School of Computer Science and Software Engineering
Monash University

Bachelor of Software Engineering Honours (2770),

Clayton Campus

Literature Review — Semester 2, 2005

Bayesian Segmentation of Dermatological Images using Mixture Models and Markov Random Fields

Ji Yong Tran,

ID:18572685

Supervisor: Dr. David Squire

Contents

1 Introduction **1**

2 Image Processing and CBIR for Dermatology **2**
2.1 Introduction 2
2.2 Content-Based Information Retrieval Systems 2
2.3 Dermatoscopy in the Medical Environment 3
2.4 Texture Features 4
2.5 Colour Features 5
2.6 Image Processing for Dermatoscopy Images 6
2.7 Content-Based Information Retrieval Tools 6
2.8 Summary 6

3 Image Segmentation **7**
3.1 Introduction 7
3.2 Clustering 8
3.2.1 K-means Clustering 8
3.2.2 Distance Similarity (Euclidean, Mahalanobis, and pixel location) . . 9
3.2.3 Fuzzy C-means Clustering 10
3.2.4 Gaussian Mixture Models 12
3.3 First Variants of the MRF models 13
3.4 Summary 14

4 Conclusions **14**

1 Introduction

Australia is a country known for its dry and sunny conditions and a high rate of skin cancer. There are many different types of skin cancer, and of the most dangerous of them all is Melanoma (Lagerstrom, 2004). Melanoma is caused from unprotected high exposure to UV radiation. Commonly, people with fair skin or skin that contains lots of freckles and moles are at high risk. Melanoma has claimed the lives of many people. It is estimated annually that over 1000 people in Australia alone die from melanoma, 7700 and 1600 people in the US and the UK respectively. The annual rate of melanoma cases world wide is steadily increasing at around 6 percent. It is an extremely alarming rate (Lagerstrom, 2004).

Fortunately, early detection of this skin cancer can improve the cure rate, and in most cases a 100% cure rate is possible (Lagerstrom, 2004). However, if the melanoma is not detected early and the lesion is more than 3 millimetres deep, the chance of survival is 59%. Most general practitioners have less experience in the full range of melanoma forms, and because of this, many melanoma cases are not diagnosed properly (Schmid-Saugeon, Guillod and Thiran, 2003). Computer technology such as the Content-Based Image Retrieval (CBIR) Systems have been adopted by the medical department to improve the global access to the knowledge of the different forms of melanoma. These tools have aided general practitioners in the detection of melanoma (Müller, Michoux, Bandon and Geissbuhler, 2004).

In recent years, there has been much research and development into the area of skin lesion analysis. The majority of skin lesion analysis involves taking image samples of the suspicious areas from a human body and then further analysing the images for signs of skin cancer. Microscopy and Dermoscopy images are two of the main images commonly used by dermatologists.

One of the most important aspects of studying dermoscopy images is the detection of skin cancer, particularly for signs of the Malignant Melanoma lesion (Schmid-Saugeon, 1999). The fundamental task in the lesion detection is to separate healthy skin from the skin lesion. The traditional pre-processing for the detection of melanoma involves manually drawing borders around the lesion areas on the image. Although it is a simple method, it is time consuming and very subjective as each individual dermatologist has a different level of experience (Guillod, Schmid-Saugeon, Guggisberg, Cerottini, Braun, Krischer, Saurat and Murat, 2002).

Current research methods have been tackling this problem through the use of image segmentation and edge detection, however the processes involved to produce accurate results are still far from perfect. Techniques such as analysing the image contours, colour, and texture for segmentation, while symmetry and the regularity of the lesions for subsequent analysis have been experimented, and yet the methods and processes involved are still lacking overall accuracy (Schmid-Saugeon, 1999; Erkol, Moss, Stanley, Stoecker and Hvatum, 2004). Other new techniques must be considered for the benefit of the current problem. This literature review will address the current work to date.

2 Image Processing and CBIR for Dermatology

2.1 Introduction

As the melanoma risk rate increases, more and more research and development have begun in an attempt to increase the early detection of the skin cancer. Research into the skin lesion segmentation area largely revolves around the digital dermotoscopy images (Schmid-Saugeon, 1999).

Extensive research into digital dermoscopy has been in progress since the last decade. Some of these digital processing techniques are contour detection (Erkol et al., 2004) and image segmentation (Schmid-Saugeon, 1999). Their purpose is to effectively pre-process dermoscopy images and eliminate human subjective measures in the detection of lesions. The visual results achieved by these digital techniques are reasonable, however the error rate incurred by these techniques are not acceptable (Schmid-Saugeon, 1999).

This review will cover various image processing techniques, features, and content-based information retrieval systems (CBIRSs) that may be of use for dermatological images. It is anticipated that the topics that follow will benefit the study of dermatological images, and as a result, CBIRSs will be developed in the future that are specialised for lesion images. Segmentation is particularly important for this application of CBIR, since we know that it is the lesion segment, not the healthy skin, that we care about.

2.2 Content-Based Information Retrieval Systems

As the advances of digital imaging technology increases steadily, more and more digital images are produced. Such a rise in digital images has increase the storage requirements for most computer workstations and image databases. The availability of digital images especially in image databases has given rise to sharing of image databases through various means such as the internet, local machines, and local intranet. Sharing of image databases requires a good retrieval system in order to be effective. Inspired by the original text querying tools, the idea of content-based information retrieval system quickly became a field of rapid research to satisfy the demands of the consumer needs (Smeulders, Worring, Santini, Gupta and Jain, 2000).

CBIRSs are systems that allow a user to search for images based on a query image or a set of textual image constraints, and in return the CBIRS will return a list of matching results from its image database. CBIRSs have been in place since the last decade, and they have been used in various digital imaging technology areas such as surveillance cameras (Huston, Campbell, Sukthankar and Pillai, 2004), facial recognition systems (Abate, Nappi, Ricciardi and Tortora, 2004), and medical radiotherapy images (Lehmann, Schubert, Keysers, Kohnen and Wein, 2003).

Generally, a CBIRS will employ various techniques and features to perform image retrieval tasks (Müller, Pun and Squire, 2004). These techniques include data classification, colour and texture pre-processing, shape analysis, and image segmentation. The features include image indexes, inverted files (Squire, Müller, Müller and Raki, 1999), and different similarity measures. Whilst those techniques are internal to the system, from the users' perspective, user interaction via an intuitive graphical user interface (GUI), the presentation of the search results, and relevance feedback to refine the search results are key factors for a good CBIRS (Smeulders et al., 2000).

