

This determination remains subjective and is based on the recognition of similarities with the previously defined classes.

Since we are attributing labels on a regular grid, it is possible to come across areas where more than one urban structures are present. We must therefore choose the main UST by majority (> 50 % coverage) or, assign to class 0 if the distinction is not possible or too ambiguous.

The results of this ground truth determination are shown on Fig. 8
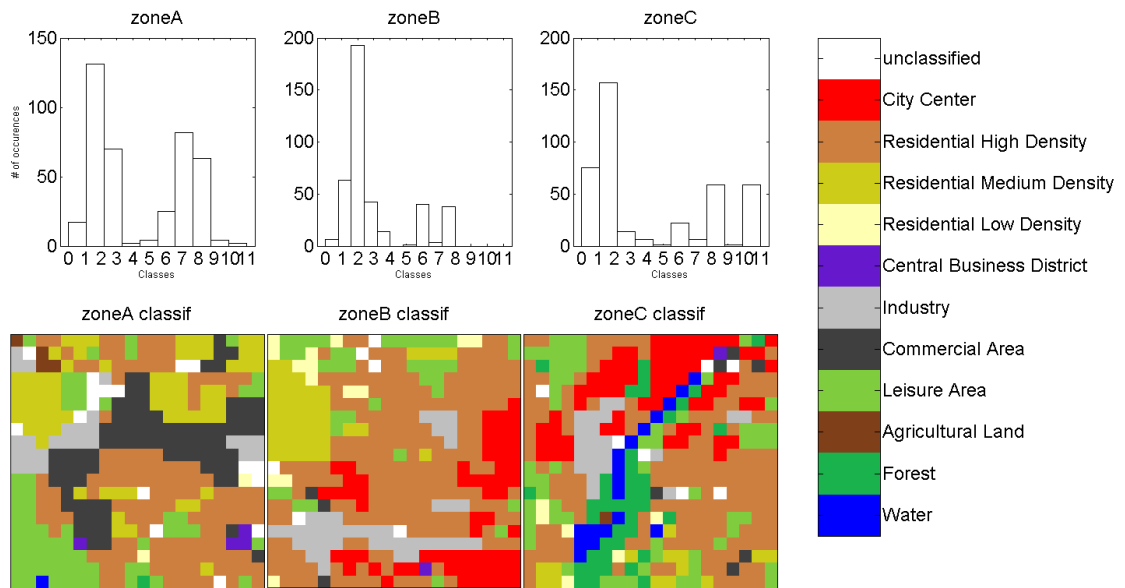
Figure 8: Ground truth grid samples

Since the class distribution was not regular in the sampled grids, we decided to mix these grids to create distinct training sets. We separate the samples to have 60% of each UST classe to the training set and 40% of each UST to the validation test set. We trained then ten classifiers using each time 70% of each UST class from the training set (from a performance analysis shown on Fig. 9), taking each time two training sets to predict the third one. We used an RBF kernel and a 5-fold cross validation to calibrate our model.

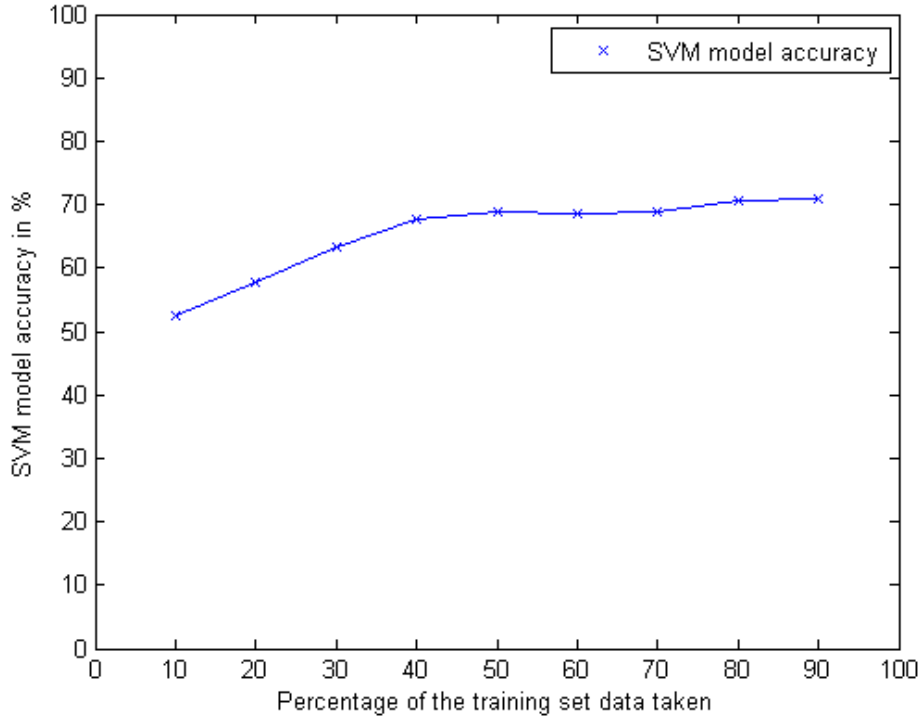Figure 9: Performance analysis

3.5 Computing land cover, spectral and spatial indices

Classifying Urban structural types requires physical and thematical information from the ground data. These data will be stored as indices and will be used as features for the SVM classifier training and for predictions. These indices are separated into two categories :

3.5.1 Spatial indices:

Spatial indices are calculated from the land cover data, from the 3D city model and from the thematic and spatial informations included in the Open Street Map vectorial layers. They are calculated on grids with a 1x1m pixel resolution. Since the three regions classified by SVM are 200x200m grids, they are then aggregated at a size of 200x200m grid pixels through averaging or other methods.

The spatial indices considered are:

27

The fractal index, which is an index based on the complexity of a shape. It is based on a perimeter/area ratio, which increases when the shape is complicated.

$$F = \frac{2\ln(0.25P)}{\ln(A)} \tag{13}$$

- Where P is the Perimeter

- and A is the Area

- and $1 \leq F \leq 2$ which means that the more complex the shape is, the greater F will be.

The Gravelius Index is an index similar to the fractal index. It is an index based on the complexity of a shape. It is based on a perimeter/area ratio but tends to be minimal if the shape is a disc.

$$G = \frac{P}{2\sqrt{\pi A}} \tag{14}$$

- Where $1 \leq G \leq \inf$

Both Gravelius and Fractal indices are used to describe the shape complexity of buildings and urban blocks which is related on the considered UST.

