Chapter 5

Texture Recognition

A model-based texture recognition system which classifies image textures seen from different distances and under different light directions is presented in this chapter. This system works on the basis of a surface model obtained by means of 4-source Colour Photometric Stereo (CPS) used to generate 2D image textures as they would have appeared if imaged under different imaging geometries. The proposed recognition system combines co-ocurrence matrices for feature extraction with a Nearest Neighbour classifier. Moreover, the recognition process allows one to guess the approximate direction of the light used to capture the test image.

5.1 Introduction

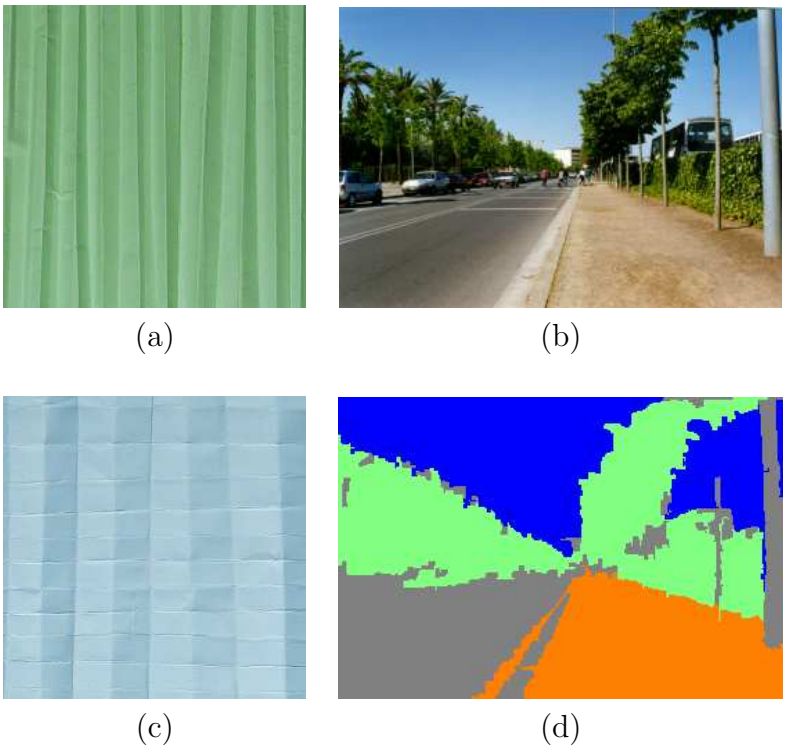
In its simplest form, image classification is the process of assigning similar images or regions of an image to the class to which they belong. It is well-known that there are two types of classification, namely *supervised* and *unsupervised*. In supervised classification, classes are specified a priori by an analyst, whereas in unsupervised classification, classes are automatically clustered into sets of prototype classes where in some cases a user specifies the number of desired categories. The recognition problem here is being posed as a classification or categorisation task where the classes are either defined by the analyst or are learned based on the similarity of patterns.

Many recognition/classification systems require a previous learning or modeling

task to obtain the *feature vectors* (represented by a set of d features) which characterise the objects of interest. Before classification can take place, some homogeneity or similarity criteria must be defined. These criteria are normally specified in terms of a set of feature measures, with each one providing a quantitative measure of a certain characteristic. The process which allows us to compute these features, called *feature extraction*, is concerned with the detection and localisation of particular image patterns which represent significant features in the image. These features are dependent on the application and are generally of two different origins: a global image property or a region of the image with a relevant measurable property.

While a global property tries to describe an image by extracting a single feature over the whole image (see Figure 5.1.(a) and Figure 5.1.(c)), a region property tries to describe the different regions of the image individually. Figure 5.1.(b) shows an example in which the main regions of the original image, such as road, sky, trees, and ground, are charactarised by a set of relevant measurable properties in order to perform their recognition (see Figure 5.1.(d)). This was done by Martí et al. [10, 110], proposing an object recognition system for outdoor scenes. Their work is based on a top-down strategy which segmented the main objects in the image, taking colour and texture features into account.

Texture is a fundamental characteristic of natural images which, in addition to colour, plays an important role in image understanding and scene interpretation. There are many researchers in image processing and computer vision areas who have considered the concept of feature vectors to cope with texture classification. In the area of database retrieval for instance, texture features are used to search an image database to find images similar to the image submitted by the user [128, 20]. In texture segmentation, many algorithms partition the image into a set of regions which are visually distinct and uniform with respect to textural properties [74, 107, 116]. In remote sensing and radar applications, texture features have been used to identify forest regions and their boundaries and to identify and analyse various crops [54, 166]. In industrial vision inspection, texture features have been used to perform the classification of different surface materials [105]. Obviously, there are many other applications in which texture is used to carry out a recognition or



(a) (b)

(c) (d)

Figure 5.1: Example of global and local image properties. In (a) and (c) a global image property is extracted in order to distinguish between the two images. In (b) each region of interest illustrated in (d) is characterised by a set of relevant measurable properties in order to perform object recognition.

classification task.

The aim of this chapter is to present a novel proposal of performing texture classification under varying imaging geometries. It will focus on the use of textural feature vectors computed over the global image to characterise the textures of interest. Note that texture features seem ideally suited to our purposes since the problem analysed is itself caused by variations in texture perception. Furthermore, this will be a supervised classification approach since prior knowledge about the analysed texture classes would be needed to perform their recognition.

Summarising, the main contributions of this chapter are:

- To propose a texture classification system to overcome the problem of classifier failure induced by varying imaging properties such as light direction and

camera distance.

- To integrate the prediction framework proposed in Chapter 4 into a texture classification system.
- To provide the capability of guessing an approximate direction of the light used to capture the test images.

The remainder of this chapter has been structured in the following way. In the next section we review the most important work on the topic of texture classification under varying imaging geometries. In Section 5.3, our model-based approach for texture classification is presented, describing first the general scheme then detailing the steps individually. Finally, some conclusions are presented in Section 5.4.

5.2 Texture Classification System

As illustrated in the introduction of this thesis (see Section 1.1), changes in the imaging geometry can significantly alter the appearance of the surface, implying significant variations in the image texture. These variations in the image may introduce critical misclassification rates in a typical texture classification system. Therefore, our main challenge is the problem of classification of textured surfaces imaged under varying geometries as well as the necessity of finding reliable methods of reducing classification errors caused by these changes in the geometry.

5.2.1 Related Work

Studying the dependence of texture on viewing and light directions is fairly new in texture research, therefore, there is only a limited amount of related work.

In computer graphics, traditional methods of 2D texture synthesis and texture mapping make no provision for the change in texture appearance with viewing and light directions. For example, in 2D texture mapping, when a single digital image of a rough surface is mapped onto a 3D object, the appearance of roughness is usually lost or distorted. Bump mapping [11, 14, 15] preserves some of the appearance

of roughness, but knowledge of the surface shape is required. Many of the problems associated with traditional texture mapping and bump mapping techniques are described by Koenderink and Van Doorn in [85].

In computer vision, very little work has been published on the topic of texture classification independent from the direction of the light. On the other hand, there are a great number of works which propose rotation and scale invariant texture classification [29, 52, 133, 108, 137, 91], as well as similar works on topics like multiscale and scale-space texture analysis [29, 17]. All these methods tend to take a texture obtained from a single view and change it for various scales and different angles, performing multiscale texture classification. In general, they assume the imaged texture is the same for all scales and rotations and, therefore, they do not consider that seeing a texture from a different physical distance and under different lighting conditions may change it entirely.

While analysing the published work on the topic of texture classification independent from the direction of light, we observed three different ways of dealing with this problem:

- The first strategy consists of extracting and using explicit separate 3D shape and surface albedo information. The colour and gradient vector of every visible surface patch describe the surface in a way independent from the light, and the classification can be done directly on the basis of this explicit information. For example, Barsky and Petrou [6] proposed an illumination-invariant classification scheme based on 5 descriptors for each surface patch obtained by means of colour photometric stereo: two gradient components and three colour components.

  McGunnigle and Chantler [113] proposed a rough surface classification scheme which extracts rotation invariant statistics from photometric estimates of the surface derivatives. Their method assumes that the surface is uniformly coloured. On the other hand, Chantler and Wu [24] proposed a novel classification scheme which is surface rotation invariant. They based it on the magnitude spectra of the surface derivatives extracted from photometric stereo.

- Another strategy to solve this problem is to study the immediate effects introduced by light direction to the observed 2D texture. Chantler [22] has shown that this effect can be described as a directional filter and, in principle, could be inverted. Recently, Chantler et al. [23] presented a formal theory demonstrating that changes in the tilt of the light direction make texture features follow super-elliptical trajectories in multi-dimensional feature spaces. Based on this work, Penirschke et al. [125] developed an illuminant rotation invariant classification scheme which uses photometric stereo for the detection of surface relief and Gabor features for feature extraction.

  In Appendix A, a study analysing the immediate effect of light direction over features extracted from the co-occurrence matrix is presented. This is used to perform a simultaneous surface texture classification and illumination tilt angle prediction.

- Finally, we can train a classifier on a wide selection of images of the same surface, obtained from various viewpoints and under various lighting conditions [33, 36, 32, 56]. Thus, the information on changes in surface appearance is explicitly built in the classifier using both the reflectance and the 3D relief information which allows it to recognise a surface correctly under novel viewing and lighting conditions. Leung and Malik [98, 99], following this strategy, developed a texture classification scheme which identifies 3D "textons" in the Columbia-Utrecht database for the purpose of light and viewpoint invariant classification. Basically, a 3D texton is an item in a vocabulary of prototypes of tiny surface patches with associated local geometric and photometric properties. More recently, Varma and Zisserman [164] proposed a new classification system which uses a distribution over textons obtained from training images as a texture model. On the other hand, Gonzalez [56] proposed a supervised statistical classification scheme which combines a bank of Gabor filters for feature extraction with a linear Bayes classifier to deal with changes in the light direction.

Figure 5.2 summarises these three different ways of dealing with the problem of

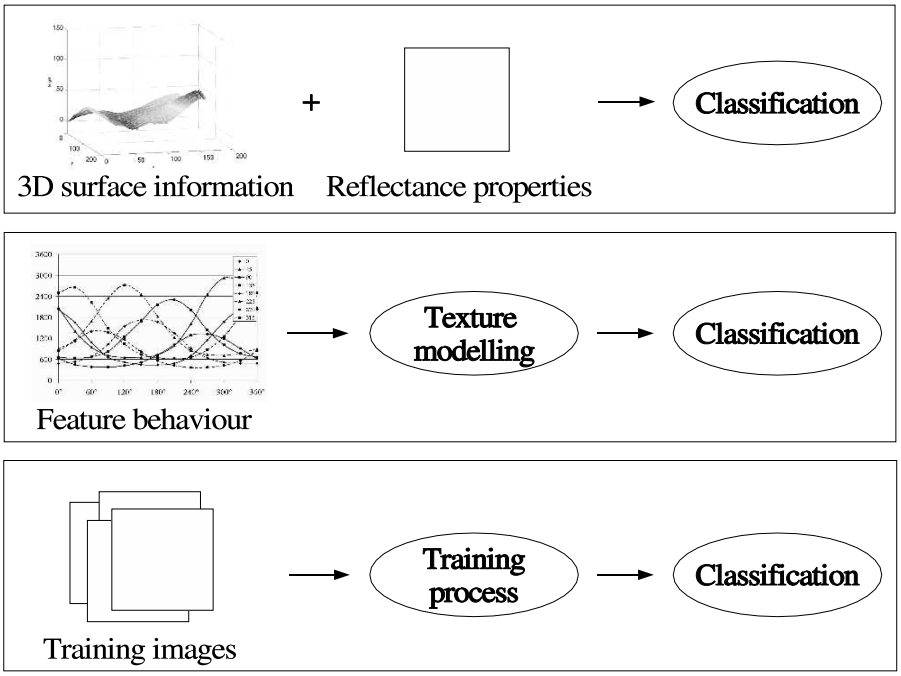


Figure 5.2: Strategies of light invariant texture classification.

classifying textures imaged with different light directions.

5.2.2 Discussion

We reviewed three different ways of dealing with the problem of texture classification independent from the direction of light. The first strategy uses the 3D shape and surface albedo information to characterise a surface texture. Although we saw that this strategy can be used to perform surface classification, it always requires the recovering of surface information. Therefore, we can not directly classify an unknown imaged texture since surface information is needed. That can be solved by using the second strategy which tries to study the immediate effects introduced by light direction over texture features. This approach requires one to work out theoretically the relationship between changes in the light direction and the values of the texture features used.

In this thesis, we have opted to follow the third strategy based on a set of training images. Roughly, we train a classifier on a selection of images. The main reason

of choosing this strategy in lieu of the others is to exploit the capability of colour photometric stereo to render images of a surface under novel lighting conditions as well as the capability of the prediction method proposed in Section 4.3.1 to create images of a surface texture under novel imaging geometries (e.g. camera distance).

The ability of performing the classification using a set of texture images is very important. Clearly, the quality of the classifier depends on the quality and size of the training set which is always finite. It is obvious that we can not think of using all the images corresponding to every single light direction in order to design a classifier.

In general, the process of designing a classifier must be inductive in the sense that the information obtained from the elements of the training set must be generalised to cover the whole feature space, implying that the classifier should be near optimal for all feasible textures, even for those it had never seen before. When tackling classification under varying lighting conditions, the problem is considerably increased. Not only should the classifier be able to discern different textures, it should also be able to perform robustly when faced with changes in the appearance of identical textures. In that case, it is not viable to train on every single light direction and store up an infinite number of training statistics to secure a more successful classification. One potential solution is a *model-based* approach. This takes advantage of the fact that it is not based on the actual surface but on various images obtained from its model. Therefore, it is a much more reasonable approach because it merely requires a finite and rather small data set to model the surface. Essentially, this strategy aims at modelling the texture variability in the feature space by means of a primary training stage.

Therefore, following this idea, we design a supervised model-based texture classification system in which a set of images under a variety of imaging configurations is used as a prior knowledge. In the next section this proposal is exhaustively described.

5.3 Model-Based Texture Classification

My approach is to integrate the surface texture information derived by colour photometric stereo, as described in Section 3.5.2, into a complete model-based texture classification system. Basically, the main idea consists of creating, by means of this surface texture information, a "virtual" database of image textures against which we compare the unknown test images in order to classify them.

In Chapter 3, we saw that photometric stereo is the technique which allows us to obtain surface texture information from only a few images of the same surface imaged under various light directions. Moreover, we have analysed various alternative techniques which also allow us to obtain 3D information about surfaces [170]: *stereo vision* [48], *optical flow* [5] and various *Shape from X* methods [65]. Over all these techniques, the photometric stereo technique was chosen because it has various advantages over all other methods. We have seen that it does not suffer from the correspondence problem like conventional stereo does, it does not make strong assumptions about the underlying surface structure like some Shape from X techniques do, and it allows us to recover both local colour and local gradient while flagging the places where some of its assumptions break down and recovery is impossible.

In supervised classification systems, a previous learning or modeling task is usually required to obtain the feature vectors which characterise the objects of interest. In our case, we use a virtual database of images as the input of this learning process in which the goal is simple: to extract those texture feature vectors which should be used as models in the classification process. It is important to note that surface information plays an important role in our proposal since it is used in a preliminary step to create image textures compatible with the imaging geometry of the test images.

Figure 5.3 illustrates the whole procedure of the proposed model-based texture classification system. Note that it is divided into two main phases:

- **Virtual database creation.** The process in which different images of each

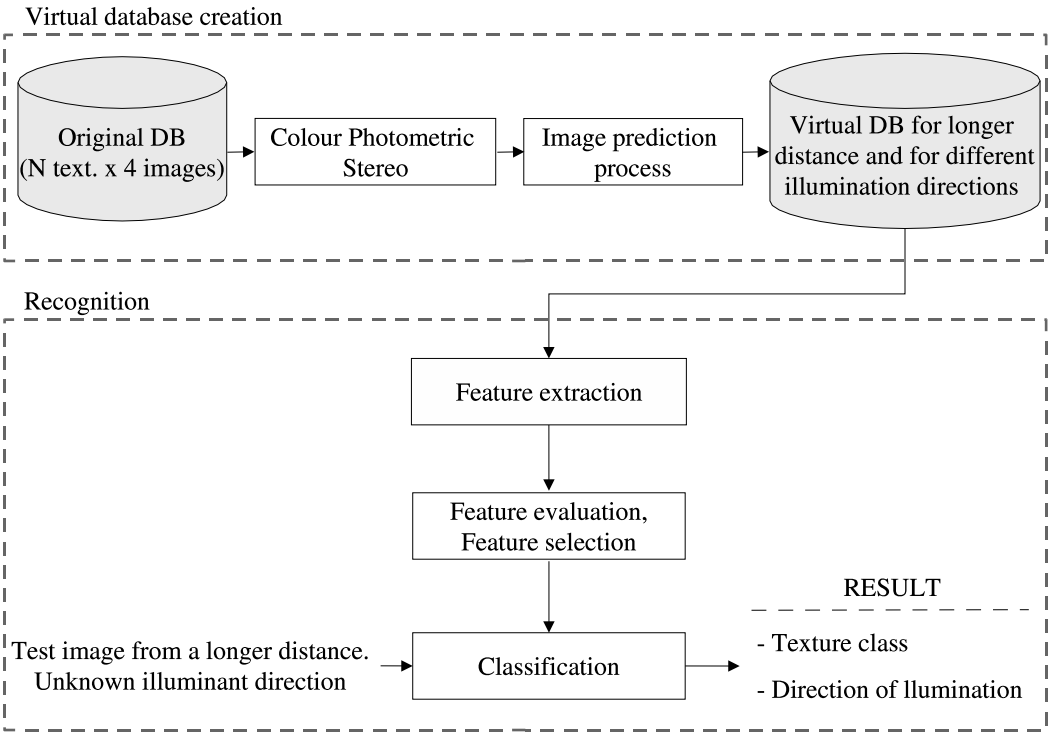


Figure 5.3: Recognition scheme. It is divided into two main phases: virtual database creation and recognition.

texture are created to be used as references in the classification process.

- **Recognition procedure.** Initially, each texture class in a learning process is modelled by a feature vector. Afterwards, in the classification process, an unknown test image is classified into the texture class to which it belongs.

We will also show how the proposed recognition procedure allows us to guess the approximate direction of the light used to capture the test images.

5.3.1 Virtual Database Creation

The virtual database creation comes in two "flavours": creation of the virtual database for test images seen from the same distance as the original images and creation of the virtual database for test images seen from a longer distance than that of the original images. When the test images are known to have been captured at

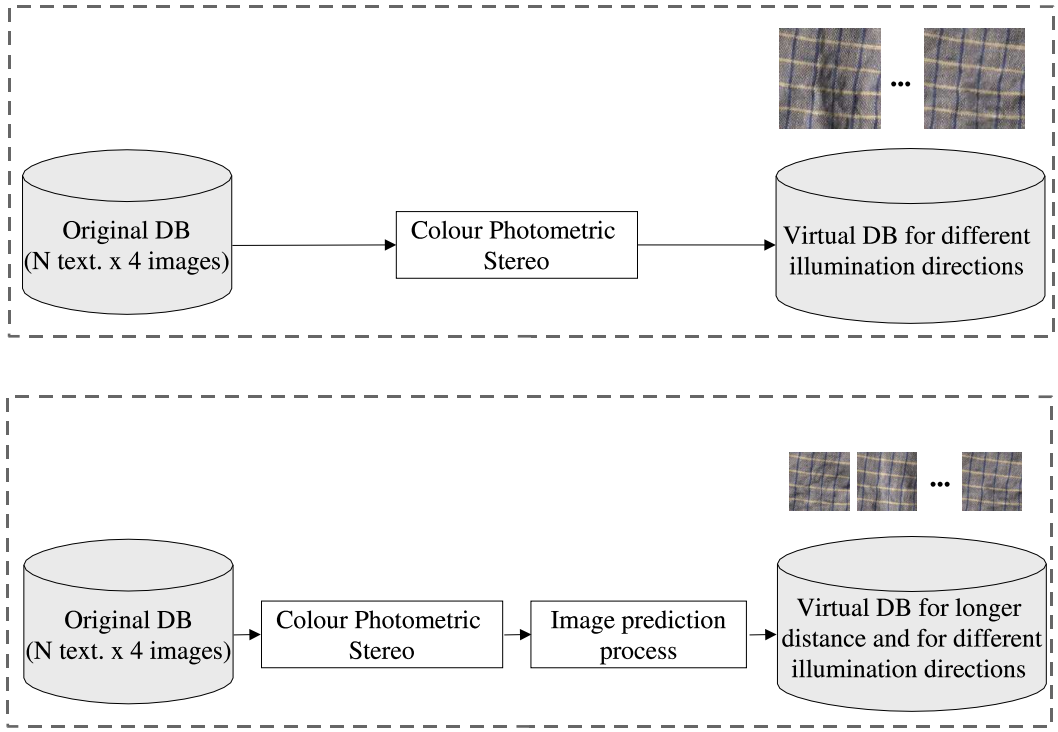


Figure 5.4: Virtual database creation. This can be performed in two different ways: creation of the virtual database for test images seen from the same distance as the training images and creation of the virtual database for test images seen from a longer distance than that of the original images.

the same distance as the originals, the creation of the virtual database is straightforward: the use of surface texture information extracted by photometric stereo allows us to directly obtain a rendering of the surface texture (see Section 3.5.4.1). However, things are less straightforward when the test images have been captured from a longer (but known or hypothesised) distance than the originals. In this case, in order to create the virtual database, we use the method presented in Section 4.3.1, which allows us to predict the appearance of a surface texture as seen from a longer distance and for various directions of light. Hence, we deal with the problem of texture recognition under varying geometries such as camera distance and light direction.

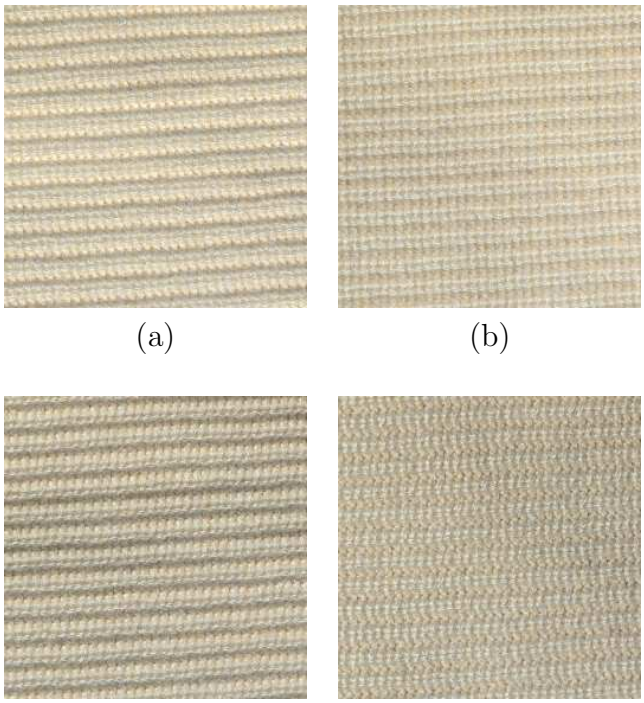
Figure 5.4 shows these two ways of performing the virtual database creation. Note that in both cases, 4-source colour photometric stereo is used to compute the detailed shape and colour of a surface which is used to create the desired images.

All the photometric stereo sets are constructed from 4 images illuminated at a fixed elevation angle and with 4 different tilt angles, 0°, 90°, 180° and 270°. This is the original database illustrated in Figure 5.4. Note that in the first case, the virtual database is created directly from the photometric information, while in the second case, the prediction method is needed. In that case, the virtual database contains a set of texture images imaged from a longer distance and under different directions of light. This will be the basis of dealing with the recognition of textures seen from a longer and known or hypothesised distance.