The sheer volume of patient visual records in the medical domain are ever increasing and CBIR systems are desired to manage these resources since they effectively allow users to store and search for visual content in large databases. Following the review by Müller, Michoux, Bandon and Geissbuhler (2004), these systems are potentially possible for clinical decision-making, medical research, and teaching. Clinical decision-making is central to the clinical experts and can cause reluctance in the acceptance of CBIRSs. The requirements a CBIRS in the medical domain needs to address are to integrate it seamlessly into the current applications with user interfaces that are familiar to the users, optimise the system for various applications to show its usefulness and performance, and finally, these systems should head towards the open-source field, such that existing tools can be reused and more effort can be spent on addressing the needs of the medical domain.

2.3 Dermatoscopy in the Medical Environment

The *ABCD* rule by Stolz, Braun-Falco, Bilek, Landthanler, Conneta and Riemann (1994) is a standard used in dermatoscopy analysis to rate and recognise malignant melanoma. The four letters stand for a particular criteria (Stolz, Braun-Falco, Bilek, Landthaler, Burgdorf and Cognetta, 2002):

- **A**symmetry: The lesion is bisected into two orthogonal axes to determine the asymmetry properties including the colour, texture, and shape characteristics. A score of 0 is given when the two axes don't have asymmetry properties, otherwise 1 or 2 points is given when asymmetry is present in one or two axes. The study found that 96% of malignant melanomas have a border score of 2. (0 - 2 points)

- **B**order: The lesion is divided into eight segments for pigment pattern analysis. Each segment with a sharp or abrupt cut-off pigment pattern receives a score of 1, and the segments with gradual or fuzzy cut-off pigment patterns receives 0 points. Malignant melanoma lesions are generally in the 3 to 8 point range. (0 - 8 points)

- **C**olour: The presence of up to six colours (white, red, light brown, dark brown, blue-gray, black) are considered. Each colour represents a property and characteristic of the lesion. 85% of the time, malignant melanomas have more than three colours. (0-6 points)

- **D**ifferential structures: The patterns and structure of the lesion are considered. These characteristics include network, structureless areas, dots, globules, and streaks. In 73% of the malignant melanoma cases, four or more structures are present. (0-5 points)

The Total Dermatoscopy Score (TDS) is $A * 1.3 + B * 0.1 + C * 0.5 + D * 0.5$, with a maximum score of 8.9. A high *ABCD* score means the lesion is more likely to be a malignant melanoma (> 5.45) (Stolz et al., 2002). As the *ABCD* rule is centered around the properties and characteristics of the lesion, knowing the complete region of the lesion is required for the task. Analysis of the *ABCD* is not possible unless segmentation is used to isolate the lesion.

Dermoscopy images are captured via the use of hand-held surface microscopes and digital cameras. These instruments have a fixed magnification rate of x10. The reflective properties of skin lesions are poor, and liquid such as mineral oil, immersion oil, liquid paraffin, ultrasound gel are often applied on the lesion before examination with the microscope. The liquid reduces the amount of reflected light by use of refractive index matching, and when it makes contact with the microscope lens, the light emitted from the

microscope is able to penetrate deeper into the skin lesion, which results in an improved skin lesion image (Menzies, Crotty, Ingvar and McCarthy, 2003).

A study conducted by Schmid-Saugeon et al. (2003) where dermatologists were involved in drawing lesion borders on skin lesion images showed that dermatologists lacked consistency in their drawings and were often biased with their subjective measures. Comparing the results with an automated image segmentation method, the automated method produced consistent and more accurate results. These studies outline the importance and need for automated lesion detection methods.

2.4 Texture Features

Texture features and extraction is a field of research for many years (Ma and Manjunath, 1996). Many texture features have been implemented in the past. Some of the commonly used features are the Gabor filter responses (Ma and Manjunath, 1996), the wavelet-based decomposition/transformations (Zarita and Lelandais, 1997), second-moment matrix, and multi-resolution filter banks (Carson, Belongie, Greenspan and Malik, 2002). Gabor filters are classified as band pass filters that allow certain frequencies of an image to pass through. It offers an optimal trade-off between location in the spatial and frequency domains. These characteristics of the Gabor filters makes it a useful filter for applications such as texture extraction, texture segmentation, edge detection, image coding and image representation (Weldon, Higgins and Dunn, 1996).

Gabor filter banks are commonly used for extracting texture features (Squire et al., 1999). It is said that spatial localisation of texture boundaries require larger Gabor bandwidths while smaller bandwidths give a better sense of texture measurements (Kato and Pong, 2001). A texture consists of a group of pixels. Fortunately, Gabor filters provide an optimal balance in both cases. Gabor filter masks of arbitrary size are used to extract texture features. These masks are orientation and scale tunable edge and line detectors (Ma and Manjunath, 1996).

The Viper system uses a bank of real circularly symmetric Gabor filters to extract texture features. These filters generate energy values from each 16×16 block of pixels in an image. Energy values are quantised into 10 bands, and responses that have energy greater than two are stored (Squire et al., 1999). These features are later used for comparison purposes. Viper stores the histogram of the mean filtered outputs, which provides a measure of the global texture characteristics of the images (Squire et al., 1999), as well as local texture features for each block.

Carson et al. (2002) states that texture features or descriptors computed at different scales can only confuse the retrieval process. While banks and multi-orientation Gabor filters are predominantly used, the scale selection problem has received the least attention (Carson et al., 2002). The novel scale selection approach by Carson et al. (2002) encapsulates the texture descriptor and groupings by use of an informative set of texture descriptors. This method, using the gradient of the image intensity values is based on edge/bar polarity stabilisation, and the texture descriptors are computed from the windowed second moment matrix (Carson et al., 2002).

In a study on CBIR retrieval performance conducted by (Ma and Manjunath, 1996), the basic Gabor features was more efficient compared to the conventional pyramid-structured wavelet transform, tree-structured wavelet transform, and multi-resolution simultaneous

autoregressive model features (Ma and Manjunath, 1996). Although there are many feature extraction methods, in the content-based retrieval task, a method that is efficient and accurate is desirable.