The Kernel Density Estimation (KDE) in geostatistics is a tool for estimating the probability density function of a random variable over space. It can be Gaussian:

$$\hat{\lambda}_k(\mathbf{x}) = \sum_{i=1}^{n} \frac{1}{\tau^2} k\left(\frac{\mathbf{x} - \mathbf{x_i}}{\tau}\right) \tag{15}$$

or Quartic:

$$\hat{\lambda}_k(\mathbf{x}) = \sum_{d_i \leq \tau} \frac{3}{\pi\tau^2}\left(1 - \frac{\mathbf{d_i^2}}{\tau^2}\right)^2 \tag{16}$$

or exponential, ... and it's main feature is its bandwidth which determines the search distance to other points. KDE provides information about the spatial neighborhood of the indices within each block. The bandwidth must be chosen accordingly to the radius of influence that we want to consider as shown in Fig. 10.
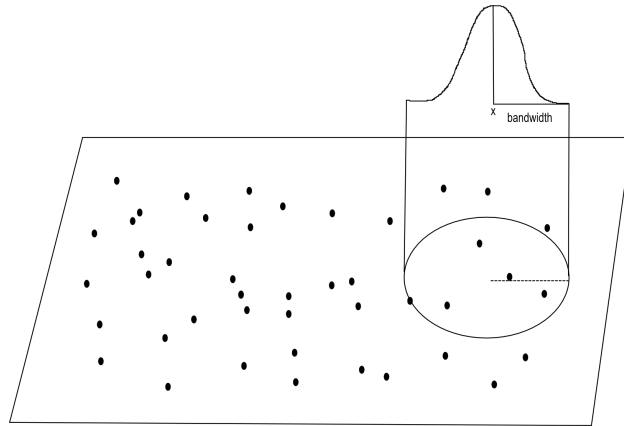
Figure 10: Kernel density estimation

7 class of land cover data at 1 meter resolution and a 3D city model provided by the Deutschen Zentrums für Luft- und Raumfahrt (DLR) were used to extract the following indices listed in Table 1 and 2.

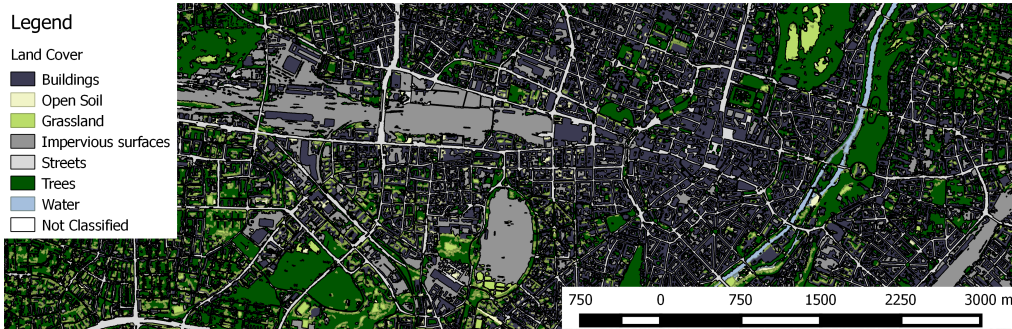| Indice Name	Definition
building coverage	% of buildings
grassland coverage	% of grassland
streets coverage	% of streets
trees coverage	% of trees
water coverage	% of water
impervious surface coverage	% impervious surfaces
opensoil coverage	% opensoil

Table 1: Spatial Indices issued from the land cover data.

Figure 11: Land cover classes

Indice Name	Definition
building area	mean area of buildings
building area max	max area of buildings
building area std	standard deviation of the area of buildings
building fractal	mean fractal indice calculated over the building shapes
building gravelius	mean gravelius indice calculated over the building shapes
building height	mean building height
building height max	max building height
building height std	standard deviation of the building height
building number	number of buildings
blocks gravelius	mean gravelius indice computed over the blocks separated from streets
blocks fractal	mean fractal indice computed over the blocks separated from the streets

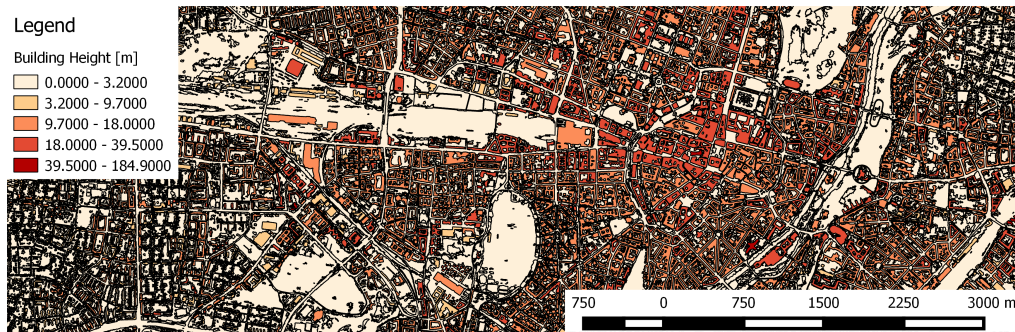Table 2: Spatial Indices issued from the 3D city model.

Figure 12: Building heights

On the Fig. 12, the building heights is shown. For the last ones, roads networks from the road land cover class were used to cut the map into urban blocks as shown in Fig. 13.
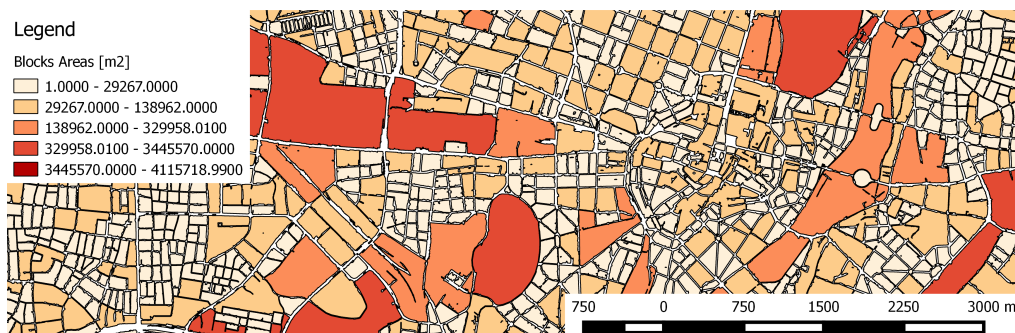

Figure 13: Urban blocks

Theses blocks could have also been a way to discretize the land use data (as in Walde et al. (2014) but their particular shapes and sizes were also interesting and could have contained mutliple land use classes, so we used the blocks shape and size as features instead.

A last set of spatial indices is extracted from the OpenStreetMap (OSM) Data

31

Indice Name	Definition
From the OpenStreetMap Data	mean values from kernel density estimation with 200m bandwith and gaussian kernel
points KDE gauss mean	all classes, all points, gaussian KDE
points KDE quart mean	all classes, all point, quartic KDE
points amenity KDE mean	class "amenity"
points highway KDE mean	class "highway"
points landuse KDE mean	class "landuse"
points place KDE mean	class "place"
points railway KDE mean	class "railway"
points tourism KDE mean	class "tourism"
points waterway KDE mean	class "waterway"
City center relation	City center proximity index computed with a kernel density estimation on one point and with 5km bandwith
city center KDE gauss mean	one point, gaussian
city center KDE quart mean	one point, quartic

Table 3: Spatial Indices Table 03.

3.5.2 Spectral indices:

Spectral indices were extracted from the Landsat 8 TM bands of two images, one acquired during summer and the other during winter. Since a pixel of Landsat 8 has a resolution of 30m to 60m and the grid chosen a resolution of 200m, multiple aggregation ways were used as indices : Mimimum, Maximum, Mean and Standard Deviation.

