

Real-Time Surface Grading of Ceramic Tiles

Fernando López García

A Thesis presented for the degree of
Doctor en Informática



Departamento de Informática de Sistemas y Computadores
Universidad Politécnica de Valencia

Spain

September 2005

Supervised by Dr. José Miguel Valiente González

Abstract

This thesis presents a case of study of the development and performance analysis of a surface grading application with real-time compliance. The application focuses on the ceramic tile industry and aims to automate the inspection process of surface grading and removing human inspectors from this subjective and tedious task. First, an overview of surface grading works is given. These works have been reported in recent years in many production areas such as ceramic tile, marble, granite and wood industries.

We then address the issue of spatial and temporal uniformity in the acquisition system. In a surface grading application it is crucial to ensure the uniform response of the system through time and space. Spatial and temporal uniformity is demonstrated and two illuminating systems (high frequency uniform fluorescents and white LED arrays) are compared from this point of view.

All the results presented for surface grading were obtained using real data from the ceramic tile industry. One of the aims of the thesis has been to build an extensive image database of ceramic tile models representing the wide range of ceramic tile surface classes. The VxC TSG database is public and can be accessed at miron.disca.upv.es/vision/vxctsg/.

Afterwards, we present a study of methodologies developed to obtain a fast and accurate approach to surface grading. From this study is extracted a method based on *soft colour-texture descriptors* computed in perceptually uniform colour spaces. The method is parameterized and the involved factors are studied using two statistical procedures; *experimental design* and *logistic regression*. Although it is not a new theoretical contribution, we have found and demonstrate that a simple set of global colour and texture statistics, together with well-known classifiers, are powerful enough to meet stringent factory requirements for real-time and performance. Two approaches from literature were also implemented, parameterized and statistically studied for comparative purposes. These methods are Colour Histograms and Centile-LBP.

Finally, we explore the method's capacity for on-line inspection in a study of real-time compliance and parallelization based on cluster and MPI technologies.

Resumen

Esta tesis presenta un caso de estudio para el desarrollo y análisis de una aplicación de gradación de superficies con restricciones de tiempo real. La aplicación se centra en la industria cerámica y su objetivo es automatizar el proceso de gradación de superficies sustituyendo a los operadores humanos en esta tarea tediosa y subjetiva. En primer término, se presenta una revisión de los trabajos de gradación de superficies presentes en la literatura. Estos trabajos han sido realizados en los últimos años en varias áreas productivas, como son las industrias del azulejo, mármol, granito y madera.

Los resultados presentados en la tesis relativos a la gradación de superficies han sido obtenidos utilizando datos reales procedentes de la industria azulejera. Uno de los objetivos de la tesis ha sido construir una extensa base de datos de imágenes de azulejos que represente el amplio rango de clases de superficie presentes en la industria azulejera. Esta base de datos se ha denominado VxC TSG y es accesible en miron.disca.upv.es/vision/vxctsg/. Previamente a la presentación de la base de datos se procede al estudio de la uniformidad espacial y temporal del sistema de adquisición. En las aplicaciones de gradación de superficies esta uniformidad es crucial. Se demuestra la uniformidad espacio-temporal al mismo tiempo que se comparan dos modernos sistemas de iluminación; los fluorescentes de alta frecuencia y los LEDs blancos.

Después se presenta un estudio de metodologías desarrolladas para obtener una aproximación rápida, fiable y precisa para la gradación de superficies. Este estudio finaliza con la presentación de un nuevo método basado en la computación de descriptores suaves de color y textura en espacios de color perceptualmente uniformes (*soft colour-texture descriptors method*). Este método es parametrizado y los factores involucrados son estudiados utilizando dos procedimientos estadísticos; el *diseño de experimentos* y la *regresión logística*. Aunque el método presentado no es una nueva contribución teórica, se demuestra que un conjunto sencillo de estadísticos globales de color y textura, junto con clasificadores bien conocidos, son suficientes para superar los requisitos solicitados en factoría relativos a la precisión, fiabilidad y capacidad de inspección en línea del sistema. Otros dos métodos proce-

dentes de la literatura son factorizados y estudiados utilizando los procedimientos estadísticos anteriormente mencionados. Este trabajo es llevado a cabo con fines comparativos.

Finalmente, se estudia la capacidad del sistema para una inspección en línea del 100% de la producción. Este estudio incluye la paralelización del método utilizando una tecnología basada en MPI y clusters.

Resum

Esta tesi presenta un cas d'estudi per al desenvolupament i anàlisi d'una aplicació de gradació de superfícies amb restriccions de temps real. L'aplicació se centra en la indústria ceràmica i el seu objectiu és automatitzar el procés de gradació de superfícies substituint als operadors humans en esta tasca tediosa i subjectiva. En primer terme, es presenta una revisió dels treballs de gradació de superfícies presents en la literatura. Estos treballs han sigut realitzats en els últims anys en diverses àrees productives, com són les indústries del taulellet, marbre, granit i fusta.

Els resultats presentats en la tesi relativa a la gradació de superfícies han sigut obtinguts utilitzant dades reals procedents de la indústria del taulellet. Un dels objectius de la tesi ha sigut construir una extensa base de dades d'imatges de taulellets que represente l'ampli rang de classes de superfície presents en la indústria del taulellet. Esta base de dades s'ha denominado VxC TSG i és accessible en miron.disca.upv.es/vision/vxctsg/. Prèviament a la presentació de la base de dades es procedix a l'estudi de la uniformitat espacial i temporal del sistema d'adquisició. En les aplicacions de gradació de superfícies esta uniformitat és crucial. Es demostra la uniformitat espai-temporal ahora que es comparen dos moderns sistemes d'il·luminació; els fluorescents d'alta freqüència i els LEDs blancs.

Després es presenta un estudi de metodologies desenvolupades per a obtindre una aproximació ràpida, fiable i precisa per a la gradació de superfícies. Este estudi porta a la presentació d'un nou mètode basat en la computació de descriptors suaus de color i textura en espais de color perceptualment uniformes (*soft colour-texture descriptors method*). Este mètode és parametrizado i els factors involucrats són estudiats utilitzant dos procediments estadístics; el disseny d'experiments i la regressió logística. Encara que el mètode presentat no és una nova contribució teòrica, es demostra que un conjunt senzill d'estadístics globals de color i textura, junt amb classificadors ben coneguts, són suficients per a superar els requisits sol·licitats en factoria relatius a la precisió, fiabilitat i capacitat d'inspecció en línia del sistema. Altres dos mètodes procedents de la literatura són factorizats i estudiats en profunditat utilitzant els mètodes estadístics mencionats anteriorment. Este treball és dut a terme amb fins comparatius.

Finalment, s'estudia la capacitat del sistema per a una inspecció en línia del 100% de la producció. Este estudi inclou la paralelització del mètode utilitzant la tecnologia basada en MPI i clusters.

Acknowledgments

Many people helped me to develop this work, thanks to all of them.

Contents

Abstract	ii
Resumen	iii
Resum	v
Acknowledgements	vii
1 Introduction	1
1.1 Motivation and goals of this work	2
1.2 Previous works on ceramic tiles	4
1.3 Thesis outline	6
2 Overview of surface grading works	9
2.1 Surface features	10
2.1.1 Colour	10
2.1.2 Texture	12
2.1.3 Colour and Texture	16
2.2 Surrey works on surface grading	17
2.3 CVC works on surface grading	25
2.4 Oulu works on surface grading	30
2.5 Other minor works	33
2.6 Conclusions	35
3 VxC TSG image database	37
3.1 Acquisition system	38

3.1.1	Scan Line Camera	39
3.1.2	Optical lenses	41
3.2	Study of spatial and temporal uniformity.	42
3.3	VxC TSG description.	46
3.4	Conclusions	51
4	On the search for a fast and accurate approach to surface grading	53
4.1	Image tessellation and local Lab statistics	54
4.2	Global Lab statistics	60
4.3	Literature methods	62
4.4	Conclusions	64
5	Extraction of <i>soft colour-texture descriptors</i> method	65
5.1	CIE Lab and CIE Luv	66
5.2	Soft colour-texture descriptors	68
5.3	Classifiers	70
5.4	Experiments and results	74
5.4.1	Experimental design	76
5.4.2	Logistic regression	78
5.4.3	Selection of best combinations	82
5.4.4	Results	82
5.5	Conclusions	88
6	Literature methods	90
6.1	Colour histograms	90
6.2	Centile-LPB	93
6.3	Experimental design and results	95
6.4	Comparison with <i>soft colour-texture descriptors</i> method	106
7	Study of real-time compliance	108
7.1	Sequential inspection process	109
7.2	Tile extraction	111
7.3	Data scaling vs real-time compliance	112

7.4	MPI-Cluster architecture	114
7.5	Parallelization experiments and results	115
7.6	Conclusions	121
8	Conclusions, discussion and further work	123
A	Image registration method for ceramic tiles	129
A.1	Introduction	129
A.2	Registration method	130
A.3	Experiments and results	134
	Bibliography	138

List of Figures

1.1	One factory prototype for the automatic inspection of integrity defects in ceramic tiles corners.	5
1.2	Samples from VxC TSG image database. From up to down; three samples of petra and marfil models, each one corresponding to a different surface grade.	7
2.1	The visible light spectrum.	10
2.2	Some samples of textures.	13
2.3	Ordered binary tree.	18
2.4	Computation of local binary pattern (LBP).	32
3.1	Acquisition system.	39
3.2	Prototype at VxC laboratory.	40
3.3	Tiles used in the study of spatial and temporal uniformity. From left to right, up to down; venice, vega, blue venice, somport, mediterranea and granito.	43
3.4	System response over 54 hours using fluorescents and LEDs respectively.	45
3.5	VxC TSG marble samples. From up to down; three samples of agata, marfil, venice and vega models, each one corresponding to a different surface grade.	48
3.6	VxC TSG granite samples. From up to down; three samples of berlin, oslo and lima models, each one corresponding to a different surface grade.	49

3.7	VxC TSG stone samples. From up to down; three samples of antique, campinya, firenze and mediterranea models, each one corresponding to a different surface grade.	50
3.8	VxC TSG stone samples. From up to down; three samples of petra, santiago and somport models, each one corresponding to a different surface grade.	51
4.1	First approach to surface grading based on tessellation and local statistics. Sorted vector of standard deviations and image splitting into two regions using squares variability.	56
4.2	Second approach to surface grading based on tessellation and local statistics. Averaged vectors of ordered standard deviation vectors corresponding to each toscana classes with a square size of 15x15 pixels.	57
4.3	Second multiresolution approach based on the computation of areas belonging to ordered standard deviation vectors in several image scales.	59
5.1	95% confidence intervals obtained from the estimated error rate for test sets with sizes; 10 (outside curves), 25, 50 , 100, 250, 2500 and 5000 (inside curves).	73
5.2	Block diagram for the selection of the best combination of factors. . .	76
5.3	Logistic regression model vs linear regression model.	81
5.4	Presence in percentage of colour spaces and classifiers in best combinations sets.	85
5.5	Computed logistic regression model for <i>soft colour-texture descriptors</i> method.	86
6.1	Computed logistic regression model for colour histograms.	99
6.2	Computed logistic regression model for centile-LBP.	100
6.3	Presence in percentage of colour spaces in best combinations of colour histograms and centile-LBP.	103
6.4	Presence in percentage of distances in best combinations of colour histograms and centile-LBP.	104

6.5	Presence in percentage of classifiers in best combinations of colour histograms and centile-LBP.	105
6.6	Timing comparison of surface grading approaches using the best combination of factors.	107
7.1	Acquired tile presenting a slightly inclination.	111
7.2	Tile extracted from background, positioned at origin (left-upper corner) and registered with a non-inclined rectangle.	112
7.3	Accuracy versus scale using the extracted <i>soft colour-texture descriptors</i> method.	113
7.4	The inspection system with a cluster processing unit for parallelization.	116
7.5	Timing evolution in Mercurio cluster using several image resolutions.	118
7.6	Timing comparison between Mercurio and Hyades clusters using several image resolutions.	119
7.7	Speedup and Efficiency in Mercurio and Hyades.	121
A.1	Line fitting process.	132
A.2	Tile rectangles obtained on several test images.	132
A.3	Transformation mapping from test tile to reference tile.	133
A.4	Tile registration examples.	137

List of Tables

1.1	Summary of surface grading literature.	3
3.1	Models of the VxC TSG image database.	46
4.1	Ground truth of ceramic tiles used in image tessellation and local Lab statistics approaches.	55
4.2	Accuracy results for the first approach based on image tessellation and local statistics.	56
4.3	Accuracy results for the second approach based on image tessellation and local statistics.	58
4.4	Accuracy results for the third approach based on image tessellation and local statistics.	58
4.5	Lab accuracy results for the fourth approach based on image tessellation and local statistics. Multiresolution approaches.	59
4.6	New models added to the image database of ceramic tiles.	61
4.7	Best accuracy results of global Lab statistics method. From left to right; mean, standard deviation, average deviation, histogram moments from 2nd to 5th, histogram moments from 6th to 10th, CIE Lab, RGB, accuracy and 95% confidence intervals. Accuracy and confidence intervals are given in %.	62
4.8	Accuracy results of colour histograms and centile-LBP. From left to right; Chi square, linear correlation and log-likelihood distances, accuracy, 95% confidence intervals. Accuracy and confidence intervals are given in %.	63

5.1	Factors involved in the design of experiments of <i>soft colour-texture descriptors</i>	75
5.2	Orthogonal sequence of experiments for the soft colour-texture descriptors method.	79
5.3	Thirty best combinations of factors. Predicted accuracy (p*100) and confidence intervals are given in %.	87
6.1	Factors involved in the design of experiments of colour histograms and centile-LBP.	95
6.2	Orthogonal sequence of experiments for the colour histograms and centile-LBP methods.	98
6.3	Thirty best combinations of factors using colour histograms. Predicted accuracy (p*100) and confidence intervals are given in %.	101
6.4	Thirty best combinations of factors using centile-LBP. Predicted accuracy (p*100) and confidence intervals are given in %.	102
6.5	Best results of surface grading approaches.	106
7.1	Timing costs of sequential inspection processes.	110
7.2	Timing costs of sequential inspection processes using new image resolutions.	114
7.3	Timing of parallelization in Mercurio cluster using different image resolutions (time in milliseconds).	117
7.4	Timing of parallelization in Hyades machine using several image resolutions (time in milliseconds).	119
A.1	Experiments with artificially introduced transformations.	135
A.2	Absolute and relative mean pixel errors.	136
A.3	Absolute and relative mean pixel errors obtained with real images.	136

Chapter 1

Introduction

In recent years, automatic inspection systems have become of paramount importance for industries with serial productive plans. These industries are commonly characterized by the following items:

- Complex processes formed by multiple stages at production lines.
- High production rates.
- High added value of every piece.
- Quality control stage for the final product.
- 100% inspection at factory rates.

The ceramic tile industry is one clear example of this type of productive systems. However, factories continue to use human inspectors to grade tile quality. These operators, located at the end of the production lines, inspect the final product and are often affected by problems such as eye fatigue, sickness or boredom. Furthermore, the different criteria of each operator regarding tile defects could produce a non-uniform quality control criterion.

In this industry the great majority of tile faults are surface defects, thus, surface inspection is an important quality control subject to automate. But, this is a complex work which is divided into several tasks due to the diversity of existing surface faults. *Surface grading* is one of the most important issues of surface inspection.

1.1 Motivation and goals of this work

There are many industries manufacturing flat surface materials that need to split their production into homogeneous series grouped by the global appearance of the final product. These kinds of products are used as wall and floor coverings. Some of them are natural products such as marble, granite or wooden boards, and others are artificial, such as ceramic tiles.

In these industries the quality control stage is crucial in remaining competitive. One of the most important quality problems is the non-uniformity of the visual aspect of the product within the same lot of a specific model. As the final product is used to form areas which are supposed to be uniform in appearance, the presence of pieces which look slightly different is considered a serious quality defect.

Currently, industries rely on human operators to perform the task of surface grading. Human grading is subjective and often inconsistent between different graders. In the area of wood inspection some observations showed low correspondence between graders. In a test of four grades, different grading operators agreed in only 60% of the samples [1]. Thus, automatic, reliable systems are needed. Also, real-time compliance is an important issue as systems need to be able to inspect global production at on-line rates.

Surface grading is related with the automatic classification of flat pieces presenting random, pseudo-random or fixed surface patterns. The aim of surface grading is to split the production into different classes sorted by their global appearance which depends on colour and texture properties.

In recent years many approaches to surface grading have been developed (see Table 1.1). Boukouvalas et al [2, 3] proposed colour histograms and dissimilarity measures of these distributions to grade ceramic tiles.

Other works consider specific types of ceramic tiles; *polished porcelain* tiles, which imitate granite. These works include texture features. Baldrich et al [4, 5] proposed a perceptual approximation based on the use of discriminant features defined by human classifiers at factory. These features mainly concerned grain distribution and size. The method included grain segmentation and features measurement. Lumbreras et al [6, 7] joined colour and texture through multiresolution decompositions on

several colour spaces. They tested combinations of multiresolution decomposition schemes (Mallat's, *à trous* and wavelet packets), decomposition levels and colour spaces (Grey, RGB, Otha and Karhunen-Loève transform). Peñaranda et al [8, 9] used the first and second histogram moments of each RGB space channel.

Kauppinnen et al [1, 10, 11] developed a method for grading wood based on the Percentile (or centile) features of histograms calculated for RGB channels. Kyllönen and Pietikäinen's approach [12] uses colour and texture features. They chose centiles for colour, and LBP (Local Binary Pattern) histograms for texture description.

Lebrun and Macaire [13] describe the surfaces of the Portuguese "Rosa Aurora" marble using the mean colour of the background and mean colour, absolute density and contrast of marble veins. They achieved good results but their approach is very dependent on the properties of this marble. Fernández et al [14] studied surface grading for granite blocks. They used the histograms of the RGB colour space (one for each channel) and a simple measure of dissimilarity; the sum of absolute differences of each bin-pairs of the histograms to be compared. Finally, Kukkonen et al [15, 16] presented a system for grading ceramic tiles using spectral images. Spectral images have the drawback of producing great amounts of data.

Table 1.1: Summary of surface grading literature.

	ground truth	features	real-time study	accuracy
Boukouvalas	ceramic tiles	colour	no	-
Baldrich	polished tiles	colour/texture	no	99%
Lumbreras	polished tiles	colour/texture	no	92.7%
Peñaranda	polished tiles	colour/texture	yes	-
Kauppinen	wood	colour	yes	72%
Kyllönen	wood	colour/texture	yes	-
Lebrun	marble	colour/texture	no	98%
Fernández	granite	colour	no	-
Kukkonen	ceramic tiles	colour	no	70%

From the literature review (see Chapter 2 for more information) we deduced that many of these approaches specialized in a specific type of surface, others did not achieve good enough accuracy or simply did not provide accuracy information, others did not carry out extensive studies of performance, and yet others did not

take into account the time restrictions of a real inspection at factory. As a result, we think surface grading is still an open issue where more contributions are possible. In this sense, the present thesis deals with some less explored aspects in terms of real-time compliance and surface grading performance.

The main thesis goal has been to develop a generic, fast and accurate surface grading method suitable for use in a wide range of surfaces, also able to comply with real-time requirements for on-line inspection at factory. This major aim is complemented with other goals, such as:

- In-depth study of the acquisition system in order to demonstrate spatial and temporal uniformity. This also involved the comparison of two modern illumination systems.
- Building of an extensive image database of ceramic tiles for surface grading. The aim of this database is to ensure extensive performance study and methods comparison.
- In-depth study of methods based on statistical tools in order to obtain objective and valid conclusions.
- Comparison with similar literature approaches.
- In-depth study of real-time compliance under real factory conditions

1.2 Previous works on ceramic tiles

The scope of the thesis is the surface grading application, but this is not an isolated work. It stems from a more extensive work performed in recent years by the VxC group for the automation of ceramic tiles inspection. VxC is a research group located at Polytechnic University of Valencia and specialized in computer vision topics.

Work on ceramic tiles began in the second part of 90's when a collaboration agreement was established between the ceramic tile company Keraben S.A. and the VxC group. The aim of this project was to develop a prototype able to automatically inspect a specific surface fault on ceramic tiles; the integrity defects at tiles corners.

At that moment, losses of materials in corners were the most important surface fault representing 90% of surface defects. These losses were due to crashes in the transportation of pieces along the production lines. In [17–19] is summarized the work carried out for the development of this prototype. Finally, two prototypes were developed. They were successfully tested at factory during six months (see Figure 1.1). The system was patented as an utility model and a second technological company was interested in its production and commercialization. This company went to bankrupt and finally the commercial system was not carried out. Later, interest on system was lost because integrity defects at factories were drastically reduced by modernizing transportation systems based on conveyor belts.

Figure 1.1: One factory prototype for the automatic inspection of integrity defects in ceramic tiles corners.



After this first work, a FEDER-CICYT project (1FD97-0999) related to automation of ceramic tiles inspection was carried out during first years of 00's. In this project we developed a methodology for the surface inspection of fixed patterned tiles [20–23]. The method performs a comparison with ideal references free of faults in order to extract the surface defects. Fast and accurate approach to registration between reference and inspected tiles was specially studied [20, 21]. Actually, the project was more ambitious and in its initial planning also covered defects detection on random and pseudo-random patterned surfaces. Although the work on these

items was started, a final method was not developed. The project also covered the issue of surface grading presented in this thesis and a very first approach was done using image tessellation and colour description of homogeneous and non-homogeneous regions [24].

Once the project was finished, work on surface grading was continued as the subject of the present thesis. A medium-size image database was collected in collaboration with Keraben S.A. and new approaches to fast surface grading derived from [24] were studied. At the end of 2003 a new FEDER-CICYT project (DPI2003-09173-C02-01) was assigned to VxC group. In this project, which is still not finished, surface grading is one of the main subjects. The performed work relative to surface grading is collected in [25–28] and the present thesis document.

1.3 Thesis outline

Chapter 2 presents an overview of surface grading works done in some industrial areas such as ceramic tile, parquet slab, woods, granite, marble. They include major works by several university groups and other isolated minor works.

The acquisition system and the study of its uniform response through time and space is described in Chapter 3. Spatial and time uniformity is of great importance in ensuring good surface grading performance [1, 2, 4, 8]. Slight changes in illumination or acquisition conditions could easily introduce surface misclassifications. Any alteration in the illumination conditions modify surface colour property giving rise to a false change in the surface class. We also present a study of different illumination systems from the point of view of spatial and time uniformity. The studied systems are; high frequency uniform fluorescents and white LED arrays. The results show that only fluorescent systems provide sufficient uniform response.

Chapter 3 also presents the VxC TSG (VxC Tiles for Surface Grading) which is an image database of ceramic tiles oriented to surface grading. Creating and compiling this database has been one important goal of the thesis (see Figure 1.2). The VxC TSG is intended to be a tool for the scientific community and future works in the field of surface grading. It is also the ground truth used in the thesis for testing

and comparing surface grading approaches. VxC TSG is an extensive image data base of ceramic tile models representing the wide range of surface classes in ceramic tiles. It is public and can be accessed at miron.disca.upv.es/vision/vxctsg/.

Figure 1.2: Samples from VxC TSG image database. From up to down; three samples of *petra* and *marfil* models, each one corresponding to a different surface grade.



Our search for a fast and accurate method for the purpose of surface grading is revised in Chapter 4. Here, we collect the previous works that finally gave rise to the *soft colour-texture descriptors* method. Work relative to this preliminary approaches has been published in [25] and [26].

In Chapter 5 we develop the in-depth statistical study performed to extract the final approach to surface grading based on *soft colour-texture descriptors*. Although the method is not a new theoretical contribution we have found and demonstrate that a simple set of global statistics softly describing colour and texture [29] computed in perceptually uniform colour spaces (CIE Lab or CIE Luv), together with well-known classifiers [30], are enough to fulfil stringent factory requirements. The two main needs of the industry are; on-line inspection at factory rates (real-time compliance) and accurate performance in surface grading. Production managers at factories will only accept an error rate close to 5% before relying on these automatic grading

systems. The method we present meets the first demand by using the simplest and fastest [to compute] colour-texture features. The second demand is met by achieving average accuracies of more than 95% in many tests carried out using the VxC TSG database. We named this approach *soft colour-texture descriptors* because it uses the less complex texture and colour descriptors known in the literature [29]. The method was extracted from a statistical procedure which is a combination of *experimental design* [31] and *logistic regression* [32] analysis. This procedure is used to determine the best combination of quantitative/categorical factors in terms of a set of experiments that maximize or minimize one response variable also involved in the experiments. We used the accuracy rate of classifications as response variable. The *soft colour-texture descriptors* method has been recently reported and accepted in [28] .

Two methods from the surface grading literature are also implemented and tested in Chapter 6 for comparison purposes. These methods are colour histograms [2, 3] and centile-LPB [1, 12]. We selected these methods from literature because they are similar to ours; they are generic solutions with low computational costs. An experimental design and logistic regression analysis was also performed using the VxC TSG database in order to determine the best combination of proposed factors providing the best accuracy results. Results show that all methods are almost equal in accuracy performance but *soft colour-texture descriptors* method achieved better results in timing costs.

Chapter 7 presents a study of real-time compliance including the parallelization of the method proposed in Chapter 5. This study is an in-depth exploration of the real-time compliance of the approach. We use the parallel architecture provided by the cluster-MPI model. The method is easily translated to this architecture and the results demonstrate that, in conjunction with standard computing technologies, the approach is able to inspect and grade more surface area per time unit than factories can produce on a production line. Work relative to real-time compliance has been published in [27].

Finally, the conclusions of all chapters and further work are summarized in Chapter 8.

Chapter 2

Overview of surface grading works

This chapter is devoted to surface grading literature. Although there are many works in literature related to surface grading, we have found they were not interconnected. These approaches were performed without establishing almost any reference among them and also does not exist a general term to describe the automatic inspection of surface materials in order to split their production into homogeneous series grouped by the global appearance. In this chapter we compile and introduce these works unifying them under the term of *surface grading* works.

We present an overview of surface grading works found in literature and performed in several industrial areas such as ceramic tile, parquet slab, wood, granite and marble. This overview include major works by several university groups and other isolated minor works. Major works were done at the University of Surrey in UK, the Computer Vision Centre at the Autonomous University of Barcelona in Spain and the Oulu University in Finland.

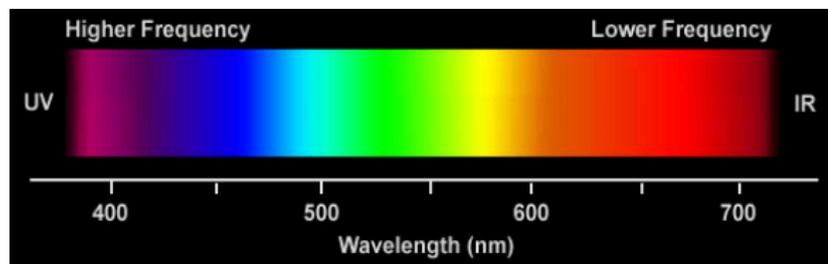
All the presented approaches used colour properties or a combination of colour and texture properties to characterize surface appearance. Therefore, previous to proper surface grading overview we present literature approach to the colour and texture properties.

2.1 Surface features

2.1.1 Colour

Colour is a sensation created in response to the excitation of our visual system by the electromagnetic radiation known as light [33,34]. More specifically, colour is the perceptual result of light from the visible region of the electromagnetic spectrum when this light meets the retina of the human eye. The visible region covers the wavelengths from 400nm to 700nm (see Figure 2.1).

Figure 2.1: The visible light spectrum.



The human retina has two kinds of receptors, rods and cones. The basic function of rods is to provide monochromatic vision under low illumination levels. The rods have a photosensitive pigment called rhodopsin. This pigment absorbs light most strongly in the blue-green region of the spectrum. This part of human vision is referred to as *scotopic vision*. Although rods are important for vision, they play no role in image reproduction. The function of the cones is to provide colour vision at normal levels of illumination. This is known as *photopic vision*. The human retina has three types of cones and each one is sensitive to a different wavelength range of the visible spectrum.

The area of science concerned with the description and specification of colour is called colorimetry [33]. As we have three types of colour receptor cells (cones), traditionally, three numerical components have been used to describe colours. Therefore, a colour can be specified by a vector with three components. The set of all colours forms a vector space called colour space or colour model. The three components of a colour can be defined in many different ways providing various colour spaces [33].

In 1931, the Commission Internationale de L'Eclairage (CIE) adopted standard colour curves for a hypothetical standard observer. These colour curves specify how a specific spectral power distribution (SPD) of an external stimulus (visible radiant light incident on the eye) can be transformed into a set of three numbers that specify the colour. The CIE colour specification system is based on the description of colour as the luminance component Y and two other components X and Z . The spectral weighting curves of X and Z have been standardized by the CIE based on statistics from experiments involving human observers [33]. The corresponding colour space is called the CIE XYZ colour space. The XYZ model is a device independent colour space that is useful in applications where consistent colour representation across devices with different characteristics is important. But, the CIE XYZ space is perceptually highly non-uniform [34]. Thus, it is not adequate for quantitative manipulations involving colour perception and is seldom used in image processing applications.

Traditionally, colour images have been specified by the red, green and blue tristimulus values. This is the RGB colour space. The red, green and blue components are called primary colours. In general, hardware devices such as video cameras, colour image scanners and computer monitors process colour information based on these primary colours. Other popular spaces in image processing are the YIQ (North American TV standard), the HSI (Hue, Saturation and Intensity), and the HSV (Hue, Saturation and Value) colour spaces used in computer graphics.

Although XYZ is used only indirectly, it has a significant role in image processing since other colour spaces can be derived from it through mathematical transforms. For example, the linear RGB colour space can be transformed to and from the CIE XYZ colour space using a linear three-by-three matrix transform. Similarly, other colour spaces, such as non-linear RGB, YIQ and HSI can be transformed to and from the CIE XYZ space, but might require complex and non-linear computations.

The CIE has also derived and standardized two other colour spaces from the CIE XYZ. These are the CIE Luv and CIE Lab colour spaces and both of them are perceptually uniform [33]. The term 'perceptual' refers to the way that humans perceive colours. The term 'uniform' means that if we move in the colour space from

one colour to another, from one coordinate to another, the perceptual difference will be related to a measure of distance, commonly the Euclidean distance, and the same distance will approximately relate to the same perceptual difference throughout the colour space. Thus, we can measure colour differences close to the human perception of colours. This makes these colour spaces useful for applications where colour difference measurement plays an important role. This is the case of the surface grading application presented in this thesis.

CIE Luv and CIE Lab are slightly different because of the different approaches to their formulation [33,34]. Nevertheless, both spaces are equally good in perceptual uniformity and provide good estimates of colour difference (distance) between two colour vectors. CIE Luv is used for industries considering additive mixing such as colour displays, TV and lighting [35], while CIE Lab is beginning to be used in applications of colour image processing [36].

We used the CIE Lab and CIE Luv colour spaces in our approaches to surface grading, and also did experiments with the RGB colour space in order to determine the advantages of using these perceptually uniform spaces. The RGB space was chosen for comparative purposes because it is often used in image processing applications [37].

2.1.2 Texture

Texture is related to some properties inherent to the surface of objects. Texture plays an important role in human vision and its analysis is of great interest in the area of computer vision. However, a formal approach or precise definition of texture does not exist. From the point of view of image processing, one general definition is: *Texture is something consisting of mutually related elements* [38]. This definition contains the two main elements of textures. Firstly, texture is formed by simple components called *texture primitives*. Secondly, texture is defined by the spatial relationships between these simple components. Some examples of textures are shown in Figure 2.2.

There are two main approaches to texture description: statistical and structural [39]. Statistical description use feature vectors of texture properties which represent

Figure 2.2: Some samples of textures.



points in a multidimensional feature space. This is suitable for statistical pattern recognition. In contrast, the structural approach is based on the analogy between texture spatial relations and the structure of a formal language. The description of a texture forms a language that can be represented by its grammar. A grammar is constructed for each texture class. Then, the recognition process becomes a syntactic analysis of the texture description.

Structural approaches are based on the theory of formal languages and they are adequate for describing strong textures [40, 41]. A texture image is defined as a structure which is made up of a large ensemble of elements which have some kind of order in their locations. This approach works well on deterministic textures (mainly artificial textures), but most natural textures are not of this type.

From a statistical point of view, textured images are complicated pictorial patterns from which sets of statistical measures are obtained to characterize them. The simplest way to statistically characterize textures is to compile *global statistics* like mean, standard deviation and histogram moments [29]. However, the most popular approximation is the *co-occurrence matrices* method [42, 43]. In this method matrices are constructed by counting the number of occurrences of pixel pairs of given grey levels at a given displacement. Statistics such as contrast, energy, entropy and

others are computed from matrices to obtain texture features. A similar approach is presented in the *sum and difference histograms* method [44]. Here, similar features are computed from one-dimensional histograms containing the sum and difference of pairs of pixels at a given displacement. Another approach in this category is the *statistical feature matrices method* [45], where three matrices of contrast, covariance, and dissimilarity are directly computed from the texture images. Each entry in a matrix contains the corresponding feature computed at different displacements. In this case, the feature vector is directly formed using all the entries of the matrices.

Stochastic models such as *Markov random fields* or *Gibbs random fields* have also been used to extract texture features [46–49]. These approaches consider textures as different realizations of random processes. The features describing each texture are the parameters of the model which is supposed to generate the given texture. There are some difficulties with these methods such as how to choose an appropriate order for the model. Recent works have extended these methods to multiresolution approaches [50].

The *mathematical morphology* approach looks for spatial repetitiveness of shapes in an image using structure primitives. These structuring elements usually consist of some simple shape, such as a square or a line. When a binary textured image is eroded by a structuring element, texture properties are present in the eroded image [51]. Different structuring elements are applied to the textured image and the number of pixels with unit value in the eroded image is counted. These numbers are used to form a feature vector that characterizes the texture. Also, another morphological texture description was derived by using the size distribution of a sequence of opening and closing granulometries [52]. The mathematical approach to texture is often successful in granulated materials, but its performance is reduced significantly in other texture types.

Another alternative for texture description is to measure its *fractal dimension* [53]. This approach was first introduced for modeling natural scenes [54]. It was reported that the fractal dimension correlates very well with a human assessment of surface roughness. Its main advantage lies in the fact that the fractal dimension is invariant to scale and to linear transformation of data. Nevertheless, the fractal

dimension on its own is not able to give complete description of natural textures.

