Norsk Regnesentral
NORWEGIAN COMPUTING CENTER

Note

Urban green structure

State of the art of classification methodology

Note no SAMBA/03/10

Authors Øivind Due Trier

Date 28 January 2010

Contract Subcontract of
Norwegian Space Centre contract no. JOP.10.08.2

Cover photo

Orthophoto made from aerial photography of the Follo district south of Oslo, on 13 May 2008. Ground resolution 10 cm. This small excerpt is from Oppegård municipality. Image owner: Geovekst.

Norsk Regnesentral

Norsk Regnesentral (Norwegian Computing Center, NR) is a private, independent, non-profit foundation established in 1952. NR carries out contract research and development projects in the areas of information and communication technology and applied statistical modeling. The clients are a broad range of industrial, commercial and public service organizations in the national as well as the international market. Our scientific and technical capabilities are further developed in co-operation with The Research Council of Norway and key customers. The results of our projects may take the form of reports, software, prototypes, and short courses. A proof of the confidence and appreciation our clients have for us is given by the fact that most of our new contracts are signed with previous customers.

Title	**Urban green structure – state of the art of classification methodology**
Authors	**Øivind Due Trier**
Quality assurance	Rune Solberg
Date	28 January 2010
Year	2010
Publication number	SAMBA/03/10

Abstract

The management of urban green structure has received much attention lately in the research community. The present project was initiated to meet the need of municipalities in Norway to develop a green structure plan. Traditional mapping has its limitations, since the land use is in focus and not the actual land cover. Therefore, other sources of information about urban and suburban green structure are being sought.

The purpose of this study is to discuss the current state of the art of classification methodology for high resolution remote sensing imagery in the context of urban green structure, and to suggest possible improvements of an existing classification algorithm, implemented in Definiens Developer. In this report, several methods are presented, with no concern of whether they can be implemented with Definiens Developer or not. Experiments are needed in order to get an indication on which of the approaches can be used.

The scope is very wide, and with no focus on budget limitations. Obviously, the cost and ambition of an improved automatic green structure classification and measurement system must be taken into account when selecting from the different methods and alternatives presented.

Keywords	Quickbird, green corridors, urban planning
Target group	Remote sensing researchers, urban planning researchers
Availability	Open
Project number	220 428
Research field	Satellite remote sensing
Number of pages	34

Contents

Excecutive summary

The management of urban green structure has received much attention lately in the research community. The present project was initiated to meet the need of municipalities in Norway to develop a green structure plan. Traditional mapping has its limitations, since the land use is in focus and not the actual land cover. Therefore, other sources of information about urban and suburban green structure are being sought.

Geodatasenteret AS, Arendal, Norway, has a contract with the Norwegian Space Centre to develop a prototype for a service for monitoring of urban green structure, with the Norwegian Computing Center as a subcontractor. In phase one of the project, Geodatasenteret developed an automatic classification method, implemented in Definiens Developer. The Norwegian Computing Center performed an evaluation of the detection result.

The purpose of this study is to discuss the current state of the art of classification methodology for high resolution remote sensing imagery in the context of urban green structure, and to suggest possible improvements of the existing classification algorithm, implemented in Definiens Developer. In this report, several methods are presented, with no concern regarding whether they can be implemented in Definiens Developer or not. Experiments are needed in order to get an indication on which of the approaches can be used.

The choice of methods is to a large extent influenced by what remote sensing and other data is available. Clearly, in order to observe changes over time, repeated acquisitions are important, preferably by the same sensor. If different sensors are being used, then, in general, a new automatic classification algorithm has to be developed. At best, an existing algorithm can be re-trained on data from the new sensor.

However, there is no guarantee that an algorithm that was trained on one single acquisition from one sensor will perform well on another image acquisition from the same sensor. Imaging conditions may have changed, due to varying atmospheric conditions, including varying presence of aerosols, haze; sun elevation, and time of the day. Also, seasonal variations within a year and between years, often referred to as phenological variation, alter the spectral and textural properties of vegetation. Also, the time since rainfall, and temperature and humidity variations adds to the variability. It is therefore important to use methods that compensate for these variations in one way or another, and to include training data that reflect the typical variations that one may encounter in the operational setting.

Another problem is the presence of clouds, which obscures parts of the scene. This means that multiple image acquisitions may be necessary to ensure full ground coverage.

Phase one of the project clearly demonstrated that Quickbird imagery is a relevant source of information for urban green structure assessment, and provided repeated acquisitions are made, one could expect to be able to monitor and quantify changes. However, shadows from trees and buildings create problems for the segmentation and classification methods. If lidar data is available, the location and extent of these shadows may be predicted, and corrected for. Cloud shadows are much larger, and they may be detected as separate classes. If multiple image acquisitions are available, even with another sensor, they could be used to confirm if it is a cloud shadow or not, and help in calibrating the expected reflectance in non-shadow

conditions. Even without any extra images, it might be possible to model the changed illumination conditions in shadows.

Very high resolution images of urban areas have green structure textures on different scales. This indicates that a multiscale approach should be considered. The multiscale approach could be applied at the pixel-level feature extraction, segmentation, segment-level feature extraction, and classification stages. One could use an image pyramid approach, in which 2×2 pixels are merged at the next coarser resolution, ore one could use a smaller number of fixed scales, e.g., three or four, corresponding to the pixel level, individual tree crown/small house level, the garden level, and/or the land use area level. With the latter approach, one needs to estimate the relevant scales.

As features, one could use texture features, the multispectral pixel values, and multispectral band ratios and indexes, all computed at all (selected) scales. At the finest scale, the multispectral pixel values could be taken from a pansharpened multispectral image of the same resolution as the panchromatic image. One alternative is to implement most or all of the texture features and multispectral features mentioned in Section 2, and to use a feature selection method to reduce the number of features. It might be that different feature subsets are selected at different scales. For feature selection, either the sequential forward floating selection method, or the adaptive floating search method could be used. Since one might expect the best feature subsets to be dependent on scale, the process must be repeated for all scales.

One possibility is to do initial segmentations independently on different scales. Next, the segment boundaries at a coarser scale can be adjusted to use the segment boundaries at a finer scale. At the finest scale, the house and road outlines from GIS could be used as recommended segment borders. At the same time, tree canopies overlapping buildings and roads should be mapped. However, perhaps it does not matter if a tree canopy is divided into two or more segments, as long as the individual parts are merged later. At a coarser scale, land use categories from GIS could be helpful, while at the same time allowing for changes that may have occurred after the GIS layers were made.

The goal of the multiscale classification is obviously to extract meaningful entities from the image. At the finest scale, one would like to extract regions corresponding to individual trees, small grass areas, individual houses, garden furniture, etc. On a coarser scale, one would like to separate residential areas from large open grass land, public parks, forests, etc.

As complementary information to the classified image, one may compute key parameters within each segment or another suitable unit, e.g., percentage tree cover, percentage grass cover, percentage grey areas, average tree distance.

In conclusion, this document reviews the state of the art of recent research relevant to the classification of urban green structure from Quickbird images, and suggests possible methods that may be used in an improved classification system. However, experiments are needed in order to get an indication on which of the approaches should be used.