The two thermal bands (10 and 11 with a ground resolution of 100m)from Landsat 8 were used as well since they are relevant indicators of the urban footprint and used also in the heat islands detection (see Kardinal Jusuf et al. (2007)). Fig.14 and 15 are 2 rgb pan-sharpened compositions of landsat images from summer and winter. These 2 compositions were not used as indices but for illustration purposes.
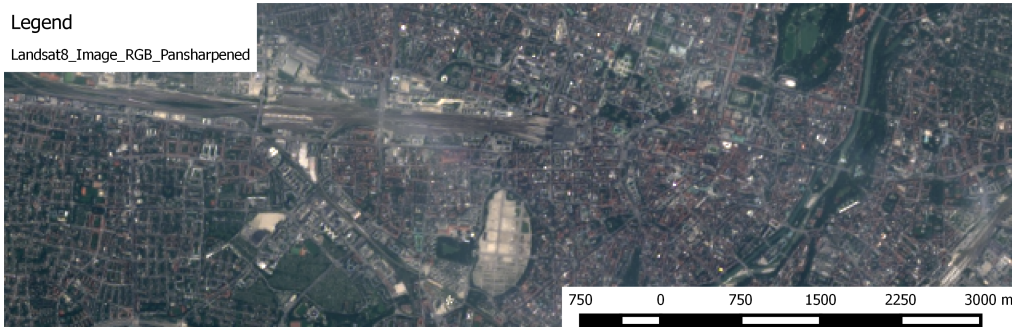
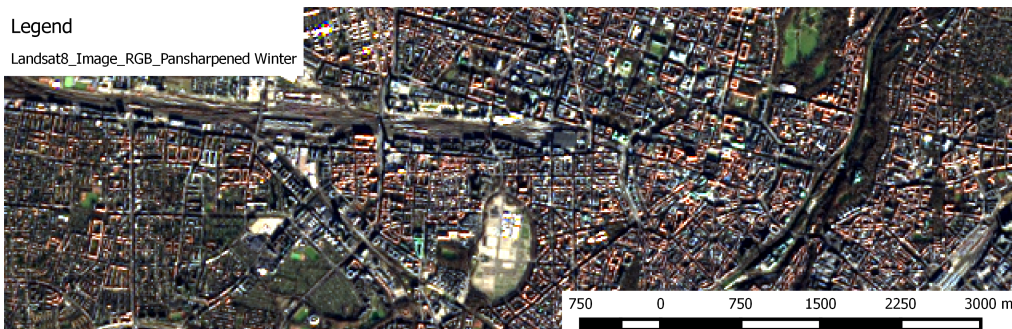Figure 14: Landsat8 pan-sharpened RGB composition Summer

Figure 15: Landsat8 pan-sharpened RGB composition Winter

3.6 Priors included in the MRF

Here, the priors included in the MRF are presented.

3.6.1 MRF1 : the smoothness prior

As we said before, the UST are not ditributed randomly in space. There is a spatial correlation of The UST and therefore a region with a UST has a higher probability to have neighbors with the same UST.

We design a 3x3 moving window algorithm which computes the Energy for each region. This energy is lower when neighboring classes are the same and higher whenever they are different. Each time, the UST minimizing the energy was chosen and we iterate several times on the whole grid as shown in Fig. 16.
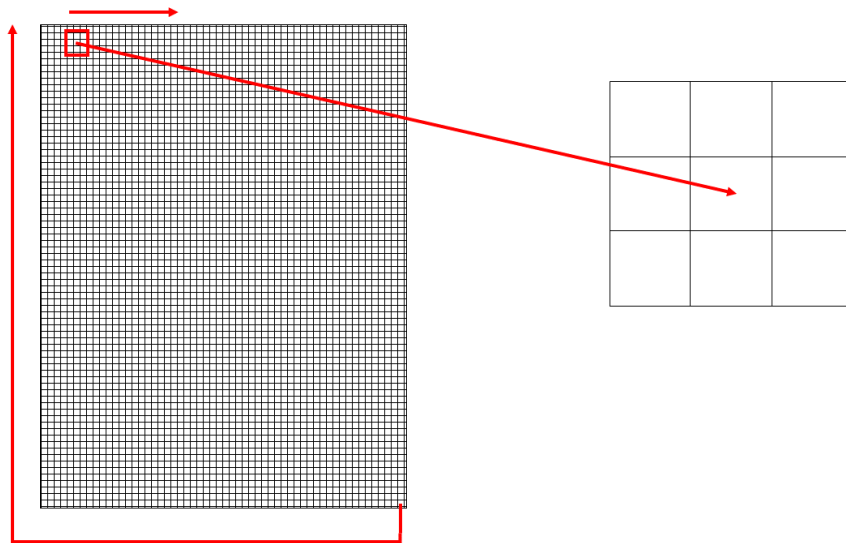
Figure 16:   3x3 moving window

3.6.2   MRF2 : the continuity prior

The second pattern that we observed was the flow continuity of water. To ensure that, we design an algorithm that was promoting the continuity but only for the water class.

We reuse the moving window but this time promoting only the water class and the presence of two neighbors to enforce continuity (one for input flow, one for output flow).

3.6.3   MRF3 : the centrality prior

The last rule that we implemented was linked to the spatial centrality of certain UST, e.g. the city center, and the non spatial centrality of others, e.g. the low density residential.

Using the previously calculated KDE city center index, with a bandwidth of 5km and a central point on the Marienplatz (historical city center of Munich), we promote the city center UST for regions with a high index value and penalize them for a low index value. We do the contrary for low density residential areas.

4 Results

This section present the main results. The discriminative nature of the features are presented in Section 4.1 The accuracies obtained in section 4.2, the classification maps in section 4.3 and a sensibility Analysis of the MRF in section 4.4

4.1 Discriminative nature of features

A different response is observed depending on the chosen class and according to the index considered. All indices were normalized to be compared. City center (Fig. 17) and industrial areas (Fig. 18) are very similar from the point of view of indices impervious surfaces and vegetation ratio as might be expected.
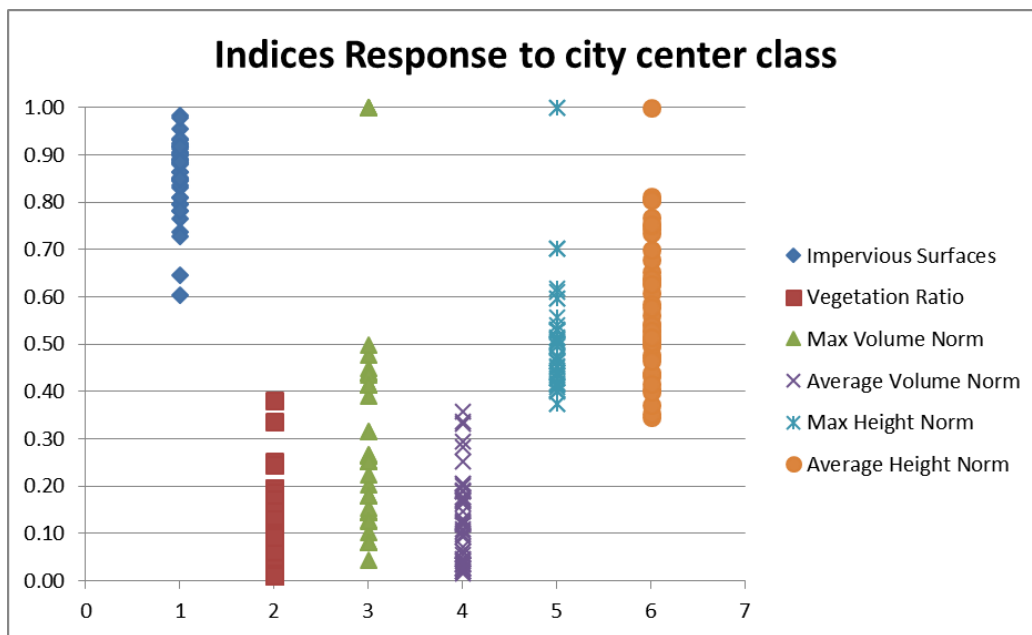