Another category of texture characterization methods is based on features computed from the *power spectrum* of the image. A two-dimensional power spectrum of a texture image often reveals texture periodicity and directionality. A coarse texture tends to generate low frequency components in its spectrum, while a fine texture have high frequency components. Stripes in one direction cause the power spectrum to concentrate near the line through the origin and perpendicular to this direction. These methods [55, 56] usually perform well in textures showing strong periodicity, but performance deteriorates when periodicity weakens.

In recent years, *wavelet theory* has become an important framework for multi-scale and texture image analysis [57, 58]. In general, the wavelets transform an image into a low resolution image and a series of detail images. The low resolution image is obtained by applying iteratively a low pass filter to the image, while the detail images are obtained applying a high pass filter at each step. The original image is blurred at each iteration, and the information lost during each operation remains in the corresponding detail image. Features such as the energy or mean deviation of the detail images are the most commonly used for texture description [59–62].

Finally, another approach to texture characterization is the *multi-channel spatial filtering*. Here, the methods try to imitate the behavior of the human vision system. There is evidence that texture discrimination in the human vision system is achieved by means of a set of parallel channels, each tuned for some specific feature. Each channel performs a specific spatial filtering operation. Therefore, the human visual system can be modeled as a set of spatial filters. The most common families of spatial filters are the *Gabor filters* [63, 64] and the *local Discrete Cosine (DCT) and Sine (DST) Transforms* [65, 66]. Gabor filters are basically directional filters, and are therefore appropriate for strongly oriented textures. On the other hand, from a theoretical point of view, local DCT and DST have better discriminatory power than Gabor filters for randomly oriented textures. However, they are not tunable and they cannot be used to capture some specific texture properties.

2.1.3 Colour and Texture

Colour-texture representation is a current topic in computer vision. Although both, colour and texture, are quite important properties of surfaces these two visual features have been usually studied separately. The study of colour-texture representations has received increasing attention in recent years.

The objective of many works have been to find co-join representations of spatial and chromatic information which capture the spatial dependence within and between the spectral bands. One of the most frequent approaches is the construction of a feature vector mixing grey level texture features and colour features [67, 68]. Another approach is to extend classical texture models, such as Markov random fields and the autocorrelation function, to deal with multichannel images [69, 70]. Other works convert RGB values into a single code from which texture measurements are computed as a grey scale image [71]. Spatio-chromatic representations are computed in [72, 73] over the smoothed Laplacian of image. Also, the structural tensor that is commonly used to represent local texture properties is extended to colour images in [74].

Finally, there are some works that have been influenced by known perceptual mechanisms of the human visual system. Here, the iteration of colour with spatial frequency of the coloured patterns is considered [75, 76]. These works take into account important conclusions from psychophysical experiments on colour texture interaction [36, 77–80]. They introduce a perceptual mechanism in order to simulate the *colour assimilation* phenomenon of the human visual system. This phenomenon consists of a spatial blurring of the colour representation when looking at colour textures with high spatial frequencies.

Other works present a complementary operator to simulate another phenomenon of the human visual system, the *colour contrast*, which appears when looking at colour textures with low spatial frequencies [4, 81, 82].

In this thesis, the proposed method for the purpose of surface grading uses statistical description representing colour and texture properties. Colour and texture are joined by creating feature vectors collecting global image statistics of both properties; mean, standard deviation and histogram moments. These global statistics

are computed separately in each channel of perceptually uniform colour spaces (CIE Lab or CIE Luv). We name this approach *soft colour-texture descriptors* method because it uses the less complex texture and colour descriptors known in literature [29]. This assertion is even more acceptable if we revise the classical approaches to texture description mentioned above. In fact, surface grading is not a complex problem of colour-texture recognition but differentiation. In Chapter 5 is demonstrated that soft colour-texture descriptors are powerful enough to well discriminate surface grades.

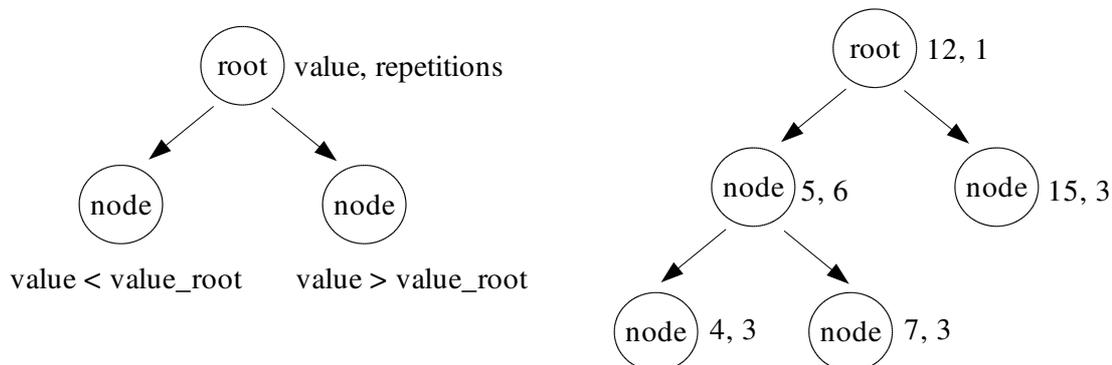
2.2 Surrey works on surface grading

Since 1995 a group of people, mainly from the Image Processing Group at the Electrical Engineering Department of the University of Surrey (UK), have been working in the area of the automatic inspection of defects and surface grading of ceramic tiles. Professor Maria Petrou has been the nexus and driving force behind all these works [2, 3, 83–94].

From the point of view of the surface grading question, the interesting part of the work is mainly that done by Boukouvalas et al [2, 3, 83–85, 89]. At a first stage they proposed using the differences between colour histograms to solve the problem of shade grading (surface grading) of multi-coloured textured surfaces (random pattern surfaces) [2, 3]. However, colour histograms are very inefficient in terms of memory requirements. A colour image acquired in RGB normally need 8 bits per colour channel at each pixel, so therefore 16Mbytes (2^{24} memory positions) are needed to store one colour histogram. However, in real images colour values tend to be clustered around just a few locations. For instance, the image of a ceramic tile may occupy only 80.000 different locations (234Kb). Apart from being highly demanding in memory, this approach is computationally intensive because in order to compare two histograms we have to parse all memory locations.

To save memory space and computational costs, they used the *binary tree* structure to store the colour histograms. A binary tree is defined as *a finite set of elements (nodes) which either is empty or consists of a root (node) with two disjoint binary trees called the left and the right subtrees of the root* [95].

Figure 2.3: Ordered binary tree.



Binary trees are frequently used to represent a set of data whose elements are retrievable through a unique key (value). If a tree is organized in such a way that for each node all values in the left subtree are less than the value of the parent node, and those in the right subtree are greater than the value of the parent node, then this tree is called *ordered binary tree* or *search tree* (see Figure 2.3). A search of a value in a tree of n elements may be performed with only $\log n$ comparisons, if the tree is balanced.

When a colour histogram is stored in a binary tree, the value of a node is a particular RGB value. This is converted to a 24 bit-integer by concatenating the R, G and B bytes. Each node also contains the number of pixels with the same RGB value (repetitions). Therefore, only RGB combinations that exist in the image are inserted in the tree, and the searching for existing nodes is very efficient.

They chose colour histograms because they are invariant to translation and rotation about an axis perpendicular to the image plane, and change only slightly with changes of viewing angle of view. In addition they are invariant to the exact spatial distribution of the coloured pixels. This property is desirable when dealing with random pattern surfaces, as often happens when dealing with ceramic tiles.

To perform the surface grading they compared the similarity (or dissimilarity) of ceramic tiles by comparing the similarity of their colour histograms. The histograms can be viewed as distributions, and, in statistics there are several methods to compare two distributions [96]. They used the *chi-square test* and the *linear correlation coefficient*.

The *chi-square* statistic is defined as

$$\chi^2 = \sum_i \frac{(N_i - n_i)^2}{n_i}$$

where N_i is the number of events observed in the i th bin, and n_i is the number expected according to some known distribution and the sum is over all bins. A large value of χ^2 indicates dissimilarity between the two distributions.

When comparing two binned data sets, with the same number of data points, the equation adopts a different form. Let R_i be the number of events in bin i for the first data set, let S_i be the number of events in the same bin for the second data set. Then the *chi-square* statistic is

$$\chi^2 = \sum_i \frac{(R_i - S_i)^2}{R_i + S_i}$$

The *linear correlation coefficient* is another test which measures the association between random variables. For pairs of quantities (x_i, y_i) , $i = 1, \dots, N$, the linear correlation coefficient r is given by

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \quad (2.1)$$

where \bar{x} is the mean of the x_i values and \bar{y} is the mean of the y_i values.

The value of r is always in the range $[-1, 1]$. The correlation is called *positive* or *direct correlation* when y tends to increase as x increases. If y tends to decrease as x increases the correlation is then called *negative* or *inverse correlation*. A value near to zero in equation 2.1 indicates poor linear correlation between the variables x and y .

Some experiments were done to test this approach. The ground truth was formed by three different models previously graded by human operators. For each model there were three different grades, and for each grade there were approximately seven samples or tiles, nearly sixty-three samples in total. Prior to colour grading, the data was spatially and temporally corrected in order to compensate the non-uniform

response of the illuminating system [2,83].

Some graphics showing model clustering when using the *chi-square* and the *linear correlation coefficient* tests were provided. They deduced from these graphics that the performance of the method was consistent with the colour grading of human experts, but they did not provide accuracy results. The *linear correlation coefficient* was finally chosen to measure similarity between histograms because it keeps the histogram differences within a well defined range. This makes it possible to select thresholds to define new surface grades.

Perceptual correction for colour grading using sensor transformations and metameric data

In a second stage they improved the method by applying two perceptual corrections. The first perceptual correction consisted of approximating the colours perceived from electronic sensors to the human perception using metameric data [2,84]. Their reasoning behind this was as follows. All the colours that can be perceived by a given observer can be computed using

$$\begin{aligned} C_1 &= \rho_1 q_{11} \alpha_1 + \rho_2 q_{12} \alpha_2 + \dots + \rho_n q_{1n} \alpha_n \\ C_2 &= \rho_1 q_{21} \alpha_1 + \rho_2 q_{22} \alpha_2 + \dots + \rho_n q_{2n} \alpha_n \\ C_3 &= \rho_1 q_{31} \alpha_1 + \rho_2 q_{32} \alpha_2 + \dots + \rho_n q_{3n} \alpha_n \end{aligned} \quad (2.2)$$

where (C_1, C_2, C_3) are the tristimulus values forming the colour, ρ_i $i = 1..n$ represents the spectral reflectance of a surface, α_i $i = 1..n$ is the spectral power distribution of the illumination, and (q_{1i}, q_{2i}, q_{3i}) $i = 1..n$ are the spectral sensitivities of the observer's sensors. The spectral range (the visible spectrum) is sampled in n equidistant positions.

From 2.2 it is deduced that, under a given illuminant, the observer will record the same tristimulus values for many different materials (surfaces). This phenomenon is called *metamerism*. Metameric colour stimuli are colour stimuli with the same tristimulus values but different spectral radiant power distributions. That is to say, they have the same spectral distributions that yield the same colour for a given set

of sensors (observer). Metameric colour stimuli are referred as *metamers*.

Two metameric colour stimuli $(\rho_1, \rho_2, \dots, \rho_n)$ and $(\rho'_1, \rho'_2, \dots, \rho'_n)$, must satisfy the following equations:

$$\begin{aligned}\rho_1 q_{11} \alpha_1 + \rho_2 q_{12} \alpha_2 + \dots + \rho_n q_{1n} \alpha_n &= \rho'_1 q_{11} \alpha_1 + \rho'_2 q_{12} \alpha_2 + \dots + \rho'_n q_{1n} \alpha_n \\ \rho_1 q_{21} \alpha_1 + \rho_2 q_{22} \alpha_2 + \dots + \rho_n q_{2n} \alpha_n &= \rho'_1 q_{21} \alpha_1 + \rho'_2 q_{22} \alpha_2 + \dots + \rho'_n q_{2n} \alpha_n \\ \rho_1 q_{31} \alpha_1 + \rho_2 q_{32} \alpha_2 + \dots + \rho_n q_{3n} \alpha_n &= \rho'_1 q_{31} \alpha_1 + \rho'_2 q_{32} \alpha_2 + \dots + \rho'_n q_{3n} \alpha_n\end{aligned}$$

As the human eye (human observer) has different spectral responses from the electronic sensors, different materials will appear as having the same colour to the human eye and other different materials will appear as having the same colour to the electronic sensors (the camera). They tried to correct this effect by introducing a perceptual correction in the system. The term 'perceptual' is referred to the attempt of making the system work as close as possible to the human vision system.

In equation 2.2, if the observer is changed, we obtain the tristimulus values (C'_1, C'_2, C'_3) that the new observer would record for the same surface, under the same illumination. And that is the aim, to compute the colour not from the point of view of the electronic sensor but from the point of view of the human eye.

$$\begin{aligned}C'_1 &= \rho_1 q'_{11} \alpha_1 + \rho_2 q'_{12} \alpha_2 + \dots + \rho_n q'_{1n} \alpha_n \\ C'_2 &= \rho_1 q'_{21} \alpha_1 + \rho_2 q'_{22} \alpha_2 + \dots + \rho_n q'_{2n} \alpha_n \\ C'_3 &= \rho_1 q'_{31} \alpha_1 + \rho_2 q'_{32} \alpha_2 + \dots + \rho_n q'_{3n} \alpha_n\end{aligned}\tag{2.3}$$

Ideally, by solving equations 2.2 (electronic sensors) for $(\rho_1, \rho_2, \dots, \rho_n)$ and substituting them into equations 2.3 it would be possible to find the stimuli this particular coloured surface would create to the second observer (human eye). However, system 2.2 is an under-determined system as in general n is much greater than 3 (typically $n = 31$). They solved this problem by assuming that they were interested only in a small subspace of the colour space which is coherent with the colour grading application (low changes in the colour appearance). They assumed that the transformation between the projections of two different sets of sensors was *locally* linear and could be expressed by a unknown 3x3 matrix T . This matrix represents the relation be-

tween the two observers. The way they used to compute the T matrix was to use pairs of corresponding triplets (C_1^j, C_2^j, C_3^j) and $(C_1^{j'}, C_2^{j'}, C_3^{j'})$ for $j = 1, 2, \dots, m$ and $m \gg 3$. The elements of the transformation matrix were determined in the least square error sense by solving the following system of equations using singular value decomposition:

$$\begin{bmatrix} C_1^{1'} & C_2^{1'} & C_3^{1'} \\ C_1^{2'} & C_2^{2'} & C_3^{2'} \\ \dots & \dots & \dots \\ C_3^{m'} & C_3^{m'} & C_3^{m'} \end{bmatrix} = \begin{bmatrix} C_1^1 & C_2^1 & C_3^1 \\ C_1^2 & C_2^2 & C_3^2 \\ \dots & \dots & \dots \\ C_3^m & C_3^m & C_3^m \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix}$$

They known the tristimulus values (C_1^j, C_2^j, C_3^j) acquired by the electronic sensor (camera's CCD), but the tristimulus values corresponding with the human vision system where unknown, and could not be computed because the reflectance functions $(\rho_1, \rho_2, \dots, \rho_n)$ were also unknown. Therefore, in order to determine the transformation matrix, they had to find a way to generate metameric reflectance functions, and they made it by using the Monte Carlo Method for generating synthetic metamers [1].

They did experiments with several sets of ceramic tiles (tile models) previously graded by human operators. Each model had three different grades or surface classes. They first extracted the transformation matrices for each set and then moved the colour data to the CIE Lab colour space. Then, for each model, they plotted the mean CIE Lab colour of each tile in this 3D colour space. The results showed better interclass distances of the clusters (grades) after the sensor transformation.

Perceptual correction for colour grading of random textures

The second perceptual correction was to simulate the spatial blurring which occurs in humans when we look texturized surfaces [2, 85]. To do so, they first removed the spatial blurring introduced by the electronic sensor, and then converted the data to a pattern-colour separable space (opponent-colours space), where they introduced blurring emulating the way the human visual system perceives colour

texture. The data was finally converted to a perceptually uniform colour space (CIE Lab), where the colour grading was performed. They reasoned as follows:

Every imaging system introduces some kind of degradation to the data it captures and a common phenomenon when dealing with fine textures is the spatial blurring that the imaging system introduces in the data. This phenomenon can be quantified in terms of how spread a point source appears to be when its image is captured. This is expressed by the *point spread function* (PSF) of the system. In order to remove this spatial blurring there are various image restoration techniques which rely on *a priori* knowledge of the PSF.

The degradation process is modeled by a function $h(x, y, x', y')$, which together with an additive noise $v(x, y)$ operates on an input image $f(x, y)$ to produce a degraded image $g(x, y)$:

$$g(x, y) = \int \int h(x, y, x', y') f(x', y') dx' dy' + v(x, y)$$

Image restoration is the process of obtaining an approximation to $f(x, y)$ given $g(x, y)$ and some knowledge of the degradation process in the form of the function $h(x, y, x', y')$. In the absence of noise the degraded image of a point source described by $f(x', y') = \delta(x' - \alpha, y' - \beta)$ would be given by $h(x, y, \alpha, \beta)$. Therefore $h(x, y, \alpha, \beta)$ is the PSF of the sensor, which in general is dependent on the position of the point (α, β) in the ideal picture.

The PSF can be computed from an image with sharp lines or step edges. In order to derive the PSF of the electronic sensor, they used an special chart with many edges and various known orientations. After obtaining the PSF they restored the image by using Wiener filtering [96, 97].

The process mentioned above is done to remove the spatial blurring introduced by the electronic sensor. In order to introduce the perceptual correction, the restored data should be spatially blurred in agreement with the blurring of the human vision system. In [98] experiments with human subjects indicate that the change in colour appearance with spatial-frequency can be explained by assuming that signals from three opponent-colour mechanisms are scaled by a gain factor that depends on the

local spatial frequency content of the image. In this respect, Zhang and Wandell proposed an extension of the CIE Lab space, based on a pattern-colour separable transformation, called *Spatial* CIE Lab (S-CIELAB) [36].

The image is then initially transformed from the CIE XYZ space to the opponent-colours space, where the three channels represent luminance, red-green and yellow-blue. The transformation is defined by:

$$\begin{aligned} O_1 &= 0.279X + 0.72Y - 0.107Z \\ O_2 &= -0.449X + 0.29Y - 0.0077Z \\ O_3 &= 0.086X - 0.59Y + 0.501Z \end{aligned}$$

Then the data of each channel are filtered by two-dimensional spatial kernels, defined as the sum of some Gaussian functions:

$$\frac{1}{\pi} \sum_i^m \frac{w_i}{\sigma_i^2} e^{-\frac{(x^2 + y^2)}{\sigma_i^2}}$$

where m is 3 for the luminance channel, and 2 for the chromatic channels and w_i and σ_i^2 are some parameters. w_i and σ_i^2 values were determined from psychological measurements of colour appearance on human subjects [36].

Finally, the blurred data are transformed back into CIE XYZ colour space and from there to CIE Lab colour space.

Grading experiments were performed using the comparison between colour histograms after transforming image data by means of the perceptual correction discussed above. The tile sets used in previous experiments were then re-graded introducing the new perceptual correction. The results showed better interclass distances of clusters (grades) after the perceptual transformation. They used the *minimum interclass distance* and the *Bhattacharyya distance* to compile interclass data before and after the correction.

In all the works there is no study about the real-time compliance in order to ensure inspection of all tiles at factory rates. Also, there is no quantitative information about the accuracy of the approach.

2.3 CVC works on surface grading

The CVC is the Computer Vision Centre of the Autonomous University of Barcelona (Spain). Several works in relation to an specific type of ceramic tile, the *polished porcelanic tile*, were carried out between 1998 and 2002 [4–7, 99].

Baldrich et al

Baldrich et al [4, 5, 100] made a perceptual approximation based on the use of discriminant features defined by human classifiers at factory. They mixed colour and texture information by means of local and global measures. They did not propose a general texture colour representation. They dealt only with textures formed by the non-oriented coloured-blobs randomly distributed on the polished tiles. The features were mainly related to grain distribution and size, thus, the method includes grain (blobs) segmentation and features measurement.

First, some human-defined characteristics for the task of surface grading, were compiled at factory:

- Fine-grained vs. coarse-grained: defines the size of the grains.
- Opened grain vs. closed grain: measure of the distance between grains of the same size (density factor).
- Light vs. dark grain colour: colour properties of a specific type of blob.
- Light vs. dark background: colour properties of the background.
- Light vs dark global colour: colour properties of the overall colour impression.

That characteristics were translated to the following global and local features:

- Global colour mean.
- Global colour standard deviation.
- Global mean colour of each type of blob.
- Global area of each type of blob.

- Local mean area of each type of blob.
- Local standard deviation of the area of each type of blob.

In order to compute this features, the coloured blobs first had to be segmented and this was performed using a supervised clustering approach. They used the K-means algorithm introducing an approximate *rgb* value for each colour center (each type of blob). The clustering was performed over the RGB space considering the Euclidean distance between *rgb* positions.

For the classification stage they used a discriminant analysis (Fisher discriminant functions) to select the prototypes providing the maximum discrimination ratio from a set of learning samples. With Fisher's approach no a priori knowledge of data is needed and it is able to select the best representation maximizing the ratio between the inter-class covariance and the intra-class covariance. A linear transform W is applied over the feature vector x of a particular image obtaining a new representation; $y = W^t x$, in a new space where discrimination capability has been maximized.

The linear transformation W that optimizes the discrimination is obtained by computing the most significant eigen vectors of the matrix $S_w^{-1} S_b$, assuring maximization of the following ratio:

$$\frac{W^t S_b W}{W^t S_w W}$$

where W^t stands for the transpose of W , S_w is the within data sparse matrix and the S_b matrix is the between class sparse matrix. They are defined as:

$$S_w = \sum_{i=1}^c \sum_{x_k \in \{L_i\}} (x_k - \mu_i)(x_k - \mu_i)^t$$

$$S_b = \sum_{i=1}^c N_i (\mu_i - \mu)(\mu_i - \mu)^t$$

where c is the number of possible classes and $\{L_i\}$ is the set of vectors that are used as learning samples in the i class. μ_i is the mean vector of the samples of the i class, N_i is the number of learning samples in the i class and μ is the global mean

vector.

From an image of a given tile the feature vector x is extracted and it is assigned to the j class if

$$|W^t x - W^t \mu_j| < |W^t x - W^t \mu_i| \quad \forall i \neq j$$

where μ_i are the prototypes of the classes.

In order to remove spatial and time variations on the image data which occur under non-constant illumination, they introduced a diagonal transform. This was a simple diagonal matrix model. It was computed using a white pattern image acquired periodically. The spatial distortions were modeled using a set of diagonal transforms $\{S_x\}$, one 3x3 matrix for each position x along the x axis where the spatial variation occurs (they used a scan-line camera). Light variations due to time were corrected in a similar way. A set of diagonal transforms were calculated $\{T_x^{t_i}\}$. This set models the distortions at time t_i referring to instant t_0 . The final set of diagonal transforms $\{D_x^{t_i}\}$ were $D_x^{t_i} = S_x T_x^{t_i}$.

They also introduced a perceptual correction based on the induction phenomenon. This phenomenon is divided in two types: chromatic assimilation and chromatic contrast. The first one implies a change in the perceived chromaticity of a given stimulus towards the chromaticity of its surround, whereas in the second the change is in the opposite direction. Chromatic assimilation was measured using a psychophysical approach of colour appearance on human subjects [36]. This approach has been used also, in Boukouvalas' works [2, 85] and other computer vision frameworks [75, 76].

Chromatic contrast is the complementary mechanism of the assimilation that takes chromaticities of regions with spatial low frequency. They defined an operator that enhances differences in transitions between lower frequency colour regions. The final goal of this operator was to produce a sharpened image for a better segmentation of texture blobs. They used a standard sharpening filter:

$$S_c(I, \gamma) = I_c - \gamma \nabla^2(I_c)$$

where I_c is the c -th channel of a colour image I of dimensions $N \times M$, $\nabla^2(I_c)$ is the Laplacian of the image channel c ($\nabla^2(I) = \partial^2 I / \partial x^2 + \partial^2 I / \partial y^2$) and γ is a constant that controls the amount of enhancement. This process is done for each channel. Nonetheless, the Laplacian operator is very noise sensitive and in order to avoid this problem the Laplacian of a Gaussian (LoG) is used.

$$S_c(I, \gamma) = I_c - LoG(I_c)$$

$$LoG(I_c) = -\frac{1}{\pi\sigma^4} \left[1 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{x^2 + y^2}{2\sigma^2}}$$

where the $LoG(I)$ expression is centered on zero and with a Gaussian standard deviation σ .

What they finally used was a modification of this common sharpening operator. Instead of operating in the RGB space they operated in the opponent space [92] which provides more perceptual approach. Also they fit the operator output in the neighborhood range of the input pixel. The expression of the new operator T is as follows:

$$T(I)_{\vec{\gamma}, w} = RGB(S(Opp(I), \vec{\gamma})_{min(I, w)}^{max(I, w)})$$

where the superindex and subindex $max(I, w)$ and $min(I, w)$ are the maximum and minimum range for each pixel inside the neighborhood w .

They tested these algorithms with a set of six different tile models and 47 classes (surface grades). The universe of samples was composed by 514 tiles. Each sample was divided in three regions which finally resulted in 1542 images. One third of the images were randomly selected for the training set and the remaining images were selected for the test set. Average accuracy results were around 94% without applying the perceptual sharpening correction and 99% when this correction was applied [4].

Lumbreras et al [6, 7, 99] developed an approach to surface grading based on multiresolution features. They combined colour and texture information through the multiresolution decomposition of each space channel in order to take as feature vector the energies and cross-correlations of the coefficient images. However, this simple approach could be used in many different ways depending on several decisions: the multiresolution decomposition scheme, the number of decomposition levels, the space for colour representation, and finally, the classification features to be computed from the decomposition. For each section they chose several options.

- Colour spaces: colour to gray conversion, raw RGB direct from the camera and frame grabber, Ohta colour space [102] (generic Karhunen-Loève transform), and Specific Karhunen-Loève transform.
- Decomposition schemes and bases: multiresolution analysis with Mallat's algorithm [58], *À trous* algorithm [103], wavelets packets [104]. Mallat's analysis and wavelet packets were performed with Daubechies orthogonal bases, whereas *à trous* decomposition used B-spline bases.
- Features: only the energy terms, all correlation signatures between decomposition levels but only within the same channel, and all the correlation signatures between channels but only within the same level.

For the classification stage they used the same approximation used by Baldrich et al. They did experiments to test the different multiresolution approaches. Sample universe comprised three models of polished porcelanic tiles. Each model was divided into eight classes or grades according to the grading operators at factory, each class contained 15 tiles. Also, two 512x512 images were captured for each tile, corresponding to the middle part of the upper and lower half. Thus, in total 720 samples. One third of these samples were selected for the training set and the remaining conformed the test set.

The results showed that no improvement in accuracy was achieved by using Ohta and specific K-L colour spaces. The best results for the three models were achieved using the RGB colour space and the correlation signatures between channels only within the same level. In this case, the accuracy in average was 92.7%.

Baldrich and Lumbreras did not study the real-time requirements of the final system at factory. Although real-time compliance was not studied, both methods seem to have significant computational costs.

2.4 Oulu works on surface grading

In the area of wood inspection, a set of works have been delivered (1999-2002) from the Machine Vision Group of the University of Oulu (Finland) [1, 10–12, 105–109]. In this case, the grading of lumber boards and parquet slabs is not related with the overall texture and colour appearance of the surface. The grade of the wood piece is assigned by detecting the wood defects (mainly knots) and then applying grade rules related to the number and types of defects found in the inspection process [1]. Therefore, from the computer vision point of view, the problem becomes a question of separating the surface into sound and faulty wood, and then classifying the defects into different types.

They focused on techniques oriented to the detection of faulty and non-faulty areas, choosing a non-segmenting approach in the sense they were not interested in a fine segmentation of defects. In the approach, images are splitted into non-overlapped rectangles which afterwards are classified as faulty or non-faulty. This coarse approximation is sufficient for the purpose of the grading task which is much closer to a global study of the appearance than an accurate splitting of the regions [11].

Kaappinen started the approach to the problem using only colour information derived from the percentile features of the RGB histograms [1, 11]. The percentiles, also called centiles, are calculated from a cumulative histogram $C_k(x)$, which is defined as a sum of all the values that are smaller than x or equal to x in the normalized histogram $P_k(x)$, corresponding to the colour channel k . Finding a value for a percentile involves finding the x when $C_k(x)$ is known, thus, requiring an inverse function of $C_k(x)$. If we denote the percentile feature with $F_k(y)$ then

$$F_k(y) = C_k^{-1}(y) = x$$

where y is a value of the cumulative histogram in the range [0%,100%].

In the classification stage they used feature vectors composed of selected sets of plain percentile features and differences of two percentile features either from the same colour channel or from two different colour channels.

First experiments for testing the method were carried out with 150 images of beech wood slabs used for training and 360 images used for grading test. To evaluate defect detection and defect recognition (classify the fault within defect classes), rectangular samples were collected from the training areas. The number of samples obtained from the training areas was 26855, 16027 of sound wood and the remainder including some kind of defect (15 classes of defects).

A set of 117 colour percentile features was calculated for the training samples. A feature selection algorithm was used to reduce the number of vector components keeping the best features for defect detection and recognition. The well-known k-NN method was used for classification [30]. Defect detection achieved a performance of 96%, whereas performance dropped to 80% in defect recognition. Finally, a grading experiment was done using the non-segmenting method together with two different sets of grading rules (UO and DTU rules). The grading performance was around 72% in both cases. This results were bellow factory requirements of a minimum grading accuracy of 85%.

After this first approach, Niskanen, Silvén and Kauppinen continued the work including texture properties [105,106]. They extended the method using the Local Binary Pattern (LBP) texture operator, previously introduced by their colleagues Ojala and Pietikäinen [108,109] (see Figure 2.4). The original 3x3 neighborhood is thresholded by the value of the center pixel. The values of the pixels in the thresholded neighborhood (Figure 2.4b) are multiplied by the weights given to the corresponding pixels (Figure 2.4c). Finally, the values of the eight pixels are summed to obtain the number of this texture unit.

In this case, for the classification task they chose a neural network based on a Self-Organizing Map (SOM) algorithm which is used to visualize and interpret large high-dimensional data sets by projecting them to a low-dimensional space that has typically one or two dimensions [110,111].

Figure 2.4: Computation of local binary pattern (LBP).

6	5	3	1	1	0	1	2	4	1	2	0
7	5	2	1		0	8		16	8		0
9	3	7	1	0	1	32	64	128	32	0	128
(a)			(b)			(c)			(d)		

$$\mathbf{LBP} = 1+2+8+32+128 = 171$$

For the experiments they used a test material consisting of pine boards. The size of the detection SOM was 22x18 nodes. The feature sets used in the test consisted of colour centiles and LBP values. The region size was 40x40 pixels because they found it gave low error rates for the test material. Defect detection accuracy was 95% whereas defect recognition was 91%.

Kyllönen and Pietikäinen [12] also combined centile and LBP features but they used a different approximation for classification. They combined the centile and LBP features in one measure of distance and then used the k-NN classifier. For the centile features they used the Euclidean distance in the feature space, and for LBP they used a log-likelihood measure to compute the dissimilarity (distance) between sample and reference histograms¹.

$$L(S, R) = - \sum_{n=0}^{N-1} S_n \ln R_n$$

where N is the number of bins. S_n and R_n are the sample and reference probabilities of bin n .

They joined these distances by simply adding them. Prior to this both distances were normalized using:

$$d = \frac{d - d_{min}}{d_{max}} \quad d = \frac{d}{d_{avg}}$$

¹Using LBP there are 2⁸ possible combinations of texture numbers. Thus, texture description of a region can be collected computing its LBP histogram.

where d_{min} , d_{max} and d_{avg} are the min, max and average values of all the distances found in the training set.

For the experiments they used the set of samples collected by Kauppinen [1, 11]. The accuracy for defect detection was 97% and 88.4% for defect recognition, which is an improvement on the results first obtained by Kauppinen (96% and 80%).

We can conclude that the inclusion of texture features, derived from the LBP operator, significantly improved the results of defect recognition but not defect detection. This is to be expected if we take into account that the success ratio using only centiles was very high (96%) for defect detection.

Finally, Niskanen, Kauppinen and Silvén completed these works presenting a study of the real-time aspects of the SOM-based inspection [107]. They focused on the classification stage because they detected it was the bottleneck in real-time inspection. More specifically, they studied how to accelerate the nearest vector search of the SOM code. Some existing methods to accelerate the search were tested; *partial distance search* (PDS), *annulus testing* (AT), *sum of components* (SOC), *dynamical hyperplanes shrinking search* (DHSS), *tree structure SOM* (TS SOM). And also they tested two own methods; *mean tree* (MT) and *focused sparse search* (FSS). It was concluded that the optimal method was the DHSS, which limits the search in the most discriminating direction.

2.5 Other minor works

Peñaranda et al [8, 9], like the CVC group, developed a surface grading application for a specific type of ceramic tile, the *polished porcelanic* tile. A porcelain tile is made up from a mixture of several proportions of grains of different colours and sizes, having the visual appearance of a random texture. Their approach consisted of calculating the histogram of each colour channel in the RGB space. Then, they used the first and second moments of each histogram (average and variance) as colour and texture features respectively. This simple approach with an in-depth study of the inspection system permitted them to fulfil the real-time requirements of on-line inspection. No accuracy results were provided, but, as far as we know,

the system is installed and working at the factory for which it was developed.

Lebrun and Macaire [13] dealt with the surface grading of tiles extracted from the Portuguese marble *Rosa Aurora*. They used four attributes to differentiate between surface classes. The first was the predominant colour which corresponds to the background colour. The mean colour of each channel in the RGB space was used to measure this property. To compute the mean they only used the half of the image pixels located around the largest colour histogram mode. Second attribute was marble vein density. This was obtained by calculating the relative area of veins. To do so, the veins had to be segmented from the background. This was performed by using a simple, automatic threshold operation in the luminance channel, the maximum entropy threshold [112]. The third attribute was vein colour which was represented by the mean colour the veins. The final characteristic involved vein contrast. The contrast was measured calculating the mean gradient in an edge image of veins [113]. They achieved good results classifying nine selected samples correctly. However, they used an approximation so related to the properties of the *Rosa Aurora* marble that is difficult to extrapolate the method to other surface types.

