Leena Lepistö

Colour and Texture Based Classification of Rock Images Using Classifier Combinations

Tampere 2006

Leena Lepistö

Colour and Texture Based Classification of Rock Images Using Classifier Combinations

Thesis for the degree of Doctor of Technology to be presented with due permission for public examination and criticism in Tietotalo Building, Auditorium TB223, at Tampere University of Technology, on the 7th of April 2006, at 12 noon.

Preliminary assessors

Professor Robert P.W. Duin
Delft University of Technology
Faculty of Electrical Engineering, Mathematics and Computer Science
The Netherlands

Professor Jussi Parkkinen
University of Joensuu
Department of Computer Science
Finland

Opponents

Professor Erkki Oja
Helsinki University of Technology
Department of Computer Science and Engineering
Finland

Professor Jussi Parkkinen
University of Joensuu
Department of Computer Science
Finland

Custos

Professor Ari Visa
Tampere University of Technology
Department of Information Technology
Finland

ii

Abstract

The classification of natural images is an essential task in current computer vision and pattern recognition applications. Rock images are a typical example of natural images, and their analysis is of major importance in the rock industry and in bedrock investigations. Rock image classification is based on specific visual descriptors extracted from the images. Using these descriptors, images are divided into classes according to their visual similarity.

This thesis investigates rock image classification using two different approaches. Firstly, the colour and texture based description of rock images is developed by applying multiscale texture filtering techniques to the rock images. The emphasis in such image description is to make the filtering for the selected colour channels of the rock images. Additionally, surface reflection images obtained from industrial rock plates are analysed using texture filtering methods. Secondly, the area of image classification is studied in terms of classifier combinations. The purpose of the classifier combination strategies proposed in this thesis is to combine the information provided by different visual descriptors extracted from the image in the classification. This is attained by using separate base classifiers for each descriptor and combining the opinions provided by the base classifiers in the final classification. In this way the texture and colour information of rock images can be combined in the classification to achieve better classification accuracy than a classification using separate descriptors.

These methods can be readily applied to automated rock classification in such fields as the rock and stone industry or bedrock investigations.

Preface

This research was carried out at the Institute of Signal Processing of Tampere University of Technology, Finland. It formed part of the DIGGER project which was jointly funded by industry and the Technology Development Centre of Finland (TEKES).

I would also like to thank the reviewers, Prof. Robert P.W. Duin and Prof. Jussi Parkkinen for their constructive comments and for giving their valuable time to review the manuscript.

I would also like to extend my thanks to my colleagues and the members of DIGGER project: Jorma Autio, Dr. Jukka Iivarinen, Juhani Rauhamaa, and Prof. Ari Visa.

I would also like to thank Alan Thompson for the language revision of this thesis. Special thanks are due to Prof. Josef Bigun for the chance to participate in his research group at Halmstad University, Sweden in autumn 2004.

To my parents go thanks for the joy of growing up with five sisters and four brothers in such an exciting and inspiring rural family background.

Finally, I am deeply grateful to my dearest Iivari. His love and support have helped make this thesis possible.

Tampere, March 2006

Leena Lepistö

> *"Kenelläkään ei ole hauskempaa,*
> *kuin mitä hän itse itselleen järjestää."*
> Tove Jansson

List of abbreviations

AR	Autoregressive
CCD	Charge coupled device
CIE	International Commission of Illumination
CMY	Cyan, magenta, yellow
CPV	Classification probability vector
CRV	Classification result vector
ECOC	Error-correcting output codes
HSI	Hue, saturation, intensity
k-NN	k-nearest neighbour
LBP	Local binary pattern
MA	Moving average
MDS	Multidimensional scaling
MPEG	Motion Picture Experts Group
PCA	Principal component analysis
RGB	Red, green, blue
SAR	Simultaneous autoregressive

Table of contents

List of publications

I. Lepistö, L., Kunttu, I., Autio, J., Visa, A., 2003. Multiresolution Texture Analysis of Surface Reflection Images, In Proceedings of 13th Scandinavian Conference on Image Analysis, LNCS Vol. 2749, pp. 4-10, Göteborg, Sweden.

II. Lepistö, L., Kunttu, I., Autio, J., Visa, A., 2003. Classification Method for Colored Natural Textures Using Gabor Filtering. In Proceedings of 12th International Conference on Image Analysis and Processing, pp. 397-401, Mantova, Italy.

III. Lepistö, L., Kunttu, I., Visa, A., 2005. Rock image classification using color features in Gabor space. Journal of Electronic Imaging, 14(4), 040503.

IV. Lepistö, L., Kunttu, I., Autio, J., Visa, A., 2003. Classification of Non-homogenous Textures by Combining Classifiers. In Proceedings of IEEE International Conference on Image Processing, Vol. 1, pp. 981-984, Barcelona, Spain.

V. Lepistö, L., Kunttu, I., Autio, J., Rauhamaa, J., Visa, A., 2005. Classification of Non-homogenous Images Using Classification Probability Vector. In Proceedings of IEEE International Conference on Image Processing, Vol. 1, pp. 1173-1176, Genova, Italy.

VI. Lepistö, L., Kunttu, I., Visa, A., 2005. Color-Based Classification of Natural Rock Images Using Classifier Combinations. In Proceedings of 14th Scandinavian Conference on Image Analysis, LNCS Vol. 3540, pp. 901-909, Joensuu, Finland.

VII. Lepistö, L., Kunttu, I., Visa, A., 2005. Rock image classification based on k-nearest neighbour voting. IEE Proceedings of Vision, Image, and Signal Processing, to appear.

x

1 Introduction

In recent years, the use of digital imaging has increased rapidly in several areas of life thanks to the decreased costs of digital camera technology and the development of image processing and analysis methods. It is nowadays common for imaging tools to be used in several fields which earlier required manual inspection and monitoring. Imaging methods are widely used in a variety of monitoring and analysis tasks in fields such as health care, security, quality control, and process inspection. Different image-based classification tasks are also routinely performed in numerous industrial manufacturing processes.

Compared to manual inspection and classification, the use of automated image analysis provides several benefits. Manual inspection carried out by people is, as might be expected, affected by human factors. These factors include personal preferences, fatigue, and the concentration levels of the individual performing the inspection task. Therefore, inspection is a subjective task, dependent on the personal inclinations of the individual inspector, with individuals often arriving at different judgments. By contrast, automated inspection by computer with a camera system performs both inspection and classification tasks dependably and consistently. Another drawback of manual inspection is the amount of manual labour expended on each task.

1.1 Computer vision and pattern recognition

Automatic image analysis carried out by computer is referred to as computer vision. In a computer vision system the human eye is replaced by a camera while the computer replaces the human brain. It can be said that the purpose of a computer vision system is to give a robot the ability to see (Schalkoff, 1989). Typically, computer vision is employed in the inspection of goods and products in industrial processes (Newman and Jain, 1995), but it can also be used in other types of image-based analysis and inspection tasks. In the process industry, significant amounts of information on the process can be acquired using computer vision. This information is utilized in process monitoring and control tasks. One typical area of the process industry employing computer vision systems is the web
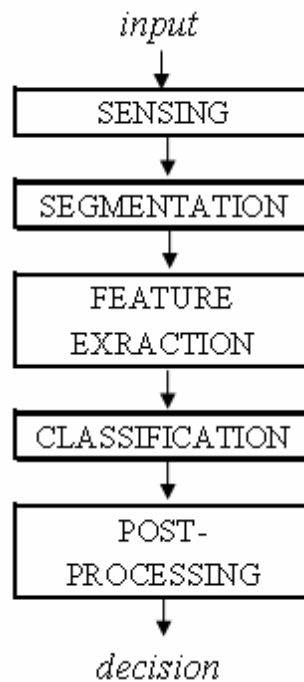
Figure 1.1. Components of a pattern recognition system (Duda et al., 2001).

material industry that includes metal, paper, plastics, and textile manufacturing (Iivarinen, 1998). In this area, computer vision is often used to detect and classify a range of defects and anomalies which occur in the production process. Quality control of products is a central task and the use of computer vision systems is also increasing in other types of manufacturing and production tasks both in industry and research. The application of different texture analysis methods is common in various visual inspection tasks (Pietikäinen et al., 1998; Kumar and Pang, 2002; Baykut et al., 2000), and colour-based applications also exist (Boukovalas et al., 1999; Kauppinen, 1999).

In contrast to inspection performed by a manual inspector, a computer vision system processes all information systematically without the inconsistencies caused by human factors. In addition to industrial quality and production control, computer vision systems are widely applied to such areas as traffic monitoring as well as a variety of security and controlling tasks. People identification based on facial features or fingerprints, recognition of handwritten characters, and medical imaging applications are examples of typical image-based recognition and classification tasks.

The main parts of a computer vision system include image acquisition, image processing and analysis. Pattern recognition methods are widely used in computer vision systems to analyze and recognize the image content. Duda et al. (2001) have described the process of pattern recognition and classification as illustrated in Figure 1.1. The process starts with sensing of certain input which, in this case refers to image acquisition. Image acquisition is nowadays mainly performed using digital imaging methods and the images are then processed using a computer. The second step in the procedure is segmentation. It is often necessary to extract a certain region of interest from the image to be used in inspection. This way, the object to be classified is isolated from the other

objects and the background of the image; a process called image segmentation. In addition to segmentation, noise reduction and image enhancement methods, such as sharpening, can also be employed. The third step is feature extraction. The purpose of the feature extractor is to characterize the object to be recognized by using measurements whose values are very similar to objects in the same category and also very different to objects in different categories (Duda et al., 2001). In image recognition and classification, certain features are extracted from the images. The features often form feature vectors, also called descriptors, which are able to describe the image content. The fourth step is classification. The idea of classification is to assign the unknown image to one of a number of categories. If predefined categories are used, the classification is said to be supervised, otherwise it is unsupervised. Finally, in the post-processing stage, the classification result can be estimated using various validation methods.

1.2   Image classification

The present study deals with the problem of image classification. In image-based pattern recognition, images are used to describe real-world objects. This thesis focuses on feature selection and classification problems arising from the classification of images, particularly rock images.

1.2.1   Feature selection

In the previous Section, it was noted that the features describing the object to be classified should be such that they distinguish between different categories. Therefore, the features should describe the desired properties of the object. On the other hand, the features should be invariant to irrelevant transformations, such as scale, translation or rotation of the object to be recognized (Duda et al., 2001).

In the case of image classification, descriptors extracted from the images are employed. The most typical visual properties used in image classification relate to the colours, textures, and shapes occurring in the images. These properties are described by calculating different kinds of descriptors based on them. In the fields of image analysis and pattern recognition, numerous descriptors have been proposed for use in the description of image content. In addition, much research has focused on the problem of image content description in the field of content-based image retrieval (Del Bimbo, 1999; Smeulders et al., 2000). In content-based image retrieval approaches, one of the goals is to describe image content by means of the visual descriptors extracted from the images. Consequently, descriptors used in retrieval approaches can also be used to characterize image content in image classification.

1.2.2   Multiple classifier systems

An unknown object is classified into one of the categories on the basis of certain properties. Normally, several different properties measured from the object are used in the decision process. One option is to employ several classifiers or experts, each dealing with different aspects of the input (Duda et al., 2002). Alternatively, different classifiers may classify the object into different categories even if the input is the same. The simplest case is when all classifiers elicit the same decision. However, when the classifiers are in disagreement, the situation is more complicated. By analogy, one may suppose that a person suffering from mysterious pains consults five doctors. Now, if four

doctors diagnose for disease A and only one for disease B, should the final diagnosis be based on a majority opinion? However, it is possible that the doctor diagnosing disease B is the only specialist in this highly specialised area of medicine and, therefore, uniquely competent to make a correct diagnosis. In this case, the medical majority would be wrong. This problem may also be viewed from a different perspective in which some of the five doctors are undecided as to a definitive diagnosis and instead propose that the patient may be suffering from either disease A or B, but that A is more likely. In this case, in addition to decisions, probabilities are also being expressed. This additional information can assist in making the final decision as to the actual nature of the disease.

The above example illustrates the difficulties in reaching a final decision based on the opinions provided by different experts. In the field of pattern recognition it has been shown that a consensus decision of several classifiers can often provide greater accuracy than any single classifier (Ho et al., 1994; Kittler et al., 1998). As a result, several strategies have been developed to obtain a consensus decision on the basis of the opinions of different classifiers in pattern classification. Some of the strategies are based on simple voting (Lam and Suen, 1997; Lin et al., 2003) whereas others consider the probabilities provided by the separate classifiers (Kittler et al., 1998).

In the case of image classification, multiple classifier systems can be used in several ways. In the present study, classifier combinations are used to classify unknown images into predefined categories. In this approach, classification based on different visual descriptors is first made separately. After this, the final decision is made by combining the results of the separate classifications. In such a procedure the opinion produced by each descriptor affects the final decision.

1.3   Rock images

The application area of this thesis is rock and stone images. In the rock industry, the visual inspection of products is essential because the colour and texture properties of rock often vary greatly, even within the same rock type. Therefore, when rock plates are manufactured, it is important that the plates used, such as in flooring, share common visual properties. In addition, visual inspection is necessary in the quality control of rock products. Traditionally, rock products have been manually classified into different categories on the basis of their visual similarity. However, in recent years the rock and stone industry has adopted computer vision and pattern recognition tools for use in rock image inspection and classification. In addition to the inspection of the visual properties of rock materials, the application of automated image-based inspection can also provide other benefits for rock manufacturers. For instance, the strength of the rock material can often be estimated by analyzing the surface structures of the rock plates.

Additionally, in the field of rock science, the development of digital imaging has made it possible to store and manage images of the rock material in digital form (Autio et al., 2004). One typical application area of rock imaging is bedrock investigation which is utilized in many areas from mining to geological research. In such analysis, rock properties are analyzed by inspecting the images collected from the bedrock using borehole imaging. Some of the essential visual features of the images obtained from bedrock samples are texture, grain structure and colour distribution of the samples. The images of the rock samples are stored into image databases to be utilized in rock

inspection. Due to the relatively large size of such databases, automated image analysis and classification methods are necessary.

1.4 Outline of thesis

This aim of this study is to contribute to research into the classification of natural rock images. The classification task of the images obtained from rock is investigated in terms of two approaches. The first is the selection of effective visual descriptors for rock images. Successful classification requires descriptors which are capable of providing an effective description of image content. In the case of rock images, the descriptors should be capable of describing the colour and texture properties of rock that is often non-homogenous. For this purpose, a multiscale texture filtering technique that is also applied to colour components of rock images is used. In addition, statistical histogram-based methods are used in image content description.

In the second approach, classifier combinations are used for rock images. The classifier combinations include different types of visual descriptors such as colour and texture, extracted from the images. This is motivated by the fact that improved classification accuracy can be achieved using classifier combinations compared to classification using separate descriptors. The organization of this thesis is as follows:

Chapter 2 provides an introduction to the field of rock imaging and image analysis. The Chapter begins with a description of the image-based rock analysis problem. The imaging methods for rock materials as well as colour and texture properties are discussed. Earlier studies on rock image analysis are reviewed.

Chapter 3 provides a brief overview of research undertaken in the field of texture and colour description.

Chapter 4 focuses on classification methods and the major classification principles are discussed. In addition, the special character of image classification is described.

The topic of classification is continued in Chapter 5 with an introduction to classifier combination methods. Previous work in this research area is reviewed and the most common classifier combination methods are presented.

Chapter 6 discusses the application of the classification methods for rock images and there are brief introductions to publications related to this thesis. The author's own contributions to the publications are presented. Conclusions arising from this study are presented in Chapter 7.

2 Rock image analysis

The application area in this thesis is rock image classification. As mentioned in the previous Chapter, automated rock image analysis is essential in the rock and stone industry as well as in rock science. However, rock like most other natural image types, such as clouds, ice or vegetation, is seldom homogenous and this often makes their classification problematic. Indeed, the colour and texture properties of rock may vary significantly even within the same rock type. In this Chapter, the special character of rock images is examined. Colour and texture features are also considered and there is finally a review of previous work conducted in rock image analysis

2.1 Image-based rock analysis

The inspection of rock materials is essential in several areas. Typical examples of these are mining, underground construction, and oil well production. In several geoscientific disciplines from remote sensing to petrography, rock inspection and analysis tasks play important roles (Autio et al., 2004). In practical rock inspection applications, rock materials have been classified according to various factors such as their mineral content, physical properties, or origin (Autio et al., 2004). In this kind of analysis, rock properties are examined by inspecting bedrock using boreholes.

In addition to bedrock investigation, another important field of rock material inspection is the construction industry. Rock is commonly used in buildings where ornamental rock plates are used for such purposes as floor and wall covering. In the rock plate manufacturing process, control of the visual properties of the plates is important. This is because visual properties such as the colour or texture of the rock plates should form a harmonic surface. In addition, cracks and other surface defects in the rock plates can be detected using imaging methods.

Visual inspection is equally important in bedrock investigation as it is in rock plate manufacture. In both cases, manual inspection of rock materials is still widely practiced. In bedrock investigation, the manual inspection of core samples obtained from boreholes is carried out by geologists to determine the mineral content of the rock (Spies, 1996). The core samples are also stored for future analysis. The storage of core samples may

easily involve hundreds of kilometres of core samples. There are several problems with this conventional way of core sample inspection. In the first place, manual inspection is subjective since classification is always dependent the personal view of the individual performing the analysis. Secondly, manual inspection is a very labour-intensive way of analyzing large amounts of rock. The third problem is storage of the core samples for future analysis tasks. Accessing a core sample of interest involves a visit to the storage site and then the desired samples must be located. Similar problems also arise with rock plate production. The conventional manual inspection of the rock plate manufacturing is labour-intensive and subjective. In addition, any documents of produced rock plates cannot be stored for future analysis.

Several problems associated with manual inspection in bedrock investigation as well as in rock plate manufacturing can be overcome by using automated image analysis. In bedrock analysis, the core samples obtained from the bedrock can be scanned into digital form using core scanning techniques. This means the rock samples can be analyzed as images, which make it possible to use automated pattern recognition and image analysis tools in the rock analysis. As a result, different rock materials can be distinguished and classified automatically on the basis of the visual properties of the rock. Automatic image analysis is a fast way of classifying large amounts of rock materials and the subjectivity problems encountered with manual classification can be avoided (Autio et al., 2004). Another significant benefit of image-based rock investigation is the storing of the rock samples. When the images are stored in digital form for future analysis tasks, the desired core samples can be easily retrieved from a digital image database. In the case of rock plate production, the image-based rock inspection makes it possible to automatically classify the rock plates according to their visual properties (Autio et al., 2004). Furthermore, the images of each plate can be stored into database which can serve as a documentary source for each plate for the manufacturer. These images can be used, for example, to construct the desired types of surfaces from the plates.

2.2   Rock imaging

Several types of image acquisition systems have been introduced in the field of rock imaging. Imaging applications are nowadays based on digital imaging, typically CCD cameras. In geoengineering applications, different types of scanners have enabled routine image acquisition in underground engineering (Autio et al., 2004). The scanners can be applied using two different principles. Core scanners are used to take images of the core samples drilled from the bedrock. They are typically horizontal scanners, which acquire the image of the cylinder-shaped core sample rotating under the camera. The core sample rotates 360° and the camera acquires the image of the surface of the cylinder. Figure 2.1 shows an example of a horizontal colour core scanner developed by DMT GmbH, Germany. In scanners of this type, special attention must be given to the stabilization of the light source and colour range of the camera system (Autio et al., 2004). Figure 2.2 presents two images of the core surface. Another image acquisition principle is borehole imaging using borehole scanners or cameras. In these applications, the borehole is imaged instead of the core samples drilled from the hole.
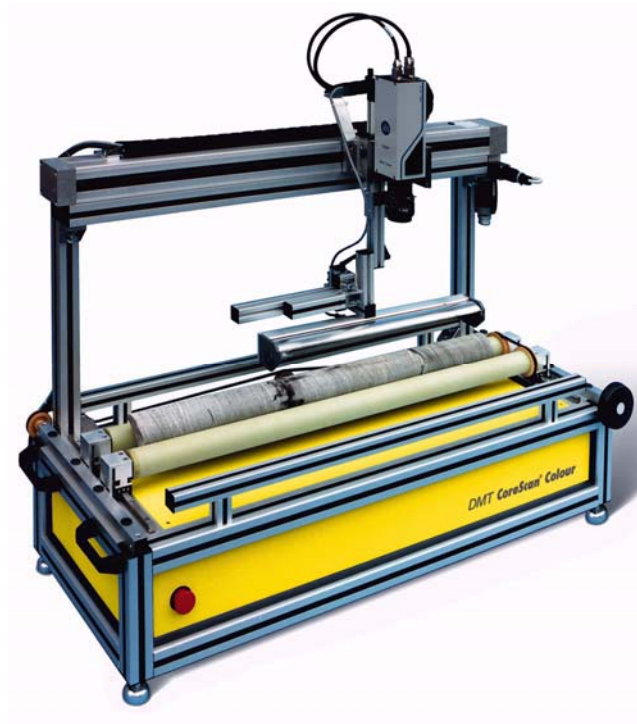
Figure 2.1. The CoreScan Colour, horizontal core scanner developed by DMT GmbH, Germany.

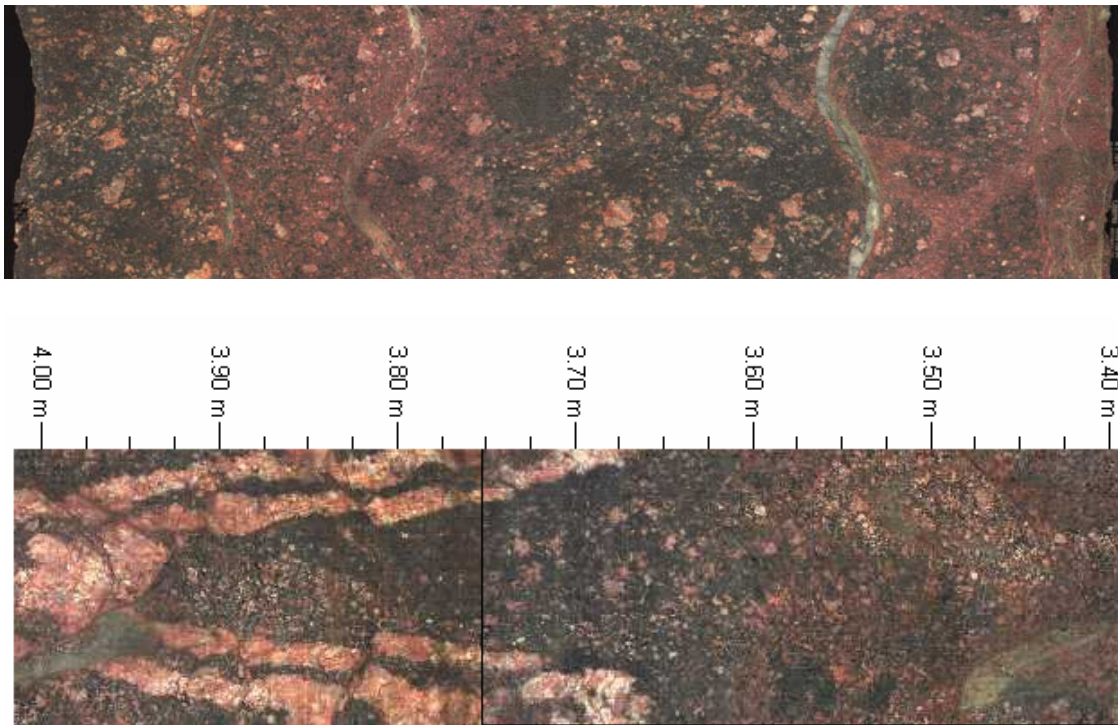
Figure 2.2. Examples of core images.
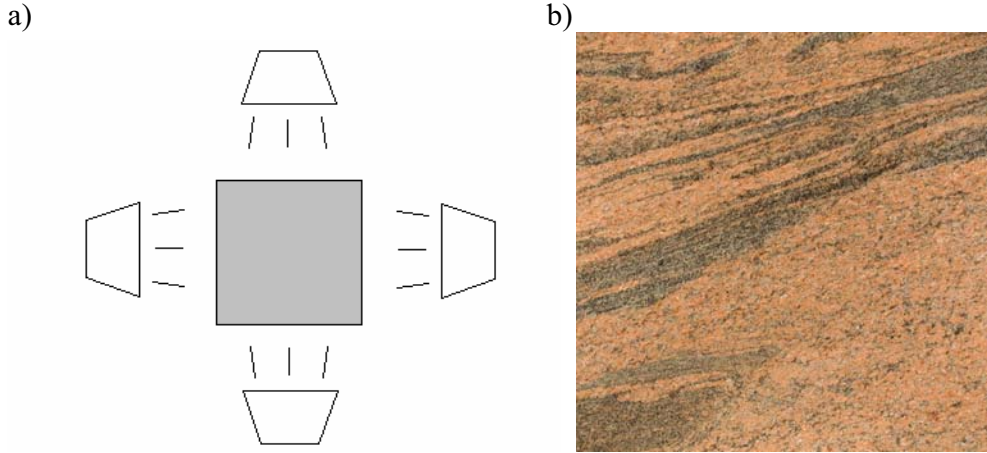
a)   b)

Figure 2.3 a) The illumination setup used in rock plate imaging, b) A sample image obtained from a rock plate.

The bedrock images used in this thesis have been obtained using horizontal core scanners. In the case of the industrial rock plates, the surface image can be acquired using a digital imaging system in which the plate is illuminated strongly enough to allow all the visual properties of rock to be acquired. To avoid light reflection on the plate surface, suitable lightning conditions can be achieved by illuminating the horizontally located square-shaped plate from each side. This allows the light to approach the plate surface horizontally. Figure 2.3a shows the lightning principle. In this kind of rock plate imaging system, the camera is located above the plate. A rock plate image is presented in Figure 2.3b.

There are also other interesting rock properties that can be measured using imaging techniques. In rock plate production it is often necessary to inspect the plate surface because when used in external walls, they must withstand a range of weather conditions (Lebrun, 2000). Cracking and other defects present in the surface of the rock plate have a significant effect on its ability to resist damage due to frost and moisture. It is, therefore, essential for a rock manufacturer to be able to inspect plate surfaces. The surface of a polished rock plate can be inspected using total reflection. According to Snell's law (Keller et al., 1993), when light reaches the surface of two different materials, it partially reflects and partially transmits. If the light approaches at an angle Θ_1 with respect to the surface normal, the angle of reflection Θ_{1r} is equal to Θ_1 (figure 2.4a). The angle of the refracted ray, Θ_2, can be defined according to Snell's law:

$$n_1 \sin \Theta_1 = n_2 \sin \Theta_2 \tag{2.1}$$

where n_1 and n_2 are constants dependent on the material. Thus the angle Θ_2 can be defined as follows:

$$\Theta_2 = \sin^{-1}\left(\frac{n_1}{n_2} \sin \Theta_1 \right) \tag{2.2}$$
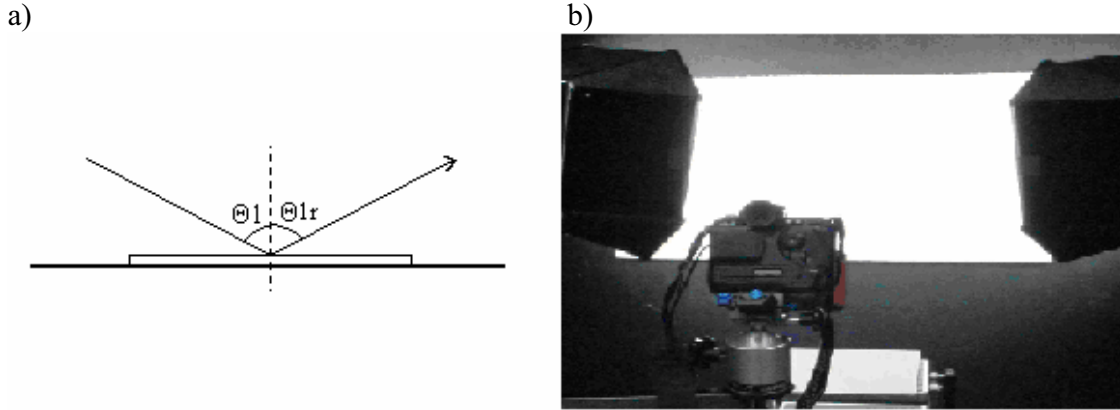
a)

b)

Figure 2.4. a) Total reflection, b) The setup for rock plate surface imaging.
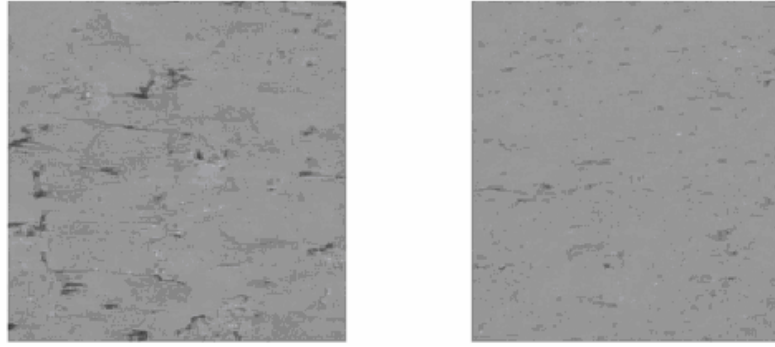

Figure 2.5. Examples of reflection images acquired from rock plate surfaces.

Sinus function cannot be given values above one. Because sinus function has a value one with an angle of 90°, it is possible to define critical angle Θ_c:

$$n_1 \sin \Theta_c = n_2 \sin 90° = n_2 \qquad (2.3)$$

$$\sin \Theta_c = \frac{n_2}{n_1} \qquad (2.4)$$

At the surface, reflection and transmission occur when the approaching angle Θ_1 is lower than critical angle Θ_c. If Θ_1 is greater than Θ_c, all light is reflected and this is referred to as total reflection (Keller et al., 1993). When light is directed against the surface at an angle Θ_1 which is higher than the critical angle, the surface acts as a mirror reflecting all the light at an angle Θ_{1r}. This can be utilized in surface inspection since light is reflected from a smooth polished surface in a different manner than from a surface containing irregularities such as cracks. Using this kind of approach, even minute cracks and defects can be detected. Figure 2.4b illustrates an imaging setup for rock plate surface inspection. In the imaging arrangement, fluorescence tubes illuminate the plate via a white vertical surface. This kind of lightning arrangement provides even illumination across the surface. Figure 2.5 shows two examples of surface reflection images of rock plates.
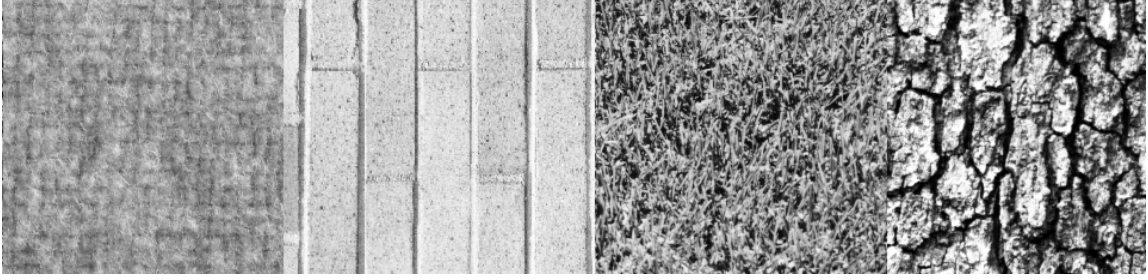
Figure 2.6. Example textures from Brodatz album (1968).

2.3 The features of rock images

2.3.1 Texture features of rock

Several types of rock properties can be estimated on the basis of texture. However, there are certain special properties of rock that can complicate analysis work and the most important of these is the non-homogeneity of the rock images. This condition is commonly expressed in the texture distribution of the rock images. This section considers the significance of the texture of rock images.

Textures

Texture is one of the most important image characteristics to be found almost anywhere in nature. It can be used to segment images into distinct objects or regions. Indeed, the classification and recognition of different surfaces is often based on texture properties. Textures can be roughly divided into two categories: deterministic and stochastic textures (Van Gool et al., 1985). A deterministic texture is composed of patterns which are repeated in an ordered manner. Almost all textures occurring in nature are stochastic and in these textures, the primitives do not obey any statistical law. Figure 2.6 presents sample textures from Brodatz album (1968) in which woollen cloth and a brick wall exemplify deterministic textures while grass and bark exemplify stochastic ones.

Even if textures occur almost everywhere, there is still no universal definition available for them. Despite this, several different definitions do exist that may be used to describe texture. Haralick (1979) describes texture as a phenomenon formed by texture primitives and their organization. In the definition of Van Gool et al. (1985), texture is defined as a structure that consists of several more or less oriented elements or patterns. Tamura et al. (1978) define texture T as a simple mathematical model:

$$T = R(t) \tag{2.5}$$

in which R corresponds to the organization of texture primitives t.

Human texture perception has been the subject of several studies. Tamura et al. (1978) investigated texture analysis from a psychological viewpoint. They proposed six significant visual properties for texture, namely coarseness, contrast, directionality, line-likeness, regularity, and roughness. According to Rao and Lohse (1993), the most essential texture properties in human perception are repetitiveness, directionality, granularity, and complexity. These properties have been studied by Liu and Picard (1996)
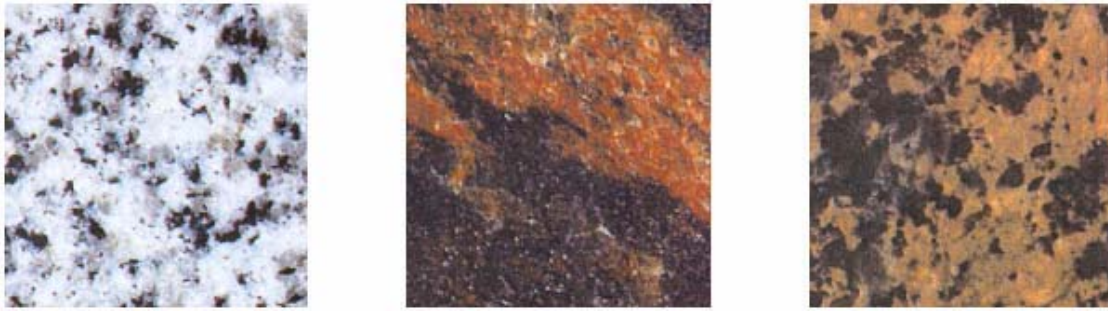
22

Figure 2.7. Examples of three different rock textures.

who proposed the three Wold features for texture description. Wold features describe periodicity, directionality, and randomness of texture. Julesz (1981) considers texture perception as a part of human visual perception. There are two basic models for texture perception suggested by Julesz (1981), feature model and frequency model. In the feature model, the texture perception obeys texture features called textons. All the patterns, lines, and orientations occurring in the texture are regarded as textons. The frequency model considers the texture image in terms of its frequency distribution.

Rock textures

Rock texture is stochastic in texture like most other natural textures. Figure 2.7 presents three different rock texture types. In addition to the stochastic nature of the rock textures, a more significant characteristic of the rock texture is non-homogeneity. In natural textures non-homogeneity is common which can make their analysis and classification somewhat complicated. The homogeneity of an image can be estimated by dividing a sample image into smaller blocks. After this certain texture features describing properties such as directionality or granularity are calculated for each block. If the features do not significantly vary between the blocks, the sample is deemed to be homogenous. Conversely, if these feature values show significant variance, the texture sample is non-homogenous. This division into blocks has been applied in (Lepistö et al., 2003a).

