

Surface Texture Recognition by Surface Rendering

X. Lladó¹, M. Petrou², and J. Martí

20th September 2004

¹X. Lladó, and J. Martí are with the Computer Vision and Robotics Group, Institute of Informatics and Applications. University of Girona, Campus de Montilivi, 17071 Girona, Spain. Tel: +34-972-418891. Fax: +34-972-418259. E-mail: {llado,joanm}@eia.udg.es

²M. Petrou is with the School of Electronics, & Physical Sciences, University of Surrey, GU2 7XH, Guildford, UK. Tel: +44-1483-689801. Fax: +44-1483-686031. E-mail: m.petrou@eim.surrey.ac.uk

Abstract

A model-based texture recognition system which classifies image textures seen from different distances and under different illumination directions is presented in this paper. The 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-occurrence matrices for feature extraction with a Nearest Neighbour classifier. The use of the co-occurrence matrices instead of filtering methods for feature extraction allows us to utilise only pixels for which valid information has been extracted by CPS.

The validity of the method is demonstrated by classifying texture images captured under different imaging geometries than the reference images in the database. Moreover, the process of recognition allows one to guess the approximate direction of the illumination used to capture the test image.

Keywords: Computer vision, image processing, texture classification, variable illumination, photometric stereo.

1 Introduction

In texture analysis one must distinguish between *image* texture and *surface* texture: image texture is what appears on the 2D image of a physical object, while surface texture refers to the variation of the physical and geometric properties of the imaged surface which give rise to the image texture. Changes in the imaging geometry can significantly alter the appearance of the surface, implying significant variations in the image texture [1]. And yet one has to perform the task of recognition from the image texture.

This work is concerned with the classification of textured surfaces imaged under varying geometries, i.e. varying distance from the sensor and varying illumination direction. Almost all work that has been published in the past on texture features invariant to rotation and scale treated texture as an innate property of a flat surface, unchanged when the camera moves or the illumination changes [2, 3, 4]. We wish to stress that *surface* texture that interests us here is distinct from *image* texture. Image texture is the result of surface relief and surface colour, and if one wishes to avoid its dependence on imaging geometry, one has to go through the fundamentals of image formation and discuss the way surface texture will appear when the imaging geometry changes. So, unless one considers texture as a pattern painted on a smooth surface, it is incorrect to deal with it ignoring the imaging geometry and, for example, use the Brodatz album to evaluate the methodology used.

The effect of varying lighting conditions on supervised texture classification was first investigated by Chantler [5]. It was observed that the direction of illumination affects the directionality of an image obtained from a given surface and indeed modifies its appearance. Nevertheless, most of the classification approaches proposed in the literature do not take into account the effect of illumination on the imaged scene, thus tacitly assume that the illumination direction is constant. Fig. 1 shows three images of the same texture illuminated from two different tilt angles and the same slant angle¹, and obtained

¹The direction of the illuminant with respect to a texture is commonly defined by two polar coordinates: tilt and slant. We define our camera axis parallel to the z -axis. Therefore, tilt is the angle that the illuminant vector makes with the x -axis when it is projected onto the x, y plane. The illuminant slant is the angle that the illuminant vector makes with the camera axis.

from two different distances. Observe how a shift in the tilt angle is clearly manifested in the recorded images. Illumination variation attenuates or accentuates the directional information of the image texture. It should not then be surprising that a classifier would have difficulty in recognising these images as belonging to the same textured surface.

It appears that there are three different ways of dealing with the problem of classifying textures imaged with different illumination directions.

- The first approach consists of extracting and using explicit separate 3-dimensional shape and surface albedo information. The colour and the gradient vector of every visible surface patch describe the surface in a way independent from illumination, 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. On the other hand, Mcgunnigle and Chantler [7] 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.
- Another approach is to study the immediate effects introduced by illumination direction to the observed 2D texture. Chantler [1] has shown that this effect can be described as a directional filter, and in principle it could be inverted. Recently, Chantler et al. [8] presented a formal theory which demonstrates that changes in the tilt of the illumination direction make texture features follow super-elliptical trajectories in multi-dimensional feature spaces. Based on their work, Penirschke et al. [9] developed an illuminant rotation invariant classification scheme. Their work is based on the use of photometric stereo for the detection of surface relief, and the use of Gabor features. The production of features by filtering, however, requires the use of all points of the surface/image. The surface gradients 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 this paper we do not use filtering but co-occurrence matrices, which allow us to work only with the pixels for which reliable information is available.

- Finally, one may train a classifier on a wide selection of images of the same surface, obtained from various viewpoints and under various illumination conditions [10, 11, 12]. Thus, the information of changes in the 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 illuminating conditions. Leung and Malik [13, 14], following this strategy, developed a texture classification scheme that identifies 3D “textons” in the Columbia-Utrecht database for the purpose of illumination 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 [15] proposed a new classification system which uses as texture model a distribution over textons obtained from training images.

Following the last strategy, in this paper we integrate the surface texture information derived by colour photometric stereo (CPS) as described in [16], into a complete model-based texture classification system. Photometric stereo [17, 18] is the technique which allows one to obtain surface texture and colour information from a few images of the same surface imaged under various illumination directions. Basically, the main idea consists of creating, by means of the surface texture information, a “virtual” database of image textures against which we compare the unknown test images in order to classify them. There are various alternative techniques which allow one to obtain 3D information about surfaces [17], like *stereo vision* [19], *optical flow* [20] and various *Shape from X* methods. The main advantage of using Shape from X methods instead of stereo vision and optical flow methods is that they allow the recovery of the local shape as well as the reflectance properties of the surface. These properties are used in different works to predict the surface appearance as a function of illumination, and also to reconstruct the 3D shape

information [21, 22, 23]. Over all these Shape from X techniques, the photometric stereo technique was chosen because it has various advantages over all other methods: it does not make strong assumptions about the underlying surface structure, like some shape from X techniques do, and it allows one to recover both local colour and local gradient while flagging the places where some of its assumptions break down and the recovery is impossible.

Our approach uses CPS to extract surface information from which a “virtual” database of reference textures is created, compatible with the imaging geometry of the test images, and used to classify the test images. The virtual database creation comes in two “flavours”: 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 training images. When the test images are known to have been captured from the same distance as the training images, the creation of the virtual database is straightforward: The use of photometric stereo allows us to define a surface as a set of facets which are normal to the gradient vectors of the surface. Using these facets and given an illuminant vector, rendering of the surface shape can be done showing the *shading* of a facet, which is calculated by taking the dot product of the illuminant vector and the gradient vector. Using this surface shape information and applying the colour map, a rendering of the surface texture can be obtained by multiplying point by point the colour map with the map of shading factors. Hence, a texture image is generated. Things are less straightforward when the test images have been captured from a longer (but known or hypothesised) distance than the training images.

The rest of this work is organised as follows: The methods we use to create the virtual database when the distance between training and testing changes are described in Section 2. In Section 3 we present the image database used for the experiments and the results of the experimental evaluation of the various components of the proposed methodology. Our conclusions are presented in Section 4.

2 Image texture prediction seen from a longer distance

In this section we describe two methods for predicting how a surface texture appears if seen from a longer distance.

First, we analyse the direct relationship between image texture information (image intensities) under two different resolutions. This leads to *direct image prediction*. Then, we present another method which predicts first how the surface itself would be approximated in a lower resolution from the original one, and then predict from it the image it would create. This leads to *image prediction via surface prediction*. We refer to this method as *surface prediction* for short.

2.1 Direct image prediction

We start by considering two grids referring to the pixels of two images of the same surface, captured from two different distances. One of them corresponds to the higher resolution image and it is finer than the other. We refer to the pixels of the coarse resolution as “superpixels” and the term “pixel” is used only for the fine resolution pixels. Each superpixel may be thought of as representing a surface patch characterised by a particular gradient vector $(p, q, -1)^T$ and a particular reflectance function $\rho(\lambda)$. A superpixel has intensity I^u in the coarse resolution image, where u indicates a particular direction of illumination.

Each superpixel is made up of several pixels. Let us say that a superpixel corresponds to a tile of size $K \times L$ of fine resolution pixels. Some of the fine resolution pixels are only partially inside a superpixel, so we do not assume that the two different resolutions are such that a superpixel is replaced by an exact and integer number of fine resolution pixels that tile it exactly.

Our problem is to predict the intensity value of a superpixel I^u , for a given direction of illumination u , given $\rho(\lambda)$, p and q for all the pixels of the fine resolution. The values of

$\rho(\lambda)$, p and q have been computed by using 4-source colour photometric stereo. Although the colour photometric stereo scheme we use can deal with non-Lambertian surfaces [16], we assume here that the surface we are dealing with is Lambertian.

