MAURICIO ANDRÉS BARRERA ACUÑA

ROTATION-INVARIANT TEXTURE CLASSIFICATION BASED ON GRAYLEVEL CO-OCCURRENCE MATRICES

SÃO PAULO
2013|

MAURICIO ANDRÉS BARRERA ACUÑA

ROTATION-INVARIANT TEXTURE CLASSIFICATION BASED ON GRAYLEVEL CO-OCCURRENCE MATRICES

Dissertation Paper submitted to the University of São Paulo's Polytechnic School in fulfillment for the award of Master of Science.

Concentration Area:
Electronic Systems

Supervisor:
Associate Professor Dr. Hae Yong Kim

SÃO PAULO
2013

To my parents, always striving for my success.

To a squirrel that lives within me.

ACKNOWLEDGEMENTS

I wish to thank my supervisor, Prof. Hae Yong Kim, for his never-ending patience, objective comments and assertive suggestions. This work shaped into a well-rounded document thanks to his constant collaboration.

To my parents, Edison and Carmen, always supportive, constantly interested and never ceasing to push me forward, I owe them the possibility to stay here and complete this work. I thank them for accompanying me through the deepest and brightest moments of this journey; even beyond the distance I could feel their constant care.

To Débora, there are no words to express her constant care and utmost preoccupation in the success of this journey. For her rock-solid patience, limitless comprehension, perfectly-timed work reminders and for the free intern hours she spent to fill in tables and data. I couldn't wish for a better companion by my side.

To my friends, the ones that were with me from the beginning, the ones that got on in the middle, and the ones that barely got to know about this work, I thank them for the meaningful words of support, appreciation and jokes that threw along the way. For those who listened to or pretended to listen to my endless explanations, thank you very much.

To the ones I leave nameless, the ones I did not even think about, and the ones I left out for some unknown reason, thank you!

ABSTRACT

Texture is one of the most primitive characteristics of objects. Digital images represent this property as local intensity variations in the image. Consequently, texture is an attribute that is innately present in virtually every digital image; mathematically describing this information leads to a myriad of different applications, from food qualities processing up to medical image analysis. It comes as no surprise that texture has been one of the most researched topics in the computer vision community, and it continues to receive a great deal of attention. One of the most classic approaches to model texture is the statistically-based co-occurrence matrix method. The present dissertation work revolves around a clever variation of the co-occurrence matrix that incorporates rotation-invariance, a very desirable property for texture classification. This variation is taken from previous work in the literature and is used to propose a robust fuzzy orthoimage classifier. Moreover, the original rotation-invariant approach is modified though a generalization and benchmarked with one of the most widely-used texture description methods in the recent literature: the Local Binary Patterns approach. The results indicate that Fuzzy Logic is a powerful tool to build texture-based classifiers that have to deal with diversely-sourced image samples; the results also indicate that the generalization proposal, which originates the here named CCM and RCM texture description methods, offers a significant performance boost that is comparable, and even better than Local Binary Patterns approach when comparing accuracy scores.

Keywords: Orthoimage, Texture Classification, Fuzzy Logic, Co-occurrence Matrix

RESUMO

A textura é uma das características mais básicas dos objetos. As imagens digitais representam esta propriedade como variações de intensidade locais na imagem. Consequentemente, a textura é um atributo que se encontra presente de forma inerente em praticamente todas as imagens digitais; descrever matematicamente esta informação produz uma série de inúmeras aplicações, desde processamento da qualidade dos alimentos até análise de imagens médicas. Não é surpreendente que a pesquisa em textura seja um dos tópicos mais pesquisados pela comunidade de visão por computador, e ainda continua recebendo muita atenção. Um dos modelos mais clássicos para modelar textura é o método estatístico da matriz de co-ocorrência. Esta dissertação gira em torno a uma modificação inteligente da matriz de co-ocorrência que inclui invariância à rotação, uma característica muito desejável para a classificação de texturas. Esta variação toma-se de um trabalho na literatura e é utilizada para propor um classificador nebuloso de imagens aéreas. Além disso, a definição matemática original é modificada através da generalização dos descritores e se realiza um estudo comparativo com um dos métodos de descrição de textura mais usados na literatura recente: o método *Local Binary Patterns*. Os resultados indicam que a Lógica Nebulosa é uma ferramenta poderosa para construir classificadores baseados em textura que precisem trabalhar com amostras tomadas de diversas fontes; os resultados também indicam que a proposta de generalização, que leva o nome de método CCM e RCM, oferece um incremento significativo no desempenho que é comparável, inclusive superior, ao método *Local Binary Patterns*, quando comparadas as taxas de acerto para ambos os métodos.

Palavras-chave: Ortoimagem, Classificação de Texturas, Lógica Nebulosa, Matriz de Co-ocorrência.

SUMMARY

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CCM	Circular Invariant Co-occurrence Matrix
COA	Center of Area Method
LBP	Local Binary Patterns
FIS	Fuzzy Inference System
FLC	Fuzzy Logic Controller
GLCM	Graylevel Co-occurrence Matrix
MOM	Mean of Maximum Method
RCM	Radial Invariant Co-occurrence Matrix
RGB	Red, Green and Blue
RMS	Root Mean Square

1. INTRODUCTION

In general, texture can be seen as the distinctive physical composition or structure of objects, especially with respect to the size, shape, and arrangement of its parts. Despite the lack of consensus on a definite characterization in the literature, it is possible to state that texture is perceived by humans and contains an array of information cues that help humans infer facts about the environment that surrounds them.

Consequently, it comes as no surprise that texture description and analysis has gathered an important deal of attention from the computer vision community, with papers presenting works expanding from the classical co-occurrence matrix approach, which aroused in the late seventies, to modern approaches appearing in publications to the current date. A diverse number of applications are based on texture descriptors. A brief literature review yields texture description applications in the most diverse fields, including medical image analysis, food industry, human features recognition and land-cover classification, among others.

Vargas et al. (2011) present an off-line handwritten signature verification system based on texture features, specifically co-occurrence and local binary patterns based descriptors. Alternatively, Zheng et al. (2006) present a comprehensive review on texture descriptors applied to the evaluation of food qualities, such as tenderness, sugar content, Ph value, etc. Another common application is face description; Ahonen and Pietikäinen (2006) apply texture description to successfully classify the publicly available Colorado State University Face Identification Evaluation System (Bolme et al., 2003). On the other hand, Nanni et al. (2010) apply texture descriptors and machine learning techniques to diverse medical image applications, such as pain state classification on neonatal facial images and cell phenotype image classification.

These and many other applications evidence that texture is a extensively used low-level image characteristic in the most diverse knowledge fields. In spite of the years texture description has been under research, it continues to be a very current research topic with a wide array of possible applications. As up-to-date image

acquisition techniques continue to improve and more knowledge areas benefit from image analysis, new techniques with more discrimination prowess and computational simplicity continue to surface.

This dissertation paper is centered on a generalization of the classical co-occurrence matrix approach designed to achieve rotation-invariant texture description. The work is divided on two main parts that describe the work developed throughout the Masters Course extent. The first part presents a successful application for aerial image classification using an early version of the rotation-invariant descriptors (Chapters 5 and 6). The second part introduces a generalization of the rotation-invariant descriptors that considerably improves the performance of the early version and yields excellent results when compared to state-of-the art methods.

1.1. Objectives

The general objective of this work is to propose a robust, reliable and computationally simple rotation-invariant texture descriptor. This work aspires to validate the applicability of the method presenting a reliable classification system for a real task: aerial image classification. It also aspires to validate the reliability and performance of the method presenting a thorough benchmarking with state-of-the-art texture description methods. These two aspirations define the specific objectives of the present work. One of the motivations behind this research is to renew the simpler methods to describe texture, seemingly discarded in the literature, which is the case of the co-occurrence matrices.

1.2. Contributions

This work presents a novel rotation-invariant texture descriptor that surpasses the classification rates achieved by highly regarded methods on the literature, when considering rotation-invariance. This descriptor is based on geometric principles and offers a very simple implementation and robust performance. Furthermore, this work

introduces an alternative approach to achieve the much desirable characteristic of rotation-invariance for aerial region classification.

1.3. Organization

This document is organized as follows:

- The second and third chapters present the basic theoretical foundation for the proposal. These chapters deal with several texture descriptors proposals in the literature while the fourth chapter deals with the foundation of Fuzzy Logic and Fuzzy Inference Systems.

- The fifth chapter describes a practical application of the basic rotation-invariant GLCM texture descriptors: a texture-based fuzzy inference system for orthoimage region classification. The sixth chapter describes the experimental procedures used to validate the system. It also presents the results and a comparison with the classic GLCM texture descriptors.

- The seventh chapter presents a novel and robust mathematical generalization for rotation-invariance that builds upon the proposal used in the fifth and sixth chapters.

- The eighth chapter describes the experimental procedures used to validate the new texture model along with benchmarking with other proposals in the literature. The ninth chapter presents a discussion of the results and further improvement considerations.

- Finally, the tenth and last chapter presents a final discussion and conclusions of the dissertation work.

2. TEXTURE DESCRIPTORS

Texture analysis is a recurring topic in digital image processing. It is directly related to the patterns an image attains from the physical characteristics of the actual object or region that is represented. Although there is no universally accepted definition for visual texture in the literature, several of these delineations share points in common.

In their comprehensive review of texture description methods, Materka and Strzelecki (1998) introduce a practical definition. They define texture as "complex visual patterns composed of entities, or subpatterns, that have characteristic brightness, color, slope, size, etc. Hence, texture can be regarded as a similarity grouping in an image". This definition clearly indicates that texture comes flavored with variations of several perceived properties, such as "lightness, uniformity, density, roughness, regularity, linearity, frequency, phase, directionality, coarseness, randomness, fineness, smoothness, granulation, etc.", as was also stated by Materka and Strzelecki. Therefore, texture is a property that contains information concerning the structural arrangement of surfaces (HARALICK et al., 1973). This definition is adopted throughout the present work.

Texture description methods attempt to represent in a numerical fashion these texture image properties. Hossain and Serikawa (2013) state that "smarter extraction of features from image textures produce better cues for image analysis, which are pivotal for object recognition, surface analysis, action recognition, disease diagnosis, etc." Every new texture description proposal seeks to better describe the aforementioned properties, which results on robust applications.

According to Zhang ad Tan (2002), the existing texture analysis methods can be divided into three main categories: statistical methods, structural methods and model based methods. Structural methods describe texture as a hierarchy of spatial arrangements. Statistical methods describe texture as measures extracted from the statistical analysis of the distribution and relationship of gray levels of an image. Finally, model based methods exploit models such as Random Markov Fields or transform operators such as Wavelets in order to describe texture.

Since the beginning of texture research, a massive set of methods has been proposed in the literature. The proposals based on the steered Hermite transform (ROMERO; RAMIREZ, 2011), Wavelet transform, Gaussian Markov random fields (PORTER; CANAGARAJAH, 1997), as well as Lahajnar and Kovačič (2003) and Porter and Canagarajah (1997) Gabor filters approach represent model-based texture description methods. On the other hand, the co-occurrence matrix (HARALICK et al., 1973) and Local Binary Patterns (OJALA et al., 2002) methods are examples of structure-based texture description. It is noteworthy of mentioning that the co-occurrence matrix method also employs first and second order statistical-based operators to calculate the final descriptors.

Including a detailed description of all these models would represent an enormous research effort on its own merit. Accordingly, this work portrays a detailed introduction solely of some of the most distinctive methods in the literature, namely, the classical co-occurrence matrix and its rotation-invariant variations, Wavelet-based features and Local Binary Patterns. The remaining of this chapter will describe the Wavelet and Local Binary Pattern approaches. Chapter 3 focuses exclusively on the classical and rotation-invariant approaches of co-occurrence matrices in additional detail.

2.1. Wavelet-based Texture Features

In their work, Porter and Canagarajah (1997) analyze three model-based texture description approaches and establish a coherent test suite to decide on the best method. From the Wavelet transform, Gabor filter and Gaussian Markov model approaches they implemented, the Wavelet-based method proved to be the best texture description method with regards to noise and rotation-invariant performance.

The Wavelet transform is a signal processing operation where the higher frequency components of a signal are represented with sharper time resolution than lower frequency components (DAUBECHIES; 1990). The wavelet transform decomposes a signal with a family of orthonormal bases obtained from the translation and dilation of a kernel function known as the mother wavelet. The mother wavelet is derived from a scaling function that is associated to a series of coefficients. These coefficients play an important role on the discrete wavelet transform, Porter and Canagarajah use coefficients derived from Daubechies' approach.

Porter and Canagarajah's method is based on the work by Chang and Kuo (1993); they introduce a tree-based decomposition of the source image based on the analysis of the energy of the resultant images obtained from the 2D wavelet transforms. The wavelet transform is successively applied to the highest energy derived decomposition until a stop criterion is reached, usually given by the smallest accepted scale (16x16 pixel subimages). Fig. 1 illustrates a 10-channel wavelet decomposition scheme. The areas in light gray represent the higher frequency channels (HH) for equally-scaled groups. These channels typically contain noise information of the image which renders them undesirable for texture analysis.
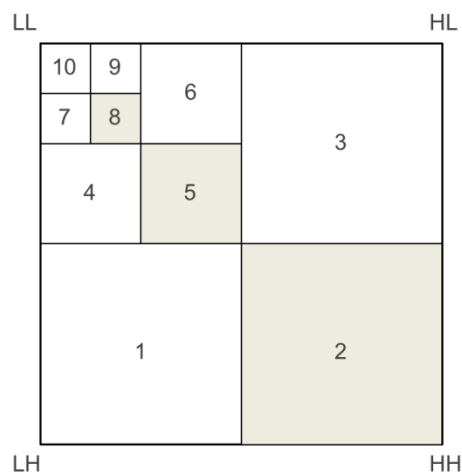
Figure 1 - Tree-based Wavelet Decomposition Scheme. Areas in gray are eliminated from the texture analysis.

On the other hand, the LH channel, for instance, represents the image information associated to low horizontal and high vertical frequencies (ZHANG, TAN; 2002).

The texture information is represented as a feature vector containing the average wavelet coefficient magnitude, as shown in Eq. 1.1.

$$e_{C_n} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} |x(i,j)| \tag{1.1}$$

where the *n*th considered channel holds dimensions of *M* x *N*, *i* and *j* represent the rows and columns of the channel and x represents the wavelet coefficient inside the channel.

The approach achieves rotation-invariance by combining pairs of channels that are opposite by the diagonal to create unique texture features. Then, according to Fig. 1, LH and HL channels at each level of decomposition (1 and 3, 4 and 6, 7 and 9) are paired together. Eq. 1.1 is applied to these pairs of channels to derive the texture descriptors.

May the reader wish for an in-depth presentation of the mathematical formulation for the Wavelet Transform, refer to Daubechies' (1990) paper.

2.2. Local Binary Patterns

The Local Binary Patterns (LBP) approach is a result of a series of works published by researchers of the Center for Machine Vision Research of the University of Oulu. In their well-known paper, Ojala et al. (2002) compile several of these previous works into a texture classification proposal that has been widely used in the literature to solve diverse problems.

In their recent work, Nanni et al. (2010) mention the Local Binary Pattern approach as "one of the most used texture descriptors in medical image analysis" and therefore introduce variants applied to that knowledge area. The face description work (AHONEN et al.; 2006) and signature verification work (VARGAS et al.; 2011) presented in the introduction chapter are examples of applications of the LBP

approach. Other approaches combine the texture description of LBP and the spatial relationship description of co-occurrence matrices, as seen on the work of Pok et al. (2005) and the more recent work of Sun et al. (2012).

The LBP method is a multiresolution approach for texture description and classification that is based on the recognition of certain uniform local binary patterns. These patterns are primitive properties of the images and represent typical texture types on them. The patterns are extracted from a local set of pixels around a circular area, represented by a center pixel g_c and P pixels g_p around a circumference with radius R. Fig. 2 presents a group of local circular regions with several P and R values.

| $P = 8, R = 1.0$ | $P = 12, R = 1.5$ | $P = 16, R = 2.0$ | $P = 24, R = 3.0$ |
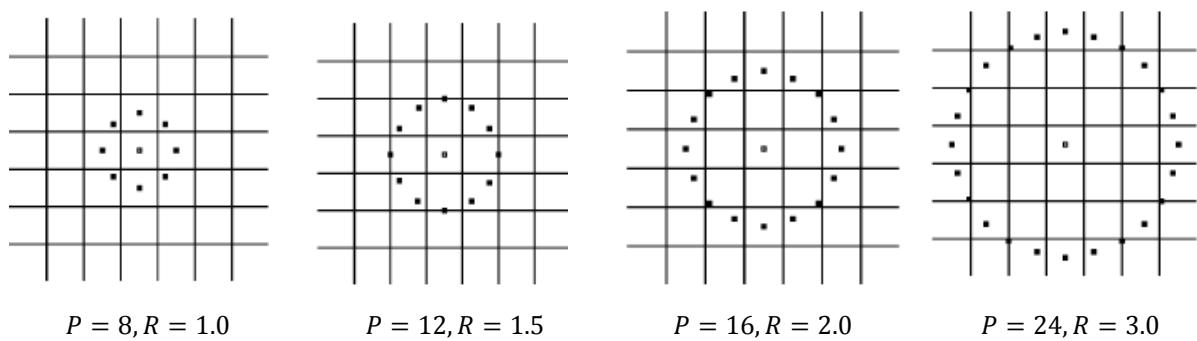
Figure 2 – Set of circular sets with varying P and R values.

The LBP method measures the local graylevel sign with respect to the center pixel. Only considering the sign of the image makes the descriptors invariant to any monotonic transformation of the grayscale of the image. The binary patterns are extracted from this sign-based analysis. In their study, Ojala et al. state that only some patterns provide useful texture information and are present in the vast majority of the images. These patterns are shown on Fig. 3.

Figure 3 - Uniform patterns for $P = 8$.

For a given *P* value, *P + 1* uniform patterns will emerge. A uniform pattern is defined as a circular succession of binary points where at most 2 binary transitions occur. Considering this, Eq. 1.2 presents the LPB descriptor definition.

$$
LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & if\ U < \ 2 \\ P+1 & otherwise \end{cases}
\tag{1.2}
$$

where U represents the number of binary transitions, $s(x)$ is a sign function with outputs of 0 and 1, and the superscript $riu2$ indicates that only the uniform patterns are used. The g_p and g_c pixels are taken according to *P* and *R*. If a pixel g_p falls outside of the center of the pixel, its value is estimated through linear interpolation.

The $LBP_{P,R}^{riu2}$ operator discards contrast information of the image. To solve this shortcoming, the authors propose a measure of local variance to complement the LBP method. This operator, named as $VAR_{P,R}$, is not invariant to grayscale variations. Eq. 1.3 presents its mathematical formulation.

$$
VAR_{P,R} = \frac{1}{P} \sum_{p=0}^{P-1} (g_p - \mu)^2
\tag{1.3}
$$

where μ represents the mean value over the set of g_p pixels. In the same fashion as the $LBP_{P,R}^{riu2}$ operator, the g_p and g_c pixels are taken according to *P* and *R*.

Both operators are used to create a histogram model of the local patterns found in an image. The $LBP_{P,R}^{riu2}$ operator yields a quantized output due to its definition. On the other hand, the $VAR_{P,R}$ operator yields a continuous output; the model histogram is created by accumulating the operator outputs of all the model images in a single distribution and dividing the entries in *B* bins.

In order to compare two texture images, the authors propose to use a histogram dissimilarity metric, the log-likelihood statistic, which is presented on Eq. 1.4.

$$L(S, M) = \sum_{b=1}^{B} S_b \log M_b \qquad (1.4)$$

where B refers to the number of bins, S_b and M_b correspond to the sample and model probabilities at bin b, respectively.

The LBP method suggests that the joint distributions of $LBP_{P,R}^{riu2}$ and $VAR_{P,R}$ are a powerful and complementary texture descriptor. The approach is said to be multiresolution given the possibility of combining histograms for different R values to create a single model. The P value controls the quantization degree around the circumference line.

3. GRAYLEVEL CO-OCCURRENCE MATRIX TEXTURE DESCRIPTORS

Haralick et al. (1973) proposed a statistical texture analysis method that has been widely applied in the literature and has become an initial benchmarking point for several texture descriptor approaches. Their method extracts first and second order statistical features from a matrix constructed with the gray-level co-occurrences in an image. The rotation-invariant descriptors proposed in this work are a variation of Haralick's descriptors.

The co-occurrence matrices for texture description have received a great deal of attention since their introduction to the scientific community. Works as recent as Clausi's (2002) study in detail the original co-occurrence matrix approach; Clausi's work studies the effect of quantization over the classification accuracy for individual statistic descriptors. Another example is Siquiera et al.'s (2013) work, which introduces a multi-scale approach for graylevel co-occurrence matrices based on the Gaussian smoothing and pyramidal decomposition. Other researches use co-occurrence matrices as the basis for new proposals for texture description; one example is the case of Liu and Yang's (2008) work of textons and co-occurrence matrices or the already mentioned work by Sun et al. (2012), where they incorporate local texture patterns to co-occurrence matrices.

The remaining of this chapter is divided into three sections that revolve around co-occurrence matrix theory and texture description applications. The first subchapter introduces the classical Haralick's approach. The second chapter presents a brief introduction on some rotation-invariant proposals for the co-occurrence matrix approach. The third and last subsection presents in detail a robust rotation-invariant approach introduced by Ito et al. (2009), which serves as the basis for the generalized descriptors proposed in this work.

3.1. Haralick's Graylevel Co-occurrence Matrix

The classical graylevel co-occurrence matrix (GLCM) texture descriptors are easily computable statistical features that rely on the graylevel spatial dependencies. The method consists on generating a matrix that captures the graylevel spatial dependencies and calculating measures of the textural characteristics or complexity of the graylevel transitions from that matrix.

Formally, an image I is a function that associates a graylevel to each pixel, that is, $I: D \rightarrow G$, where D is the domain of the image and $G = \{1,2,\dots,L_g\}$ is the set of graylevels. The graylevel co-occurrence matrix $P_{(\Delta x, \Delta y)}$ is a square matrix of size $L_g \times L_g$ and parameterized by an offset $(\Delta x, \Delta y)$ such that:

$$P_{(\Delta x,\Delta y)}(i,j) = \sum_{p=1}^{n}\sum_{q=1}^{m} \begin{cases} 1, & \text{if } I(p,q) = i \text{ and } I(p+\Delta x, q+\Delta y) = j \\ 0, & \text{otherwise} \end{cases} \quad (3.1)$$

In this equation, *n* refers to the number of columns and *m* refers to the number of rows of the image. The selection of the offset $(\Delta x, \Delta y)$ relates to an angle parameter θ. This angle refers to the direction of the pixel chosen to make the gray-level comparison. Typical values for θ are 0°, 45°, 90° and 135°, which relate to the $(\Delta x, \Delta y)$ offsets $(1,0)$, $(1,1)$, $(0,1)$ and $(-1,1)$ respectively. Before extracting the texture descriptors from the matrix, the GLCM is normalized. To achieve this, each element is divided by the sum of all the elements of the matrix. The normalized GLCM is represented as $p(i,j)$.

In their original work, Haralick et al. (1973) indicated 14 texture descriptors that could be extracted from the GLCM. Baraldi and Parmiggiani (1995) examined these 14 GLCM texture descriptors and concluded that some of them were more statistically significant. These descriptors are the homogeneity, contrast, entropy, correlation, energy and variance.

Eq. 3.2 - 3.7 present the expressions for the homogeneity, contrast, entropy, correlation, energy and variance, respectively.

$$f_{hom} = \sum_{i=1}^{L_g} \sum_{j=1}^{L_g} \frac{1}{1+(i-j)^2} p(i,j) \tag{3.2}$$

$$f_{con} = \sum_{n=0}^{L_g-1} n^2 \left\{ \sum_{i=1}^{L_g} \sum_{j=1}^{L_g} p(i,j) \right\}_{|i-j|=n} \tag{3.3}$$

$$f_{ent} = -\sum_{i=1}^{L_g} \sum_{j=1}^{L_g} p(i,j) \log(p(i,j)) \tag{3.4}$$

$$f_{corr} = \sum_{i=1}^{L_g} \sum_{j=1}^{L_g} \frac{(ij)p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{3.5}$$

$$f_{ener} = \sum_{i=1}^{L_g} \sum_{j=1}^{L_g} \{p(i,j)\}^2 \tag{3.6}$$

$$f_{var} = \sum_{i=1}^{L_g} \sum_{j=1}^{L_g} (i-\mu)^2 p(i,j) \tag{3.7}$$

where $p(i,j)$ is the normalized co-occurrence matrix; $p_x(i,j) = \sum_{j=1}^{L_g} p(i,j)$ and $p_y(i,j) = \sum_{i=1}^{L_g} p(i,j)$ are the marginal distributions; μ_x and μ_y are the mean value of the marginal distributions; σ_x and σ_y are their standard deviations; and $\mu = \sum_{i=1}^{L_g} \sum_{j=1}^{L_g} i\, p(i,j)$.