Texture properties are significant in rock image analysis. Based on texture, it is possible to estimate several types of rock properties. For example, the visual properties of rock plates manufactured in the building industry are dependent on the texture of the rock surface. Texture directionality and granularity are very important texture properties in rock texture analysis. This is because several types of rock textures have strong directionality and granular size of the rock texture also often varies. The orientation and the strength of directionality is important, for example, in obtaining a harmonic rock plate surface. All the plates should be similarly oriented to achieve an impression of a visually regular surface. In addition the granular sizes of the plates should not vary greatly. Texture directionality and granularity are important factors in terms of human texture perception and therefore have a significant effect on the visual properties of a surface constructed of rock. In bedrock investigation directionality and granularity play a major role in the recognition of different rock types. Certain rock properties such as strength of
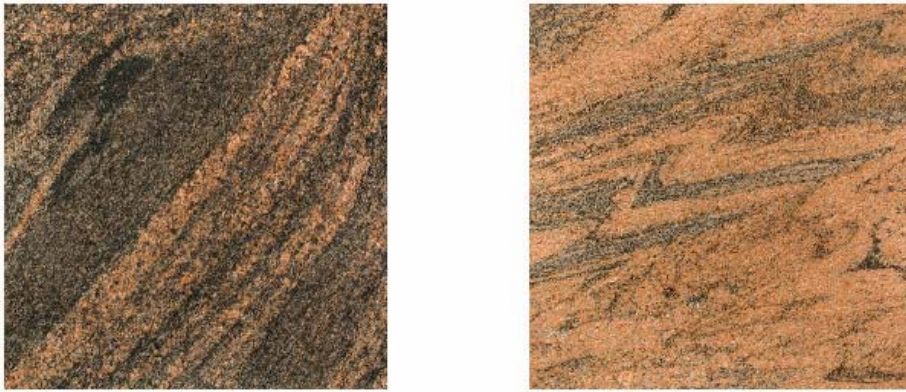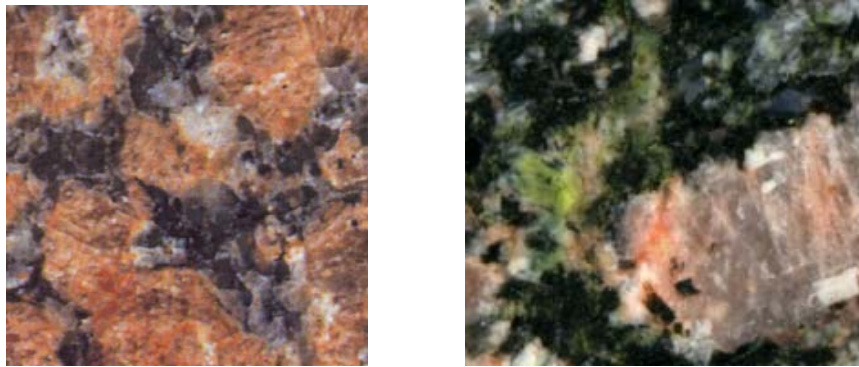
Figure 2.8. Directional rock textures.

Figure 2.9. Rock textures with different grain structures.

the rock can also be estimated on the basis of directionality and granularity. In some applications, the granular size and detection of grains of a certain size and colour are also important (Lepistö et al., 2004).

Texture homogeneity is often expressed in terms of directionality or granularity. Figure 2.8 shows two examples of rock textures with different directionalities. In the first, the orientation is quite regular, but there are changes in the strength of directionality. In the second texture sample, texture directionality is clearly non-homogenous. Textures with varying grain structures are shown in Figure 2.9.

2.3.2 Colour features of rock

In addition to texture, colour is one of the basic characteristics used in image content description (Del Bimbo, 1999). Gonzales and Woods (1993) present two basic motivations for the use of colour description in image analysis. First, in automatic image analysis, colour is a powerful descriptor that often simplifies object identification and extraction from the scene. Second, in image analysis performed by human beings, the motivation for colour is that the human eye can discern thousands of colour shades and intensities, compared to about only two-dozen shades of grey. The use of colour information is, therefore, essential in several areas of image analysis. The present Section discusses the significance of colour in rock image analysis.

Colour

The history of image analysis begins in the 17th century, when Sir Isaac Newton conducted experiments with light. He observed that a beam of sunlight can be divided into a continuous spectrum of colours ranging from violet at one end to red at the other (Gonzales and Woods, 1993). The spectrum of light can be divided into six broad regions: violet, blue, green, yellow, orange, and red. The colour perceived by the human eye in an object is determined by the nature of light reflected from the object. Visible light is a narrow band in the spectrum of electromagnetic energy, whose wavelengths are varying between 400 and 700 nm.

Achromatic light does not include colour and hence its only attribute is its intensity (Gonzales and Woods, 1993). Intensity is often described by means of its scalar measure, grey level. However, colour information is also necessary in several recognition and analysis tasks. In these cases, chromatic colour is considered. Chromatic light is coloured and can be described in terms of three basic properties: radiance, luminance, and brightness. Radiance refers to the total energy that flows from the light source and is usually measured in watts (W). Luminance characterizes the energy perceived by the observer from the light source. The unit of luminance is lumen (lm). Brightness corresponds to the intensity of achromatic light. It is a subjective descriptor that is almost impossible to measure (Gonzales and Woods, 1993). In addition to brightness, other measures describing colour are hue and saturation. Hue is associated with the dominant wavelength in a mixture of light waves. The combination of hue and saturation is referred to as chromaticity of light, and therefore a colour may be characterized by its brightness and chromaticity (Gonzales and Woods, 1993).

All colours can be presented as variable combinations of the three primary colours red (R), green (G), and blue (B) (Wyszecki and Stiles, 1982). The primary colours were standardized in 1931 by the International Commission of Illumination, CIE[1]. The exact wavelengths for the primary colours defined by CIE were 700 nm for red, 546.1 nm for green, and 435.8 nm for blue. It is possible to make additional secondary colours, magenta, cyan, and yellow by adding the primary colours. The secondary colour model containing these three colours is referred to as the CMY model. The RGB colour model can be presented in a Cartesian coordinate system as shown in Figure 2.10. In the colour cube of Figure 2.10, each colour appears in its primary spectral components of red, green, and blue. In the cube, the primary colours are at three corners and the secondary colours are at the other three corners. Black is at the origin and white is at the corner farthest from the origin. The grey level extends from black to white along the diagonal while colours are points on or inside this cube, defined by vectors extending from the origin.

In addition to the RGB colour space, several other colour models have been introduced (Gonzales and Woods, 1993; Wyszecki and Stiles, 1982). Another colour space used in this study is HSI colour space. In the HSI model, hue, saturation, and intensity are considered separately. This is beneficial for two reasons (Gonzales and Woods, 1993). First, the intensity component (I) is decoupled from the colour information in the image. Second, the hue (H) and saturation (S) components are intimately related to the way in which human beings perceive colour.

---

[1] Commission Internationale de l' Eclairage
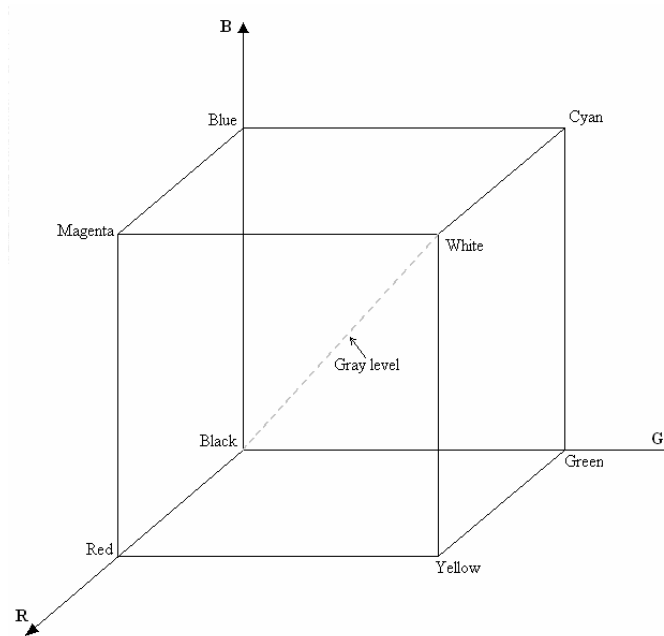
Figure 2.10. The colour cube of RGB system.
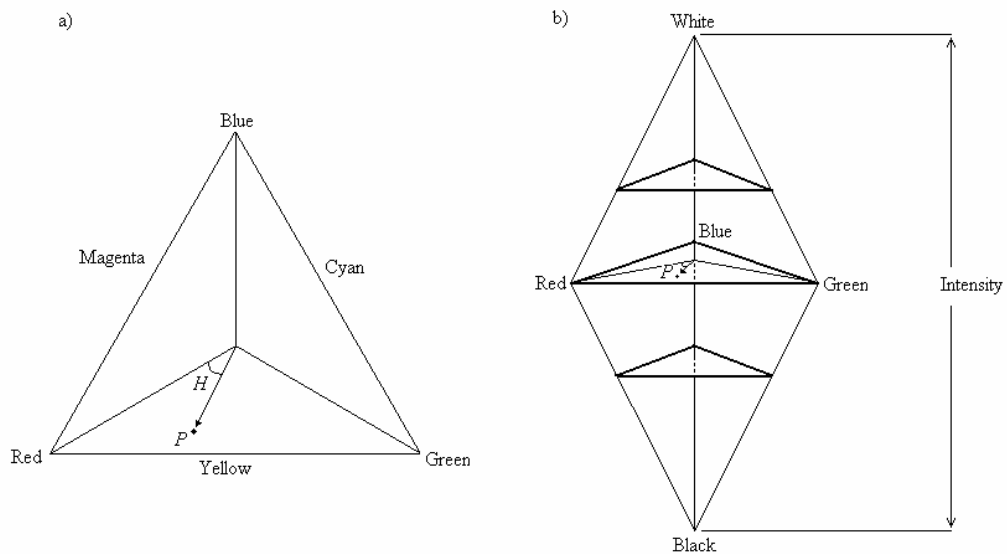
Figure 2.11. HSI colour model

The colour components of the HSI model are defined with respect to the colour triangle (Gonzales and Woods, 1993) presented in Figure 2.11a. In this figure, the hue value of colour point P is the angle of the vector shown with respect to the red axis. Thus when hue is 0°, the colour is red, when it is 60°, the colour is yellow and so on. The saturation is proportional to the distance between P and the centre of the triangle, so that the farther P is from the triangle centre, the more saturated is the colour. When an intensity component is added, the model shown in Figure 2.11a, a three dimensional, pyramid-like structure is obtained (Gonzales and Woods, 1993), as shown in Figure

2.11b. The hue value of colour point P is determined by its angle with respect to the red axis. Any point on the surface of this structure represents a purely saturated colour. The intensity in the model is measured with respect to a line perpendicular to the triangle and passing through its centre.

The RGB colour model defined with respect to the unit cube presented in Figure 2.10 can be converted to the HSI model shown in Figure 2.11 (Gonzales and Woods, 1993) according to the following equations:

$$H = \cos^{-1}\left\{\frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2+(R-B)(G-B)]^{1/2}}\right\} \quad (2.6)$$

$$S = 1 - \frac{3}{(R+G+B)}[\min(R,G,B)] \quad (2.7)$$

$$I = \frac{1}{3}(R+G+B) \quad (2.8)$$

The RGB model is a non-uniform colour model because the differences in this colour space do not directly correspond to the colour differences as perceived by humans (Wyszecki and Stiles, 1982). Also HSI colour space does not represent the colours on a uniform space. For this reason, the CIE has introduced perceptually uniform colour spaces, L*a*b* and L*u*v* (Wyszecki and Stiles, 1982). These models have been defined in order to make easier the evaluation of perceptual distances between colours (Del Bimbo, 1999).

Colour of rock

The colour of the rock has significance for the visual appearance of rock materials used in buildings as well as for the recognition of rock types. In recognition tasks, the colour is one of the most important characteristics for describing rock properties such as strength. Figure 2.12 shows sample images of typical Finnish rock types widely used in the rock industry. In colour-based rock description, the problem is similar to that in texture description; colour distribution is often non-homogenous. This can be seen, for example, in samples 3 and 4 in Figure 2.12, in which the red and black colours of the rock image are unevenly distributed. As a result, this kind of image cannot be characterized using, for example, the mean colour of the sample. However, statistical distributions such as histograms are able to describe these kinds of image.

When considering the visual properties of rock, selection of the colour space is essential. In addition to conventional RGB colour space, another colour space such as HSI, may often provide improved colour description since it is closer to human colour perception than is RGB space. In the case of bedrock inspection, images are sometimes also obtained using additional invisible light wavelengths which can be used to discriminate between certain minerals or chemical elements (Autio et al., 2004).
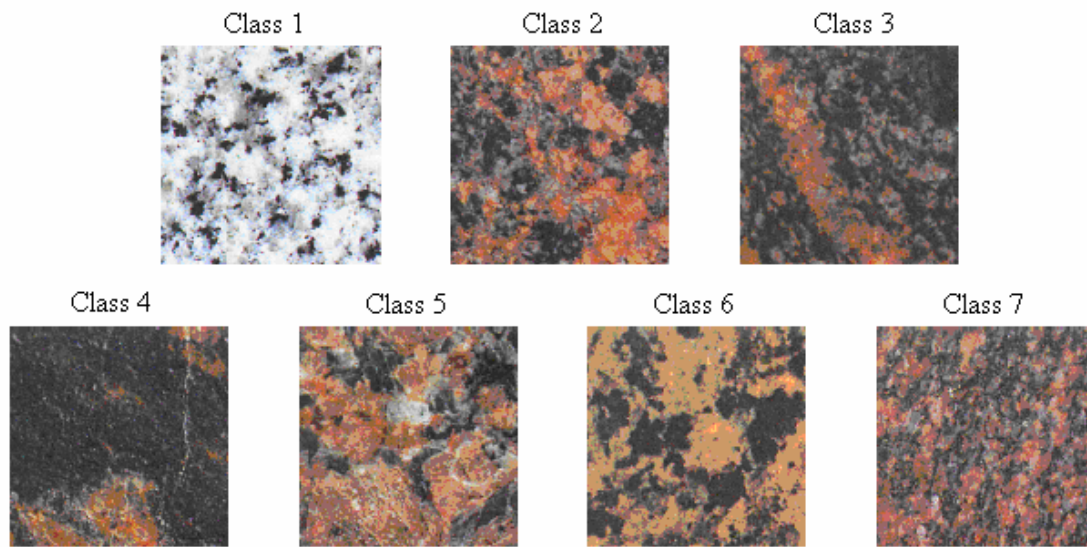
Figure 2.12. Example images of seven rock types used in rock industry.

2.4   Previous work in rock image analysis

The last decade has seen a growth of interest in the application of imaging methods to rock analysis. There have also been various kinds of studies on rock image analysis published in various conferences and journals and research has been carried out in a variety of fields such as bedrock investigation, quality control of rock, stone, and ceramic products, as well as mining. Lindqvist and Åkesson (2001) present a literature review of image analysis applied to geology in which the central areas are rock structure and texture analysis utilizing image analysis methods.

Luthi (1994) has proposed a technique for texture segmentation of borehole images using filtering methods. In Singh et al. (2004), texture features for rock image classification are compared. In this comparison, the best classification performance was achieved using Law's masks and co-occurrence matrices. The co-occurrence representation of rock texture was used in maximum-likelihood classification by Paclík et al. (2005). The textural features computed from the co-occurrence matrix were also used in rock image analysis in the study of Duarte and Fernlund (2005). In this study, entropy and textural correlation were found to be the most significant descriptors in the characterization of granite samples. Autio et al. (1999) employed co-occurrence matrix and Hough transform to describe the texture properties of rock. Lepistö et al. (2003a), employ contrast and entropy extracted from the co-occurrence matrix in the classification of non-homogenous rock samples. Tobias et al. (1995) proposed a texture analysis method based on the co-occurrences in the visual inspection of ceramic tile production. Texture directionality was used in the rock image classification of Lepistö et al. (2003b) in which directional histograms were formed for the rock samples using filtering with directional masks. The grain structure of rock was analyzed (Lepistö et al., 2004) by finding grains of a selected colour and size from the images. In this analysis morphological tools were also employed. Bruno et al. (1999) have analyzed the granular size of the rock texture using morphological image analysis tools.

In the production of rock and ceramic materials, colour-based image analysis tools have been utilized in several studies. One application area has been the quality control of ceramic tile production (Lebrun, 2001; Lebrun and Macaire, 2001). The use of colour analysis of ceramic tiles has also been studied by Boukovalas et al. (1997) in which the imaging of tiles and colour analysis in RGB colour space are discussed. Kukkonen et al. (2001) have applied spectral representation of colour to measure the visual properties of ceramic tiles. The tiles are classified based on their colour using self-organizing maps. Boukovalas et al. (1999), use RGB histograms used in the recognition of tile colour. Lebrun et al. (2000) have studied the influence of weather conditions on the rock materials using colour analysis. Mengko et al. (2000) have studied the recognition of minerals from the images. This recognition method uses colour features in RGB and HSI colour spaces. Lebrun et al. (1999), deal briefly with the surface reflection imaging and analysis of rock materials.

In mining, image analysis has been applied to such fields as rock material recognition and stone size estimation (Crida and Jager, 1994, 1996). In Salinas et al. (2005), rock fragment sizes are estimated by means of a computer vision system utilizing segmentation, filtering and morphological operations.

3 Texture and colour descriptors

In the literature, a wide variety of descriptors have been proposed for identifying image content. The descriptors are used to characterize the different properties in images such as textures, colours, and shapes. In this thesis the texture and colour properties of the rock images are selected as the characterising descriptors and the present Chapter provides an overview of them both.

3.1 Texture descriptors

Numerous techniques have been proposed for texture description. Tuceryan and Jain (1993) have divided texture description methods into four main categories: statistical, geometric, model-based, and signal processing methods (see Figure 3.1). These categories are briefly reviewed in this Section, though the central focus is on the signal processing methods.

3.1.1 Statistical methods

The use of statistical methods is common in the texture analysis. These techniques are based on the description of the spatial organization of the image grey levels. On the basis of the grey level distribution, it is possible to calculate several types of simple statistical features. When the features are defined in terms of single pixel values (such as mean or variance), they are called first order statistics. However, if the statistical measures are defined for the relationship of two or more pixel values, they are referred to as second- and higher order statistics.

Statistical methods have been used since the 1950s when Kaizer (1955) studied aerial photographs using autocorrelation function. An example of further use of correlation function is the work of Chen and Pavlidis (1983), in which correlation was applied to texture segmentation. Grey level co-occurrence matrix developed by Haralick (1973) has been a popular tool in texture analysis and classification. Co-occurrence matrix estimates the second order joint probability density functions $g(i,j \mid d, \Theta)$. Each $g(i,j \mid d, \Theta)$ is the probability of going from grey level i to grey level j, when the intersample spacing is d

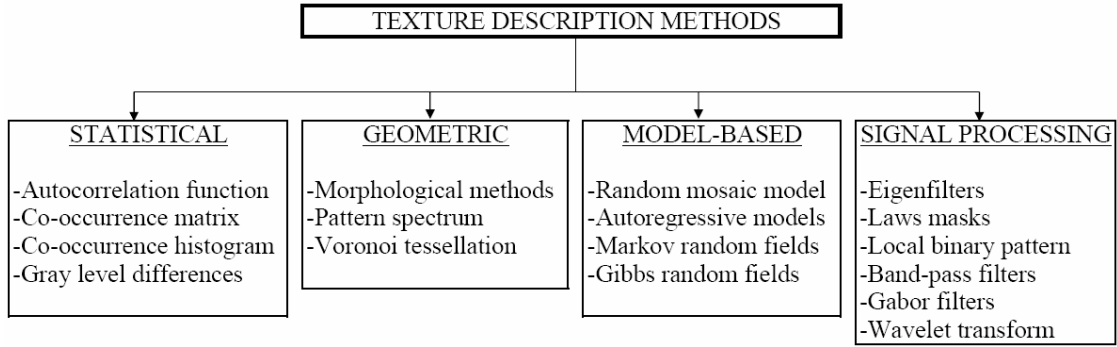| TEXTURE DESCRIPTION METHODS			
STATISTICAL	**GEOMETRIC**	**MODEL-BASED**	**SIGNAL PROCESSING**
-Autocorrelation function -Co-occurrence matrix -Co-occurrence histogram -Gray level differences	-Morphological methods -Pattern spectrum -Voronoi tessellation	-Random mosaic model -Autoregressive models -Markov random fields -Gibbs random fields	-Eigenfilters -Laws masks -Local binary pattern -Band-pass filters -Gabor filters -Wavelet transform

Figure 3.1. Main categories of texture description methods.

and the direction is Θ. These probabilities create the co-occurrence matrix $\mathbf{M}(i,j \mid d, \Theta)$. It is possible to extract a number of textural features from the matrix Haralick (1973). Contrast, entropy, and energy include commonly used texture features extracted from the co-occurrence matrix:

$$contrast = \sum_{i,j}(i-j)^2 \mathbf{M}(i,j \mid d,\Theta) \qquad (3.1)$$

$$entropy = -\sum_{i,j}\mathbf{M}(i,j \mid d,\Theta)\log\mathbf{M}(i,j \mid d,\Theta) \qquad (3.2)$$

$$energy = \sum_{i,j}\mathbf{M}(i,j \mid d,\Theta)^2 \qquad (3.3)$$

Valkealahti and Oja (1998) introduced a co-occurrence histogram that is a simpler variation of the co-occurrence matrix. Grey level difference method (Weszka et al., 1976) measures the differences between pixel grey level values at a certain displacement in the texture and presents these differences as a table. Based on this table, several textural features can be calculated. Unser (1986) presents the grey level differences as a histogram. Ojala et al. (2001) have used signed grey level differences instead of absolute differences. In this case, the mean luminance of texture has no influence on the texture description and local image texture is also better described than in the case of absolute differences.

3.1.2 Geometric methods

In geometric texture analysis methods, textures are characterized by means of texture primitives and their spatial organization. Texture primitives are extracted from the image using such techniques as edge detection algorithms or morphological tools. An example of the use of the morphological techniques in texture description is the work of Wilson (1989). Pattern spectrum developed by Dougherty et al. (1992) also uses morphological methods in texture description. Asano (1999) has presented an application of the pattern spectrum to extract the texture primitives from the image. In this application, the size and shape of the primitives are measured using morphological tools. The structure and organization of the primitives has also been characterized by means of Voronoi tessellation (Tüceryan and Jain, 1990). In these approaches, the texture primitives are combined into regions of similar textures that are referred as Voronoi polygons.

3.1.3   Model-based methods

Model-based texture analysis methods model the mathematical process describing the texture. Random mosaic model (Schacter et al., 1978; Ahuja and Rosenfeld, 1981) is a common method in this area in which the pixels in the texture are merged into regions based on their grey level distributions. On the basis of these regions, it is possible to calculate different statistical measures describing the texture. In texture characterization, time series models (Deguchi and Morishita, 1978) have also been proposed. These models include autoregressive (AR), moving average (MA), and their combination (ARMA). Mao and Jain (1992) have applied simultaneous autoregressive models (SAR) to texture segmentation and classification. In the SAR model, the relationships between texture pixels and their neighbourhoods are modelled using statistical parameters. These relationships are also utilized in Markov Random fields, in which texture is regarded as an independent stationary process. The random fields are used in unsupervised texture segmentation by both Manjunath and Chellappa (1991), and Kervrann and Heitz (1995). The model-based texture analysis methods also include Gibbs random fields (Besag, 1974). Elfadel and Picard (1994) have further developed the Gibbs model and presented a new texture feature, Aura feature. In addition, the Wold features presented in (Liu and Picard, 1996) are based on random fields.

3.1.4   Signal processing methods

Methods based on signal processing are nowadays popular tools in texture analysis. In most of these methods the texture image is submitted to a linear transform, filter, or filter bank, followed by some energy measure (Randen and Husøy, 1999). The first filtering-based approaches were introduced in the beginning of the 1980s. Eigenfilters (Ade, 1983) and Law's masks (Law, 1980) are some of the early filtering approaches. In the Eigenfilters, a covariance matrix is defined for the 3x3 neighbourhoods of each texture pixel. Texture identification is based on the eigenvalues calculated from the covariance matrices. In Law's method, convolution masks of different orientations are applied to the texture image. In the 1990s Ojala et al. proposed a new spatial filtering method, local binary pattern (LBP). In the LBP method, texture properties are characterized by means of the spatial organization of the texture neighbourhoods (Ojala et al., 1996). Based on the neighbourhood, a LBP number is defined for each texture pixel. The LBP numbers are presented as a histogram that describes the texture.

Methods based on Fourier transform utilize the frequency distribution of texture. This field has been researched since the mid 1970s when Dyer and Rosenfeld (1976) used Fourier transform in texture description. The Fourier transform of an image $f(x,y)$ can be defined as:

$$F(u,v) \equiv \int\limits_{-\infty}^{\infty} \int e^{-i2\pi(ux+vy)} f(x,y) dx dy \qquad (3.4)$$

The Fourier power spectrum is $|F|^2 = FF^*$, in which * denotes complex conjugate (Weszka et al., 1976). In practice, the images are in digital form and therefore, discrete Fourier transform is employed (Dyer and Rosenfeld (1976)):
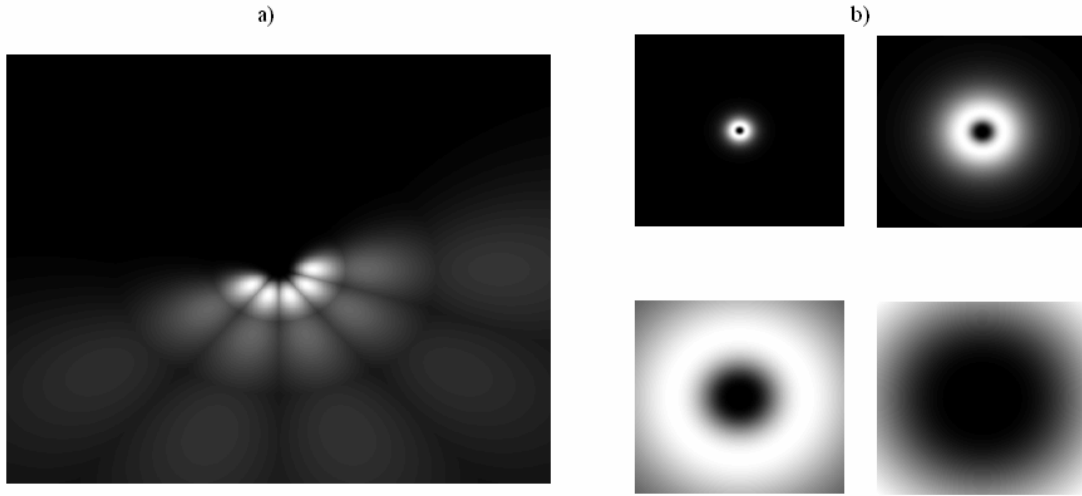
Figure 3.2. Different texture filters. a) A set of Gabor filters at three scales and five orientations b) four ring-shaped Gaussian filters at different scales.

$$F(u,v) = \frac{1}{n^2} \sum_{x,y=0}^{n-1} f(x,y) e^{-i2\pi(ux+vy)/n} \qquad (3.5)$$

where f and F are n by n arrays (assuming that the images are square-shaped). In this case, the power spectrum is also of the form $|F|^2$. In texture analysis, the Fourier power spectrum can be utilized in several manners. For example, the radial distribution of the spectrum values is sensitive to texture coarseness or granularity in f. Hence, the granularity can be analyzed by selecting ring-shaped regions from the spectrum. Similarly, the angular distribution of the spectrum is sensitive to the directionality of texture in f (Weszka et al., 1976). Coggins and Jain (1985) employed ring- and wedge-shaped filters to extract features related to texture coarseness and directionality. The purpose of the filtering approaches is to estimate the energy in the spectrum at a specific local region. One of the most popular approaches in this area is Gabor filtering. Gabor filter is a Gaussian-shaped local band-pass filter that covers a certain radial frequency and orientation. It is typical that an image is filtered using a bank of Gabor filters of different orientations and radial frequencies, often referred to as scales. An example of this kind of approach is the work of Jain and Farrokhia (1991), in which Gabor filter banks were used to texture segmentation. Figure 3.2 shows different texture filters. Manjunath and Ma (1996) suggest simple texture features for texture image retrieval. In their method, the mean and standard deviation of the transform coefficients at each scale and orientation are used as texture features. Bigün and du Buf (1994) use complex moments of the local power spectrum as texture features. In the work of Kruizinga and Petkov (1999), Grating cell operators for Gabor filtering are used. The Grating cells are selective to orientation but they do not react to single lines or edges in the texture. These approaches have also given good results when compared to other Gabor methods in texture discrimination and segmentation in (Grigorescu et al., 2002). A review and comparison of the various filtering methods in texture classification is provided by Randen and Husøy (1999) who

conclude that no single filtering method can outperform all others with every kind of image.

In addition to Fourier transform, wavelet transform (Chui, 1992) has received major research interest in recent years. Wavelet transform approaches use filter banks with particular filter parameters and sub-band decompositions (Randen and Husøy, 1999). Mallat (1989) was the first to apply wavelet transform to texture characterization since when wavelet-based texture description has been a popular research area. Wavelet packet transform (Laine and Fan, 1993) has also been widely used in texture description. Wavelet frame introduced by Unser (1995) is a translation invariant version of wavelet transform. It is an over-complete wavelet representation and is more effective in texture edge localization than other wavelet-based approaches (Randen and Husøy, 1999).

3.2   Colour descriptors

Colour distribution is a typical characteristic used in the image classification and can be described using statistical methods. Different moments are examples of simple statistical measures. Stricker and Orengo (1995) have used colour moments to describe image colour distribution. The moments include mean ($\bar{x}$), variance ($\hat{\sigma}_x^2$) and skewness (S):

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3.6}$$

$$\hat{\sigma}_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{3.7}$$

$$S = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{3/2}} \tag{3.8}$$

The histogram is probably the most commonly used statistical tool for the description of image colour distribution. It is a first order statistical measure that estimates the probability of occurrence of a certain colour in the image. Hence, the histogram is a normalized distribution of pixel values. If the number of colour levels in an image is n, the histogram H can be expressed as a vector of length n. The i:th component of the vector is defined as:

$$H(i) = \frac{N_i}{N} \qquad i = 0, 1, 2, ..., n-1 \tag{3.9}$$

where N refers to the total number of pixels and N_i to the number of pixels of colour i. Image histogram has been widely used in the description of image colour content. In (Swain and Ballard, 1991), a colour histogram is used as a feature vector for describing the image content. The benefit of the image histogram is its computational lightness and low dimensionality, which is equal to the number of colour levels in the image. The main drawback is that the histogram ignores the spatial relationships of the colours in the image. For this reason, a variety of second order statistical measures have been

introduced for image description. Several statistical measures utilize the correlation function in image description. Huang et al. (1997) introduce a correlogram that describes the relationships of pixel pairs at a distance d in the image. The correlogram is formed in the same manner as the co-occurrence matrix, but in the case of the correlogram it is usual that several values of d are used. If an image has N colour levels, the size of the correlogram is N^2 at each value of d. Hence, the correlogram is computationally a relatively expensive method and because of this, it is usual that the autocorrelogram is employed instead. The autocorrelogram (Huang et al., 1997) is a subset of correlogram that gives the probability that the pixels at distance d in the image are of the same colour. The size of the autocorrelogram is N.

In addition to statistical methods for colour description, other colour description methods have also been proposed. Colour naming system is an approach in which basic colour names are used to describe the colour content of the images (Del Bimbo, 1999).

3.2.1 Coloured textures

The common texture analysis methods have been developed for grey level images. However, the colour that is often present in the texture image is also an important characteristic describing the image content. In many cases, it is practical to present the texture and colour properties of an image using a single descriptor. In the case of coloured textures, it is usual that different texture analysis methods, such as filters, are employed. Thai and Hailey (2000) propose a spatial filtering method that is based on Fourier transform in colour texture analysis. This method uses RGB colour space. Palm et al. (2000) have used HSI colour space in colour texture analysis. They present hue and saturation components as polar coordinates and apply Fourier transform to them. In colour texture analysis, the selection of the colour space is essential. Paschos (2001) has compared RGB, L*a*b*, and HSI colour spaces in colour texture analysis that is based on Gabor filtering. The experimental results indicate that HSI space gives the best classification results.

In addition to the filtering-based methods, statistical methods are also employed in colour texture analysis. In the covariance method (Lakmann, 1997), a covariance matrix is computed for the colour channels. Paschos (1998) has studied the coloured textures using the chromaticity of colour. The idea behind his approach is that the correlation function has been calculated for the chromatic components of the image. Paschos and Radev (1999), make use of chromaticity-based moments in the classification of coloured textures. Valkealahti and Oja (1998b) have applied the statistical texture analysis methods for coloured textures. In their approach, multidimensional co-occurrence histograms have been used for the colour texture analysis. Co-occurrence matrix can also be calculated for coloured texture images. In the work of Shim and Choi (2003), co-occurrence matrix is used to describe the spatial relationships between hue levels in the image.

4 Image classification

The previous Chapter reviewed the various techniques presented in the literature for image content description that can be used in image recognition and classification. However, in addition to good descriptors, an effective image classification system needs an appropriate classifier. The present Chapter concerns the field of the classification, and beginning with the theoretical background, the classification problem is discussed in some detail. This topic is continued with a consideration of Bayesian decision theory and nonparametric classification. Finally, the special character of image classification problem is examined.

4.1 Classification

A pattern to be classified consists of one or several features. In image classification, it is usual that a pattern is characterized using a feature vector containing n features. Such a vector is often referred as a feature vector of n dimensions. If f_i represents the ith feature, the vector can be expressed as $\mathbf{S}=(f_1, f_2,\ldots, f_n)^T$. This way, the feature vector represents the pattern in n dimensional feature space (Duda et al., 2001). A pattern class is a family of patterns that share some common properties (Gonzales and Woods, 1993). For example, in colour-based image classification, images sharing similar colour properties belong to the same class. This means that the colour histograms of n bins can be used as colour descriptors, and the images that have similar histograms are assigned to same classes. The classes can be denoted ω_1, ω_2, …, ω_m, where m is the number of these classes. Hence, the problem in classification is to assign the unknown sample pattern to one of the classes. It should be noted that in the classification problems discussed in this Chapter, only supervised classification is discussed. This means that the classes are predefined. In classification problems in general, the patterns in the feature space should be assigned to classes as accurately as possible. For this purpose, several types of classification methods have been proposed.
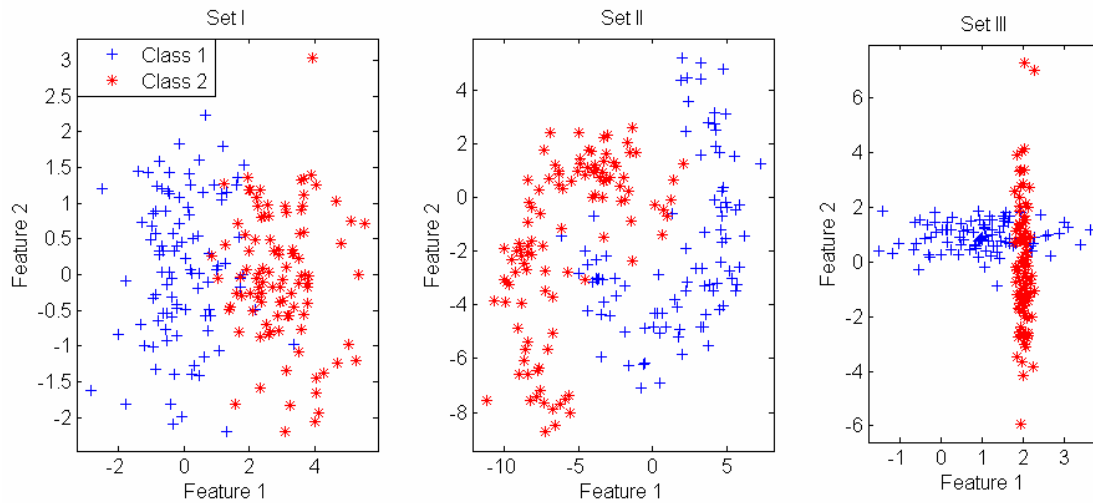
Figure 4.1. Three classification examples containing two classes in two-dimensional feature spaces.