Let us consider a facet of the surface of size $R_A \times R_A$ centred at position (α_0, β_0) and a sensor cell of size $X \times Y$ centred at position (x_0, y_0) . Let us define the following quantities:

- $\mathcal{L}_A(\lambda)$: The energy per unit time per unit area reaching the imaged surface from a point illuminating source at a certain imaging setup denoted by A .
- $N_{\alpha\beta}$: Normal vector to the imaged surface at point (α, β) .
- $\mathcal{G}(N_{\alpha\beta}, u_A)d\alpha d\beta$: Geometric factor determining the amount of incident radiation from the illuminating source at direction u_A , received by an infinitesimal tile of size $d\alpha d\beta$.
- $\rho_{\alpha\beta}(\lambda)$: The fraction of the incident radiation that is reflected by infinitesimal tile $d\alpha d\beta$ at wavelength λ .
- $\mathcal{G}(N_{\alpha\beta}, S_A(x, y))dx dy$: Geometric factor affecting the fraction of reflected radiation reaching the area of the sensor element $dx dy$ centred at (x, y) from surface point (α, β) with the direction of the particular sensor element defined by unit vector $S_A(x, y)$.
- $\frac{1}{2\pi d_A^2}$: Fraction of the reflected radiation reaching a unit area at distance d_A . The quantity $2\pi d_A^2$ is the area of the hemi-sphere over which the opaque point (α, β) is expected to reflect radiation.
- $\mathcal{S}(x, y, \lambda)$: Sensitivity of sensor element (x, y) .
- Δt_A : Time interval during which the aperture of the sensor was open when the image at imaging setup A was being captured.
- k_A : A factor encompassing all photomultiplier and quantization processes used to convert the recorded energy into grey values.

- $rec(x - x_0, y - y_0)$: A rectangular of size $X \times Y$ of sensor surface centred at sensor point (x_0, y_0) .
- $rec(\alpha - \alpha_0, \beta - \beta_0)$: A rectangular of size $R_A \times R_A$ of the imaged surface centred at point (α_0, β_0) .

Then the grey level recorded by sensor $X \times Y$ centred at (x_0, y_0) is:

$$I_A(x_0, y_0) = \frac{\Delta t_A k_A}{2\pi} \int_x \int_y \int_\alpha \int_\beta rec(x - x_0, y - y_0) rec(\alpha - \alpha_0, \beta - \beta_0) \mathcal{G}(N_{\alpha\beta}, S_A(x, y)) \mathcal{G}(N_{\alpha\beta}, u_A) \frac{1}{d_A^2(\alpha, \beta, x, y)} \int_0^\infty \mathcal{S}(x, y, \lambda) \rho_{\alpha\beta}(\lambda) \mathcal{L}_A(\lambda) d\lambda dx dy d\alpha d\beta \quad (1)$$

Let us assume that over the rectangle $rec(x - x_0, y - y_0)$ the properties of the sensor are uniform i.e. $\mathcal{S}(x, y, \lambda) = \mathcal{S}(x_0, y_0, \lambda)$, that over rectangle $rec(\alpha - \alpha_0, \beta - \beta_0)$ the properties of the surface are uniform, i.e. $\rho_{\alpha\beta}(\lambda) = \rho_{\alpha_0\beta_0}$, $N_{\alpha\beta} = N_{\alpha_0\beta_0}$ and that all points of the surface and all points of the sensor are effectively in the same relative orientation defined by unit vector $S_A(x, y) = S_A(x_0, y_0)$ and at the same relative distance² $d_A(\alpha, \beta, x, y) = d_A(\alpha_0, \beta_0, x_0, y_0)$. We may also assume that $S_A(x_0, y_0)$ and $d_A(\alpha_0, \beta_0, x_0, y_0)$ are virtually the same for all surface patches and all sensor elements so we may drop the dependence on (x_0, y_0) and (α_0, β_0) . Then we may write:

$$I_A(x_0, y_0) = \frac{\Delta t_A k_A}{2\pi d_A^2} R_A^2 XY \mathcal{G}(N_{\alpha_0\beta_0}, S_A) \mathcal{G}(N_{\alpha_0\beta_0}, u_A) \int_0^\infty \mathcal{S}(\lambda) \rho_{\alpha_0\beta_0}(\lambda) \mathcal{L}_A(\lambda) d\lambda \quad (2)$$

For a Lambertian surface

$$\mathcal{G}(N_{\alpha_0\beta_0}, S_A) = 1 \quad (3)$$

and

$$\mathcal{G}(N_{\alpha_0\beta_0}, u_A) = \frac{N_{\alpha_0\beta_0} \cdot u_A}{|N_{\alpha_0\beta_0}| |u_A|} = \frac{N_{\alpha_0\beta_0} \cdot u_A}{|N_{\alpha_0\beta_0}|} \quad (4)$$

Let us now assume a different imaging set up, call it B , for which the light which reaches sensor (x_0, y_0) comes from an extended rectangle of size $R_B \times R_B$. If $rec(\alpha -$

²This is equivalent to assuming that the sensor is at infinity with respect to the imaged surface.

$\alpha_0, \beta - \beta_0$) is rather extended, with variable shape and reflectance properties, we must use the following equation for the recorded brightness by sensor (x_0, y_0) :

$$I_B(x_0, y_0) = \frac{\Delta t_B k_B XY}{2\pi} \int_{\alpha} \int_{\beta} \text{rec}(\alpha - \alpha_0, \beta - \beta_0) \mathcal{G}(N_{\alpha_0\beta_0}, S_B) \mathcal{G}(N_{\alpha_0\beta_0}, u_B) \frac{1}{d_B^2(\alpha, \beta)} \int_0^{\infty} \mathcal{S}(\lambda) \rho_{\alpha\beta}(\lambda) \mathcal{L}_B(\lambda) d\lambda d\alpha d\beta \quad (5)$$

We may still assume that no matter how extended rectangle $\text{rec}(\alpha - \alpha_0, \beta - \beta_0)$ is, all its parts are to a high accuracy at the same distance from the sensor, so $d_B(\alpha, \beta) = d_B$. Further, we may still assume Lambertianity, so the two geometric factors take again the form of equations 3 and 4. Then:

$$I_B(x_0, y_0) = \frac{\Delta t_B k_B XY}{2\pi d_B^2} \int_{\alpha} \int_{\beta} \text{rec}(\alpha - \alpha_0, \beta - \beta_0) \frac{N_{\alpha\beta} \cdot u_B}{|N_{\alpha\beta}|} \int_0^{\infty} \mathcal{S}(\lambda) \rho_{\alpha\beta}(\lambda) \mathcal{L}_B(\lambda) d\lambda d\alpha d\beta \quad (6)$$

Let us assume that this big rectangle of size $R_B \times R_B$ may be analysed as a superposition of smaller rectangles, each one of uniform properties for which equation 2 may be written. Then the integral over α and β in equation 6 may be replaced by a sum over all these facets:

$$I_B(x_0, y_0) = \frac{\Delta t_B k_B XY}{2\pi d_B^2} \sum_{k=1}^K \sum_{l=1}^L \frac{N_{kl} \cdot u_B}{|N_{kl}|} A_{kl} \int_0^{\infty} \mathcal{S}(\lambda) \rho_{kl}(\lambda) \mathcal{L}_B(\lambda) d\lambda \quad (7)$$

where A_{kl} is the area of facet (k, l) inside rectangle $R_B \times R_B$. This formula is quite general: it allows us to predict the intensity value of a superpixel from the information we have on its constituent pixels of the fine resolution, even when seen by a different imaging set up, i.e. different sensor, under illumination with different spectral characteristics and different orientation than those under which the original images from which the surface information was extracted, were captured.

If we assume that the spectral properties of the sensor and the light remain constant,

then we may redefine the reflectance function of the surface as

$$\tilde{\rho}_{kl} \equiv \int_{\lambda} \mathcal{S}(\lambda) \rho_{kl}(\lambda) \mathcal{L}_B(\lambda) d\lambda \quad (8)$$

At the same time we may assume that the calibration constant k_B is chosen so that all constant factors cancel out. Then, equation 7 is simplified to:

$$I_B(x_0, y_0) = \sum_{k=1}^K \sum_{l=1}^L A_{kl} \frac{N_{kl} \cdot u_B}{|N_{kl}|} \tilde{\rho}_{kl} \quad (9)$$

This is the equation we used for our experiments.