2.5 Colour Features

Colour images are generally stored in the RGB colour space. The RGB colour space consists of three primary colour components, red, green, and blue that range from 0 to 255 or 0.0 to 1.0. The RGB format is not a strong model for calculating distances between colour pixels since it is not a uniform colour space, and so, it will not perform well for colour segmentation (Schmid-Saugeon, 1999). Perceptually, small changes in a colour pixel co-ordinate should result in small perceptual distances. However, the RGB colour space does not have this property (Smith and Chang, 1996). The basic greyscale colour space has a uniform property, with a greyscale colour component that ranges from 0 to 255, black to white.

Calculating distance measures for colour pixels requires the use of a uniform colour space. Usually, the HSV or the L^*u^*v (Smith and Chang, 1996) colour spaces are preferred for colour distance calculations, especially for image segmentation. Due to the difficult computations of the L^*u^*v colour space, HSV is preferred as a better perceptually uniform colour space that is simple to compute (Squire et al., 1999).

HSV stands for Hue, Saturation, and Value. The model of this colour space is typically viewed as a cone. From the tip of the cone straight down to the base is the V component, and that varies the brightness of the current colour (0 to 1, dark to bright). From the centre of the cone, rotating around the V axis is the Hue, and it determines what the colour is (0 to 360). The radius from the centre of the cone to the perimeter of the cone determines the Saturation or shade of the colour (Bourke, 2000). In Cartesian co-ordinates, HSV can be represented as $(SV \cos 2\pi H, SV \sin 2\pi H, V)$, which makes it possible for finding colour differences (Belongie, Carson, Greenspan and Malik, 1998). HSV provides an intuitive colour system that corresponds to the human perception.

Colour histograms are often used for content-based retrieval of images (Swain and Ballard, 1990; Squire et al., 1999; Carson et al., 2002). It stores the frequencies of each colour in an image, restricted to a global set of colours determined by the image bit rate. Histograms offer a simplistic method to compare the global features of images. This is the basis of histogram intersection, where two histograms are compared by calculating the difference between each pair of bins (Squire et al., 1999). As there are thousands and millions of colours in an image, it is not feasible to capture every colour in the histogram. This problem is solved by quantising colours into colour bins or smaller colour sets (Squire et al., 1999). The downside of histogram intersection or in general, similarity by colour descriptors only is that it does not take into account the perceptual similarity between bins (Squire et al., 1999).

For a particular CBIRS query, a user may query for a tiger, and because a tiger has a similar orange colour to an orange fruit, the CBIRS using colour descriptors alone will return both images, which is not what the user expects. Global colour features of an image alone are not accurate for the content-based retrieval of images. Segmentation of images and then extracting each segments' local features including colour, texture, and spatial properties is a better approach for similarity comparison (Squire et al., 1999). However, as more features are introduced, the computational complexity rises, which leads to the non-trivial problem of labeled graph matching (Squire et al., 1999).

Blobworld uses the $L^*a^*b^*$ (Smith and Chang, 1996) perceptually uniform colour space to extract colour features (Carson et al., 2002). The colour features are smoothed to prevent over-segmentation where small regions of an image that are insignificant are dissolved into the region that encapsulates it (Carson et al., 2002). This smoothing approach leads to *blobs*, and the cost of similarity comparison is reduced with desirable results.

2.6 Image Processing for Dermatoscopy Images

Most Dermatoscopic images contain features that can cause inaccurate segmentation results (Schmid-Saugeon, 1999). Commonly, a dermatoscopic image will contain unwanted hair, noise, and bubbles that were formed through the application of the immersion oil used to render the lesion. A diagnostic program called "DullRazor" (Lee, Ng, Gallagher, Coldman and McLean, 1997) can be applied on the images to remove unwanted hair, while the Median filter can be used to remove noise and lessen the effect of bubbles in an image (Schmid-Saugeon, 1999).

"DullRazor" examines the properties of hair structures, produces and merges fixed size masks for each colour component (RGB) to filter out hair from an image. A predefined threshold value is used to determine the hair pixels. Pixel values that lie on a hair region are filled up with its neighbouring pixels after the filtering stage. This blends the removed hair region with the background texture. The preprocessing process improves the overall segmented image result (Lee et al., 1997).

Schmid-Saugeon (1999) developed a symmetry quantification scheme to calculate symmetry maps for dermatology images. It uses a combination of texture, color, and shape symmetry values to distinguish the benign and malignant lesions after the image segmentation phase. Although the results were reasonably effective in the separation of the two common lesions, it was stated by Schmid-Saugeon that the symmetry property itself is not sufficient to aid in the diagnosis of skin cancer.

2.7 Content-Based Information Retrieval Tools

Currently, there are many CBIRSs available in the computer vision field. Two of the well known systems are the Blobworld and the Viper systems. Blobworld uses a tree indexing system to classify images. Images are segmented into objects (blobs or regions) including the colour and texture descriptors which provide a faster and more accurate method for searching similar images compared to the conventional methods that don't use image segmentation (Carson, Thomas, Belongie, Hellerstein and Malik, 1999).

The Viper system is a fully flexible system that allows for extensions to suit different needs. It employs Query By Example by allowing users to refine their query results through relevance feedback. Viper utilises both local and global image colour and spatial frequency features, extracted at several scales. It has access to more than 80000 features including mode colour of various regions, the quantised average energies of the outputs of Gabor filters at various orientations and scales, and image features are stored in inverted files which improves the query response time as the database files are searched efficiently (Squire et al., 1999).

2.8 Summary

The need for CBIRSs and further development is steadily becoming an area of great importance. CBIRSs are querying systems that allow users to search for similar visual content

based on a query. Current CBIRSs employ digital image processing techniques to perform querying tasks. Whilst they're important, a good CBIRS is one that is used frequently. The essential formula for the success of CBIRS is a good user interface, presentation of results, and allow for relevance feedback for users to refine their search criteria based on the results that are shown to them. The current systems available to date include the Blobworld and Viper systems.

CBIRSs have slowly found their way into the medical domain, and have successfully been used for various applications. These systems have helped the medical environment manage their resources through clinical decision-making, medical research, and teaching. CBIRSs must be introduced seamlessly into the medical domain while being practical in order to gain user acceptance.

