Dublin City University
School of Electronic Engineering

Thesis submitted for the Degree of
Doctor of Philosophy

Unsupervised Segmentation of Natural Images Based on the Adaptive Integration of Colour-Texture Descriptors

by

Dana E. Ilea

danailea@eeng.dcu.ie

Supervisor: Prof. Paul F. Whelan

September 2008

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _____ (Candidate)

ID No.: 55139931

Date: September 19th 2008

Acknowledgements

I would like to thank Professor Paul F. Whelan for his guidance and support during the progression of my research project that has commenced in October 2005. Special thanks go to Dr. Ovidiu Ghita for his constructive comments concerning the theoretical issues associated with the development of this work. I would also like to thank all the members of the Vision Systems Group for their positive feedback in relation to my project. I am most grateful to my parents for their permanent encouragement and tremendous support.

I would also wish to express my gratitude to the School of Electronic Engineering and Science Foundation Ireland (SFI) for supporting financially this research project.

Table of Contents

List of Figures

List of Tables

Publications Resulting from this Research

D. E. Ilea and P. F. Whelan, "CTex - An Adaptive Unsupervised Segmentation Algorithm Based on Colour-Texture Coherence", *IEEE Transactions on Image Processing*, vol. 17, no. 10, pp. 1926-1939, 2008.

D. E. Ilea and P. F. Whelan, "Colour Image Segmentation Using a Spatial K-Means Clustering Algorithm", *Proceedings of the Irish Machine Vision and Image Processing Conference (IMVIP 2006)*, pp. 146-153, Dublin City University, Ireland, 30 August – 1 September, 2006.

D. E. Ilea and P. F. Whelan, "Colour Image Segmentation Using a Self-initialising EM Algorithm", *Proceedings of the International Conference on Visualization, Imaging and Image Processing (VIIP 2006)*, Palma de Mallorca, Spain, 28-30 August 2006.

D. E. Ilea and P. F. Whelan, "Automatic Segmentation of Skin Cancer Images Using Adaptive Colour Clustering", *Proceedings of the China-Ireland International Conference on Information and Communications Technologies (CIICT 2006)*, pp. 348-351, Hangzhou, China, 18 -19 October, 2006.

D. E. Ilea and P. F. Whelan, "Adaptive Pre-Filtering Techniques for Colour Image Analysis", *Proceedings of the International Machine Vision and Image Processing Conference (IMVIP 2007)*, pp. 150-157, National University of Ireland, Maynooth, *IEEE* Computer Society Press, 5 -7 September 2007.

D. E. Ilea, O. Ghita and P. F. Whelan, "Evaluation of Local Orientation for Texture Classification", *Proceedings of the 3rd International Conference on Computer Vision Theory and Applications (VISAPP 2008)*, pp. 357-364, Funchal, Madeira – Portugal, 22 - 25 January, 2008.

D. E. Ilea, P. F. Whelan and O. Ghita, "Performance Characterization of Clustering Algorithms for Colour Image Segmentation", *Proceedings of the International Conference on Optimization of Electrical and Electronic Equipments (OPTIM 2006)*, May 18-19, 2006, Brasov, Romania.

O. Ghita, P. F. Whelan and D. E. Ilea, "Multi-resolution Texture Classification Based on Local Image Orientation", *Proceedings of the International Conference on Image Analysis and Recognition (ICIAR 2008)*, pp. 688-696, Póvoa de Varzim, Portugal, 25-27 July, 2008.

M. Lynch, D. Ilea, K. Robinson, O. Ghita and P. F. Whelan, "Automatic Seed Initialization for the Expectation-maximization Algorithm and its Application in 3D Medical Imaging", *Journal of Medical Engineering and Technology*, vol. 31, no. 5, pp. 332 – 340, September/October 2007.

Title: Unsupervised Segmentation of Natural Images Based on the Adaptive Integration of Colour-Texture Descriptors

Author: Dana E. Ilea

Abstract

This thesis presents the development of a theoretical framework capable of encompassing the colour and texture information in a robust image descriptor that can be applied to the identification of coherent regions in complex natural images. In the suggested approach, the colour and texture features are extracted explicitly on two independent channels and the main emphasis of this work was placed on their adaptive inclusion in the process of data partition. The proposed segmentation framework consists of several computational tasks including adaptive filtering, colour segmentation, texture extraction and the adaptive integration of the colour and texture features in the segmentation process in an unsupervised manner. In this regard, an important contribution of this work is represented by the development of a multi-space colour segmentation scheme where the key element was the inclusion of the Self Organising Map network that was applied to compute the optimal parameters required by data clustering algorithms. The second component of the proposed segmentation framework deals with the extraction of texture features and in this study the performance of several texture descriptors when applied to the segmentation of synthetic and natural images was analysed. To this end, several texture descriptors are evaluated including multi-channel texture decomposition filtering based on Gabor and isotropic filter banks, and multi-resolution approaches that analyse the texture at micro-level. The most important contribution of this work resides in the adaptive inclusion of the colour and texture features in a compound mathematical descriptor with the aim of identifying the homogenous regions in natural images. This colour-texture integration is performed by a novel clustering algorithm that is able to enforce the spatial continuity during the data assignment process. To demonstrate the efficiency of the proposed colour-texture segmentation scheme, a comprehensive quantitative and qualitative performance evaluation has been carried out on natural image databases. The experimental results indicate that the proposed framework is accurate in capturing the colour and texture characteristics even when applied to the segmentation of complex natural images.

Chapter 1

Introduction

The objective of this chapter is to introduce the motivation for the investigation of an adaptive composite image descriptor that can be used in the development of robust colour-texture segmentation algorithms. The major objectives of this thesis will be discussed along with the major contributions that emerged from this study. Finally, an outline of this thesis will be presented.

1.1 Problems and Motivation

Image segmentation represents one of the most important areas of research in the fields of image processing and computer vision. This is motivated by the fact that image segmentation is one task that is commonly employed in the development of high-level image analysis tasks such as scene understanding, object recognition and image retrieval and its accuracy has a decisive impact on the overall performance of the developed vision systems. Over the past decades, the field of image segmentation has developed extensively and due to improvements and advances in the computer and photographic technologies, the research world has been given the possibility and the means to address real life problems including the segmentation of biomedical, natural, industrial and aerial images.

The aim of the segmentation process is to divide an input image into several disjoint coherent regions that are strongly related to the imaged objects. Over

several decades an abundance of image segmentation algorithms have been proposed in the literature, where the segmentation task has been formulated in terms of the accurate extraction and modelling of two fundamental attributes of digital images, namely colour and texture. Although image segmentation is one of the fundamental areas of research, still represents a challenging topic for the computer vision community due to the difficulty in extracting precise colour and texture models that can locally adapt to variations in the image content. In particular, the segmentation of natural images proved to be a very difficult task, since these images exhibit significant inhomogeneities in colour and texture as they are often characterised by a high degree of complexity, randomness and irregularity. Due to these problems, most of the developed algorithms have been developed in conjunction with well-defined applications and in general they require a significant amount of user interaction.

Colour and texture play an important role in human perception as these attributes are employed in discerning between real-world objects. Most natural images exhibit both colour and texture information and the computer vision researchers developed a large spectrum of mathematical models whose aims were to sample the local and global properties of these fundamental image descriptors. Although there is no definition of texture that is widely accepted, this image descriptor can be regarded as a function of spatial variation of the pixel intensities in the image. Thus, texture can be globally defined using intuitive terms such as coarse, fine, smooth, granulated, rippled, etc. and it is assumed that similar patterns describing texture units are placed in the image in a periodic manner. The area of texture analysis is vast and many approaches have been developed to sample the texture properties at micro and macro level, but in spite of the enormous research effort, texture analysis is still an open issue. This is motivated by the fact that textures that are present in natural images are not uniform and they are scale and rotation dependent. With the proliferation of modern digital cameras, colour information started to be actively investigated by the vision researchers. This is motivated by the fact that the use of colour and texture information has strong links with human perception and practice indicated that texture alone image descriptors are not

sufficient to robustly characterise the image information. Thus, in recent years colour-texture analysis has received a significant interest from the vision community, and as a result, a large number of approaches has been developed with the aim of obtaining robust image segmentation.

There is no doubt that the inclusion of texture and colour information in the segmentation process leads to improved performance when compared with the performance of the algorithms that evaluate either the colour or the texture image information. But in spite of this obvious observation that the use of colour and texture collectively is a key issue in achieving robust segmentation, there is still a large degree of confusion in the computer vision community in regard to the optimal approach of combining these image attributes in a joint coherent image descriptor. Thus, the extraction and the optimal inclusion of colour and texture features in the segmentation process represent the major topics that will be addressed in this investigation.

1.2 Objectives of this Research

The central aim of this research work is the development of a theoretical framework that is able to include the colour and texture information in an adaptive fashion in the process of image segmentation. As indicated earlier, colour-texture analysis is an active research topic and many approaches have been proposed to address the problem of robust image segmentation. It is useful to note, that the vast majority of proposed approaches attempted to analyse the colour and texture information in a simplistic manner and as a result the developed algorithms proved to be limited to certain application areas and moreover their performances are highly dependent on the optimal selection of various parameters (a detailed literature survey that describes the most relevant texture-alone, colour-alone and colour-texture segmentation approaches is provided in Chapter 2).

Segmentation of natural images is in particular difficult since both colour and texture attributes are not uniformly distributed within image areas defined by similar objects and often the strength of texture and colour can vary

considerable from image to image. In addition to this, complications added by the uneven illumination, image noise, perspective and scale distortions make the process of identifying the homogenous regions in the image extremely difficult. All these challenges form the set of macro objectives that will be addressed in this thesis and they can be summarised as follows:

- To develop a generic theoretical framework that allows the integration of various colour and texture features without any level of supervision.
- To investigate the development of novel colour extraction techniques that are able to adapt to the image content.
- To develop a statistical framework that enforces spatial constraints during the data partitioning process.
- To develop pre-processing adaptive filtering algorithms that are able to compensate for image noise and uneven illumination.
- To conduct a large number of experiments to produce quantitative and qualitative results when the algorithm is applied to standard databases of natural images. Also the proposed algorithm will be benchmarked against state of the art colour-texture segmentation schemes.

These macro objectives generated the major and minor contributions of this research work. One of the major theoretical contributions of this work is located in the development of an unsupervised colour segmentation algorithm that involves a statistical analysis of the input image in a multi-space colour representation. The main novelty of this work resides in the inclusion of the Self-Organising Maps (SOMs) to determine the dominant colours and the number of clusters in the image.

Another contribution of this thesis resides in the development of a new texture analysis method that is based on the evaluation of the image orientation at micro and macro level. This new texture descriptor is evaluated in detail and its performance is compared against that offered by standard statistical and signal processing texture analysis techniques.

The most important contribution of this work resides in the adaptive inclusion of the colour and texture features in a compound mathematical

descriptor with the aim of identifying the homogenous regions in natural images. The proposed colour-texture integration is performed by a novel clustering algorithm that is able to enforce the spatial continuity during the data assignment process.

Minor theoretical contributions can be found in the area of adaptive filtering and implementation of statistical measures that are applied to quantify the similarity between the segmentation result and multiple ground truth data.

1.3 Overview of the Proposed Colour-Texture Segmentation Framework

The main computational components of the proposed segmentation framework (referred in this thesis to as CTex) are depicted in Figure 1.1. The colour–texture segmentation framework investigated in this study analyses the colour and texture information on separate channels. In this regard, colour segmentation is the first major component of the proposed framework and involves the statistical analysis of data using multi-space colour representations. The first step of the colour segmentation involves filtering the input data using a new gradient-boosted forward and backward (GB-FAB) anisotropic diffusion algorithm that is applied to eliminate the influence of the image noise and improve the local colour coherence. The selection of the number of clusters and the initial cluster centres is one of the most difficult problems that have to be addressed in the implementation of statistical data partitioning schemes. To address this problem, the first stream of the colour segmentation algorithm extracts the dominant colours and the optimal number of clusters from the first colour representation of the image using an unsupervised procedure based on a Self Organising Map (SOM) network. The second stream of the proposed colour segmentation scheme analyses the image in a complementary colour representation where the number of clusters calculated from the first colour representation performs the synchronisation between the two computational streams of the algorithm. In the final stage of the colour segmentation process, the clustered results obtained for each colour

representation form the input for a multi-space clustering that outputs the final colour segmented image.

The second major component of the proposed CTex framework involves the extraction of the texture features from the original image. In this research several statistical and signal processing texture extraction techniques have been evaluated with the aim of identifying the texture extraction scheme that returns optimal results when applied to the segmentation of natural images.

The resulting colour segmented image and texture data are the inputs of a novel Adaptive Spatial K-Means (ASKM) framework that outputs the colour-texture segmented result.

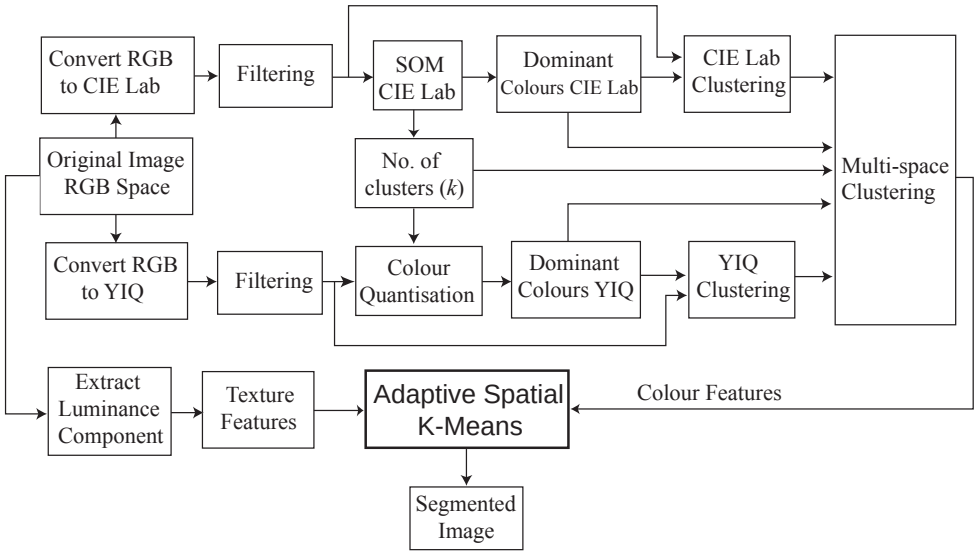


Figure 1.1 Outline of the proposed CTex colour-texture image segmentation framework.

1.4 Thesis Organisation

The thesis is organised as follows.

Chapter 2 reviews the existing colour, texture and colour-texture segmentation techniques and the main conclusions resulting from the literature survey are discussed.

Chapter 3 presents the theoretical issues related to the development of adaptive pre-filtering techniques for colour image analysis. In this chapter several non-linear adaptive smoothing techniques are investigated where the main emphasis is placed on the development of a new strategy that improves the feature preservation of the anisotropic diffusion algorithms in image areas defined by low colour contrast. A large number of experiments were conducted to evaluate the performance of the filtering algorithms evaluated in this study.

Chapter 4 describes the proposed colour features extraction algorithm. In this chapter each computational component of the developed algorithm is discussed in detail with the main focus being on the description of the novelty aspects such as the inclusion of a SOM network that is applied to estimate the parameters required by data partitioning algorithms in an adaptive manner from input data. To justify the effectiveness of the proposed techniques, a comprehensive quantitative and qualitative evaluation was conducted.

Chapter 5 presents the texture extraction methods evaluated in this study and introduces a new texture analysis technique based on the evaluation of the image orientation. The performance of the proposed texture extraction technique is evaluated against that offered by standard texture analysis techniques.

Chapter 6 highlights the mathematical framework behind the integration of the two fundamental image attributes, colour and texture. In this chapter several issues related to the optimal inclusion of the colour and texture features into a spatially coherent clustering strategy are discussed in detail. Chapter 6 continues with the presentation of a comprehensive quantitative and qualitative evaluation of the colour-texture extraction algorithms discussed in this thesis. The experiments were conducted on databases containing natural images and the results indicate that the proposed technique is accurate in capturing the

colour and texture characteristics present in complex natural images. The experimental data also show that the proposed colour-texture segmentation technique outperforms one standard state of art colour-texture segmentation algorithm.

Chapter 7 presents a summary of this research work and outlines the contributions resulting from this investigation. This chapter also provides a discussion on possible extensions that can be made to the proposed segmentation framework and several future directions of research are advanced.

Appendix A defines the Probabilistic Rand Index that is used for the quantification of the segmentation results. In Appendix B, experimental results using additional similarity metrics are presented and discussed.

Chapter 2

Literature Survey

The aim of this chapter is to present the background of this research work. The main trends in the individual areas of texture and colour image analysis will be presented, followed by a detailed discussion on the existing colour-texture segmentation approaches with the main focus being on the optimal inclusion of the texture and colour features in the segmentation process.

2.1 Texture Analysis for Image Segmentation

Texture-based image segmentation represents a major field of research in the area of computer vision that has been intensively investigated for more than three decades. This has been motivated by the fact that the robust identification of texture primitives in digital images plays a key role for many applications domains such as image segmentation, image retrieval, remote sensing and analysis of biomedical data. Taken into consideration the large spectrum of applications based on texture analysis, an impressive number of approaches has been published in the computer vision literature. As indicated in several reviews on texture-based segmentation [1, 2], the proposed techniques can be classified into four major categories: statistical, model-based, signal processing and structural.

Statistical methods are based on the evaluation of the spatial distributions and relationships between the pixel intensities in the image. Relevant statistical texture analysis techniques include the autocorrelation function [3], texture

energy features [4], grey-level co-occurrence matrices [3] and Local Binary Patterns [5]. The texture analysis techniques based on autocorrelation function (ACF) evaluates the rate of decay of this function with the increase of the spatial shift with respect to the origin of the input signal. The rate of decay of the ACF samples the amount of regularity and the fineness/coarseness character of the textures present in the image. In this regard, if the analysed texture shows a high degree of complexity (coarseness) the ACF values will decrease slowly with distance. Conversely, for fine textures the drop in the ACF values will be more pronounced with distance [1]. This technique is attractive for its low computational cost, but performs modestly when applied to texture segmentation tasks. This is motivated by the fact that the ACF values capture only the global texture characteristics and as a result many textures may be described by the same ACF pattern.

Texture energy features involve the convolution of the input image with small kernels such as Level, Edge, Spot and Ripple, where the features are computed by summing the absolute values of the filtering results for each pixel in the image. Although the energy features offer a simple approach to texture analysis, there is no information available in regard to the relative position of the pixels to each other in the image and these features have shown only limited discriminative power when applied to the analysis of non-synthetic textures. The co-occurrence matrices tend to solve this issue as they analyse textures based on the joint probability distribution of pairs of pixels. These texture descriptors have been first studied by Haralick [3] and are based on the evaluation of the spatial relationships between pixels in the textural pattern. In the experiments conducted by Weszka [6] it has been shown that the co-occurrence matrices outperform autocorrelation features when they were evaluated for texture classification tasks. This conclusion is motivated by the fact that co-occurrence matrices encompass the local relationships between the pixels that form the texture unit as opposed to the autocorrelation function that evaluates only the global distribution of the pixel values in the texture unit. Although the texture analysis based on co-occurrence matrices proved to be an effective technique, able to discriminate a large number of synthetic and natural textures, the major weakness associated with this approach is that it is

built on the assumption that the textural patterns are repetitive and uniformly distributed within the image.

A non-parametric approach that analyses texture at micro level based on the calculation of the Local Binary Patterns (LBP) has been introduced by Ojala and Pietikainen [7]. This approach attempts to decompose the texture into small texture units where the texture features are represented by the distribution of the LBP values that are calculated for each pixel in the image. In their experiments they demonstrated that this texture descriptor is efficient in discriminating natural and synthetic textures, but it cannot adapt to scale and rotation variation. Sensitivity to texture rotation may be useful for some applications such as surface inspection, but this property of the LBP is a considerable drawback when applied to the segmentation of natural images. To address this issue, Ojala et al [5] proposed a new multi-resolution rotational invariant LBP texture descriptor whose performance was evaluated on standard texture databases.

Based on the review of relevant papers on statistical texture analysis it can be concluded that these methods return adequate results when applied to synthetic images, but their performance is limited when applied to natural images, unless these images are defined by uniform textures. It is also worth mentioning that these techniques are more often used for texture classification, rather than for texture segmentation. Recently, more sophisticated approaches for statistical texture analysis have been proposed and these include the work of Kovalev and Petrou [8], Elfadel and Picard [9] and Varma and Zisserman [10].

In the category of model based methods, Markov Random Field (MRF) models are often employed for texture analysis [11, 12]. The MRF models statistically capture the spatial dependence between pixels and assume that the probability of a pixel taking a certain intensity value depends only on the intensity values of the pixels situated in its neighbourhood. When applied to texture analysis, the performance of the MRF-based techniques proved to be superior to that of the algorithms based on co-occurrence matrices. Also, as opposed to co-occurrence matrices, the MRF texture features are invariant to rotation and they can be further extended to cover scale invariance [13].

However, the main drawback of the MRF models, shared also by statistical texture analysis techniques, is their inability to adapt to local distortions in textures that are very common in natural images. It should also be mentioned that these techniques are computationally intensive and they are not capable of providing accurate texture models for natural images where textures have a high degree of complexity and inhomogeneity.

Another important category of texture analysis techniques is represented by the signal processing methods. These techniques were developed as a consequence of the psychophysical investigations [14] that indicated that the human brain performs a frequency analysis of the image perceived by the retina. Based on this information, the signal processing techniques formulate the texture extraction in terms of the frequency information associated with the texture primitives present in digital images. Representative methods that belong to this category are: spatial domain filtering, Fourier analysis and Gabor and Wavelet analysis. Spatial domain filtering methods extract the edge information (using standard detectors such as Roberts and Sobel) with the aim of measuring the edge density per unit area, relying on the assumption that fine textures are characterized by a higher density of edges per unit area than coarse textures. Building on this concept, Unser and Eden [15] generated a set of multi-resolution images that are obtained after the application of a spatial filter bank to the input image that is followed by the application of an iterative Gaussian smoothing algorithm. From the multi-resolution image stack they calculated a set of energy features that were further reduced by applying the Karhunen-Loève transform [117]. The experimental data provided in their paper indicates that the proposed technique is able to produce accurate segmentation, but the segmentation results are influenced by the appropriate selection of the window size where the texture features are calculated.

Early signal processing methods attempted to analyse the image texture with respect to its Fourier spectrum that offers information about the directionality and periodicity of repeated textured patterns. These techniques have been primarily applied to image classification tasks and the experiments indicate that their performance in texture discrimination is poor. These results were motivated by the fact that the spatial information plays no role in the extraction

of the texture features. This problem is addressed by the Wavelet transform [16, 17, 18] that allows the calculation of the texture features in the spatial/frequency domain. The Wavelet transform analyses the signal by multiplying it with a window function, an architecture that provides a certain degree of spatial localisation and allows texture analysis at multiple scales. Texture segmentation using wavelets has been investigated in [19] where the authors conducted a number of classification experiments using scaled and rotated Brodatz textures. In their tests they achieved 85.40% correct classification when the developed algorithm has been applied to 50 texture classes.

A distinct category of signal processing texture analysis techniques is based on filtering the input image using a bank of multi-channel narrow band Gabor filters. This approach has been first applied by Bovik et al [20], where they used quadrature Gabor filters to segment images defined by oriented textures. The main conclusion resulting from their investigation is that the spectral difference sampled by narrow band filters provides sufficient information for texture discrimination. Jain and Farrokhnia [21] followed a similar approach and developed a multi-channel Gabor filtering technique that was applied for image segmentation. In their paper, each filtered image was subjected to a non-linear transform and the energy was calculated within a pre-defined window around each pixel in the image. The energy features were afterwards clustered using a standard algorithm to obtain the segmented image. This approach was further advanced by Randen and Husoy [22] while noting that filtering the image with a bank of Gabor filters or filters derived from Wavelet transform is computationally intensive. In [22] they proposed a new methodology to compute optimised filters for texture discrimination that requires a reduced number of filters than the standard implementation developed by Jain and Farrokhnia [21]. A different segmentation strategy is proposed by Hofmann et al [23] where the texture segmentation is formulated as a data clustering problem. In their approach the dissimilarities between pairs of textured regions are computed from a multi-scale Gabor filtered image representation. The resulting unsupervised segmentation scheme was successfully applied on both Brodatz textures and natural images.

From the review provided in this section it can be concluded that in spite of the enormous research effort dedicated to the development of optimal texture-based image segmentation strategies, the problem of robust texture extraction is still an open issue.

Structural methods require an exact description of the texture primitives and their spatial distribution within the image. These methods were primarily developed for texture synthesis and since this is a topic beyond the scope of this investigation these approaches are not analysed in this literature survey. Some relevant structural methods include the work of Haralick [3] and Serra [24].

2.2 Colour Analysis for Image Segmentation

Colour is another important characteristic of digital images and started to be investigated more recently than texture. Colour information allows a better description of the objects in digital images than the greyscale representation and practice indicates that its inclusion in the segmentation process considerably improves the accuracy of the segmented result. According to several reviews on colour-based segmentation [25, 26] the existent algorithms can be divided into three main categories namely feature-based, area-based (or region-based) and physics-based segmentation techniques.

The feature-based colour segmentation algorithms are developed based on the assumption that colour is locally homogeneous in the image and the task of segmentation can be viewed as that of grouping the pixels into regions that satisfy a coherence criterion. According to Skarbek and Koschan [27], the feature-based segmentation techniques can be further subcategorised into histogram-based methods and clustering techniques. The histogram-based segmentation approaches attempt to identify apparent peaks and valleys in the colour histograms and they provide a coarse segmentation that is typically refined by applying additional post-processing steps. One of the first colour segmentation algorithms has been proposed by Ohlander et al [28]. Their algorithm is based on a multi-dimensional histogram thresholding scheme where the threshold values employed to partition the input image were selected

in conjunction with the peaks of the histograms calculated from the input image that has been converted into the RGB, YIQ and HSI colour spaces. A different approach has been suggested by Tseng et al [29], where a circular histogram thresholding approach for colour image segmentation was proposed. Their algorithm constructs a circular histogram from the hue component of the HSI colour space that is afterwards smoothed using a scale-space filtering approach. The segmentation process is implemented by a thresholding operation that is applied to partition the histogram into a predefined number of regions using a criterion that maximises the between-class variance and minimises the within-class variance. The authors compared the performance of their algorithm with that offered by the 3D K-Means clustering algorithm and they found that the performance of their segmentation algorithm was marginally lower than that offered by the clustering technique. One major advantage of this algorithm is its low computational cost and in their experiments the authors demonstrated that their algorithm was 100 times faster than the 3D K-Means clustering scheme. Boukouvalas et al [30] proposed a related algorithm that was developed to grade randomly textured multicoloured ceramic tiles. Their approach employs a simple comparison between the colour histograms calculated for the tiles under investigation and tiles that have been previously graded by human experts. The grading process was carried out using simple similarity measures such as chi-square statistics and linear correlation. Since their implementation addressed an industrial application, a number of pre-processing steps were applied to compensate for the non-uniform illumination conditions. The tile grading algorithm was tested on 81 images and the experimental data indicated that the proposed algorithm was feasible to be used in the implementation of an industrial inspection system.

The main weakness of the histogram-based colour segmentation algorithms resides in the fact that the colour information is globally analysed and as a result they are not able to adapt to the local variations in colour information that are often present in natural images. For instance, the peaks in the colour histogram are not always statistically relevant and the segmentation results returned by the histogram-based techniques are often over-segmented. In general, histogram-based methods provide only coarse information about the

image and this is caused in part by the fact that the spatial relationships between adjacent pixels in the image are not used in the segmentation process.

Clustering methods [31] have been widely applied in the development of colour segmentation algorithms and the most popular techniques include the K-Means [32], Fuzzy C-Means [33], Mean Shift [34] and Expectation Maximisation [35, 36, 37]. The principle behind clustering techniques consists of reducing the high dimensionality (number of colours) of the input data to a more compact representation where the number of clusters (or colour prototypes) is specified *a-priori*.

K-Means is one of the simplest clustering algorithms and due to its low computational cost it has been used in many applications such as data compression and vector quantisation [38]. Historically, this clustering technique has been applied to the segmentation of greyscale images and later to the segmentation of colour images represented in different colour spaces [39]. Although the application of the clustering algorithms to colour segmentation proved to be successful, one of their main problems is the fact that no spatial constraints are imposed during the data partitioning process. To address this issue, Luo et al [40] proposed a novel spatially constrained K-Means algorithm that has two distinct computational stages. In the first stage, a spatially constrained region growing is applied at each iteration that is followed by the clustering process that is implemented by the standard K-Means algorithm. In the second stage the algorithm applies a merging procedure in order to reduce over-segmentation. The authors conclude that the explicit inclusion of spatial information in the segmentation process improves the overall performance of the image segmentation algorithm. In spite of its popularity, the standard K-Means clustering algorithm has several disadvantages, where the most important is that the clustering process is based on a rigid space partitioning strategy that consists of the minimisation of a global objective function. To alleviate this problem, Bezdek [33] proposed a fuzzy clustering scheme (also known as Fuzzy C-Means) that is closely related to the K-Means algorithm where it is assumed that all elements of the data belong to all clusters with a degree of membership that is controlled by a fuzziness parameter (also referred to as fuzzifier factor). This clustering strategy has been involved in the

implementation of many applications, but its main drawback consists in the difficulty to select the optimal value for the fuzziness parameter and this fact restricted the application of this clustering algorithm to segmentation tasks where no *a-priori* knowledge is available.

A distinct category of clustering methods is based on the probabilistic partitioning of the data represented in a given colour space with respect to a predefined number of finite mixtures that are usually modelled using Gaussian Mixtures Models (GMM). One of the most popular probabilistic space partitioning techniques is the Expectation-Maximisation (EM) procedure [35, 36] whose aim is the optimisation of the space partitioning process by maximising the likelihood between the input data and an initial set of GMMs that are *a-priori* defined. A recent approach [41] proposes a probabilistic framework for automatic colour image segmentation where the global colour statistics (that are modelled using GMMs) are combined with local image descriptors in order to preserve the spatial and colour coherence. This approach was tested on a number of synthetic and natural images and its performance was compared against that offered by several state of art colour segmentation algorithms. The experimental data indicates that the proposed method outperformed the standard colour segmentation schemes.

The main disadvantage associated with clustering and probabilistic space partitioning algorithms is that their performances are highly dependent on the optimal selection of a relative large number of parameters that have to be known *a-priori* (such as the initial number of clusters (K-Means), fuzziness parameter (Fuzzy C-Means) or the initial estimation and the number of GMMs (EM)). In practice, the selection of these parameters is either carried out experimentally or by developing global techniques such as those based on histogram analysis. To circumvent the development of complex procedures that are applied to estimate the optimal set of clustering parameters, Comaniciu and Meer [34] developed a non-parametric procedure called Mean Shift that is applied to estimate the density gradients of the pattern distributions. The main advantage of this technique is that it can locally adapt to the image content (does not require any initialisation stage) and the experimental data indicates that this algorithm produces superior results than the standard clustering

techniques when applied to natural images affected by noise and uneven illumination.

Another important disadvantage associated with standard clustering algorithms resides in their inability to include the spatial relationships between the data points in the space partitioning process. To address this limitation, Pappas [42] proposed a new technique based on the standard K-Means algorithm where the spatial coherence is enforced during the cluster assignment process. The spatial constraints are imposed by the Gibbs Random Field models, while the local intensity variations are sampled by an iterative procedure that consists of averaging over a sliding window whose size decreases with the increase in the number of iterations. This algorithm has been initially applied to the segmentation of greyscale images and it has been later extended to cover multi-dimensional data by Chen et al [43].

The second category of colour segmentation techniques is represented by the region-based approaches which are arguably considered the most investigated segmentation schemes. Their main advantage over feature-based methods is that the spatial coherence between adjacent pixels (or image regions) in the input image is enforced during the segmentation process.

As indicated in the review on colour segmentation written by Lucchese and Mitra [26], the area-based techniques can be further divided into split and merge, region growing and edge-based segmentation techniques. The split and merge methods start in general with an inhomogeneous partition of the image and they agglomerate the initial partitions into disjoint regions with uniform characteristics. There are two distinct stages that characterize these techniques. In the first phase (splitting) the image is hierarchically divided into sub-blocks until a homogeneity criterion is met, while in the second stage (merging) the adjacent regions that have similar properties are joined, usually using a region adjacency graph (RAG) data structure. Building on this, Round et al [44] proposed a split and merge strategy for colour segmentation of pigmented lesions in dermoscopy images. In their implementation, a quad-tree representation of the input image was achieved by iterative splitting, while the merging phase was implemented using a simple agglomerative process based on a graph adjacency procedure. A similar approach was proposed by Celenk

[45] where a split and merge based hierarchical clustering method was developed for colour image segmentation. In the splitting phase, the input image is uniformly divided into $n \times n$ non-overlapping rectangular partitions, where the size of the initial partitions is a user-defined parameter. For each resulting region, the K-Means algorithm is applied to classify pixels into two classes and then a merging procedure is applied in order to group the patterns resulting from the split stage.

An important limitation of the split and merge techniques resides in the fact that the initial partition resulting from the split stage is formed by rectangular regions. Thus, the result obtained after the application of the merge stage has a blocky structure and cannot accurately capture the shape of the imaged objects. To compensate for this problem, the image resulting from the merge stage is further processed by applying refinement techniques where the pixels situated on the border are re-classified using some similarity criteria. Also, it is useful to mention that the minimum size of the region resulting from the splitting process is an important parameter that influences the overall performance of the segmentation algorithm. In this regard, if the region size is too small, the features calculated in the region under analysis will have low statistical relevance and this has a negative influence towards the decisions that will be made in the merging stage. On the other hand, if the region size is set to large values, the statistical relevance of the features in the region is increased but this is achieved at the expense of missing small and narrow objects in the input image. To address this issue, region growing techniques have been proposed as an alternative to split and merge strategies. In general, the region growing techniques are iterative schemes that start with a selection of initial seeds that are expanded at each iteration based on some homogeneity criteria. The main advantage of these techniques is that during the seed growing process the locality is preserved, but the performance of these approaches is highly dependent on the appropriate selection of the initial seeds. To address this, Tremeau and Borel [46] proposed a colour segmentation algorithm that combines both region growing and region merging processes. In this way, the first step of their algorithm is based on a region growing procedure that takes into account the colour similarity and spatial proximity. The regions resulting

from region growing are merged in accordance with a global homogeneity criterion that is applied to obtain spatially coherent regions. One of the drawbacks of this approach is the subjective selection of the threshold parameters that were employed to evaluate the colour similarity between adjacent pixels and the experimental results proved that these thresholds parameters are image dependent. Also, the performance of this algorithm is poor when applied to images with shadows and highlights. Shih and Cheng [47] proposed a similar segmentation technique where the seeded region growing process is performed in the YCbCr colour space using a three-stage algorithm. In the first stage the initial seeds are automatically selected from the input data by enforcing the condition that the seed pixel must have high similarity with its neighbours (i.e. the maximum Euclidean distance to its eight neighbours must be less than a threshold value). The second step involves the application of a seeded region-growing algorithm, while the last step aims to eliminate the over-segmentation by applying a region merging procedure. The experimental data presented in their paper indicates that this algorithm shares the same limitations as the algorithm developed by Tremeau and Borel [46]. Similar region growing colour segmentation techniques are represented by the works of Cheng and Sun [48], Deng and Manjunath [49] and Moghaddamzadeh and Bourbakis [50]. Based on the survey of relevant region growing techniques it can be concluded that they place a high level of confidence on the appropriate initialisation of a number of threshold parameters that are applied to evaluate the local homogeneity and, as a result these segmentation techniques are not able to adapt to problems caused by image noise, shadows and uneven illumination.

The edge-based segmentation methods aim to detect the boundaries between image objects. In this regard, a large number of approaches formulated the segmentation process as that of fusing the gradient-based edge detection with the information provided by histogram-based or clustering schemes. Colour snakes (active contours) are a distinct category of edge-based segmentation techniques whose aim is to iteratively deform an initial contour by minimising energies that relate to the intrinsic properties of the contour and those dependent on the colour information (image data) in order to capture the

outline of the object of interest in the image. These techniques achieve accurate performance, but the main problem that restricts the application of colour snakes to segmentation tasks is the high level of supervision required to select the initial contour. In addition to this, active contours approaches are implemented using iterative procedures that are computationally intensive and in general they are better suited to be applied to tasks where strong knowledge in regard to the shape of the objects of interest is available.

The third category of colour image segmentation techniques is represented by the physics-based techniques [51, 52] whose aim is to eliminate the over-segmentation caused by uneven illumination, shadows and highlights. These techniques require a significant amount of *a-priori* information in regard to the illumination model and reflectance properties of the scene objects and as a result they have been developed in conjunction with well-defined applications. These techniques are not evaluated in this literature survey since the aim of this research is the development of generic (unsupervised) image segmentation schemes.

2.3 Colour-Texture Image Segmentation

As indicated in the reviews provided in sections 2.1 and 2.2, a significant amount of research has been dedicated to the development of algorithms where the colour and texture features where analysed alone. Most of these algorithms were designed in conjunction with particular applications and they return appropriate segmentation results only when they are applied to images that are composed of scenes defined by regions with uniform characteristics. Segmentation of natural images is by far a more difficult task, since natural images exhibit significant inhomogeneities in colour and texture. Thus, the complex characteristics associated with natural images forced researchers to approach their segmentation using features that locally sample both the colour and texture attributes. The use of colour and texture information collectively has strong links with the human perception, but the main challenge and difficulty reside in the manner these fundamental image attributes are combined in a coherent colour-texture image descriptor.

This section will present a comprehensive review of the existent segmentation approaches based on the way the colour and texture features are integrated in the segmentation process.

One of the first colour texture segmentation algorithms was proposed by Panjwani and Healey [53]. They followed a region-based approach that uses colour Gaussian Markov Random Field (GMRF) models which take into consideration not only the spatial interactions within each of the colour bands but also the interactions between colour planes. The parameters of the GMRF are estimated using maximum likelihood methods and the segmentation algorithm is divided into two main steps. The first step of the algorithm performs region splitting that is applied to recursively split the image into square regions until a uniformity criterion is upheld. The second step implements an agglomerative clustering which merges regions with similar characteristics in order to form texture boundaries. At each step of the merging phase the conditional likelihood of the image is maximised. Experiments were performed on natural colour images and the authors conclude that the use of joint colour-texture models for unsupervised segmentation improves the segmented result when compared to colour alone or texture alone methods. Still they conclude that the availability of *a-priori* image knowledge would improve the effectiveness of the random field models when used in the context of unsupervised segmentation.

Paschos and Valavanis [54] proposed an algorithm that complements the texture information with the chrominance components of the image. The colour space used in their implementation is the *xyY*, where *Y* represents the luminance component that is separated from the chrominance values *xy*. The algorithm first calculates the *xy* chromaticity maps and then performs a spatial averaging procedure. Then, the texture is sampled using the autocorrelation of the image and the corresponding directional histogram. The main aim of this paper was to emphasise the importance of the chromatic content in the process of texture description.

Other researchers adopted different strategies regarding the inclusion of texture and colour in the segmentation process. Tan and Kittler [55] developed an image segmentation algorithm that separately extracts the colour and texture

attributes. Their approach evaluates the image data on two channels, one for texture representation and one for colour description. The texture information is extracted by applying a local linear transform, while the colour is sampled by the moments calculated from the colour histograms. The extraction of colour and texture features separately is extremely appealing since the contribution of colour and texture can be easily quantified in the segmentation process.

Manduchi [56] proposed a similar method that computes the colour and texture features separately and then combines them using a Bayesian framework. Although their technique is general, it requires knowledge in regard to the posterior distributions that are applied to initiate the segmentation process. Jolly and Gupta [57] followed the same approach where the colour and texture features were extracted on separate channels. To compute the texture features, multi-resolution autoregressive models were used and each colour vector is defined by the colour components of each pixel in two colour spaces. In each separate feature space, the maximum likelihood is calculated and the final segmentation is obtained by combining the two likelihoods using some pre-defined fusion criteria. The authors proved that the extraction of the colour and texture features individually help preserve their strength and accuracy. The proposed method was applied to the segmentation of mosaic and aerial images and the experimental results demonstrated that the use of combined colour-texture features significantly improves the segmentation results. However, the calculation of the parameters of the conditional density function (employed in the computation of the maximum likelihoods) requires a significant amount of user intervention.