These descriptors are combined into feature arrays that describe the texture over a certain image. Usually, texture image classification is done by the direct comparison of the sample image and test image feature arrays by the means of a particular distance or similarity measurement. It is noteworthy to state that this texture description approach is not rotation invariant; the texture model described here only takes into account the texture orientation of a single angle, rendering the descriptors highly sensible to any rotation to the capture angle of the image.

3.2. Rotation-invariant Approaches for the Co-occurrence Matrix

The graylevel co-occurrence matrices are parameterized using pixel-to-pixel distance and an angle. This angle parameter has been exploited to achieve rotation-invariance for the GLCM method. This section briefly presents two straightforward extensions of the co-occurrence matrix using the angle parameter. The first method is taken from Maillard's (2003) classification-based texture analysis methods review and the second method is taken from Zhang and Tan's (2002) review of invariant texture analysis methods. The two depicted methods describe the general approach that is usually followed to generate rotation-invariant GLCM descriptors.

3.2.1. Multi-angled GLCM Feature Extraction

Maillard's reported method calculates several angled GLCM as a starting point for rotation-invariance. He presents a calculation example with matrices computed for a three pixels distance scale. The method states that several co-occurrence matrices should be calculated for a given set of angles, then, the method starts by composing a set of co-occurrence matrices $S_{GLCM} = \{P_{\theta_1}, P_{\theta_2}, \dots, P_{\theta_n}\}$, where n represents the number of considered angles. For instance, take a set of four angles: 0°, 45°, 90° and 135° calculated for a three pixel distanced matrix. Following the notation introduced in section 3.1, the four matrix set $\{P_{(3,0)}, P_{(3,3)}, P_{(0,3)}, P_{(-3,3)}\}$ is computed to represent all angles.

The statistical features introduced in section 3.1 are computed from each co-occurrence matrix and stored into arrays. The rotation-invariant texture descriptors are defined as the mean and standard deviation over each set of angle-related features. For instance, the statistical feature f was calculated from all the matrices of set S_{GLCM}. Then the rotation invariant features $\bar{f}$ and f_σ are calculated as the arithmetic mean and standard deviation of the feature array $f_{GLCM} = \{f_{\theta_1}, f_{\theta_2}, \dots, f_{\theta_n}\}$.

3.2.2. Polar plots and polarograms

As stated by Zhang and Tan (2002), a polarogram is a tool for image texture analysis used to get invariant texture features. In more formal terms, a polarogram is a "polar plot of a texture statistics as a function of orientation". This definition leads to the forthright deduction that centers the graylevel co-occurrence matrix's angle parameter as the center of this rotation-invariant approach.

Let $D(a)$ be a displacement array of fixed magnitude d and variable orientation a. A polarogram is defined in Eq. 3.8.

$$P_f(a) = f\big(C_{D(a)}\big) \qquad (3.8)$$

where the function f argument $C_{D(a)}$ is the graylevel co-occurrence matrix for displacement $D(a)$, which corresponds to the displacement $(\Delta x, \Delta y)$ offsets defined in section 3.1. The election of $(\Delta x, \Delta y)$ will define the orientation a.

The shape and size of the polarogram define the rotation-invariant texture features. The rotation-invariance is guaranteed given that a rotation on the image domain translates into a rotation of the polarogram. Several polarograms associated to different displacements $D(a)$ can be computed to model a determined texture.

3.3. Ito et al.'s Circular and Radial GLCM Approach

The rotation-invariant approaches for the co-occurrence matrix presented so far attempt to capitalize on the angle parameter of the original GLCM definition. None of them proposes a variation on the definition *per se.* There is, nonetheless, an approach that redefines the GLCM to achieve rotation-invariance from its core. This definition was introduced by Ito et al. (2009) and had the contribution of the advisor of the present dissertation paper.

Ito et al. introduced two types of co-occurrence matrix: circular and radial. The statistical descriptors calculated from these matrices have the desirable rotation-invariance characteristic. Both matrices exploit the geometric properties of circumferences and radii lines to achieve the desired invariance characteristics. The following subsections 3.3.1 and 3.3.2 present in detail the definition of the circular and radial co-occurrence matrices.

3.3.1. Circular Co-Occurrence Matrix

As its name indicates, the circular co-occurrence matrix is based on measurements over circular shapes. Ito et al. employ such geometry in order to exploit the innate rotation-invariance this shape offers.

The circular co-occurrence matrix is calculated from the mean grayscale occurrences of over two concentric circular rings. These circular rings are measured by their radii; the internal ring holds a radius r_i and the external ring holds a radius r_o. Fig. 4 shows an example of the geometrical representation of the rings used to calculate the circular co-occurrence matrix.
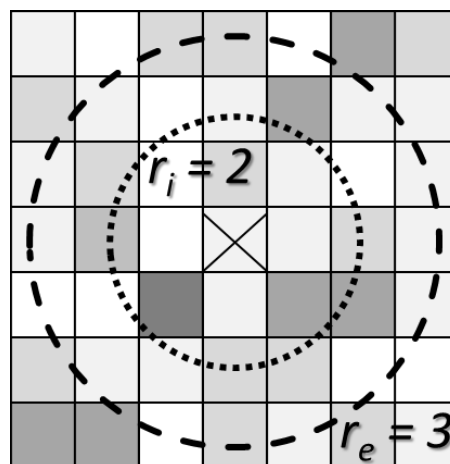
Figure 4 - Calculation of average grayscale using circular rings with radii $r_i = 2$ and $r_o = 3$.

Now, let's denote $F_r(p, q)$ as the mean gray-level on the circle ring centered at (p, q) with radius r. Let's then define the circular co-occurrence matrix $P_{(r_i, r_o)}$ with internal radius r_i and external radius r_o as shown on Eq. 3.9.

$$P_{(r_i, r_e)}(i, j) = \sum_{p=1}^{n} \sum_{q=1}^{m} \begin{cases} 1, & if\ F_{r_i}(p, q) = i\ \text{and}\ F_{r_o}(p, q) = j \\ 0, & otherwise \end{cases} \tag{3.9}$$

The radii values used in the specific example of Fig. 4, $r_i = 2$ and $r_o = 3$, yielded the best results in the experimental setup found on the original work of Ito et al. and are consequently used to model texture on the application presented in the fifth and sixth chapters.

3.3.2. Radial Co-Occurrence Matrix

The radial co-occurrence matrix is based on measurements over radial lines. The radial line geometry is not rotation-invariant by definition. However, the statistical descriptors extracted from the radial co-occurrence matrix are rotation-invariant. The radial co-occurrence matrix is calculated from the weighted average of the pixel values that belong to a set of diversely-angled radial lines. These radial lines are concentric and equally spaced by an angle θ and circumscribed inside a circumference of radius r. Fig. 5 shows an example of the geometrical representation of the radial lines used to calculate the radial co-occurrence matrix.

Now, let's denote $F_{(r,\theta)}(p, q)$ as the mean gray-level on radial line centered at (p, q) with radius r and inclination θ. The radial co-occurrence matrix with radius r and N uniformly-spaced radial lines is described as:

$$P_{(r, N)}(i, j) = \sum_{p=1}^{n} \sum_{q=1}^{m} \sum_{l=0}^{N-1} \begin{cases} 1, & \text{if}\ F_{\left(r, \frac{2\pi l}{N}\right)}(p, q) = i\ \text{and}\ F_{\left(r, \frac{2\pi(l+1)}{N}\right)}(p, q) = j \\ 0, & \text{otherwise} \end{cases} \tag{3.10}$$
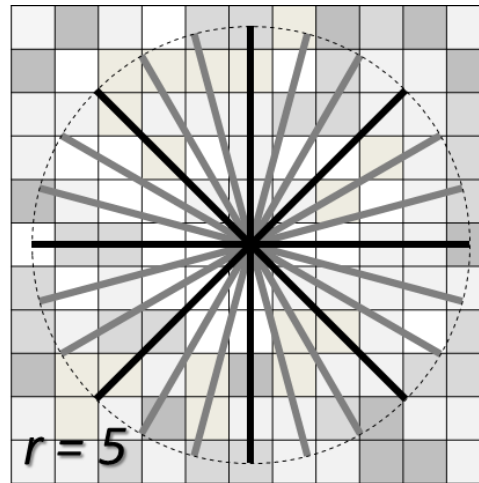
Figure 5 - Calculation of average grayscale on N=8 directions with radius r=5. We used 4 radial lines to compute the average in each direction.

In Eq. 3.10, m and n represent the indexes of the pixels of the image. The sum (l+1) must be computed modulus N. In order to enhance the robustness of the method, it is possible to use the weighted mean of some number of neighboring radial lines, instead of using the mean of only one radial line. This is roughly equivalent to low-pass filtering the image. The specific example of Fig. 5 illustrates the use of the mean of four radial lines (external two black lines and central two gray lines) to compute each $F_{(r,\theta)}(p,q)$.

The parameter values r = 5 and N = 8 yielded the best classification results in the experimental setup found on the original work of Ito et al. and are consequently used to model texture on the application presented in the fifth and sixth chapters.

3.3.3. Matrix Normalization and Descriptor Definition

Ito et al. normalize the circular and radial co-occurrence matrices in order to extract the statistical descriptors. The co-occurrence matrix normalization was first proposed by Haralick et al. in their original work. The matrices are normalized by dividing each

element by the sum of all the matrix elements. A normalized co-occurrence matrix $p(i,j)$ is given by the following expression:

$$p(i,j) = \frac{P(i,j)}{\sum_{i=1}^{L_g} \sum_{j=1}^{L_g} P(i,j)} \tag{3.11}$$

After normalizing both matrices, the statistical descriptors can be extracted from each matrix. The same statistical descriptors presented in Section 3.1 can be used here.

4. FUZZY LOGIC AND FUZZY INFERENCE SYSTEMS

The aerial image classification application presented in the two forthcoming chapters is based on a classifier designed as a Fuzzy Inference System. This chapter introduces the base concepts of the Fuzzy Theory needed to understand the design of Fuzzy Expert Systems.

4.1. Principles of Fuzzy Logic

Fuzzy Logic is based on the well-known concept of Fuzzy Sets, introduced by Zadeh (1965) in his reputable paper. Intuitively, Fuzzy Logic generalizes the basic membership notion of Classical Logic, which dictates that an element belongs or does not belong to a certain class. This characteristic of classical sets is referred to on Fuzzy Theory as "crisp" membership functions.

A Zadeh well puts it, "a fuzzy set is a class of objects with a continuum of grades of memberships. Such a set is characterized by a membership […] function which assigns to each object a grade of membership ranking between zero and one". Such continuous-valued membership functions, usually represented as μ_f, allow for a completely broader description of phenomena. For instance, one could say that the water is "moderately hot" or that the pressure level is "slightly low" and model those facts just by defining a proper membership function to a set, for example, "temperature" or "pressure level" and assigning suitable values. Another definition by Zadeh (1988) perfectly portrays this line of thought when he states that Fuzzy Logic is "the logic underlying approximate, rather than exact, modes of reasoning".

Zadeh (1988) also presents another fundamental concept in Fuzzy Logic: linguistic variables. As it name suggests, a linguistic variable holds values that are words; for instance, the labels "temperature" and "pressure level" of the paragraph just above are typical examples of linguistic variables. The membership functions in Fuzzy Logic represent the degree of attachment that an object has to a linguistic variable.

Lee (1990a) presents a suitable mathematical definition for the concepts so far presented. Formally, a fuzzy set *F* in an universe of discourse *U* is characterized by a membership function μ_f in the interval [0,1], where $U \rightarrow [0,1]$, is represented by the expression:

$$F = \sum_{i=1}^{n} \mu_f(u_i)/u_i \tag{4.1}$$

where u is a generic element of the set, U represents a discrete set of values that hold the scale and units the generic object is measured as.

Usually, linguistic variables as represented as fuzzy numbers. A fuzzy number is a fuzzy set *F* in a universe of discourse *U*, in which *F* is normal and convex (LEE, 1990a). There are several types of fuzzy numbers in the literature. Fig. 6 presents two of the most typical fuzzy numbers used in the literature.

(a) Triangular Fuzzy Number   (b) Trapezoidal Fuzzy Number
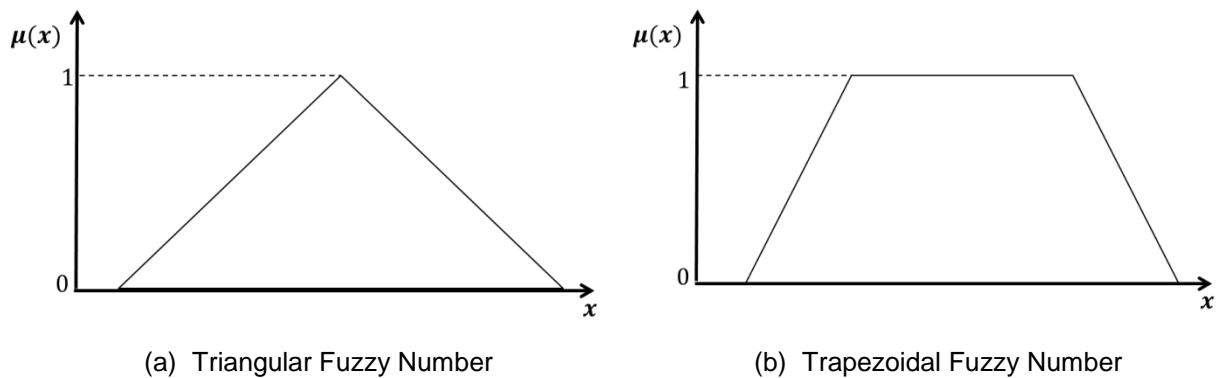
Figure 6 - Graphical representation of the triangular and trapezoidal fuzzy numbers.

4.2. Fuzzification and Fuzzy Rules

Fuzzification refers to the action of taking input variables from the real-world domain into the fuzzy membership domain. Fuzzification includes creating a sound set of linguistic labels, represented by appropriate fuzzy numbers, to represent a linguistic variable and its universe of discourse. A linguistic label is, typically, an adjective that

represents a quality of a linguistic variable. For instance, the linguistic variable "temperature" can be said to be "cold", "warm" or "hot". The election of the complete set of linguistic labels and fuzzy numbers is determinant when creating a fuzzy expert system.

Fuzzy systems can model expert knowledge in the form of a rule set. Fuzzy rules establish relationships between input fuzzy variables to determine the value of an output fuzzy variable; they usually follow the structure that follows:

$$R_1: if\ x\ is\ A_1 and\ y\ is\ B_1\ then\ z\ is\ C_1$$
$$R_2: if\ x\ is\ A_2 and\ y\ is\ B_2\ then\ z\ is\ C_2$$
$$\dots \dots \dots \dots \dots$$
$$R_n: if\ x\ is\ A_n and\ y\ is\ B_n\ then\ z\ is\ C_n$$

where *n* represents the number of rules, *x*, *y* and *z* are linguistic variables (the first two being input variables and the last being output variables); A,B and C are linguistic labels of the linguistic variables *x,* y and *z*; and the logical operator *and* could also be any other logical operator, for instance, the operator *or*.

The result of the output fuzzy variables *z* is obtained through fuzzy implication; this refers to the principles followed to derive an output fuzzy variable out several input variables. Mamdani's (1977) implication rule is one of the most known and widely applied fuzzy implication principles. Mamdani's principle states that the output of two input variables related by an *and* operator is taken as the minimum value of both membership functions; for the case of the *or* operator, the output will be the maximum value of both membership functions.

4.3. Fuzzy Inference Systems and Defuzzification

A Fuzzy Inference System (FIS) is a system that models expert knowledge using Fuzzy Logic Principles. Lee (1990a) defines the constituent components of a Fuzzy Logic Controller (FLC), which can be regarded as an application of a FIS. He states

that there are four main components for a FLC: a Fuzzification interface, a knowledge base, a decision-making logic, and a Defuzzification interface.

The Fuzzification interface measures the values of the input variables and fuzzifies the data into linguistic variables. The knowledge base comprises the knowledge of the application domain and refers to the fuzzy rule set. The decision making logic is related to the implication method, with Mamdani's approach presented as an example. These three components were already introduced in the last paragraphs.

The final component is the Defuzzification interface. The output fuzzy variables obtained from the decision-making component still remain in the fuzzy domain. The defuzzification interface refers to operations in the fuzzy domain that derive a crisp data point in the universe of discourse; this output can be interpreted as a decision for a FIS or a control action for a FLC.

In the second part of his paper, Lee (1990b) presents some of the most commonly used defuzzification methods. Two of these are the Mean of Maximum Method (MOM) and the Center of Area Method (COA). The MOM approach generates an output crisp response based on the mean value of the output functions that hold the highest membership value; this is expressed in Eq. 4.2.

$$z_0 = \sum_{i=1}^{M} \frac{u_i}{M} \tag{4.2}$$

where u_i is the value at which the membership function is maximum, and M is the number of such values.

The COA approach calculates the centroid, or center of gravity, of the possibility distribution of the output variable. This distribution is composed by the union of the contribution of all the rules. Eq. 4.3 shows its mathematical description for a discrete universe.

$$z_0 = \frac{\sum_{i=1}^{N} \mu_z(u_i) \cdot u_i}{\sum_{i=1}^{N} \mu_z(u_i)} \tag{4.3}$$

where N is the number of quantization levels, $\mu_z(u_i)$ is the value of the possibility distribution at the i-eth level.

These four components are the main constituents of a Fuzzy Inference System and are the result of continuous research in the area for several decades.

5. ROTATION-INVARIANT ORTHOIMAGE FUZZY CLASSIFIER

Satellite and aerial images have a very wide range of applications. They are relevant in numerous fields such as agriculture, geology, land-use, regional planning and disaster prevention. Aerial images are taken from low-altitude airplanes. These images are usually geometrically corrected in order to obtain an accurate representation of the Earth's surface, which is achieved by eliminating camera distortions. These images are called Orthorectified Aerial Images, or orthoimages.

One particularly relevant orthoimage application is land-use classification. Typically, land-use changes over time as new areas are for instance, urbanized, deforested or natural events dramatically change the landscape. Remote sensing techniques and aerial images can be used to complement ground observations to achieve an accurate inventory on a region's land-use (ANDERSON et al., 1976). That region could refer to a state or a whole nation.

Intuitively, one can perceive that different land-use types possess different grayscale variation patterns. For instance, a crowded city region of a metropolis is expected to show high contrast variations due to the buildings array and the sun illumination. On the other side, sea regions may be softer in contrast, especially in off-shore regions that display few wave patterns. Taking these facts into consideration, texture comes as a natural choice to represent the land-use information found in orthoimages.

This chapter describes the formulation of an orthoimage classifier build solely using the texture descriptors from Subchapter 3.3 and the Fuzzy Logic principles of Chapter 4. Subsequently, Chapter 6 presents the experimental approach and a comparison with the traditional GLCM approach. The application is motivated by the will to establish whether the modified co-occurrence matrix is suitable for real applications.

5.1. Initial Descriptor and Orthoimage Considerations

In order to build the orthoimage region classifier of this work, the only descriptors used are the homogeneity, contrast and entropy descriptors. A Fuzzy Inference System usually benefits from a reduced input variable set; an initial mapping of the descriptor values for the considered orthoimages brought forth the fact that these three descriptors are the most discriminative for the application in mind, since their data distribution falls in somewhat separated intervals.

The descriptors are extracted from the normalized circular and radial co-occurrence matrices. Consequently, six separate rotation-invariant descriptors are obtained, three related to the circular co-occurrence matrix and three related to the radial version. Each pair of descriptors (e.g.: circular and radial contrast) is taken to derive three new features using the root of mean square of each pair:

$$f_i^{rms} = \sqrt{\frac{(f_i^{cir})^2 + (f_i^{rad})^2}{2}} \tag{5.1}$$

where $i \in \{hom, con, ent\}$; f_i^{cir} denotes the three circular descriptors and f_i^{rad} denotes the three radial descriptors. The three root mean square texture features f_i^{rms} are used in the remaining of Chapters 5 and 6.

Now, let us consider three different land cover types for the orthoimage region classification problem: sea, forest and city. A set of orthoimages containing samples of these three region types was acquired and used to build a training database. These images were extracted from the Google Maps and Microsoft Bing Maps services. Grayscale images were obtained from the color ones by adding the RGB components with weights 0.3, 0.59 and 0.11, respectively (BALA and BRAUN; 2004).

A total of 56 images, with an average size of approximately 360x240 pixels, were gathered for training purposes. These training orthoimages were divided in 20x20 pixel blocks, obtaining over 12000 training blocks. In order to validate the desired

rotation-invariance characteristic, 90 additional test images were gathered; 6 rotated versions of each test image were generated, obtaining a total of 630 images. Bilinear interpolation was used to rotate the images in angles of $360°n/7$, where $n \in \{1, 2, ..., 6\}$.

The three RMS features f_i^{rms} were calculated for each 20x20 pixels block of the training images. This block size maintains the textural characteristics of the three regions. The blocks are extracted without overlapping and without considering pixels outside the image. As previously stated on Subchapter 3.3, a radii of 2 and 4 pixels were used for the circular co-occurrence matrix; besides, a radius of 5 pixels and a set of angles varying in steps of 45° were used for the radial co-occurrence matrix. Furthermore, the total graylevel count for the input images was 128 gray-levels; the input images were quantized from 256 to 128 levels by a straightforward division of the original image graylevels by two and subsequent rounding of the result to the lower nearest integer.

The data obtained applying the descriptors to the training images was analyzed to generate the referred FIS; this analysis is presented in the following subchapter. It's worth mentioning that there are some works that use tailored FIS for satellite images classification, for example (BINAGHUI et al., 1997), (WEI, 2010). Machine learning techniques have also been used to achieve partial rotation invariance and these approaches have been applied to the traditional GLCM (SALEM; NASRI, 2011). The following subchapters present the data analysis and the subsequent FIS development.

5.2. Training Data Analysis

The training data consists of three RSM descriptors per image block. The training images were labeled as sea, forest or city, accordingly. Consequently, 3 descriptor values were obtained for each annotated 20x20 pixel block.

After obtaining the raw training data, the next taken step was to plot the data distribution for each type of descriptor (homogeneity, contrast and entropy) separated by the three region labels (sea, forest and city). The distribution plots show the data intervals of each descriptor type for the three annotated regions. Instead of the scattered data points, the plots only show the first quartile, third quartile and arithmetic mean of each training descriptor to visualize the data. Fig. 7 presents the distribution plots of each one of the three RMS descriptors.