Figure 4.1 shows three examples of two-class classification problems in which the patterns are presented in a two-dimensional feature space[2]. In all the examples, the number of patterns is 100 in both classes. In the first data set, the classes are spherical, Gaussian distributed datasets with the same variance. As presented in the figure, the means of the classes in the first dimension are relatively distant, which indicates that the classes are not significantly overlapping. The classes in data set II obey banana-shaped distribution, which is a more demanding classification task. This is because the means of the classes are close together. As shown in this example, the pattern classes are not always spherically shaped in the feature space. Instead, their shapes can be complicated, even if the classes are not significantly overlapping. The third dataset has two classes which are clearly overlapping in the feature space. The pattern distributions obey Highleyman classes (Duin et al., 2004).

Validation

The set of known samples in the supervised classification is usually called the training set. The selection of these samples should occur randomly from the population (van der Heijden et al., 2004). These samples are used as prototype samples in the classification, and the *a priori* knowledge of their classes is used in the classification of the unknown samples. The unknown samples form a testing set. It is usual that roughly one third of the available data is used as training set, and the remaining two thirds serves as testing set. The classification performance can be estimated by defining the error rate or the classification rate, which correspond respectively to the number of misclassified or correctly classified samples in the testing set. These values are often presented as percentages.

The estimation of classification performance is often referred to as validation. Instead of using the above mentioned division into training set and testing set, other validation methods for classification also exist (van der Heijden et al., 2004). In the cross-validation

---

[2] The examples presented in figures 4.1-4.5 have been generated with PRTools 4, a Matlab toolbox for pattern recognition (Duin et al., 2004).

38

method, the available data is randomly partitioned into L equally sized subsets. Each subset is used as a test set in turn whereas the rest of the data is used as a training set. The final error rate is evaluated as an average of these L classifications. In the case of leave-one-out method, only one sample is regarded as an unknown sample in turn, and all the other samples in the data set serve as training data. This way, all the samples in the dataset are classified. These two methods, however, are computationally expensive with large datasets.

4.1.1 Pattern classification in the feature space

Several classification methods use decision (or discriminant) functions (Duda et al, 2001). For m pattern classes, the problem is to find m decision functions $d_1(\mathbf{S})$, $d_2(\mathbf{S})$, …, $d_m(\mathbf{S})$. If a pattern $\mathbf{S}$ belongs to class ω_i, then

$$d_i(\mathbf{S}) > d_j(\mathbf{S}) \quad j = 1, 2, ..., m; \ j \neq i \tag{4.1}$$

Hence, an unknown pattern $\mathbf{S}$ belongs to the ith pattern class if $d_i(\mathbf{S})$ yields the largest numerical value. The decision boundary that separates the class ω_i from ω_j is given by values of $\mathbf{S}$ for which $d_i(\mathbf{S}) = d_j(\mathbf{S})$. Alternatively, the decision boundary for the values of $\mathbf{S}$ can be defined by:

$$d_i(\mathbf{S}) - d_j(\mathbf{S}) = 0 \tag{4.2}$$

The decision boundaries in the feature space can be found in several different ways. It is usual that some prototype patterns are used as training data. The classes of these patterns are known in advance, and therefore, the unknown patterns can be compared to the prototype patterns.

Distance metrics

In order to make a comparison between two patterns, a method is needed for evaluating the similarity (or dissimilarity) between them. For dissimilarity measurement, several types of distance functions have been proposed, known also as distance metrics (Santini and Jain, 1999; Duda et al., 2001). Duda et al. (2001) have presented four properties for distance metrics D between vectors $\mathbf{a}$ and $\mathbf{b}$:

1. Nonnegativity: $D(\mathbf{a},\mathbf{b}) \geq 0$
2. Reflexivity: $D(\mathbf{a},\mathbf{b}) = 0$, if and only if $\mathbf{a} = \mathbf{b}$
3. Symmetry: $D(\mathbf{a},\mathbf{b}) = D(\mathbf{b},\mathbf{a})$
4. Triangle inequality: $D(\mathbf{a},\mathbf{b}) + D(\mathbf{b},\mathbf{c}) \geq D(\mathbf{a},\mathbf{c})$

A number of different distance metrics for different kinds of data has been proposed and a review of these is presented by Santini and Jain (1999). Minkowski metrics is one general class of metrics for n-dimensional patterns:

$$L_k(\mathbf{a},\mathbf{b}) = \left(\sum_{i=1}^{n} |a_i - b_i|^k \right)^{1/k} \tag{4.3}$$
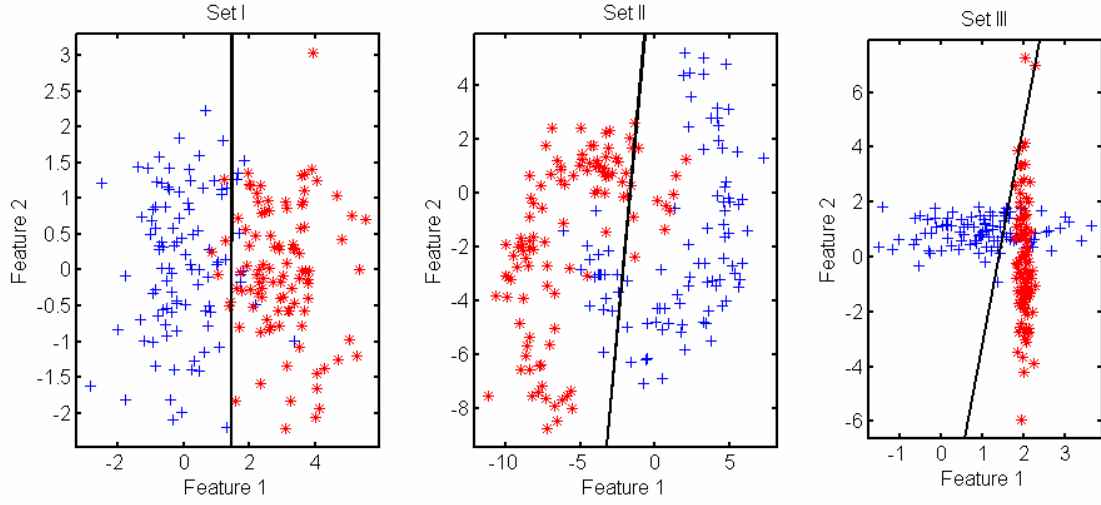
Figure 4.2. Minimum distance classification in three datasets.

Minkowski metrics is also referred to as L_k norm. The most popularly used L_k metrics are L_1 norm, which is also called Manhattan distance or city block distance:

$$L_1(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{n} |a_i - b_i| \tag{4.4}$$

and L_2 norm, also known as Euclidean distance:

$$L_2(\mathbf{a}, \mathbf{b}) = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2} \tag{4.5}$$

Minimum distance classifier

A simple method for evaluating the decision boundaries in the feature space is to use a minimum distance classifier (van der Heijden et al., 2004), in which each pattern class is represented by a prototype vector that is the mean vector of the patterns in this particular class. Each pattern is then assigned to the class represented by their nearest prototype vector in the feature space. The minimum distance classifier works well when the distance between means is large compared to the spread or randomness of each class with respect to its mean. This kind of approach, however, is not effective when the classes are overlapping in the feature space. Figure 4.2 demonstrates the minimum distance classification in the three datasets shown in Figure 4.1. In this example, the decision boundaries are drawn on the basis of minimum distance to the mean of both classes. It is obvious that in the case of simple spherical distributions with only small overlap, the minimum distance method is usable. However, in the case of banana-shaped classes of dataset II, this classification method has lower performance. This is because the method does not take into account the shapes of the pattern classes, only their mean is significant. Minimum distance classifier is also unable to classify the overlapping datasets of set III.

The reason for misclassification is that the means of the classes do not provide any information on the spatial organizations of the patterns in the feature space.

4.1.2 Bayesian decision theory

Bayesian decision theory (Duda et al., 2001) approaches the classification problem by means of probabilities. Hence, it is assumed that there is some *a priori probability* that the sample belongs to one particular class. In the case of a two-class problem, classes ω_1 and ω_2, have probabilities $P(\omega_1)$ and $P(\omega_2)$ which sum to one. In the general case with m classes, it is possible to write:

$$\sum_{i=1}^{m} P(\omega_i) = 1 \tag{4.6}$$

These probabilities[3] reflect the prior knowledge of how probable these two classes are for the unknown sample. If this is the only available information, it is reasonable to use the following simple decision rule: Decide ω_1 if $P(\omega_1) > P(\omega_2)$; otherwise decide ω_2. It is obvious that this kind of classification is not useful, because the decision is always the same despite the fact that both classes can be present within the patterns. However, there is normally additional information available than only probabilities. The features describing the unknown sample are such that they are selective to particular categories and, therefore, can be used to describe the sample. Let us assume that we use a measurement f as a feature describing the sample. If f is considered as a continuous random variable, its distribution can be expressed as $p(f|\omega_i)$. This is referred as a *conditional probability density function* and it expresses the probability that the class is ω_i at a certain value of f.

Let us assume that we know the *a priori probabilities* $P(\omega_i)$ and conditional probability densities $p(f|\omega_i)$ for i=1, 2. Also the feature value f for an unknown sample is known. Based on the two probabilities provided, it is possible to form a joint probability density of finding a pattern that is of class ω_i and has a feature value f. This can be presented in two ways: $p(\omega_i | f) = P(\omega_i | f) p(f) = p(f | \omega_i) P(\omega_i)$, (Duda et al., 2001). Based on this, the Bayes formula:

$$P(\omega_i | f) = \frac{p(f | \omega_i) P(\omega_i)}{p(f)} \tag{4.7}$$

can be defined (Duda et al., 2001). Where, in the case of two classes:

$$p(f) = \sum_{i=1}^{2} p(f | \omega_i) P(\omega_i) \tag{4.8}$$

In the case of multiclass problem, this can be written as:

---

[3] In this study, P and p denote probability mass function and probability density function, respectively.

$$p(f) = \sum_{i=1}^{m} p(f \mid \omega_i) P(\omega_i) \tag{4.9}$$

According to the Bayes formula, it is possible to convert the *a priori probability $P(\omega_i)$* to *a posteriori probability $P(\omega_i \mid f)$*, which is the probability of class ω_i given that f has been measured (Duda et al., 2001). This *a posteriori probability* can be used to make a classification decision. It is obvious that for an observation f for which $P(\omega_1 \mid f)$ is greater than $P(\omega_2 \mid f)$, the decision is ω_1. This kind of decision minimizes the classification error, and it is known as the Bayesian decision rule (Duda et al., 2001).

It is also possible to define the decision boundaries using the Bayesian decision rule. The decision boundaries can be obtained from the discriminant functions presented in equation (4.1). In the Bayesian classification, the discriminant functions are selected such that they minimize the classification error (Duda et al., 2001). Figure 4.3 illustrates the Bayesian classification in the three datasets. In these classification examples, two kinds of Bayesian classifiers are employed. The first one, Bayes Normal-1, refers to linear Bayes normal classifier and the second one, Bayes Normal-2 refers to quadratic Bayes normal classifier (Duin et al., 2004). The linear Bayes normal classifier assumes that the classes are spherical in the feature space and they have the same variance (Duda et al., 2001). The classifier makes a linear decision boundary between the classes. Therefore, the decision boundary is like that of the minimum distance classifier. In the case of a quadratic classifier, the decision surface is a hypersurface in the feature space. In a two-dimensional case, the decision boundary is a quadratic curve, such as an ellipse or hyperbola. The difference between these two classifiers is clearly visible in the classification of dataset III, in which the quadratic classifier is able to distinguish between the classes by means of a hyperbola-shaped decision boundary. However, the decision boundaries of these two classifiers are relatively similar in the first two datasets. Figure 4.4 shows the scatter diagrams with contour plots of the conditional probability densities of quadratic Bayesian normal classifier. The probability densities are also shown as three-dimensional surfaces.
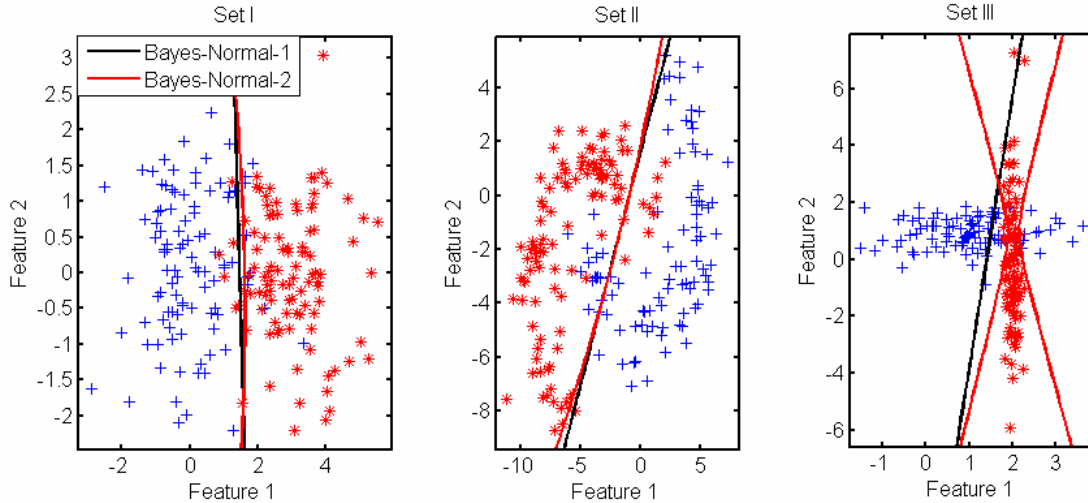

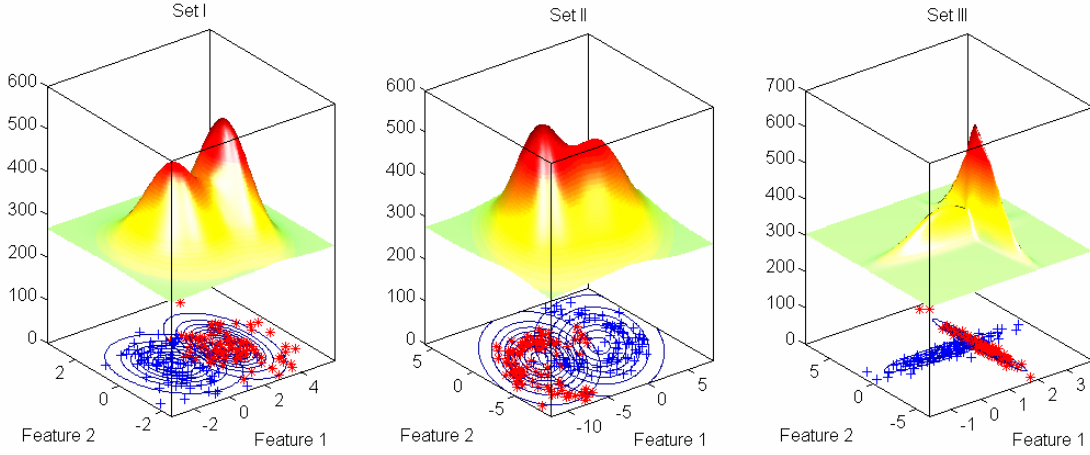Figure 4.3. Bayesian classification in three datasets.

Figure 4.4. The scatter plots of the classes with contour plots and three-dimensional surfaces describing the conditional probability densities of quadratic Bayesian normal classifier.

4.1.3   Nonparametric classification

In the Bayesian decision theory, the basic assumption is that the prior knowledge about the probability distributions is available. In the case of nonparametric classification, these distributions are not used (van der Heijden et al., 2004).

Parzen windows

The minimum distance classifier introduced in Section 4.1.1 selects one prototype pattern to represent the whole class. In contrast to this approach, it is also possible to estimate the densities of the patterns representing different classes in the region $\mathcal{R}$ around the unknown pattern. In this kind of density estimation, Parzen window estimates (Duda et al., 2001) can be employed. In the case of two-dimensional feature space, the region is a circle drawn around the unknown sample. Generally, in an n-dimensional space, this region is a hypercube of n dimensions with edge length h. The volume of the hypercube is then $V = h^n$ and it can be defined by means of a window function φ. Let us assume that we have a set of N patterns, $\mathbf{S}_1, \mathbf{S}_2, \dots \mathbf{S}_N$. Each pattern has n dimensions. If the hypercube is centred at $\mathbf{S}$, $\varphi((\mathbf{S}\text{-}\mathbf{S}_i)/h)$ is equal to unity if $\mathbf{S}_i$ falls within the hypercube. The number of samples in this hypercube is therefore given by:

$$k = \sum_{i=1}^{N} \varphi\left(\frac{\mathbf{S} - \mathbf{S}_i}{h}\right) \tag{4.10}$$

Based on this, it is possible to estimate the probability density function (Duda et al., 2001):

$$p(\mathbf{S}) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{V}\varphi\left(\frac{\mathbf{S} - \mathbf{S}_i}{h}\right) \tag{4.11}$$
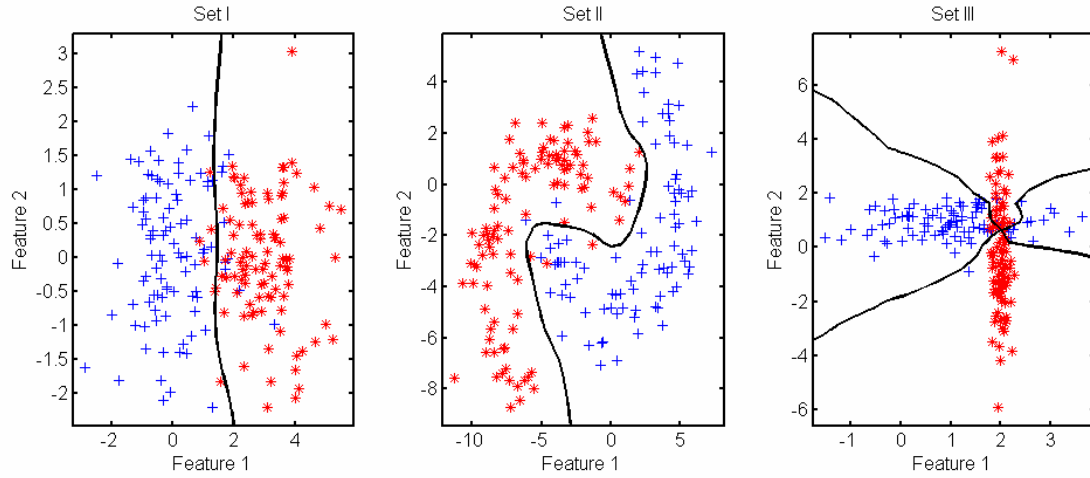
Figure 4.5. Parzen window classifications for three datasets.

Figure 4.5 presents the Parzen window classification in three datasets. In this experiment, optimum size of window function has been estimated for each classification (Duin et al., 2004). The obtained decision boundaries show that this kind of nonparametric classification approach is able to distinguish between the classes in all three datasets.

Nearest neighbour classifiers

A problem with a Parzen window estimate is the selection of the region size. However, this size can be decided by growing the region (or a cell) around the sample pattern **S** until it captures k samples in the neighbourhood. This is known as k-nearest neighbour classifier (k-NN classifier). Hence, the nearest neighbour classifiers use the training set directly and do not explicitly estimate the probability densities. If the density around **S** is high, this region is relatively small. If the density is low, the region grows larger, but stops soon after finding an area of higher density (Duda et al., 2001). In both cases, the density estimation can be written as:

$$p_n(\mathbf{S}) = \frac{k/N}{V}$$
(4.12)

It is also possible to estimate the *a posteriori probability* $P(\omega_i \,|\, \mathbf{S})$, for a k-NN classifier. As presented above, a cell of volume V is placed around **S**, and k samples are captured from the neighbourhood. If k_i of these samples turn out to be of class ω_i, the estimate for joint probability $p(\mathbf{S}, \omega_i)$ can be written as:

$$p(\mathbf{S}, \omega_i) = \frac{k_i/N}{V}$$
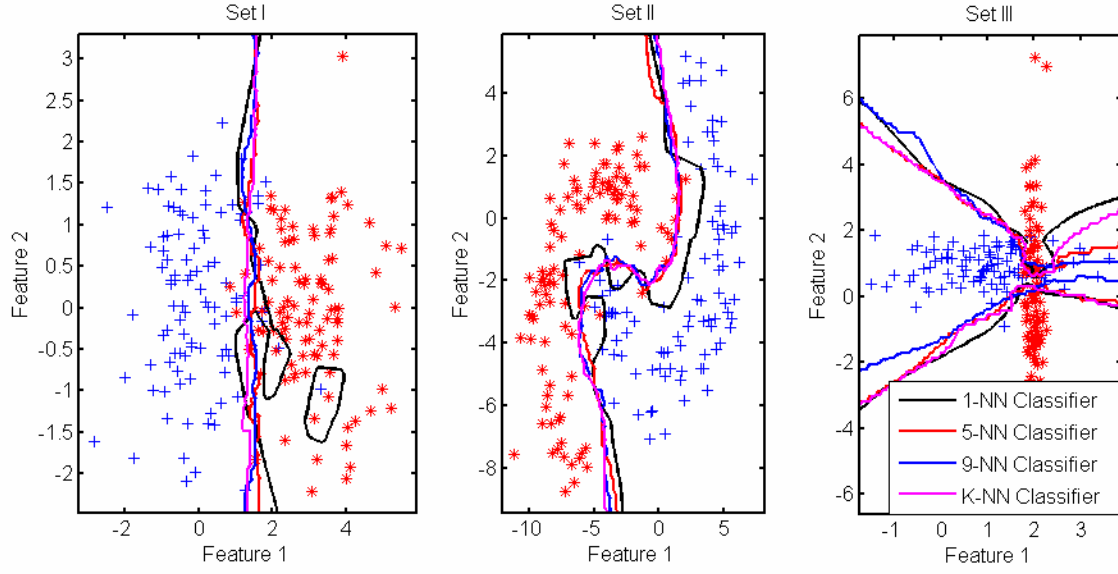(4.13)

and an estimate for $P(\omega_i |\, \mathbf{S})$ is:

Figure 4.6. *k*-nearest neighbour classification using three values for *k*=1,5, and 9. In addition, the optimized number of nearest neighbours, *K*, is used.

$$P(\omega_i \mid \mathbf{S}) = \frac{p(\mathbf{S}, \omega_i)}{\sum_{j=1}^{m} p(\mathbf{S}, \omega_j)} = \frac{k_i}{k} \qquad (4.14)$$

(Duda et al., 2001). In other words, the fraction of the samples in the cell labelled as ω_i can be considered as the *a posteriori probability* for this particular class. Consequently, the class that is the most frequently represented in the cell is selected. Figure 4.6 shows *k*-NN classification of the three datasets. In these classifications, the numbers of nearest neighbours (*k*) are selected to be 1, 5, and 9. In addition, the value of *k* has been optimized by minimizing the leave-one-out error in classification (Duin et al., 2004). This optimized classifier is marked as *K*-NN classifier in the figure. The optimized values of *k* are 30, 11, and 3, in the datasets I, II, and III, respectively.

4.2 Image classification

Real-world image classification is a relatively complicated problem. A typical characteristic of image classification tasks is that the image content is expressed in terms of numerical features. These features should describe the visual appearance of the image as accurately as possible. The feature values usually form relatively high dimensional feature vectors that are referred to as image descriptors. Some of the descriptor types were introduced in Chapter 3. It is also very common that the descriptors are overlapping in the feature space and this makes the classification challenging. In the present Section, the special challenges of image classification are discussed.

4.2.1 Dimensionality

Since the descriptors are often high dimensional, the feature spaces are difficult to visualize. The high dimensionality of the feature vector may also yield a decreased classification result, especially with a small amount of training data. This is known as "the curse of dimensionality" (Duda et al., 2001), which means that the demand for a large number of training data samples increases exponentially with increasing dimensions of the feature space.

The number of the dimensions in an image classification task is dependent on the descriptor types used. For example, the colour histograms are often evaluated for the whole range of colours in the image. Hence, the typical number of bins in the histogram is 256. In addition, if the histograms are defined for three colour channels[4], the number of dimensions is 768. When the second order statistical measures, such as correlograms, are employed as descriptors, the dimensionality can be even higher.

With texture descriptors, the dimensionality may also be a problem. The texture features that are presented in the form of a histogram are typically high dimensional. Gabor filtering with multiple scales and orientations is also able to produce a remarkable number of coefficients. Nevertheless, this problem can be solved by using, for example, mean or standard deviation of the coefficients as texture descriptors (Manjunath and Ma, 1996). The features extracted from the co-occurrence matrix are also all one dimensional (see Section 3.1.1). However, some of the texture descriptors of recently introduced MPEG-7 standard (Manjunath et al., 2002) also have quite high dimensionality. For example, the number of dimensions in *Homogenous Texture* and *Edge Histogram* descriptors is 62 and 80, respectively.

Reducing dimensionality

If high dimensionality of the feature space is a problem, it is possible to decrease the number of the dimensions. This is also known as feature reduction (van der Heijden et al., 2004). In feature reduction, features with low significance in the classification are removed.

There have been several techniques proposed for feature reduction. Probably the best known is principal component analysis (PCA) which is an unsupervised method for selecting the "right" features from the data (Duda et al., 2001). In this feature reduction principle, the high dimensional feature vector $\mathbf{S}_n$ of n dimensions is transformed to a k-dimensional vector $\mathbf{S}_k$. This is carried out by calculating an n-dimensional mean vector $\boldsymbol{\mu}$ and n x n covariance matrix $\boldsymbol{\Sigma}$ for the dataset. Next, the eigenvectors and eigenvalues are computed, and k eigenvectors having the largest eigenvalues are selected. Then, an n x k matrix $\mathbf{A}$ is formed. The columns of $\mathbf{A}$ contain the k selected eigenvectors. Finally, the data is represented by projecting the data onto the k-dimensional subspace according to:

$$\mathbf{S}_k = \mathbf{A}^t(\mathbf{S}_n - \boldsymbol{\mu}) \qquad (4.15)$$

(Duda et al., 2001). The PCA yields a k-dimensional linear subspace of the feature space. Other approaches to feature reduction include nonlinear component analysis (NLCA), which is more suitable for data with complicated interactions of features (Duda et al.,

---

[4] For RGB or HSI color spaces, for example.

2001). Multidimensional scaling (MDS) is sometimes also a better alternative than PCA (van der Heijden et al., 2004), particularly in cases when the data needs to be inspected in two or three dimensions. In such cases PCA often discards too much information (van der Heijden et al., 2004).

4.2.2 Overlapping classes

In most of the examples presented in the literature concerning classification problems, the classes form relatively uniform and separated clouds in the feature space. This is also the case with the examples presented at the beginning of this Chapter. However, in several kinds of real-world classification problems, the situation is more complicated. In image classification, the visual features employed in the classification task can be spread in the feature space in many difficult ways. For example, the image classes are often overlapping in the feature space. This is typical for images that have relatively similar content, but still belong to different classes. This may occur when the textures or colours of the image classes are similar.

However, non-homogenous content in an image may also cause difficulties when an image belongs to class A, but some regions in this image have, for example, colour or texture that is typical of class B. In this case the feature values of the image are located between these classes in the feature space. Another typical situation may occur when images of the same class have variations in their colours or textures. In such cases, the feature values may be scattered over the larger region in the feature space.

The problem with overlapping image classes can be seen in Figure 4.7, in which feature values of a set of rock images are shown in two dimensions. The image set contains 336 rock samples, which are divided into four classes by geologists. In the first figure, the features are contrast and entropy calculated from the grey level co-occurrence matrix. In the second figure, the features are mean hue and mean grey level of the images. Figure 4.8 presents three sample images of each class in the image set. The classification problem presented in this example is typical of rock image classification: the feature values are overlapping in the feature space, and they are also scattered over a large area in this space. This can be seen in the hue-intensity plot, in which class 1 samples are overlapping with all the other classes. It should be noted that in real image classification problems the number of dimensions is usually higher than two as in this example.
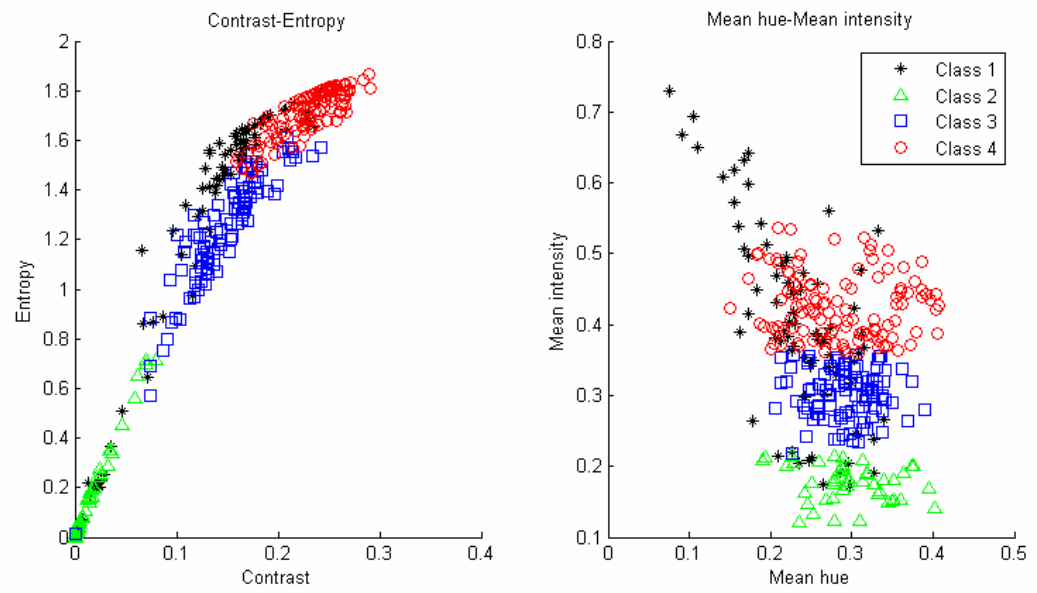
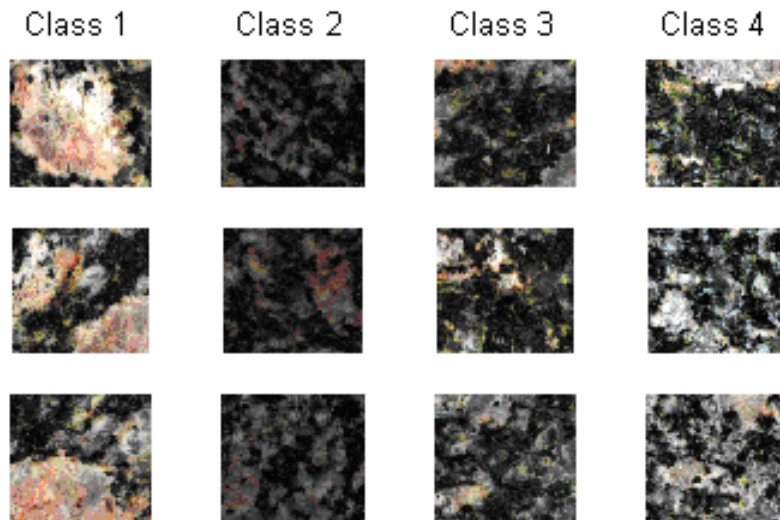Figure 4.7. Features calculated from the four rock image classes presented in two feature spaces.

Figure 4.8. Three sample images from each class.

5 Combining classifiers

A traditional approach in classification is the use of a single classifier, which assigns a class label for each feature vector describing the image content. It was noted in the previous Chapter that the decision functions produced by different classification principles differ from each other. This makes the classification accuracy somewhat varied. Furthermore, with real-world images it is common that the feature patterns are non-homogenous, noisy, and overlapping, which may cause variations in the decision boundaries of different classifiers. For these reasons, different classifiers may classify the same image in different ways. However, it has been observed that a consensus decision of several classifiers can yield improved performance compared to individual classifiers (Alkoot and Kittler, 1999; Kittler et al., 1998; Kuncheva, 2002). This is motivated by the fact that the features or classifiers of different types are able to complement one another in classification performance (Ho et al., 1994).

This Chapter considers the area of classifier combinations. The idea behind classifier combinations is to combine multiple classifiers into one classification system (Barandela et al., 2003). Classifier combinations[5] have become a rapidly growing field of research. The number of publications in this area is very large with more than a hundred of journal articles and an huge number of conference papers. Indeed entire conferences have been devoted to the topic (see Roli et al., 2004; Oza et al, 2005, for example). Classifier combinations have been applied to several kinds of classification tasks such as facial recognition (Lu et al., 2003), handwritten characters or numerals (Xu et al., 1992; Cao et al., 1995), person identification (Brunelli and Falavigna, 1995), speech recognition (Bilmes and Kirchhoff, 2003), fingerprint verification (Jain et al., 1999), to name just a few.

---

[5] In the literature, this research area is also referred to as classifier ensembles, multiple classifier systems, multiple expert fusions, mixtures of experts, committees of learners.

5.1 Base classification

According to Kittler et al (1998), the main motivation behind classifier combinations is that instead of using a single decision-making theme, the classification can be made by combining the individual opinions of separate base classifiers to derive a consensus decision. Ideally, the combination method should take advantage of the strengths of the individual classifiers, avoid their weaknesses, and improve the classification accuracy (Ho et al., 1994).

In a multiple classifier system, it is common that there are several base classifiers that are combined using a particular classifier combination strategy. It is obvious that a combination of base classifiers with identical errors does not improve the classification and hence, the base classifiers with decorrelating errors are preferred. Consequently, the base classifiers should differ from each other in some manner. This kind of classifier combination can be formed in several different ways. Duin (2002) presents six ways in which a consistent set of base classifiers can be generated:

1. *Different initializations*. Different initializations may yield different classifiers if the classifier training is initialization dependent. Examples of such classifier combinations can be found in the literature dealing with neural network classifiers (Hansen and Salamon, 1990).

2. *Different parameter choices*. There are some parameters to be selected in most of the classifier principles. For instance, in k-NN classification the value of k needs to be selected or in Parzen classifier, the window function needs to be selected. By using different parameter values, it is possible to obtain differently behaving classifiers.

3. *Different architectures.* In several kinds of classifiers, the architecture can be selected. For instance, the size of neural networks in the base classifiers can be varied.

4. *Different classifiers.* It is possible to use the same feature space and the same training set, but different classifiers in the base classification. An example of this approach is presented in the experimental parts of the works by Kittler et al., (1996) and Kittler et al., (1997), in which base classifiers of different classification principles are combined using selected combination rules.