Kukkonen et al [15, 16] tested the use of accurate spectral colour representation to grade ceramic tiles. They used a spectral scan-line camera [114] manufactured by SPECIM (Spectral Imaging Ltd). The spatial scanning resolution was of 97 lines per tile, and the spectral resolution was of 237 channels from the range of 451 nm to 700 nm (a bandwidth of 1.05 nm per channel). A spatial resolution of 330 x 97 pixels was used, having each pixel a rectangle size of 0.85 mm x 3.18 mm. Therefore, the colour representation of each pixel was a vector composed by 237 equidistant samples in the visible light spectrum. For the classification stage they used a neural network called the Self-Organizing Map (SOM) [110]. The spectra vectors of tiles were the input data for the neural network. They did experiments with five classes or grades of a brown tile model. In each class there were five tiles, 25 samples in total. Three tiles of each class were used to train the neural network and the remaining tiles for testing. An accuracy rate of 70% was achieved with this method. They also did experiments using only the RGB mean colour of each tile and the k-NN

classifier. In this case, the percentage of success rose to 90%. The spectral method did not achieve good results. Furthermore, spectral images have the drawback of producing great amounts of data for processing which is not suitable for real-time requirements.

Fernández et al [14] studied the surface grading of granite tiles originating from the *Rosa Porriño* variety located in Galicia. They only used colour information to discriminate between surface classes. The basis of the method were the histograms of each RGB channel and a simple measure of histogram similarity; the sum of absolute differences of each bin-pairs. They carried out some experiments with a small set of samples (only six tiles with three surface classes). No accuracy information was provided but in the paper they admit that texture information would be necessary to improve the results.

2.6 Conclusions

Many works on the issue of surface grading have been reported in recent years, but many of them were very specialized in a specific type of surface, others did not achieve good enough accuracy, and yet others did not take into account the time restrictions of a real inspection at factory. As a result, we think surface grading is still an open issue where more contributions are possible. In this sense, the present thesis deals with some less explored aspects in terms of real-time compliance and surface grading performance.

From the literature review we can deduce that there are no extensive experiments of grading performance in the area of ceramic tiles. Only for a specific kind of ceramic tile, the polished porcelanic tile, have there been extensive studies of grading performance (CVC group and Peñaranda). There is only one work dealing with generic surfaces (Surrey group) but they used only surface colour property, and no accuracy study was given.

The Oulu group carried out a large work in the area of wood inspection, but this work focuses more on separating good and faulty wood areas than on accomplishing the grading task. Grading results are not sufficiently good. Other minor works deal

with very specific types of surfaces such as *Rosa Aurora* marble (Lebrun) and *Rosa Porriño* granite (Fernández).

We can also see that there is a lack in the literature of real-time compliance in the methods. Only two works pay attention to the time requirements of factory production lines (Peñaranda and Oulu group).

In our work, we focus on the ceramic tile industry where there is a large demand for automatic grading. As far we know, we present the most extensive study of surface grading performance in the area of ceramic tiles. We use the VxC TSG image database which is a wide representation of a typical factory catalog with many types of surfaces, such as imitation marble, imitation granite, and stone. Both, texture and colour properties are used to successfully discriminate surface grades. In addition, we present an in-depth study of real-time compliance. The real-time approach is based on the use of features with low computational cost and parallel processing techniques.

Chapter 3

VxC TSG image database

In this chapter we present the VxC TSG image database (VxC Tiles for Surface Grading). Building this database has been one important goal of the present thesis. The VxC TSG is based on samples taken from the ceramic tile industry and is comprised of 14 ceramic tile models, 42 surface grades and 960 pieces. It was built in the VxC laboratory in collaboration with Keraben S.A. and is an extensive image database of ceramic tiles representing the wide range of surface classes in the ceramic tile industry. VxC TSG is the ground truth used in the experiments of Chapters 5 and 6, and is also intended to be a tool for the scientific community working on surface grading. It is public and available at miron.disca.upv.es/vision/vxctsg/.

Before describing the image database itself, we describe the acquisition system used to capture the digital images of tiles. We also present a study of the uniform response of the system through time and space. This study is a subgoal of thesis work. Spatial and temporal uniformity are of great importance in order to ensure surface grading performance [2–5, 9, 83]. Slight changes in illumination or acquisition conditions can easily produce different grades for the same surface and then misclassifications. In order to overcome this problem we chose high quality components for the acquisition system; camera, illumination, and optics. In the literature many of the modern components were not available and system variability had to be compensated using data transformation algorithms. Our goal in this issue has been to demonstrate that modern acquisition components are able to meet spatial and temporal requirements without needing any transformation of the original data.

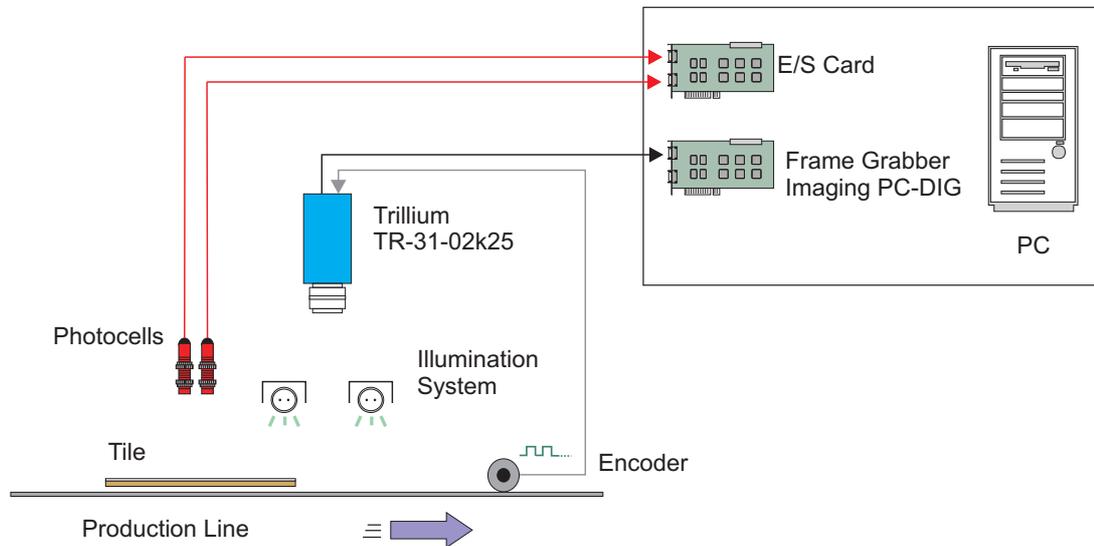
The study of spatial and temporal uniformity was carried out comparing two modern illumination systems; uniform high frequency fluorescents and arrays of white LEDs.

3.1 Acquisition system

The acquisition system (see Figure 3.1) compromises the following high quality components:

- One Dalsa Trillium camera (TR-31-02k25). This is a colour scan line camera with 3 CCDs (RGB). It provides 2048 pixels of horizontal resolution and a maximum acquisition rate of 11kHz. This acquisition rate is more than we need in worst factory conditions. As we use a resolution of 3.2 pixels per millimetre only 884 lines per second are needed (see Chapter 7).
- One Nikon optics (35mm, 1:2.0 mm).
- One Coreco-Imaging PC-DIG frame grabber with 4Mb of internal RAM and 100MB/s of PCI transfer rate. We need only 111.3 milliseconds to transfer an image from the camera to the PC memory in the worst case (2048x1900 RGB images).
- High frequency and uniform fluorescents (Mercrom FXC2372-2). This illumination system has two special high frequency fluorescent lamps (60kHz) with uniform illuminance throughout its length. To overcoming variations with time, the power supply is automatically regulated by a photoresistor located near the fluorescents. The high frequency provides 135.8 luminance peaks for each scanned line, thus, dark acquisitions are not possible.
- Alternatively, another illumination system formed by two arrays of white LEDs (DCM Sistemas PRL 350). A priori, LEDs are supposed to be uniform in time because they use constant DC power, and also they are supposed to be spatially uniform as they are arranged in line equidistantly.

Figure 3.1: Acquisition system.



In the VxC laboratory there is a prototype system which is a replica of a section of a typical ceramic tile production line. The prototype comprises a cabin integrated in the production line as the acquisition unit (see Figure 3.2) and one PC with the frame grabber and the I/O card as the processing unit. This prototype with the above mentioned components has been used to capture the digital images for the VxC database.

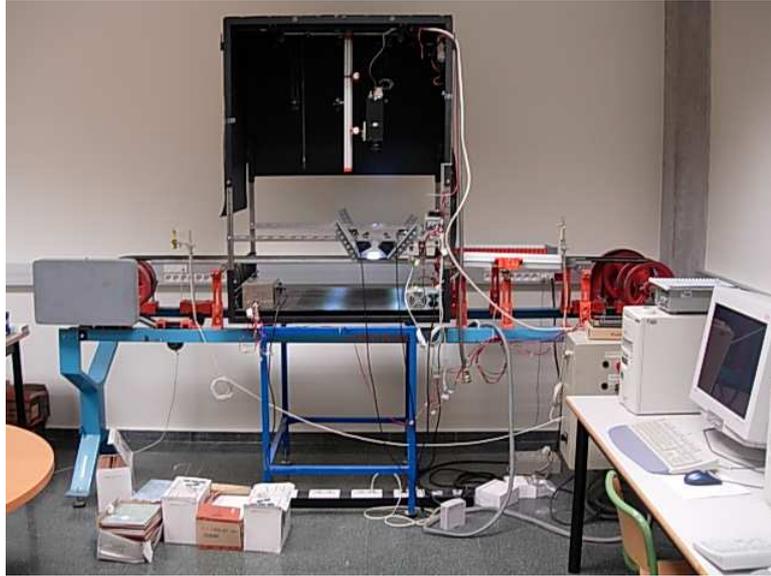
3.1.1 Scan Line Camera

We chose a scan line camera with advanced properties to provide high quality acquisition and overcome some common acquisition problems.

It has a 3 CCD sensor with a precisely-aligned beam-splitting prism to separate red, green and blue inputs. Colour sensitivity is much better than using 1 CCD camera because in 1 CCD cameras the responses on red, green and blue channels are mixed in the same CCD.

The camera is able to calibrate itself to improve colour balance and image flatness. It performs a video correction that operates on a pixel-by-pixel basis and implements a two point correction for each pixel. This correction reduces or eliminates image distortion caused by the following factors:

Figure 3.2: Prototype at VxC laboratory.



- Fixed Pattern Noise (FPN).
- Photo Response Non Uniformity (PRNU).
- Colour imbalance.
- Lens and light source non-uniformity.

The two point correction is implemented for each pixel on the CCD using:

$$V_{output} = PRNU(pixel) * V_{input} + FPN(pixel)$$

where V_{output} is the output pixel value, V_{input} is the input pixel value from the CDD, $PRNU(pixel)$ is the PRNU correction coefficient for this pixel and $FPN(pixel)$ is the FPN coefficient for this pixel.

The calibration algorithm is performed in two steps. The fixed offset (FPN) is determined first by performing a calibration with no light (Dark Calibration). This calibration determines how much offset to subtract per pixel in order to obtain flat output when the CDD is not exposed. The Dark Calibration is carried out covering the lenses or/and closing the iris to the maximum.

White Light Calibration is performed next to determine the multiplication factors required to bring each pixel to the required value (balance target, usually 95% of saturation) for flat white output. The white light calibration also sets the analog gains in the camera appropriately to balance all the channels (colour balance). White light calibration is more complex than dark calibration because the camera attempts to create a flat white image. This calibration corrects PRNU effects as well as non-uniform lighting and lens vignetting affects.

White light calibration requires a clean, white reference. The quality of the reference is important for proper calibration. White paper is often not sufficient because the grain in the white paper will distort the correction. Usually a white plastic reference, which is more uniform, achieves better balancing.

3.1.2 Optical lenses

The camera was equipped with a Nikon optics (35mm, 1:2.0 mm). This is a common high quality lens, but as with most lenses, it can be affected by two major optical colour aberrations; *vignetting* and *chromatic aberration* [115]. At present, some companies can supply aberration-free lenses but they are made on demand, do not completely remove the aberrations and are very expensive.

Vinegting is an unintended darkening of the image corners and is inherent to the lens design. Chromatic aberration appears because common lenses refract light differentially as a function of wavelength. Short (blue appearing) wavelengths are refracted more than long (red appearing) wavelengths. Thus, chromatic aberrations are introduced.

Some tests at the laboratory, carried out to prepare and calibrate the acquisition system, showed that the vignetting affect grewed as we used large openings of the iris. We achieved images free of vignetting by using small openings in combination with the camera calibration process. With respect to chromatic aberration, it is inherent to the lenses and is not time or camera dependent. Thus, the chromatic aberration introduced is constant in all acquisitions and therefore is not relevant when we compare colour differences rather than absolute colour values, as it occurs in surface grading application.

3.2 Study of spatial and temporal uniformity.

Spatial and temporal uniformity is crucial to ensuring surface grading performance. Slight changes in illumination or acquisition conditions can easily introduce different grades for the same surface and then misclassifications. In the surface grading literature this question has been addressed compensating the system variability with data transformation algorithms.

In [2,3] Boukouvalas et al captured a set of images of the same plain tile in all four possible orientations. From these images they determined the spatial variation of the illumination by averaging the four images and fitting the data with low-order two-dimensional polynomial. The coefficients of this polynomial were computed using *least square error fitting*. Temporal variability of the illumination was determined by capturing a sequence of images of the same plain tile next to a reference surface over a period of time. From these images a set of points (I_R, I_T) was collected representing the mean intensities of the reference surface and the tile, respectively. The intensity change of the tile and the reference surface was locally described by a *linear function*, the slope of this function was computed with *least square error fitting* again and used to overcome temporal variability in the illumination.

In [4, 5] Baldrich et al modeled the global variability of the acquisition system including illumination and sensor affects. They developed a method based on colour constancy techniques using a diagonal matrix model. This was computed using a white pattern image acquired periodically. Spatial distortions were modeled using a set of diagonal transforms $\{S_x\}$, one 3x3 matrix for each position x along the x axis where the spatial variation occurs (they used a scan-line camera). Light variations due to time were corrected in a similar way. A set of diagonal transforms were calculated $\{T_x^{t_i}\}$. This set modeled the distortions at time t_i referring to instant t_0 . The final set of diagonal transforms was $\{D_x^{t_i}\}$ where $D_x^{t_i} = S_x T_x^{t_i}$.

When these works were reported many of the modern acquisition components were not available and system variability had to be compensated using data transformation methods. Our approach to this question has been to demonstrate that modern acquisition components are sufficiently stable to meet spatial and temporal uniformity requirements without transforming the original data.

We carried out an experiment to determine the reliability of the acquisition system in relation to spatial and temporal uniformity. This experiment also compared two different modern illumination systems; uniform high frequency fluorescents and arrays of white LEDs.

For each illumination system we captured repeatedly the images of six tiles, each one corresponding to a different model. The tiles were chosen trying to cover a wide range of surface types and colours (see Figure 3.3). The complete set of tiles was acquired at random moments over 54 hours. We extended the experiment over 54 hours (two days and six hours) because this is the mean period at factories when they produce a specific model, and we wanted to study the spatial and temporal uniformity for a complete surface grading session. In total, the set of tiles was captured 23 times. Environmental conditions were holded constant using an air conditioner system for temperature and a closed cabin for illumination.

Figure 3.3: Tiles used in the study of spatial and temporal uniformity. From left to right, up to down; venice, vega, blue venice, somport, mediterranea and granito.



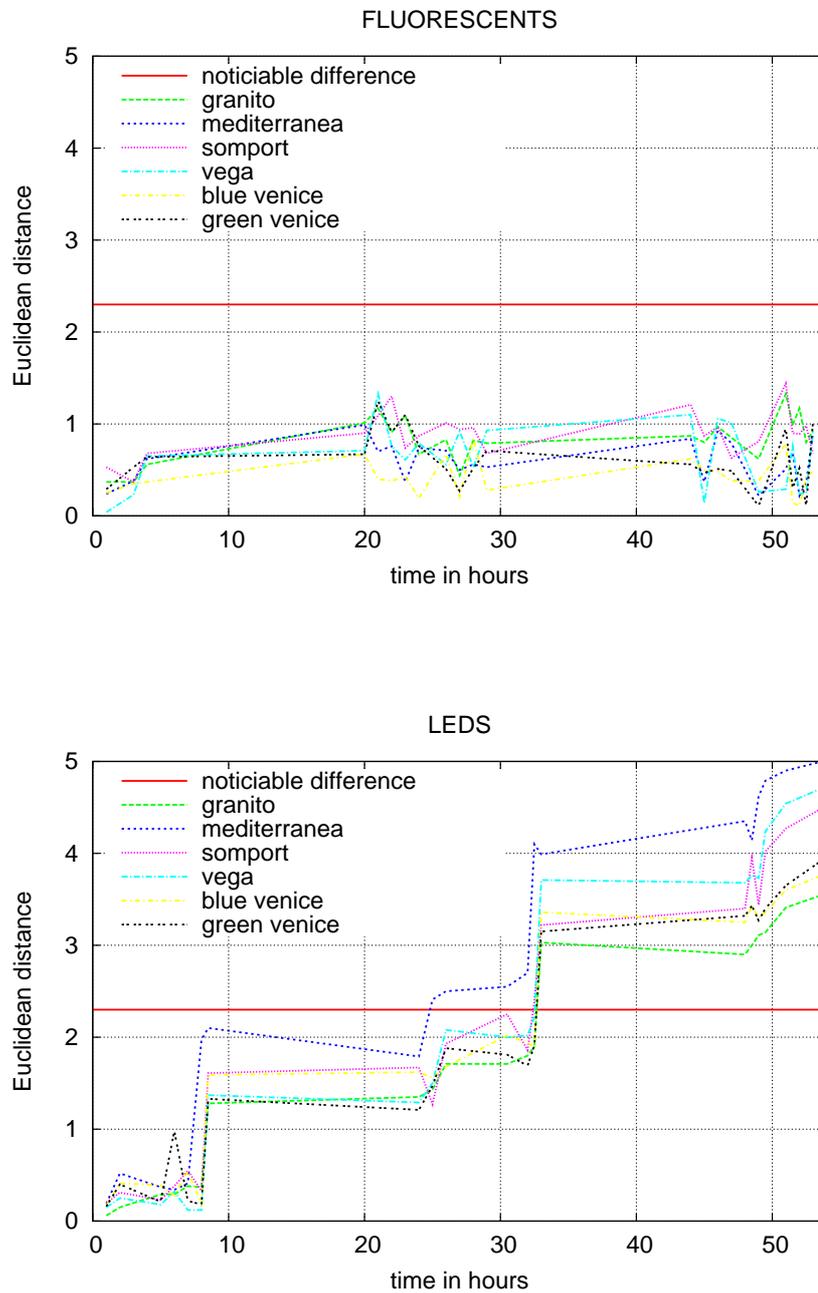
In order to study the temporal response we measured the mean CIE Lab colour of each piece. And also, in order to study spatial response we randomly oriented

the pieces in each capture. The CIE Lab is a perceptually uniform colour space and we can measure the perceptual difference between two colours using the Euclidean distance in this space [33]. Thus, colour differences can be measured in a very similar way to the human perception of colours.

In [116] Mahy and Oosterlink established that in CIE Lab a noticeable difference of colour [for humans] begin at 2.3 or greater Euclidean distances. From this assertion we can consider a system sufficiently stable if there is no Euclidean distance above 2.3 when we calculate all the Euclidean distances between the first sample and the rest. Figure 3.4 shows the system response for each tile over the 54 hours when using fluorescents and LEDs respectively. In the results of fluorescents there was no distance above 2.3, and all of them remain significantly far away from this limit. Distances using LEDs did not remain under the *noticeable difference* showing a clear degradation of the system with time, the *noticeable difference* was surpassed after approximately 33 hours. LEDs experiment was repeated using a better performance power supply but again the *noticeable difference* was exceeded after approximately 33 hours.

The goal of this study has been to determine whether or not the acquisition system is stable enough for the surface grading purpose. The conclusion of the study is that uniform high frequency fluorescents comply with spatial a temporal uniformity, while the arrays of white LEDs are not appropriate when temporal uniformity is required. A great part of the success of fluorescents is due to the power supply circuit which is auto-regulated taking in account emitted light by using a photoresistor located near the fluorescents. Thus, the system can respond immediately to illumination changes and stabilize luminous power with time.

Figure 3.4: System response over 54 hours using fluorescents and LEDs respectively.



3.3 VxC TSG description.

Table 3.1: Models of the VxC TSG image database.

	classes	tiles/class	size (cm)	pattern	aspect
agata	13, 37, 38	16	33x33	fixed	marble
antique	4, 5, 8	14	23x33	pseudo-random	stone
berlin	2, 3, 11	24	20x20	random	granite
campinya	8, 9, 25	30	20x20	pseudo-random	stone
firenze	9, 14, 16	20	20x25	random	stone
lima	1, 4, 17	24	20x20	random	granite
marfil	27, 32, 33	14	23x33	pseudo-random	marble
mediterranea	1, 2, 7	30	20x20	random	stone
oslo	2, 3, 7	24	20x20	random	granite
petra	7, 9, 10	28	16x16	random	stone
santiago	22, 24, 25	28	19x19	random	stone
somport	34, 35, 38	28	19x19	random	stone
vega	30, 31, 37	20	20x25	fixed	marble
venice	12, 17, 18	20	20x25	pseudo-random	marble

The image database has been built in collaboration with Keraben S.A. which is a large ceramic tile company (2nd in Spain by total turnover) located at Nules, province of Castellón. Together with the R+D staff we studied their catalog and chose a set of models intended to be a good representation of the wide variety of surface classes that factories can produce. A catalog of 700 models is common in these companies. But, in spite of this great number of models, almost all of them imitate one of the following mineral textures; marble, granite or stone.

Due to the way that tiles are produced there are three basic types of surface patterns; fixed, random and pseudo-random. Fixed pattern models are produced by impressing the fix decorative motives on to the enameled tile surfaces. This is done through silk-screen machines. Printing rollers are used to made pseudo-random models. The patterns on the rollers are fixed but the part of each roller that is impressed over the tile is randomly chosen by the moment that tile begin to pass under the rollers. For random models several techniques may be used depending on the model to be produce. One of these techniques consists of spreading pigments over the tile using sponges. From the point of view of surface grading purposes fixed

and pseudo-random pattern models can be considered subsets of random pattern models.

The database is formed by the digital images of 960 tiles acquired from 14 different models, each one with three different surface classes (see Table 3.1). The classes were given by specialized graders at the factory. Every model has two close classes and one class distant to them. Surface classes (grades) are represented by numbers and close numbers mean close classes. Thus, the database include “difficult to discriminate” cases in each model. The tiles were collected directly from the end of the production line, just before the tiles are graded and packed, or from stock stored at the factory.

Figures 3.5, 3.6, 3.7 and 3.8 show VxC TSG samples itemized by aspect. Quality of images was improved enhancing brightness and contrast independently in every model.

In the case of granite models there were only six tiles per surface grade. The size of these pieces was large, 50x50cm. To increase the number of samples in these models we decided to take four sub-samples of each piece with a size of 20x20cm. This could be done because grains and tile colour were uniformly distributed through all the surface of tiles.

Figure 3.5: VxC TSG marble samples. From up to down; three samples of agata, marfil, venice and vega models, each one corresponding to a different surface grade.



Figure 3.6: VxC TSG granite samples. From up to down; three samples of berlin, oslo and lima models, each one corresponding to a different surface grade.



Figure 3.7: VxC TSG stone samples. From up to down; three samples of antique, campinya, firenze and mediterranea models, each one corresponding to a different surface grade.

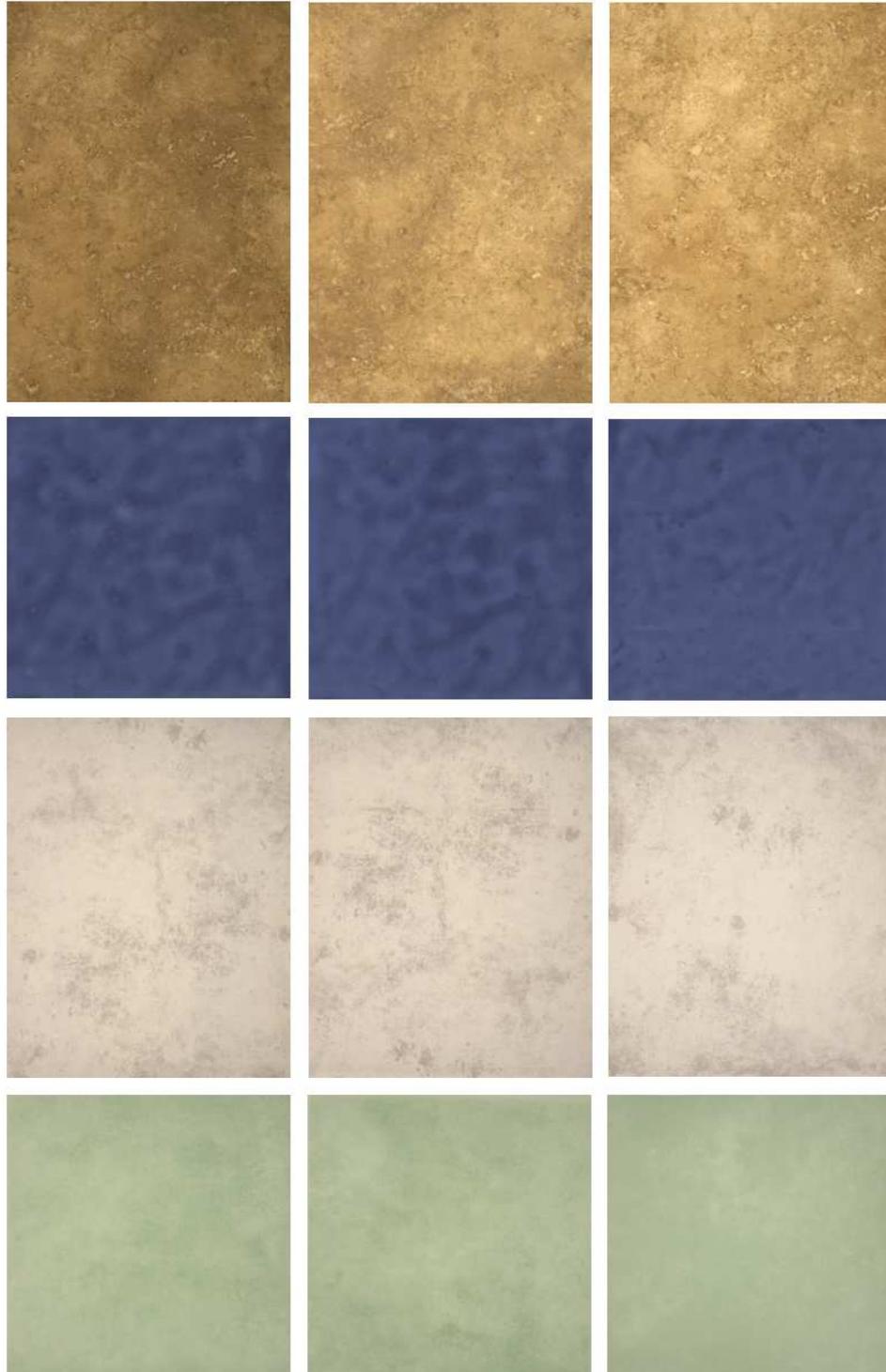


Figure 3.8: VxC TSG stone samples. From up to down; three samples of petra, santiago and somport models, each one corresponding to a different surface grade.



3.4 Conclusions

In this chapter we have presented the VxC TSG database for surface grading which has been one important goal of this thesis. The acquisition system, based on high quality components, has been described and also a study about the spatial and temporal uniformity of the system is performed. This study has been a thesis subgoal planned at the beginning of the thesis work. Uniform high frequency fluorescents and arrays of white LEDs are two modern illumination systems that have been compared from the point of view of spatial and temporal uniformity. The conclusion of the study is that the acquisition system using uniform high frequency fluorescents

comply with spatial a temporal uniformity, while the arrays of white LEDs do not provide temporal uniformity. LEDs clearly degrade the system with time surpassing the *noticeable difference* limit after 33 hours of use.

Finally, and extensive image database of ceramic tiles for the purpose of surface grading has been compiled. This database has been built in collaboration with the R+D staff of Keraben S.A. intending to be representative of the wide range of surface classes present in ceramic tile industry. Furthermore, the database is available for the scientific community working on surface grading at miron.disca.upv.es/vision/vxctsg/. The VxC database has already been used partially in [130].

Chapter 4

On the search for a fast and accurate approach to surface grading

This chapter presents the first approaches we developed to solve the question of fast and reliable surface grading of flat pieces decorated with random patterns. In the first works, we used image tessellation and simple local statistics of colour to describe surface appearance. The statistics were computed in a perceptually uniform colour space, the CIE Lab. These first works did not achieve the minimum accuracy requested at factory (95% of success ratio). Finally, we proposed a method based on global colour and texture statistics, also computed in CIE Lab. This method achieved accuracy compliance. CIE Lab was used to provide accuracy and perceptual approach in colour difference computation. Experiments with RGB were carried out to study CIE Lab reliability. These approaches were tested on a medium sized image database of ceramic tiles. This database was the antecedent of the VxC TSG image database presented in the previous chapter. Global statistics in CIE Lab were also compared with two other methods from the literature; colour histograms [2, 3] and centile-LBP [1, 12].

After experiments and comparison we concluded that a simple collection of global colour and texture statistics in the CIE Lab space was powerful enough to well discriminate surface grades. The average success rate was over 95% in most tests, improving on the methods in the literature and achieving factory compliance. The approach based on global statistics in CIE Lab is the antecedent of the *soft colour-*

texture descriptors method developed in the next chapter. Both global statistics in CIE Lab and *soft colour-texture descriptors* are basically the same method but the latter was extracted as a result of an extensive study based on statistical tools and VxC TSG image database. Work presented in this chapter was reported in [25, 26].

4.1 Image tessellation and local Lab statistics

The methods corresponding to this section split the image into squares of $N \times N$ pixels. For each square two simple statistics, the mean and the standard deviation, are computed in each CIE Lab colour space channel. As we will see, this local data is used in several ways to perform the surface grading.

CIE Lab was designed to be perceptually uniform. The term 'perceptual' is referred to the way that humans perceive colours, and 'uniform' implies that perceptual difference between two coordinates (two colours) will be related to a measure of distance, which commonly is the Euclidean distance. Thus, colour differences can be measured in a way close to the human perception of colours.

The images of the ground truth (image data base) were acquired in RGB, and therefore needed to be converted to CIE Lab coordinates using standard RGB to CIE Lab transformation [33].

The experiments in this section we carried out using an image data base formed by the digital RGB images of 276 tiles acquired from five different models, each with three different surface classes or grades (see Table 4.1) given by specialized graders at factory. For each model there were two close classes and one distant class. The models were chosen to represent the great variety of models that factories can produce. Almost all factory models imitate one of the following mineral textures; marble, granite or stone. In this initial image data base there were no models imitating granite. This type of tiles were added latter in the global Lab statistics experiments.

Digital images of tiles were acquired using an spatially and temporally uniform illumination system. Spatial and temporal uniformity is important in surface grading [1,2,4,8] because variations on illumination can produce different shades for the same

Table 4.1: Ground truth of ceramic tiles used in image tessellation and local Lab statistics approaches.

	classes	tiles/class	size (cm)	pattern	aspect
agata	13, 37, 38	16	33x33	fixed	marble
firenze	9, 14, 16	20	20x25	random	stone
toscana	13, 18, 19	16	33x33	random	stone
vega	30, 31, 37	20	20x25	fixed	marble
venice	12, 17, 18	20	20x25	pseudo-random	marble

surface and consequently, misclassification. The illumination system was formed by two special high frequency fluorescent lamps with uniform illumination along its length. To overcome variations through time, the power supply was automatically regulated by a photoresistor located near the fluorescents.

The first approach based on tessellation and local statistics uses the standard deviation to sort squares from low to high variability. Then, a simple algorithm seeks a slope exceeding a given threshold in the sorted vector of standard deviations. The image is divided into two regions defined by squares variability; low and high textured regions (see Figure 4.1). Then, the mean colour vector of both regions are computed and used to classify the tiles. The hypothesis is that tile surface can be divided into two general regions, one with an homogeneous aspect, and another with a textured aspect. Each region seems to have a different general colour and these two colours could be enough to characterize the tile tone or grade.

The classification results are presented in Table 4.2. We used the well known k-NN classifier [30] with k factor equal to 1 and 3. Samples were divided into training and test sets, 30% of samples were used for training and 70% for test. The results show better performance for CIE Lab space. The method does not achieve accuracy compliance and also it requires two non-automatic parameters; the standard deviation threshold and the square size. Table 4.2 only shows the estimated values of standard deviation, but the $N \times N$ size also had to be estimated. This parameters were heuristically studied for each tile model. This non-automatic procedure in parameter estimation is an important drawback.

A post-study of images and square sizes revealed that in several tile models more

Figure 4.1: First approach to surface grading based on tessellation and local statistics. Sorted vector of standard deviations and image splitting into two regions using squares variability.

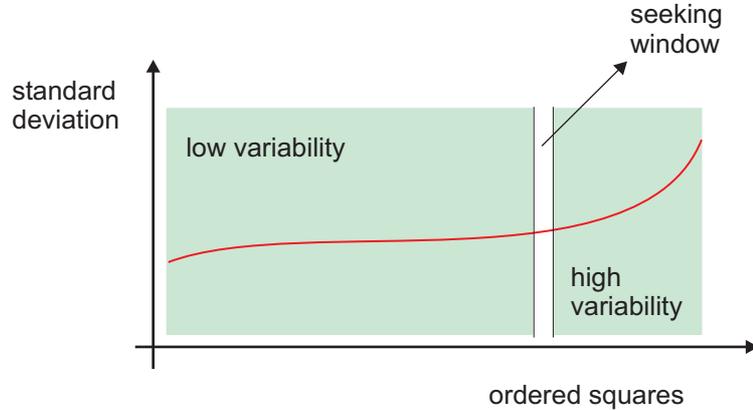


Table 4.2: Accuracy results for the first approach based on image tessellation and local statistics.