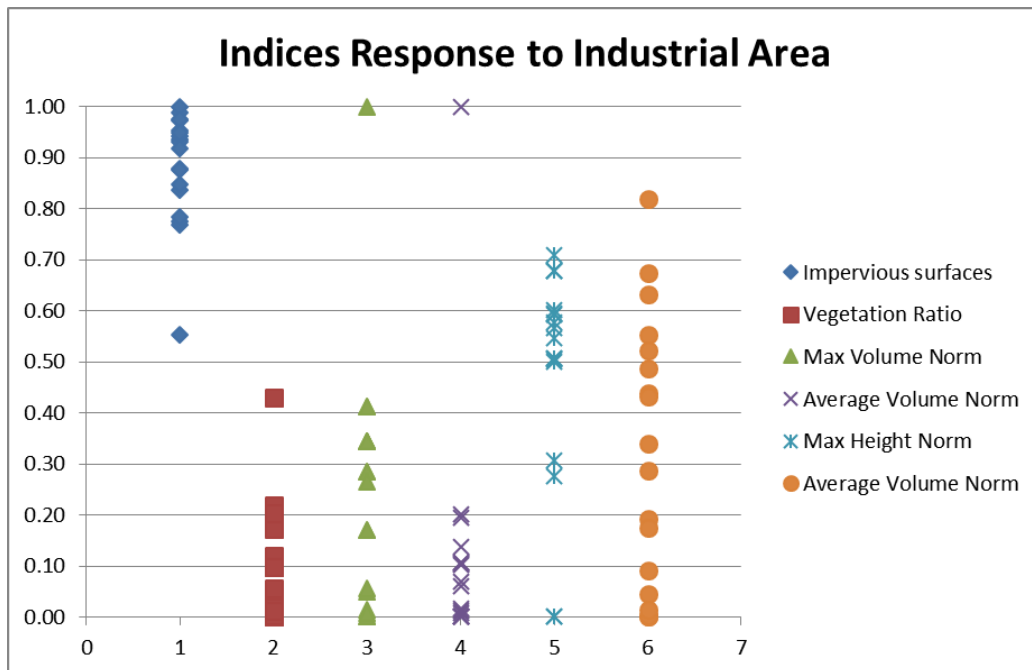Figure 17: City center response to selected indices

Figure 18: Industrial area response to selected indices

They also show a different response to other indices, which is a good thing because this means that they are not dependent. In fact, the more the classes differ in their responses to the indices, the easier it will be then to differentiate those thereafter.

4.2 Accuracies

Method	OA		Kappa	
	mean	std	mean	std
SVM	0.69	0.01	0.59	0.01
MRF 1	0.68	0.01	0.58	0.02
MRF 2	-	-	-	-
MRF 3	0.66	0.01	0.55	0.01

Table 4: Accuracy measurements, overall accuracy in % and Kappa coefficient

As shown in Table 4, the overall accuracy of the classification is about 70% which is not really high for using a SVM. SVM accuracies are usually higher

than 80%. This is due to the fact that the classification problem is pretty complex. Involving twelve classes on spatially different regions. The samples for certain UST are not numerous enough to assess a high accuracy classification of these UST. Meanwhile, this accuracy is already sufficient to recognize most of the UST distribution on the city. The Kappa coefficient is about 0.6 which means that the model slightly agree with the reference values.

Markov Random Fields overall accuracies are just lower than 70%. Theses accuracies are just a bit lower than the predicted ones from the SVM but their spatial coherency is higher and offers a better visual representation (see the classification maps in Section 4.3. The Kappa coefficient is a bit lower than 0.6 which means that the model slightky agree with the reference values.

The accuracy of the continuity prior (MRF2) could not be calculated since the water class was not available enough on the classified map. The classified map is in appendix B

4.3   Classification Maps

On the following pages, you will find the UST classified maps from Munich. The Fig. 19 is one of the SVM outputs. The Fig. 20 is the output of the MRF model with the smoothness prior and the Fig. 21 is the output of the MRF model with the centrality prior.

A transparent version of this classification result overlayed on a landsat 8 rgb composition has been added in appendix C.
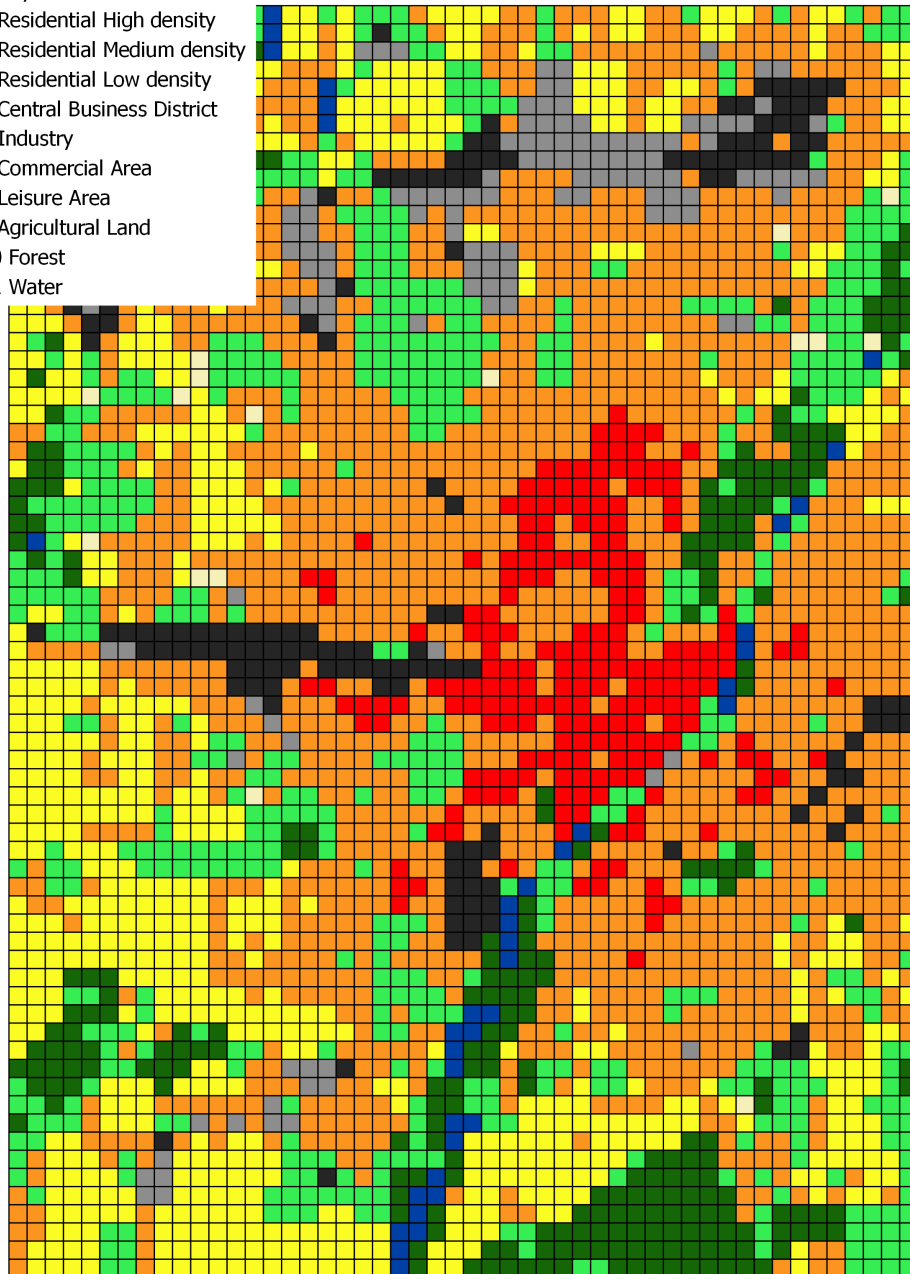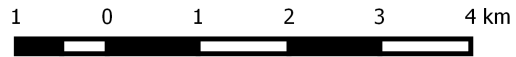
Figure 19: SVM classified map using the first classifier