The *ABCD* rule is a scoring system for rating and recognising malignant melanoma. It is often used on dermoscopy images. Dermoscopy images are acquired through the use of microscopes with immersion oil and digital cameras. These images are studied by dermatologists, and it has been shown that these experts lacked consistency in the process of marking skin lesion borders, which suggests that their analysis is subjective. Automated skin lesion methods are more consistent, and has gained interest in the computer vision research area.

Whilst CBIRSs employ sophisticated image processing techniques, some of the essential fundamental features required are the texture features extracted by using Gabor filters, wavelet-based decomposition/transformations, and scale selection; colour features including colour histograms for global colour feature comparisons, colour bins and colour sets for a simpler representation of the full colour set, and the HSV colour space for a better perceptual uniform distance measure; and finally image processing techniques for skin lesions such as "DullRazor" for pre-processing of dermoscopy images by removing unwanted hair, and the symmetry quantification scheme by Schmid-Saugeon for distinguishing benign and malignant melanoma from segmented dermoscopy images.

3 Image Segmentation

3.1 Introduction

The purpose of image segmentation as the name suggests, is to break an image up into meaningful segments. Segmented images improves image analysis especially for image querying and retrieval (Winter and Nastar, n.d.). Image segmentation has been around since the 1960s, where it was originally applied to binary images, and later when imaging technology improved, it progressed to greyscale and eventually colour images (Jain, Murty and Flynn, 1999).

Segmentation is classified into two main categories, supervised and unsupervised. In cases where the different types of features is small, supervised segmentation is used where the *a priori* knowledge of the features is known (Ruzon, 1997). In cases where the different types of features is large or the assumptions of the features cannot be made, unsupervised segmentation must be used.

Algorithms that use supervised learning are provided with a set of training data where each data sample is labeled according to their specific features. Using the training data, algorithms can use it as a guideline to classify new data inputs based on the set of labels in the training data set. Supervised learning has prior knowledge of the possible classes

for the data sample it uses. Unsupervised learning does not use training data, instead data are classified based on the features present in the total data sample set (Duda, Hart and Stork, 2000).

Image segmentation has been used in a variety of digital domains. Being able to separate image regions into meaningful segments has aided the task of information retrieval and object character recognition, especially for the purpose of content-based image retrieval systems. Image segmentation is a generic technique that can be applied to most forms of images, some of which have been specialised to specific images types such as multisprectral medical images, LANDSAT satellite images, and more recently magnetic resonance images (MRI) and dermatoscopic images (Jain et al., 1999).

Some forms of image segmentation are called texture segmentation. Essentially, it involves extracting the texture features of the image and then using segmentation methods such as the clustering technique to segment an image into textured segments (Jain et al., 1999). There are many different approaches to segment images, and this review will focus mainly on clustering and the Markov Random Field (MRF) techniques. The following will describe the techniques in detail.

3.2 Clustering

Image segmentation via the use of data clustering is a very effective technique (Jain et al., 1999). When it is used with the appropriate distance and probability distribution functions, segments/clusters of an image can be properly distinguished. Often, the most difficult task in data clustering in general is determining the optimal number of clusters (Deng and Clausi, 2004). Specifying too few clusters will generate incomplete image segments, and specifying too much will generate redundant segments. See figure 1 for an example of the cluster threshold problem.

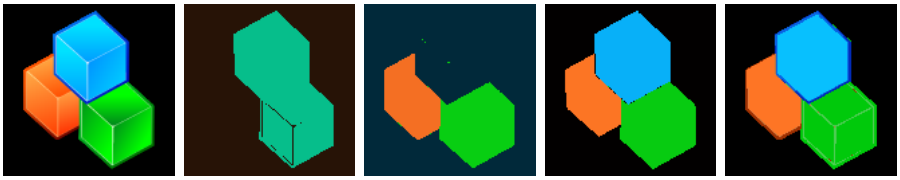

Figure 1: From left to right: the original image *(source image from: http://ij-plugins.sourceforge.net/plugins/clustering/)*, clustered image using 2 clusters (poor), clustered image using 3 clusters (close but one key cluster is missing), clustered image using 4 clusters (optimal), and clustered image using 10 clusters (artifacts are noticeable).

3.2.1 K-means Clustering

K-means (MacQueen, 1967) is one of the common approaches towards unsupervised data clustering. This data clustering technique is not restricted to images data only, in fact it is commonly used in data mining. Given a set of data, it will group them into different distinctive groups which would allow for easier and more effective analysis of data (Jain et al., 1999).

K-means attempts to group data points to K clusters based on the cluster's mean value, such that each cluster's data set share a common feature and each cluster is dissim-

ilar from each other (Jain et al., 1999). The naive approach to find the distance between data points in the feature space is to use the Euclidean distance. Although it is a faster way to determine distances between data points, it assumes that clusters are shaped in hyper-spheres only, which is quite restrictive.

The K-means algorithm is as follows:

1. Initialise K cluster centroid values.
2. Assign all data points to the closest cluster centroid.
3. Using the data in each cluster set, calculate the new cluster centroids.
4. If the new cluster centroid values have changed, return to step 2.

The K-means technique is sensitive to the number of clusters (K) given, the distance similarity criterion, and the initial cluster centres (Jain et al., 1999). Extensive research has been conducted in the past and present to solve the selection of the initial cluster centres. Techniques such as observing the image histogram for peaks is a way to find the initial cluster centres as well as the optimal amount of K clusters needed.

Similarly, a variant of the K-means algorithm, ISODATA (Ball and Hall, 1965) uses split and merge techniques to automatically determine the optimal number of clusters for a data set. Using pre-specified threshold values and any number of initial partition sets, the underlying technique will split the partitions that have a variance greater than the variance threshold, and merge partitions together that have a small difference in their partition mean values. It is an iterative process that will repeat the merge and splitting process until all the partitions satisfy the pre-specified threshold values (Jain et al., 1999).

A similar approach that uses the split method only was devised by Ray and Turi (1999) where validity measures that incorporated the inter and intra cluster distances were used. It initially starts off with one cluster containing all the data. The iterative process begins and repeatedly until the maximum number of clusters is reached, it splits the cluster with the maximum variance, assigns the two new clusters with a new mean value (based on the original cluster mean), and then uses the K-means algorithm to rebuild the clusters. At each stage, it calculates the total intra and inter cluster distances of each cluster, and uses those values to calculate the validity measure. The iterative step, K that yields the most accurate validity measure is considered as the optimal number of clusters required to segment an image.