As it will be explained in more detail in the experimental section, the virtual database will contain 4 representative images for each texture class. Each image will correspond to a different direction of light from that used for training and testing. Note that reference and training images could be the same, however we want to add robustness to the features we will extract, on the basis that the test images are bound to be of different illumination than the reference images. It is important to remember that another objective of my work, apart from performing texture classification, is to provide an estimation of the light direction used to capture the test images. Specifically, we focus on the tilt angle estimation, assuming that all the images have been captured with the same elevation angle. With the idea of performing an accurate evaluation of the illuminant tilt angle estimation, we will create the images of the virtual database under 4 illuminant tilt angles which will be different from those used for the training and test images. This way, we will ensure that there is no correspondence between the light directions in the images in the virtual databases and the training and test images. Figure 5.5 shows the 4 images created in the virtual database for one texture.

5.3.2 Recognition Procedure

This section details each step in the recognition procedure. Basically, as it is shown in Figure 5.6, our recognition procedure is divided into two main phases: the learning process and the classification process. In the learning process, each texture class presented in the virtual database is modelled by a texture feature vector. In order



(a) (b)

(c) (d)

Figure 5.5: Example of images created in the virtual database. 4 images corresponding to 4 different illuminant tilt angles: 10°, 100°, 190° and 280°, (a)-(d) respectively.

to carry out this learning task, it is necessary to perform first a feature extraction process in which different texture features are computed for each image texture. Afterwards, a feature selection and evaluation process will allow us to choose from all the computed features those which could discriminate between the different classes best. When the learning process is finished, the classification process starts. This extracts the feature vectors for the unknown image textures (test images) and assigns them, by means of a classifier, to one of the classes in the virtual database, at the same time estimating the light direction in the test images.

In the remainder of this section, we describe how feature extraction, feature selection, feature evaluation and classification can be performed. Moreover, we detail the solution adopted for each of these processes.

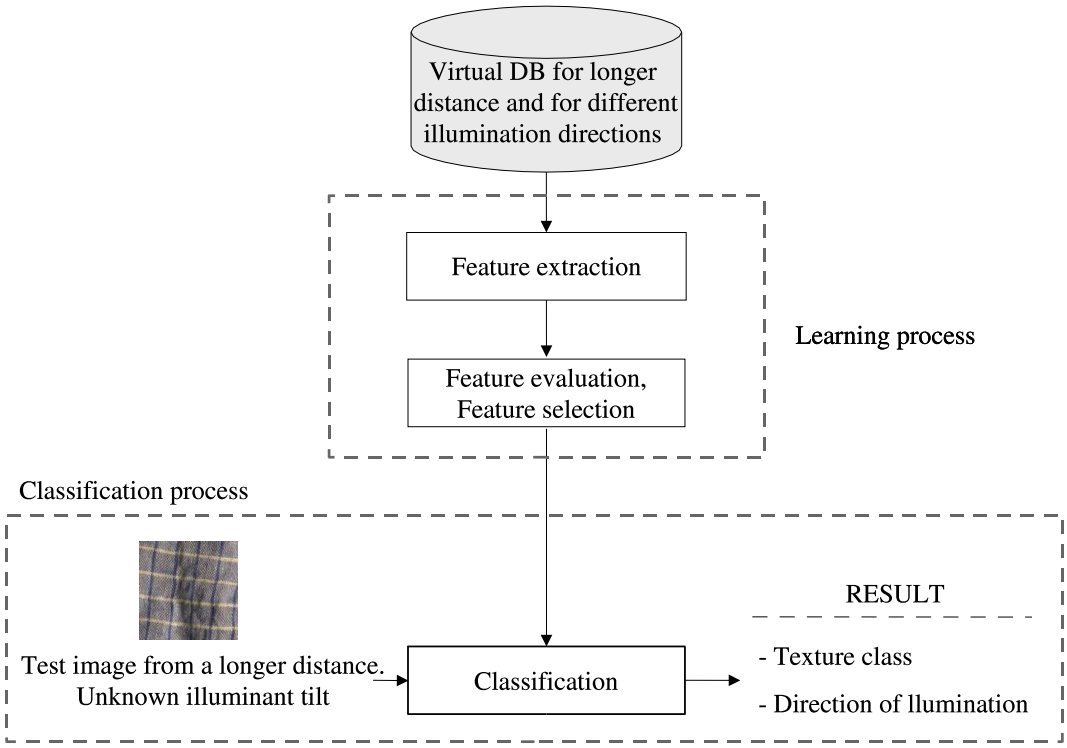


Figure 5.6: Recognition procedure. This procedure is divided into two main phases: learning process and classification process.

5.3.2.1 Texture Feature Extraction

Haralick provided a classic survey of texture measures [60]. Moreover, many types of representations and texture features have been proposed in the past few decades. Essentially, two major categories of texture measure methods have been identified: structural and statistical. This same basic classification was later adopted by other authors like Wechsler [168], Van Gool et al. [163], Grau [57], and Al-Janobi [4].

In structural methods, texture is considered as the repetition of some basic primitive patterns with certain rules of placement which are formally defined by grammars of various types. Nevertheless, since natural textures are not very regular, structural techniques are not very popular as argued by Wang and Liu [165]. On the other hand, statistical methods are based on the characterisation of the stochastic properties of the spatial distribution of grey levels in the image. Statistical approaches include a wide number of methods to extract texture measures. For example:

- **Laws's texture energy filters.** With the main desire of producing a computationally efficient method, Laws developed a coherent set of texture energy masks [90]. These masks allow us to achieve texture characterisation in two steps: first, the image is convolved with a set of small sized masks, second, statistics are computed from the outputs of these convolutions.

- **Random field models.** It is well-known that a number of random field models have been used for modelling and synthesising texture. If a model proves to be capable of representing and synthesising a range of textures, then its parameters may provide a suitable feature set for the classification of textures. Popular random field models include fractals, autoregressive models and Markov random fields. An extensive review of these approaches can be found in [139].

- **Frequency domain methods.** Some authors argue that many naturally occurring textures exhibit a combination of regularity such as approximate periodicity and variation which is hard to model using straightforward repetition or traditional statistical techniques. Hence, features related to the local spectrum have been proposed in the literature and used for the purpose of texture classification and/or segmentation. In most of these studies the relationship with the local spectrum is established through features obtained by filtering with a set of the two-dimensional Gabor filters to highlight frequency bands of two-dimensional spectra. This filter is linear and local and is characterised by a preferred orientation and spatial frequency, which cover the spatial frequency domain adequately. Roughly speaking, it acts as a local band-pass filter with certain optimal joint localisation properties in both the spatial domain and the spatial frequency domain [38].

- **Co-occurrence matrix.** The grey-level co-occurrence matrices (GLCM) are essentially two-dimensional histograms of the occurrence of pairs of grey-levels for a given displacement vector. Formally, the co-occurrence of grey levels can be specified as a matrix of relative frequencies P_{ij}, in which two pixels separated by a distance d along direction θ are found in the image, one with gray

level i and the other with gray level j. These GLCMs depend on the angular relationship between neighbouring pixels as well as on the distance between them. GLCMs are not generally used as features, rather a large number of textural features derived from the matrix have been proposed starting with the original fourteen features described by Haralick et al. [61].

Many researchers have attempted to carry out comparative studies to evaluate the performance of textural features. Weszka et al. [169] compared features derived from GLCMs on terrain images and found that co-occurrence features obtained the best result. Moreover, a theoretical comparison of four types of texture measures which Weszka et al. investigated was reported by Conners and Harlow [31]. They measured the amount of texture context information contained in the intermediate matrices of each algorithm and their conclusions were similar to those obtained by Weszka et al. Focusing on frequency domain methods, Pichler et al. [132] compared wavelet transforms with adaptive Gabor filtering feature extraction and reported superior results using the Gabor technique, although its higher computational cost was a drawback. Recently, Singh and Sharma [149] compared five different texture analysis methods in terms of recognition ability. They performed the experiments on the image benchmarks Meastex [1] and Vistex [3], concluding that features extracted from co-occurrence matrices are the best.

Summarising, the results of comparing the relative merits of the different types of features have been inconclusive and not a single method has emerged as being acceptable in all cases [138]. Comparative works have resulted in different, and sometimes contradictory, conclusions. The reason can be found in the use of different test images and evaluation or classification methods as well as some aspects related to code implementation.

It is important to note that in a previous work related to illuminant invariant texture classification, Penirschke et al. [125] developed a classification scheme based on the use of photometric stereo for the detection of surface relief and the use of Gabor features as texture measures. The production of features by filtering, however, requires the use of all points of the surface/image. The surface gradient of some of

these points may have been wrongly calculated by the photometric stereo technique and their inclusion in the feature extraction process may affect the performance of the classifier. That is why in our approach, we do not use filtering but co-occurrence matrices which allow us to work only with the pixels for which reliable information is available. In addition to this advantage, the co-occurrence matrix is an intuitive measure of texture and it is straightforward to compute. For all of these reasons it has been chosen in our recognition system as a feature extraction method.

Co-occurrence Matrix

As noted above, the co-occurrence matrix can be described as a matrix conveniently representing the occurrences of pairs of grey levels of pixels separated by a certain distance d and lying in a certain direction (angle θ) in the image.

Suppose that two grey values are separated by distance d along angle θ in an image which has m gray values. Typically, the angles are quantised in intervals of 45°: horizontal, first diagonal, vertical, and second diagonal (0°, 45°, 90°, and 135° respectively). Also, suppose that these grey values occur in this configuration of d and θ, N times in the image. The co-occurrence matrix is therefore the $m \times m$ matrix containing the number of occurrences N for each combination of grey values in the image. Note that a new co-occurrence matrix can be created for each different choice of d and θ.

Let us illustrate this with a simple example. Using the 4×4 image of Figure 5.7, we compute the co-occurrence matrices $P_{0°,1}$ and $P_{135°,1}$. Note that the subscripts indicate angle θ and distance d, respectively. As shown in the figure, the image has only 4 distinct grey levels, so the co-occurrence matrix is 4×4. For every horizontal grey level pair (horizontal in the first case in which we are computing $\theta = 0$) the number of instances of this pair in the entire image is counted up and then recorded in the co-occurrence matrix at the index corresponding to the two grey levels. Therefore, the number 6 at $P_{0°,1}(2,2)$ means that occurence $(2,2)$ is found 6 times in the original image. The elements of the co-occurrence matrix are usually normalised by dividing each entry by the total number of pixel pairs. This

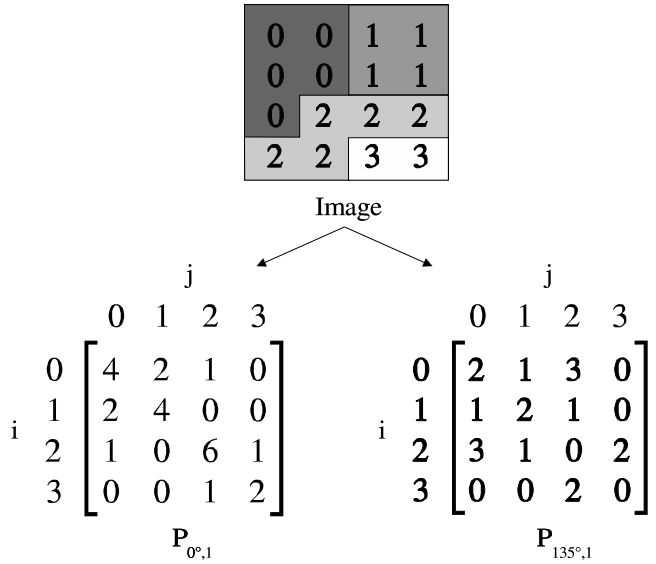


Figure 5.7: Example of co-occurrence matrix computation. The co-occurrence matrices at distance 1 and angle 0° and 135° are computed over the original image.

way, all values are between 0 and 1 and may be thought of as probabilities.

In general, a smooth texture gives a co-occurrence matrix with high values along the diagonals if d is small compared with the texture spatial variation (distance and angle). This is because the pairs of points at distance d should have similar grey levels. Conversely, if the texture is macrotexture and distance d is comparable to the texture scale, then the grey levels of points separated by distance d should often be quite different, so the values in the co-occurrence matrix should be spread out relatively uniformly.

Many researchers used statistics based on the co-occurrence matrix proposed by Haralick et al. [61] in their work [31, 46, 169]. In our approach, we will also use three of these tipical texture features shown in Table 5.1.

The main drawback of the GLCM technique is the dependence on the parameters used. The number of matrices needed to obtain good texture features is related to the angle and distance used. This number can be potentially enormous. As it will be seen in the experimental section, the co-ocurrence matrices will be used to extract features as contrast, homogeneity and energy for 20 different values of distance d

Table 5.1: Some of the most known texture measures computed from the co-occurrence matrix. $P_{\theta,d}(i,j)$ is the probability that grey level j follows grey level i at distance d and angle θ.

Contrast= $\sum_{i,j}(i-j)^2 P_{\theta,d}(i,j)$	This is the measure of the amount of local variation in the image. A low value results from uniform images whereas images with greater variation produce a high value.
Homogeneity= $\sum_{i,j} \frac{P_{\theta,d}(i,j)}{1+\|i-j\|}$	This is high when the GLCM concentrates along the diagonal. This occurs when the image is locally homogeneous.
Energy= $\sum_{i,j}[P_{\theta,d}(i,j)]^2$	This is also known as Uniformity or Angular second moment. It is high when the GLCM has few entries of large magnitude and low when all entries are almost equal.

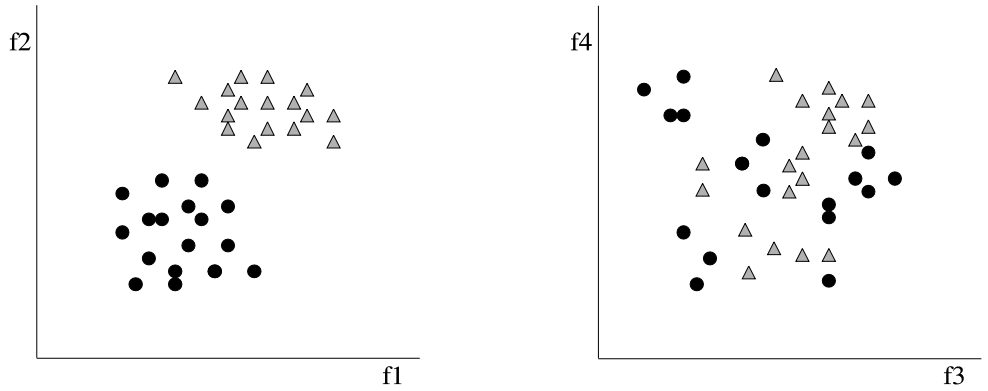
and for 4 different angles: 0°, 45°, 90°, 135°. This means that the co-ocurrence matrices are implemented in an anisotropic way. Note also that we compute in all 240 texture features (3 features × 4 directions × 20 distances).

Obviously, this is a large number of features to be used for a classification process. Moreover, some of these features can introduce redundant information, implying a decrease in classification accuracy. Therefore, we propose from all of these features, to select those which could best discriminate between the different classes.

5.3.2.2 Feature Selection

The aim of this section is to introduce the feature selection process as well as summarise the most relevant methods.

The problem of choosing the best features to discriminate between the different classes is known as feature selection. In the last few decades, this has been analysed from various points of view: statistics [47, 115], pattern recognition [39, 82, 75] and machine learning [16, 86]. These areas define the problem of feature selection as follows: given a set of d features, select a subset of size m (where $m < d$)



(a) (b)

Figure 5.8: Illustrative example of feature spaces. (a) Feature space in which the samples of each class are clearly separate. (b) Feature space in which the samples of each class overlap.

which obtains the highest value of a criterion function, assuming that a higher value indicates a better feature subset. In other words, choosing those features which allow feature vectors belonging to different classes to occupy compact and disjoint regions in a d-dimensional feature space. Figure 5.8 shows two examples of feature spaces. Note that the feature space in Figure 5.8.(a), composed of features f1 and f2, is preferred over that in Figure 5.8.(b), composed of features f3 and f4. The reason is simple: the first feature space has more discriminative power than the second in which the samples of the two classes overlap.

The simplest way of finding the best feature set is to perform an exhaustive search. However, this may be too costly and practically prohibitive even for a medium sized-feature set size. In order to solve this problem, different methods have been proposed in the literature in an attempt to reduce computational complexity by compromising performance. Obviously, these methods affect how well the patterns fit reality, and these become a central issue in the learning process. The importance of feature selection in a broader sense is due to its potential for speeding up the processes of learning and classifying objects, reducing the cost of recognition and, in some cases, improving the quality of classification.

Many authors have proposed different classifications of feature selection methods.

For instance, Siedlecki and Sklandsky [146] discussed the evolution of feature selection methods and grouped them into past, present, and future categories. Doak [40] identified three categories of search algorithms: exponential, sequential and randomised. Jain and Zongker [75] presented a taxonomy of available feature selection algorithms, dividing them into those based on statistical pattern recognition techniques and those using artificial neural networks. In a more recent survey, Kudo and Sklansky [88] focused on a comparative study of algorithms for large scale feature selection.

In what follows, we briefly review the best known feature selection methods, classifying them into three broad categories according to the generation procedure: those which perform a complete search, heuristic search, or a stochastic search.

- **Complete search.** The first algorithm which can be used to carry out a complete search and find an optimal subset is the exhaustive search. This consists of exploring all possible feature subsets (feature combinations) and selecting the best one. However, it has exponential complexity in the number of features and is frequently prohibitively expensive to use.

  Some heuristic algorithms, which introduce backtracking in the search, are used to reduce this exhaustive search without jeopardising the chances of finding the optimal subset. For instance, the Branch and Bound (B&B) algorithm, proposed by Narendra and Fukunaga [117], can be used to find the optimal subset of features much more quickly than through an exhaustive search. However, the main drawback of this algorithm is the requirement of a criterion function which has to be monotone. This means that the performance of a feature subset should improve whenever a feature is added to it. Furthermore, this algorithm is still impractical for problems with very large feature sets because the worst case complexity of this algorithm is exponential.

- **Heuristic search.** In each iteration of this generation procedure, all remaining features yet to be selected or rejected are considered for selection or rejection. There are many variations of this simple process and all of them

find a suboptimal solution. The search space order is, in general, polynomial, so these procedures are very simple to implement and very fast.

In this generation procedure we find popular methods such as Sequential Forward Selection (SFS) and Sequential Backward Selection (SBS) [81]. These methods begin with a single solution and iteratively add or remove features until some termination criterion is achieved. These are the most commonly used methods for performing feature selection. SFS, also called the "bottom-up" method, starts from the empty set and generates a new subset in each iteration by adding the feature which best improve the quality of the selected subset (that is measured by some evaluation function). SBS, also called the "top-down" method, starts from the complete feature set and generates a new subset in each iteration by discarding a feature selected by some evaluation function. Both methods suffer from the so-called "nesting effect". This means that in the case of the "top-down" search, the discarded features can not be reselected, while in the case of the "bottom-up" search, the features once selected can not be later discarded. Note that since they do not examine all possible subsets, these algorithms are not guaranteed to produce the optimal result. With the aim of preventing this problem, Stearns [155] developed the "Plus l-take away r" (PTA(l,r)) search method. This method goes forward l stages by adding l features by SFS and goes backward r stages by deleting r features by SBS and repeating this process. The main drawback of this method is that there is no theoretical way of predicting the values of l and r to achieve the best feature subset. Research in this direction was concluded by introducing the generalisation of SFS and SBS (GSFS/GSBS) and the GPTA(l-r) algorithms proposed by Kittler [81]. Pudil et al. [134, 135] updated this study by introducing the two "floating" selection methods, SFSS and SFBS. Besides avoiding the nesting of features, one of their distinctive characteristics is that, during the backtracking process, the values of the criterion function are always compared only with those related to the same cardinality of the feature subset. Somol et al. [154] presented a more sophisticated version of "classical" floating search algorithms (adaptative floating methods) which attempt to remove

some of their potential deficiencies and facilitate the finding of a solution even closer to the optimal one. The adaptative floating search is called "adaptive" because of its ability to adjust the limit under which the actual generalisation level can be automatically set.

- **Stochastic search.** This search procedure is based on a random generation. Although the search space is exponential, these methods typically search fewer subsets by setting a maximum number of possible iterations.

  The best known method is based on the use of a genetic algorithm (GA). In 1989 Siedlecki and Sklansky introduced the use of GA for feature selection [147]. Since then, many works have used this strategy to perform feature selection [160, 55, 175, 161, 145]. In GA, a feature subset is represented by a binary string with length n, called a *chromosome*, with a zero or one in position i denoting the absence or presence of feature i. Each chromosome is evaluated for fitness through an optimisation function in order to survive to the next generation. Unlike classical hill-climbers, it does not evaluate and improve a single solution but, instead, analyses and modifies a "population" of solutions at the same time. The optimisation process is carried out in cycles called generations. During each generation, the population of chromosomes is maintained and evolved by two operators: (i) crossover: where parts of two different parent chromosomes are mixed to create an offspring, and (ii) mutation: where the bits of a single parent are randomly perturbed to create a child, increasing the variability of the population. In each generation only a few of the best chromosomes survive to the next cycle of reproduction. When the maximum number of cycles is achieved, a set of feature subsets is provided. The main drawback of this feature selection method is that it needs a proper assignment of values of different parameters: initial population size, crossover rate and mutation rate.

We have seen that quite a few methods have been proposed for feature selection. However, choosing the proper method for a particular problem is a difficult task, as Pudil and Novovicová argued in [136]. The best choice depends on a large number of

conditions related to the user's knowledge of the problem. The following questions describe perhaps the most important aspects we must keep in mind when choosing a feature selection method: What exactly is your aim? What is the dimensionality of your problem? What is your a priori knowledge? Do you have a criterion for subset evaluation appropriate to your knowledge of the problem? In Table 5.2, we provide recommendations for each feature selection method with the aim of facilitating the election of one of them for a particular problem.

As it will be seen in the experimental section, we will use a heuristic search algorithm in order to perform our feature selection process. Basically, the reason for this choice can be found in the reduction of the search cost to that of polynomial complexity. Moreover, the heuristic search implies a fast generation procedure and, as it has been mentioned, a simple implementation.

5.3.2.3 Feature Evaluation

The process of choosing an appropriate criterion function is known as feature evaluation. The main goal of this process is to measure the "goodness" of a subset produced by some generation procedure. In this context, goodness means the capability of a feature subset to distinguish the different class labels and the ability of providing compact and maximally distinct descriptions for every class. Geometrically, this constraint can be interpreted to mean that this feature takes on (i) nearly identical values for all samples in the same class and (ii) different values for all samples of the other classes. This is illustrated in Figure 5.8.

Feature evaluation has been studied for many years and different measures have been proposed. Ben-Bassat [12] grouped the existing criterion functions up to 1982 into three categories: information measures, distance measures, and dependence measures. Another viewpoint was introduced by Langley [123], who grouped different feature selection methods into two broad groups, *wrapper* and *filter*, based on their dependence on the feature evaluation used.

Table 5.2: Feature selection methods.

Method	Description	Comments
Exhaustive Search	Evaluates all possible subsets.	Guaranteed to find the optimal subset; not feasible for even moderately large values of m and d.
Branch and Bound Search (B&B)	Only a fraction of all possible feature subsets need to be enumerated to find the optimal subset.	Guaranteed to find the optimal subset provided the criterion function satisfies the monotonicity property; the worst case complexity of this algorithm decision is exponential.
Sequential Forward Selection (SFS)	Selects the best single feature and then add one feature at a time which in combination with the selected features maximizes the criterion function.	Once a feature is retained, it can not be discarded; computationally attractive; it examines only $(d-1)$ possible subsets.
Sequential Backward Selection (SBS)	Starts with all the d features and successively deletes one feature at a time.	Once a feature is selected, it cannot be brought back into the optimal subset; requires more computations than sequential forward selection.
"Plus l-take away r" PTA(l,r) Selection	First enlarge the feature subset by l features using forward selection and then delete r features using backward selection.	Avoids the problem of feature subset "nesting" encountered in SFS and SBS methods; need to select values of l and $r(l > r)$.
Generalized Sequential Forward Selection (GSFS) and Generalized Sequential Backward Selection (GSBS)	SFS and SBS are generalized in such a way that a number (n) of features are evaluated at the same time and the best n features subset is chosen for addition or deletion.	Their performance is a little better than SFS and SBS; they are effective when the number of features is very small; they are very time consuming.
Sequential Forward Floating Search (SFFS) and Sequential Backward Floating Search (SBFS)	A generalization of "plus-l take away-r" methods; the values of l and r are determined automatically and updated dynamically.	Provides close to optimal solution at an affordable computational cost for small-scale and medium-scale problems; in order to avoid excessive computation, the maximum level of backtracking could be constrained.
Adaptive Sequential Forward Floating Search (ASFFS) and Adaptive Sequential Backward Floating Search (ASBFS)	They are called adaptive because of their ability to adjust the limit under which the actual generalization level can be automatically set.	Adaptive Sequential Floating Search yields better results than classical Sequential Floating Search; eliminates the lack of an explicitly specified termination condition.
Genetic Algorithms (GA)	A feature subset is represented by a binary string with length n, called a chromosome, with a zero or one in position i denoting the absence or presence of feature i; a population of chromosomes is maintained and evolved by two operators of crossover and mutation.	It is very useful to find a compromise between maximum criterion value and minimum size subset in large scale; Needs proper assignment of values to different parameters: iterations, initial population size, crossover rate, and mutation rate.