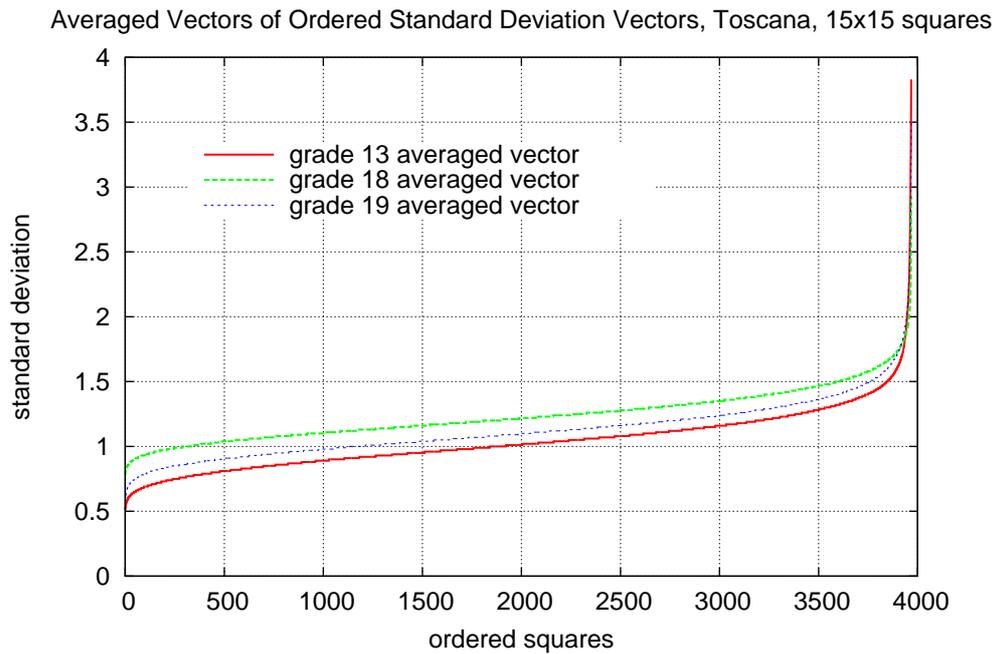
	std deviation	Lab	RGB
agata	0.5	97.0%	87.9%
firenze	0.5	85.7%	78.6%
toscana	1.0	75.8%	66.7%
vega	0.6	90.2%	100%
venice	1.0	88.1%	78.6%
mean		87.4%	82.4%

than two regions and mean colours could be defined in relation to colour variability of squares. Thus, the method could be extended dividing images into more regions using more standard deviation thresholds, but we thought it would be complex and not interesting because parameters (standard deviation thresholds and squares size) have to be determined non-automatically.

In a second approximation, we explored the discriminant properties of the ordered standard deviation vector. Figure 4.2 shows the averaged vectors of the ordered standard deviation vectors belonging to each class (grade) of the toscana model using 15x15 squares. Each mean vector was computed using the half of the samples of each class. Classification was performed using the square root of the *least-square-error* (LSE) as a measure of distance between test tile vector and the

reference mean vector of each class. Thus, the sorted standard deviation vector is used as a signature of local variability. The distances between references and test signatures are used to determine the surface grade or class.

Figure 4.2: Second approach to surface grading based on tessellation and local statistics. Averaged vectors of ordered standard deviation vectors corresponding to each toscana classes with a square size of 15x15 pixels.



This method is simpler than previous one because no threshold parameter is needed for the standard deviation, only the squares resolution had to be studied. This method improves the first approach but still does not reach accuracy compliance. Tests using squares resolutions from 5x5 to 300x300 were performed. Table 4.3 collects the best accuracy results among the different square resolutions for each model. RGB obtains the worst results.

An evolution of this second approach is carried out using the mean colour of squares instead of the standard deviation to compute the distance between signatures. The vector is again sorted by the standard deviation but values correspond to the mean colour of squares. In this case, we used pure colour information plus the variability represented by the order imposed with the standard deviation sorting.

Table 4.3: Accuracy results for the second approach based on image tessellation and local statistics.

	square size	Lab	RGB
agata	80x80	95.8%	91.7%
firenze	30x30	86.7%	76.7%
toscana	30x30	83.3%	79.2%
vega	70x70	100%	100%
venice	25x25	83.3%	86.7%
mean		89.8%	86.9%

The results of this third approach are shown in Table 4.4. The method achieves very good performance for three models but it drops significantly in the remaining two models. After studying the visual properties of these groups of models, we concluded that the fundamental difference was the variability level. Tiles are formed by homogeneous and texturized (non-homogeneous) regions. In the first group of models, homogeneous regions cover great areas of tiles while in the second group texturized areas cover almost all the tile surface. In the second group, the standard deviation values did not correlate with the mean colours of squares while they did in the first group. Once again RGB space performed worse than CIE Lab.

Table 4.4: Accuracy results for the third approach based on image tessellation and local statistics.

	Lab	RGB
agata	100%	100%
firenze	100%	96.7%
toscana	58.3%	54.2%
vega	100%	40%
venice	66.7%	76.7%
mean	85.0%	73.5%

Finally, we developed several multiresolution approaches. The idea was to use some kind of multiresolution approach with the methods previously studied. The image scales we used were: 1.0, 0.5, 0.25, 0.12 and 0.06. The first multiresolution approximation is quite simple, it is a voting system. Each sample is classified in every scale, and the final class is the most voted class out of all the scales. This

approach is carried out using second and third previous approaches.

The second multiresolution approximation is an extension of the approach based on the ordered standard deviation vector. We used an absolute measure, the area of the ordered standard deviation vector (see Figure 4.3). This measure was computed in all scales, and therefore a feature vector of 5 areas was obtained for each sample. Classification was performed using k-NN. Experiments were carried out using only CIE Lab.

Figure 4.3: Second multiresolution approach based on the computation of areas belonging to ordered standard deviation vectors in several image scales.

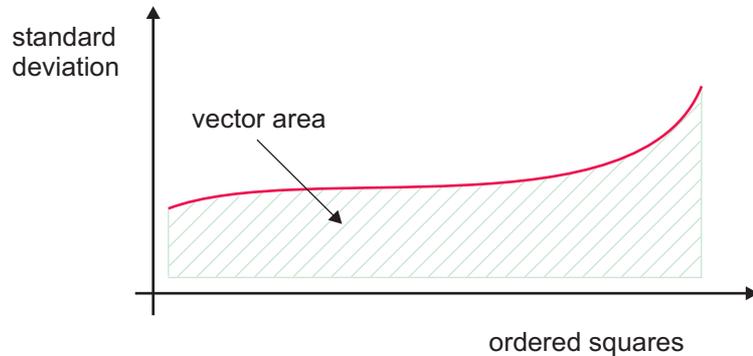


Table 4.5: Lab accuracy results for the fourth approach based on image tessellation and local statistics. Multiresolution approaches.

	standard deviation	mean colour	std dev. areas
agata	100%	100%	100%
firenze	90.0%	100%	90.0%
toscana	83.3%	58.3%	95.8%
vega	100%	100%	96.6%
venice	86.7%	66.7%	93.3%
mean	92.0%	85.0%	95.1%

Table 4.5 shows multiresolution results. The first multiresolution approach achieves only a slight improvement using the ordered standard deviation vector. However, the second multiresolution approach based on the areas of the ordered standard deviation vectors performs well exceeding in 0.1% factory compliance (95%).

In spite of the good result achieved by the second multiresolution approach, the drawback of these methods is they need at least one parameter to be estimated for every tile model. In next section we present a method which is parameter independent and discriminates surface grades very well.

4.2 Global Lab statistics

This method is even simpler than the previous approaches. Here, a set of global statistical features describing colour and soft texture properties are collected. The features are computed in a perceptually uniform colour space, the CIE Lab. These statistics form a feature vector used in the classification stage where the well known k-NN method [30] was chosen as classifier.

We propose several statistical features for describing surface appearance. For each channel we chose the mean, the standard deviation and the average deviation $ADev(z)$.

$$ADev(z) = \frac{1}{L} \sum_{i=1}^L |z_i - m|$$

where z is the random variable, L size of the data set and m the mean value of z values.

Also, by computing the histogram of each channel we are able to calculate histogram moments. Histogram moments are considered soft texture descriptors [29]. We defined two blocks of histogram moments; one from 2nd to 5th and the other from 6th to 10th. The n th moment of z about the mean is defined as

$$\mu_n(z) = \sum_{i=1}^L (z_i - m)^n p(z_i)$$

where z is the random variable, $p(z_i)$, $i = 1, 2, \dots, L$ the histogram, L the number of different variable values and m the mean value of z .

The experiments were carried out using and extended image data set which included three new tile models for polished porcelanic tiles were added (see Table 4.6).

Table 4.6: New models added to the image database of ceramic tiles.

	classes	tiles/class	size (cm)	pattern	aspect
berlin	2 ,3, 11	24	20x20	random	granite
lima	1, 7, 17	24	20x20	random	granite
oslo	2, 3, 7	24	20x20	random	granite

The experiments were carried out for CIE Lab and RGB spaces. Classification was made using half of the samples as training set and the remaining half as test set. Values of 1, 3, 5 and 7 were used for the k factor of the k-NN classifier.

The performance results of several statistics sets are shown in Table 4.7. The accuracy rates were computed as the average accuracy ratios achieved over all models. More combinations of statistics were tested, but only the most prominent are presented. The last two columns corresponds to the averaged accuracy rate and the 95% confidence intervals [122] respectively. The table is divided into two blocks, the first one corresponds with CIE Lab experiments. Here, the majority of sets have confidence intervals under the minimum accuracy rate of 95% which is the factory performance requirement. The best choice was to use the mean colour plus the standard deviation. Histogram moments did not introduce any improvement. The second block collects the results of RGB which presents significant less discriminative power than CIE Lab.

Table 4.7: Best accuracy results of global Lab statistics method. From left to right; mean, standard deviation, average deviation, histogram moments from 2nd to 5th, histogram moments from 6th to 10th, CIE Lab, RGB, accuracy and 95% confidence intervals. Accuracy and confidence intervals are given in %.

mean	std dev	ave dev	2-5 ms	6-10 ms	Lab	RGB	acc.	c.i. 95%
x					x		86.8	[83.6, 89.7]
x	x				x		98.9	[97.7, 99.6]
x		x			x		97.0	[95.3, 98.4]
x	x		x		x		96.8	[95.1, 98.3]
x	x		x	x	x		96.7	[94.8, 98.1]
x						x	86.6	[83.4, 89.6]
x	x					x	92.1	[89.4, 94.3]
x		x				x	92.7	[90.1, 94.9]
x	x		x			x	94.1	[91.7, 96.0]
x	x		x	x		x	93.3	[90.8, 95.4]

4.3 Literature methods

We selected two methods from the literature for comparison purposes: colour histograms [2, 3] and centile-LBP [1, 12]. We chose these methods because they are similar to ours, both are generic solutions with low computational costs. For in-depth information about these approaches review Chapters 2 and 6.

We should point out that the centile-LBP method is not used in literature directly for wood grading but for surface segmentation into sound wood and knocks. However, we use the method as a global surface grader achieving good results.

Colour histograms are 3D histograms which are compared using dissimilarity measures. In [2,3] they used the *chi square test* and the *linear correlation coefficient* to measure histograms dissimilarities.

Centiles [1,12] are colour features calculated from the cumulative histograms of colour channels corresponding to a given colour space. A total number of 171 centiles are compiled to describe the colour property of surfaces. The Local Binary Pattern (LBP) is a local texture operator related to each image pixel's neighbourhood. This operator provides a number for each pixel (texture unit) in the range [0, 255], then a histogram collects the LBP texture description of an image.

In [12] centile and LBP features were combined in one measure of distance and then the k-NN classifier was used. For Centile features they used the Euclidean distance in the feature space. For LBP they used a log-likelihood measure. Both distances were normalized using the min and max values of all the distances found in the training set and then joined by simply adding them together.

Experiments for colour histograms and centile-LBP were carried out. Once again, classification was made using the half of the samples for training and the remaining half for testing. In centile-LBP experiments the original *log-likelihood* formula, the *chi square test* and the *linear correlation coefficient* were used for measuring histograms differences.

Table 4.8: Accuracy results of colour histograms and centile-LBP. From left to right; Chi square, linear correlation and log-likelihood distances, accuracy, 95% confidence intervals. Accuracy and confidence intervals are given in %.

	Chi	Corr.	Log	Lab	RGB	acc.	c.i. 95%
Colour Histo.	x			x		90.3	[87.4, 92.8]
Colour Histo.		x		x		88.5	[85.4, 91.2]
Colour Histo.	x				x	88.9	[85.8, 91.5]
Colour Histo.		x			x	87.6	[84.5, 90.5]
Centile-LBP	x			x		94.4	[92.2, 96.4]
Centile-LBP		x		x		94.9	[92.6, 96.7]
Centile-LBP			x	x		91.3	[88.5, 93.6]
Centile-LBP	x				x	94.7	[92.4, 96.5]
Centile-LBP		x			x	95.4	[93.4, 97.2]
Centile-LBP			x		x	93.3	[90.8, 95.4]

The results of Table 4.8 show that centile-LBP achieves the best accuracy rates when using RGB, but neither method achieves factory compliance because none of their confidence intervals are over the min accuracy rate (95%) requested at factory. Compared with Table 4.7 Lab Statistics presents significant improvement in performance and also is the only method with confidence intervals completely surpassing the min factory accuracy limit.

4.4 Conclusions

In this chapter we have presented several approaches to surface grading. First methods were based on image splitting into equally sized squares and simple local statistics computed for each square. Only one of these methods achieved the accuracy compliance requested at factory. An important drawback of all these methods is that they needed to non-automatically estimate at least one parameter for every tile model. Next, a method based on global statistics with no parameter estimation was presented. Many combinations of the proposed statistics on colour and texture achieved accuracy compliance clearly exceeding the minimum requested at factory (95%). All the proposed methods use a perceptually uniform colour space, the CIE Lab. Experiments using RGB were performed and this space showed less discriminative power. Thus, perceptual approach based on CIE Lab appears as a good choice.

For comparison purposes we selected two methods from the literature (colour histograms and centile-LBP) and performed experiments using the same image database. The results achieved by both approaches were worst than global Lab statistics and also did not reach factory compliance.

Global statistics computed in CIE Lab is the antecedent of *soft colour-texture descriptors* method which is in-depth developed in next chapter. Both, global statistics in CIE Lab and *soft colour-texture descriptors* are basically the same method. Once we found a fast method able to comply with factory requests, we decided to study the approach in-depth, adding new factors such as new colour spaces (CIE Luv and Grey scale), classifiers (leaving-one-out) and testing all the possible combinations of soft colour-texture descriptors (mean, standard deviation, and histogram moments from 2nd to 5th). To do so, we used statistical tools to manage the large amount of resulting experiments and achieve objective and valid conclusions. We also used an extensive image database, the VxC TSG.

Chapter 5

Extraction of *soft colour-texture descriptors* method

In previous chapter we found a successful new approach to surface grading based on global statistics of colour and texture computed in a perceptually uniform colour space, the CIE Lab. This approach achieved factory compliance in accuracy performance and also improved literature methods (colour histograms and centile-LBP). Basing on this method in this chapter we present an extensive study in order to extract a definitive method. The study includes more factors such as new colour spaces (CIE Luv and Grey scale) and classifiers (leaving-one-out) and also all the possible combinations of soft colour-texture descriptors (mean, standard deviation, and histogram moments from 2nd to 5th) are tested in conjunction with the other factors.

The resulting method from this in-depth study is named *soft colour-texture descriptors* method. The method is extracted and validated using a statistical procedure based on two statistical tools; *experimental design* [31] and *logistic regression* analysis [32]. These tools in conjunction provide a way to determine the best combination of quantitative/categorical factors related with a set of experiments. The best combination is achieved by seeking to maximize or minimize one response variable also involved in the experiments. In our case this output variable was the classification accuracy rate.

Although the method is not a new theoretical contribution, we demonstrate that

a simple set of global statistics of colour and texture, together with well-known classifiers, are powerful enough to comply hard factory requirements for real-time and performance. The two main requests of the industry are on-line inspection at factory rates (real-time compliance) and a high performance surface grading system. Production managers at factories will only accept an error rate close to 5% before relying on these automatic grading systems. The method meets the first demand by using the simplest and fastest [to compute] colour-texture features [29]. The second demand is met by achieving average accuracies over 95% in many of the tests carried out using the VxC TSG database described in Chapter 3.

5.1 CIE Lab and CIE Luv

The CIE (Commission Internationale de L'Eclairage) derived and standardized two perceptually uniform colour spaces from the CIE XYZ; the CIE Luv and the CIE Lab. The term 'perceptual' refers to the way that humans perceive colours. The term 'uniform' means that if we move in the colour space from one colour to another (from one coordinate to another) the perceptual difference will be related to a measure of distance, commonly the Euclidean distance, and the same distance will be approximately related to the same perceptual difference in all the colour space. Thus, we can measure colour differences close to the human perception of colours which makes these spaces useful for applications where colour difference measurement plays an important role, as is the case of the surface grading application presented in this thesis.

In fact, both spaces are only approximately uniform. However, them and their colour-difference formulae are the best approximation to perceptually uniform spaces and perceptual colour difference computation available at the moment [33, 116]. They are by far much more uniform than the XYZ and RGB colour spaces.

CIE Luv and CIE Lab are slightly different because of the different approaches to their formulation [33, 34]. Nevertheless, both spaces are equally good in perceptual uniformity and provide good estimates of colour difference (distance) between two colour vectors. Both spaces are used in colorimetry. The CIE Luv is mainly used for

industries considering additive mixing such as colour displays, TV and lighting [35], while the CIE Lab and the CMC difference formula have found wider acceptance in colour control industries. They are commonly used in image processing applications involving colour [117].

In CIE Lab and CIE Luv the L component is the same and represents lightness. It extends from 0 (black) to 100 (white). Also, both spaces have the same opponent colour axes approximately representing red-green versus yellow-blue, which are respectively a and b in the CIE Lab and u and v in the CIE Luv.

The database images were acquired originally in RGB , so conversion to Lab/Luv coordinates is needed. This conversion is done through the CIE XYZ colour space [33].

Linear conversion from RGB to XYZ :

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Nonlinear conversion from XYZ to CIE Lab :

$$L = \begin{cases} 116(Y/Y_n)^{1/3} - 16 & \text{if } Y/Y_n > 0.008856 \\ 903.3(Y/Y_n) & \text{otherwise} \end{cases}$$

$$a = 500(f(X/X_n) - f(Y/Y_n))$$

$$b = 200(f(Y/Y_n) - f(Z/Z_n))$$

where

$$f(t) = \begin{cases} t^{1/3} & \text{if } Y/Y_n > 0.008856 \\ 7.787t + 16/116 & \text{otherwise} \end{cases}$$

Nonlinear conversion from XYZ to CIE Luv :

$$L = \begin{cases} 116(Y/Y_n)^{1/3} - 16 & \text{if } Y/Y_n > 0.008856 \\ 903.3(Y/Y_n) & \text{otherwise} \end{cases}$$

$$u = 13L(w - w_n)$$

$$v = 13L(v - v_n)$$

where

$$w = 4X/X + 15Y + 3Z$$

$$w_n = 4X_n/X_n + 15Y_n + 3Z_n$$

$$v = 9X/X + 15Y + 3Z$$

$$v_n = 9X_n/X_n + 15Y_n + 3Z_n$$

X_n , Y_n , and Z_n are the values of X , Y and Z for the illuminant (reference white point). We followed the ITU-R Recommendation BT.709 [132, 133], and used the illuminant D_{65} , where $[X_n Y_n Z_n] = [0.95045 \ 1 \ 1.088754]$.

In both spaces colour difference is calculated using the Euclidean distance:

$$\Delta E_{Lab} = \sqrt{(\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2}$$

$$\Delta E_{Luv} = \sqrt{(\Delta L)^2 + (\Delta u)^2 + (\Delta v)^2}$$

More sophisticated colour-difference formulae has been proposed [118]; CMC(l:c), BFD(l:c), CIE94 and CIEDE2000. Some of these formulae were tested in previous tests against the Euclidean distance and no improvement was achieved in surface grading performance. Therefore, we used the Euclidean distance which is less expensive in computing costs.

5.2 Soft colour-texture descriptors

The mean colour vector was chosen as the colour feature of surfaces while texture description was done using simple statistical features. In addition, the standard deviation gives a measure of how the colour components vary along the image.

$$\sigma(z) = \sqrt{\frac{\sum_{i=1}^L (z_i - m)^2}{L - 1}}$$

where z is the random variable, L size of the data set and m the mean value of z

values.

By computing the image histogram for each component of the colour space we can calculate the histogram moments. Second, third, fourth and fifth moments are the simplest and fastest [to compute] approach for describing texture properties [29]. The n th moment of z about the mean is defined as:

$$\mu_n(z) = \sum_{i=1}^L (z_i - m)^n p(z_i)$$

where z is the random variable, $p(z_i)$, $i = 1, 2, \dots, L$ the histogram, L the number of distinct variable values and m the mean value of z .

$$m = \sum_{i=1}^L z_i p(z_i)$$

We call these set of features *soft colour-texture descriptors* because they 'softly' compile colour and texture properties from the whole image without using 'hard' approaches to colour or texture description. Colour histograms can easily collect 80,000 bins (different colours) which are all used to compute histogram dissimilarities. Centile-LBP approach uses 171 centile measures to compile colour property, and LBP histograms of 256 components to collect texture property (see Chapter 6). We can consider that these approaches use 'hard' colour and texture descriptors in comparison to our method which only uses the mean, standard deviation and histogram moments from 2nd to 5th to compile colour and texture properties (a maximum feature vector of 18 components). By comparison we named the proposed method *soft colour-texture descriptors*. This assertion is even more acceptable if we revise classical approaches to texture description in the literature (see texture subject in Chapter 2). In the surface grading application we do not strictly compare textures but global appearance differences, thus, the complex methods which compile precise texture description are not needed as results obtained in present chapter confirm.

5.3 Classifiers

We used statistical pattern recognition for classification. Here, the samples are formed by d -dimensional vectors which components x_1, x_2, \dots, x_d are observed features. Each sample belongs to one class w_1, w_2, \dots, w_c where c is the total number of classes. The samples belonging to a class will be distributed in the d dimensional space following an specific conditional probability function of density for this class; $p(x|w_i)$.

From this point of view, if we want to assign one class to a new sample, minimizing the error rate, we should know the conditional probability $P(w_i|x)$, also known as *a posteriori* probability, for each class. With this information, we can minimize the error risk assigning to x the class w with the greatest *a posteriori* probability.

$$w = \max_{w_i=w_1, \dots, w_c} \{P(w_i|x)\}$$

This criterion is known as the *Bayes Rule for the Minimum Error Rate*, and is the basics of most statistical classification methods. There are different statistical classifiers according to the approach used to estimate the *a posteriori* probabilities. In fact, some methods estimate the conditional function density of a class; $p(x|w_i)$. Then, in order to achieve the function of *a posteriori* probability the *Bayes formula* is required:

$$P(w_i|x) = \frac{p(x|w_i)P_i}{p(x)}$$

where $p(x)$ is the probability of x happening, and P_i is the probability of obtaining a sample of class w_i (also called *a priori probability*). As the denominator does not depend on the class, the *Bayes Rule* can be rewritten as:

$$w = \max_{w_i=w_1, \dots, w_c} \{p(w_i|x)P_i\}$$

It is possible to calculate the mean error probability (also called the *error rate*) over a data set X if we use the *Bayes Rule for the Minimum Error Rate*. The error

rate would be:

$$E = \int \left[1 - \max_{w_i=w_1, \dots, w_c} \{P(w_i|x)\} \right] p(x) d(x) \quad (5.1)$$

where $x \in X$ and $\{w_i : i = 1, \dots, c\}$ are the c classes contained in X .

This expression is very useful because it provides a way to compute the better error rate that we can achieve from a data set without depending on the type of classifier used. However, the expression 5.1 is not easy to calculate because we need to know the *a posteriori* probabilities of the classes.

In the case of the k nearest neighbours (k-NN) [30], is possible to achieve the *a posteriori* probability with the following formula.

$$\hat{P}(w_i|x) = \frac{k_i}{k}$$

where k is the total number of x neighbours used by the classifier, and k_i is the number of them belonging to class w_i . The error estimation from the expression 5.3 using $\hat{P}(w_i|x)$ provides a pessimistic estimation, although it is very appropriate when the k-NN classifier meets convergence conditions [30]. These conditions impose:

1. $k \rightarrow \infty$
2. $k/n \rightarrow 0$ when $n \rightarrow \infty$

where n is the number of elements in the data set used to design the k-NN classifier. In practice, the number of samples are not usually large enough to meet convergence conditions. Thus, the error rate has to be estimated by another method.

When estimating the $P(w_i|x)$ it is possible to use:

$$\hat{P}(w_i|x) = \frac{\sum_{u \in \theta_{k_i}} \frac{1}{d(u,x)}}{\sum_{v \in \theta_k} \frac{1}{d(v,x)}} \quad (5.2)$$

where θ_k is set of neighbours of x , θ_{k_i} is set of x neighbours with class w_i , and $d(.,.)$ is the distance between two samples [120]. Equation 5.2 uses the distance between neighbours, this information is useful when the design of the k-NN classifier do not comply with convergence conditions, as occurs in most cases.

In practice, when estimating the error rate ($accuracy\ rate = 1 - error\ rate$) the most frequently used method is the *count of errors*. This method needs two data sets; the training set and the test set. Both are build from the collected set of samples which is commonly called the *sample universe*. The idea is to design a classifier from the training set and then estimate the error rate from the accumulated errors when classifying the samples of the test set. Ideally, the training and test sets should be as large as possible and independent of each other. Nevertheless, normally the size of the universe of samples is not enough to comply with these ideal conditions. Thus, the design of the training and test sets is particularly important.

There are several approaches to design the training and test sets in order to estimate the error rate; *resubstitution*, *hold out*, *leaving-one-out*, *N-fold cross validation* [30] and *bootstrapping* [121]. For the methods of *resubstitution*, *hold out*, *leaving-one-out* and *N-fold cross validation* the estimated error rate (\hat{E}) would follow a binomial distribution.

$$P(k) = \binom{n}{k} E^k (1 - E)^{n-k} \quad (5.3)$$

where n is the number of samples of the test set, k is the number of errors made by the classifier and E the real error rate, which is unknown.

Is it possible to calculate confidence intervals for the estimated error rate ($\hat{E} = k/n$) using the expression 5.3 [122]. Figure 5.1 shows the confidence intervals at 95% for different sized test data sets. Estimation variance can be computed by means of the following expression [123]:

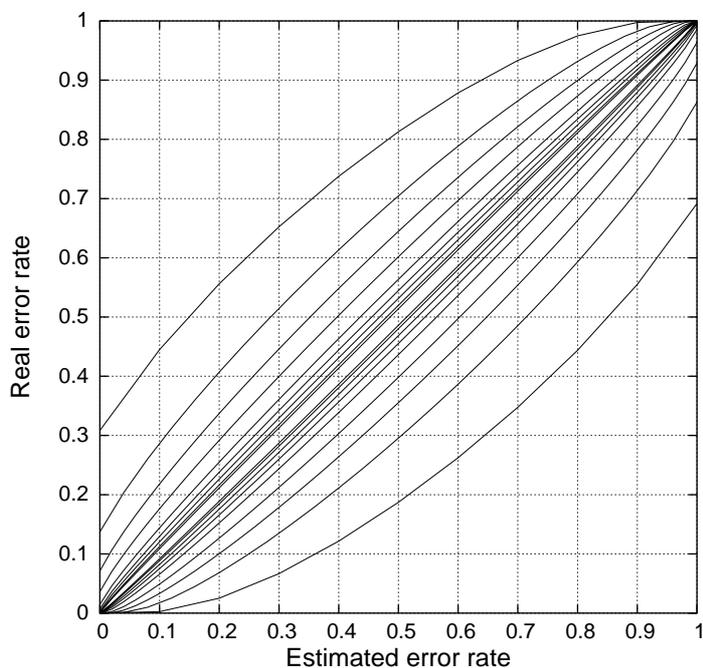
$$var(\hat{E}) = \frac{E(1 - E)}{n}$$

As expected, as larger the test set, the better estimation of the error rate.

Now, we revise briefly the characteristics of each method for estimating the error rate:

- *Resubstitution*, the training and the test sets are the same set; the original universe of samples. The error rate estimation is optimistic because the samples

Figure 5.1: 95% confidence intervals obtained from the estimated error rate for test sets with sizes; 10 (outside curves), 25, 50, 100, 250, 2500 and 5000 (inside curves).



used for the design of the classifier are also used to test it.

- *Hold-out*, the universe of samples is splitted into two separated sets. The desirable proportion between them varies from one author to other. Some defend the use of more samples in the training set [124], while others propose more samples in the test set [123]. In any case, the error rate estimation will be pessimistic because all the information is not used neither for training or testing. For this reason [30] proposes redesigning the classifier, once the error rate has been obtained, using samples from both sets and verifying the error rate using resubstitution.
- *Leaving-one-out*, this is a attempt to use the data as efficiently as possible. One sample is extracted from the universe of samples, the resulting set is now used as the training set and the extracted sample as the test set. This is done for all samples present in the original data set. The final error rate estimation will be the percentage of failed classifications. This method achieves an unbiased error

rate and it is recommended for small data sets where the previous approaches achieve poor estimations [125].

- *N-fold cross validation*, the universe of samples is split into N subsets of approximately the same size. In the case of *stratified K-fold cross validation* [123], the proportion among classes should be held in each subset. \hat{E} is achieved by computing the mean of the error rates collected for each subset when using a classifier designed without taking into account this subset. With this method the estimation variance is less than with leaving-one-out. Also, the estimation is less pessimistic than using hold-out when the size of the subsets is not very small [123]. If N is equal to the number of samples in the universe of samples then the appropriate method is leaving-one-out.
- *Bootstrap*, in this method b subsets are built with the same size from the universe of samples. Each data set is generated by random extraction with replacement, in other words, without eliminating the chosen samples from the universe of samples. The error rate estimation will be the mean of the b error rates. The variance of this estimation can be achieved from the subsets error variances [123]. This approach needs a large number of subsets to be effective (around 100 subsets or more).

In our experiments we used the k-NN classifier with hold-out and leaving-one-out error estimation because the universe of samples for each ceramic tile model was not very large. The size varies from 42 to 90 samples (a mean of 68 samples per model). The N-fold cross validation and the Bootstrap methods need larger universes of samples, and were therefore not appropriate for our application.

5.4 Experiments and results

In order to study the feasibility of the *soft colour-texture descriptors* we carried out a statistical design of experiments. Our aim was to test several factors to determine the combination of them providing the best accuracy results. These factors were

Table 5.1: Factors involved in the design of experiments of *soft colour-texture descriptors*.

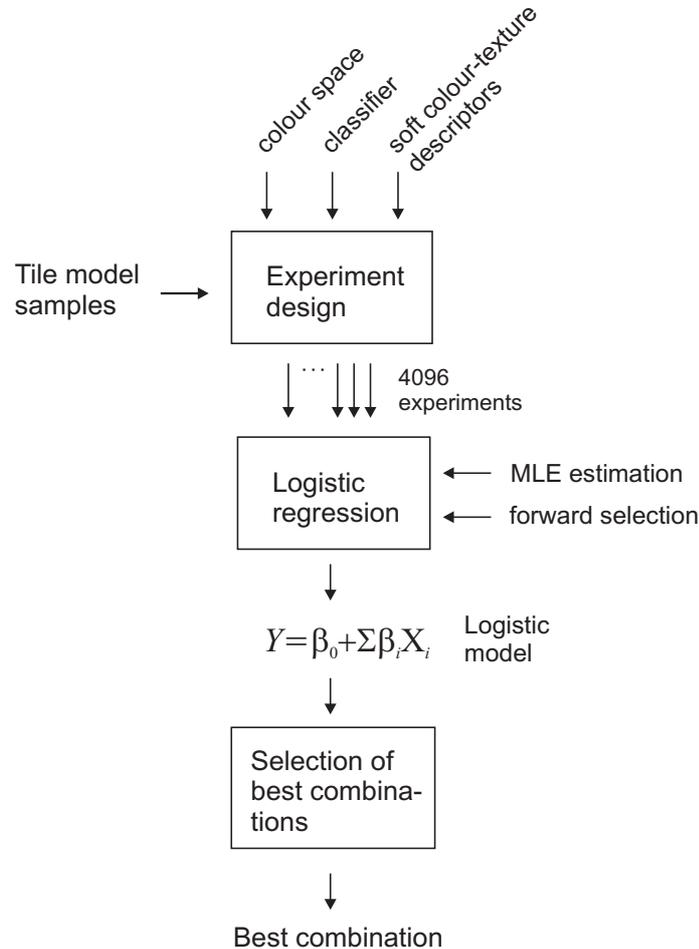
Factors	Values
Colour space	CIE Lab CIE Luv RGB Grey Scale
Classifier	k-NN with k=1,3,5,7 k-loo with k=1,3,5,7
Soft colour-texture descriptors	mean standard deviation 2nd to 5th moments

related with colour spaces, classifiers, and sets of soft colour-texture descriptors as it is shown in Table 5.1.

The chosen factors and their possible values defined 4096 different classification experiments for each tile model. The ground truth (the VxC TSG image database) was formed by 14 tile models, thus, a total number of 57.344 experiments had to be carried out. We decided to use a statistical tool, the *experimental design* [31, 119], in order to manage the large amount of experiments and results. This tool, in combination with the *logistic-regression* [32], provides a methodology for finding the best combination of factors involved in a set of experiments to maximize or minimize one response variable. In our case, we sought to maximize classification accuracy rates. This methodology follows the plan presented in the block diagram in Figure 5.2.

We added the RGB and Grey spaces to the colour space factor in order to test the goodness of using the perceptually uniform colour spaces in comparison to a common non perceptually uniform space (RGB) and when no colour information is provided (Grey scale).

Figure 5.2: Block diagram for the selection of the best combination of factors.



5.4.1 Experimental design

When we want to perform a complex experiment or set of experiments efficiently we need a scientific approach to experiment planning. Statistical design of experiments refers to the process of planning the experiment so that appropriate data can be collected for analysis with statistical methods and thus lead to objective and valid conclusions. The statistical approach to planning experiments is called *design of experiments* or *experimental design*.

An experiment is a test or series of tests in which changes are made to the input variables of a process in order to observe and identify the reasons for the changes that may be observed in the output response. The general approach to planning and conducting an experiment is called *experimentation strategy*. There are several

strategies for planning an experiment:

- Best-guess approach. From a set of factors and their possible values, one combination is tested. Depending on the results, another test is done varying one or two factors while maintaining the values of the remaining factors. This approach can be continued almost indefinitely.
- One-factor-at-a-time. This method consists of selecting a starting combination of factors values and then successively varying each factor over its range with the other factors held constant at the starting conditions. After all the tests are performed, a series of graphs are constructed showing how the response variable is affected by varying each factor with all other factors held constant. The optimal combination of factors is determined using these graphs. Here, the iteration between factors is excluded.
- Factorial. In this case the factors are varied together instead of one at a time to collect the affects of iteration between factors. This enables the experimenter to investigate the individual effects of each factor (or the main effects) and determine whether the factors interact with each other. This is an important and useful feature in factorial experiment because it makes the most efficient use of the experimental data providing useful information on both the factors and their iterations.