2.2 Surface prediction

As our objective is to describe surface information, we need to understand what happens with the surface shape information if the distance of the camera changes. We must answer the following question: What will the normal vectors be if the distance of the camera is changed leading to a new image in which every pixel is the union of several old pixels? This question is answered by deriving the relationship between the normal vectors when they are calculated for two different image resolutions.

The proposed strategy in order to recover the normal vectors of the coarse resolution is based on the prediction of the facet by which the surface is approximated locally in this resolution. First, we perform a surface integration in the fine resolution using the gradient vectors recovered by photometric stereo. Next, we reconstruct the facet of the superpixel by using the information recovered in the fine resolution. We predict the planar facet of the superpixel, computing the plane passing through the facets of the fine resolution in the least square error sense. The normal vector of the recovered facet is the normal vector of the superpixel.

Note that surface integration is necessary in order to know the height $z = S(x, y)$ (up to an additive constant) at each point of the surface in the fine resolution (*height map*). If we define a surface by $(x, y, S(x, y))$ where $S(x, y)$ is the height at point (x, y) , the normal

as a function of (x, y) is

$$N(x, y) = \frac{1}{\sqrt{p^2 + q^2 + 1}}(p, q, -1)^T \quad (10)$$

where the partial derivatives of $S(x, y)$ give us the components of the gradient vector $p = \frac{\partial S}{\partial x}$ and $q = \frac{\partial S}{\partial y}$. There are two types of surface integration approaches [24], *local integration along paths* and *global integration*. Path integration techniques use local calculations of height increments by curve integrals. The main idea is to use the gradient vectors, which give the change in surface height with a small step, to recover a height map of the surface by summing these changes in height along some path. In global techniques, the surface integration is treated as an optimisation problem. In other words, surface integration is considered as a variational problem where a certain functional has to be minimised. Usually, some representation of the unknown surface, e.g. in terms of the Fourier basis functions, and an integrability condition, e.g. the smoothness constraint [25, 22], are used to constrain the global optimisation process. Obviously, this global integration technique works well only if the surface is smooth and continuous.

The photometric stereo surface reconstruction yields also the set of points where the recovery of the gradient vectors is impossible. In general, these are places which are in shadow in more than one of the four images used in the photometric stereo set. Such points are likely to occur at places where the surface has deep “ravines” and they may be associated with the places where the surface is not differentiable and continuous. Therefore, as we need to preserve as much as possible the surface roughness, and we want to perform surface integration using only points with available gradient vectors, we apply a local integration technique which uses different paths to obtain the final surface reconstruction. Thus, when we reconstruct the surface we stop the piecewise integration at the boundaries where the surface is not differentiable and continuous. Note that we predict the normal vectors only of the superpixels for which surface integration in the fine resolution is possible. Therefore, the surface shape we reconstruct and the image intensity we predict consist of image patches and not full images.

3 Experiments

The experiments we have performed have as purpose to check various aspects of the theory, namely:

1. The accuracy of CPS to use images captured with certain illumination direction in order to predict images referring to the same camera distance, but different illumination directions (section 3.2).
2. The accuracy of *image prediction* using a photometric set captured at distance A to predict images captured from distance B ($>$ distance A). The predictions will be compared with real images captured from distance B and also with images produced from a photometric set captured from distance B (section 3.3).
3. 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 and with the same light orientations as that used for the set at distance A (section 3.4).
4. The accuracy of the classifier when photometric sets at distance B are used to produce virtual model images for the same distance B , in terms of which images captured from distance B are to be classified (section 3.5).
5. The accuracy of the classifier when photometric sets at 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 (section 3.6).

3.1 Experimental data

Twenty five physical texture samples were used throughout the experimental trials presented in this work. For each texture two photometric sets composed of 4 images each were available. The two sets were captured from two different distances, distance A and

distance B (longer than A). All photometric stereo sets consisted of 4 images illuminated at a slant angle of 55° and with 4 different tilt angles 0° , 90° , 180° and 270° . As well as the 4 images used in the photometric sets, different images for each surface and from the two distances were captured for testing. We call these images tA and tB . The lights or the object position were not changed when the distance of the sensor changed. The first thirteen surfaces were captured using 12 illuminant tilt angles between 0° and 360° incremented in steps of 30° . The remaining ones were captured using 24 illuminant tilt angles between 0° and 360° incremented in steps of 15° . All surfaces were illuminated at a slant angle of 55° .

Virtual images constructed from photometric sets of the same distance are referred to as images AA and BB . Virtual images constructed from photometric set A for distance B are referred to as predictions AB . In order to distinguish between the direct image prediction method of section 2.1 and the image prediction method via the surface prediction of section 2.2, we shall refer to them as prediction $imaAB$ and $surAB$, respectively.

Fig. 2 shows one image for each surface texture captured from distance A and one from distance B . These images include two major subsets of surface textures:

- One subset of surfaces consists of a wide variety of relatively smooth surfaces which may be further divided into *isotropic surfaces* and *directional surfaces*.
- The other subset of surfaces consists of a variety of very rough surfaces for which the assumption on which photometric stereo is based (i.e. relatively smooth surface with low roughness) is violated. We do not expect CPS to work well for such surfaces, but we included them in order to test to the extreme the proposed method.

The description of a surface may be made 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. $AASR$ is calculated as

$$AASR = \frac{1}{2n} \sum_x \sum_y |p(x, y)| + |q(x, y)| \quad (11)$$

where n is the number of samples for which the (p, q) values are known. For other purposes, in which 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 work we use both descriptors to characterise the surface shape and roughness of each texture, namely the estimated probability density functions (PDFs) for the surface partial derivatives p and q (i.e. the normalised histograms of these quantities), and the *AASR* parameter. The p -map and q -map 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.

Analysing all the PDFs we identified three different types of gradient distribution, corroborating the diversity of the dataset. For instance, let us consider three textures (T5, T14, and T15) which provide representative examples of each texture class defined above (see first column of Fig. 3).

We observed that some surfaces, like the first example (T5), which was formed by a fracture process, are essentially random textures. Note that the PDFs for this texture could be modelled by Gaussian distributions (see first row of Fig. 3). Similar behaviour can be observed for instance for textures T1, T6, and T10. The second example, which corresponds to a directional texture, has one gradient component which is close to a typical distribution of a surface with sinusoidal height profile (see the PDF of p in the second row of Fig. 3). 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 which are 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 it is shown in the third example of Fig. 3. Other textures which follow this behaviour are textures T17, T18, T20, and T21.

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 subsets 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).

3.2 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 illumination directions. We compare the set of test images captured from distance B (tB) with the corresponding images for the same illumination conditions created from photometric stereo information (BB).

To quantify the difference between a captured colour image and a generated one, we use the mean square error of colour differences computed over all pixels. In order to compute the colour difference between the predicted and true colour values for a pixel, we use again the Euclidean metric in conjunction with the Luv colour space. 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 1 we give the average mean square error and its standard deviation for each type of surface we used and for all twenty five textures together, for

all 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 always the same image texture captured under a particular illumination direction is used as the reference image. Fig. 4 shows an illustrative example of the MSE obtained when comparing the test images tB with the images generated by photometric stereo BB , and also 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. The testing has been carried out on texture T21 which is one of the cases with the least accurate generated images with an average MSE of 12.7424. The average MSE for the naive case is however 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 the 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, should be preferable for the classification of textures under varying lighting conditions, to the use of a fixed reference image.

3.3 Experiment 2: Accuracy of image prediction when the distance changes

This experiment has as purpose to evaluate the accuracy of image prediction, using a photometric set captured at distance A , to predict how an image will appear from distance B . Using the image prediction method of section 2.1, the image intensities are directly predicted from the photometric information extracted at distance A . We call these predicted images $imaAB$. However, using the surface prediction method of section 2.2 only the surface gradient vectors are directly predicted but not the image intensities. We generate predicted images by using the predicted gradient vectors at distance B and the average reflectance function for each surface tile, and assuming that our sensor has narrow sensitivity (i.e. $\mathcal{S}(\lambda)$ is a delta function in equation 6) while the illuminant is white (i.e. $\mathcal{L}(\lambda)$ is a constant in equation 6). 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 $surAB$.

Some results obtained over three textures are shown in Fig. 5, 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 , and the third curve shows the results of comparing the image created by use of the information extracted by photometric stereo at 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.