- The idea behind the wrapper approach is simple: a classifier is used as a criterion function in order to obtain a metric (classification accuracy) for guiding the feature subset selection.
- On the other hand, the filter approach attempts to assess the merits of features from the data alone and the selection is performed independently of the classifier.

The filter approach normally uses a criterion function which is simple and fast to compute, so it is generally computationally more efficient. Its major drawback is that an optimal selection of features may not be independent of the classifier used. Moreover, the wrapper approach provides a better estimate of accuracy for a feature subset, but involves a computational overhead by executing the classifier.

A recent research made by Dash and Liu [37] considering the latest developments, divided the evaluation functions into five categories: distance, information (or uncertainty), dependence, consistency and classifier error rate.

- The first category called distance, is also known as separability, divergence, or discrimination measure, where, for a two class problem, a feature A is preferred to another B if A induces a greater difference between both classes. If the difference is zero, it is impossible to distinguish the two classes. An example is the Euclidean distance measure.

- The information (or uncertainty) measures determine the information gained from a feature. Therefore, one feature is preferred to another if the information gain is greater. An example of this category is the entropy measure.

- The dependence measures or correlation measures qualify the ability of predicting the value of one variable from the value of another. The coefficient is a classical dependence measure and can be used to find the correlation between a feature and a class. If the correlation of feature A with a class is higher than the correlation of feature B, then feature A is preferred to B. A slight variation of this is to determine the dependence of a feature on other features. This value indicates the degree of redundancy of the feature.

- The consistency measures find the smallest sized subset which satisfies the acceptable inconsistency rate set by the user. This rate is usually calculated as the sum of inconsistency counts, divided by the total number of samples. Two samples are considered inconsistent if their attribute values are the same (they match).

Table 5.3: A comparison of evaluation functions provided by Dash and Liu [37].

Evaluation function	Generality	Time complexity	Accuracy
Distance measure	Yes	Low	–
Information measure	Yes	Low	–
Dependence measure	Yes	Low	–
Consistency measure	Yes	Moderate	–
Classifier error rate	No	High	Very High

- The last category is the classifier error rate. This corresponds with the idea of wrapper methods in which the accuracy of the classifier is used as an evaluation measure.

Obviously, an optimal subset in feature selection is always relative to a certain criterion function. Hence, choosing an optimal subset using one criterion function may not be the same as using another one. Therefore, a good evaluation function is a key factor in choosing the best feature subset for a recognition process. In their work Dash and Liu compared the evaluation functions using different properties; (*i*) generality: how suitable is the selected subset for different classifiers, (*ii*) time complexity: time taken for selecting the subset of features, and (*iii*) accuracy: how accurate is the prediction using the selected subset. Table 5.3 shows a comparison of evaluation functions realised by Dash and Liu. The "–" in the last column means that nothing can be concluded about the accuracy of the corresponding evaluation function. In all of these cases, the accuracy of the evaluation functions depends on the data set and the classifier used. They discovered an unsurprising trend, the more time spent, the higher the accuracy. The table also shows us which measure should be used under different circumstances such as with time constraints, when given several classifiers to choose from, etc.

Classification error rate and distance measures are usually chosen as a criterion function in many works. It is difficult to estimate the correct recognition rate of a classifier on the basis of a limited number of training samples. This is one reason why most comparative studies on feature selection used distance as a criterion function.

Nevertheless, a recent study made by Kudo and Sklansky [88] compares feature selection algorithms using a recognition rate measure. On the other hand, if it is known which classifier will be used in the problem under consideration, the best criterion is, in general, the classification error rate because it makes feature selection procedures specific for the classifier and the sizes of the training and test sets used. As it will be seen in the experimental section, the classification error rate will be used as a feature evaluation in our feature selection process.

5.3.2.4 Classifiers

Up to now, we have described the learning process of our recognition system. We have seen that each texture class presented in the virtual database is modelled by a texture feature vector extracted from the co-occurrence matrices. The necessity of performing feature selection/evaluation in order to choose the best feature subset which distinguishes the different classes best has also been stated. At this point, when the learning process is finished, the classification process proper begins. The objective of this process is to assign, by means of a classifier, the feature vectors of the unknown image textures (test images) into one of the classes of the virtual database.

It is well-known that a classifier can be designed using a large variety of possible approaches. In practice, the choice of a classifier is a difficult problem and is often based on which classifier happens to be available or best known to the user. As Jain et al. presented in [73], there are three different approaches to design a classifier.

- The simplest and most intuitive approach to classifier design is based on the concept of similarity: patterns which are similar should be assigned to the same class. Therefore, once a good metric has been established to define similarity, patterns can be classified by template matching or the minimum distance classifier using a few prototypes per class. The choice of metric and prototypes is crucial to the success of this approach.

- The second main concept used for designing pattern classifiers is based on the probabilistic approach. These classifiers are commonly known as Bayesian

classifiers. They are based on probabilistic information on the populations from which a sample of training data is to be randomly drawn. Randomness in sampling is assumed and it is necessary for a better representation of the sample of the underlying population probability function. There are different approaches to Bayesian classifiers. See [114] for a survey of these classifiers. Among them, the optimal Bayes decision rule, used in many works, assigns the unknown patterns to the class with the maximum posterior probability, taking the class conditional probabilities and the a priori probabilities into account.

- The third category of classifiers is to construct decision boundaries. A classical example of this type of classifier is the linear discriminant analysis (LDA) [114] which uses linear boundaries between data distributions to discriminate between samples. Another example is the single-layer perceptron where the separating hyperplane is iteratively updated as a function of the distance of the misclassified patterns from the hyperplane. It is important to note that neural networks themselves can lead to many different classifiers depending on how they are trained.

  A decision tree is a special type of classifier which is trained by an iterative selection of individual features which are most salient at each node of the tree. The criteria for feature selection and tree generation include the information content, the node purity or Fisher's criterion (a linear discrimination which minimises the MSE between the classifier output and the desired labels). During classification, only the features needed for the test pattern under consideration are used, so feature selection is implicitly built-in. The main advantage of the tree classifier, apart from its speed, is the possibility of interpreting the decision rule in terms of individual features.

Nearest Neighbour Classifier

Considering the first approach described above, we use the Nearest Neighbour classifier (NN) to perform our texture classification.

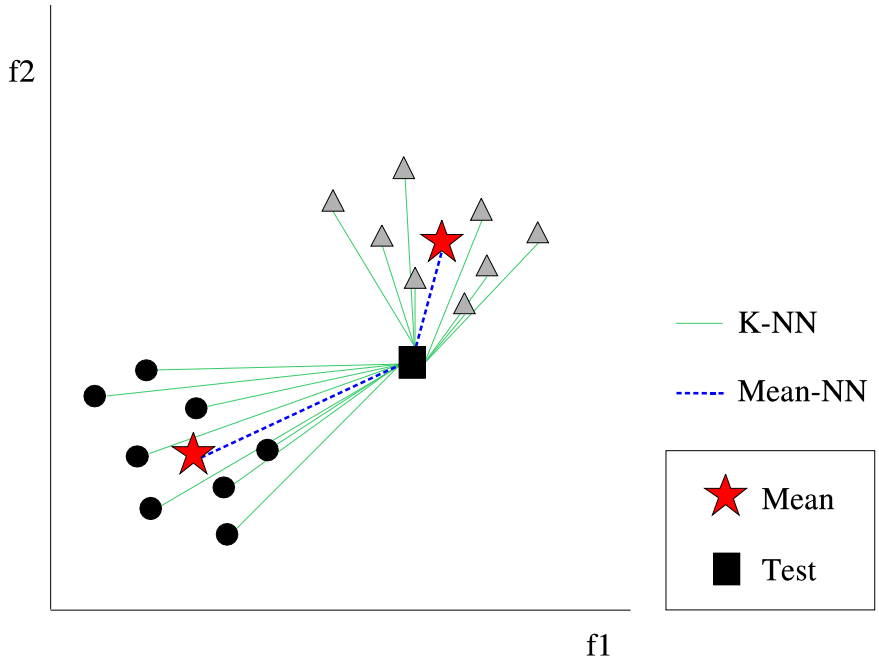


Figure 5.9: Example of K-nearest neighbour (K-NN) and mean nearest neighbour (mean-NN) classifier. Note that the test sample is classified in both cases to the same class.

Nearest neighbour methods have been used as an important pattern recognition tool [47, 89, 114]. The one-nearest neighbor (1-NN) classifier is the most natural classification method we can think of. It consists of comparing an unknown observation x with all the N cases in the training set. N distances between a pattern vector x and all the training patterns are calculated and the label information contained in the training set with which the minimum distance results is assigned to the incoming pattern x. The K-nearest neighbour (K-NN) is the same as the 1-NN rule except that the algorithm finds K nearest points within the points in the training set from the unknown observation x and assigns the class of the unknown observation to the majority class in the K points. Note that another criterion has to be defined in case of a tie. Another common approach is the nearest mean classifier (mean-NN) in which each class is represented by a single prototype (feature vector) which is the mean vector of all the training samples in that class. Figure 5.9 intuitively illustrates how the nearest neighbour classifiers (K-NN, Mean-NN) classify a test sample.

The nearest neighbour classifiers are simple and robust. The most straightforward one-nearest neighbour decision rule (1-NN) is conveniently used in many works as a benchmark for all the other classifiers since it appears to provide a reasonable classification performance in most applications. The 1-NN classifier only has to define the distance metric used to find the nearest neighbour. In general, Manhatan or Euclidian distance are commonly used. The main drawback of this classifier is the high computational cost.

In terms of typical texture classification the different approaches of nearest neighbour classifiers perform in a similar way. Nevertheless, when dealing with texture classification independent from light direction, things are different. Changes in the imaging geometry and lighting conditions can significantly alter the appearance of the surface, implying significant variations in the image texture and, therefore, in the feature vectors. This fact causes the data to suffer high intraclass variation. In other words, the feature values for the same texture class are not stable. Hence, the data distribution for each class does not make a compact cluster. It is this intraclass variation which introduces critical misclassification rates in the texture classification process. In that case, it is reasonable to think that the mean-NN is not an adequate approach since the variation of a texture can provide mean values which do not reflect the nature of the data (i.e. one texture with more than one cluster). Analysing these facts, in our experimental results we use a 1-NN classifier in which an unknown test image is compared with all the cases contained in the virtual database and is classified with the one which has minimum distance. At the same time it performs texture classification, this classifier allows us to provide an estimation of the illumination direction of the test images directly. It is important to remember that for each texture class, different images corresponding to different light directions are available in the virtual database. Therefore, using the 1-NN classifier, as well as assigning the test image to the texture class to which it belongs, the light direction of the unknown test image is approximated by one of those contained in the virtual database. This will be specified in more detail in the experimental section related to texture classification and light estimation.

5.4 Summary

A novel methodology of texture classification under varying imaging geometries (i.e. light direction and camera distance) was presented in this chapter. Basically, the main contributions of this method are: (i) to overcome the problem of classifier failure induced by varying imaging properties, (ii) to integrate the texture prediction framework proposed in Chapter 4 into a texture classification system, and (iii) to provide the capability of guessing an approximate direction of the light used to capture the test images.

This chapter has been structured in two blocks. In the first, works on the topic of texture classification under varying imaging geometries were reviewed. Moreover, we discussed the different ways of realising this texture classification system, concluding that a model-based approach is a potential solution to the problem. In the second block, our model-based approach for texture classification was presented, describing the general scheme and detailing each step individually.

Roughly speaking, the main purpose of our method has been to train a classifier on a selection of images of the same surface obtained under different imaging geometries. Hence, variability is modelled for each surface texture. The main reason for this choice is our interest in exploiting the capability of colour photometric stereo to render images of a surface under novel lighting conditions as well as the capability of the image prediction method proposed in Section 4.3.1 of creating images of a surface texture under novel imaging geometries such as camera distance.

The whole procedure of the proposed **model-based texture classification system** was divided into two main phases:

- **Virtual database creation.** The aim of this phase is to create, by means of the surface texture information and the image prediction method, a "virtual" database of image textures against which we compare the unknown test images in order to classify them.
- **Recognition procedure.** The recognition procedure was divided into two steps: the learning process and the classification process.

- The learning process has the goal of modelling each texture class by means of a representative feature vector.
- The classification process has the goal of classifying an unknown test image into the texture class it belongs.

In this chapter, we have seen how the learning process requires the application of other processes. First, the feature extraction process is needed to extract different texture features for each image in the virtual database. Different techniques to extract textural properties were reviewed and, from them, the **co-occurrence matrix** was chosen to compute the texture features used in our classification system. We also explained that it is very important to choose from all the computed features those which can distinguish the different classes best. For this reason, we analysed and discussed various methods which allow us to perform the feature selection and feature evaluation process.

When the learning process is finished the classification process starts, extracting first the feature vectors for the unknown image textures (test images) and assigning them by means of a classifier into one of the classes of the virtual database. We described how the **Nearest Neighbour classifier** is used to perform the texture classification and, at the same time, to provide the approximate direction of the light used to capture the test images.

In the next chapter, we will test and evaluate this texture classification approach over a large set of surface textures with different properties (i.e. smooth surfaces, rough surfaces and directional surfaces). Different experiments will be performed and, after discussing them, we will extract the corresponding conclusions.

Chapter 6

Experiments

The proposed prediction methods, as well as the model-based texture classification system, are tested and evaluated in this chapter. A set of real surface textures containing a wide variety of relatively smooth and very rough surfaces are used in this thesis as our image database (experimental data). Different experiments and error measures are used in order to carry out an exhaustive evaluation. The validity of the texture classification system is demonstrated by classifying texture images captured under imaging geometries different from the reference images in the database. The process of recognition allows us also to guess the approximate direction of the light used to capture the test images.

6.1 Introduction

The aim of this chapter is to assess the performance of the methodology presented in this thesis. As it has been seen in previous chapters, there are certain properties of the theory which must be accurately analysed. For instance, the ability of the colour photometric stereo technique to use images captured with a certain light direction in order to predict novel images referring to the same camera distance, but with different light directions. It is also important to analyse the abilities of the prediction methods presented in Section 4.3 in order to predict the surface/image information when seen from a longer distance. An exhaustive evaluation of these prediction methods is presented in this chapter. As it will be seen, different experiments and

error measures are used to extract conclusions and discuss them.

On the other hand, we must also evaluate the model-based texture classification system proposed in Section 5.3. Texture classification results obtained on a broad set of real surface textures are presented and analysed. Moreover, results related to the estimation of the direction of the light are also provided and discussed.

In order to obtain all these experimental results, it is necessary to define which experimental data will be used. It is well-known that there are many texture image databases available to perform and evaluate a typical texture classification system: the Brodatz album [18], the Meastex database [1] and the Vistex database [3]. Although there are various databases of images to test texture classificaton algorithms invariant to the direction of illumination, there is no database of images to test algorithms for changes of camera distance. For example, the "Photometric Image Databases" from the Texture lab at Heriot-Watt University [2] are not suitable to our purposes because they do not provide photometric sets of images captured from different distances. Moreover, the "Columbia-Utrecht database" established by Dana et al. [35] is also unsuitable because the light was held constant while the viewpoint and orientation of the samples were varied during data capture. Obviously, in order to perform the set of experiments which allows us to evaluate and extract conclusions for our proposals, an appropriate experimental data set is required. Therefore, acknowledging the lack of a texture database which accomplishes our requirements, we opted to build our own image database which provided all the information needed for our purposes.

The main contributions of this chapter can be summarised with the following points:

- The presentation of the image database of textured surfaces used in this thesis. This database contains the required images to apply the photometric stereo technique correctly, as well as those images which should be used as training and test images.

- The evaluation of results achieved by our prediction methods (direct image prediction and surface prediction), providing a comparison between them.

- The evaluation of results obtained on texture classification under varying geometries, such as light direction and camera distance.

The remainder of this chapter is structured in the following blocks. In the following section we describe the experimental data. In Section 6.3 and 6.4 all the experimental trials are analysed and discussed individually. Finally, conclusions are drawn in Section 6.5.

6.2 Experimental Data

Twenty five physical texture samples were used throughout the experimental trials presented in this thesis. Basically, we used real textures commonly found in our every day life. For example, simple textures based on different cloth, paper, etc, and other textures formed by repeating primitives of foods such as beans, lentils, spaghetti, chips, etc.

For each texture sample two photometric sets composed of 4 images each were available. The two sets were captured at two different distances; distance A and distance B (longer than A). All photometric stereo sets consisted of 4 images lit at an elevation angle of 55° with 4 different tilt angles 0°, 90°, 180° and 270°. Figure 6.1 illustrates the scheme of the platform used to capture the experimental data. In addition to the 4 images used in the photometric sets, different images for each surface taken at the two distances were captured for testing. We call these sets of images tA and tB. Thirteen surfaces were captured using 12 illuminant tilt angles between 0° and 360° incremented by steps of 30°. The remaining surfaces were captured using 24 illuminant tilt angles between 0° and 360° incremented by steps of 15°. All surfaces were lit at an elevation angle of 55°. It is important to state that the lights or the object position were not changed when the distance of the sensor was changed. For fifteen texture samples, in addition to these images, another set of images including the photometric images and the test images were captured at distance C (longer than B).

Virtual images constructed from photometric sets of the same distance are re-

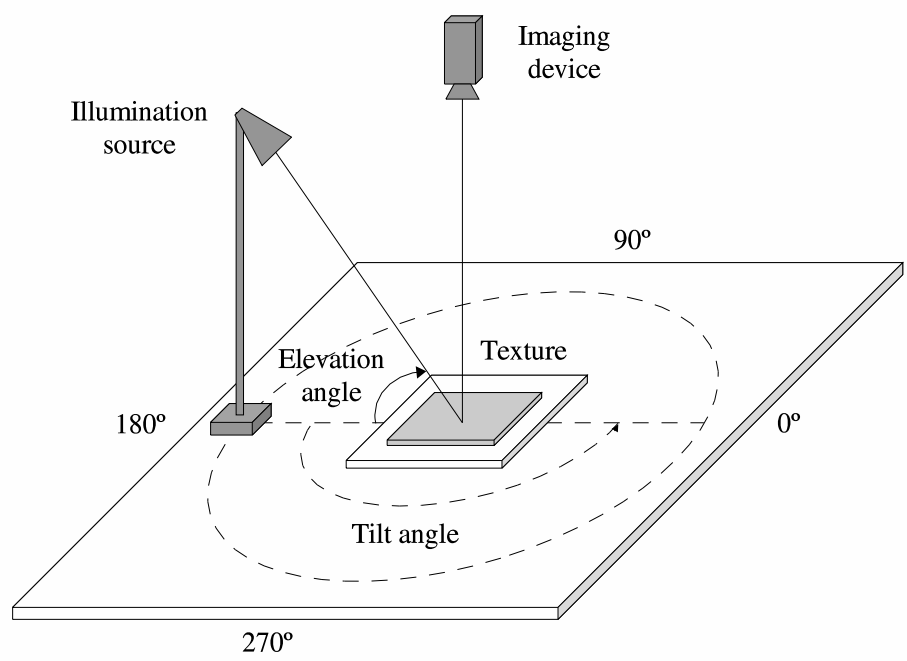


Figure 6.1: Scheme of the platform used to capture the experimental data. Different images for each texture were captured under different illuminant tilt angles and a fixed elevation angle of 55°.

ferred to as images AA, BB and CC. Virtual images constructed from photometric set A for distance B are referred to as predictions AB. In a similiar way, virtual images constructed from photometric set A for distance C are referred to as predictions AC. In order to distinguish between the direct image prediction method of Section 4.3.1 and the image prediction method via the surface prediction of Section 4.3.2, we shall refer to them as prediction $imaAB$ and $surAB$ respectively. Figure 6.2 summarises the notation described above.

Figure 6.3 shows one image for each surface texture captured from distance A and one from distance B. These images include two major groups of surface textures:

- One group of surfaces consists of a wide variety of relatively smooth surfaces which may be further divided into:
  - Isotropic surfaces and
  - Directional surfaces.

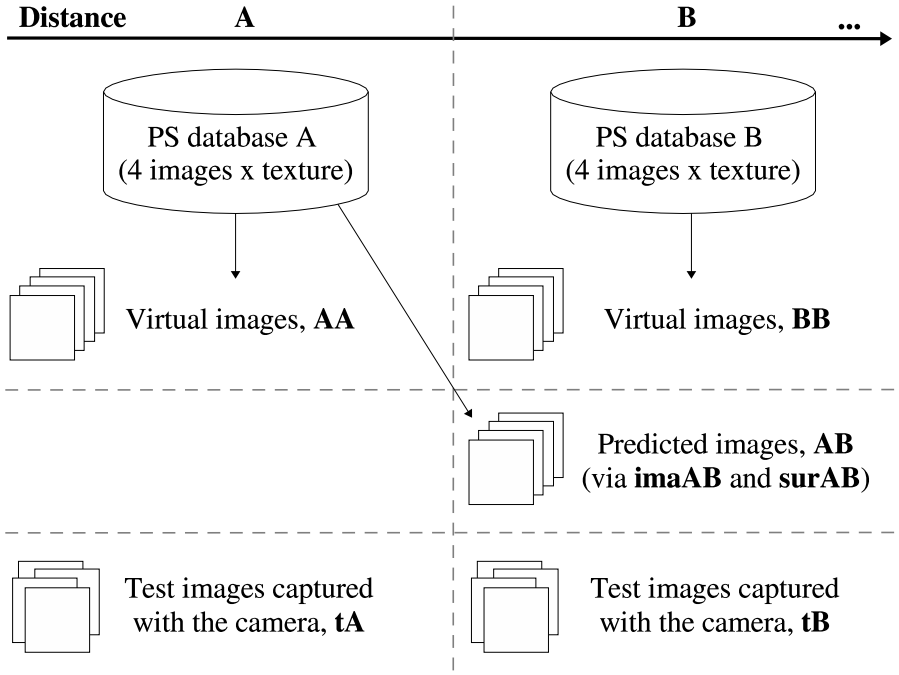


Figure 6.2: Virtual images constructed at the same distance are referred to as images AA and BB. Virtual images contructed from photometric set A with distance B are referred to as predictions AB. Images captured for testing at distances A and B are called tA and tB respectively.

- The other group of surfaces consist of a variety of very rough surfaces, for which the assumption on which photometric stereo is based is violated, (i.e. relatively smooth surface with low roughness). We do not expect CPS to work well for such surfaces, but we included them in order to test the proposed method to the extreme.