In our design of experiments we chose the factorial approach, which is the most appropriate when dealing with several factors [31]. More specifically, we used a complete factorial design. In this case, the experimenter selects a fixed number of 'levels' or 'versions' for each factor and then carries out experiments with all the possible combinations. If there are l_1 levels for the first factor, l_2 for the second, and l_k for the k^{th} factor, then this gives a $l_1 \times l_2 \times \dots \times l_k$ factorial design. For example, a factorial design $2 \times 3 \times 5$ is formed by $2 \times 3 \times 5 = 30$ single experiments.

All the factors in our design of experiments were categorical factors (non quantitative). The colour space could adopt the lab, luv, rgb or grey values. The classifier could be 1-nn, 3-nn, 5-nn, 7-nn, 1-loo, 3-loo, 5-loo or 7-loo (classifiers derived from

the k-NN approach [30]). And, the soft colour-texture descriptors were binary factors; '1' present, '0' not present. Each single experiment was the classification of a set of test samples belonging to an specific tile model. The sequence of experiments for each tile model was as shown in Table 5.2. The factors were varied in a nested way using the algorithm 1. By doing this, we defined an experimental design with orthogonal factors effects achieving independence between factors, iterations and experiments. Independent experiments lead to reliable results and conclusions when using statistics methods.

Algorithm 1 Nested variation of factors for an orthogonal sequence of experiments.

```

for colour_space in [ lab luv rgb grey ]
do
  for classifier in [ 1-nn 3-nn 5-nn 7-nn 1-100 3-100 5-100 7-100 ]
  do
    for mean in [ 0 1 ]
    do
      for standard_deviation in [ 0 1 ]
      do
        for 2nd moment in [ 0 1 ]
        do
          for 3rd moment in [ 0 1 ]
          do
            for 4th moment in [ 0 1 ]
            do
              for 5th moment in [ 0 1 ]
              do
                EXPERIMENT
              done
            done
          done
        done
      done
    done
  done
done
done
done
done
done
done
done
done
done

```

5.4.2 Logistic regression

The experimental design is used as the preliminary stage for modeling the behavior of a process which is characterized with k input factors and one output variable, also called the response variable. Commonly, once the experiments has been carried out according to the experimental design, a linear regression method is used to define a predictive model of the process. For two factors the linear regression model

Table 5.2: Orthogonal sequence of experiments for the soft colour-texture descriptors method.

Colour space	Classifier	Mean	Std dev	2nd Mnt	3rd Mnt	4th Mnt	5th Mnt	accuracy %
lab	1-nn	0	0	0	0	0	1	x
lab	1-nn	0	0	0	0	1	0	x
lab	1-nn	0	0	0	0	1	1	x
-	-	-	-	-	-	-	-	-
lab	1-nn	1	1	1	1	1	1	x
lab	3-nn	0	0	0	0	0	1	x
lab	3-nn	0	0	0	0	1	0	x
lab	3-nn	0	0	0	0	1	1	x
-	-	-	-	-	-	-	-	-
lab	3-nn	1	1	1	1	1	1	x
-	-	-	-	-	-	-	-	-
lab	7-loo	0	0	0	0	0	1	x
lab	7-loo	0	0	0	0	1	0	x
lab	7-loo	0	0	0	0	1	1	x
-	-	-	-	-	-	-	-	-
lab	7-loo	1	1	1	1	1	1	x
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
luv	1-nn	0	0	0	0	0	1	x
luv	1-nn	0	0	0	0	1	0	x
luv	1-nn	0	0	0	0	1	1	x
-	-	-	-	-	-	-	-	-
luv	1-nn	1	1	1	1	1	1	x
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
grey	7-loo	1	1	1	1	1	1	x

corresponds to:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

where y is the response variable, the β 's are parameters whose values are to be determined, x_1 is the variable that represents factor 1, x_2 is the variable that represents factor 2 and $x_1 x_2$ represents the interaction between factors x_1 and x_2 . The estimation of the β 's parameters is done using the *least square error fitting*.

The linear regression model for two factors can be generalized to k factors as follows:

$$y = \beta_0 + \sum_{i=1}^I \beta_i x_i + \sum_{i=1}^I \sum_{j=1, j \neq i}^J \beta_{ij} x_i x_j + \dots + \sum_{i=1}^I \sum_{j=1, j \neq i}^J \dots \sum_{k=1, k \neq i, j}^K \beta_{ij\dots n} x_i x_j \dots x_k$$

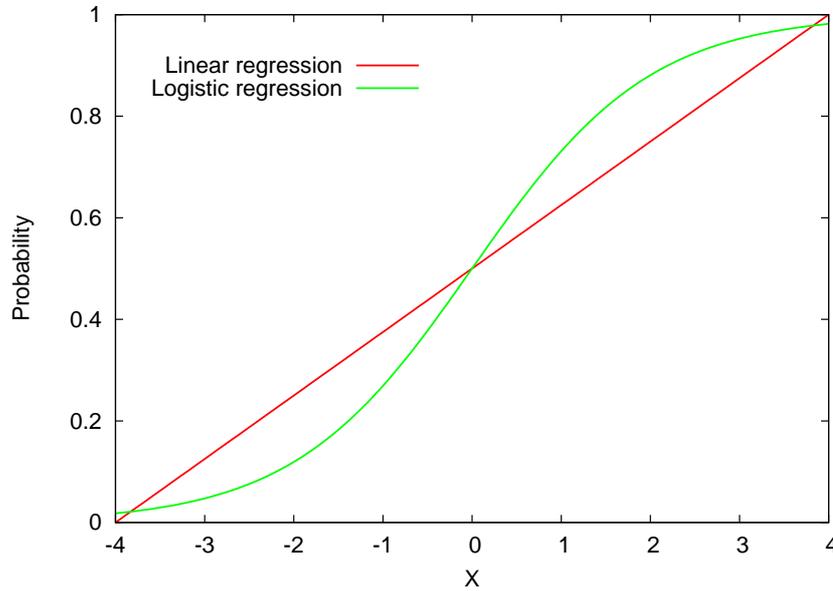
This expression is simplified if we consider the interactions between factors as new factors.

$$y = \beta_0 + \sum \beta_i X_i \tag{5.4}$$

However, in our experiments the response variable y is an accuracy percentage or probability. In these cases, the linear regression model does not correspond to a normal distribution for the response variable, which is desirable, but to a uniform distribution (see Figure 5.3), and also, it could give estimated values out of the range $[0,1]$. In order to solve these problems a logistic regression model is commonly used. This model approximates the probability response to a normal distribution (the logistic S-shaped distribution is similar to the standard normal distribution) and also forces the estimated probabilities to lie between 0 and 1. Furthermore, the logistic model is easier to work with in most applications.

With the logistic regression model the response variable p is converted into $\log\left(\frac{p}{1-p}\right)$ and the expression 5.4 is transformed to:

Figure 5.3: Logistic regression model vs linear regression model.



$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \sum \beta_i X_i$$

Now, the expected probability or percentage can be defined as:

$$p = \frac{e^{\beta_0 + \sum \beta_i X_i}}{1 + e^{\beta_0 + \sum \beta_i X_i}} \quad (5.5)$$

The β parameters of the logistic regression approach can be estimated using the *Maximum Likelihood Estimation* (MLE) or the *Weighted Least Squares* (WLS). MLE is used as an alternative to non-linear least squares (WLS) for nonlinear equations. In addition to the parameter estimation method, there are several approaches for computing the logistic model depending on the number of factors taken into account. The simplest way to determine the logistic model is to consider all the factors and their iterations. Then, we are forced to calculate all the β parameters for all factors (or iteration of factors) even if they are not significant in relation to the response variable. There are two other approaches which do not consider all the factors and iterations, these are the *stepwise methods*:

- *Forward selection*, the process of forward selection begins with an initial model. Then, new terms with greater significance are added to the model while the global level of significance remains below a given level of [first class] risk (usually 5-10%). The process continues until there is no factor or iteration which when added to the model keeps it under the risk level.
- *Backward elimination*, now the starting point is the complete logistic model with all the factors and iterations. Factors or iterations are eliminated one by one using a lower significance order. The significance of each factor is studied (p -value computed) and the less significant (greater p -value) is eliminated. This process continues until the global significance of the model reaches a value below a previously determined level of risk.

Many software implementing statistics methods include options for computing the most common regression methods. In our case, we used the statistical package Statgraphics v5.1 for computing the logistic regression models. These models were computed using MLE estimation and Forward selection.

5.4.3 Selection of best combinations

The procedure for selecting the best combination of factors is the following:

1. For each input in the design of experiments (see Table 5.2) we compute the output variable y using the achieved regression model $y = \beta_0 + \sum \beta_i X_i$. Then we compute the predicted accuracy rate $p = \frac{e^y}{1+e^y}$.
2. The best combination of factors will be the one which the greatest predicted accuracy rate.

5.4.4 Results

We should point out that we used the experimental design in combination with the logistic regression model not to model the response of a process, but to find the best combination of factors to achieve the maximum accuracy rates in a surface grading application. This is another usage of these methods [31, 32, 119].

Diagram in Figure 5.2 shows the steps to be taken to make an experimental design and select the best combination of factors for each tile model. However, the experimental designs of all models can be grouped into a single one if we re-define the output response y as the achieved mean accuracy computed over all models. We carried out the 4096 experimental design of each model, 57.344 experiments in total, and grouped the results in a new 4096 experimental design where the output response y was the mean accuracy of all models.

We used the Statgraphics v5.1 software and computed the logistic regression model (see Figure 5.5). The high adjusted percentage of deviance explained by model (81%) indicated the model was very good. After this and following the diagram (Figure 5.2), we carried out the selection of best combinations as it is explained in section 5.4.3 and obtained a table similar to 5.2 with the predicted accuracy rates for each experiment or combination of factors. We sorted the results using the predicted accuracy rate (equation 5.5). The 30 best combinations are shown in Table 5.3. From the complete sorted table of results we extracted the following summary of results.

1. Best combination corresponds to CIE Lab, 1-loo and all soft colour-texture descriptors. This combination achieved an accuracy rate of 97.4% .
2. A number of 237 combinations (5,8% of total combinations) achieved an accuracy rate over 95%.
3. All the combinations over 95% used CIE Lab or CIE Luv spaces.
4. Best combination using RGB space achieved 93.6% of accuracy, although its presence in best combinations was very low (Figure 5.4).
5. Best combination using Grey space achieved 90.8% of accuracy, although its presence in best combinations was almost null (Figure 5.4).
6. Best combinations make wide use of CIE Lab and CIE Luv spaces (Figure 5.4).
7. The classifiers from best to worst are 1-loo, 3-loo, 5-loo, 7-loo, 1-nn, 3-nn, 5-nn and 7-nn (Figure 5.4).

8. The very best combinations use all or almost all soft colour-texture descriptors (Table 5.3).

From this summary of results we can conclude:

1. The method we extract from the design of experiments and logistic regression methods uses CIE Lab colour space, 1-100 classifier and all soft colour-texture descriptors.
2. The perceptually uniform colour spaces CIE Lab and CIE Luv achieve by far the best performance.
3. RGB colour space provides less performance, and although it reaches some accuracy rates around 93%, the percentage using RGB in sets of best combinations is almost null.
4. Grey space provides the worst performance as expected.
5. Best classifiers are the derived from the leaving-one-out method. Therefore, we should use 1-NN method using as more training samples as possible.
6. Using all soft colour-texture descriptors gives the best discriminant power.

Figure 5.4: Presence in percentage of colour spaces and classifiers in best combinations sets.

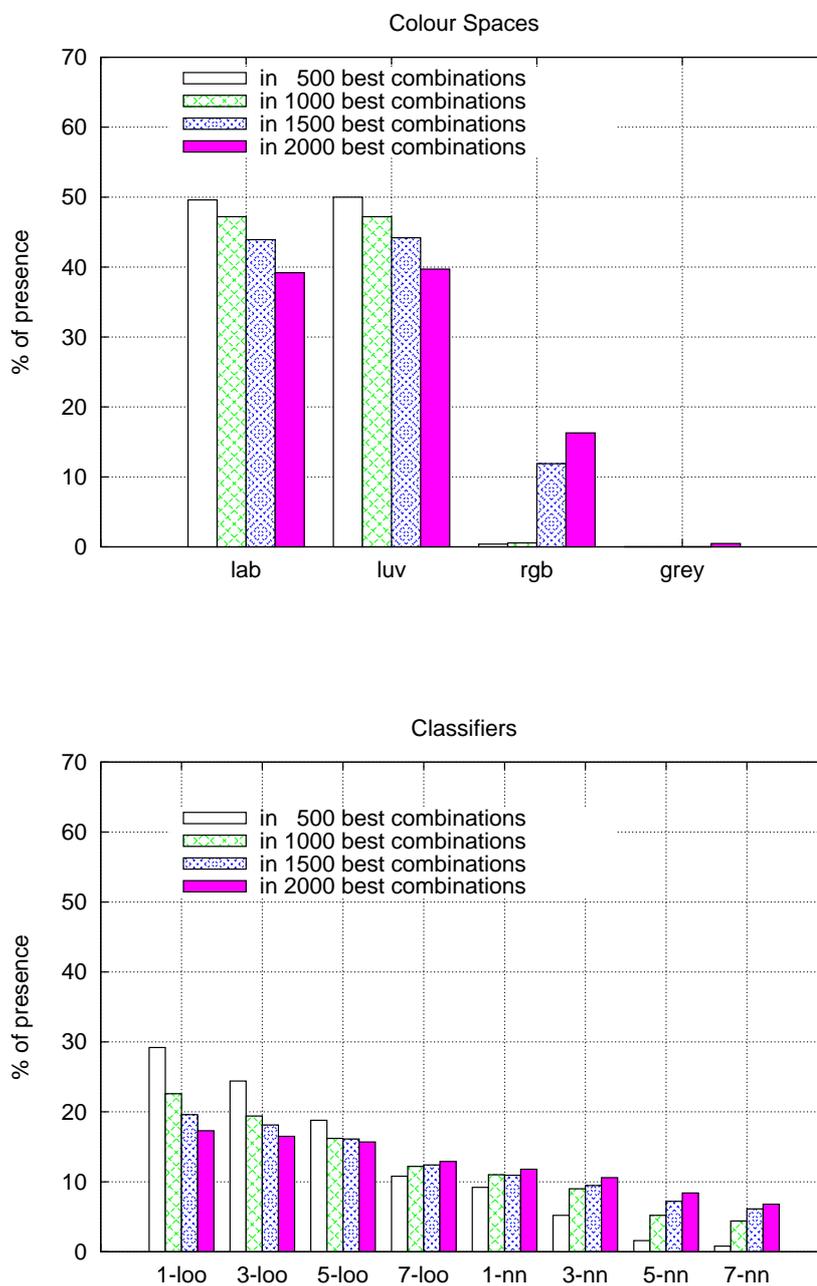


Figure 5.5: Computed logistic regression model for *soft colour-texture descriptors* method.

```

Dependent variable: accuracy/100
Sample sizes: n_tests
Factors: Spa, Clas, Mean, Stdv, M2, M3, M4, M5

Estimated Regression Model (Maximum Likelihood)
-----
Parameter                Estimate          Standard          Estimated
                        Error              Odds Ratio
-----
CONSTANT                  2.01982          0.0111142
Spa=grey                  -0.13145         0.0142337        0.876823
Spa=lab                   0.706523        0.0162755        2.02693
Spa=luv                   0.759474        0.0165152        2.13715
Clas=1loo                 0.66783         0.0129068        1.95
Clas=1nn                  0.392598        0.0145836        1.48082
Clas=3loo                 0.644807        0.0128619        1.90562
Clas=3nn                  0.337767        0.0144461        1.40181
Clas=5loo                 0.601846        0.0127805        1.82549
Clas=5nn                  0.238795        0.014215         1.26972
Clas=7loo                 0.463897        0.0125395        1.59026
Mean=0                    -0.667277       0.00357258       0.513104
Stdv=0                    -0.0650154      0.00647503       0.937053
M2=0                      -0.094827       0.00347998       0.90953
M3=0                      -0.260792       0.00349194       0.770441
M4=0                      -0.212335       0.00348728       0.808694
M5=0                      -0.239989       0.00348982       0.786636
Spa=grey*Clas=1lo       -0.269689       0.0176259        0.763617
Spa=grey*Clas=1nn       -0.184656       0.0199877        0.831391
Spa=grey*Clas=3lo       -0.298305       0.0175457        0.742075
Spa=grey*Clas=3nn       -0.160635       0.0198475        0.851603
Spa=grey*Clas=5lo       -0.255156       0.0174863        0.774796
Spa=grey*Clas=5nn       -0.105141       0.019625         0.900198
Spa=grey*Clas=7lo       -0.193629       0.0172447        0.823963
Spa=lab*Clas=1loo       0.213016        0.0212053        1.2374
Spa=lab*Clas=1nn        -0.0656253      0.0230624        0.936482
Spa=lab*Clas=3loo       0.152144        0.0209448        1.16433
Spa=lab*Clas=3nn        -0.114159       0.0226799        0.892116
Spa=lab*Clas=5loo       -0.00596527     0.0203991        0.994052
Spa=lab*Clas=5nn        -0.136091       0.0222191        0.872763
Spa=lab*Clas=7loo       -0.0445569     0.0198791        0.956421
Spa=luv*Clas=1loo       0.125391        0.0213999        1.13359
Spa=luv*Clas=1nn        -0.109549       0.0233727        0.896238
Spa=luv*Clas=3loo       -0.0273054     0.0209092        0.973064
Spa=luv*Clas=3nn        -0.144568       0.0230209        0.865396
Spa=luv*Clas=5loo       -0.0892126     0.0206141        0.914651
Spa=luv*Clas=5nn        -0.222829       0.0224077        0.800252
Spa=luv*Clas=7loo       -0.168171       0.0200229        0.845209
Spa=grey*Stdv=0         0.000901265    0.00872695       1.0009
Spa=lab*Stdv=0          -0.110032       0.0104852        0.895805
Spa=luv*Stdv=0          -0.0489122     0.0104693        0.952265
-----

Percentage of deviance explained by model = 80.9518
Adjusted percentage = 80.8989

```

N° Exp.	Space	Classifier	Mean	Std dev	2nd Mnt	3rd Mnt	4th Mnt	5th Mnt	p*100	c.i. 95%
620	lab	1loo	1	1	1	0	1	1	96.60	[95.26, 97.68]
636	lab	1loo	1	1	1	0	1	1	96.60	[95.26, 97.68]
720	lab	3loo	1	0	1	1	1	1	96.60	[95.26, 97.68]
736	lab	3loo	1	0	1	1	1	1	96.60	[95.26, 97.68]
1646	luv	1loo	1	1	1	1	0	1	96.64	[95.38, 97.76]
1662	luv	1loo	1	1	1	1	0	1	96.64	[95.38, 97.76]
1608	luv	1loo	1	0	0	1	1	1	96.66	[95.38, 97.76]
1624	luv	1loo	1	0	0	1	1	1	96.66	[95.38, 97.76]
623	lab	1loo	1	1	1	1	1	0	96.67	[95.38, 97.76]
639	lab	1loo	1	1	1	1	1	0	96.67	[95.38, 97.76]
622	lab	1loo	1	1	1	1	0	1	96.75	[95.50, 97.85]
638	lab	1loo	1	1	1	1	0	1	96.75	[95.50, 97.85]
1776	luv	3loo	1	1	1	1	1	1	96.76	[95.50, 97.85]
1792	luv	3loo	1	1	1	1	1	1	96.76	[95.50, 97.85]
744	lab	3loo	1	1	0	1	1	1	96.86	[95.63, 97.94]
760	lab	3loo	1	1	0	1	1	1	96.86	[95.63, 97.94]
592	lab	1loo	1	0	1	1	1	1	96.87	[95.63, 97.94]
608	lab	1loo	1	0	1	1	1	1	96.87	[95.63, 97.94]
1616	luv	1loo	1	0	1	1	1	1	96.95	[95.75, 98.02]
1632	luv	1loo	1	0	1	1	1	1	96.95	[95.75, 98.02]
1640	luv	1loo	1	1	0	1	1	1	97.00	[95.75, 98.02]
1656	luv	1loo	1	1	0	1	1	1	97.00	[95.75, 98.02]
616	lab	1loo	1	1	0	1	1	1	97.10	[95.87, 98.11]
632	lab	1loo	1	1	0	1	1	1	97.10	[95.87, 98.11]
752	lab	3loo	1	1	1	1	1	1	97.13	[95.87, 98.11]
768	lab	3loo	1	1	1	1	1	1	97.13	[95.87, 98.11]
1648	luv	1loo	1	1	1	1	1	1	97.27	[96.12, 98.28]
1664	luv	1loo	1	1	1	1	1	1	97.27	[96.12, 98.28]
624	lab	1loo	1	1	1	1	1	1	97.36	[96.25, 98.36]
640	lab	1loo	1	1	1	1	1	1	97.36	[96.25, 98.36]

Table 5.3: Thirty best combinations of factors. Predicted accuracy (p*100) and confidence intervals are given in %.

5.5 Conclusions

In this chapter we have extracted a final approach to surface grading. This approach is based on the use of *soft colour-texture descriptors* in conjunction with the perceptually uniform colour space CIE Lab and the well known classifier 1-leaving-one-out (1-NN method using as more training samples as possible). The method has been extracted from two statistical tools; the *design of experiments* and the *logistic regression*. These tools provided us a method with which to study and determine the best combination of factors to maximize the accuracy rate. The studied factors have been several colour spaces, classifiers, and sets of soft colour-texture descriptors. A total of 57.344 independent experiments were carried out exploring all the combinations of factors over all the models of the VxC TSG database.

The extracted method achieves 97.4% predicted mean accuracy and 97.6% measured mean accuracy of all models. The computed confidence interval is [96.25%, 98.36%] which clearly surpasses the factory requirement for performance. In addition, many tests using the perceptually uniform spaces CIE Lab or CIE Luv, the leaving-one-out classifiers and several combinations of soft colour-texture descriptors surpassed the factory minimum accuracy of 95%.

The results show that RGB, a common colour space but non perceptually uniform, provide less discriminative power. Perceptually uniform colour spaces are closer to the human perception of colours, therefore, it can be considered logical if they provide better performance when carrying out visual tasks done by humans. As expected, the worst results come from the Grey space which only makes use of light intensity omitting colour information. From the in-depth statistical study it is also concluded that all the soft colour-texture descriptors contribute to achieve accuracy performance and also the classifiers derived from the leaving one out method perform better, thus we should select the 1-NN method using as more training samples as possible.

Finally, we report that *soft colour-texture descriptors* method is a good choice when a deep colour-texture description is not needed. This is the case of the surface grading application where the pieces are graded taking into account the global appearance of colour and texture. Local description is not needed, hard description

of colour and texture is not necessary. Furthermore, *soft colour-texture descriptors* method is a good choice for real-time compliance because it uses the less expensive features of texture and colour from the point of view of computational costs.

Chapter 6

Literature methods

This chapter deals with the implementation and in-depth study of two methods from the literature of surface grading. These methods are the *colour histograms* [2,3] and *centile-LBP* [1,12]. We performed this study for comparison purposes with the method presented in previous chapter based on *soft colour-texture descriptors*. We chose them because they are similar to ours, both are generic solutions with low computational costs. Also, a review of the literature presented in Chapter 2 shows that no more methods are available for comparison because the rest of them deal with specific surface types or use hard colour-texture description with expensive computational costs.

Again we use the statistical tools of experimental design and logistic regression to in-depth study the methods using several factors. Experiments included the factors of colour space, classifier and inter-histograms distance. In order to perform a proper comparison we also used the VxC TSG image database, the ground truth used in previous chapter.

6.1 Colour histograms

Basically colour histograms are 3D histograms where each axis represents one colour space channel. Colour histograms are compared using dissimilarity measures between histograms. This approach to surface grading was developed by Boukouvalas et al [2,3]. They proposed to use the differences between colour histograms to solve

the problem of shade grading (surface grading) of multi-coloured textured surfaces (random pattern surfaces). However, colour histograms are inefficient in terms of memory requirements. A colour image acquired in RGB normally need 8 bits per colour channel at each pixel, so 16Mbytes (2^{24} memory positions) are needed to store one colour histogram. Nevertheless, in real images colour values tend to be clustered around a few locations. For instance, the image of a ceramic tile may occupy only 80.000 different locations (234Kb). Apart from being highly demanding in memory, this approach is computationally intensive because in order to compare two histograms all memory locations have to be parsed.

To save memory and computational costs binary trees are used to store the colour histograms. Binary trees are frequently used to represent a set of data whose elements are retrievable through a unique key (value). If a tree is organized in such a way that for each node all values in the left subtree are less than the value of the parent node, and those in the right subtree are greater than the value of the parent node, then this tree is called *ordered binary tree* or *search tree*. A search of a value in a tree of n elements may be performed with only $\log n$ comparisons, if the tree is balanced.

When a colour histogram is stored in a binary tree, the value of a node is a particular RGB value. This is converted to a 24 bit-integer by concatenating the R, G and B bytes. Each node also contains the number of pixels with the same RGB value (repetitions). Therefore, only RGB combinations that exist in the image are inserted in the tree, and the searching of existing nodes is very efficient.

Colour histograms are invariant to translation and rotation about an axis perpendicular to the image plane, and change only slightly under changes of angle of view. And also they are invariant to exact spatial distribution of the coloured pixels. This property is desirable when dealing with random pattern surfaces, as often occurs when we deal with ceramic tile models. The method based on *soft colour-texture descriptors* is also invariant to translation, rotation and exact spatial distribution of coloured pixels.

To perform the surface grading, the similarity (or dissimilarity) of ceramic tiles is compared using the similarity of their colour histograms. As the histograms can be

viewed as distributions we can use statistical methods to compare two distributions [96]. Boukouvalas et al used the *chi-square test* and the *linear correlation coefficient*.

The *chi-square* statistic is defined as:

$$\chi^2 = \sum_i \frac{(N_i - n_i)^2}{n_i}$$

where N_i is the number of events observed in the i th bin, and n_i is the number expected according to some known distribution and the sum is over all bins. A large value of χ^2 indicates dissimilarity between the two distributions.

When comparing two binned data sets, with the same number of data points, the equation adopts a different form. Let R_i be the number of events in bin i for the first data set, let S_i be the number of events in the same bin for the second data set. Then the *chi-square* statistic is:

$$\chi^2 = \sum_i \frac{(R_i - S_i)^2}{R_i + S_i}$$

The *linear correlation coefficient* is another test which measures the association between random variables. For pairs of quantities (x_i, y_i) , $i = 1, \dots, N$, the linear correlation coefficient r is given by:

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}}$$

where \bar{x} is the mean of the x_i values and \bar{y} is the mean of the y_i values.

The value of r is always in the range $[-1, 1]$. The correlation is called *positive* or *direct correlation* when y tends to increase as x increases. If y tends to decrease as x increases the correlation is then called *negative* or *inverse correlation*. A value near to zero in the previous equation indicates poor linear correlation between the variables r and y .

See Chapter 2 for more information about colour histograms.

6.2 Centile-LPB

This method proceeds from the area of wood inspection. In this case, the grading of lumber boards and parquet slabs is not related to the overall texture and colour appearance of the surface. The grade of the wood piece is assigned by detecting the wood defects (mainly knots) and then applying grade rules related to the number and classes of defects found in the inspection process [1]. Thus, from the point of view of the computer vision, the problem becomes a question of separating the surface into sound and faulty wood, and classifying the defects into their different types. However, we should point out that we use this method not in the literature sense but to compile global description of colour and texture and directly grade surfaces.

Kauppinen started the approach to the problem using only colour information which was derived from the percentile features of the RGB histograms [1, 11]. The percentiles, also called centiles, are calculated from a cumulative histogram $C_k(x)$, which is defined as a sum of all the values that are smaller than x or equal to x in the normalized histogram $P_k(x)$, corresponding to the colour channel k . Finding a value for a percentile involves finding the x when $C_k(x)$ is known, thus, requiring an inverse function of $C_k(x)$. If we denote the percentile feature with $F_k(y)$ then

$$F_k(y) = C_k^{-1}(y) = x$$

where y is a value of the cumulative histogram in the range $[0\%, 100\%]$.

In the classification stage feature vectors composed of selected sets of plain centile features are used, and also differences of two centile features either from the same colour channel or from two different colour channels are used. Specifically, in implementing the method we used the centiles from 5 to 95 step 5 of each channel, differences inside channels between correlative positions separated 5 steps, and also, inter-channel differences between the same positions every 5 steps. A total number of 171 features were computed for a three dimensional channel. Finally, Kauppinen used the well-known k-NN method to perform the classification.

After this first approach, Niskanen, Silvén and Kauppinen continued the work

including texture properties [105,106]. They extended the method using the Local Binary Pattern (LBP) texture operator, previously introduced by their colleagues Ojala and Pietikäinen [108,109]. The original 3x3 neighborhood is thresholded by the value of the center pixel. The values of the pixels in the thresholded neighborhood are multiplied by the weights given to the corresponding pixels. Finally, the values of the eight pixels are summed to obtain the number of this texture unit. Niskanen et al used Self Organizing Maps (SOM) combining Centile and LBP properties in order to carry out the classification task.

Kyllönen and Pietikäinen [12] also combined centile and LBP features but they used a different approximation for classification. They combined the centile and LBP features in one measure of distance and then used the k-NN classifier. For the centile features they used the Euclidean distance in the feature space, and for LBP they used a log-likelihood measure to compute the dissimilarity (distance) between sample and reference histograms. When using LBP there are 2^8 possible combinations of texture numbers. Thus, the texture description of a region can be collected computing its LBP histogram.

The log-likelihood measure to compute the dissimilarity was:

$$L(S, R) = - \sum_{n=0}^{N-1} S_n \ln R_n$$

where N is the number of bins. S_n and R_n are the sample and reference probabilities of bin n .

They joined these distances by simply adding them. Prior to this, both distances were normalized using:

$$d = \frac{d - d_{min}}{d_{max}}$$

$$d = \frac{d}{d_{avg}}$$

where d_{min} , d_{max} and d_{avg} are the min, max and average values of all the distances found in the training set.

We chose this last approximation of centile-LBP to compare with our method because SOMs are relatively complex systems for classification which also need sophisticated training procedures. See Chapter 2 for more details about the centile-LBP method.

6.3 Experimental design and results

We used the factors of colour space and classifier in the same way as they were used for the *soft colour-texture descriptors*, and added a new factor related to the measurement of dissimilarity between histograms, the distance measure factor (see Table 6.1). In both methods the histogram representation of data is used and so a measure of dissimilarity is needed. We chose the distances used in the methods in the literature; the *chi-square statistic*, the *linear correlation coefficient* and the *log-likelihood measure*.

For each tile model 96 independent experiments are defined and applying the design of experiments to the ground truth formed by the 14 models of the VxC TSG database, a total number of 1344 experiments had to be performed.

Table 6.1: Factors involved in the design of experiments of colour histograms and centile-LBP.

Factors	Values
Colour space	CIE Lab CIE Luv RGB Grey Scale
Classifier	k-NN with k=1,3,5,7 k-loo with k=1,3,5,7
Distance measure	Chi-square statistic Linear correlation coefficient Log-likelihood measure

Once again, all the factors in our design of experiments were categorical factors (non quantitative). The colour space could adopt the lab, luv, rgb or grey values.

The classifier could be 1-nn, 3-nn, 5-nn, 7-nn, 1-loo, 3-loo, 5-loo or 7-loo. In addition, the measure of distance among histograms could be the chi-square test, the linear correlation coefficient or the log-likelihood measure proposed for the centile-LBP method. In each single experiment a set of test samples belonging to an specific tile model was classified using reference sets of samples. The sequence of experiments for each tile model was carried out as shown in Table 6.2. The factors were varied in a nested way using an algorithm similar to the algorithm 1, shown in the previous chapter. Again, an experimental design with orthogonal factors effects was defined achieving independence between factors, iterations and experiments.

As in the previous experimental design of *soft colour-texture descriptors*, the experiments of all models were grouped in a single one redefining the output response as the achieved mean accuracy computed over all models. The 96 tests of the experimental design were carried out for each model, 1.344 experiments in total, and the results grouped into a new 96 experimental design where the output response was the mean accuracy of all models. Then, we computed the logistic regression models (see Figures 6.1 and 6.2). Both models achieved a very high adjusted percentage of deviance explained by model (96% and 97.4%), which indicates good regression models.

Then, we carried out the selection of best combinations and obtained a table similar to 6.2 with the predicted accuracy rates for each experiment. The results were sorted using the predicted accuracy rate. The 30 best combinations of each approach are shown in Tables 6.3 and 6.4.

From the results and figures we extract the following conclusions.

Colour histograms

1. Best accuracy result (97.8%) is achieved using RGB colour space, Chi-square distance and one leaving-one-out classifier. This result from the experimental design meets the proposed method in literature.
2. The overall best performance of colour spaces is for RGB followed by CIE Lab.
3. Chi-square statistic is clearly the best distance, followed by linear correlation.

The log-likelihood measure does not work properly for colour histograms.

4. Best classifiers stem from the leaving-one-out approach. Therefore, we should select the 1-NN method using as more training samples as possible.

Centile-LBP

1. Best accuracy result (98.3%) is achieved using CIE Lab colour space, linear correlation distance for LBP histograms and one leaving-one-out classifier. In this case, the method extracted from the experimental design does not coincide with the proposed in literature (RGB colour space, log-likelihood distance and k-NN classifier).
2. CIE Lab followed by RGB are the colour spaces giving best results.
3. All the distances show good behavior, but the linear correlation gives the best performance.
4. Best classifiers are the derived from the leaving-one-out approach. Again, we should select the 1-NN method using as more training samples as possible.

Table 6.2: Orthogonal sequence of experiments for the colour histograms and centile-LBP methods.