5. *Different training sets.* If the same feature space is employed, the base classification can be carried out by taking different samples from the feature set to be used for training (Skurichina and Duin, 2002). A popular solution in this field has been bagging (Breiman, 1996), which involves selecting several training sets by sub-sampling the dataset randomly using bootstrapping (Duda et al., 2001). A classifier is constructed for each training set, and finally the classifier outputs are combined by majority voting. Another common algorithm, boosting (Freund and Shaphire, 1995), also manipulates the training data but emphasizes the training of samples which are difficult to classify. This is done by weighting the training set objects according to their performance in the previous classification. The classification begins with equal weighting, but after each classification the training sets are assigned new performance related weightings. This kind of classification produces a sequence of classifiers and weights. The final decision is made using a majority vote or a weighted majority vote of the classifiers

(Skurichina and Duin, 2002). Stacked generalization (Wolpert, 1992) splits the training set into several partitions. The final decision is based on the guesses of the base classifiers that have been taught with different parts of the training sets.

6. *Different feature sets.* The base classifiers may use separate feature sets as their inputs. These feature sets may describe different properties of the object to be classified, such as in character recognition (Ho et al., 1994; Xu et al., 1992). In these kinds of classifier combinations, an important application is the combination of physically different descriptions of the objects. An example of these applications is person identification (Brunelli and Falavigna, 1995), in which acoustic and visual features are combined using classifier combination. The image classification methods presented in this thesis combine classifiers employing separate feature sets (visual descriptors) extracted from the images.

It has been observed that randomly selected feature sets can also be effective in classifier combinations. In (Bryll et al., 2003), the features in the feature space are randomly selected for use in base classification. Ho (1998) introduces a random subspace method. In this method, the classifiers are constructed for randomly selected subspaces of a high-dimensional feature space. The classifications carried out in the subspaces are usually combined by majority voting (Skurichina and Duin, 2002). Further, the selection of the feature sets can also be based on cluster analysis (Cao et al., 1995). It is also possible to consider each dimension of the feature space separately. Guvenir and Sirin (1996) partitioned the feature space into each dimension and applied voting to reach a consensus decision based on the classifications at each dimension.

5.2 Classifier combination strategies

Once the base classifiers have been constructed, it is necessary to combine their opinions using some combination strategy. Various approaches to this have been proposed in the literature. The selection of the classification strategy is dependent on the information provided by the base classifiers. If only the class labels are available, different kinds of voting-based methods are possible. However, if the *a posteriori* probabilities of the base classifiers are available, different linear combination rules can be used. It is also possible to use the outputs of the base classifiers as features in the final classification. In this section, different classifier combination strategies are described.

5.2.1 Strategies based on probabilities

Classifier combination methods that use the *a posteriori* probabilities provided by the base classifiers have been the subject of much research (Alkoot and Kittler, 1999; Kittler et al., 1996, 1998; Kuncheva, 2002). The methods are also known as fixed combining rules (Duin, 2002). These strategies utilize the fact that the base classifier outputs are not just class numbers, but that they also include the confidence of the classifier (Duin, 2002). The confidence can be expressed as probability $P_i(\mathbf{S})$ of pattern $\mathbf{S}$ with respect to class ω_i ($i=1,2,\ldots,m$), where m is the number of classes (Duin, 2002):

$$P_i(\mathbf{S}) = \text{Prob}(\omega_i \mid \mathbf{S}) \tag{5.1}$$

However, if several classifiers are combined, the probabilities need to be defined for each C classifier. It is assumed that each classifier represents a particular feature type and the feature vector used by j:th classifier is denoted as f_j. Hence the pattern $\mathbf{S}$ is represented by C feature vectors. If the outcome of the jth classifier ($j=1,2,\ldots, C$) is denoted as O_j, the probability depends on the outcome O_{ij} of this classifier for class ω_i (Duin, 2002):

$$P_{ij}(\mathbf{S}) = \text{Prob}(\omega_i \mid O_{ij}) \tag{5.2}$$

The classification is carried out by assigning the pattern $\mathbf{S}$ to the class with the highest confidence. According to the Bayesian theory, $\mathbf{S}$ is assigned to class ω_i if the *a posteriori* probability of that class is maximum. Hence, $\mathbf{S}$ is assigned to class ω_i if:

$$P(\omega_i | f_1,\ldots,f_C) = \max_n P(\omega_n | f_1,\ldots,f_C) \tag{5.3}$$

It is possible to simplify the Bayesian decision rule of equation (5.3) by rewriting the *a posteriori* probability $P(\omega_n | f_1,\ldots,f_C)$ using the Bayes theorem (Kittler et al., 1998):

$$P(\omega_n | f_1,\ldots,f_C) = \frac{p(f_1,\ldots,f_C \mid \omega_n)P(\omega_n)}{p(f_1,\ldots,f_C)} \tag{5.4}$$

where $p(f_1,\ldots, f_C)$ is the unconditional measurement joint probability density which can be expressed in terms of the conditional measurement distributions (Kittler et al., 1998). For this reason, only the numerator terms of equation (5.4) are used in the following, and the joint probability of the classifiers can be represented by $p(f_1,\ldots,f_C|\omega_n)$. Based on this, it is possible to obtain several classifier combination rules (Kittler et al., 1996, 1998; Duin, 2002). A detailed discussion of the rules can be found in (Kittler et al., 1998).

Product rule is an important classifier combination rule. This rule can be obtained from the joint probability given by the base classifiers, $p(f_1,\ldots,f_C|\omega_n)$. The product rule assigns the pattern $\mathbf{S}$ to class ω_i, if:

$$p^{-(C-1)}(\omega_i)\prod_{j=1}^{C} P(\omega_i|f_j) = \max_{n=1}^{m} p^{-(C-1)}(\omega_n)\prod_{j=1}^{C} P(\omega_n|f_j) \tag{5.5}$$

Sum rule can be obtained from equation (5.5) (Kittler et al., 1998) by assuming that the *a posteriori* probabilities $P(\omega_n|f_l)$ computed by the respective classifiers do not deviate dramatically from the prior probabilities. This rule assigns the pattern $\mathbf{S}$ to class ω_i, if:

$$(1 - C)P(\omega_i) + \sum_{j=1}^{C} P(\omega_i|f_j) = \max_{n=1}^{m}\left[(1 - C)P(\omega_n) + \sum_{j=1}^{C} P(\omega_n|f_j) \right] \tag{5.6}$$
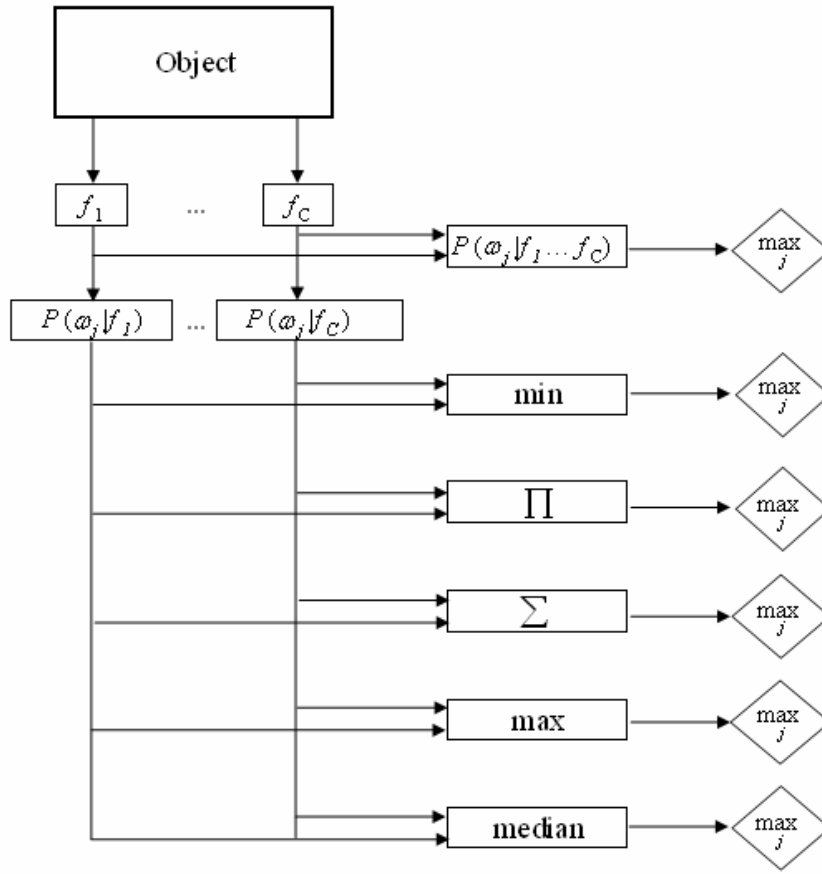
Figure 5.1. Outline of probability-based classifier combination schemes (Kittler et al., 1998).

Maximum rule selects the classifier that is most confident of itself. According to the maximum rule, **S** is assigned to class ω_i, if:

$$\max_{j=1}^{C} P(\omega_i | f_j) = \max_{n=1}^{m} \max_{j=1}^{C} P(\omega_n | f_j) \tag{5.7}$$

It is also possible to define the minimum rule, which selects the classifier that has the least objection against a certain class. In this rule, **S** is assigned to class ω_i, if:

$$\min_{j=1}^{C} P(\omega_i | f_j) = \max_{n=1}^{m} \min_{j=1}^{C} P(\omega_n | f_j) \tag{5.8}$$

The median rule assigns the pattern **S** to a class whose average *a posteriori* probability is at maximum. In this rule, the median is used instead of the mean to estimate the average because the mean can be affected by outliers within the input patterns and this could lead to an incorrect decision (Kittler et al., 1998). The median rule assigns the pattern **S** to class ω_i, if:

$$\operatorname*{med}_{j=1}^{C} P(\omega_i | f_j) = \max_{n=1}^{m} \operatorname*{med}_{j=1}^{C} P(\omega_n | f_j) \qquad (5.9)$$

These rules have been experimentally evaluated in (Alkoot and Kittler, 1999). A theoretical study of several of these classification rules is provided by Kuncheva (2002). The combination methods presented and their relationships are set out in Figure 5.1.

5.2.2 Voting-based strategies

Voting-based techniques are simple methods for combining classifiers. The basic idea behind these methods is to make a consensus decision based on the base classifier opinions using voting. Hence, the class labels provided by the base classifiers are regarded as votes, and the final class is decided to be the class that receives the majority or most of the votes. The benefit of these methods is that the decision can be made solely on the basis of the class labels provided by the base classifier. For this reason, no additional internal information, such as *a posteriori* probabilities, is required from the base classifiers. Hence the base classifiers can be regarded as "black boxes" (Lin et al., 2003).

Voting-based classifier combination strategies have been used extensively in the field of pattern recognition (Lam and Suen, 1997; Lin et al., 2003; Xu et al., 1992). The methods can be divided into two classes, majority voting and plurality voting. Majority voting (Lam and Suen, 1997) requires the agreement of more than half the participants in order to make a decision. If a majority decision cannot be reached, the sample is rejected. Plurality voting (Lin et al., 2003), on the other hand, selects the sample which has received the highest number of votes. As a result, the problem arising with the rejected samples can be avoided because all the samples can be classified.

5.2.3 Strategies employing the class labels

In addition to the voting and the probability-based classifier combination methods, various classifier combination methods have been proposed that utilize the base classifier outputs in other ways than voting. In the most common case, these outputs are the class labels given by the base classifiers, though in certain cases methods such as probability distributions are employed.

An early approach in this field was the use of class rankings (Ho et al., 1992, 1994). Here, rankings of classes are used instead of unique class choices. The classifier ranks a given set of classes with respect to an input pattern. A classifier strongly believes that the input pattern belongs to the class ranked at the top, but also that the other classes in the rank can be significant. In *highest rank* method (Ho et al., 1992, 1994), C classifiers are applied to rank a set of classes for each input pattern and thus each class receives C ranks. The highest of these C ranks is assigned to that class as its score. The set of classes is then sorted according to these scores to yield a combined class ranking for input to the pattern. *Borda count* (Ho et al., 1992, 1994) is a generalization of majority vote which also uses rankings. The Borda count of a class can be expressed as the sum of the number of the classes ranked below it by each classifier. The consensus ranking is given by arranging the classes so that their Borda counts are in descending order. The magnitude

of this count for each class measures the strength of agreement by the classifiers that the input pattern belongs to that class.

A number of combination methods have also been proposed in which the base classifier outputs have been used as new features. Wolpert (1992) presents the principle of *stacked generalization* in which each of the base classifiers employs separate parts of the training set. The outputs of the base classifiers form a new feature space in which the final classification is carried out. This principle shares similarities with the CRV method presented in the present thesis. The main difference is that in the proposed CRV method, the base classifiers use distinct visual descriptors as their inputs, not different parts of the training set. *Error-correcting output codes* (ECOC) (Dietterich and Bakiri, 1995) have received a certain amount of research interest. In these methods, a combination of binary classifiers produces a bit string that describes the object to be classified. This binary string is used as a code word that describes the sample. In these methods, the objective is to break up a multi-class problem into several two-class problems which are the tasks of each binary base classifier.

6 Applications in rock image classification

This study deals with the application field of rock image classification. The author has been involved in a research project which investigates rock image analysis. The project has been carried out in co-operation with industry and the author has been responsible for the development of visual description and classification methods for different kinds of rock images. For this purpose, various approaches have been presented in the publications attached to this thesis. In this Chapter, the main contributions of this work are presented.

6.1   Rock image classification methods

All the publications consider the use of colour and texture information in rock image description for classification purposes. However, the methodology introduced in the publications can be divided into two parts, texture filtering and classifier combinations.

6.1.1   Texture filtering

A variety of different kinds of filtering-based methods have been introduced for texture description, as presented in Chapter 3. In this thesis, however, texture description has been applied to natural rock images. The first three publications deal with this area of research. Texture analysis using Gabor filtering is applied to the surface inspection of industrial rock plates. In this way it is possible to detect the orientation and strength of surface cracking. This is essential because cracking in the rock surface is an important characteristic indicating, for instance, the extent to which it can withstand frost and moisture.

Another field of texture analysis discussed in this thesis concerns coloured textures. Colour and texture are two essential visual properties describing the rock type. In common texture analysis methods, image processing operations, such as filtering, are usually carried out only for the intensity (grey level) channel of the texture image. However, it is often beneficial to include colour information of the image in the texture

description. This is possible by making the texture description for other colour components and not solely intensity.

6.1.2 Combining colour and texture descriptors by classifier combinations

In rock image classification, it is necessary to combine different kinds of visual descriptors in a particular way. In this study, the descriptors characterize the texture and colour content of rock images. These descriptors are typically high dimensional and the feature values are spread in many difficult ways in the feature space. For this reason, it is often problematic to combine several kinds of descriptors into a single feature vector to be used in the classification. Instead, the descriptors can be first considered separately by using an individual base classifier to classify the input image based on a single descriptor. After this, the opinions of the separate base classifiers are combined to reach the final decision. For this purpose, various classifier combination strategies have been introduced in Publications IV-VII.

6.2 Overview of the publications and author's contributions

This thesis includes five publications presented at conferences and two journal articles. This section provides a short overview of the content of each publication and the author's own contribution to each publication is explained.

Publication I is related to the surface inspection of industrial rock plates. In this paper, a texture analysis method is presented for the analysis of surface reflection of polished rock surface. In addition, an image acquisition method that employs total reflection of the rock surface is proposed. The orientation of surface cracking can be inspected using a bank of oriented Gabor filters. The filtering results can be also used to measure the homogeneity of the surface.

The author suggested the idea of using the Gabor filtering method for the surface inspection of the rock plates. For this reason, she is responsible for the analysis methods presented in this paper. The idea of surface reflection imaging was originally provided by Mr. Autio, while image pre-processing and certain implementation work were carried out by Mr. Kunttu. Prof. Visa supervised the work.

Publication II is a study of texture filtering applied to colour channels of rock images. In this paper, a bank of Gabor filters is applied to the colour channels of the images in HSI colour space. Using this kind of method, the rock images can be classified using multiple scales and orientations. Two sets of industrial rock plate images have been used as testing material in the experiments. The experimental results show that by using the colour information, classification accuracy is improved compared to conventional grey level texture filtering.

The idea of combining the colour information of rock images with texture description was suggested and developed by the author. The implementation of the methods was carried out by Mr. Kunttu, and Mr. Autio was responsible for the geological input in the paper. Prof. Visa supervised the work.

Publication III also concerns the combination of colour and texture in filtering-based texture description. In this paper, colour information is applied to the texture description

of rock by using a set of Gaussian band-pass filters. In this case, the filters are ring-shaped, which means that texture orientation is not measured. This yields low dimensional texture description. In addition, the borehole rock images used in the experiments do not have clear orientations that could be used as classifying characteristics. Filtering is applied to the rock images in RGB and HSI colour spaces. The experimental results show that the proposed visual descriptors outperform, for example, several MPEG-7 colour and texture descriptors in classification accuracy.

This study follows on from Publication II by employing orientation-insensitive filters for the analysis of coloured rock images. The article was written by the author who had a central role in the filter design and in the selection of the colour spaces.

Publication IV presents the principle of classification result vector (CRV) for rock image classification. This is a method for combining separate base classifiers employing different visual descriptors. In this method, the unknown sample image is first classified by employing a separate base classifier for each colour and texture descriptor. The class labels provided by the base classifiers are then combined into a feature vector that describes the image in the final classification. This vector is called classification result vector. In the final classification the images are classified using their CRV's as feature vectors.

The CRV method was invented by the author. Mr. Kunttu was responsible for most of the implementation of the classifiers and Prof. Visa supervised the work.

Publication V introduces a method for combining base classifiers using their probability distributions. In this method, the probability distributions provided by separate base classifiers employing different visual descriptors are combined into a classification probability vector (CPV). The CPV method is an extension of the CRV method. The main difference between these two is that CPV uses the probability distributions provided by the base classifiers as features in the final classification, whereas CRV employs only the class numbers for this purpose. The experimental results show that the CPV gives better classification accuracy than many other classifier combination approaches for rock images and also for paper defect images.

This method was also developed by the author while the implementation issues were overseen by Mr. Kunttu. Mr. Autio and Mr. Rauhamaa were responsible for applications for rock and paper defect image classification, respectively. Prof. Visa supervised the work.

Publication VI presents an application example of the CPV method. This paper proposes an approach to multilevel colour description of rock images. Such a description is obtained by combining separate base classifiers that use image histograms at different quantization levels as their inputs. The base classifiers are combined using the CPV method. The experimental results obtained with rock images show that an accurate colour-based classification can be achieved with the method.

The idea of using the CPV method for making multilevel colour description was proposed by the author. Mr. Kunttu oversaw implementation and Prof. Visa supervised the study.

Publication VII presents a method for combining various visual descriptors in rock image classification. In this method, the descriptors extracted from an image are used by k-nearest neighbour base classifiers and then a final decision is made by combining the nearest neighbours in each base classification. The total numbers of the neighbours representing each class are used as votes in the final classification. Experimental results with rock image classification indicate that the proposed classifier combination method is more accurate than conventional plurality voting.

This voting-based approach was invented and reported by the author. Again, the method was implemented mainly by Mr. Kunttu and Prof. Visa supervised the work.

7 Conclusions

With the rapid development of digital imaging tools, imaging applications have been adopted in many areas in which inspection and monitoring have been done manually. The application area of this thesis is an example of this change. Formerly, rock samples were inspected manually in the rock industry as well as in geological research. It was not until fairly recently that imaging and image processing methods made it feasible to start developing automatic approaches for the visual inspection and recognition of rock. Compared to several other goods and materials that are inspected by computer vision systems on the production line, rock material is a significantly more demanding analysis task. This is because rock is a natural material whose visual properties are often varying. For example, in terms of classification, the division of the rock images into classes is a difficult task even for a geological expert. Indeed, individual experts often classify the same rock image set in different ways.

There are several advantages to using an automatic rock image classifier. For instance, the amount of manual labour can be significantly reduced and the subjective nature of the classification can be eliminated. With bedrock investigations, especially, the visual information obtained from the boreholes can be effectively utilized with appropriate image analysis and recognition tools. It is for these reasons that the area of rock image analysis is a significant application area of image analysis and pattern recognition.

The goal of this thesis was to develop methods and techniques for the classification of natural rock images. In image classification, visual descriptors extracted from the images are used to describe image content. In this study, description methods were developed to characterize colour and texture content of the rock images. In addition, the classification procedure of the rock images was considered and the focus of this part of the research was on classifier combinations. The reason for this is that different types of visual descriptors can be easily combined in the classification by using classifier combinations.

The main contributions of this work include new multiscale filtering techniques which combine colour and texture information of the rock. These techniques improve filtering-based texture classification compared to conventional texture filtering which is applied

only to the intensity channel of the image. Another main contribution is a novel method for the analysis and measurement of the rock plate surface structures by means of surface reflection imaging and texture analysis. The third main contribution is a combination of visual descriptors of the rock images in the classification by using classifier combinations. It has been shown that classification of rock images can be easily improved by combining separate classifiers employing distinct visual descriptors. For this purpose, several classifier combination methods have been introduced.

The methods introduced in this thesis are all directly applicable to practical rock image classification problems. They can, therefore, be used whenever rock image classification systems are being constructed.

Bibliography

Ade, F., 1983. Characterization of Textures by 'Eigenfilters'. Signal Processing 5(5), 451-457.

Ahuja, N., Rosenfeld, A., 1981. Mosaic Models for Textures. IEEE Transactions on Pattern Analysis and Machine Intelligence 3(1), 1-11.

Alkoot, F.M. Kittler, J., 1999. Experimental evaluation of expert fusion strategies. Pattern Recognition Letters 20(11-13), 1361-1369.

Asano, A., 1999. Texture Analysis Using Morphological Pattern Spectrum and Optimization of Structuring Element. Proceedings of International Conference on Image Analysis and Processing, Venice, Italy, pp. 209-214.

Autio, J., Lukkarinen, S., Rantanen, L., Visa, A. 1999. The Classification and Characterization of Rock Using Texture Analysis by Co-occurrence Matrices and the Hough Transform. Proceedings of International Symposium on Imaging Applications in Geology, pp. 5-8.

Autio, J., Lepistö, L., Visa, A., 2004. Image analysis and data mining in rock material research. Materia, (4), 36-40.

Baykut, A., Atalay, A., Aytul, E., Guler, M., 2000. Real-time defect inspection of textured surfaces. Real-Time Imaging 6(1), 17-27.

Besag, J., 1974. Spatial Interaction and Statistical Analysis of Lattice Systems. Journal of Royal Statistical Society 36, 192-236.

Bigün, J., du Buf, J.M.H., 1994. N-folded symmetries by complex moments in Gabor space. IEEE Transactions on Pattern Analysis and Machine Intelligence 16(1), 80-87.

Bilmes, J.A., Kirchhoff, K., 2003. Generalized rules for combination and joint training of classifiers. Pattern Analysis & Applications 6(3), 201-211.

Bibliography

Boukovalas, C., Kittler, J. Marik, R., Petrou, M., 1997. Automatic Color Grading of Ceramic Tiles Using Machine Vision. IEEE Transactions on Industrial Electronics 44(1), 132-135.

Boukovalas, C., Kittler, J., Marik, R., Petrou, M., 1999. Color Grading of Randomly Textured Ceramic Tiles Using Color Histograms. IEEE Transactions on Industrial Electronics 46(1), 219-226.

Breiman, L., 1996. Bagging predictors, Machine Learning 26(2), 123-140.

Brodatz, P Texture: A photographic Album for Artists and Designers, Reinhold, New York, 1968.

Brunelli, R., Falavigna, D., 1995. Person identification using multiple cues. IEEE Transactions on Pattern Analysis and Machine Intelligence 17(10), 955-966.

Bruno, R., Persi Paoli, S., Laurenge, P., Coluccino, M., Muge, F., Ramos, V., Pina, P., Mengucci, M., Chica Olmo, M., Serrano Olmedo, E., 1999. Image Analysis on Ornamental Stone Standard's Characterization. Proceedings of International Symposium on Imaging Applications in Geology, pp. 29-32.

Bryll, R. Gutirrez–Osuna, R., Quek, F., 2003. Attribute bagging: improving accuracy of classifiers ensembles by using random feature subsets. Pattern Recognition 36(6), 1291-1302.

Cao, J., Ahmadi, M., Shridhar, M., 1995. Recognition of handwritten numerals with multiple feature and multistage classifier. Pattern Recognition 28(2), 153-160.

Chen, P.C., Pavlidis, T., 1983. Segmentation by Texture Using Correlation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 5, 64-69.

Chui, C.K., 1992. An Introduction to Wavelets. Wavelet Analysis and Its Applications. Vol. 1, Academic press, London.

Coggins, J.M., Jain, A.K., 1985. A Spatial Filtering Approach to Texture Analysis. Pattern Recognition Letters 3(3), 195-203.

Crida, R.C., De Jager, G., 1994. Rock Recognition Using Feature Classification. Proceedings of the IEEE South African Symposium on Communications and Signal Processing, Stellenbosch, South Africa, pp. 152-157.

Crida, R.C., De Jager, G., 1996. Multiscalar Rock Recognition Using Active Vision. Proceedings of International Conference on Image Processing, Lausanne, Switzerland, Vol. 2, pp. 345-348.

Deguchi, K., Morishita, I., 1978. Texture Characterization and Texture-Based Image Partition Using Two-Dimensional Linear Estimation Techniques. IEEE Transactions on Computer 27(8), 739-745.

Del Bimbo, A., 1999. Visual Information Retrieval, Morgan Kaufmann Publishers, San Fransisco, California.

64

Bibliography

Dietterich, T.G., Bakiri, G., 1995. Solving multiclass learning problems via error-correcting output codes. Journal of Artificial Intelligence Research 2, 263-286.

Dougherty, E.R., Newell, J.T., Pelz, J.B., 1992. Morphological Texture-Based Maximum-Likelihood Pixel Classification Based on Local Gramulometric Moments. Pattern Recognition 25(10), 1181-1198.

Duarte, M.T., Fernlund, J.M.R., 2005. Analysis of meso textures of geomaterials through Haralick parameters. Proceedings of the 2nd Iberian Conference on Pattern Recognition and Image Analysis, LNCS 3523, Estoril, Portugal, pp. 713-719.

Duda, R.O., Hart, P.E., Stork, D.G., 2001. Pattern Classification. 2nd edition, John Wiley & Sons.

Duin, R.P.W., 2002. The combining classifier: to train or not to train. Proceedings of 16th International Conference on Pattern Recognition, Quebec, Canada, Vol. 2, pp. 765-770.

Duin, R.P.W., Juszczak, P., Paclik, P., Pekalska, E., de Ridder, D., Tax, D.M.J., 2004. PRTools4, A Matlab Toolbox for Pattern Recognition, Delft University of Technology.

Dyer, C.R., Rosenfeld, A., 1976. Fourier Texture Features: Suppression of Aperture Effects. IEEE Transactions on Systems, Man, and Cybernetics 6(10), 703-706.

Elfadel, I.M., Picard, R.W., 1994. Gibbs Random Fields, Co-occurrences and Texture Modelling. IEEE Transactions on Pattern Analysis and Machine Intelligence 16(1), 24-37.

Freund, Y., Shaphire, R.E., 1995. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences 55(1), 119-139.

Gonzalez, R.G., Woods, R.E., 1993. Digital Image Processing. Addison-Wesley Publishing Company.

Gricorescu. S.E., Petkov, N., Kruizinga, P., 2002. Comparison of Texture Features Based on Gabor Filters. IEEE Transactions on Image Processing 11(10), 1160-1167.

Guvenir, A., Sirin, I., 1996. Classification by feature partition. Machine Learning 23, 47-67.

Hansen, L.K., Salamon, P., Neural network ensembles. IEEE Transactions on Pattern Analysis and Machine Intelligence 12(10), 993-1001.

Haralick, R.M., Shanmugam, K. Dinstein, L., 1973. Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics 3(6), 610-621.

Haralick, R.M., 1979. Statistical and Structural Approaches to Texture. Proceedings of the IEEE 67(5), 786-804.

Ho, T.K., Hull, J.J. Srirari, S.N., 1992. On multiple classifier systems for pattern recognition. Proceedings of 11th International Conference on Pattern Recognition, Hague, the Netherlands, pp. 84-87.

Bibliography

Ho, T.K., Hull, J.J. Srirari, S.N., 1994. Decision combination in multiple classifier systems. IEEE Transactions on Pattern Analysis and Machine Intelligence 16(1), 66-75.

Ho, T.K., 1998. The random subspace method for construction decision forests. IEEE Transactions on Pattern Analysis and Machine Intelligence 20(8), 832-844.

Huang, J., Kumar, S.R., Mitra, M., Zhu, W.J., 1997. Image Indexing Using Color Correlograms. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Juan, Puerto Rico, pp. 762-768.

Iivarinen, J., 1998. Texture Segmentation and Shape Classification with Histogram Techniques and Self-Organizing Maps. Acta Polytechnica Scandinavica, Matematics, Computing and Management in Engineering Series No. 95. Doctor of Science (Technology) Thesis, Espoo.

Jain, A.K., Farrokhnia, F., 1991. Unsupervised texture segmentation using Gabor filters. Pattern Recognition 24(12), 1167-1186.

Jain, A.K., Prabhakar, S., Chen, S., 1999. Combining multiple matchers for a high security fingerprint verification system. Pattern Recognition Letters 20 (11-13), 1371-1379.

Julesz, B., 1981. Textons, the Elements of Texture Perception, and Their Interactions, Nature 290, 91-97.

Kaizer, H., 1955. A Quantification of Textures on aerial Photographs. Tech. Note No. 121, A 69484, Boston University Research Laboratories, Boston University.

Kauppinen, H., 1999. Development of color machine vision method for wood surface inspection. Dr. tech. dissertation, university of Oulu, Finland.

Keller, F. J., Gettys, W.E., Skove, M.J., 1993. Physics, Classical and Modern. 2nd edition, McGraw-Hill inc.

Kervrann, J., Heitz, F., 1995. A Markov Random Field Model-Based Approach to Unsupervised Texture Segmentation Using Local and Global Spatial Statistics. IEEE Transactions on Image Processing 4(6), 856-862.

Kittler, J., Hatef, M., Duin, R.P.W., 1996. Combining Classifiers. Proceedings of the 13th International Conference on Pattern Recognition, Vienna, Austria, Vol. 2, pp. 897-901.

Kittler, J., Hojjatoleslami, A., Windeatt, T., 1997. Strategies for combining classifiers employing shared and distinct pattern representations, Pattern Recognition Letters 18(11-13), 1373-1377.

Kittler, J., Hatef, M., Duin, R.P.W., Matas, J., 1998. On combining classifiers. IEEE Transactions on Pattern Analysis and Machine Intelligence 20(3), 226-239.

Kruizinga, P., Petkov, N., 1999. Non-linear operator for oriented texture. IEEE Transactions on Image Processing 8(10), 1395-1407.

Bibliography

Kukkonen, S., Kälviäinen, H., Parkkinen, J., 2001. Color features for quality control in ceramic tile industry. Optical Engineering 40(2), 170-177.

Kumar, A., Pang, G., 2002. Defect detection in textured materials using Gabor filters. IEEE Transactions on Industry Applications 38(2), 425-440.

Kuncheva, L.I., 2002. A theoretical study on six classifier fusion strategies. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(2) 281-286.

Laine, A., Fan, J., 1993. Texture Classification by Wavelet Packet Signature. IEEE Transactions on Pattern Analysis and Machine Intelligence 15(11), 1186-1191.

Lakmann, R., Priese, L., 1997. A Reduced Covariance Color Texture Model for Micro-Textures. In Proceedings of 10th Scandinavian Conference on Image Analysis, Lappenranta, Finland.

Lam, L., Suen, C.Y., 1997. Application of majority voting to pattern recognition: An analysis of the behavior and performance. IEEE Transactions on Systems, Man, and Cybernetics, Part A, 27(5) 553-567.

Laws, K.I., 1980. Rapid texture identification. In Proceedings of SPIE Conference on Image Processing and Missile Guidance, pp. 376-380.

Lebrun, V., 1999. Development of Specific Image Acquisition Techniques for Field Imaging -Applications to Outcrops and Marbles. Proceedings of International Symposium on Imaging Applications in Geology, pp. 165-168.

Lebrun, V., Toussaint, C., Pirard, E. 2000. On the Use of Image Analysis for Quantitative Monitoring of Stone alteration. Weathering 2000 International Conference, Belfast.

Lebrun, V., 2001. Quality Control of Ceramic Tiles by Machine Vision. Asian Ceramics: Manufacturing equipment guide spec. pub.

Lebrun, V., Macaire, L., 2001. Aspect Inspection of Marble Tiles by Colour Line Scan Camera. Proceedings of QCAV.

Lepistö, L., Kunttu, I., Autio, J. Visa, A. 2003a. Rock image classification using non-homogenous textures and spectral imaging. WSCG Short papers proceedings, Plzen, Czech Republic, pp. 82-86.

Lepistö, L., Kunttu, I., Autio, J. Visa, A, 2003b. Retrieval of Non-Homogenous Textures Based on Directionality. Proceedings of 4th European Workshop on Image Analysis for Multimedia Interactive Services, London, UK, pp. 107-110.

Lepistö, L., Kunttu, I., Autio, J. Visa, A, 2004. Rock Image Retrieval and Classification Based on Granularity. Proceedings of 5th International Workshop on Image Analysis for Multimedia Interactive Services, Lisbon, Portugal.

Lin, X., Yacoub, S., Burns, J., Simske, S., 2003. Performance analysis of pattern classifier combination by plurality voting. Pattern Recognition Letters 24(12), 1959-1969.

Bibliography

Lindqvist, J.E., Åkesson, U., 2001. Image analysis applied to engineering geology, a literature review. Bulletin of Engineering Geology and the Environment 60(2), 117-122.

Liu, F., Picard, R.W., 1996. Periodicity, Directionality, and Randomness: Wold Features for Image Modeling and Retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence 18(7), 722-733.

Lu, X., Wang, Y., Jain, A.K., 2003. Combining classifiers for face recognition. Proceedings of International Conference on Multimedia and Expo, Baltimore, Maryland, USA, Vol. 3, pp. 13-16.

Luthi, S.M., 1994. Textural segmentation of digital rock images into bedding units using texture energy and cluster labels. Mathematical Geology 26(2), 181-196.

Mallat, S.G., 1989. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 11(7), 674-693.

Manjunath, B.S., Chellappa, R., 1991. Unsupervised Texture Segmentation Using Markov Random Field Models. IEEE Transactions on Pattern Analysis and Machine Intelligence 13(5), 478-482.

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), 837-842.

Manjunath, B.S., Salembier, P., Sikora, T., 2002. Introduction to MPEG-7 Multimedia content description interface. John Wiley & Sons, UK.

Mao, J., Jain A. K., 1992. Texture Classification and Segmentation Using Multiresolution Simultaneous Autoregressive Models. Pattern Recognition 25(2), 173-188.

Mengko, T.R., Susilowati, Y. Mengko, R. Leksono, B.E., 2000. Digital Image Processing Technique in Rock Forming Minerals Identification. Proceedings of IEEE Asia-Pacific Conference of Circuits and Systems, pp. 441-444.

Newman, T., Jain, A., 1995. A survey of automated visual inspection. Computer Vision and Image Understanding 61(2), 231-262.

Ojala, T., Pietikäinen, M., Harwood, D., 1996. A Comparative Study of Texture Measures with Classification Based on Feature Distributions. Pattern Recognition 29(1), 51-59.

Ojala, T., 1997. Nonparametric Texture Analysis Using Spatial Operators, with Applications in Visual Inspection, Acta Universitatis Ouluensis Technica C 105, Doctor of Science (Technology) Thesis, Oulu, Finland.