Observing these 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 between them. In table 1 we give an overall quantitative assessment for each method by computing the average MSE and its standard deviation over all textures and all tilt angles. The average MSE is similar for each prediction approach, although the image prediction method, which predicts directly the pixel intensities, 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 1). Other surface properties such as directionality or specularity may contribute to the errors too. For example, in the directional texture T14 of Fig. 5 some orientations of light source provoke more difficulties than others, and therefore the error of image prediction may significantly vary depending on the illumination direction.

3.4 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 at distance B .

Using the surface prediction method described in section 2.2, the gradient vectors are directly predicted from the photometric information extracted at distance A . However, using the image prediction method (see section 2.1) only the intensity values can be predicted but not the gradient vectors. Therefore, to make possible the surface shape evaluation, photometric stereo has been applied to these predicted images (four images corresponding to four directions of the illumination) in order to compute the gradient vectors.

The first column of Fig. 6 shows three examples of image predictions $imaAB$ with a particular direction of illumination. Note that the image intensity we predict consists of image patches and not full images since there are points which photometric stereo can not correctly recover. We flag these points with white values.

Before we perform the evaluation it is necessary to solve the problem of localising which region of the original set at 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 directly to the original set, and results obtained with our prediction AB . When computing 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 exactly the region of the original set. After that, the PDFs of p and q are used in order to compare the surface shape information.

Each plot of Fig. 6 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 better the original PDF distributions. On the other 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 2 we confirm this conclusion by providing an overall quantitative assesment of all these histogram comparisons, namely the average MSE of each histogram and its standard deviation, computed over all textures, as well as the average MSE with which the gradient components p and q of individual pixels are computed. Both these measures indicate that better results are obtained with the surface prediction approach. The reason of this may be attributed to 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 it was mentioned earlier, we used the *absolute average slope ratio* ($AASR$) 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 at distance B . Table 3 gives an overall quanti-

tative assessment over all textures computing the average MSE of the $AASR$ parameter obtained using 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 an important influence in the accuracy of the surface shape predictions. For rougher surfaces the error of the predictions is increased (see tables 2 and 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 ($AASR = 0.3960$) is 0.1019 and 0.1573 respectively, while for texture T5, a relatively smooth surface ($AASR = 0.1003$) is 0.0470 and 0.0558, respectively. Observe that sometimes the error obtained by directional surfaces is closer to that obtained from rough surfaces. That is due to the directional textures T13 and T14 which have an important degree of roughness ($AASR$ equal 0.2476 and 0.1973, respectively).

3.5 Experiment 4: 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 at 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 subset of surface textures (isotropic, directional, and rough surfaces).

For each surface texture 4 images were rendered using a slant angle of 55° and 4 illuminant tilt angles: 10° , 100° , 190° and 280° . This is the virtual database of images used as references for classification. It is composed of 100 texture images (25 surfaces \times 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 in two steps: the learning process and the classification process. The learning process has the goal of modeling each texture class of the virtual database by means of a representative feature vector, while the classification process has the goal of classifying an unknown test image into the texture class 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-occurrence matrices [26] 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). We used all these distances to ensure that images containing very big texels were properly characterised. The pixels labeled as un-reconstructed points (points the shape and colour of which were unreliably calculated by photometric stereo) were not used in the computation of the co-occurrence matrices. The co-occurrence matrices were implemented in an anisotropic way, using 256 grey levels. We used four different directions: 0° , 45° , 90° , 135° , so that we had in all 240 texture features ($3 \text{ features} \times 4 \text{ directions} \times 20 \text{ distances } d$). Among all the computed features, those which could discriminate between the different classes best were chosen. We used the Sequential Forward Selection (SFS) algorithm [27] and a set of training images in order to select the best feature set for discrimination. The feature evaluation was performed 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 had in all 600 rendered images = $25 \text{ textures} \times 8 \text{ illuminations} \times 3 \text{ images}$.

When the learning is finished, the classification process starts by extracting first the feature vectors for the unknown test (real) images, and classifying them by means of the Nearest Neighbour classifier into one of the classes of the virtual database. At the same time, this classification process allows one to approximately estimate the illuminant tilt

angle of the test image by identifying it with the one of its nearest neighbour in the virtual database.

Due to the large size of the captured images, we produced from each one of them different subimages of size 133×133 , to be used for testing. The test set consists of 9 real images for each surface texture and each illuminant tilt angle. We have in all 2664 images ($13 \text{ textures} \times 8 \text{ illuminations} \times 9 \text{ images} + 12 \text{ textures} \times 16 \text{ illuminations} \times 9 \text{ images}$). Note that different tilt angles are used for training and testing, as in a real situation one does 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 in the “best” case, in which the 4 reference images of each texture are real images captured with the camera, and not images rendered by photometric stereo. We also compare the results with those obtained in the “naive” case, in which just a single captured image of one illuminant direction is used to characterise each texture class.

Feature selection was performed separately for each case. Table 4 lists the selected features for each case. In general, the selected feature sets include a variety of features (C : contrast, H : homogeneity, E : energy), computed for a variety of distances and a variety of directions.

As well as classifying the test images captured from distance B into the corresponding texture class, we may also classify the illuminant tilt angle under which they had been captured, into one of the 4 illuminant tilt angles of the images in the virtual database. Note that 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 correctly in 70.15% of the cases by the model-based approach. However, in the “best” case an accuracy of 97.06% was obtained (see table 5). 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, and not so much on the 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.

Using our model-based approach we obtained a 97.04% accuracy of texture classification. However, in the “best” case the texture classification accuracy was 100%. For the “naive” case the texture classification accuracy was 87.73% (see table 5). Analysing the missclassification error of the model-based approach (2.96%), we concluded that 84.98% of it was produced by the subset of very rough surfaces, while the remaining 15.02% of it was produced by the isotropic surfaces. This indicates that the missclassification errors of the model-based approach were mainly due to the image generation by colour photometric stereo information. The results demonstrate that the model-based approach reduces significantly the texture classification errors caused by changes in illuminant direction compared with the “naive” case.

3.6 Experiment 5: 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 at distance A are used to produce model images for distance B .

The virtual database of images corresponding to the longer distance B is generated using the image prediction method described in 2.1. As for experiment 4, the virtual database is composed of 100 image textures. In this experiment 600 virtual 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 we classified 89.57% of them into the correct texture class. We had 100% correct texture classification in the “best” case, when reference and test images were captured from distance B (note that this case corresponds exactly to the “best” case used in experiment 4). The “naive” case here is the case when one uses as the reference image one that was captured from a different distance than that of the test image. In such a 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 subset of very rough surfaces, while the isotropic surfaces contributed only 7.39% of this error.

Table 5 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 4, we conclude that the performance of the classifier is decreased due to the error introduced by the image prediction method of section 2.1. However, we also demonstrate that the model-based approach increases significantly the accuracy of the texture classification compared with the “naive” case. In the same table we also present the percentage of test images for which the illumination tilt angle under which they had been captured was correctly predicted.

4 Summary and Conclusions

Two different prediction methods were proposed, one which allows one to directly predict the image intensities (direct image prediction), and another which allows one to predict first the surface shape information 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 of the system to predict the *image texture* a particular *surface texture* will create when seen from a longer than the original distance. The direct image prediction method produced in general smaller errors. The major part of this error was due to the photometric stereo and not due to the step dealing with the change of distance. On the other hand, the surface prediction method produced the best shape predictions. Several error measures were used in order to evaluate the surface shape predictions: the absolute average slope ratio, 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 also the MSE per pixel of the gradient components p and q .

As a result of these experiments one may conclude that surface roughness plays an

important role in the accuracy of image and shape prediction. The rougher the surface is, the larger the errors of the predictions. Other surface properties such as directionality or specularity may also contribute to the errors.

However, even if the predicted images are not perfect, they may still make significant difference in the accuracy of a texture classifier. It was shown that when using such generated images to extract features in terms of which textures were classified, the results were significantly better than those obtained by the “naive” case where the features used were extracted from a single real image. The classification results were not much inferior to those obtained by the “best” case when test images and training images were captured under exactly the same imaging geometries. The improvement applied also to the cases for which the assumptions made by the photometric stereo technique we used were violated and the photometric stereo results were not as accurate as one might have wished. In all cases an estimate of the unknown illumination orientation under which the test image was captured could also be obtained, although not with very high accuracy.

Although we have 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 changes in the camera direction. If we have the information of the surface relief, and the surface albedo, and we use the general equations we derived, then we could render the surface to create model textures as required.

Acknowledgments

This work was supported by UdG grant BR00/03, and by EPSRC grant GR/M73361.