Colour space	Classifier	Distance	accuracy %
lab	1-nn	chi-square	x
lab	1-nn	correlation	x
lab	1-nn	log-likelihood	x
lab	3-nn	chi-square	x
lab	3-nn	correlation	x
lab	3-nn	log-likelihood	x
-	-	-	-
lab	7-loo	chi-square	x
lab	7-loo	correlation	x
lab	7-loo	log-likelihood	x
-	-	-	-
-	-	-	-
luv	1-nn	chi-square	x
luv	1-nn	correlation	x
luv	1-nn	log-likelihood	x
-	-	-	-
luv	7-loo	chi-square	x
luv	7-loo	correlation	x
luv	7-loo	log-likelihood	x
-	-	-	-
-	-	-	-
-	-	-	-
grey	1-nn	chi-square	x
grey	1-nn	correlation	x
grey	1-nn	log-likelihood	x
-	-	-	-
grey	7-loo	chi-square	x
grey	7-loo	correlation	x
grey	7-loo	log-likelihood	x

Figure 6.1: Computed logistic regression model for colour histograms.

```

Logistic Regression

Dependent variable: accuracy/100
Sample sizes: n_tests
Factors: Spa, Dis, Clas

Estimated Regression Model (Maximum Likelihood)
-----
Parameter                Estimate      Standard      Estimated
                        Error          Odds Ratio
-----
CONSTANT                 0.588225     0.0827639
Spa=grey                 0.875956     0.109761     2.40117
Spa=lab                  1.17319      0.120174     3.23228
Spa=luv                  1.24737      0.117719     3.48117
Dis=chi                  2.0869       0.115529     8.05992
Dis=corr                 1.82857      0.107692     6.22498
Clas=1loo               0.227155     0.104175     1.25502
Clas=1nn                -0.150119    0.113644     0.860606
Clas=3loo               0.325216     0.104668     1.38433
Clas=3nn                 0.125589     0.11613      1.13382
Clas=5loo               0.400251     0.10472      1.4922
Clas=5nn                 0.0639772    0.115304     1.06607
Clas=7loo               0.343234     0.103363     1.4095
Spa=grey*Dis=chi        -2.24924     0.0960474    0.105479
Spa=grey*Dis=corr       -2.08892     0.0880021    0.12382
Spa=lab*Dis=chi         -1.51507     0.106083     0.219793
Spa=lab*Dis=corr        -1.47667     0.095813     0.228398
Spa=luv*Dis=chi         -2.01041     0.103336     0.133934
Spa=luv*Dis=corr        -1.96702     0.093971     0.139873
Spa=grey*Clas=1loo      0.407846     0.142735     1.50358
Spa=grey*Clas=1nn       0.290179     0.146772     1.33667
Spa=grey*Clas=3loo      0.422044     0.143303     1.52508
Spa=grey*Clas=3nn       0.114563     0.149548     1.12138
Spa=grey*Clas=5loo      0.411459     0.14192      1.50902
Spa=grey*Clas=5nn       0.0141489    0.147173     1.01425
Spa=grey*Clas=7loo      0.313169     0.136576     1.36775
Spa=lab*Clas=1loo       0.248035     0.157121     1.2815
Spa=lab*Clas=1nn        -0.0438777    0.158413     0.957071
Spa=lab*Clas=3loo       0.255287     0.158025     1.29083
Spa=lab*Clas=3nn        -0.231686     0.160787     0.793195
Spa=lab*Clas=5loo       0.171913     0.155418     1.18757
Spa=lab*Clas=5nn        -0.125283     0.16093      0.882248
Spa=lab*Clas=7loo       0.20717      0.151614     1.23019
Spa=luv*Clas=1loo       0.551315     0.158477     1.73553
Spa=luv*Clas=1nn        0.0748652     0.154739     1.07774
Spa=luv*Clas=3loo       0.371197     0.155113     1.44947
Spa=luv*Clas=3nn        -0.00367895   0.158814     0.996328
Spa=luv*Clas=5loo       0.27189      0.151869     1.31244
Spa=luv*Clas=5nn        -0.0908977    0.156081     0.913111
Spa=luv*Clas=7loo       0.284234     0.147342     1.32874
Dis=chi*Clas=1loo      0.900621     0.142402     2.46113
Dis=chi*Clas=1nn       0.0995474     0.132334     1.10467
Dis=chi*Clas=3loo       0.702057     0.139227     2.0179
Dis=chi*Clas=3nn        0.0583487     0.135531     1.06008
Dis=chi*Clas=5loo       0.508544     0.13549      1.66287
Dis=chi*Clas=5nn        0.0595543     0.133344     1.06136
Dis=chi*Clas=7loo       0.318586     0.129355     1.37518
Dis=corr*Clas=1loo     0.74998      0.131465     2.11696
Dis=corr*Clas=1nn       0.19488      0.128234     1.21517
Dis=corr*Clas=3loo     0.621296     0.130495     1.86134
Dis=corr*Clas=3nn       0.00917647    0.129328     1.00922
Dis=corr*Clas=5loo     0.364569     0.126336     1.43989
Dis=corr*Clas=5nn       0.0317999     0.127666     1.03231
Dis=corr*Clas=7loo     0.134646     0.120905     1.14413
-----

Percentage of deviance explained by model = 98.7064
Adjusted percentage = 96.0086

```

Figure 6.2: Computed logistic regression model for centile-LBP.

```

Logistic Regression

Dependent variable: accuracy/100
Sample sizes: n_tests
Factors: Spa, Dis, Clas

Estimated Regression Model (Maximum Likelihood)
-----
Parameter                Estimate          Standard      Estimated
                        Error              Odds Ratio
-----
CONSTANT                  2.12407          0.0905594
Spa=grey                  -0.50628         0.114461     0.602733
Spa=lab                   0.188014        0.129802     1.20685
Spa=luv                   -0.306623       0.117977     0.735928
Dis=chi                   0.105861        0.0370227    1.11167
Dis=corr                  0.147246        0.0373627    1.15864
Clas=1loo                 1.45345         0.148945     4.27785
Clas=1nn                  0.79341         0.152151     2.21092
Clas=3loo                 1.26328         0.140771     3.537
Clas=3nn                  0.514671        0.140611     1.67309
Clas=5loo                 1.14671         0.136281     3.14783
Clas=5nn                  0.340032        0.134495     1.40499
Clas=7loo                 0.661778        0.121283     1.93824
Spa=grey*Clas=1loo       -0.239453       0.186357     0.787058
Spa=grey*Clas=1nn        -0.311246       0.190007     0.732534
Spa=grey*Clas=3loo       -0.38216        0.174727     0.682386
Spa=grey*Clas=3nn        -0.148233       0.179055     0.86223
Spa=grey*Clas=5loo       -0.466824       0.16865      0.62699
Spa=grey*Clas=5nn        -0.124236       0.172121     0.883171
Spa=grey*Clas=7loo       -0.120554       0.155215     0.886429
Spa=lab*Clas=1loo        0.120522        0.224951     1.12809
Spa=lab*Clas=1nn         -0.334304       0.213777     0.715836
Spa=lab*Clas=3loo        0.198711        0.215086     1.21983
Spa=lab*Clas=3nn         -0.166218       0.202645     0.846862
Spa=lab*Clas=5loo        -0.00792229     0.201052     0.992109
Spa=lab*Clas=5nn         -0.15237        0.194394     0.858671
Spa=lab*Clas=7loo        0.19113         0.183204     1.21062
Spa=luv*Clas=1loo        -0.868065       0.182459     0.419763
Spa=luv*Clas=1nn         -0.502609       0.192278     0.60495
Spa=luv*Clas=3loo        -0.652074       0.176162     0.520964
Spa=luv*Clas=3nn         -0.328953       0.181608     0.719677
Spa=luv*Clas=5loo        -0.674476       0.170968     0.509423
Spa=luv*Clas=5nn         -0.192762       0.176328     0.824679
Spa=luv*Clas=7loo        -0.324074       0.157786     0.723197
-----

Percentage of deviance explained by model = 98.5128
Adjusted percentage = 93.2502

```

Table 6.3: Thirty best combinations of factors using colour histograms. Predicted accuracy (p*100) and confidence intervals are given in %.

N° Exp.	Space	Distance	Classifier	p*100	c.i. 95%
79	grey	chi	5loo	93.23	[91.44, 94.75]
49	rgb	chi	1nn	93.24	[90.72, 95.96]
52	rgb	chi	7nn	93.55	[91.20, 95.74]
39	luv	corr	5loo	93.90	[92.26, 95.41]
78	grey	chi	3loo	94.00	[92.38, 95.50]
16	lab	corr	7loo	94.26	[96.62, 95.68]
51	rgb	chi	5nn	94.26	[91.92, 96.26]
77	grey	chi	1loo	94.47	[92.85, 95.87]
32	luv	chi	7loo	94.57	[92.97, 95.96]
50	rgb	chi	3nn	94.58	[92.16, 96.43]
64	rgb	corr	7loo	94.76	[93.21, 96.14]
38	luv	corr	3loo	95.32	[93.81, 96.60]
15	lab	corr	5loo	95.48	[94.05, 96.78]
31	luv	chi	5loo	95.66	[94.17, 96.87]
63	rgb	corr	5loo	96.01	[94.65, 97.23]
8	lab	chi	7loo	96.09	[94.77, 97.32]
37	luv	corr	1loo	96.18	[94.77, 97.32]
30	luv	chi	3loo	96.48	[95.14, 97.59]
14	lab	corr	3loo	96.49	[95.14, 97.59]
56	rgb	chi	7loo	96.57	[95.26, 97.68]
13	lab	corr	1loo	96.57	[95.26, 97.68]
62	rgb	corr	3loo	96.65	[95.38, 97.76]
61	rgb	corr	1loo	96.75	[95.50, 97.85]
7	lab	chi	5loo	96.81	[95.50, 97.85]
55	rgb	chi	5loo	97.30	[96.12, 98.28]
29	luv	chi	1loo	97.32	[96.12, 98.28]
6	lab	chi	3loo	97.38	[96.25, 98.37]
54	rgb	chi	3loo	97.59	[96.50, 98.54]
5	lab	chi	1loo	97.61	[96.50, 98.54]
53	rgb	chi	1loo	97.82	[96.50, 98.54]

Table 6.4: Thirty best combinations of factors using centile-LBP. Predicted accuracy (p*100) and confidence intervals are given in %.

N° Exp.	Space	Distance	Classifier	p*100	c.i. 95%
1	lab	chi	1-nn	94.67	[92.41, 96.60]
56	rgb	chi	7-loo	94.74	[93.21, 96.14]
65	rgb	log	1-nn	94.87	[92.65, 96.77]
9	lab	corr	1-nn	94.88	[92.65, 96.77]
64	rgb	corr	7-loo	94.95	[93.45, 96.33]
77	grey	chi	1-loo	94.97	[93.45, 96.33]
85	grey	corr	1-loo	95.16	[93.69, 96.51]
49	rgb	chi	1-nn	95.36	[93.14, 97.11]
57	rgb	corr	1-nn	95.54	[93.39, 97.27]
24	lab	log	7-loo	95.95	[94.53, 97.14]
71	rgb	log	5-loo	96.34	[95.01, 97.50]
8	lab	chi	7-loo	96.34	[95.01, 97.50]
16	lab	corr	7-loo	96.48	[95.14, 97.59]
55	rgb	chi	5-loo	96.70	[95.38, 97.76]
70	rgb	log	3-loo	96.73	[95.50, 97.85]
63	rgb	corr	5-loo	96.83	[95.63, 97.94]
23	lab	log	5-loo	96.93	[95.63, 97.94]
54	rgb	chi	3-loo	97.05	[95.88, 98.12]
62	rgb	corr	3-loo	97.17	[96.00, 98.20]
7	lab	chi	5-loo	97.23	[96.00, 98.20]
69	rgb	log	1-loo	97.28	[96.12, 98.28]
15	lab	corr	5-loo	97.34	[96.12, 98.28]
53	rgb	chi	1-loo	97.55	[96.37, 98.45]
61	rgb	corr	1-loo	97.64	[96.50, 98.54]
22	lab	log	3-loo	97.76	[96.63, 98.62]
6	lab	chi	3-loo	97.98	[96.88, 98.77]
21	lab	log	1-loo	97.99	[96.88, 98.77]
14	lab	corr	3-loo	98.06	[97.01, 98.87]
5	lab	chi	1-loo	98.19	[97.14, 98.95]
13	lab	corr	1-loo	98.26	[97.27, 99.03]

Figure 6.3: Presence in percentage of colour spaces in best combinations of colour histograms and centile-LBP.

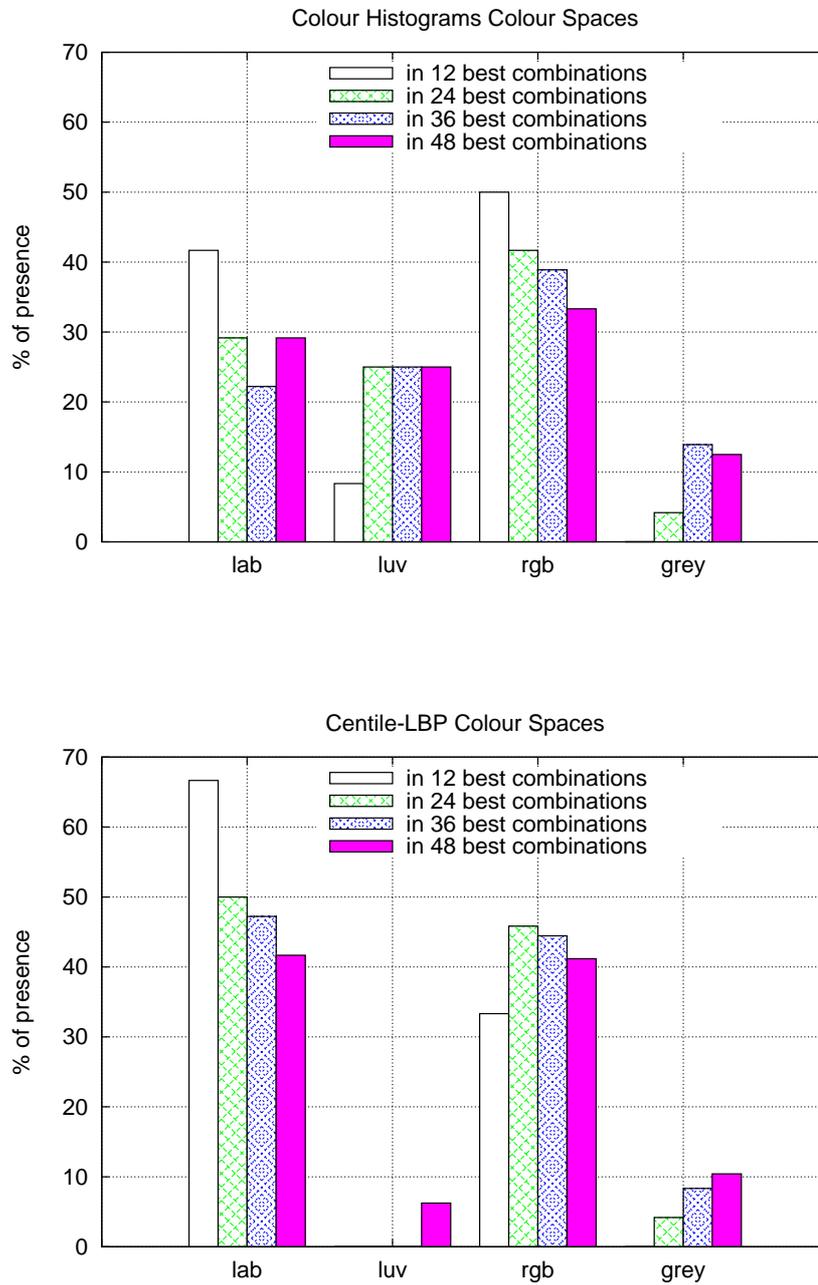


Figure 6.4: Presence in percentage of distances in best combinations of colour histograms and centile-LBP.

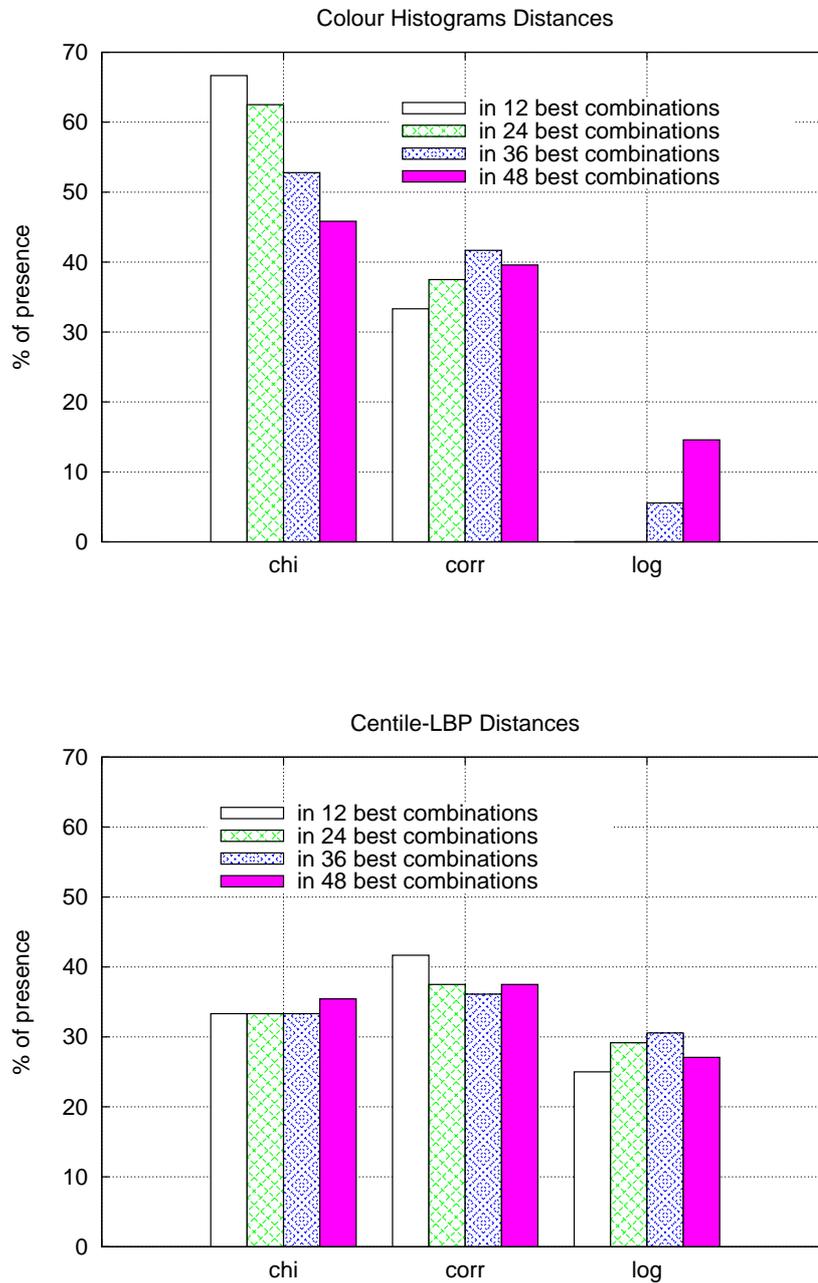


Figure 6.5: Presence in percentage of classifiers in best combinations of colour histograms and centile-LBP.

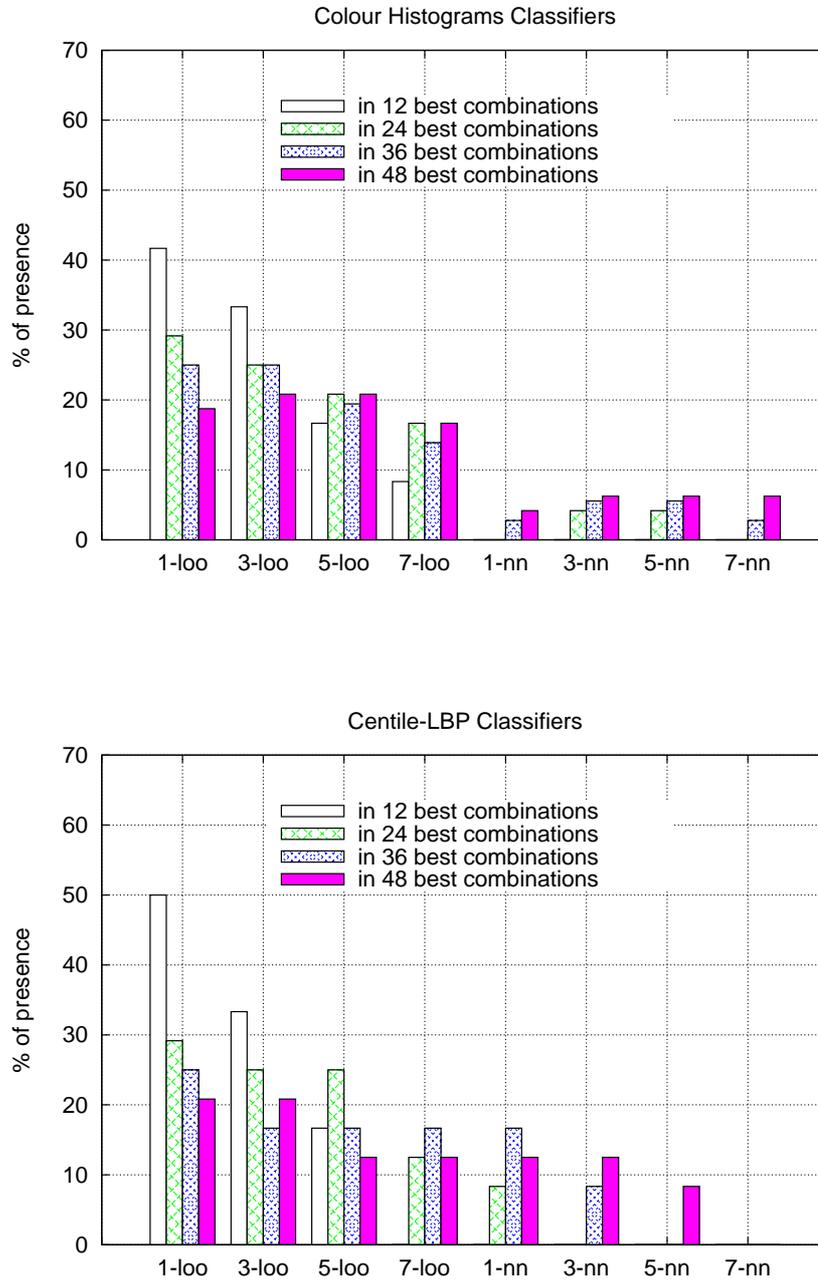


Table 6.5: Best results of surface grading approaches.

	factors	predicted accuracy	c.i. 95%	measured accuracy
Soft colour-texture descriptors	CIE Lab, 1-loo, all descriptors	97.36%	[96.25%, 98.38%]	96.70%
Colour histograms	RGB, 1-loo, Chi square	97.82%	[96.50%, 98.54%]	98.67%
Centile-LBP	CIE Lab, 1-loo, linear correlation	98.26%	[97.27%, 99.03%]	98.25%

6.4 Comparison with *soft colour-texture descriptors* method

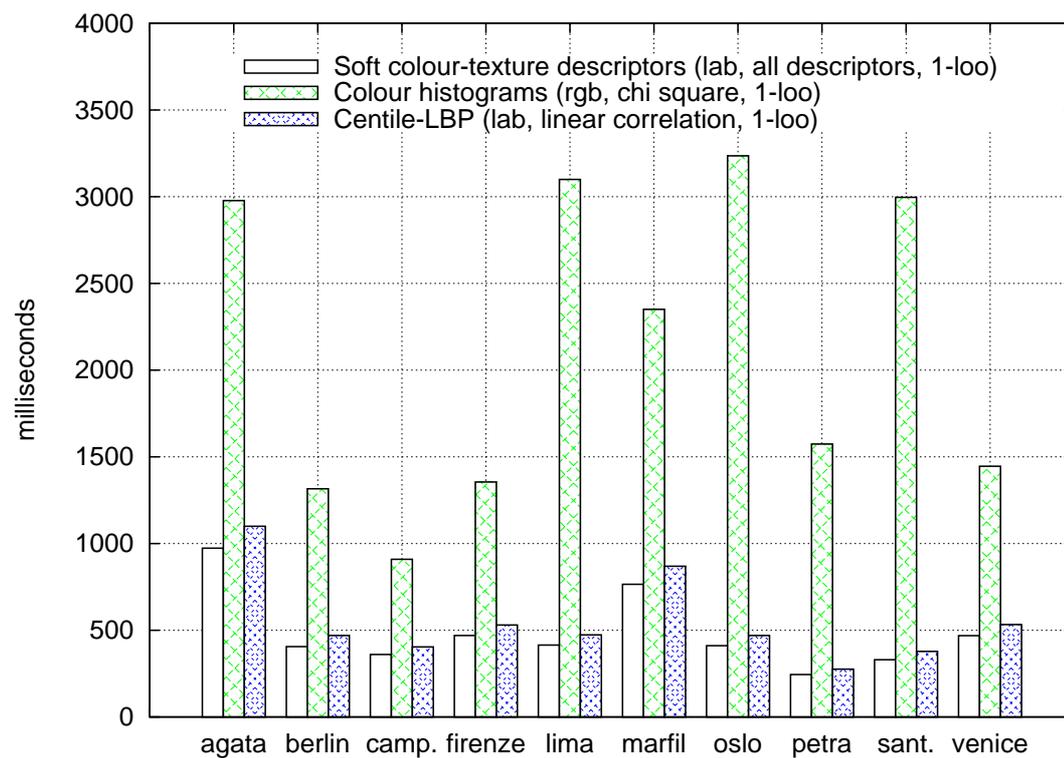
In all methods the achieved performance is very good and quite similar, predicted accuracy varies in less than 1%. Also, for all of them confidence intervals and predicted accuracy exceed factory demands of 95% (see Table 6.5).

Perceptually uniform colour spaces, CIE Lab and CIE Luv, work fine with *soft colour-texture descriptors* while RGB and CIE Lab provide good behavior in colour histograms and centile-LBP approaches respectively.

It is in timing costs where the differences arise among methods. In Figure 6.6 they are compared by timing costs (measured in a common PC) for ten of the fourteen tile models. The soft colour-texture descriptors method provides the best performance in timing costs, closely followed by centile-LBP. The colour histograms approach compile by far the worst timing despite this method does not need to translate the image data, originally in RGB, into CIE Lab space. Also, this method presents irregular timing for the same data size. The berlin, lima and oslo models share the data size (tile and image size) but the method achieves significant timing differences among them. This effect is due to the use of binary trees to store the colour histograms of images. Those images with larger number of different colours will need larger trees and more time to compute the differences between histograms. This timing dependence related to data values does not appear in the other two

methods whose computational costs only depend on the image size and algorithm; $\Theta(n) + C$ where n is the image size and C is a constant related with the algorithm used for implementing the approach.

Figure 6.6: Timing comparison of surface grading approaches using the best combination of factors.



Chapter 7

Study of real-time compliance

This chapter concerns the study of the real time compliance of the inspection system working with the proposed method¹ based on *soft colour-texture descriptors*. In the previous chapter we concluded this approach is faster than the other two methods in the literature. Also, the method achieves very good surface grading performance (97.4%) differing from the centile-LBP result by only 0.9 (centile-LBP is the best in accuracy among the compared approaches).

First, we studied the sequential timing for the worst possible case at factory. This study was done using a modern PC (Pentium IV at 3,2GHz) and it was concluded that parallelization could be necessary in some cases. The parallelization procedure was carry out using a cluster-MPI scheme.

Before parallelization, a preliminary study on the influence of data scaling on accuracy results was performed to determine the real image resolution needed for surface grading. The original images were acquired with a 3.2 pixels per millimetre resolution because the system was also designed to detect small surface defects. Nevertheless, we suspected this resolution might be excessive for the surface grading purpose. The experiments using lower image resolutions demonstrate we were right and a resolution of 0.8 pixels per millimetre is sufficient to surpass the minimum accuracy limit (95%) requested at factory.

¹CIE Lab colour space, 1-loo classifier and all soft colour-texture descriptors.

7.1 Sequential inspection process

First, we carried out a study for determining the real-time compliance using only one computer and then the need for parallelization. We used a modern PC² to study the timing costs of the inspection process, which is divided into four subprocesses:

1. Image acquisition: the tile image was captured using a Dalsa Trillium scan line camera connected to a Coreco-Imaging PC-DIG grabber (see Chapter 3 for more details). Image acquisition includes the time needed to transfer image data from grabber to the computer memory, 111 milliseconds in the worst case corresponding to 50x50cm pieces (2048x1900 RGB images).
2. Tile extraction: this corresponds to segmentation of the tile surface from the background and also tile repositioning and reorientation to facilitate computation of its features (see next section 7.2 and Appendix A).
3. Computation of features: the computation of all soft colour-texture descriptors of the global image.
4. Surface grade classification.

The worst case at factory corresponds with the inspection of 50x50cm pieces (the larger manufactured tiles) which can be produced at 20 pieces per minute. Thus, there are three seconds to perform all the inspection processes for each tile.

As the image acquisition is done using a scan line camera, the image is only completed after all the piece has passed under the camera. The pieces are 20 centimeters apart. As the production ratio is 20 pieces per minute, 140 meters pass under the inspection system every 60 seconds, then, the spatial separation of 0.20 meters between tiles corresponds with 857 milliseconds. If we subtract the time needed to transfer the image from the grabber to the computer memory (111 milliseconds), we have 746 milliseconds to perform the other three inspection processes.

Table 7.1 shows the measured timing costs of the inspection processes. The total time needed to inspect a piece for surface grading clearly exceeds the 3 second time

²Pentium IV at 3,2 GHz. 1GB of memory at 400 MHz.

Table 7.1: Timing costs of sequential inspection processes.

Processes	milliseconds
Acquisition	2254
Tile extraction	1093
Computation of features	2520
Classification	0
Total	5867

limit imposed by the worst case. Therefore, in this cases, parallelization is needed to provide 100% inspection at on-line rates.

Parallelization can be applied using two schemes:

1. By dividing the image data into n sub-images (as many as there are computation nodes involved in the parallel architecture). Each node performs the tile extraction and computation of features to its corresponding sub-image and then returns the computed features.
2. By transferring complete images to each node. In this case, each node extracts the tile surface and computes the features of complete images. When a node finalizes, it returns the computed features to the Master node and classification is then performed. The Master node is also assigned the task of acquiring tile images (see Figure 7.4). When the Master has a new image to inspect it uses a simple algorithm to determine which node is free in the cluster.

Parallelization cannot be applied to acquisition task neither classification. These two tasks will be performed only by one node. Also, if we study the approach used for tile extraction (see next section and Appendix A), it is easy to realize that in the first scheme for parallelization tile extraction is a nonsense operation when using sub-images. Tile extraction cannot be performed with partial images. Thus, in the first case, the parallelization is only possible for the computation of features. But, although parallelization would reduce the time for the computation of *soft colour-texture descriptors*, it is not enough to achieve real-time compliance because the

Figure 7.1: Acquired tile presenting a slightly inclination.



time for acquisition and tile extraction exceeds the 3 second limit. Therefore we chose the second approach to parallelization because in this case complete images are used and thus tile extraction can be included in parallelization.

7.2 Tile extraction

Once the image has been acquired, the tile must be extracted from the background and fitted to a non-inclined square so the features can be properly computed. When tiles pass under the acquisition system, a perfect alignment of the piece with the scan line camera is not assured. Usually the images present an aspect similar to that shown in Figure 7.1, where the tile is slightly rotated or inclined.

We used an image registration methodology in order to perform tile extraction and its adjustment to a non-inclined rectangle. This methodology was taken from a development that we performed in previous works for the detection of defects in fixed pattern ceramic tiles [20–23]. The registration method is explained in detail in Appendix A of present document.

The aim of the tile extraction process is to obtain images of tiles free of background and inclination (see Figure 7.2). After this, the image is ready for the

Figure 7.2: Tile extracted from background, positioned at origin (left-upper corner) and registered with a non-inclined rectangle.



computation of the soft colour-texture descriptors which are only concerned with tile surface.

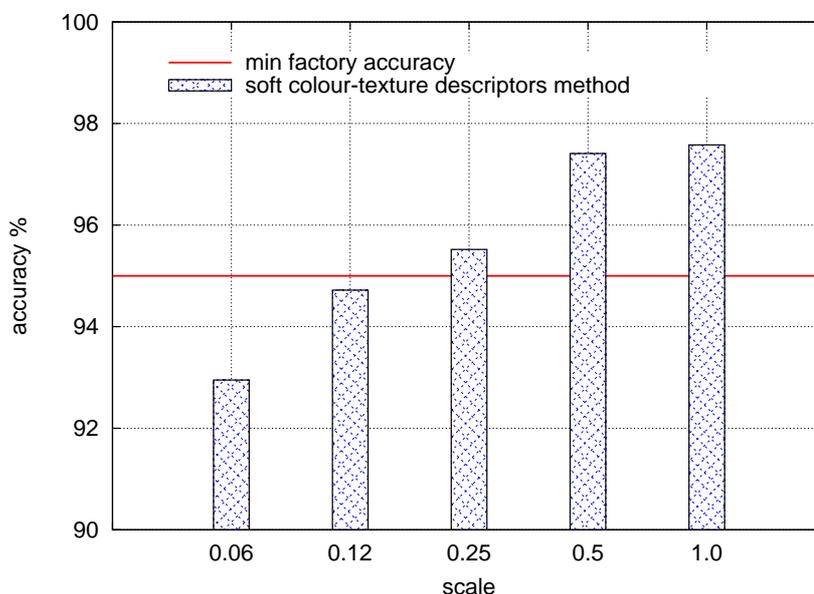
7.3 Data scaling vs real-time compliance

Once we extracted the best surface grading method based on *soft colour-texture descriptors*, we planned a study on the influence of the image resolution on accuracy as a first approximation to the issue of real-time compliance.

For the experiments in Chapters 5 and 6 we used the original resolution of the VxC TSG images, 3.2 pixels per millimetre. This high resolution was selected because the system was also designed to detect small surface defects. However, we thought this resolution might be excessive for the purpose of surface grading. For surface grading we need measures of global appearance rather than fine local information.

As the data size is a primary factor in the computational costs, we studied the evolution of accuracy using smaller image resolutions. We repeated classification of the VxC TSG models using the method extracted in Chapter 5 and different image resolutions or scales. We used the scales 1.0, 0.50, 0.25, 0.12, 0.06 and measured the accuracy over all models computing the achieved mean accuracy.

Figure 7.3: Accuracy versus scale using the extracted *soft colour-texture descriptors* method.



The influence of the scale on accuracy is shown in Figure 7.3. Scales 1.0 and 0.25 are only separated by a loose of 2% in accuracy. The 0.25 scale exceeds the factory limit using an amount of data 16 times smaller (more than one order of magnitude). Thus, the resolution could be reduced from 3.2 to 0.8 pixels per millimetre without a significant loose of discriminant power (95.5%) complying with the factory requirement for grading performance. Also interesting is the result obtained using the 0.5 scale. Here, with an amount of data four times smaller, the accuracy almost remains equal (only droops 0.2%, from 97.6% to 97.4%). A resolution of 1.6 pixels per millimetre can be also a good choice to achieve real-time compliance.