(a) Contrast

(b) Homogeneity
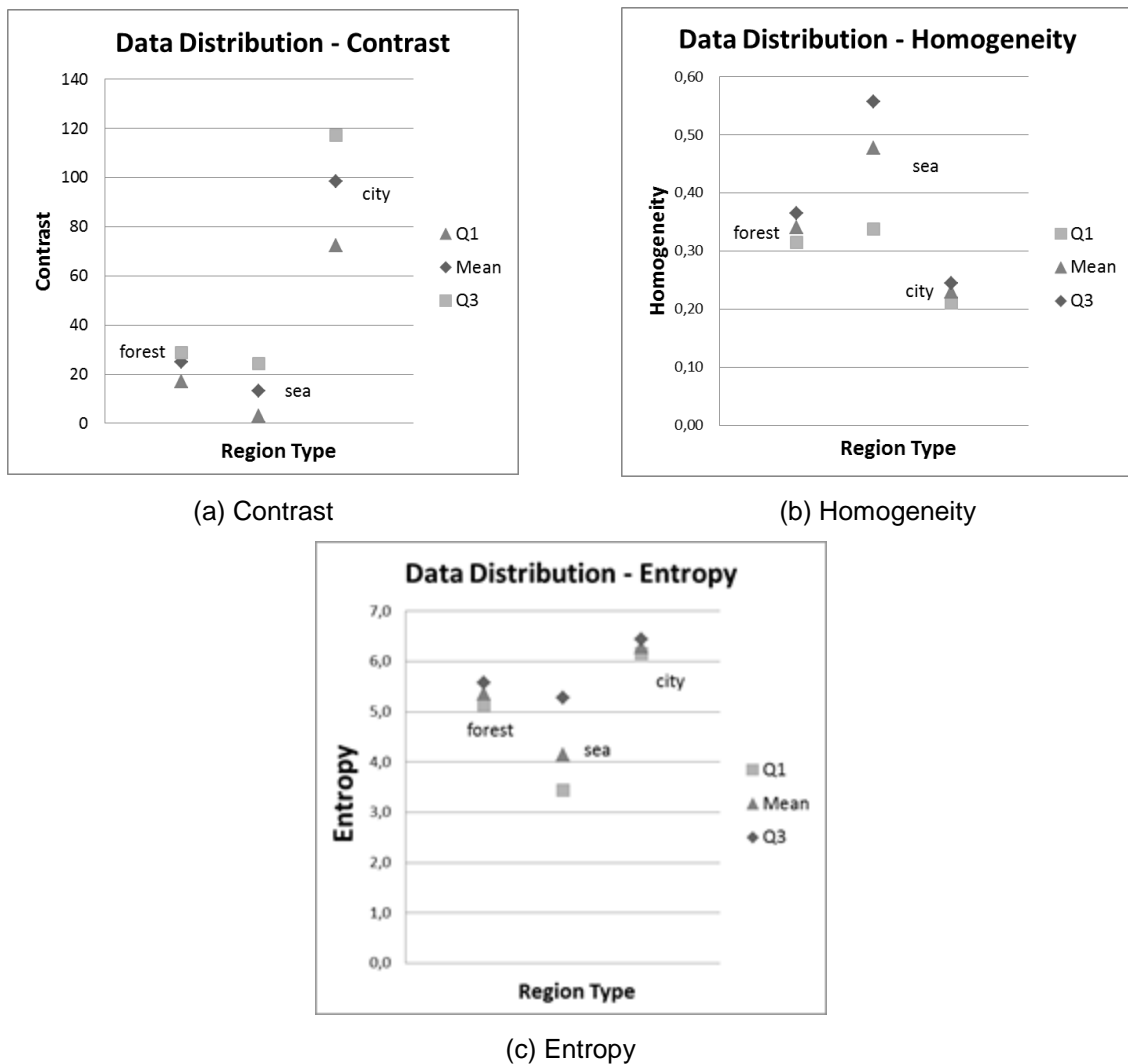
(c) Entropy

Figure 7 – Data Distribution of the Rotation-Invariant RMS Features

The plots of Fig. 7 show that the distribution of the contrast, homogeneity and entropy features falls into several well-defined intervals. If one considers each region-type as a data class, it can be stated that the values for each RMS feature are somewhat clustered. Even so, it is not possible to linearly separate each cluster from

another. This impossibility comes from the fact that the clusters overlap over each feature domain. This cluster overlapping, however, leads to the choice of Fuzzy Logic principles to build a classification system. The overlapping can be treated as fuzziness, that is, soft transitions between clusters.

5.3. Deriving a Fuzzy Inference System

It is now clear that the data distribution of the descriptors falls into several well-defined intervals. Intuitively, one can associate "labels" or "degrees" to every single cluster. Let's make this idea clearer; for instance, the data points for a forest region would fall into a "low" contrast level and a "medium" homogeneity level. In practical terms, fuzzy intervals are associated to every region-descriptor pair. This is the very first step that is taken in order to build a FIS.

Continuing with this reasoning, the RMS texture features are associated to input linguistic variables and linguistic labels are assigned to the data clusters. Three output linguistic variables are defined. Each output variable refers to the possibility that an input region belongs respectively to a sea, forest or city-type. All output variables fall into the range [0,100]. Then, a group of trapezoidal membership functions are defined from the data intervals and associated to the clusters of the rotation-invariant features. The universe of discourse for each input variable was set as the maximum possible range of values each feature takes.

Each input linguistic variable holds two or three linguistic labels. This means that a certain linguistic label could be associated to the training vectors of one or two region-types. Table 1 summarizes the associations between the linguistic labels and region types for the homogeneity, contrast and entropy features. The information presented in Table 1 can be visually verified using Fig. 7.

Feature / Linguistic Label	Low	Medium	High
Homogeneity	City	Forest	Sea
Contrast	Forest/Sea	--	City
Entropy	Sea	Forest	City

Table 1 - Region Clusters Associated to Linguistic Labels per RMS Feature.

Continuing with the FIS definition, $\mu(x)$ is denoted as the membership function associated to any input linguistic label; x belongs to the universe of discourse of the corresponding input linguistic variable and ρ is an array that contains all the training data values associated to the linguistic label. The equation below is proposed to model the membership function of any input linguistic label:

$$\mu(x) = \begin{cases} (x - \mu_\rho + 2\sigma_\rho)/(Q_1 - \mu_\rho + 2\sigma_\rho) & , & \mu_\rho - 2\sigma_\rho \leq x < Q_1 \\ 1 & , & Q_1 \leq x < Q_3 \\ (\mu_\rho + 2\sigma_\rho - x)/(\mu_\rho + 2\sigma_\rho - Q_3) & , & Q_3 < x < \mu_\rho + 2\sigma_\rho \\ 0 & , & \text{otherwise} \end{cases} \tag{5.2}$$

where σ_ρ is the standard deviation of ρ, μ_ρ is the arithmetic mean of ρ, Q_1 is the first quartile of ρ and Q_3 is the third quartile of ρ. This equation generates trapezoidal membership functions. Fig. 8 illustrates a general representation of the trapezoidal functions that Eq. 5.2 produces.
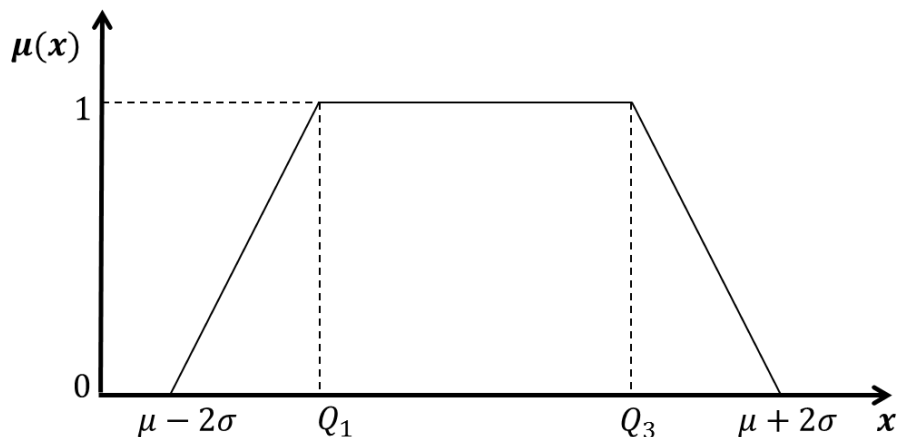
Figure 8 – General Trapezoidal Functions for the Input Linguistic Variables.

The last step to define the initial FIS elements remains in the definition of membership functions of the output variables. These are defined as three triangular

fuzzy numbers associated, respectively, to the linguistic labels "not likely", "likely" and "very likely". These considerations build up the fundamentals for a Fuzzy Inference System.

5.4. Fuzzy Inference System Description

Subchapter 5.3 holds the basis to derive a fuzzy inference system from the training data. The aforementioned considerations are applied into the FIS design. These constitute the complete system for rotation-invariant orthoimage classification.

The FIS consists of a set of input linguistic variables, a set of output linguistic variables, a set of rules that express the relationship between both types of linguistic variables and a defuzzification method. The rules are applied under a specific fuzzy inference method. The input and output linguistic variables were defined in the previous subchapter. The knowledge base is represented by 15 fuzzy rules that characterize the relations between the feature value intervals and the associated outputs. These rules are applied under Mamdani's inference method (LEE, 1990); the defuzzification method of choice is that of the center of gravity (Eq. 4.3).

Fig. 9 plots all three input linguistic variables and output linguistic variables. The plots also depict the linguistic labels attached to each linguistic variable. In addition, Table 2 presents the complete set of rules for the knowledge database. The rules are designed to exploit the combinations between the data cluster of the regions and the associated features. The complete Fuzzy Inference System is built around the input and output variables of Fig. 9, the fuzzy rule set of Table 2, the choice of Mamdani's inference method and the use of the centroid defuzzification method.

After applying the defuzzification method to the output variables, the system delivers a specific percentage value for each of them. This percentage refers to the likelihood that an input 20x20 pixel block has of belonging to each one of the possible region-types. This likelihood value presents is the resulting classification tool for the land-use image blocks.
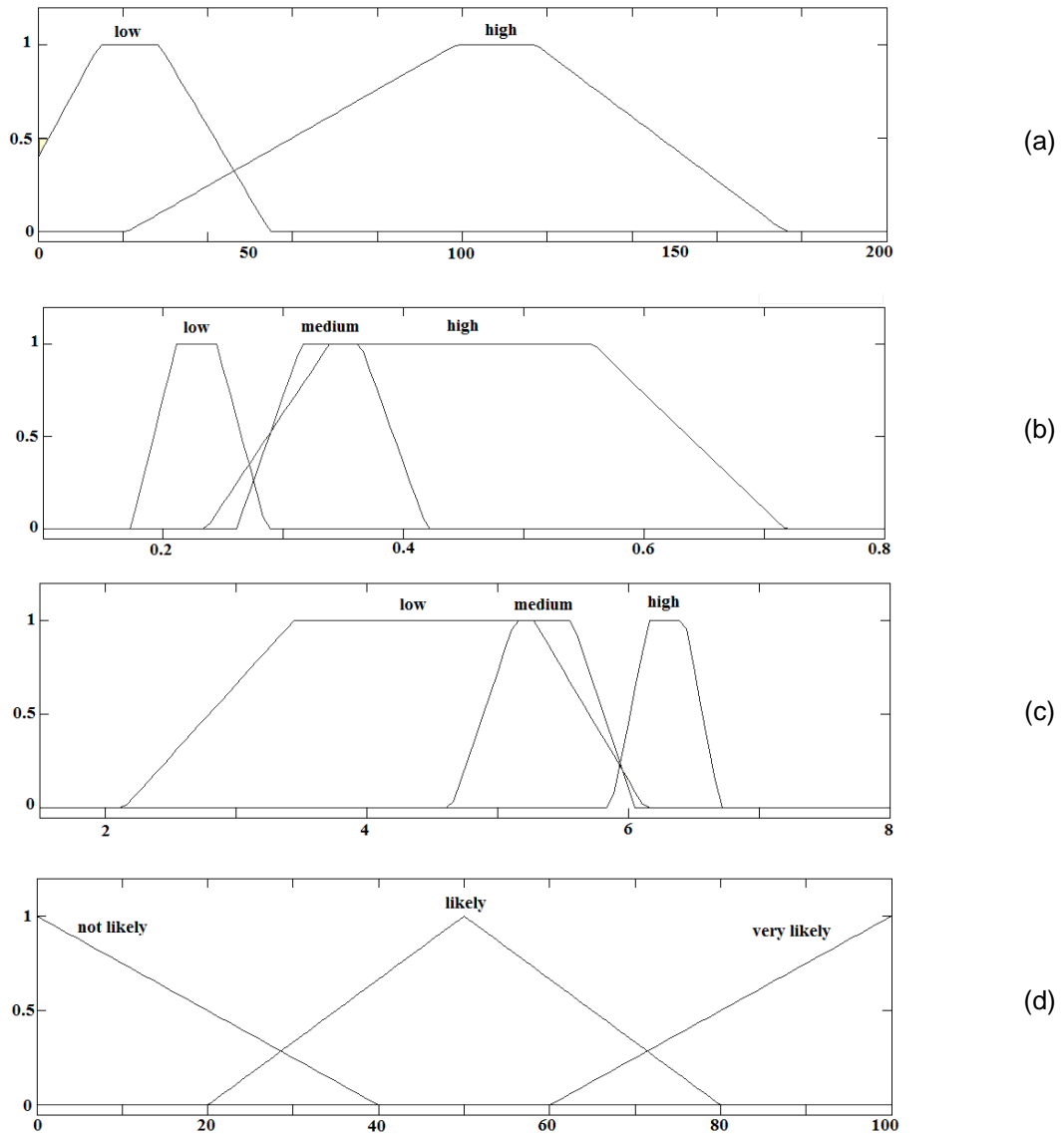
Figure 9 - Input and Output Linguistic Variables. From top to bottom: (a) contrast, (b) homogeneity, (c) entropy and (d) general output (city, forest, sea).

Rule	Conditional Statement					
1		Homogeneity is medium		Entropy is medium		Forest is very likely
2		Homogeneity is high		Entropy is low		Sea is very likely
3		Homogeneity is low		Entropy is high		City is very likely
4	**IF**	Contrast is low	**AND**	Entropy is medium	**THEN**	Forest is likely
5		Contrast is low		Homogeneity is medium		Forest is likely
6		Contrast is low		Entropy is low		Sea is likely
7		Contrast is low		Homogeneity is high		Sea is likely

Rule	Conditional Statement		
8	Contrast is high	Entropy is high	City is likely
9	Contrast is high	Homogeneity is low	City is likely
10	Contrast is high	Entropy is high	Forest is not likely
11	Contrast is high	Homogeneity is low	Forest is not likely
12	Contrast is high	Entropy is high	Sea is not likely
13	Contrast is high	Homogeneity is low	Sea is not likely
14	Contrast is low	Entropy is low	City is not likely
15	Contrast is low	Homogeneity is high	City is not likely

Table 2 - Complete Fuzzy Rule Set.

5.5. General Structure of the Classifier

The system classifies a whole image into of the three defined region types. The general scheme of the classifier is summarized in the flowchart of Fig. 10. The training flow is illustrated with dashed arrows and the application flow with solid arrows.
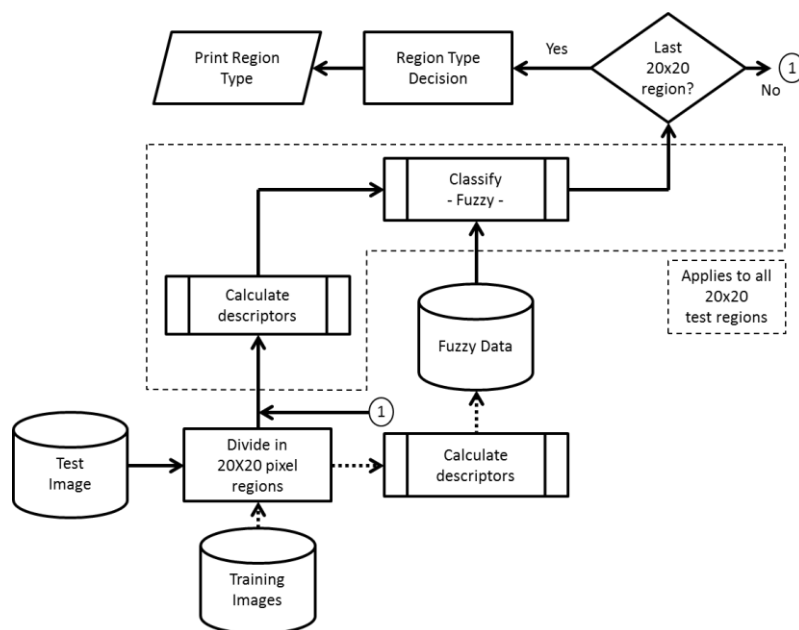
Figure 10 - Texture Analysis Scheme for Rotation-Invariant Region Classification.

The training flow refers to the steps needed to train and build the classifier. This section initiates with the extraction of 20x20 pixel blocks from the training images. Then, it continues with the calculation of the texture features for each training block and generation of the Fuzzy Inference System as presented in the previous subchapters.

The application flow refers to the steps needed to classify an input image. This section also initiates with the extraction of 20x20 pixel blocks from the input image. It continues with the calculation of the three RMS texture features for each block. The array with these three features is used as the input of the fuzzy classifier. The region type that yields the highest possibility is associated to each block. If two or more regions yield the same and highest possibility, the block is associated to a null class. The final classification of the test image follows a simple rule: the region type with the highest number of occurrences is elected as the land type of the test image.

6. ORTHOIMAGE CLASSIFIER: EXPERIMENTS AND RESULTS

This chapter aims to introduce the experimental approach to test the validity of the rotation-invariant descriptors and the performance of the rotation-invariant orthoimage texture classifier. The following subchapters detail the test database, performance metrics, test procedures and results. Two test situations are considered: the classification of complete images and individual image blocks.

6.1. Test Database and Test Procedure

The system described in Subchapter 5.4 was implemented using Matlab. Overall, 90 images representing city, forest and sea regions were gathered, being 30 images for each region-type with average size of 400x400 pixels. Fig. 11 shows some image examples for the three region-types. Several of them contain small secondary regions, as seen in Fig. 11(b), 11(d), 11(e) and 11(f).

(a) City I	(c) Forest I	(e) Sea I
(b) City II	(d) Forest II	(f) Sea II
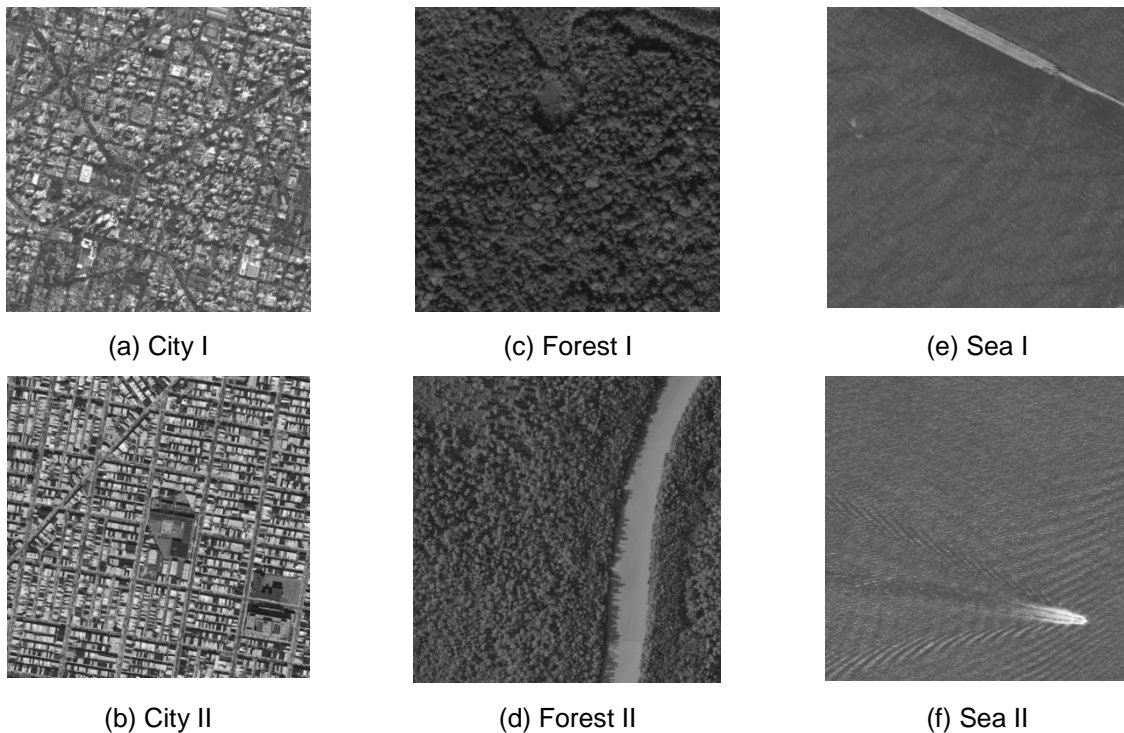
Figure 11 - Examples of Non-Rotated Test Orthoimages.

Six rotated versions were generated from the initial 90 images. The complete test database holds, accordingly, a total of 630 test images. The rotated images have an average size of 260x260 pixels; the images were extracted from the central area of the rotations in order to eliminate the diagonal borders created. This operation is illustrated in Fig. 12. For all the tests, the images are quantized from 256 to 128 grayscale levels.

(a) Original Image – No Rotation

(b) Counterclockwise Rotation. The extracted image is indicated by the white inner square.
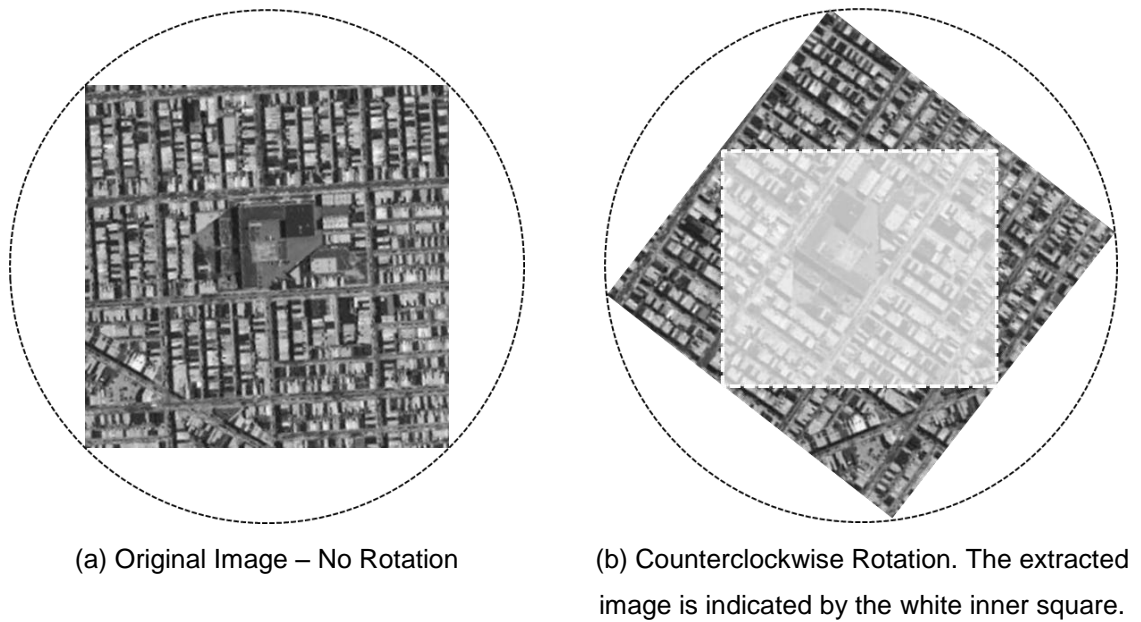
Figure 12 - Extraction of Rotated Test Images.

The test procedures serve the purpose of assessing the performance of the rotation-invariant region classifier and the influence of the rotation-invariant texture descriptors. To assess the performance of the region classifier, the correct classification rate is measured for two different cases. The first case concerns the correct classification of rate the test database images. The second case concerns the correct classification rate of individual 20x20 pixel blocks extracted from the test database images. Additionally, the influence of the rotation-invariant descriptors is measured by a performance comparison. The results for the correct classification rate of the proposed system (Fig. 12) are contrasted with a similar system constructed with the classical GLCM descriptors (Subchapter 3.1).

For the first case, the correct classification rate is defined as the number of images correctly classified divided by the total number of images in the test database. An image is said to be correctly classified if the output of the system of Fig. 12 matches the actual image type.

For the second case, the correct classification rate is defined as the number of blocks correctly classified divided by the total number of blocks contained in the test database. A block is said to be correctly classified if the output of the dashed section of Fig. 12 matches the actual block type.

6.2.   Complete Image Classification

This subchapter presents the classification results for the performance assessment of complete images. Initially, the original 90 non-rotated orthoimages were classified. For this case, no misclassifications were obtained, that is, the correct classification rate was 100%. The scheme was also tested for the remaining 540 rotated orthoimages. Again, there were no misclassifications. The total correct classification rate considering all 630 images was 100%.

To validate the effect of the rotation-invariant descriptors, the experiments were repeated using the classic horizontal grayscale co-occurrence matrix. The same 630 image database was used to assess the performance rate. The benchmark system uses the homogeneity, contrast and entropy descriptors calculated from the traditional GLCM. For the traditional descriptors, the GLCM was calculated with an offset $(\Delta x, \Delta y)$ of $(1,0)$, which corresponds to a comparison angle θ of 0°. A Fuzzy Inference System was generated following the same analysis of Chapter 5 and the classical GLCM descriptors. From the tests over the complete database, a correct classification rate of only 25.25% was obtained.

6.3. Classification of Individual Region Blocks

This subchapter presents the classification results for the performance assessment of individual 20x20 pixels region blocks. Considering all the test images, around 127000 blocks were classified. The classification rate was calculated for the blocks associated to the 90 non-rotated images and the 540 rotated images.

A correct block classification rate of 85.23% was obtained for the non-rotated images and a correct block classification rate of 85.24% was obtained for the rotated images. The total correct classification rate for all the test blocks was of 85.24%.

The classification results of individual blocks are illustrated below using some of the test images. To make the results visible, the masks depicted in Fig. 13 are superimposed over the blocks.

(a) Forest Mask   (b) Sea Mask   (c) City Mask   (d) Null Mask
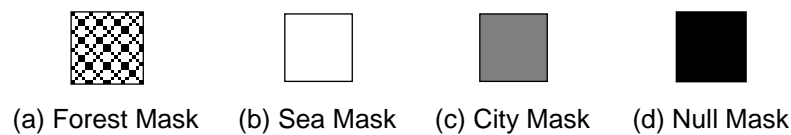
Figure 13 - Block Classification Masks.