As it was seen in Section 2.3.2, the description of a surface can be stated in different ways. For example, a single parameter may be sufficient to characterise a surface for some purposes. This is the case of the *absolute average slope ratio* ($AASR$) which provides an easy way to characterise the degree of roughness of a given surface texture. Remember that $AASR$ is calculated as

$$AASR = \frac{1}{2NM} \sum_{x}^{N-1} \sum_{y}^{M-1} |p(x,y)| + |q(x,y)|$$

where $N \times M$ is the number of points for which the (p, q) values are known. For

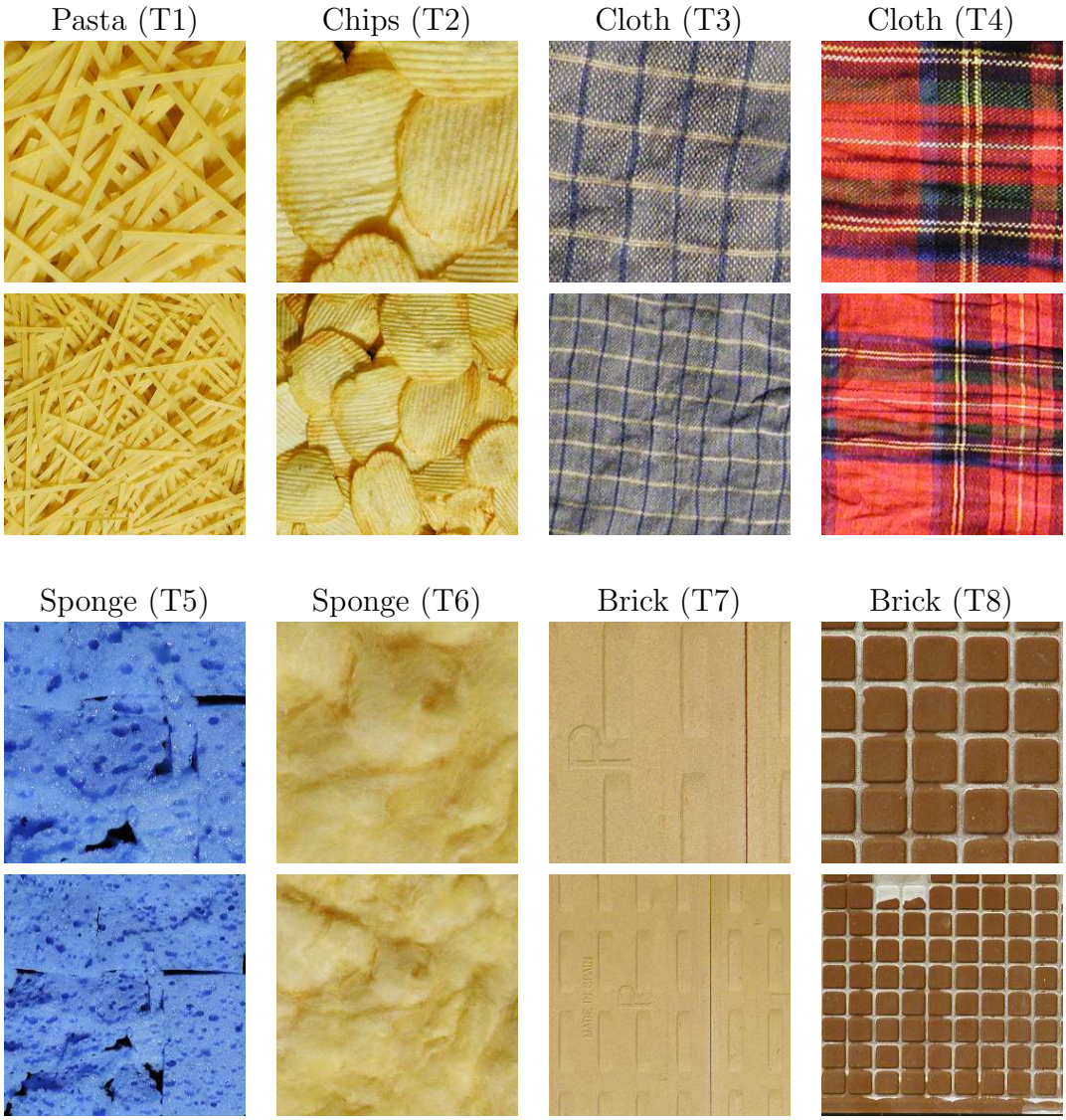


Figure 6.3: One image from distance A and one from distance B of each of the twenty five sample textures.

Figure 6.3: (continued).

Figure 6.3: (continued).

other purposes in which a more accurate description is required, statistical models such as the histograms of the values of the components of the gradient vectors may provide better descriptors. In this chapter we use both descriptors to characterise the surface shape and roughness of each texture, namely the $AASR$ parameter and the estimated probability density functions (PDFs) for the surface partial derivatives p and q (i.e. the normalised histograms of these quantities). The p-map and q-maps of the surface textures are estimated using CPS. The histograms are 256 point discrete approximations of the PDFs in the range $[-1, 1]$. These histograms represent statistical models of the surfaces. Moreover, as there is a linear relationship between surface gradient and surface height, the characteristics observed here are also valid for describing the surface height map.

After analysing all the PDFs, we identified three different types of gradient distribution corroborating the diversity of the dataset. For instance, consider three textures (T5, T14, and T15) which provide representative examples of each texture class defined above (see first column of Figure 6.4).

We observed that some surfaces, are essentially random textures as in the first example (T5), which was formed by a fracture process. Note that the PDFs for this texture could be modelled by Gaussian distributions (see first row of Figure 6.4). Similiar behaviour has been observed, for instance, for textures T1, T6, and T10. The second example, which corresponds to a directional texture, has one gradient component close to a typical distribution of a surface with a sinusoidal height profile (see the PDF of p in the second row of Figure 6.4). The q component concentrates all its values close to 0 with a maximum probability density of 0.1148. This distribution is shown out of scale in order to maintain the same scale for the y-axis so as to allow comparison between histograms. We observed that all the directional surfaces, such as textures T11, T12, and T13, show the same behaviour. Other textures (very rough isotropic surfaces) do not generally fit any particular distribution. They look like normal distributions which have been flattened out, presumably because the estimation is affected by severe shadowing, as shown in the third example of Figure 6.4. Other textures which follow this behaviour are textures T17, T18, T20, and T21.

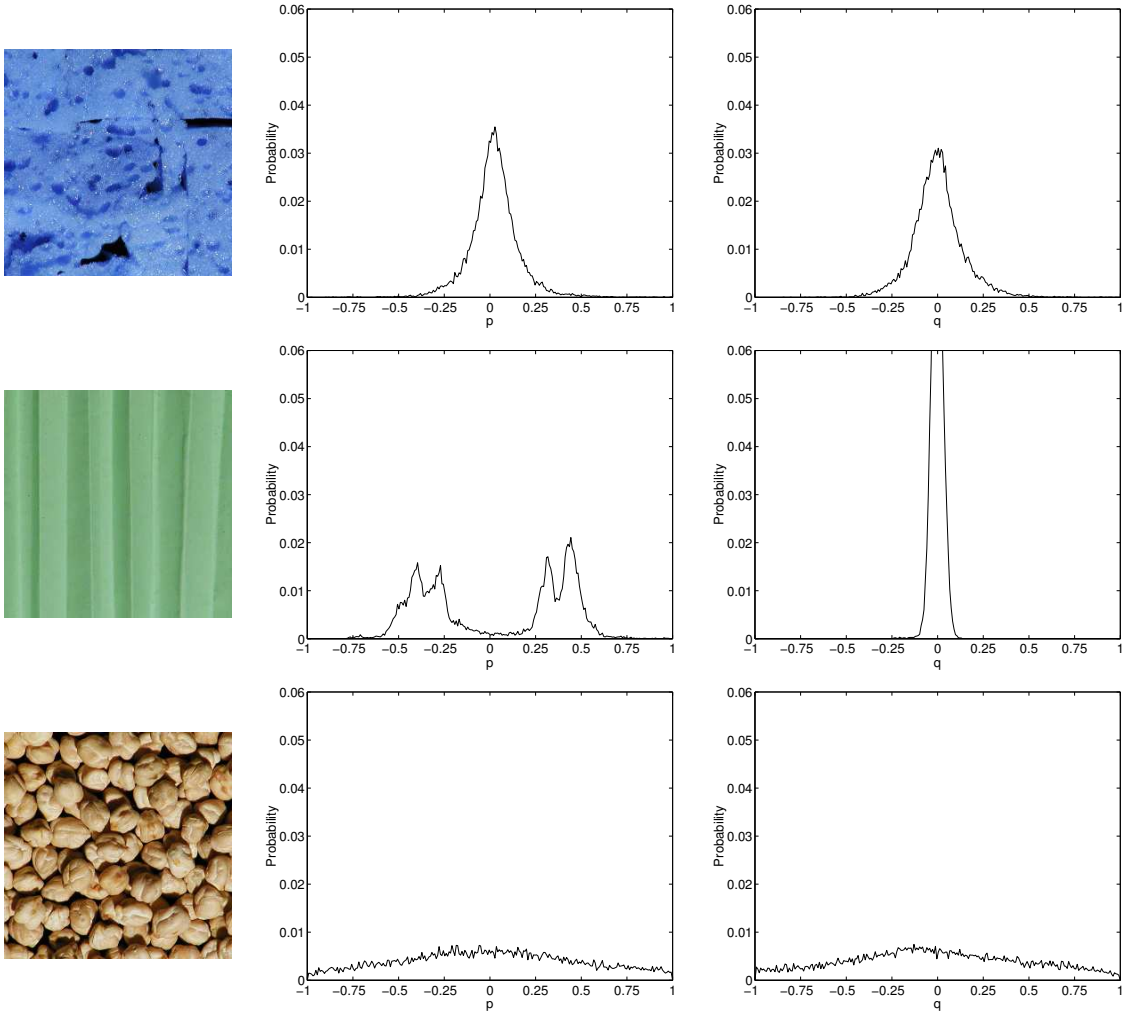


Figure 6.4: One representative image of an isotropic (T5), a directional (T14) and a very rough isotropic surface (T15). The second and third columns show the PDF representations of the surface gradients p and q.

Moreover, by analysing the $AASR$ parameter, we can also establish a simple classification scheme of surfaces based on their degree of roughness. Relatively smooth surfaces have small $AASR$ values. For example, textures T1, T5, and T12 have $AASR$ values of 0.0897, 0.1003, and 0.1129 respectively. On the other hand, very rough surfaces have large $AASR$ values. Textures T15, T20, and T22, have $AASR$ values of 0.3960, 0.3530, and 0.2470 respectively. Hence, this parameter allows us to distinguish between a smooth and a rough surface in a simple way.

Considering both surface descriptors, the $AASR$ and the estimated PDFs, we have classified our textures into one of the three groups described earlier: isotropic surfaces (from texture T1 to texture T10), directional surfaces (from texture T11 to texture T14), and very rough surfaces (from texture T15 to texture T25).

Another aspect we analysed is how chromatically distinct the surface textures are. To quantify this we used the Euclidean metric in conjunction with the Luv colour space. First, we calculate the average Luv colour value from each surface. Next, using these values, we find the distance between all pairs of surface textures, choosing the minimum value. This value indicates a measure of how distinct the surfaces are. We find that the minimum distance is 14.3471, corresponding to textures T1 and T11. We also observe that the texture most chromatically distinct from the others is texture T5, with a minimum distance value of 1273.2258.

6.3 Texture Prediction Results

After describing the experimental data, this section will focus on different experimental trials related to the texture prediction described in Chapter 4. All experiments performed have been to check various aspects of our methodology, namely:

1. To check the accuracy of the photometric stereo technique to use images captured with a certain light direction in order to predict images referring to the same camera distance but with different light directions. The results of this experiment allow us to demonstrate the validity of the photometric stereo technique in order to create virtual images of the same surface seen under different light directions.

2. To check the accuracy of *image prediction* using a photometric set captured at distance A to predict images captured at distance B ($>$ distance A). The predictions will be compared with real images captured at distance B and with images produced from a photometric set captured at the same distance B. The results of this experiment allow us to evaluate how well an *image* can be predicted from one resolution to another.

3. To check the accuracy of *surface shape prediction* using a photometric set captured at distance A to predict the surface as it would appear at the resolution of distance B and compare it with the surface reconstructed from a photometric set captured at distance B with the same light orientations as for the set at distance A. The results of this experiment are to be used as a benchmark of how well a *surface* can be predicted from one resolution to another under ideal conditions where the lighting does not change in orientation.

4. To check the accuracy of image prediction using photometric sets captured from distances A and B to predict images captured from distance C ($A < B < C$). The results of this experiment allow us to analyse the effect of using two different resolutions in the original sets to perform the same prediction when seen from distance C.

5. To check the accuracy of surface shape prediction using photometric sets captured from distances A and B to predict the surface as it will appear from the longer distance C. The results of this experiment are to be used as a benchmark of how well a surface shape can be predicted from different distances and from original sets with different resolution.

6.3.1 Experiment 1: Accuracy of Photometric Stereo

This experiment analyses the accuracy of the photometric stereo technique for creating virtual images referring to the same camera distance but under different light directions. In order to examine the performance of the photometric technique, we have decided to compare the set of test images captured from distance B (tB) with the corresponding images for the same lighting conditions created from photometric stereo information (BB). Figure 6.5 shows one image of the set tB and the corresponding image generated from photometric stereo BB for textures T3 and T22. Note that to evaluate the quality of the generated images visually is not an easy task since the differences between the original and the generated images are difficult to perceive.

In order to perform this evaluation, we propose to quantify the difference between

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

Figure 6.5: (a) and (c) Images captured with camera (tB). (b) and (d) Corresponding images generated by photometric stereo (BB).

a captured colour image and a generated one using the mean square error of colour differences computed over all the pixels. To compute the colour difference between the predicted and true colour values for a pixel, we use the Euclidean metric in conjunction with the Luv colour space [159]. This way the estimated error in colour reflects the perceived difference in colour since the Luv space is assumed to be perceptually uniform.

In the first row of Table 6.1 we give the average mean square error and its standard deviation for each type of surface used and for all twenty-five textures together for all the images captured for testing with different tilt angles (i.e. over 13 textures $\times$ 12 tilt angles + 12 textures $\times$ 24 tilt angles = 444 test images). From all these tests we observed that, in general, the MSE is larger when the degree of surface roughness increases. For example, rough surfaces such as textures T16, T20, and T21 had an average MSE for all tilt angles of 13.9003, 13.5633 and 12.7423 respectively, while relatively smooth surfaces such as T3, T7 and T10 had an average MSE of 5.7843, 4.1196 and 4.2268. We conclude that the photometric stereo technique introduces errors and the generated images are not perfect. However, for many textures we may consider these results as acceptable. We must remember that our goal is to use generated images as models in the classification process, therefore, the key question is whether the generated images are accurate enough to enhance the classification performance with respect to a naive classification system in which the same image texture captured under a particular light direction is always used as the reference image. Figure 6.6 shows an illustrative example of the MSE obtained when

Table 6.1: Quantitative assessment for each approach over all tilt angles. Average MSE and its standard deviation for the colour difference between the predicted and true values, using the Euclidean metric in conjunction with the Luv colour space.

	Isotropic		Directional		Rough		Overall	
Approach	Avg	Std	Avg	Std	Avg	Std	Avg	Std
tB vs BB	7.2346	2.9259	6.7650	1.8441	11.9886	2.9665	9.2512	3.6690
tB vs $imaAB$	8.5669	3.5692	7.5840	3.0038	12.4888	2.5783	10.1352	3.6501
tB vs $surAB$	8.6240	3.6385	7.6447	3.0654	12.6897	2.5231	10.2526	3.7084

comparing the test images tB with the images generated by photometric stereo BB, along with the MSE obtained in the naive case, in which only the captured image under the tilt angle 0° is used as a model to perform the comparison with all tB images. This testing was carried out on texture T21 which is one of the cases with the least accurate generated images with an average MSE of 12.7424. This means that is an example in which photometric stereo introduces a large error in image generation. We observe that the average MSE for the naive case is much higher, i.e. 20.1357, with a maximum MSE of 29.0195.

Note that the result of this experiment allows us to show how the use of a single image as the reference image produces a larger difference between images. This will imply errors in the computation of features and, therefore, errors in classification. Hence, we demonstrated that a model-based system using virtual images generated by photometric stereo, even if they are not very accurate, would be preferable over the use of a fixed reference image for the classification of textures under varying lighting conditions.

6.3.2 Experiment 2: Accuracy of Image Prediction when the Distance Changes

The purpose of this experiment is to evaluate the accuracy of image prediction using a photometric set captured from distance A and to predict how an image will appear from distance B. Using the image prediction method in Section 4.3.1, the image

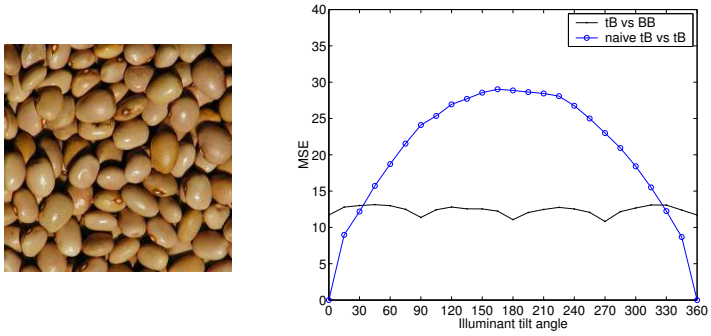


Figure 6.6: Mean square errors for texture T21 under varying tilt angles. MSE for the evaluation tB vs BB and the naive case tB vs BB, in which the process of surface recovery and rendering has been bypassed and the image at tilt=0° is compared directly with images tB.

intensities are directly predicted from the photometric information extracted from distance A. Therefore, we call these predicted images $imaAB$. However, using the surface prediction method in Section 4.3.2, only the surface gradient vectors are directly predicted but not the image intensities. As it was explained in Section 4.3.2, we generate predicted images by using the predicted gradient vectors at distance B and the average reflectance function for each surface tile assuming that our sensor and light remain spectrally unchanged. We call these predicted images $surAB$. Therefore, in this experiment we perform a comparison between the test images tB and the predicted images $imaAB$ and $surAB$.

Several results obtained over three textures are shown in Figure 6.7, where we plot the MSE over all the tilt angles. Note that for each texture, three different comparisons are shown: two curves show the results of comparing the predictions $imaAB$ and $surAB$ with the test images tB captured with the camera at distance B. The third curve shows the results of comparing the image created by using the information extracted by photometric stereo for distance B and rendering, with the test images tB. This curve presents the error produced exclusively by the photometric stereo technique, as no change in distance is involved.

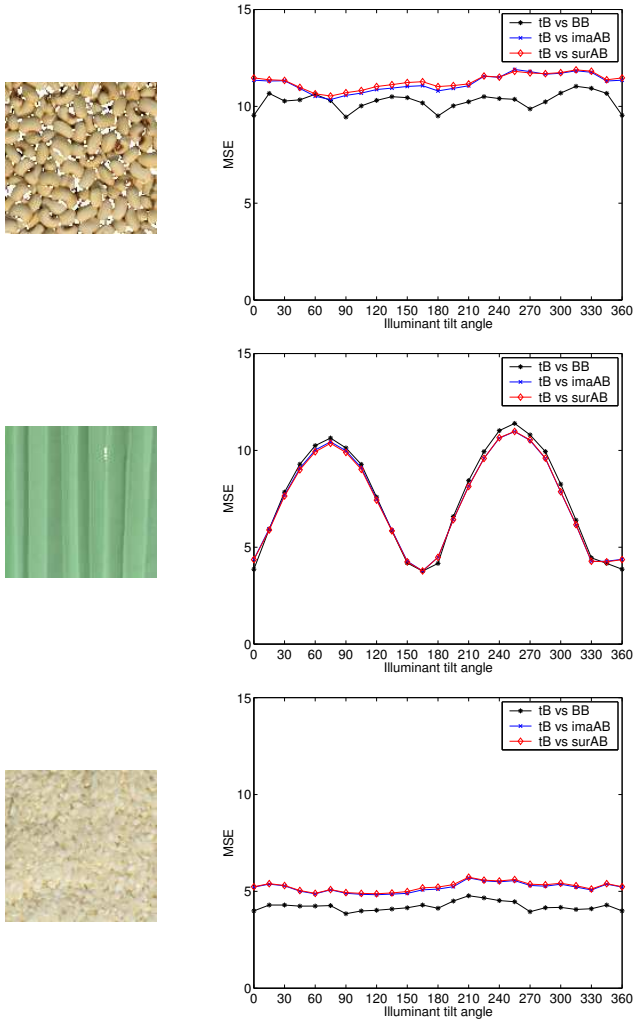


Figure 6.7: Accuracy of image prediction for three surface textures under varying tilt angles (textures T22, T14, and T10). The curve tB *vs* BB compares real images captured from distance B with those created using the information extracted by photometric stereo at distance B as well. The curves tB *vs* $imaAB$ and tB *vs* $surAB$ compare predicted images with data from distance A with real images captured from distance B.

Observing the results obtained over the twenty-five textures we conclude that in almost all cases, the performance of both prediction AB methods is very similar, producing very small differences. In Table 6.1 we give an overall quantitative assessment for each method by computing the average MSE and its standard deviation over all textures and tilt angles. The average MSE is similar for each prediction ap-

proach, although the image prediction method, which predicts the pixel intensities directly, gives, in general, smaller errors in the image. Note that most of the error can be accounted for as being produced by the photometric stereo technique and not by the step dealing with the distance change.

We also observe that surface roughness has an influence on the accuracy of the image predictions. For rougher surfaces, the error of the prediction is increased (see Table 6.1). Other surface properties, such as directionality or specularity, may contribute to the errors as well. For example, in the directional texture T14 in Figure 6.7, some orientations of the light source provoke more difficulties than others and, therefore, the error of image prediction may vary significantly depending on the light direction.

6.3.3 Experiment 3: Accuracy of Surface Shape Prediction

The goal of this experiment is to perform an evaluation of surface shape prediction, comparing the predicted gradient vectors AB with those obtained using the original photometric set for distance B.

Using the surface prediction method described in Section 4.3.2, the gradient vectors are directly predicted from the photometric information extracted for distance A. However, using the image prediction method (see Section 4.3.1), only the intensity values can be predicted but not the gradient vectors. Therefore, to make the surface shape evaluation possible, photometric stereo was applied to these predicted images (four images corresponding to four directions of the light) in order to compute the gradient vectors.

The first column of Figure 6.8 shows three examples of image predictions $imaAB$ with a particular light direction. Note that the image intensity we predict consists of image patches and not full images since there are points which photometric stereo can not recover correctly. We flag these points with white values.