The previous timing table (Table 7.1) is up-dated to take into account the new image resolutions as shown in Table 7.2.

With the improvement introduced using the new resolutions the parallelization approach is not needed if a resolution of 0.8 pixels per millimetre is chosen. Nevertheless, if we prefer to give priority to the grading performance or add more surface inspection tasks, a parallelization study would be useful.

Table 7.2: Timing costs of sequential inspection processes using new image resolutions.

Processes	3.2 (pixels/mm)	1.6 (pixels/mm)	0.8 (pixels/mm)
Acquisition	2254	2254	2254
Tile extraction	1093	250	47
Computation of features	2520	630	150
Classification	0	0	0
Total time (in milliseconds)	5867	3134	2451

7.4 MPI-Cluster architecture

For parallelization we chose the architecture based on the union of the cluster hardware and the MPI (Message Passing Interface) software. From the point of view of parallel architectures, a cluster is a collection of complete computers with dedicated interconnects. Typically, all the machines in the cluster are set up to function identically, they have the same instruction set and operating system. Cluster building technology is usually standard; microprocessors, and interconnection networks. Clusters are more cheaper than specialized parallel systems and they provide easy and affordable access to parallel processing.

Clusters were first used to serve multiprogramming workloads. In multiprogramming clusters, a single front-end machine usually acts as an intermediary between a collection of compute servers and a large number of users at terminals on remote machines. However, clusters are increasingly being used as parallel machines, often called *networks of workstations* (NOWs) [126]. A major influence on clusters has been the increase in popular domain software, such as PVM (Parallel Virtual Machine) [127] and MPI (Message Passing Interface) [128]. This allows users to farm jobs over a collection of machines or to run a parallel program on a number of machines connected by a local area network.

Both, PVM and MPI are software libraries for parallelization using clusters and C (also Fortran) standard programs. We decided to use the MPI software because it

is oriented to achieving higher communication performance when the same kind of hosts are connected [129] (homogeneous networks). This is our case, and also much of the parallelization scheme is based on network transfer performance.

Including a cluster to increase surface inspection capabilities means that the previous inspection system architecture has to be re-designed (see Chapter 3) as shown in Figure 7.4.

The machine connected with the camera and sensors through the frame grabber and I/O card is the Master node and, once the image of one tile is acquired, it manages the distribution of work among the nodes of the cluster and also collects the results. When the soft colour-texture descriptors are available the Master node performs the classification and finally grade the ceramic tile.

When the Master node has a new tile image it uses a simple algorithm to determine which node is free in the cluster. Then, the image is sent to this node which carries out the tile extraction and computation of features. Finally, the node returns the computed features to the Master which performs the final classification or surface grading.

7.5 Parallelization experiments and results

For the first experiment we used the 'Mercurio' cluster located at the GAP Laboratory (Parallel Architectures Group) in the Polytechnic University of Valencia. This machine is formed by 21 nodes, one of which is the Master node. Each node is equipped with a Pentium III bi-processor at 1GHz and 1GByte of RAM memory. All nodes are interconnected using a high performance Ethernet network with a 1Gb bandwidth. The software for parallelization installed in the cluster was the MPI version 1.2.

The experiment consisted in measuring computing times when using a growing number of nodes. The parallel algorithm corresponds with the second scheme of parallelization (see section 7.1), based on the distribution of whole images to the free nodes in the cluster. The nodes extract the tile and compute of features. The soft colour-texture descriptors are sent to the Master node which finally performs

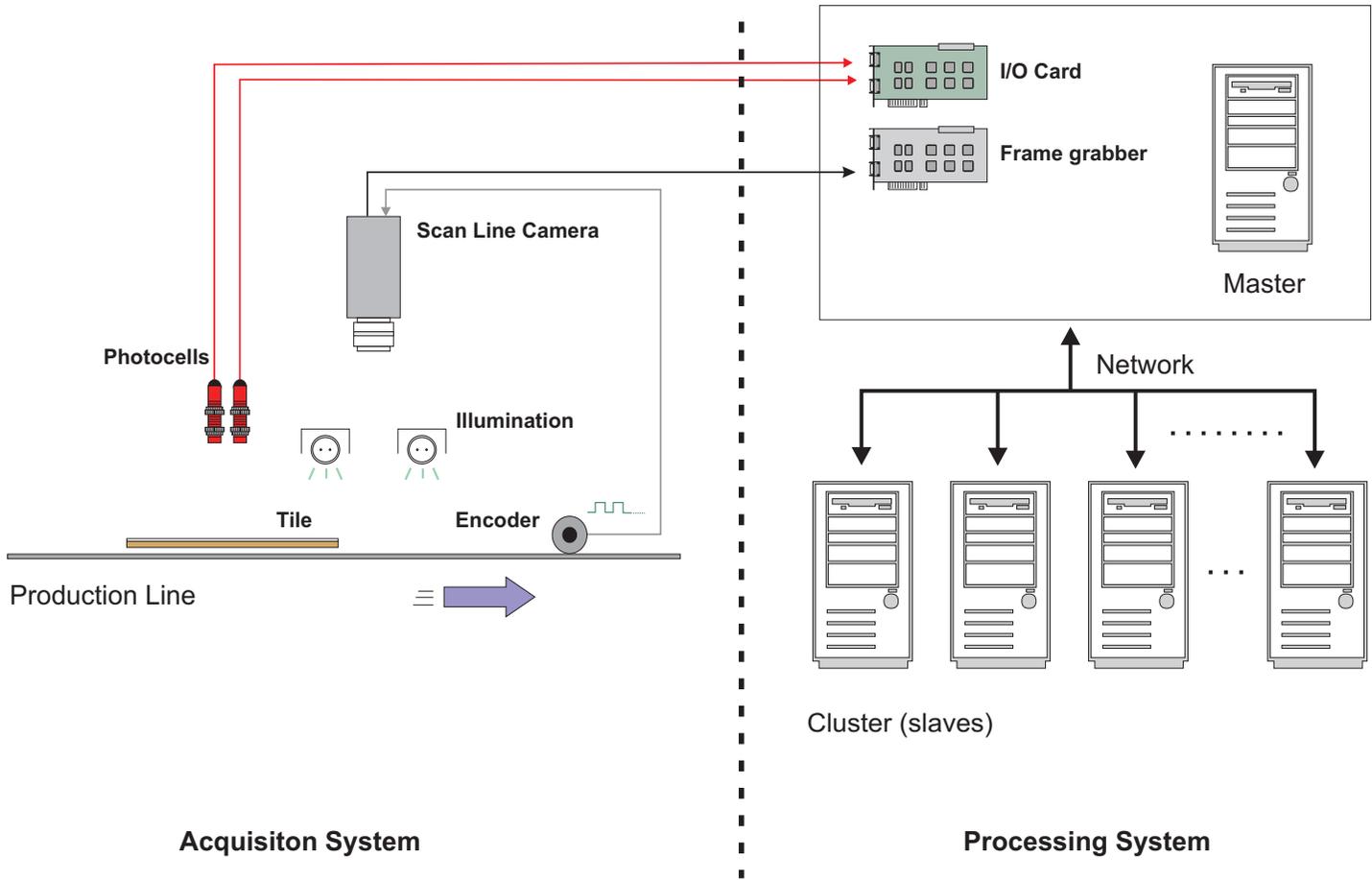


Figure 7.4: The inspection system with a cluster processing unit for parallelization.

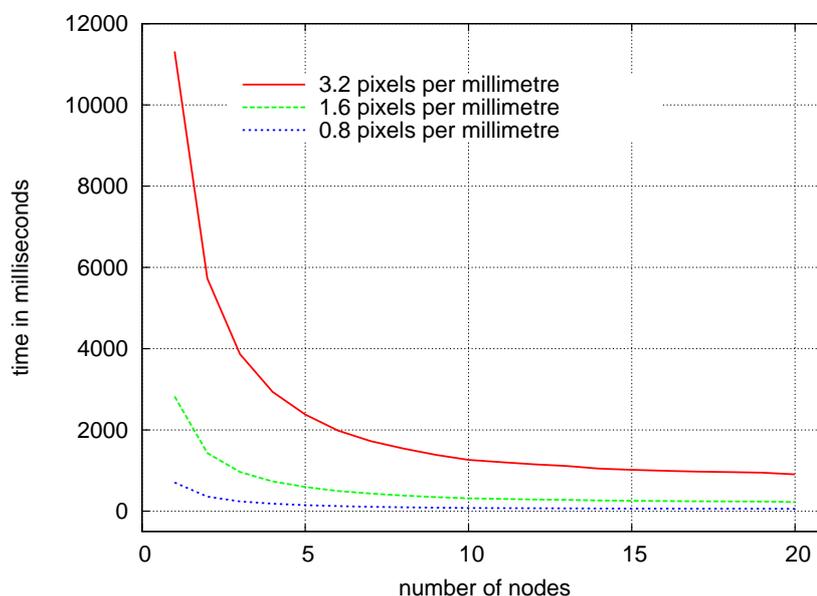
Table 7.3: Timing of parallelization in Mercurio cluster using different image resolutions (time in milliseconds).

Number of Nodes	3.2 (pixels/mm)	1.6 (pixels/mm)	0.8 (pixels/mm)
1	11321	2825	704
2	5727	1431	357
3	3867	964	240
4	2935	732	183
5	2379	596	148
6	1989	496	124
7	1725	434	108
8	1541	385	96
9	1386	346	86
10	1261	315	78
11	1207	301	75
12	1152	288	72
13	1113	278	69
14	1048	262	65
15	1018	254	63
16	995	248	62
17	974	243	60
18	963	240	60
19	946	236	59
20	907	226	56

the surface grade classification. Table 7.3 and Figure 7.5 show the results for image resolutions of 3.2, 1.6 and 0.8 pixels per millimetre.

Figure 7.5 shows that parallelization reaches saturation when we use more than 10 nodes. We achieve success in parallelization when the required time is under 746 milliseconds which is the time remaining after the image acquisition at the Master node. From this point of view, the 3.2 resolution does not succeed because the time needed using 20 nodes is 907 milliseconds. Using more nodes could lead to achieve times under the limit of 746 milliseconds, but a cluster of 20 nodes (or more) is far too expensive, both economically and spatially. The prototype at factory should be equipped with a reasonable number of nodes to take into account economic and spatial costs. The best relation between accuracy and number of nodes is achieved using four nodes and a resolution of 1.6 pixels per millimetre. This configuration is

Figure 7.5: Timing evolution in Mercurio cluster using several image resolutions.



under the time limit (746 milliseconds) and uses a reduced number of nodes. More nodes, or a resolution of 0.8 pixels/mm, can be chosen if we plan to introduce more inspection tasks in addition to surface grading. The computational costs of these new tasks should be studied in conjunction with the surface grading application.

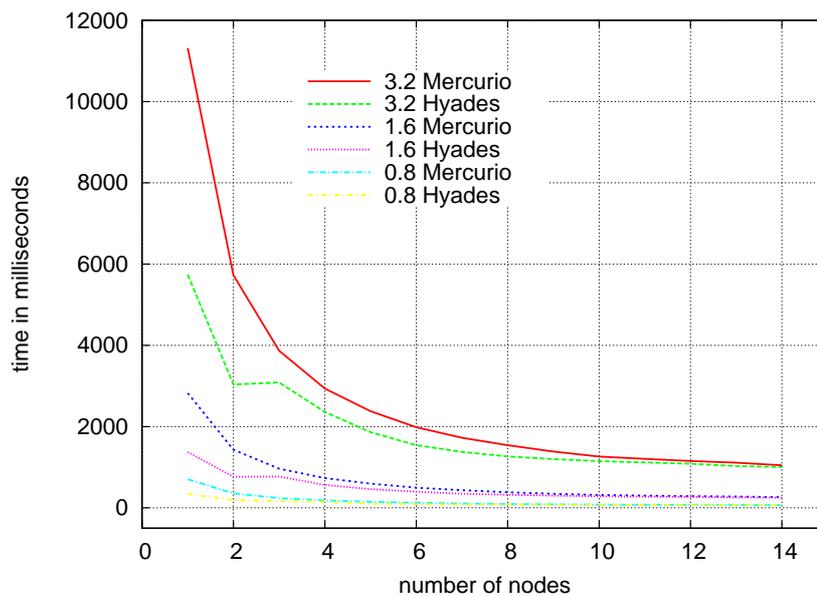
In a second experiment we tested the surface grading application using a high performance cluster; the 1350 IBM cluster. This machine is called 'Hyades' and is located at the Computing Centre of the Polytechnic University of Valencia. This cluster, used for supercomputing tasks, has 60 nodes equipped with Intel Xeon bi-processors at 2,4 GHz and 1GB of RAM memory. The nodes are interconnected using a Myrinet which provides a bandwidth of 2Gb per second. This cluster is an IBM machine made for supercomputing applications and is much more expensive than Mercurio. With this experiment we tried to determine if it is worthwhile to use high performance clusters for the application of surface grading. Only 14 nodes were available when we carried out our experiments, but that was enough for comparison purposes (see Table 7.4).

From the timing comparison of Figure 7.6 we can conclude that Hyades performs significantly better when using a reduced number of nodes, however, the timing

Table 7.4: Timing of parallelization in Hyades machine using several image resolutions (time in milliseconds).

Number of Nodes	3.2 (pixels/mm)	1.6 (pixels/mm)	0.8 (pixels/mm)
1	5739	1373	341
2	3037	765	193
3	3130	770	166
4	2362	563	147
5	1862	461	101
6	1542	399	91
7	1376	345	84
8	1266	321	85
9	1201	302	74
10	1151	285	71
11	1118	276	68
12	1084	269	66
13	1030	256	64
14	1003	251	62

Figure 7.6: Timing comparison between Mercurio and Hyades clusters using several image resolutions.

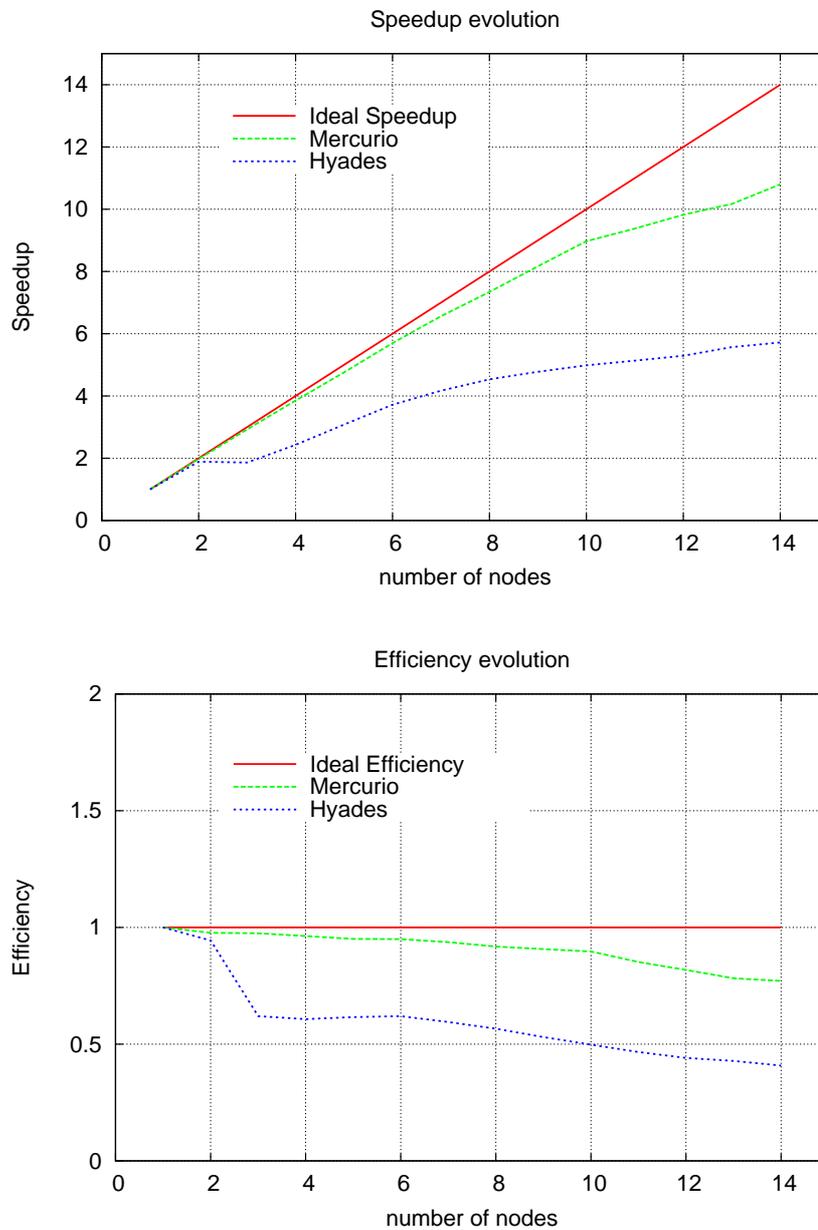


costs of both machines quickly become similar from the third node on. Individually a Xeon bi-processor clearly performs better than a Pentium III bi-processor in performance, but this individual advantage does not transfer to the cluster scheme of parallelization.

Mercurio and Hyades machines were compared also using the Speedup and Efficiency which are two classical measures for characterizing parallelization results (see Figure 7.7). Speedup is the relation between the time needed to carry out the task in only one node and the time needed for the same task using n nodes. The Efficiency is the relation between the real and the ideal Speedup. The ideal Speedup using 2 nodes is 2, two nodes should double system speed, three nodes should treble and so on.

Figure 7.7 shows clear advantage of Mercurio cluster which achieves better performance in Speedup and Efficiency. We conclude that high performance equipment Hyades does not provide enough benefits to justify the investment costs.

Figure 7.7: Speedup and Efficiency in Mercurio and Hyades.



7.6 Conclusions

In this chapter we have studied the real-time compliance of the surface grading method proposed in this thesis, which is also the fastest among the compared methods. The method was extracted in Chapter 5 and is based on *soft colour-texture descriptors* using the CIE Lab colour space, the 1-leaving-one-out classifier and all

the descriptors.

We carried out all the studies with reference to the worst case at factory; 50x50cm tiles at a ratio of 20 tiles per minute. This gives 3 seconds to inspect and grade each tile. Firstly, we studied the method's timing requirements with a standard modern computer. The result was unsuccessful because the computing time required for only one tile was of 5.867 seconds which exceeds the limit.

Secondly, we studied the influence of the image resolution on accuracy. The original resolution was high, 3.2 pixels per millimetre, because the system was designed not only for surface grading but also for detecting small defects. We tested lower resolutions and achieved an accuracy over 95% for resolutions of 1.6 and 0.8 pixels per millimetre. The corresponding computing times were of 3134 and 2451 milliseconds respectively. Thus, 0.8 resolution complies with real-time requirements.

Although parallelization is not strictly needed we decided to perform this study because it could be of interest if we want to prioritize accuracy performance or plan to add more inspection tasks to the system. We chose the cluster-MPI scheme because it provides easy, affordable access to parallel processing. We first carried out the experiments with a medium-level performance cluster called Mercurio. Mercurio is composed of Pentium III bi-processors at 1GHz using a standard 1Gb Ethernet for interconnection. The timing results showed that the best relation between number of nodes and accuracy was achieved using four nodes and a resolution of 1.6 pixels per millimetre. Four is a reasonable number of nodes for the cluster if we think in terms of economic and spatial costs. The space occupied by the inspection system at factory is an important issue because more space makes the system less operative. Finally we repeated the surface grading experiments in a high performance machine called Hyades which was composed of Xeon bi-processors at 2.4 GHz and one Myrinet network at 2Gb per second. In timing, Hyades only achieved clear better performance using less than 3 nodes. Mercurio surpassed Hyades in two classical measures of parallelization performance; Speedup and Efficiency. The difference in cost between both systems and the achieved results led us to conclude that it is not worthwhile investing in a high performance cluster.

Chapter 8

Conclusions, discussion and further work

In this thesis we have presented a case of study of the development and performance analysis of a surface grading application with real-time compliance. The application focuses on the ceramic tile industry and aims to automate the inspection process of surface grading and removing human inspectors from this subjective and tedious task. From the overview of surface grading works we concluded that many of them were very specialized in a specific type of surface, others did not achieve good enough accuracy, and yet others did not take into account the real-time restrictions of a factory inspection. Therefore, we thought surface grading was still an open issue where more contributions were possible. In this sense, this thesis dealt with less explored aspects related to real-time compliance and surface grading performance.

In Chapter 3 we addressed the question of spatial and temporal uniformity in the acquisition system. We performed a study of spatial and temporal uniformity comparing two modern illumination systems; fluorescents and white LED arrays. The results showed that only fluorescent systems were able to provide sufficient uniform response for the application of surface grading. In this chapter was also presented the VxC TSG (VxC Tiles for Surface Grading) which is an image database of ceramic tiles oriented to surface grading. Creating and compiling this database was one of the thesis goals. It has been the ground truth used for testing and comparing surface grading approaches presented in Chapters 5 and 6. VxC TSG is an extensive

image data base of ceramic tile models representing the wide range of surface classes in ceramic tiles, and also, it is intended to be a tool for the scientific community and future works in the field of surface grading (*miron.disca.upv.es/vision/vxctsg/*).

In Chapter 4, first works seeking for a fast and accurate method of surface grading were revised. Here, we presented the preliminary works that finally gave rise to the *soft colour-texture descriptors* method. First, we studied methods based on image tessellation and local statistics of colour. These methods did not achieved good enough accuracy and only one of them, using multiresolution techniques, achieved factory compliance exceeding 95% of minimum accuracy. However, all these methods have an important drawback, they need to non-automatically estimate at least one parameter for every tile model. Then, we proposed a generic method with no parameter estimation needed. This method used global statistics of colour and texture computed in a perceptually uniform colour space, the CIE Lab. Global statistics in CIE Lab achieved factory compliance in accuracy improving also two literature methods; colour histograms and centile-LBP. In these works we used a medium-sized image database of ceramic tiles, the antecedent of VxC TSG.

Global statistics computed in CIE Lab is the antecedent of *soft colour-texture descriptors* method which was in-depth studied in chapter 5. Both, global statistics in CIE Lab and *soft colour-texture descriptors* are basically the same method. Once we found a fast method able to comply with factory requests, we decided to in-depth study the approach using several factors such as colour spaces (CIE Lab, CIE Luv, RGB and Grey scale), classifiers (k-NN and leaving-one-out) and testing all the possible combinations of soft colour-texture descriptors (mean, standard deviation, and histogram moments from 2nd to 5th). To do so, statistical tools were used in order to manage the large amount of experiments and achieve objective conclusions. These statistical tools were the *design of experiments* and the *logistic regression*. They provided a procedure with which to study and determine the best combination of factors to maximize the accuracy rate. The extracted method achieved 97.4% of predicted mean accuracy and 97.6% of measured mean accuracy over all models. The computed confidence interval was [96.25%, 98.36%] which clearly surpasses the factory requirement for performance. From the in-depth statistical study it

was also concluded that RGB, a common colour space non-perceptually uniform, provides less discriminative power, all the soft colour-texture descriptors contribute to achieve accuracy performance, and also the classifiers derived from the leaving one out method perform better.

In Chapter 6, two methods from the surface grading literature were implemented and tested for comparison purposes. These methods were colour histograms and centile-LBP. We chose these methods from literature because they were similar to ours; they are generic solutions with low computational costs. With these methods was also performed a statistical analysis using again *experimental design* and *logistic regression*. The results showed that all approaches are almost equal in accuracy performance if we compare with *soft colour-texture descriptors* method. However, *soft colour-texture descriptors* method achieved better results in timing costs. Colour histograms achieved 97.8% in mean predicted accuracy and 98.3% the centile-LBP method. Predicted accuracies and confidence intervals of both approaches exceeded factory demands of 95%.

The centile-LBP method from literature did not use the overall texture and colour appearance of surfaces to grade wood boards. The grade of wood pieces was assigned by detecting the wood defects and then applying grade rules related to the number and classes of defects found. Therefore, the problem became a question of separating the surface into sound and faulty wood, and classifying the defects into their different types. However, we should point out that we used this method not in the literature sense but to compile global description of colour and texture and then directly grade surfaces. Thus, the fact that centile-LBP achieved the best performance in accuracy and also closely followed *soft colour-texture descriptors* method in timing costs is an interesting result.

Finally, in Chapter 7 we presented a study of real-time compliance including the parallelization of the method extracted in Chapter 5. The study was performed having as reference the worst case at factory; 50x50cm tiles at a ratio of 20 tiles per minute. Thus, we had 3 seconds to inspect and grade each tile. Firstly, we studied the timing requirements of the method using a standard modern computer. The result was unsuccessful because the computing time required for only one tile

was of 5.867 seconds surpassing the 3 seconds limit. Then, we performed a study about the influence of the image resolution on accuracy. The original resolution was high, 3.2 pixels per millimetre, because the system was designed not only for surface grading but also for detecting small defects. We tested lower resolutions and achieved an accuracy over the limit of 95% for the resolutions of 1.6 and 0.8 pixels per millimetre. The corresponding computing times were of 3124 and 2451 milliseconds respectively. Therefore, 0.8 resolution complied real time requirements.

Although parallelization was not strictly needed we decided to perform this study because it could be of interest if we want to prioritize accuracy performance or plan to add more inspection tasks to the system. We chose the cluster-MPI scheme because it provides easy, affordable access to parallel processing. We first carried out the experiments with a medium-level performance cluster called Mercurio. Timing results showed that the best relation between number of nodes and accuracy was achieved using four nodes and a resolution of 1.6 pixels per millimetre. We repeated the surface grading experiments in a high performance machine called Hyades. In timing, Hyades only achieved clearly better performance using less than 3 nodes. Mercurio surpassed Hyades in two classical measures of parallelization performance; Speedup and Efficiency. The difference in cost between both systems and the achieved results led us to conclude that it was not worthwhile to invest in a high performance cluster.

In previous paragraphs we have summarized thesis conclusions. Now, we proceed to discuss some issues:

- At the beginning of the thesis, in Chapter 1, is established that the different criteria of each operator (human grader) regarding tile defects could produce a non-uniform quality control criterion. This assertion seems to be corroborated with a study in the area of wood grading where in a test of four grades different human graders agreed only in 60% of the samples [1]. However, we achieved very high accuracy performance in experiments using an image database of tiles previously graded by human operators at factory. Thus, the idea of a non-uniform quality control criterion of human operators cannot be established as a general rule, at least in the surface grading of ceramic tiles.

- In Chapter 4 the method based on global statistics in CIE Lab improved literature methods in accuracy. Also, literature approaches (colour histograms and centile-LBP) did not achieve factory compliance because their confidence intervals performed under the minimum factory demand on accuracy (95%). However, when we carried out the in-depth statistical studies using the large VxC TSG image database, all the methods were very good and quite similar in performance achieving factory compliance. We think this was due to the following factors: a larger image database was used, the acquisition system was better calibrated for the acquisition of VxC TSG and also the images were improved for each tile model using brightness and contrast control.
- We used perceptually uniform colour spaces (CIE Lab and CIE Luv) because we thought they will perform better than non-uniform spaces (RGB) as they are closer to human perception of colours. However, from the experiments, although perceptually uniform spaces showed good behaviour, they do not always perform better (p.e. colour histograms perform better using RGB).
- In this thesis we have presented and tested general methods for the purpose of surface grading. Nevertheless, the ground truth has been comprised only of samples proceeding from the ceramic tile area and it could be interesting to test the approaches using other surface types.
- The studied approaches are able to differentiate surface grades in 'a posteriori' study once the ceramic tiles have been graded. However, at factory is more interesting a method able to automatically determine when a 'new grade' appears at production line. This would involve the use of thresholds with a surface grade measure (i.e. distance in a space of features describing colour and texture).

Further work would include the following items:

- Extend the image database adding more types of surface, i.e. natural surfaces such as marble, granite or wood and also other artificial stuff such as textile fabrics.

- Test the studied methods with the new surface types.
- Study and develop methods able to detect changes of surface grades at production lines.
- Develop a prototype to test the system and methods under real conditions at factory.

Appendix A

Image registration method for ceramic tiles

A.1 Introduction

This appendix presents an image registration method carried out for the purpose of ceramic tiles inspection. Tile registration is the first step in detecting surface faults when a 'compare with a reference' technique is used. The presented method is based on edge detection techniques, used to extract the bounding rectangle defined by tiles. This method only uses a reduced set of border tile points, obtained with the minimum cross entropy thresholding algorithm, and fits them to straight lines using the least-squares method. An outlier detection test is included to eliminate the influence of background noisy points or tile border defects. The bounding tile rectangle is then registered with respect to a reference tile using simple geometrical mapping. Several experiments were made to show the feasibility of the method from the point of view of registration quality.

In the application of surface grading the registration method is applied to perform the tile extraction from background and fit the tile boundary into a non-inclined rectangle. This is necessary for proper computation of colour-texture features. As a reference we used a rectangle with the dimensions of the original tile, with the upper-left corner positioned at origin and without any inclination.

A.2 Registration method

In [20, 21] we studied several approaches for tile edge detection and boundary rectangle extraction. The proposed approach used tile border pixels to adjust the four straight lines that compose the tile rectangle. These border points are obtained by separating the tile from the image background using an optimal threshold level. To do so, the minimum cross entropy thresholding algorithm [131] was selected due to the accuracy of the obtained results.

For any image this algorithm selects the histogram threshold that minimizes the cross entropy between the thresholded image and the original image, in other words, the threshold which minimizes the 'error' between both images.

For every threshold t

$$\eta_1(t) = \frac{\sum_{f_i < t} f_i}{N_1}$$

$$\eta_2(t) = \frac{\sum_{f_i \geq t} f_i}{N_2}$$

where f_i is the histogram value for the i level in the original image and $g_i \in (\eta_1, \eta_2)$ is the thresholded image. The cross entropy between both images can be computed as:

$$\eta(t) = \sum_{f_i < t} f_i \log \left(\frac{f_i}{\eta_1(t)} \right) + \sum_{f_i \geq t} f_i \log \left(\frac{f_i}{\eta_2(t)} \right)$$

The optimum threshold level is selected by:

$$t_o = \min(\eta(t))$$

Using N_P evenly spaced horizontal and vertical scan lines, the outermost points of each tile border are selected creating four sets of N edge points (P_{UP} , P_{DOWN} , P_{LEFT} , P_{RIGHT}). For each set of lateral points, a straight line is fitted using the least-squares method.

Let $P_{UP} = \{p_0(x_0, y_0), \dots, p_{N-1}(x_{N-1}, y_{N-1})\}$ be the set of N_P edge points of the upper side. A line $y = a_{UP} + b_{UP}x$ can be fitted using the standard formulation:

$$y = a_{UP} + b_{UP}x = E(y) - E(x) \frac{\sigma_X^2}{\sigma_{XY}^2} + \frac{\sigma_X^2}{\sigma_{XY}^2} \cdot x$$

Proceeding in a similar way, the remaining lateral lines can be computed (a_{DOWN} , b_{DOWN}), (a_{LEFT} , b_{LEFT}), (a_{RIGHT} , b_{RIGHT}). During this stage, an outlier detection test is applied to eliminate the outlying points that sometimes appear due to the background noise or border tile defects. This process works as follows:

Repeat

1. From the point set P_{UP} computes the line parameters a_{UP} , b_{UP} .
2. For each point $p_i(x_i, y_i)$ computes its distance to the line:

$$D_i = |y_i - a_{UP} + b_{UP}x_i|.$$

If $D_i \leq k\sigma_y^2$ then do nothing, else extract $p_i(x_i, y_i)$ from P_{UP} .

Until (no point is extracted)

Those points lying further than one and a half standard deviations ($k=1.5$) from the computed line are considered outliers, and discharged the line parameters being computed again. The fitting process is repeated, usually two or three iterations, until all the remaining points fulfill the condition.

Figure A.1 shows the progressive approximation of the fitted line to the correct edge points, for the left side. After three steps, the furthest edge points are discharged and the correct line fitting is produced using the remaining points. Figure A.2 shows the results obtained after applying this procedure on several tile images. Some points lying outside the tile area can be seen.

Once the bounding rectangle is obtained, the four corner co-ordinates are computed and the inverse geometric mapping is applied as follows:

$$\begin{pmatrix} x_i^r \\ y_i^r \end{pmatrix} = S \cdot \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \begin{pmatrix} d_x \\ d_y \end{pmatrix}$$

where $p_i(x_i, y_i)$ and $p_i^r(x_i^r, y_i^r)$ $i = 1..4$ are the test and reference tile corners, d_x

Figure A.1: Line fitting process.

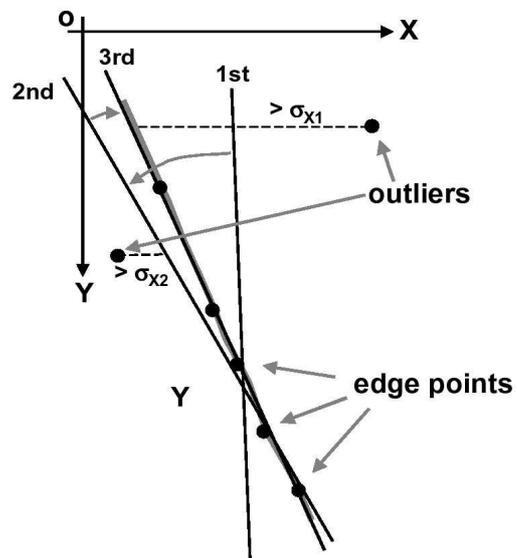


Figure A.2: Tile rectangles obtained on several test images.