Another implementation has been proposed by Carson et al [58]. They developed a Blobworld technique that has been applied to the segmentation of natural images in perceptual regions. The central part of this algorithm is represented by the inclusion of the anisotropy, polarity and contrast features in a multi-scale texture model. The colour features are independently extracted and are given by the three colour components of the CIE Lab image obtained after filtering the input image with a Gaussian operator. For automatic colour-texture image segmentation, the authors proposed to model the joint distribution of colour, texture and position features using Gaussian Mixture

Models (GMMs). The main advantage of the Blobworld algorithm consists in its ability to segment the image in compact regions, therefore, it has been included in the development of a content-based image retrieval system.

A similar approach has been developed by Chen et al [43] where the implementation of an algorithm for segmentation of natural images into perceptual regions is detailed. In their approach, the local colour features are extracted using a spatially adaptive clustering algorithm [42], while the texture features are extracted on a different channel using a multi-scale frequency decomposition procedure. The colour-texture features are combined using a region growing algorithm that results in a crude segmentation that is afterwards post-processed using a border refinement procedure. A disadvantage of this algorithm, resides in the fact that the number of classes required by the clustering algorithm has to be manually selected.

An implicit integration of the colour and texture features is given by Shafarenko et al [59] where they proposed a bottom-up segmentation approach developed for the segmentation of randomly textured colour images. Their algorithm starts with a watershed transform that is applied to the image data converted to the LUV colour representation. This procedure results in over-segmentation and to compensate for this problem the resulting regions are merged according to a colour contrast measure until a termination criteria is met (regions separated by a minimum contrast are merged using the Euclidian metric in the LUV colour space). Although the authors assert that the proposed technique is completely automatic and returns good segmentation results, the experimental data indicates that the method has been specifically designed for processing granite and blob like images.

Ojala and Pietikainen [7] proposed a segmentation algorithm based on texture description with feature distributions. The texture information is sampled by the joint distribution of the Local Binary Pattern and Contrast features. The segmentation algorithm consists in a region based method and is divided into three steps. In the first step, hierarchical splitting is applied to partition the image into roughly uniform regions, while in the second step adjacent regions are merged according to their similarity that was sampled using the G-statistics measure. Because the image returned from the merging

stage has a blocky structure, the last step performs a pixelwise classification. This technique returns good results when applied to typical mosaic images, but it cannot correctly segment small and narrow objects that are often encountered in natural images. It also had difficulties with the localisation of the boundaries between the objects present in the image. The Ojala and Pietikainen [7] algorithm was modified by Chen and Chen [60] where the feature distributions (namely the colour histogram and local edge pattern histogram) were extracted from the image that has been first subjected to a quantisation procedure. The segmentation algorithm is based on the three stages described above, the main difference being the application of the histogram intersection similarity measure to evaluate the closeness between two distributions. A similar colour-texture segmentation approach was proposed by Nammalwar et al [61] where the colour and texture were collectively used in a split and merge segmentation scheme. In their paper, the texture features are calculated using the Local Binary Pattern technique, while colour features are extracted using the standard K-Means clustering procedure. The proposed method was tested on mosaic and natural images and the experimental data shows that the inclusion of the colour distribution in the merging process proved to be the key issue in achieving accurate segmentation. This algorithm presents the same drawbacks as the method proposed by Ojala and Pietikainen.

Mirmehdi and Petrou [62] approached the segmentation of colour images from a perceptual point of view. In their paper, they calculate a multi-scale perceptual image tower that is generated by mimicking a human observer when looking at the input image from different distances. The first stage of their algorithm deals with the extraction of the core colour clusters and the segmentation task is defined as a probabilistic process that hierarchically reassigns the non-core pixels starting from the coarsest image in the tower to the image with the highest resolution. The main limitation of this algorithm is that the colour and texture features are not explicitly used and this causes problems in the analysis of their contribution in the overall segmentation process.

Hoang et al [63] proposed a framework to measure the local colour-texture content in the image and they applied their algorithm to the segmentation of

synthetic and natural images. The proposed approach converts the RGB image into a Gaussian colour model and on each colour channel a set of Gabor filters was applied to extract the primary colour-texture features. Because a large number of filters were applied to sample the textural properties of the image, they applied a Principal Component Analysis (PCA) technique to reduce the dimension of the feature space from sixty to four. The resulting feature vectors are used as inputs for a K-Means algorithm that is applied to provide the initial segmentation that is further refined by a region-merging procedure. The main advantage of their algorithm is the application of the standard multi-band filtering approach and the representation of the colour image in the wavelength Fourier space. This algorithm shares the same problems as the segmentation scheme proposed by Mirmehdi and Petrou [62], namely the ambiguous inclusion of the texture and colour features in the segmentation process.

Deng and Manjunath [49] proposed a different colour-texture segmentation technique (referred to as JSEG) that consists of two independent computational stages: colour quantisation and spatial segmentation. During the first stage, the colour information from the input image is quantised into a representative number of classes without enforcing the spatial relationship between pixels. The aim of this process is to map the image into a structure where each pixel is assigned a class label. The next stage of the algorithm enforces the spatial composition of the class labels using a segmentation criterion (J value) that samples the local homogeneity. The main merit of this paper is the use of colour and texture information in succession and the authors argue that this approach is beneficial since it is difficult to analyse the colour similarity and spatial relationships between the neighbouring pixels at the same time. One disadvantage of this method is that the J values cannot distinguish adjacent regions with similar texture patterns and different colour contrast and moreover its performance is dependent on the optimal selection of internal parameters. To address these limitations, different colour-texture segmentation approaches based on JSEG have been proposed. In this regard, Wang et al [64] suggested additional measures to integrate directional operators into the J measure and Wang et al [65] replaced the colour quantisation phase of the JSEG algorithm with an adaptive Mean Shift based clustering. Zheng et al [66] followed the

same idea and combined the quantisation phase of the JSEG algorithm with fuzzy connectedness. All these improvements led to an increase in the number of parameters that have to be optimised and the performances of these JSEG-based algorithms were only marginally better when compared to the original implementation.

Liapis and Tziritas [67] proposed an image retrieval algorithm where the colour and texture are separately extracted and afterwards combined using the Bhattacharya distance. In their implementation, the texture features were extracted using a Discrete Wavelet Frame analysis and modelled using Laplacian distributions, while the colour features were defined by the 2D histograms of the chromaticity components of the CIE Lab (the histogram was modelled according to a Gaussian distribution and quantised to 1024 chromaticity bins). They tested the proposed retrieval system on images from the Brodatz album and on natural images from the VisTex database and Corel Photo Gallery. The experimental results have shown that the proposed technique outperformed several state of art image retrieval techniques, but the authors indicated that their algorithm produces inaccurate results when applied to images defined by random or chaotic patterns.

A different segmentation approach was recently proposed by Shi and Funt [68] where they used a quaternion colour-texture representation. The algorithm consists of three computational stages. During the first stage, feature vectors are generated by applying a Quaternion Principal Component Analysis to the training data obtained from a set of sub-windows taken from the input image. In the second step these vectors are grouped using a K-Means clustering algorithm, while in the final stage a region merging is applied to join the regions defined by similar textures. The authors state that the use of a quaternion representation to measure the RGB colour-textures is advantageous as both intra- and inter-channel relationships between neighbouring pixels are taken into account. Although interesting, the performance of this colour-texture segmentation scheme is highly dependent on the appropriate selection of the size of the sub-windows where the colour-texture features are calculated and also on several user defined parameters (like the merging threshold and the number of initial clusters).

Freixenet et al [69] proposed to combine the region and boundary information for colour-texture segmentation. To achieve this, they employed an active region growing approach where the initial seeds are sampled from the regions obtained as a combination of perceptual colour and texture edges. To this end, the combined colour-texture properties of the regions were modelled by the union of non-parametric kernel density estimators and classical co-occurrence matrices where initial seeds compete for feature points by minimising an energy function that take both region and boundary information into account. A similar approach has been recently proposed by Sail-Allili and Ziou [70] where the colour-texture information is sampled by compound Gaussian Mixture Models and the region boundaries are estimated in conjunction with a polarity measure. Similar to the algorithm proposed by Freixenet et al [69], colour-texture segmentation is defined in terms of minimising an energy function. A related technique was proposed by Luis-Garcia et al [71], where the local structure tensors and the original image components where combined within an energy minimisation framework to accomplish colour-texture segmentation. Since these methods were based on energy minimisation techniques they have the advantage of enforcing strong geometrical constraints during boundary propagation, but it is useful to note that this advantage is achieved at the expense of increasing the level of supervision, as several parameters need to be specified *a-priori* to control the evolution of the algorithm at each iteration.

As indicated in the literature survey provided in this section, a large number of approaches were proposed to integrate the colour and texture features to achieve robust segmentation, but most of the developed techniques attempted to exploit in a simplistic manner the complementary character of the texture and colour. However from this literature survey it can be concluded that the use of texture and colour features in the image segmentation process leads to improved performance (when compared to the performance of the colour or texture alone segmentation algorithms), but the main issue that has not been addressed is how to include the texture and colour information coherently in a robust image descriptor. Another issue resulting vividly from the review detailed in Section 2.3 is that the performance of the developed algorithms is

highly dependent on the optimal selection of user defined parameters and this issue represents a major drawback since the application domain of these algorithms is limited.

2.4 Conclusions

From the literature survey detailed in this chapter it can be concluded that the adaptive inclusion of the colour and texture features in the segmentation process represents the key issue in achieving accurate results. Although several attempts have been made towards the development of a robust unsupervised colour-texture segmentation scheme, still there is no colour-texture framework that is widely accepted as generic by the computer vision community.

Thus, the main objective of this thesis is to detail the development of a theoretical image segmentation framework that is capable of adaptively exploiting the complementary character of colour and texture and accurately extract the coherent regions in the image. While most of the developed colour-texture algorithms involve a high level of supervision, one of the main goals of this work was to completely eliminate any supervision during the segmentation process. As mentioned in Chapter 1, the proposed segmentation framework extracts the colour and texture attributes in parallel on two different channels with the purpose of maintaining the strength of each feature and quantifying precisely their influence in the segmentation process. In the following chapters of this thesis, the extraction of the colour and texture features will be detailed and the discussion will be continued with their adaptive integration into a novel spatially coherent clustering strategy.

Chapter 3

Adaptive Pre-Filtering Techniques for Colour Image Analysis

Data smoothing is a fundamental operation in the field of computer vision that is carried out in order to pre-process an input image for further analysis. In the context of colour image segmentation, the purpose of the data smoothing process is to eliminate or reduce the level of image noise and improve the local colour homogeneity. The literature on data smoothing is vast and the existing algorithms can be divided in two major categories: linear and non-linear [72, 73]. Some representative traditional linear smoothing techniques are the mean (average) filtering and Gaussian smoothing. Average-based filtering methods simply replace the intensity value for each pixel in the image with the mean value of its neighbours. This approach reduces the noise level, but it is worth mentioning that this solution is ill-suited since the data smoothing is achieved at the expense of severe edge attenuation that leads to poor feature preservation. Gaussian smoothing is similar to mean filtering but the main difference is that the weights assigned to the pixels situated in the neighbourhood of the central pixel are modelled by a Gaussian function and as a result the attenuation produced after the application of the local averaging operation is not as severe as that generated by the mean filtering strategy. To address the limitations associated with linear filtering schemes, a large number of non-linear smoothing techniques were proposed. The most popular non-linear filtering methods include median filters [72], statistical approaches based on non-parametric estimators [74, 75] and more recent developments based on non-linear diffusion. Among non-linear smoothing techniques, the anisotropic

diffusion received a special attention from the vision community [76, 77, 78, 79, 80, 81] since this approach offers the optimal trade-off in smoothing efficiency, removal of weak texture and feature preservation.

In this chapter three non-linear advanced filtering techniques are analysed in the context of adaptive data smoothing of colour images. The first technique evaluated in this thesis is the bilateral filtering, while the second is the anisotropic diffusion that was originally developed by Perona and Malik [76]. Finally, the Forward and Backward (FAB) anisotropic diffusion will be detailed, while the main emphasis will be placed on the optimisation of the FAB anisotropic diffusion by the inclusion of a boost function that increases the values of the medium gradients that are generated by the low colour contrast.

The aim of this chapter is to evaluate several techniques for adaptive colour smoothing that are able to deal with multi-dimensional data since the application of one-dimensional (1D) filters on each colour channel will generate spurious colours in the output data.

Because the analysed filtering schemes have the purpose of reducing the complexity of the image data before the application of the colour feature extraction procedure, their performance will be measured by evaluating the region homogeneity and the edge preservation strength.

3.1 Bilateral Filtering for Colour Images

The bilateral filtering technique was proposed by Tomasi and Manduchi [82] with the aim of smoothing colour data while preserving the most important edges present in the input image. The basic idea that lies behind this approach is to model the spatial and intensity information between the data points during the filtering process. As opposed to traditional linear smoothing strategies such as Gaussian filtering, the spatial averaging in the bilateral filtering process is reformulated in order to assign weights to the pixels situated in the neighbourhood of the central pixel by calculating a function that measures the pixel closeness in the spatial domain. The other component of the bilateral filtering measures the similarity of the intensity values of the pixel

under analysis with the intensity values of the pixels situated in its neighbourhood. In this fashion, the smoothing process is more pronounced in regions defined by noise and weak textures and this process is halted for locations in the image data with step intensity discontinuities. The bilateral filtering is applied to the entire data in a non-iterative manner by replacing the value of each pixel under analysis with the value calculated using the following expression:

$$h(x) = k^{-1}(x) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\tau) c(\tau, x) s(f(\tau), f(x)) d\tau \tag{3.1}$$

In the above equation x is the central pixel located at positions (i, j), τ is the neighbouring pixel at positions (m,n) in the image, while f defines the original colour image. Function $c(\tau, x)$ measures the spatial closeness between the central pixel x and the neighbouring pixels τ, while function $s(f(\tau), f(x))$ measures the similarity between the intensity values of the central pixel and the neighbouring pixels. The component $k(x)$ is a normalisation term that enforces that the sum of the weights for all pixels in the neighbourhood around the pixel x to be 1:

$$k(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(\tau, x) s(f(\tau), f(x)) d\tau \tag{3.2}$$

The spatial closeness and the intensity similarity between two data points are sampled using the Gaussian function in conjunction with the Euclidian distance in the colour space as follows:

$$c(\tau, x) = e^{\frac{-\|\tau - x\|^2}{2\sigma_d^2}} \tag{3.3}$$

$$s(f(\tau), f(x)) = e^{\frac{-\|f(\tau) - f(x)\|^2}{2\sigma_r^2}} \tag{3.4}$$

where σ_d is the standard deviation that controls the smoothing process in the spatial domain and its value is chosen with respect to the strength of the filtering process. Thus, large values of σ_d imply more spatial blurring as

intensities from more distant image locations have a larger contribution in equation (3.1).

In a similar fashion, σ_r controls the smoothing process with respect to the intensity variation and the smoothing is more pronounced for larger values. These parameters control to what extent the spatial and intensity information is preserved during the smoothing process and the optimal set of parameters can be set in conjunction with the level of noise present in the input image. For instance, if the objective is the preservation of very narrow details in the input data, then it will be best to set the parameter σ_d to low values since the decay of $c(\tau, x)$ will be more pronounced with distant pixels and only a small number of pixels will have a significant contribution in equation (3.1). Using a similar approach, if the purpose it to preserve the intensity discontinuities in the colour data, the value of σ_r needs to be set to low values. This can be clearly observed in Figure 3.1 where it is illustrated the effect of the bilateral filtering on a natural image when different combinations of parameters σ_d and σ_r are employed in the smoothing process.

For images depicted in Figure 3.1 (b-e) σ_d is kept constant to 3 and σ_r is varied, while for images shown in Figure 3.1 (f-i) σ_d is kept constant to 10 and σ_r is varied. For small values of σ_r edges are preserved; when this parameter is set to high values, the filtered result will be similar to that achieved by a typical Gaussian smoothing. In order to obtain a filtered image with crisp details, both σ_d and σ_r should be set to small values. In Figure 3.1 it can be observed that the best smoothing/edge-preservation criterion is obtained for the image depicted in Figure 3.1 (b) where the object borders are not blurred and the weak textures are eliminated. In cases when the input image is corrupted with noise, then the values of the parameters σ_d and σ_r should be modified as shown in Figure 3.2 (the original image depicted in Figure 3.2 (a) has been corrupted with Gaussian noise with a standard deviation of 30 intensity levels on each colour channel).

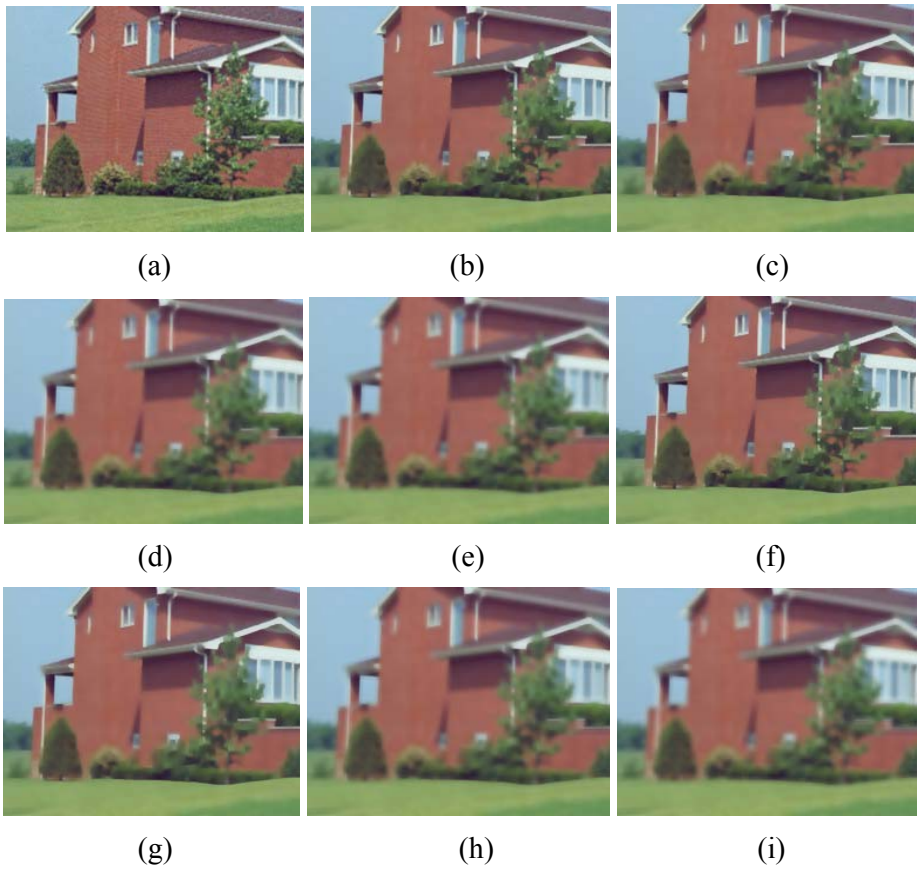
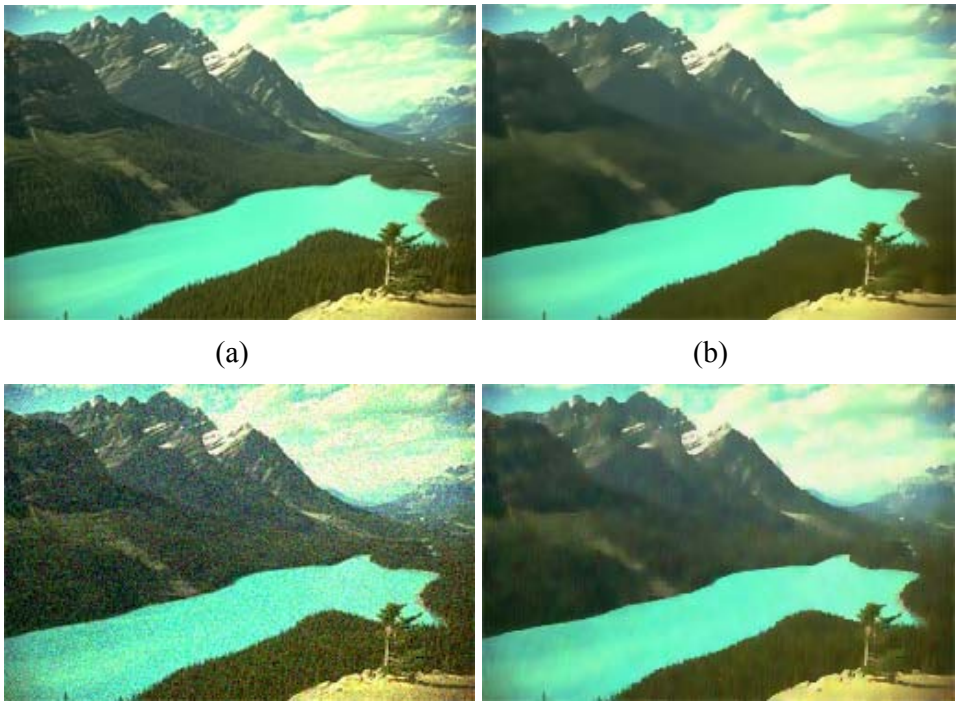


Figure 3.1 (a) Original natural image. (b-i) Smoothed images when different combinations of parameters σ_d and σ_r are employed in the bilateral filtering process. (b) σ_d =3 and σ_r =20. (c) σ_d =3, σ_r =30. (d) σ_d =3, σ_r =60. (e) σ_d =3, σ_r =90. (f) σ_d =10, σ_r =20. (g) σ_d =10, σ_r =30. (h) σ_d =10, σ_r =60. (i) σ_d =10, σ_r =90.



(a) (b)

(c) (d)

Figure 3.2 (a) Original natural image. (b) Image obtained after the application of bilateral filtering with the following parameters σ_d =3 and σ_r =20. (c) The original image shown in (a) was corrupted with Gaussian noise (standard deviation of 30 intensity levels on each colour channel). (d) Image obtained after the application of bilateral filtering with the following parameters σ_d =5 and σ_r =30. It could be observed that good feature preservation is obtained even for image areas characterized by low signal to noise ratio.

3.2 Perona-Malik Anisotropic Diffusion for Colour Images

The second technique analysed in this study was originally developed by Perona and Malik (PM) with the purpose of implementing an optimal, feature preserving smoothing strategy [76]. In their paper, smoothing is formulated as a diffusive process and is performed within the image regions and suppressed at regions boundaries. This non-linear smoothing procedure can be defined in

terms of the derivative of the flux function as illustrated in the following equation:

$$\frac{\partial I(x,y,t)}{\partial t} = div[D(|\nabla I(x,y,t)|)\nabla I(x,y,t)] \tag{3.5}$$

In equation (3.5), $I(x, y)$ defines the image data, $\nabla I(x,y,t)$ is the gradient operator at position (x, y) and iteration t, $|\nabla I(x,y,t)|$ is the magnitude of the gradient operator, $D(.)$ represents the diffusion function and *div* is the divergence operator. The central part of this formulation is played by D that controls the smoothing process and it is implemented using an exponential function as follows:

$$D(|\nabla I(x,y,t)|) = e^{-\left(\frac{|\nabla I(x,y,t)|}{d}\right)^2}, d > 0. \tag{3.6}$$

The most important parameter of the PM anisotropic filtering is the diffusion parameter d that controls the level of smoothing. The value of d is chosen in conjunction with the level of noise and the smoothing effect is more pronounced for higher values of d. It can be observed that the diffusion function $D(.)$ is bounded in the interval (0, 1] and decays with the increase of the gradient value ∇I. The smoothing strategy described in equation (3.5) can be implemented in the discrete domain using an iterative formulation as follows:

$$I(x,y,t+1) = I(x,y,t) + \lambda \sum_{j=1}^{4} [D(|\nabla_j I(x,y,t)|)\nabla_j I(x,y,t)] \tag{3.7}$$

In equation (3.7), $I(x, y, t+1)$ is the pixel intensity value at position (x, y) and iteration $(t+1)$, $\nabla_j I$ is the gradient operator implemented as the 4-connected nearest-neighbourhood differences (see equation (3.8)) and λ is the contrast parameter that takes values in the interval (0, 0.25] [76].

$$\begin{aligned}
\nabla_1 I(x,y) &\equiv I(x-1,y) - I(x,y) \\
\nabla_2 I(x,y) &\equiv I(x+1,y) - I(x,y) \\
\nabla_3 I(x,y) &\equiv I(x,y-1) - I(x,y) \\
\nabla_4 I(x,y) &\equiv I(x,y+1) - I(x,y)
\end{aligned} \tag{3.8}$$

Although the PM anisotropic diffusion is an efficient smoothing procedure, it is useful to note that it has several drawbacks, namely problems with stability and the weak response of the diffusion function to medium gradients. The stability problems were caused by the fixed contrast parameter λ that produces intensity offsets between the input and output data. The problems generated by the inability of the exponential function D to model the diffusion process for low and medium gradients received less attention from the vision community. This is surprising since the efficiency of the smoothing procedure is greatly influenced by the optimal use of the gradient data. To eliminate the limitations associated with the original PM formulation, the Forward and Backward (FAB) anisotropic diffusion has been proposed [83, 84].

3.3 Forward and Backward Anisotropic Diffusion for Colour Images

The goal of the FAB diffusion function is to highlight the medium and large gradients that are noise independent. This can be achieved through a conceptually simple procedure that consists of reversing the diffusion process. This can be implemented by applying two diffusions simultaneously. The forward diffusion acts upon the low gradients that are usually caused by noise, while the backward is applied to reverse the diffusion process when dealing with medium gradients. The combination of these two processes results in a coupled forward and backward diffusion procedure, where the diffusion function D_{FAB} has both negative and positive values.

This can be observed in Figure 3.3 where it is shown that the D_{FAB} diffusion function (defined in equation (3.9)) becomes negative for medium values of the gradient:

$$D_{FAB}(|\nabla I(x,y,t)|) = 2e^{-\left(\frac{|\nabla I(x,y,t)|}{d_1(t)}\right)^2} - e^{-\left(\frac{|\nabla I(x,y,t)|}{d_2(t)}\right)^2} \tag{3.9}$$

In equation (3.9), $d_1(t=0)$ defines the initial diffusion parameter, ∇I represents the image gradient, $d_1(t)$ and $d_2(t)$ are the time dependent parameters that control the forward and backward diffusion respectively and t is the time or iteration step.

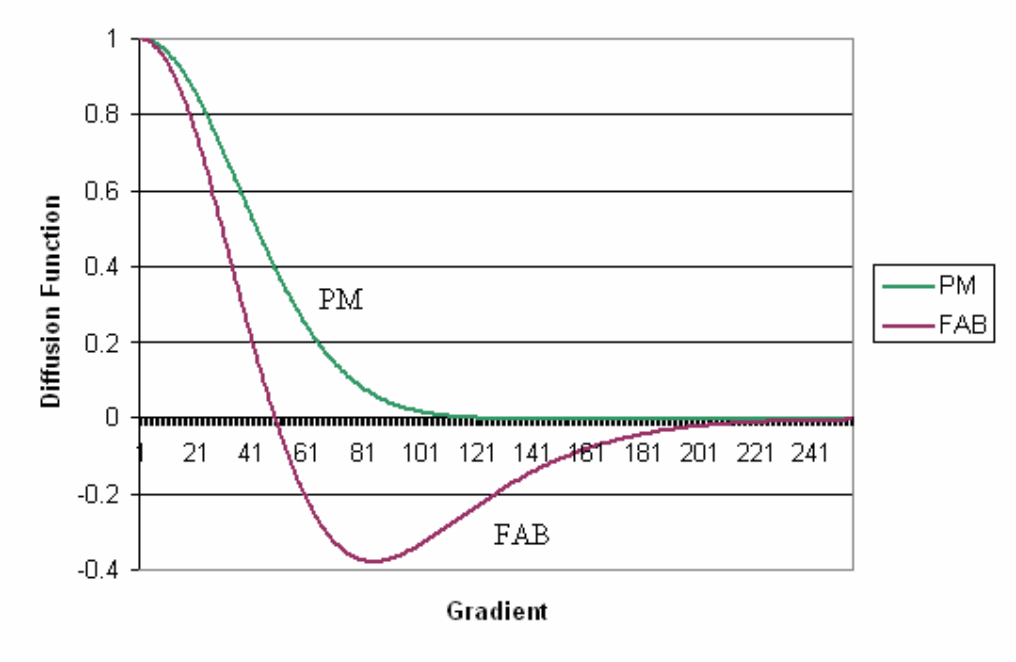


Figure 3.3 Comparison between the Forward and Backward (FAB) diffusion function and the standard Perona-Malik (PM) diffusion function. Parameters are set as follows, $d = 40$ (PM) and $d_1(t=0) = 40$, $d_2(t=0) = 80$.

Since the D_{FAB} function is defined by two parameters, the problems associated with stability are more difficult to control. To address these problems, Smolka and Plataniotis [84] proposed the inclusion of a time dependent cooling

procedure where the values of the diffusion parameters are progressively reduced with the increase in the number of iterations:

$$d_i(t+1) = d_i(t) \cdot \gamma \;,\; i = 1,2 \text{ and } d_i(t+1) < d_i(t) \,,\, \gamma \in (0,1] \qquad (3.10)$$

In equation (3.10) γ is a fixed parameter that takes values in the interval (0, 1]. If γ=1 no cooling takes place (diffusion parameters take the same values at each iteration). The cooling process is faster for lower values of γ. Since the diffusion parameters are lowered at each iteration, this will decrease the strength of the smoothing process and for low values of the diffusion parameters this process is halted.

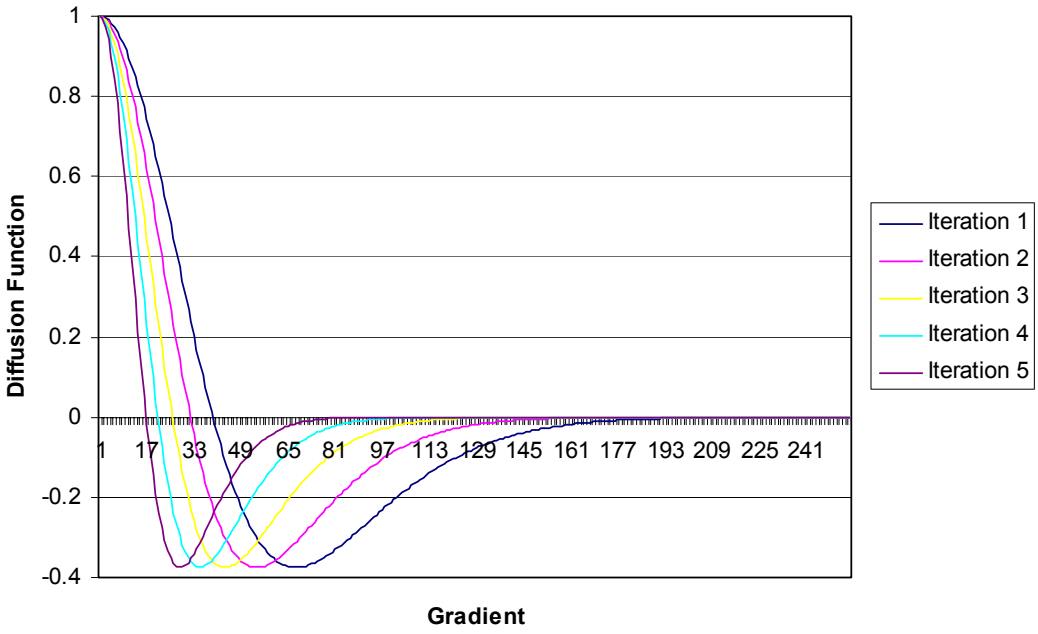


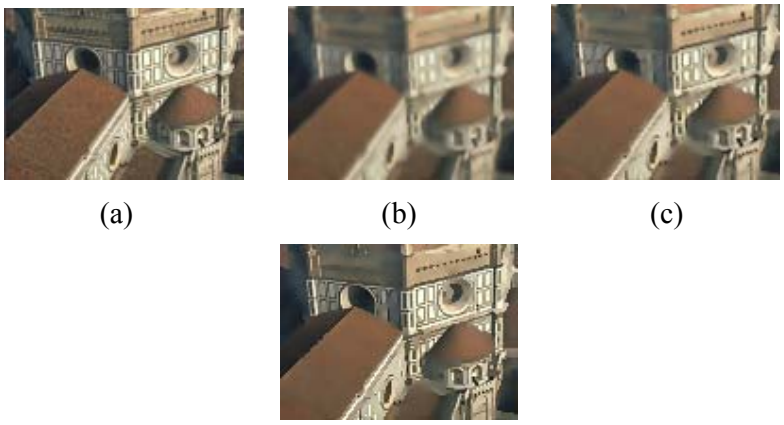
Figure 3.4 The effect of the D_{FAB} cooling process. Note that the position where the curve intersects the x axis is lowered at each iteration which implies less smoothing. The parameters are $d_1(t=0) = 40$, $d_2(t=0) = 80$.

Since the smoothing process will naturally cool itself with the increase in the number of iterations, the FAB anisotropic diffusion does not require the implementation of complex stopping rules or the use of a fixed parameter to

indicate the number of iterations, as is the case with the PM algorithm. The cooling process is illustrated in Figure 3.4 where the D_{FAB} curves for five consecutive iterations are depicted.

3.4 Gradient-Boosted Forward and Backward Anisotropic Diffusion

While the FAB anisotropic diffusion eliminates some of the problems associated with the standard PM anisotropic diffusion, the experimental results show that the smoothed data is still blurred especially around regions defined by medium gradients. This difference can be observed in Figure 3.5 (b) and (c) where the PM and FAB filtered results are illustrated when applied to an image with a high amount of detail.



(a) (b) (c)

(d)

Figure 3.5 (a) Original image. (b) PM filtered image (d = 40). (c) FAB filtered image. ($d_1(t=0) = 40$, $d_2(t=0) = 80$). (d) GB-FAB filtered image.

To further reduce the level of blurriness in the filtered data, in this thesis the inclusion of a boost function to amplify the medium gradients and obtain much crisper image details (see Figure 3.5 (d)) is proposed [85]. This is achieved by

replacing the gradient value in equation (3.9) with a boosted value which is calculated as follows:

$$|\nabla I(x,y,t)| \leftarrow |\nabla I(x,y,t)|(1+2e^{-\frac{||\nabla I(x,y,t)|-m|}{d_1(t)}}) \tag{3.11}$$

where m is the median value of the gradient data. The boost function is depicted in Figure 3.6 and typical results with and without boosting are illustrated in Figures 3.5 and 3.7. It can be observed that the application of the gradient boosting procedure generated a smoothed image with crisper details (see the area around the cart's wheel in the right hand side of the image).

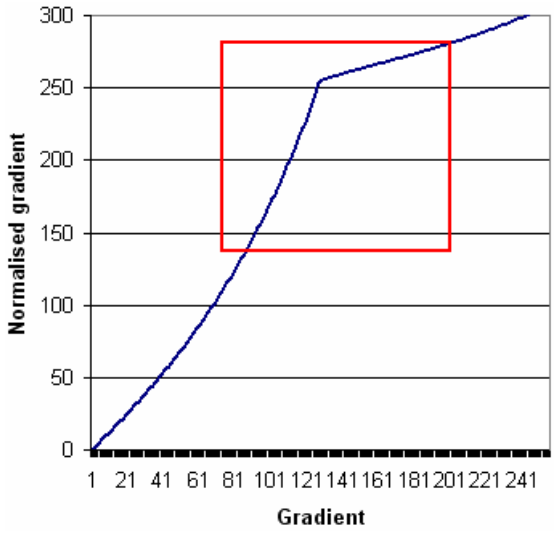


Figure 3.6 Gradient boosting function. Note the amplification of the gradients with medium values - marked in the box.

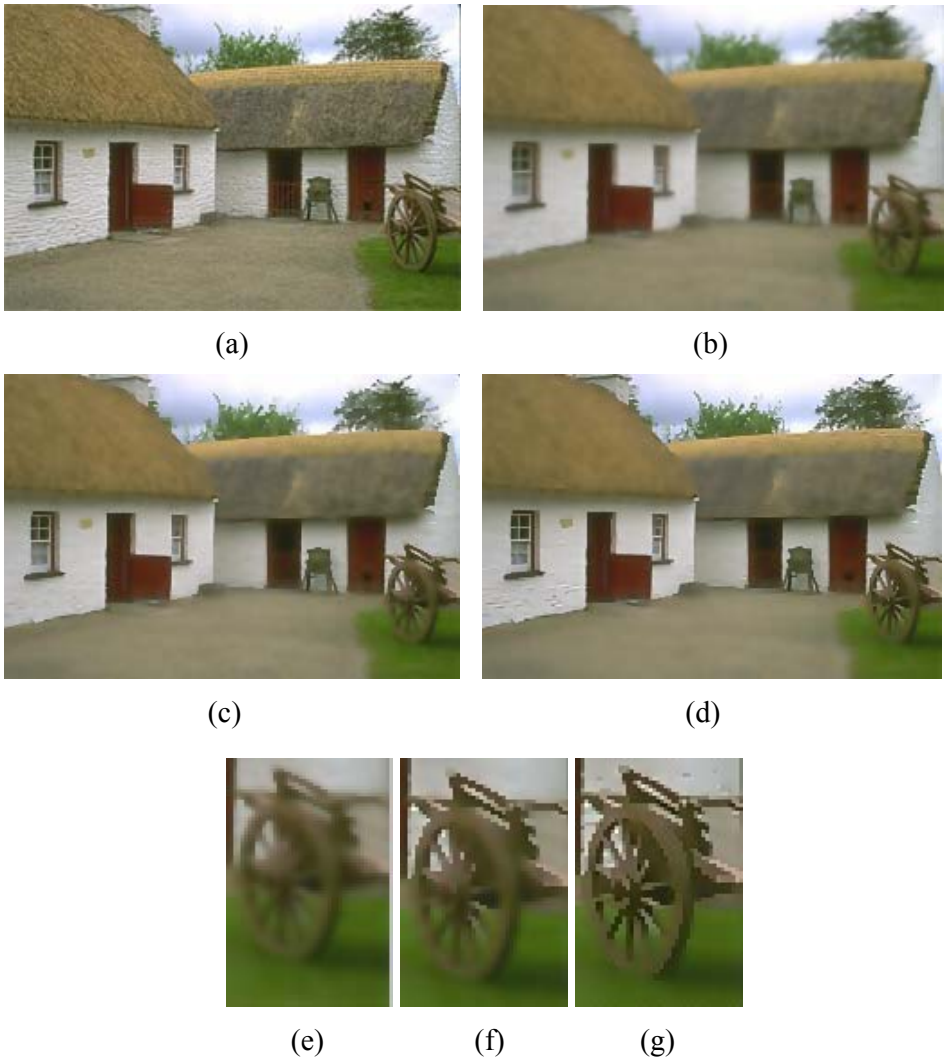


Figure 3.7 Smoothing results when the anisotropic diffusion is applied to the image depicted in (a). (b) PM filtered image (d = 40). (c) FAB filtered image – no gradient boosting (($d_1(t=0)$ = 40, $d_2(t=0)$ = 80). (d) GB-FAB filtered image ($d_1(t=0)$ = 40, $d_2(t=0)$ = 80). (e-g) Close-up details for the results depicted in (b), (c) and (d) respectively.

3.5 Experiments and Results

The smoothing strategies discussed in the previous sections of this chapter have been applied to a large number of natural images. Examples are shown in Figure 3.8 where the bilateral filtering, Perona-Malik anisotropic diffusion and GB-FAB anisotropic diffusion have been applied to several natural images.

From a visual evaluation of these results it can be observed that the highest amount of blur is produced by the PM algorithm, followed by the bilateral filtering where features are generally better preserved but still important image details such as edges are attenuated. The best performance is achieved by the GB-FAB anisotropic diffusion as it enhances the important image features and efficiently smoothes the colour inhomogeneities and weak textures in the input image.

In order to quantitatively evaluate the performance of the adaptive smoothing algorithms, we need to asses their ability in reducing the image noise, improve colour homogeneity while preserving the important discontinuities present in the colour data (the linear filtering techniques such as mean and Gaussian filtering mentioned in the introductory section are not evaluated in this study, since their performance when applied to colour images is poor). The edge preservation is evaluated by plotting the intensity data at selected image locations before and after the application of the filtering techniques analysed in this chapter. Experimental results are shown in Figures 3.9 and 3.10. From the graphs shown in these figures, it can be concluded that the gradient boosted (GB-FAB) anisotropic diffusion clearly outperforms the other evaluated smoothing techniques, as edges are enhanced and the interiors of the regions are smooth (they do not show any oscillations in intensity data). It can also be observed that the edge localisation is significantly improved when images are filtered with the GB-FAB smoothing strategy. The evaluation of the edge localisation is an accurate measure when analysing smoothing techniques for colour image pre-processing as these graphs give clear indications in regard to both feature preservation and local region smoothing.