The scope is very wide, and with no concern regarding budget limitations. Obviously, the cost and ambition of an improved automatic green structure classification and measurement system must be taken into account when selecting from the different methods and alternatives presented. There is no individual technique that stands out as the single best solution; rather, a combination of techniques could be used.

1 Introduction

The management of urban green structure has received much attention lately in the research community (e.g., see Caspersen and Olafsson, 2009; Vallejo et al., 2009; Uy and Nakagoshi, 2008; Nilsson et al., 2007). The present project was initiated to meet the need of municipalities in Norway to develop a green structure plan. Traditional mapping has its limitations, since the land use is in focus and not the actual land cover. Therefore, other sources of information about urban and suburban green structure are being sought.

Geodatasenteret AS, Arendal, Norway, has a contract with the Norwegian Space Centre to develop a prototype for a service for monitoring of urban green structure, with the Norwegian Computing Center as a subcontractor. In phase one of the project, Geodatasenteret developed an automatic classification method, implemented in Definiens Developer. The Norwegian Computing Center performed an evaluation of the detection result.

A municipality is interested in a green structure plan for several reasons:

1. To map current status of green areas and their changes over time. For example, what happens with the vegetation in public parks over time, even if the mapped land use does not change?

2. To maintain biological diversity. Different species or groups of species use different varieties of green structure as corridors. For example, small birds would avoid open areas, and need a corridor of trees to move safely. In open areas, they would expose themselves to predators.

3. Green structures are being used for recreation.

4. Vegetation converts carbon dioxide to oxygen, reduces noise, and has aesthetical value. Vegetation also binds water, reducing the prospect of floods after heavy rainfall.

5. If accurate, the green structure map can be used in overlays

The green structure includes private gardens. Although not accessible to the public, private gardens containing trees contributes to items 2 and 4 above.

Forest and farmland are not in the focus of this study, since they are well mapped, and the land cover aligns well with the land use classification of traditional mapping.

The purpose of this study is to discuss the current state of the art of classification methodology for high resolution remote sensing imagery in the context of urban green structure, and to suggest possible directions for improvement. Earlier in this project, a prototype automatic classifier has been developed in Definiens Developer, which uses a hierarchy of classification rules to classify a Quickbird image into three vegetation classes and non-vegetation:

1. Trees

2. Grass

3. Little vegetation

4. Grey areas

5. Water, unknown, or missing data

The classification result has been validated, and the conclusions were (Trier, 2009):

1. Segmentation borders were very rugged

2. Leaving the segmentation problems aside, the classifier had a 9% misclassification rate in the two class problem 'grey area' versus 'green area'. This is a very good basis for further improvement.

3. It is unclear if the automatic method can be used on another image with different light conditions, e.g., with the presence of clouds and light haze.

In this report, several methods are presented, with no concern of whether they can be implemented with Definiens Developer or not.

The rest of the report is organized as follows: Section 2 reviews the current methodology, relevant to, but not restricted to, classification of high resolution remote sensing imagery of urban green structure. These methods are discussed in section 3 in the context of improving the current classification algorithm. The report ends with a conclusion in section 4, and a list of references to the research literature.

2 Current methodology

This section presents recent research relevant to the classification of vegetated areas in urban landscapes.

2.1 Multispectal pixel values

In low to medium resolution images (e.g., MODIS and Landsat), the multispectal pixel values themselves have been widely used as features for classification. As an alternative, spectral band ratios or other band combinations have been used to eliminate varying illumination conditions from one image acquisition to another, or within an image. However, for high resolution images (e.g., Quickbird), the individual pixel values themselves can not always be reliably classified without considering the context, or local neighborhood, e.g., the texture in a local neighborhood. Some sort of segmentation might be needed before classification can take place.

2.2 Spectral mixing

Tooke et al. (2009) use spectral mixing analysis to extract vegetation characteristics in urban areas from Quickbird images. Lidar data is used to estimate shadows in the Quickbird image. A principal component analysis was performed to convert the four-band input space to three dimensions, which were interpreted as 'vegetation', 'high albedo substrate' and 'dark' endmembers in a linear mixing model:

$$R_i = \sum_{j=1}^{n} f_j r_{ij} + \varepsilon_i$$

and

$$0 \leq \sum_{j=1}^{n} f_j \leq 1$$

where R_i is the total pixel reflectance for input image band i, f_j is the endmember image fraction, r_{ij} is the reflectance of image endmember j at band i, n is the number of endmembers, and ε_i is the residual error for band i. In Tooke et al.'s (2009) approach, the number of possible endmembers equals the number of bands minus one.

Decision trees were used to classify pixels into four vegetation classes and a non-vegetation class:

1. Manicured grass cover
2. Mixed ground vegetation, including wild, herbaceous and dry.
3. Evergreen trees
4. Deciduous trees
5. Non-vegetation (including shadow and soil)

The threshold values used in the decision trees were calibrated from field observations.

Small and Lu (2006) used spectral mixing analysis to estimate urban vegetation abundance from Landsat images. Three endmember classes were used: vegetation, substrate, and dark. Calibration and verification was done with Quickbird images.

Stow et al. (2003) used NDVI thresholding followed by spectral mixing analysis to map irrigated vegetation from Ikonos images. The spectral mixing analysis had the following endmembers: (1) green vegetation, (2) soil and impervious materials, and (3) shade. A simple, 3 x 3 variance metric was used to measure texture in the 1 m resolution panchromatic band, this was used to separate medium-to-high vegetation (trees and shrubs) from low vegetation (grass). A decision tree classifier was implemented in the Expert classifier module of the Erdas Imagine software.

Leboeuf et al. (2007) used the shadow fraction to estimate the above ground biomass in boreal forests in Canada from Quickbird images. There was a positive correlation between biomass and shadow fraction, which can be expected when the forest is not too dense.

2.3 3D models from overlapping aerial photos

Iovan et al. (2008) uses 20 cm ground resolution multispectral aerial images to detect, characterize and model vegetation in urban areas, in a four step approach. The images have four bands (blue, green, red and near infrared) and 60% overlap both within-strip and between-strip. The first step is to extract vegetation areas, using supervised classification with a support vector machines classifier. The recognition performance is much better than using NDVI (98.5% versus 87.5%), but the support vector machines classifier has to be re-trained each time the image acquisition conditions change. The second step is based on a digital surface model, which is obtained from the overlapping images. Local height variance was computed using an 11 × 11 neighborhood, and the vegetated areas from step 1 were divided into grass and trees by thresholding the local height variance image. In step three, individual tree crowns are delineated, using as seed points the local maxima of a smoothed version of the digital surface model. In step four, simple parameters such as tree height and crown diameter were extracted. In the final step, texture features are used for supervised classification of tree crowns into tree species using support vector machines. The following texture measures were computed within each delineated tree crown: mean, standard deviation, range, angular second moment, contrast, correlation, entropy, and inverse difference moment. Several of these can be computed from the grey level co-occurrence matrix (Haralick et al., 1973).