38

Legend

MRF1 Beta=0.5

- ☐ 0 Non Attributed
- 🟥 1 City Center
- 🟧 2 Residential High density
- 🟨 3 Residential Medium density
- ◻ 4 Residential Low density
- 🟪 5 Central Business District
- ⬛ 6 Industry
- ◼ 7 Commercial Area
- 🟩 8 Leisure Area
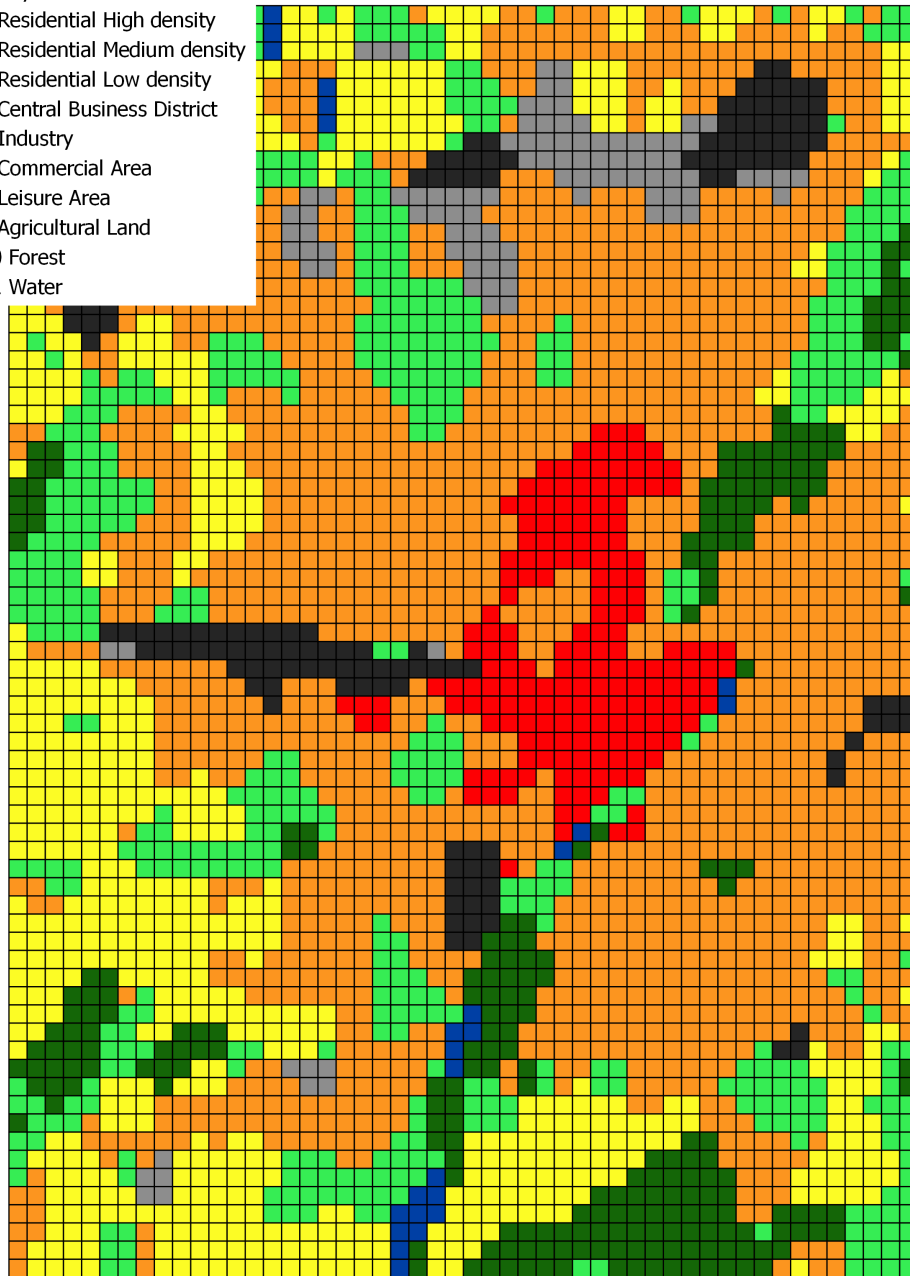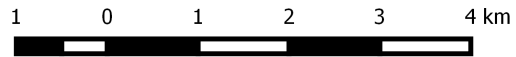- 🟫 9 Agricultural Land
- 🟩 10 Forest
- 🟦 11 Water

Figure 20: MRF using the smoothness prior

39
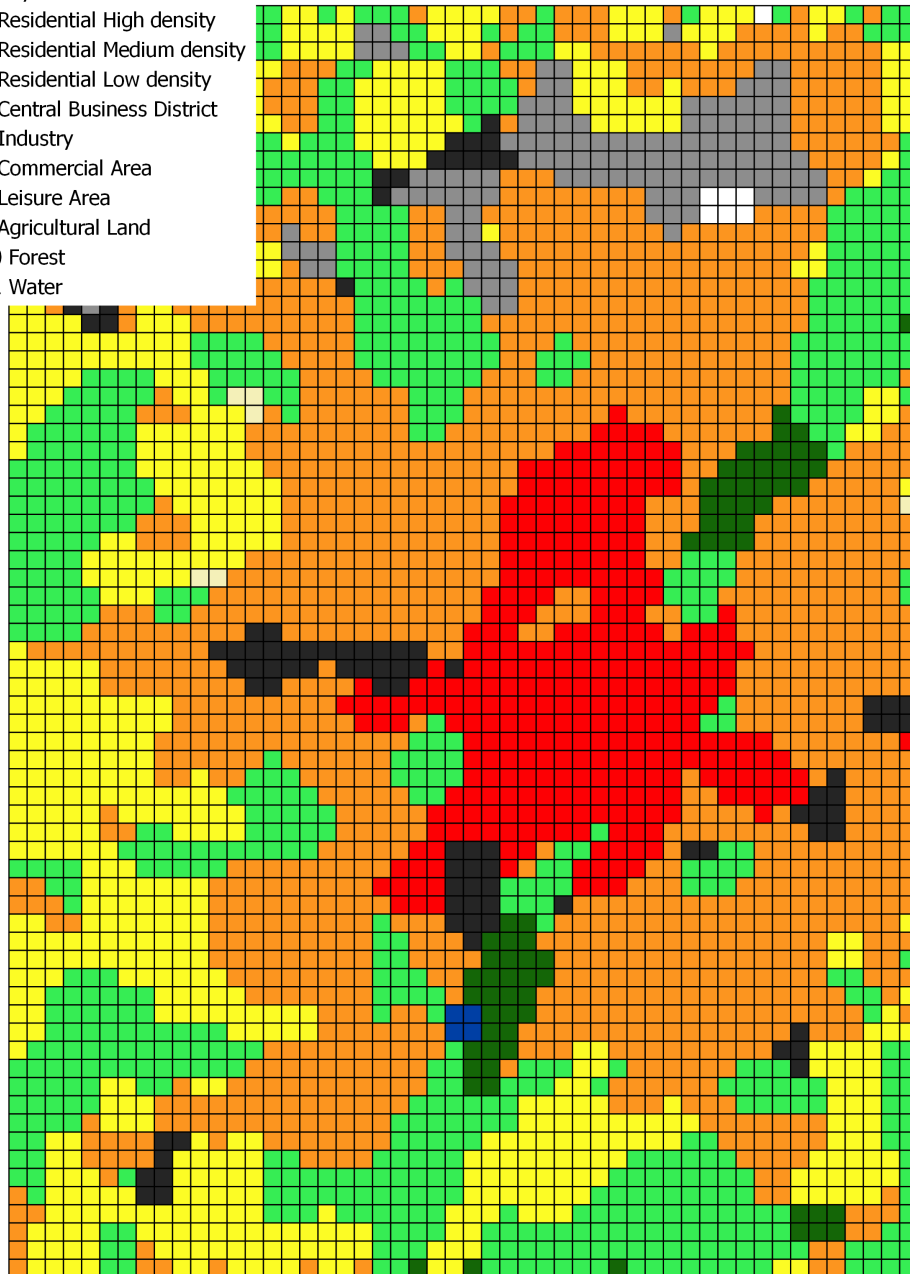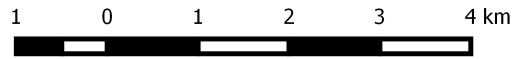
Figure 21: MRF using the smoothing and centrality prior