Figure 3.8 Additional results. First row (a-c) Natural images. Second row (d-f) Bilateral filtering results. Third row (g-i) PM anisotropic diffusion results. Fourth row (j-l) GB-FAB anisotropic diffusion results.

In the next experiment, to globally determine the efficiency of intra-region smoothing performance of the analysed filtering strategies the standard deviation from all colour channels is calculated. In order to achieve this, for every pixel in the image the standard deviation in a 5×5 neighbourhood is computed. The resulting values are sorted according to their magnitude and 25% of the highest values were eliminated (as they are likely to correspond the image edges), as well as 25% of the lowest values were eliminated (as they belong to homogeneous regions and as a consequence they do not present

significant discontinuities). In order to get a quantitative estimation, the Root Mean Square (RMS) values of the standard deviations are calculated. The experimental results shown in Table 3.1 indicate that the bilateral and standard (PM) anisotropic filtering generate images with the lowest RMS values of the standard deviation values. This indicates that these techniques generate smoother images but this is achieved especially at the expense of the attenuation of edges associated with medium gradients.

The experimental results illustrated in Table 3.1 indicate that the gradient boosted GB-FAB scheme achieves an appropriate level of smoothing but not at the expense of poor feature preservation. The parameters involved in these experiments are set to the following default values: for bilateral filtering $\sigma_d = 3$ and $\sigma_r = 20$, for PM anisotropic diffusion $k = 40$, while for the proposed GB-FAB smoothing algorithm the parameters are set to $d_1(t=0) = 40$, $d_2(t=0) = 80$ and $\gamma = 0.8$.

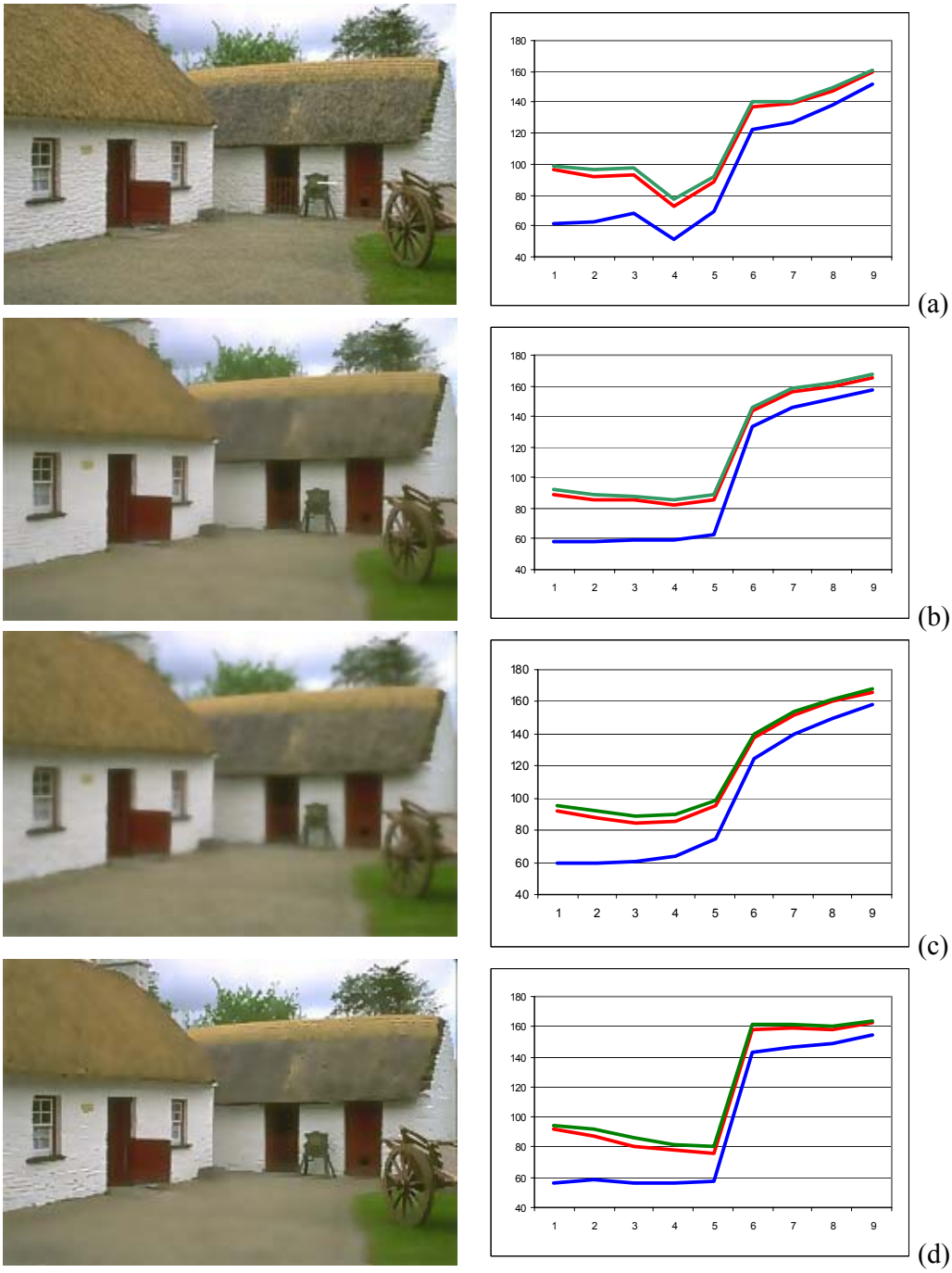


Figure 3.9 Analysis of feature preservation. (a) Original image (the data plotted in these graphs is marked with a white line in the chair area). (b) Bilateral filtering (σ_d =3 and σ_r =20). (c) PM anisotropic diffusion. (d) Gradient-Boosted (GB) FAB anisotropic diffusion. (In the graphs displayed on the right hand side of the diagram the *x*-axis depicts the pixel position on the white line, while on the *y*-axis the pixel's RGB values are plotted).

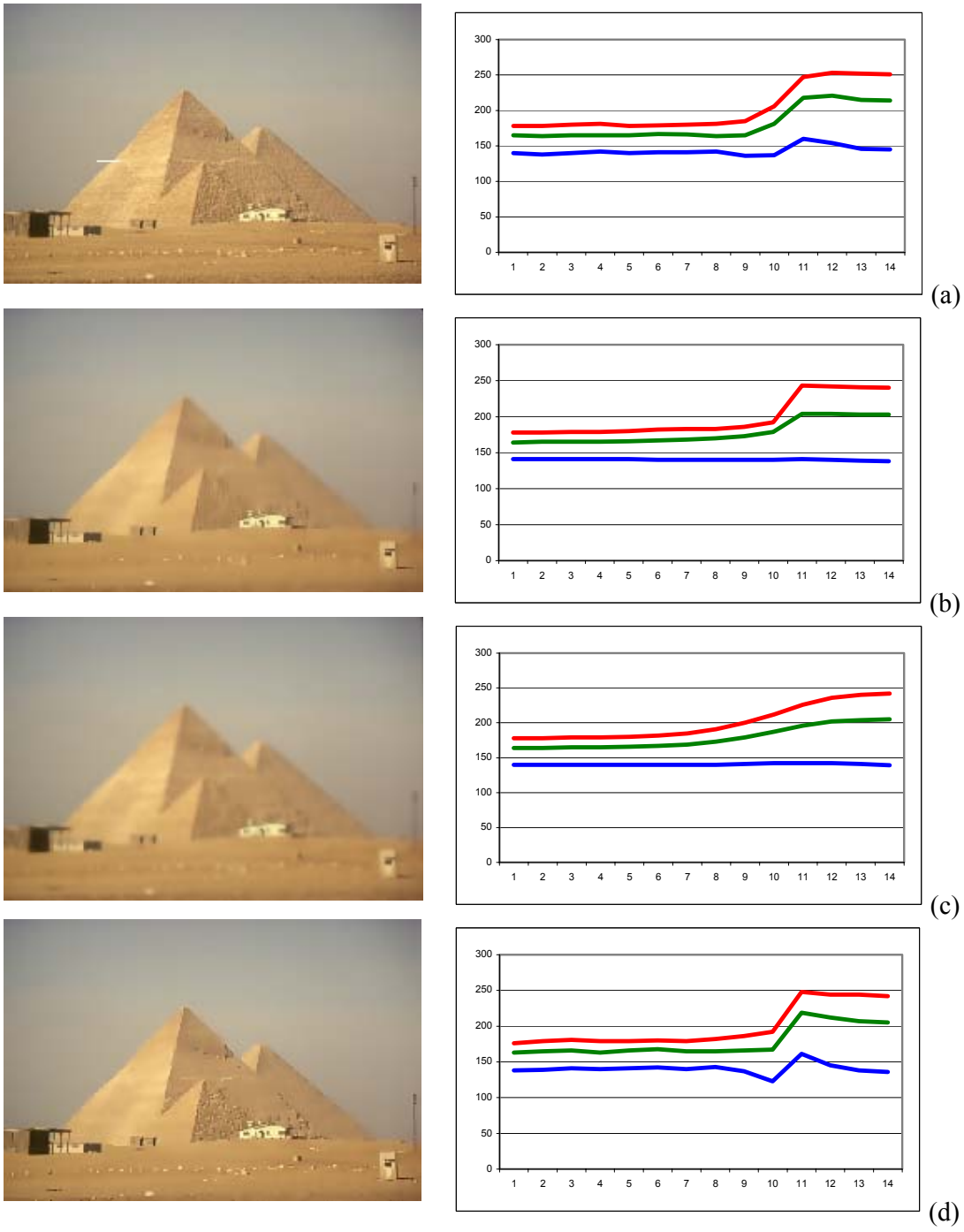


Figure 3.10 Analysis of feature preservation. (a) Original image (the data plotted in these graphs is marked with a white line). (b) Bilateral filtering (σ_d =3 and σ_r =20). (c) PM anisotropic diffusion. (d) Gradient-Boosted (GB) FAB anisotropic diffusion. (In the graphs displayed on the right hand side of the diagram the *x*-axis depicts the pixel position on the white line, while on the *y*-axis the pixel's RGB values are plotted).

TABLE 3.1

The RMS of the standard deviation values for the original and filtered images used in the experiments performed in this chapter.

Figure Number	**Original**	**Bilateral**	**PM**	**FAB**	**GB-FAB**
Figure 3.1(a)	21.93	11.54	12.96	15.03	17.09
Figure 3.2(a)	17.25	7.96	9.28	10.90	12.17
Figure 3.8(a)	31.45	17.01	18.30	20.51	26.55
Figure 3.8(b)	18.04	8.26	9.43	10.63	13.60
Figure 3.8(c)	15.92	6.18	6.94	7.77	10.03
Figure 3.9(a)	21.85	9.44	10.48	8.96	15.21
Figure 3.10(a)	6.57	2.45	2.77	2.83	4.32

3.6 Conclusions

The aim of this chapter was to describe the implementation of three feature preserving smoothing schemes where the main emphasis was placed on evaluating their performances when applied to colour images. The bilateral filtering, PM anisotropic diffusion, FAB anisotropic diffusion and the new GB-FAB anisotropic diffusion have been analysed in this chapter in the context of colour pre-processing. Since the standard smoothing techniques offer inadequate feature preservation for edges produced by medium gradients, in this thesis a new approach to boost the medium gradient data is proposed. The experimental data confirms that the inclusion of gradient boosting in the

implementation of the FAB anisotropic diffusion generates images with crisper details where the contrast in the colour data is preserved. The experimental results also indicate that the proposed GB-FAB smoothing outperforms the others analysed schemes as it enhances the important image features while reducing the level of noise.

The GB-FAB anisotropic diffusion scheme will be included in the development of an adaptive colour segmentation algorithm that is detailed in the next chapter and proved to be an important factor in reducing the level of over-segmentation caused by inhomogeneities, shadows and weak textures.

Chapter 4

Colour Features Extraction

One of the major contributions of this thesis consists of a new formulation for the extraction of colour features that involves a statistical analysis of the input image in multi-space colour representations. In the proposed method, two complementary colour spaces are employed in an advanced statistical clustering process, where the key component is the inclusion of the Self Organising Map (SOM) network that is applied to automatically compute the optimal parameters required by the data clustering algorithms. A colour saliency measure that samples the contrast between the neighbouring regions is applied for parameter optimisation.

The first section of this chapter provides an overview of the overall computational scheme of the colour extraction algorithm. Section 4.2 describes the K-Means clustering technique and the associated limitations, while in Section 4.3 the automatic initialisation procedure proposed for the detection of the optimal number of clusters and the corresponding dominant colours (initial seeds) is detailed. Section 4.4 presents the multi-space computational architecture, while the experimental results presented in Section 4.5 indicate that the proposed technique is accurate in capturing the colour data characteristics when applied to complex natural images.

4.1 Overview of the Colour Segmentation Algorithm

The main computational steps of the multi-space colour extraction algorithm are illustrated in Figure 4.1 (a discussion in regard to the selection of the colour spaces used in the colour segmentation algorithm presented in this chapter is provided in section 4.5.1). The original image is first converted to the perceptually uniform CIE Lab colour space and then pre-filtered with the GB-FAB anisotropic diffusion algorithm proposed by Ilea and Whelan in [85] that is applied to eliminate the image noise, weak textures and improve the local colour coherence (see Chapter 3). From the filtered CIE Lab converted image, the dominant colours (initial seeds) and the number of clusters (k) are automatically extracted. The initialisation step consists of an unsupervised classification procedure where the key element is the inclusion of a Self Organising Map (SOM) network. The filtered image is clustered using a K-Means algorithm where the cluster centres are initialised with the dominant colours and the number of clusters (k) calculated during the previous step. As illustrated in Figure 4.1, the second data stream of the colour segmentation algorithm analyses the input image converted to the YIQ colour space. The YIQ converted image is subjected to a pre-filtering procedure where the proposed GB-FAB algorithm is employed. The filtered YIQ image is clustered with a K-Means clustering algorithm where the cluster centres are initialised with the dominant YIQ colours obtained through a colour quantisation procedure and the number of clusters is set to k (that has been calculated during the previous automatic SOM classification step). The clustered CIE Lab and YIQ images are concatenated to generate an intermediate image that will be further subjected to a six-dimensional (6D) multi-space K-Means clustering that outputs the colour segmented image. In order to optimise the clustering parameters (the initial seeds and the number of clusters), an image colour saliency measure is applied to sample the contrast between the neighbouring regions. Based on this measure, the optimal set of parameters for the clustering algorithms involved in the multi-space segmentation scheme is obtained for the image that maximises the saliency measure.

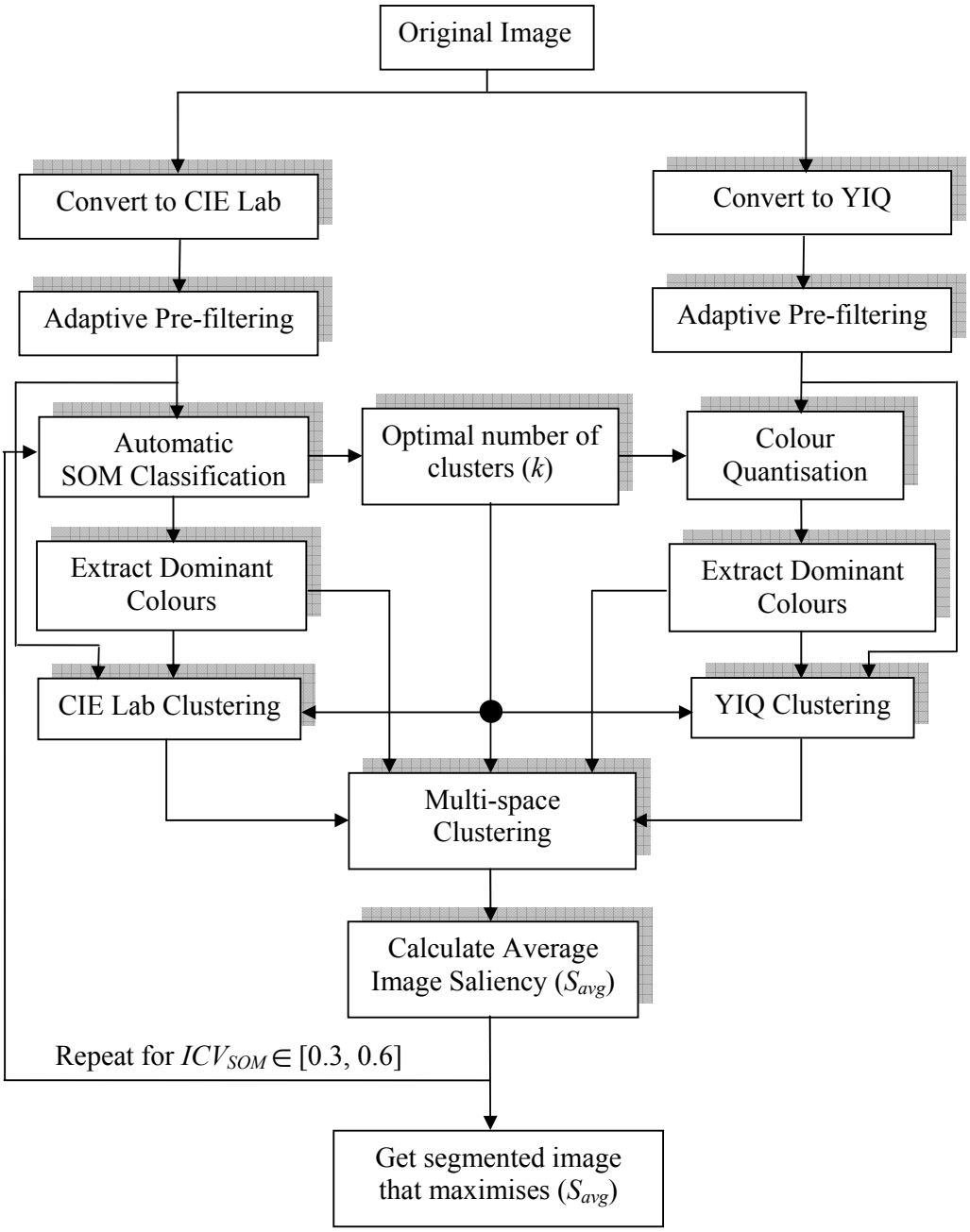


Figure 4.1 Overall computational scheme of the proposed multi-space colour segmentation algorithm.

4.2 Statistical Clustering for Colour Image Segmentation

Clustering algorithms have been widely applied to the segmentation of colour images due to their simplicity and low computational cost. These techniques perform a statistical analysis of the input data and have the purpose of reducing its size and complexity by constructing clusters based on the mutual similarity of image pixels. Therefore, the ideal approach is to use the statistical partitioning techniques to extract the colour information because the large number of colour components existing in the original image is reduced in the segmented image and only features that are strongly related to the image objects will be preserved.

K-Means [32] is a simple, non-hierarchical clustering analysis method that addresses the problem of identifying clusters of data points in a multi-dimensional space. The goal of K-Means is to iteratively partition the dataset into a number of clusters k, where the feature points are exchanged between clusters based on a pre-defined metric (typically the Euclidian distance) to satisfy the criteria of minimising the variation within each cluster and maximising the variation between the resulting k clusters. In its standard formulation, the K-Means algorithm consists of four main steps:

1. Initialisation – determine the number of clusters (k) and select the initial cluster centres from the input data represented in the N-dimensional feature space.
2. Generate data partitions by assigning each feature point to the nearest cluster centre by minimising the within cluster variation defined in equation (4.1).
3. Recalculate the new centres for clusters receiving new data points and for clusters losing data points (the new centres are computed as the mean of all members of the cluster under analysis).
4. Repeat steps 2 and 3 until no elements are exchanged between clusters.

As indicated earlier, the assignment of data points into clusters is achieved by minimising an objective function J that is defined as the sum of squared

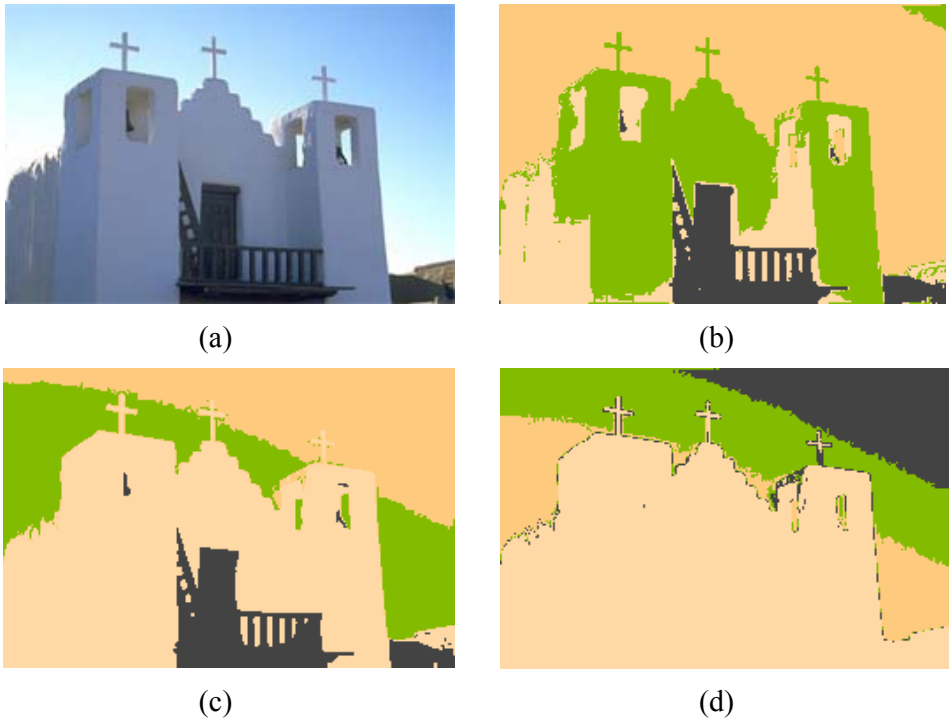
differences that samples the closeness between the data points and the cluster centres.

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i - c_j \right\|^2 \tag{4.1}$$

where n is the total number of pixels in the image, k is the number of clusters while $\left\| x_i - c_j \right\|$ is the distance (defined as the Euclidian metric) between the data point x_i and the cluster center c_j. It is important to mention that in equation (4.1) the data point x is typically defined as a vector x= [x_R, x_G, x_B] that describes the colour intensity of the pixel under analysis. This formulation of the clustering process is straightforward and requires a low computation cost, a fact that allows the algorithm to efficiently run on large datasets. However, in spite of its advantages it is useful to note that the performance of the K-Means clustering algorithm is highly dependent on two critical conditions:

- The improper initialisation of the cluster centres that will force the algorithm to converge to local minima and produce erroneous results. For most data partitioning algorithms, the cluster centres are initialised either using a starting condition specified *a-priori* by the user or by applying a random procedure that selects the cluster centres from input data. The random initialisation is often employed in the initialisation of K-Means clustering, but it proved to be an inappropriate solution as different results are obtained by selecting different initial centres and the algorithm will not be able to determine optimal partitioning. This is illustrated in Figure 4.2, when a randomly initialised K-Means algorithm has been applied three times in succession to a natural image. It can be observed that the three segmentations are dissimilar and only Figure 4.2 (c) is close to an accurate result. Hence the random initial selection is far from being a reliable solution.
- The difficulty in determining the optimal number of clusters k. This parameter is usually selected *a-priori* by the user, an approach that presents critical drawbacks since the performance of the algorithm is highly influenced by the appropriate selection of this parameter. This

can be observed in Figure 4.3 where k was varied in the interval [2, 9]. The results show that low values of k may lead to under-segmentation and important image information is lost, while large values better preserve the image details, but this is achieved at the expense of over-segmentation. Therefore different images will require different settings for the parameter k and to avoid erroneous partitions this parameter must be optimally determined.



(a) (b)

(c) (d)

Figure 4.2 Colour segmentation results when the K-Means clustering algorithm is applied three times (b-d) to the original image depicted in (a) using a random initialisation procedure. The number of clusters k is manually set to 4. It can be observed that the algorithm produces different segmentations every time it is executed.

The statistical space-partitioning scheme proposed in this dissertation automatically addresses the problems associated with standard clustering algorithms, with the main focus being on the description of a novel technique based on the Self Organising Maps (SOM) that is applied to identify the optimal set of clustering parameters.

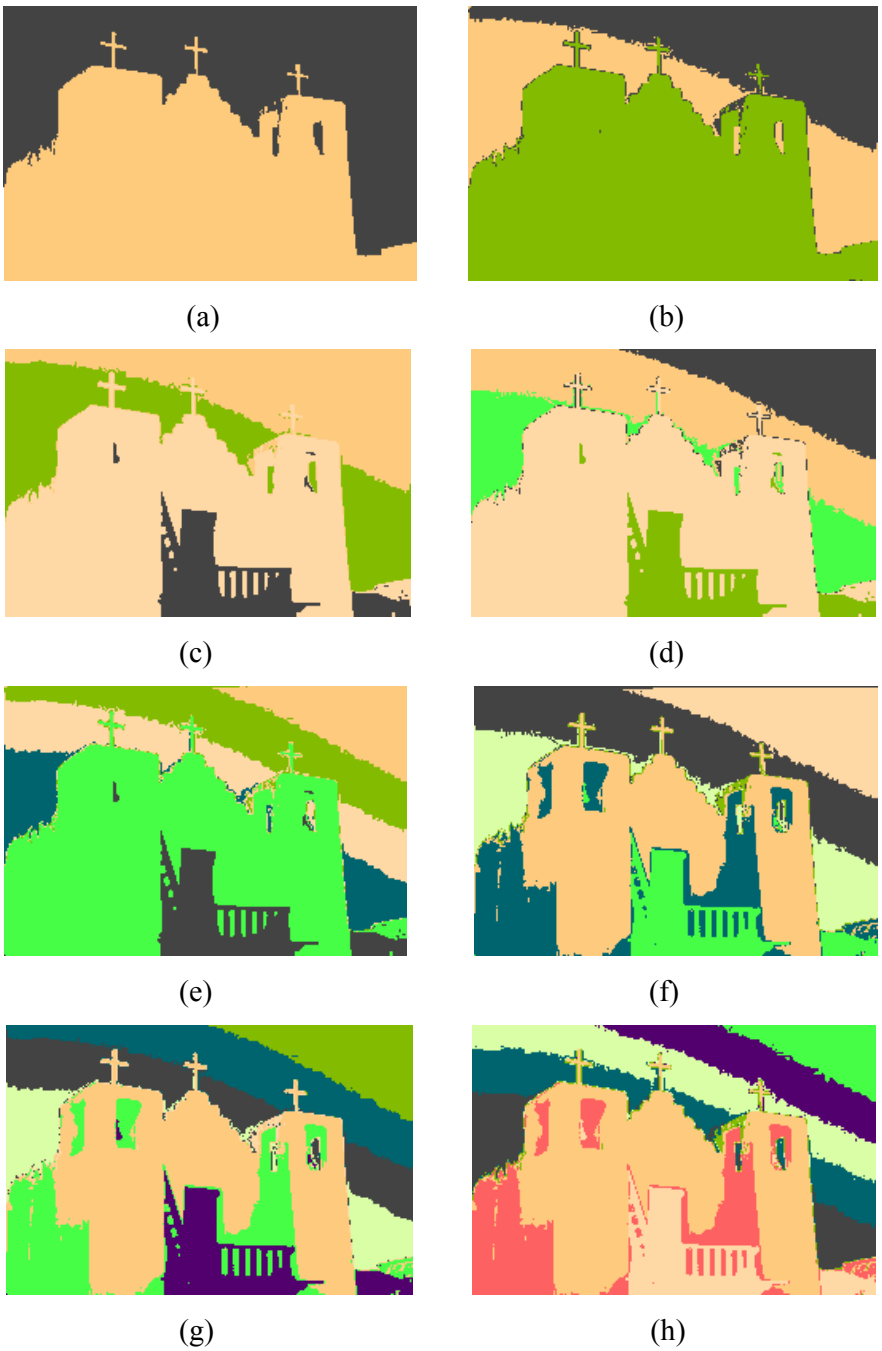


Figure 4.3 Colour segmentation results when the K-Means clustering algorithm is applied to the image depicted in Figure 4.2 (a). The algorithm is initialised using randomly selected values from the input image and the number of clusters k is manually set to the following values: (a) $k = 2$. (b) $k = 3$. (c) $k = 4$. (d) $k = 5$. (e) $k = 6$. (f) $k = 7$. (g) $k = 8$. (h) $k = 9$. For visualisation purposes, the images are shown in pseudo colours.

4.3 Initialisation of the Cluster Centres

As indicated in Figure 4.2, an accurate selection of initial cluster centres prevents the clustering algorithm to converge to local minima, hence producing erroneous decisions. As it has been discussed in the previous section, the most common initialisation procedure selects the initial cluster centres randomly from the input data [86, 32]. This procedure is far from optimal because it does not prevent the initialisation of the clustering algorithm on outliers and in addition to this the segmentation results will be different any time the algorithm is executed on the same data. To improve the random initialisation procedure, some authors applied the clustering algorithms in a nested sequence [87, 88], but the experiments indicated that this solution is not any better than the random initialisation procedure, as an improper initial selection of the cluster centres will propagate towards the final result and produce non-optimal image partitions. The selection of optimal seeds that are used to initialise the clustering algorithms is not a new topic and it is useful to note that a number of techniques are documented in the computer vision literature. Pena et al [89] carried out an empirical evaluation of four early initialisation methods for K-Means clustering and analysed the random, Forgy [86], MacQueen [32] and Kaufman [90] approaches. Their experiments indicated that the random and Kaufman methods outperformed the other two with respect to the effectiveness and robustness of the clustering process. At present, the research trend assumes that the cluster centres can be found in advance, and some recent initialisation approaches include a deterministic method applied for the initialisation of the K-Means algorithm based on the Principal Component Analysis [91], the Cluster Centres Initialisation Algorithm (CCIA) proposed by Khan and Ahmad [92] and the method developed by Kim and Lee [93] that is applied to extract the most vivid and distinguishable colours as the initial cluster centres. Most of these techniques have been developed to solve the initialisation problem for normalised data with high dimensionality and they are not directly applicable to colour segmentation problems. It is also useful to mention that for most of the initialisation schemes proposed to date, the number of clusters parameter has to be manually selected in advance. Therefore, a generic method for the

automatic detection of both the optimal number of clusters and the initial colour seeds has not been fully developed so far in the context of colour image segmentation.

4.3.1 Dominant Colours Extraction. Automatic Detection of the Cluster Centres

4.3.1.1 Dominant Colours Extraction Using the SOM Initialisation Procedure

The performance of the clustering algorithms is highly influenced by the selection of the initial colour seeds and the number of clusters *k*. In this thesis an efficient solution to automatically detect the dominant colours and the optimal number of clusters in the image using a classification procedure based on the Self Organising Maps (SOM) is proposed. A set of input vectors is trained using a SOM network in order to obtain a lower dimensional representation of the input image in the form of a feature map that maintains the topological relationship and metric within the training set. The SOM networks were first introduced by Kohonen [94] and became popular due to their ability to learn the classification of a training set without any external supervision. In this implementation, a two-dimensional (2D) SOM network that is composed of nodes or cells (see Figure 4.4 (a)) has been created. Each node N_i ($i \in [1, M]$, where M is the number of nodes in network) has assigned a 3D weight vector (w_i) that matches the size of each element of the input vector. It is useful to mention that the training dataset represented by the input image is organised as a 1D vector V_j (j=1…n, where n is the total number of pixels in the image) in a raster scan manner. Each element of the training set V_j is defined as a 3D colour vector whose components are the normalised values of the pixel in the image. As illustrated in Figure 4.4 (a), each element V_j of the input data is connected to all nodes in the 2D SOM network.

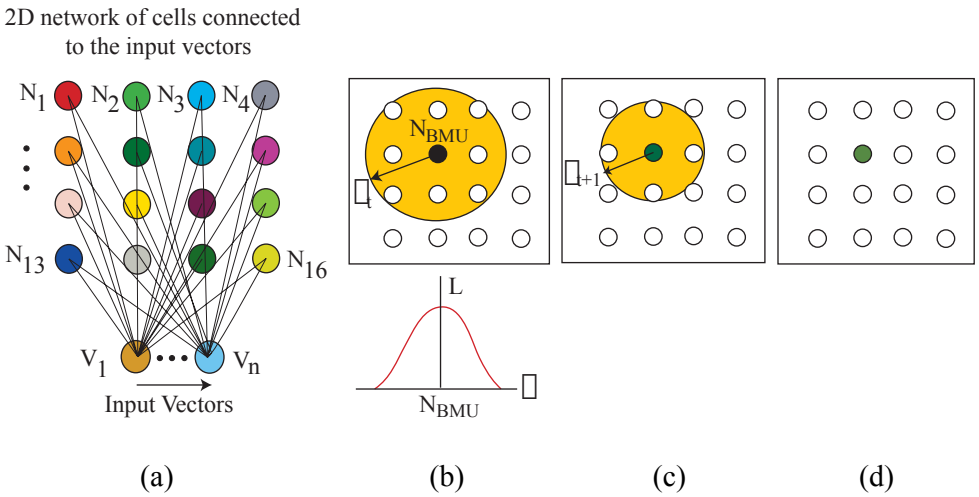


Figure 4.4 (a) A 2D SOM network. (b) The neighbourhood of N_{BMU} at iteration t. The learning process of each cell's weight follows a Gaussian function, i.e. it is stronger for cells near node N_{BMU} and weaker for distant cells. (c, d) The radius $\nu(t)$ is progressively reduced until it reaches the size of one cell (N_{BMU}).

In line with other clustering schemes, before starting the training procedure we need to initialise the weights w_i belonging to all cells in the network. In practice, the random initialisation is usually adopted when working with SOM networks [95, 96], an approach motivated by the fact that after several hundreds of iterations the corresponding values of the initial random weights will change in accordance to the colour content of the image. This procedure has been applied in [97] where the authors initialise the SOM network by randomly picking colour samples from the input image. But the random selection of the starting condition is sub-optimal since the algorithm can be initialised on outliers. Therefore, in this thesis the weights of the nodes in the SOM network are initialised with the dominant colours that are represented by the peaks (P_i) of the 3D colour histogram that is calculated from the image that has been subjected to colour quantisation [36]. This is achieved by applying a linear colour quantisation procedure that consists of linearly re-sampling the number of colours on each colour axis. It has been experimentally demonstrated in [36, 98] that a quantisation level of 8 is sufficient to sample the statistical relevant peaks in the 3D histogram. Thus, the quantised version

of the input image is re-mapped so that the initial number of grey-levels in all colour bands (256×256×256) is now reduced to 8×8×8. After constructing the 3D histogram in the quantised colour space, the peaks P_i related to the desired number of dominant colours are selected by applying a quicksort algorithm.

$$P_i = \arg\max_{i \in [1,M]} \left(Colour\_Histogram \right) \tag{4.2}$$

In equation (4.2) M denotes the number of cells in the network, while P_i is the histogram peak that initialises the weight w_i of the SOM network. The implementation of the colour quantisation-based initialisation procedure is detailed in the following pseudo-code sequence:

```
int ColourHistogram[noColours][noColours][noColours];
int levels[noColours];

//initialise the colour histogram, compute the quantisation levels
InitialiseHistogram(ColourHistogram);
ComputeQuantisationLevels(levels, noColours);


for (int i=0; i<Image.RowNo(); i++)
{
  for (int j=0; j<Image.ColNo(); j++)
  {
      int L = (int)Image.Plane[0].Val[i][j];
      int a = (int)Image.Plane[1].Val[i][j];
      int b = (int)Image.Plane[2].Val[i][j];

      //compute the quantised value for each colour component
      int quant_L = ExtractQuantisedValue(levels,noColours, L);
      int quant_a = ExtractQuantisedValue(levels,noColours, a);
      int quant_b = ExtractQuantisedValue(levels,noColours, b);

      ColourHistogram [quant_L][quant_a][quant_b]++;
  }
}
int ColourPeaks[noClusters];
//apply quicksort to extract the dominant colours
ExtractDominantColours (ColourPeaks, ColourHistogram);
```

Taking into consideration that the size of the SOM lattice is four by four (*i.e.* M=16 cells), the first 16 highest histogram peaks are sufficient to accurately sample the dominant colours in the image. Once the initialisation is completed ($w_i \leftarrow P_i, i \in [1,16]$), the classification procedure is iteratively applied and consists of assigning the input vectors to the cell in the network whose corresponding weight values are most similar. The node in the SOM network that returns the smallest Euclidean distance is declared the BMU (Best Matching Unit):

$$BMU = \arg\min_{i \in [1,16]} \left\| V_j - w_i \right\|, \quad j \in [1, n] \tag{4.3}$$

The weights for all cells that are inside the circular neighbourhood of the BMU must be updated. The radius of the neighbourhood (ν) has a large initial value (ν_0) and is set to cover all nodes in the lattice, but with the increase in the number of iterations it suffers an exponential decay controlled by the following expression:

$$\nu(t) = \nu_0 \cdot e^{-\frac{t}{\lambda}} \tag{4.4}$$

where t is the iteration step and λ is a time constant that is usually set as the ratio between the maximum number of iterations and the initial radius of the neighbourhood ν_0. As indicated earlier, with the increase in the number of iterations, the radius ν is progressively reduced until it reaches the size of the BMU cell. This process is repeated for all vectors V_j in the training set and it is graphically illustrated in Fig. 4.4 (b-d).

The most important step of the SOM algorithm is represented by the process of adjusting the weights of the cells situated in the BMU's neighbourhood that must satisfy the following condition: $\left\| N_{BMU} - N_i \right\| \leq \nu(t)$, where ||.|| defines the Euclidean distance between the cell N_i and the BMU.

The weights of the cells are adjusted in order to minimise the distance to the input vector that has generated the BMU, using the following learning rule:

$$\begin{aligned} w_i(t+1) &= w_i(t) + \psi(t) \cdot L(t) \cdot [V_j(t) - w_i(t)], \quad \text{if } \|N_{BMU} - N_i\| \le \nu(t) \\ w_i(t+1) &= w_i(t), \qquad\qquad\qquad\qquad\quad \text{if } \|N_{BMU} - N_i\| > \nu(t) \end{aligned} \tag{4.5}$$

Equation (4.5) indicates that if the cell is situated inside the BMU's neighbourhood its weights are updated, otherwise the weights are left unchanged. The process that updates the weights of the cells situated inside the BMU's neighbourhood is controlled by two functions, $\psi(t)$ and $L(t)$. The function $L(t)$ represents the learning rate and its influence is reduced with the increase in the number of iterations t, as follows:

$$L(t) = L_0 \cdot e^{-\frac{t}{\lambda}} \tag{4.6}$$

where L_0 is a small constant value that is set in conjunction with the size of the SOM network. The function ψ is implemented using a Gaussian function and controls the strength of the updating process. Thus, the weights of the cells situated closer to the BMU are more aggressively updated while for cells situated far from BMU the learning process is less pronounced.

$$\psi(t) = e^{-\frac{\|N_{BMU} - N_i\|^2}{2\nu^2(t)}} \tag{4.7}$$

The SOM algorithm is iterated until convergence (radius ν reaches the size of N_{BMU}) and the final weights of the 2D network are the initial dominant colours of the input image. Figure 4.5 (column b) illustrates the sixteen initial dominant colours (displayed as a bar graph) resulting after the SOM classification procedure for a set of natural images contained in the Berkeley database [99]. These initial colours will be further reduced until they reach their optimal value (column c) using a multi-step technique that will be described in the following section.

4.3.1.2 Selection of the Optimal Number of Clusters

To obtain the optimal number of clusters, in this thesis a multi-step technique that progressively reduces the number of dominant colours resulting after the SOM classification procedure is proposed. In the first step, the pixels in the image are mapped to the final weights of the cells in the SOM network based on the minimum Euclidean distance (see equation (4.8)).

$$V_j \leftarrow w_g,\ g = \arg\min_{i\in[1,16]} \left\| V_j - w_i \right\|,\quad j \in [1,n] \tag{4.8}$$

The resulting colour map can be viewed as a preliminary clustering of the input image. In the second step a confidence map is constructed, where the cumulative smallest distances between the weights of the SOM network and pixels in the image are recorded (see equation (4.9)). For all pixels V_j labelled with w_i, a confidence value is defined as follows:

$$confidence(w_i) = \frac{\sum_{j\in D_{w_i}} \left\| V_j - w_i \right\|^2}{no\_pixels\_labelled(w_i)},\ i \in [1,16] \tag{4.9}$$

where D_{w_i} is the image domain defined by the pixels V_j that are closest to the weights w_i. The confidence map returns a weighted measure between the variance within the cluster and the number of pixels in the cluster. The lower its value is, the more reliable the estimate w_i is. The confidence map calculated for the example depicted in Figure 4.5 (a1) is shown in Table 4.1.