2.4 The use of low resolution satellite images

There are several products based on low resolution imagery, such as MODIS (250 m resolution). The 500 m resolution MODIS vegetation continuous field tree cover product is a yearly product that gives a percentage canopy cover for each pixel. Montesano et al. (2009) validated the year 2005 MODIS vegetation continuous field tree cover product in the taiga-tundra transition zone in the northern hemisphere. This area includes the northern limits of the boreal forest. The study area was divided into seven regions, with the Nordic countries and the Kola peninsula being one region. The evaluation was done using Quickbird images. The MODIS vegetation continuous field tree cover product was found to be over-estimating tree cover values in areas with low percent tree cover. Further, in the Nordic countries plus Kola peninsula region, the MODIS product did not perform well. We think that this does not necessarily mean that MODIS is not suitable for estimating percentage tree cover in the Nordic countries. Rather, one will

have to use an alternative algorithm to derive percentage tree cover from MODIS images in the Nordic countries.

2.5 Vegetation indexes

As an alternative to using the multispectral pixel values as features directly, one can use band ratios or other combinations of the spectral bands.

Eklundh et al. (2009) mapped insect damages in pine forests from MODIS time series. They estimated leaf area index from the wide dynamic range vegetation index, WDRVI. WDRVI is based on the normalized difference vegetation index, NDVI. NDVI is popular for extracting vegetated areas; however, it reaches a saturation level, so the WDRVI is better for, e.g., leaf area index estimation in dense forest. NDVI is computed from the red (R) and near infrared (NIR) bands as

$$NDVI = \frac{NIR - R}{NIR + R}$$

WDRVI can be computed from NDVI as

$$WDRVI = \frac{(\alpha + 1)NDVI + (\alpha - 1)}{(\alpha - 1)NDVI + (\alpha + 1)},$$

or, equivalently, directly from the near infrared and red bands (Gitelson, 2004) as

$$WDRVI = \frac{\alpha \cdot NIR - R}{\alpha \cdot NIR + R}$$

The parameter value for α must be selected. α is usually between 0.05 and 0.2, and $\alpha = 0.2$ is a good starting point.

Jin and Sader (2005) used the normalized difference moisture index, NDMI, to detect forest disturbances in Landsat images. NDMI is based on the near infrared band (Landsat band 4) and the short wave infrared band (Landsat band 5):

$$NDMI = \frac{NIR(4) - SWIR(5)}{NIR(4) + SWIR(5)}$$

There are many other spectral vegetation indexes (e.g., see Tucker, 1979; Perry and Lautenschalger, 1984; Ji and Peters, 2007), including green NDVI (GNDVI) and green atmospherically resistant vegetation index (GARI) (Gitelson et al, 1996), and visible atmospherically resistant index (VARI) (Gitelson et al., 2002). GNDVI uses the green (G) band in place of the red band used in NDVI:

$$GNDVI = \frac{NIR - G}{NIR + G}$$

GARI uses atmospherically corrected versions of the blue (B), green (G), red (R), and near infrared (NIR) bands:

$$GARI = \frac{NIR'' - (G'' - \lambda(B'' - R''))}{NIR'' + (G'' - \lambda(B'' - R''))}$$

λ is a parameter that controls the atmospheric correction. VARI uses only visible bands, and the original spectral bands:

$$VARI = \frac{G - R}{G + R - B}$$

VARI was found to be very little affected by atmospheric effects (Gitelson et al., 2002).

2.6 Textural features

A number of textural features have been described in the literature, including

1. Wavelet filter features
2. Gabor filter features
3. Granulometry features
4. Priority sequence Gaussian Markov random field
5. Fractal features
6. Grey level co-occurrence matrix features
7. Local statistics

Due to the curse of dimensionality effect (e.g., see Jain and Chandrasekaran, 1982), one will often have to select a subset of the above texture measure methods. Some studies have investigated which texture measures perform the best in different applications.

Ohanian and Dubes (1992) evaluated the performance of four different classes of textural features: grey level co-occurence features, Markov random field features, Gabor filter features, and fractal features. For the test images they used, the gray level co-occurrence features performed the best, followed by fractal features, but they point out that there is no universal best features. Rather, feature selection has to be performed for each specific application or problem. Clausi and Deng (2005) noted that grey level co-occurrence features captured high frequency information in the textures well, whereas Gabor filters are better at capturing lower and mid frequency information in the textures. Solberg and Jain (1997) used feature selection (Whitney, 1971; also, see Section 2.11 below for alternatives) to combine features derived from different texture models: grey level co-occurrence, local statistics, fractal features, and parameters of lognormal field models.

Aksoy et al. (2010) detected strips of tree vegetation in agricultural landscapes from Quickbird images. First, non-vegetation areas were eliminated by thesholding of the NDVI image. Next, pixel-based image classification was done. For this, two different multiscale texture measures were used on the panchromatic band: Gabor wavelet features (Manjunath and Ma, 1996) and granulometry features (Soille, 2002). In addition, the multispectral values of the pansharpened

image were used as features. The sequential backward selection algorithm (Marill and Green, 1963; Duda et al., 2000) was used to reduce the number of features from 20 to 9. The pixel-based image classification divided the vegetated areas into woody and non-woody. Then, skeletonizing was used to identify elongated objects with thickness between two threshold values.

To obtain rotation invariance, at each pixel, Aksoy et al. (2010) used the maximum value of all Gabor wavelet filters with different orientations at a given scale, and used six scales, resulting in six features at each pixel.

For the granulometry features, Aksoy et al. (2010) used five different disk structuring element radii: 1, 3, 5, 7, and 9; combined with granulometry by closing and granulometry by opening, this gave 10 features at each pixel.

Chellappa and Chatterjee (1985) introduced Gaussian Markov random field texture features. Zhao et al. (2007) used what they called "priority sequence Gaussian Markov random field"-based texture features for classification of land cover types in Ikonos and Spot-5 images. They observed that residential areas are difficult to detect by spectral pixel values as features, because of high variance and low correlation. The lowest order variance texture feature was effective for residential area extraction.

2.6.1 Wavelet filter features

Lucieer and van der Werff (2007) extracted wavelet texture features from the panchromatic band of a Quickbird image, and combined them with the multispectral pixel values for land cover classification. A 16 x 16 pixels neighborhood in the panchromatic band, centered on a multispectral pixel, was used for wavelet decomposition, based on the Daubechies wavelet (Daubechies, 1988; Antonini et al., 1992). The highest resolution wavelet decomposition on the 16 x 16 pixels subimage has 64 coefficients. From these, the following features were computed: standard deviation, skewness, kurtosis, entropy, and energy.

2.6.2 Gabor filter features

Ohanian and Dubes (1992) used four directions and four frequencies, resulting in 16 Gabor filters. The averages of the filtered images were used as features, 16 in total. The Gabor filters are expected to discriminate textures with different orientations and/or different frequencies.