4.4   Sensibility Analysis
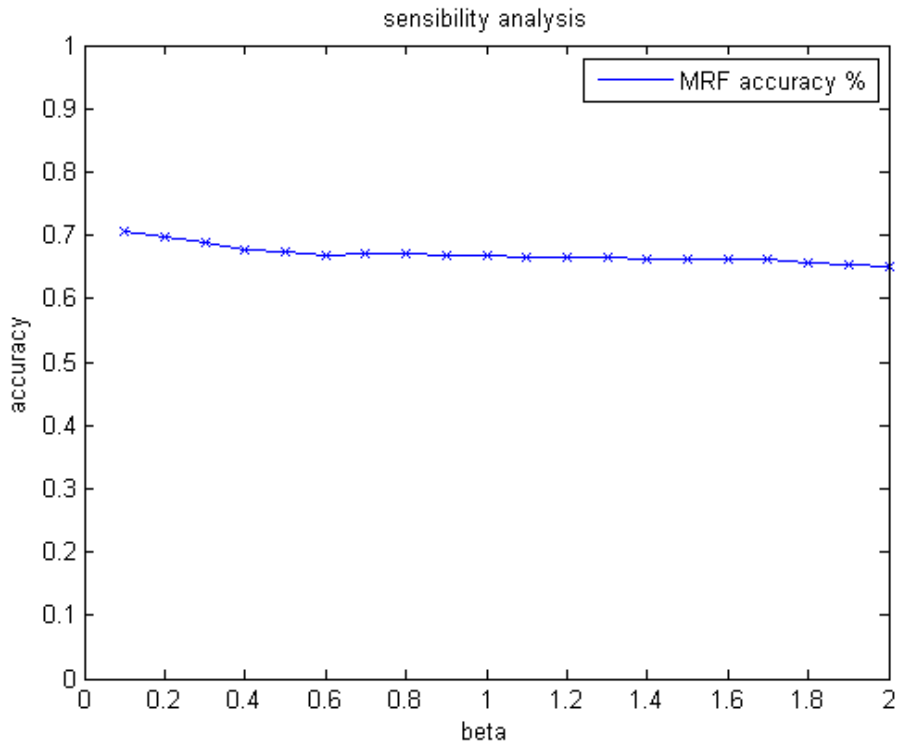
Figure 22:   Sensibility analysis

Increasing beta increases the spatial smoothing of the MRF and makes the overall accuracy lightly decrease (as shown in Fig. 22. It is visible that it has just a small influence on the Kappa which lightly decreases with increasing beta. When beta is equal to zero, there is no spatial smoothing and if the class occurence probability is the same for each class, it means that nothing changes. Too much increase beta does not bring much since the spatial smoothing is rapidly maximum which means that increasing the beta beyond this value will not bring anything to the classification.

5 Conclusion

The goal of this work is to map Urban Structural Types on the city of Munich (Germany). To achieve this goal, we firstly define eleven UST using litterature sources. To classify the whole map, we decide to use Support Vector Machine (SVM) Classification. On the city of Munich, we map a grid of 3500 (70x50) squared regions of 200m side. On this grid, we choose 3 patches of 400 regions each, to be determined as a ground truth by visual interpretation using available satellite and aerial imagery. From this ground truth, we used 60% of it as training set and 40% as validation set for the SVM.

To describe UST, we use several indices, computed from the land cover data, from the 3D city model and from OSM data (e.g. land cover class coverage, building heights, areas or numbers, urban blocks shape indices or even kernel density estimation of points of interest from OSM. Theses indices are then aggregated to the mesh size of 200x200m to be used as SVM features for the UST classification.

After applying SVM, we used MRF modelling to increase the spatial consistency of our results and apply some urban knowledge rules as spatial priors. We define a spatial smoothness prior, a continuity prior and a centrality prior on our data using iterated conditional modes, which is a MRF model that iterates over the whole grid to minimize the energy.

Finally we obtain a classified map of UST on 200m squared regions which is predicted with an overall accuracy of 70%. The model slightly agree with the reference values provided from the validation test with a Kappa of 0.6. The different MRF models are not improving the overall accuracy but the spatial coherency has been improved. The overall accuracy is not as high as expected from SVM but the result seems to be sufficient to detect the spatial distribution of UST which is what we want to achieve.

This opens possibilities to describe the urban footprint of the city, to detect the key areas for urban planification and to better understand the city dynamics. An adjacency analysis could be carried out to detect potential conflicts between UST, e.g. industrial zones near to residential or leisure areas. This proposed processing chain could be carried out on mutltiple cities for comparison and automatised to extract in an easier way these UST.

References

(2009). *Markov Random Field Modeling in Image Analysis*. Advances in Pattern Recognition. London: Springer London. URL: `http://link.springer.com/10.1007/978-1-84800-279-1`.

Banzhaf, E., & Hofer, R. (2008). Monitoring urban structure types as spatial indicators with CIR aerial photographs for a more effective urban environmental management. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *1*, 129–138. URL: `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4652571`. doi:10.1109/JSTARS.2008.2003310.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory* (p. 144–152). ACM. URL: `http://dl.acm.org/citation.cfm?id=130401`.

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*, 1–27. URL: `http://dl.acm.org/citation.cfm?doid=1961189.1961199`. doi:10.1145/1961189.1961199.

Herold, M., Liu, X., & Clarke, K. C. (2003). Spatial metrics and image texture for mapping urban land use. *Photogrammetric Engineering and Remote Sensing*, *69*, 991–1002. URL: `http://www.asprs.org/a/publications/pers/2003journal/september/2003_sep_991-1001.pdf`.