3.2.2 Distance Similarity (Euclidean, Mahalanobis, and pixel location)

The Euclidean distance formula:

$$d(X, Y) = \sqrt{(X - Y)^T (X - Y)}, \tag{1}$$

where $(X - Y)$ is a matrice of the difference between the two data vectors X and Y. As mentioned previously, the Euclidean distance measure assumes that data is modeled as hyper-spheres only, which is quite restrictive as the full orientation of the clusters of data points are not considered (Jain et al., 1999).

The Mahalanobis Distance method models the distance between colour pixels more accurately compared to the Euclidean distance model (Jain et al., 1999). It models clusters

as hyper-ellipsoids rather than hyper-spheres, and it measures distances from each cluster centre using a metric specific to that cluster (it calculates the distances using the variance values of a cluster's data set). That is, the distance between colour pixels are calculated using the variance values between each colour component, which effectively considers the orientation of the data points in the sample space.

The Mahalanobis distance formula:

$$d_{M_j} = \sqrt{(X-Y)^T \sum_j^{-1} (X-Y)}, \tag{2}$$

where $(X-Y)$ is a matrice of the difference between the two data vectors X and Y with respect to cluster j, and $\sum_j$ is the covariance matrix for cluster j (Jain et al., 1999).

Analogous to the distance measures above, in the domain of digital images, the location of each sample vector in the image space is also important for segmentation. In data clustering for images, a sample vector may have similar colour values to a particular cluster's mean. However, if the sample vector's pixel position in the image is closer to another cluster's mean position, then the sample vector chosen cannot be assigned to the cluster with the closest average colour, but to the cluster with the closest average position. This is particularly useful for reducing the impact of noise in the segmentation phase (Jain et al., 1999). One approach to this problem is to include the pixels' co-ordinates as features.

3.2.3 Fuzzy C-means Clustering

Another variant of the K-means clustering is the Fuzzy C-means (FCM) clustering algorithm. It was first proposed by Dunn (1973) and later improvised by Bezdek (1981). Data points in the non-fuzzy clustering methods can only belong to a single cluster only (Wolfram, 2005), which is a form of hard clustering.

In the the FCM approach, data points can belong to more than one cluster based on some predefined membership rules that determine the degree of membership within each cluster for each data point. In the basic form of K-means clustering, each data points' distance from its cluster centre contribute equally to its centroid location. This makes the cluster centroid values vulnerable to outliers. The FCM method reduces the impact of outliers by reducing the influence of each outlier for those clusters that are least likely to have generated them (Schmid-Saugeon, 1999).

The description of Fuzzy C-means algorithm below closely follows that in Lim and Lee (1990):

1. Initialise the membership matrix $U = [u_{ij}] \ \forall \ i, j$, where u_{ij} denotes the fuzzy membership value of the ith data point at cluster j, $1 \leq i \leq N$, $1 \leq j \leq C$, N is the number of data points, and C is the number of clusters. The initialisation can be a set of random membership values.

2. Calculate the new cluster centre vectors at step K using the membership matrix $U^{(k)}$, $C^{(k)} = [c_j]$,

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m . x_i}{\sum_{i=1}^{N} u_{ij}^m}, \tag{3}$$

   where m the fuzzy factor is a real number greater than 1.0.

3. Update the fuzzy membership matrix, $U^{(k)}$ and $U^{(k+1)}$,

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}, \tag{4}$$

   where $\| * \|$ is any norm function.

4. Stop the iterative loop if $max_{ij} \| u_{ij}^{(k+1)} - u_{ij}^{(k)} \| < \epsilon$, otherwise goto step 2. $\epsilon \in (0, 1]$ is the stopping criteria value.

The following conditions must remain true for a constrained fuzzy membership:

1. $0 \leq u_{ij} \leq 1, \ \forall \ i, j$
2. $0 < \sum_{i=1}^{N} u_{ij} < N, \ \forall \ j$
3. $\sum_{j=1}^{C} u_{ij} = 1, \ \forall \ i$

The fuzzy c-means clustering technique has been widely used in fuzzy data statistical analysis and in particular, magnetic resonance images with significant success (Ahmed, Yamany, Mohamed, Farag and Moriarty, 2002). Although it does sound like a much better approach towards clustering, it does have its drawbacks. It suffers from the same problem encountered in the K-means algorithm where *a priori* knowledge of the number of clusters is required for a successful segmentation. Incorrect pixel classification is often seen due to the frequent overlapping of clusters in color space. Whilst the accuracy of the FCM technique is questionable, it has an exhaustive computational burden problem (Lim and Lee, 1990).

As the Mahalanobis distance measure has been used in the the K-means clustering method to detect clusters based on their orientation of data points, this has also been implemented in the fuzzy c-means approach. In Schmid-Saugeon's (1999) thesis research, he has shown that using the Mahalanobis distance provides an effect way to measure the orientation of the fuzzy cluster points. He calls this the Orientation Sensitive Fuzzy C-means (OSFCM) algorithm, which is basically the same as the existing Fuzzy C-means clustering method except the distance similarity is based on Mahalanobis.

Replacing the equations at step three of the original fuzzy c-means algorithm, the new equations consists of a fuzzy co-variance matrix for each cluster j:

$$F_j = \frac{\sum_{i=1}^{N} u_{ij}^m (x_i x_i^T - c_j c_j^T)}{\sum_{i=1}^{N} u_{ij}^m} \tag{5}$$

and the new membership matrix is calculated as follows:

$$u_{ij} = \left(1 + \sum_{k=1, k \neq j}^{C} \left(\frac{(x_i - c_j)^T F_j (x_i - c_j)}{(x_i - c_k)^T F_k (x_i - c_k)} \right)^{\frac{1}{m-1}} \right)^{-1} \tag{6}$$

Schmid-Saugeon (1999) applied the fuzzy c-means and the orientation sensitive fuzzy c-means algorithms on a set of dermoscopic images, and has shown that the orientation sensitive version is more reliable and accurate in segmenting the lesion from the healthy skin. The OSFCM is aimed at reducing the impact of data misclassification.