A Gabor filter in 2D is a sinusiodal plane wave, with frequency θ and orientation ρ, modulated by a Gaussian envelope (Ohanian and Dubes, 1992; Manthalkar et al., 2003):

$$h(x, y; \theta, \rho) = \cos(2\pi\theta u) e^{-\frac{1}{2}\left(\frac{u^2}{\sigma_x^2} + \frac{v^2}{\sigma_y^2}\right)},$$

where

$u = x\cos\rho$ and $v = y\sin\rho$.

We will then use n different filters

$$h_j = h(x, y; \theta_j, \rho_j), \quad 0 \le j < n$$

resulting in n filtered images

$$I_j = I * h_j, \quad 0 \le j < n$$

where '*' denotes convolution, and I is the input image.

Ohanian and Dubes (1992) used the average absolute deviation of the zero-mean filtered images I_j as features:

$$f_j = \frac{1}{N^2} \sum_{x=1}^{N} \sum_{y=1}^{N} \left| I_j(x,y) \right|$$

Ohanian and Dubes (1992) computed the features for 32 x 32 subimages, using four orientations 0°, 45°, 90°, and 135°, and four frequencies 1/16, 1/8, 1/4, and 1/2.

2.6.3 Granulometry features

Granulometric features can be extracted from the panchromatic band by grey scale morphological opening and closing operations. Granulometry by opening in effect removes bright grains smaller than the structuring element, whereas granulometry by closing removes similarly small dark grains. Aksoy et al. (2009) used disk structuring elements of radii 1, 3, 5, 7, and 9. Local granulometric features were computed as the sum within sliding windows, for each of the opened and closed images. We suggest that mean values be used instead of sums. For the Quickbird images, we would need to consider if the number of grey level values should be reduced prior to applying the opening and closing operations.

2.6.4 Grey level co-occurrence features

For the grey level co-occurrence features, Ohanian and Dubes (1992) used a distance of one pixel only, and four angles (0, 45, 90 and 135 degrees). Four features for each angle were extracted:

1. angular second moment
2. contrast
3. correlation
4. entropy

Solberg and Jain (1997) used the same four angles, but a 9 x 9 window (instead of a 3 x 3 window) for the computation of the co-occurrence matrix. They used the following additional features:

5. cluster shade
6. inertia
7. inverse difference moment

The grey level co-occurrence features are expected to capture high frequency information in a texture.

A grey level co-occurrence matrix looks at all pixels in an image that are separated by a fixed distance d at a fixed orientation α, and counts the number of occurrences for all possible pairs of grey level values, (i,j). These counts are divided by the total number of pixel pairs separated by the distance d at the direction α, $N(d,\alpha)$. For texture analysis, the image in question is usually a small subimage, so, to get reliable estimates, the number of grey levels is usually reduced to a very few, e.g., $N_g = 8$. The dimension of the grey level co-occurrence matrix is then $N_g \times N_g$. The elements of the $N_g \times N_g$ matrix $P(d,\alpha)$ are defined as

$$P(i,j;d,\alpha) = \frac{p(i,j;d,\alpha)}{N(d,\alpha)}$$

It is common to use only one distance, $d = 1$, and four directions $\alpha = 0^\circ$, 45°, 90°, and 135°. To obtain rotation invariance, the average over all directions can be taken, resulting in only one co-occurrence matrix P, with elements:

$$P(i,j) = \frac{1}{n_\alpha} \sum_{i=1}^{n_\alpha} P(i,j;1,\alpha_i),$$

where n_α is the number of angles.

Ohanian and Dubes (1992) computed features from four directional co-occurrence matrices $P(1,\alpha)$, while Solberg and Jain (1997) used the single, rotation invariant one, P.

From the co-occurrence matrix, a number of features can be extracted. Haralick et al. (1973) suggested 14. Solberg and Jain (1997) used four of these, and two additional ones. In the expressions below, P can be either one of the directional co-occurrence matrices, or the rotation invariant one. In the context of urban green structure, it makes sense to use rotation invariance.

1. angular second moment

   $$\sum_{i,j} P^2(i,j)$$

2. contrast

   $$\sum_{n=0}^{N_g-1} n^2 \sum_{|i-j|=n} P(i,j)$$

3. inverse difference moment

   $$\sum_{i,j} \frac{p(i,j)}{1+(i-j)^2}$$

4. entropy

   $$-\sum_{i,j} P(i,j)\log(P(i,j))$$

5. inertia

$$\sum_{i,j}(i-j)^2 P(i,j)$$

6. cluster shade

$$\sum_{i,j}(i+j-\mu_x-\mu_y)^3 P(i,j)$$

where

$$\mu_x = \sum_{i=0}^{N_g-1} i \sum_{j=0}^{N_g-1} P(i,j)$$

and

$$\mu_y = \sum_{j=0}^{N_g-1} j \sum_{i=0}^{N_g-1} P(i,j)$$

7. correlation

$$\frac{1}{\sigma_x \sigma_y}\left(\sum_{i,j} ijp(i,j) - \mu_x \mu_y\right)$$

2.6.5 Markov random field features

An image pixel and its eight neighbors can be modeled as a Markov random field. The grey tone of image pixel *t* is modeled as a random variable X_t. The conditional probability distribution of an observed value x_t for X_t is $P(X_i=x_i \mid R_t)$, where R_t is the grey levels of the neighboring pixels. Below, *G* is the number of grey levels in the image.

$$P(x_i \mid R_t) = \frac{(G-1)!}{x_t!(G-1-x_t)!} q^{x_t} (1-q)^{G-1-x_t}$$

where

$$q = \frac{e^{-T}}{1+e^{-T}}$$

and

$$T = \sum_r \theta_r (x_{t+r} + x_{t-r})$$

It is common to reduce *G*, the number of grey levels in the image, to a small number in order to obtain numerical stability. Ohanian and Dubes (1992) used $G = 4$. Also, the number of neighbors is usually kept small. Ohanian and Dubes (1992) used eight neighbors, meaning that *r* varied from 1 to 4 in the sum in the equation for *T* above. The parameters θ_r are estimated and used as features.

If a larger number of grey levels is desired, say, G>8, then one may use Gaussian Markov random field, instead of the discrete Markov random field described above.

2.6.6 Gaussian Markov random field

The following description is based on Zhao et al. (2007). In a Gaussian Markov random field model, the grey level value of a pixel *t* has a conditional probability density function

$$p(f(t) \mid f_R(t)) = \frac{1}{\sqrt{2\pi\nu}} e^{-\frac{1}{2\nu}\left[f(s)-\mu-\sum_{r\in R}\beta(r)(f(s+r)-\mu)\right]^2}$$

where

$f_R(t) = \{f(t+r) \mid r \in R\}$ is the set of grey level values of the neighbor pixels of *t*, μ is the mean value of the entire image, $\beta(r)$ are the modeled parameters, and ν is the conditional variance.