References

- [1] M. Chantler, “Why illuminant direction is fundamental to texture analysis,” *IEE Proceedings-Vision Image and Signal Processing* 142, pp. 199–206 (1995).

- [2] M. Leung and A. Peterson, “Scale and rotation invariant texture classification,” in *International Conference on Acoustics, Speech and Signal Processing*, pp. 461–465 (1992).
- [3] R. Manthalkar, P. Biswas, and B. Chatterji, “Rotation and scale invariant texture classification using gabor wavelets,” in *International Workshop on Texture Analysis and Synthesis*, pp. 87–90, (Copenhagen, Denmark) (2002).
- [4] 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, pp. 590–603 (2003).
- [5] M. 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 (1994).
- [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, pp. 10–19 (2001).
- [7] G. McGunnigle and M. Chantler, “Rough surface classification using point statistics from photometric stereo,” *Pattern Recognition Letters* 21, pp. 593–604 (2000).
- [8] M. Chantler, M. Schmidt, M. Petrou, and M. G., “The effect of illuminant rotation on texture filters: Lissajous’s ellipses,” in *European Conference on Computer Vision*, 3, pp. 289–304, (Copenhagen, Denmark) (2002).
- [9] A. Penirschke, M. Chantler, and M. Petrou, “Illuminant rotation invariant classification of 3d surface textures using lissajous’s ellipses,” in *International Workshop on Texture Analysis and Synthesis*, pp. 103–107, (Copenhagen, Denmark) (2002).
- [10] K. Dana and S. Nayar, “Histogram model for 3d textures,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 618–624 (1998).

- [11] K. Dana, B. Van Ginneken, S. Nayar, and J. Koenderink, “Reflectance and texture of real-world surfaces,” *ACM Transactions on Graphics* 18, pp. 1–34 (1999).
- [12] O. Cula and K. Dana, “Recognition methods for 3d textured surfaces,” in *SPIE Conference on Human Vision and Electronic Imaging VI*, 4299, pp. 209–220 (2001).
- [13] T. Leung and J. Malik, “Recognising surfaces using three-dimensional textons,” in *International Conference on Computer Vision*, pp. 1010–1017 (1999).
- [14] T. Leung and J. Malik, “Representing and recognising the visual appearance of materials using three-dimensional textons,” *International Journal of Computer Vision* 43, pp. 29–44 (2001).
- [15] M. Varma and A. Zisserman, “Classifying images of materials: achieving viewpoint and illumination independence,” in *European Conference on Computer Vision*, 3, pp. 255–271 (2002).
- [16] 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* , pp. 1239–1252 (2003).
- [17] R. Woodham, “Reflectance map techniques for analyzing surface defects in metal castings,” Tech. Rep. AI-TR-457, MIT A.I. Laboratory (1978).
- [18] R. Woodham, “Photometric method for determining surface orientation from multiple images,” *Optical Engineering* 19(1), pp. 139–144 (1980).
- [19] O. Faugeras, *Three-Dimensional Computer Vision*, The MIT Press, Cambridge, Massachusetts (1993).
- [20] J. Barron, D. Fleet, S. Beauchemin, and T. Burkitt, “Performance of optical flow techniques,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 236–242, (Champaign, Illinois, USA) (1992).

- [21] Van Ginneken, B. and Stavridi, M. and Koenderink, J.J., “Diffuse and specular reflectance from rough surfaces,” *Applied Optics* 37(1), pp. 130–139 (1998).
- [22] 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), pp. 203–222 (1999).
- [23] X. Liu, Y. Yu, and H. Shum, “Synthesizing bidirectional texture functions for real-world surfaces,” in *SIGGRAPH’2001*, pp. 97–106, (Los Angeles, CA) (2001).
- [24] D. Forsyth and J. Ponce, *Computer Vision - A modern approach*, Prentice Hall, Upper Saddle River, N.J. (2002).
- [25] R. Frankot and R. Chellappa, “A method for enforcing integrability in shape from shading algorithms,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 10, pp. 439–451 (1988).
- [26] R. Haralick, K. Shanmugan, and I. Dunstein, “Textural features for image classification,” *IEEE Transactions on Systems, Man, and Cybernetics* 3(6), pp. 610–621 (1973).
- [27] J. Kittler, “Feature set search algorithms,” in *Pattern Recognition and Signal Processing*, pp. 41–60, (Sijthoff and Noordhoff, Alphen aan den Rijn, The Netherlands) (1978).

Biographies

Dr. Xavier Lladó. Xavier Lladó received the B.S. degree in Computer Science from the University of Girona in 1999 before joining the Computer Vision and Robotics Group. He also received the PhD in Computer Engineering from the University of Girona. At present, he is a Post-doctoral Research Assistant in the Department of Computer Science

at the Queen Mary, University of London. His current interest are in the field of image processing and computer vision, focusing on Colour and Texture analysis, Shape from Shading, and Non-rigid Structure from Motion.



Prof. Maria petrou. Maria Petrou studied Physics at the Aristotle University of Thessaloniki, Greece, Applied Mathematics in Cambridge and she did her PhD in the Institute of Astronomy in Cambridge, UK. She has been working on image processing and computer vision since 1986. She has been the Professor of Image Analysis since 1998 and leads a group of 20 researchers on this topic in Surrey University. She has published more than 250 scientific papers, on Astronomy, Remote Sensing, Computer Vision, Machine Learning, Colour analysis, Industrial Inspection, Medical Signal and Image Processing. She has co-authored a book “Image Processing: the fundamentals” published by John Wiley in 1999 and reprinted in 2000 and 2003, and numerous book chapters. She is a Fellow of the Royal Academy of Engineering, Fellow of IEE and Fellow of IAPR. She has served as the Chairman of the Technical Committee for Remote Sensing of IAPR, the Chairman of the British Machine Vision Association (BMVA), as an Associate Editor of IEEE Transactions on Image Processing, as the Newsletter Editor of IAPR and is currently the treasurer of IAPR and an Honorary Editor of IEE Electronics Letters. A full list of publications and other details can be found in <http://www.ee.surrey.ac.uk/Personal/M.Petrou>



Dr. Joan Martí. Joan Martí received the BSc degree in Physics from the Autonomous University of Barcelona (UAB) in 1986. He received the PhD degree in Computer Vision from the Polytechnical University of Catalonia (UPC) in 1998. He is currently a senior lecturer in Computer Vision at the University of Girona (UdG), in the Institute of Informatics and Applications. His research interests include knowledge-based systems, segmentation of biomedical images, and colour and texture- based segmentation for object recognition purposes.



List of Tables

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.	31
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. (2) the average MSE per pixel of gradient components p and q and their standard deviation.	31
3	Quantitative assessment using the $AASR$ parameter. Average MSE and its standard deviation for each prediction approach ($imaAB$ and $surAB$). .	32
4	Feature sets selected for experiments 4 and 5 (best, model-based and naive cases).	32
5	Texture and illuminant classification rates obtained for the best, model-based and naive cases, when photometric sets at distance B are used to create images at distance B , and also when the direct image prediction method is used to generate images at distance B from the photometric sets at distance A	33

Table 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.

Approach	Isotropic		Directional		Rough		Overall	
	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

Table 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. (2) the average MSE per pixel of gradient components p and q and their standard deviation.

	Prediction	Isotropic		Directional		Rough		Overall	
		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

Table 3: Quantitative assessment using the *AASR* parameter. Average MSE and its standard deviation for each prediction approach (*imaAB* and *surAB*).

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

Table 4: Feature sets selected for experiments 4 and 5 (best, model-based and naive cases).