Fig. 14 and Fig. 15 show the classification results for two images taken form the test database. Fig. 14(b) shows the block classification for a city-type test image. There was one null classification and six misclassifications. Note that the classifier detected a forest-region on top of the visible stadium. For this image, the correct classification rate considering the misclassified blocks is 98.15%.

Fig. 15(a) shows an image of a city region. There is a park and a lake to the left of the image; a river and an island stand to the right. According to Fig. 15(b), the main city region was correctly classified (mid-gray area). The lake was classified as sea (clear-gray area). Some areas of the park were classified as forest (grid area). The two city-like spots to the bottom right were classified as such (mid-gray area). Finally, the river at right had some correctly classified areas and a considerable portion was

not classified ("null mask"). For this image, the correct classification rate considering the misclassified and null blocks is 83.33%.

(a) Original Image   (b) Classified Image
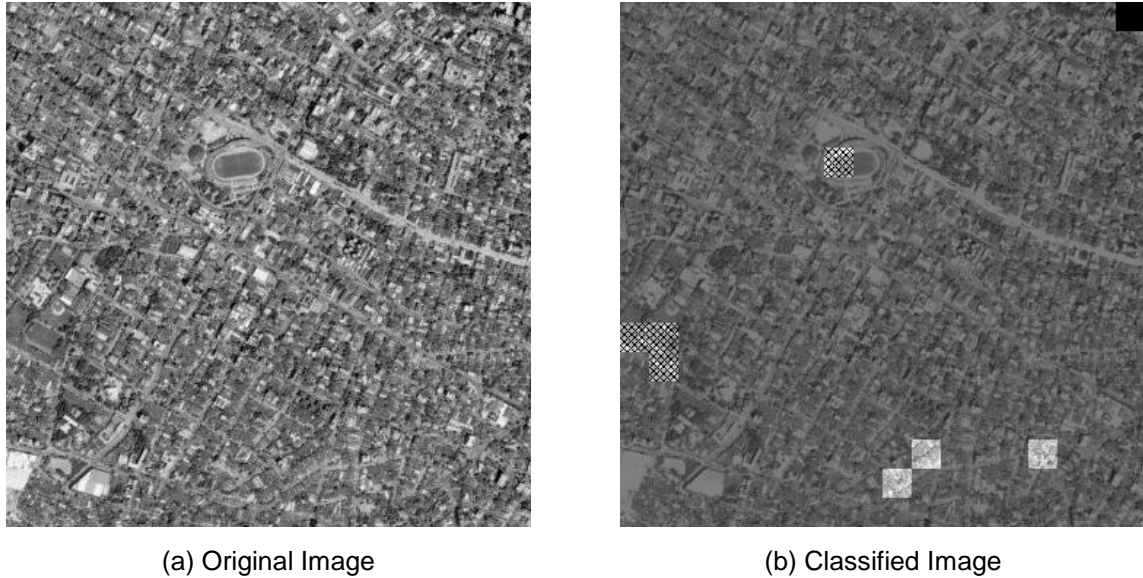
Figure 14 - Block Classification Results of a City Region

(a) Original Image   (b) Classified Image
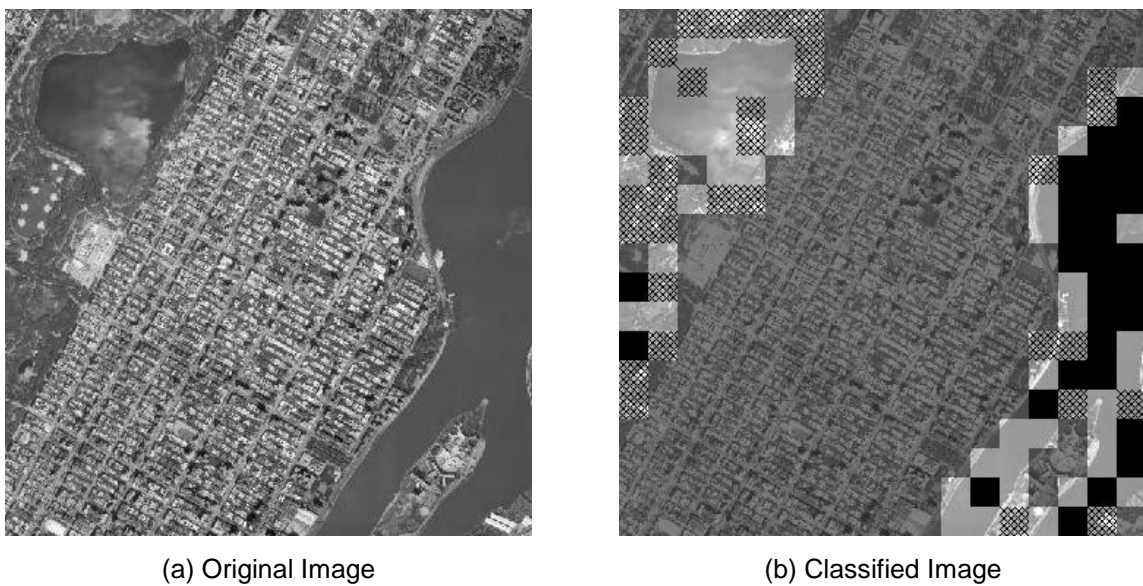
Figure 15 - Block Classification of a City Mixed Region

As a means of taking a qualitative discrimination assessment of the system, the block classification results for an image outside the test database are shown below. The image was selected to equally represent two of the considered region-types. Fig. 16(a) illustrates a coast city image with two regions (city and sea). In Fig. 16(b), the majority of the city portion was correctly classified (mid-gray area), while there were

some misclassifications. The sea region was correctly classified. The empty area in the upper left corner does not belong to any of the considered region types. The classifier associated it with sea and forest type regions. For this image, the correct classification rate considering the misclassified and null blocks is 88.77%

(a) Original Image   (b) Classified Image
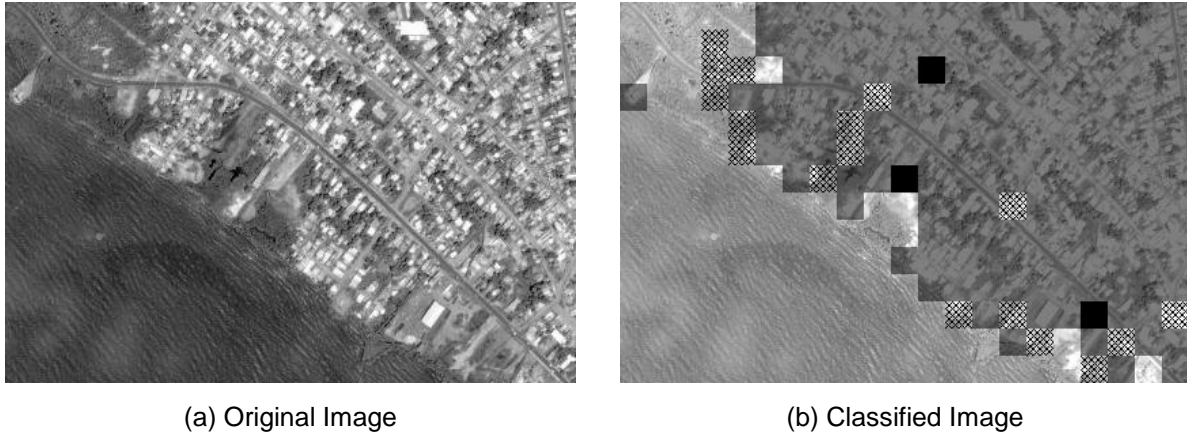
Figure 16 – Block Classification of a City and Sea Mixed Region

6.4.   Results Discussion

The experimental results indicate that the proposed rotation-invariant features describe robustly the texture information, even for rotated samples. Moreover, the inclusion of fuzzy logic provides a good platform for orthoimage classification. The complete image correct classification percentages for the rotation-invariant approach (100%) and classical approach (25.25%) clearly show that the descriptors greatly enhance the performance for classifying rotated images.

The rotation-invariant features offer another advantage: it is possible to use a smaller training database. This is evidenced by the fact that the proposal coped much better with the unseen textures than the system based on the classic co-occurrence matrix. The proposed approach maintained the block correct classification percentages of the non-rotated and rotated images nearly identical (around 85.2%). In order to achieve a similar performance, the classic system would need extra training data with all possible rotation angles.

The block classification results show that it the proposed system may serve as a basis for orthoimage segmentation in the three considered regions. The database images were obtained from two sources: Google Maps and Microsoft Bing Maps. As a consequence, the proposed system correctly classifies orthoimages derived from aerial picture sources captured by different devices. It is important to highlight that the results were achieved using only the texture information derived from a single grayscale channel.

The results presented on Chapter 6 serve as strong evidence that point out the strength of Ito et al.'s variant of the classical co-occurrence matrix. Even with the excellent results of the orthoimage classification experience, the current description of Ito et al.'s approach is not at its best. The remaining of this work focuses on a generalization of these descriptors and a comprehensive benchmark study with one of the most used texture descriptor method in the literature: the Local Binary Patterns approach.

7. RETHINKING THE CIRCULAR AND RADIAL TEXTURE DESCRIPTORS

The formulation proposed by Ito et al. has proven to be a very reasonable scheme for rotation-invariant texture description. The orthoimage application derived from the original descriptors certainly indicates that a more robust classification scheme, in this case using Fuzzy Inference Systems, can effectively bring an optimal performance for real applications.

Despite the positive points of the radial and circular descriptors, their mathematical definition can be improved to deliver more robustness. This section proposes a rethinking of the cited descriptors, conveying a generalization delineation that can enhance the texture discrimination capabilities.

One of the main issues present in the current definition of the descriptors that may directly influence the performance is that of the pixel information acquisition. The pixel information is gathered in a hard way, in other words, the selection for the circular and radial pixel compositions is binary – a yes or no situation - and results in an approximation for the pixel weights. Additionally, the data of the areas adjacent to the defined circles and radial lines are completely ignored. To fully understand these statements consider the situations portrayed in Fig. 17.

(a) Pixel Selection   (b) Adjacent Pixel Area
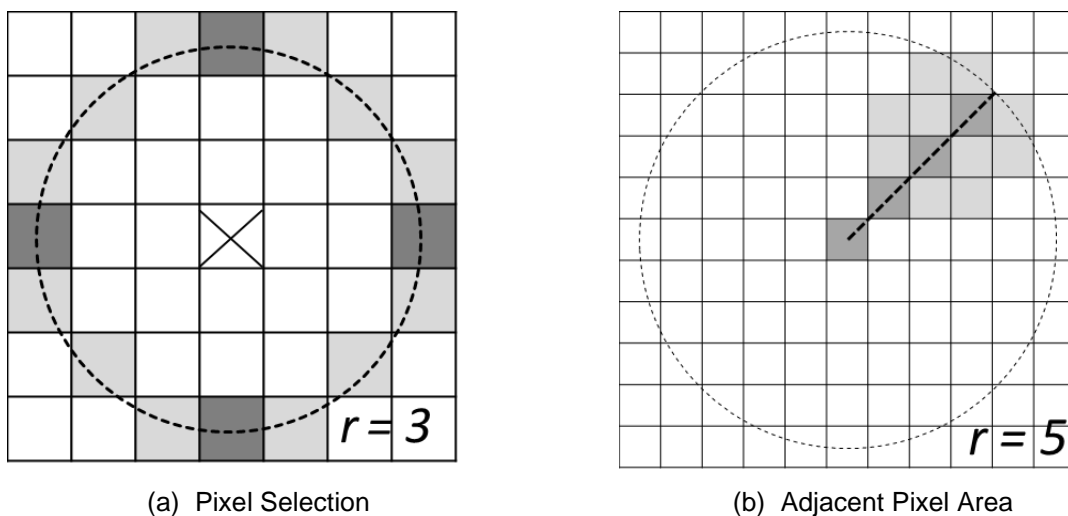
Figure 17 - Issues with the circular and radial geometry definitions

Fig. 17(a) illustrates the pixels that would belong to a circumference with a radius length of 3 pixels. In Ito's model, these pixels are awarded equal weights in the mean graylevel calculations. Then again, the model incurs in approximation; in a formal manner, if the pixels are geometrically represented by dots lying in the center of each square, the only four pixels that fully belong into the circumference are the ones highlighted in dark grey. The other twelve pixels in light grey do belong into the circle; they should be awarded different weights. This situation also applies for the radial descriptors.

Now take the single radial line of 45° illustrated in Fig. 17(b). Every dark grey pixel effectively belongs into the radial line. Maintaining the radial geometry, there are several light grey pixels around the line. These pixels contain information that might be useful to build the co-occurrence matrices but are not being considered. This situation also applies for the circular descriptors.

Fig. 17 demonstrates the aforementioned shortcoming. Textural and structural information is directly related to the spatial relationships between pixels in a given area. Consequently, addressing this issue would be a reasonable approach to improve the texture modeling of the current proposal.

The following subsections describe in detail the scheme centered on improving the circular and radial graylevel co-occurrence matrix texture descriptors.

7.1. Initial Considerations

7.1.1. Computational Cost

While the mathematical definition of the circular and radial descriptors is unquestionably the main component to be enhanced, the implementation of these modifications comes at a computational cost. The goal remains to improve the efficiency of the proposal and maintaining the calculations cost to the minimum.

To achieve this, the new circular and radial geometry definitions will be mapped into kernels and then applied to the desired images using a moving-window approach. Accordingly, the kernels will function as the input for a filtering operation over the image and the resulting images will be compared to generate the wanted graylevel co-occurrence matrices. In formal terms, a processed image I' is obtained from an original image I and a normalized kernel k according to the following expression.

$$I'(u,v) = \sum_{i,j} I(u-i, v-j) \cdot k(i,j) \qquad (7.1)$$

where u and v represent the indexes of the columns and rows of the input and output images, i and j represent the indexes of the rows and columns of the kernel.

In terms of the algorithm complexity, this approach brings desired benefits. Following the usual practice for spatial filtering algorithms, large enough kernels can be applied over the image using calculations in the frequency domain via the Fourier transform. Otherwise, the simpler calculations – multiplication and sums – will be employed. The Fourier option greatly diminishes the computing time for large kernels.

In this way, a simple tactic allows the geometrical definition for the descriptors to be calculated only once (kernel calculations) and be extensively applied over the required set of problems (2D filtering), saving additional computational costs. Furthermore, typical kernel configurations can be calculated beforehand and made available for the algorithm.

7.1.2. Smoothing the Pixel Information Acquisition

Following the main principles exposed in the introduction of Chapter 7, this subsection presents the basis for the circular and radial descriptor formulation.

There is one function that has a wide range of applications, alternating from models in statistics (probability distributions) to applications in image processing (blurring

filters). This well-known mathematical figure is the Gaussian function. In general terms, the Gaussian function can be described by the following expression:

$$g(x) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{7.2}$$

where μ represents the point of maximum amplitude of the curve and σ is the standard deviation, parameter associated to the width of the curve.

The Gaussian function serves as an excellent means to provide a smooth scale to a given parameter. The definition of the circular and radial kernels is obtained from an ingenious combination of the Gaussian function and the geometrical features of each descriptor type. The specific Gaussian to be incorporated holds specific values for and μ and σ, which are specifically defined for each descriptor. The following figure illustrates three fixed Gaussians with μ=0 and varying σ values.
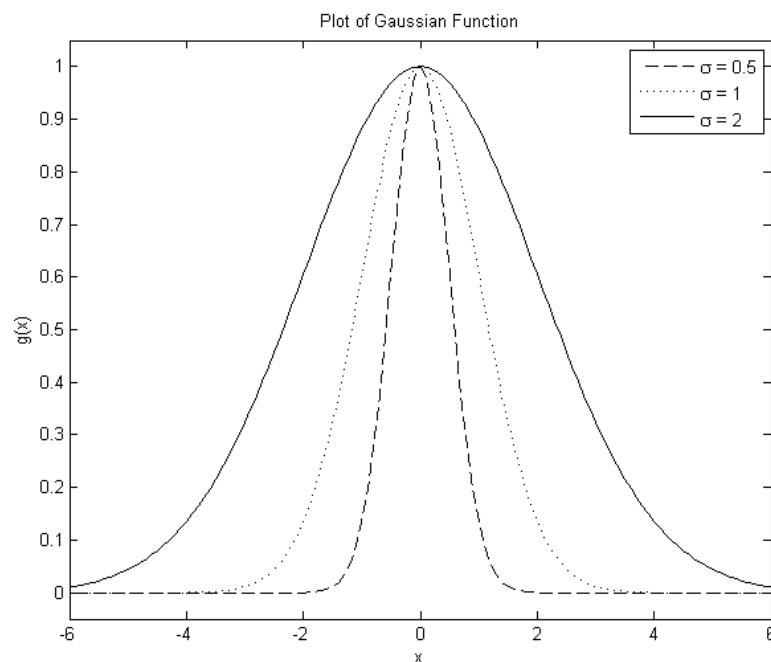
Figure 18 - Gaussian Function for different σ values.

Fig. 18 clearly illustrates the width control the σ parameter offers. Smaller values shrink the curve, while larger values stretch it. The slope is also influenced by this behavior, being steeper for low σ values and gently decaying for larger values. Another important characteristic of the Gaussian function is evidenced in the graph:

the Gaussian function falls to values near 1% of the maximum magnitude after a distance of ± 3σ measured from the origin. This indicates that the main slopes will always fall inside this range.

The following subsections will present the mathematical definition for the new circular and radial descriptors, based on the kernel and Gaussian approach here presented. The final subsection comments on the final considerations regarding the statistical formulae of the descriptors.

7.2. Circular Kernel and Circular Graylevel Co-occurrence Matrix

The circular descriptors model texture exploiting the relationship between the mean graylevels over two concentric circular structures. These circular structures are defined over two radii lengths, r_i and r_o for the inner and outer circles, respectively. The circular descriptors definition is associated, therefore, to a pair of kernels that represent the inner and outer circles structures. Take the circular geometry presented on Fig. 19.
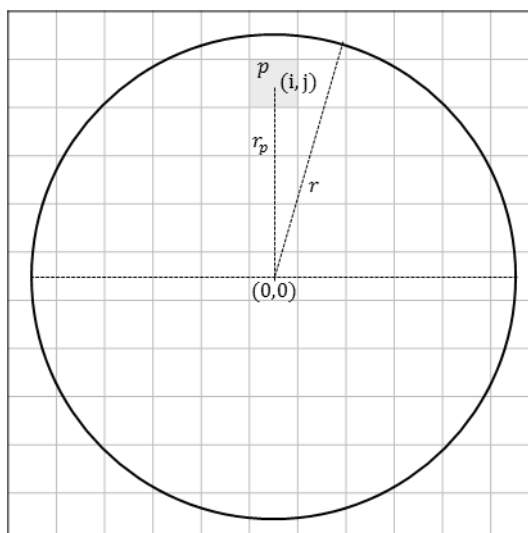
Figure 19 – A circumference over an 11x11 pixels area.

A generic circular kernel is defined considering the setup of Fig. 19. Note that the Cartesian coordinate system's origin is located over the center of the circumference.

This type of kernel, labeled $K_c^{r,\sigma}(i,j)$, is a square matrix with parameters r and σ, representing the radius of the circular structure and the smoothing parameter. The pair (i,j) represents any pixel p inside the kernel, where $i,j \in [-t, +t]$. A generic circular kernel is defined according to the following expressions:

$$t = round(r + 3\sigma) \quad and \quad l = 2t + 1 \tag{7.3}$$

$$K_c^{r,\sigma}(i,j) = e^{-\frac{\left(r - \sqrt{i^2 + j^2}\right)^2}{2\sigma^2}} \tag{7.4}$$

Eq. 7.3 defines the circular kernel size and Eq. 7.4 defines the contents of the kernel. The two inner and outer kernels are obtained by modifying the parameter r. Note that the $\sqrt{i^2 + j^2}$ of Eq. 7.4 refers to the distance from the origin to the considered pixel, in other words, the r_p value of Fig. 19.

Let's analyze both expressions separately. Eq. 7.3 simply states that the kernel is constructed with a size two times that of the radius and an additional pixel room given by a factor of σ. The relationship of the size and the radius is straightforward, but the relationship with the σ value merits clarification. Basically, the factor of 3 indicates that the kernel will hold information from the pixels located at 3σ pixels of distance measured from the circumference; this decision is linked to the fact that the Gaussian function holds significant information up to $\pm 3\sigma$. In practice, the 3σ factor is rounded to the next even integer after the $2t$ calculation.

The more complex expression of Eq. 7.4 is the conjunction of the Gaussian definition and the geometry of the circumference applied to the kernel construction. In an intuitive manner, this equation defines pixels as a value ranging from 1 to near-zero, where 1 indicates that the pixel fully belongs to the circumference. In this sense, the kernel can be understood as a pixel-wise membership function. The parameter σ controls the amount of pixels around the circumference that will be considered for the mean graylevel calculations. The next figure portrays a group of $K_c^{r,\sigma}$ kernels with different parameter combinations scaled to 256 graylevels.

(a) $K_c^{5,0.7}$ (b) $K_c^{5,1.0}$ (c) $K_c^{5,1.2}$

(d) $K_c^{10,0.7}$ (e) $K_c^{10,1.0}$ (f) $K_c^{10,1.2}$
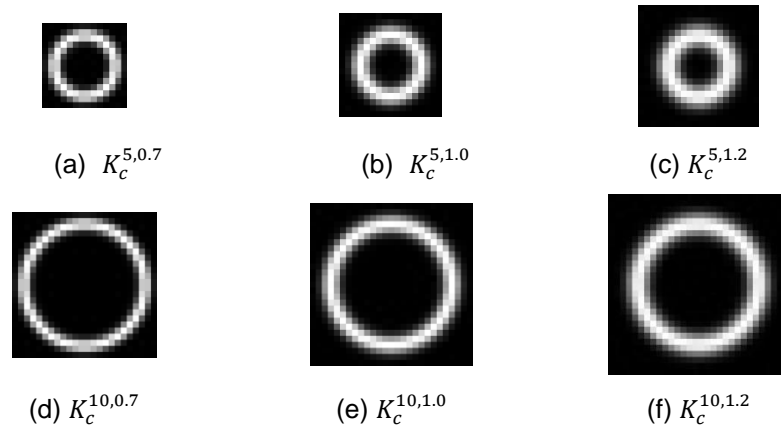
Figure 20 - Circular Kernels with radius of 5 and 10 pixels and σ values from 0.7 to 1.2 pixels.

The kernels of Fig. 20 illustrate practical applications of Eq. 7.3 and Eq. 7.4. The effects of the σ parameter are evident: for the same radius size, the kernel size increases with σ and is sufficient to hold all the relevant data. Also, higher σ values yield a wider pixel range around the circumference, indicating that a greater amount of structure information around the circumference is taken into account.

The circular kernel $K_c^{r,\sigma}$ must be normalized before applying it to the target images. The normalization indicates a simple division of every element of the kernel by the total sum of the elements. A normalized kernel is denoted by $k_c^{r,\sigma}$.

The circular graylevel co-occurrence matrix is defined from the pixel-wise comparison of the pair of images $I_{r_i}(u,v)$ and $I_{r_o}(u,v)$. These images are calculated from Eq. 7.1, which is applied to the target image $I(u,v)$ and a pair of normalized kernels $k_c^{r_i,\sigma}$ and $k_c^{r_o,\sigma}$. The restraint $r_i < r_o$ must be fulfilled, where r_i and r_o represent the inner and outer circumferences.

Given this, the circular graylevel co-occurrence matrix $P_c^{r_i,r_o,\sigma}$ is described by the following expression, a straightforward re-formulation of the original.

$$P_c^{r_i,r_o,\sigma}(i,j) = \sum_{p=1}^{u}\sum_{q=1}^{v}\begin{cases}1, & \text{if } I_{r_i}(p,q) = i \text{ and } I_{r_o}(p,q) = j \\ 0, & \text{otherwise}\end{cases} \qquad (7.5)$$

where u and v are the number of rows and columns of the images $I_{r_i}(u,v)$ and $I_{r_o}(u,v)$; also, $i,j \in 0 \ldots L_g - 1$, remembering that L_g represents the total graylevels of the input image.

7.3. Radial Kernel and Radial Graylevel Co-occurrence Matrix

The radial descriptors model texture exploiting the relationship between the mean graylevels over several radial line structures. These radial structures are defined over a radius r and a set of equidistant angles. The radial descriptors definition is associated to several kernels that represent each radial line structure.