At the same time, the grey level value $f(t)$ of the pixel *t* can be represented as a linear combination of its neighbors and additive noise:

$$f(t) - \mu = \sum_{r\in R} \beta(r)(f(t+r) - \mu) + e(t)$$

where $\{e(t)\}$ is a zero-mean Gaussian noise sequence with the following correlation structure:

$$Cov[e(t), e(t+r)] = \begin{cases} \nu, & r = (0,0) \\ -\beta(r)\nu, & r \in R \\ 0, & otherwise \end{cases}$$

The neighborhood *R* can be split into two regions R^+ and R^-, so that if $r \in R^+$, then $-r \in R^-$, $r \notin R^-$, and $-r \notin R^+$. Also, $\beta(r) = \beta(-r)$. Then,

$$f(s) - \mu = e(s) + \sum_{r\in R^+} \beta(r)\left[(f(t+r) - \mu) + (f(t-r) - \mu)\right]$$

The parameters $\beta(r)$ and the conditional variance ν, can be used as texture features. The least square estimate for the vector $\boldsymbol{\beta}$ of $\beta(r)$ values is

$$\hat{\beta} = \left[\sum_t Q(t)Q^T(t)\right]^{-1}\left[\sum_t Q(t)(f(t) - \mu)\right]$$

where Q(t) is a column vector

$$Q(t) = col\left[(f(t+r) - \mu) + (f(t-r) - \mu) \mid r \in R^+\right]$$

The conditional variance ν is estimated as

$$\hat{v} = \frac{1}{M \times N} \sum_t \left[(f(t) - \mu) - Q^T(s)\hat{\beta} \right]^2$$

The sums over t are taken on an $M \times N$ subimage centered on the pixel t.

2.6.7 Priority sequence Gaussian Markov random field

Zhao et al. (2007) introduced *priority sequence* Gaussian Markov random fields as a means to estimate the parameters $\beta(r)$ and v in a step-by-step fashion. For a small subimage centered on a pixel t, they ordered the pixels by distance from the central pixel. For example, by using a 5×5 neighborhood, orders up to 5 could be used (Figure 1).

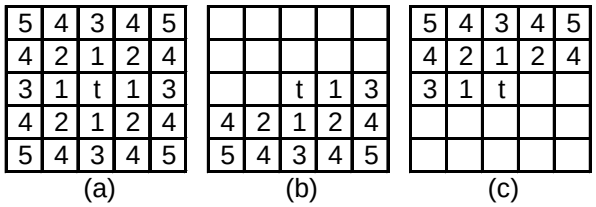

Figure 1. (a) Neighborhood orders of pixels surrounding a pixel t., (b) the region R^+, (c) the region R^-.

In the case of a 5×5 neighborhood, one could then choose to use orders up to 4, and the parameter vector β is then divided into four groups: β_1, β_2, β_3, and β_4. Each group β_i corresponds to a group of neighbors, R_i^+ (Figure 2).

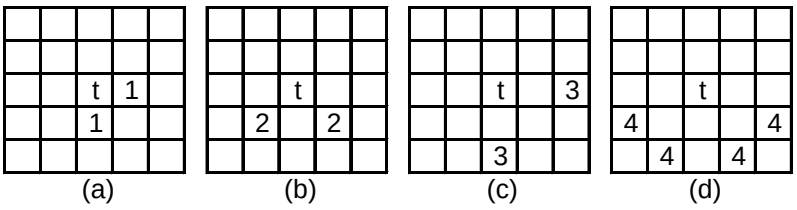

Figure 2. The four groups of neighborhoods. (a) R^+_1, (b) R^+_2, (c) R^+_3, and (d) R^+_4.

Then, for group k, $Q_k(t)$ is the column vector

$$Q_k(t) = col\left[(f(t+r) - \mu) + (f(t-r) - \mu) \mid r \in R_k^+ \right]$$

For $k = 1$, the estimated parameters are

$$\overline{\beta}_1 = \left[\sum_t Q_1(t) Q_1^T(t) \right]^{-1} \times \left[\sum_t Q_1(t)(f(t) - \mu) \right]$$

$$\overline{v}_1 = \frac{1}{M \times N} \sum_t \left[f(t) - \mu - Q_1^T(t)\overline{\beta}_1 \right]^2$$

For $1 < k \leq m$, where m is the total number of groups,

$$\overline{\beta}_k = \left[\sum_t Q_k(t) Q_k^T(t) \right]^{-1} \times \left[\sum_t Q_k(t)(f(t) - \mu + Z) \right]$$

$$\overline{v}_k = \frac{1}{M \times N} \sum_t \left[f(t) - \mu - \sum_{i=1}^{k} Q_i^T(t) \overline{\beta}_i \right]^2$$

where

$$Z = \sum_{i=1}^{k-1} \left[\frac{1}{m} (f(t) - \mu) - Q_i^T(t) \overline{\beta}_i \right]$$

The estimated values of the parameters β_1, β_2, …, β_m are then combined to form the estimated parameter vector $\boldsymbol{\beta}$. The estimates of v_i for increasing orders are based on increasing neighborhood sizes, so all of them are valid as features. Zhao et al. (2007) found the lowest order variance v_1 to be useful for the discrimination between residential areas and other land use types.

2.6.8 Fractal features

The fractal dimension of a grey level image can be computed when the image is regarded as a surface with height equal to the grey level value in each pixel. Another fractal measurement is lacunarity.

The fractal dimension can be estimated by several different methods:

1. power spectrum method
2. variation method (Dubuc et al, 1989)
3. box counting method (Keller et al., 1989)
4. ε-blanket method (Peleg et al., 1984)

Ohanian and Dubes (1992) found the power spectrum method to give the best estimate for the fractal dimension, whereas Huang et al (1994) found the variation method (Dubuc et al, 1989) to be the best. The variation method computes the maximum variation, $V(\varepsilon)$, of the image surface, within squares of side ε. The fractal dimension D is estimated as 3-s, where s is the slope of the least square fit of the data $\{\ln(\varepsilon), \ln(V(\varepsilon))\}$, as ε varies (Solberg and Jain, 1997). Ohanian and Dubes (1992) noted that the methods for computing the fractal dimension should ideally be applied on larger subimages than the size of 32 × 32 pixels that they used. Further, they observed that it could be beneficial to use more than one estimate of the fractal dimension in the set of textural features.

Lacunarity can be estimated from the box counting method (Keller et al., 1989). We will be using subimages of size $n \times n$, and the actual range of grey levels of the entire image should be re-mapped to n values from 0 to n-1 (regarded as height values in the context of a fractal surface). Let A be the volume in $\mathcal{R}^3$, composed of cubes of unit size, between the height z=0 and the re-mapped image. Let $N(L)$ be the number of boxes of side L needed to completely contain

A, and *M*(*L*) the average mass density within a box of side *L*. The lacunarity can then be estimated as

$$C(L) = \frac{N(L) - M(L)}{N(L) + M(L)}$$

Ohanian and Dubes (1992) used C(3) as the lacunarity feature for 32 × 32 subimages (possibly reduced to 30 × 30 pixels when estimating lacunarity).