Experiments	Selected Features (<i>feature</i> _{distance,direction})
Best case (Exp. 4, 5)	$C_{10,90^\circ}, H_{22,0^\circ}, H_{1,90^\circ}, H_{40,90^\circ}, C_{4,0^\circ}, H_{25,135^\circ}, C_{37,45^\circ}, E_{43,45^\circ}, H_{55,135^\circ},$ $C_{1,0^\circ}, H_{46,90^\circ}, H_{28,90^\circ}, E_{52,45^\circ}$
Model-based (Exp. 4)	$C_{4,90^\circ}, H_{1,0^\circ}, H_{52,0^\circ}, H_{13,90^\circ}, E_{52,90^\circ}, H_{43,45^\circ}, E_{52,0^\circ}, H_{25,45^\circ}, H_{10,0^\circ},$ $C_{34,45^\circ}, H_{52,135^\circ}, H_{7,90^\circ}, H_{10,90^\circ}, H_{7,45^\circ}, E_{55,135^\circ}, C_{16,90^\circ}, H_{19,90^\circ}$
Naive case (Exp. 4)	$C_{37,0^\circ}, H_{16,90^\circ}, E_{31,45^\circ}, C_{37,0^\circ}, C_{1,135^\circ}, H_{1,90^\circ}, H_{43,0^\circ}, C_{16,45^\circ}, C_{19,45^\circ},$ $C_{22,90^\circ}, C_{19,135^\circ}, C_{25,0^\circ}, H_{1,135^\circ}, C_{22,45^\circ}, H_{37,0^\circ}, C_{46,135^\circ}, C_{31,45^\circ}, C_{10,45^\circ}$
Model-based (Exp. 5)	$H_{22,90^\circ}, H_{49,90^\circ}, C_{16,0^\circ}, C_{13,90^\circ}, H_{28,0^\circ}, C_{22,135^\circ}, H_{4,0^\circ}, H_{55,135^\circ}, H_{13,0^\circ},$ $H_{19,45^\circ}, H_{4,90^\circ}, H_{4,45^\circ}, H_{52,90^\circ}, H_{46,45^\circ}, E_{49,0^\circ}, E_{7,135^\circ}$
Naive case (Exp. 5)	$H_{19,90^\circ}, E_{49,45^\circ}, C_{46,45^\circ}, E_{52,90^\circ}, H_{28,90^\circ}, H_{52,45^\circ}, C_{19,45^\circ}, E_{52,90^\circ}, C_{28,90^\circ},$ $C_{40,90^\circ}, E_{1,0^\circ}, H_{37,90^\circ}, H_{25,90^\circ}, H_{22,135^\circ}, C_{16,90^\circ}$

Table 5: Texture and illuminant classification rates obtained for the best, model-based and naive cases, when photometric sets at distance B are used to create images at distance B , and also when the direct image prediction method is used to generate images at distance B from the photometric sets at distance A .

Texture	Experiment 4					Experiment 5		
	Best case		Model-based		Naive case	Model-based		Naive case
	Texture	Illuminant	Texture	Illuminant	Texture	Texture	Illuminant	Texture
T1	100%	100%	100%	75.01%	46.67%	100%	77.87%	0%
T2	100%	100%	100%	31.29%	100%	100%	77.87%	35.55%
T3	100%	100%	100%	88.20%	100%	100%	73.56%	0%
T4	100%	100%	100%	75.01%	73.33%	93.33%	28.07%	0%
T5	100%	100%	100%	81.26%	64.44%	100%	83.40%	53.33%
T6	100%	100%	100%	100%	76.29%	100%	100%	0%
T7	100%	100%	100%	81.26%	100%	87.41%	100%	85.76%
T8	100%	97.04%	100%	75.71%	46.67%	100%	61.27%	100%
T9	100%	100%	88.89%	14.63%	73.33%	100%	17.00%	46.67%
T10	100%	100%	100%	66.68%	97.04%	100%	55.73%	100%
T11	100%	100%	100%	91.67%	46.67%	100%	39.13%	0%
T12	100%	100%	100%	57.66%	100%	100%	77.87%	23.61%
T13	100%	100%	100%	100%	100%	100%	92.62%	40.40%
T14	100%	100%	100%	100%	100%	100%	77.87%	5.92%
T15	100%	100%	100%	63.91%	100%	100%	100%	26.67%
T16	100%	87.41%	91.85%	41.00%	100%	95.56%	60.65%	65.93%
T17	100%	95.56%	99.56%	68.07%	100%	87.41%	81.56%	44.44%
T18	100%	94.07%	59.70%	59.74%	100%	48.15%	82.17%	73.11%
T19	100%	94.07%	100%	61.83%	100%	100%	58.81%	85.67%
T20	100%	82.96%	85.93%	76.40%	100%	9.63%	23.76%	0%
T21	100%	98.52%	100%	80.57%	93.33%	23.70%	35.44%	0%
T22	100%	95.56%	100%	72.24%	100%	100%	77.87%	71.84%
T23	100%	97.04%	100%	84.73%	100%	94.07%	78.48%	0%
T24	100%	93.33%	100%	63.91%	75.56%	100%	75.41%	0%
T25	100%	91.11%	100%	43.09%	100%	100%	33.60%	0%
Overall	100%	97.06%	97.04%	70.15%	87.73%	89.57%	66.80%	34.35%

List of Figures

1	Three images of the same surface texture captured using different illuminant tilt angles, (a) and (b), and different distance of the camera, (b) and (c).	35
2	One image for distance A and one for distance B of each of the twenty five sample textures.	37
3	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	38
4	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	39
5	Accuracy of image prediction for 3 surface textures under varying tilt angle (textures T22, T14, and T10). Curve tB vs BB compares real images captured from distance B with those created using the information extracted by photometric stereo from distance B too. Curves tB vs $imaAB$ and tB vs $surAB$ compare predicted images with data from distance A , with real images captured from distance B	40
6	Accuracy of the surface shape predictions for three surface textures (textures T5, T14, and T15). The second and third columns show the distributions of the surface gradients p and q obtained by applying directly photometric stereo to the original images at distance B , by applying photometric stereo to the images predicted by the method of section 2.1 ($imaAB$), and by applying the surface prediction method of section 2.2 ($surAB$).	41

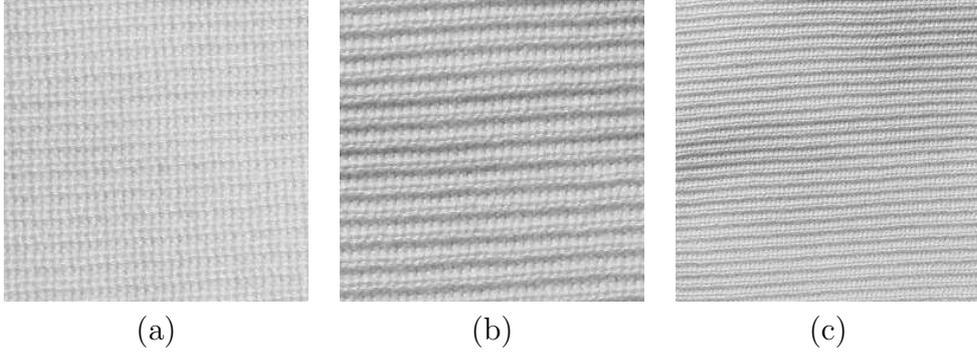
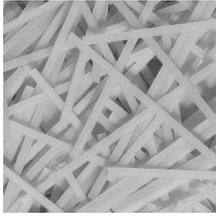
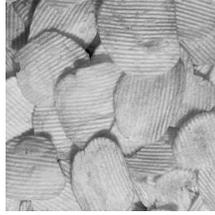
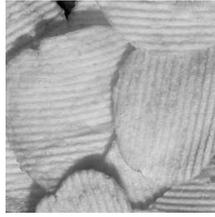


Figure 1: Three images of the same surface texture captured using different illuminant tilt angles, (a) and (b), and different distance of the camera, (b) and (c).

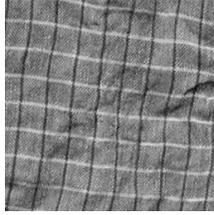
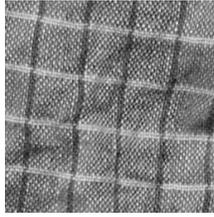
Pasta (T1)



Chips (T2)



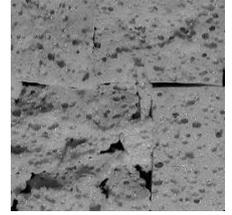
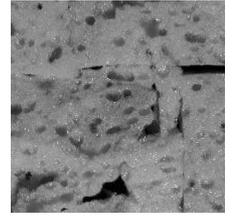
Cloth (T3)



Cloth (T4)



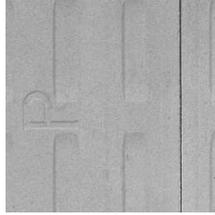
Sponge (T5)



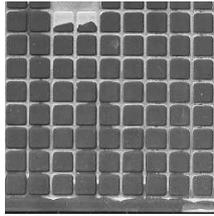
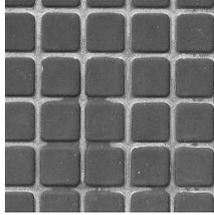
Sponge (T6)



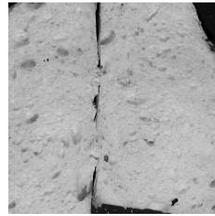
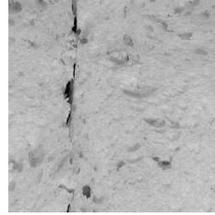
Brick (T7)