The last step determines the final number of clusters by evaluating the inter-cluster variability. To achieve this, the similarity matrix is constructed where the Euclidean distances between the weights of any neighbouring nodes in the SOM network are stored. If this distance is smaller than a pre-defined inter-cluster threshold ICV_{SOM}, then the node that has the highest confidence value is eliminated. This process is iteratively applied until the distance between the weights of all adjacent nodes in the SOM network is higher than the pre-defined threshold value ICV_{SOM}.

Figure 4.5 (Column a) Original natural images from Berkeley database [99]. (Column b) The 16 dominant colours resulting after the application of the SOM classification procedure described in Section 4.3.1.1. (Column c) The optimal dominant colours determined after the application of the cluster optimisation procedure described in Section 4.3.1.2. For these tests, the ICV_{SOM} is set to 0.3. The final number of clusters k calculated for each image are: (a1) $k = 5$. (a2) $k = 9$. (a3) $k = 7$. (a4) $k = 6$.

TABLE 4.1

Confidence map corresponding to the 16 dominant colours depicted in Figure 4.5 (b1).

Seed	*Confidence value*	*No. of samples*
c_1	0.075193	14683
c_2	0.073174	15849
c_3	0.133122	5688
c_4	0.170826	1118
c_5	0.069306	15044
c_6	0.075331	15165
c_7	0.095346	17003
c_8	0.206337	1209
c_9	0.066002	14819
c_{10}	0.109216	15633
c_{11}	0.078680	16581
c_{12}	0.162184	17403
c_{13}	0.065022	13282
c_{14}	0.186783	18886
c_{15}	0.116970	12933
c_{16}	0.092707	15756

In Ilea and Whelan [100] the inter-cluster variability threshold is set to 0.3, a value that proved to be appropriate for all images analysed in our study. But in order to eliminate any supervision, in Section 4.4.1 a colour saliency measure that samples the contrast between the neighbouring regions of the clustered image is applied for the automatic optimisation of the ICV_{SOM} parameter. In the next section of this chapter the multi-space architecture adopted for colour feature extraction is discussed.

4.4 The Multi-Space Colour Image Segmentation Architecture

The selection of the proper colour spaces is of high importance since they provide data representations where the clustering algorithms will be applied and where the initialisation conditions are deduced from. Usually, the original image is represented in the RGB colour space, but RGB is a perceptually non-uniform colour space and one of its limitations is the fact that the chrominance and intensity components are not explicitly defined. To compensate for its limitations, many researchers converted the original RGB image to different colour representations (the most used being CIE Lab, YIQ and HSI) and their selection proved to be application dependent. Each of these colour spaces has distinctive characteristics in representing the colour information in the image. Therefore, in this thesis a multi-space colour image segmentation architecture is proposed that takes advantage of the complementary character of pairs of colour spaces. As it can be observed in Figure 4.1, the segmentation algorithm analyses the colour information on two different streams of data. In the first stream, the original image is converted to a colour space and afterwards is processed, while in the second stream the original image is converted to a complementary colour representation. The colour spaces analysed within the framework described in this thesis are: the original image colour space representation (RGB), a perceptually equalised colour space (CIE Lab) and a colour space (YIQ) that encodes the original image information into two separate components: luminance and chrominance. A detailed discussion on different colour image representations can be found in [13, 101].

As mentioned before, the idea behind the multi-space architectural design was adopted in order to take advantage of the property of complementary colour spaces in representing the colours in the input image. Colours with high degrees of similarity in the RGB space may be difficult to distinguish, while the YIQ and CIE Lab representations may provide a stronger discrimination. In addition, the shadows and local inhomogeneities are generally better modelled in YIQ and CIE Lab than in the RGB colour space.

The reason behind the proposed multi-space colour segmentation approach is that the use of complementary colours spaces is highly advantageous for the accuracy of colour feature extraction.

In this thesis, a set of quantitative experiments were conducted in order to determine the pair of colour spaces that returns the best segmentation when used within the multi-space colour segmentation scheme. A detailed description of these experiments is provided in Section 4.5.1. The numerical results indicate that in the context of colour segmentation, the best results are obtained using the CIE Lab and YIQ colour spaces. This conclusion is justified because CIE Lab and YIQ are colour spaces with complementary characteristics, the first being a perceptually uniform colour space, while the second has the advantage of separating the luminance from the chrominance components.

In the remainder of this section the multi-space architectural design will be discussed. As illustrated in Figure 4.1 the first data channel starts with the conversion of the original image into the CIE Lab colour space that will be further subjected to a filtering procedure. From the filtered image, the dominant colours and the optimal number of clusters (k) are extracted using the SOM based procedure described in section 4.3.1. These parameters are used to initialise the K-Means clustering algorithm that is applied to the filtered CIE Lab converted image.

The YIQ converted image goes through similar operations as the previously image that is analysed in the CIE Lab colour representation. Initially it is filtered and then it is further processed using a K-Means clustering algorithm. The key issue associated with the extraction of the colour features from the YIQ image is the fact that the parameter k that selects the number of clusters for the K-Means algorithm is set to the same value that has been obtained after the application of the SOM procedure to the image represented in the CIE Lab colour space (see Section 4.3.1). Thus, the parameter k performs the synchronisation between the CIE Lab and YIQ data channels by forcing the K-Means algorithms applied to the CIE Lab and YIQ images (see Figure 4.1) to return the same number of clusters k. The dominant colours from the YIQ image that are used to initialise the clusters centres for the K-Means algorithm

are determined using a colour quantisation procedure, the same as the one applied to initialise the weights of the SOM network and which has been detailed in section 4.3.1.1. The dominant colours are the highest k peaks of the three-dimensional (3D) image histogram and will be used to initialise the K-Means clustering algorithm that is applied to the filtered YIQ image during the analysis of the second data stream.

The next step of the colour extraction algorithm proposed in this thesis consists in the concatenation of the colour features calculated from the CIE Lab and YIQ clustered images. Each pixel in the image is defined by a six-dimensional (6D) vector whose components are the pixel's values of the clustered CIE Lab and YIQ images. In the final step, the CIE Lab-YIQ data is clustered with a 6D K-Means algorithm where the number of clusters is again set to k and the cluster centres are initialised with the dominant colours that were used to initialise the two K-Means algorithms that have been applied to cluster the CIE Lab and YIQ images.

As opposed to the approach that analyses the CIE Lab and YIQ images on parallel data streams, an alternative implementation would directly cluster the 6D CIE Lab-YIQ data. The approach proposed in this thesis is motivated by the fact that the initialisation of the 6D SOM network using the procedure based on colour quantisation is unreliable since the 6D histogram calculated from the concatenated CIE Lab-YIQ data is sparse and the peaks are not statistically relevant. The adopted approach circumvents this issue while it attempts to find the optimal result by fusing the clustered CIE Lab and YIQ images which have a reduced dimensionality that is sampled by the parameter k (as opposed to the high dimensionality of the original CIE Lab-YIQ data).

The multi-space clustering algorithm generates the final colour segmented image. Figures 4.6 to 4.8 illustrate the performance of the developed multi-space colour segmentation algorithm when compared to the results obtained when the input image is analysed in the CIE Lab and YIQ colour representations. For these tests, the ICV_{SOM} parameter was automatically detected using the optimisation technique that will be described in the following section. Additional experimental results are shown in Figure 4.9.

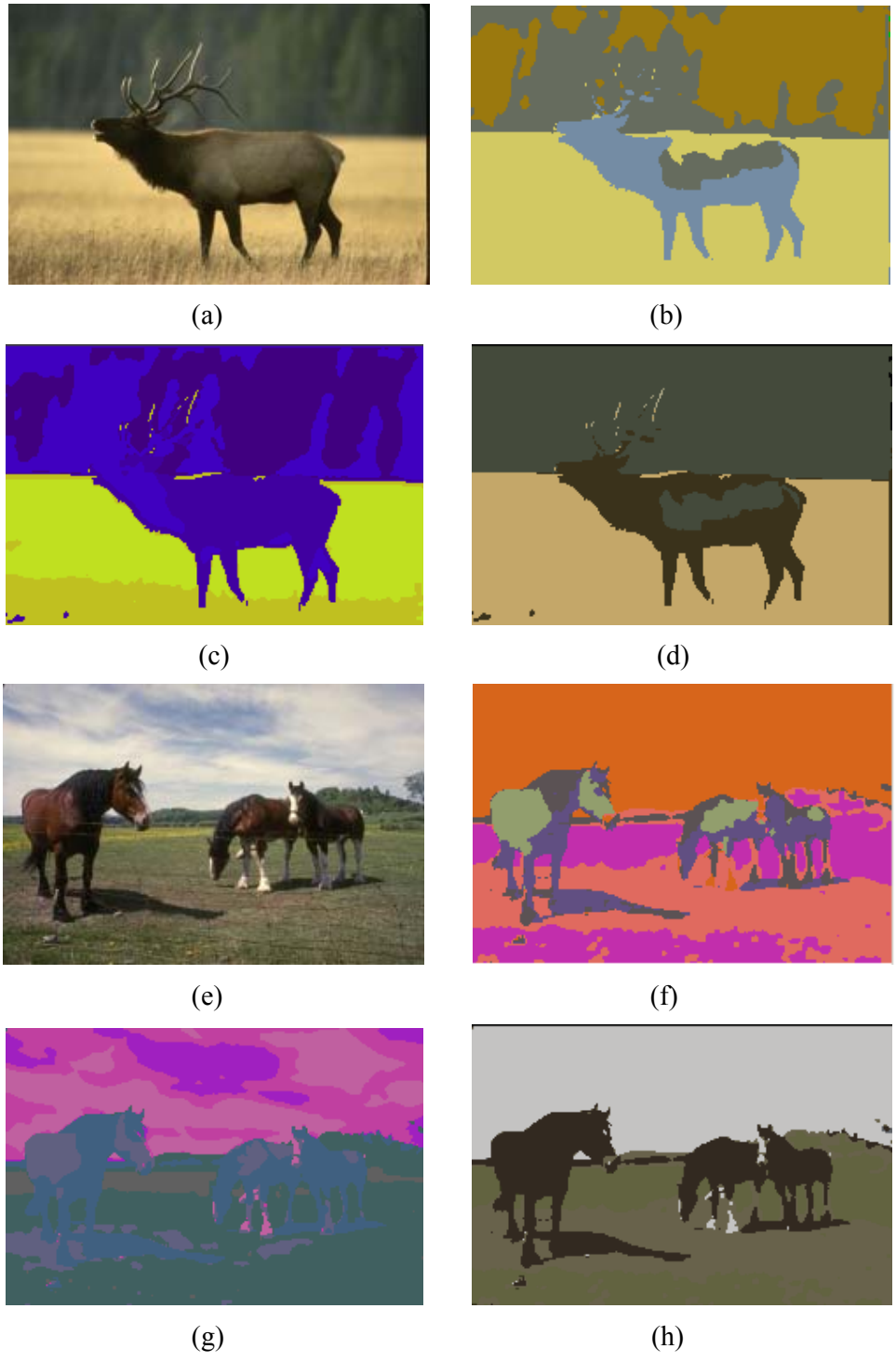


Figure 4.6 (a, e) Original images [99]. (b, f) The clustered image in the CIE Lab colour space. (c, g) The clustered image in the YIQ colour space. (d) The final multi-space colour segmentation result (final number of clusters is 5). (h) The final multi-space colour segmentation result (final number of clusters is 7).

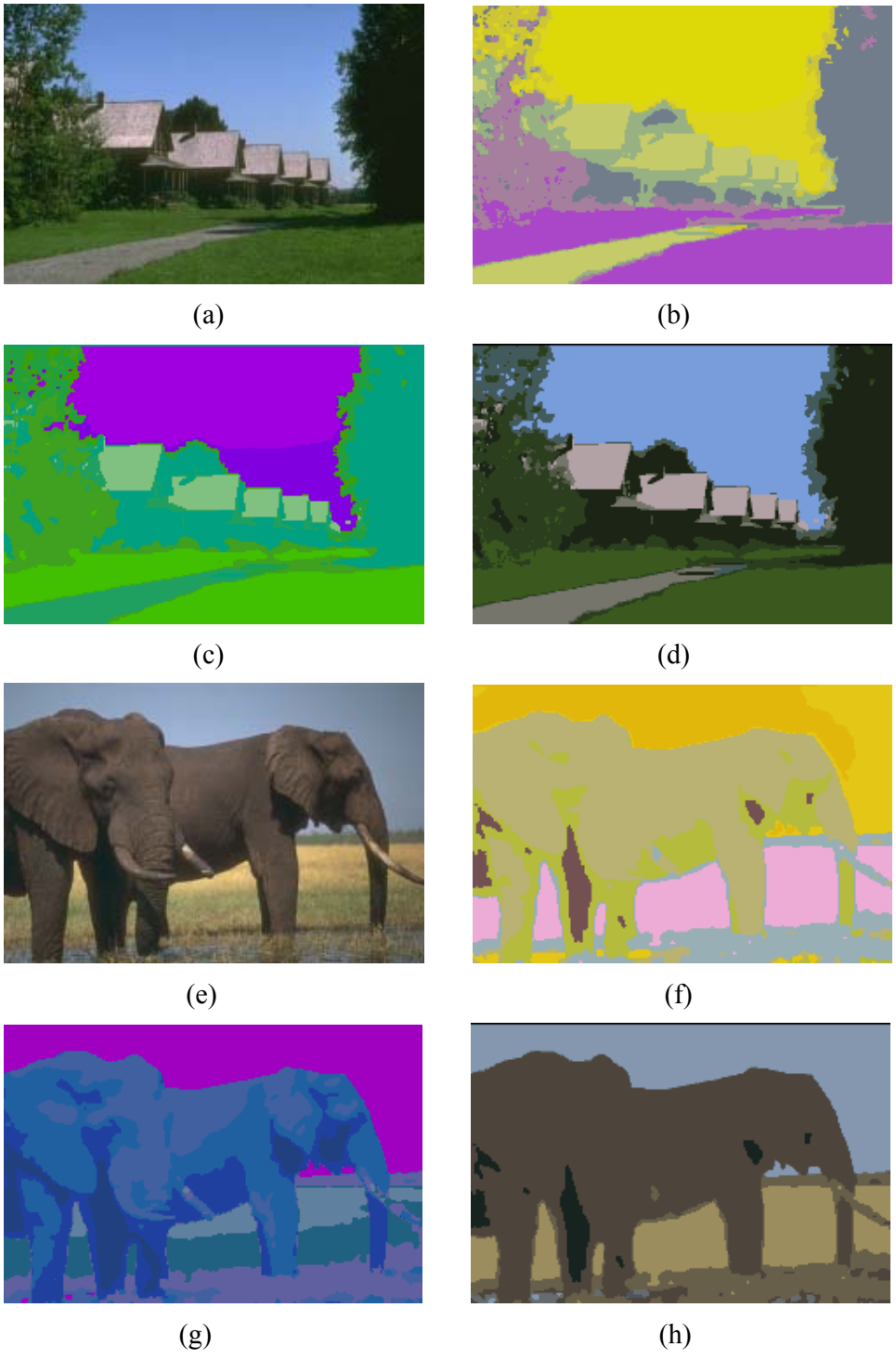


Figure 4.7 (a, e) Original images [99]. (b, f) The clustered image in the CIE Lab colour space. (c, g) The clustered image in the YIQ colour space. (d) The final multi-space colour segmentation result (final number of clusters is 9). (h) The final multi-space colour segmentation result (final number of clusters is 8).

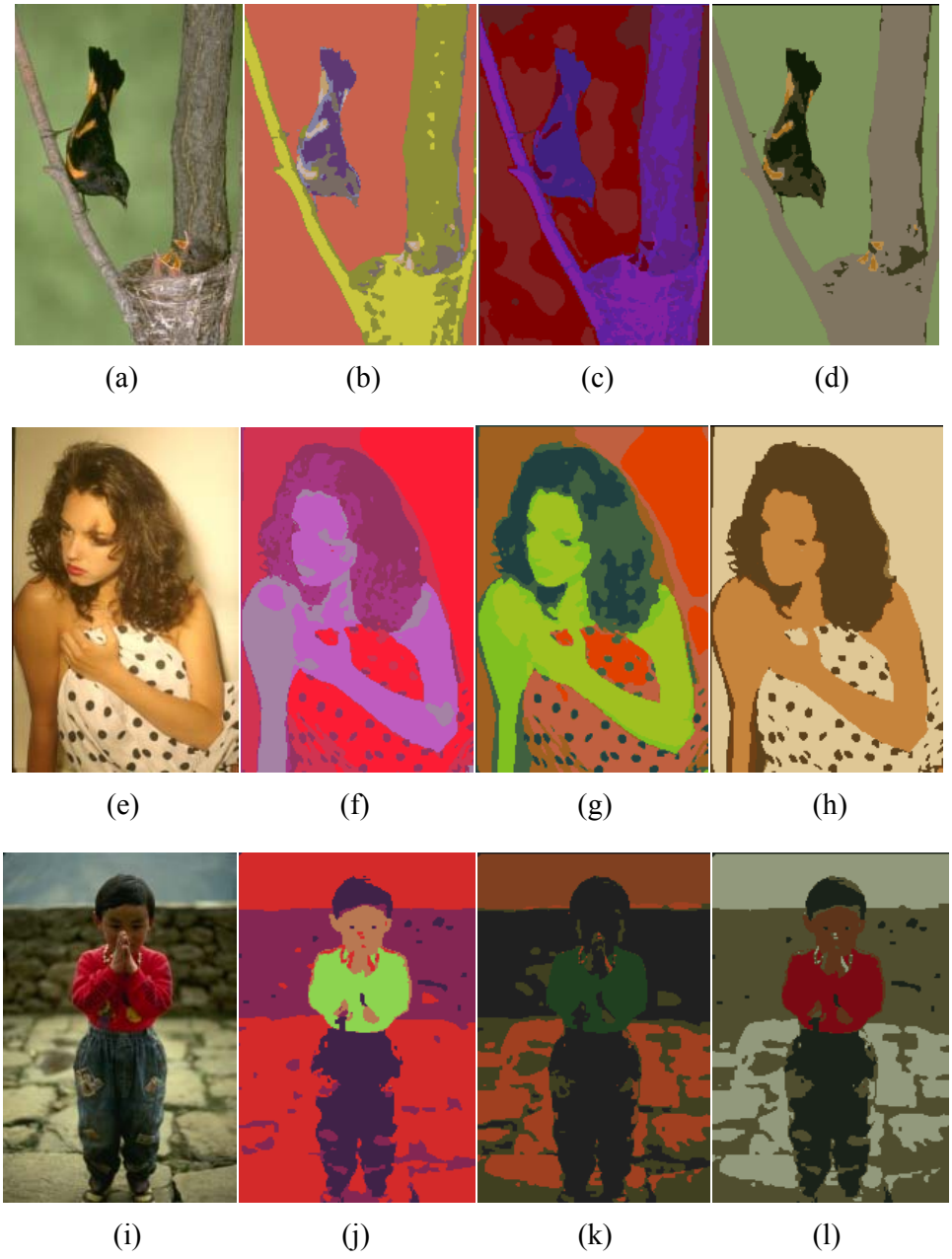


Figure 4.8 (a, e, i) Original images [99]. (b, f, j) The clustered image in the CIE Lab colour space. (c, g, k) The clustered image in the YIQ colour space. (d) The final multi-space colour segmentation result (final number of clusters is 9). (h) The final multi-space colour segmentation result (final number of clusters is 7). (l) The final multi-space colour segmentation result (final number of clusters is 5).

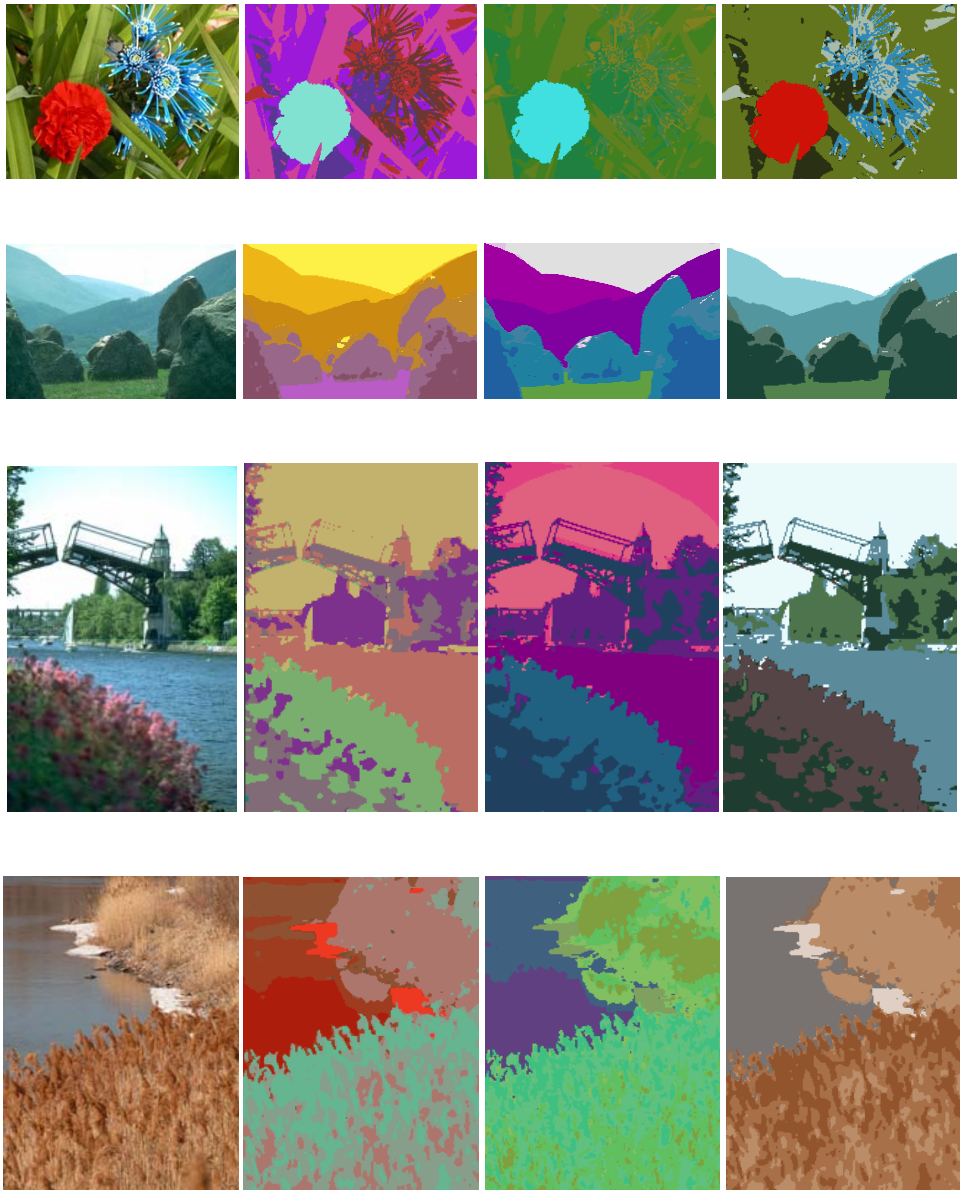


Figure 4.9 Performance of the developed multi-space colour segmentation algorithm (4th column) when compared to the results obtained when the input image is analysed in CIE Lab (2nd column) and YIQ (3rd column) colour representations. The natural images (1st column) are from Berkeley [99] and McGill [102] databases and exhibit complex colour-texture characteristics.

4.4.1 Automatic Optimisation of the ICV_{SOM} Parameter

In section 4.3.1.2 the ICV_{SOM} parameter was introduced for the detection of the optimal parameters used to initialise the clustering algorithms. After the first step of the automatic SOM initialisation process (described in section 4.3.1.1) a number of 16 dominant colours strongly related to the colour content in the image are obtained. During the second step of the initialisation process, the 16 colour seeds are progressively reduced in conjunction with an inter-cluster variability threshold (ICV_{SOM}) as discussed in section 4.3.1.2. In other words, the dominant colours resulting from the SOM map are processed using a confidence measure where the Euclidean distance between any two dominant colours returned by the SOM network is higher than a pre-defined ICV_{SOM} threshold. Using this procedure, the colour seeds and the optimal number of clusters (k) used to initialise the K-Means algorithm are determined.

It has been quantitatively demonstrated in [103] that the maximisation of the border contrast (this implies the maximisation of the average saliency S_{avg} defined in sub-section 4.4.1.1) leads to improved region stability. A high region stability in the context of image segmentation means that the resulting regions in the segmented data are not erroneously divided due to spurious texture or uneven illumination. Based on this observation, the optimal set of parameters is obtained for the clustered image where the average saliency measure is maximised.

In order to eliminate the supervision that involves the manual selection of the ICV_{SOM} parameter, an approach that automatically determines this parameter with respect to the maximisation of a region saliency measure that samples the contrast between the neighbouring regions in the clustered image is proposed. In this regard, clustered results for differing values of the ICV_{SOM} parameter are generated and the optimal result is selected as the clustered image that maximises the average saliency measure (S_{avg}).

This optimisation approach is graphically illustrated in Figure 4.10. The computational steps in Figure 4.10 are repeated for different values of the $ICV_{SOM} \in [0.3,0.6]$ parameter and the image with the highest S_{avg} returns the best segmentation. This range of values have been established based on

experimentation since ICV_{SOM} values lower than 0.3 consistently lead to over-segmentation, while values larger than 0.6 produce under-segmented results. Hence, the calculation of the saliency measure for parameters outside the range [0.3, 0.6] will just increase the computational cost of the algorithm without any increase in the performance of the colour segmentation algorithm.

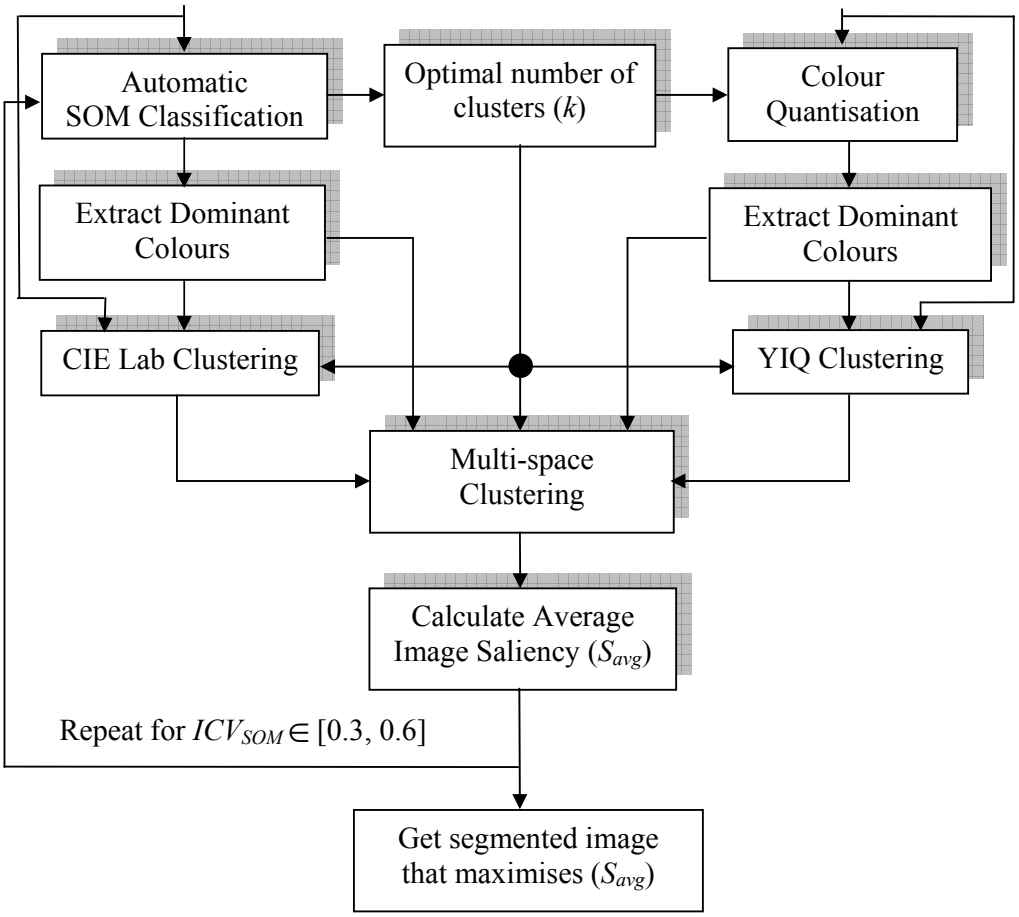


Figure 4.10 Automatic calculation of the ICV_{SOM} parameter. This diagram is part of the overall computational scheme of the colour segmentation algorithm depicted in Figure 4.1. In order to automatically determine the optimal value of the ICV_{SOM} parameter, all computational steps displayed in this diagram are iteratively applied for different values of ICV_{SOM} and the average saliency S_{avg} of the final clustered image is calculated. The image that returns the maximum S_{avg} corresponds to the optimal segmentation.

The procedure illustrated in Figure 4.10 can be summarised in the following sequence:

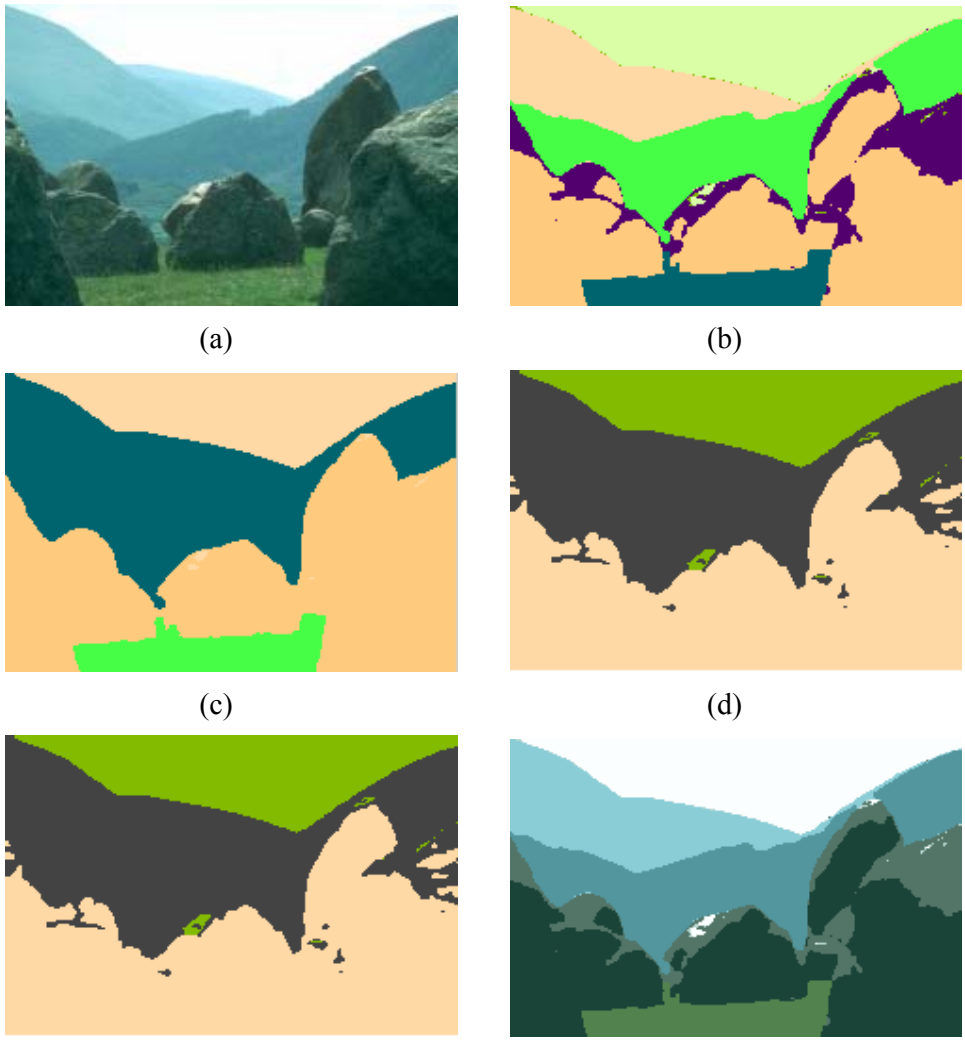
1. Apply the SOM classification to the CIE Lab filtered image. Calculate the 16 colour seeds.
2. for (ICV_{SOM} = 0.3; ICV_{SOM} <=0.6; ICV_{SOM}+0.1)
   {
   a) Reduce the number of colour seeds in agreement to the ICV_{SOM} and calculate the number of clusters k;
   b) Initialise the cluster centres and the number of clusters k for the K-Means scheme with parameters resulting from step 2a);
   c) Perform clustering on the CIE Lab image;
   d) Perform clustering on the YIQ converted image;
   e) Apply the multi-space clustering algorithm;
   f) Compute the average colour saliency S_{avg} resulting after the multi-space clustering algorithm.
   }
3. Determine the clustered image that returns the maximum average saliency (S_{avg}).

4.4.1.1 Definition of the Colour Saliency Measure

The colour saliency (*CS*) is calculated for every cluster R_i, ($i \in [1,k]$ where k is the number of clusters) of the segmented image and is defined as the average colour difference with respect to the 4–connected neighbouring (*N4*) regions [103]:

$$CS(R_i) = \frac{1}{NBP(R_i)} \sum_{(x,y)\in BP(R_i)} \frac{1}{N_{diff\_N4}(x,y)} \times \sum_{\{R_j(p,q)\mid(p,q)\in N4(x,y)\}} \left\| \overline{Col}(R_i) - \overline{Col}(R_j) \right\| \tag{4.10}$$

In equation (4.10), $NBP(R_i)$ is the total number of border pixels of the region R_i, $BP(R_i)$ is the set of pixels that defines the boundary of region R_i and $N_{diff\_N4}(x, y)$ represents the number of pixels in the 4-connected neighbourhood of the current pixel (x, y) that belong to a region different than R_i. In equation (4.10), $\overline{Col}(R_i)$ is the 3D vector describing the colour of the region R_i whose value is calculated as the mean of all pixels contained in the region R_i and (p,q) are the neighbouring pixels that belong to the regions R_j that are adjacent to the pixel (x, y). It can be noted that the saliency of the region R_i is calculated over all boundary pixels (x, y) and is defined as the sum of absolute differences between the pixels in the 4-connected neighbourhood that belong to the region R_i and those that belong to different regions. The sum of absolute differences depicted in equation (4.10) is weighted by $N_{diff\_N4}(x, y)$ to avoid the local maxima generated by the regions that are defined by single pixels. The average image saliency S_{avg} is calculated as the mean of the saliency values obtained for all regions in the image R_i ($i = 1 \ldots k$). The step-by-step execution of the developed parameter selection scheme is depicted in Figures 4.11 to 4.16 and it can be observed that the optimal clustering parameters vary significantly from image to image based on the colour image content.



(a) (b) (c) (d) (e) (f)

Figure 4.11 Colour image segmentation with parameter optimisation. The cluster centres of the K-Means algorithm are automatically initialised using the colour seeds resulting from the SOM classification and the number of clusters k is calculated in agreement with an inter-cluster variability parameter (ICV_{SOM}). (a) Natural image [99]. (b) Clustered image 1 (for visualisation purposes the clustered images are shown in pseudo colours), ICV_{SOM} = 0.3, S_{avg}= 2.33, number of clusters k = 8; (c) Clustered image 2, ICV_{SOM} = 0.4, S_{avg} = 1.10, k= 6; (d) Clustered image 3, ICV_{SOM} = 0.5, S_{avg} = 0.77, k = 4; (e) Clustered image 4, ICV_{SOM} = 0.6, S_{avg} = 0.77, k = 4. Image (b) generates the maximum S_{avg} = 2.33 and the number of clusters k = 8. (f) The final colour segmented image where the pseudo colours were replaced with the mean values of the corresponding colour values from the original image.

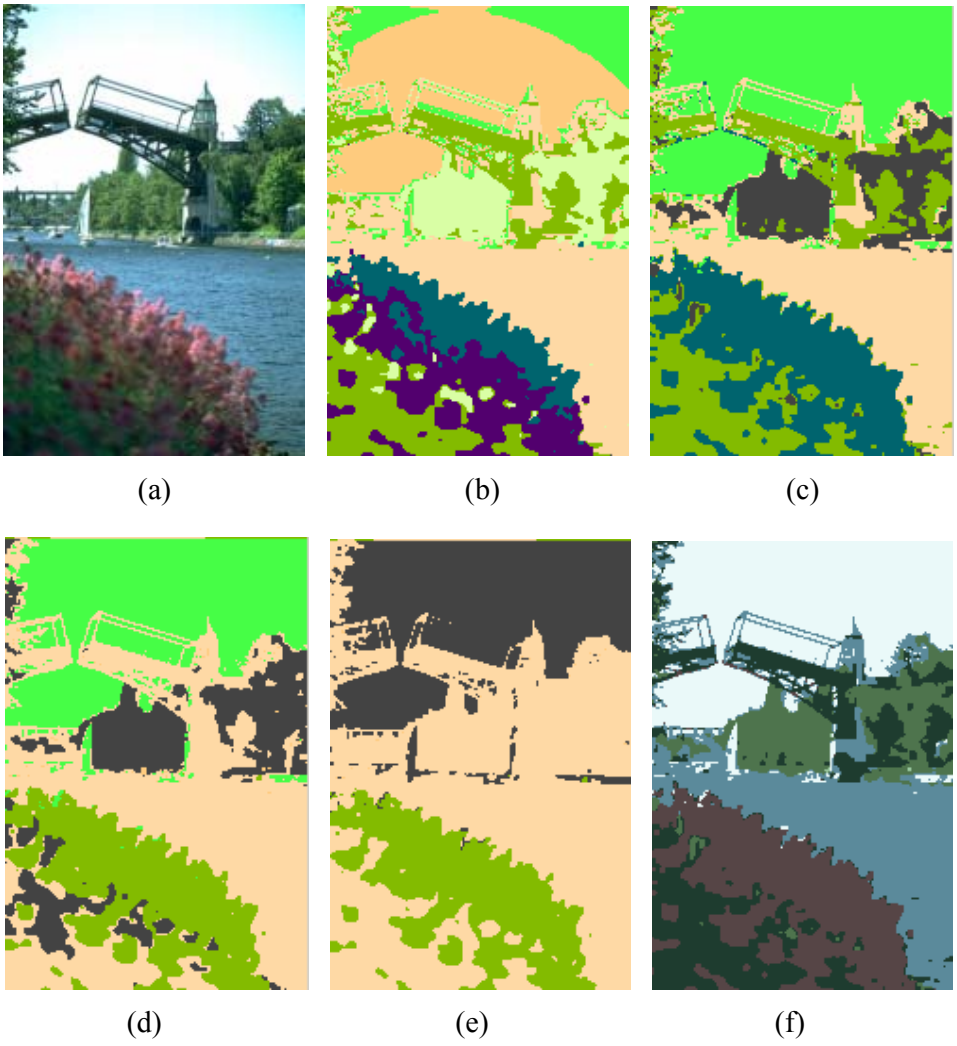


Figure 4.12 Colour image segmentation with parameter optimisation. (a) Natural image [99]. (b) Clustered image 1, $ICV_{SOM} = 0.3$, $S_{avg} = 0.30$, resulting number of clusters $k = 8$; (c) Clustered image 2, $ICV_{SOM} = 0.4$, $S_{avg} = 4.04$, $k = 6$; (d) Clustered image 3, $ICV_{SOM} = 0.5$, $S_{avg} = 1.84$, $k = 5$; (e) Clustered image 4, $ICV_{SOM} = 0.6$, $S_{avg} = 0.24$, $k = 4$. Image (c) generates the maximum $S_{avg} = 4.04$ and the number of clusters automatically detected is $k = 6$. (f) The final colour segmented image where the pseudo colours were replaced with the mean values of the corresponding colours from the original image.