Ojala, T., Valkealahti, K., Oja, E., Pietikäinen, M., 2001. Texture Discrimination with Multidimensional Distributions of Signed Gray Level Differences, Pattern Recognition 34(3), 727-739.

Bibliography

Oza, N.C., polikar, R., Kittler, J., Roli, F. (Eds.), 2005. Proceedings of 6th International Workshop on Multiple classifier systems, LNCS 3541, Seaside, CA, USA.

Paclík, P., Verzakov, S., Duin, R.P.W., 2005. Improving the maximum-likelihood co-occurrence classifier: a study on classification of inhomogenous rock images. Proceedings of 14th Scandinavian Conference on Image Analysis, Joensuu, Finland, LNCS 3540, pp. 998-1008.

Palm, C., Keysers, D., Lehmann, T., Spitzer, K., 2000. Gabor Filtering of Complex Hue/Saturation Images for Color Texture Classification. In Proceedings of Joint Conference on Information Sciences – International Conference on Computer Vision, Pattern Recognition, and Image Processing, Vol. 2, pp. 45-49.

Paschos, G., 1998. Chromatic Correlation Features for Texture Recognition. Pattern Recognition Letters 19(8), 643-650.

Paschos, G., Radev, I., 1999. Image Retrieval Based on Chromaticity Moments. In Proceedings of International Conference on Image Analysis and Processing, Venice, Italy, pp. 904-908.

Paschos, G., 2001. Perceptually Uniform Color Spaces for Color Texture Analysis: An Empirical Evaluation. IEEE Transactions on Image Processing 10(6), 932-937.

Pietikäinen, M., Ojala, T., Silven, O., 1998. Approaches to texture-based classification segmentation and surface inspection. In: Chen, C.H., Pau, L.F., Wang, P.S.P. (Eds.), Handbook of Pattern Recognition and Computer Vision (2nd edition), World Scientific Publishing Company, pp. 711-736

Randen, T., Husøy, J.H., 1999. Filtering for Texture Classification: A Comparative Study. IEEE Transactions on Pattern Analysis and Machine Intelligence 21(4), 291-310.

Rao, A.R., Lohse G.L., 1993. Towards a Texture Naming System: Identifying Relevant Dimensions of Texture. Proceedings of IEEE Conference on Visualization, San Jose, California, USA, pp. 220-227.

Roli, F., Kittler, J., Windeatt, T. (Eds.), 2004. Proceedings of 5th International Workshop on Multiple classifier systems, LNCS 3077, Cagliari, Italy.

Salinas, R.A., Raff, U., Farfan, C., 2005. Automated estimation of rock fragment distributions using computer vision and its application in mining. IEE proceedings on Vision, Image, and Signal Processing 152(1), 20050810.

Santini, S., Jain, R., 1999. Similarity measures. IEEE Transactions on Pattern Analysis and Machine Intelligence 21(9), 871-883.

Schacter, B.J., Rosenfeld, A., Davis, L.S., 1978. Random Mosaic Models for Textures. IEEE Transactions on Systems, Man, and Cybernetics 8(9), 694-702.

Schalkoff, R.J., 1989. Digital image processing and computer vision. John Wiley & Sons.

Bibliography

Shim, S.-O., Choi, T.S., 2003. Image indexing by modified color co-occurrence matrix. Proceedings of International Conference on Image Processing, Barcelona, Spain, Vol. 3, pp. 493-496.

Singh, M., Javadi, A., Singh, S., 2004. A comparison of texture features for the classification of rock images. Proceedings of 5th International Conference on Intelligent Data Engineering and Automated Learning, LNCS 3177, Exeter, UK, pp. 179-184.

Skurichina, M., Duin, R.P.W., 2002. Bagging, boosting, and the random subspace method for linear classifiers. Pattern Analysis and Applications 5, 121-135.

Smeulders, A. W. M., Worring, M., Santini, S. Gupta, A., Jain, R. 2000. Content-Based Image Retrieval at the End of the Early Years. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22, 1349-1380.

Spies, B.R., 1996. Electric and electromagnetic borehole measurements: a review. Surveys in Geophysics 17(4), 517-556.

Stricker, M., Orengo, M., 1995. Similarity of Color Images. Proceedings of SPIE Storage and Retrieval for Image and Video Databases III, Vol. 2420.

Swain, M., Ballard, D., 1991. Color Indexing. International Journal of Computer Vision 7(1), 11-32.

Tamura, H., Mori, S., Yamawaki, T., 1978. Textural features corresponding to visual perception. IEEE Transactions on System, Man, and Cybernetics, SMC-8 (6).

Thai, B., Healey, G., 2000. Optimal Spatial Filter Selection for Illumination-Invariant Color Texture Discrimination. IEEE Transactions on System, Man, and Cybernetics – Part B 30(4), 610-616.

Tobias, O.J., Seara, R. Soares, F.A.P., Bermudez, J.C.M., 1995. Automatic Visual Inspection Using the Co-occurrence Approach. Proceedings of the 38th Midwest Symposium on Circuits and Systems, Vol. 1, pp. 154-157.

Tüceryan, M., Jain, A.K., 1990. Texture Segmentation Using Voronoi Polygons. IEEE Transactions on Pattern Analysis and Machine Intelligence 12(2), 211-216.

Tüceryan, M., Jain, A.K., 1993. Texture Analysis. In: Chen, C. H., Pau, L. F., and Wang, P. S. P., editors, Handbook of Pattern Recognition and Computer Vision, pp. 235-276, World Scientific.

Unser, M., 1986. Sum and Difference Histograms for Texture Classification. IEEE Transactions on Pattern Analysis and Machine Intelligence 8(1), 118-125.

Unser, M., 1995. Texture Classification and Segmentation Using Wavelet Frames. IEEE Transactions on Image Processing 4(11), 1549-1560.

Valkealahti, K., Oja, E., 1998a. Reduced Multidimensional Co-Occurrence Histograms in Texture Classification. IEEE Transactions on Pattern Analysis and Machine Intelligence 20(1), 90-94.

Bibliography

Valkealahti, K., Oja, E., 1998b. Reduced Multidimensional Histograms in Color Texture Description. In Proceedings of 14th international Conference on Pattern Recognition, Vol. 2, pp. 1057-1061.

van der Heijden, F., Duin, R.P.W., de Ridder, D., Tax, D.M.J., 2004. Classification, parameter estimation and state estimation. John Wiley & Sons.

Van Gool, L., Dewaele, P., Oosterlinck, A. 1985. Survey texture analysis Anno 1983. Computer Vision, Graphics, and Image Processing 29, 336-357.

Visa, A., 1990. Texture Classification and Segmentation Based on Neural Network Methods. Doctor of Science (Technology) Thesis, Espoo, Finland.

Weszka, J.S., Dyer, C.R., Rosenfeld, A., 1976. A Comparative Study of Texture Measures for Terrain Classification. IEEE Transactions on Systems, Man, and Cybernetics 6, 269-285.

Wilson, S.S., 1989. Vector Morphology and Iconic Neural Networks. IEEE Transactions on Systems, Man, and Cybernetics 19(6), 1636-1644.

Wolpert, D.H., 1992. Stacked generalization. Neural Networks 5(2), 241-260.

Wyszecki, G., Stiles, W.S., 1982. Color Science, Concepts and Methods, Quantitative Data and Formulae. 2nd edition, John Wiley & Sons.

Xu, L., Krzyzak, A., Suen, C.Y., 1992. Methods for combining multiple classifiers and their applications to handwriting recognition. IEEE Transactions on Systems, Man, and Cybernetics 22(3), 418-435.

Bibliography

Publications

Publication I

Lepistö, L., Kunttu, I., Autio, J., Visa, A., 2003. Multiresolution Texture Analysis of Surface Reflection Images.

Multiresolution Texture Analysis of Surface Reflection Images

Leena Lepistö[1], Iivari Kunttu[1], Jorma Autio[2], and Ari Visa[1]

[1] Tampere University of Technology, Institute of Signal Processing
P.O. Box 553, FIN-33101 Tampere, Finland
{Leena.Lepisto, Iivari.Kunttu, Ari.Visa}@tut.fi
http://www.tut.fi
[2] Saanio & Riekkola Consulting Engineers
Laulukuja 4, FIN-00420 Helsinki, Finland
Jorma.Autio@sroy.fi
http://www.sroy.fi

Abstract. Surface reflection can be used as one quality assurance procedure to inspect the defects, cracking, and other irregularities occurring on a polished surface. In this paper, we present a novel approach to the detection of defects based on analysis of surface reflection images. In this approach, the surface image is analyzed using texture analysis based on Gabor-filtering. Gabor-filters can be used in the inspection of the surface in multiple resolutions, which makes it possible to inspect the defects of different sizes. The orientation of the defects and surface cracking is measured by applying the Gabor-filters in several orientations. A set of experiments were carried out by using surface reflection images of polished rock plates and the orientation of the surface cracking was determined. In addition, the homogeneity of the rock surface was measured based on the Gabor features. The results of the experiments show that Gabor features are effective in the measurement of the surface properties.

1 Introduction

During recent years, the number of industrial imaging systems has increased remarkably. The purpose of these digital imaging solutions is often the control of quality or production. In these solutions, analysis of the image data is made by using some image processing method. One typical application for the image analysis system is to detect and analyze the defects occurring in the production. Image analysis is used to detect possible malfunctioning as soon as possible to minimize the economic losses. Another application is classification of the products in different categories.

Surface reflection is a phenomenon that can be utilized in the detection of defects and microfracturing on different surfaces. In addition, based on the surface reflection, other surface properties can also be analyzed. These properties can be for example uniformity and smoothness of the surface.

The reflection image obtained from the surface can be analyzed using methods and tools developed in the field of texture analysis. The majority of the texture analysis

methods are based on the statistical textural features, some texture model, or application of the signal processing tools. An example of the statistical tools is co-occurrence matrix [3], whereas Multiresolution autoregressive model (MRSAR) [10] represents the model-based features. Recently, the methods based on signal processing have been widely used in the analysis of textures. Commonly used filtering-based signal processing methods use wavelets [4] or Gabor features [9]. The benefit of the wavelet-based methods is that they measure the textural properties in multiple resolutions. Manjunath and Ma [9] have made a comparison between multiple oriented Gabor filters, other wavelet-features, and MRSAR-model. In this comparison, the best results were achieved using Gabor-filtering.

In rock industry, the digital imaging tools are used in the quality and production control of rock. Rock is a commonly used as ornamental stones in the building industry, where rock plates are used e.g. to cover the floors and walls of buildings. When the plates are used in the external walls, they are required to tolerate different weather conditions. The cracking and other defects occurring in the surface of the rock plates have significant effect on their strength and ability to bear frost and moisture. Therefore, it is beneficial for a rock manufacturer to be able to classify and assure the quality of the plates. Important properties of the rock surface are strength and directionality of the surface cracking. The homogeneity of the surface reflection is also important, because the smoothness of the polished surface can be determined based on its homogeneity.

In the rock image analysis, it has been made several studies about the rock texture images. Autio et al. [1] have researched rock texture characterization and classification using co-occurrence matrix and texture directionality. In [6] we presented a classification system for non-homogenous rock images using textural and spectral features. The texture directionality was used in rock image classification in [7]. All these studies concern the images of rock texture, whereas the number of the researches about the surface reflection is very limited. However, in [5] surface reflection model for rock is presented. This study was focused mainly on image acquisition and the methods used for the reflection image analysis were not discussed.

In this paper, we use multiscale texture analysis methods for the inspection of the surface reflection images. We use the textural properties to distinguish between homogenous and non-homogenous surfaces. For testing purposes, we used a set of industrial rock plates, whose surface reflection was used in testing of analysis methods.

2   Analysis of Surface Reflection Images

When the light approaches a surface at an angle Θ_1 to the normal of the surface, a part of it reflects from the surface at the same angle Θ_2 at the opposite side of the normal (figure 1a). If the angle Θ_1 exceeds a certain critical angle Θ_c, all the approaching light reflects from the surface. This phenomenon is called total reflection, in which the surface acts like a mirror reflecting the light at angle Θ_2. This reflection can be utilized in the surface inspection, because the light reflects from the smooth surface in
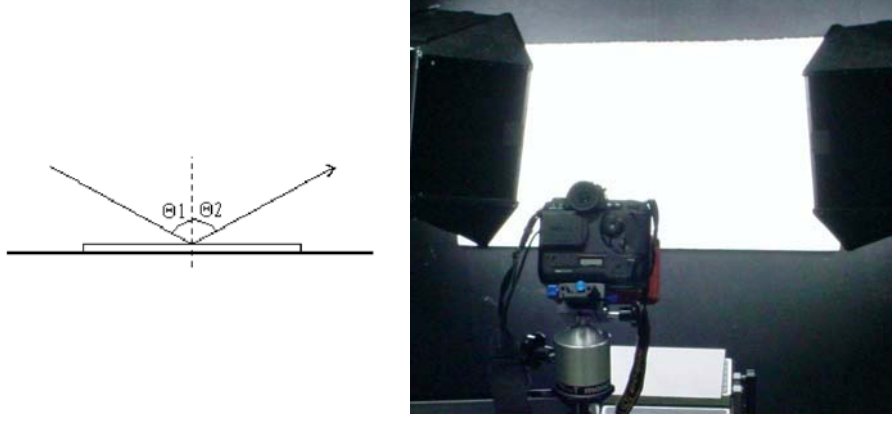
Fig. 1. a) The surface reflection model, b) The imaging arrangement of the rock plates

a different way than from the cracking and other defects. Using a sophisticated digital camera system, this reflection pattern can be acquired into digital form, in which it can be processed and analyzed.

2.1 Multiresolution Texture Analysis

Directionality of the cracking and the other surface properties can be described using texture analysis. By means of multiscale texture analysis, the surface structures can be analyzed in the several resolutions. This is essential, because the size of the surface cracks and defects may vary strongly.

The analysis methods presented in this paper use the texture representation based on the Gabor filters. The experimental results of [9] show that Gabor filtering gives the best texture classification compared to the other, commonly used wavelet-based texture analysis tools. Gabor filters can also be considered as tunable edge and line (bar) detectors [8]. This property can be utilized in the description of the surface cracking. The Gabor-features used in this study are based on the work of Manjunath and Ma [9]. They have used a bank of Gabor filters to characterize the texture properties. This filter bank can be used in multiple scales and orientations. Gabor function is based on the gaussian wavelet function [2]. A two dimensional Gabor function $g(x,y)$ and its Fourier transform $G(u,v)$ can be written as [9]:

$$g(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y} \right) \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi j W x \right] \quad (1)$$

$$G(u, v) = \exp\left\{ -\frac{1}{2}\left[\left(\frac{(u-W)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2} \right) \right] \right\} \quad (2)$$

where $\sigma_u=1/2\pi\sigma_x$ and $\sigma_v=1/2\pi\sigma_y$. Let $g(x,y)$ be a mother Gabor wavelet, then filter at multiple rotations and scales can be obtained by appropriate dilations and rotations of $g(x,y)$ through the generating function:

$$g_{mn}(x,y) = a^{-m}G(x',y'), \quad a > 1, \, m,n=\text{integer}$$

$$x'= a^{-m}(x\cos\theta + y\sin\theta), \text{ and } y'= a^{-m}(-x\sin\theta + y\cos\theta) \quad (3)$$

where $\Theta=n\pi/K$, K is the total number of orientations, and a^{-m} is an energy factor that ensures that energy is independent of m [9]. When we have an image $\mathbf{I}(x,y)$, its Gabor wavelet transform is defined as [9]:

$$W_{mn}(x,y) = \int \mathbf{I}(x_1,y_1)g_{mn}*(x - x_1, y - y_1)dx_1 dy_1 \quad (4)$$

where * indicates the complex conjugate.

2.2   Textural Description of the Surface Reflection Images

The multiscale texture representation presented in section 2.1 can be used to characterize the surface reflection image. An effective textural descriptor for a texture region is the mean of the transform coefficient magnitudes for the scale m and the orientation n [9]:

$$\mu_{mn} = \int\int |W_{mn}(x,y)|dxdy \quad (5)$$

Using μ_{mn}, the distribution of the orientations occurring in the image can be formed for each scale. Mean value μ_{mn} can be defined for a set of scales $[M]=s_1, s_2, \dots , s_k$, and for the orientations $[N]=\Theta_1, \Theta_2, \dots , \Theta_l$. Then the dominating orientation Θ_D at the scale m is the orientation, in which μ_{mn} has its maximum value. When μ_{mn} is defined for all k scales and l orientations of the sets $[M]$ and $[N]$, a feature vector for a texture region can be defined as:

$$\mathbf{F} = [\mu_{00},\mu_{01},\mu_{02},\dots,\mu_{kl}] \quad (6)$$

Using this feature vector, texture properties of a texture region can be compared with the properties of the other regions.

In addition to the orientation of the surface cracking, also homogeneity of the surface is essential in the rock surface analysis. The mean of the transform coefficients μ_{mn} can be used also in the measurement of the texture homogeneity. In this case, the reflection image is divided into B subimages (or blocks). The textural properties of i:th block are measured by calculating a feature vector $\mathbf{F}_i$ for it. Then the average feature vector $\mathbf{F}_{ave}$ of the all B blocks in the image is defined.
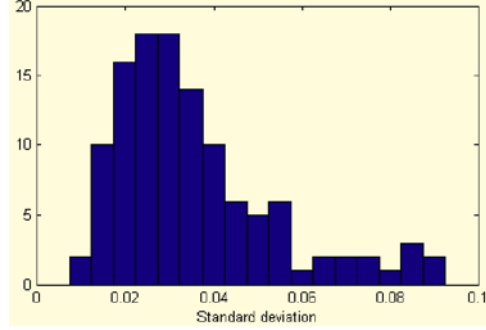
Fig. 2. The distribution of the standard deviation (*std*) values of the test set images
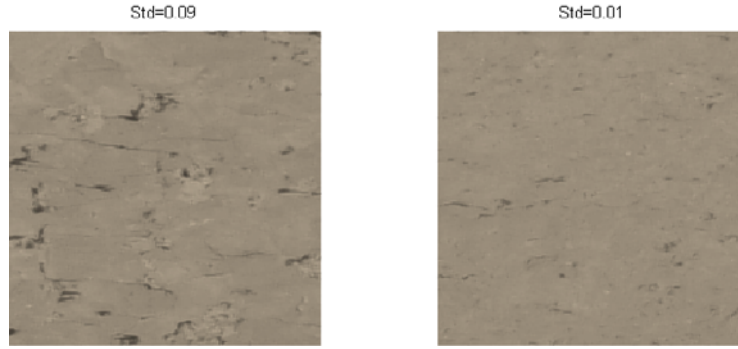
Fig. 3. Example images of non-homogenous and homogenous reflection images

The homogeneity of the image can be measured by means of the deviation between $\mathbf{F}_i$:s and $\mathbf{F}_{ave}$. Hence, in homogeneity measurement we use standard deviation (*std*):

$$std = \frac{1}{B}\sqrt{\sum_{i=1}^{B} D_i^2} \ , \quad \text{in which} \quad D_i = \sum_{ii=1}^{kl} \mathbf{F}_{ave}(ii) - \mathbf{F}_i(ii) \tag{7}$$

Standard deviation is a measure that describes the texture homogeneity.

3   Experiments

For testing purposes we had 118 polished industrial rock plates. The surface of each plate was photographed using a digital camera combined with polarization filter. The imaging arrangement is presented in figure 1b. In this arrangement, a plane of fluorescence tubes illuminated the plate via a white vertical surface. Using this lightning method, the plate surface was evenly illuminated.

We applied to the obtained surface reflection images the texture analysis methods presented in section 2.2. An area of 500x500 pixels was selected from the middle of the plate to represent the surface of the plate. For this region, Gabor wavelet transform of equation 4 was defined using a set of four scales and six orientations. Using these scales and orientations, the transform coefficients μ_{mn} were calculated. Based on these coefficients, the dominating orientation of each plate was defined. Based on the manual inspection of the reflection images, the results of the orientation measurements were valid.

Another goal was to measure the homogeneity of the surface. For this purpose, the images were divided into 25 blocks so that the size of each block was 100x100 pixels. The feature vector $\mathbf{F}_i$ of equation 6 was defined for each block of each image. Then the standard deviation (*std*) between the block feature vectors (equation 7) was calculated for each image. The manual inspection of the images showed that there is a clear relation between the *std*-value and the homogeneity of the surface image. Hence, the low value of *std* means that the surface is homogenous whereas large *std*-value indicates non-homogeneity of the surface. The distribution of the *std*-values of the images is presented in figure 2. Based on this distribution, the thresholds for *std*-values of homogenous and non-homogenous rock plates can be defined. In figure 3, examples of non-homogenous and homogenous surface reflection images are presented.

4   Discussion

In this paper we presented a novel method for the analysis of the surface reflection images. This method is based on the multiscale texture analysis, which makes it possible to analyze the orientations and other textural properties in multiple resolutions. The benefit of this approach is that the defects and microcracks of different sizes can be found.

The application field of this method lies in the rock and stone industry. The surface inspection by image analysis is beneficial in the quality control of the rock plates. It is fast and accurate compared to traditional visual methods. The directionality of surface cracking of the rock plate is an important feature to inspect. On the other hand, the homogeneity of the surface reflection indicates the average smoothness and void content of the plate. Therefore, we applied the surface analysis method to two purposes: directionality and homogeneity measurement. The experimental results show that both of these properties can be inspected from the images using the methods presented in this paper. The method made it possible to classify the test material accurately based on both of these two properties.

The surface reflection method proved to be feasible in the analysis of the rock plate surfaces. The method for the image acquisition is straightforward and the texture analysis tools were able to analyze the desired features in the images. The obtained results show that this method has great potential in rock and stone industry. However, these methods can also be applied to other surface inspection tasks.

Acknowledgment

References

1. Autio, J., Luukkanen, S., Rantanen, L., Visa, A.: The Classification and Characterization of Rock Using Texture Analysis by Co-occurrence Matrices and the Hough Transform. International Symposium on Imaging Applications in Geology, Belgium (1999) 5-8
2. Chui, C.K.: An Introduction to Wavelets. Wavelet Analysis and Its Applications, Vol. 1. Academic press, London (1992)
3. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-3, 6 (1973) 610-621
4. Laine, A. Fan, J.: Texture Classification by Wavelet Packet Signature. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, 11 (1993) 1186-1191
5. Lebrun, V.: Development of Specific Image Acquisition Techniques for Field Imaging - Applications to Outcrops and Marbles-. International Symposium on Imaging Applications in Geology, Belgium (1999) 165-168
6. Lepistö, L., Kunttu, I., Autio, J., Visa, A.: Rock Image Classification Using Non-Homogenous Textures and Spectral Imaging. WSCG Short papers proceedings, WSCG'2003, Plzen, Czech Republic (2003) 82-86
7. Lepistö, L., Kunttu, I., Autio, J., Visa, A.: Retrieval of Non-Homogenous Textures Based on Directionality. Proceedings of the 4th European Workshop on Image Analysis for Multimedia Interactive Services, London, UK (2003) 107-110
8. Manjunath, B.S., Chellappa, R.: A Unified Approach to Boundary Detection. IEEE Transactions on Neural Networks, Vol. 4, 1 (1993) 96-108
9. Manjunath, B.S., Ma, W.Y.: Texture Features for Browsing and Retrieval of image Data. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, 8 (1996) 837-842
10. Mao, J., Jain, A.K.: Texture Classification and Segmentation using Multiresolution Simultaneous Autoregressive Models. Pattern Recognition, Vol. 25, 2 (1992) 173-188

Publication II

Lepistö, L., Kunttu, I., Autio, J., Visa, A., 2003. Classification Method for Colored Natural Textures Using Gabor Filtering.

Classification Method for Colored Natural Textures Using Gabor Filtering

Leena Lepistö[1], Iivari Kunttu[1], Jorma Autio[2], and Ari Visa[1],

[1]Tampere University of Technology
Institute of Signal Processing
P. O. Box 553, FIN-33101 Tampere, Finland

[2]Saanio & Riekkola Consulting Engineers
Laulukuja 4
FIN-00420 Helsinki, Finland

Leena.Lepisto@tut.fi

Abstract

In texture analysis the common methods are based on the gray levels of the texture image. However, the use of color information improves the classification accuracy of the colored textures. In the classification of non-homogenous natural textures, human texture and color perception are important. Therefore, the color space and texture analysis method should be selected to correspond to human vision.

In this paper, we present an effective method for the classification of colored natural textures. The natural textures are often non-homogenous and directional, which makes them difficult to classify. In our method, the multiresolution Gabor filtering is applied to the color components of the texture image in HSI color space. Using this method, the colored texture images can be classified in multiple scales and orientations. The experimental results show that the use of the color information improves the classification of natural textures.

1. Introduction

The analysis of color and texture are both essential topics in image analysis and pattern recognition. Usually texture and color content of the image have been analyzed separately. Hence the conventional texture analysis methods use only the gray level information of the texture image. However, color of the texture image provides a significant amount of information about the image content. Therefore, in several recent studies, color has been taken into account in the analysis of texture images. The experimental results show that the use of the color in texture analysis has improved the texture classification results.

In most of the studies concerning color texture analysis, the commonly used texture analysis methods have been applied to the colored textures. One of the most popular texture analysis methods is based on wavelets [6]. Gabor filtering [9], [11] is a wavelet-based method that provides a multiresolution representation of texture. In the comparison of Manjunath and Ma [9], Gabor filtering method proved to be the most effective wavelet-based method in the texture classification. Gabor filtering has also been the basis of many color texture analysis methods, such as [3], [5], [10].

The choice of the color space is essential in the color texture analysis. The use of RGB color space is common in the image processing tasks. However, it does not correspond to the color differences perceived by humans [13]. In the work of Paschos [10], Gabor filtering was applied to the classification of color textures. He compared the color texture classification in RGB, L*a*b*, and HSI color spaces. The best classification result was obtained using HSI color space.

Most of the natural textures are non-homogenous. Classification of natural non-homogenous textures is significantly more difficult than classification of homogenous textures such as the commonly used texture image set presented by Brodatz [2].

In this type of texture images, there can be variations in directionality, granularity, and other textural features. Color is also an essential feature of natural texture images. Color levels may vary significantly within these images. In [7] we presented a method for the classification of colored natural non-homogenous textures. In this method, the texture image was divided into blocks and the textural and spectral features of these blocks formed a feature histogram. The classification of the texture images was based on these histograms. Many natural texture types, like rock texture [1], are directional. Directionality is also one of the most remarkable dimensions in human texture perception [12]. Therefore, directionality can be used as a classifying feature between natural textures. In [8] we presented a directionality-based method for the retrieval of non-homogenous textures. Because Gabor filtering method can be used in multiple orientations, it is a powerful tool for the analysis of the directional textures. Therefore, it is a suitable method for the classification of these kinds of textures. In addition to its ability to describe

Figure 1. An example image of non-homogenous rock texture.

the texture directionality, Gabor filtering can be used in multiple scales, which is a desirable property in the classification of non-homogenous natural textures.

In this paper, we present an effective method for the use of multiscale Gabor filtering in the classification of colored natural textures. In our method, we apply a bank of Gabor filters presented in [9]. This filter bank is used for the texture images in HSI color space. Compared to the conventionally used gray level Gabor filtering, our approach gives clearly better classification result without significantly increasing the computational cost.

In section 2, the classification of colored natural textures is discussed. In the same section, our method for color-based Gabor filtering is presented. Section 3 is the experimental part of this work. The classification experiments are made using two databases of colored rock textures. The results are discussed in section 4.

2. Classification of colored natural textures

In non-homogenous natural textures, one or more of the texture properties are not constant in the same texture sample. An example image of non-homogenous rock texture is presented in figure 1. The texture sample presented in this figure is strongly non-homogenous in terms of directionality, granularity, and color. The homogeneity of a texture sample can be measured by dividing the sample into blocks. If the texture or color properties do not vary between the blocks, the texture is homogenous. On the other hand, if these feature values have significant variance, the texture sample is non-homogenous. In our previous approach [7], this division into blocks was applied.

In this section, we present a method for the classification of natural non-homogenous textures. Because texture directionality can be used as a classifying feature between the texture samples [8], we use a bank of multiple oriented Gabor filters. The benefit of this approach is also the fact that the textures are considered

in multiple scales, which makes the classification more accurate. In our approach, we apply the filter bank for the colored texture images.

In the beginning of this section, the previous work in the Gabor filtering of color textures is presented. After that, we present our approach to this purpose.

2.1 Gabor filtering of colored textures

Gabor filtering is a wavelet-based method for texture description and classification. Gabor filters extract local orientation and scale information of the texture. These features have been shown to correspond to human visual system [11]. In the classification and retrieval, the common approach is to use a bank of multiple oriented Gabor filters in multiple scales [9].

In the color texture analysis, several Gabor-based methods have been presented. In [10], the Gabor filters have been applied to different color bands of the texture images. After that, the classification is made by calculating distances between the transform coefficients. The work of Jain and Healey [5] is based on the opponent features that are motivated by color opponent mechanisms in human vision. The unichrome opponent features are used in texture image classification. Multichannel Gabor illuminant invariant texture features (MII) combine the ideas of color angle and Gabor filtering [3]. In this approach, MII is calculated as the color angles between the image color bands convolved with two different Gabor filters.

2.2 Our approach

Our approach to the classification of the colored natural textures is based on the Gabor filtering in HSI color space. This color space is selected to be the basis of our texture analysis, because it corresponds to the human visual system [13]. Also the comprehensive comparison presented in [10] proved that HSI color space gives the best result in the classification of color textures.

Manjunath and Ma [9] have introduced a method for the classification of the gray level textures. They have used a bank of Gabor filters to extract features that characterize the texture properties. The Gabor filter bank is used in multiple scales m and orientations n. The feature vector is formed using the mean μ_{mn} and standard deviation σ_{mn} of the magnitude of the transform coefficients. If the number of scales is M and the number of orientations is N, the resulting feature vector is of the form:

$$f = \left[\mu_{00}\sigma_{00}, \mu_{01} \ldots \mu_{MN}\sigma_{MN}\right] \qquad (1)$$

Figure 2. Example images of the textures in the testing database I.

Figure 3. Example images of the textures in the testing database II.

This approach has proved to be effective in the classification and retrieval of different types of gray level textures [9]. Also in the case of natural non-homogenous textures, this method has given reasonably good results. However, when the color information of the texture image is added to this method, the classification results can be improved as shown in the experimental part of this paper.

In our approach, we define the feature vector f for each color channel of the texture image:

$$f_H = \left[\mu_{00}^H \sigma_{00}^H, \mu_{01}^H \ldots \mu_{MN}^H \sigma_{MN}^H\right]$$
$$f_S = \left[\mu_{00}^S \sigma_{00}^S, \mu_{01}^S \ldots \mu_{MN}^S \sigma_{MN}^S\right] \quad (2)$$
$$f_I = \left[\mu_{00}^I \sigma_{00}^I, \mu_{01}^I \ldots \mu_{MN}^I \sigma_{MN}^I\right]$$

Then the feature vectors can be combined to a single vector that characterizes all the color channels:

$$f_{HSI} = \left[f_H, f_S, f_I\right] \quad (3)$$

When μ_{mn} and σ_{mn} are calculated for M scales, N orientations, and C color channels, the size of the resulting feature vector is $2*M*N*C$. In classification, the feature vectors of texture samples i and j are compared using the distance measure:

$$d(i,j) = \sum_m \sum_n d_{mn}(i,j) \quad (4)$$

where

$$d_{mn}(i,j) = \left|\frac{\mu_{mn}^{(i)} - \mu_{mn}^{(j)}}{\alpha(\mu_{mn})}\right| + \left|\frac{\sigma_{mn}^{(i)} - \sigma_{mn}^{(j)}}{\alpha(\sigma_{mn})}\right| \quad (5)$$

in which $\alpha(\mu_{mn})$ and $\alpha(\sigma_{mn})$ are the standard deviations of the respective features over the whole database [9].
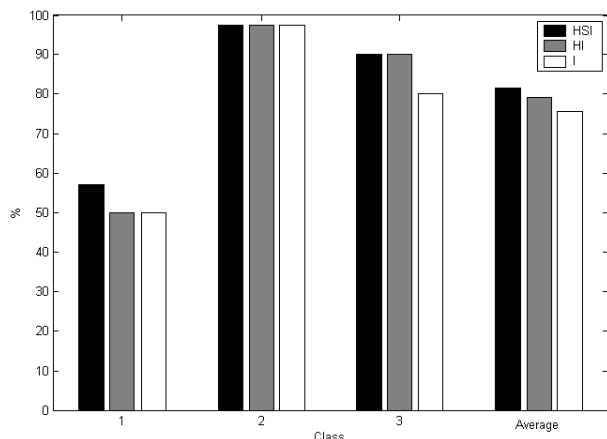
3. Experiments

In this section, our approach to the classification of the colored natural textures is tested. For testing purposes, we used two testing databases. These databases consisted of colored rock textures, which are typical natural textures.
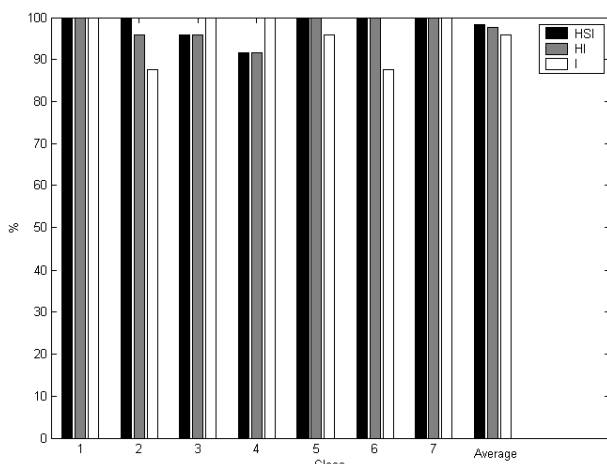
3.1 Testing databases

The two testing databases used contained several types of non-homogenous rock textures. Test-set I (figure 2), consisted of 64 industrial rock images. The size of the images was 714x714 pixels and their texture was strongly directional. The texture samples were divided manually in three visually similar classes. The second testing database, test set II (figure 3), consisted of 168 rock texture samples. The size of each sample image was 500x500 pixels. The test set II represented seven different rock texture types, and there were 24 samples from each texture class. Also in this test set, most of the texture types were directional and non-homogenous.

3.2 Classification

In classification, k-nearest neighbor classification principle [4] was used. The validation of the classification experiments was made using leave one out validation method [4]. In this method, each sample is left out from the database in turn, whereas the rest of the images form the testing database. This classification is repeated for the whole image database, and the average classification rate can be defined as the mean value of these classification experiments. In the classification, the value of k was selected to be 3.

Figure 4. The mean classification results in each class of the testing database I.

Figure 5. The mean classification results in each class of the testing database II.

The classification experiments were made for hue, saturation, and intensity channels (HSI) as well as for hue and intensity channels (HI). For comparison, classification result was calculated also for intensity channel (I), which represents the gray level of the image. In the experiments, we used four scales and six orientations, as in [9].

Table 1 presents the average classification results. These results show that the best classification rate was achieved using HSI color space, whereas hue and intensity channels (HI) gave the second best result. In both testing databases, the classification based on intensity component (gray level) gave the lowest classification result. The mean classification results in each class of both testing databases are presented in figures 4 and 5. The computational characteristics of the methods are presented in table 2. In this table, the feature vector lengths and the classification times are presented for the testing databases. The computation was made

using Matlab on a PC with 804 MHz Pentium III CPU and 256 MB primary memory.