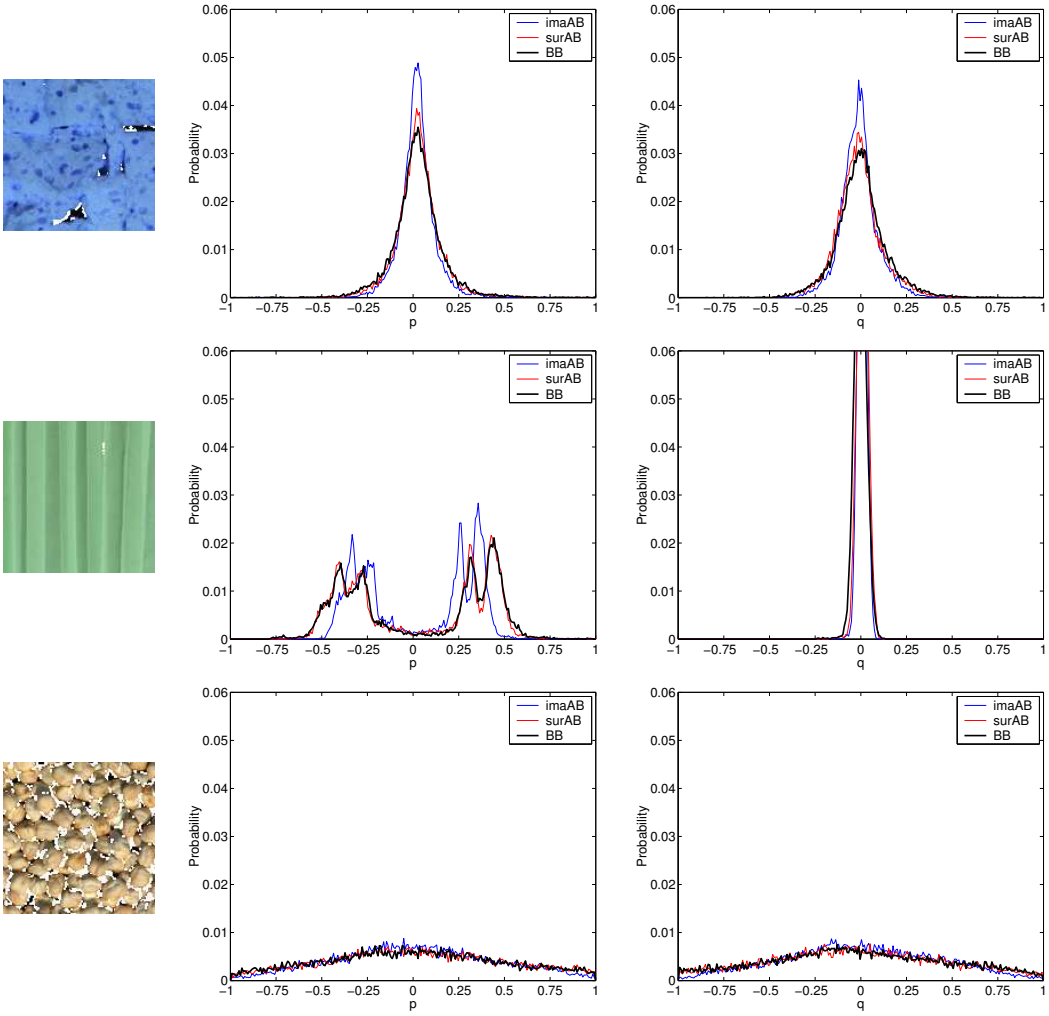


Figure 6.8: Accuracy of the surface shape predictions for three surface textures (T5, T14, and T15). Second and third column show the distributions of the surface gradients p and q obtained by applying photometric stereo directly to the original images for distance B, by applying photometric stereo to the images predicted by the method in Section 4.3.1 ($imaAB$) and by applying the surface prediction method in Section 4.3.2 ($surAB$).

Before we perform the evaluation, it is necessary to solve the problem of localising which region of the original set of distance B corresponds exactly to the region of the prediction AB. We do this by computing the correlation of surface shape (gradient components p and q) between results obtained by applying photometric stereo to the original set directly and results obtained with our prediction AB. When computing

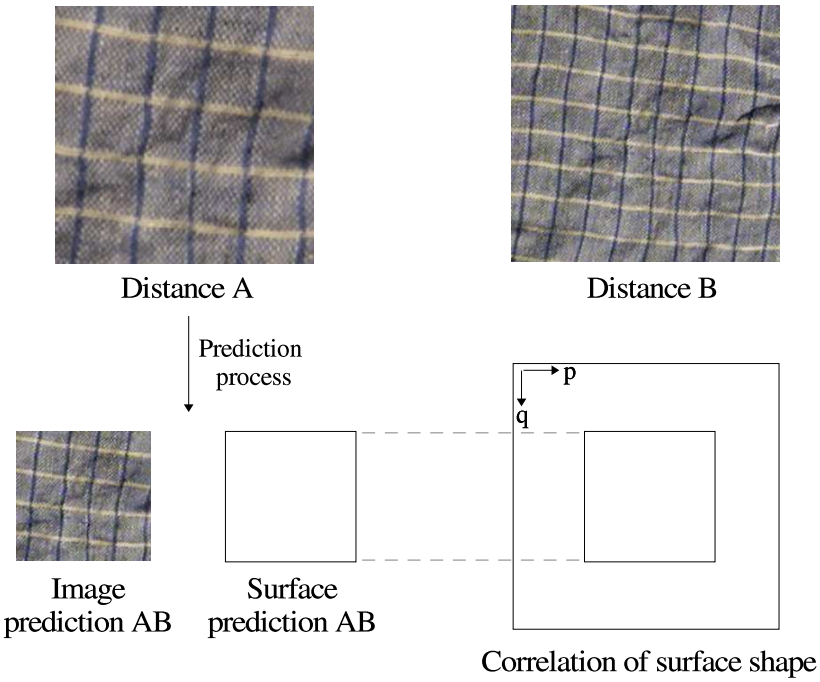


Figure 6.9: Correlation scheme. The correlation of surface shape (gradient components p and q) between results obtained by applying photometric stereo directly to the original set of distance B and results obtained with our prediction AB.

the correlation function, we exclude all points which are flagged as not reconstructed. The correlation method is applied separately for the gradient components p and q, obtaining a set of possible relative shifts between the corresponding images from p and another set of possible relative shifts between the corresponding images from q. Then, the common shifting which maximises both correlations of p and q is chosen, localising the region of the original set exactly. Figure 6.9 summarises the scheme of this correlation process. After that, the PDFs of p and q are used in order to compare the surface shape information.

Each plot in Figure 6.8 shows the surface distributions obtained by the prediction methods and those obtained by applying photometric stereo to the original images captured from distance B.

Analysing the PDFs of this figure, we observe that the results obtained by the surface prediction method fit the original PDF distributions better. On the other

Table 6.2: Quantitative assessment of shape predictions $imaAB$ and $surAB$. Two quantitative measures are used: (1) the average MSE of the PDFs and their standard deviation and (2) the average MSE per pixel of gradient components p and q and their standard deviation.

			Isotropic		Directional		Rough		Overall	
	Prediction		Avg	Std	Avg	Std	Avg	Std	Avg	Std
PDF error	$imaAB$	p	0.0008	0.0002	0.0018	0.0008	0.0017	0.0016	0.0013	0.0011
	$surAB$	p	0.0005	0.0001	0.0008	0.0003	0.0012	0.0017	0.0008	0.0011
	$imaAB$	q	0.0009	0.0002	0.0023	0.0007	0.0017	0.0018	0.0014	0.0012
	$surAB$	q	0.0005	0.0001	0.0011	0.0003	0.0012	0.0018	0.0009	0.0011
Error per pixel	$imaAB$	p	0.0708	0.0592	0.0623	0.0699	0.0849	0.0371	0.0757	0.0507
	$surAB$	p	0.0620	0.0389	0.0627	0.0507	0.0914	0.0274	0.0751	0.0377
	$imaAB$	q	0.0679	0.0476	0.0507	0.0464	0.0998	0.0330	0.0792	0.0442
	$surAB$	q	0.0605	0.0384	0.0572	0.0491	0.0896	0.0268	0.0728	0.0373

hand, surface information extracted from the predicted images ($imaAB$) introduces more error in the predicted gradient vectors. In general, the gradient values are smaller than those obtained by the surface prediction method. That can be clearly observed, for instance, in the q distribution of the second texture. In Table 6.2 we confirm this conclusion providing an overall quantitative assesment over all these histogram comparisons. We computed the average MSE of each histogram and its standard deviation over all textures. We also included the average MSE per pixel of the gradient components p and q. In both cases, better results are obtained with the surface prediction approach. The reason for these results can be found in the error introduced by the image prediction ($imaAB$), which is propagated when photometric stereo is applied to the generated images in order to recover shape information.

As mentioned earlier, the *absolute average slope ratio* ($AASR$) provides another way of characterising the degree of roughness of a given surface texture with a single parameter. We used this ratio as an alternative measure to evaluate our surface shape predictions. We compared the $AASR$ of our predictions AB with the estimated values computed using the shape information extracted by the photometric set captured from distance B. Table 6.3 gives an overall quantitative assessment over all textures computing the average MSE of the $AASR$ parameter obtained using

Table 6.3: Quantitative assessment using the $AASR$ parameter. Average MSE and its standard deviation for each prediction approach ($imaAB$ and $surAB$).

	Isotropic		Directional		Rough		Overall	
Prediction	Avg	Std	Avg	Std	Avg	Std	Avg	Std
$imaAB$	0.0273	0.0095	0.0382	0.0257	0.0369	0.0170	0.0334	0.0161
$surAB$	0.0061	0.0035	0.0129	0.0125	0.0124	0.0097	0.0095	0.0085

both prediction approaches. Note that the values obtained by the surface prediction method ($surAB$) are again better than those obtained by the photometric stereo approach applied to the predicted images ($imaAB$).

From this experiment we conclude that the surface prediction method provides the best shape estimation. Moreover, we observe that, in general, surface roughness has a strong influence on the accuracy of the surface shape predictions. For rougher surfaces the error of the predictions is increased (see Tables 6.2 and 6.3). For instance, the MSE per pixel of the gradient components p and q for texture T15, which is a surface with a high degree of roughness ($AASR$ of 0.3960), are 0.1019 and 0.1573 respectively, while for texture T5, a relatively smooth surface ($AASR$ of 0.1003), are 0.0470 and 0.0558, respectively.

6.3.4 Experiment 4: Accuracy of Image Prediction from Different Distances

The purpose of this experiment is to evaluate the accuracy of the direct image prediction method proposed in Section 4.3.1, using a photometric set captured at distance A and another at distance B to predict images captured at the longer distance C ($A < B < C$). The main goal of this experiment is to analyse the effect of using two different resolutions in the original sets to perform the same prediction when seen from distance C. Two different prediction distances (AC and BC) are used in the image prediction method.

Note that now three different sets of images taken from distances A, B, and C are used. In our image database we have fifteen textures which were captured for

(a) (b) (c)

Figure 6.10: Three images of the same surface texture seen from distances A, B, and C, (a), (b) and (c) respectively.

these three distances: textures T1, T2, T3, T4, T10, T11, T12, T13, T14, T15, T16, T17, T18, T19 and T20. Figure 6.10 shows 3 images of texture T15 seen from distances A, B and C. Over these fifteen textures, we have performed a comparasion between the test images tC and the predicted images obtained from distances A and B. To quantify the difference between a captured colour image (test image) and a predicted colour image, we use again the mean square error of colour differences computed over all pixels.

Observing the results obtained over the fifteen textures, we conclude that in almost all cases, the performance of both predictions is similar, although the predictions AC (those obtained from the information extracted at the closest distance A) give, in general, smaller errors in the image. In Table 6.4 we give an overall quantitative assessment for each prediction (AC and BC) by computing the average MSE and its standard deviation over all textures and all tilt angles which confirm this conclusion. The reason for these results can be found in the superior resolution of the images from distance A and therefore in the superior amount of information available to perform the predictions. In fact, with prediction AC, more surface patches at distance A are used to compute the intensity value corresponding to a surface path of distance C. On the other hand, there is less information available from distance B to predict the corresponding value for distance C. Note that the major error is again produced by the photometric stereo technique and not by the inclusion of the distance prediction.

Table 6.4: Overall quantitative assessment for each prediction over all 15 textures and all tilt angles. Average MSE and its standard deviation for the colour difference between the predicted and true values using the Euclidean metric in conjunction with the Luv colour space.

Overall assessment	Avg.	Std
tC vs CC	10.4412	3.5115
tC vs AC	11.4239	3.3770
tC vs BC	13.2937	3.9568

6.3.5 Experiment 5: Accuracy of Surface Shape Prediction from Different Distances

The goal of this experiment is to perform an evaluation of the surface shape prediction method described in Section 4.3.2, comparing the predicted gradient vectors AC and BC with those obtained using the original photometric set for distance C. In this experiment, we have considered the same experimental setup as for previous experiment. Therefore, we use the set of fifteen textures taken from distances A, B, and C.

Before we perform the surface evaluation, we have seen that it is necessary to solve the problem of localising which region of the original set of distance C corresponds exactly to the region of the predictions AC and BC. As explained in experiment 3 in Section 6.3.3, we do this by computing the correlation of surface shape (gradient components p and q) between results obtained by applying photometric stereo to the original set directly and results obtained with our predictions.

By analysing the PDFs for all the predictions, we observed that results obtained by prediction AC fit the original PDF distributions better. On the other hand, surface shape information extracted from distance B introduces more error in the predicted gradient vectors. In general, the gradient values are smaller than those obtained from distance A. In Table 6.5 we confirm this conclusion by providing an overall quantitative assesment over all these histogram comparisons. As in the third experiment, we computed over all the textures the average MSE of each histogram and its standard deviation. Moreover, we also included the average MSE per pixel

Table 6.5: Overall quantitative assessment for all 15 textures of shape predictions AC and BC. Two quantitative measures: (1) the average MSE of the PDFs and their standard deviation and (2) the average MSE per pixel of gradient components p and q and their standard deviation.

		p		q	
	Pre.	Avg.	std	Avg.	std
PDF error	AC	0.0008	0.0005	0.0006	0.0002
	BC	0.0010	0.0004	0.0011	0.0006
Pixel error	AC	0.0794	0.0525	0.0873	0.0449
	BC	0.1130	0.0606	0.1343	0.0638

of the gradient components p and q. In both cases, better results are achieved when the prediction is carried out from the original distance A. This is due to the fact that for prediction AC more surface patches are used to compute the desired normal vector corresponding to the surface patch of distance C.

We also used the $AASR$ ratio to evaluate the surface shape predictions. We compared the $AASR$ of our predictions AC and BC. Table 6.6 gives an overall quantitative assessment for all the textures, computing the average MSE of the $AASR$ parameter obtained using both predictions. Note that the values obtained by the surface predictions AC are again better than those obtained by the predictions BC.

As a result of this experiment we conclude that prediction AC, which has a larger distance prediction but also the most resolution among the original images, provides better shape estimation than prediction BC. We demonstrated that the accuracy of the original information used to perform the predictions has a strong influence on the results of our predictions.

6.4 Texture Classification Results

After describing the experimental trials performed on texture prediction, we will focus on the experimental trials related to the texture classification system proposed in Chapter 5. All experiments performed have been for the purpose of checking

Table 6.6: Overall quantitative assessment for all 15 textures of the $AASR$ parameter. Average MSE and its standard deviation for prediction AC and BC.

Prediction	Avg.	std
AC	0.0091	0.0078
BC	0.0263	0.0114

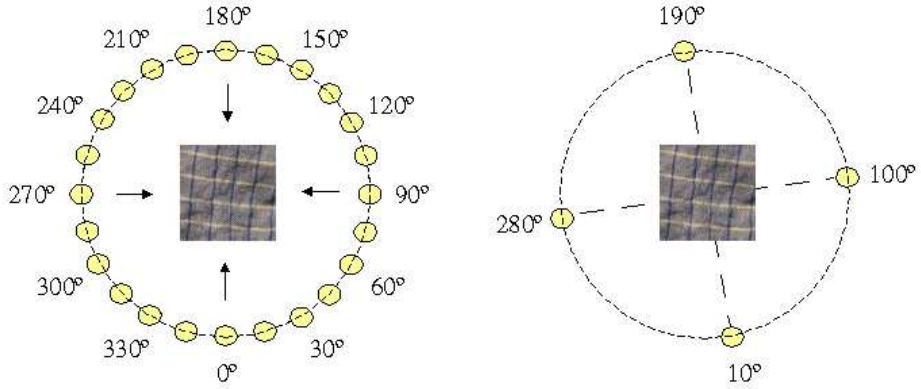
various aspects of the theory, namely:

1. To check the accuracy of the classifier when photometric sets of distance B are used to produce virtual model images for the same distance, in terms of which images captured from distance B are to be classified. This experiment provides a quantitative evaluation of our model-based texture classification when we use model images and test images from the same distance B and when the test images have been captured under different light directions from those used by the photometric stereo.

2. To check the accuracy of the light direction estimation provided by our model-based classification, giving a measure of how well the light direction in the test images can be estimated when reference and test images have been captured from the same distance.

3. To check the accuracy of the classifier when photometric sets from distance A are used to produce virtual model images for distance B, in terms of which images captured from distance B are to be classified. The results of this experiment are to be used to show how well our system can classify test images captured from the longer distance B and under different light directions.

4. To check the accuracy of the light direction estimation provided by our model-based classification, giving a measure of how well the light direction in the test images can be estimated when reference and test images have been captured from different distances.



(a) (b)

Figure 6.11: Illumination setup. (a) Illuminant tilt angles used to capture the test images. (b) Illuminant tilt angles used to create the virtual database. The unknown illuminant tilt angle of the test image has to be classified to the nearest one among these four angles used for the creation of the virtual database.

6.4.1 Experiment 1: Accuracy of Classification when Photometric and Testing Images are Captured from the Same Distance

This experiment analyses the accuracy of the texture classifier when photometric sets from distance B are used to produce model images for the same distance B. For this experiment, and all successive ones, we always use 25 texture classes in the classification process. However, in order to show which textures are more difficult to classify, we present the classification errors separately for each group of surface textures (isotropic, directional, and rough surfaces).

For each surface texture, 4 images were rendered using an elevation angle of 55° and 4 illuminant tilt angles: 10°, 100°, 190° and 280° (see Figure 6.11.b). This is the virtual database of images used as references for classification. It is composed of 100 texture images (25 surfaces × 4 illuminant tilt angles). The four illumination tilt angles used are different from those used for the test images. Therefore, when classification is performed we do not have exact correspondence between the tilt angles of the images in the virtual database and the test images.

When the virtual database is created, the recognition procedure starts. This procedure is divided into two steps: the learning process and the classification process. The goal of the learning process is to model each texture class of the virtual database by means of a representative feature vector. The goal of the classification process is to classify an unknown test image into the texture class to which it belongs.

The learning process starts by feature extraction, i.e. computing a representative feature vector for each texture image in the virtual database. Co-ocurrence matrices [61] are used to extract as features the contrast, homogeneity and energy for 20 different values of a distance d (distances between $[1 \dots 55]$ incremented in steps of 3). The pixels labeled as un-reconstructed points (points the shape and colour of which are unreliably calculated by photometric stereo) are not used in the computation of the co-ocurrence matrices. The co-ocurrence matrices are implemented in an anisotropic way. We analyse 4 different directions: 0°, 45°, 90° and 135° so that we have in all 240 texture features (3 features $\times$ 4 directions $\times$ 20 distances d). From all the computed features, those which can distinguish best between the different classes are chosen. We use the Sequential Forward Selection (SFS) algorithm [81] and a set of training images in order to select the best feature set for discrimination. The feature evaluation is performed by applying the Nearest Neighbour classifier over the set of virtual training images. This training set is composed of 3 virtual images for each surface texture and each illuminant tilt angle. We have in all 600 rendered images = 25 textures $\times$ 8 illuminations $\times$ 3 images.

When the learning is finished, the classification process starts by first extracting the feature vectors for the unknown test (real) images and classifying them by means of the Nearest Neighbour classifier into one of the classes in the virtual database. At the same time, this classification process allows us to approximate the illuminant tilt angle of the test image by identifying it with one of its nearest neighbours in the virtual database.

Due to the large size of the captured images, we produced from each one different subimages of size 133 $\times$ 133 to be used for testing. The test set consists of 9 real images for each surface texture and each illuminant tilt angle. We have

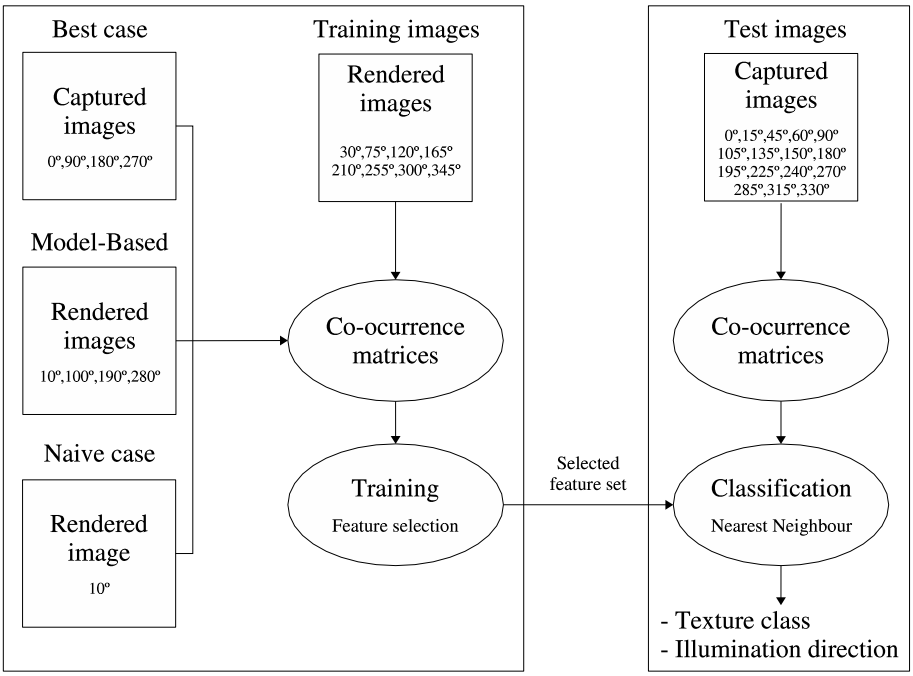


Figure 6.12: Model-based, best case, and naive case approaches to texture classification.

in all 2664 images (13 textures × 8 illuminations × 9 images + 12 textures × 16 illuminations × 9 images). Note that different tilt angles are used for training and testing; as in a real situation we do not know the true tilt angle of the test image.

In this experiment, we compare the results of our model-based approach with the results obtained by the "best" case in which the 4 reference images of each texture are real images captured with a camera and not images rendered by photometric stereo. We also compare the results with those obtained by the "naive" case in which just a single captured image of one light direction (tilt angle 0°) is used to characterise each texture class. Figure 6.12 illustrates the classification system configuration for these three different cases.

Using our model-based approach, the system obtained a 97.04% accuracy of texture classification. However, in the "best" case in which original images were used as references for the purpose of classification, the texture classification accuracy was 100%. For the "naive" case in which only one light direction was used for training, the texture classification accuracy was 87.73% (see Table 6.7). Analysing the mis-

Table 6.7: Texture and illuminant classification rates obtained for the best, model-based and naive cases when photometric sets from distance B are used to create images for distance B.

	Experiments 1 and 2				
	Best case		Model-based		Naive case
Texture	Texture	Illuminant	Texture	Illuminant	Texture
T1	100%	100%	100%	75.01%	46.67%
T2	100%	100%	100%	31.29%	100%
T3	100%	100%	100%	88.20%	100%
T4	100%	100%	100%	75.01%	73.33%
T5	100%	100%	100%	81.26%	64.44%
T6	100%	100%	100%	100%	76.29%
T7	100%	100%	100%	81.26%	100%
T8	100%	97.04%	100%	75.71%	46.67%
T9	100%	100%	88.89%	14.63%	73.33%
T10	100%	100%	100%	66.68%	97.04%
T11	100%	100%	100%	91.67%	46.67%
T12	100%	100%	100%	57.66%	100%
T13	100%	100%	100%	100%	100%
T14	100%	100%	100%	100%	100%
T15	100%	100%	100%	63.91%	100%
T16	100%	87.41%	91.85%	41.00%	100%
T17	100%	95.56%	99.56%	68.07%	100%
T18	100%	94.07%	59.70%	59.74%	100%
T19	100%	94.07%	100%	61.83%	100%
T20	100%	82.96%	85.93%	76.40%	100%
T21	100%	98.52%	100%	80.57%	93.33%
T22	100%	95.56%	100%	72.24%	100%
T23	100%	97.04%	100%	84.73%	100%
T24	100%	93.33%	100%	63.91%	75.56%
T25	100%	91.11%	100%	43.09%	100%
Overall	100%	97.06%	97.04%	70.15%	87.73%

classification error of the model-based approach (2.96%), we concluded that 84.98% of this error was produced by the group of very rough surfaces, while the remaining 15.02% was produced by the isotropic surfaces. This indicates that the missclassification errors of the model-based approach are mainly due to the image generation by colour photometric stereo information. Moreover, the results demonstrate that the model-based approach significantly reduces the texture classification errors caused by changes in light direction compared with the "naive" case.

6.4.2 Experiment 2: Accuracy of Illuminant Estimation when Photometric and Testing Images are Captured from the Same Distance

The purpose of this experiment is to evaluate the accuracy of the light direction estimation provided by our model-based classification when reference and test images have been captured from the same distance B.