The original descriptor definition presents the mean graylevel calculation as a weighted average of a set of line clusters. The base for the redefined radial descriptors no longer lies in a group of radii, but in a group of circular sectors. Take the diagram of Fig. 21, which represents one such sector and the geometrical parameters needed to describe the associated kernel.
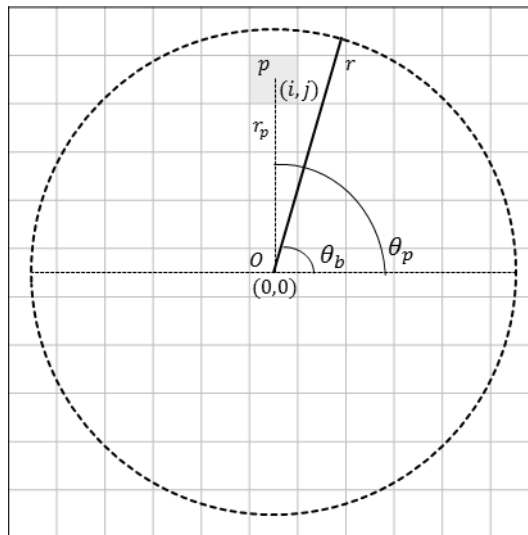
Figure 21 - A circular sector over an 11x11 pixel area.

A generic radial kernel is defined considering the geometry of Fig. 21. Note that the Cartesian system origin is located on the center of the kernel (point O).

This type of kernel, labeled $K_{rad}^{r,n,N,m}(i,j)$, is a square matrix parameterized by r, n and N and m. These parameters represent the radius of the circular sector, the index of the circular sector, the number of circular sectors and the quantity of standard deviations of Gaussian function that should be mapped into the circular sector with angle $2\pi/N$. The pair (i,j) represents any pixel inside the kernel, where $i,j \in [-r,+r]$. A generic radial kernel is defined according to the following set of expressions:

$$l = 2r + 1 \tag{7.6}$$

$$\sigma_r = \frac{\pi}{mN} \tag{7.7}$$

$$G_r = exp(-(atan2(j,i) \ominus (2\pi n/N))^2/2\sigma_r^2) \tag{7.8}$$

$$G_b = exp\left(-\left(r - \sqrt{i^2 + j^2}\right)^2/2\sigma_b^2\right) \tag{7.9}$$

$$K_{rad}^{r,n,N,m}(i,j) = \begin{cases} G_r & , if\ (i,j) \neq (0,0)\ and\ \sqrt{i^2 + j^2} \leq r \\ G_r * G_b & \sqrt{i^2 + j^2} > r \\ 0 & , otherwise \end{cases} \tag{7.10}$$

where l represents the square kernel size and r, n, m and N are the same as previously defined. The two Gaussian functions, G_r and G_b, refer to a radial smoothing function and a smoothing function around the border. The variable σ_r controls the effect of the G_r function, which is directly determined by the parameter m. The variable σ_b controls the effect of the G_b Gaussian and is defined as $0.2r$. The operator $\ominus$ introduced in Eq. 7.8 is angular subtraction and is defined as: $a \ominus b = MIN[mod(a - b, 2\pi), 2\pi - mod(a - b, 2\pi)]$. Finally, the function $atan2(y,x)$ calculates the arctangent considering the signs and magnitudes of the x,y points. Note that the standard deviation σ_r is measured in radians and the standard deviation σ_b is measured in pixels.

Eq. 7.6 states that the radial kernel size is directly defined by the radius length. On the other hand, Eq. 7.7 serves as intermediate calculation for Eq. 7.8, which defines the radial kernel inside the circle area. Finally, Eq. 7.9 represents the same expression introduced in the circular kernel definition. Eq. 7.8 and Eq. 7.10 justify further clarification.

Let's start with Eq. 7.8. This equation calculates the weight of a radial kernel pixel based on angular variables. Take the expression left to the $\ominus$ operator; the arctangent of the j, i pair is equivalent to the angle θ_p of the sample pixel p of Fig. 21. Now consider the expression right to the $\ominus$ operator; it is equivalent to the bisector angle θ_b of each circular sector (refer to Fig. 21). This simply means that the Gaussian function awards values to the pixels considering the angular distance from the pixel to the bisector radius r. Thus, the pixels directly over the bisector radius are given weights of 1, and the other pixels are awarded lower weights according to the σ parameter and their distance.

The σ_r parameter is calculated in Eq. 7.7. In this expression, the π/N fraction is exactly equivalent to $\theta_s/2$, that is, half of the central angle of the circular sector (see Fig. 21). The intuitive significance of the m parameter is better understood by highlighting a simple fact: the m parameter modifies the value of σ following an inverse relation.

For instance, suppose that $m = 1$, this would indicate that the pixels at the lateral borders of the circular sector (line segments OA and OB, see Fig. 21) are given weights approximately 0.6. Now, for $m = 2$, the weights for the same pixels would be approximately 0.14; for $m = 3$ they would be approximately 0.01; finally, for $m = 0.5$ they would approximately be 0.8. This means that m controls the area of the circular sector that is considered in the kernels; higher m values yield lower areas, and vice-versa.

Eq. 7.10 incorporates both Gaussian definitions into a simple expression. This expression outlines the contents of the radial kernel inside and outside the defined circle. The inner part of the kernel maps angular distance values and the outer part of

the kernel applies a smoothing operation over the mapping of those same angular distance values. Fig. 22 illustrates several examples of radial kernels. The effect of the parameter m is clearly evidenced in this figure.

(a) $K_{rad}^{20,0,6,2.0}$ (b) $K_{rad}^{20,1,6,2.5}$ (c) $K_{rad}^{20,2,6,3.0}$

(d) $K_{rad}^{20,3,6,3.5}$ (e) $K_{rad}^{20,4,6,4.0}$ (f) $K_{rad}^{20,5,6,4.5}$
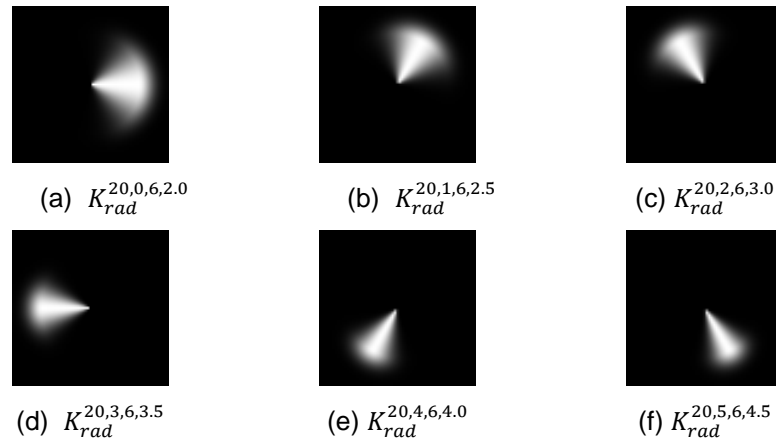
Figure 22 - Set of Radial Kernels with $r = 20$ pixels, $N = 6$ radial lines and m varying from 2.0 to 4.5.

The kernels of Fig. 22 clearly illustrate the results of applying Eq. 7.6 to Eq. 7.10. The kernels are a discrete geometric representation of Fig. 21 along with the Gaussian implementation discussed throughout this subchapter. The parameter m varied from 2.0 to 4.5 in order to illustrate the stated effect. Each column of Fig. 22 presents opposing circular sectors; the top row holds radial kernels with a lower m value than the opposing angle counterparts of the bottom row. It is evidenced that the increase of the number of m values decreases the weight of the pixels in the vicinity of the bisector radius. Furthermore, it is possible to confirm that the bisector radius holds the pixels with the highest weights. Fig. 23 presents a composition of a different set of radial kernels $\{K_{rad}^{20,0,6,4.0}, K_{rad}^{20,1,6,4.0}, ..., K_{rad}^{20,5,6,4.0}\}$ in one single image, illustrating the circular-based geometry derived from a complete set of kernels.
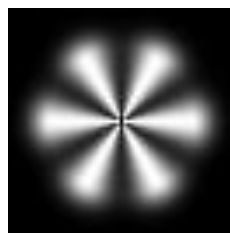
Figure 23 - Composition of 6 Radial Kernels with $r = 20$ pixels and $m = 4.0$.

The radial kernels $K_{rad}^{r,n,N,m}$ must be normalized before applying them to the target images. The normalization indicates a simple division of every element of each kernel by the total sum of its elements. A normalized radial kernel is denoted by $k_{rad}^{r,n,N,m}$.

The radial graylevel co-occurrence matrix is defined from the pixel-wise comparison of a set of images $\{I_0(u,v), I_1(u,v), \dots, I_n(u,v)\}$, where $n \in 0,1,\dots,N-1$. These images are calculated from Eq. 7.1, which is applied to the target image $I(u,v)$ and a set of normalized kernels $\{k_{rad}^{r,0,N,m}, k_{rad}^{r,1,N,m}, \dots, k_{rad}^{r,N-1,N,m}\}$.

Given this, a radial graylevel co-occurrence matrix $P_{rad}^{r,N,m}$ is described in the following expression, a straightforward re-formulation of the original.

$$P_{rad}^{r,N,m}(i,j) = \sum_{p=1}^{u}\sum_{q=1}^{v}\sum_{n=0}^{N-1}\begin{cases}1 & if\ I_n(u,v) = i\ and\ I_{n+1}(u,v) = j\ and\ n < N-1 \\ 1 & if\ I_{N-1}(u,v) = i\ and\ I_0(u,v) = j\ and\ n = N-1 \\ 0 & otherwise\end{cases} \quad (7.11)$$

where u and v are the number of rows and columns of the set of I_n images; also, $i,j \in 0\dots L_g - 1$, remembering that L_g represents the total graylevels of the input image.

7.4. Matrix normalization and Statistical Descriptor Considerations

The statistical descriptors defined for the original graylevel co-occurrence matrix are calculated from the newly defined matrices of Eq. 7.5 and Eq. 7.11. As in the original method, the circular and radial co-occurrence matrices must be normalized before applying the descriptor formulations. The co-occurrence matrix normalization is applied as in Eq. 3.11 purely by substituting $P(i,j)$ for $P_c^{r_i,r_o,\sigma}(i,j)$ and $P_{rad}^{r,N,m}(i,j)$ in the circular and radial case, respectively.

Henceforth, the redefined normalized circular co-occurrence matrix is denoted as $CCM_{r_i,r_o,\sigma}$, where CCM refers to "Circular Co-occurrence Matrix Method". For the

radial case, the redefined normalized notation is stated as $RCM_{r,N,m}$; likewise, RCM refers to "Radial Co-occurrence Matrix Method".

The set of Equations Eq. 3.2 to Eq. 3.7 are applied to the CCM and RCM matrices, with the exception of the contrast and correlation descriptors. For this work, related definitions are used for these two descriptors. The new definitions are similar in spirit but simpler in complexity. The complete descriptor definition is found on the comprehensive and peer-reviewed GLCM Tutorial maintained by Hall-Beyer (2007). The whole set of texture descriptors are presented in Eq. 7.12 to Eq. 7.17. The descriptors remain the same, following the results of Baraldi and Parmiggiani (1995).

$$f_{hom} = \sum_{i=0}^{L_g-1} \sum_{j=0}^{L_g-1} \frac{1}{1+(1-j)^2} p(i,j) \tag{7.12}$$

$$f_{con} = \sum_{i=0}^{L_g-1} \sum_{j=0}^{L_g-1} p(i,j) \cdot (i-j)^2 \tag{7.13}$$

$$f_{ent} = -\sum_{i=0}^{L_g-1} \sum_{j=0}^{L_g-1} p(i,j) \cdot \ln p(i,j) \tag{7.14}$$

$$f_{corr} = \sum_{i=0}^{L_g-1} \sum_{j=1}^{L_g-1} p(i,j) \cdot \frac{(i-\mu_x)(j-\mu_y)}{\sigma_x \sigma_y} \tag{7.15}$$

$$f_{ener} = \sum_{i=0}^{L_g-1} \sum_{j=0}^{L_g-1} p^2(i,j) \tag{7.16}$$

$$f_{var} = \sum_{i=0}^{L_g-1} \sum_{j=0}^{L_g-1} (i-\mu)^2 \cdot p(i,j) \tag{7.17}$$

Here, $p(i,j)$ represents a normalized co-occurrence matrix; μ_x and μ_y are the mean value of the marginal distributions; and σ_x and σ_y are their standard deviations.

8. MEASURING THE CCM AND RCM PERFORMANCE

Chapter 7 introduced a well-thought redefinition of Ito et al.'s variants of the co-occurrence matrix method. This chapter introduces a comprehensive test suite to validate the performance of the CCM and RCM methods individually and in conjunction. Accordingly, three main tests are introduced, two of them focused on benchmarking – with the Ito's variants and with the Local Binary Patterns method (OJALA et al., 2002) – and the last aimed at testing a simple luminance and contrast invariance proposition.

8.1. Benchmarking with Ito et al.'s approach

In their paper, Ito et al. (2009) propose a classification problem of 18 texture types. These images were rotated 6 times following a scheme identical to that of Fig. 12, obtaining 108 images for the test database. The original eighteen 512x512 pixels images were taken from the University of Southern California's (USC) texture database. The rotated versions have an area of 256x256 pixels. From the total 18 images, 12 are taken from the extensively-used Brodatz (1966) texture database. Fig. 24 portrays the 18 texture types labeled as they appear in the USC online database.

The main objective of this subchapter lays on establishing a straightforward comparison with Ito's approach. In other words, the goal is to measure in a tangible manner the effect of the new descriptor formulation, instead of analyzing in detail the properties of the CCM and RCM descriptors. In consequence, this section benchmarks the proposed approach with the broadest experiment description found in Ito's original paper.

A description of the experiment approach follows suite. The 256x256 pixels upper quadrants of the non-rotated images of Fig. 24 are used to train the classifier. The

training data consists of an array of 12 descriptors, the same ones introduced in Subchapter 3.1, per texture type.
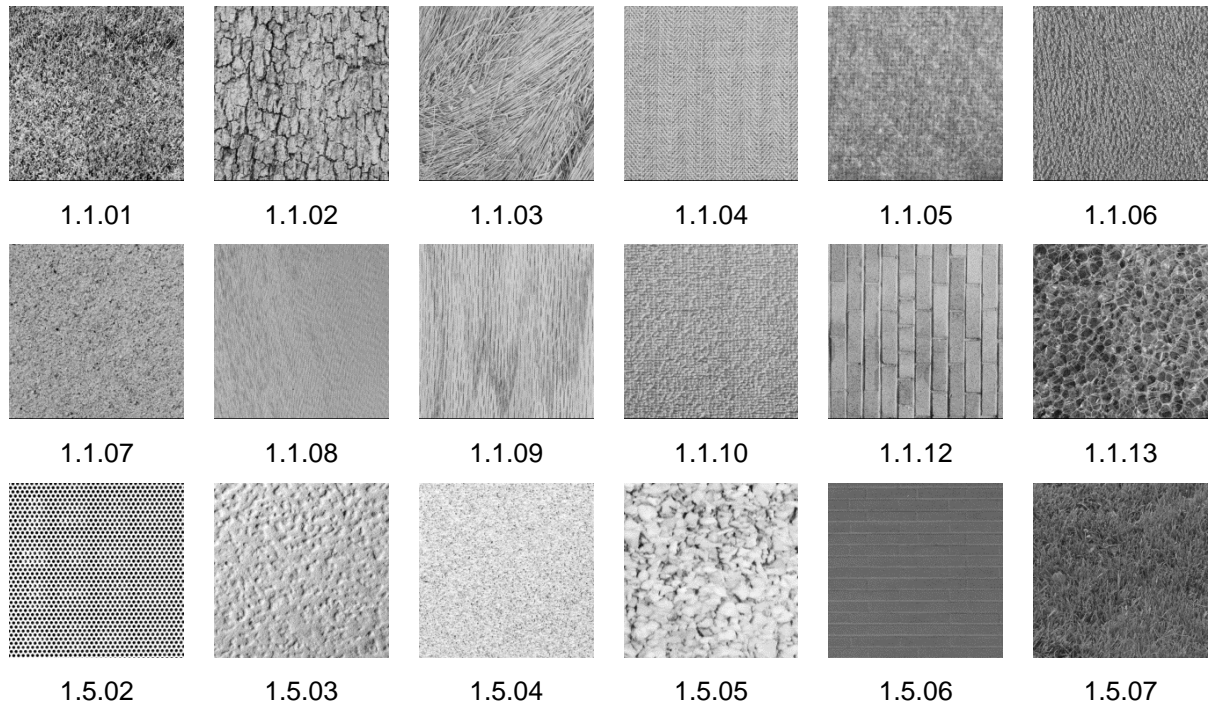
Figure 24 - Original Training Images from Ito et al.'s approach

This indicates that the training matrix holds 18 twelve-element arrays, bearing in mind the 18 texture types considered in the experiment. The test database is comprised of 108, 256x256 pixels rotated images generated from the 18 texture types. The employed classification scheme remains to be a simple Euclidean distance classifier, where the texture type belongs to the training array that yields the smaller distance when compared to the specific test array of analysis.

The experiment suite proposed by Ito et al. varies the configuration parameters of their circular and radial matrices in a controlled manner and closed value interval. The paper presents some combinations of the r_i and r_o circular matrix parameters, varying in the intervals [2,3] and [3,5], respectively. The radial matrix parameter r varies takes the pair of values 4 and 5; the radial matrix parameter N is fixed as 8 for all the experiments. In their paper, these parameter variations are combined to test the method. Another controlled parameter is the graylevel quantization of the co-

occurrence matrices. For the experiments, the matrices are calculated with several values: 16, 32, 64, 128 and 256.

This experiment setup was replicated using the CCM and RCM methods. The same radii parameters were used for both, the CCM and RCM, compared with Ito at al.'s circular and radial matrix approach. The only difference comes in the new parameters the reformulation introduces. For the CCM, the σ parameter is fixed as 1.0 for all the experiments. For the RCM, the m parameter is fixed as 1.4 and the N parameter is fixed as 7.

The adopted comparison lies in the mean classification rate for the complete set of experiments. The classification rate for a single experiment is defined as the total amount of correctly classified images divided by the total amount of test images.

Table 3 and Table 4 present the correct classification rate for each experiment combination. Table 3 deals with the results for Ito et al.'s proposal, and Table 4 deals with the results of the CCM and RCM proposal. Both tables display the variable descriptor parameters and their arrangements with the graylevel quantization that were used to obtain the shown classification rates.

Radial	Circular		Graylevels				
r	r_i	r_o	16	32	64	128	256
4	2	3	85.19%	90.74%	86.11%	86.11%	87.96%
4	2	4	93.52%	94.44%	95.37%	96.30%	95.37%
4	2	5	93.52%	96.30%	99.07%	99.07%	100%
4	3	4	93.52%	92.59%	86.11%	86.11%	86.11%
5	2	3	85.19%	89.81%	88.89%	87.96%	87.04%
5	2	4	94.44%	94.44%	97.22%	97.22%	96.30%
5	2	5	93.52%	95.37%	98.15%	99.07%	99.07%
5	3	4	93.52%	92.59%	89.81%	86.11%	87.04%

Table 3 - Correct Classification Rates for Ito et al.'s method

CCM	RCM		Graylevels				
r	r_i	r_o	16	32	64	128	256
4	2	3	100%	98.15%	100%	100%	100%
4	2	4	100%	100%	100%	100%	100%
4	2	5	100%	100%	100%	100%	100%
4	3	4	100%	100%	100%	100%	100%
5	2	3	100%	97.22%	100%	100%	100%
5	2	4	100%	100%	100%	100%	100%
5	2	5	100%	100%	100%	100%	100%
5	3	4	100%	100%	100%	100%	100%

Table 4 - Correct Classification Rates for the CCM and RCM.

For both cases, a total of forty individual experiments were completed. The mean correct classification rate for the data contained in Table 3 is 92.4%. On the other hand, the mean correct classification rate for the data contained in Table 4 is 99.9%.

Both numbers show the evident performance improvement that the CCM and RCM pair brings to the descriptor formulation. Keeping the same scale arrangements, the configuration with the reformulated matrices brings an evident benefit to this classification problem.

This experience clearly shows that the proposed reformulation has so far presented superb results for the texture description problems. Next subchapter deals with a far more demanding set and a complete comparison with a well-known method in the literature.

8.2. Benchmarking with the Local Binary Patterns approach

Thus far this work has introduced an up-front comparison with the method that served as a basis for the CCM and RCM approach. However, this subchapter presents a more comprehensive assessment of the performance and rotation-invariant characteristics of the proposed method. The Local Binary Patterns (LBP) method has been chosen as a reasonable contrasting method due to its popularity,

the fact that it is a relative modern method and its extremely wide range of applications and modifications. All these points were emphasized on the literature review of Subchapter 2.2.

In their original paper, Ojala et al. (2002) borrow a thorough evaluation scheme proposed by Porter and Canagarajah (1997) based on the Brodatz Texture Database. The Brodatz database has become a *de facto* standard for texture description algorithms due to its extensive use in literature; it has become a virtually obligatory database to benchmark new methods with older and more established ones. This database is taken from scanned copies of Brodatz original work (1966), a book originally aimed at artists. In their work, Porter and Canagarajah present three rotation-invariant proposals and experimentally find the best and most discriminative one: the wavelet-based descriptors (briefly introduced on Subchapter 2.1). Consequently, the selection of this particular classification problem leads to a direct comparison with yet another texture description method.

The remaining subsections present in great detail the problem description and derived experiment delineation.

8.2.1. Describing the Classification Problem

The classification scheme to be described in this subsection is specifically designed to evaluate the rotation-invariant characteristics of the texture descriptors used to solve it. The scheme database is the same as the one introduced by Porter and Canagarajah and used by Ojala et al. Even so, the training characteristics vary in these two papers.

The classification problem is comprised of 16 texture types and 10 classification sessions. As stated before, the 16 texture types were extracted from the Brodatz database. All the texture types are represented by 7 samples each; all the sample images have a pixel area of 180x180 and were extracted from larger source images.

Fig. 25 shows all 16 texture types with the same labels given by Brodatz himself in his book.

canvas	*cloth*	*cotton*	*Grass*
leather	*matting*	*paper*	*pigskin*
raffia	*rattan*	*reptile*	*sand*
straw	*weave*	*wood*	*wool*
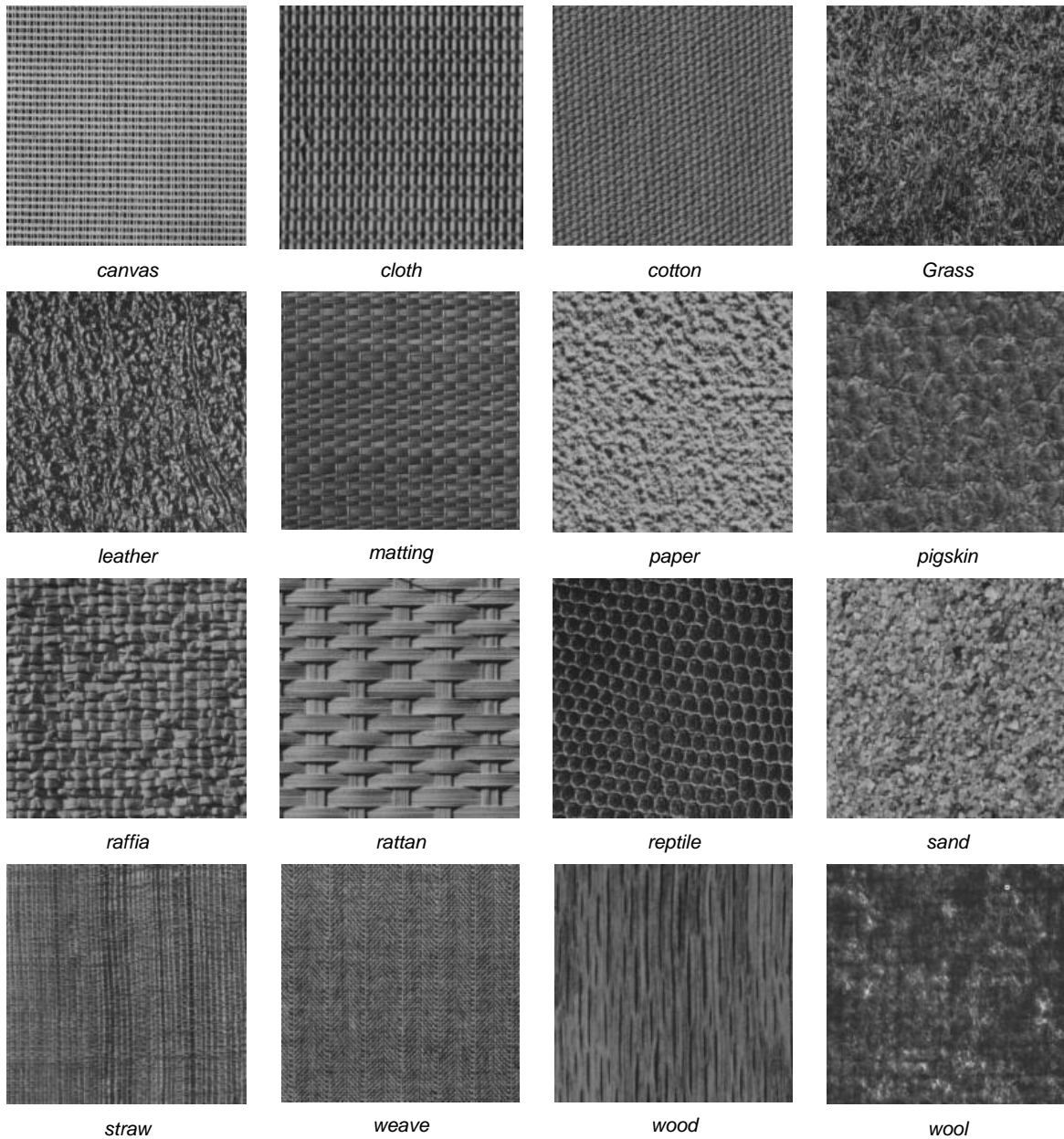
Figure 25 - Sixteen Texture Types for the LBP Benchmarking

In order to test the rotation-invariance of the descriptors, a set of rotated images was generated from the every sample of every texture type. Ten rotation angles are considered: 0°, 20°, 30°, 45°, 60°, 70°, 90°, 120°, 135° and 150°. The original samples correspond to the 0° angle. The images were taken from the Outex Texture Database (2013) and correspond to the exact same images used in Ojala et al.'s

paper. The 16 texture types, with 7 samples per type, rotated with images of 10 different angles yields a test database of 1120 images in total.