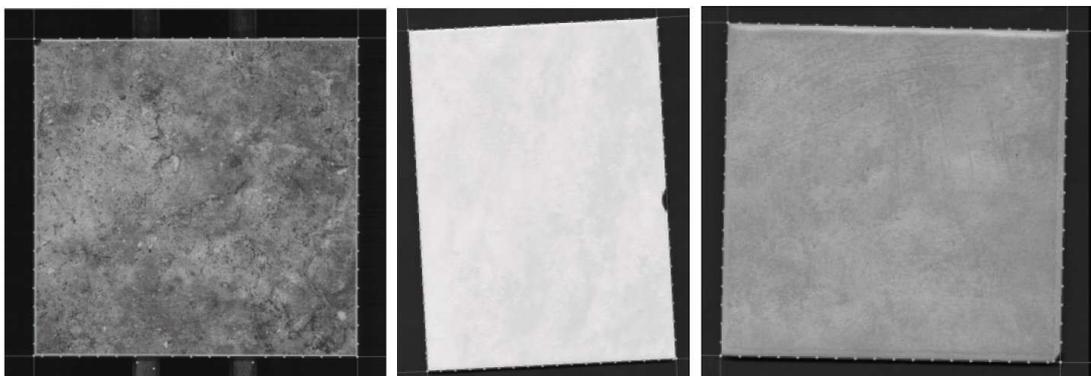
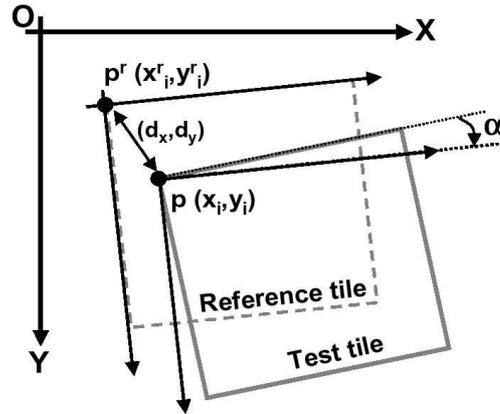


Figure A.3: Transformation mapping from test tile to reference tile.



and d_y are the displacements between the top left points in the two tiles, α is the rotation angle, and S is the scale factor. (O, X, Y) is the image reference coordinate system (see Figure A.3).

Then, it is possible to determine the parameters of the transformation mapping by minimizing the sum of the squared errors:

$$E = \sum_{i=1}^4 [x_i^r - S(x_i \cos \alpha - y_i \sin \alpha) - dx]^2 + [y_i^r - S(x_i \sin \alpha - y_i \cos \alpha) - dy]^2$$

Replacing $(S \cdot \cos \alpha)$ by 'a' and $(S \cdot \sin \alpha)$ by 'b':

$$E = \sum_{i=1}^4 [x_i^r - S(a \cdot x_i - b \cdot y_i) - dx]^2 + [y_i^r - S(b \cdot x_i - a \cdot y_i) - dy]^2$$

Deriving E with respect to a , b , dx , and dy and equaling to zero, the following linear system of equations appears:

$$\begin{bmatrix} \sum_i(x_i^2 + y_i^2) & 0 & \sum_i x_i^2 & \sum_i y_i^2 \\ 0 & \sum_i(x_i^2 + y_i^2) & -\sum_i y_i^2 & \sum_i x_i^2 \\ \sum_i x_i^2 & -\sum_i y_i^2 & n & 0 \\ \sum_i y_i^2 & \sum_i x_i^2 & 0 & n \end{bmatrix} \begin{bmatrix} a \\ b \\ dx \\ dy \end{bmatrix} = \begin{bmatrix} \sum_i(x_i^r x_i + y_i^r y_i) \\ \sum_i(x_i^r x_i - y_i^r y_i) \\ \sum_i x_i^r \\ \sum_i y_i^r \end{bmatrix}$$

from which the optimal parameters of the geometrical transformation can be obtained.

This approach saves a great deal of computing time as it directly calculates the exact inverse transformation needed to registrate test tiles onto reference tiles.

A.3 Experiments and results

Two image data sets were used to validate the method. The first data set was constructed in controlled conditions. Five colour images of different tile models were acquired in the laboratory and were used as the 'fault free' reference data set. For each reference image, a set of geometrical transformations were artificially performed, with x and y displacement ranging from -10 to +10 pixels (step 0.5) and rotations ranging from -2.0° to $+2.0^\circ$ (step 0.25°). As result, a set of $N = 28577$ test images was produced for each tile model.

After applying the registration method, the displacements and rotation parameters (dx , dy , α) were obtained for each test image i , and the mean square errors with respect to the original ones (d_x , d_y , α') were computed as follows:

$$E_x = \sum_{i=1}^N (d_x(i) - d_x(i'))^2 / N$$

$$E_y = \sum_{i=1}^N (d_y(i) - d_y(i'))^2 / N$$

$$E_\alpha = \sum_{i=1}^N (\alpha(i) - \alpha(i'))^2 / N$$

where E_x and E_y are measured in pixels, and E_α in degrees.

Table A.1: Experiments with artificially introduced transformations.

Tile model	E_x (pixels)	E_y (pixels)	E_α (degrees)
firenze	0.191	0.268	0.014
biscuit	0.229	0.166	0.011
toscana	0.230	0.357	0.030
venice	0.214	0.195	0.013
austin	0.291	0.360	0.024
Mean	0.231	0.269	0.018

The achieved results (Table A.1) showed the good behaviour of this method, with a mean in the displacement errors lower than 0.3 pixels and a mean in the rotation errors lower than 0.02° . We used a reduced set of $N_P = 20$ edge points. The scale factor error E_S was negligible ($S = 1$).

A similarity measure was also introduced to compare the resulting registered test image with respect to the reference. The similarity measure E_P used in these experiments was the sum of the absolute differences of both images, that is the mean grey level error per pixel. For comparison purposes, a relative measure ε_P , normalized with the image dynamic range, was also computed as follows:

$$E_P^i = \sum_{x=1}^m \sum_{y=1}^n |f_i(x, y) - f_{:ref}(x, y)| / (m \cdot n)$$

$$E_P = \left(\sum_{i=1}^N E_P^i \right) / N$$

$$\varepsilon_P = 100 \cdot \left(\frac{\sum_{i=1}^N E_P^i}{\Delta f_i} \right) / N$$

where $f_i(x, y)$ is the test image i and $f_{:ref}(x, y)$ is the reference image, (m, n) is the image size, and Δf_i is the dynamic range of the image i . Because all tiles used in this experiment were fault-free, the computed errors were only due to registration errors.

Table A.2 shows the achieved results. These results show that lower relative pixel errors are obtained in the case of bright tiles. Dark tile models produce higher mean errors on the registration parameters and, consequently, greater relative pixel

Table A.2: Absolute and relative mean pixel errors.

Tile model	Δf_i	E_P	$\varepsilon_P(\%)$
firenze	256	3.790	1.49
biscuit	222	3.789	1.67
toscana	184	3.278	1.78
venice	170	3.306	1.92
austin	119	3.498	2.94

Table A.3: Absolute and relative mean pixel errors obtained with real images.

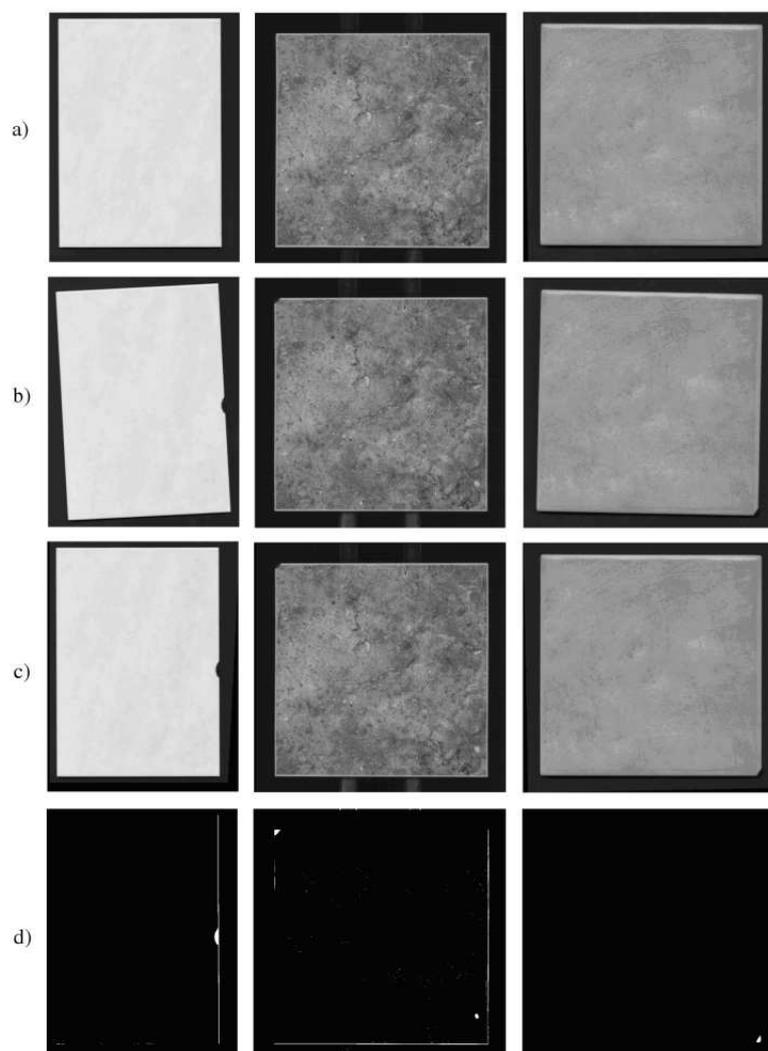
Tile model	Δf_i	E_P	$\varepsilon_P(\%)$
firenze	255	2.337	0.917
biscuit	222	2.211	0.996
toscana	184	2.363	1.284
venice	170	2.343	1.378
austin	119	1.792	1.506

errors. However, the absolute mean pixel error is nearly constant in all cases. This is because higher registration error involves more points that are different in both images. This is, however, compensated by lower grey level differences in the case of dark tiles.

The second image data set was obtained under real conditions on the laboratory line prototype. Eleven tile samples of five fault-free tile models were used. For each model the first tile passing under the camera was considered as the reference tile, and the ten remaining tiles were used as test tiles. As there was no previous knowledge about the translation or rotation parameter of the test tiles with respect to the reference tile, the above similarity measurements were used to compare them. The achieved results are shown in Table A.3.

These results can not be extrapolated due to the use of a reduced set of samples. The method has a stable behaviour with different tile models under real conditions. Figure A.4 shows the results obtained applying the registration method to faulty tiles (spots, sticks). To visualise the quality of the registration, Figure A.4.d shows the error map, that is the difference between the reference tile (A.4.a) and the registered test tiles (A.4.b-c), with a threshold of 50. It can be seen how the defects

Figure A.4: Tile registration examples.



appear clearly in the error map. Continuous lateral errors also appear, in some cases, probably due to small differences in size between reference and test tiles.

Bibliography

- [1] H. Kauppinen. *Development of a color machine vision method for wood surface inspection*. Phd Thesis, Oulu University, 1999.
- [2] C. Boukouvalas. *Colour shade grading and its applications to visual inspection*. Phd Thesis, University of Surrey, 1996.
- [3] C. Boukouvalas, J. Kittler, R. Marik and M. Petrou. *Color grading of randomly textured ceramic tiles using color histograms*. IEEE Trans. Industrial Electronics, 46(1), pp. 219-226, 1999.
- [4] R. Baldrich. *Perceptual approach to a computational colour texture representation for surface inspection*. Phd Thesis, Autonomous University of Barcelona, 2001.
- [5] R. Baldrich, M. Vanrell and J.J. Villanueva. *Texture-colour features for tile classification*. EUROPTO/SPIE Conference on Colour and Polarisation Techniques in Industrial Inspection, Munich (Germany), 1999.
- [6] F. Lumbreras, R. Baldrich, M. Vanrell, J. Serrat, and J.J. Villanueva. *Multiresolution colour texture representation for tile classification*. Recent Research Developments in Optical Engineering, Vol. 2, pp. 213-228, India, 1999.
- [7] F. Lumbreras, R. Baldrich, M. Vanrell, J.Serrat and J.J. Villanueva. *Multiresolution texture classification of ceramic tiles*. VIII Simposium Nacional de Reconocimiento de Formas y Análisis de Imágenes, Vol. 1, pp. 145-152, Bilbao, 1999.

- [8] J.A. Peñaranda and M.A. Serna. *Classification system for pieces of Porcelanatto based on computer vision*. European Symposium on Productivity in Manufacturing, SPIE Vol. 2249, Automated 3D and 2D Vision, pp. 339-348, 1994.
- [9] J.A. Peñaranda, L. Briones and J. Florez. *Colour machine vision system for process control in ceramics industry*. European Symposium on Lasers and Optics in Manufacturing, SPIE Vol. 3101, New Image Processing Techniques and Applications: Algorithms, Methods and Components II, pp. 182-192, 1997.
- [10] H. Kauppinen, H. Rautio and O. Silven. *Nonsegmenting defect detection and SOM-based classification for surface inspection using color vision*. Proc. EUROPTO Conf. on Polarization and Color Techniques in Industrial Inspection, SPIE Vol. 3826, pp. 270-280, Munich, Germany, 1999.
- [11] H. Kauppinen. *Two stage defect recognition method for parquet slab grading*. Proc. 15th International Conference on Pattern Recognition, pp. 803-806, Barcelona, Spain, 2000.
- [12] J. Kyllönen and M. Pietikäinen. *Visual inspection of parquet slabs by combining color and texture*. Proc. IAPR Workshop on Machine Vision Applications (MVA'00), Tokyo, Japan, pp. 1987-1992, 2000.
- [13] V. Lebrun and L. Macaire. *Aspect inspection of marble tiles by colour line-scan camera*. Proc. 5th International Conference on Quality Control by Artificial Vision, Le Creusot, France, 2001.
- [14] A. Fernández, E. González and X. Leiceaga. *Caracterización de la apariencia visual de superficies pulidas de granito "Rosa Porriño" mediante procesamiento digital de imagen*. Proc XIV Congreso Internacional de Ingeniería Gráfica, pp. 330-335, Santander, Spain, 2002.
- [15] S. Kukkonen, H. Kälviäinen and J. Parkkinen. *Visual Inspection by Spectral Features in Ceramics Industry*. Proc. of EUROPTO Conference on Polarisation and Colour Techniques in Industrial Inspection, SPIE Vol. 3826, pp. 64-75, Munich, Germany, 1999.

- [16] S. Kukkonen, H. Kälviäinen and J. Parkkinen. *Color Features for Quality Control in Ceramic Tile Industry*. Optical Engineering, 40(2), pp. 170-177, 2001.
- [17] Fernando López, José-Miguel Valiente and Floreal Acebrón. *Sistema automático de inspección óptica de defectos de integridad en piezas cerámicas; Estudio comparativo de métodos de reconocimiento de formas*. VII Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA'97), pp. 509-518, Málaga , 1997.
- [18] José-Miguel Valiente, Floreal Acebrón and Fernando López. *An automatic visual inspection system for ceramic tile manufacturing defects*. Proceedings of the IASTED International Conference on Signal Processing and Communications (SPC'98), pp. 257-260, Canarias (Spain), 1998.
- [19] José-Miguel Valiente, Floreal Acebrón and Fernando López. *A ceramic tile inspection system for detecting corner defects*. Proceedings of the IX Spanish Symposium on Pattern Recognition and Image Analysis (SNRFA'2001), pp. 213-218, Castellón (Spain), 2001.
- [20] Fernando López, Floreal Acebrón, José-Miguel Valiente and Enrique Pérez. *A study of registration methods for ceramic tile inspection purposes*. Proceedings of the IX Spanish Symposium on Pattern Recognition and Image Analysis (SNRFA'2001), Vol. I, pp. 145-150, Castellón (Spain), 2001.
- [21] José-Miguel Valiente, Fernando López, Floreal Acebrón and Enrique Pérez. *An image registration method for ceramic tile inspection purposes*. V International Conference on Quality Control by Artificial Vision (QCAV2001), pp. 498-503, Le Creusot (France), 2001.
- [22] Floreal Acebrón, Fernando López, José-Miguel Valiente and José-Ramón Navarro. *Surface defect detection on fixed ceramic tiles*. Proceedings of the Second IASTED International Conference on Visualization, Imaging, and Image Processing (II VIIP), pp. 35-40, Málaga (Spain), 2002.
- [23] José-Miguel Valiente, Floreal Acebrón, Fernando López and José-Ramón Navarro. *Sistema de inspección visual automática de azulejos de patrón fijo*. XIV

- Congreso Internacional de Ingeniería gráfica (INGEGRAF02), pp. 309-318, Santander (Spain), 2002.
- [24] Fernando López, José-Miguel Valiente, José-Ramón Navarro and Floreal Acebrón, *Inspección automática de tonalidad en procesos industriales: aplicación a la clasificación de tonos en azulejos*, XIV Congreso Internacional de Ingeniería gráfica (INGEGRAF02), pp. 288-294, Santander (Spain), 2002.
- [25] Fernando López, José-Miguel Valiente, Ramón Baldrich and María Vanrell. *Fast surface grading using color statistics in the CIE Lab space*. 2nd Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA2005). Lecture Notes in Computer Science, Vol. 3523, pp. 666-673, 2005.
- [26] Fernando López, José-Miguel Valiente and Ángel Rodas. *On the search for an accurate approach to surface grading*. Primer Congreso Español de Informática (CEDI2005) - I Workshop sobre Reconocimiento de Formas y Análisis de Imágenes (AERFAI), pp. 183-190, Granada (Spain), 2005.
- [27] Fernando López, José-Miguel Valiente and Gabriela Andreu. *A study of real-time compliance and parallelization for the purpose of surface grading*. Primer Congreso Español de Informática (CEDI2005) - XVI Jornadas de Paralelismo (JP'2005), pp. 83-90, Granada (Spain), 2005.
- [28] Fernando López, José-Miguel Valiente and José-Manuel Prats. *Surface grading using soft colour-texture descriptors*. X Iberoamerican Congress on Pattern Recognition (CIARP2005), Lecture Notes in Computer Science, Vol. 3773, pp. 13-23, 2005.
- [29] R.C. Gonzalez and P. Wintz. *Digital image processing*. Addison-Wesley, 2nd Edition, 1987.
- [30] P.A. Devijver and J. Kittler. *Pattern recognition: a statistical approach*. Prentice-Hall, London, 1982.
- [31] D.C. Montgomery. *Design and analysis of experiments*. John Wiley and Sons, 4th Edition, New York, 1997.

- [32] R. Christensen. *Log-linear models and logistic regression*. Springer-Verlag, 2nd Edition, New York, 1997.
- [33] G. Wyszecki and W.S. Stiles. *Color science: concepts and methods, quantitative data and formulae*. Wiley, New York, 1982.
- [34] C.A. Poynton. *A technical introduction to digital video*. Prentice Hall, 1996.
- [35] S.J. Sangwine and R.E.N. Horne. *The colour image processing handbook*. Chapman & Hall, 1998.
- [36] X. Zhang and B.A. Wandell. *A spatial extension of CIE Lab for digital color image reproduction*. In Society for Information Display Symposium, San Diego, 1996.
- [37] K.N. Plataniotis and A.N. Venetsanopoulos. *Color image processing and applications*. Springer, 2000.
- [38] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis and Machine Vision*. Chapman & Hall, 1993.
- [39] R.M. Haralick. *Statistical and structural approaches to texture*. Proceedings of 4th International Joint Conference on Pattern Recognition, pp. 45-60, 1979.
- [40] S.Y. Lu and K.S. Fu. *A syntactic approach to texture analysis*. Computer Graphics and Image Processing, 7:303-330. 1978.
- [41] L. Carlucci. *A formal system for texture languages*. Pattern Recognition, 4:53-72, 1972.
- [42] R.M. Haralick, K. Shanmugam and I. Dinstein. *Texture features for image classification*. IEEE Trans. Systems Machines and Cybernetics, 6(3):609-622, 1973.
- [43] R. Connors and C. Harlow. *A theoretical comparison of texture algorithms*. IEEE Trans. Pattern Analysis Mach. Intell., 3:25-39, 1980.
- [44] M. Unser. *Sum and difference histograms for texture classification*. IEEE Trans. Pattern Analysis Mach. Intell., 8(1):118-125, 1986.

- [45] C.M. Wu and Y.C. Chen. *Statistical feature matrix for texture analysis*. CVGIP Graphical Models and Image Processing, 54:407-419, 1992.
- [46] M. Hassner and J. Sklansky. *The use of markov random fields as models of texture*. Computer Graphics and Image Processing, 12:357-370, 1980.
- [47] H. Derin and W. Cole. *Segmentation of textured images using gibbs random fields*. Computer Vision Graphics and Image Processing, 35:72-98, 1986.
- [48] H. Derin and H. Elliot. *Modelling and segmentation of noisy and textured images using gibbs random fields*. IEEE Trans. Pattern Analysis Mach. Intell., 9(1):39-55, 1987.
- [49] L. Wang and J. Liu. *Texture segmentation using MRMRMF modelling*. Pattern Recognition Letters, 21(2000):189-200, 2000.
- [50] C.S. Won and H. Derin. *Unsupervised segmentation of noisy and textured images using markov random fields*. CVGIP: Graphical Models and Image Processing, 54:308-328, 1992.
- [51] J. Serra and G. Verchery. *Mathematical morphology applied to fibre composite materials*. Film Sci. Tech., 6:141-158, 1973.
- [52] Y. Chen and E.R. Dougherty. *Grey-scale morphological granulometric texture classification*. Optical Engineering, 33(8):2713-2722, 1994.
- [53] B.B. Mandelbrot. *The fractal geometry of nature*. Freeman, New York, 1983.
- [54] A. Pentland. *Fractal-based descriptions of natural scenes*. IEEE Trans. Pattern Analysis Mach. Intell., 6:661-674, 1984.
- [55] C.H. Chen. *A study of texture classification using spectral features*. In 6th ICPR, pp. 1074-1077, 1982.
- [56] S.S. Liu and M.E. Jerningan. *Texture analysis and discrimination in additive noise*. Computer Vision, Graphics and Image Processing, 49:52-67, 1990.

- [57] I. Daubechies. *Orthonormal bases of compactly supported wavelets*. Communications on pure and applied mathematics, 41:909-996, 1988.
- [58] S.G. Mallat. *A theory for multiresolution signal decomposition: the wavelet representation*. IEEE Trans. Pattern Analysis Mach. Intell., 11(7):674-693, 1989.
- [59] A. Kundu and J.L. Chen. *Texture classification using qmf bank-based subband decomposition*. CVGIP: Graphical Models and Image Processing, 54(5):369-384, 1992.
- [60] T. Chang and C.C.J. Kuo. *Texture analysis and classification with tree-structured wavelet transform*. IEEE Trans. Image Processing, 2(4):429-441, 1993.
- [61] A. Laine and J. Fan. *Texture classification by wavelet packet signatures*. IEEE Trans. Pattern Analysis Mach. Intell., 15(11):1186-1190, 1993.
- [62] M. Unser. *Texture classification and segmentation using wavelets frames*. IEEE Trans. Image Processing, 4(11):1549-1560, 1995.
- [63] A. Bovik, M. Clark and W. Geisler. *Multichannel texture analysis using localized spatial filters*. IEEE Trans. Pattern Analysis Mach. Intell., 12(1):55-73, 1990.
- [64] D. Dunn, W.E. Higgins and J. Wakeley. *Texture segmentation using 2-d gabor elementary funtions*, IEEE Trans. Pattern Analysis Mach. Intell., 16(2):130-149, 1994.
- [65] M. Unser, *Local linear transforms for texture measurements*, Signal Processing, 11:61-79, 1986.
- [66] G.F. McLean. *Vector quantization for texture classification*. IEEE Trans. Sys., Man., Cyben., 23(3):637-649, 1993.
- [67] C. Carson, S. Belongie, H. Greenspan and J. Malik. *Region based image querying*, In CPR Workshop on Content-Based access of Image and Video Libraries, 1997.

- [68] S.C. Tan and J. Kittler. *Colour texture classification using features from colour histogram*. In Proceedings of the 8th Scandinavian Conference on Image Processing, 1993.
- [69] D.K. Panjwani and G. Healey. *Markov random fields for unsupervised segmentation of textured color images*. IEEE Trans. Pattern Analysis Mach., Intell., 17(10):939-954, 1995.
- [70] G. Healey and L. Wang. *Illumination-invariant recognition of texture in color images*. Journal of the Optical Society of America, 12(9):1877-1883, 1995.
- [71] A. Gagalowicz, S. De Ma and C. Tournier-Lasserre. *Efficient models for color textures*. In 8th ICPR, pp. 412-414, 1986.
- [72] T. Caelli and D. Reye. *On the classification of image regions by colour, texture and shape*. Pattern Recognition, 26(4):461-470, 1993.
- [73] G.D. Finlayson, S.S. Chatterjee and B.V. Funt. *Color angular indexing*, In European Conference on Computer Vision. pp. 16-27, 1996.
- [74] J. Weickert. *Coherence-enhancing diffusion of colour images*. Image and Vision Computing, 17:201-212, 1999.
- [75] M. Mirmehdi and M. Petrou. *Segmentation of color textures*. IEEE Trans. on Pattern Analysis Mach. Intell., 22(2):142-159, 2000.
- [76] M. Petrou, M. Mirmehdi and M. Coors. *Perceptual smoothing and segmentation of colour textures*. In 5th European Conference on Computer Vision, pp. 623-639, 1998.
- [77] K.H. Baeuml and B. Wandell. *Color appearance of mixture gratings*. Vision Research, 36(18):2849-2864, 1996.
- [78] A. Poirson and B. Wandell. *The appearance of color patterns: Pattern-color separability*. Journal of the Optical Society of America, 10(12):2458-2470, 1993.
- [79] B.A. Wandell. *Color appearance: the effects of illumination and spatial pattern*, Technical report, Psychology Department, Standford University, 1995.

- [80] X. Zhang, J.E. Farrell and B.A. Wandell. *Applications of a spatial extension to CIELab*. In SPIE, 1997.
- [81] R. Baldrich, R. Toledo, E. Valveny and M. Vanrell. *Perceptual Colour Image Segmentation*. Proceedings of the Second IASTED International Conference Visualization, Imaging and Image Processing VIIP 2002: 145-150, 2002.
- [82] R. Baldrich, M. Vanrell, M. Benavente and A. Salvatella. *Color Enhancement based on perceptual sharpening*. Proceedings of the IEEE International Conference on Image Processing, 2003.
- [83] C. Boukouvalas, J. Kittler, R. Marik and M. Petrou. *Automatic color grading of ceramic tiles using machine vision*. IEEE Transactions on Industrial Electronics, 44(1), pp. 132-135, 1997.
- [84] C. Boukouvalas and M. Petrou. *Perceptual correction for colour grading using sensor transformations and metameric data*. Machine Vision and Applications, 11(2), pp. 96-104, 1998.
- [85] C. Boukouvalas and M. Petrou. *Perceptual correction for colour grading of random textures*. Machine Vision and Applications, 12(3), pp. 129-136, 2000.
- [86] K.Y. Song, M. Petrou and J. Kittler. *Texture crack detection*. Machine Vision and Applications, 8:63-76, 1995.
- [87] K.Y. Song, J. Kittler and M. Petrou. *Defect detection in random colour textures*. Image and Vision Computing, 14:617-625, 1996.
- [88] C. Boukouvalas, J. Kittler, R. Marik, M. Mirmehdi and M. Petrou. *Ceramic tile inspection for colour and structural defects*. Proceedings of AMPT95, pp. 390-399, 1995.
- [89] C. Boukouvalas, J. Kittler, R. Marik and M. Petrou. *Automatic grading of textured ceramic tiles* Proceedings of the SPIE. 2423:248-256, 1995.
- [90] M. Petrou and C. Boukouvalas. *Colour quantification for industrial inspection*. Proceedings IWISPO'96. Third International Workshop on Image and Signal

- Processing on the Theme of Advances in Computational Intelligence. Elsevier, Amsterdam, Netherlands, pp. 611-614. 1996.
- [91] C. Boukouvalas, F. De Natale, G. De Toni, J. Kittler, R. Marik, M. Mirmehdi, M. Petrou, P. Le Roy, R. Salgari and G. Vernazza. *An Integrated System for Quality Inspection of Tiles*. International Conference on Quality Control by Artificial Vision, QCAV'97, pp.49-54, 1997.
- [92] C. Boukouvalas, F. De Natale, G. De Toni, J. Kittler, R. Marik, M. Mirmehdi, M. Petrou, P. Le Roy, R. Salgari and G. Vernazza. *Automatic system for surface inspection and sorting of tiles*. Journal of materials processing technology, Elsevier, 82:179-188, 1998.
- [93] M. Petrou, M. Mirmehdi and M. Coors. *Perceptual smoothing and segmentation of colour textures*. Lecture Notes in Computer Science, Springer Verlag, 1406:623-639, 1998.
- [94] C.E. Costa and M.Petrou. *Automatic registration of ceramic tiles for the purpose of fault detection*. Machine Vision and Applications, 11:225:230, 2000.
- [95] N. Wirth. *Algorithms + Data Structures = Programs*. Prentice-Hall, 1976.
- [96] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery. *Numerical recipes in C*. Cambridge University Press, 2nd Edition, 1992.
- [97] A. Rosenfeld and A.C. Kak. *Digital picture processing*. Vol. 1, 2nd Edition, Academic Press, New York.
- [98] K.H. Bäumel and B.A. Wandell. *Color appearance of mixture gratings*. Vision Research, Vol. 36, 18:2849-2864, 1996.
- [99] F. Lumbreras, J. Serrat, R. Baldrich, M. Vanrell, and J.J. Villanueva. *Color Texture Recognition Through Multiresolution Features*. Proceedings of the International Conference on Quality Control by Artificial Vision, Vol. 1, pp. 114-121, France, 2001.

- [100] R. Baldrich and M. Vanrel. *Ceramic tile classification based on colour similarity measurements*. In Workshop on European Scientific and Industrial Collaboration, pp. 307-311, Girona (Spain), 1998.
- [101] M.J. Swain and D.H. Ballard. *Color indexing*. International Journal on Computer Vision, 7(1):11-32, 1991.
- [102] Y.I. Ohta, T. Kanade and T. Sakai. *Color information for region segmentation*. Computer Graphics and Image Processing, 13(3):222-241, 1980.
- [103] J.L. Starck and A. Bijaoui. *Filtering and deconvolution by the wavelet transform*. Signal Processing, 35, pp.195-211., 1994.
- [104] M.V. Wickerhauser. *Adapted wavelet analysis from theory to software*. A.K.Peters, 1994.
- [105] O. Silven, M. Niskanen and H. Kauppinen. *Wood inspection with non-supervised clustering*. Machine Vision and Applications, 13(5-6), pp. 275-285, 2003.
- [106] M. Niskanen, O. Silvén and H. Kauppinen. *Color and texture based wood inspection with non-supervised clustering*. Proc. 12th Scandinavian Conference on Image Analysis, pp. 336-342, Bergen, Norway, 2001.
- [107] M. Niskanen, H. Kauppinen and O. Silvén. *Real-time aspects of SOM-based visual surface inspection*. Proc. SPIE Vol. 4664 Machine Vision Applications in Industrial Inspection X, pp. 123-134, San Jose, 2002.
- [108] T. Ojala, M. Pietikäinen and D. Harwood. *A comparative study of texture measures with classification based on feature distributions*. Pattern Recognition, 29, pp. 51-59, 1996.
- [109] T. Ojala and M. Pietikäinen. *Unsupervised texture segmentation using feature distributions*. Pattern Recognition, 32(1999):477-486, 1999.
- [110] T. Kohonen. *Self-Organization and associative memory*. Springer-Verlag, Berlin, Germany, 1988.

- [111] H. Kauppinen, O. Silven and T. Piirainen. *Self-Organizing map based user interface for visual surface inspection*. Proc. EUROPTO Conf. on Polarization and Color Techniques in Industrial Inspection, SPIE Vol. 3826, pp.270-280, Munich, Germany, 1999.
- [112] J.N. Kapur, P.K. Sahoo, A.K.C. Wong and Y.C. Chen. *A new method for grey-level picture thresholding using the entropy of the histogram*. Computer Vision Graphics and Image Processing, 29, pp.273-285, 1985.
- [113] S. Di Zenzo. *A note on the gradient of a multi-image*. Computer Vision Graphics and Image Processing, 33, pp.116-125, 1986.
- [114] E. Herrala, T. Hyvärinen, O. Voutilainen and J. Lammasniemi. *An optoelectronic sensor system for industrial multipoint and imaging spectrometry*. Sensors and Actuators, A 61, pp.335-338, 1997.
- [115] P. Mouroulis. *Visual Instrumentation: Optical Design & Engineering Principles*. McGraw-Hill, 1999.
- [116] M.L. Mahy and E.A. Oosterlink. *Evaluation of uniform color spaces developed after the adoption of CIELAB and CIELUV*. Color Res. and Appli, 19(2):105-121, 1994.
- [117] G. Sharma and H.J. Trussell. *Digital color imaging*. IEEE Transactions on Image Processing, 6(7):901-932, 1997.
- [118] CIE technical report. *Methods to derive colour differences for images*. draft version 0.4, 2004.
- [119] G.E.P. Box, W.G. Hunter and J.S. Hunter. *Statistics for experimenters: An introduction to design, data analysis, and model building* John Wiley & Sons, 1978.
- [120] J. Arlandis, J.C. Pérez-Cortes, and J. Cano. *Rejection strategies and confidence measures for a k-NN classifier in an OCR task*. 16th. International Conference on Pattern Recognition ICPR-2002, Québec, Canada, 2002.

- [121] B. Efron. *Estimating the error rate of a prediction rule: Improvements on cross-validation*. Journal of the American Statistical Association, 78:316-331, 1983.
- [122] R.O. Duda and P.E. Hart. *Pattern classification and scene analysis*. John Wiley and Sons, New York, 1973.
- [123] R. Kohavi. *A study of cross-validation and bootstrap for accuracy estimation and model selection*. In: IJCAI, pp.1137-1145, 1995.
- [124] K. Fukunaga. *Statistical Pattern Recognition*. Academic Press, 2nd Edition, 1990.
- [125] C. Goutte. *Note on free lunches and cross-validation*. Neural Computation, 9:1211-1215, 1997.
- [126] D.E. Culler and J.P. Singh. *Parallel computer architecture: A hardware/software approach*. Morgan Kaufmann, 1999.
- [127] A. Geist et al. *Parallel Virtual Machine: A user's guide and tutorial for networked parallel computing*. MIT Press, 1994.
- [128] <http://www.mpi-forum.org/docs/mpi-11-html/mpi-report.html>. *MPI: A message passing interface standard*. University of Tennessee. 1995.
- [129] A. Geist, J.A. Kohl and P.M. Papadopoulos. *PVM and MPI : A Comparison of Features*. Calculateurs Paralleles, Vol. 8, No. 2, 1996.
- [130] X. Xie, M. Mirmehdi and B. Thomas. *Inspecting colour tonality on textured surfaces*. Proceedings of the 1st International Conference on Image Analysis and Recognition, pp. 810–817. Springer LNCS 3212, 2004.
- [131] D.M. Tsai. *A fast thresholding selection procedure for multimodal and unimodal histograms*. Pattern Recognition Letters. Vol. 16, No. 6, pp. 653-665, 1995.
- [132] International Telecommunication Union (<http://www.itu.int/ITU-R/index.html>).

- [133] Y. Ohno. *CIE Fundamentals for color measurements*. International Conference on Digital Printing Technologies, Vancouver, Canada, 2000.