2.7 Local statistics

In a local neighborhood *N* centered on the pixel (*x*,*y*), one can compute local statistics. Many of these are based on computing moments

$$E(I^k(x,y)) = \frac{1}{n}\sum_{(i,j)\in N} I^k(i,j)$$

and central moments

$$E((I(x,y) - \mu)^k) = \frac{1}{n-1}\sum_{(i,j)\in N} (I(i,j) - \mu)^k$$

In both the above, n is the number of pixels in the neighborhood N. μ is the mean value in the local neighborhood, and is the first moment:

$$\mu = E(I(x,y))$$

We will also use σ, the standard deviation, which squared is the second central moment:

$$\sigma^2 = E((I(x,y) - \mu)^2)$$

Solberg and Jain (1997) used the following features based on local statistics:

1. power-to-mean ratio

   $$\sigma / \mu$$

2. skewness

   $$\frac{E((I(x,y) - \mu)^3)}{\sigma^3}$$

3. kurtosis

   $$\frac{E((I(x,y) - \mu)^4)}{\sigma^4}$$

4. contrast

$$\frac{1}{n-1}\sum_{(i,j)\in N}\frac{(I(x,y)-I(i,j))^2}{\mu^2}$$

5. homogeneity

$$\frac{1}{n-1}\sum_{(i,j)\in N}\frac{1}{1+\frac{(I(x,y)-I(i,j))^2}{\mu^2}}$$

2.8 Hyperspectral images

Xiao et al. (2004) used airborne hyperspectral 3.5 m resolution images to detect tree types and tree species in an urban environment. With this resolution, some tree canopies were smaller than one pixel. The sensor has 224 bands in the range 400-2400 nm. Of these, 131 were used, avoiding bands with wavelengths above 1800 nm and water absorbing bands. Spectral mixing analysis was used to separate land cover surface into vegetation, bare soil, pavement, building, water, etc. Spectral mixing analysis was further used to subdivide vegetation into shrub, grass and tree; tree into broadleaf and conifer; broadleaf into evergreen and deciduous; and, finally, the evergreen, deciduous and conifer tree types into their respective tree species.

2.9 Multisensor image fusion

Hyde et al. (2006) studied combinations of lidar, interferometric X-band SAR, Landsat ETM+ and Quickbird imagery for the mapping of forest canopy height and biomass. Lidar was the best single sensor. The combination of all sensors gave the best result. The best combination of two sensors was lidar and Landsat ETM+. Hyde et al. (2006) did not give any reason why Landsat ETM+ was a better companion to lidar than Quickbird. The lidar data was collected at 7 km flying height and with approximately 20m footprint diameter. This resolution matches well with the 15m panchromatic and 30m multispectral resolutions of Landsat ETM+, in contrast with the 0.6 m and 2.4 m resolutions of Quickbird.

2.10 Multi-scale segmentation

Lamonaca et al. (2008) used multi-scale segmentation in eCognition, to map structural complexity in Quickbird images of beech forests in Italy. They used three segmentation levels, starting with large segments in level 1 and small segments in level 3. The smallest segments in level 3 corresponded to the canopy gap resulting from one tree that has died. They observe that forests segments that are homogeneous at a coarse scale, are heterogeneous when analyzed internally at a finer scale, and *vice versa*.

2.11 Feature selection methods

In pattern recognition problems, one is often faced with the problem of reducing the number of measurements, or features, of each pattern or object to classify. At a first glance, one might think that it is beneficial to have many rather than few measurements of each object or pattern. However, in order to estimate a large number of parameters in a statistical classifier, neural network classifier or support vector machine classifier, a large training set is needed to estimate the parameters reliably. Also, in non-parametric classifiers (e.g., k nearest neighbor) and in decision tree classifiers, it is beneficial to identify the features which have the most discriminative power.

As mentioned above, Aksoy et al. (2010) used the sequential backward selection algorithm (Duda et al., 2000). Jain and Zongker (1997) evaluated the following methods, and found the sequential forward floating selection method to be performing almost as well as the "optimal" branch-and-bound method, while having a much lower computation time.

1. Branch-and-bound (Narendra and Fukunaga, 1977)
2. Node pruning (Mao et al., 1994)
3. Genetic algorithm (Siedlecki and Sklansky, 1989)
4. Sequential forward selection (Whitney, 1971)
5. Sequential backward selection (Marill and Green, 1963)
6. Plus-*r* take-away-*l* (Stearns, 1976)
7. Sequential forward floating selection (Pudil et al., 1994)
8. Sequential backward floating selection (Pudil et al., 1994)
9. Max-min algorithm (Backer and Schipper, 1977)

Somol et al. (1999) have proposed an adaptive floating search method, which is in essence a variation of Plus-*r*-take-away-*l*, where the parameters *r* and *l* are determined dynamically by the algorithm.

The sequential forward floating selection method can be summarized as follows (Somol et al., 1999).

1. Add the most significant feature to the current subset of size k. Let $k = k + 1$.
2. Conditionally remove the least significant feature in the current subset.
3. If the current subset is the best subset of size $k - 1$ found so far, let $k = k - 1$ and go to step 2. Else, return the conditionally removed feature back into the current subset, and go to step 1.

The adaptive floating search, in the sequential forward variety, is more complicated to describe, and we refer to (Somol et al., 1999).

3 Discussion

In the previous section, a number of relevant image analysis and pattern recognition techniques that could be used for automatic classification of urban green structure. There is no individual technique that stands out as the single best solution; rather, a combination of techniques could be used.

The choice of methods is to a large extent influenced by what remote sensing and other data is available. Clearly, in order to observe changes over time, repeated acquisitions are important, preferably by the same sensor. If different sensors are being used, then, in general, a new automatic classification algorithm has to be developed. At best, an existing algorithm can be re-trained on data from the new sensor.

However, there is no guarantee that an algorithm that was trained on one single acquisition from one sensor will perform well on another image acquisition from the same sensor. Imaging conditions may have changed, due to varying atmospheric conditions, including varying presence of aerosols, haze; sun elevation, and time of the day. Also, seasonal variations within a year and between years, often referred to as phenological variation, alter the spectral and textural properties of vegetation. Also, the time since rainfall, and temperature and humidity variations adds to the variability. It is therefore important to use methods that compensate for these variations in one way or another, and to include training data that reflect the typical variations that one may encounter in the operational setting.

Another problem is the presence of clouds, which obscures parts of the scene. This means that multiple image acquisitions may be necessary to ensure full ground coverage.

The choice of imaging sensor is dependent on many factors, including

1. The time frequency of repeated acquisitions
2. The prospect of cloud free acquisitions
3. The size of the area to be covered
4. The prospect of performing atmospheric correction in a reliable way
5. The ground pixel size of the image in relation to the sizes of the phenomena on the ground one wishes to observe

The alternatives include aerial orthophotos with ground resolution between 10 and 50 cm, very high resolution satellites like Quickbird (0.6 m panchromatic, 2.4 m visible and near infrared (VNIR)), Landsat (15 m pan, 30 m multispectral (MS), MODIS (250 m red and near infrared, 500 m MS), and the upcoming Sentinel-2 (10 m MS).