The classification task consists of 10 train-and-test rounds. Each round differs from one another in the training data used. For each training sequence, the classifier is trained with one sample of only a particular rotation angle. That said, the first training sequence is completed with a zero-degree sample of each texture type, the second with a twenty-degree sample of each texture type, and so forth, until the tenth and last sequence is completed with the last rotation angle samples. As a consequence, the 1120 test images are classified 10 times each, which accounts for a total of 11200 classification instances. Fig. 26 summarizes the complete classification test suite as a graphical representation.

10 TRAINING DATABASES

1	2	10
16 textures x 1 sample $\theta = 0°$	16 textures x 1 sample $\theta = 20°$	16 textures x 1 sample $\theta = 150°$

...

CLASSIFIER   CLASSIFIER   CLASSIFIER

CLASSIFY

CLASSIFICATION ACCURACY (%)

Mean of the 10 individal results

TEST DATABASE
16 textures x 7 samples x 10 rotation angles
1120 images
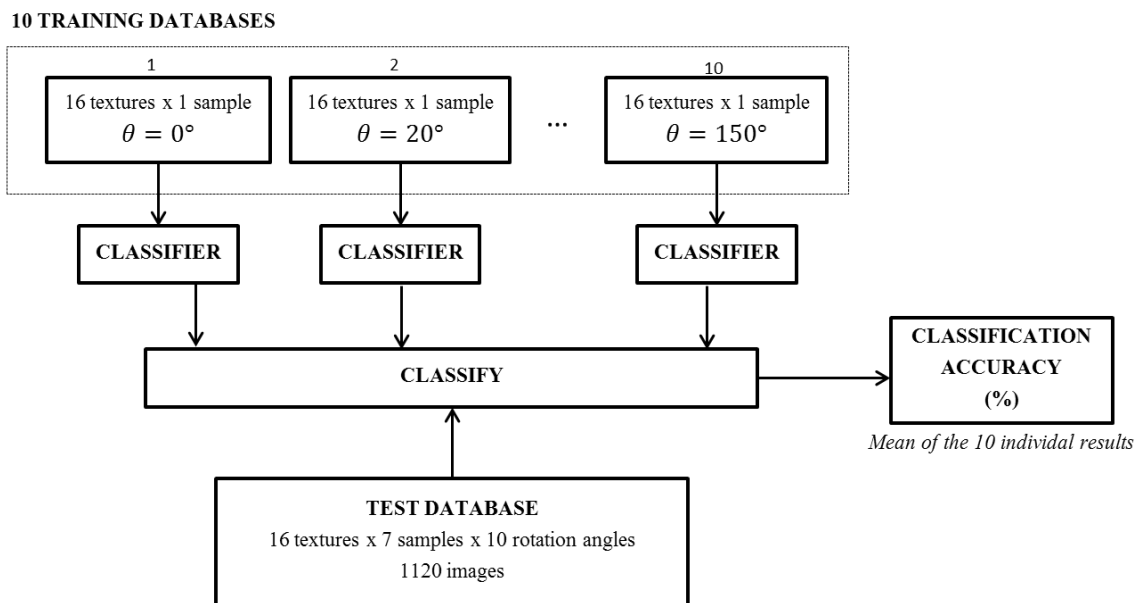
Figure 26 - Test Suite for the LBP benchmarking

This classification proposal measures the rotation-invariance of the descriptors while considering several important aspects: first of all, this suite considers inter-sample classification, that is, the analysis of non-seen texture images; secondly, the scheme builds up classifiers with every single available rotation angle, in other words, it tests whether the features are able to maintain their rotation-invariance regardless of the angle of the input (training) images; and last, it validates the robustness of the approach by training a classifier with only one sample per texture type, which is, very

little training data. Considering these facts, one can state that this is a well-rounded and sound test suite for rotation-invariance.

There is one small practical difference with the classification experiments implemented in this work and the ones of Ojala et al.'s paper that should be noted. The original test suite has an additional eighth sample used for training. Unfortunately, this eighth sample per texture image is delivered in the form of random chunks of 10x10 pixel images, that is, it is incomplete. After direct inquiry with the Outex database manager, it was confirmed that these images were lost. In order to maintain the same amount of test images (1120) per round, the first sample of the test database is used to train and the tests are undertaken using the complete original test database.

This inconvenient incorporates two fundamental alterations to the test suite properties. First of all, the inter-sample classification falls from 7 different samples to 6. Secondly, for every classification round, the 16 images used to train the classifier are guaranteed to be correctly classified. Even so, the main characteristics of the experiment are maintained: the fall of the inter-sample value is acceptable and including the 16 training images in the test suite would only incur in a mean classification difference of 0.0013%, which is irrelevant for the intended analysis.

8.2.2. Choosing a Similarity Measurement

In an IEEE paper centered on similarity metrics, Santini and Jain (1999) make a very interesting point that summarizes the motivations behind this subsection. They state that "measuring meaningful image similarity is a dichotomy that rests on two elements: finding a set of features which adequately encodes the characteristics that we intend to measure and endowing the feature space with a suitable metric". The CCM and RCM descriptors are intended to provide the wanted set of features to solve the introduced classification problem. Meanwhile, this section intends provide the feature space with an appropriate similarity measurement that manages to retain a simple description principle.

Aksoy and Haralick (2001) present an analysis of similarity measures applied to texture descriptors for image retrieval. In their work, they stress the importance of feature normalization prior to attempting classification. This consideration is entirely coherent with the CCM and RCM feature space, as some features are defined, for instance, in the closed $[0, 1]$ interval, while others are defined inside the $[0, \infty]$ interval. The multi-dimensional feature space of concern is, then, also multi-scaled.

In the cited work, Aksoy and Haralick introduce five different normalization approaches, where the goal is to make all the features fall into the closed interval $[0, 1]$. Afterwards, they define four different similarity metrics. They iterate some normalization-similarity combinations and apply them to an image retrieval problem. The main conclusion of their work lies in the fact that there is no optimal normalization-similarity combination and that an iterative and empirical process is required to choose the best pair. Nonetheless, they highlight the fact that the normalization schemes highly increase the performance of the testes systems. These considerations are the starting point to decide on a similarity measurement and normalization scheme.

In the Aksoy and Haralick approach, the main interest to take the features to a $[0, 1]$ range was to turn them into random variables. By achieving this, the authors could define probability-based classifiers such as a two-class Bayesian classifier, multivariate normal density functions or fitted probability distributions to build the decision components of their system. For this work, choosing a fairly simple Similarity Measurement it is deemed important; this design choice reflects the will to measure the performance of the CCM and RCM texture descriptors without giving too much weight on complexity of the measurement. All the probability-based schemes were discarded and a simpler distance-based measurement was chosen. This also indicates that normalization does not necessarily have to yield $[0, 1]$ range values.

This sequence of deliberations leads to the introduction of the chosen base similarity metric: the Minkowsky L_p metric. This metric is defined on Eq. 8.1.

$$L_p(x, y) = \left(\sum_{i=1}^{q} |x_i - y_i|^p \right)^{1/p} \tag{8.1}$$

where $p \geq 1$, $x, y \in \mathbb{R}^q$, q represents the total amount of features; x_i and y_i are the i-th components of the train and test feature vectors, respectively. The L_p metric is known by other denominations for certain p values; for instance, L_1 corresponds to the city-block distance and L_2 corresponds to the Euclidean distance. For this work, the election of p is purely empirical and is based on the value that yields the best results. For a conceptual-based selection of p, refer to Aksoy and Haralick's (2001) paper.

The Minkowsky metric is taken as the basis for the proposed Similarity Measurement. The modification to be introduced does not comply with the definition of a metric, thus, the expression is labeled as a similarity measurement instead of a similarity metric. The normalization scheme adopted for this work is based on a ratio of the test and training feature arrays. The normalization will guarantee that all features have the same relevance for the similarity calculations. The proposed normalization scheme consists on comparing the test to train features ratio with the unit. The idea underlying in the classification becomes much clearer through the analysis of Eq. 8.2 and Fig. 27.

$$\rho_p(x, y) = \left(\sum_{i=1}^{q} |1 - y_i/x_i|^p \right)^{1/p} \tag{8.2}$$
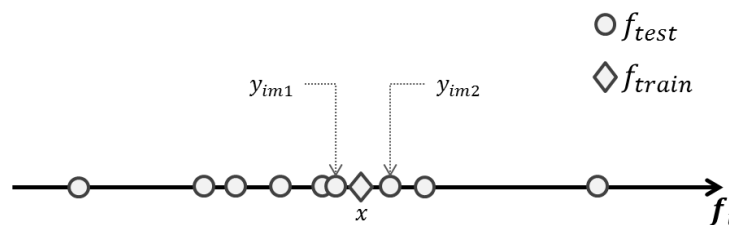
Figure 27 – Tran and Test Feature Space Example for any feature f_i.

The $1 - y_i/x_i$ expression simply indicates that the similarity of each test feature is measured as a function of its correspondence with the train feature of the same type.

Note that x_i refers to the train data and y_i refers to the test data point for the same feature type f_i. The training data point x_i is fixed as the denominator of the fraction, thus, it is possible to control division-by-zero cases.

Now take the situation of Fig. 27. The axis holds one train data point and several test data points for a given feature f_i. The train value is shown as the diamond-shaped marker and the test points are shown as the circle-shaped markers. Each data point is associated to a single and separated texture image. Suppose that $q = 1$ and $p = 2$; also, let's define the single training point as $x = 125$. Now, assume that the similarity of two texture images with $y_{im1} = 130$ and $y_{im2} = 120$ is measured using Eq. 8.2. For this case, both similarity measurements, $\rho_2(125,130)$ and $\rho_2(125,120)$, yield the same value, 0.0016. Note that both test values are at a distance of 5 units to the left and to the right of the train value. This indicates that the $\rho_p(x,y)$ measurement is symmetrical around the train features. That is, test feature values located to identical distances from the train feature yield the same similarity results; this is a very desirable characteristic. If the test and train values are identical, the similarity yields a 0 value, which agrees with perfect similarity for a distance-based proposal (zero distance).

The expression of Eq. 8.2 is used throughout the remaining of Chapter 8. The p value is set as 2, given that it produced the best results; that is, $\rho_2(x,y)$ is used for all the remaining experiments. Moreover, the q value varies with each experiment. It is set as 6 for the experiments concerning one type of invariant matrix (CCM or CRM). It is set as 12 for the experiments concerning both types of invariant matrix at the same time. It holds other values when indicated.

8.2.3. Classification Results – Local Binary Patterns

This section presents the results reported by Ojala et al. in their Local Binary Patterns (2002) paper. The results correspond to the classification problem presented in Section 8.2.1.

Ojala et al. present the classification results of a single-resolution and multiresolution approach of the LBP descriptor alone, the VAR descriptor alone and the joint histogram of the LBP and VAR descriptors used together. For the formal definition of the LBP and VAR descriptors, as well as the employed similarity metric, refer to Subchapter 2.2.

When the authors state that they perform multiresolution analysis, they refer to the calculation of the LPB or VAR operators with two or more radii values and their joint use to perform image classification. In this sense, the greater majority of the analysis performed in the Ojala et al.'s paper is multiresolution. Another important fact worth-highlighting are the reasons behind the VAR (contrast-based) operator. This operator is used for two main reasons, the first being to complement the main LBP operator due to its lack of contrast information and second, because the classification problem is not illumination-invariant (the VAR operator is grayscale dependent).

For all the tests, the $LBP_{P,R}^{riu2}$ and $VAR_{P,R}$ parameters P and R were varied considering the sets $P = \{8, 16, 24\}$ and $R = \{1, 2, 3\}$. Table 5 shows the complete experiment results discriminated by configuration type.

Table 5 depicts the classification results for the 1120 images of the test database considering each one of the training angles. The average classification rate for a specific configuration is shown on the rightmost column. It is easily found that the best classification rate for the proposed problem was 99.7%, an outstanding result.

OPERATOR	P,R	BINS	TRAINING ANGLE										AVERAGE
			0°	20°	30°	45°	60°	70°	90°	120°	135°	150°	
$LBP_{P,R}^{riu2}$	8,1	10	68.7	86.4	84.7	76.4	85.0	84.3	69.4	84.4	76.3	84.8	80.1
	16,2	18	96.2	99.0	98.6	98.9	98.5	99.1	97.6	98.6	98.7	97.5	98.3
	24,3	26	98.7	98.9	90.9	97.6	99.2	98.2	100	98.7	96.7	98.0	98.5
	8,1+16,2	10+18	94.3	99.5	99.8	99.8	98.5	97.2	92.9	99.6	99.2	99.2	98.0
	8,1+24,3	10+26	96.2	99.6	99.4	98.6	99.4	98.9	97.2	99.5	98.3	99.4	98.7
	16,2+24,3	18+26	97.7	100	99.8	99.2	99.3	100	99.6	99.4	98.5	98.4	99.2
	8,1+16,2+24,3	10+18+26	97.6	100	100	100	100	100	98.5	100	98.6	99.8	99.4
$VAR_{P,R}$	8,1	128	92.7	96.6	94.6	94.0	95.6	96.9	93.9	94.2	94.6	95.6	94.9
	16,2	128	89.9	84.5	86.2	90.5	87.3	85.6	91.0	89.8	90.8	88.5	88.4
	24,3	128	85.4	86.4	85.7	84.4	85.4	85.6	86.0	86.7	86.3	85.9	85.8
	8,1+16,2	128+128	97.5	96.9	98.8	99.0	97.9	97.7	97.5	99.1	98.8	97.9	98.,1
	8,1+24,3	128+128	95.2	97.0	98.7	98.9	97.5	98.5	96.1	99.5	99.0	97.9	97.8
	16,2+24,3	128+128	88.3	86.5	86.8	86.9	85.5	86.5	89.3	86.9	87.5	87.1	87.1
	8,1+16,2+24,3	128+128+128	94.9	94.6	97.0	98.3	96.2	96.2	95.0	98.2	98.1	97.3	96.6

OPERATOR	P,R	BINS	TRAINING ANGLE										AVERAGE
			0°	20°	30°	45°	60°	70°	90°	120°	135°	150°	
	8,1	10/16	99.1	94.2	95.7	97.3	95.2	94.4	99.3	96.0	97.3	95.6	96.4
	16,2	18/16	100	99.5	99.4	99.4	99.4	99.6	100	99.5	99.5	99.7	99.6
	24,3	26/16	95.8	95.0	96.2	97.4	96.0	95.5	95.6	97.2	97.9	98.9	96.5
$LBP_{P,R}^{riu2}/VAR_{P,R}$	8,1+16,2	10/16+18/16	100	99.3	99.1	99.2	99.3	99.2	100	99.3	99.3	99.4	99.4
	8,1+24,3	10/16+26/16	99.8	99.8	99.6	99.8	99.6	99.8	99.6	99.7	99.8	99.9	99.7
	16,2+24,3	18/16+26/16	97.2	98.9	98.9	99.8	99.6	99.9	97.3	99.6	99.8	99.9	99.1
	8,1+16,2+24,3	10/16+18/16+26/16	100	99.7	99.5	99.8	99.6	99.7	99.8	99.6	99.8	99.9	99.7

Table 5 - Correct Classification Rates for the LPB and VAR Operators

As it was stated before, the proposed test database was taken from the original work of Porter and Canagarajah. The wavelet descriptors, briefly presented on Subchapter 2.1, were chosen as the best classification performance scheme on their work. The classification principle was slightly different to the one presented for the LBP descriptors. The 0°, 30°, 45° and 60° angles were used to train the classifier, while the 20°, 70°, 90°, 120°, 135° and 150° angles were used to test the classifier. The best result reported by Porter and Canagarajah was 95.8% of correct classification rate for the wavelet-based descriptors.

The classification scheme found on Porter and Canagarajah's work is a simpler problem to the one used for the LBP operator. Due to the different classification principle, a direct comparison with their method is not considered in this work. Nonetheless, it is clear that the LBP's results were considerably better, given the quantitative rates and the more complex classification setup.

8.2.4. Classification Results – CCM Method

The first round of experiments for the reformulated matrices is completed with the CCM method. The expression of Eq. 8.2 is used to classify the images with a ratio-based comparison considering six test and train CCM features; this means that $q = 6$ for the similarity measurement. The train features are taken from the first sample of each texture type, considering the defined training angle. The features are taken from the expressions of Eq. 7.10 to Eq. 7.15. The CCM matrices for all the experiments are calculated considering the original 256 graylevels.

The tests are divided in two phases: discovering a good scale and refining the smoothing factor. This implies that the first round of tests for the CCM method relies on maneuvering between the (r_o, r_i) pairs to find an appropriate scale and the second round relies on finding the best σ for the optimal detected scale. The first round of tests fixes the σ parameter as 1.0 and the second round fixes the (r_o, r_i) as the best value found on the initial round.

Table 6 illustrates the results for the first round of CCM experiments. The (r_o, r_i) values correspond to all the accepted combinations for the sets $r_o \in \{3,4,5,6\}$ and $r_i \in \{2,3,4,5\}$, considering the $r_i < r_o$ restriction. The average classification results are shown on the rightmost column, following the same presentation structure as that of Table 5.

There is a motivation behind the choice of iterating around the CCM (r_o, r_i) pairs in advance. During the course of development of the descriptors, it was noticed that these scale-associated parameters had a major impact on the application results. This fact, indeed, can be explained by mere logic and sense: the proposed descriptors are a means of representing the gray-level variations among a circular area; thus, directly controlling the parameters that define the area where the graylevel occurrences are taken from, certainly will have an impact on the descriptor performance. This impact can be directly seen in the results of Table 6.

| OPERATOR | r_o, r_i | σ | TRAINING ANGLE | | | | | | | | | | AVERAGE |
			0°	20°	30°	45°	60°	70°	90°	120°	135°	150°	
	3, 2	1.0	98.6	99.6	99.7	99.4	98.8	97.7	98.9	99.7	99.4	98.8	99.0
	4 ,2	1.0	97.1	96.3	95.4	94.9	94.6	95.0	97.0	95.4	94.9	94.,6	95.5
	4, 3	1.0	89.0	86.5	88.8	89.8	89.1	90.3	88.7	88.8	89.8	89.1	89.0
	5, 2	1.0	86.9	86.9	81.3	80.4	78.4	84.6	86.7	81.3	80.4	78.4	82.5
$CCM_{r_i, r_o, \sigma}$	5, 3	1.0	88.8	85.5	84.9	84.8	82.9	84.5	89.0	84.9	84.9	82.9	85.3
	5, 4	1.0	89.5	88.1	87.6	86.3	86.2	87.1	89.1	87.6	86.3	86.2	87.4
	6, 2	1.0	83.6	82.1	79.0	77.6	78.1	82.6	83.8	79.0	77.6	78.1	80.1
	6, 3	1.0	81.0	77.6	75.8	76.4	79.4	79.9	81.6	75.8	76.4	79.4	78.3
	6, 4	1.0	80.7	84.8	83.6	85.2	86.4	86.0	81.3	83.6	85.2	86.4	84.3
	6, 5	1.0	86.4	83.0	81.5	86.0	86.3	83.2	86.0	81.5	86.0	86.3	84.6

Table 6 - Correct Classification Rates for the CCM Operator with fixed σ

Table 6 clearly shows that the CCM does offer a strong rotation-invariant approach for the proposed test suite. The best scale for the considered values lies on the $(3, 2)$ radii pair. As the external radius value increments, the accuracy of the classifier

diminishes, while some increments in the internal radius values offer some compensation. A greater discussion of the results of this section is considered on Chapter 9.

Having ascertained that the $(3, 2)$ radius pair offers the best classification results, it is taken as the starting point of the second round of experiments. Table 7 shows the classification results for a varying σ parameter.

OPERATOR	r_o, r_i	σ	TRAINING ANGLE										AVERAGE
			0°	20°	30°	45°	60°	70°	90°	120°	135°	150°	
	3, 2	0.5	94.3	95.7	95.7	93.6	90.1	91.6	94.2	95.6	93.7	90	93.4
	3, 2	0.6	95.9	97.1	98.1	96.8	96.3	94.6	96.1	97.9	97.1	96.0	96.6
	3, 2	0.7	96.6	98.4	98.7	98.2	97.1	96.1	96.6	98.8	98.5	97.2	97.6
	3, 2	0.8	97.3	99.2	99.5	99.2	98.4	96.8	97.8	99.4	99.2	98.3	98.5
	3, 2	0.9	97.8	99.8	99.8	99.6	98.8	97.2	98.3	99.8	99.6	98.8	99.0
$CCM_{r_i, r_o, \sigma}$	3, 2	1.0	98.6	99.6	99.7	99.4	98.8	97.7	98.9	99.7	99.4	98.8	99.0
	3, 2	1.1	99.3	98.9	99.4	99.1	97.8	97.7	99.5	99.4	99.1	97.9	98.8
	3, 2	1.2	98.9	97.4	98.3	97.7	97.1	97.1	99.3	98.3	97.7	97.0	97.9
	3, 2	1.3	97.4	94.6	96.6	96.3	96.5	96.9	97.2	96.6	96.4	96.5	96.5
	3, 2	1.4	96.7	93.4	94.8	94.0	94.7	95.4	96.1	94.7	94.3	94.7	94.9
	3, 2	1.5	94.6	91.6	93.0	92.2	93.7	94.6	94.2	92.9	92.5	93.7	93.3

Table 7 - Correct Classification Rates for the CCM Operator with fixed r_i and r_o

The σ parameter controls the smoothing level of the CCM operator; in other words, it controls the spread of the circumference in the neighboring pixels; the higher its value, the larger the considered pixel vicinity. This parameter, thus, is also associated to scale, but in a clearly lesser magnitude. As a result, the parameter may be regarded as a fine-tuning tool to improve the optimal scale's descriptor's performance. The mean accuracy of the CCM classifier reported on Table 7 confirms these remarks. The σ parameter is incremented considering 0.1 steps from 0.5 to 1.5. This small increment is chosen to fine-tune the descriptors. The $CCM_{2,3,\sigma}$ classifier configuration slowly increments its accuracy until it reaches its peak classification rate at $\sigma = 0.9$. After that point, it slowly decreases its classification rate.

This set of tests clearly indicates that, under a proper configuration, the CCM approach does offer a very good classification accuracy using only one scale definition. The best result of 99.0% is attributed to the $CCM_{2,3,0.9}$ operator using the full six-descriptor train and test arrays.

8.2.5. Classification Results – RCM Method

This section introduces the tests for the RCM method. The same similarity measurement, feature space and matrix quantization (256 levels) considered for the CCM classifier tests apply for the RCM. The tests are divided in the same phases: discovering a good scale and refining the smoothing parameter. This implies that the first round of tests for the RCM method relies on maneuvering between the (r, N) pairs to find an appropriate scale and the second round relies on finding the best m for the best detected scale. The first round of tests fixes the m parameter as 3.5 and the second round fixes the (r, N) pair as the best value found on the initial round.