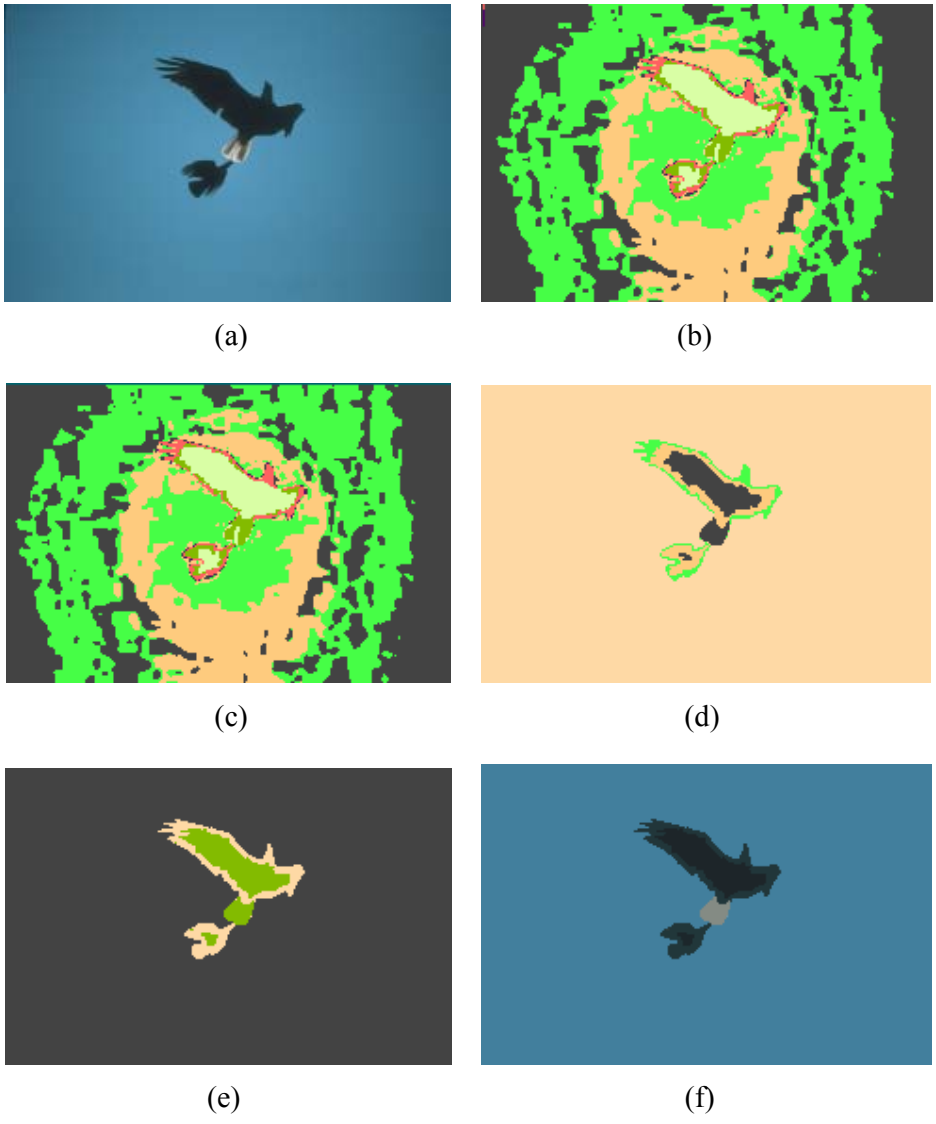
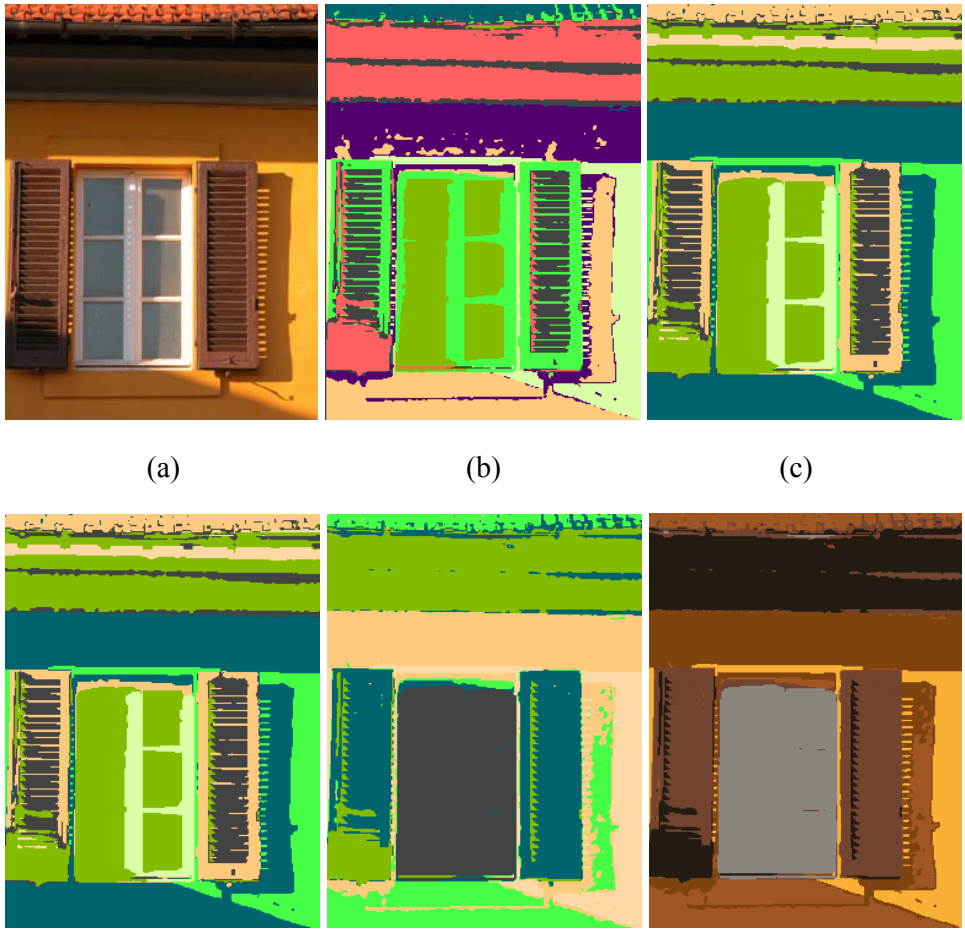


Figure 4.13 Colour image segmentation with parameter optimisation. (a) Natural image [99]. (b) Clustered image 1, $ICV_{SOM} = 0.3$, $S_{avg} = 0.23$, resulting number of clusters $k = 9$; (c) Clustered image 2, $ICV_{SOM} = 0.4$, $S_{avg} = 0.23$, $k = 9$; (d) Clustered image 3, $ICV_{SOM} = 0.5$, $S_{avg} = 0.83$, $k = 5$; (e) Clustered image 4, $ICV_{SOM} = 0.6$, $S_{avg} = 1.33$, $k = 4$. Image (e) generates the maximum $S_{avg} = 1.33$ and the number of clusters automatically detected is $k = 4$. (f) The final colour segmented image where the pseudo colours were replaced with the mean values of the corresponding colours from the original image.



(a) (b) (c)

(d) (e) (f)

Figure 4.14 Colour image segmentation with parameter optimisation. (a) Natural image [102]. (b) Clustered image 1, $ICV_{SOM} = 0.3$, $S_{avg} = 0.24$, resulting number of clusters $k = 9$; (c) Clustered image 2, $ICV_{SOM} = 0.4$, $S_{avg} = 0.39$, $k = 7$; (d) Clustered image 3, $ICV_{SOM} = 0.5$, $S_{avg} = 0.39$, $k = 7$; (e) Clustered image 4, $ICV_{SOM} = 0.6$, $S_{avg} = 0.44$, $k = 6$. Image (e) generates the maximum $S_{avg} = 0.44$ and the number of clusters automatically detected is $k = 6$. (f) The final colour segmented image where the pseudo colours were replaced with the mean values of the corresponding colours from the original image.

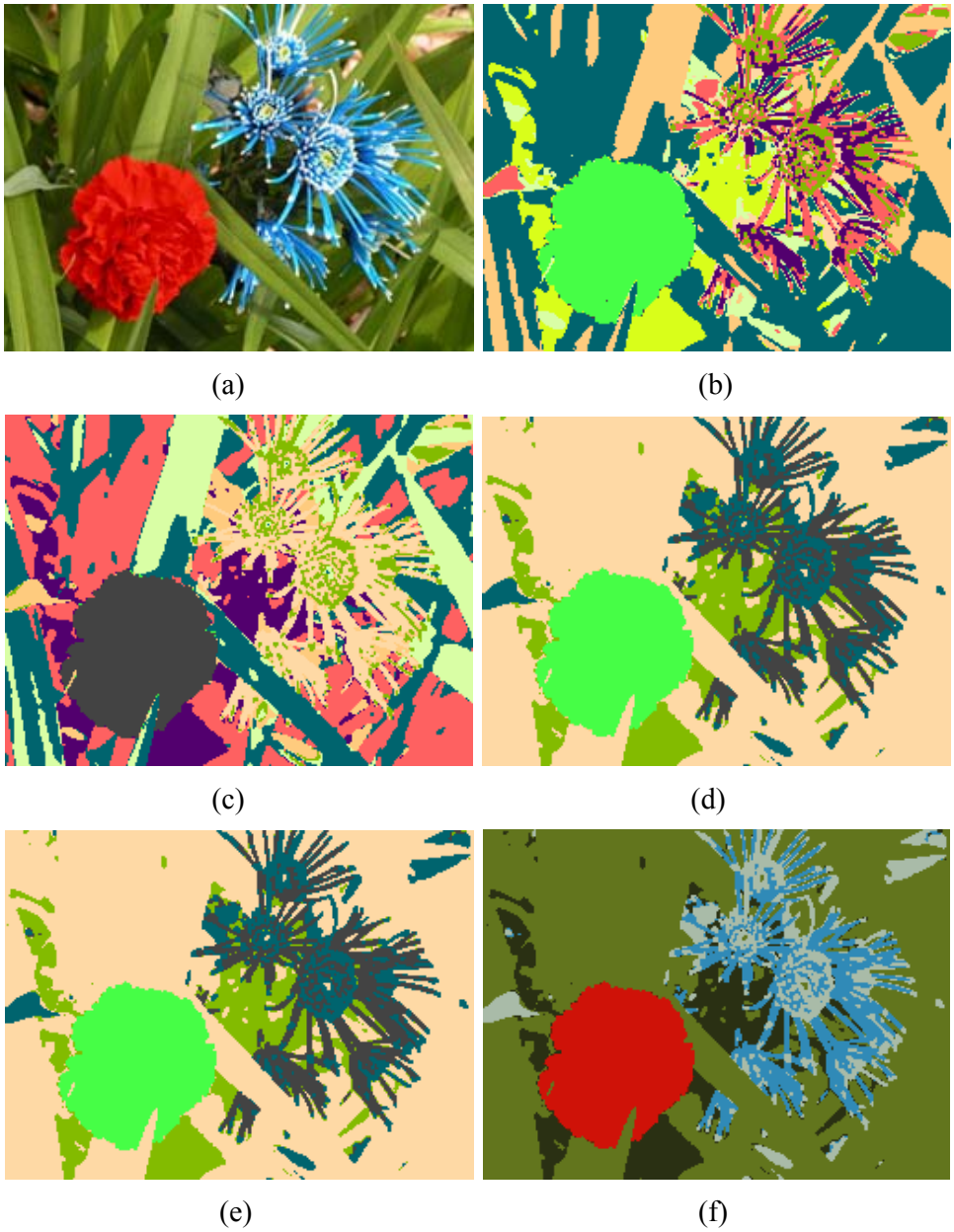


Figure 4.15 Colour image segmentation with parameter optimisation. (a) Natural image [102]. (b) Clustered image 1, $ICV_{SOM} = 0.3$, $S_{avg} = 0.35$, resulting number of clusters $k = 10$; (c) Clustered image 2, $ICV_{SOM} = 0.4$, $S_{avg} = 0.31$, $k = 9$; (d) Clustered image 3, $ICV_{SOM} = 0.5$, $S_{avg} = 0.38$, $k = 6$; (e) Clustered image 4, $ICV_{SOM} = 0.6$, $S_{avg} = 0.38$, $k = 6$. Image (d) generates the maximum $S_{avg} = 0.38$ and the number of clusters automatically detected is $k = 6$. (f) The final colour segmented image where the pseudo colours were replaced with the mean values of the corresponding colours from the original image.

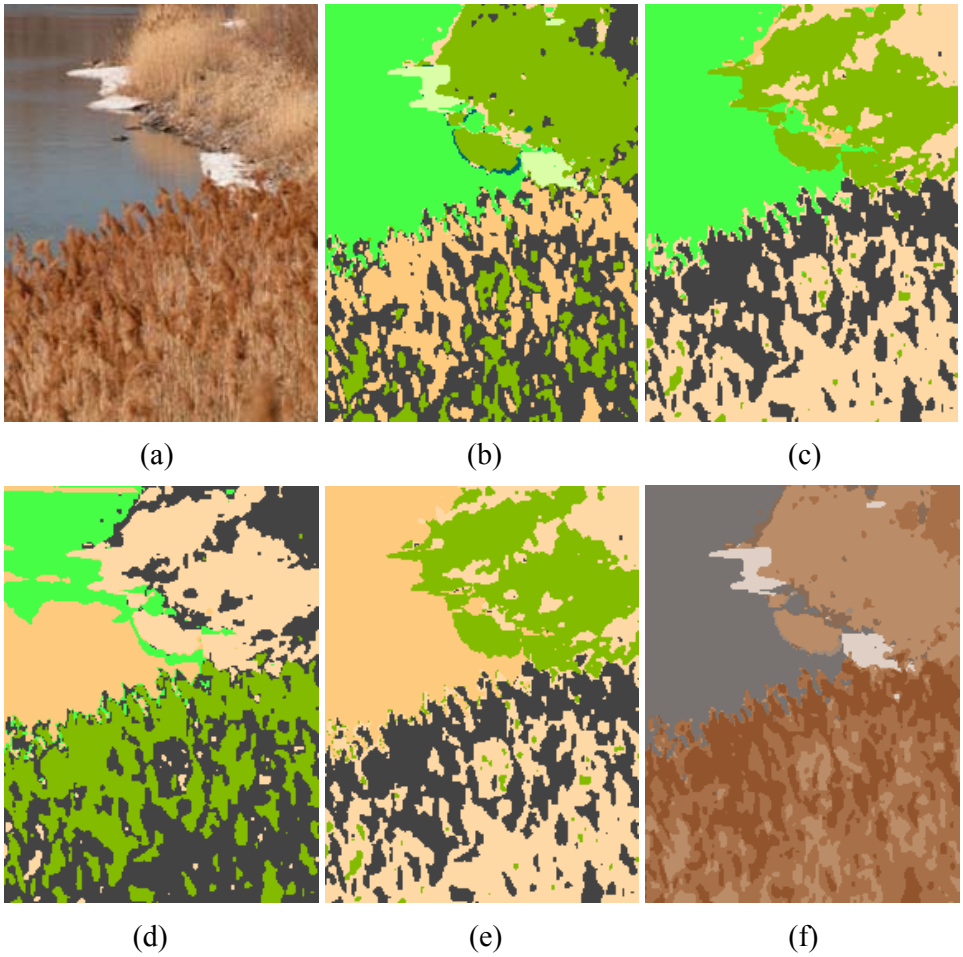


Figure 4.16 Colour image segmentation with parameter optimisation. (a) Natural image [102]. (b) Clustered image 1, $ICV_{SOM} = 0.3$, $S_{avg} = 0.40$, resulting number of clusters $k = 8$; (c) Clustered image 2, $ICV_{SOM} = 0.4$, $S_{avg} = 0.12$, $k = 5$; (d) Clustered image 3, $ICV_{SOM} = 0.5$, $S_{avg} = 0.10$, $k = 5$; (e) Clustered image 4, $ICV_{SOM} = 0.6$, $S_{avg} = 0.09$, $k = 4$. Image (b) generates the maximum $S_{avg} = 0.40$ and the number of clusters automatically detected is $k = 8$. (f) The final colour segmented image where the pseudo colours were replaced with the mean values of the corresponding colours from the original image.

4.5 Experiments and Results

A large number of experiments were carried out to asses the performance of the proposed colour segmentation algorithm. These tests were conducted on Berkeley [99] and McGill [102] natural image databases and the results were

quantitatively and qualitatively evaluated. The first set of tests was conducted in order to determine the performance of the multi-space segmentation technique when analysed in conjunction with complementary pairs of colour spaces. The purpose of this investigation was to determine the pair of colour spaces that returns the best results in the context of colour features extraction. Moreover, in order to illustrate the validity of the proposed scheme, the second set of tests compares the results produced by the proposed colour-segmentation algorithm with those returned by the Mean Shift [34] segmentation benchmark.

4.5.1 Optimal Selection of Complementary Pairs of Colour Spaces

In order to deduce the pair of complementary colour spaces that returns the best colour segmentation performance, a quantitative evaluation of the results returned by the proposed algorithm when used in conjunction with different pairs of colour spaces has been carried out. The proposed multi-space colour segmentation algorithm was tested on one hundred complex natural images from the Berkeley database [99] in order to evaluate its performance with respect to the identification of homogenous colour regions. The Berkeley database includes images with complex colour-texture characteristics, fuzzy borders and low image contrast and also allows a quantitative evaluation of the segmented result as it provides a set of manual segmentations for each colour image. The existence of multiple ground truth data for every image is motivated by the fact that the segmentation of natural images has a subjective character and the images that form the ground truth set correspond to the perception of different subjects. Samples of images from the Berkeley database and their corresponding ground truths segmentations are illustrated in Figure 4.17. The quantitative measurements were produced using the Probabilistic Rand Index (PR) [104] that performs a comparison between the obtained segmentation result and the multiple ground truth segmentations by evaluating the pairwise relationships between the pixels in the segmented result and the ground truth data.

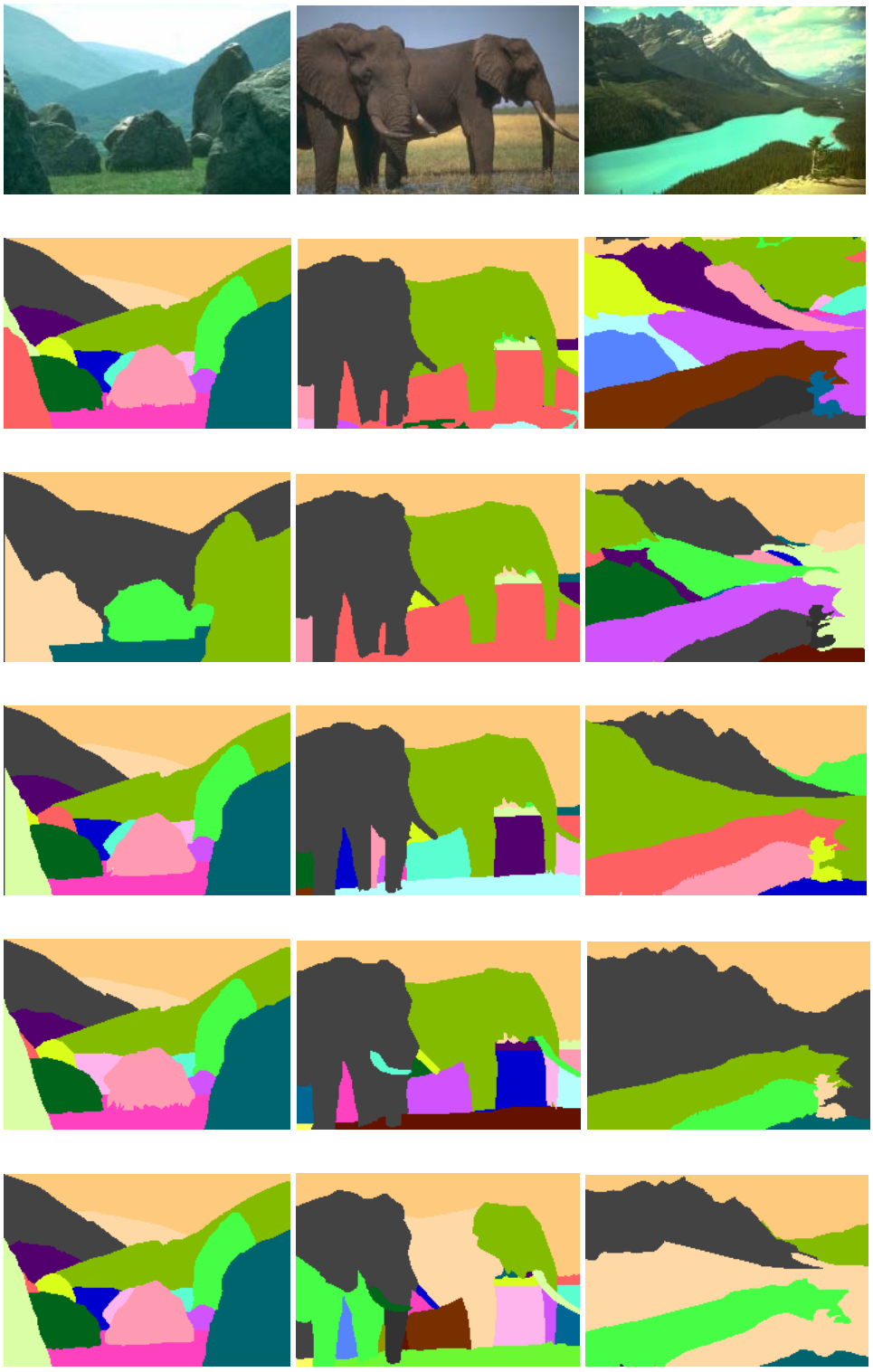


Figure 4.17 Three natural images from the Berkeley database (first row) and their corresponding set of ground truth segmentations.

The PR index measures the agreement between the segmented result and the manually generated ground truths and takes values in the range [0, 1]. A higher PR value indicates a better match between the segmented result and the ground truth data. The PR Index is defined in Appendix A.

TABLE 4.2

Performance evaluation of the proposed Multi-Space Colour Segmentation (MSCS) algorithm when used in conjunction with different pairs of complementary colour spaces. The experiments were conducted on natural images from the Berkeley database.

Method	PR $\text{Index}_{\text{mean}}$	PR $\text{Index}_{\text{standard\_deviation}}$
MSCS (RGB and YIQ)	0.720	0.14
MSCS (RGB and CIE Lab)	0.724	0.15
MSCS (YIQ and CIE Lab)	0.737	0.14
MSCS (YIQ and RGB)	0.734	0.14
MSCS (CIE Lab and YIQ)	0.748	0.13
MSCS (CIE Lab and RGB)	0.742	0.14

Numerical results when the proposed colour segmentation algorithm was used in conjunction with pairs of complementary colour spaces are depicted in Table 4.2. The analysed colour-spaces are: RGB, YIQ and CIE Lab. As expected, the mean values of the PR index are close for all experiments (see Table 4.2) and this indicates that the algorithm shows accurate performance when used in conjunction with different pairs of complementary colour spaces. The explanation of these results is that no colour space has a clear advantage over the others. Still, as shown in Table 4.2, the highest PR value and thus the best performance is obtained when the proposed colour segmentation algorithm is used in conjunction with the CIE Lab and YIQ colour image representations.

4.5.2 Performance Evaluation of the Proposed Colour Segmentation Algorithm

In order to obtain a comprehensive quantitative evaluation of the proposed colour segmentation algorithm, this technique has been applied to the entire Berkeley natural image database [99] that contains 300 colour images where for each image a set of manual segmentations is provided. For a numerical estimation of the obtained results, the Probabilistic Rand Index (PR) was calculated for every image.

Colour segmentation results and the calculated PR values are illustrated in Figure 4.18. Additional results are illustrated in Figure 4.19 when the proposed multi-space colour segmentation algorithm has been applied to several natural images with inhomogeneous colour-texture characteristics. It can be noted that the shapes of the objects follow the real boundaries present in the original images and the small and narrow objects are not suppressed during the colour segmentation process (see the crosses placed on top of the building in Figure 4.18 (b)). The segmentation results illustrated in Figure 4.19 indicate that the colour information alone is not sufficient to describe the regions characterised by complex textures such as the building's windows in Figures 4.19 (b) and (d), the cloth textures in Figures 4.19 (f) and (h), the corn's kernels and the starfish textures in Figure 4.19 (j) and (l). Therefore, a technique that complements the colour segmentation with texture features has been proposed. The texture feature extraction procedures that will be included within the proposed CTex colour-texture segmentation framework will be described in the next chapter of this thesis.

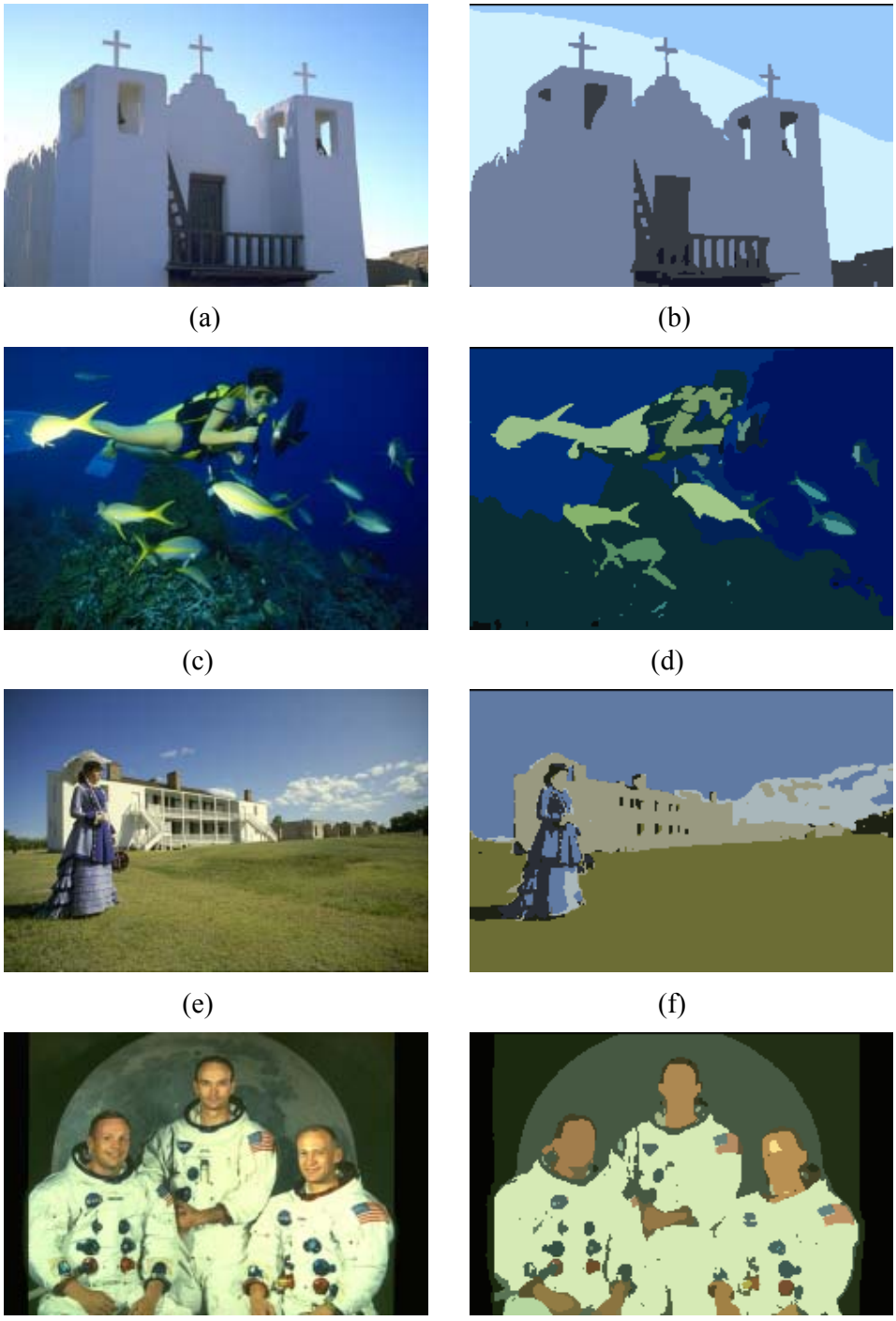


Figure 4.18 First row (a, c, e, g) Original images [99]. Second row (b, d, f, h) Multi-space colour segmentation results. (b) PR = 0.74. (d) PR = 0.73. (f) PR =0.93. (h) PR = 0.85.

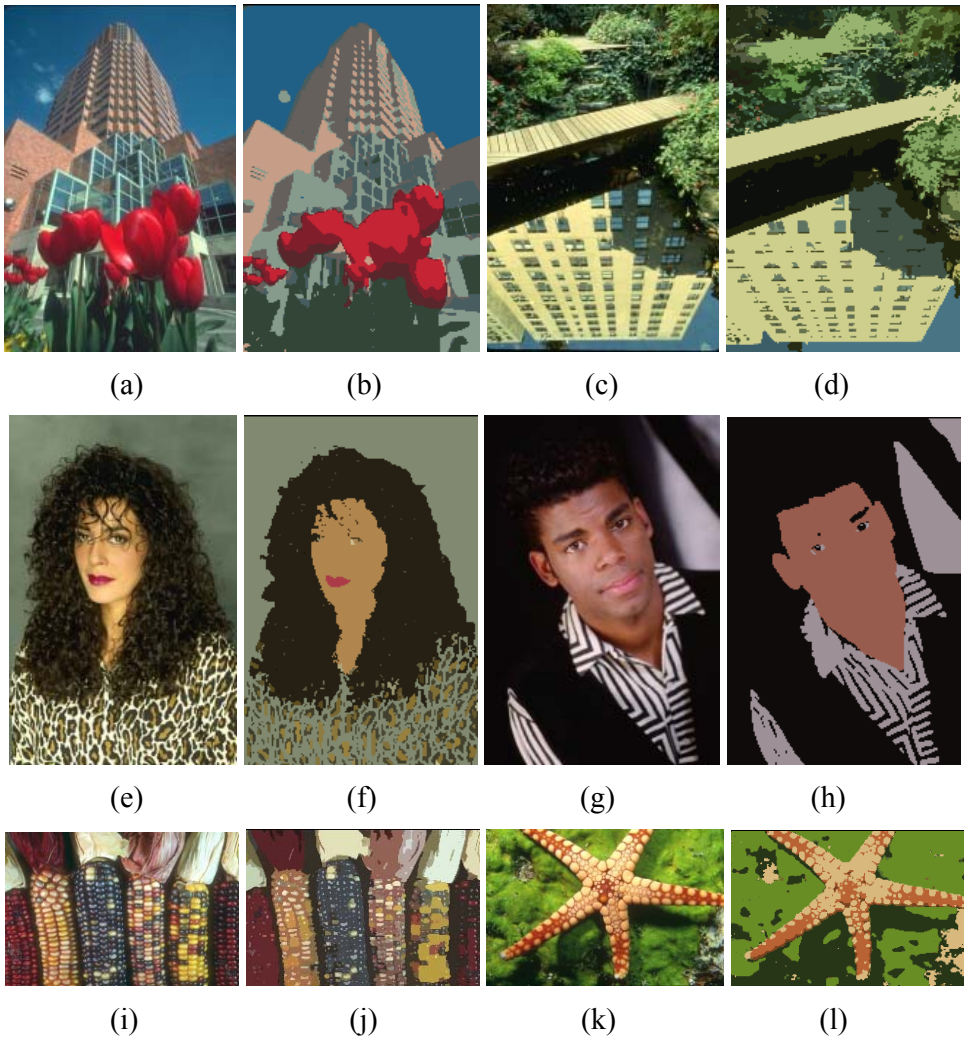


Figure 4.19 Colour segmentation results. (a, c, e, g, i, k) Original natural images sampled from the Berkeley database [99] showing complex colour-texture characteristics. (b, d, f, h, j, l) Multi-space colour segmentation results.

4.5.2.1 Performance Evaluation for the Proposed MSCS and Mean Shift Segmentation Algorithms

In order to illustrate the validity of the proposed technique, we have compared our segmentation results with those obtained using the Mean Shift colour segmentation algorithm [34]. Mean Shift is a standard colour segmentation benchmark and for these experiments the implementation made

available online by the authors (http://www.caip.rutgers.edu/riul/research/code.html) was used.

Mean Shift is a nonparametric density estimator which iteratively computes the nearest mode of a sample distribution. It is important to mention that the performance of the Mean Shift algorithm strongly depends on several parameters that must be manually selected, namely the spatial bandwidth (σ_s), the colour bandwidth (σ_c) and the minimum region area measured in pixels (M). In all experiments conducted in this study, the values of these parameters have been conservatively selected based on the guidelines provided in [34] and were set to the following values: σ_s=3, σ_c=3 and M=10.

In Table 4.2 are depicted the means and standard deviations of the PR values calculated when the proposed Multi-Space Colour Segmentation (MSCS) and Mean Shift (MS) algorithms were applied to the entire Berkeley database (300 images). As it can be observed in Table 4.3, the MSCS algorithm achieved a mean of 0.75, while the mean value for Mean Shift is 0.69.

TABLE 4.3

Performance evaluation of the proposed Multi-Space Colour Segmentation (MSCS) and Mean Shift algorithms conducted on the entire Berkeley database.

Method	PR Index$_{mean}$	PR Index$_{standard_deviation}$
Mean Shift	0.69	0.16
MSCS	0.75	0.13

Colour segmentation results when the proposed algorithm and Mean Shift technique are applied to natural images from the Berkeley database are illustrated in Figure 4.20.

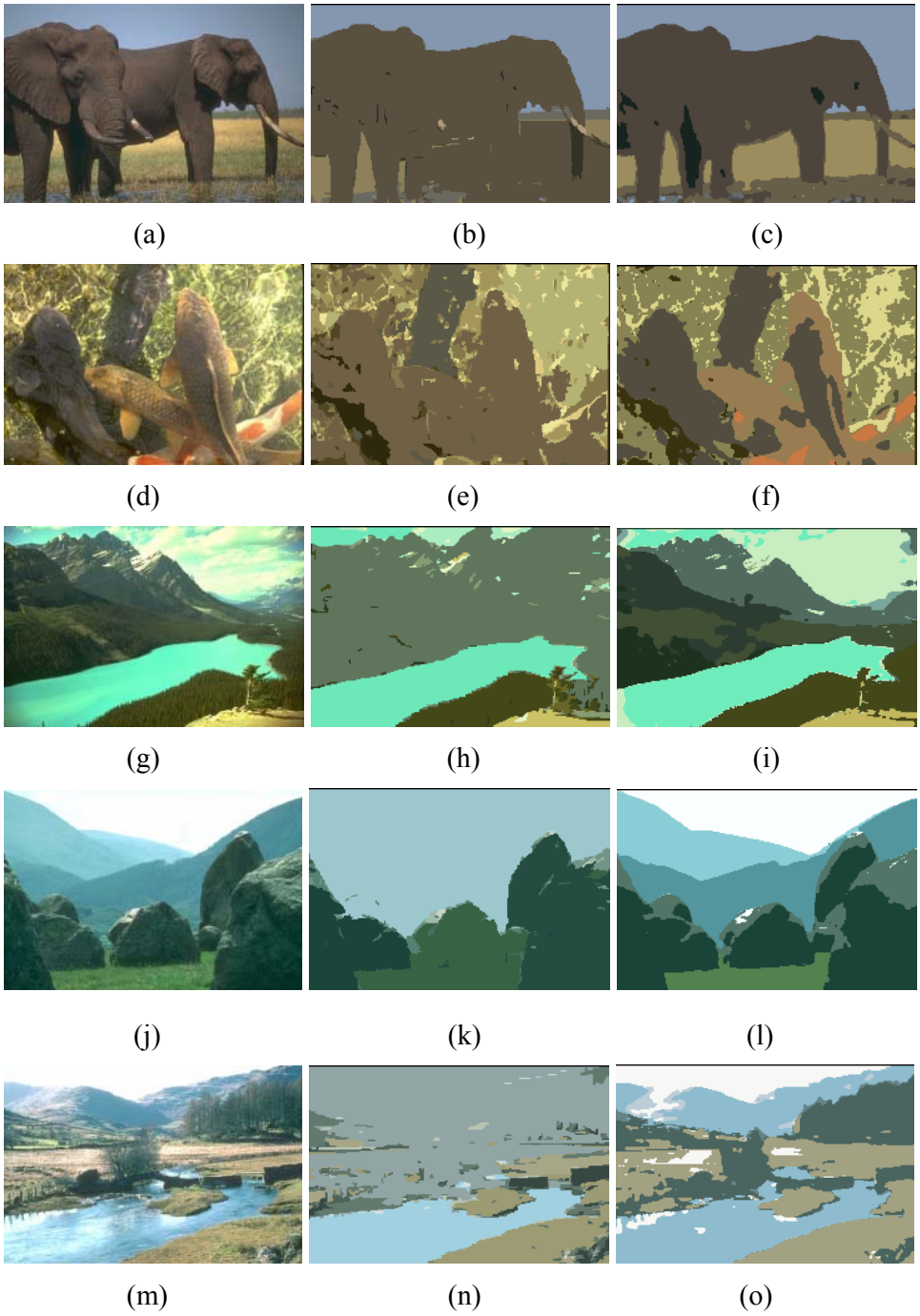


Figure 4.20 Colour segmentation results. (a, d, g, j, m) Natural images from the Berkeley database. (b, e, h, k, n) Mean Shift results: (b) PR = 0.62. (e) PR = 0.67. (h) PR = 0.68. (k) PR = 0.77. (n) PR = 0.71. (c, f, i, l, o) Multi-space colour segmentation (MSCS) results: (c) PR = 0.81. (f) PR = 0.84. (i) PR = 0.87. (l) PR = 0.90. (o) PR = 0.89.

4.6 Conclusions

In this chapter, an important contribution of this thesis that consists of a new formulation for the automatic extraction of colour features is presented. To achieve this, the opponent characteristics of pairs of complementary colour spaces have been used where the key component was the inclusion of an automatic SOM based procedure in the computation of the dominant colours and the estimation of the optimal number of clusters. The selection of the pair of complementary colour spaces that returns the best segmentation results has been experimentally deduced by performing a quantitative evaluation on natural images from the Berkeley database. In order to illustrate the validity of the proposed colour segmentation scheme, the obtained experimental results have been compared with those returned by the Mean Shift colour segmentation algorithm when applied to the entire Berkeley database. The numerical results indicate that the proposed technique outperforms the Mean Shift clustering strategy and is able to return consistent segmentation results when applied to large collections of natural images. Because colour alone is not sufficient to accurately segment complex textured image regions, in this thesis the colour information is complemented with texture features. In the next chapter several texture extraction techniques will be presented and evaluated.

Chapter 5

Texture Features Extraction

The second major component of the proposed segmentation framework deals with the extraction of the texture information that is used to complement the colour features in the segmentation process.

There are a large number of possible approaches that can be applied to the extraction of the texture information and the most relevant are described in the review presented in Chapter 2. Because the segmentation framework proposed in this thesis is unsupervised, the methods based on texture models or methods developed for texture synthesis are not evaluated in this study. Therefore, in this chapter several statistical and signal processing based texture extraction techniques will be described and evaluated. These techniques include: the Local Binary Pattern (LBP) operators (standard and rotational invariant forms), Gabor and isotropic filter banks and a proposed approach [105] that samples the texture information by evaluating the distribution of the texture orientations at micro and macro level. It is important to mention that the texture features are independently extracted from the luminance component of the input image. A number of experiments were conducted in order to provide a numerical evaluation of the above-mentioned texture extraction techniques when applied to the segmentation of mosaic images.

5.1 The Local Binary Pattern (LBP) Operator

The first method investigated in this study is based on the Local Binary Pattern (LBP) operator that was developed by Ojala and Pietikainen [7]. The LBP approach attempts to decompose the texture into small texture units and

the texture features are defined by the distribution of the LBP values calculated for each pixel in the image under analysis. Conceptually, the LBP operator is a powerful texture descriptor as it analyses the texture at micro level, but at the same time characterizes the textured regions at macro level by the distribution of the LBP values. In order to further improve the robustness of the LBP distributions and compensate for the illumination offsets present in textures, the LBP operator is used in conjunction with the contrast operator C.

As indicated earlier, a region in an image can be defined by its textural characteristics that are given by the joint LBP/C distribution. The main advantage of analysing the texture using the distribution of LBP/C values is given by the fact that they can be used to discriminate textures in the input image regardless the region size. In this fashion, the dissimilarities between two or more textured regions can be calculated by using histogram intersection metrics such as Kolmogorov-Smirnov, Kullback-Leibler or χ^2 (chi-square test).

In order to evaluate the spatial distribution of the pixels at local level, the LBP values are calculated in the 3×3 neighbourhood around each pixel in the image.