Table 1. The average classification rates.

Color space	Database I	Database II
HSI	87.5 %	98.2 %
HI	85.9 %	97.6 %
I	84.4 %	95.8 %

Table 2. The computational characteristics of the methods.

Color space	Vector length	Classification time	
	$2*M*N*C$	DB I	DB II
HSI	144	1.8 sec	9.1 sec
HI	96	1.7 sec	8.9 sec
I	48	1.4 sec	8.4 sec

4. Discussion

In this paper, we studied the color properties of the natural textures. The color information of the texture images is an essential feature in the classification of them. The results of the texture classification can be improved by combining the color information of the texture image in the classification.

Rock texture is an example of natural textures. Rock texture images are often non-homogenous. The non-homogeneity of these textures may appear as variations in directionality, granularity and color. Because Gabor filters have proved to be effective in the classification of directional textures, they were selected also to this study. Gabor filters can also be used to analyze the texture images in multiple scales, which is desirable in practical applications. This is due to the variations in the granular size of the rock textures. The multiscale texture representation is also essential, when human texture perception is considered.

The objective of this work was to include the color information of the textures to the Gabor-based texture classification. In this way, the classification results obtained from the Gabor filtering could be further improved. The color space selected to be HSI, which corresponds to human color vision. In our texture classification method, the Gabor filtering is applied to the selected color channels separately. Then the feature vectors of each channel are combined into a single vector, which is used in the classification.

Compared to the commonly used, gray level based Gabor filtering, our method provides improved classification results. These results are achievable at a reasonable computational cost. Good computational efficiency is the benefit of the presented method when compared to the other color texture analysis methods.

5. Acknowledgment

6. References

[1] J. Autio, S. Lukkarinen, L. Rantanen, and A. Visa, "The Classification and Characterisation of Rock Using Texture Analysis by Co-occurrence Matrices and the Hough Transform", International Symposium of imaging Applications in Geology, pp. 5-8, Belgium, May. 6-7 1999.

[2] P. Brodatz, *Texture: A photographic Album for Artists and Designers,* Reinhold, New York, 1968.

[3] J. F. Camapum Wanderley and M. H. Fisher, "Multiscale Color Invariants Based on the Human Visual System", *IEEE Transactions on Image Processing*, Vol. 10, No. 11, Nov. 2001, pp. 1630-1638.

[4] R.O. Duda, P.E. Hart, and D.G. Stork, *Pattern Classification*, 2nd edition, John Wiley & Sons, New York, 2001.

[5] A. Jain and G. Healey, "A Multiscale Representation Including Opponent Color Features for Texture Recognition", *IEEE Transactions on Image Processing*, Vol. 7, No. 1, Jan. 1998, pp. 124-128.

[6] A. Laine and J. Fan, "Texture Classification by Wavelet Packet Signature", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 11, Nov. 1993, pp. 1186-1191.

[7] L. Lepistö, I. Kunttu, J. Autio, and A. Visa, "Rock Image Classification Using Non-Homogenous Textures and Spectral Imaging", *WSCG SHORT PAPERS proceedings, WSCG'2003,* Plzen, Czech Republic, Feb. 3.-7. 2003, pp. 82-86.

[8] L. Lepistö, I. Kunttu, J. Autio, and A. Visa, "Retrieval of Non-Homogenous Textures Based on Directionality", *Proceedings of 4th European Workshop on Image Analysis for Multimedia Interactive Services,* London, UK, Apr. 9.-11. 2003. pp. 107-110.

[9] B. S. Manjunath and W. Y. Ma, "Texture Features for Browsing and Retrieval of image Data", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 8, Aug. 1996, pp. 837-842.

[10] G. Paschos, "Perceptually Uniform Color Spaces for Color Texture Analysis: An Empirical Evaluation", *IEEE Transactions on Image Processing*, Vol. 10, No. 6, Jun. 2001, pp. 932-937.

[11] M. Porat and Y. Y. Zeevi, "The Gabor Scheme of Image representation in Biological and Machine Vision", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 4, Jul. 1988, pp. 452-468.

[12] A. R. Rao and G. L. Lohse, "Towards a Texture Naming System: Identifying Relevant Dimensions of Texture", *Proceedings of IEEE Conference on Visualization*, San Jose, California, Oct. 1993, pp. 270-227.

[13] G. Wyszecki and W.S. Stiles,. *Color Science, Concepts and Methods, Quantitative Data and Formulae*, 2nd Edition, John Wiley & Sons, Canada, 1982.

Publication III

Lepistö, L., Kunttu, I., Visa, A., 2005. Rock image classification using color features in Gabor space.

Rock image classification using color features in Gabor space

Leena Lepistö
Iivari Kunttu
Ari Visa
Tampere University of Technology
Institute of Signal Processing
P.O. Box 553
FI-33101 Tampere, Finland
E-mail: Leena.Lepisto@tut.fi

Abstract. *In image classification, the common texture-based methods are based on image gray levels. However, the use of color information improves the classification accuracy of the colored textures. In this paper, we extract texture features from the natural rock images that are used in bedrock investigations. A Gaussian bandpass filtering is applied to the color channels of the images in RGB and HSI color spaces using different scales. The obtained feature vectors are low dimensional, which make the methods computationally effective. The results show that using combinations of different color channels, the classification accuracy can be significantly improved.* © 2005 SPIE and IS&T. [DOI: 10.1117/1.2149872]

1   Introduction

The division of natural images like rock, stone, clouds, ice, or vegetation into classes based on their visual similarity is a common task in many machine vision and image analysis solutions. Classification of natural images is demanding, because in the nature the objects are seldom homogenous. For example, when the images of rock surface are inspected, there are often strong differences in directionality,[1] granularity, or color of the rock, even if the images represented the same rock type. These kinds of variations make it difficult to classify these images accurately. In the current rock imaging applications, rock images analysis is used in, e.g., bedrock investigations. Therefore, effective inspection methods are required to classify the rock images.

Texture is an important feature in the content-based image classification. Also in the analysis of natural images, texture plays a remarkable role. Rao and Lohse[2] indicated that most important perceptual dimensions in the natural texture discrimination are repetitiveness, directionality, and granularity. The directionality and granularity of nonhomogenous natural textures have been discussed in our earlier work.[1,3] In addition to texture, color is also an essential feature of natural images. In this study, we combine the color information to the textural features of rock images.

Gabor filtering provides a multiresolution representation of texture. In the comparison of Manjunath and Ma,[4] Gabor filtering method proved to be the most effective method in the texture classification. Gabor filtering has also been the basis of many color texture analysis methods.[5,6] In this paper, we present an efficient approach to the classification of colored rock textures. The method is based on the bandpass filtering in Gabor space that is applied to different color channels of the images. In Sec. 2, we present the principle of our method. In Sec. 3, the method is used to classify rock images obtained from the boreholes. The results are discussed in Sec. 4.

2   Color Filtering in Gabor Space

Gabor filtering is a method for texture description and classification. In most cases, the filters are used to extract orientation and scale information from the local spectrum of the texture image. The local spectrum is the Fourier transform of a window function which is multiplied with the Fourier transform of the image. The filters are used to estimate the selected frequency band of the image using a Gaussian as a smoothing window function. It has been shown that Gabor features correspond to human visual system.[7]

The filters as texture analysis tools are usually applied to gray-level (intensity) images in the Gabor space. In our previous approach,[8] we showed that Gabor filters applied to color channels of the rock texture images can improve the classification accuracy of these images. In this paper, we use rock textures that are not directional or their directionality cannot be regarded as classifying feature. Therefore, we do not utilize the orientation of the texture. We apply filters of different scales to the color channels of the texture images. This way, the obtained feature vectors are shorter than in Ref. 8, which makes the feature extraction and classification significantly faster. Hence, instead of using filters of multiple scales and orientations, we use a filter bank that works independent on the orientation at a selected scale.

In this work, we use ring-shaped bandpass filters whose amplitude responses are presented in Fig. 1. The cross section of the ring is a Gaussian function. The feature vector is formed using the mean μ_m and standard deviation σ_m of the magnitude of the transform coefficients. This is repeated at each scale m. If the number of scales is M, the resulting feature vector is of the form:

$$f = [\mu_1 \sigma_1, \mu_2 \sigma_2 \ldots \mu_M \sigma_M]. \tag{1}$$

A comprehensive comparison presented in Ref. 6 revealed that HSI color space gives the best result in the classification of color textures. This comparison, however, used quite homogenous textures and oriented Gabor filters. In this paper, we compare the results obtained in RGB space with those obtained from HSI space in rock texture filtering without orientation. In our approach, we define the feature vector f for each color channel of the texture image:

$$f_H = [\mu_1^H \sigma_1^H, \mu_2^H \sigma_2^H \ldots \mu_M^H \sigma_M^H]$$
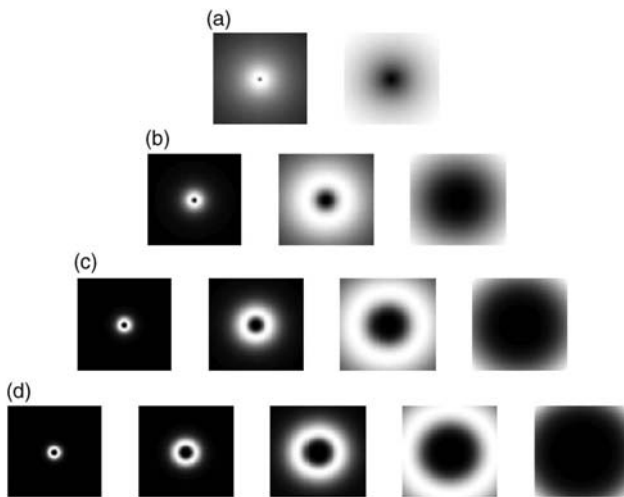
$$f_S = [\mu_1^S \sigma_1^S, \mu_2^S \sigma_2^S \ldots \mu_M^S \sigma_M^S]$$

$$f_I = [\mu_1^I \sigma_1^I, \mu_2^I \sigma_2^I \ldots \mu_M^I \sigma_M^I]. \tag{2}$$

Then the feature vectors can be combined to a single vector that characterizes all the color channels:

Fig. 1 The filters used at (a) two, (b) three, (c) four, and (d) five scales.

Fig. 2 Three examples from each class of the rock images in the testing database.

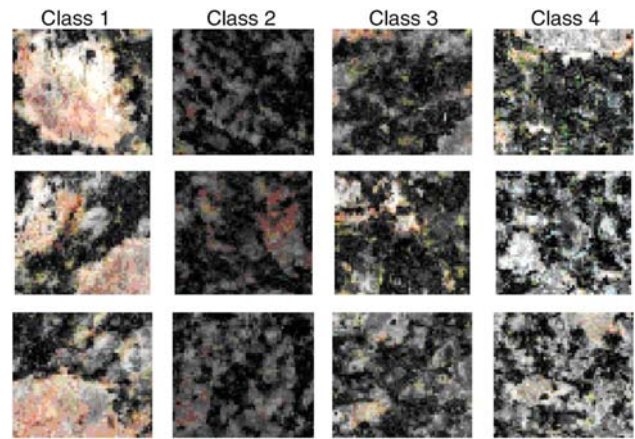$$f_{HSI} = [f_H, f_S, f_I] \tag{3}$$

in which each component is normalized by removing its mean and dividing by its standard deviation. When μ_m and σ_m are calculated for M scales and C color channels, the size of the resulting feature vector is 2^*M^*C, which yields to quite short feature vectors, especially when the number of scales is low. The same procedure is followed with the experiments in HSI and RGB color spaces.

3 Experiments Using Rock Images

3.1 Nonhomogenous Rock Images

In the field of rock science, the development of digital imaging has made it possible to store and manage the images of the rock material in digital form. Rock represents typical example of nonhomogenous natural image type. This is because there are often strong differences in directionality, granularity, or color of the rock texture, even if the images represented the same rock type.[8] In bedrock investigation, rock properties are analyzed by inspecting the images that are collected from the bedrock. Different rock layers can be recognized from the borehole images based on the color and texture properties of rock. Therefore, there is a need for an automatic classifier that is capable of classifying the rock images into visually similar classes.

As a testing database, we use a set of rock images that consists of 336 images, which are obtained by dividing large borehole images into parts. These images are manually divided into four classes by an expert. The division is based on their color and texture properties. Figure 2 presents three example images from each of the four classes. The images show that there is directionality in some of the classes, but the orientations vary within the same classes. Therefore, directionality is not used in the classification. In classes 1–4, there are 46, 76, 100, and 114 images in each class, respectively.

3.2 Classification

The database of rock images is classified using feature vectors of Eq. (2) in different color channels. The number of scales varyies between two and five. In all cases, the filter centers have been located uniformly to the frequency band such that the frequency channels of the filters cover the whole frequency area. The corresponding filters are presented in Fig. 1. In classification, we have used k-nearest neighbor (k-NN) classification principle. The selection of the k-NN classifier is due to its robustness with nonhomogenous feature distributions of the rock images. With this type of database, the selected classification algorithm is also fast. The selection of value 5 for k was based on preliminary experiments. In the experiments, a leave-one-out validation method was employed. The distance measure in the classification was Euclidean distance. In preliminary experiments, Euclidean distance outperformed slightly L_1-norm that is other common distance metrics for texture features.

Table 1 The average classification rates (%) in each class in RGB and HSI color spaces.

	Dimensions	RGB color space					HSI color space				
		1	2	3	4	Ave	1	2	3	4	Ave
2 scales	12	60.5	93.5	86.0	82.5	80.1	47.4	84.8	81.0	81.6	74.1
3 scales	18	59.2	95.7	87.0	86.0	81.3	56.6	87.0	80.0	82.5	76.5
4 scales	24	55.3	100.0	86.0	84.2	80.4	52.6	80.4	85.0	79.8	75.3
5 scales	30	53.9	100.0	88.0	86.0	81.3	52.6	82.6	81.0	77.2	73.5

Table 2 The average classification rates (%) in each class using MPEG-7 color and texture descriptors.

MPEG-7 Descriptor	Dimensions	1	2	3	4	Ave
Homogenous Texture Descriptor	62	40.8	93.5	48.0	73.7	61.3
Color Layout Descriptor	12	35.5	97.8	83.0	71.1	70.3
Color Structure Descriptor	256	56.6	89.1	85.0	84.2	78.9
Scalable Color Descriptor	256	48.7	76.1	85.0	86.8	76.2

The average classification results are presented as percentages in Table 1. The results are presented for RGB and HSI color spaces, respectively. In Table 1, the average classification rates are presented for each of the four classes separately and as average value. The dimensionality of each descriptor type is also mentioned in the table. To compare our results to other commonly used visual descriptors, we have calculated the classification results for the testing database using some MPEG-7 texture and color descriptors.[9] This comparison is presented in Table 2. We selected the homogenous texture descriptor to represent texture description, because it is based on Gabor filtering in gray-level texture images.[9] This descriptor is based on the method presented in Ref. 4, and it uses Gabor filters in five scales and six orientations. Because this paper considers also color information of the rock images, we have included also MPEG-7 color descriptors for comparison. Color structure descriptor and color layout descriptor employ HMMD color space[9] whereas scalable color descriptor uses HSI color space.

3.3 Results

The classification rates presented in Table 1 show that the rock texture filtering in RGB color space produces slightly better classification results than that in HSI color space. There are remarkable differencies in the classification performance between the classes. Class 1 is especially difficult to classify for all the features. This is due to the nonhomogenous nature of the class 1 images. These images are very varying in terms of their color distributions and texture properties. In class 2, RGB color space gives clearly better results than HSI space. On the other hand, in classes 3 and 4, RGB space is only slightly better. When the comparison with MPEG-7 visual descriptors is considered (Table 2), one can see that these descriptors are outperformed by the proposed methods. Homogenous texture descriptor uses directional Gabor filtering, which yields poorer classification results than the proposed methods. This is due to the fact that especially in classes 1 and 3 the textures are randomly oriented and therefore directionality cannot be regarded as a classifying feature. In addition, color information of texture has not been utilized in this descriptor. The performance of color descriptors is also lower than in the case of the proposed methods. This is natural, because they consider only color distribution of the images and not their texture content.

Computational complexity is always a central matter with practical image classification tasks. Compared to conventional Gabor filtering,[4,9] the proposed method is somewhat lighter because it does not calculate the filter responses for different orientation. On the other hand, the filtering is repeated for three color channels instead of one. Therefore the computational cost is dependent on the number of scales and color channels. In fact, the computational complexity can be estimated by comparing the dimensionality of the descriptors. In Tables 1 and 2, the dimensionality of each method is presented. For example, by using two scales the dimensionality of the proposed method is 12, which yields to a classification rate of 80.1% in RGB space. This can be regarded as a good result with such a low dimensional descriptor.

4 Discussion

In this paper, we showed that the classification of natural rock texture images can be improved by combining the color information to the texture description. We used bandpass filtering that was applied to the images in different color spaces. This way it is possible to analyze colored texture images in multiple scales, which is desirable in practical applications. This is due to the variations in the granular size of the rock textures. In the practical solutions, the computational cost is always an essential matter. In the presented approach the filtering is a straightforward operation that is repeated for the selected color channels. The obtained feature vectors are relatively short, which makes online classification possible.

References

1. L. Lepistö, I. Kunttu, J. Autio, and A. Visa, "Retrieval of non-homogenous textures based on directionality," *Proc. 4th European Workshop Image Analysis for Multimedia Interactive Services*, pp. 107–110 (2003).
2. A. R. Rao and G. L. Lohse, "Towards a texture naming system: identifying relevant dimensions of texture," *Proc. IEEE Conf. Visualization*, pp. 270–227 (1993).
3. L. Lepistö, I. Kunttu, J. Autio, and A. Visa, "Rock image retrieval and classification based on granularity," *Proc. 5th Inlt. Workshop Image Analysis for Multimedia Interactive Services* (2004).
4. B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(8), 837–842 (1996).
5. J. F. Camapum Wanderley and M. H. Fisher, "Multiscale color invariants based on the human visual system," *IEEE Trans. Image Process.* **10**(11), 1630–1638 (2001).
6. G. Paschos, "Perceptually uniform color spaces for color texture analysis: an empirical evaluation," *IEEE Trans. Image Process.* **10**(6), 932–937 (2001).
7. M. Porat and Y. Y. Zeevi, "The Gabor scheme of image representation in biological and machine vision," *IEEE Trans. Pattern Anal. Mach. Intell.* **10**(4), 452–468 (1998).
8. L. Lepistö, I. Kunttu, J. Autio, and A. Visa, "Classification method for colored natural textures using Gabor filtering," *Proc. 12th Intl. Conf. Image Analysis Processing*, pp. 397–401 (2003).
9. B. S. Manjunath, P. Salembier, and T. Sikora, *Introduction to MPEG-7 Multimedia Content Description Interface*, John Wiley & Sons, UK (2002).

Publication IV

Lepistö, L., Kunttu, I., Autio, J., Visa, A., 2003. Classification of Non-homogenous Textures by Combining Classifiers.

CLASSIFICATION OF NON-HOMOGENOUS TEXTURE IMAGES BY COMBINING CLASSIFIERS

Leena Lepistö[1], Iivari Kunttu[1], Jorma Autio[2], and Ari Visa[1]

[1]Tampere University of Technology
Institute of Signal Processing
P. O. Box 553, FIN-33101 Tampere, Finland

[2]Saanio & Riekkola Consulting Engineers
Laulukuja 4
FIN-00420 Helsinki, Finland

E-Mail: Leena.Lepisto@tut.fi

ABSTRACT

Most of the natural textures are non-homogenous. In non-homogenous texture images, the textural features may have strong variations. These variations cause errors in the classification of these images. In this paper we present a novel method for classification of the non-homogenous textures. The classification method is based on the combination of separate classifiers. The outputs of the separate classifiers are collected into a classification result vector (CRV). This vector is used in the final classification of the texture samples. Using this method, the classification errors caused by variations of feature values can be minimized. The method is tested using non-homogenous rock texture images. The results show that our method is suitable for classifying non-homogenous texture samples. It also gives better classification results than the commonly used methods for combining classifiers.

1. INTRODUCTION

Classification of textures is an important task in image processing and pattern recognition. In the texture classification, similar texture images are classified in predefined classes. When the classification criteria for the textures are defined, it is essential to consider the human texture perception. Rao and Lohse [13] have indicated that three most important perceptual dimensions of the natural texture discrimination are repetitiviness, directionality, and granularity. These properties of the textures have been widely used in the texture classification.

A number of different methods for the texture classification have been developed. One of the most popular statistical tool in the texture classification is co-occurrence matrix [4]. Also model-based features have been a common approach. Multiresolution autoregressive model (MRSAR) [11] is an example of these models.

During the last decade, signal processing tools have been widely used in the analysis of textures. Most of these tools are based on some filtering technique. Commonly used filtering-based methods use wavelets [7] or Gabor features [10]. Manjunath and Ma [10] have made a comparison between multiple oriented Gabor filters, other wavelet-features, and MRSAR-model. In this comparison, the best results were achieved using Gabor-filtering. Also Randen and Husoy [12] have made a comprehensive comparison between currently used filtering techniques in the texture classification.

The majority of the methods developed in this field use a set homogenous texture samples as a testing database. A commonly used testing database consists of Brodatz textures [2], which are typical homogenous textures. However, most of the natural textures are strongly scale dependent, non-homogenous and directional. In the non-homogenous textures, features like directionality and granularity are not constant in every part of texture sample. Also color of the texture sample may have variations.

In this work we present a new approach to the classification of the non-homogenous textures. Our classification method is based on the combination of separate classifiers. In our approach, we extract from the non-homogenous texture samples several features that characterize their color and textural content. The classification of the samples is made based on the features separately. After that, the feature vector of each sample is formed by means of the separate classification results. The final classification is made using these feature vectors. In this way the non-homogenous textures can be classified effectively.

In section two we define the classification problem of the non-homogenous textures. The special character of the non-homogenous textures and the means for the classification of them are presented. At the end of section two, our approach is presented. Section three is the experimental part of this work. In that section, the classification experiments and results are presented. In section four the results of our method are discussed.
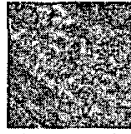
Figure 1. An example of non-homogenous texture.

2. CLASSIFICATION OF NON-HOMOGENOUS TEXTURES

Most of the natural textures are non-homogenous. In these textures one or more of the texture properties are not constant in the same texture sample. An example image of a non-homogenous rock texture is presented in figure 1. The texture sample presented in this figure is strongly non-homogenous in terms of directionality, granularity, and color. The homogeneity of a texture sample can be measured by dividing the sample into blocks. If the texture or color properties do not vary between the blocks, the texture is homogenous. On the other hand, if these feature values have significant variance, the texture sample is non-homogenous. In our previous approach [9], this division into blocks was applied.

2.1. Classification methods

A common approach to the texture image classification is to extract some features from the texture sample. These features are based on some specific texture analysis method, such as [4], [7], [10], [11]. The features form a feature vector and the classification is made using some classification principle. Different types of classifiers are presented in [3]. The commonly used classification principles for textures are presented by Randen and Husoy [12]. In case of the non-homogenous textures, commonly used texture classifiers are not suitable approach, because the feature values may have strong variations in different parts of the texture sample. This may cause errors in classification. In [9], we divided the texture samples into blocks. The textural and color features were calculated from each block separately. Then the feature values of each sample were presented as feature histograms. The classification was made using these feature histograms.

One approach to solve the classification problem of non-homogenous textures is to combine separate

classifiers. In [5] several methods are presented for this purpose. Kittler et al. [6] have made a comparison between different methods for combining classifiers. In their comparison, the best results were achieved using majority vote and averaging. In voting-based approach, the majority of the separate classification results define the final classification result. Averaging means that the outputs of the classifiers are averaged to obtain the final result. Averaging can be made using either mean or median rule [6].

2.2. Combining separate classifiers using classification result vector

We use a novel approach to the classification. In this approach, the results of the separate classifiers are used to form a feature vector, classification result vector (CRV), for a texture sample. The classification system is presented in figure 2. The classification procedure presented in figure 2 consists of five steps:

1. Feature extraction: Certain textural or color features are extracted from the image.
2. Separate classification: The image is classified using each feature separately.
3. Result of the separate classification: The result of each classification is the number of class, to which the classifier assigns the sample image. These results are marked with R1, R2, ... ,Rn.
4. Classification result vector, CRV: The separate classification results are combined by forming a feature vector. This vector consists of classification results R1, R2, ... ,Rn.
5. Final Classification of the sample image based on the CRV.

In this approach, the features extracted from the texture image are used only in classification only in step two. In case of the non-homogenous textures, the classification based directly on the features does not give the best possible result. However, the separate classifiers classify similar samples in similar way. Therefore, in our approach we use the outputs of the separate classifiers as a new feature vector, CRV. In contrary to the voting and averaging, in CRV method the individual features do not affect the final classification result. This is an important property in the classification of non-homogenous textures.
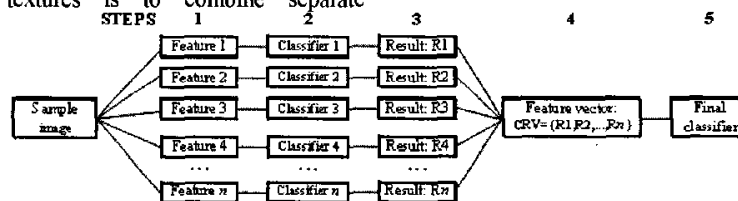
STEPS   1   2   3   4   5
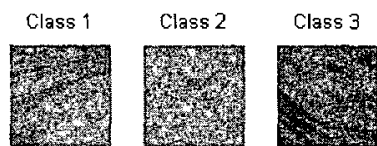

Figure 2. Classification process in five steps.

Class 1   Class 2   Class 3

Figure 3. Example samples of test set I textures.
Class1   Class2   Class3
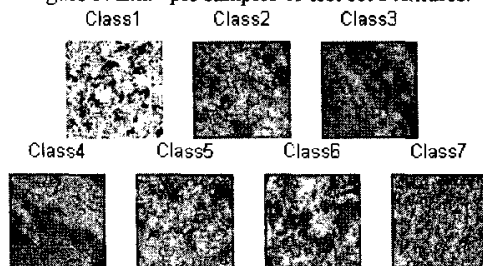

Class4   Class5   Class6   Class7

Figure 4. Example samples of test set II textures.

3. EXPERIMENTS

In the experimental part of this work we test our classification method using non-homogenous rock texture images. We make a practical comparison between our classification method (CRV) and commonly used majority voting and averaging methods for combining classifiers.

3.1. Rock texture images

Rock texture is an example of natural textures. In many cases it is non-homogenous. There is a practical need for methods for classification of rock textures, because today rock and stone industry uses digital imaging for rock analysis. Using image analysis tools, visually similar rock texture images can be classified. For this purpose, effective classification method for non-homogenous rock textures is needed. In [8] we developed some tools for the rock texture retrieval based on texture directionality and co-occurrence features. In addition to the textural features, the color of the rock images was considered in [9].

In this work we used two rock image databases. Test set I (figure 3), consisted of 64 industrial rock images. The size of the images was 714x714 pixels and their texture properties were non-homogenous in terms of color and texture properties. Especially directionality was strong and varying in these samples. The images in test set I were divided into three visually similar classes. The second testing database, test set II (figure 4), consisted of 168 rock texture samples. The size of each image was 500x500 pixels. The test set II consisted of seven different rock texture types, and there were 24 samples from each class. Also in this test set, most of the texture types were directional and non-homogenous.

3.2 Feature extraction

In [9] we showed that HSI-color space is suitable for characterizing the color of the rock samples. An usual way to characterize the image color content is to use the image histogram. The histogram represents the color distribution of the image and therefore, it is a suitable measure to be used in this work. In this work, we used hue and gray level histograms of the rock samples to characterize their color and gray level distribution. In histogram-based classification, the distance was L_1-norm.

For testing purposes, we selected several types of textural features. These features characterize the most important textural properties and their calculation is based on the different principles. In this way we obtain several different types of texture descriptors for a sample image.

Gray level co-occurrence matrix [4] is one of the most popular statistical tools in texture classification. The repetitiveness and granularity of the texture can be measured using this method. In [8] we found that contrast, entropy and energy were the most effective of co-occurrence features in rock texture classification. In classification, the feature vectors containing these three features can be compared using Euclidean distance.

Gabor features have proved to be effective in different texture classification tasks. The features are rotation variant and therefore, they can be used in classification of directional textures. In this study, the Gabor features were calculated from the texture images using the method presented in [10]. In this method a Gabor filter bank is used in four scales and six orientations. The distance between the Gabor-features was defined using similarity measure presented in [10].

A tool for texture directionality measurement is directional histogram [1],[8]. In this method, a set of directional masks is used to filter the texture image. The directionality is found based on the projections of these masks on the image. The directions form a distribution that is called directional histogram. The directional histogram can also be calculated using Hough transform [1]. In this work, we used directional histograms calculated both filtering method and Hough transform method. The distance measure for directional histograms was L_1-norm.

3.3 Classification

In the combination of the classifiers, the classification principle can be selected depending on the application. In our work we used nearest neighbor classifier [3] and leave one out validation principle [3] both in step 2 and 5.

First the classification was made for each feature separately (step 2), then the classification results were collected to the feature vector (CRV) as in step 5. In the

	Test set I	Test set II
Separate Features	Classification rate	Classification rate
Hue-histogram	90.63%	89.28%
Gray-histogram	85.94%	91.07%
Co-occurrence	59.38%	66.66%
Gabor-features	89.06%	95.83%
Directionality (filtering)	85.94%	79.76%
Directionality (Hough)	82.81%	67.86%
Classifier combinations		
CRV	92.19%	99.40%
Majority voting	90.63 %	81.55%
Mean	89.06 %	61.31 %
Median	87.50 %	91.67 %

Table 1. The mean results of classification.

CRV-based classification, the similarity between the CRV:s was indicated by calculating the number of similar values in the vectors. The mean results of the nearest neighbor classification using both test sets are presented in table 1. The table presents mean classification result using separate classifiers (step 2). In the lower part of the table 1, the mean classification results using combined classifiers are presented. In this case, classification results obtained from CVR-classification are compared to the results of the majority voting and averaging (mean and median). The computing time for CRV-based classification of test set II using CRV was 1.3 sec. using Matlab on a PC with 804 MHz Pentium III CPU and 256 MB primary memory.

4. DISCUSSION

In this paper we presented a method for supervised classification of the non-homogenous texture images. This method is based on a novel way to combine separate classifiers. We tested this method using two sets of non-homogenous rock texture images. The results presented in table 1 show that there are significant differences in the classification ability of separate classifiers, which are based on the individual features. This fact is mainly due to the non-homogenous nature of the texture images, which makes it difficult to make a correct classification between them. Our solution to this problem is the use of the combination of the separate classifiers. For this purpose we presented a new method, classification result vector (CRV). In this method, the outputs of the separate classifiers are used as features in a new feature space. Then the final classification is made in this new feature space. The results obtained using the classifier combinations gave better classification results than the separate classifiers. Compared to the commonly used classifier combination techniques, majority voting and averaging, CRV gives the best classification results.

Because the most of the natural textures are non-homogenous, the analysis of them is an important part of the texture image analysis. In this work we used rock textures, which are typical natural textures having strong

differences in textural properties and color. The results obtained in this work show that combining classifiers is an effective way to classify the non-homogenous texture images with great potential in industrial applications.

6. REFERENCES

[1] J. Autio, S. Lukkarinen, L. Rantanen, and A. Visa, "The Classification and Characterization of Rock Using Texture Analysis by Co-occurrence Matrices and the Hough Transform," *International Symposium on Imaging Applications in Geology*, pp. 5-8, Belgium, May 6-7, 1999.

[2] P. Brodatz, *Texture: A photographic Album for Artists and Designers*, Reinhold, New York, 1968.

[3] R.O. Duda, P.E.Hart, and D.G. Stork, *Pattern Classification*, 2nd ed., John Wiley & Sons, New York, 2001.

[4] R.M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 3, No. 6, 1973.

[5] A.K. Jain, R.P.W. Duin, and J. Mao, "Statistical Pattern Recognition: A Review," *IEEE Transactions, on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, pp. 4-37, 2000.

[6] J. Kittler, M. Hatef, and R.P.W. Duin, "Combining Classifiers", *Proceedings of 13th International Conference on Pattern Recognition*, Vol. 2, pp. 897-901, Aug. 1996.

[7] A. Laine and J. Fan, "Texture Classification by Wavelet Packet Signature," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 11, pp. 1186-1191, 1993.

[8] L. Lepistö, I. Kunttu, J. Autio, and A. Visa, "Comparison of Some Content-Based Image Retrieval Systems with Rock Texture Images," *Proceedings of 10th Finnish AI Conference*, Oulu, Finland, pp. 156-163, Dec. 16-17, 2002.

[9] L. Lepistö, I. Kunttu, J. Autio, and A. Visa, "Rock Image Classification Using Non-Homogenous Textures and Spectral Imaging," *WSCG SHORT PAPERS Proceedings, WSCG '2003*, Plzen, Czech Republic, pp. 82-86, Feb. 2003.

[10] B.S. Manjunath and W.Y. Ma, "Texture Features for Browsing and Retrieval of image Data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 8, pp. 837-842, 1996.

[11] J. Mao and A.K. Jain, "Texture Classification and Segmentation using Multiresolution Simultaneous Autoregressive Models," *Pattern Recognition*, Vol. 25, No. 2, pp. 173-188, 1992.

[12] T. Randen and J.H. Husoy, "Filtering for Texture Classification: A Comparative Study," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 4, pp. 291-310, 1999.

[13] A.R. Rao and G.L. Lohse, "Towards a Texture Naming System: Identifying Relevant Dimensions of Texture," *Proceedings of IEEE Conference on Visualization*, San Jose, California, pp. 220-227, Oct 1993.

Publication V

Lepistö, L., Kunttu, I., Autio, J., Rauhamaa, J., Visa, A., 2005. Classification of Non-homogenous Images Using Classification Probability Vector. In Proceedings of IEEE International Conference on Image Processing,.

CLASSIFICATION OF NON-HOMOGENOUS IMAGES USING CLASSIFICATION PROBABILITY VECTOR

Leena Lepistö[1], Iivari Kunttu[1], Jorma Autio[2], Juhani Rauhamaa[3], and Ari Visa[1]

[1]Tampere University of Technology P.O. Box 553 FI-33101 Tampere, Finland	[2]Saanio & Riekkola Consulting Engineers Laulukuja 4 FI-00420 Helsinki, Finland	[3]ABB Oy Process Industry P.O. Box 94 FI-00381 Helsinki, Finland

E-Mail: Leena.Lepisto@tut.fi

ABSTRACT

Combining classifiers has proved to be an effective solution to several classification problems. In this paper, we present a classifier combination strategy that is based on classification probability vector, CPV. In this approach, each visual feature extracted from the image is first classified separately, and the probability distributions provided by separate classifiers are used as a basis of final classification. This approach is particularly suitable for images with non-homogenous and overlapping feature distributions.

1. INTRODUCTION

Classification of real-world images is an essential task in image processing. It is usual that features are spread in many difficult ways. In most of the real classification problems the input patterns can be noisy, non-homogenous and overlapping. Different classifiers may classify the same sample in different ways and hence there are differences in the decision surfaces. However, it has been found that a consensus decision of several classifiers can give better accuracy than any single classifier [1],[2],[7],[8]. Therefore, combining classifiers has become a popular research area.