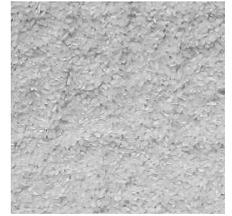
Brick (T8)



Bread (T9)



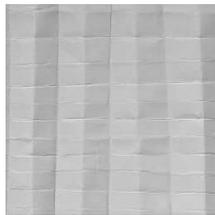
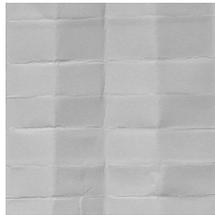
Rice (T10)



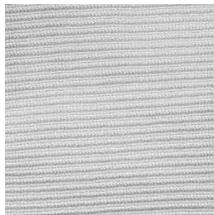
Pasta (T11)



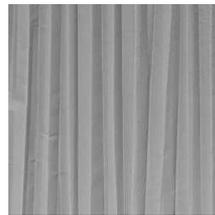
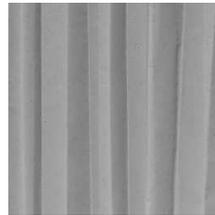
Paper (T12)



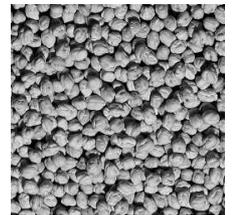
Cloth (T13)



Paper (T14)



Chickpeas (T15)



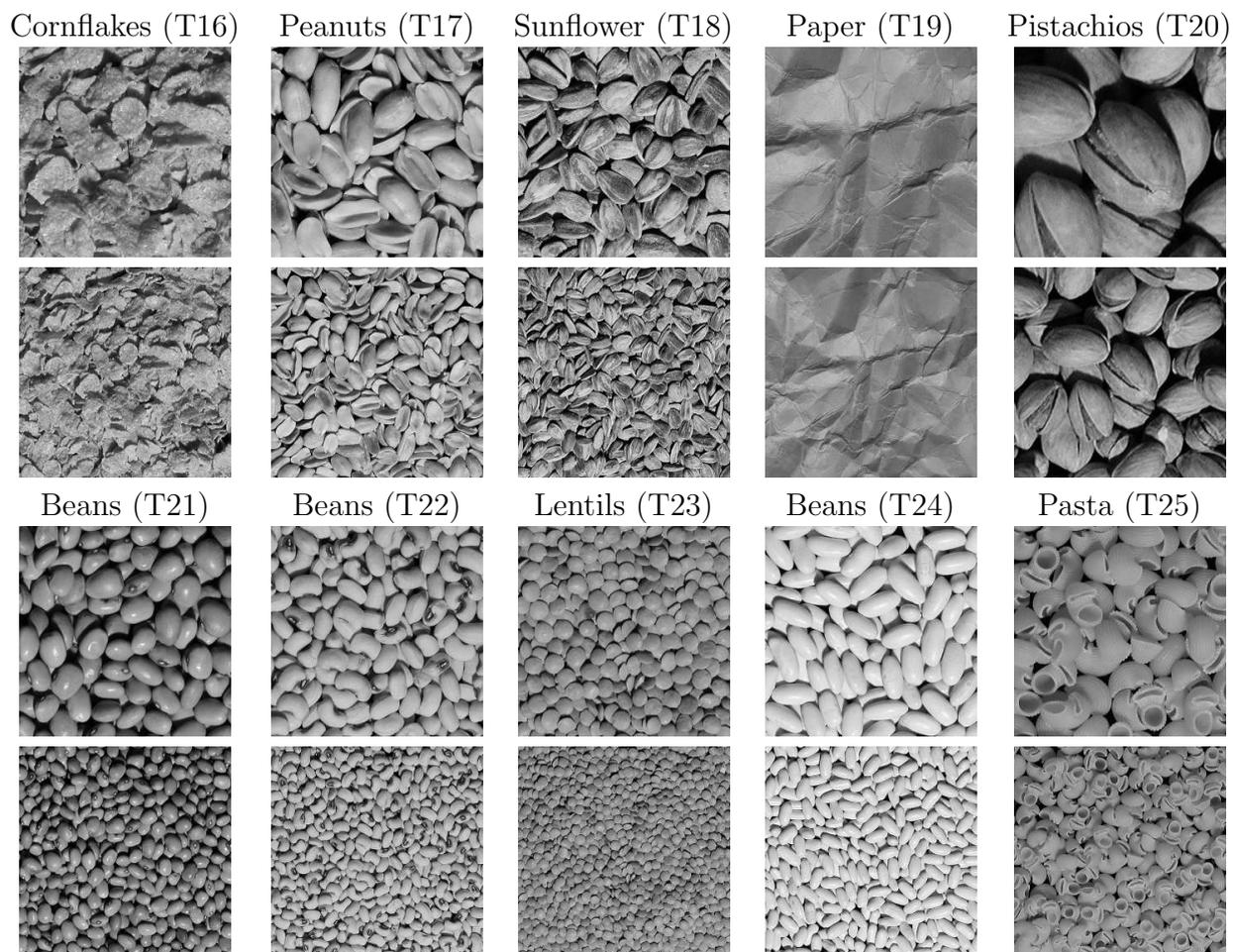


Figure 2: One image for distance A and one for distance B of each of the twenty five sample textures.

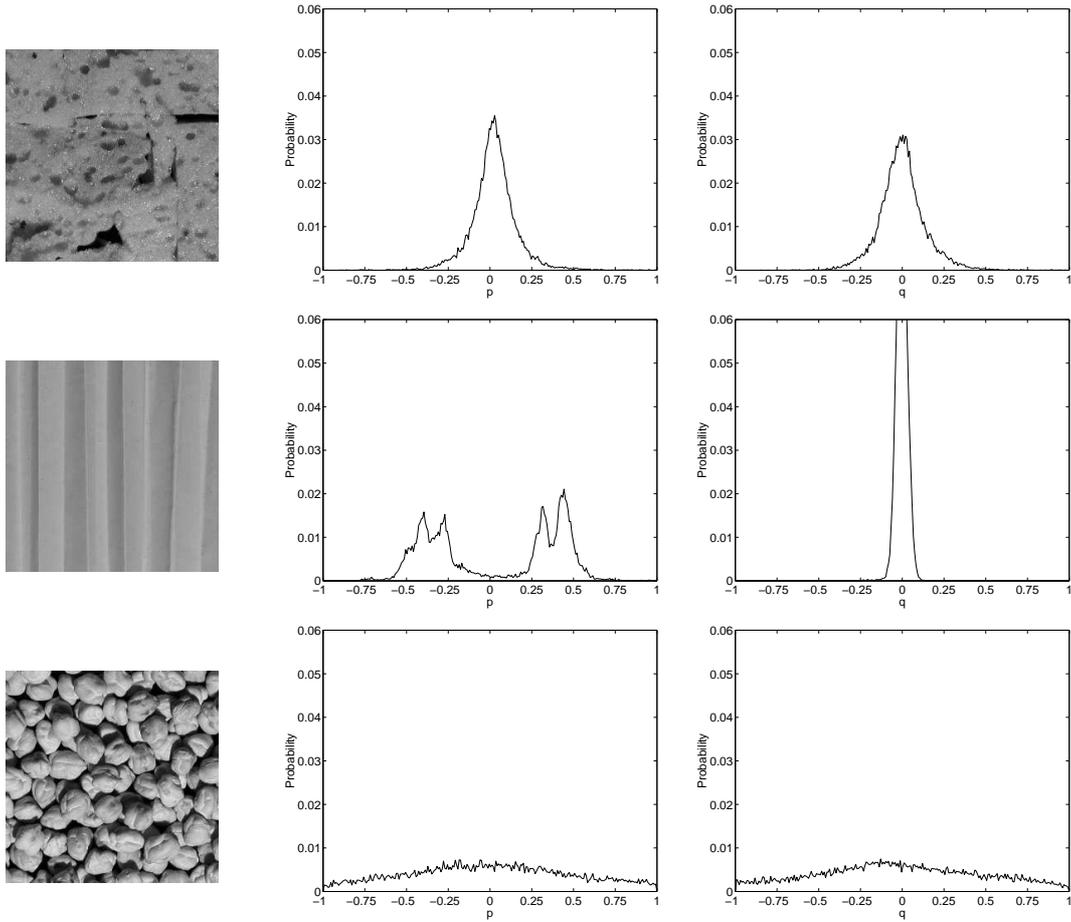


Figure 3: One representative image of an isotropic (T5), a directional (T14) and a very rough isotropic surface (T15). The second and third columns show the PDF representations of the surface gradients p and q .