As well as classifying the test images captured from distance B into the corresponding texture class, we can also classify the illuminant tilt angle under which they were captured, into one of the 4 illuminant tilt angles in the images in the virtual database. Note that by using only 4 virtual images for each surface, we can estimate the illuminant tilt angle with an accuracy of $\pm 45^\circ$ only.

With this in mind, in this experiment, the illuminant tilt angle could be estimated corectly in 70.15% of the cases by the model-based approach. However, in the "best" case, in which captured images are used as references, an accuracy of 97.06% was obtained (see Table 6.7). So, the accuracy of the illuminant tilt angle classification is considerably lower compared with the "best" case. This is presumably due to the errors introduced by the process of virtual image generation. These errors have a major influence on the illumination classification but less on texture classification because the differences between features extracted from images of the same texture under different tilt angles are smaller than the differences between features extracted from different texture classes.

6.4.3 Experiment 3: Accuracy of Classification when Photometric and Testing Images are Captured from Different Distances

This experiment analyses the accuracy of the texture classifier when photometric sets from distance A are used to produce model images for distance B.

Now, the virtual database of images corresponding to the longer distance B is generated using the direct image prediction method described of Section 4.3.1. As discussed in the experiment in Section 6.3.2, this image prediction method produces

Figure 6.13: One image predicted for each of the twenty five textures. The points flagged with white are those which cannot be correctly recovered by the photometric stereo technique.

better results than those obtained by the image prediction via the surface prediction method proposed in Section 4.3.2.

As for experiment 6.4.1, the virtual database is composed of 100 image textures (25 surfaces $\times$ 4 illuminant tilt angles). Figure 6.13 shows a predicted image for each of the 25 surface textures. In the experiment, 600 images are used for training, while 2664 images are used for testing.

After applying the feature selection algorithm and choosing the appropriate feature set, we apply the classifier to the unknown test images. Using the model-based approach, the system classified 89.57% of them into the correct texture class (see Table 6.8). We had 100% correct texture classification in the "best" case when reference and test images were captured from distance B (note that this case cor-

responds exactly to the "best" case used in experiment 6.4.1). The "naive" case here is the case when we use as the reference image one captured from a different distance from that of the test image. In this case, the classifier achieved only an accuracy of 34.35%.

Analysing the missclassification error of the model-based approach (10.43%), we concluded that 92.61% of this error was produced by the group of very rough surfaces while the isotropic surfaces contributed only 7.39% of this error.

Table 6.8 summarises the obtained texture classification rates using the model-based approach, the "best" case and the "naive" case. Comparing these results with those obtained in experiment 1 in Section 6.4.1, we conclude that the performance of the classifier is decreased due to the error introduced by the image prediction method in Section 4.3.1. However, we also demonstrate that the model-based approach increases the accuracy of the texture classification significantly compared with the "naive" case.

Note that results obtained with the "naive" case justify the creation of the virtual database of images seen at the longer distance as well as the specific learning process for these images. In [103] we presented similar experimental results related to the naive texture classification. We tried to classify test images captured from distance B using the features extracted from the virtual database of images from distance A. In this "naive" case, we confirmed the conclusion stated above since we classified correctly only 21.83% of the test images.

6.4.4 Experiment 4: Accuracy of Illuminant Estimation when Photometric and Testing Images are Captured from Different Distances

The goal of this experiment is to evaluate the accuracy of the light direction estimation provided by the model-based classification when reference and test images are from different distances.

Table 6.8: Texture and illuminant classification rates obtained for the best, model-based and naive cases when the direct image prediction method is used to generate images for distance B from the photometric sets from distance A.

	Experiments 3 and 4				
	Best case		Model-based		Naive case
Texture	Texture	Illuminant	Texture	Illuminant	Texture
T1	100%	100%	100%	77.87%	0%
T2	100%	100%	100%	77.87%	35.55%
T3	100%	100%	100%	73.56%	0%
T4	100%	100%	93.33%	28.07%	0%
T5	100%	100%	100%	83.40%	53.33%
T6	100%	100%	100%	100%	0%
T7	100%	100%	87.41%	100%	85.76%
T8	100%	97.04%	100%	61.27%	100%
T9	100%	100%	100%	17.00%	46.67%
T10	100%	100%	100%	55.73%	100%
T11	100%	100%	100%	39.13%	0%
T12	100%	100%	100%	77.87%	23.61%
T13	100%	100%	100%	92.62%	40.40%
T14	100%	100%	100%	77.87%	5.92%
T15	100%	100%	100%	100%	26.67%
T16	100%	87.41%	95.56%	60.65%	65.93%
T17	100%	95.56%	87.41%	81.56%	44.44%
T18	100%	94.07%	48.15%	82.17%	73.11%
T19	100%	94.07%	100%	58.81%	85.67%
T20	100%	82.96%	9.63%	23.76%	0%
T21	100%	98.52%	23.70%	35.44%	0%
T22	100%	95.56%	100%	77.87%	71.84%
T23	100%	97.04%	94.07%	78.48%	0%
T24	100%	93.33%	100%	75.41%	0%
T25	100%	91.11%	100%	33.60%	0%
Overall	100%	97.06%	89.57%	66.80%	34.35%

We used the same experimental setup as that used in experiment 3 in Section 6.4.3. Using the model-based approach, the system estimated the illuminant tilt angle with an accuracy of 66.80% while the light classification accuracy of the "best" case[1] was 97.06% (see Table 6.8). Comparing the results of this experiment with those obtained when the reference and testing images were captured from the same distance, we conclude that the accuracy of the tilt angle classification is de-

[1]Note that for the "best" case, the texture and light classification rates are the same as those obtained in experiment 1 of Section 6.4.1. In both experiments, we used the same reference and test images from distance B.

creased due to the error introduced by the step of the algorithm when dealing with the change in distance.

6.5 Summary and Conclusions

The prediction methods presented in Section 4.3 and the model-based texture classification system proposed in Section 5.3 were tested and evaluated in this chapter.

After analysing different texture databases we observed the lack of a database to accomplish our principal requirements: textures seen under various directions of light and various distances from the camera. Due to this fact, we opted to build our **own image database** to perform the experimental trials presented in this thesis. The image database contains real surface textures which can be classified into relatively smooth surfaces (which may be further divided into isotropic surfaces and directional surfaces) and very rough surfaces. For each surface texture, as well as the 4 required images to apply the photometric stereo technique, the images which should be used as training and test images are available in our image database.

The experimental trials presented in this chapter are divided into two blocks: one related to texture prediction, and the other to texture classification. In the first block we performed different experimental trials in order to evaluate the two prediction methods in Section 4.3:

- One which allows us to predict image intensities directly (direct image prediction).
- Another which allows us to predict the surface shape information first and then the image intensities (image prediction via surface prediction).

Both methods were tested and evaluated over a set of twenty-five surface textures, demonstrating the ability to predict the image texture a particular surface texture will create when seen from a distance longer than the original. The direct image prediction method produces, in general, smaller errors. We also concluded that most errors can be accounted for as being produced by the photometric stereo technique

and not by the step dealing with distance change. On the other hand, the surface prediction method produces the best shape predictions. Several error measures were used in order to evaluate the surface shape predictions: the absolute average slope ratio ($AASR$) which measures the degree of roughness of a surface, the MSE of the estimated probability density functions for the surface partial derivatives p and q, and the MSE per pixel of the gradient components p and q.

As a result of these experiments we concluded that surface roughness plays an important role in the accuracy of image and shape prediction. The rougher the surface, the larger the error of the prediction. Other surface properties, such as directionality or specularity, may also contribute to errors.

Another set of experimental trials related to the texture classification system were presented in the second block. We observed that even if the predicted images are not perfect, they may still make a significant difference in the accuracy of a texture classifier. We compared the results of our approach with those obtained by the "best" case where the reference and test images of each texture were real images captured under exactly the same imaging geometries and by the "naive" case where only a single real image was used as reference. We saw how the results of our texture classification approach were significantly better than those obtained by the "naive" case, while the results were not much inferior to those obtained by the "best" case. An improvement was also obtained in cases for which the assumptions made by the photometric stereo technique were violated and the photometric stereo results were not as accurate as we might have wished. In all cases, an estimate of the unknown light orientation under which the test image was captured was also obtained, although not with very high accuracy.

Although we presented results of our texture classification system when seen from different distances and under different illuminant tilt angles, it is interesting to mention that our approach could be used in other situations. For instance, changes in the illuminant slant angle and in the camera direction. If we have information on the surface relief and the surface albedo and we use the general equations derived in Section 4.3, then we could render the surface to create model textures as required.

Chapter 7

Conclusions

Conclusions extracted from this thesis work are presented. Moreover, possible further work is analysed considering the different directions in which this research line could go. Finally, publications related to this thesis are listed.

7.1 Contributions

In this thesis we have presented a model-based texture classification system able to classify textures when seen from different imaging geometries such as the distance from the camera and the direction of light. The main contributions of this thesis work are summarised in the following points:

- **Texture prediction framework.** We presented a general framework for describing textures when seen from different imaging geometries. The 4-source colour photometric stereo was used in order to obtain the reflectance and shape information of the surface from a close distance. From this surface information, the proposed prediction framework allows us to predict how the surface texture will look even when seen under different imaging conditions: different sensor, distance from the sensor (the camera positioned at distances at least as long as the distance used for capturing the training images), different spectral characteristics of the light source, and different directions of light. This is based on the assumption of Lambertian surfaces, but it can easily be

generalised to other types of surface.

Two different prediction methods have been proposed, one which allows us to predict image intensities directly (direct image prediction), and another which allows us to predict the surface shape information first and then the image intensities (image prediction via surface prediction). We observed that the direct image prediction method produces, in general, smaller errors in the images while the surface prediction method produces the best shape predictions. We extracted unsurprising conclusions: (i) a large amount of the error in the predictions can be accounted for as being produced by the photometric stereo technique, and (ii) surface roughness plays an important role in the accuracy of image and shape prediction. The rougher the surface, the larger the error of the prediction.

- **Model-based texture classification system.** A model-based classification system able to classify textures when seen from different distances and under different directions of lights was presented. The main motivation of this methodology has been to overcome the problem of classifier failure induced by varying these imaging properties. Moreover, we have exploited the capability of the classification system to guess an approximate direction of the light for the unkown test images. The whole procedure of the proposed texture classification system was divided into two main phases. In the first phase, *"virtual" database creation*, a virtual database of textured images under novel imaging geometries was created by using the surface texture information and the proposed texture prediction framework. The second phase, the *recognition procedure*, was divided into two steps: the learning process in which each texture class of the virtual database is modelled by means of a representative texture feature vector and the classification process in which an unknown test image is classified into the texture class to which it belongs. In the learning process, the images of the virtual database were used to extract texture features derived from the co-occurrence matrices. Moreover, a feature selection process was used to chose the best feature models for each texture class. On the other hand, classification was achieved by determining the nearest neighbour

in the feature space populated with feature vectors from the virtual database computed during training.

We saw, in the experimental trials, how the texture classification results of our approach are significantly better than those obtained with a typical texture classification system in which the variability produced by changes in light direction and camera distance are not considered.

- **A novel texture database.** Different databases were considered, for example, the Brodatz album [18], the "Photometric Image Databases" from the Texture lab at Heriot-Watt University [2], and the "Columbia-Utrecht database" established by Dana et al. [35]. None of them could fulfill all the requirements needed for our purposes. Therefore, we opted to build our own image database, providing for each surface texture different sets of images captured from different distances of the camera and under different directions of light. This image database is of public domain and we think it can be useful for the research community to check and test techniques related to scale invariant and light invariant texture analysis and classification. It can be downloaded from *http://eia.udg.es/∼llado/database.html*.

7.2 Future Work

The design of a texture classification system invariant to the imaging geometries involves the consideration of a wide number of questions. In addition to the various solutions which have been adopted and described in this thesis, different ideas and different approaches have been dicussed throughout this thesis work. Moreover, a large number of ideas have remained as undeveloped tasks which need to be further analysed and worked on in depth and we therefore suggest them as future work.

Basically, we have divided this future work into two blocks. The first is composed of ideas which can directly improve our texture classification approach. The second suggests some ideas for further work to be considered as future research lines.

Further Work on the Proposed Classification System

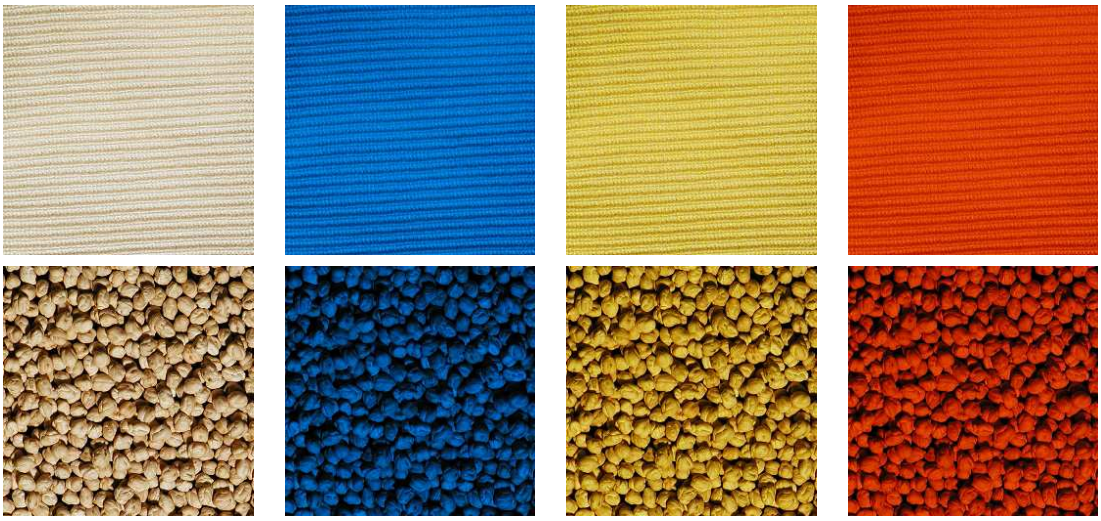
- **Different sensors and light sources.** We propose as further work to analyse the effect of parameters such as the sensitivity of the sensor and the spectral properties of the light source which have been assumed to be the same for all the experimental trials presented in this thesis. Figure 7.1 illustrates the effect of varying the spectral properties of the light source on two textures. Three different colour filters (blue, yellow and red) have been used in order to obtain these images.

- **Changes in slant angles.** Obviously, large slant angles will intensify the level of shadowing, providing more information for the photometric stereo technique but also deterioring the precision of the surface estimation. On the other hand, low slant angles (close to a perpendicular to the surface plane) will provide less information since the perceived image textures would be more similar. It could be very interesting to analyse what happens in all of these cases.

- **Tolerance in the estimate of the guessed distance of the camera.** What if we do not known exactly the longer distance from which the test image was captured? How accurately do we have to guess it for the performance of the classifier not to deteriorate significantly? These are important aspects which could be studied.

- **Different texture features.** As it has been noted, there are different texture analysis techniques which allow the extraction of texture information from an image. We propose a future project to consider different types of features and compare them with those extracted from co-occurrence matrices.

- **Extend our texture database.** We also suggest, as further work, to extend our texture database. We propose to include more surface textures captured under different imaging geometries i.e. camera distance, light directions (including different slant angles), different spectral properties of the light source and using cameras with different sensitivity.



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

Figure 7.1: Images of two textures captured under varying the spectral properties of the light source. (a) Images captured with the original light source, (b)-(d) images captured using three different colour filters: blue, yellow, and red, respectively.

Further Work along this Research Line

- **Analyse the texture features behaviour.** In Appendix A we presented a first step in order to deal with the illumination invariant classification analysing the behaviour of the features extracted from the co-occurrence matrices. This has been used to perform a simultaneous surface texture classification and illumination tilt angle prediction. We think that this is an interesting idea to continue this research line. This can provide important knowledge and open new possibilities.

- **Viewpoint independent texture classification.** Recently, Lazebnik et al. [91, 92] introduced a texture representation suitable for recognising images of textured surfaces under a wide range of transformations, including viewpoint changes and non-rigid deformations. Obviously, camera direction is an important imaging parameter which has not been considered in this thesis. We think this subject should be studied in depth and exploited, starting from the work presented here.

- **Applications.** Some of the ideas proposed in this thesis can be used in different Computer Vision applications. We would like to mention some possible applications closely related to the research developed within the Computer Vision and Robotics group at the University of Girona. Considering two of the basic research lines of this group, 3D perception and Underwater robotics, it could be interesting to study the use of photometric stereo in order to achieve 3D reconstruction. Moreover, and more ambitiously, we could attempt to work it into an underwater reconstruction basically focusing on the idea of recovering distances and depth maps of the underwater floor.

7.3 Related Publications

Ideas and projects of this work have evolved from initial stages to obtain the final version presented in this thesis. The following publications are a direct consequence of the research carried out during the elaboration of this thesis and give an idea of the progress achieved.

Journals

- X. Lladó, M. Petrou, and J. Martí. Surface texture recognition by surface rendering. *IEEE Transactions on Image Processing* **IEEE T-IP**, in second revision 2003.

International Conferences

- X. Lladó, A. Oliver, M. Petrou, J. Freixenet, and J. Martí. Simultaneous surface texture classification and illumination tilt angle prediction. In *British Machine Vision Conference* **BMVC 2003**, volume 2, pages 789-798, Norwich, England, September 2003.

- X. Lladó, J. Martí, and M. Petrou. Classification of textures seen from different distances and under varying illumination direction. In *IEEE International Conference on Image Processing* **ICIP 2003**, Barcelona, Spain, September 2003.

- X. Lladó, and M. Petrou. Classifying textures when seen from different distances. In *IAPR International Conference on Pattern Recognition* **ICPR 2002**. Quebec, Canada, August 2002.

- X. Lladó, and M. Petrou. Predicting surface texture when seen from different distances. In *International Workshop on Texture Analysis and Synthesis* **ECCV-Texture 2002**, pages 83-86, Copenhagen, Denmark, June 2002.

- X. Lladó, J. Martí, and M. Petrou. Image texture prediction using colour photometric stereo. In *Lecture Notes in Artificial Intelligence* **LNAI 2002**, pages 355-363. Springer-Verlag, 2002.

- X. Lladó, J. Pagès, J. Martí, J. Batlle, and C. Dragoste. Quality control by using surface shape analysis. In *IEEE-TTTC International Conference on Quality Control, Automation and Robotics* **QCAR 2002**, volume 1, pages 444-449, Cluj-Napoca, Romania, May 2002.

- J. Freixenet, X. Lladó, J. Martí and X. Cufí. Use of decision trees in colour feature selection. Application to object recognition in outdoor scenes". In *IEEE International Conference on Image Processing* **ICIP 2000**, volume 3, pages 496-499, Vancouver, Canada, September 2000.

- X. Lladó, J. Martí, J. Freixenet, and Ll. Pacheco. A novel criterion function in feature evaluation. Application to the classification of corks. In *International Conference on Quality Control, Automation and Robotics* **QCAR 2000**, volume 2, pages 275-280. Cluj-Napoca, Romania, May 2000.

National Conferences

- X. Lladó, J. Martí, J. Freixenet, and J. Batlle. Automated industrial inspection using photometric stereo. In *Catalonian Conference on Artificial Intelligence* **CCIA 2001**, pages 269-275, Barcelona, Spain, October 2001.

- J. Martí, J. Freixenet, X. Lladó, X. Muñoz, and X. Cufí. A new approach to natural object labelling in colour aerial images. In *Spanish Symposium on Pattern Recognition and Image Analysis* **SNRFAI 2001**. volume 1, pages 67-72, Benicàssim, Spain, May 2001.

- X. Lladó, J. Martí. J. Freixenet, and V. Ila. Object characterization in outdoor scene analysis. In *Spanish Symposium on Pattern Recognition and Image Analysis* **SNRFAI 2001**, volume 2, pages 113-118, Benicàssim, Spain, May 2001.

- X. Lladó, J. Martí, J. Freixenet, and X. Muñoz. Feature selection using genetic algorithms. In *Catalonian Conference on Artificial Intelligence* **CCIA 2000**, pages 152-156, Vilanova i la Geltrú, Spain, October 2000.

- X. Lladó, J. Freixenet, J. Martí, J. Forest, and J. Salvi. Object recognition in outdoor scene using a bottom-up strategy. In *Catalonian Conference on Artificial Intelligence* **CCIA 1999**, pages 299-307, Girona, Spain, October 1999.

Technical Reports

- X. Lladó, J. Martí, J. Freixenet, and J. Batlle. A review of feature selection and feature evaluation. Technical Report 02-19-RR, Institute of Informatics and Applications. University of Girona, 2002.

- X. Lladó, and J. Martí. Surface texture prediction using colour photometric stereo. Technical Report 01-14-RR, Institute of Informatics and Applications. University of Girona, 2001.

Appendix A

Simultaneous Surface Texture Classification and Illumination Tilt Angle Prediction

In this Appendix we investigate the effect of the illuminant tilt rotation over surface textures by analysing a set of image texture features extracted from the co-occurrence matrix. From the behaviour of each texture feature, a simple method, able to predict the illuminant tilt angle of test images, is developed. Moreover, the method is also used to perform a texture classification invariant to the illuminant tilt angle rotation. This study, and experimental results over different real textures, show that the illumination tilt angle can be accurately predicted as part of the texture classification process.

A.1 Introduction

Very little work has been published on the topic of illumination invariant texture classification. One strategy to solve this problem is to study the immediate effects introduced by light direction on the observed 2D texture. This was done recently by Chantler et al. [23] who presented a formal theory which demonstrates that changes in the tilt of the light direction make texture features follow super-elliptical trajectories in multi-dimensional feature spaces. Based on this work, Penirschke et al. [125] developed an illuminant rotation invariant classification scheme which

uses photometric stereo for the detection of surface relief and Gabor features for feature extraction. In other applications of Computer Vision, a correct estimation of light can play an important role. For instance, light estimation is apparent in all applications which use photometric measurements which obviously depend on light. Weber and Cipolla [167] focus their attention on reconstruction problems and the estimation of light-source.

In this Appendix, we investigate the effect of the illuminant tilt rotation over surface textures by analysing a set of texture features extracted from the co-occurrence matrix. From the behaviour of each texture feature, we develop a simple method able to predict the illuminant tilt angle of unknown test images. Moreover, we use this prediction method to perform the classification of textures under varying illuminant tilt angles.

The remainder of this Appendix is organised as follows. In section A.2 we analyse the behaviour of texture features in the feature space. The method to predict the illuminant tilt angle is explained in section A.3. In section A.4 the performace of the method is evaluated in two experimental trials: illuminant tilt angle prediction and texture classification. Finally, the Appendix ends with conclusions and further work.

A.2 Co-occurrence Matrix Behaviour

The co-occurrence matrix [61] is a well known method used to extract texture features. In this work, the co-ocurrence matrices are implemented in an anisotropic way. That is, we analyse 4 different directions: 0°, 45°, 90°, and 135° computing for each the contrast feature for a fixed distance 1.

A set of 15 different textures have been used throughout the experimental trials presented in this work. Figure A.1 shows one image of each texture captured under a specific direction of light. Note that almost all are isotropic textures (see the first 10 textures). Therefore, the only anisotropy in the image is induced by the anisotropic light. It is this anisotropy we wish to detect in order to identify the light direction.

Figure A.1: One image of each of the fifteen sample textures.

We have included in the database some anisotropic textures for comparison. In order to detect anisotropy in an image we use the rotationally sensitive co-occurrence matrix from which we construct features.

Figures A.2.(a), A.2.(c) and A.2.(e) illustrate the behaviour of these 4 features for textures 1, 9, and 15 respectively. Each plot shows how one output of one contrast feature varies when it is applied to the same texture sample, but under varying illuminant tilt angles (we use steps of 30° degrees from 0° to 360°).

Note that the graph of the anisotropic texture 15, which consists of long rough structures (see figure A.1), shows some strange behaviour at first glance: three of the features behave in a very similar way (i.e. note the three similar curves in figure A.2.(e)) while one feature is totally constant. The catastrophic behaviour of this feature is due to the fact that the direction used to compute the corresponding co-occurence matrix coincides with the direction of the long structures which make up the imaged surface. In this particular case, the images captured under all the tilt angles have no significant intensity changes with respect to the axis of the local structures. Therefore, we obtain constant feature values for all tilt angles. On the other hand, when the long rough structures are lit from lateral directions, we expect some invariant behaviour as one side of these structures remains lit through a broad range of incident angles. This is the reason three of the features behave in a very similar way.