The goal of combining classifiers is to form a consensus decision based on opinions provided by different base classifiers. Duin [5] presented six ways, in which consistent set of base classifiers can be generated. In the base classifiers, there can be differences in initializations, parameter choices, architectures, classification principle, training sets, or feature sets. Combined classifiers have been applied to several classification tasks, for example to face recognition [12], person identification [3] and fingerprint verification [6]. Kittler et al. [7] presented a theoretical framework for combining classifiers.

In image classification, a number of visual descriptors are used to classify the images based on their content. In the images, there are different types of visual features, like color, texture and shape. The feature space is typically high dimensional. Also the categories of images are often overlapping in the feature space. A common approach is to combine all the selected descriptors into a single feature vector. The similarity between these vectors is defined using some distance metric and the most similar images are then classified (labeled) to the same category. However, when different types of features are combined into the same feature vector, some large-scaled features may dominate the distance, while the other features do not have the same impact on the classification. Especially, in the case of high dimensional descriptors combination into the same vector can be problematic. Therefore it is often more reasonable to consider each descriptor separately.

In image classification, separate classifiers can be used to classify each visual feature individually [9]. The final classification can be obtained based on the combination of separate base classification results. In this way, each feature has its own effect on the final classification, independent of its scaling. Hence the non-homogenous properties of individual features do not necessarily effect directly on the final classification. In several image classification problems, this approach can be used to improve the classification [9],[10].

In this paper, we present a novel method for the classification of non-homogenous real-world images using combined classifiers. The proposed combination method is based on the probabilities provided by base classifiers employing separate visual descriptors. The final classification is then carried out using the probability distribution provided by the base classifiers. The rest of this paper is organized as follows. Section two presents the idea of classifier combinations in image classification and the proposed method, classification probability vector (CPV). In section three, the method is tested using databases of real non-homogenous images. The results are discussed in section four.
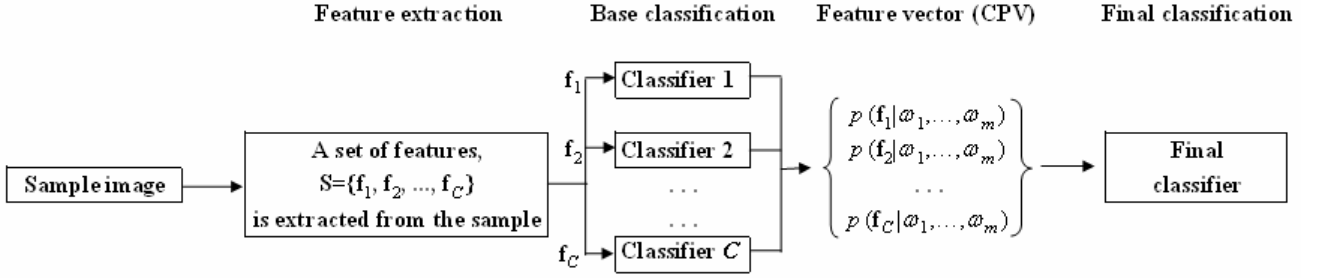
Figure 1. The outline of the proposed classifier combination scheme.

2. CLASSIFIER COMBINATIONS IN IMAGE CLASSIFICATION

2.1. Methods for Combining Classifiers

The general methods for combining classifiers can be roughly divided into two categories, voting-based methods and the methods based on probabilities. The voting-based techniques are popularly used in pattern recognition [11]. Voting has proved to be a simple and effective method for combining classifiers in several classification problems. Furthermore, the voting-based methods do not require any additional information, like probabilities, from the base classifiers [11]. Lepistö et al. [9] presented a method for combining classifiers using classification result vectors (CRV). In this approach, the class labels provided by the base classifiers are used as a feature vector in the final classification, and hence the result is not based on direct voting. CRV method outperformed voting method in image classification experiments [9],[10]. In [10] an unsupervised variation of CRV was presented.

Recently, the probability-based classifier combination strategies have been popularly used in pattern recognition. In these techniques, the final classification is based on the a posteriori probabilities of the base classifiers. Kittler et al. [7] presented several common strategies for combining base classifiers. These strategies are e.g. product rule, sum rule, max rule, min rule, and median rule. In [7], the best experimental results were obtained using sum and median rules. Theoretical comparison of the rules has been carried out in [8]. Also Alkoot and Kittler [1] have made a comparison between the classifier combination strategies.

2.2. Classification probability vector method

Our previous method for combining classifiers combined the outputs of the base classifiers into a feature vector that is called classification result vector (CRV) [9]. However, CRV method uses only the class labels provided by the base classifiers, and ignores their probabilities. In this paper, we use the probability distributions of the separate base classifiers as features in the final classification.

In general, in the classification problem a pattern S is to be assigned to one of the m classes ($\omega_1,\ldots,\omega_m$) [7]. We assume that we have C classifiers each representing a particular descriptor and we denote the feature vector used by each classifier by $\mathbf{f}_i$. Then each class ω_k is modeled by the probability density function $p(\mathbf{f}_i|\omega_k)$. The well-known Bayesian decision theory [4],[7] defines that S is assigned to class ω_j if the a posteriori probability of that class is maximum. However, the probabilities of all the other classes than ω_j have also significance in classification. They are particularly interesting when the pattern S is located near the decision surface. Therefore, we focus on the whole probability distribution $p(\mathbf{f}_i|\omega_1,\ldots,\omega_m)$ provided by each classifier. Hence, if the probability is defined for each m class, the obtained probability distribution is a C by m matrix for each pattern S. This matrix is used as a feature vector in the final classification, and it is called classification probability vector (CPV). In the final classification, the images with similar CPV:s are assigned into same classes. The outline of the CPV method is presented in figure 1.

In contrary to the common probability-based classifier combinations [7], CPV method uses the whole probability distribution as a feature vector in the final classification. The CPV method utilizes the fact that the separate base classifiers classify similar samples in the similar way, which leads to a similar probability profile. The final classification is based merely on the similarity between the probabilities of the base classifiers. Hence, in contrary to voting, in the CPV method the base classifier outputs (class labels) do not directly affect the final classification result.

When image classification is considered, the CPV method has several advantages. CPV method considers each visual descriptor of the images in the base classifiers separately. In the final classification, the probability distributions are employed instead of features. This way, the individual features do not directly affect the final classification result. Therefore, classification result is not sensitive to variations and non-homogeneities of single images.
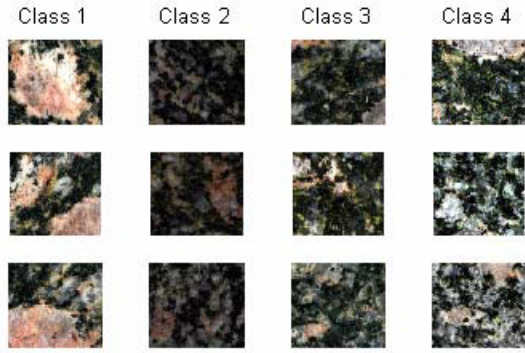
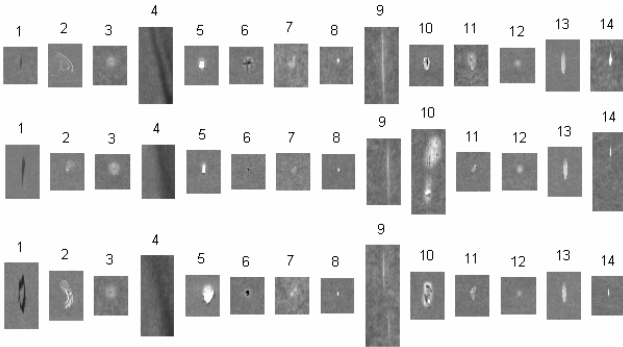Figure 2. Three example images of each class of rock images in testing database I.

Figure 3. Three example images of each class of paper defect images in testing database II.

3. EXPERIMENTS

3.1. Testing databases

The experiments in this paper are focused on non-homogenous image data that is represented by two real image databases. The database I consists of rock images. Rock texture is an example of natural textures. In many cases it is non-homogenous [9]. One application field for rock imaging is geological research work, in which the rock properties are inspected using borehole imaging. Different rock layers can be recognized from the borehole images based on the color and texture properties of rock. Therefore, there is a need for an automatic classifier that is capable of classifying the borehole images into visually similar classes. Testing database I consists of 336 images that are obtained by dividing large borehole images into parts. These images are manually divided into four classes by an expert. The division is based on their color and texture properties. Figure 2 presents an example image of each four class. Testing database II contains defect images that are collected from paper manufacturing process using a paper inspection system [14]. The reason for the collection of the defect image databases in the paper industry is the practical need of controlling the quality and production [14]. The accurate classification of the defect images into classes based on

the defect type they represent is also essential. The defects can be for example holes, wrinkles, or different kinds of dirt spots. The defects in the images are typically very varying in terms of their size, shape, and gray level, which make them non-homogenous. The testing database II consists of 1204 paper images, which represent 14 defect classes. Example images of each defect class are presented in figure 3.

3.2. Classification experiments

For the database I, we used five descriptors as input features $\mathbf{f}_i$. The descriptors were color layout, homogenous texture and edge histogram descriptors of MPEG-7 standard [13] as well as color and gray level histograms. In the case of the database II, MPEG-7 color layout, scalable color, color structure homogenous texture and edge histogram descriptors were used. Furthermore, defect shapes were described using Fourier descriptors. The classification was made for each feature separately and the separate classifications were combined using different combination strategies. The CPV approach was compared to CRV [9], sum rule, max rule, median rule, and majority voting, which have given good results in [7]. Product rule was not included into comparison, because the probability estimates of k-NN classifiers are sometimes zero, which may corrupt the result. Also bagging and boosting algorithms were not tested, because they sub-sample the same feature set and therefore they cannot be applied to classifier combination problems with separate feature sets. The classification principle was selected to be k-nearest neighbor (k-NN) method. Barandela et al. [2] found that nearest neighbor principle is efficient and accurate method to be used in classifier combinations. Classification results were obtained using leave-one-out validation [4]. Euclidean distance was used as distance metrics in the base classification as well as in the final classification of CPV:s.

The classification results are presented in figures 4 and 5 in which the mean classification results are presented using different combination methods. The results are presented for k varying between 7 and 15. The reason for using relatively high values of k is that the probability distributions used by CPV are able to accurately distinguish between image classes only when k is relatively high. However, the presented classification rates performed by CPV are not achieved by any other combination method in the comparison at any value of k.

The results of figures 4 and 5 show that CPV clearly outperforms the other combination strategies in both image sets. Furthermore, the computational cost of CPV is not significant, because the probability distributions are used as feature vectors as themselves. Hence, in contrary to the other probability-based methods, any mathematical operation is not applied to probability distributions.
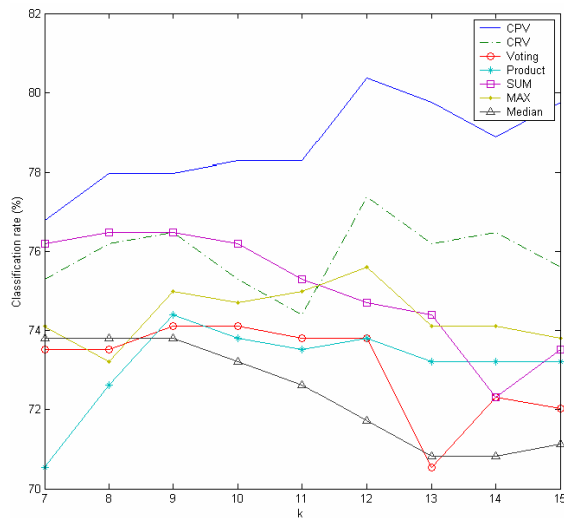
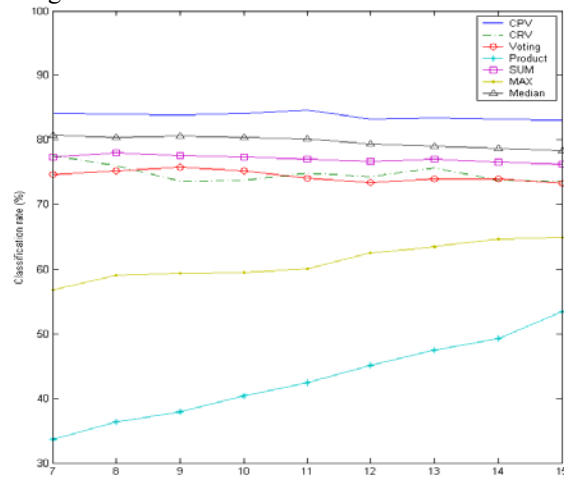Figure 4. The classification rate in the database I.

Figure 5. The classification rate in the database II.

4. DISCUSSION

In this paper, we presented a method for combining classifiers in the classification of non-homogenous real-world images. In the image classification, it is often beneficial to combine different descriptors to obtain the best possible classification result. Therefore, classifiers employing separate feature sets can be combined. We used natural rock images and industrial defect images as testing material. Due to their non-homogenous nature, the classification of these images is a difficult task.

In our method, the feature vector that describes the image content is formed using the probability distributions of separate base classifiers. The probabilities provided by the base classifiers form a new feature space, in which the final classification is performed. Hence the final classification depends on the metadata of the base classification, not the image features directly. This way the non-homogeneities of individual features do not have direct impact on the final result. The usability of the

proposed methods has been proved in the comparison with other probability-based classifier combination methods. The results show that CPV method clearly outperforms the other methods in the comparison.

5. ACKNOWLEDGMENT

6. REFERENCES

[1] F.M. Alkoot and J. Kittler, "Experimental evaluation of expert fusion strategies", *Pattern Recognition Letters*, Vol. 20, 1999, pp. 1361-1369.

[2] R. Barandela, J.S. Sánchez, and R.M. Valdovinos, "New applications of ensembles of classifiers", *Pattern Analysis & Applications*, Vol. 6, 2003, pp. 245-256.

[3] R. Brunelli and D. Falavigna, "Person Identification Using Multiple Cues", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17 No. 10, 1995, pp. 955-966.

[4] R.O. Duda, P.E. Hart, and D.G. Stork, *Pattern Classification*, 2nd ed., John Wiley & Sons, New York, 2001.

[5] R.P.W. Duin, "The Combining Classifier: to Train or Not to Train", *Proceedings of 16th International Conference on Pattern Recognition,* Vol. 2, 2002, pp. 765-770.

[6] A.K. Jain, S. Prabhakar, and S. Chen, "Combining Multiple Matchers for a High Security Fingerprint Verification System", *Pattern Recognition Letters*, Vol. 20, 1999, pp. 1371-1379.

[7] J. Kittler, M. Hatef R.P.W. Duin, and J. Matas, "On Combining Classifiers", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 3, 1998, pp. 226-239.

[8] L. I. Kuncheva, "A Theoretical Study on Six Classifier Fusion Strategies", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 2, 2002, pp. 281-286.

[9] L. Lepistö, I. Kunttu, J. Autio, and A. Visa, "Classification of Non-homogenous Textures by Combining Classifiers", *Proceedings of IEEE International Conference on Image Processing,* Vol. 1, 2003, pp. 981-984.

[10] L. Lepistö, I. Kunttu, J. Autio, and A. Visa, "Combining Classifiers in Rock Image Classification – Supervised and Unsupervised Approach", *Advanced Concepts for Intelligent Vision Systems*, 2004, pp. 17-22.

[11] X. Lin, S. Yacoub, J. Burns and S. Simske, "Performance analysis of pattern classifier combination by plurality voting", *Pattern Recognition Letters*, Vol. 24, 2003, 1959-1969.

[12] X. Lu, Y. Wang, and A. K. Jain, "Combining Classifiers for Face Recognition", *Proceedings of International Conference on Multimedia and Expo*, Vol. 3, 2003, pp. 13-16.

[13] B.S. Manjunath, P. Salembier, and T. Sikora, *Introduction to MPEG-7 Multimedia content description interface*, John Wiley & Sons, UK, 2002.

[14] J. Rauhamaa and R Reinius, "Paper web imaging with advanced defect classification", *Proceedings of TAPPI Technology Summit*, 2002.

Publication VI

Lepistö, L., Kunttu, I., Visa, A., 2005. Color-Based Classification of Natural Rock
Images Using Classifier Combinations.

Color-Based Classification of Natural Rock Images Using Classifier Combinations

Leena Lepistö, Iivari Kunttu, and Ari Visa

Tampere University of Technology, Institute of Signal Processing,
P.O. Box 553, FI-33101 Tampere, Finland
{Leena.Lepisto, Iivari.Kunttu, Ari.Visa}@tut.fi
http://www.tut.fi/

Abstract. Color is an essential feature that describes the image content and therefore colors occurring in the images should be effectively characterized in image classification. The selection of the number of the quantization levels is an important matter in the color description. On the other hand, when color representations using different quantization levels are combined, more accurate multilevel color description can be achieved. In this paper, we present a novel approach to multilevel color description of natural rock images. The description is obtained by combining separate base classifiers that use image histograms at different quantization levels as their inputs. The base classifiers are combined using classification probability vector (CPV) method that has proved to be an accurate way of combining classifiers in image classification.

1 Introduction

Image classification is an essential task in the field of image analysis. The classification is usually based on a set of visual features extracted from the images. These features may characterize for example colors or textures occurring in the images. Most of the real-world images are seldom homogenous. Especially, different kinds of natural images have often non-homogenous content. The division of natural images like rock, stone, clouds, ice, or vegetation into classes based on their visual similarity is a common task in many machine vision and image analysis solutions. In addition to non-homogeneities, the feature patterns can be also noisy and overlapping. Due to these reasons, different classifiers may classify the same image in different ways. Hence, there are differences in the decision surfaces, which lead to variations in classification accuracy. However, it has been found that a consensus decision of several classifiers can give better accuracy than any single classifier [1],[8],[9]. This fact can be easily utilized in the classification of real-world images.

The goal of combining classifiers is to form a consensus decision based on opinions provided by different base classifiers. Duin [5] presented six ways, in which consistent set of base classifiers can be generated. In the base classifiers, there can be differences in initializations, parameter choices, architectures, classification principle, training sets, or feature sets. Combined classifiers have been applied to several classification tasks, for example to face recognition [15], person identification [3] and fingerprint verification [7]. Theoretical framework for combining classifiers is provided in [8].

In the image classification, several types of classifier combination approaches can be used. In our previous work [11],[13] we have found that different feature types can be easily and effectively combined using classifier combinations. In practice, this is carried out by making the base classification for each feature type separately. The final classification can be obtained based on the combination of separate base classification results. This has proved to be particularly beneficial in the case of non-homogenous natural images [11],[13]. Hence, the non-homogenous properties of individual features do not necessarily affect directly on the final classification. In this way, each feature has its own affect on the classification result.

Rock represents typical example of non-homogenous natural image type. This is because there are often strong differences in directionality, granularity, or color of the rock texture, even if the images represented the same rock type [11]. Moreover, rock texture is often strongly scale-dependent. Different spatial multiscale representations of rock have been used as classification features using Gabor filtering [12]. However, the scale dependence of the rock images can be used also in another way, using color quantization. It has been found that different color features can be found from the rock images using different numbers of quantization levels. Hence, by combining the color representation at several levels, a multilevel color representation can be achieved. For this kind of combination, a classifier combination method can be used. In this paper, we present our method to make a classifier combination that is used to produce this multilevel color representation. The rest of this paper is organized as follows. Section two presents the main principle of classifier combinations as well as our method for that purpose. In section three, the principle of multilevel color representation is presented. The classification experiments with natural rock images are presented in section four. The obtained results are discussed in section five.

2   Classifier Combinations in Image Classification

The idea of combining classifiers is that instead of using single decision making theme, classification can be made by combining opinions of separate classifiers to derive a consensus decision [8]. This can increase classification efficiency and accuracy. In this section, methods for combining separate classifiers are presented. Furthermore, we present our approach to make a probability-based classifier combination.

2.1   Methods for Combining Classifiers

The general methods for combining classifiers can be roughly divided into two categories, voting-based methods and the methods based on probabilities. The voting-based techniques are popularly used in pattern recognition [10],[14]. In the voting-based classifier combinations, the base classifier outputs vote for the final class of an unknown sample. These methods do not require any additional information, like probabilities from the base classifiers. Voting has proved to be a simple and effective method for combining classifiers in several classification problems. Also in the comparisons with the methods presented by Kittler et al., voting-based methods have given relatively accurate classification results [8]. Lepistö et al. [11] presented a method for combining

classifiers using classification result vectors (CRV). In this approach, the class labels provided by the base classifiers are used as a feature vector in the final classification, and hence the result is not based on direct voting. CRV method outperformed voting method in the classification experiments. In [13] an unsupervised variation of CRV was presented and compared to other classifier combinations.

Recently, the probability-based classifier combination strategies have been popularly used in pattern recognition. In these techniques, the final classification is based on the a posteriori probabilities of the base classifiers. Kittler et al. [8] presented several common strategies for combining base classifiers. These strategies are e.g. product rule, sum rule, max rule, min rule, and median rule. All these rules are based on the statistics computed based on the probability distributions provided by the base classifiers. In [8], the best experimental results have been obtained using sum and median rules. Theoretical comparison of the rules has been carried out in [9]. Also Alkoot and Kittler [1] have compared the classifier combination strategies.

2.2   Classification Probability Vector Method

Our previous method for combining classifiers combined the outputs of the base classifiers into a feature vector that is called classification result vector (CRV) [11]. However, CRV method uses only the class labels provided by the base classifiers, and ignores their probabilities. In this paper, we use the probability distributions of the separate base classifiers as features in the final classification.

In general, in the classification problem a pattern S is to be assigned to one of the m classes ($\omega_1,\ldots,\omega_m$) [8]. We assume that we have C classifiers each representing a particular feature type and we denote the feature vector used by each classifier by f_i. Then each class ω_k is modeled by the probability density function $p(f_1 \mid \omega_k)$. The priori probability of occurrence of each class is denoted $P(\omega_k)$. The well-known Baeysian decision theory [4],[8] defines that S is assigned to class ω_j if the a posteriori probability of that class is maximum. Hence, S is assigned to class ω_k if:

$$P(\omega_j \mid f_1,\ldots,f_C) = \max_k P(\omega_k \mid f_1,\ldots,f_C) \qquad (1)$$

However, the probabilities of all the other classes than ω_j have also significance in classification. They are particularly interesting when the pattern S is located near the decision surface. Therefore, we focus on the whole probability distribution $p(f_i \mid \omega_1,\ldots,\omega_m)$ provided by each classifier. Hence, if the probability is defined for each m class, the obtained probability distribution is a C by m matrix for each pattern S. This matrix is used as a feature vector in the final classification, and it is called
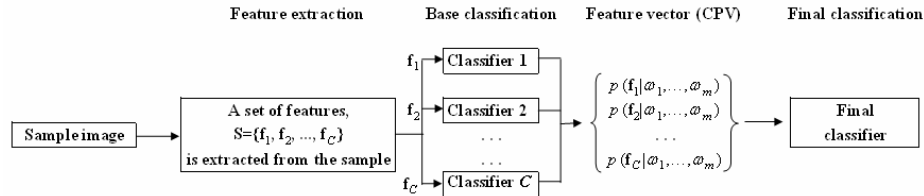

Fig. 1. The outline of the CPV classifier combination method

classification probability vector (CPV). In the final classification, the images with similar CPV's are assigned into same classes. The outline of the CPV method is presented in figure 1.

The common probability-based classifier combinations [8] are used to calculate some statistics based on the probability distributions provided by the base classifiers. In contrary to them, CPV method uses the whole probability distribution as a feature vector in the final classification. The CPV method utilizes the fact that the separate base classifiers classify similar samples in the similar way, which leads to a similar probability profile. The final classification is based merely on the similarity between the probabilities of the base classifiers. Hence, in contrary to voting, in the CPV method the base classifier outputs (class labels) do not directly affect the final classification result.

When image classification is considered, the CPV method has several advantages. CPV method considers each visual feature of the images in the base classifiers separately. In the final classification, the probability distributions are employed instead of features. This way, the individual features do not directly affect the final classification result. Therefore, classification result is not sensitive to variations and non-homogeneities of single images.

3   Multilevel Color Representation Using Quantization

3.1   Color Image Representation

In digital image representation, an image has to be digitized in two manners, spatially (sampling) and in amplitude (quantization) [6]. The use of spatial resolutions is common in different texture analysis and classification approaches, whereas the effect of quantization is related to the use of image color information. The quantization can be applied to different channels of a color image. Instead of using common RGB color space, the use of HSI space has found to be effective, because it corresponds to the human visual system [16].

A common way of expressing the color content of an image is the use of image histogram. Histogram is a first-order statistical measure that expresses the color distribution of the image. The length of the histogram vector is equal to the number of the quantization levels. Hence the histogram is a practical tool for describing the color content at each level. Histogram is also a popular descriptor in color-based image classification, in which images are divided into categories based on their color content.

3.2   Multilevel Classification

The classifier combination tools presented in section two provide a straightforward tool for making a histogram-based image classification at multiple levels. Hence the histograms at selected quantization levels and color channels are used as separate input features. Each feature is then classified separately at base classification. After that, the base classification results are combined to form the final classification. This way the final classifier uses multilevel color representation as classifying feature.
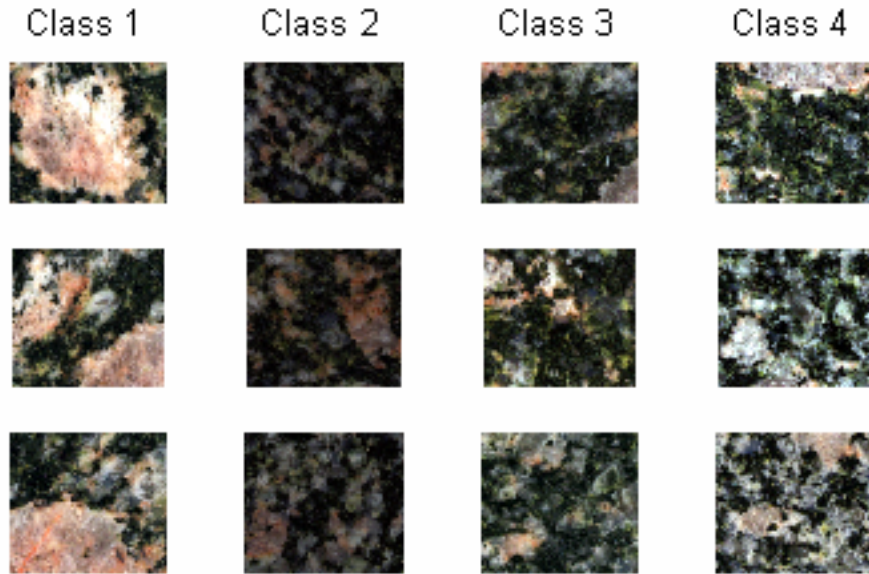
Fig. 2. Three example images from each rock type in the testing database

4   Experiments

In this section, we present the classification experiments using rock images. The purpose of the experiments is to show that an accurate multilevel color representation is achievable using classifier combinations. We also compare the CPV method to other classifier combination approaches.
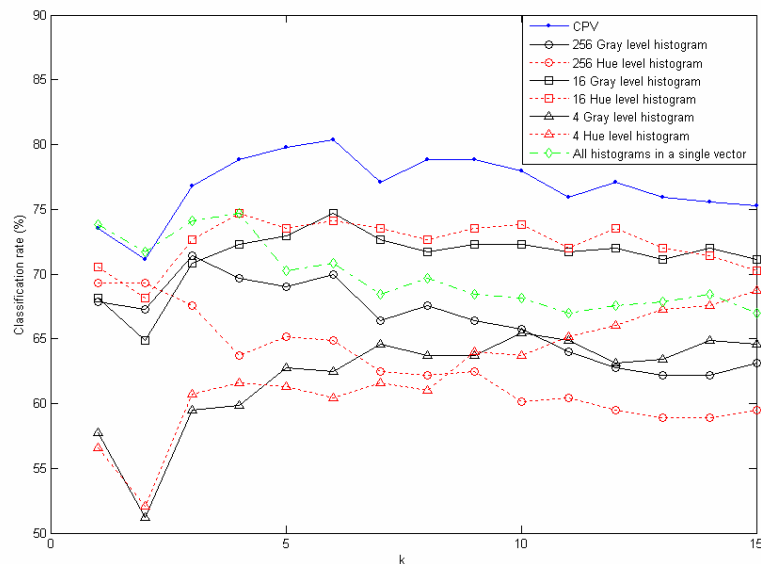
4.1   Rock Images

The experiments in this paper are focused on non-homogenous natural image data that is represented by a database of rock images. There is a practical need for methods for classification of rock images, because nowadays rock and stone industry uses digital imaging for rock analysis. Using image analysis tools, visually similar rock texture images can be classified. Another application field for rock imaging is geological research work, in which the rock properties are inspected using borehole imaging. Different rock layers can be recognized and classified from the rock images based on e.g. the color and texture properties of rock. The degree of non-homogeneity in rock is typically overwhelming and therefore, there is a need for an automatic classifier that is capable of classifying the borehole images into visually similar classes. The testing database consists of 336 images that are obtained by dividing large borehole images into parts. These images are manually divided into four classes by an expert. In classes 1-4, there are 46, 76, 100, and 114 images in each class, respectively. Figure 2 presents three example images of each four class.

4.2   Classification Experiments

In the experimental part, the principle for classification was selected to be k-nearest neighbor (k-NN) method in base classification and final classification. Barandela et al. [2] have proved that nearest neighbor principle is efficient and accurate method to be used in classifier combinations. Classification results are obtained using leave one out validation principle [4]. The distance metrics for the comparison of histograms in the base classification was selected to be L_1 norm. In the CPV method, the final classifier used L_2 norm (Euclidean distance) to compare CPV:s.

The histograms were calculated for the database images in HSI color space. In the classification experiments we used hue (H) and intensity (I) channels, which have been proved to be effective in color description of rock images [12]. The hue and intensity histograms were calculated for the images quantized to 4, 16, and 256 levels. Hence the number of the features was six. The classification was carried out using values of k varying between 1 and 15. In the first experiment, the classification rate of CPV method was compared to that of separate base classifiers which use different histograms. In this comparison, also the classification accuracy all the histograms combined into a single feature vector was tested. The average classification rates are presented in figure 3 as a function of k. The second experiment measured the classification accuracy of different classifier combination strategies compared to CPV method. The idea of this experiment was to combine the six base classifiers that use different histograms as input features. In this case, CPV was compared to the most usual probability-based classifier combinations, sum, max and median rules [8]. Product rule is not included into comparison, because the probability estimates of

Fig. 3. The average classification rates of the rock images using base classifiers that use different histograms and the classifier combination (CPV)

k-NN classifiers are sometimes zero, which may corrupt the result. In addition to the selected probability-based combination methods, we used also majority voting and our previously introduced CRV method [11] in the comparison. Figure 4 presents the results of this comparison with k varying between 1 and 15.
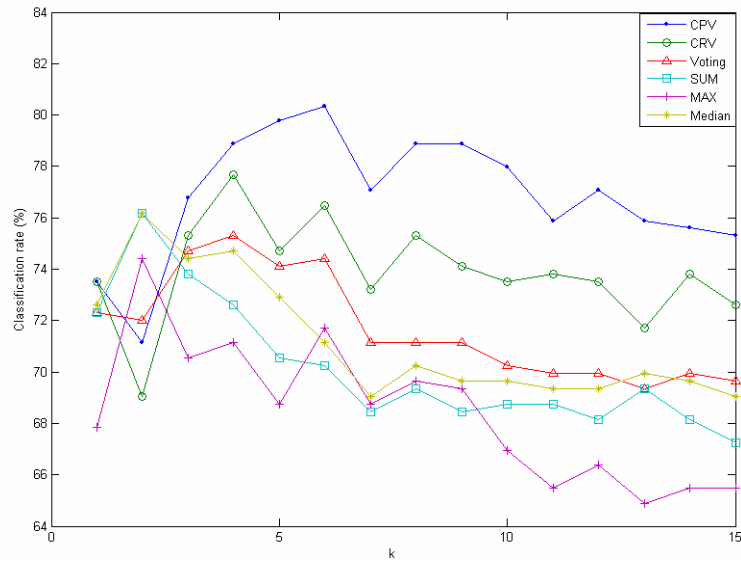
Fig. 4. The average classification rates of the rock images using different classifier combinations

4.3 Results

The results presented in figure 3 show that using CPV the classification accuracy is clearly higher than that of any single base classifier. The performance of CPV is also compared to the alternative approach, in which all the histograms are collected into a single feature vector. This vector is very high dimensional (552 dimensions), and its performance is significantly lower than in the case of CPV. This observation gives the reason for the use of classifier combinations in the image classification. That is, different features can be combined by combining their separate base classifiers rather than combining all the features into a single high dimensional feature vector in classification. This way, also the "curse of dimensionality" can be avoided.

The results of the second experiment presented in figure 4 show that CPV method outperforms the other classifier combinations in the comparison with a set of rock images. Also the CRV [11] method gives relatively good classification performance. Only with small values of k CPV is not accurate one. This is due to the probability distributions used by CPV are able to effectively distinguish between image classes only when more than three nearest neighbors are considered in k-NN algorithm.

5 Discussion

In this paper, we presented a method for combining classifiers in the classification of real rock images. Due to their non-homogenous nature, the classification of them is a difficult task. We presented a method for an effective multilevel color representation using our classifier combination strategy, classification probability vector (CPV). In CPV method, the feature vector that describes the image content is formed using the probability distributions of separate base classifiers. The probabilities provided by the base classifiers form a new feature space, in which the final classification is made. Hence the final classification depends on the metadata of the base classification, not the image features directly. This way the non-homogeneities of individual features do not have direct impact on the final result.

In the color-based image classification, like in image classification in general, it is often beneficial to combine different visual features to obtain the best possible classification result. Therefore, classifiers that use separate feature sets can be combined. In this study this feature combination approach was applied to color histograms with different numbers of bins. By combining the histograms using classifier combinations, a multilevel color representation was achieved. The experimental results showed that this representation outperforms any single histogram in classification. Furthermore, CPV method also gives better classification accuracy than any other classifier combination in the comparison.

Acknowledgement

References

1. Alkoot, F.M., Kittler, J.: Experimental evaluation of expert fusion strategies, Pattern Recognition Letters, Vol. 20 (1999) 1361-1369
2. Barandela, R., Sánchez, J.S., Valdovinos, R.M.: New applications of ensembles of classifiers, Pattern Analysis & Applications, Vol. 6 (2003) 245-256
3. Brunelli, R., Falavigna, D.: Person Identification Using Multiple Cues, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17 (1995) 955-966
4. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification, 2nd ed., John Wiley & Sons, New York (2001)
5. Duin, R.P.W.: The Combining Classifier: to Train or Not to Train, In: Proceedings of 16th International Conference on Pattern Recognition, Vol. 2 (2002) 765-770
6. Gonzales, R.C., Woods, R.E.: Digital Image Processing, Addison Wesley, 1993.
7. Jain, A.K., Prabhakar, S., Chen, S.: Combining Multiple Matchers for a High Security Fingerprint Verification System, Pattern Recognition Letters, Vol. 20 (1999) 1371-1379
8. Kittler, J., Hatef, M., Duin, R.P.W., Matas J.: On Combining Classifiers, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20 (1998) 226-239
9. Kuncheva, L.I.: A Theoretical Study on Six Classifier Fusion Strategies, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24 (2002) 281-286

10. Lam, L., Suen, C.Y.: Application of majority voting to pattern recognition: An analysis of the behavior and performance, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 27 (1997) 553-567
11. Lepistö, L., Kunttu, I., Autio, J., Visa, A.: Classification of Non-homogenous Textures by Combining Classifiers, Proceedings of IEEE International Conference on Image Processing, Vol. 1 (2003) 981-984
12. Lepistö, L., Kunttu, I., Autio, J., Visa, A.: Classification Method for Colored Natural Textures Using Gabor Filtering, In: Proceedings of 12th International Conference on Image Analysis and Processing (2003) 397-401
13. Lepistö, L., Kunttu, I., Autio, J., Visa, A.: Combining Classifiers in Rock Image Classification – Supervised and Unsupervised Approach, In: Proceedings of Advanced Concepts for Intelligent Vision Systems, (2004) 17-22
14. Lin, X., Yacoub, S., Burns, J., Simske, S.: Performance analysis of pattern classifier combination by plurality voting, Pattern Recognition Letters, Vol. 24 (2003) 1959-1969
15. Lu, X., Wang, Y., Jain, A.K.: Combining Classifiers for Face Recognition, In: Proceedings of International Conference on Multimedia and Expo, Vol. 3 (2003) 13-16
16. Wyszecki, G., Stiles, W.S.: Color Science, Concepts and Methods, Quantitative Data and Formulae, 2nd Edition, John Wiley & Sons (1982)

Publication VII

Lepistö, L., Kunttu, I., Visa, A., 2006. Rock image classification based on *k*-nearest
neighbour voting.