Figure 5.1 The 3×3 neighbourhood of the central pixel g_c.

The LBP value of the texture unit can be calculated by thresholding the values of the neighbouring pixels $g_0 \ldots g_7$ (see Figure 5.1) with respect to the value of the central pixel g_c as follows:

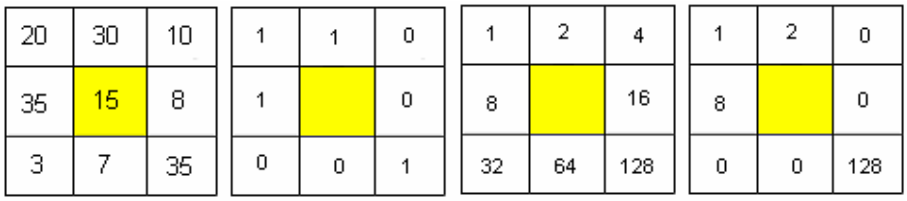
$$T = \{t(g_0 - g_c), \ldots, t(g_{P-1} - g_c)\}, \qquad t(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{5.1}$$

In equation (5.1) T is the texture unit, g_c is the grey value of the central pixel, g_P are the pixels adjacent to the central pixel, P defines the number of pixels in the neighbourhood and $t(.)$ is the threshold operation. For a 3×3 neighbourhood, the value of P is 8. To encode the spatial arrangement of the binary values in the texture unit T, the LBP value is calculated by multiplying the elements of the vector T with the binomial weights and summing the result as follows:

$$LBP = \sum_{i=0}^{P-1} t(g_i - g_c) \cdot 2^i \tag{5.2}$$

Because the LBP values are inadequate to measure the greyscale variations, they are often used in conjunction with the contrast measure C that is defined as the difference between the average grey-level of the pixels with values 1 and the pixels with values 0 contained in the texture unit T.

$$C = avg_{(g_i \geq g_c)}(g_i) - avg_{(g_i < g_c)}(g_i) \tag{5.3}$$



(a) (b) (c) (d)

LBP = 1+2+8+128=139

$$C = \frac{20+30+35+35}{4} - \frac{10+8+3+7}{4} = 23$$

Figure 5.2 The computational steps required to calculate the LBP and C values. (a) The 3×3 neighbourhood. (b) The result obtained after the application of the threshold operation using equation (5.1). (c) The binomial weights corresponding to each position in the 3×3 neighbourhood. (d) The final result is obtained by multiplying the elements of the matrix (b) with the binomial weights shown in (c).

An example that illustrates the computational steps required to compute the Local Binary Pattern (LBP) value and the contrast measure C is shown in Figure 5.2.

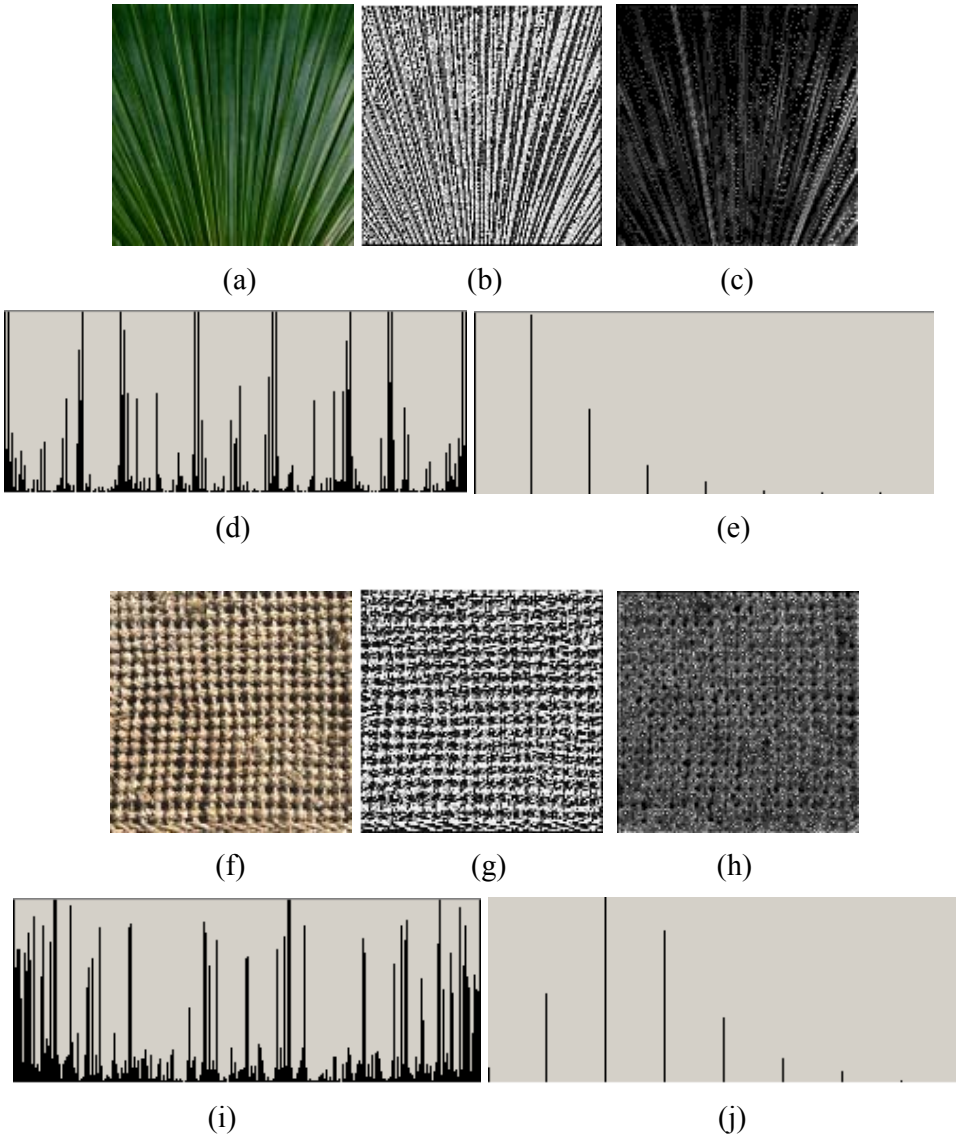


Figure 5.3 The LBP and contrast C distributions associated with different VisTex textures [106]. (a) Oriented and (f) isotropic textures. (b, g) LBP (texture) images. (c, h) Contrast images (*bins* = 8). In these graphs on the *x*-axis are shown the LBP (d, i) and C (e, j) values while on the *y*-axis are plotted the number of elements in each bin.

The distribution of the LBP/C values calculated over an image region represents the texture spectrum that can be defined as a joint histogram of size (256 + *bins*), where the first 256 bins are required by the distribution of the LBP values and *bins* represents the number of bins employed to sample the contrast measure. Figure 5.3 illustrates the LBP and contrast C distributions for two different texture images: one oriented (Figure 5.3 (a)) and one isotropic (Figure 5.3 (f)). Based on the experiments performed by Ojala and Pietikainen [7], the best results are obtained when the contrast distribution is quantised into 4 to 16 bins. The optimal selection of the number of bins is a difficult issue since for low values of *bins* the histogram will lack resolution, while for high values of *bins* the histogram will become sparse and unstable. This can be observed in Figure 5.4 where the contrast (C) distributions at different quantisation levels for a VisTex texture image are shown. Based on experimentation it has been demonstrated that a quantisation of the contrast measure in 8 bins returns the best results.

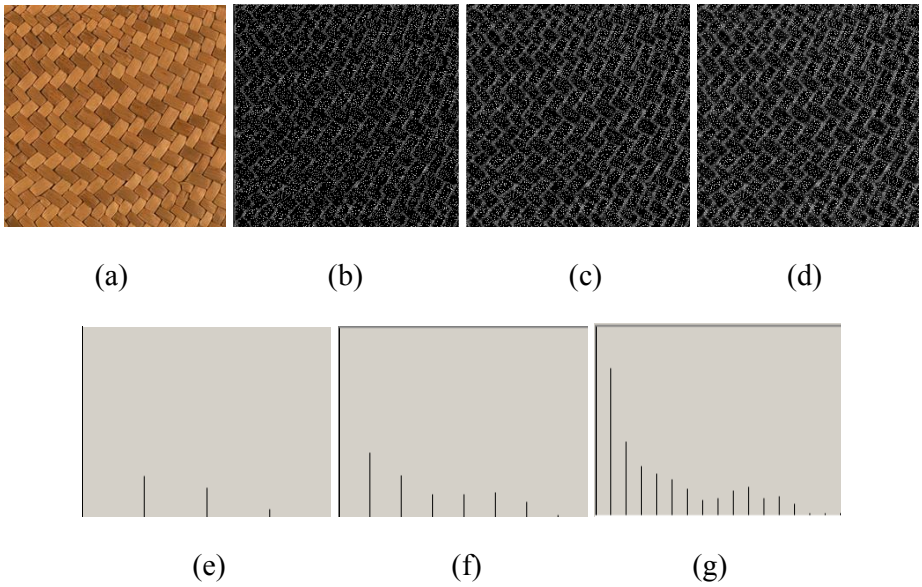


Figure 5.4 The contrast (C) distributions calculated at different quantisation levels. (a) Original VisTex image. (b-d) The Contrast images when the quantisation level is set to 4, 8 and 16 respectively. (e-f) The distribution of the contrast values when the quantisation level is set to 4, 8 and 16 respectively.

5.2 The Rotation Invariant Local Binary Pattern Operator

The LBP values calculated for each texture unit using equation (5.2) are sensitive to texture orientation. This is motivated by the fact that the elements of the texture unit are multiplied with the binomial weights that uniquely encode the position of each pixel in the 3×3 neighbourhood. To address this problem, Ojala et al [5] proposed to modify the LBP descriptor in order to achieve rotation invariance. To remove the sensitivity to rotation, the texture descriptor has to be calculated within a circular neighbourhood and the texture should be evaluated in terms of uniformity. Since the pixels in the image are organised as a discrete matrix, the pixels situated within the circular neighbourhood are not positioned exactly on the image grid and their values are calculated by using bilinear interpolation. To enforce the concept of uniformity in the calculation of the LBP values, Ojala et al [5] introduced the term "uniform patterns" that is defined as the number of transitions between 0 and 1 in the LBP mask obtained after thresholding the pixels from the circular neighbourhood with the intensity value of the central pixel. In this way, they defined a pattern as uniform if the binary LBP pattern has maximum two transitions; otherwise the pattern is classed as non-uniform. For instance if the LBP value is calculated in a circular 8 pixels neighbourhood, the binary patterns 00000000, 00000001, 00000011, 00001000,…, 11111111 are classed as uniform while patterns such as 01100101, 10101000, 01101101,…, 10101010 are classed as non-uniform (see Figure 5.5).

As indicated earlier, the idea behind the uniformity concept is to group the non-uniform patterns in a distinct class while the values for uniform patterns are given by the total number of elements with a value 1 in the binary LBP pattern.

In the original implementation [5], the rotational invariant LPB descriptors (LBP^{ri}) are calculated for neighbourhoods with different sizes (see Figure 5.6) as follows:

$$T_r(g_m, g_n) = \left| t(g_m - g_c) - t(g_n - g_c) \right| \tag{5.4}$$

$$U(LBP_{P,R}) = T_r(g_{p-1}, g_0) + \sum_{p=1}^{P-1} T_r(g_p, g_{p-1}) \tag{5.5}$$

$$LBP_{P,R}^{ri} = \begin{cases} \sum_{p=0}^{P-1} t(g_p - g_c) & if \quad U(LBP_{P,R}) \le 2 \\ P+1 & Otherwise \end{cases} \tag{5.6}$$

where P is the number of pixels in the LBP mask, R is the radius of the mask, *ri* indicates that the LBP value is rotational invariant, $T_r(g_m,g_n)$ defines the transition between pixels with indexes m and n and U is the function that evaluates the uniformity of the binary LBP pattern. To improve its discriminative power, the $LBP_{P,R}^{ri}$ value is complemented with the contrast measure that is calculated as the variance of the pixels situated in the LBP mask.

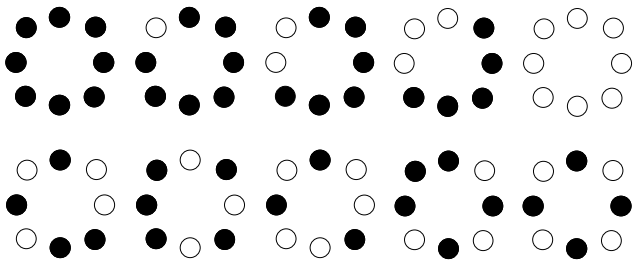
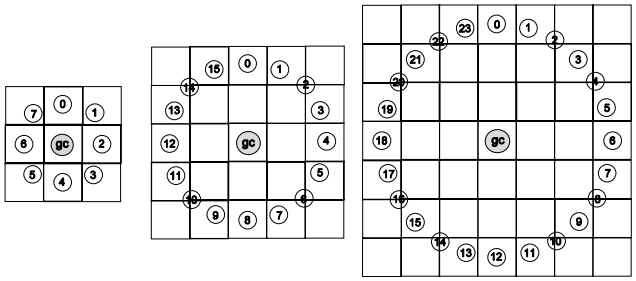


Figure 5.5 LBP masks resulting after the application of the thresholding operation. The pixels that return the value 0 in equation (5.1) are marked in the diagram with a black disc while the pixels that generate the value 1 are marked with a white disc. (Top row) Examples of uniform patterns (maximum two transitions in the binary pattern). (Bottom row) Examples of non-uniform patterns.



(a) (b) (c)

Figure 5.6 $LBP^{ri}_{P,R}$ pattern sizes where: (a) P=8, R=1. (b) P=16, R=2. (c) P=24, R=3.

5.3 Multi-channel Texture Decomposition Using Gabor Filtering

There has been a widely accepted consensus among vision researchers that filtering an image with a large number of oriented band pass filters such as Gabor represents an optimal approach to analyse textures [20, 21, 22]. This approach implements a multi-channel texture decomposition and is achieved by filtering the input image with a two-dimensional (2D) Gabor filter bank that was initially suggested by Daugman [107] and later applied to texture segmentation by Jain and Farrokhnia [21]. The 2D Gabor function that is used to implement the even-symmetric 2D discrete filters can be written as:

$$G_{\sigma,f,\varphi}(x,y) = \exp\left(-\frac{x'^2+y'^2}{2\sigma^2}\right)\cos\left(2\pi f x'+\varphi\right) \tag{5.7}$$

$$B = \log_2\left(\frac{\pi\sigma f + \sqrt{\ln(2)/2}}{\pi\sigma f - \sqrt{\ln(2)/2}}\right) \tag{5.8}$$

where $x' = x\cos\theta + y\sin\theta$, $y' = -x\sin\theta + y\cos\theta$ and B is the bandwidth of the filter. In equation (5.7), the parameter σ represents the scale of the Gabor filter, θ is the orientation and f is the frequency parameter that controls the

number of cycles of the cosine function within the envelope of the 2D Gaussian (φ is the phase offset and it is usually set to zero to implement 2D even-symmetric filters).

Gabor filter banks are constructed using band pass filters where the parameters σ, θ and f determine the sub-band that is covered by each filter in the spatial-frequency domain. The parameters of the Gabor filters are chosen to optimise the trade-off between spectral selectivity and the size of the bank of filters. Typically, the central frequencies are selected to be one octave apart and for each central frequency is constructed a set of filters corresponding to four (0^0,45^0,90^0,135^0) or six orientations (0^0,30^0,60^0,90^0,120^0,150^0). To illustrate the spectral selectivity of the Gabor filters, in Figure 5.7 a number of Gabor filters with different values for σ, θ and f parameters are illustrated. Figures 5.8 and 5.9 show the textures features extracted from a natural image, when the Gabor filters are calculated using four and six orientations.

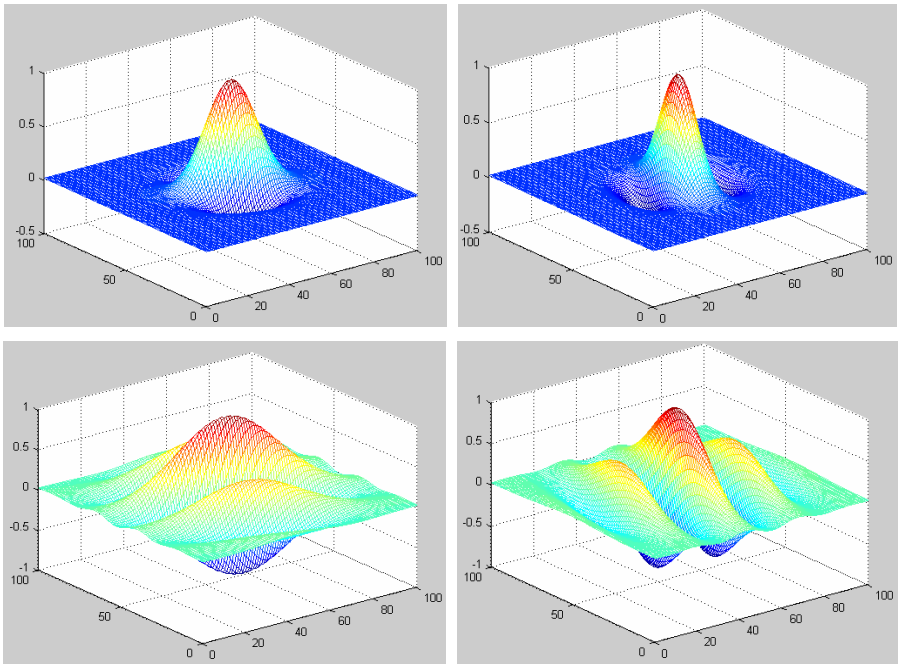
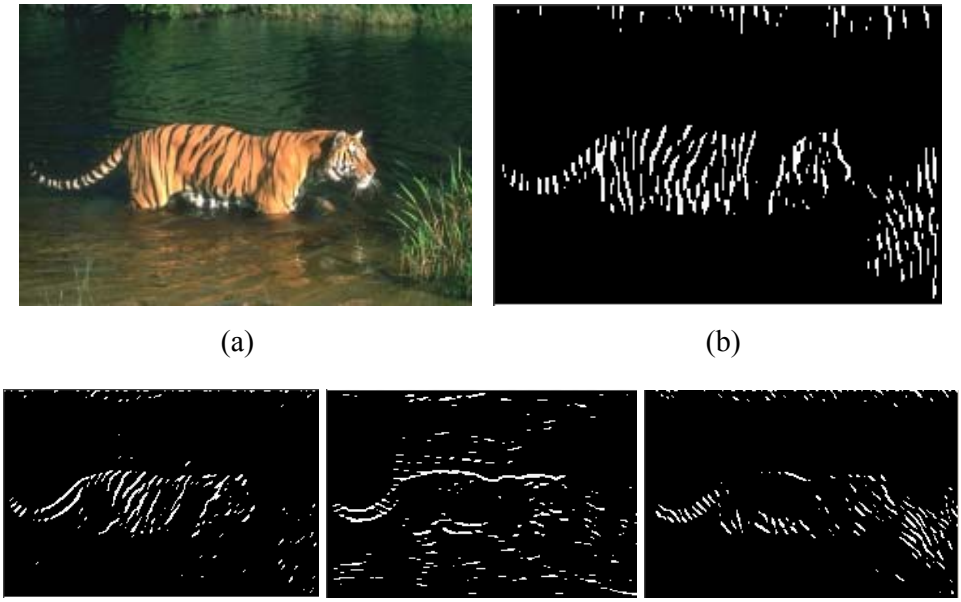


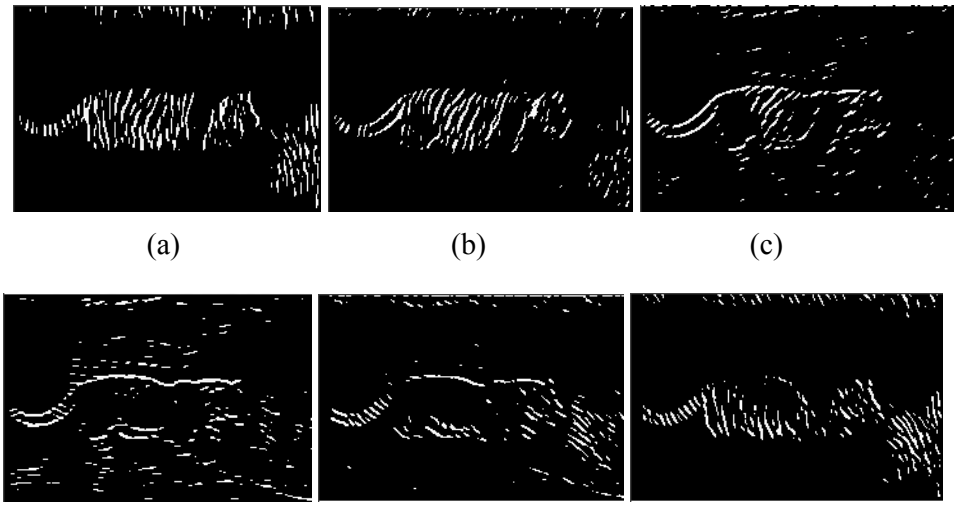
Figure 5.7 Two-dimensional (2D) Gabor filters for 30^0 and 120^0 orientations. (Top row) scale σ = 1.0, central frequency f= 1.5/2π. (Bottom row) Scale σ = 2.0, central frequency f= 2.5/2π.



(a) (b)

(c) (d) (e)

Figure 5.8 (a) Original image. (b-e) The texture features extracted from the natural image using a Gabor filter bank with four orientations: (b) 0^0. (c) 45^0. (d) 90^0 and (e) 135^0 (filter size = 9; $\sigma = 3.0$; $f = 1.5/2\pi$).



(a) (b) (c)

(d) (e) (f)

Figure 5.9 The texture features extracted for the natural image depicted in Figure 5.8 (a) using a Gabor filter bank with six orientations. (a) 0^0. (b) 30^0. (c) 60^0 and (d) 90^0. (e) 120^0. (f) 150^0 (filter size = 9; $\sigma = 3.0$; $f = 1.5/2\pi$).

5.4 Multi-channel Texture Decomposition Using Isotropic Filters

The main disadvantage associated with approaches based on multi-channel texture decomposition using a large number of Gabor filter banks consists in the high computational cost. Therefore, to limit the number of orientations in the Gabor filter bank it would be advantageous if the spectral filters would have isotropic characteristics, namely rotation invariant filters. Based on this idea, Schmid [108] developed a set of rotational invariant Gabor-like filters, also referred to as Symmetrical Filters (S-Filters) that were applied to construct texture models for image retrieval. The S-Filters proposed in [108] are defined as follows:

$$S_{\tau,\sigma}(x,y) = S_o(\tau,\sigma) + \cos\left(\frac{\pi\tau\sqrt{x^2+y^2}}{\sigma}\right)\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{5.9}$$

In equation (5.9) τ is the frequency parameter that controls the number of cycles of the cosine function within the Gaussian envelope, σ is the scale parameter and the term S_0 is added to remove the DC component of the 2D S-Filters. The parameters τ and σ can be adjusted to capture a particular sub-band in the spatial-frequency domain as illustrated in Figure 5.11. Figure 5.10 shows the textures features extracted from the natural image depicted in Figure 5.8 (a), when the S-Filters are calculated using different (σ, τ) parameters.

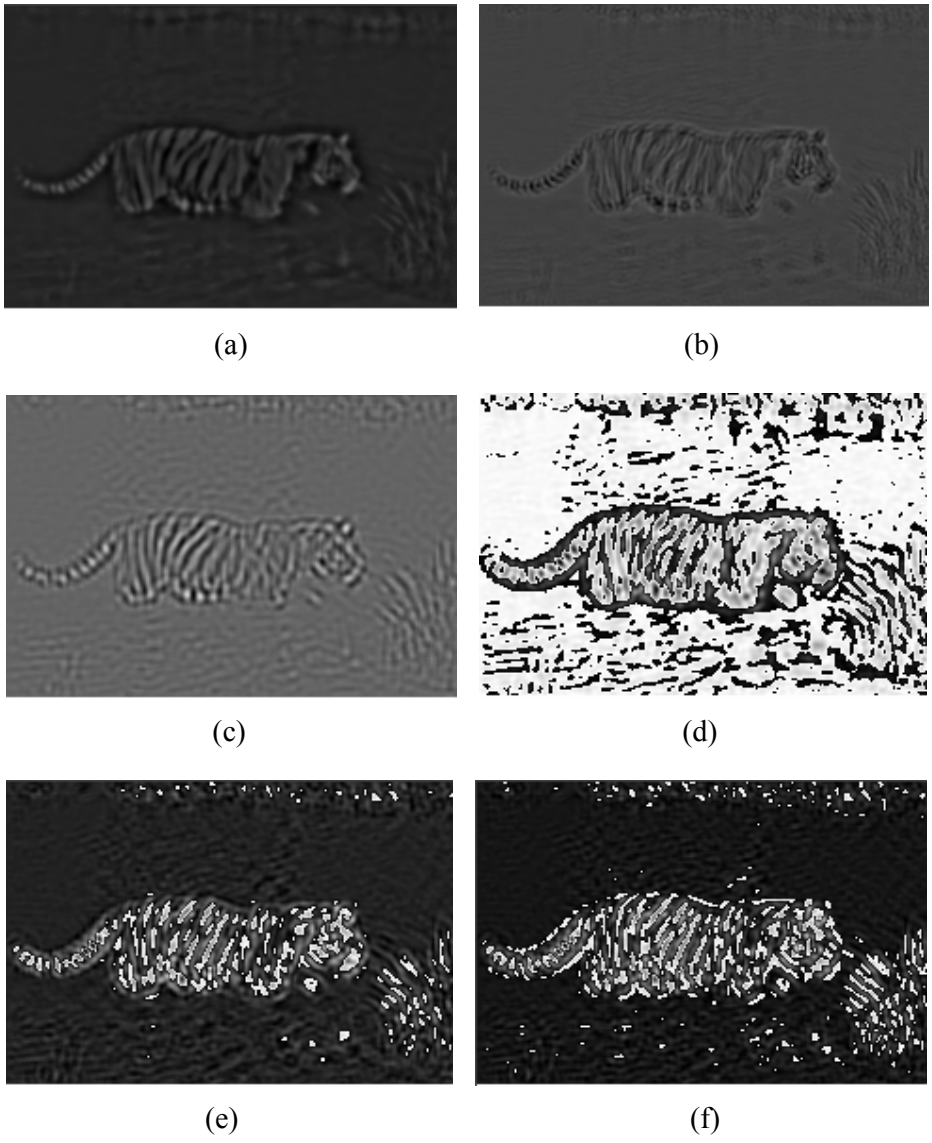


Figure 5.10 The texture features extracted for the natural image depicted in Figure 5.8 (a) using a bank of S-Filters with the following parameters (σ, τ): (a) σ =1.0, τ = 0.0. (b) σ = 2.0, τ = 1.0. (c) σ = 4.0, τ = 1.0. (d) σ = 6.0, τ = 2.0. (e) σ = 8.0, τ = 3.0. (f) σ = 10.0, τ = 4.0.

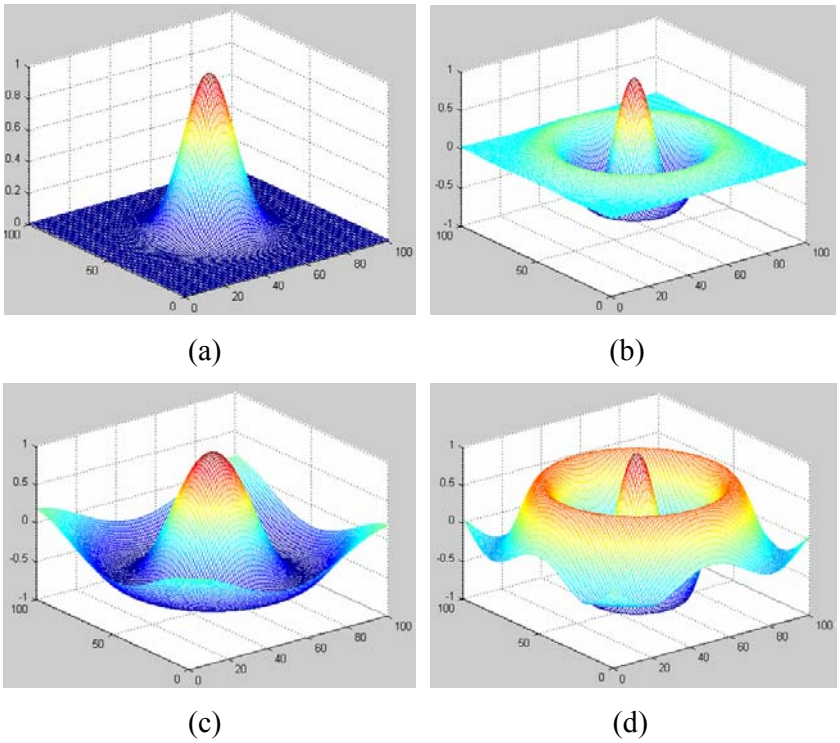


Figure 5.11 Two-dimensional (2D) S-Filters constructed using the following (σ, τ) parameters: (a) (1, 0). (b) (2, 1). (c) (4, 1) and (d) (4, 2).

5.5 Texture Features Extraction Using Local Image Orientation Distributions

In this thesis a new hybrid statistical-structural approach where the texture is described in terms of the distribution of orientations calculated at micro-level for all pixels in the image [105] is proposed. In this regard, the orientation for each pixel in the image is extracted using the partial derivatives of the Gaussian function while the main focus is centred on the evaluation of the local dominant orientation.

5.5.1 Estimation of the Edge Orientation

The local orientation in the image is typically obtained by calculating the first derivatives in two orthogonal directions [109] and the edge orientation can simply be determined by using the expression illustrated in equation (5.10):

$$\nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right], \qquad \Theta = arctg\left(\frac{\partial f / \partial y}{\partial f / \partial x}\right) \tag{5.10}$$

In equation (5.10) ∇ defines the gradient operator and Θ is the edge orientation. To circumvent the problems caused by image noise, the partial derivatives in the x and y directions are calculated using filters that implement the derivatives of the Gaussian function. This is achieved by filtering the image with 1-dimensional (1D) operators that are calculated using the expression illustrated in equation (5.11).

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}, \qquad \frac{\partial G}{\partial x} = -\frac{x}{\sigma^2} G(x) \tag{5.11}$$

where σ is the scale of the Gaussian function (see Figure 5.12 for a plot of the derivative of the Gaussian function for different scale parameters).

The main advantage of the calculation of the partial derivatives using the derivative of the Gaussian resides in the fact that the Gaussian has a smoothing effect and the scale parameter controls the amount of noise reduction. After the calculation of the partial derivatives, the weak edge responses were eliminated by applying a non-maxima suppression procedure (similar to that described in [110]) and the edge orientation is determined for each pixel using the expression illustrated in equation (5.10).

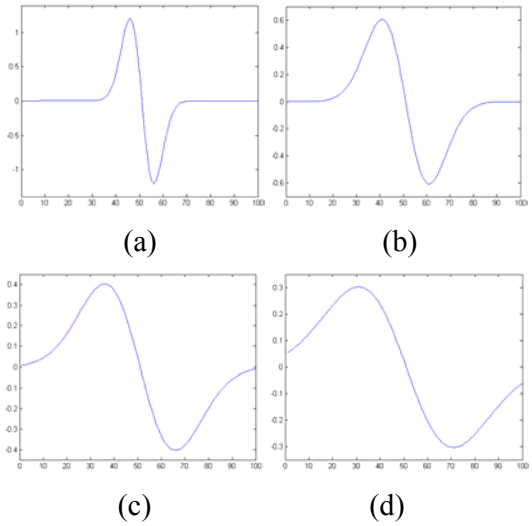


Figure 5.12 The derivative of Gaussian function. (a) $\sigma = 0.5$. (b) $\sigma = 1.0$. (c) $\sigma = 1.5$. (d) $\sigma = 2.0$.

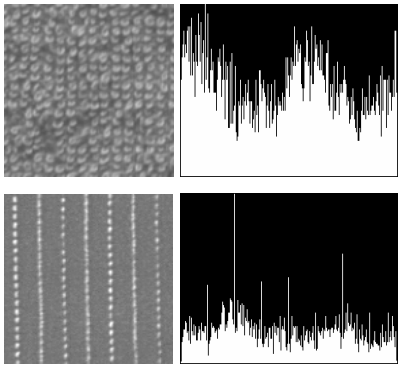


Figure 5.13 Distributions of edge orientations calculated for two textures (top - isotropic and bottom - oriented) from the Outex database [111].

The distributions of the edge orientations calculated for two different textures from the Outex texture database [111] are illustrated in Figure 5.13. It can be observed that these distributions provide the primary discrimination between textures. Its main disadvantage is the fact that the distribution of edge orientations is calculated over the entire image and is not able to robustly sample the texture orientation at micro-level. Therefore, we propose to evaluate the dominant orientation of the texture calculated at micro-level for all texture units that are defined as the local neighbourhood around each pixel in the

image. The distribution of the dominant orientations calculated for all texture units is employed to capture the local orientation of texture at macro-level.

5.5.2 Estimation of the Dominant Texture Orientation at Micro and Macro-Level

The problem of analysing the texture orientation at a given observation scale is not a straightforward task as the orientation of textures may be isotropic at macro-level but having a strong orientation at micro-level. This problem has been addressed in the paper by Germain et al [112] where they analysed a texture that is generated using sinusoidal curves having the same period. Thus, when the texture is evaluated at micro-level in a small neighbourhood (for example 3×3 or 5×5) the texture appears to have a strong orientation, but when it is analysed in large neighbourhoods the texture appears to be isotropic. This is illustrated in Figure 5.14.

The next problem is how to extract the dominant orientation of the texture information in a local $k \times k$ neighbourhood. In this implementation, the orientation of the texture is determined by constructing the histogram of orientations for all pixels in the local neighbourhood and the dominant orientation is selected as the dominant peak in the histogram as follows:

$$H_\Theta(x,y) = \bigcup_{i \in D} h_\Theta(x,y,i), \qquad h_\Theta(x,y,i) = \sum_{x-(k/2)}^{x+(k/2)} \sum_{y-(k/2)}^{y+(k/2)} \delta(\Theta(x,y),i)$$

$$\text{where } D \in [0, 2\pi] \ , \ \delta(i,j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{5.12}$$

$$\Theta_d = \arg\max(H_\Theta) \tag{5.13}$$

In equations (5.12) and (5.13), Θ is the local orientation, i is the orientation bin, D defines the orientation domain, $H_\Theta(x, y)$ is the distribution of the local orientations and Θ_d is the dominant texture orientation in the neighbourhood

$k \times k$. The dominant orientation calculated at macro-level ($H_{\Theta d}$) is simply estimated by the distribution of the dominant orientations that are calculated over the region of interest as follows:

$$H_{\Theta_d} = \bigcup_{i \in D} \int_{\Gamma} \delta(\Theta_d^{k \times k}(x, y), i) d\Gamma \tag{5.14}$$

where Γ is the image domain. In equation (5.14) it should be noted that the texture orientation is sampled at a pre-defined observation scale that is controlled by the size of the neighbourhood $k \times k$.

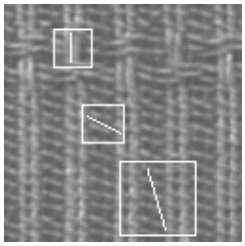


Figure 5.14 The variation of the texture orientation at different observation scales.

The dominant orientation in the image is a powerful local texture descriptor, but it is not robust in sampling the difference between textures that are subjected to illumination variation and image inversion. Thus, the local texture orientation is augmented with measures such as local orientation coherence and contrast that are calculated in the local neighbourhood $k \times k$ where the dominant orientation of the texture has been estimated. In the implementation detailed in this thesis, the contrast measure is sampled by the mean grey-scale value calculated in the $k \times k$ neighbourhood and the orientation coherence (Θ_c) is calculated using the weighted standard deviation of the edge orientation of all the pixels in the neighbourhood as follows:

$$\Theta_c(i, j) = \sqrt{\frac{1}{k^2} \sum_{m,n \in (k \times k)} (I_x^2(m, n) + I_y^2(m, n)) \cdot (\Theta(m, n) - \Theta_{avg})^2} \tag{5.15}$$

In equation (5.15) (*i, j*) are the coordinates of the pixel of interest, I_x and I_y are the partial derivatives calculated for all pixels with coordinates (*m*, *n*) in the window *k*×*k* and Θ_{avg} is the average edge orientation calculated for all pixels in the window *k*×*k*. The feature vector for the edge orientation technique is formed by the three distributions, namely the dominant orientation (Θ_d), the contrast value (C) and the orientation coherence (Θ_c) that are calculated over the entire image at different observation scales (the observation scales are varied in small increments to sample the oriented or isotropic character of texture). It is useful to mention that these three measures define the texture composition at micro-level while their distributions calculated over the image regions define the features that describe the texture at macro-level. This process is illustrated in Figure 5.15.

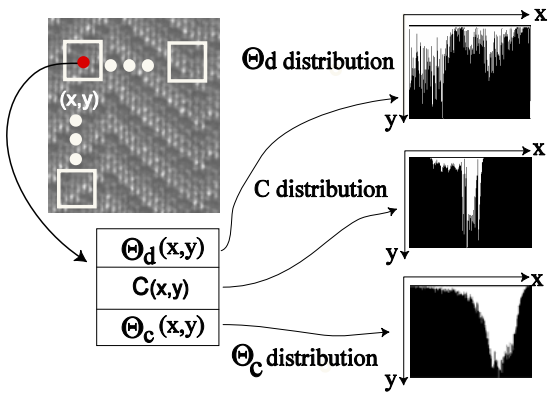


Figure 5.15 The calculation of the dominant orientation, contrast and orientation coherence distributions.

To sample the texture information at different observation scales, the orientation, contrast and orientation coherence distributions are calculated for different values of the window parameter *k*. The size of the local window is varied from 3×3 to 11×11. Figure 5.16 illustrates the shape of the macro-level distributions when the dominant orientation is calculated for all pixels in the image depicted in Figure 5.16 (a) in neighbourhoods of varying sizes. It can be

noticed that the distribution of the local dominant orientations in the image shows apparent peaks when calculated in small neighbourhoods (3×3) and shows an even distribution when calculated in larger neighbourhoods such as 7×7 and 11×11. This is motivated by the fact that the dominant orientation calculated in small neighbourhoods samples the local orientation of the texture, while with the increase in the size of the neighbourhood the dominant orientation tends to sample the isotropic character of the texture.

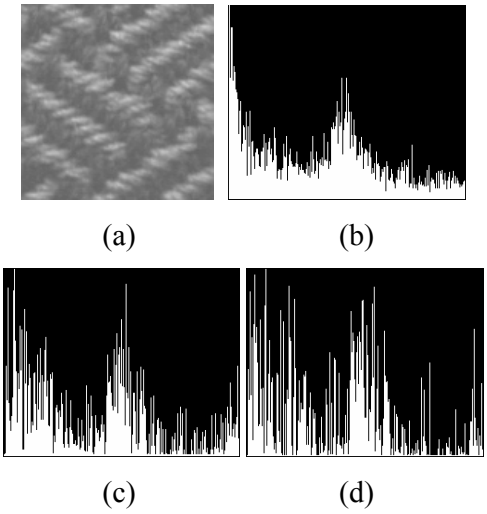


Figure 5.16 The distribution of the dominant orientations when the window parameter k is varied. (a) Input texture image from Outex database [111]. Distribution of the dominant orientations calculated for texture units in (b) 3×3, (c) 7×7 and (d) 11×11 neighbourhoods.

5.6 Experiments and Results

5.6.1 Experimental Setup

Since the ground truth data associated with complex natural images is difficult to estimate and its extraction is highly influenced by the subjectivity of the human operator, the texture evaluation was performed on synthetic data where the ground truth is unambiguous. Therefore, the proposed CTex algorithm was applied on a database of 33 mosaic images (image size

184×184) that were created by mixing textures from VisTex [106] and Photoshop databases. The mosaics used in these experiments consist of various texture arrangements that also include images where the borders between different regions are irregular. The suite of 33 mosaic images is depicted in Figure 5.17.

Figure 5.17 The database of 33 mosaic images. These images are labelled from 01 to 33 starting from the upper left image in a raster scan manner.

In order to evaluate only the contribution of the texture information, the colour component of the CTex segmentation algorithm (see Figure 1.1) was not included in the ASKM clustering technique (see Chapter 6).