Moreover, observing the feature behaviour for isotropic textures, we can see that the contrast feature at distance 1 has a symmetric behaviour approximately between ranges $[0°, 180°]$, and $[180°, 360°]$. Therefore, it is difficult to distinguish whether the illuminant tilt angle is in one or the other range. That is, by measuring one of these features, which is tilt angle sensitive, we may be able to identify the illuminant tilt angle only up to an ambiguity of $\pm 180°$.

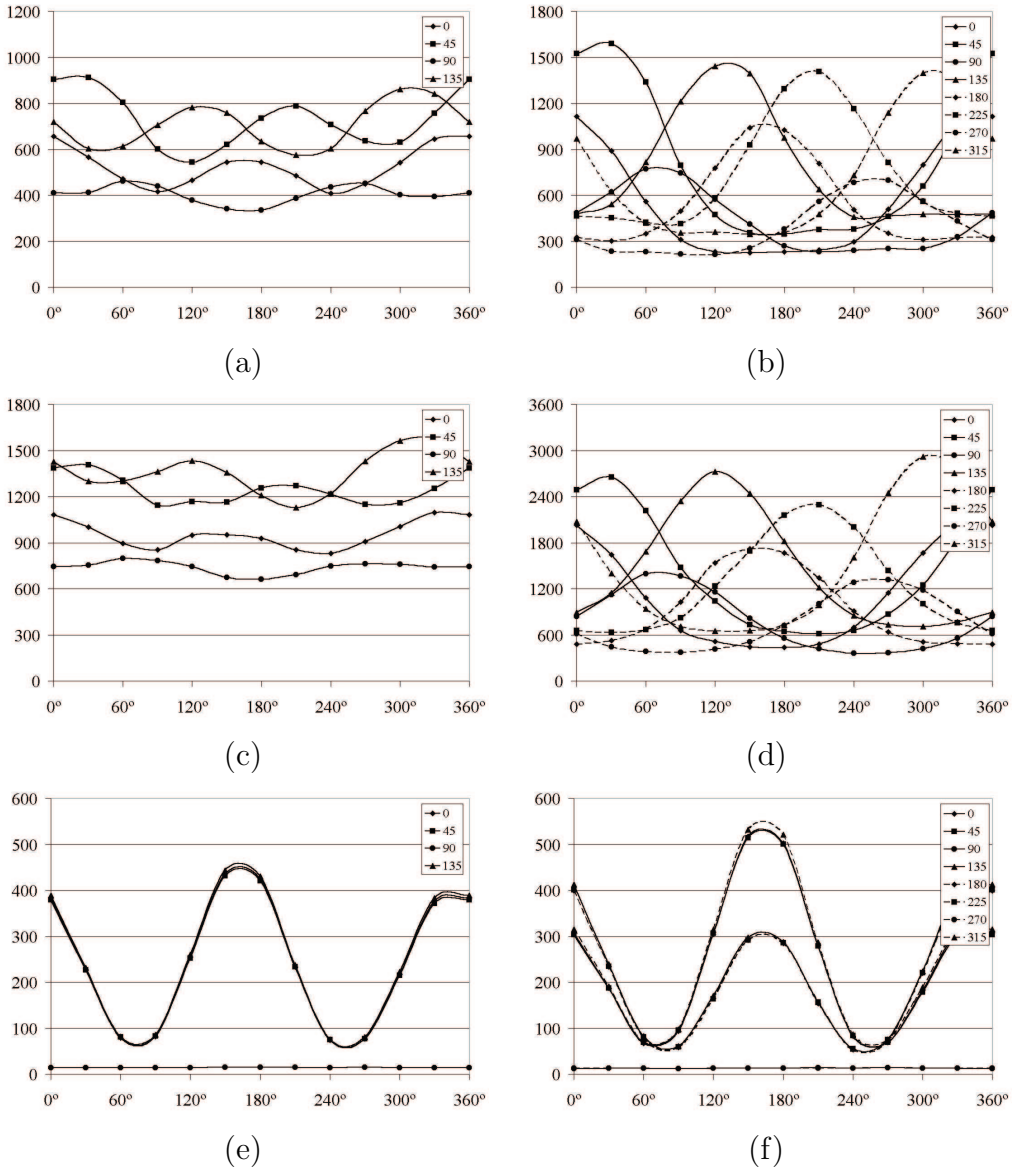


Figure A.2: Feature behaviour for isotropic textures 1 (first row), and 9 (second row), and anisotropic texture 15 (third row). Each plot shows how one output of one feature varies when it is applied to the same physical texture sample, but under varying illuminant tilt angles. (a), (c), and (e) Co-ocurrence matrices computed for 4 different directions. (b), (d), and (f) Co-ocurrence matrices computed for 8 different directions.

A.2.1 Removing the Ambiguity from the Estimation of the Tilt Angle

With the aim of improving the illuminant tilt angle prediction, allowing the estimation over the whole tilt rotation of 360°, we propose to introduce the use of 4 new directions in the computation of the co-occurrence matrix. Therefore, we compute the co-occurrence matrix using 8 different directions: 0°, 45°, 90°, 135°, 180°, 225°, 270° and 315° . Note that in this approach the co-occurrence matrix obtained for 180°, 225°, 270°, and 315° are the transposed matrices of 0°, 45°, 90°, and 135° respectively. Hence, computing the classical contrast feature for these 8 matrices we only obtain 4 different values since the contrast feature gives us the same value for a matrix and its transposed.

As our objective is to distinguish the sense of the directions used in the co-occurence matrix, we propose to compute the contrast feature from the upper triangular matrix only. We do that with the idea of counting the pairs of pixels in which the intensity value increases (transitions of darker pixels to brighter pixels).

This can be seen in figures A.2.(b) and A.2.(d) where we plot the variation of these 8 new features for the isotropic textures 1 and 9. Note that for each tilt angle, the maximum value of the contrast feature is attained by the feature which has been computed from the co-occurrence matrix with the same orientation angle as the tilt angle. This is because we always find more transitions of darker pixels to brighter pixels when the orientation of the light source "coincides" with the orientation of the co-occurence matrix. It is important to clarify that we use the term "coincide" when we refer to the same angle, but under two reference systems: one which defines the direction used in the co-occurence matrix, and one which defines the direction of the incident light.

In contrast, the anisotropic textures do not follow this behavior (see figure A.2.(f)), although they still exhibit an approximately symmetric behaviour over the two ranges of values: namely $[0°, 180°]$, and $[180°, 360°]$.

The feature behaviour described in this section will be used as a model to predict the tilt angle of an unknown texture image. Each texture model is created using a

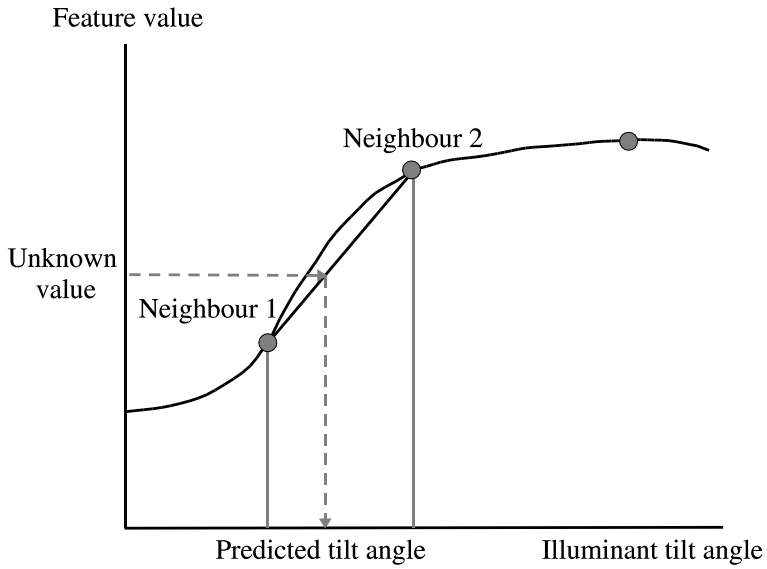


Figure A.3: Prediction scheme. In a first step, we find the two feature vectors (neighbour 1 and 2) closest to the analysed feature vector. Then, the tilt angle is obtained in the current interval applying a linear interpolation individually for each single feature.

small number of different tilt angles. Specifically, we have used 12 tilt angles 30° apart. Hence, a texture model is composed of 12 vectors of 8 features each (12 tilt angles, 8 directional co-occurrence matrices for each).

A.3 Illuminant Tilt Angle Prediction

The process of predicting the illuminant tilt angle, given an unkown test image, starts by extracting a representative feature vector for this image which implies the computation of 8 texture features. Note that the model of each texture is composed of 12 different feature vectors, one for each reference tilt angle.

When the feature vector for the unknown image is obtained, the prediction consists of looking for the most similar feature vector among the model textures. Then, a simple method of three steps is proposed:

- First, we obtain a first approximation of the predicted angle with one of the known angles used to describe the model texture. The nearest neighbour

classifier is used to find the closest feature vector.

- After that, we localise the angle interval which contains the test feature vector. We use again the nearest neighbour classifier to find the second most similar feature vector of the same model texture (see figure A.3). In this second step we ensure that the angle interval is of 30°. This means that the second known angle is ±30° with respect to the first approximation.

- Next, the exact tilt angle is found in the current interval applying linear interpolation for each single feature. The results are averaged to produce the final predicted tilt angle. The tilt angles computed from individual features for which the linear interpolation provides slopes close to zero are not considered in the averaging process.

A.4 Experimental Trials

The proposed prediction method was tested with 15 different textures. The first 10 were isotropic textures, the remaining five were anisotropic. For each one, 4 complete sets of 24 images corresponding to the illuminant tilt angles between 0° and 360° incremented in steps of 15° are available. All these images are lit at an elevation angle of 55°. From these images we create 2 different sets: one used for modeling the light behaviour, and another for testing the classification process.

The first set is composed of 15 textures×12 light directions = 180 images. The 12 illuminant tilt angles used for modeling are 0°, 30°, 60°, 90°, 120°, 150°, 180°, 210°, 240°, 270°, 300°, and 330°. On the other hand, the testing set is composed of 15 textures×12 light directions×3 images = 540 test images. The 12 illuminant tilt angles used for testing are different from the reference tilt angles mentioned above. Figure A.4 illustrates the configuration used in our experimental trials.

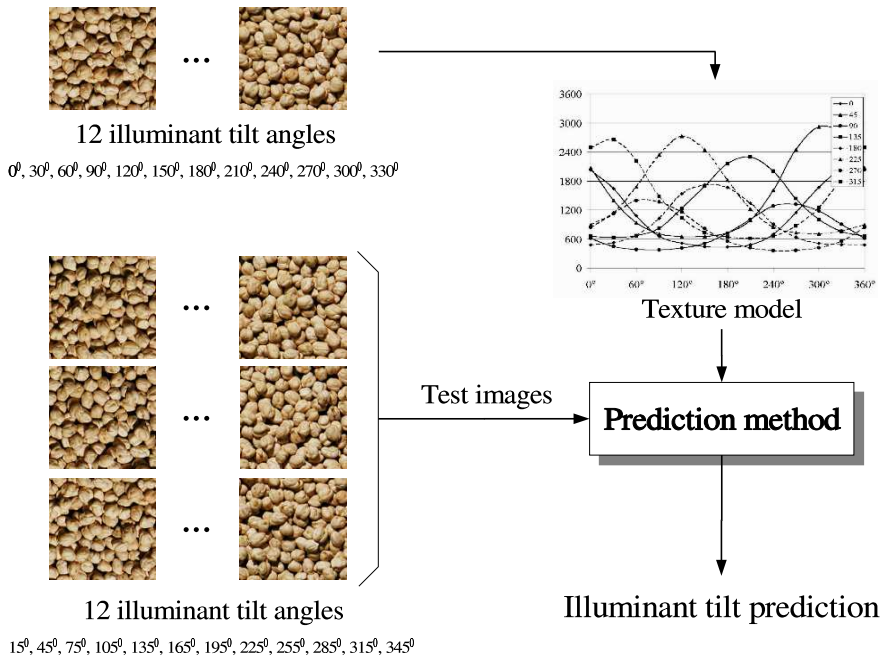


Figure A.4: Experimental setup. Four different surface patches are available for each texture. One is used for modelling the feature behaviour, the remaining ones for testing.

A.4.1 Accuracy of Tilt Angle Prediction

The purpose of this experiment is to evaluate the accuracy of the illuminant tilt angle prediction. After computing the model behaviour, we individually apply for each texture, the prediction method for all the corresponding test images.

Figure A.5 shows the error distributions for textures 1, 9, and 15 of the absolute tilt angle difference between our predictions and the correct values. Each plot has been computed for the 36 test images available for each texture and shows the percentage of times a particular error value was observed. Note that for isotropic textures 1 and 9, we predict the tilt angle for almost all the test images with an error of a few degrees. However, for texture 15, the errors are significantly larger.

From the error distributions of all 15 textures we conclude that for isotropic textures the illuminant tilt angle may be predicted quite accurately. However, poor results are obtained for anisotropic textures. We confirm this conclusion providing

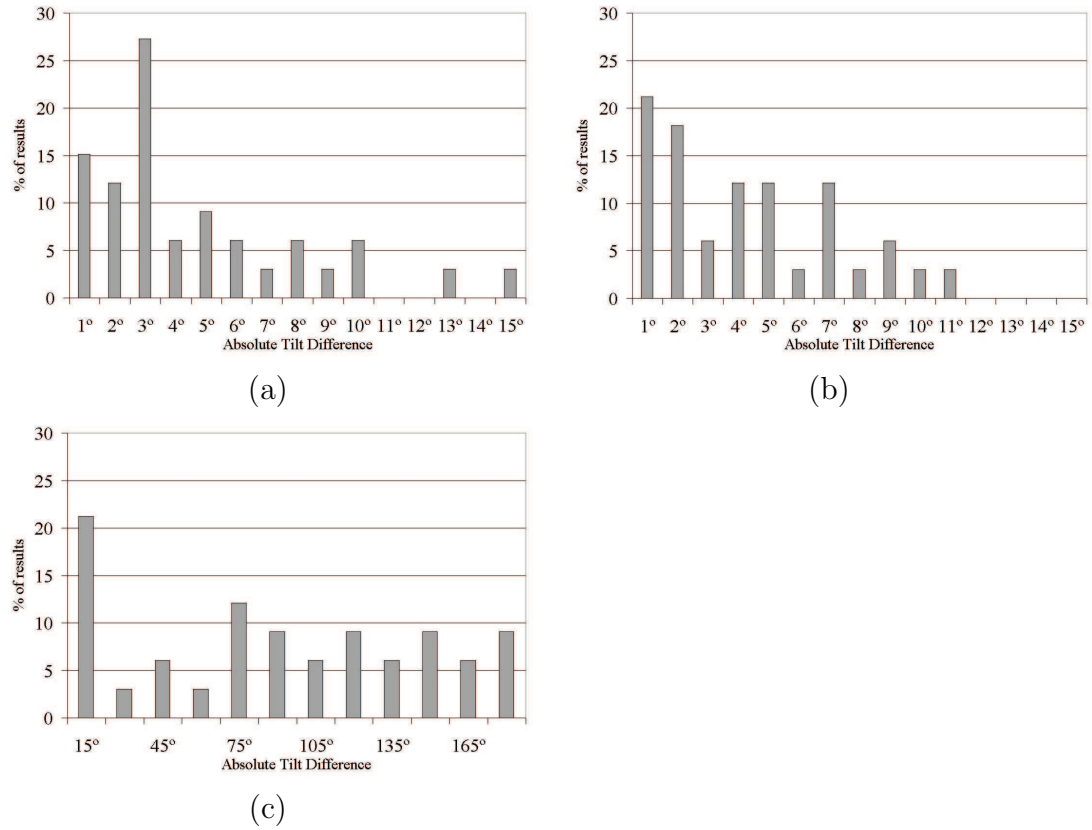


Figure A.5: Error in the tilt angle prediction for textures 1, 9, and 15, (a), (b), and (c) respectively.

an overall quantitative assessment over all these predictions (see table A.1). We have computed over all the textures the average MSE of the tilt angle prediction and its standard deviation. In the same table we present a quantitative assessment for the isotropic and anisotropic textures separately. It is important to note that we predict the tilt angle for isotropic textures with an average error of 6°. Nevertheless, the prediction error increases when anisotropic textures are considered. This is because the anisotropy of the image cannot be solely attributed to the light direction and, therefore, the detected anisotropy does not give us the clear and unambiguous information which is needed by our prediction method.

Table A.1: Overall quantitative assessment over all 15 textures of illuminant tilt angle predictions. Average MSE and its standard deviation for the tilt angle difference between the predicted and true values.

Isotropic		Anisotropic		Overall	
Avg	Std	Avg	Std	Avg	Std
5.96	1.85	59.61	22.57	23.84	28.86

A.4.2 Accuracy of Texture Classification

This experiment analyses the accuracy of the texture classification when our feature behaviour models are used as references for classification.

The method described in section A.3 may be used not only to predict the illuminant tilt angle of a test image, but also to classify the unkown test image into one of the texture classes present in the database. Specifically, the first step in which the nearest neighbour classifier is used to find the closest feature vector to the model, allows us to perform texture classification as well.

Table A.2 summarises the obtained texture classification results when all fifteen models are used in the classification process. The texture classification accuracy is 82.63%, while the illuminant tilt angle is predicted with an average MSE of 24.04° and standard deviation of 43.07°. We have repeated the same experiment, but using only the isotropic textures, achieving a texture classification accuracy of 79.09%. In this case, the illuminant tilt angle is predicted with an average MSE of 5.88° and standard deviation of 4.80°. Note that when using all fifteen textures we obtain better classification results than those using only isotropic textures. That is because isotropic textures have similar feature behaviour. For instance, test images of textures 2 and 4 have been misclassified. In contrast, as is shown in figure A.2.(f), anisotropic textures have different feature behaviour. This fact causes an improvement in the classification rate, when anisotropic textures are included in the reported results. However, it is important to notice that accurate tilt angle predictions are only obtained for isotropic textures.

Table A.2: Texture classification rates and MSE of the tilt angle prediction obtained over the isotropic textures and over all fifteen textures.

	Texture classification	Tilt angle MSE	
		Avg	Std
Isotropic textures	79.09%	5.85	4.80
All textures	82.63%	24.06	43.07

A.5 Conclusions

We presented a simple method able to predict the illuminant tilt angle of unknown test images. This method is based on behaviour models of texture features extracted from the co-occurrence matrix. It works under the assumption that the only anisotropy in the image is induced by the anisotropic light. The experimental results over different real textures, including some anisotropic textures for comparison, show that the illumination tilt angle can be accurately predicted. As well as predicting the illuminant tilt angle, this method is used to perform texture classification. The results show that anisotropic textures may be classified more accurately but their illuminat tilt angle may not be predicted so well, while isotropic textures cause more confusion to the classifier but allow us to predict the direction from which the imaged surfaces were lit very accurately, as long as reference images for 12 different tilt angles of each surface are available. Such reference images may be created either by direct image capture when creating the reference database, or from surface and colour information recovered by 4 source colour photometric stereo [8] and subsequent image rendering [104].

Bibliography

[1] Maestex database. *http://www.cssip.uq.edu.au/staff/meastex/meastex.html*.

[2] Photometric image databases. Photometric Image Databases maintained by the Texture lab. of the Heriot-Watt University, Edinburgh. *http://www.cee.hw.ac.uk/texturelab/database/*.

[3] Vistex database. Vision Texture database maintained by the Vision and Modelling group at the Massachusetts Institute of Technology, MIT Media Lab. *http://www-white.media.mit.edu/vismod/imagery/VisionTexture/vistex.html*.

[4] A. Al-Janobi. Performance evaluation of cross-diagonal texture matrix method of texture analysis. *Pattern Recognition*, 34:171–180, 2001.

[5] J.L. Barron, D.J. Fleet, S.S. Beauchemin, and T.N. Burkitt. Performance of optical flow techniques. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 236–242, Champaigne, Illinois, USA, 1992.

[6] S. Barsky and M. Petrou. Classification of 3d rough surfaces using colour and gradient information recovered by colour photometric stereo. *SPIE Proceedings. Visualization and Optimization Techniques*, 4553:10–19, October 2001.

[7] S. Barsky and M. Petrou. Colour photometric stereo: Simultaneous reconstruction of local gradient and colour of rough textured surfaces. In *International Conference on Computer Vision*, volume II, pages 600–605, 2001.

[8] S. Barsky and M. Petrou. The 4-source photometric stereo technique for 3-dimensional surfaces in the presence of highlights and shadows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, November 2003.

[9] R. Basri and D.W. Jacobs. Photometric stereo with general, unknown lighting. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages II:374–381, 2001.

[10] J. Batlle, A. Casals, J. Freixenet, and J. Martí. A review on strategies for recognizing natural objects in colour images of outdoor scenes. *Image and Vision Computing*, 18(6–7), May 2000.

[11] B. G. Becker and N.L. Max. Smooth transitions between bump rendering algorithms. In *Computer Graphics SIGGRAPH*, pages 183–190, Anaheim, CA, 1993.

[12] M. Ben-Bassat. *Pattern recognition and reduction of dimensionality. In: Handbook of Statistics*. P.R. Krishnaiah and L.N. Kanal, North Holland, 1982.

[13] J.M. Bennet and L. Mattson. *Introduction to surface roughness and scattering*. Optical Society of America, Washington D.C., 1989.

[14] J.F. Blinn. Models of light reflection for computer synthesized pictures. In *Computer Graphics SIGGRAPH*, volume 11, pages 192–198, 1977.

[15] J.F. Blinn. Simulation of wrinkled surfaces. In *Computer Graphics SIGGRAPH*, pages 286–292, 1978.

[16] A. L. Blum and P. Langley. Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97(1–2):245–271, 1997.

[17] A. Bradley, P. Jackway, and B. Lovell. Classification in scale-space: Applications to texture analysis. In *Information Processing in Medical Imaging*, page 375.

[18] P. Brodatz. *Textures, A photographic album for artists and designers*. Dover Publications, New York, 1966.

[19] M.W. Burke. *Image Acquisition: Handbook of Machine Vision Engineering.* Volume 1. New York. Chapman and Hall, 1996.

[20] C. Carson, S. Belongie, H. Greenspan, and J. Malik. Blobworld: Color and texture-based image segmentation using em and its application to content-based image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(8):1026–1038, August 2002.

[21] M.J. Chantler. *The effect of variation in illuminant direction on texture classification. Ph.D. thesis, Dept. Computing and Electrical Engineering, Heriot-Watt University.* PhD thesis, Department of Computing and Electrical Engineering, Heriot-Watt University, August 1994.

[22] M.J. Chantler. Why illuminant direction is fundamental to texture analysis. *IEE Proceedings-Vision Image and Signal Processing*, 142(4):199–206, August 1995.

[23] M.J. Chantler, M. Schmidt, M. Petrou, and McGunnigle G. The effect of illuminant rotation on texture filters: Lissajous's ellipses. In *European Conference on Computer Vision*, volume 3, pages 289–304, Copenhagen, Denmark, May 2002.

[24] M.J. Chantler and J. Wu. Rotation invariant classification of 3d surface textures using photometric stereo and surface magnitude spectra. In *British Machine Vision Conference*, pages 486–495, 2000.

[25] R. Chellappa and S. Chatterjee. Classification of textures using gaussian markov random fields. *IEEE Transactions on Acoustics Speech and Signal Processing*, 33:959–963, 1985.

[26] C. Chen, L. Pau, and P. Wank. *Handbook of Pattern Recognition and Computer Vision.* World Scientific, 1992.

[27] P.H. Christensen and L.G. Shapiro. Three-dimensional shape from colour photometric stereo. *International Journal of Computer Vision*, 13(2):213–227, 1994.