Aerial photographs are traditionally acquired at cloud free conditions. In the process of making orthophotos, which is a geo-corrected mosaic of individual photos, illumination corrections are applied to produce an image with no visible abrupt changes in color across mosaic edges. Aerial photographs may also be acquired at complete overcast, provided the clouds are at high

elevation so that the airplane can fly below the clouds. Atmospheric correction of a single Quickbird image is difficult, since the methods are based on infrared channels that are not present. Landsat images have such channels to some extent, and MODIS have dedicated channels for this purpose. The upcoming Sentinel-2 satellite will also have dedicated channels for atmospheric correction. A technique that might help is to find MODIS acquisitions acquired within a few hours of the Quickbird acquisition to estimate the atmospheric correction to be applied to the Quickbird image. An alternative that could be considered is to use the ideas of the VARI index (Gitelson et al., 2002), which was found to be very little influenced by atmospheric effects, to provide a rough estimate of the atmospheric disturbance.

The previous project phase (Trier, 2009) clearly demonstrated that Quickbird imagery is a relevant source of information for urban green structure assessment, and provided repeated acquisitions are made, one could expect to be able to monitor and quantify changes.

As we have seen in the previous phase of the project (Trier, 2009), shadows from trees and buildings create problems for the segmentation and classification methods. If lidar data is available, the location and extent of these shadows may be predicted, and corrected for. Cloud shadows are much larger, and they may be detected as separate classes. If multiple image acquisitions are available, even with another sensor, they could be used to confirm if it is a cloud shadow or not, and help in calibrating the expected reflectance in non-shadow conditions. Even without any extra images, it might be possible to model the changed illumination conditions in shadows, i.e., direct sunlight is blocked, allowing only indirect light to illuminate the ground, objects, and vegetation. This could then be compensated for, in effect removing the cloud shadows. In a more simplistic approach, one could increase the brightness of all channels by the same amount, or to the amount of cloud shadow free areas.

Some authors have eliminated non-vegetation areas at an early stage. This can be done using, e.g., NDVI, or another vegetation index. However, small non-vegetated areas within vegetated areas should not be removed, as these are part of the texture patterns of residential areas. Whether this step is needed, or if it is actually counterproductive, must be decided through initial experiments. A variation of this step can also be used for automatic cloud removal, provided a good cloud index is used.

The previous chapter mentioned a number of feature extraction methods. In principle, all of them could be used to assign features to individual pixels, or be computed for segments. For segments, one could either use the average of the values for each pixel, or one could re-compute the values within each segment.

Very high resolution images of urban areas have green structure textures on different scales. This indicates that a multiscale approach should be considered. The multiscale approach could be applied at the pixel-level feature extraction, segmentation, segment-level feature extraction, and classification stages. One could use an image pyramid approach, in which 2×2 pixels are merged at the next coarser resolution, or one could use a smaller number of fixed scales, e.g., three or four, corresponding to the pixel level, individual tree crown/small house level, the garden level, and/or the land use area level. With the latter approach, one needs to estimate the relevant scales.

As features, one could use texture features, the multispectral pixel values, and multispectral band ratios and indexes, all computed at all (selected) scales. At the finest scale, the

multispectral pixel values could be taken from a pansharpened multispectral image of the same resolution as the panchromatic image.

Two vegetation indexes that could be used in Quickbird images are NDVI and VARI. If we have Landsat imagery available, one could use the normalized difference moisture index, NDMI, but one should consider the extra effort needed to process another image type for the sake of only one extra feature.

Some of the texture feature extraction methods are multiscale by nature. However, these multiscale texture measures are often extracted from the same neighborhood size in the finest resolution. Since we really are interested in textures at very different scales, e.g., single pixels, single tree canopies, and contiguous residential areas, an alternative is to fix the texture measurement methods at the finest scale, and apply it on an image pyramid. Alternatively, if using a few select scales, the features may be computed for all intermediate scales up to the next coarser image scale.

One alternative is to implement most or all of the texture features and multispectral features mentioned in Section 2 above, and to use a feature selection method to reduce the number of features. It might be that different feature subsets are selected at different scales. For feature selection, either the sequential forward floating selection method (Pudil et al., 1994), or the adaptive floating search method (Somol et al., 1999) could be used. Since one might expect the best feature subsets to be dependent on scale, and on whether it is segmentation of pixels or classification of segments, the process must be repeated for all combinations.

One possibility is to do initial segmentations independently on different scales. Next, the segment boundaries at a coarser scale can be adjusted to use the segment boundaries at a finer scale. At the finest scale, the house and road outlines from GIS could be used as recommended segment borders. At the same time, tree canopies overlapping buildings and roads should be mapped. However, perhaps it does not matter if a tree canopy is divided into two or more segments, as long as the individual parts are merged later. At a coarser scale, land use categories from GIS could be helpful, while at the same time allowing for changes that may have occurred after the GIS layers were made.

The goal of the multiscale classification is obviously to extract meaningful entities from the image. At the finest scale, one would like to extract regions corresponding to individual trees, small grass areas, individual houses, garden furniture, etc. On a coarser scale, one would like to separate residential areas from large open grass land, public parks, forests, etc.

As complementary information to the classified image, one may compute key parameters within each segment or another suitable unit, e.g., percentage tree cover, percentage grass cover, percentage grey areas, average tree distance.

4 Conclusions

This document has reviewed the state of the art of recent research relevant to the classification of urban green structure from Quickbird images, and suggested possible methods that may be used in an improved classification system. Experiments are needed in order to get an indication on which of the approaches should be used.

The scope has been very wide, and with no concern regarding budget limitations. Obviously, the cost and ambition of an improved automatic green structure classification and measurement system must be taken into account when selecting from the different methods and alternatives presented above. There is no individual technique that stands out as the single best solution; rather, a combination of techniques could be used.

Acknowledgements

This research was supported by the Norwegian Space Centre.

References

Aksoy, S., Akçay, G., Wassenaar, T., 2010. Automatic mapping of linear woody vegetation features in agricultural landscapes using very high resolution imagery. *IEEE Transactions on Geoscience and Remote Sensing*, to appear.

Antonini, M., Barlaud, M., Mathieu, P., Daubechies, I., 1992. Image coding using wavelet transform. *IEEE Transactions on Image Processing* 1 (2), pp. 205-220.

Backer, E., Schipper, J. A. D., 1977. On the max-min approach for feature ordering and selection. *The Seminar on Pattern Recognition*, Liège University, Sart-Tilman, Belgium, p. 2.4.1.

Caspersen, O. H., Olafsson, A. S., 2010. Recreational mapping and planning for enlargement of the green structure in greater Copenhagen. *Urban Forestry and Urban Greening*, to appear.

Chellappa, R., Chatterjee, S., 1985. Classification of textures using Gaussian Markov random fields. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 33 (4), pp. 959-963.

Clausi, d. a., Deng, H., 2005. Design-based texture feature fusion using Gabor filters and co-occurrence probabilities. *IEEE Transactions on Image Processing* 14 (7), pp. 925-936.

Daubechies, I., 1988. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics* 41 (7), pp. 909-996.

Duda, R. O., Hart, P. E., Stork, D. G., 2000. *Pattern Classification*. Wiley, Hoboken, NJ.