Kardinal Jusuf, S., Wong, N., Hagen, E., Anggoro, R., & Hong, Y. (2007). The influence of land use on the urban heat island in singapore. *Habitat International*, *31*, 232–242. URL: `http://linkinghub.elsevier.com/retrieve/pii/S0197397507000148`. doi:10.1016/j.habitatint.2007.02.006.

Moser, G., & Serpico, S. B. (2010). Contextual remote-sensing image classification by support vector machines and markov random fields. In *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International* (p. 3728–3731). IEEE. URL: `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5651182`.

Moser, G., Serpico, S. B., & Benediktsson, J. A. (2013). Land-cover mapping by markov modeling of Spatial–Contextual information in

very-high-resolution remote sensing images. *Proceedings of the IEEE*, *101*, 631–651. URL: `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6304904`. doi:10.1109/JPROC.2012.2211551.

Pauleit, S., & Duhme, F. (2000). Assessing the environmental performance of land cover types for urban planning. *Landscape and urban planning*, *52*, 1–20. URL: `http://www.sciencedirect.com/science/article/pii/S0169204600001092`.

PoncetMontanges, A. (2014). *SIE Project*. Technical Report. URL: `https://documents.epfl.ch/users/p/po/poncet/www/Project_SIE_ArnaudPoncetMontanges.pdf`.

Puissant (2012). URBAN MORPHOLOGY ANALYSIS BY HIGH AND VERY HIGH SPATIAL RESOLUTION REMOTE SENSING. (p. 524). Rio de Janeiro - Brazil. URL: `http://mtc-m18.sid.inpe.br/col/sid.inpe.br/mtc-m18/2012/05.18.17.44/doc/138.pdf?metadatarepository=&mirror=urlib.net/www/2011/03.29.20.55`.

Sirmacek, B., Taubenbock, H., Reinartz, P., & Ehlers, M. (2012). Performance evaluation for 3-d city model generation of six different DSMs from air- and spaceborne sensors. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *5*, 59–70. URL: `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6122037`. doi:10.1109/JSTARS.2011.2178399.

Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A., & Dech, S. (2012). Monitoring urbanization in mega cities from space. *Remote Sensing of Environment*, *117*, 162–176. URL: `http://linkinghub.elsevier.com/retrieve/pii/S0034425711003427`. doi:10.1016/j.rse.2011.09.015.

Taubenböck, H., Esch, T., & Roth, A. (2006). An urban classification approach based on an object-oriented analysis of high resolution satellite imagery for a spatial structuring within urban areas. In *First Workshop of the EARSeL Special Interest Group on Urban Remote Sensing" Challenges and Solutions*. URL: `http://elib.dlr.de/22591/1/Session7_Taubenboeck.pdf`.

Taubenböck, H., Klotz, M., Wurm, M., Schmieder, J., Wagner, B., Wooster, M., Esch, T., & Dech, S. (2013). Delineation of central business districts in mega city regions using remotely sensed data. *Remote Sensing of*

Environment, *136*, 386–401. URL: `http://linkinghub.elsevier.com/retrieve/pii/S0034425713001739`. doi:10.1016/j.rse.2013.05.019.

Tuia, D. (2009). *Advanced kernel methods for remote sensing image classification*. Ph.D. thesis Université de Lausanne Lausanne.

Vapnik, V. (1998). *Statistical Learning Theory*. New York: Wiley. URL: `http://read.pudn.com/downloads161/ebook/733192/Statistical-Learning-Theory.pdf`.

Walde, I., Hese, S., Berger, C., & Schmullius, C. (2014). From land cover-graphs to urban structure types. *International Journal of Geographical Information Science*, (pp. 1–26). URL: `http://www.tandfonline.com/doi/abs/10.1080/13658816.2013.865189`. doi:10.1080/13658816.2013.865189.

Wickop, E. (1998). *Environmental Quality Targets for Urban Structural Units in Leipzig with a View to Sustainable Urban Development*. Springer Berlin Heidelberg.

A The UST classes

1 - City Center

This area is defined by multi-floors buildings (at least 3 floors) contiguous wholesale separate blocks (with the possibility of an inner courtyard), adjacent to roads and located in a highly impermeable zone (70-90% degree of imperviousness).

2 - Residential High Density

This area contains multi-floors buildings (at least 3 floors) to rectangular shapes, with regular alignment and common areas (gardens / playgrounds), near roads and located in a moderately impermeable zone (25-70% of degree of imperviousness) (25-60% degree of green spaces).

3 - Residential Medium Density

This area includes 2 floors houses on average, spaced from each other by gardens and vegetation. There is a presence of small buildings (garages). This area is slightly impermeable (10-40% degree of imperviousness) and moderately vegetated (40-70% degree of green spaces).

4 - Residential Low Density

This area consists of simple two-floors villas and less, spaced by gardens and vegetation zones. It is a very low impervious area (0-20% imperviousness) and highly vegetated (50-90% degree of green spaces).

5 - Central Business District

Very tall buildings (more than ten floors) close main roads and highways, composed often with original forms contrasting from the urban texture. They are located in very impermeable zones (70-90% degree of imperviousness).

6 - Industry

Large multi-floors buildings (at least 2 floors) with a rectangular shape elongated and very near from highways or railway lines with substantial parking. Presence of structures such as tanks, drop zones, equipment etc. These areas are located in highly impermeable zones (70-90% degree of imperviousness).

7 - Commercial Area

Large multi-floors buildings (at least 2 floors), with rectangular shapes, adjacent to large areas of parking and close to transport links, situated in a highly impermeable zone (70-90% degree of impermeability).

8 - Leisure Area

Parks and public gardens, schools, stadiums, etc. . Include only few buildings (sports complexes, for example). Located in low permeability zones (10-40% degree of imperviousness) except for car parks and sports complexes, generally well vegetated (40-80% degree of green spaces).

9 - Agricultural Land

Very low presence of buildings (only farms and isolated dwellings); there is a presence of specific agricultural structures (fields). Areas weakly impervious or permeable (0-20% imperviousness).

10 - Forest

Little or no buildings, high forest cover, weakly impervious or permeable (0-20% imperviousness), highly vegetated (70-100% degree of green spaces).

11 - Water

Little or no buildings, strong presence of water (water bodies, small lakes, rivers, ...)

0 - Not Attributed

It was not possible to determine the urban structural type for that cell. Happens when no UST is predominant on a region.

The **degree of imperviousness** is defined as the ratio between total surface and impermeable surface as well as the degree of green areas is the ratio between the vegetation covered surface and the total surface.