[28] J.J. Clark. Active photometric stereo. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 29–34, 1992.

[29] F.S. Cohen, Z. Fan, and M.A. Patel. Classification of rotated and scaled textured images using gaussian markov random field models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:192–202, 1991.

[30] E.N. Coleman and R. Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Computer Graphics and Image Processing*, 18(4):309–328, April 1982.

[31] R.W. Conners and C.A. Harlow. A theoretical comparison of texture algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(3):204–222, May 1980.

[32] O.G. Cula and K.J. Dana. Recognition methods for 3d textured surfaces. In *SPIE Conference on Human Vision and Electronic Imaging VI*, volume 4299, pages 209–220, January 2001.

[33] K.J. Dana and S.K. Nayar. Histogram model for 3d textures. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 618–624, June 1998.

[34] K.J. Dana and S.K. Nayar. Correlation model for 3d texture. In *International Conference on Computer Vision*, pages 1061–1067, 1999.

[35] K.J. Dana, B. Van Ginneken, S.K. Nayar, and J.J. Koenderink. Reflectance and texture of real-world surfaces. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 151–157, 1997.

[36] K.J. Dana, B. Van Ginneken, S.K. Nayar, and J.J. Koenderink. Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics*, 18(1):1–34, January 1999.

[37] M. Dash and H. Liu. Feature selection for classification. *Intelligent Data Analysis*, 1(3), 1997.

[38] J. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America*, 2(7):1160–69, July 1985.

[39] P. A. Devijver and J. V. Kittler. *Pattern Recognition. A Statistical Approach.* Prentice-Hall, Englewood Cliffs, NJ, 1982.

[40] J. Doak. An evaluation of feature selection methods and their application to computer security. Technical report, Davis, CA: University of California, Department of Computer Science, 1992.

[41] J. Dong and M.J. Chantler. Capture and synthesis of 3d surface texture. *International Workshop on Texture Analysis and Synthesis*, pages 41–45, June 2002.

[42] M.S. Drew. Shape from color. Technical Report CSS/LCCR TR 92-07, Simon Fraser University School of Computing Science, 1992.

[43] M.S. Drew. Optimization approach to dichromatic images. *Journal of Mathematical Imaging and Vision*, 3:189–205, 1993.

[44] M.S. Drew. Direct solution of orientation-from-color problem using a modification of pentland's light source direction estimator. *Computer Vision and Image Understanding*, 64:286–299, 1996.

[45] M.S. Drew. Photometric stereo without multiple images. *Human Vision and Electronic Imaging*, 3016:369–308, February 1997.

[46] J.M.H. Du Buf, M. Kardan, and M. Spann. Texture feature performance for image segmentation. *Pattern Recognition*, 23(3/4):291–309, 1990.

[47] O. Duda and P. Hart. *Pattern Classification and Scene Analysis.* John Wiley & Sons, Inc, 1973.

[48] O.D. Faugeras. *Three-Dimensional Computer Vision.* The MIT Press, Cambridge, Massachusetts, 1993.

[49] P. Favaro and S. Soatto. Learning shape from defocus. In *Proceedings of European Conference on Computer Vision*, volume 2 of *Lecture Notes in Computer Science*, pages 735–745, Copenhagen, Denmark, 2002. Springer.

[50] G.D. Finlayson, J. Dueck, B.V. Funt, and M.S. Drew. Colour eigenfaces. In *International Workshop on Image and Signal Processing*, Manchester, England, November 4-7 1996.

[51] D. Forsyth and J. Ponce. *Computer Vision - A modern approach*. Prentice Hall, Upper Saddle River, N.J., 2002.

[52] S. Fountain, T. Tan, and K. Baker. A comparative study of rotation invariant classification and retrieval of texture images. In *British Machine Vision Conference*, 1998.

[53] R.T. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10:439–451, 1988.

[54] S. Fukuda and H. Hirosawa. A wavelet-based texture feature set applied to classification of multifrequency polarimetric sar images. *IEEE Transactions on Geoscience and Remote Sensing*, 37:2282–2286, May 1999.

[55] E. S. Gelsema. Special issue on genetic algorithms. *Pattern Recognition Letters*, 16(8), 1995.

[56] A. Gonzalez. Model-based texture classification under varying illumination. Technical report, Research Memorandum RM/02/8, Dept. of Computing and Electrical Engineering. Heriot-Watt University, Edinburgh, September 2002.

[57] A. Grau. *Mètode d'Extracció Multiparamètrica de Característiques de Textura Orientat a la Segmentació d'Imatges*. PhD thesis, Departament d'ESAII-UPC, Universitat Politècnica de Catalunya, May 1997.

[58] G.M. Haley and B.S. Manjunath. Rotation-invariant texture classification using modified gabor filters. In *IEEE International Conference on Image Processing*, pages 262–265, 1995.

[59] G.M. Haley and B.S. Manjunath. Rotation-invariant texture classification using a complete space-frequency model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(2):255–269, February 1999.

[60] R.M. Haralick. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67:786–804, 1979.

[61] R.M. Haralick, K.S. Shanmugan, and I. Dunstein. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6):610–621, 1973.

[62] X.D. He, K.E. Torrance, F.X. Sillion, and D.P. Greenberg. A comprehensive physical model for light reflection. *Computer Graphics*, 25(4):175–186, 1991.

[63] G. Healey and T.O. Binford. Local shape from specularity. *Computer Vision, Graphics and Image Processing*, 42:62–86, 1988.

[64] G. Healey and R. Jain. Depth recovery from surface normals. In *IAPR International Conference on Pattern Recognition*, pages 894–896, July 1986. Montreal, Canada.

[65] K.P. Horn. *Obtaining shape from shading information,* in The Psychology of Computer Vision. P.H. Winston, Ed. McGraw-Hill, New York, 1975.

[66] K.P. Horn. Understanding image intensities. *Artificial Intelligence*, 8:201–238, 1977.

[67] K.P. Horn. Height and gradient from shading. *International Journal of Computer Vision*, 5:37–75, 1990.

[68] K.P. Horn and M.K. Brooks. The variational approach to shape from shading. *Computer Vision, Graphics and Image Processing*, 33:174–208, 1986.

[69] K. Ikeuchi. Determining surface orientation of specular surfaces by using the photometric stereo method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3(6):661–669, November 1981.

[70] K. Ikeuchi. Determining a depth map using a dual photometric stereo. *International Journal of Robotics Research*, 6(1):15–31, 1987.

[71] Y. Iwahori, Y. Watanabe, R.J. Woodham, and A. Iwata. Self-calibration and neural network implementation of photometric stereo. In *IAPR International Conference on Pattern Recognition*, pages IV: 359–362, 2002.

[72] Y. Iwahori, R. Woodham, M. Bhuiyan, and N. Ishii. Neural network based photometric stereo for object with non-uniform reflectance factor. In *International Conference on Neural Information Processing*, volume III, pages 1213–1218, 1999.

[73] A.K. Jain, R. P.W. Duin, and J. Mao. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):4–37, January 2000.

[74] A.K. Jain and F. Farrokhnia. Unsupervised texture segmentation using gabor filters. *Pattern Recognition*, 24:1167–1186, December 1991.

[75] A.K. Jain and D. Zongker. Feature selection: Evaluation, application, and small sample performace. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2):153–158, February 1997.

[76] R. Jain, R. Kasturi, and B.G. Schunck. *Machine Vision*. McGraw-Hill, 1995.

[77] G Kay and T. Caelli. Estimating the parameters of an illumination model using photometric stereo. *Graphical Models and Image Processing*, (5):365–388, September 1995.

[78] S.C. Kee, K.M. Lee, and S.U. Lee. Illumination invariant face recognition using photometric stereo. *Machine Vision and Applications*, (7):1466–1474, July 1999.

[79] B. Kim and P. Burger. Depth and shape from shading using the photometric stereo method. *Computer Vision, Graphics and Image Processing*, 54(3):416–427, November 1991.

[80] B. Kim and R. Park. Shape from shading and photometric stereo using surface approximation by legendre polynomials. *Computer Vision and Image Understanding*, 66(3):255–270, June 1997.

[81] J. Kittler. Feature set search algorithms. In *Pattern Recognition and Signal Processing*, pages 41–60, Sijthoff and Noordhoff, Alphen aan den Rijn, The Netherlands, 1978.

[82] J. Kittler. Feature selection and extraction. In T. Y. Young and K. S. Fu, editors, *Handbook of Pattern Recognition and Image Processing*, pages 59–83, Orlando, FL, 1986. Academic Press.

[83] R. Klette and K. Schlüns. Height data from gradient fields. *SPIE Proceedings Machine Vision Applications, Architectures, and Systems Integration*, 2908:204–215, November 1996.

[84] G.J. Klinker. *A Physical Approach to Color Image Understanding*. A.K. Peters, 1993.

[85] J.J. Koenderink and A.J. Van Doorn. Illuminance texture due to surface mesostructure. *Journal of the Optical Society of America - Optics Image Science and Vision*, 13(3):452–463, 1996.

[86] R. Kohavi and G. H. John. Wrappers for feature subset selection. *Artificial Intelligence*, 97(1–2):273–323, 1997.

[87] L.L. Kontsevich, A.P. Petrov, and I.S. Vergelskaya. Reconstruction of shape from shading in colour images. *Journal of the Optical Society of America*, 11(3):1047–1052, March 1994.

[88] M. Kudo and J. Sklansky. Comparison of algorithms that select features for pattern classifiers. *Pattern Recognition*, 33(1):25–41, 2000.

[89] L. I. Kuncheva and L. C. Jain. Nearest neighbor classifier: Simultaneous editing and feature selection. *Pattern Recognition Letters*, 20:1149–1156, 1999.

[90] K.I. Laws. Textured image segmentation. Technical Report 940, Image Processing Institute, University of Southern California, Los Angeles, California, 1980.

[91] S. Lazebnik, C. Schmid, and J. Ponce. Affine-invariant local descriptors and neighborhood statistics for texture recognition. In *International Conference on Computer Vision*, Nice, France, October 2003.

[92] S. Lazebnik, C. Schmid, and J. Ponce. A sparse texture representation using affine-invariant regions. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume II, pages 319–324, Madison, WI., June 2003.

[93] K.M. Lee and C.C.J. Kuo. Shape reconstruction from photometric stereo. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 479–484, 1992.

[94] K.M. Lee and C.C.J. Kuo. Shape from shading with a generalized reflectance map model. *Computer Vision and Image Understanding*, 67(2):143–160, August 1997.

[95] S.W. Lee and R. Bajcsy. Detection of specularity using color and multiple views. *Image and Vision Computing*, 10:643–653, 1992.

[96] M. Leung and A.M. Peterson. Scale and rotation invariant texture classification. In *International Conference on Acoustics, Speech and Signal Processing*, pages 461–465, 1992.

[97] T. Leung and J. Malik. On perpendicular texture: why do we see more flowers in the distance? In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 807–813, June 1997.

[98] T. Leung and J. Malik. Recognising surfaces using three-dimensional textons. In *International Conference on Computer Vision*, pages 1010–1017, September 1999.

[99] T. Leung and J. Malik. Representing and recognising the visual appearance of materials using three-dimensional textons. *International Journal of Computer Vision*, 43(1):29–44, June 2001.

[100] X. Liu, Y. Yu, and H. Shum. Synthesizing bidirectional texture functions for real-world surfaces. In *SIGGRAPH'2001*, pages 97–106, Los Angeles, CA, August 2001.

[101] X. Lladó, J. Martí, and M. Petrou. Surface texture prediction using colour photometric stereo. Technical Report 01-14-RR, Institute of Informatics and Applications. University of Girona, 2001.

[102] X. Lladó, J. Pagès, J. Martí, J. Batlle, and C. Dragoste. Quality control by using surface shape analysis. In *International Conference on Quality Control, Automation and Robotics*, Cluj-Napoca, Romania, May 2002.

[103] X. Lladó and M. Petrou. Classifying textures when seen from different distances. In *IAPR International Conference on Pattern Recognition*, pages 83–86, August 2002.

[104] X. Lladó, M. Petrou, and J. Martí. Surface texture recognition by surface rendering. *IEEE Transactions on Image Processing*, in second revision, 2003.

[105] F. Lumbreras, R. Baldrich, M. Vanrell, J. Serrat, and J.J. Villanueva. Multiresolution colour texture representation for tile classification. In *Simposium Nacional en Reconocimiento de Formas y Análisis de Imagen*, volume I, pages 145–152, Bilbao, Spain, May 1999.

[106] J. Malik and P. Perona. Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America*, 7(5):923–932, 1990.

[107] B.S. Manjunath and R. Chellappa. Unsupervised texture segmentation using markov random field models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(5):478–482, 1991.

[108] R. Manthalkar, P.K. Biswas, and B.N. Chatterji. Rotation and scale invariant texture classification using gabor wavelets. In *International Workshop on Texture Analysis and Synthesis*, pages 87–90, Copenhagen, Denmark, June 2002.

[109] J. Mao and A.K. Jain. Texture classification and segmentation using multiresolution simultaneous autoregressive models. *Pattern Recognition*, 25(2):173–182, 1992.

[110] J. Martí, J. Freixenet, J. Batlle, and A. Casals. A new approach to outdoor scene description based on learning and top-down segmentation. *Image and Vision Computing*, 19(14):1041–1055, December 2001.

[111] G. McGunnigle. *The Classification of Textured Surfaces Under Varying Illuminant Direction. Ph.D. thesis, Dept. Computing and Electrical Engineering, Heriot-Watt University.* PhD thesis, Department of Computing and Electrical Engineering, Heriot-Watt University, June 1998.

[112] G. McGunnigle and M.J. Chantler. Rotation invariant classification of rough surfaces. *IEE Proceedings-Vision Image and Signal Processing*, 146(6), December 1999.

[113] G. McGunnigle and M.J. Chantler. Rough surface classification using point statistics from photometric stereo. *Pattern Recognition Letters*, 21:593–604, 2000.

[114] E. Micheli-Tzanakou. *Supervised and Unsupervised Pattern Recognition. Feature extraction and computational intelligence.* CRC Press. New York, Washington, D.C., 2000.

[115] A. J. Miller. *Subset Selection in Regression.* Chapman and Hall, 1990. Mos H517.23 M647S.

[116] X. Muñoz. *Image segmentation integrating colour, texture and boundary information.* PhD thesis, Department of Electronics, Informatics and Automation. University of Girona, February 2003.

[117] P.M. Narendra and K. Fukunaga. A branch and bound algorithm for feature subset selection. *IEEE Transactions on Computers*, 26(9):917–922, September 1977.

[118] S. K. Nayar, K. Ikeuchi, and T. Kanade. Surface reflection: Physical and geometrical perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):611–634, 1991.

[119] S.K. Nayar and R. Bolle. Reflectance based object recognition. *International Journal of Computer Vision*, 1996.

[120] S.K. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Transactions on Robotics and Automation*, 6(4):418–431, August 1990.

[121] S.K. Nayar, K. Ikeuchi, and T. Kanade. Shape from interreflections. In *International Conference on Computer Vision*, pages 2–11, December 1990.

[122] M. Oren and S.K. Nayar. Generalization of Lambert's reflectance model. *Computer Graphics*, 28(Annual Conference Series):239–246, 1994.

[123] Langley P. Selection of relevant features in machine learning. In *AAAI Fall Symposium on Relevance*, pages 140–144, New Orleans, 1994.

[124] J. Pagès, J. Salvi, and C. Matabosch. Implementation of a robust coded structured light technique for dynamic 3d measurements. In *IEEE Int. Conference on Image Processing*, Barcelona, Spain, September 2003.

[125] A. Penirschke, M.J. Chantler, and M. Petrou. Illuminant rotation invariant classification of 3d surface textures using lissajous's ellipses. In *International Workshop on Texture Analysis and Synthesis*, pages 103–107, Copenhagen, Denmark, June 2002.

[126] A.P. Pentland. Local shading analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(2):170–187, March 1984.

[127] A.P. Pentland. Shading into texture. *Artificial Intelligence*, 29:147–170, 1986.

[128] A.P. Pentland, R.W Picard, and S Sclaroff. Photobook: Content-based manipulation of image databases. *International Journal of Computer Vision*, 18(3):233–254, 1996.

[129] M. Petrou and S. Barsky. Shadows and highlights detection in 4-source colour photometric stereo. In *IEEE International Conference on Image Processing*, pages 967–970, October 2001.

[130] M. Petrou, S. Barsky, and M. Faraklioti. Texture analysis of 3d surface roughness. *Pattern Recognition and Image Analysis*, 2(3):616–632, 2001.

[131] B.T Phong. Illumination for computer generated pictures. *Communications of the ACM*, 18(6):311–317, 1975.

[132] O. Pichler, A. Teuner, and D.J. Hosticka. A comparison of texture feature extraction using gabor filter, pyramidal and tree structured wavelet transforms. *Pattern Recognition*, 29(5):733–742, 1996.

[133] M. Pietikainen, T. Ojala, and Z. Xu. Rotation-invariant texture classification using feature distributions. *Pattern Recognition*, 33(1):43–52, 2000.

[134] P. Pudil, F. Ferri, J. Novovicova, and J. Kittler. Floating search methods for feature selection with nonmonotonic criterion functions. *Pattern Recognition*, 2:279–283, October 1994.

[135] P. Pudil, J. Novovicova, and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15(11):1119–1125, November 1994.

[136] P. Pudil and J. Novovičová. Novel methods for feature subset selection with respect to problem knowledge. *IEEE Intelligent Systems & Their Applications*, 13(2), March/April 1998.

[137] C. Pun and M. Lee. Log-polar wavelet energy signatures for rotation and scale invariant texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5):590–603, May 2003.

[138] T. Randen and J.H. Husoy. Filtering for texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21:291–310, 1999.

[139] T.R. Reed and J.M.H. Du Buf. A review of recent texture segmentation and feature extraction techniques. *Computer Vision, Graphics and Image Processing - Image Understanding*, 57(3):359–372, 1993.

[140] V. Rodehorst. Vertiefende analyse eines gestalts-constraints von aloimonos und shulman. Technical report, CV-Bericht 8, Institut für Technische Informatik, TU Berlin, 1993.

[141] J. Salvi, J. Batlle, and E. Mouaddib. A robust-coded pattern projection for dynamic 3d scene measurement. *International Journal of Pattern Recognition Letters*, (19):1055–1065, September 1998.

[142] A.S. Sanderson, L. Weiss, and S.K. Nayar. Structured highlight inspection of specular surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(1):44–55, January 1988.

[143] K. Schluns and O. Witting. Photometric stereo for non-lambertian surfaces using colour information. In *International Conference on Computer Analysis of Images and Patterns*, pages 444–451, Budapest, Hungary, September 1993.

[144] S.A. Shafer. Using color to separate reflection components. *Color Resolution Applications*.

[145] J. Sherrah. *Automatic Feature Extraction for Pattern Recognition*. PhD thesis, University of Adelaine. South Australia, 1998.

[146] W. Siedlecki and J. Sklansky. On automatic feature selection. *International Journal of Pattern Recognition and Artificial Intelligence*, 2:197–220, 1988.

[147] W. Siedlecki and J. Sklansky. A note on genetic algorithms for large-scale feature selection. *Pattern Recognition Letters*, 10:335–347, November 1989.

[148] W. Silver. *Determining Shape and Reflectance Using Multiple Images*. PhD thesis, MIT, Cambridge, MA., 1980.

[149] S. Singh and M Sharma. Texture analysis experiments with meastex and vistex benchmarks. *Lecture Notes in Computer Science*, (2013):417–424, 2001.

[150] M.L. Smith. The analysis of surface texture using photometric stereo acquisition and gradient space domain mapping. *Image and Vision Computing*, 17:1009–1019, 1999.

[151] M.L. Smith, T. Hill, and G. Smith. Surface texture analysis based upon the visually acquired perturbation of surface normals. *Image and Vision Computing*, 15:949–955, 1997.

[152] M.L. Smith, G. Smith, and T. Hill. Gradient space analysis of surface defects using a photometric stereo derived bump map. *Image and Vision Computing*, 17:321–332, 1999.

[153] F. Solomon and K. Ikeuchi. Extracting the shape and roughness of specular lobe objects using four light photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:449–454, November 1996.

[154] P. Somol, P. Pudil, J. Novovicova, and P. Paclík. Adaptative floating search methods in fetaure selection. *Pattern Recognition Letters*, 20:1156–1163, 1999.

[155] S.D. Stearns. On selecting features for pattern classifiers. In *IAPR International Conference on Pattern Recognition*, pages 71–75, 1976. Coronado, CA.

[156] H. Tagare and R. deFiguiredo. A theory of photometric stereo for a class of diffuse non-lambertian surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(2):133–152, February 1991.

[157] H. Tagare and R. deFiguiredo. Simultaneous estimation of shape and reflectance map from photometric stereo. *CVGIP: Image Understanding*, 55(3):275–286, May 1992.

[158] K.E. Torrance and E.M. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America*, 57(9):1105–1114, 1967.

[159] D.C. Tseng and C.H. Chang. Color segmentation using perceptual attributes. In *IAPR International Conference on Pattern Recognition*, volume A, pages 228–231, 1992.

[160] H. Vafaie and K. De Jong. Robust feature selection algorithms. In *International Conference on Tools with Artificial Intelligence*, pages 356–363, Los Alamitos, CA, USA, November 1993. IEEE Computer Society Press.

[161] H. Vafaie and K. De Jong. Feature space transformation using genetic algorithms. *IEEE Intelligent System*, 13:57–65, 1998.

[162] Van Ginneken, B. and Stavridi, M. and Koenderink, J.J. Diffuse and specular reflectance from rough surfaces. *Applied Optics*, 37(1):130–139, 1998.

[163] L. Van Gool, P. Dewaele, and A. Oosterlinck. Texture analysis anno 1983. *Computer Vision, Graphics and Image Processing*, 29:336–357, 1985.

[164] M. Varma and A. Zisserman. Classifying images of materials: achieving viewpoint and illumination independence. In *European Conference on Computer Vision*, volume 3, pages 255–271, May 2002.

[165] L. Wang and J. Liu. Texture classification using multiresolution markov random field models. *Pattern Recognition Letters*, 20:171–182, 1999.

[166] D. We and J. Linders. A new texture approach to discrimination of forest cleancut, canopy and burned area using airborne c-band sar. *IEEE Transactions on Geoscience and Remote Sensing*, 37(1):555–563, 1999.

[167] M. Weber and R. Cipolla. A practical method for estimation of point lightsources. In *British Machine Vision Conference*, pages 471–480, Manchester, CA, September 2001.

[168] H. Wechsler. Texture analysis - a survey. *Signal Processing*, 2:271–282, 1980.

[169] J. Weszka, C. Dyer, and A. Rosenfeld. A comparative study of texture measures for terrain classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 6(4):269–285, 1976.

[170] R. Woodham. Reflectance map techniques for analyzing surface defects in metal castings. Technical Report AI-TR-457, MIT A.I. Laboratory, June 1978.

[171] R. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.

[172] R. Woodham. Gradient and curvature from the photometric-stereo method, including local confidence estimation. *Journal of the Optical Society of America*, 11(11):3050–3068, November 1994.

[173] R. Woodham, Y. Iwahori, and R. Barman. Photometric stereo: Lambertian reflectance and light sources with unknown direction and strength. Technical Report 91-18, University of British Columbia, Laboratory for Computational Intelligence., 1991.

[174] Z. Wu and L. Li. A line-integration based method for depth recovery from surface normals. *Computer Vision, Graphics and Image Processing*, 43:53–66, 1988.

[175] J. Yang and V. Honavar. Feature subset selection using a genetic algorithm. *IEEE Intelligent Systems & Their Applications*, 13(2), March/April 1998.

[176] A. Yuille, D. Snow, R. Epstein, and P. Peter, Belhumeur. Determining generative models for objects under varying illumination: shape and albedo from multiple images using svd and integrability. *International Journal of Computer Vision*, 35(3):203–222, 1999.

[177] Z. Zhang. The matching problem: The state of the art. Technical Report 2146, Institut National de Recherche en Informatique et en Automatique, December 1993.