Eklundh, L., Johansson, T., Solberg, S., 2009. Mapping insect defoliation in Scots pine with MODIS time-series data. *Remote Sensing of Environment* 113, pp. 1566-1573.

Gitelson, A. A., 2004. Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation. *Journal of Plant Physiology* 161, pp. 165-173.

Gitelson, A. A., Kaufman, Y. J., Merzlyak, M. N., 1996. Use of green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment* 58, pp. 289-298.

Gitelson, A. A., Kaufman, Y. J., Stark, R., Rundquist, D., 2002. Novel algorithms for remote estimation of vegetation fraction. Remote Sensing of Environment 80, pp. 76-87.

Haralick, R. M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics* 3 (6), pp. 610-621.

Huang, Q., Lorch, J. R., Dubes, R. C., 1994. Can the fractal dimension of images be measured? *Pattern Recognition* 27, pp. 339-349.

Iovan, C., Boldo, D., Cord, M., 2008. Detection, characterization and modeling vegetation in urban areas from high resolution aerial imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and remote Sensing* 1 (3), pp. 206-213.

Jain, A. K., Chandrasekaran, B., 1982. Dimensionality and sample size considerations in pattern recognition practice. In *Handbook of Statistics*, Krishnaiah, P. R., Kanak, L. N., Eds. North-Holland, Amsterdam, The Netherlands. Vol. 2, pp. 835-855.

Jain, A. K., Zongker, D., 1997. Feature selection: evaluation, application, and small sample performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19 (2), pp. 153-158.

Ji, L., Peters, A. J., 2007. Performance evaluation of spectral vegetation indices using a statistical sensitivity function. *Remote Sensing of Environment* 106, pp. 59-65.

Jin, S., Sader, S. A., 2005. Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. *Remote Sensing of Environment* 94, pp. 364-372.

Kudo, M., Sklansky, J., 2000. Comparison of algorithms that select features for pattern classifiers. *Pattern Recognition* 33 (1), pp. 25-41.

Leboeuf, A., Beaudoin, A., Fournier, R. A., Guindon, L., Luther, J. E., Lambert, M.-C., 2007. A shadow fraction method for mapping biomass of northern boreal black spruce forests using Quickbird imagery. *Remote Sensing of Environment* 110, pp. 488-500.

Lucieer, A., van der Werff, H., 2007. Panchromatic wavelet texture features fused with multispectral bands for improved classification of high-resolution satellite imagery. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium,* 23-28 July 2007, pp. 5154-5157.

Manjunath, B. S., Ma, W. Y., 1996. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18 (8), pp. 837-842.

Mao, J., Mohiuddin, K., Jain, A. K., 1994. Parsimonious network design and feature selection through node pruning. *Proceedings of the 12th IAPR International. Conference on Pattern Recognition, 1994. Vol. 2 - Conference B: Computer Vision & Image Processing*., pp. 622-624.

Marill, T., Green, D. M., 1963. On the effectiveness of receptors in recognition systems. *IEEE Transactions on Information Theory* 9, pp. 11-17.

Montesano, P. M., Nelson, R., Sun, G., Margolis, H., Kerber, A., Ranson, K. J., 2009. MODIS tree cover validation for the circumpolar taiga-tundra transition zone. *Remote Sensing of Environment* 113, pp. 2130-2141.

Narendra, P. M., Fukunaga, K., 1977. A branch and bound algorithm for feature subset selection. *IEEE Transactions on Computers* 26 (9), pp. 917-922.

Nilsson, K., Åkerlund, U., Konijnendijk, C. C., Alekseev, A., Caspersen, O. H., Guldager, S., Kuznetsov, E., Mezenko, A., Selikhovkin, A., 2007. Implementing urban greening aid projects – the case of St. Petersburg, Russia. *Urban Forestry and Urban Greening* 6, pp. 93-101.

Ohanian, P. P., Dubes, R. C., 1992. Performance evaluation for four classes of textural features. *Pattern Recognition* 25 (8), pp. 819-833.

Peleg, S., Naor, J., Hartley, R., Avnir, D., 1984. Multiple resolution texture analysis and classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, pp. 518-523.

Perry jr, C. R., Lautenschlager, L. F., 1984. Functional equivalence of spectral vegetation indices. *Remote Sensing of Environment* 14, pp. 169-182.

Pudil, P., Novovičová, J., Kittler, J., 1994. Floating search methods in feature selection. *Pattern Recognition Letters* 15 (11), pp. 1119-1125.

Siedlecki, W., Sklansky, J., 1989. A note on genetic algorithms for large-scale feature selection. *Pattern Reecognition Letters* 10 (5), pp. 335-347.

Small, C., Lu, J. W. T., 2006. Estimation and vicarious validation of urban vegetation abundance by spectral mixture analysis. *Remote Sensing of Environment* 100, pp. 441-456.

Soille, P., 2002. *Morphological Image Analysis*, 2nd ed. Springer Verlag, New York.

Somol, P., Pudil, P., Novovičová, J., Paclík, P., 1999. Adaptive floating search methods in feature selection. *Pattern Recognition Letters* 20, pp. 1157-1163.

Stearns, S. D:, 1976. On selecting features for pattern classifiers. *Proceedings of the Third Joint International Conference on Pattern Recognition*, Coronado, CA, pp. 71-75.

Stow, D., Coulter, L., Kaiser, J., Hope, A., Service, D., Schutte, K., Walters, A., 2003. Irrigated vegetation assessment for urban environments. *Photogrammetric Engineering and Remote Sensing* 69 (4), pp. 381-390.

Tooke, T. R., Coops, N. C., Goodwin, N. R., Voogt, J. A., 2009. Extracting urban vegetation characteristics using spectral mixture analysis and decision tree classifications. *Remote Sensing of Environment* 113, pp. 398-407.

Trier, Ø. D., 2009. *Urban green structure – validation of automatic classification*. Note no. SAMBA/39/09, Norwegian Computing Center, Oslo, Norway.

Tucker, C. J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment* 8, pp. 127-150.

Uy, P. D., Nakagoshi, N., 2008. Application of land suitability analysis and landscape ecology to urban greenspace planning in Hanoi, Vietnam. *Urban Forestry and Urban Greening* 7, pp. 25-40.

Vallejo jr, B. M., Aloy, A. B., Ong, P. S., 2009. The distribution, abundance and diversity of birds in Manilla's last greenspaces. *Landscape and Urban Planning* 89, pp. 75-85.

Withney, A. W., 1971. A direct method of nonparametric measurement selection. *IEEE Transactions on Computers* 20, pp. 1100-1103.

Xiao, Q., Ustin, S. L., McPherson, E. G., 2004. Using AVIRIS data and multiple-masking techniques to map urban forest tree species. *International Journal of Remote Sensing* 25 (24), pp. 5637-5654.

Zhao, Y., Zhang, L., Li, P., Huang, B., 2007. Classification of high spatial resolution imagery using improved Gaussian Markov random field-based texture features. *IEEE Transactions on Geoscience and Remote Sensing* 45 (5), pp. 1458-1468.