Table 8 illustrates the results for the first round of RCM experiments. The (r, N) values correspond to all the combinations for the sets $r \in \{3,4,5,6\}$ and $N \in \{5,6,7,8,\}$. The average classification results are shown on the rightmost column.

| OPERATOR | r | N | m | TRAINING ANGLE | | | | | | | | | | AVERAGE |
				0°	20°	30°	45°	60°	70°	90°	120°	135°	150°	
	3	5	3.5	90.5	89.0	94.0	90.7	91.7	92.9	93.3	93.7	94.2	92.5	92.3
	3	6	3.5	73.2	82.9	83.9	85.3	81.3	79.3	79.3	76.2	75.5	74.2	79.1
	3	7	3.5	86.6	85.8	89.1	89.2	90.1	88.5	87.7	90.8	90.2	83.8	88.2
	3	8	3.5	83.6	87.9	91.6	89.7	88.7	85.2	84.8	88.8	84.0	86.1	87.0
	4	5	3.5	93.7	88.4	96.1	94.1	92.8	93.9	94.6	95	95.7	95.3	93.9
	4	6	3.5	75.2	77.2	75.9	78.3	76.0	75.7	70.7	75.3	76.3	74.8	75.5
	4	7	3.5	89.5	90.4	92.1	92.6	91.6	89.5	89.5	93.3	92.6	89.6	91.0
$RCM_{r,N,m}$	4	8	3.5	84.2	89.3	93.6	89.6	88.8	86.2	84.9	91.0	89.3	89.8	88.7
	5	5	3.5	92.6	92.0	93.8	92.1	90.4	93.4	94.3	89.1	95.0	95.3	92.8
	5	6	3.5	76.1	71.6	70.2	70.8	74.4	71.9	72.0	79.3	83.3	80.0	74.9
	5	7	3.5	92.9	92.3	94.4	96.3	95.0	92.1	91.7	94.9	95.1	94.6	93.9
	5	8	3.5	87.8	90.7	96.0	94.8	91.7	89.9	87.9	91.9	94.2	91.5	91.6
	6	5	3.5	92.8	92.0	91.9	91.0	94.6	95.8	93.0	91.5	92.9	95.9	93.1
	6	6	3.5	60.3	66.6	68.9	73.4	75.2	77.1	72.4	69.7	77.0	81.1	72.2
	6	7	3.5	90.1	90.8	91.6	91.8	91.9	91.1	92.8	91.6	92.8	92.7	91.7
	6	8	3.5	85.7	86.3	88.7	89.6	89.9	89.7	90.4	89.2	88.8	88.6	88.7

Table 8 - Correct Classification Rates for the RCM Operator with fixed m

Coherent with the remarks made for the CCM approach, the inspection of the (r, N) parameters is based on the selection of a proper scale. For the RCM case, the r parameter controls the circular area to be analyzed in the form of the length of the circular sectors. The N parameter defines the total area to be analyzed for each

circular sector; a higher N number leads to smaller circular sectors and vice-versa. Again, the impact of varying the scale-related parameters is evidenced on Table 8.

The reported accuracy for the RCM classifier slowly increases until it reaches its maximum at an (r, N) pair of $(6,5)$. From that point onward, increments in N show the general tendency to lower the classifier's accuracy. For a same N value, the variations in r do not affect the accuracy results in the same manner as the variations in N. This prevalence of the N parameter may be explained in the fact that a higher value for this parameter may introduce redundancy in the co-occurrence matrix; let us remind that the grayscale pairs are compared N times. A more detailed discussion of these results is introduced on Chapter 9.

Table 9 summarizes the results for the fine-tuning variations of the m parameter. Similar to the CCM tests, m is varied in 0.25 steps from 2.5 up to 4.5. The (r, N) is fixed at $(6,5)$, as shown in the last table results. The classification behavior is shown below.

OPERATOR	r	N	m	TRAINING ANGLE										AVERAGE
				0°	20°	30°	45°	60°	70°	90°	120°	135°	150°	
	6	5	2.50	88.7	88.8	89.3	88.9	90.1	89.4	88.0	87.5	88.6	91.3	89.0
	6	5	2.75	88.8	89.2	89.9	88.8	91.0	90.7	90.3	89.3	89.1	92.0	89.9
	6	5	3.00	89.6	90.0	90.4	89.2	92.5	92.5	91.4	90.3	90.8	92.9	91.0
	6	5	3.25	91.6	91.3	91.1	89.9	93.2	94.4	92.4	91.2	91.8	94.5	92.1
$RCM_{r,N,m}$	6	5	3.50	92.8	92.0	91.9	91.0	94.6	95.8	93.0	91.5	92.9	95.9	93.1
	6	5	3.75	93.1	92.4	92.6	92.5	95.5	96.6	93.4	91.5	93.9	96.5	93.8
	6	5	4.00	92.7	92.1	92.9	93.5	95.4	96.7	93.1	91.5	94.6	97.1	94.0
	6	5	4.25	91.8	91.3	92.4	93.2	95.0	95.6	92.4	91.4	94.7	97.3	93.5
	6	5	4.50	91.6	90.6	91.3	92.6	94.0	94.8	92.1	90.7	95.4	96.9	93.0

Table 9 - Correct Classification Rates for the RCM Operator with fixed r and N

The m parameter controls the smoothing level of the RCM operator; specifically, it controls the spread of considered pixels from the bisector line of the circular sectors; the higher its value, the lower the vicinity of the bisector radius that is considered. A similar behavior to that of the CCM case is expected for the fine-tuning of the RCM method. Table 9 confirms one more time these remarks. The $CCM_{6,5,m}$ configuration slowly increments its accuracy until it reaches its peak classification at $m = 4.0$. Further increments of the m value do not improve the mean classification rate.

While not as accurate as the CCM method, this set of tests clearly indicates that, under a proper configuration, the RCM approach does offer a very good classification accuracy using only one scale definition. The best result of 94.0% is attributed to the $RCM_{6,5,4.0}$ operator using the full six-descriptor train and test arrays.

8.3. Relevant Feature Selection and Joint CCM/RCM Analysis

So far, the classifier definition in this work has been presented under the supposition that the whole set of features provides relevant information to model the considered textures and solve the classification problem in hand. Nonetheless, this supposition may not necessarily hold true for all the features.

According to Julesz (1965), there is evidence that human perception of texture can be modeled using second-order statistics. Many researchers have focused on the study of second-order statistics to mode texture, including approaches such as the spatial-autocorrelation method, the covariogram method, the semi-variogram method and the central component of this work, the co-occurrence matrices (MAILLARD, 2003). These statistics-based methods do offer a catch, according to Materka and Strzelecki (1998), although methods based on second-order statistics have shown to achieve higher discrimination rates than some transform-based and structural methods, the texture in these images is discriminated if "they differ in second order moments".

Consequently, some of the second-order texture features extracted from the CCM and CRM matrices may be similar for two or more texture types. Including these characteristics may introduce redundant information into the classification scheme and deter the accuracy of the classifier. This section introduces a simple method to detect the least discriminant features and also presents the results of the joint use of the CCM and RCM methods.

8.3.1. The Problem of Irrelevant Features

The problem of selecting the most relevant features and discarding those that do not provide discriminant information has been extensively researched. In their literature review of such methods, Blum and Langley (1997) build upon what they call the "problem of irrelevant features". They indicate that at conceptual level, the two main learning tasks are to select the features to be used and to decide how to combine those features. In their paper, they review several methods in the literature to automatically discover and select these features.

Behind these methods, there lies a fundamental notion: the meaning of relevance. Blum and Langley indicate that the definitions for relevance are dependent of the application domain. They present five separate definitions, from which the fifth one is completely applicable to the problem in hand. The definition of relevance as a notion of "incremental usefulness", directly reproduced from their paper, indicates that:

> Given a sample of data S, a learning algorithm L, and a feature set $\mathcal{A}$, feature x_i is *incrementally useful to L with respect to $\mathcal{A}$* if the accuracy of the hypothesis that L produces using the feature set $\{x_i\} \cup \mathcal{A}$ is better than the accuracy achieved using just the feature set $\mathcal{A}$.

In other words, this definition indicates that a feature is relevant to the feature subset only if the learning algorithm that employs these features has a better accuracy than without that feature. In consequence, feature relevance is directly related to adding or removing features to the current feature subset.

This feature relevance definition is directly applied in an empirical manner to detect the features that offer a better accuracy to the classification problem. Although it escapes the scope of this work, several approaches to select feature subsets can be found in the literature, such as applications of Bayesian optimization (INZA et al., 1999), wrapper algorithms (KOHAVI and JOHN, 1997) and models based on support vector machines (MALDONADO; WEBER; BASAK, 2011). For general introductory reading on the subject, Blum and Langley's (1997) paper is recommended.

8.3.2. CCM and RCM Analysis – Selecting Relevant Features

As it was introduced on the last section, the feature relevance is treated as a problem of adding or removing features from the feature subspace. For the CCM and RCM texture descriptor operators, the total amount of defined features is 12. Such a low number of features allows for the definition of an empirical method that selects the relevant feature subset. The feature space $\mathcal{A}$ will contain the complete set of 12 features, 6 CCM and 6 CRM descriptors. Then, the thought-to-be irrelevant features are removed from the feature space and a final feature subset $\mathcal{A}'$ is defined to efficiently solve the classification problem.

The feature relevance analysis is executed from the feature data associated to the training angle 0°. The empirical iterative method is defined as follows: the classifier accuracy is calculated using all 12 descriptors. Then, the classification procedure is repeated removing one feature at a time. These results are examined to find the irrelevant descriptors. For simplicity, a simpler notation is employed for the forthcoming tests. This notation is introduced in Table 10. As the results of the relevance analysis will be directly related to removing certain features, please refer to the following table to pair a specific feature type with the statistical property it describes and the specific method it was computed from.

METHOD	CCM	CCM	CCM	CCM	CCM	CCM
FEATURE	Homog.	Contrast	Entropy	Correl.	Energy	Variance
NOTATION	f_0	f_1	f_2	f_3	f_4	f_5
METHOD	RCM	RCM	RCM	RCM	RCM	RCM
FEATURE	Homog.	Contrast	Entropy	Correl.	Energy	Variance
NOTATION	f_6	f_7	f_8	f_9	f_{10}	f_{11}

Table 10 - Feature Notation for the Significance Tests

The data for the tests is taken from the feature values of the $CCM_{2,3,0.9}$ and $RCM_{6,5,4.0}$ operators, as this configuration offered the best individual classification results. From now on, both descriptor types are used together to solve the problem. The interest to apply a relevance analysis is based on the hypothesis that some descriptors may be introducing redundant or irrelevant information to the classifier. This hypothesis is put

to a test with the forthcoming analysis comparison. The results are conveniently summarized on Table 11.

OPERATOR	REMOVED FEATURES	TRAINING ANGLE										AVERAGE
		0°	20°	30°	45°	60°	70°	90°	120°	135°	150°	
	none	99.0	99.7	100	100	100	99.6	98.7	98.0	100	100	99.5
	f_0	98.8	99.7	100	100	100	99.6	98.6	97.4	99.9	100	99.4
	f_1	96.7	97.2	97.9	97.6	97.8	97.7	95.9	94.7	97.2	97.9	97.1
	f_2	99.1	99.9	100	100	100	99.6	98.8	97.7	100	100	99.5
	f_3	99.1	99.9	100	100	100	99.6	98.8	97.6	100	100	99.5
$CCM_{2,3,0.9}$	f_4	98.8	99.5	99.6	100	99.6	99.3	98.5	97.3	99.6	99.6	99.2
and	f_5	99.3	99.9	100	100	100	100	99.0	97.9	100	100	99.6
$RCM_{6,5,4.0}$	f_6	100	100	100	100	100	99.6	100	100	100	100	100
	f_7	99.1	99.9	100	100	100	99.6	98.8	97.7	100	100	99.5
	f_8	99.0	99.9	100	100	100	99.6	98.8	97.7	100	100	99.5
	f_9	98.92	99.6	100	100	100	99.6	98.7	97.5	99.6	99.8	99.4
	f_{10}	98.8	99.2	99.6	99.9	99.6	99.3	98.1	96.3	99.0	99.6	98.9
	f_{11}	99.1	100	100	100	100	99.6	98.8	97.8	100	100	99.5

Table 11 - Feature Relevance Analysis

The results of Table 11 are analyzed in a simple manner. The mean accuracy obtained applying the complete set of features, 99.5%, is taken as a basis for comparison. The accuracy results associated to removing each single feature are compared to this basis, if the classification result improves, the incremental usefulness hypothesis is fulfilled and the descriptor is said to be irrelevant. Two descriptors fulfill this hypothesis, being f_5 and f_6. There is another set of features that do not fulfill the incremental usefulness hypothesis but do have an interesting characteristic: removing these descriptors neither deters nor improves the accuracy. These features are f_2, f_3, f_7, f_8 and f_{11}. The best classification result of Table 11 corresponds to the removal of f_6 and is measured as approximately 100.0%. Table 12 presents the results for removing two or more descriptors simultaneously.

OPERATOR	REMOVED FEATURES	TRAINING ANGLE										AVERAGE
		0°	20°	30°	45°	60°	70°	90°	120°	135°	150°	
	none	99.0	99.7	100	100	100	99.6	98.7	98.0	100	100	99.5
$CCM_{2,3,0.9}$	f_2,f_6	100	100	100	100	100	99.6	100	100	100	100	100
and	f_5,f_6	99.6	100	100	100	100	100	99.6	100	100	100	99.9
$RCM_{6,5,4.0}$	f_6,f_7	100	100	100	100	100	99.6	100	100	99.9	100	100
	f_6,f_8	100	100	100	100	100	99.6	100	100	100	100	100
	f_2,f_6,f_8	100	100	100	100	100	99.6	100	100	100	100	100

Table 12 - Joint Feature Relevance Analysis

The first row of Table 12 presents the results obtained from removing the two detected irrelevant features. The mean correct classification rate for this case is approximately 100%. The other rows present of the results obtained from removing the irrelevant features with some of the features that do not alter the results. Removing any of the combinations of features presented in Table 12 results in an slight improvement of the accuracy of the classifier. Removing the feature f_6 always leads to the biggest improvement, while removing any of the "status-quo" features f_2, f_7 and f_8 does not affect the improvements. This can be best evidenced in the example of the last row of Table 12, where removing the two irrelevant features f_6, alongside the two "status-quo" features f_2 and f_8, keeps the accuracy at almost 100%.

So far, Subchapter 8.2 and Subchapter 8.3 have introduced a robust classification problem, a coherent selection of a similarity metric, an iterative selection of an optimal scale, an iterative fine-tuning of the operators and a selection of the most relevant features to define a good configuration to solve the problem in hand. All the procedures are very simple in nature and easy to implement.

The best configuration for the rotation-invariant classification problem is given by the use of the $CCM_{2,3,0.9}$ and $RCM_{6,5,4.0}$ operators and the feature subset $\mathcal{A}' = \{f_0, f_1, f_3, f_4, f_5, f_7, f_9, f_{10}, f_{11}\}$, that is, nine different features, six CCM and four RCM. The accuracy of this configuration reaches an exact value of 99.96% (approximated to 100% in the tables), which improves upon the best results of the LBP method (99.7%) and also bests the results of the wavelet-based method (95.8%), without overlooking the different classification principle used by Porter and Canagarajah (1997).

8.4. An Alternative for Grayscale Invariance

Thus far, the development of Chapter 8 has shown that the proposed CCM and RCM operators offer considerably good rotation-invariant properties. Nonetheless, several methods in the literature also offer grayscale invariance. This subchapter presents one simple modification to the images that adds grayscale invariance to an already excellent rotation-invariant approach. As the proposed modification directly modifies the images, it was not introduced in Chapter 7. The modifications recommended in this chapter do not offer a definite descriptor approach for grayscale invariance, but, may perfectly serve application situations where grayscale invariance is a requirement.

Due to the nature of the CCM and RCM methods, it is possible to normalize the images before applying the kernels and calculating the co-occurrence matrices to eliminate any monotonic transformations of the grayscale information. The normalized image is quantized into the domain of the original. Eq. 8.3 shows the mathematical expression used to normalize the image.

$$I_{norm} = \frac{I - \mu_I}{\sigma_I} \qquad (8.3)$$

where I is the input image, μ_I is the mean over the image and σ_I is the standard deviation over the image.

The output images I_{norm} are represented by floating-point data. The images are quantized in order retrieve the values to the original or lower grayscale count. Fig. 28 shows an area-normalized histogram of the I_{norm} pixel values for the complete 1120 image set used to benchmark the CCM and RCM methods.
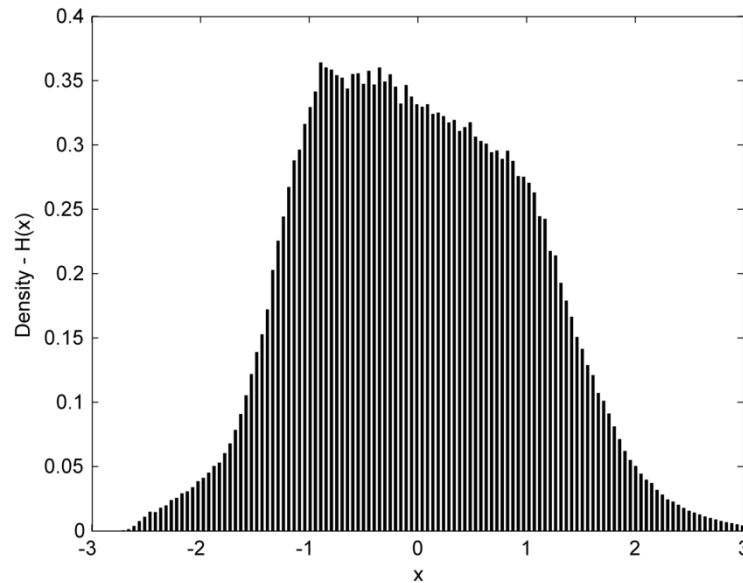
Figure 28 – Histogram of the Normalized Image Set

Fig. 28 clearly shows that the pixel values fall into the ± 3 interval. The shape of the histogram reminds of a Gaussian curve; the percentage of normalized pixels that are inside the ± 3 interval is 99.8%, for the ± 2.5 interval the percentage is 99.2% and for the ± 2 interval the percentage is 96.3%. The quantization procedure consists on a mapping from a defined interval into a number of bins. The interval space is uniformly separated in equally spaced regions defined by the ratio of the interval length to the number of bins; the interval spaces are numbered from 0 up to $bins - 1$. The mapping simply relates to associating every normalized value to the bin number of the interval space it falls on. The quantized images are used instead of the originals to perform the train and classification procedures. For the quantized images, the number of bins is treated as the number of graylevels for the co-occurrence matrix.

In order to test this grayscale and rotation-invariant approach, the original test database is redefined to contain brightness and contrast variations. The database is modified following the scheme presented in Table 13.

ANGLE	0°	20°	30°	45°	60°	70°	90°	120°	135°	150°
BRIGHTNESS (%)	0	5	15	-5	-15	0	0	0	0	15
CONTRAST (%)	0	0	0	0	0	5	15	-5	-15	15

Table 13 – Brightness and Contrast Modifications

The brightness and contrast variations of Table 13 are incorporated to each sample of each texture type of the image database, according to the rotation angle of the original image. The variations include a $\pm 5\%$ and $\pm 15\%$ alterations of the brightness and contrast level of the images, as well as a joint $+15\%$ variation of the brightness and contrast levels. Fig. 29 shows the image variations for the first sample of the first texture type (canvas texture).

The same experiments used for the LBP benchmarking method are implemented for the grayscale variation of the image database. The operators $CCM_{2,3,0.89}$ and $RCM_{6,5,1.22}$ are used in classification since the image scale has not changed. The relevance analysis introduced in Subchapter 8.3 is considered for the tests with the normalized and quantized images. The number of bins is set to 256 for all the experiments, taken from a normalized interval of ± 2.5.

| Original | +5% Brightness | +15% Brightness | -5% Brightness | -15% Brightness |

| +5% Contrast | +15% Contrast | -5% Contrast | -15%Contrast | +15% Bright./Contrast |
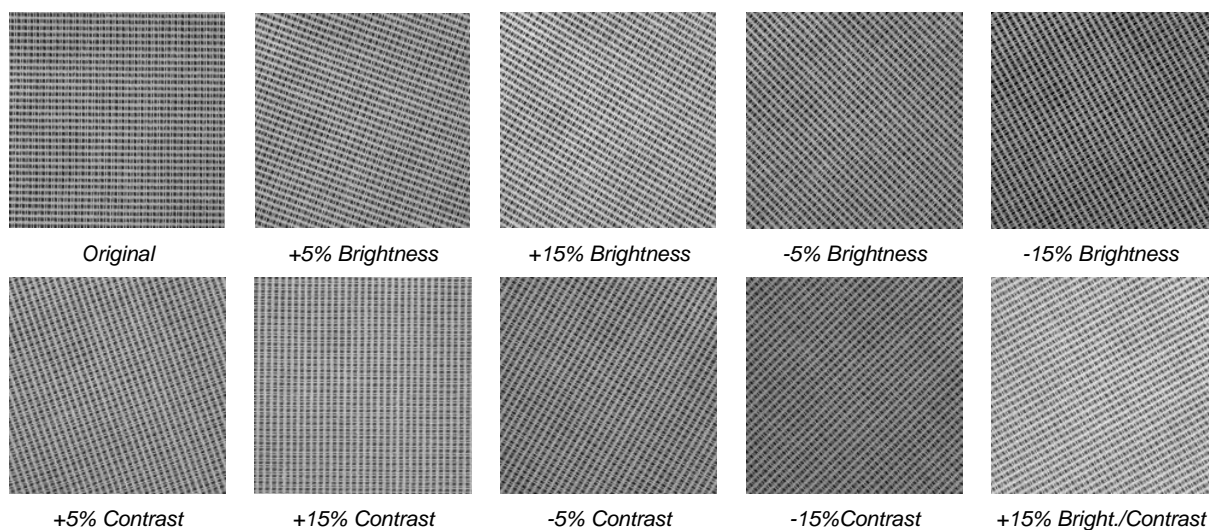
Figure 29 – Brightness and Contrast Variations for the Canvas Texture Type

Table 14 shows the results for the mean accuracy obtained for the twelve-features set and the results associated to the relevance analysis of individual features.

The best mean accuracy of the classifier, considering all the descriptors, is 82.5%, a considerable decrease in performance when compared to the original image database. The relevance analysis indicates that features f_5 and f_9 correspond to irrelevant features; the set of features f_2, f_6, f_7, f_8 and f_{11} correspond to already named "status-quo" features. Feature f_9 offers a very significant increase in the

classifier performance, consequently, the joint relevant analysis of Table 15 presents combinations of this feature with the other irrelevant feature and the "status-quo" features.

OPERATOR	REMOVED FEATURES	TRAINING ANGLE										AVERAGE
		0°	20°	30°	45°	60°	70°	90°	120°	135°	150°	
	none	92.1	94.4	93.8	92.2	91.6	92.7	92.3	92.1	92.6	91.1	92.3
	f_0	90.8	92.4	92.3	90.2	90.1	91.3	91.0	90.6	88.9	90.0	90.8
	f_1	87.1	90.5	90.4	89.6	87.9	88.9	90.5	88.8	86.9	86.2	88.7
	f_2	92.1	94.4	93.8	92.2	91.7	92.6	92.3	92.1	90.6	91.0	92.3
	f_3	91.1	92.7	92.4	90.9	90.4	91.9	91.5	90.9	89.4	90.4	91.2
$CCM_{2,3,0.9}$	f_4	91.0	93.8	93.8	91.8	91.1	92.1	92.1	91.9	90.5	91.0	91.9
and	f_5	91.8	93.0	94.8	93.2	92.8	93.8	94.0	92.5	91.7	93.6	93.1
$RCM_{6,5,4.0}$	f_6	92.1	93.9	93.4	91.3	91.3	92.1	92.3	91.3	90.2	90.9	91.9
	f_7	92.1	94.4	93.8	92.2	91.7	92.7	92.3	92.1	90.6	91.0	92.3
	f_8	92.1	94.4	93.8	92.3	91.7	92.6	92.3	92.1	90.6	91.0	92.3
	f_9	96.1	98.8	98.7	97.7	97.2	97.2	97.6	98.7	97.1	96.9	97.6
	f_{10}	91.1	93.8	93.6	91.3	90.0	90.5	92.6	92.1	89.4	89.9	91.4
	f_{11}	92.1	94.4	93.9	92.2	91.7	92.6	92.3	92.1	90.6	91.0	92.3

Table 14 – Feature Relevance Analysis for the Grayscale Invariant Variation

OPERATOR	REMOVED FEATURES	TRAINING ANGLE										AVERAGE
		0°	20°	30°	45°	60°	70°	90°	120°	135°	150°	
	none	92.1	94.4	93.8	92.2	91.6	92.7	92.3	92.1	92.6	91.1	92.3
	f_5, f_9	97.4	99.5	99.2	99.4	99.3	99.6	99.6	99.4	99.6	99.7	99.3
$CCM_{2,3,0.9}$	f_2, f_9	96.1	98.8	98.7	97.9	97.1	97.2	97.5	98.7	97.1	96.9	97.6
and	f_6, f_9	97.1	98.7	98.3	97.1	97.0	97.8	97.5	98.0	96.7	96.8	97.5
$RCM_{6,5,4.0}$	f_7, f_9	96.1	98.8	98.7	97.7	97.1	97.2	97.6	98.7	97.1	96.9	97.6
	f_8, f_9	96.1	98.7	98.7	97.8	97.2	97.2	97.6	98.7	97.1	96.9	97.6
	f_9, f_{11}	96.1	98.8	98.7	97.7	97.2	97.3	97.6	98.7	97.1	96.9	97.6

Table 15 – Joint feature Relevance Analysis for the Grayscale Invariant Variation

The results of Table 15 show a considerable improvement over the complete feature set classifier. The main contribution is obtained by removing the f_9 feature. Almost all combinations do not offer any improvement over the results obtained by removing f_9 alone. There is one combination that offers great results; he best classification rate is 99.3%, given by removing the two irrelevant features f_5 and f_9.