The segmentation accuracy of the CTex algorithm (using only texture features) is estimated by calculating the Probabilistic Rand index (PR) (see Appendix A). As already mentioned in Chapter 4, the PR index performs a comparison between the obtained segmentation result and the ground truth segmentation by evaluating the pairwise relationships between pixels.

The construction of the texture vectors is illustrated in Figure 5.18. It can be noticed that the feature vectors are defined either by the LPB/C joint distributions (size 256 + *bins*, where 256 is the size of the LBP distribution and *bins* is the quantisation level of the contrast measure) or by the joint

distributions calculated from the responses obtained after filtering the image with the multi-channel filter banks (the intensity values of the filtered images were normalised in the interval [0, 255] so the size of the feature vector is 256×number of filters in the filter bank). For the proposed edge orientation technique, the texture vector is defined by three distributions, namely the dominant orientation (Θ_d), the contrast value (C) and the orientation coherence (Θ_c).

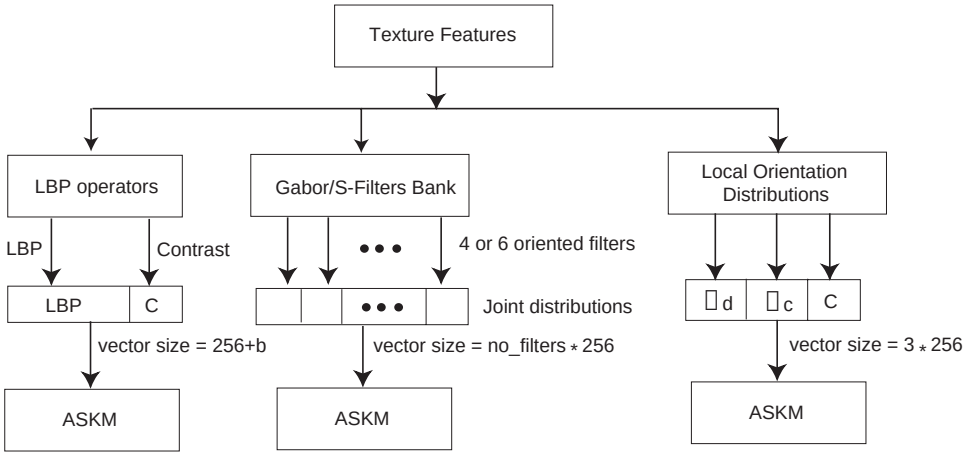


Figure 5.18 Outline of the texture segmentation process.

5.6.2 Results Returned by the LBP Technique

The first set of tests evaluates the accuracy of the CTex algorithm when used in conjunction with the standard Local Binary Pattern (LBP) and the rotation invariant $LBP^{ri}_{8,1}$, $LBP^{ri}_{16,2}$, $LBP^{ri}_{24,3}$ texture descriptors. As indicated in section 5.6.1, the experiments were performed on a database consisting of mosaic images and the numerical results are illustrated in Table 5.1 (the *bins* parameter indicates the quantisation level for the contrast measure).

TABLE 5.1

Quantitative results when the LBP-based texture descriptors were evaluated in the CTex segmentation framework (texture only).

Method	PR $Index_{mean}$	PR $Index_{standard\_deviation}$
LBP/C Bins = 8	0.84	0.12
$LBP^{ri}_{8,1}$	0.80	0.11
$LBP^{ri}_{16,2}$	0.82	0.09
$LBP^{ri}_{24,3}$	0.82	0.12

5.6.3 Results Returned by the Gabor Filtering Technique

In order to evaluate the multi-channel texture decomposition scheme based on Gabor filtering, the input image has been filtered with a small bank of filters with four (0^0, 45^0, 90^0, 135^0) and six (0^0, 30^0, 60^0, 90^0, 120^0, 150^0) orientations. The central frequency and the scale parameters were also varied. The standard deviation (scale) parameter was set to the values 1.0, 2.0 and 3.0 respectively, while the central frequency parameter was varied by setting it to the following values $1.5/2\pi$, $2.0/2\pi$ and $2.5/2\pi$, respectively.

The experimental tests were conducted on the mosaic database depicted in Figure 5.17 and the mean segmentation errors and the corresponding standard deviations of the PR values are depicted in Table 5.2.

TABLE 5.2

Quantitative results when the Gabor filtering (GF) technique was evaluated in the CTex segmentation framework (texture only).

Scale (σ)	Method	PR $Index_{mean}$	PR $Index_{standard\_deviation}$
σ = 1.0	GF f= 1.5/2π, 4 angles	0.46	0.24
	GF f= 2.0/2π, 4 angles	0.61	0.17
	GF f= 2.5/2π, 4 angles	0.81	0.12
	GF f= 1.5/2π, 6 angles	0.50	0.26
	GF f= 2.0/2π, 6 angles	0.62	0.18
	GF f= 2.5/2π, 6 angles	0.81	0.12
σ = 2.0	GF f= 1.5/2π, 4 angles	0.65	0.17
	GF f= 2.0/2π, 4 angles	0.83	0.10
	GF f= 2.5/2π, 4 angles	0.85	0.08
	GF f= 1.5/2π, 6 angles	0.65	0.17
	GF f= 2.0/2π, 6 angles	0.84	0.09
	GF f= 2.5/2π, 6 angles	0.85	0.08
σ = 3.0	GF f= 1.5/2π, 4 angles	0.78	0.13
	GF f= 2.0/2π, 4 angles	0.85	0.08
	GF f= 2.5/2π, 4 angles	0.85	0.11
	GF f= 1.5/2π, 6 angles	0.79	0.12
	GF f= 2.0/2π, 6 angles	0.84	0.08
	GF f= 2.5/2π, 6 angles	0.86	0.08

5.6.4 Results Returned by the S-Filtering Technique

The 2D S-Filters are constructed using equation (5.9) where the scale σ and frequency τ are the parameters that model the spectral sensitivity of these filters. In the original implementation [108], the filter set has been created by varying the scale σ between 1 and 10 and frequency τ between 0 and 4. Following this idea, for these experiments a filter set using the following pairs for (σ, τ) has been implemented: (1,0), (2,1), (4,1), (6,2), (8,3), (10,4). The PR index mean and standard deviation values are shown in Table 5.3.

TABLE 5.3

Quantitative results when the S-Filtering technique was evaluated in the CTex segmentation framework (texture only).

Method	PR $Index_{mean}$	PR $Index_{standard\_deviation}$
S-Filtering	0.83	0.07

5.6.5 Results Returned by the Local Orientation-based Distributions Texture Descriptor

In [105, 113] a texture descriptor based on the evaluation of the dominant image orientation calculated at micro and macro-level is proposed. In this section, experimental results that quantify the performance of the image orientation based texture descriptor in the segmentation process are provided.

TABLE 5.4

Quantitative results for the local orientation based texture extraction technique when the window size is varied.

Scale (σ)	Window size	PR $Index_{mean}$	PR $Index_{standard\_deviation}$
$\sigma = 0.5$	3×3	0.83	0.12
	7×7	0.82	0.11
	11×11	0.82	0.12
$\sigma = 1.0$	3×3	0.81	0.12
	7×7	0.81	0.12
	11×11	0.81	0.11

For these experiments the value of the parameter σ (that sets the scale of the derivative of the Gaussian function) is set to 0.5 and 1.0. The experimental results illustrated in Table 5.4 indicate that the optimal results are obtained when the scale parameter σ is set to 0.5. These results are in agreement with those reported by Ilea et al [105]. There are two reasons behind the selection of

this value for the σ parameter. The first is motivated by the fact that with the increase in the value of the scale parameter the edges derived from weak textures are eliminated and the second reason consists in the requirement to increase the size of the derivative of the Gaussian filters with the increase of the scale parameter σ (the convolution with large kernels generates windowing errors when the partial derivatives are calculated).

The feature vectors for the edge orientation technique is formed by three distributions (see Figure 5.18) namely, the dominant orientation, the contrast and the orientation coherence. The experiments were conducted on a mosaic database when the size of the texture unit $k \times k$ is varied. Quantitative results are shown in Table 5.4.

The last experiment was conducted to evaluate whether the sampling of the texture orientation at different resolutions improves the overall segmentation results. Thus, the texture orientation is sampled by the dominant orientation, contrast and orientation coherence distributions that are calculated for texture units with differing sizes in the range 3×3 to 11×11. The experimental results are shown in Table 5.5.

TABLE 5.5

Quantitative results when the multi-resolution local texture orientation image descriptors were evaluated in the CTex segmentation framework.

Scale (σ)	Window size	PR Index$_{mean}$	PR Index$_{standard_deviation}$
σ = 0.5	3×3 + 7×7	0.82	0.12
	3×3 + 11×11	0.83	0.11
	7×7 + 11×11	0.82	0.12

5.6.6 Discussion on the Reported Results

The results illustrated in Table 5.1 show that that the LBP/C operator provides better discrimination in its standard form than the rotation invariant $LBP^{ri}_{8,1}$, $LBP^{ri}_{16,2}$, $LBP^{ri}_{24,3}$ descriptors. The LBP/C operator returned the highest PR values for 21 out of 33 mosaic images, while the $LBP^{ri}_{8,1}$ operator

returned the lowest PR values for 13 images out of 33. The drop in segmentation accuracy for rotation invariant LBP descriptors is caused by the relative weak discriminative power offered by the uniform patterns in sampling the texture characteristics.

The results depicted in Table 5.2 indicate that the best results are obtained when the texture features are extracted using Gabor filter banks with six orientations and the central frequency is set to large values. These results are justified since by lowering the value of the central frequency the high frequency components from the texture spectrum are filtered out and this process is somewhat similar to the uniformity concept enforced in the calculations of the rotation invariant LBP distributions. This confirms that the local orientation is an important texture property that is better sampled when the central frequency of the Gabor filters is set to large values. This is motivated by the fact that with the increase of the central frequency, the Gabor filters approximate oriented operators.

The S-Filter bank is formed by a set of rotational invariant filters and they are useful in sampling the isotropic texture characteristics at different scales. The segmentation results shown in Table 5.3 indicate that the performance of this rotational invariant texture analysis technique is similar to that offered by the rotational invariant LBP/C descriptors.

The most important finding resulting from the experimental data illustrated in Tables 5.4 and 5.5 is that the distribution of the dominant orientations that are calculated at micro-level for all texture units in the image is an appropriate measure for texture description. The aim of the experiments conducted in section 5.6.5 was to evaluate the performance of the local orientation-based texture descriptor at different resolutions and to evaluate whether the implementation of multi-resolution descriptors leads to the development of improved texture descriptors. The experimental data indicates that optimal performance is obtained when the texture orientation is sampled in small texture units and these results are motivated by the fact that the texture orientation is best analysed at micro-level. The overall performance of the orientation-based texture analysis technique closely matches the performance achieved by the standard LBP/C texture descriptor.

5.7 Conclusions

The aim of this chapter was to evaluate the performance of a number of statistical and signal processing texture analysis techniques when applied to image segmentation. The techniques evaluated in this chapter are: the LBP/C operators, multi-channel texture decomposition based on filtering the image with Gabor and symmetrical filter banks and a new texture analysis technique based on the evaluation of the image orientation at local level. The main novelty associated with this work resides in the evaluation of the dominant orientation in the image at different resolutions. The experimental data reinforced the concept that texture is an important attribute of digital images and it also indicates that the local orientation is the dominant texture feature.

The selection of the texture descriptors discussed and evaluated in this chapter is justified since they analyse the texture at micro-level in small neighbourhoods while the distributions of the texture features calculated from texture units describe the texture at macro-level. This approach provides an appropriate framework to analyse and discriminate between textures present in the image and furthermore the distributions of texture features offer a facile integration with colour features using statistical clustering techniques. In the next chapter the integration of the colour and texture features using an Adaptive Spatial K-Means Clustering (ASKM) technique will be detailed.

Chapter 6

Colour and Texture Integration

The major contribution of this thesis is a novel methodology that adaptively integrates the colour and texture features in a compound mathematical descriptor with the aim of identifying the homogeneous regions in the image. The integration step is performed by a novel Adaptive Spatial K-Means (ASKM) clustering algorithm that enforces the spatial continuity during the data assignment process.

An outline of the proposed **C**olour-**Tex**ture segmentation framework (CTex) is illustrated In Figure 6.1. The colour and texture attributes are independently extracted on two channels. Thus, the colour segmented image is obtained using the multi-space colour segmentation algorithm described in Chapter 4, while the texture features are extracted from the luminance component of the input image. In the example presented in Figure 6.1 the texture information is extracted using a multi-channel texture decomposition technique based on Gabor filters with four orientations. As illustrated in Figure 6.1, the ASKM technique is the strongest component of the CTex framework and performs the integration of the two extracted features. The ASKM colour-texture integration algorithm will be further described in this chapter.

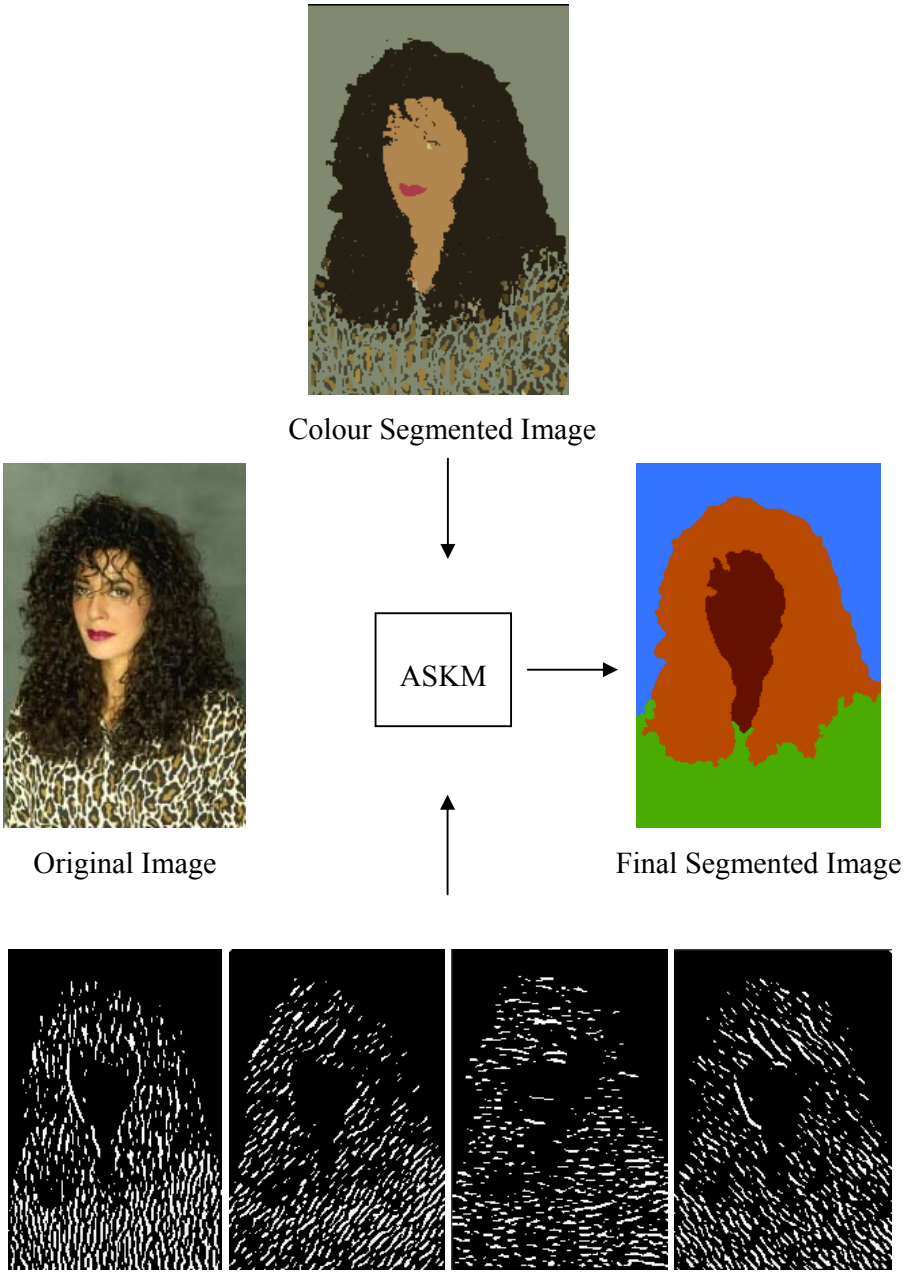


Figure 6.1 An outline of the CTex segmentation framework. The extracted colour and texture features are integrated by the Adaptive Spatial K-Means (ASKM) algorithm in order to obtain the final segmented image.

6.1 Adaptive Spatial K-Means Clustering Algorithm

The inclusion of the texture and colour features in an adaptive fashion is a difficult task [98, 114] since these attributes are not constant within the image. Thus, the application of standard clustering techniques to complex data such as natural images leads to over-segmentation since the spatial continuity is not enforced during the space partitioning process. The main aim of this chapter is to describe the development of a space-partitioning algorithm that is able to return meaningful results even when applied to complex natural scenes that exhibit large variations in colour and texture. To achieve this objective, a new clustering strategy called Adaptive Spatial K-Means (ASKM) whose implementation can be viewed as a generalisation of the K-Means algorithm is proposed. The ASKM technique attempts to minimise the errors in the assignment of the data-points into clusters by adaptively sampling the local texture continuity and the local colour smoothness in the image. The inputs of the ASKM algorithm are: the colour segmented image, the texture images and the final number of clusters k that has been established using the SOM based procedure (see section 4.3.1). The main idea behind ASKM is to minimise an objective function J based on the fitting between the local colour and local texture distributions calculated for each data-point (pixel) in the image and the colour and texture distributions calculated for each cluster. This approach is motivated by the fact that the colour-texture distribution enforces the spatial continuity in the data partitioning process since the colour and texture information are evaluated in a local neighbourhood for all pixels in the image. The local colour distribution for the data-point at location (x, y) is calculated as follows:

$$H_C^{s\times s}(x,y) = \bigcup_{b\in[1,k]} h_C^{s\times s}(x,y,b), \text{where}$$

$$h_C^{s\times s}(x,y,b) = \sum_{p=\left(x-\frac{s}{2}\right)}^{x+\frac{s}{2}} \sum_{q=\left(y-\frac{s}{2}\right)}^{y+\frac{s}{2}} \delta\big(C(p,q),b\big) \text{ and} \tag{6.1}$$

$$\delta(i,j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

In equation (6.1) $H_C^{s\times s}(x,y)$ is the local colour distribution calculated from the colour segmented image C in the neighbourhood of size $s\times s$ around the data-point at position (x, y) and k is the number of clusters. In equation (6.1) the union operator $\cup$ defines the concatenation of the individual histogram bins $h_C^{s\times s}(x,y,b)$, ($b\in[1,k]$) that are calculated from the colour segmented image C.

The local texture distribution $H_T^{s\times s}(x,y)$ is obtained by concatenating the distributions $H_{T_j}^{s\times s}(x,y)$ as follows:

$$\begin{aligned} H_{T_j}^{s\times s}(x,y) &= \bigcup_{b\in[0,255]} h_{T_j}^{s\times s}(x,y,b), \\ h_{T_j}^{s\times s}(x,y,b) &= \sum_{p=\left(x-\frac{s}{2}\right)}^{x+\frac{s}{2}} \sum_{q=\left(y-\frac{s}{2}\right)}^{y+\frac{s}{2}} \delta\left[T_j(p,q),b\right] \end{aligned} \tag{6.2}$$

$$H_T^{s\times s}(x,y) = \bigcup_{j\in[1,\alpha]} H_{T_j}^{s\times s}(x,y) = [H_{T_1}^{s\times s}, H_{T_2}^{s\times s}, ..., H_{T_\alpha}^{s\times s}] \tag{6.3}$$

where T_j is the j^{th} texture image. It is important to note that the pixel values of the texture images T_j are normalised in the range [0, 255].

In order to accommodate the colour-texture distributions in the clustering process, the global objective function of the standard K-Means algorithm is replaced with the formulation shown in equation (6.4). The aim of the ASKM algorithm is the minimisation of the objective function J that is composed of two distinct terms that impose the local coherence constraints. The first term optimises the fitting between the local colour distribution for the data-point under analysis and the global colour distribution of each cluster, while the second term optimises the fitting between the local texture distributions for the same data-point with the global texture distribution of each cluster.

$$J = \sum_{x=1}^{width} \sum_{y=1}^{height} \left\{ \sum_{i=1}^{k} \left[\min_{s \in [3\times3,\ldots,25\times25]} KS\left(H_C^{s\times s}(x,y), H_C^i\right) + \right.\right. \\ \left.\left. + \min_{s \in [3\times3,\ldots,25\times25]} KS\left(H_T^{s\times s}(x,y), H_T^i\right) \right] \right\} \tag{6.4}$$

In equation (6.4), k is the number of clusters, $s \times s$ defines the size of the local window, $H_C^{s\times s}(x,y)$ and $H_T^{s\times s}(x,y)$ are the local colour and the local texture distributions calculated for the pixel at position *(x, y)* respectively and H_C^i and H_T^i are the colour and texture distributions for the cluster with index i respectively. The similarity between the local colour-texture distributions and the global colour-texture distributions of the clusters is evaluated using the Kolmogorov-Smirnov (*KS*) metric:

$$KS\left(H_a, H_b\right) = \sum_{i \in [0, hist\_size]} \left| \frac{h_a(i)}{n_a} - \frac{h_b(i)}{n_b} \right| \tag{6.5}$$

where n_a and n_b are the number of data-points in the distributions H_a and H_b respectively. The main advantage of the *KS* metric over other similarity metrics such as chi-square and the Kullback-Leibler divergence [115] is the fact that the KS metric is normalised and the result of the comparison between the distributions H_a and H_b is bounded in the interval [0, 2] (experimental results using additional similarity metrics are presented in Appendix B of this thesis). The fitting between the local colour-texture distributions and global colour-texture distributions of the clusters is performed adaptively for multiple window sizes in the interval [3×3] to [25×25]. The evaluation of the fitting between the local and global distributions using a multi-resolution approach is motivated by the fact that the composition of the texture in the image is not constant and the algorithm adjusts the window size until it is achieved the best fit value. It is important to note that the global colour-texture distributions H_C^i and H_T^i are updated after each iteration and the algorithm is executed until convergence.

6.2 Post-Processing Re-labelling Procedure

Although the ASKM algorithm returns an accurate partition of the input image into regions with uniform characteristics, it also contains small spurious regions that are caused by the strong transitions in colour information. These undesired regions are eliminated by applying a post-processing step which consists of an iterative re-labelling process that re-assigns these regions to the neighbouring regions based on a colour similarity criterion. The re-labelling procedure can be summarised as follows:

1. Label all regions in the image and generate a list of neighbours for each region.
2. Calculate the mean colour for each region.
3. Calculate the colour similarities between all adjacent regions that are recorded as the Euclidean distance between their mean colours.
4. Re-label the regions smaller than 100 pixels to the neighbouring region with the most similar colour properties and update the colour mean of the region resulting after the merging process.
5. Repeat the steps 3 and 4 until no regions are re-labelled.

6.3 Experiments and Results

The proposed CTex segmentation framework has been tested on more than 300 complex natural images in order to evaluate its performance with respect to the identification of the colour-texture homogenous regions. To achieve this goal, CTex has been applied to Berkeley [99] and McGill [102] natural image databases that include images characterised by non-uniform textures, fuzzy borders and low image contrast. The experiments were conducted in order to obtain quantitative and qualitative evaluations of the performance of the CTex colour-texture segmentation framework with respect to the correct identification of homogeneous regions in the image and the level of image detail. In order to illustrate its validity, the obtained segmentation results were compared against those obtained using the JSEG segmentation algorithm

developed by Deng and Manjunath [49]. JSEG is a standard colour-texture segmentation benchmark and its computational steps were described in the literature review provided in Chapter 2. Although JSEG has a very different computational architecture than CTex, this comparison is appropriate since both algorithms include the colour and texture attributes in the segmentation process. For CTex, the parameters required by the anisotropic diffusion and colour segmentation method were discussed in the previous chapters and were left to the default values in all experiments. In these experiments the JSEG implementation made available online by the authors (http://vision.ece.ucsb.edu/segmentation/jseg/software/) was used. JSEG involves three parameters that have to be specified by the user (the colour quantisation threshold, the scale and the merge threshold) and these parameters have been set to the values suggested by the authors (255, 1.0 and 0.4 respectively).

As demonstrated in the experimental section of Chapter 5, the texture extraction techniques that returned the best results are: the standard LBP/C operator, Gabor filter banks with six orientations and the proposed texture extraction method based on the distribution of the dominant orientations. In the following experiments, the CTex algorithm was independently quantified when the colour information was used in conjunction with each of these three texture descriptors.

While the tests conducted on McGill database only allowed a qualitative evaluation (since no ground truth data is available), the tests performed on the Berkeley database allowed both qualitative and quantitative evaluations since this database provides a set of manual segmentations for each image. The quantitative measurements were carried out using the Probabilistic Rand (PR) index [104]. As indicated in section 4.5, the PR index measures the agreement between the segmented result and the manually generated ground truths and takes values in the range [0, 1], where a higher PR value indicates a better match between the segmented result and the ground truth data.

Table 6.1 depicts the mean and the standard deviation of the PR values that are calculated when the CTex and JSEG algorithms were applied to all 300 images from the Berkeley database. CTex was analysed in conjunction with the

texture features extracted using Gabor filters (CTex - GF), standard LBP/C operator (CTex - LBP/C) and image orientation-based (CTex - IO) texture descriptors. When the texture features are extracted using a Gabor filter bank that samples the texture in six orientations (0^0, 30^0, 60^0, 90^0, 120^0, 150^0), the scale and frequency parameters were set to $\sigma = 3.0$ and $f=2.5/2\pi$ respectively. When CTex is evaluated in conjunction with the standard LBP/C texture descriptor, the contrast measure is quantised in 8 bins. The texture features based on local image orientation distributions were calculated for 3×3 texture units.

In all experiments conducted in this section the colour features are extracted using the algorithm detailed in Chapter 4.

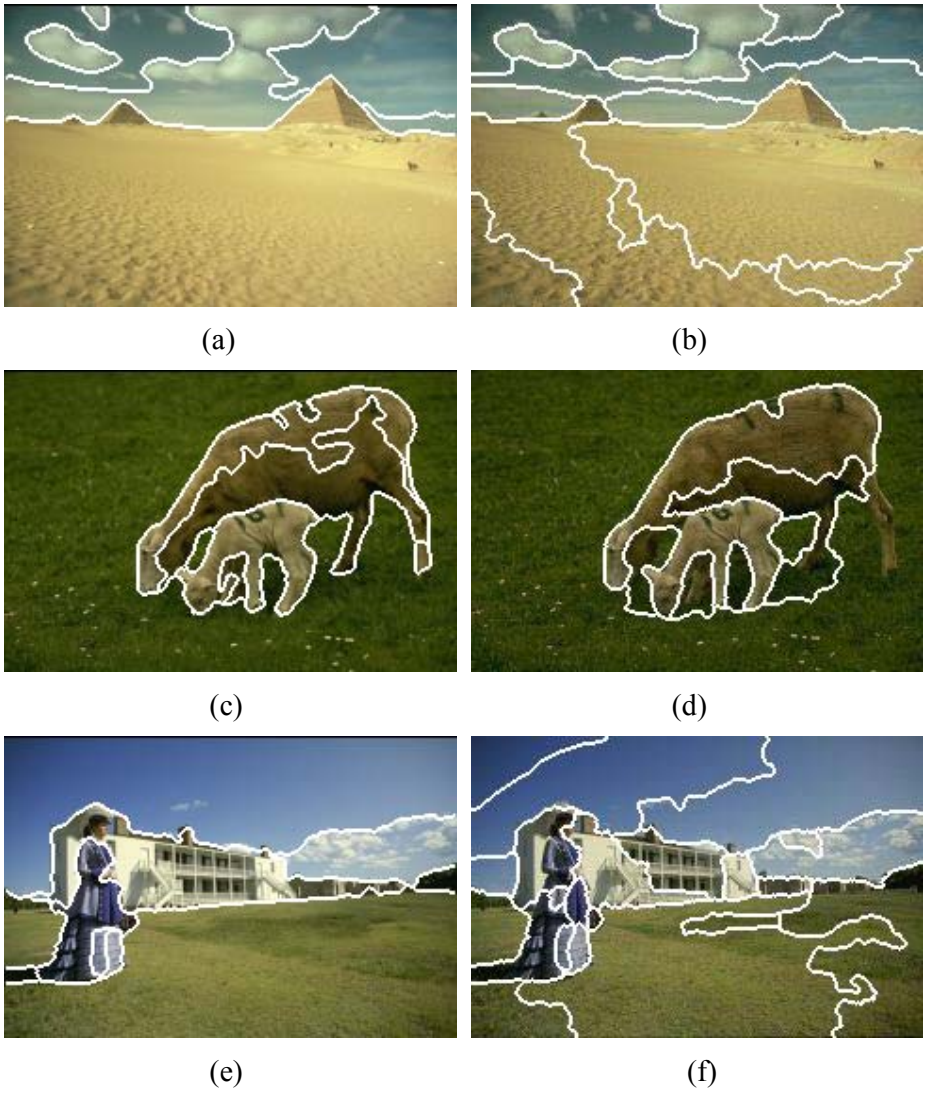
TABLE 6.1

Performance evaluation of the CTex segmentation algorithm when the texture features are extracted using three different techniques. For comparison purposes, the results obtained for JSEG are also included. The CTex and JSEG algorithms were applied to the entire Berkeley database (300 images).

Method	PR $\text{Index}_{\text{mean}}$	PR $\text{Index}_{\text{standard\_deviation}}$
JSEG	0.77	0.12
CTex - GF	0.81	0.10
CTex - LBP/C	0.79	0.11
CTex - IO	0.79	0.11

As shown in Table 6.1, the CTex colour-texture segmentation algorithm achieved the highest PR $\text{Index}_{\text{mean}}$ value of 0.81 when the texture features were extracted using a Gabor filter bank with six orientations (CTex - GF) and the mean value for JSEG is 0.77. The experimental results shown in Table 6.1 indicate that CTex also outperforms JSEG when the texture features are extracted using the LBP/C and local image orientation methods. In the remainder of this section, the CTex technique that returned the best results (when texture features are extracted using a bank of Gabor filters calculated for

six orientations with parameters $\sigma = 3.0$ and $f = 2.5/2\pi$) will be further examined. The segmentation results returned by CTex - GF will be compared against those returned by JSEG.



(a) (b)

(c) (d)

(e) (f)

Figure 6.2 Segmentation of natural images using CTex (first column) and JSEG (second column) algorithms. The recorded PR values are: (a) PR = 0.89. (b) PR = 0.69. (c) PR = 0.94. (d) PR = 0.85. (e) PR = 0.93. (f) PR = 0.79. The segmentation borders for both algorithms were superimposed on the original image.

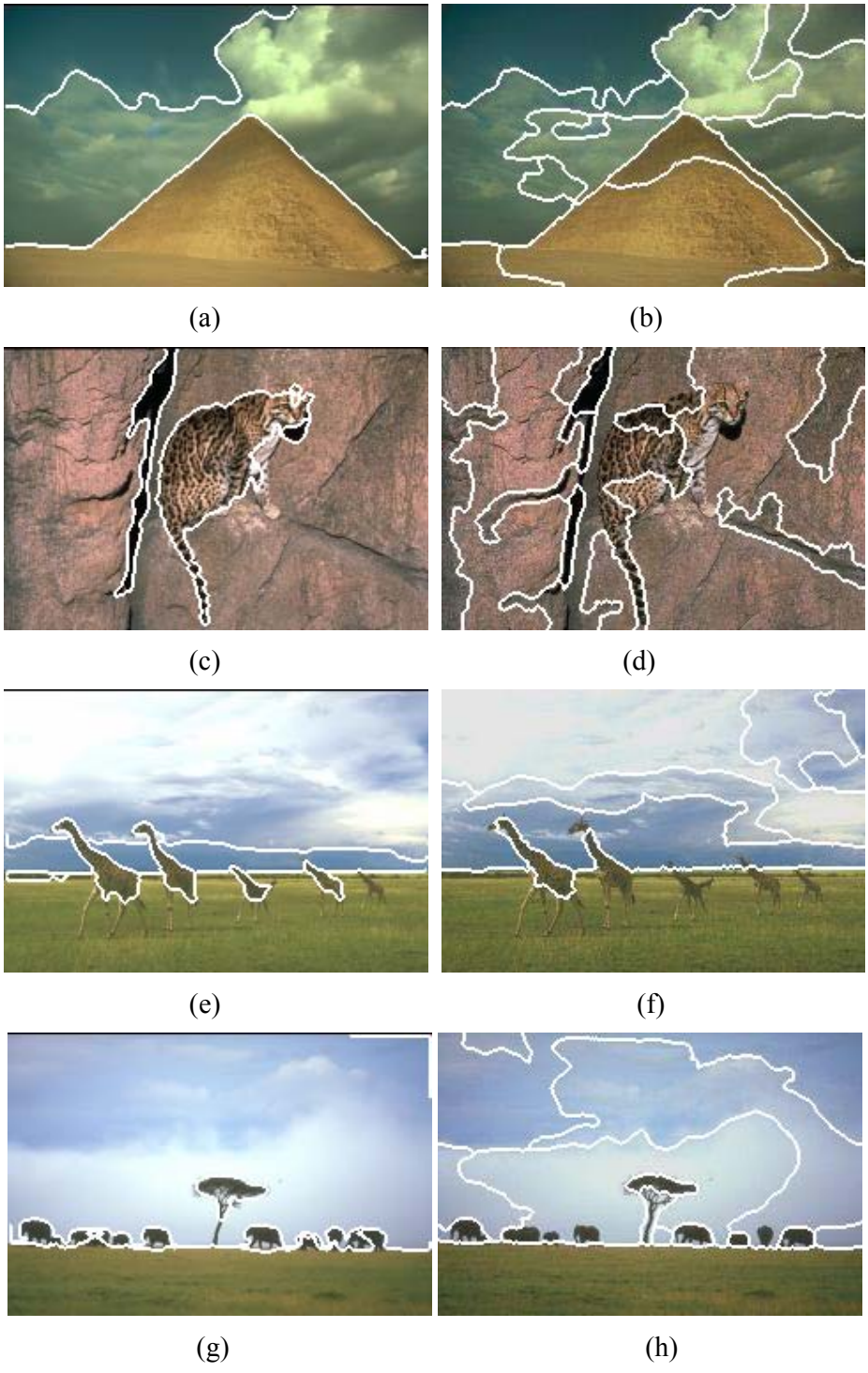


Figure 6.3 Segmentation of natural images using CTex (first column) and JSEG (second column) algorithms. The recorded PR values are: (a) PR = 0.78. (b) PR = 0.72. (c) PR = 0.65. (d) PR = 0.62. (e) PR = 0.88. (f) PR = 0.68. (g) PR = 0.93. (h) PR = 0.66.

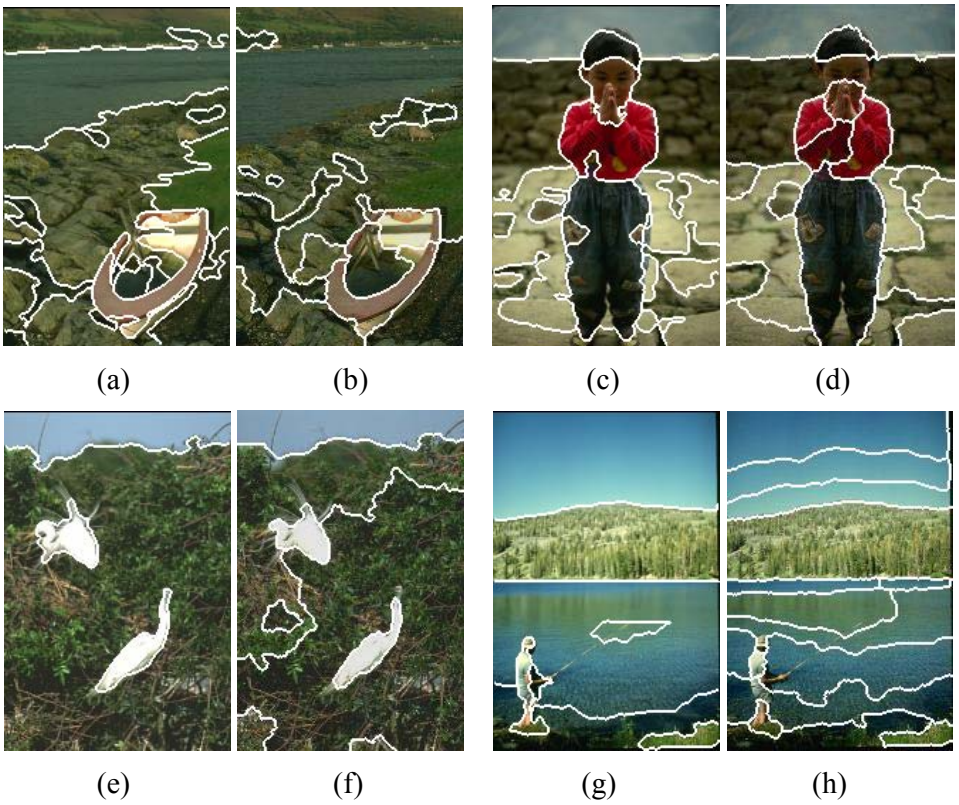


Figure 6.4 Segmentation of natural images using CTex (first and third columns) and JSEG (second and fourth columns) algorithms. The recorded PR values are: (a) PR = 0.80. (b) PR = 0.57. (c) PR = 0.92. (d) PR = 0.90. (e) PR = 0.91. (f) PR = 0.65. (g) PR = 0.91. (h) PR = 0.82.

The relative small difference between the quantitative results recorded by CTex and JSEG (as shown in Table 4.1) is motivated by the fact that the ground truth images from the Berkeley database are in general under-segmented since the manual annotation of these images was performed to highlight only the perceptual uniform regions. This testing scenario favoured the JSEG algorithm while the goal of the CTex framework is to achieve image segmentation at a high level of image detail. The experimental results indicate that the JSEG algorithm performs well in the identification of the image regions defined by similar colour-texture properties, but it fails to accurately determine the object borders between the regions that are characterised by similar colour compositions. This can be observed in Figures 6.2 to 6.4 where a

number of segmentations results achieved after the application of CTex and JSEG on natural images from Berkeley database are illustrated. The PR value recorded by each segmented image is also given. The results depicted in Figures 6.2 to 6.4 indicate that although the overall performance of the JSEG algorithm is good, it has difficulty in the identification of the image regions defined by low colour contrast (see the cat's body in Figure 6.3 (d), the small animals in Figures 6.3 (f) and (h), the land and water in Figure 6.4 (b) or the child's head in Figure 6.4 (d)) and small and narrow details (see the feet regions in Figure 6.2 (d) or the background buildings in Figure 6.2 (f)). JSEG also has the tendency to over-segment the regions defined by similar colour-texture characteristics as observed in Figure 6.2 (b) and (f), in Figure 6.3 (b), (f) and (h) and Figure 6.4 (h). These results also indicate that the CTex technique was able to produce consistent results with respect to the border localisation of the perceptual regions and the level of image detail (see Figures 6.2 (a), (c) and (e) and Figures 6.4 (a), (c) and (g)). These results show that the CTex algorithm is better able to handle the local inhomogeneities in texture and colour than JSEG. Also CTex produced more consistent results when applied to images defined by low colour contrast (see Figures 6.3 (c) and (e) and Figures 6.4 (a) and (c)).

The elimination of the small and narrow image details in the segmented results by the JSEG algorithm is caused by two factors. The first is generated by the fact that the region growing that implements the segmentation process performs the seed expansion based only on the J values that sample the texture complexity rather than a texture model and the spatial continuity is evaluated in relative large neighbourhoods. The second factor that forces the JSEG algorithm to eliminate the small regions from the segmented output is related to the procedure applied to determine the initial seeds for the region growing algorithm. In the original implementation proposed by Deng and Manjunath [49] the initial seeds correspond to minima of local J values and to prevent the algorithm to be trapped in local minima the authors imposed a size criterion for the candidate seed region. In contrast to this approach, the algorithm detailed in this thesis (CTex) evaluates the colour and texture information using explicit models (distributions) and the spatial continuity is enforced during the adaptive

integration of the colour and texture features in the ASKM framework. As illustrated in equation (6.4), the ASKM algorithm is able to adjust the size of the local colour and texture distributions to the image content and this is an important advantage of the CTex algorithm while the level of image detail in the segmented output is preserved. Additional segmentation results of natural images from the Berkeley database are illustrated in Figures 6.5 to 6.7. For clarity purposes, in all segmented images included in this section, the segmentation borders for both algorithms were superimposed on the original image.