Rock image classification based on *k*-nearest neighbour voting

Leena Lepistö[1], Iivari Kunttu, and Ari Visa

Tampere University of Technology, Institute of Signal Processing
P.O. Box 553, FI-33101 Tampere, Finland
E-mail: {Leena.Lepisto, Iivari.Kunttu, Ari.Visa@tut.fi}

ABSTRACT

Image classification is usually based on various visual descriptors extracted from the images. The descriptors characterizing for example image colours or textures are often high dimensional and their scaling varies significantly. In the case of natural images, the feature distributions are often non-homogenous and the image classes are also overlapping in the feature space. This can be problematic, if all the descriptors are combined into a single feature vector in the classification.

In this paper, we present a method for combining different visual descriptors in rock image classification. In our approach, *k*-nearest neighbour classification is first carried out for each descriptor separately. After that, the final decision is made by combining the nearest neighbours in each base classification. The total numbers of the neighbours representing each class are used as votes in the final classification. The experimental results with rock image classification indicate that the proposed classifier combination method is more accurate than the conventional plurality voting.

---

[1]Corresponding author. tel: +358 3 3115 4964, fax: +358 3 3115 4989

1. Introduction

The division of natural images such as rock, stone, clouds, ice, or vegetation into classes based on their visual similarity is a common task in many machine vision and image analysis solutions [1],[2],[3]. Classification of natural images is demanding, because in the nature the objects are seldom homogenous. For example, when the images of rock surface are inspected, there are often strong differences in directionality, granularity, or colour of the rock, even if the images represented the same rock type. In addition to non-homogeneities, the feature patterns can be also noisy and overlapping, which may cause variations in the decision surfaces of different classifiers. Due to these reasons, different classifiers may classify the same image in different ways. These kinds of variations and non-homogeneities make it difficult to classify these images accurately using a single classifier. On the other hand, non-homogenous feature distributions of certain image types may also improve classification of these images in certain classifiers. Thus the fact that different classifiers often give varying decisions can be utilized in the classification. It has been found that a consensus decision of several classifiers can often give better accuracy than any single classifier [4],[5],[6],[7]. This fact can be easily utilized in the classification of real-world images.

In the image classification, a number of visual descriptors are used to classify the images based on their content. The most typical descriptor types characterize colours, textures and shapes occurring in the images. The feature space is typically high dimensional and image categories are often overlapping in the feature space. A common approach in the image classification is to combine all the selected descriptors into a single feature vector. In the case of rock images, examples of this kind of classification are [2],[8]. The similarity between these vectors is defined using some distance metric and the most similar images are then classified (labelled) to the same category. However, when different types of descriptors are combined into the same feature vector, some large-scaled features may dominate the distance, while the other features do not have the same influence on the classification. It is clear that the effect of scaling can be minimized by using feature normalization. Especially, in the case of high dimensional descriptors combination into the same vector can be problematic and may yield to remarkable drawbacks in classification performance. This is known as "the curse of dimensionality" [9]. In addition, high dimensional descriptors may have stronger impact on the distance than low dimensional ones, even if the features

2

were normalized. Therefore it is often more reasonable to consider each visual descriptor separately. This way, the descriptors are not needed to normalize to any particular scale.

In images with non-homogenous content, visual features defined for a particular image vary in the extent of non-homogeneity of this sample image. In our previous work [10],[11],[12] we have found that different visual descriptors (feature sets) obtained from non-homogenous images can be easily and effectively combined using classifier combinations employing k-nearest neighbour (k-NN) classifiers. This is because the k-NN principle is robust to the variations and non-homogeneities in the dataset [13]. Furthermore, the k-NN method is simple and fast method and it is easy to implement **[14].** It has also been indicated to be suitable method for feature spaces with overlapping classes [11],[12]. Furthermore, it has been found that the employment of k-NN base classifiers in classifier combinations is motivated, when separate feature sets are employed [13]. This is because by selecting appropriate feature sets for the base classification, diverse and accurate base classification can be achieved [13]. Our experiments with non-homogenous natural images have shown that diverse base classification results can be achieved also with quite similar input descriptors. One reason for this can be the non-homogenous content of the input images. In practice, the classification is carried out by making the base classification for each descriptor separately. The final classification can be obtained based on the combination of separate base classification results. This way, each descriptor has its own impact on the classification result independent of its scaling and dimensionality. This has proved to be beneficial in the case of non-homogenous natural texture images [11],[12].

The use of classifier combinations has been a subject of an intensive research work during last ten years. Popular solution on this field has been bagging [15], which manipulates the training data sets with sub-sampling. Another common algorithm, boosting [16], also manipulates the training data, but it emphasizes the training of samples which are difficult to classify. These methods, however, sub-sample the same feature set and therefore they cannot be applied to classifier combination problems with separate descriptors. Recently, the probability-based classifier combination strategies have received much attention in the field of pattern recognition

[4],[5],[6],[7]. In these techniques, the final classifier makes the decision based on the a posteriori probabilities provided by base classifiers.

In voting-based techniques, the final decision is made based on the outputs of the base classifiers by voting. Hence, voting-based methods do not require any further training in the final classification, as most other combination methods. This makes them simple and computationally effective methods. In addition, the risk of overtraining can be avoided. Voting has found to be accurate and effective method for combining classifiers in several classification problems [17],[18],[19],[20]. Voting-based methods can be divided into two classes, majority voting and plurality voting. Majority voting [19] requires the agreement of more than half of the participants to make a decision. If majority decision cannot be reached, the sample is rejected. On the other hand, plurality voting selects the sample which has received the highest number of votes. The comparisons of Lin et al. [20] have indicated that plurality voting is more efficient of these two techniques. Furthermore, using plurality voting, the problem rising with the rejected samples can be avoided because all the samples can be classified. In literature, there are several examples on the feature selection in classifier combinations. In [21], the features to be used in base classification are randomly selected. Also in [22], the feature space is partitioned into each dimension, and voting is applied to reach a consensus decision based on the classifications at each dimension.

In this paper, we consider different visual descriptors extracted from the images as inputs for the base classifiers. Thus the descriptors are not divided into one dimensional features, as in [22]. Instead of that, we use the n-dimensional visual descriptors as themselves as inputs of the base classifiers. In our approach, we combine the method of plurality voting with k-nearest neighbour (k-NN) classification principle, which in fact is also based on voting. This is because in the k-NN principle, the class of unknown sample is decided based on the most frequent class within k nearest neighbour samples of the sample in the feature space. In our approach, the votes of each k-NN base classifier are used to make the final voting. This kind of approach is able to improve the accuracy of "uncertain" k-NN base classifiers. For example, let us assume that a 5-NN classifier has to decide the class label of an unknown sample that is located at the decision boundary of classes A and B in the feature space. If two neighbours vote for class A and three neighbours vote for class B, then the 5-NN classifier decides that the class label of the sample is B. This

decision, however, is a weak one, because all the neighbours didn't agree. Therefore, in the proposed classifier combination approach all the neighbour opinions are used as votes in the final voting. In contrary to this approach, conventional voting uses only the class labels provided by the base classifiers to make the decision. Hence, instead of the plain class labels, in the proposed approach the base classifiers provide a set of votes that are employed as a basis of the final voting. Therefore, k votes provided by each base classifier are in fact weights for final voting. This kind of weighted voting approach is particularly suitable for the combination of the base classifiers that exhibit different accuracies. This is typical in the classification of natural rock images.

The experiments presented in this paper use the proposed classifier combination method to combine colour and texture descriptors extracted from the rock images. In the case of colour-based classification, histograms are used as input descriptors. In [10] we presented that an efficient multilevel colour-based rock image classification can be achieved by combining classifiers that employ histograms of different quantization levels and colour channels. The combination rule presented in [10], however, was not based on the voting as in this paper. The second of the experiments is related to texture analysis. In [2], we showed that an effective multiresolution texture representation for non-homogenous rock images can be obtained using band-pass filters in Gabor space. Using ring-shaped Gaussian filters at different frequency bands; the frequency information of rock texture can be obtained. The problem with the multiresolution texture representation is to find a method how to combine the filter responses obtained from the different frequency bands. In [2], this problem was solved by combining the feature vectors of each band into a single feature vector. In the experiment presented in this paper, we show that the proposed voting-based classifier combination is a useful way of combining the filtering results in classification and gives more accurate classification result than the method used in [2].

The rest of this paper is organized as follows. Section two describes briefly the area of rock image analysis. In section three, the idea of the proposed classification method is presented. Section four presents the classification experiments, in which the proposed method is tested using a database of rock images. Section five includes discussion and conclusions.

2. Rock image analysis

2.1. Non-homogenous natural images

Most of the real-world images are somehow non-homogenous, which means that there are clearly visible changes in their visual features [11]. Typical examples of these images are natural images like rock, stone, clouds, ice, or vegetation. In these images, the colour and texture properties may vary strongly, even if the sample images represented same image type. Rock is a typical example of natural images that is often non-homogenous. For example, rock sample presented in figure 1 is strongly non-homogenous in terms of directionality, granularity, and colour. The homogeneity of a sample image can be measured by dividing the sample into blocks. If the visual properties do not vary between the blocks, the sample is homogenous. On the other hand, if these feature values have significant variance, the texture sample is non-homogenous. In our previous approach [23], this division into blocks was applied.

2.2. Bedrock imaging and classification

In the field of rock science, the development of digital imaging has made it possible to store and manage the images of the rock material in digital form. One typical application area of the rock imaging is bedrock investigation. In this kind of analysis, rock properties are analyzed by inspecting the images which are collected from the bedrock using borehole imaging.

The borehole images can be obtained from the core samples drilled from the bedrock using core scanning techniques [1]. An example of an image that is scanned from the core surface is presented in figure 2. The purpose of the core sample classification is to find interesting sections of rock material. The classification tasks can be based on e.g. mineral content, physical properties or origin of the core samples. The essential visual features used in the classification can be for example texture, grain structure and colour distribution of the samples. The current core scanning techniques are able to produce high resolution images of rock material. The core images can be also acquired by using additional invisible light wavelengths, which can be used to discriminate between certain minerals or chemical elements [1]. Therefore, the number of images obtained from a deep drilled core is remarkable. The images of the core samples are stored into image databases, which can be utilized in the rock inspection. Due to relatively large sizes of these databases, automated image classification methods are necessary.

3. Combining *k*-NN classifiers by voting

In general, in the classification problem an unknown sample image *S* is to be assigned to one of the *m* classes ($\omega_1,\dots,\omega_m$) [9]. We assume that we have *C* classifiers each representing a particular visual descriptor extracted from the image. We denote the feature vector of the descriptor used by *i*:th classifier by f_i. Then each class ω_n is modelled by the probability density function $p(f_i|\omega_n)$. The probability of occurrence of *n*th class is denoted $P(\omega_n)$. The well-known Bayesian decision theory [9] defines that *S* is assigned to class ω_j if the a posteriori probability of that class is maximum. Hence, *S* is assigned to class ω_j if:

$$P(\omega_j|f_1,\dots,f_C) = \max_n P(\omega_n|f_1,\dots,f_C) \tag{1}$$

Based on this, base classifier gives a label ω_j to an unknown sample image. When we have several base classifiers using different descriptors as their input feature vectors, each classifier provides a class label for the unknown image.

3.1. k-NN classification

In this paper, the classifier principle is selected to be *k*-nearest neighbour (*k*-NN) method.

In the *k*-NN classifier, a pattern *S* is to be assigned to one of the *m* classes ($\omega_1,\dots,\omega_m$). The algorithm finds the *k* nearest patterns in the feature space [9]. Let $N_i=(\varepsilon_1,\dots,\varepsilon_m)_i$ denote the number of these patterns in each *m* class in feature space f_i. Thus the probability set for a sample pattern *S* can be expressed as:

$$P(\omega_1,\dots,\omega_m \mid f_i) = {(\varepsilon_1,\dots,\varepsilon_m)_i}\Big/{k} \tag{2}$$

in which

$$\sum_{i=1}^{m} \varepsilon_i = k \tag{3}$$

In *k*-NN classification, the pattern *S* is assigned to the class that has the highest probability according to Equation (2). In our approach, however, we utilize the number of nearest patterns in each *m* class ($\varepsilon_1,\dots,\varepsilon_m$) as votes in the final classification. Hence, if the same value of *k* is used for all classifiers, the set *N* is the same as the probability set *P*.

3.2. Voting

Plurality voting means that the sample that has received most votes is chosen [20]. In conventional plurality voting, each label ω_j provided by C base classifiers equals to a vote that is used in final decision. Hence, the class that receives most of C votes is selected.

In our approach, we do not use class labels but we take the set of nearest patterns N used in each C k-NN classification. This approach is called k-NN voting. The number of nearest patterns in each class is used as votes. Hence the total number of votes used in final classification is $C*k$. The votes received by each class can be combined by summing the sets N of each C base classifier:

$$(V_1,...,V_m) = \sum_{i=1}^{C} N_i = \sum_{i=1}^{C} (\varepsilon_1,...,\varepsilon_m)_i \qquad (4)$$

Hence the set of votes $(V_1,…,V_m)$ expresses the number of votes received by each class. Based on this, the final decision is reached using plurality voting. The outline of the k-NN voting approach is presented in figure 3.

4. Classification experiments

In this section, the performance of the proposed classifier voting-based classifier combination principle is examined using rock images. As presented in section 2, the geological research work uses borehole imaging to inspect the properties of rock. Different rock layers can be recognized from the borehole images based on the local colour and texture properties. Therefore, there is a need for an automatic classifier that is capable of classifying the borehole images into visually similar classes. A testing database used in this paper consists of 336 images that are obtained by dividing large borehole images into parts. These images are manually divided into four classes by an expert. The division is based on their colour and texture properties. Especially the granular size and colour distribution of the images are essential visual features that distinguish between the rock types. Figure 4 presents three example images of each four class. In classes 1-4, there are 46, 76, 100, and 114 images in each class, respectively.

We made two experiments with the testing database. In the first experiment, we tested the proposed classifier combination principle to the colour-based rock image classification using histograms. The second experiment concentrated on the texture

properties of the images. In both experiments, the proposed classifier combination method was compared to the plurality voting and the results given by the separate descriptors. In all the experiments leave-one-out cross validation principle [9] was employed. In the leave-one-out method, only one sample is regarded as an unknown sample in turn, and all the other samples in the data set serve as training data. This way, all the samples in the dataset are classified.

It is widely accepted that a classifier combination is capable of improving the classification rate of the base classifiers only if there is diversity between the base classifiers [13],[14]. This means that the base classifiers do not always agree. In other words, if all the base classifiers make the same error, their combination is not meaningful and provides no improvement. Moreover, several empirical and theoretical studies have found that the classifier combination is most successful when the errors of the base classifiers are as uncorrelated as possible [13],[24]. For this reason, the correlation between the base classifier errors is an issue of interest also in the experiments presented in this paper. On the other hand, when the visual descriptors extracted from an image are used as input features, there is often some correlation between these features. This is because for example individual colour descriptors or texture descriptors extracted from the same image are probably correlated. The results presented in figures 5 and 7 report varying results for each base classifier, which indicates that the errors are not equal, and classifier combination is able to improve the base classification results. The degree of agreement of the individual base classifier decisions can be estimated by investigating error residuals of the base classifiers. In practice, the residual function is formed by taking the difference between the classifier outputs and the real class numbers of the samples in the whole test set. Then the errors of different classifiers can be compared by counting the number of the differences between their residual functions. For this purpose, Hamming distance [25] can be employed. In our experiments, we have normalized the distances with the total number of the test set samples. This way, the normalized distance is between 0 and 1 such that 1 means that the residuals are totally different and 0 implies that the functions are equal. The normalized distances for the both experiments for $k=5$ are presented in tables 1 and 2.

4.1. Multilevel colour classification of the rock images

The results of previous classification experiments with rock images [10] have shown that the selection of quantization levels is an essential matter in histogram-based classification. Histogram is a first-order statistical measure that expresses the colour distribution of the image. The length of the histogram vector (the number of histogram bins) is equal to the number of the quantization levels. Hence the histogram is a practical tool for describing the colour content at each level. Histogram is also a popular descriptor in colour-based image classification, in which images are divided into categories based on their colour content. The problem with the use of multilevel colour representation using multiple histograms is the fact that when several histograms are combined into a single feature vector the resulting feature space is very high dimensional.

In this paper, we use the proposed voting-based classifier combination method to form a multilevel colour classifier for images in the testing database. The histograms were calculated for the database images in HSI colour space. In the classification experiments we used hue and intensity (grey level) channels, which have been shown to be effective in colour description of rock images [8]. The hue and intensity histograms were calculated for the images quantized to 16, 32, and 64 levels. The quantization level numbers above 64 were not used, because they gave lower base classification results in preliminary experiments. Hence the number of the descriptors was six. In the base classification, the similarity between the histograms was evaluated using L_1-norm (Manhattan distance). The classification was carried out using values of k varying between 1 and 10. The average classification rates using separate base classifiers as well as classifier combinations are presented in figure 5 as a function of k. In addition, also the classification accuracy all the histograms combined into a single feature vector is tested.

The results presented in figure 5 reveal that the proposed k-NN voting principle outperforms clearly the plurality voting, which in fact gives lower classification rate than the best of the histograms used in base classification (32 grey level histogram). When the histogram descriptors are considered, the results show that intensity histograms outperform those calculated for hue channels. However, the voting results are higher when they are evaluated for hue and intensity rather than merely intensity. This is the reason why both hue and intensity histograms were selected to be used in

the experiments. The approach that combines all the histograms into a single feature vector gives about the same classification rate as plurality voting.

4.2. Multiresolution texture classification of the rock images

In texture-based classification experiment, band-pass filters of five scales were used [2]. Amplitude responses of the five ring-shaped Gaussian filters are presented in figure 6. The intensity components of the rock texture images were filtered using the each five filters. The feature vector of each five scale was formed using the mean and standard deviation of the magnitude of the transform coefficients at the selected scale. These vectors were used as input descriptors for the base classifiers. The similarity between these descriptors was defined using L_2-norm (Euclidean distance). The classification was carried out using values of k varying between 1 and 10. The average classification rates are presented in figure 7 as a function of k.

Also in the case of this experiment, the proposed k-NN voting principle gives better results than the plurality voting. As figure 7 reveals, the plurality voting has not been able to outperform the best base classification result (filtering at scale 4). In this experiment, also the conventional approach to combine all the features into a single feature vector [2] was tested. This approach was able to give slightly better classification result than plurality voting when k=5.

5. Discussion and conclusions

In this paper, a method for image classification was presented. In the image classification, it is often beneficial to combine different visual descriptors to obtain improved classification result. Therefore, classifiers employing separate descriptors can be combined. When each base classifier uses a single visual descriptor as its input, all the descriptors have the same influence on the final classification independent of their scaling and dimensionality.

The classifier combination method presented in this paper is based on voting. Voting-based classifier combination methods are simple and general methods to be used in all kinds of classification problems. Voting is also computationally an effective way of combining separate classification results, because the final classification result is decided merely based on the votes provided by the base classifiers and no classifier training is required in the final classification. In this paper, we have introduced a novel way of making a consensus decision using voting. In the

proposed approach, the base classifiers use k-NN classification principle, and the number of nearest neighbours in each individual base classifier are used as votes in the final classifier. This kind of approach outperforms the conventional plurality voting, in which only the base classifier outputs (class labels) are regarded as votes. On the other hand, the proposed voting-based approach is as simple and fast method as plurality voting.

In conclusion, the experimental results show that the proposed voting-based method is able to give accurate results in the case of natural rock images, which are quite challenging classification task. The obtained results also indicate that the proposed classifier combination method provides an accurate and efficient way of combining separate visual descriptors in practical image classification tasks.

6. Acknowledgments

7. References

[1] J. Autio, L. Lepistö, and A. Visa, "Image analysis and data mining in rock material research," *Materia*, (4), 36-40, (2004).

[2] L. Lepistö, I. Kunttu, and A. Visa, "Rock image classification using color features in Gabor space," To be published in *Journal of Electronic Imaging*.

[3] Visa, A., Iivarinen, J., Evolution and evaluation of a trainable cloud classifier. *IEEE Transactions on Geoscience and Remote Sensing*, 35(5), (1997).

[4] J. Kittler, M. Hatef, R.P.W. Duin, and J. Matas, "On combining classifiers," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(3), 226-239, (1998).

[5] F.M. Alkoot and J. Kittler, "Experimental evaluation of expert fusion strategies," *Pattern Recognition Letters* 20, 1361-1369 (1999).

[6] L. I. Kuncheva, "A theoretical study on six classifier fusion strategies," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(2), 281-286, 2002.

[7] R. P. W. Duin, "The combining classifier: to train or not to train," *Proceedings of 16th International Conference on Pattern Recognition,* vol. 2, 765-770 (2002).

[8] L. Lepistö, I. Kunttu, J. Autio, and A. Visa, "Classification method for colored natural textures using Gabor filtering," *Proceedings of 12th International Conference on Image Analysis and Processing*, 397-401, (2003).

[9] R.O. Duda, P.E. Hart, and D.G. Stork, *Pattern Classification*, 2nd ed., John Wiley & Sons, New York (2001).

[10] L. Lepistö, I. Kunttu, and A. Visa, "Color-based classification of natural rock images using classifier combinations," *Proceedings of 14th Scandinavian*

Conference on Image Analysis, Joensuu, Finland, LNCS Vol. 3540, pp. 901-909, (2005).

[11]L. Lepistö, I. Kunttu, J. Autio, and A. Visa, "Classification of non-homogenous texture images by combining classifiers," *Proceedings of IEEE International Conference on Image Processing,* Barcelona, Spain, Vol. 1, 981-984, (2003).

[12]L. Lepistö, I. Kunttu, J. Autio, J. Rauhamaa, and A. Visa, Classification of non-homogenous images using classification probability vector, *Proceedings of IEEE International Conference on Image Processing,* Genova, Italy, Vol. 1, 1173-1176, (2005).

[13]C. Domeniconi, B. Yan, "Nearest neighbor ensemble", *Proceedings of 17th International Conference on Pattern Recognition* (2004)

[14]R. Barandela, J.S. Sánchez, R.M. Valdovinos, "New applications of ensembles of classifiers," *Pattern Analysis & Applications* 6, 245-256, 2003.

[15]L. Breiman, "Bagging predictors," *Machine Learning,* 26(2), 123-140, (1996).

[16] Y. Freund and R.E. Shaphire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences* 55(1), 119-139, (1995).

[17]L. Xu, A. Krzyzak, C.Y. Suen, "Methods for combining multiple classifiers and their applications to handwriting recognition," *IEEE Transactions on Systems, Man, and Cybernetics* 22(3), 418-435 (1992).

[18] T.K. Ho, J.J. Hull, S.N. Srirari, "Decision combination in multiple classifier systems," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16 (1), 66-75 (1994).

[19] L. Lam and C.Y. Suen, "Application of majority voting to pattern recognition: An analysis of the behavior and performance," *IEEE Transactions on Systems, Man and Cybernetics,* Part A 27(5), 553-567, (1997).

[20] X. Lin, S. Yacoub, J. Burns, and S. Simske, "Performance analysis of pattern classifier combination by plurality voting," *Pattern Recognition Letters* 24, 1959-1969, (2003).

[21] R. Bryll, R. Gutirrez–Osuna, and F. Quek, "Attribute bagging: improving accuracy of classifiers ensembles by using random feature subsets," *Pattern Recognition* 36, 1291-1302, (2003).

[22]A. Guvenir and I. Sirin, "Classification by feature partition," *Machine Learning* 23, 47-67, (1996).

[23]L. Lepistö, I. Kunttu, J. Autio, and A. Visa, "Rock image classification using non-homogenous textures and spectral imaging." *WSCG Short papers proceedings*, 82-86, (2003).

[24]J.A. Bilmes, K. Kirchhoff, "Generalized rules for combination and joint training of classifiers." Pattern Analysis & Applications 6, 201-211, (2003).

[25]N. Gaitanis, G. Kapogianopoulos, D.A. Karras, "Pattern classification using a generalized Hamming distance metric." In Proceedings of International Joint Conference on Neural Networks, pp. 1293-1296, (1993).
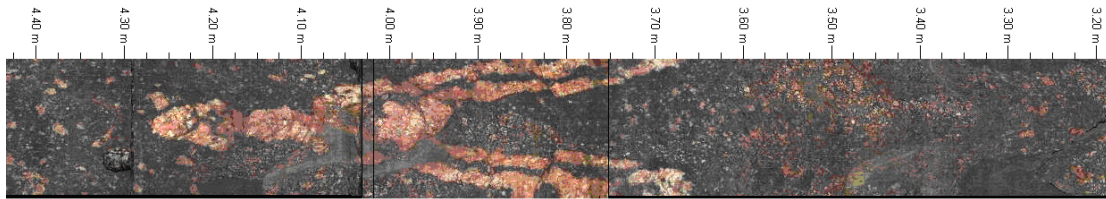
Figure 1. An example of non-homogenous rock texture image.

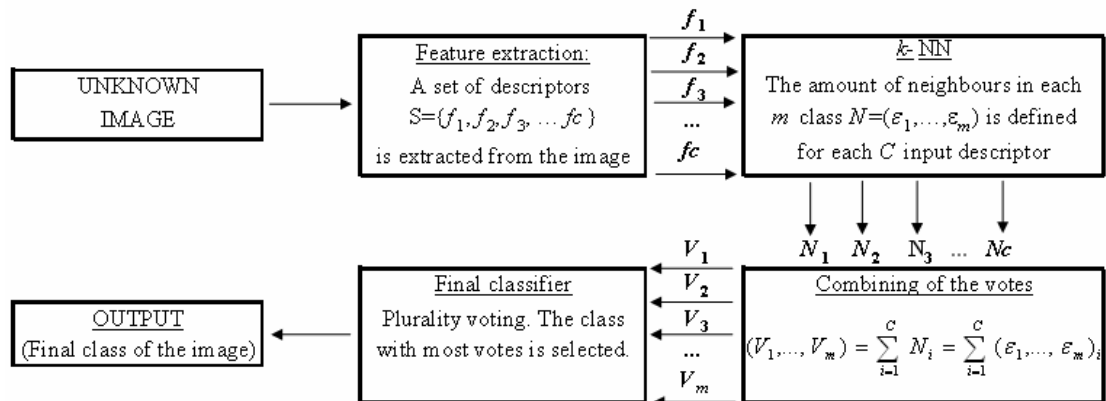Figure 2. An example of borehole image used in bedrock investigation.

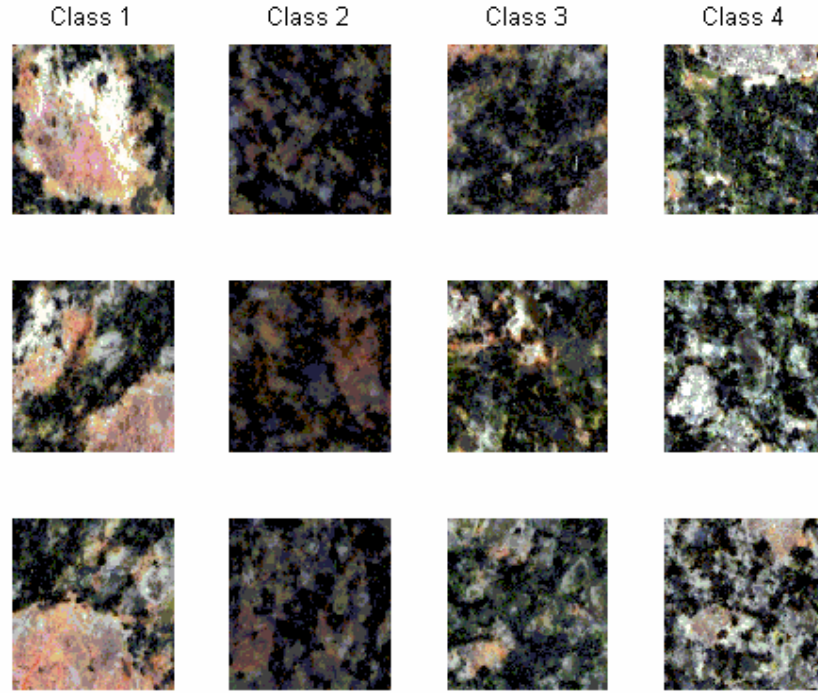Figure 3. The outline of the proposed classifier combination scheme.

14

Figure 4. Three examples from each class of the rock images in the testing database.
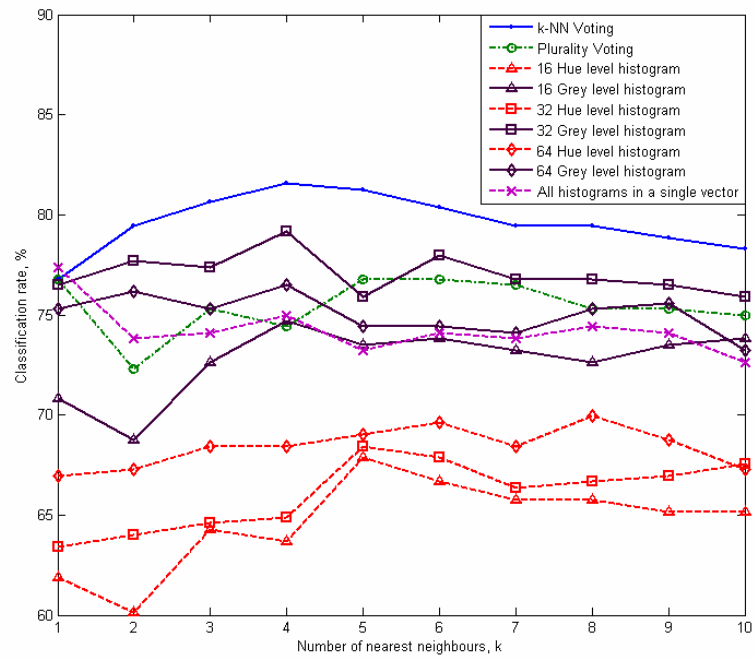

Figure 5. Average classification rates of the rock images using base classifiers that use different histograms as their input descriptors. The base classifiers are combined using the proposed k-NN voting method and plurality voting. The results are also compared to an approach, in which all the histograms are combined into a single feature vector.
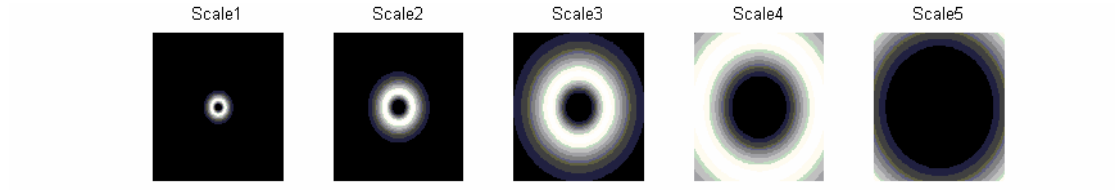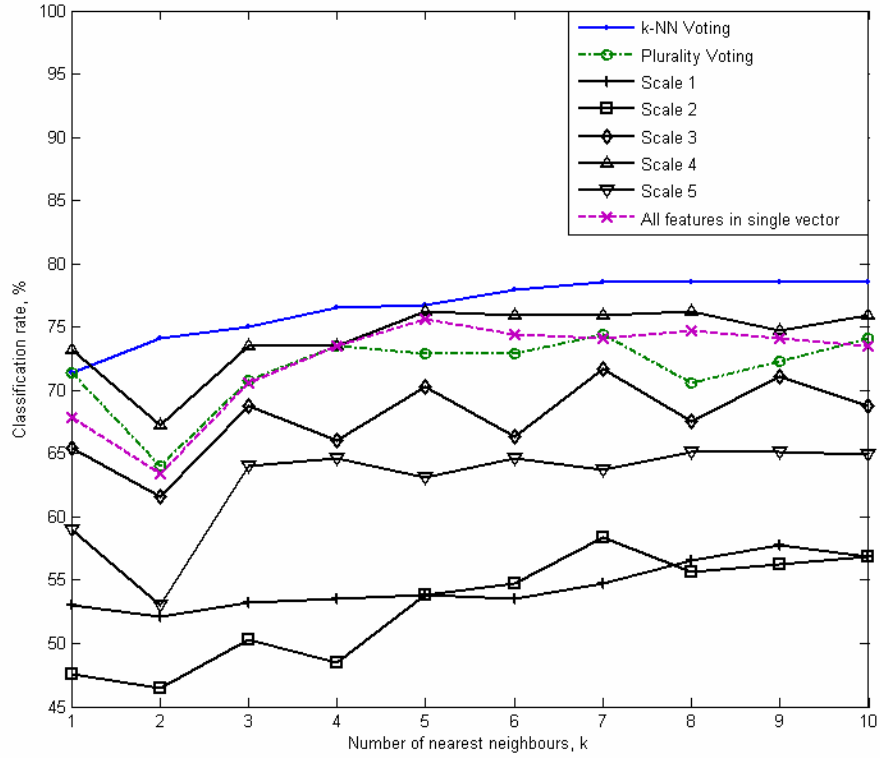
Figure 6. The filters at five scales.

Figure 7. Average classification rates of the rock images using base classifiers that use feature vectors obtained from band-pass filtering at different scales as their input descriptors. The base classifiers are combined using the proposed *k*-NN voting method and plurality voting. The results are also compared to an approach, in which the filtering results are combined into a single feature vector.

Table 1. The distance matrix for the error residuals of individual base classifiers in the histogram-based classification experiment

	Hue 16	Grey 16	Hue 32	Grey 32	Hue 64	Gray 64
Hue 16	0					
Grey 16	0.38	0				
Hue 32	0.11	0.38	0			
Grey 32	0.36	0.13	0.35	0		
Hue 64	0.18	0.39	0.12	0.37	0	
Grey 64	0.37	0.17	0.36	0.10	0.38	0

Table 2. The distance matrix for the error residuals of individual base classifiers in the filtering-based classification experiment

	Scale 1	Scale 2	Scale 3	Scale 4	Scale 5
Scale 1	0				
Scale 2	0.39	0			
Scale 3	0.47	0.41	0		
Scale 4	0.49	0.47	0.34	0	
Scale 5	0.49	0.47	0.38	0.34	0

Tampereen teknillinen yliopisto
PL 527
33101 Tampere

Tampere University of Technology
P.O. Box 527
FIN-33101 Tampere, Finland