For the sake of comparison, Table 16 shows the mean accuracy results for the modified database using the original, non-normalized images to train and classify. The best feature combinations from Section 8.3.2 are used to test the classifier.

OPERATOR	REMOVED FEATURES	TRAINING ANGLE										AVERAGE
		0°	20°	30°	45°	60°	70°	90°	120°	135°	150°	
$CCM_{2,3,0.9}$	none	70.5	69.3	51.3	65.5	51.3	65.5	36.7	68.9	44.7	47.5	56.4
and	f_5, f_6	66.9	66.3	42.0	60.6	33.5	66.6	47.9	64.6	40.6	57.3	54,6
$RCM_{6,5,4.0}$	f_2, f_6, f_8	72.0	71.9	48.8	64.6	36.9	71.3	51.5	69.8	44.8	67.2	59.9

Table 16 – Feature Relevance Analysis for the Non-normalized Images

The results, as expected, indicate that using images with brightness or contrast variations compromises the performance of the CCM and RCM operators. The best mean accuracy level using the non-normalized images was of 59.9%, an evidently poor result when compared to the accuracy obtained by the simple inclusion of normalized images.

A closer inspection of the individual accuracy results per texture class indicates that some classes present more classification errors due to the inclusion of the normalized image scheme. After normalizing the images, some texture types present similar co-occurrence matrix representations, deterring the classification capabilities of the CCM and RCM operators. Even so, the alternative of image normalization and careful selection of the quantization intervals greatly solves the grayscale invariance problem. This approach may be enough to tackle practical applications with less texture types or different texture classes.

9. REVIEWING THE EXPERIMENT RESULTS

This dissertation work has presented numerous facts, applications, hypotheses, experiments and results centered on a rotation-invariant texture approach. Chapter 5 and 6 presented a successful application of Ito et al's version. Chapter 7 then presents a generalization for the descriptors that naturally leads to the CCM and RCM method introduction. Chapter 8 cohesively presents a methodical experiment setup that proves with certainty that the CCM and RCM methods improve upon Ito et al.'s approach and offer a better performance that the widely used Ojala et al.'s Local Binary Patterns Method. This Chapter presents a clear-cut discussion of the results obtained after the method proposals and elucidation of the experimental suits contained from Chapters 5 to 8.

9.1.   Findings for the Orthoimage Fuzzy Classifier Tests

The Orthoimage fuzzy classifier of Chapters 5 and 6 was initially conceived as a means of verifying whether the approach from Ito et. al could serve as a factual method to describe texture in a more compelling application scenario. The original test suite in their paper included a fairly high number of texture types – eighteen – but a fairly low number of texture samples per type – only six test images -.

Even though the orthoimage fuzzy classifier test suite only considered three region types, it considered a higher sample of 30 images per texture type. The most pronounced difference from the aerial-image case to the original Ito et al.'s experiment suite comes with the rich diversity of texture types provided by orthoimages taken from different sources and different geographical areas. The problem discussed achieves, then, a completely new and more difficult dimension.

The results of the proposed classifier, an accuracy of 100% for the 90 test images, show that Ito et al.'s approach serves as robust texture description method. Even so, the accuracy results of the classifier may be questioned. The fact is that the success

of the method not only comes from the texture description approach, but also from the texture fuzzy modeling. Fuzzy Logic comes as a very natural alternative to model a large amount of different texture types that are *somewhat* similar.

The tests from Chapters 5 and 6 offer several main findings, which can be summarized as: (i) Ito et al.'s method offers an excellent proposal for rotation-invariance; (ii) the aerial image (orthorectified or not) classification problem can be treated with as little information as only one channel (grayscale); (iii) Fuzzy Logic is an excellent tool to model classifiers that have to deal with texture classes that contain several sample types from different sources and contrast variations; (iv) a straightforward extension of the learning point (ii) can be stated as the very plausible improvement in performance for aerial image classification that comes from incorporating the infrared and spectral bands usually captured in these type of applications; and (v) it is worthy to pursue improvement efforts over the used rotation-invariant texture descriptors.

9.2. Findings for the CCM and RCM Benchmarking with Ito et al.'s Approach

Subchapter 8.1 presents the first benchmarking experiment for the CCM and RCM methods proposed in Chapter 7. Although simple in complexity, these tests offer a convincing and pivotal proof of the main hypothesis proposed in the introduction of Chapter 7, furthered elaborated on Section 7.1.2 and finally developed into the CCM and RCM methods. This hypothesis simply states that the original configuration presented by Ito et al. fails to completely describe the textural properties inside the circular and radial areas due to the hard selection of the pixel weights.

The benchmark consisted on replicating Ito et al.'s most comprehensive experiment, maintaining the scale (radii values) and adding the new parameters that represent the new formulation. The boost of the mean classifier accuracy from 92.4% to 99.9% clearly indicates that the new description greatly improves upon the original proposal.

The single and fundamental finding of the benchmark experiments is that the smoothing-based redefinition that gives birth to the CCM and RCM methods significantly enhances the texture description and discrimination capabilities of the original circular and radial co-occurrence matrices.

9.3. Findings for the CCM and RCM Benchmarking with the Local Binary Patterns Approach

The experiments used to benchmark the CCM and RCM operators with the Local Binary Patterns method is the most comprehensive test suite presented in this work. The various instances of the test suite allow extracting useful information about the CCM and RCM characteristics, as well as its performance quality against a widely used method in the literature. This subchapter is divided in individual analysis of the circular and radial descriptors, continuing with general findings for the whole setup.

9.3.1. Findings for the CCM Method

The CCM operator is parameterized by three variables: the inner and outer radii, r_i and r_o respectively, and the σ smoothing parameter. All three variables are measured in pixels. The experiments of Section 8.2.4 allow inferring the behavior of these parameters.

The first finding, evidenced in the behavior of the classifier accuracy, refers to the importance of the three parameters. Modifying the (r_i, r_o) pair holds a bigger influence in the classifier accuracy than modifying the σ parameter, as seen on Table 6 and Table 7. This means that variations of (r_i, r_o) alter the texture description in a significant manner. This effect is based on the association of each parameter to the description of the texture characteristics of the image.

Variations of the (r_i, r_o) values change the mean grayscale level measured over the circles. Notice that comparing pairs of circles that are much bigger than the texture details of the image leads to losing textural information because the texture details repeat all over the circular area. On the other hand, radii values smaller than the textural details of the image lead to high-entropy information, that is, the circles capture very different mean graylevel pairs and provide confusing texture descriptions. The σ parameter has a direct influence over pixels considered in the circular areas. Large σ values lead to including more pixels into the mean level calculations. Consequently, the σ value can serve as an excellent fine-tuning tool to get the best information details for a certain scale. Nonetheless, great care should be taken to avoid setting too large or too low σ values; the first case leads to an effect similar to that of large radii, and the latter leads to the same definition of the original circular descriptors.

Choosing the appropriate parameters for the CCM features leads to superb classification results, given that the best classification accuracy for the tests reached 99.0% with this operator alone. Nevertheless, this is not to say that the proposed CCM descriptor is the optimal one. One possible future modification could be defining three or more smoothed circular rings and consider two or more pairs of mean graylevel comparisons, instead of one.

9.3.2. Findings for the RCM Method

The RCM operator is also parameterized by three variables: the circular sector radius r, the number of circular sectors N and the m smoothing parameter. The first variable is measured in pixels and the other two are dimensionless. The N parameter is self-explanatory; the m parameter modifies the number of σ values that are included from the circular sector's bisector radius to any of the circumscribing radii. The experiments of Section 8.2.5 allow inferring the behavior of these parameters.

As in the case of the CCM method, variations of the r and N parameters have the most significant impact in the accuracy of the classifier. Analogously, these two

parameters define the scale of the operator, that is, they define the amount of texture detail that is considered in the mean graylevel comparisons. Too large values of r lead to losing textural detail and too small values lead to high-entropy information. Consecutively, large values of the N parameter lead to redundant textural information (too many circular sectors) and small values lead to losing textural details (too wide circular sectors). The m plays a similar role to the CCM's σ parameter; large m values tend to transform the circular sector into a single radial line, and small m values tend to give equal weights to the circular sector's pixels, modifying the low-pass characteristics of the descriptor.

As evidenced for the CCM operator, choosing the appropriate parameters for the RCM features leads to excellent classification results, given that the best classification accuracy for the tests reached 94.0% with this operator alone. In general, the RCM-based classifier yields a lower accuracy level than the CCM-based approach. The RCM matrix definition is not rotation-invariant *per se*, meaning that a pair of circular sectors does not possess an intrinsically circular geometry. The statistical features extracted from the matrix are rotation-invariant, but, the decrease of performance may be associated with the stated fact.

One possible future modification for the RCM descriptors could be defining mean grayscale level comparisons between non-adjacent circular sectors or between more than two circular sectors and filling this information into a single co-occurrence matrix.

9.3.3. Findings for the Complete Benchmarking and Further Improvements

Subchapters 8.3 and 8.4 provide the best found CCM and RCM joint setup to solve the classification problem and a direct comparison with the LBP approach. The joint CCM and RCM classifier is based on the principles of feature significance.

These subchapters provide very clear evidence that supports the following experimental findings: (i) the incremental significance principle for feature relevance

offers a very simple and straightforward approach to choose optimal feature subspaces; (ii) the joint application of the most relevant CCM and RCM features improves upon the individual operators application, demonstrating the complementary nature of both operators; (iii) the joint CCM/RCM classifier obtains better results than the LBP by employing only nine features derived from one scale per operator, the LBP classifier uses two scales for each one of the their two operators (LBP and VAR); (iv) the proposed alternative for adding grayscale invariance offers a reasonable option for the applications that require this characteristic; (v) the classification result almost reaches a perfect 100% score.

These findings show that the proposed rotation-invariant texture descriptors offer great textural description characteristics, alongside an easy-to-implement computational kernel-based approach. All the benchmarking experiments for the CCM and RCM operators were implemented in C++ and the Cekeikon Library (CEKEIKON), a proprietary library maintained by this work's supervisor, Prof. Hae Yong Kim, that incorporates the OpenCV digital image processing library (OPENCV). The algorithms developed for this work are meant purely to develop conceptual tests and are not in any way optimized. The mean processing time to calculate the twelve $CCM_{2,3,0.9}$ and $RCM_{6,5,4.0}$ features for all 1120 images was calculated; the value obtained was $38.3\ ms$. With this non-optimized algorithm, the calculation of, say, all twelve features for a 20fps video of 180x180 pixels is perfectly conceivable. Further optimization of the algorithm can lead to more demanding real-time applications. The algorithms were executed on an Intel Core i5 processor using a single core running at 2.53GHz.

As it was stated on Chapter 7, the CCM and RCM methods can produce co-occurrence matrices of a size given by the number of graylevels. All the tests for the benchmark experiments, with the exception of Ito et al.'s (2009) benchmarking, were done using the original set of graylevels, that is, a grayscale of 256 levels. For the LBP method benchmarking database, a reduction on the number of graylevels did not produce better results, and lower (64 levels and below) grayscales actually deterred the accuracy of the classifier. This effect simply signifies that the co-occurrence matrix derived from the CCM and RCM methods for this particular database was better represented with the pairs of mean graylevel occurrences over

the complete 256 set of possible values. Consequently, as these results did not bring significant improvements, they were left out of the previous chapters.

Now, considering future research, the two main steps that serve as possible directions include a grayscale invariance method embedded in the descriptor definition, and a scale invariance method. For the latter, good starting points include the concepts of Gaussian smoothing and pyramid decomposition.

The main motivation behind these concepts lies in the generation of several differently-scaled images from a source image. These concepts have been applied by Siquiera, Schwartz and Pedrini (2013) to propose a multi-scale co-occurrence matrix approach, and by Montoya-Zegarra et al. (2008) to propose a general approach to achieve scale invariance in any texture-description model. An additional scale-invariant approach could be to exploit the innate characteristics of the CCM and RCM operators; for instance, a "bad scale" could be detected by iteratively increasing the operator's radii until the mean graylevel of the circular rings (CCM) or the circular sectors (RCM) over the complete image present little variation. In other words, the texture details would be lost at that point. Brief experiences applying this last concept show that as the radii increases, the mean graylevel variation tends to decrease.

Finally, another possible future research principle would be to consider other groups of statistics, including other second-order statistics and third-order statistics, to describe the information found in the CCM and RCM matrices.

10.   FINAL DISCUSSION AND CONCLUSIONS

This dissertation paper revolves around several principles of texture description. Texture is a distinctive characteristic of objects that is directly reproduced on the images that capture their information. Therefore, texture may be relevant for virtually any image in multiple knowledge fields. Considering this, rotation-invariance comes as a very desirable property for texture descriptor models, as the capture of images are usually obtained under limited controlled environments.

These specifics are the main motivation behind the continuous research on texture features in the literature and one of the fundamental motivations behind this work. This paper retrieves the spotlight and directs it towards a variation of Haralick et al.'s classical co-occurrence matrix method, obtaining excellent results comparable with other widely used methods.

This work solidly grounds the CCM and RCM operators as functional, practical and easy-to-implement texture descriptors that offer excellent rotation-invariant robustness, as well as a conceptually simple extension to grayscale invariance that manages to maintain very good results. The core improvement hypothesis that leads to the definition of the CCM and RCM operators was proven to be correct.

The early version of the proposed descriptors was successfully implemented as an orthoimage region classifier. The rotation-invariance properties of the descriptors allow obtaining excellent classification results using less train data and diversely rotated images. By straightforward extension, it can be stated that the CCM and RCM offer an excellent base to build applications in the aerial images domain.

This work attempts to institute the CCM an RCM operators as desirable texture descriptors for a wide arrange of practical applications, with the possibility of being used in union with other methods in the literature. In their current definition, the recommended course of action to obtain an optimal configuration is to adjust the scale-related parameters, then adjusting the smoothing parameters and finally performing a simple relevance test, as indicated on Chapter 8.

10.1. Publications

The contents of Chapter 5 and Chapter 6 were published in the Fourth International Conference on Geographic Object-based Image Analysis (GEOBIA 2012), which was held in Rio de Janeiro, Brazil, from May 7th to May 9th of 2012. The article is titled "Texture-based Fuzzy Inference System for Rotation-Invariant Classification of Aerial Orthoimage Regions". The paper was accepted for an oral presentation.

REFERENCES

AHONEN, T.; PIETIKÄINEN, M. Face Description with Local Binary Patterns: Application to Face Recognition. **IEEE Trans. Pattern Anal. Mach. Intell**, v. 28, n. 12, p. 2037-2041, 2006.

AKSOY, S.; HARALICK, R. Feature normalization and likelihood-bases similarity measures for image retrieval**. Pattern Recognition Letters**, v. 22, p. 563-582, 2001.

ANDERSON, J.; HARDY, E.; ROACH, J.; WITMER, R. A Land Use and Land Cover Classification System for Use with Remote Sensor Data. **US Geological Survey Professional Paper**, v. 964, 1976.

BALA, R.; BRAUN, K.M. Color-to-grayscale conversion to maintain discriminability. In: SPIE-IS&T Electronic Imaging. 2004, San Jose, California. SPIE Proceedings Vol. 5293, p. 196-202.

BARALDI, A.; PARMIGGIANI, F. An Investigation of the Textural Characteristics Associated with Gray Level Co-occurrence Matrix Statistical Parameters. **IEEE T. Geosci. Remote**, v. 33, n. 2, p. 293-304, 1995.

BINAGHI, E.; MADELLA, P.; MONTESANO, M.; RAMPINI, A. Fuzzy Contextual Classification of Multisource Remote Sensing Images. **IEEE T. Geosci. Remote**, v. 35, n. 2, p. 326-340, 1997.

BLUM. A; LANGELY, P. Selection of relevant features and examples in machine learning. **Artificial Intelligence**, v. 97, p. 245-271, 1997.

BRODATZ, P. **Textures: A Photographic Album for Artists and Designers**. New York: Dover Publications, 1966, 112p.

BOLME, D. S. The CSU Face Identification Evaluation System: Its Purpose, Features, and Structure. **Lecture Notes in Computer Science**, v. 2626, p. 304-313, 2003.

CEKEIKON. Produced by Hae Yong Kim. Available at: <http://www.lps.usp.br/~hae/software/cekeikon3.html>. Accessed on April 21st, 2013.

CHANG, T.; KUO, C.-C. J. Texture Analysis and Classification withTree-Structured Wavelet Transform. **IEEE Trans. Image Process**, v. 2, n. 4, p. 429-441, 1993.

CLAUSI, D. A. An analysis of co-occurrence texture statistics as a function of grey level quantization. **Remote Sensing**, v. 26, n. 1, p 45-62, 2002.

DAUBECHIES, I. The Wavelet Transform, Time-Frequency Location and Signal Analysis. **IEEE Trans. Inf. Theory**, v. 36, n. 5, p. 961-1005, 1990.

HALL-BEYER, M. **The GLCM Tutorial Home Page**, v. 2.10, February 2007. Available at: <http://www.fp.ucalgary.ca/mhallbey/tutorial.htm>. Accessed on: April 21st, 2013.

HARALICK, R.; SHANMUGAM, K.; DINSTEIN, I. Textural Fatures for Image Classification. **IEEE Trans. Syst., Man, Cybern., Syst**, v. 3, n. 6, p. 610-621, 1973.

HOSSAIN, S.; SERIKAWA, S. Texture Databases – A Comprehensive Survey. **Pattern Recognition Letters**, 2013, In press. DOI: http://dx.doi.org/10.1016/j.patrec.2013.02.009

INZA, I. et al. Feature Subset Selection by Bayesian network-based optimization. **Artificial Intelligence**, v. 123, p. 157-184, 2000.

ITO, R.; KIM, H.; SALCEDO, W. Classificação de Texturas Invariante a Rotação Usando Matriz de Co-ocorrência. In: INTERNATIONAL INFORMATION AND TELECOMMUNICATION TECHNOLOGIES SYMPOSIUM, 8. 2009, Florianópolis.

JULES, B. Texture and Visual Perception. **Scientific American**, v. 212, n. 2, p. 38-48, 1965.

KOHAVI, R.; JOHN, G. H. Wrappers for feature subset selection. **Artificial Intelligence**, v. 97, p. 273-324, 1997.

LAHAJNAR, F.; KOVAČIČ, S. Rotation-invariant texture classification. **Pattern Recognition Letters**, v. 24, p. 1151-1161, 2003.

LEE, C. Fuzzy Logic in Control Systems: Fuzzy Logic Controller, Part I. **IEEE T. Syst. Man Cyb.**, v. 20, n. 2, p. 419-435, 1990.

_____. Fuzzy Logic in Control Systems: Fuzzy Logic Controller, Part II. **IEEE T. Syst. Man Cyb.**, v. 20, n. 2, p. 419-435, 1990.

LIU, G.; YANG, J. Image retrieval based on the texton co-occurrencematrix. **Pattern Recognition**, v. 41, p. 3521-3527, 2008.

MAILLARD, P. Comparing Texture Analysis Methods through Classification. **Photogrammetric Engineering & Remote Sensing**, v.69, n. 4, p. 357-367, 2003.

MALDONADO, S.; WEBER, R.; BASAJ, J. Simultaneous feature selection and classification using kernel-penalized support vector machines. **Information Sciences**, v. 181, p. 115-128, 2011.

MAMDANI, E. H. Application of Fuzzy Logic to Approximate Reasoning Using Linguistic Synthesis. **IEEE Trans. Comput.**, v. 26, n. 12, p. 1182-1191, 1977

MATERKA, A.; STRZELECKI, M. Texture Analysis Methods – A Review. **COST B11 report,** 1998.

NANNI, L.; LUMINI, A.; BRAHNAM, S. Local binary patterns variants as texture descriptors for medical image analysis. **Artificial Intelligence in Medicine**, v. 49, p. 117-125, 2010.

OJALA, T.; PIETIKÄINEN, M,; MÄENPÄÄ, T. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. **IEEE Trans. Pattern Anal. Mach. Intell.**, v. 24, n. 7, p. 971-987, 2002.

OPEN SOURCE COMPUTER VISION. Available at: <http://opencv.org/>. Accessed on April 21st, 2013.

OUTEX Texture Database: Contributed Test Suite *Contrib_TC_00001*. Maintained by the University of Oulu. Available at: <outex.oulu.fi/db/classification/tmp/Contrib_TC_00001.tar.gz>. Accessed on: April 2nd, 2013.

POK, G.; RYU, K.H.; LYU, J. Rotation and Gray-Scale Invariant Classification of Textures Improved by Spatial Distribution of Features. **Lecture Notes in Computer Science,** v. 3588, p. 250-259, 2005.

PORTER, R.; CANAGARAJAH, N. Robust rotation-invariation texture classification: wavelet, Garbor filter and GMRF based schemes. **IEE Proc -Vis Image Signal Process**, v. 144, n. 3, 1997.

PUN, C.; LEE, M. Log-Polar Wavelet Energy Signatures for Rotation and Scale Invariant Texture Classification. **IEEE T. Pattern Anal.**, v. 25, n. 5, p. 590-603, 2003.

ROMERO, A.; RAMIREZ, B. Rotation-invariant texture features from the steered Hermite transform. **Pattern Recognition Letters**, v. 32, p. 2150-2162, 2011.

SALEM, Y.; NASRI, S. Rotation Invariant Texture Classification using Support Vector Machines. In: INTERNATIONAL CONFERENCE ON COMMUNICATIONS, COMPUTING AND CONTROL APPLICATIONS. 2011, Tunisia.

SANTINI, S; JAIN, R. Similarity Measures. **IEEE Trans. Pattern Anal. Mach. Intell.**, v. 21, n. 9, p. 871-883, 1999.

SIQUEIRA, F. R.; SCHWARTZ, W. R.; PEDRIN, H. Multi-Scale Gray Level Co-Occurrence Matrices for Texture Description. **Neurocomputing.** 2013. In press. DOI:
http://homepages.dcc.ufmg.br/~william/papers/paper_2012_NeurocomputingB.pdf

SONG, C.; YANG, F.; Li, P. Rotation Invariant Texture Measured by Local Binary Pattern for Remote Sensing Image Classification. In: INTERNATIONAL WORKSHOP ON EDUCATION TECHNOLOGY AND COMPUTER SCIENCE, 2. 2010, Qingdao.
SUN, X. et al. Multi-scale local pattern co-occurrence matrix for textural image classification. In: IEEE WORLD CONGRESS ON COMPUTATIONAL INTELLIGENCE, 2012, Brisbane.

VARGAS, J.F. et al. Offline signature verification based on grey level information using texture features. **Pattern Recognition**, v. 44, p. 375-385, 2011.

WEI, W. Research on Remote Sensing Image Classification Based on Neuro-Fuzzy and Texture Analysis. In: CHINESE CONTROL CONFERENCE, 29. 2010, Beijing.

ZADEH, L.A. Fuzzy Sets. **Information and Control**, v. 8, p. 338-353, 1965.

_____. Fuzzy Logic. **Computer**, v. 21, n. 4. p. 83-93, 1988

ZHANG, J; TAN, T. Brief review of invariant texture analysis methods. **Pattern Recognition**, v. 35, p. 735-747, 2002.

ZHENG, C.; SUN, D.; ZHENG, L. Recent applications of image texture for evaluation of food qualities – a review. **Trends in Food Science & Technology**, v. 17, p. 113-128, 2006.