3.2.4 Gaussian Mixture Models

Clustering can be improved by using Gaussian Mixture Models (GMM) to cluster data to sets that are most likely generated by a particular Gaussian model, while the Expectation Maximisation (EM) algorithm is used to find the parameters of the GMMs (Jain et al., 1999). EM computes the probabilities of cluster memberships by maximising the log-likelihood of the data that were generated by the Gaussian mixture model (Dempster, Laird and Rubin, 1977). Other possible probability distributions are the log-normal and the Poisson distributions, all of which provide a statistical measure of the image information.

Another important note to mention is the EM method can be applied to continuous or categorised data. Categorisation of data will allow for weighting of data, thus areas of an image that are considered less important can be successfully separated from the key areas.

The following describes in detail, how the EM algorithm can be applied to GMMs (Moore, 2004):

1. Initialise the parameters for the mixture of Gaussians:

$$\lambda_0 = \{\mu_1^{(0)}, \mu_2^{(0)}, \cdots, \mu_K^{(0)}, p_1^{(0)}, p_2^{(0)}, \cdots, p_K^{(0)}\}, \tag{7}$$

   where K is the number of classes, μ_k is the mean of class ω_k, and p_k is the probability that μ_k appears in class ω_k.

2. Expectation step, calculate the expected classes for each data point for each class:

$$p(\omega_i|x_k, \lambda_t) = \frac{p(x_k|\omega_i, \lambda_t)p(\omega_i|\lambda_t)}{p(x_k|\lambda_t)} = \frac{p(x_k|\omega_i, \mu_i^{(t)}, \sum_i^{(t)})p_i^{(t)}}{\sum_{j=1}^{K} p(x_k|\omega_j, \mu_j^{(t)}, \sum_j^{(t)})p_j^{(t)}}, \tag{8}$$

   where t represents the tth iteration.

3. Maximisation step, calculate the maximum likelihood μ given each datas' class membership distributions:

$$\mu_i^{(t+1)} = \frac{\sum_k p(\omega_i|x_k, \lambda_t)x_k}{\sum_k p(\omega_i|x_k, \lambda_t)} \tag{9}$$

$$\sum_i^{(t+1)} = \frac{\sum_k p(\omega_i|x_k, \lambda_t)[x_k - \mu_i^{(t+1)}][x_k - \mu_i^{(t+1)}]^T}{\sum_k p(\omega_i|x_k, \lambda_t)} \tag{10}$$

$$p_i^{(t+1)} = \frac{\sum_k p(\omega_i|x_k, \lambda_t)}{R}, \tag{11}$$

   where R is the number of data points or records.

4. Repeat steps 2 to 3 until the EM parameters $\mu_k^{(t)}$ and $p_k^{(t)}$ are stable.

As the EM algorithm guarantees a local optimum, the initialisation of the EM GMM parameters is crucial for good results. Attempts to improve the EM initialisation include using the K-means algorithm to find the initial class mean values (Moore, 1999).

Just like the K-means algorithm, the EM with GMM can be applied to most forms of data related tasks such as data mining and statistical data analysis (Moore, 2004). EM and GMM have been used in various fields such as image texture and colour classification for image database retrieval (Permuter, Francos and Jermyn, 2003), and data compression including image (So and Paliwal, 2002) and speech compression (Rao, 2003). EM with GMM clustering via the use of probabilistic constraints has been experimented on dermatological image data with some success (Law, Topchy and Jain, 2005).

The EM algorithm in general suffers from its high rate of data accesses required to calculate the GMM parameters. However, faster implementations of the EM algorithm such as the use of multi-resolution kd-trees has proved that EM is practical for its applications (Moore, 1999). The other benefits of using EM with GMM is it has low complexity, rate independence, and scalability which is good for large data sets (So and Paliwal, 2002).

3.3 First Variants of the MRF models

In the past, stochastic models using the Bayesian model, Gibbs distribution, and the Markov Random Field models have been applied to the field of digital image restoration of degraded images (Geman and Geman, 1984). The image restoration process is closely related to image segmentation and boundary detection.

Throughout the years, MRFs have gained high interests in the field of computer vision. It has been utilised for a large range of vision related problems some of which include image segmentation, surface reconstruction, edge detection, texture analysis, active contours, data fusion, and visual integration (Li, 1995). More recently, it has been employed for modeling Genetic Algorithms (Brown, Garmendia-Doval and McCall, 2001).

It is shown that the MRF models are a powerful stochastic tool for modeling joint probability distributions between spatial image pixels (pixels that are related in the feature space) (Deng and Clausi, 2004). MRF models are good at extracting textures from an image while preserving the spatial relationships between regions of an image. The MRF models consists of two layers, a region labeling component via MRFs and a feature modeling component via GMMs, with a constant weighting factor to combine the two. Many researchers have applied this MRF scheme to various image segmentation schemes.

Alternatively, a recent study conducted by Deng and Clausi (2004) has shown that the use of an improved MRF model in image segmentation is able to produce sound results for different types of images particularly in synthetic aperture radar sea ice images. Although the sea ice images are not related to the dermatological images, they do however contain large quantities of rapid changing textures that are similar to the characteristics of the dermatological images.

The new variant of the MRF model proposed by Deng and Clausi (2004) introduces a variable weighting factor to combine the two essential components of the MRF model. The segmentation results achieved by this new scheme is significantly better than the previous MRF model scheme.

As the variable weighting factor MRF model is based on unsupervised data, that is no training data is present, the weighting factors cannot be determined beforehand. To determine the weighting factors, the estimation of the mean and standard deviation of each class in an image must be calculated. This can be calculated accurately using the EM algorithm.

Although the new MRF model technique is quite effective in image segmentation, it is still not clear whether it will perform better than the previous image segmentation techniques for dermatological images. However, given that the proposed MRF model by Deng and Clausi (2004) is suitable for images with similar characteristics to dermatological images, it may be possible for the new technique to pre-process dermatological images (remove unwanted hair), as well as segmenting the images in one go. This can effectively improve the task of content-based image retrieval, especially for lesion images.

3.4 Summary

Image segmentation is the process of breaking up an image into meaningful segments, and provides a more practical way to analyse images. Image segmentation opens paths for effective information retrieval and object character recognition, which is commonly viewed as a fundamental building block of a good content-based image retrieval system. The broad field of image segmentation is a general purpose domain that is applicable to most forms of digital images. There are many approaches to image segmentation, and this review has covered a few of the most commonly studied techniques.