Figure 6.5 Segmentation of natural images from Berkeley database using the CTex algorithm.

Figure 6.6 Segmentation of natural images from Berkeley database using the CTex algorithm. These images exhibit complex colour-texture characteristics.

Figure 6.7 Segmentation of natural images from Berkeley database using the CTex algorithm. These images exhibit complex colour-texture characteristics.

It is useful to note that some small erroneous regions are generated by our approach that are caused by the incorrect adaptation of the window size (see equation (6.4)) when dealing with small image regions defined by step transitions in colour and texture. This problem can be addressed by applying a post-processing merging procedure, but this will lead to a reduction in the level of image detail in the segmented data. However, this problem is difficult to tackle taken into consideration the unsupervised nature of the CTex algorithm, but the experimental results indicate that the proposed segmentation framework is robust in determining the main homogeneous regions in complex natural images. One potential solution to address this problem is to include in the ASKM formulation a new regularisation term that penalises the weak

discontinuities between adjacent regions by calculating a global spatial continuity cost.

Due to the large number of distributions that have to be calculated during the ASKM data clustering process and the iterative procedure employed in the optimisation of the ICV_{SOM} parameter, the computational complexity of the CTex algorithm is higher than that associated with JSEG. For instance, CTex segments a natural image of size 240×160 in 140 seconds (texture features extracted using a bank of Gabor filters with six orientations), while JSEG requires approximately 6 seconds, but it is worth mentioning that the implementation of the CTex algorithm has not been optimised with respect to the minimisation of the computational cost. The experiments have been conducted using a 2.4 GHz AMD X2 4600 PC and running Windows XP.

6.4 Conclusions

This chapter presents a novel algorithm called Adaptive Spatial K-Means (ASKM) where the colour and texture features are adaptively integrated into a clustering strategy that enforces the spatial constraints during the assignment of the data into image regions with uniform texture and colour characteristics. The proposed colour-texture composite descriptor represents a major contribution of this research work and proved to be accurate in the identification of the image regions with homogenous characteristics.

An important aim of this chapter was concentrated on the performance evaluation of the CTex colour-texture segmentation algorithm. To achieve this goal, the CTex scheme has been examined on the Berkeley standard natural image database that allowed a comprehensive quantitative and qualitative evaluation. For comparative purposes, the CTex algorithm has been benchmarked against one of the standard colour-texture segmentation schemes (JSEG). The experimental data indicates that CTex outperformed JSEG in both quantitative and qualitative evaluations when the segmentation algorithms examined in this chapter were applied to complex natural images.

Chapter 7

Contributions and Further Directions of Research

The aim of this chapter is to emphasise the novelties and contributions that emerged during the three years' duration of this research programme, and to provide concluding remarks that resulted from this investigation. In the final part of this chapter potential future directions of research will be examined.

7.1 Contributions

Segmentation of natural images is a challenging task because the colour and texture attributes are not uniformly distributed and as a result the process of identifying the homogeneous regions in the image is extremely difficult. Thus, the aim of this thesis was the development of a theoretical framework that adaptively integrates the colour and texture information in the segmentation process. During this investigation a number of major and minor contributions resulted in the areas of feature extraction, data pre-processing, statistical data analysis and feature integration. Also an important objective of this work was to perform a quantitative and qualitative evaluation of the proposed colour-texture segmentation algorithm when applied to annotated natural image databases.

The major contribution of this thesis is the overall colour-texture segmentation (CTex) framework. The framework has been developed in a modular fashion and this allowed quantifying the contribution of each component in the segmentation process. The resulting colour-texture

segmentation scheme is generic and can be used in conjunction with various colour and texture features without any level of supervision.

Other major contributions resulting from this research are in the areas of colour segmentation and adaptive feature integration. In this thesis, a novel colour segmentation algorithm has been proposed where the contributions are located in the analysis of data in a multi-space colour representation and the automatic selection of the initial parameters required by data partitioning algorithms. The key contributions associated with the proposed colour segmentation algorithm can be summarised as follows:

- Evaluation of the input data in a multi-space colour representation.
- Inclusion of a Self Organising Map network to identify the optimal number of clusters and the initial cluster centres.
- Application of a colour saliency measure to reduce the level of over-segmentation.
- Quantitative evaluation of the proposed algorithm when applied to the annotated Berkeley database.

The most important theoretical contribution of this work is located in the integration of the colour and texture features with the aim of producing a mathematical formulation that is able to accurately sample the colour and texture attributes in natural images. The integration of colour-texture features is performed by an adaptive spatial coherent clustering strategy (ASKM). The main idea behind the ASKM is to minimise an objective function based on the fitting between the local colour and texture distributions calculated for each pixel in the image at different resolutions and the colour and texture distributions calculated for each cluster. The experimental data proved that the devised ASKM clustering algorithm is efficient in finding the homogeneous regions in the image that are coherent with respect to the texture and colour information.

Another important contribution of this thesis resides in the development of a new texture analysis method that is based on the evaluation of the image orientation at micro and macro level. The experimental results demonstrated

that the augmentation of the dominant orientation with the orientation coherence and contrast measures significantly improves the discriminative power of the texture descriptor. This new texture descriptor has been evaluated in detail and its performance benchmarked against statistical and signal processing texture analysis techniques.

Minor contributions can be found in the optimisation of colour data pre-processing techniques. Thus, in Chapter 3 the optimisation of the forward and backward (FAB) anisotropic diffusion filtering is presented. The main novelty associated with the proposed filtering strategy is the inclusion of a gradient boosting function whose aim is the improvement of feature preservation around image data defined by low colour contrast.

Additional minor contributions can be found in the implementation of statistical measures that are applied to quantify the similarity between the segmentation results returned by the colour-texture segmentation algorithms and multiple ground truth data.

7.1.1 Summary of the Contributions

Major contributions resulting from this research:

- Overall colour-texture segmentation framework (CTex).
- Multi-space colour segmentation algorithm.
- Automatic initialisation of the clustering parameters.
- Integration of colour and texture features by an adaptive spatial coherent clustering strategy (ASKM).

Publications associated with these contributions:

> D. E. Ilea and P. F. Whelan, "CTex - An Adaptive Unsupervised Segmentation Algorithm Based on Colour-Texture Coherence", *IEEE Transactions on Image Processing*, vol. 17, no. 10, pp. 1926-1939, 2008.
>
> D. E. Ilea and P. F. Whelan, "Colour Image Segmentation Using a Spatial K-Means Clustering Algorithm", *Proceedings of the Irish Machine Vision and Image Processing Conference (IMVIP 2006)*, pp. 146-153, Dublin City University, Ireland, 30 August – 1 September, 2006.

D. E. Ilea and P. F. Whelan, "Automatic Segmentation of Skin Cancer Images Using Adaptive Colour Clustering", *Proceedings of the China-Ireland International Conference on Information and Communications Technologies (CIICT 2006)*, pp. 348-351, Hangzhou, China, 18 -19 October, 2006.

- Development of a new texture descriptor based on the analysis of image orientation at micro and macro level.

Publication associated with this contribution:

D. E. Ilea, O. Ghita and P. F. Whelan, "Evaluation of Local Orientation for Texture Classification", *Proceedings of the 3rd International Conference on Computer Vision Theory and Applications (VISAPP 2008)*, pp. 357-364, Funchal, Madeira – Portugal, 22 - 25 January, 2008.

Minor contributions resulting from this research:

- Colour data pre-processing.
- Application of the colour saliency measure to reduce the level of over-segmentation.
- Evaluation of several non-parametric similarity metrics when used in the context of colour-texture segmentation.
- Implementation of a statistical measure to quantify the segmentation results.

Publications associated with these contributions:

D. E. Ilea and P. F. Whelan, "Adaptive Pre-Filtering Techniques for Colour Image Analysis", *Proceedings of the International Machine Vision and Image Processing Conference (IMVIP 2007)*, pp. 150-157, National University of Ireland, Maynooth, *IEEE* Computer Society Press, 5 -7 September 2007.

D. E. Ilea and P. F. Whelan, "Colour Image Segmentation Using a Self-initialising EM Algorithm", *Proceedings of the International Conference on Visualisation, Imaging and Image Processing (VIIP 2006)*, Palma de Mallorca, Spain, 28-30 August 2006.

7.2 Future Work

While the objectives of this research have been fully accomplished, there is future work that can be carried out based on the ideas advanced in this thesis. The major future directions of research can be focused on a number of areas. For instance, the inclusion of multi-scale texture analysis techniques in the CTex framework could generate a more elaborate texture representation that would help the algorithm to adapt to problems caused by local non-uniformities in the texture information. Also the addition of a regularisation term in the ASKM formulation that penalises the weak discontinuities between adjacent regions during the data partitioning process would help in eliminating the spurious regions caused by uneven illumination and shadows.

Other directions of future research can be focused on improving the computational time associated with the CTex algorithm. The most computationally intensive component of CTex is the ASKM clustering algorithm and the implementation of multi-scale colour and texture models would alleviate the computational time required to perform the fitting between the local and cluster distributions at different resolutions.

Further work may involve the analysis of the proposed CTex framework when applied to the segmentation of medical data such as dermoscopy images, industrial inspection, or included in the development of content based image retrieval systems.

7.3 Conclusions

The aim of this thesis was to present a theoretical framework (referred to as CTex) that includes the colour and texture information in an adaptive fashion in the image segmentation process. As indicated in the literature review detailed in Chapter 2, most of the developed colour-texture segmentation algorithms are reliant on the manual selection of a large number of parameters and as a result they were generally developed in conjunction with particular applications. One of the major objectives of this research work has been the development of a generic unsupervised colour-texture segmentation framework

that is able to accurately partition natural images into regions with coherent colour-texture properties.

Segmentation of natural images is a complex task since the colour and texture attributes are not uniformly distributed within image areas defined by similar properties. To tackle these problems, in this thesis the colour and texture features are calculated on different channels in order to maximise and evaluate their contribution in the segmentation process. Thus, the first step of CTex addresses the optimal extraction of the colour information using a statistical clustering algorithm that evaluates the input data in a multi-space colour representation. While standard clustering algorithms require a high level of user intervention to select the initial conditions such as the number of clusters and the initial cluster centres, in this thesis an automatic procedure based on Self Organising Maps (SOM) has been developed to determine the initial parameters required by a multi-dimensional clustering scheme. The developed colour segmentation scheme has been quantitatively and qualitatively evaluated on standard databases containing natural images and the results have shown that the proposed algorithm is able to accurately partition the image data with respect to the colour content in the image. The experimental results also indicated that colour alone is not sufficient to accurately partition complex natural images and a more complete image descriptor needs to encompass the colour and texture attributes in an adaptive fashion. In this thesis, several texture analysis techniques have been numerically evaluated with the aim of identifying the texture extraction schemes that return optimal results.

The major objective of this research was the development of a theoretical framework that integrates the texture and colour features in a spatial coherent manner. To this end, an adaptive spatially coherent K-Means clustering algorithm has been proposed where the colour-texture distributions are applied to identify the homogeneous regions in the image with respect to the colour and texture information. The developed CTex algorithm has been comprehensively tested on natural images and the experimental results indicate that the proposed algorithm outperforms the JSEG colour-texture segmentation benchmark.

In conclusion, the experimental data has reinforced the concepts presented in this thesis and has demonstrated that the proposed CTex framework offers a generic and robust computational architecture that addresses the colour-texture segmentation of natural images.

Appendix A

The Probabilistic Rand Index

The Probabilistic Rand index (PR) was proposed in [104] with the aim of obtaining a quantitative evaluation of the segmentation result when compared to one or more ground truth (manual) segmentations of the same image.

Let S_{test} be the segmented image that will be compared against the manually labelled set of ground truth images $\{S_1, S_2, \ldots, S_G\}$ (where G defines the total number of manually segmented images). The segmentation result is quantified as appropriate if it correctly identifies the pairwise relationships between the pixels as defined in the ground truth segmentations. In other words, the pairwise labels $l_i^{S_{test}}$ and $l_j^{S_{test}}$ (corresponding to any pair of pixels x_i, x_j in the segmented image S_{test}) are compared against the pairwise labels $l_i^{S_G}$ and $l_j^{S_G}$ in the ground truth segmentations and vice versa. Based on this principle, the PR index is defined as follows:

$$PR(S_{test}, \{S_{1\ldots G}\}) = \frac{1}{\binom{N}{2}} \sum_{\substack{i,j \\ i \neq j}} \left[\mathrm{I}\left(l_i^{S_{test}} = l_j^{S_{test}}\right) \cdot p_{ij} + \mathrm{I}\left(l_i^{S_{test}} \neq l_j^{S_{test}}\right) \cdot (1 - p_{ij}) \right] \quad \text{(a1)}$$

In equation (a1) N is the total number of pixels in the image, $\mathrm{I}\left(l_i^{S_{test}} = l_j^{S_{test}}\right)$ denotes the probability that the pair of pixels x_i and x_j have the same label in S_{test} and p_{ij} represents the mean pixel pair relationship between the ground truth images.

$$p_{ij} = \frac{1}{G} \sum_{g=1}^{G} \mathrm{I}\left(l_i^{S_g} = l_j^{S_g}\right) \quad \text{(a2)}$$

The PR index takes values in the interval [0, 1] and a higher PR value indicates a better match between the segmented result and the ground truth data. The PR index takes the value 0 when there are no similarities between the segmented result and the set of manual segmentations and it takes the value 1 when all segmentations are identical.

Appendix B

Experimental Results Using Additional Similarity Metrics

In Chapter 6.1, the Kolmogorov-Smirnov (KS) metric was used to evaluate the similarity between the colour and texture distributions. In this appendix additional experimental results are presented when the proposed CTex algorithm has been used in conjunction with other similarity metrics such as Kullback-Leibler and χ^2-test.

B.1 Definitions

a) Kolmogorov-Smirnov Similarity Metric (*KS*)

The *KS* metric is a non-parametric test originally proposed in [116] and it is used to evaluate the similarity between two independent distributions H_a and H_b. It is defined as the sum of the absolute differences between two normalised distributions (the number of pixels contained in each bin is divided by the total number of pixels that form the distribution) and it is calculated using the following equation:

$$KS(H_a, H_b) = \sum_{i \in [0, hist\_size]} \left| \frac{h_a(i)}{n_a} - \frac{h_b(i)}{n_b} \right| \quad \text{(b1)}$$

In equation (b1), n_a and n_b represent the number of data-points in the distributions H_a and H_b respectively and $h_a(i)$ and $h_b(i)$ represent the number of

pixels contained by the bin with index i in the H_a and H_b distributions respectively.

b) The Kullback-Leibler Similarity Metric (*KL*)

The *KL* metric measures the difference between two probability distributions H_a and H_b as follows:

$$KL(H_a, H_b) = \sum_{i \in [0, hist\_size]} \frac{h_a(i)}{n_a} \cdot \log\left(\frac{h_a(i)}{h_b(i)} \cdot \frac{n_b}{n_a}\right) \tag{b2}$$

In equation (b2), n_a and n_b represent the number of data points in the distributions H_a and H_b respectively and $h_a(i)$ and $h_b(i)$ represent the number of pixels contained by the bin with index i in the H_a and H_b distributions respectively.

c) The χ^2-test

The χ^2-test is defined by:

$$\chi^2(H_a, H_b) = \sum_{i \in [0, hist\_size]} \frac{\left(\frac{h_a(i)}{n_a} - \frac{h_b(i)}{n_b}\right)^2}{\frac{h_b(i)}{n_b}} \tag{b3}$$

B.2 Experimental Results

The quantitative measurements were carried out by applying the proposed CTex colour-texture segmentation algorithm to the entire Berkeley natural image database. To numerically evaluate the obtained results, the mean value and standard deviation of the Probabilistic Rand index (PR) were computed. For these tests the CTex algorithm was independently used in combination

with the above discussed similarity measures: KS, KL and χ^2-test. It is important to note that the CTex algorithm was analysed in conjunction with the texture features extracted using Gabor filters that sample the texture in six orientations (0^0, 30^0, 60^0, 90^0, 120^0, 150^0) and the scale and frequency parameters were set to $\sigma = 3.0$ and $f = 2.5/2\pi$ respectively.

The experimental results are shown in Table B1.

TABLE B.1

Performance evaluation of the CTex segmentation algorithm when used in conjunction with different similarity measures: KS, KL and χ^2-test. In these experiments the entire Berkeley database (that contains 300 natural images) was used.

Method	PR $\text{Index}_{\text{mean}}$	PR $\text{Index}_{\text{standard\_deviation}}$
CTex (when used in conjunction with KS metric)	0.81	0.10
CTex (when used in conjunction with KL metric)	0.78	0.12
CTex (when used in conjunction with χ^2-test)	0.80	0.10

As shown in Table 6.1, the CTex colour-texture segmentation algorithm achieved the highest PR index mean value and the lowest standard deviation when the similarity of the colour and texture distributions were calculated using the KS metric.

It is also important to mention that the KS metric involves a lower computational cost than the other two measures analysed in this study. The computational times required by CTex to segment a natural image of size 240×160 are illustrated in Table B.2. The experiments have been conducted using a 2.4 GHz AMD X2 4600 PC and running Windows XP.

TABLE B.2

Computational complexity of the CTex segmentation algorithm when used in conjunction with different similarity metrics: KS, KL and χ^2-statistics.

Method	Computational Time
CTex (when used in conjunction with KS metric)	140 seconds
CTex (when used in conjunction with KL metric)	300 seconds
CTex (when used in conjunction with χ^2-test)	254 seconds

References

[1] M. Tuceryan and A. K. Jain, *Texture Analysis*, Handbook of Pattern Recognition and Computer Vision, C.H. Chen, L.F. Pau and P.S.P Wang (eds.) World Scientific Publishing, 1998.

[2] A. Materka and M. Strzelecki, Texture Analysis Methods – A Review, *Technical Report*, University of Lodz, Cost B11 Report, 1998.

[3] R. M. Haralick, Statistical and Structural Approaches to Texture, *Proceedings of the IEEE*, vol. 67, no. 5, pp. 786-804, 1979.

[4] K. L. Laws, Rapid Texture Identification, *Proceedings of the SPIE Conference on Image Processing for Missile Guidance*, vol. 238, pp. 376-380, 1980.

[5] T. Ojala, M. Pietikainen and T. Maenpaa, Multiresolution Grey-scale and Rotation Invariant Texture Classification with Local Binary Patterns, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, 2002.

[6] J. S. Wezska, C. R. Dyer and A. Rosenfeld, A Comparative Study of Texture Measures for Terrain Classification, *IEEE Transactions on Systems, Man and Cybernetics*, vol. 6, no. 4, pp. 269-285, 1976.

[7] T. Ojala and M. Pietikäinen, Unsupervised Texture Segmentation Using Feature Distributions, *Pattern Recognition*, vol. 32, no. 3, pp. 477-486, 1999.

[8] V. A. Kovalev and M. Petrou, Multidimensional Co-occurrence Matrices for Object Recognition and Matching, *CVGIP: Graphical Model and Image Processing*, vol. 58, no. 3, pp. 187-197, 1996.

[9] I. M. Elfadel and R. W. Picard, Gibbs Random Fields, Co-occurrences and Texture Modelling, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 1, pp. 24-37, 1994.

[10] M. Varma and A. Zisserman, Unifying Statistical Texture Classification Frameworks, *Image and Vision Computing*, vol. 22, no. 14, pp. 1175-1183, 2004.

[11] B. S. Manjunath and R. Chellappa, Unsupervised Texture Segmentation Using Markov Random Fields Models, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 5, pp. 478-482, 1991.

[12] G. C. Cross and A. K. Jain, Markov Random Field Texture Models, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 5, no. 1, pp. 25-39, 1983.

[13] P. F. Whelan and D. Molloy, *Machine Vision Algorithms in Java: Techniques and Implementation*, Springer-Verlag, London, 2001.

[14] M. A. Georgeson, Spatial Fourier Analysis and Human Vision, *Tutorial Essays in Psychology, A Guide to Recent Advances*, *Chapter 2,* N. S. Sutherland (ed.), vol. 2, Lawrence Earlbaum Associates, Hillsdale, N.J., 1979.

[15] M. Unser and M. Eden, Multi-resolution Feature Extraction and Selection for Texture Segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 717-725, 1989.

[16] S. G. Mallat, A Theory of Multi-resolution Signal Decomposition: The Wavelet Representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674-693, 1989.

[17] M. Unser, Texture Classification and Segmentation Using Wavelet Frames, *IEEE Transactions on Image Processing*, vol. 4, no. 11, 1995.

[18] F. Mourougaya, P. Carré and C. Fernandez-Maloigne, Wavelets Texture Features: A New Method for Sub-band Characterisation, *Proceedings of the IEEE International Conference on Image Processing*, vol. 1, pp. 445-448, 2005.

[19] R. Manthalkar, P. K. Biswas and B. N. Chatterji, Rotation and Scale Invariant Texture Features Using Discrete Wavelet Packet Transform, *Pattern Recognition Letters*, vol. 24, no. 14, 2455-2462, 2003.

[20] A. C. Bovik, M. Clark and W. S. Geisler, Multi-channel Texture Analysis Using Localised Spatial Filters, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 55-73, 1990.

[21] A. K. Jain and F. Farrokhnia, Unsupervised Texture Segmentation Using Gabor Filters, *Pattern Recognition*, vol. 24, no. 12, pp. 1167-1186, 1991.

[22] T. Randen and J. H. Husoy, Texture Segmentation Using Filters with Optimised Energy Separation, *IEEE Transactions on Image Processing*, vol. 8, no. 4, pp. 571-582, 1999.

[23] T. Hofmann, J. Puzicha and J. M. Buhmann, Unsupervised Texture Segmentation in a Deterministic Annealing Framework, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 803-818, 1998.

[24] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, 1984.

[25] H. D. Cheng, X. H. Jiang, Y. Sun and J. L. Wang, Colour Image Segmentation: Advances and Prospects, *Pattern Recognition*, vol. 34, no. 12, pp. 2259-2281, 2001.

[26] L. Lucchese and S. K. Mitra, Colour Image segmentation: A State-of-the-art Survey, *Proceedings of the Indian National Science Academy*, vol. 67A, no. 2, pp. 207-221, New Delhi, India, 2001.

[27] W. Skarbek and A. Koschan, Colour Image Segmentation – A Survey, *Technical Report*, University of Berlin, 1994.

[28] R. Ohlander, K. Price and D. R. Reddy, Picture Segmentation Using a Recursive Region Splitting Method, *Computer Graphics and Image Processing*, vol. 8, pp. 313-333, 1978.

[29] D. C. Tseng, Y. F. Li and C. T. Tung, Circular Histogram Thresholding for Colour Image Segmentation, *Proceedings of the 3rd International Conference on Document Analysis and Recognition*, vol. 2, pp. 673-676, Montreal, Canada, 1995.

[30] C. Boukouvalas, J. Kittler, R. Marik and M. Petrou, Colour Grading of Randomly Textured Ceramic Tiles Using Colour Histograms, *IEEE Transactions on Industrial Electronics*, vol. 46, no. 1, pp. 219-226, 1999.

[31] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*, Prentice-Hall, 1998.

[32] J. B. MacQueen, Some Methods for Classification and Analysis of Multivariate Observations, *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, vol. 1, pp. 281-297, 1967.

[33] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.

[34] D. Comaniciu and P. Meer, Mean Shift. A Robust Approach Toward Feature Space Analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603-619, 2002.

[35] J. A. Blimes, A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixed and Hidden Markov Models, *Technical Report*, University of California, Berkeley, TR-97-021, 1998.

[36] D. E. Ilea and P. F. Whelan, Colour Image Segmentation Using a Self-initializing EM Algorithm, *Proceedings of the International Conference on Visualisation, Imaging and Image Processing (VIIP 2006)*, Spain, 28-30 August 2006.

[37] M. A. Figueiredo and A. K. Jain, Unsupervised Learning of Finite Mixture Models, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 3, pp. 381-396, 2002.

[38] B. Kovesi, J. M. Boucher and S. Saoudi, Stochastic K-Means Algorithm for Vector Quantisation, *Pattern Recognition Letters*, Elsevier Science, vol. 22, no. 6, pp. 603-610, 2001.

[39] A. R. Weeks and G. E. Hague, Colour Segmentation in the HSI Colour Space Using the K-Means Algorithm, *Proceedings of the SPIE - Nonlinear Image Processing VIII*, pp. 143-154, San Jose, CA, 1997.

[40] M. Luo, Y. F. Ma and H. J. Zhang, A Spatial Constrained K-Means Approach to Image Segmentation, *Proceedings of the Joint Conference of the 4th International Conference on Information, Communications and Signal Processing and the Fourth Pacific Rim Conference on Multimedia*, vol. 2, pp. 738- 742, Singapore, 2003.

[41] Y. W. Tai, J. Jia and C. K. Tang, Soft Colour Segmentation and its Applications, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 9, 2007.

[42] T. N. Pappas, An Adaptive Clustering Algorithm for Image Segmentation, *IEEE Transactions on Image Processing*, vol. 14, no. 4, pp. 901-914, 1992.

[43] J. Chen, T. N. Pappas, A. Mojsilovic and B. E. Rogowitz, Adaptive Perceptual Colour-Texture Image Segmentation, *IEEE Transactions on Image Processing,* vol. 14, no. 10, pp. 1524-1536, 2005.

[44] A. J. Round, A. W. Duller and P. J. Fish, Colour Segmentation for Lesion Classification, *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2, pp. 582-585, Chicago, 1997.

[45] M. Celenk, Hierarchical Colour Clustering for Segmentation of Textured Images, *Proceedings of the 29th South-Eastern Symposium on System Theory*, pp. 483-487, 1997.

[46] A. Tremeau and N. Borel, Region Growing and Merging Algorithm to Colour Segmentation, *Pattern Recognition*, vol. 30, no. 7, pp. 1191-1203, 1997.

[47] F. Y. Shih and S. Cheng, Automatic Seeded Region Growing for Colour Image Segmentation, *Image and Vision Computing*, vol. 23, no. 10, pp. 877–886, 2005.

[48] H. D. Cheng and Y. Sun, A Hierarchical Approach to Colour Image Segmentation Using Homogeneity, *IEEE Transactions on Image Processing*, vol. 9, no. 12, pp. 2071-2082, 2000.

[49] Y. Deng and B. S. Manjunath, Unsupervised Segmentation of Colour-Texture Regions in Images and Video, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 8, pp. 800-810, 2001.

[50] A. Moghaddamzadeh and N. Bourbakis, A Fuzzy Region Growing Approach for Segmentation of Colour Images, *Pattern Recognition*, vol. 30, no. 6, pp. 867-881, 1997.

[51] G. Healey, Using Colour for Geometry-Insensitive Segmentation, *Optical Society of America*, vol. 22, no. 1, pp. 920-937, 1989.

[52] G. Healey, Segmenting Images Using Normalised Colour, *IEEE Transactions on Systems, Man and Cybernetics*, vol. 22, no. 1, pp. 64-73, 1992.

[53] D. K. Panjwani and G. Healey, Markov Random Fields Models for Unsupervised Segmentation of Textured Colour Images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 10, pp. 939-954, 1995.

[54] G. Paschos and K. P Valavanis, Chromatic Measures for Colour Texture Description and Analysis, *Proceedings of the IEEE International Symposium on Intelligent Control*, pp. 319 – 325, 1995.

[55] T. S. C. Tan and J. Kittler, Colour Texture Analysis Using Colour Histogram, *IEEE Proceedings of Vision, Image and Signal Processing*, vol. 141, no. 6, pp.403-412, 1994.

[56] R. Manduchi, Bayesian Fusion of Colour and Texture Segmentations, *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2, pp. 956 – 962, 1999.

[57] M. P. Dubuisson-Jolly and A. Gupta, Colour and Texture Fusion: Application to Aerial Image Segmentation and GIS Updating, *Image Vision Computing*, no. 18, pp. 823–832, 2000.

[58] C. Carson, S. Belongie, H. Greenspan and J. Malik, Blobworld: Image Segmentation Using Expectation-Maximisation and its Application to Image Querying, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1026-1038, 2002.

[59] L. Shafarenko, M. Petrou and J. Kittler, Automatic Watershed Segmentation of Randomly Textured Colour Images, *IEEE Transactions on Image Processing*, vol. 6, no. 11, pp. 1530-1544, 1997.

[60] K. M. Chen and S. Y. Chen, Colour Texture Segmentation Using Feature Distributions, *Pattern Recognition Letters*, vol. 23, no. 7, pp. 755-771, 2002.

[61] P. Nammalwar, O. Ghita and P. F Whelan, Integration of Feature Distributions for Colour Texture Segmentation, *Proceedings of the International Conference on Pattern Recognition*, vol. 1, pp. 716 – 719, 2004.

[62] M. Mirmehdi and M. Petrou, Segmentation of Colour Textures, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 2, pp. 142-159, 2000.

[63] M. A. Hoang, J. M. Geusebroek and A. W. Smeulders, Colour Texture Measurement and Segmentation, *Signal Processing*, vol. 85, no. 2, pp. 265-275, 2005.

[64] Y. G. Wang, J. Yang and Y. C. Chang, Colour-texture Image Segmentation by Integrating Directional Operators into JSEG Method, *Pattern Recognition Letters*, vol. 27, no. 16, pp. 1983-199, 2006.

[65] Y. Wang, J. Yang and Y. Zhou, Colour-texture Segmentation Using JSEG Based on Gaussian Mixture Modelling, *Journal of Systems Engineering and Electronics*, vol. 17, no. 1, pp. 24-29, 2006.

[66] Y. Zheng, J. Yang, Y. Zhou and Y. Wang, Colour-Texture Based Unsupervised Segmentation Using JSEG with Fuzzy Connectedness, *Journal of Systems Engineering and Electronics*, vol. 17, no. 1, pp. 213-219, 2006.

[67] S. Liapis and G. Tziritas, Colour and Texture Image Retrieval Using Chromaticity Histograms and Wavelet Frames, *IEEE Transactions on Multimedia*, vol. 6, no. 5, pp. 676-686, 2004.

[68] L. Shi and B. Funt, Quaternion Colour Texture Segmentation, *Computer Vision and Image Understanding*, vol. 107, no. 1-2, pp. 88-96, 2007.

[69] J. Freixenet, X. Muñoz, J. Martí and X. Lladó, Colour Texture Segmentation by Region-Boundary Cooperation, *Proceedings of the European Conference on Computer Vision*, vol. 2, pp. 250-261, 2004.

[70] M. Saïd Allili and D. Ziou, Globally Adaptive Region Information for Automatic Colour–Texture Image Segmentation, *Pattern Recognition Letters*, vol. 28, no. 15, pp. 1946-1956, 2007.

[71] R. Luis-García, R. Deriche and C. Alberola-López, Texture and Colour Segmentation Based on the Combined Use of the Structure Tensor and the Image Components, *Signal Processing*, vol. 88, no. 4, pp. 776-795, 2008.

[72] M. Sonka, V. Hlavac and R. Boyle, *Image Processing, Analysis and Machine Vision,* 2nd edition, PWS Boston, 1998.

[73] C. Vertan, M. Zamfir, E. Zaharescu, V. Buzuloiu and C. Fernandez-Maloigne, Non-linear Colour Image Filtering by Colour to Planar Shape Mapping, *Proceedings of the IEEE International Conference on Image Processing,* vol. 1, pp. 885-888, 2003.

[74] B. Smolka, R. Lukac, A. Chydzinski, K. N. Plataniotis and K. Wojciechovski, Fast Adaptive Similarity Based Impulse Noise Reduction Filter, *Real Time Imaging*, vol. 9, no. 4, pp. 261-276, 2003.

[75] K. Tang, J. Astola and Y. Neuovo, Non-linear Multivariate Image Filtering Techniques, *IEEE Transactions on Image Processing*, vol. 4, no. 6, pp. 788-797, 1995.

[76] P. Perona and J. Malik, Scale-space and Edge Detection Using Anisotropic Diffusion, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629-639, 1990.

[77] N. Nordstrom, Biased Anisotropic Diffusion. A Unified Regularisation and Diffusion Approach to Edge Detection, *Image and Vision Computing*, vol. 8, no. 4, pp. 18-27, 1990.

[78] J. Weickert, B. M. ter Haar Romeny and M. A. Viergever, Efficient and Reliable Schemes for Non-linear Diffusion Filtering, *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 398-410, 1998.

[79] M. J. Black, G. Sapiro, D. H. Marimont and D. Heeger, Robust Anisotropic Diffusion, *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 421-432, 1998.

[80] O. Ghita, K. Robinson, M. Lynch and P. F. Whelan, MRI Diffusion-Based Filtering: A Note on Performance Characterisation, *Computerised Medical Imaging and Graphics*, vol. 29, no. 4, pp. 267-277, 2005.

[81] A. Maalouf, P. Carré, B. Augereau and C. Fernandez-Maloigne, Bandelet-Based Anisotropic Diffusion, *Proceedings of the IEEE International Conference on Image Processing*, vol. 1, pp. 289-292, 2007.

[82] C. Tomasi and R. Manduchi, Bilateral Filtering for Gray and Colour Images, *Proceedings of the IEEE International Conference on Computer Vision*, pp. 839-846, India, 1998.

[83] G. Gilboa, N. Sochen and Y. Y. Zeevi, Forward-and-Backward Diffusion Processes for Adaptive Image Enhancement and Denoising, *IEEE Transactions on Image Processing*, vol. 11, no. 7, pp. 689-703, 2002.

[84] B. Smolka and K. N. Plataniotis, On the Coupled Forward and Backward Anisotropic Diffusion Scheme for Colour Image Enhancement, *Lecture Notes in Computer Science*, *Springer Verlag,* vol. 2383, pp. 70-80, 2002.

[85] D. E. Ilea and P. F. Whelan, Adaptive Pre-Filtering Techniques for Colour Image Analysis, *Proceedings of the International Machine Vision & Image Processing Conference (IMVIP 2007)*, pp. 150-157, National University of Ireland, Maynooth, *IEEE* Computer Society Press, 5 -7 September 2007.

[86] E. W. Forgy, Cluster Analysis of Multivariate Data: Efficiency vs. Interpretability of Classifications, *Biometrics*, vol. 21, pp. 768-769, 1965.

[87] U. M. Fayyad, C. A. Renia and P. S. Bradley, Initialisation of Iterative Refinement Clustering Algorithms, *Proceedings of the International Conference of Knowledge Discovery and Data Mining*, pp. 194-198, 1998.

[88] J. F. Lu, J. B. Tang, Z. M. Tang and J. Y. Yang, Hierarchical Initialisation Approach for K-Means Clustering, *Pattern Recognition Letters*, no. 29, no. 6, pp. 787-795, 2008.

[89] J. M. Pena, J. A. Lozano and P. Larranaga, An Empirical Comparison of Four Initialisation Methods for the K-Means Algorithm, *Pattern Recognition Letters*, vol. 20, no. 10, pp. 1027-1040, 1999.

[90] L. Kaufmann and P. J. Rousseeuw, *Finding Groups in Data. An Introduction to Cluster Analysis*, John Wiley & Sons, Inc., Canada, 1990.

[91] T. Su and J. Dy, A Deterministic Method for Initializing K-Means Clustering*, Proceedings of the IEEE International Conference on Tools with Artificial Intelligence*, pp. 784-786, 2004.

[92] S. Khan and A. Ahmad, Cluster Centre Initialisation Algorithm for K-Means Clustering, *Pattern Recognition Letters*, vol. 25, no. 11, pp. 1293-1302, 2004.

[93] D. W. Kim, K. H. Lee and D. Lee, A Novel Initialisation Scheme For the Fuzzy C-Means Algorithm for Colour Clustering, *Pattern Recognition Letters*, vol. 25, no. 2, pp. 227-237, 2004.

[94] T. Kohonen, *Self-Organising Maps*, Berlin Heidelberg, New York, 3rd Edition Springer Verlang, 2001.

[95] G. Dong and M. Xie, Colour Clustering and Learning for Image Segmentation Based on Neural Networks, *IEEE Transactions on Neural Networks*, vol. 16, no. 4, pp. 925-936, 2005.

[96] J. Vesanto and E. Alhoniemi, Clustering of the Self-Organizing Map, *IEEE Transactions on Neural Networks*, vol. 11, no. 3, 2000.

[97] S. H. Ong, N. C. Yeo, K. H. Lee, Y. V. Venkatesh and D. M. Cao, Segmentation of Colour Images Using a Two-stage Self-organising Network, *Image and Vision Computing*, vol. 20, no. 4, pp. 279-289, 2002.

[98] D. E. Ilea and P. F. Whelan, Automatic Segmentation of Skin Cancer Images Using Adaptive Colour Clustering, *Proceedings of the China-Ireland International Conference on Information and Communications Technologies*, pp. 348-351, Hangzhou, China, October 18th -19th, 2006.

[99] D. Martin, C. Fowlkes, D. Tal and J. Malik, A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics, *Proceedings of the IEEE International Conference on Computer Vision*, pp. 416-425, 2001.

[100] D. E. Ilea and P. F. Whelan, CTex - An Adaptive Unsupervised Segmentation Algorithm Based on Colour-Texture Coherence, *IEEE Transactions on Image Processing*, vol.17, no. 10, pp.1926-1939, 2008.

[101] E. J. Giorgianni, T. E. Madden, *Digital Color Management: Encoding Solutions*, Addison-Wesley, 1998.

[102] A. Olmos and F. A. A. Kingdom, McGill Calibrated Colour Image Database, http://tabby.vision.mcgill.ca, 2004.

[103] G. Heidemann, Region Saliency as a Measure for Colour Segmentation Stability, *Image and Vision Computing*, vol. 26, no. 2, pp. 211-227, 2008.

[104] R. Unnikrishnan and M. Hebert, Measures of Similarity, *Proceedings of IEEE Workshop on Computer Vision Applications*, vol. 1, pp. 394 – 394, 2005.

[105] D. E. Ilea, O. Ghita and P. F. Whelan, Evaluation of Local Orientation for Texture Classification, *Proceedings of the 3rd International Conference on Computer Vision Theory and Applications (VISAPP 2008)*, pp. 357-364, Funchal, Madeira – Portugal, 22 - 25 January, 2008.

[106] Vision Texture (VisTex) Database, Massachusetts Institute of Technology, MediaLab. Available online at: http://vismod.media.mit.edu/vismod/imagery/VisionTexture/vistex.html

[107] J. G. Daugman, Complete Discrete 2D Gabor Transforms by Neural Networks for Image Analysis and Compression, *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 36, no. 7, pp. 1169-1179, 1988.

[108] C. Schmid, Constructing Models for Content Based Image Retrieval, *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 39-45, 2001.

[109] M. Kass and A. Witkin, Analysing Oriented Patterns, *Computer Vision, Graphics, and Image Processing*, vol. 37, no. 3, pp. 362-385, 1987.

[110] J. Canny, A Computational Approach to Edge Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679-698, 1986.

[111] T. Ojala, T. Maenpäa, M. Pietikainen, J. Viertola, J. Kyllonen and S. Huovinen, Outex – A New Framework for Empirical Evaluation of Texture Analysis Algorithms, *Proceedings of the International Conference on Pattern Recognition*, vol. 1, pp. 701–706, 2002.

[112] C. Germain, J. P. Da Costa, O. Lavialle and P. Baylou, Multiscale Estimation of Vector Field Anisotropy Application to Texture Characterisation, *Signal Processing*, vol. 83, pp. 1487-1503, 2003.

[113] O. Ghita, P. F. Whelan and D. E. Ilea, Multi-resolution Texture Classification Based on Local Image Orientation, *Proceedings of the International Conference on Image Analysis and Recognition (ICIAR 2008)*, pp. 688-696, Póvoa de Varzim, Portugal, 25-27 July, 2008.

[114] D. E. Ilea and P. F. Whelan, Colour Image Segmentation Using a Spatial K-Means Clustering Algorithm, *Proceedings of the Irish Machine Vision & Image Processing Conference (IMVIP 2006)*, pp. 146-153, Dublin City University, Ireland, 30 August – 1 September, 2006.

[115] Y. Rubner, J. Puzicha, C. Tomasi and J. M. Buhmann, Empirical Evaluation of Dissimilarity Measures for Colour and Texture, *Computer Vision and Image Understanding*, vol. 84, no. 1, pp. 25-43, 2001.

[116] D. Geman, S. Geman, C. Graffigne and P. Dong, Boundary Detection by Constrained Optimisation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 609-628,1990.

[117] S. Watanabe, Karhunen-Loeve Expansion and Factor Analysis, *Transactions of the 4th Prague Conference on Information Theory*, pp. 635-660, Prague, 1965.