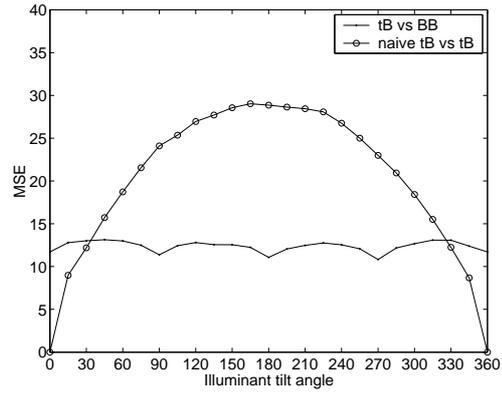
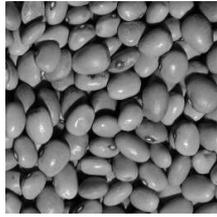


Figure 4: Mean square errors for texture T21 under varying tilt angles. MSE for the evaluation tB vs BB and the naive case tB vs BB , in which the process of surface recovery and rendering has been bypassed, and the image at tilt= 0° is compared directly with images tB .

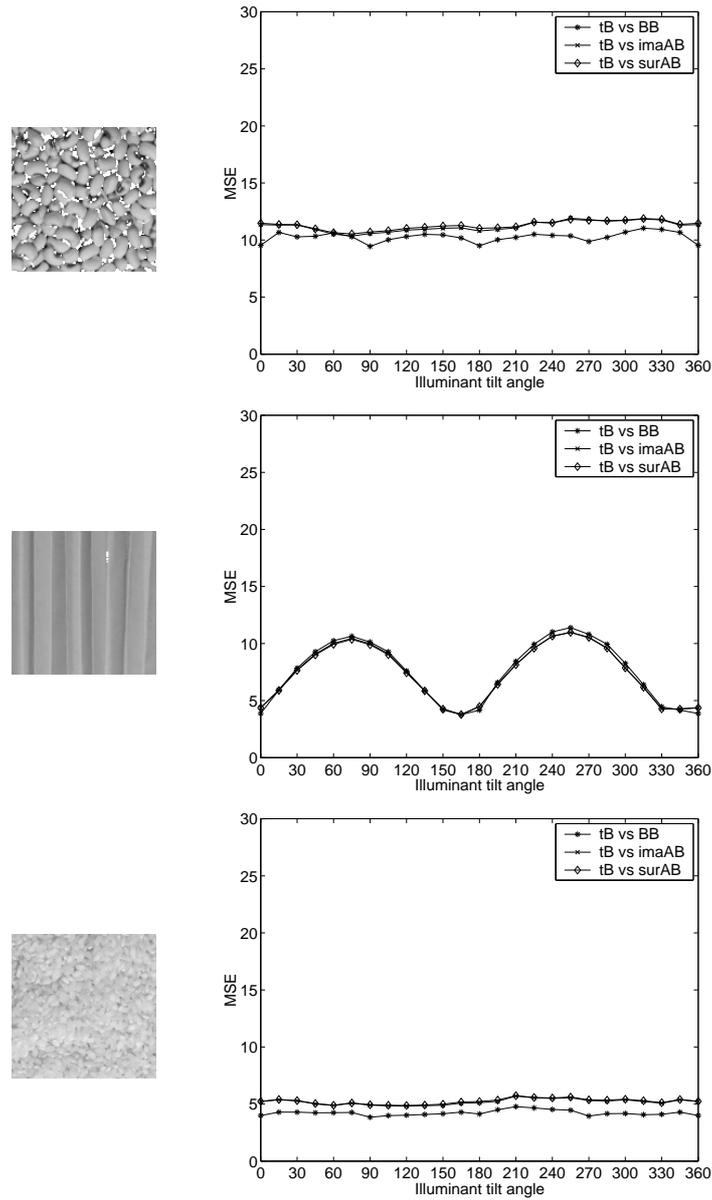


Figure 5: Accuracy of image prediction for 3 surface textures under varying tilt angle (textures T22, T14, and T10). Curve tB vs BB compares real images captured from distance B with those created using the information extracted by photometric stereo from distance B too. Curves tB vs $imaAB$ and tB vs $surAB$ compare predicted images with data from distance A , with real images captured from distance B .

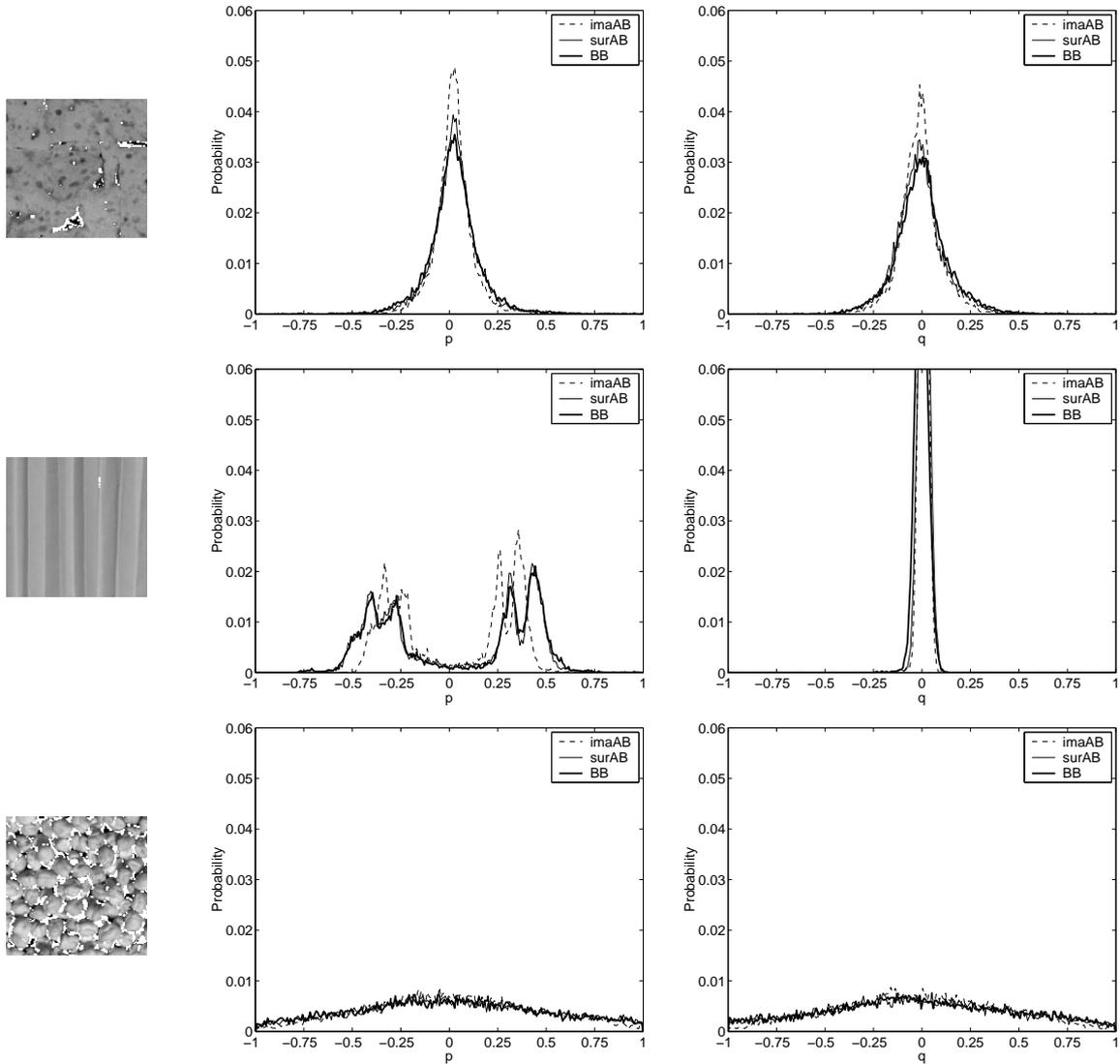


Figure 6: Accuracy of the surface shape predictions for three surface textures (textures T5, T14, and T15). The second and third columns show the distributions of the surface gradients p and q obtained by applying directly photometric stereo to the original images at distance B , by applying photometric stereo to the images predicted by the method of section 2.1 ($imaAB$), and by applying the surface prediction method of section 2.2 ($surAB$).