Clustering is a well known technique in data mining, especially for data analysis. It is a general purpose technique that has lead to the development of several different clustering algorithms in order of increasing complexity and accuracy, the K-means, Fuzzy C-means, and Expectation Maximisation of Gaussian Mixture Models. These algorithms have been applied to images with significant success. The fundamental problem associated with these algorithms is choosing the right number of clusters and their initial cluster centroids. Whilst those are important, the distance measure such as the Euclidean or the Mahalanobis, and the spatial information of the data points in an image are critical factors that need to be considered.

Finally, the last segmentation technique covered in this review is Markov Random Fields. MRFs are a powerful stochastic tool for modeling joint probability distributions between image pixels. This essential property of MRFs is suitable for texture segmentation. The MRF segmentation model consists of a region labeling component via MRFs and a feature modeling component via GMMs. A constant factor is normally used to combined the two, but until recently, a new algorithm proposed by showed that a variable weighting factor to combine the two components offered desirable results. MRFs are powerful, as such, they have been used in a large variety of applications.

4 Conclusions

In this review, content-based information retrieval systems, texture and colour features, image processing for dermatological images, image segmentation including K-means, Fuzzy C-means, Gaussian Mixture Models with Expectation Maximisation, and the variants of Markov Random Field based segmentation have been discussed. These essential features are the basis of a good CBIRS, and if they are applied properly to a particular domain such as dermatological images, specialised CBIRSs can be developed.

It is envisioned that the topics discussed in this review will lead to further research into the study of dermatological images, and in the near future, effective lesion CBIRSs will be available to the world. Image Segmentation is important for studying dermatological images, and it is the foundation for building CBIRSs that are specialised for lesion images.

References

Abate, A. F., Nappi, M., Ricciardi, S. and Tortora, G. (2004). FACES: 3D FAcial reConstruction from anciEnt Skulls using content based image retrieval, *Journal of Visual Languages & Computing* **15**: 373–389.

Ahmed, M. N., Yamany, S. M., Mohamed, N., Farag, A. A. and Moriarty, T. (2002). A Modified Fuzzy C-Means Algorithm for Bias Field Estimation and Segmentation of MRI Data, *IEEE Transactions on Medical Imaging* **21:3**.

Ball, G. H. and Hall, D. J. (1965). ISODATA, a novel method of data analysis and classifcation, *Technical report*, Stanford University.

Belongie, S., Carson, C., Greenspan, H. and Malik, J. (1998). Color- and Texture-Based Image Segmentation Using EM and Its Application to Content-Based Image Retrieval, *Technical report*, University of California, Berkeley.

Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York. Citation from Jain, A. K., Murty, M. N. and Flynn, P.J, "Data clustering: a review", 1999.

Bourke, P. (2000). HSV colour space, Online.
URL: *http://astronomy.swin.edu.au/˜pbourke/colour/hsv/*

Brown, D. F., Garmendia-Doval, A. B. and McCall, A. W. (2001). Markov Random Field Modelling of Royal Road Genetic Algorithms, *Selected Papers from the 5th European Conference on Artificial Evolution*, Springer-Verlag, pp. 65–76.

Carson, C., Belongie, S., Greenspan, H. and Malik, J. (2002). Blobworld: Image Segmentation Using Expectation-Maximisation and Its Application to Image Querying, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24:8**: 1–13.

Carson, C., Thomas, M., Belongie, S., Hellerstein, J. M. and Malik, J. (1999). Blobworld: A System for Region-Based Image Indexing and Retrieval, *Third International Conference on Visual Information Systems*, University of California, Berkeley, pp. 1–8.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM algorithm, *Journal of the Royal Statistical Society* **39**: 1–38. Citation from Jain, A. K., Murty, M. N. and Flynn, P.J, "Data clustering: a review", 1999.

Deng, H. and Clausi, D. A. (2004). Unsupervised image segmentation using a simple MRF model with a new implementation scheme, *Pattern Recognition* **37**: 2323–2335.

Duda, R. O., Hart, P. E. and Stork, D. G. (2000). *Pattern Classification (2nd ed.)*, Wiley Interscience.

Dunn, J. C. (1973). A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters, *Journal of Cybernetics* **3**: 32–57. Citation from Jain, A. K., Murty, M. N. and Flynn, P.J, "Data clustering: a review", 1999.

Erkol, B., Moss, R. H., Stanley, R. J., Stoecker, W. V. and Hvatum, E. (2004). Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes, *Skin Research and Technology* **11**: 17–26.

Geman, S. and Geman, D. (1984). Stochastic relaxation and Gibbs distributions and the Bayesian restoration of images, *IEEE-PAMI* **6**: 721–741.

Guillod, J., Schmid-Saugeon, P., Guggisberg, D., Cerottini, J. P., Braun, R., Krischer, J., Saurat, J.-H. and Murat, K. (2002). Validation of segmentation techniques for digital dermoscopy, *Skin Research and Technology* **8**: 240–249.

Huston, L., Campbell, J., Sukthankar, R. and Pillai, P. (2004). Forensic Video Reconstruction, *Proceedings of the ACM 2nd international workshop on Video surveillance & sensor networks*, pp. 20–28.
URL: *www.cs.cmu.edu/˜rahuls/pub/vssn2004-rahuls.pdf*

Jain, A. K., Murty, M. N. and Flynn, P. J. (1999). Data clustering: a review, *ACM Computing Surveys (CSUR)* **31**: 1–60.

Kato, Z. and Pong, T.-C. (2001). A Markov Random Field Image Segmentation Model Using Combined Color and Texture Features, *Proceedings of the 9th International Conference on Computer Analysis of Images and Patterns*, pp. 547–554.

Lagerstrom, R. (2004). Melanoma Identification - A few facts about melanoma, Online. (last accessed April 26, 2005).
URL: *http://www.cmis.csiro.au/iap/RecentProjects/melanoma_facts.htm*

Law, M. H. C., Topchy, A. and Jain, A. K. (2005). Model-based Clustering With Probabilistic Constraints, *Technical report*, Michigan State University.

Lee, T., Ng, V., Gallagher, R., Coldman, A. and McLean, D. (1997). DullRazor: a software approach to hair removal from images, *Comput Biol Med.* **6**: 533–543.