$$I = \frac{S_{imp}}{S_{tot}} \tag{17}$$

B MRF 2 Continuity Prior
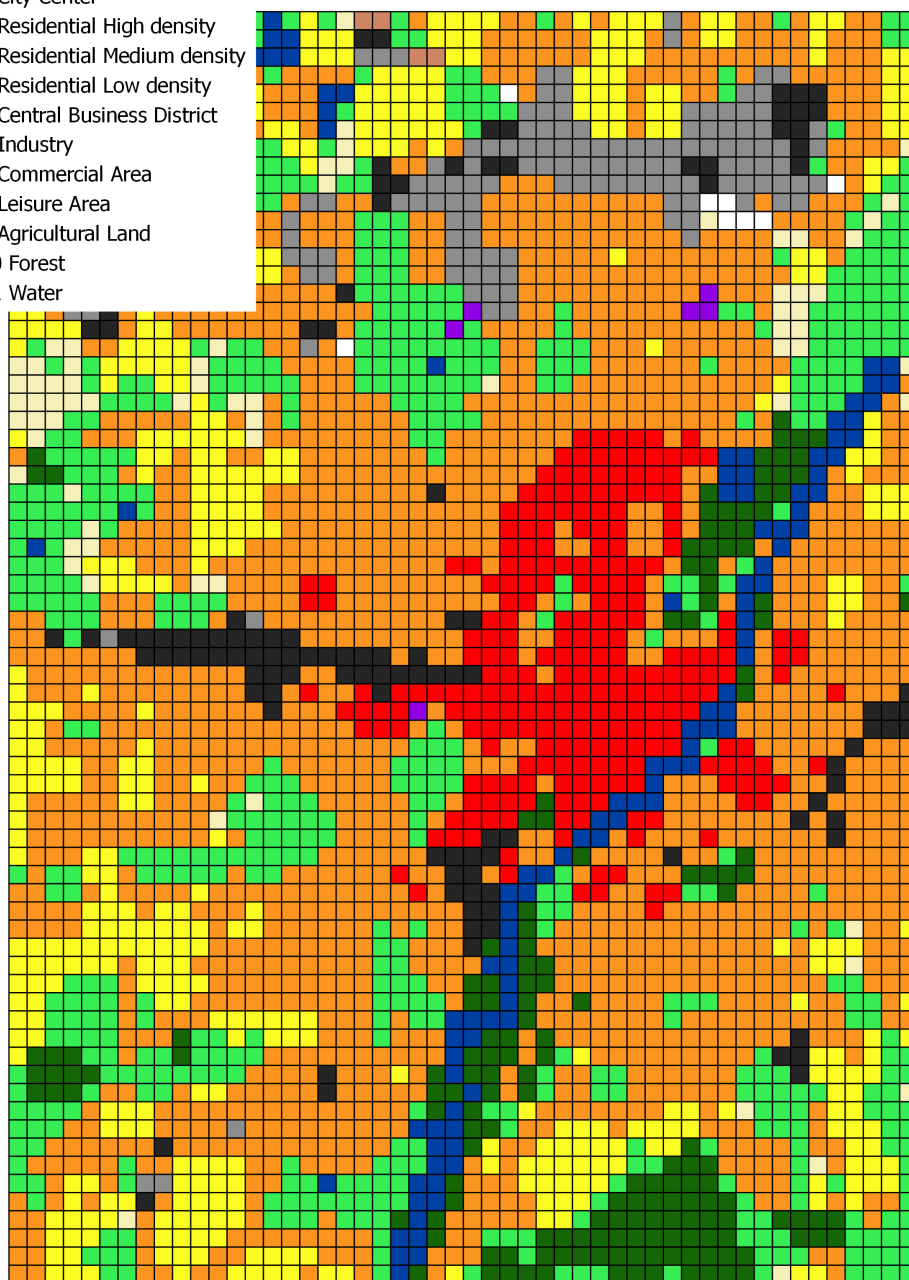
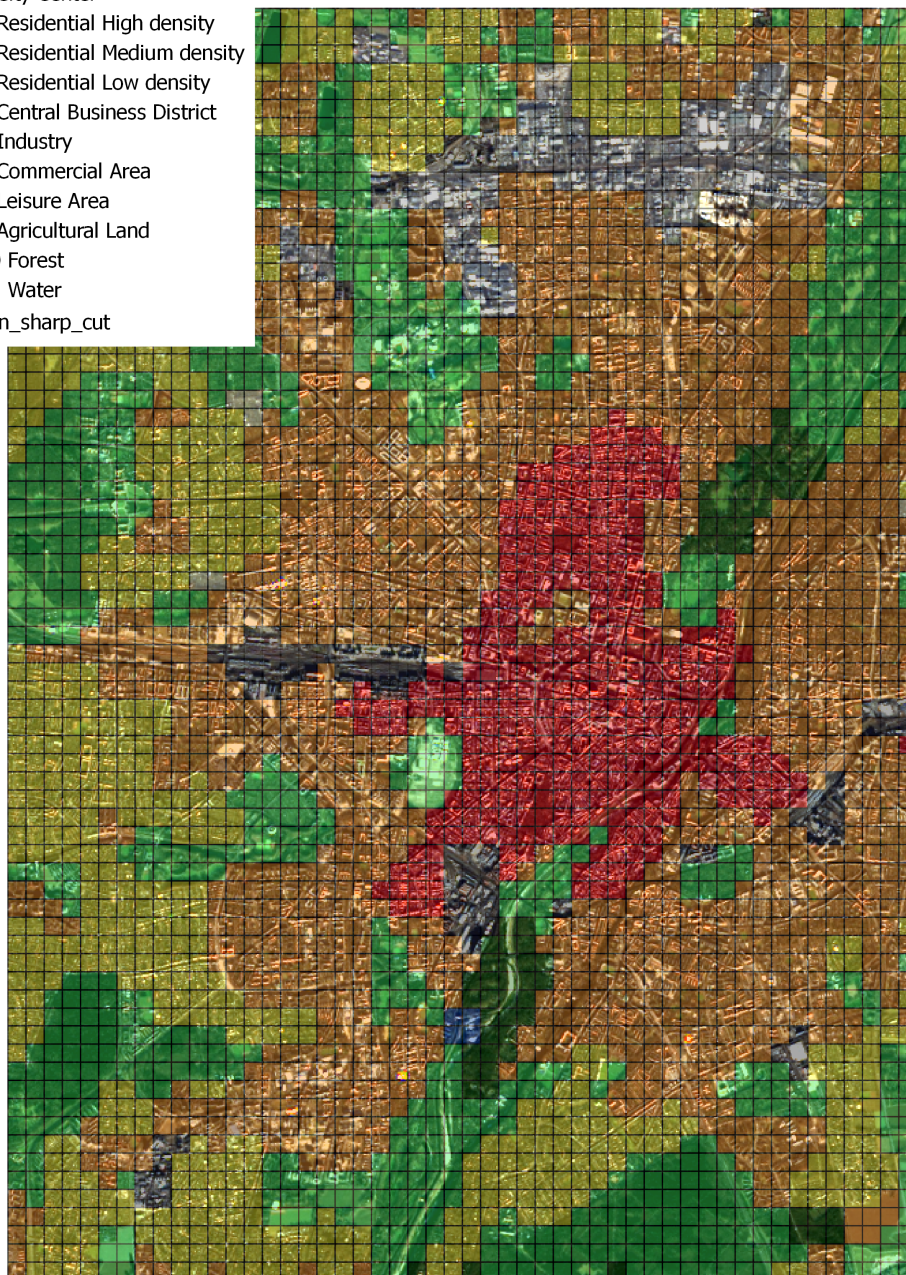Figure 23: MRF using the continuity prior

C   MRF 3 Transparent

Figure 24:   MRF using the centrality prior with transparent Landsat 8 image

D   Green tea

Merci à Gilian Milani pour son "Non, c'est hyper bien!" en réponse à la question "C'est hyperspectral?" de Devis Tuia.

Merci à Solange Duruz pour son fameux gigot d'agneau.

Merci à Devis Tuia pour ses cerises de Savièse.

Merci à Estelle Rochat, Timothee Produit, Stephane Joost, Kevin Leempoel, Emanuele Strano, Marc Soutter, François Golay, Véronique Boillat, Ivo, Jessy, et tous les autres pour ces bons moments passés au LASIG.

Merci à Rama Chellappa pour son "Data don't generalise, models do!"