Lehmann, T. M., Schubert, H., Keysers, D., Kohnen, M. and Wein, B. B. (2003). The IRMA code for unique classification of medical images, *Proceedings SPIE*, pp. 109–117.
URL: *libra.imib.rwth-aachen.de/irma/ps-pdf/SPIE_2003-5033(x)-56.pdf*

Li, S. Z. (1995). *Markov Random Field Modeling in Computer Vision*, Springer-Verlag, New York Berlin Heidelberg Tokyo.
URL: *http://www.ntu.ac.sg/˜szli/MRF_Book/MRF_Book.html*

Lim, Y. W. and Lee, S. U. (1990). On the Color Image Segmentation Algorithm based on the Thresholding and the Fuzzy C-means Techniques, *Pattern Recognition* **23**: 935–952.

Ma, W. and Manjunath, B. (1996). Texture features and learning similarity, *Computer Vision and Pattern Recognition* **5**: 425–430.

MacQueen, J. B. (1967). Some Methods for classification and Analysis of Multivariate Observations, *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, University of California, Berkeley, pp. 281–297. Citation from Jain, A. K., Murty, M. N. and Flynn, P.J, "Data clustering: a review", 1999.

Menzies, S. W., Crotty, K. A., Ingvar, C. and McCarthy, W. H. (2003). *An Atlas of Surface Microscopy of Pigmented Skin Lesions: Dermoscopy*, McGraw-Hill Australia.

Moore, A. (1999). Very Fast EM-based Mixture Model Clustering using Multiresolution KD-trees, *in* M. Kearns and D. Cohn (eds), *Advances in Neural Information Processing Systems*, Morgan Kaufman, 340 Pine Street, 6th Fl., San Francisco, CA 94104, pp. 543–549.

Moore, A. W. (2004). Clustering with Gaussian Mixtures, Online tutorial notes, Carnegie Mellon University.
URL: *http://www.cs.cmu.edu/˜awm/tutorials/gmm14.pdf*

Müller, H., Michoux, N., Bandon, D. and Geissbuhler, A. (2004). A Review of Content-Based Image Retrieval Systems in Medical Applications - Clinical Benefits and Future Directions, *International Journal of Medical Informatics* **73**: 1–30.

Müller, H., Pun, T. and Squire, D. M. (2004). Learning from user behaviour in image retrieval: Application of market basket analysis, *International Journal of Computer Vision* **56**(1–2): 65–77. (Special Issue on Content-Based Image Retrieval).
URL: *http://dx.doi.org/10.1023/B:VISI.0000004832.02269.45*

Permuter, H., Francos, J. and Jermyn, I. H. (2003). Gaussian Mixture Models of Texture and Colour for Image Database Retrieval, *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Hong Kong, pp. 1–4.

Rao, B. D. (2003). Speech Coding using Gaussian Mixture Models, *Technical Report 02-062*, Department of Electrical and Computer Engineering, University of California, San Diego CA 92122.

Ray, S. and Turi, R. H. (1999). Determination of Number of Clusters in K-Means Clustering and Application in Colour Image Segmentation, *Proceedings of the 4th International Conference on Advances in Pattern Recognition and Digital Techniques*, Narosa Publishing House, New Delhi, India, pp. 137–143. (invited paper).

Ruzon, M. A. (1997). Texture Segmentation An Introductory Primer, Online tutorial notes. (last accessed April 26, 2005).
URL: *http://ai.stanford.edu/~ruzon/tex_seg/node6.html*

Schmid-Saugeon, P. (1999). *Segmentation and symmetry measure for image analysis: application to digital dermatoscopy*, PhD thesis, Department of Engineering.

Schmid-Saugeon, P., Guillod, J. and Thiran, J.-P. (2003). Towards a computer-aided diagnosis system for pigmented skin lesions, *Computerized Medical Imaging and Graphics* **27**: 65–78.

Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A. and Jain, R. (2000). Content-Based Image Retrieval at the End of the Early Years, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22:12**: 1–32.

Smith, J. R. and Chang, S.-F. (1996). Tools and techniques for color image retrieval, *Storage & Retrieval for Image and Video Databases IV*, Vol. 2670 of *IS&T/SPIE Proceedings*, San Jose, CA, USA, pp. 426–437.
URL: *ftp://ftp.ctr.columbia.edu/CTR-Research/advent/public/papers/96/smith96b.ps.gz*

So, S. and Paliwal, K. K. (2002). Efficient Block Coding of Images using Gaussian Mixture Models, *Proceedings of the Fourth Australasian Workshop on Signal Processing and Applications*, pp. 1–4.

Squire, D. M., Müller, W., Müller, H. and Raki, J. (1999). Content-based query of image databases, inspirations from text retrieval: inverted files, frequency-based weights and relevance feedback, *The 11th Scandinavian Conference on Image Analysis (SCIA'99)*, Kangerlussuaq, Greenland, pp. 143–149.

Stolz, W., Braun-Falco, O., Bilek, Landthanler, M., Conneta, A. B. and Riemann, A. (1994). ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma, *Eur J Dermatol* **4**: 521–7.

Stolz, W., Braun-Falco, O., Bilek, P., Landthaler, M., Burgdorf, W. H. C. and Cognetta, A. B. (2002). *Color Atlas of Dermatoscopy*, Blackwell Publishing, Wissenchafts-Verlag Berlin.

Swain, M. J. and Ballard, D. H. (1990). Indexing via color histograms, *Proceedings of the DARPA Image Understanding Workshop*, Pittsburgh, PA, USA, pp. 623–630.

Weldon, T. P., Higgins, W. E. and Dunn, D. F. (1996). Gabor filter design for multiple texture segmentation, *Optical Engineering* **35:10**: 2852–2863.

Winter, A. and Nastar, C. (n.d.). Differential Feature Distribution Maps for Image Segmentation and Region Queries in Image Databases, *IEEE Workshop on Content-based Access of Image and Video Libraries (CBAIVL'99)*, Fort Collins, Colorado, USA.

Wolfram (2005). 1.12 Fuzzy Clustering, Online manual. (last accessed April 28, 2005). **URL:** *http://documents.wolfram.com/applications/fuzzylogic/Manual/12.html*

Zarita, R. and Lelandais, S. (1997). Wavelets and high order statistics for texture classification, *Scandinavian Conference on Image Analysis* **8**: 95–102.