Image Texture Prediction using Colour Photometric Stereo

Xavier Lladó[1], Joan Martí[1], and Maria Petrou[2]

[1] Institute of Informatics and Applications, University of Girona,
17071, Girona, Spain
{llado,joanm}@eia.udg.es
[2] Centre for Vision, Speech, and Signal Processing, University of Surrey,
GU2 5XH, Guildford, Surrey, England
m.petrou@ee.surrey.ac.uk

Abstract. The purpose of this work is to analyse what happens to the surface information when the image resolution is modified. We deduce how the surface texture appears if seen from different distances. Using Colour Photometric Stereo a method for predicting how surface texture looks like when changing the distance of the camera is presented. We use this technique on the recognition of textures seen from different distances. Real sets of images have been used in order to evaluate the performance of the recognition system.

1 Introduction

The main motivation for this paper is the problem of description of multicoloured surfaces invariant to the geometry. Recognition of 3-dimensional surface textures from 2-dimensional images is difficult. The 2-dimensional texture in the image, the *image* texture, is produced by variation in both surface reflectance and surface relief. The latter two constitute the *surface* texture. While the reflectance properties are intrinsic to the surface, the surface relief produces a pattern of shadings that depends strongly on the direction of the illumination [1]. Thus, the *image* texture created by a 3D *surface* texture changes drastically with the imaging geometry.

This paper uses *Colour Photometric Stereo* (CPS), as described in [2] and [3], to compute the detailed shape and colour of a rough surface when seen by a camera at the zenith of the surface. Photometric stereo is based on the fact that image intensities depend on the surface orientation and its reflectance. Hence, if several images are taken from the same viewing position but with different lighting directions, variation of pixel intensities in these images will be due to changes in the relative positions of the illuminant and the surface [4]. This constraint permits us to calculate the normal vectors, which represent the surface orientation of any point on the surface, and the reflectance factor or albedo, which describes the reflection properties of the surface.

We assume that in a database of textures, we have all the information concerning the surface texture constructed from the photometric stereo set. We

then assume that we are given the image of one of these textures captured by a camera at a different (longer) distance and with unknown direction of illumination. From the original information in the database, we predict how each surface would look like when seen from the new distance, for various directions of illumination. Thus we create a "virtual" database of image textures against which we compare the unknown image texture in order to classify it. Recognition of the texture allows us also to guess the approximate orientation of the illumination under which the image was captured. The image texture classifier we use is based on the co-ocurrence matrices [5].

The rest of this work is organised as follows. In section 2 and 3 the image prediction method is explained. In section 4, the method is validated using real sets of images in the context of texture recognition. Finally, the work ends with conclusions and future work.

2 Image prediction process

Our purpose is to predict how surface texture looks like if seen from a different distance. In order to do this, it is necessary to understand what happens with the colour and surface shape information if the distance of the camera is changed.

We shall start by considering two grids referring to the pixels of two images of the same surface, captured from two different distances. One of them must correspond to the higher resolution image and it must be finer than the other. Let us indicate by indices ij a pixel of the coarse grid. This pixel is made up from several pixels of the fine resolution grid, some of which contribute to it only partially. Let us for the moment ignore by how much each pixel of the fine resolution contributes to pixel ij of the coarse resolution, and let us simply say that "superpixel" ij corresponds to a tile of size $K \times L$ of fine resolution pixels. We shall refer to the pixels of the coarse resolution as "superpixels" and the term "pixel" will be used only for the fine resolution pixels. Each superpixel may be thought of as representing a surface patch characterised by a particular gradient vector $(p_{ij}, q_{ij}, 1)^T$ and a particular reflectance function $\rho_{ij}(\lambda)$. The superpixel will have intensity I_{ij}^u in the coarse resolution image, $u = 1, 2, 3$ or 4, each corresponding to a different direction of the illumination.

Each superpixel corresponds to a tile of pixels. We wish to keep track of the superpixel to which a pixel contributes. So, we shall give to every pixel three sets of indices: one tells us to which tile it belongs, one tells us where about in the tile it is, and one tells us its location in the fine resolution grid. Let us indicate by indices mn the position of pixels in the fine resolution grid. So, a pixel that contributes to superpixel ij will have indices $ij; kl\ mn$, where $k = 1, 2, ..., K$ and $l = 1, 2, ..., L$. Any other quantity associated with pixel $ij; kl\ mn$ will be indicated by the same notation as for superpixel ij. That is, pixel $ij; kl\ mn$ corresponds to a surface patch with gradient vector $(p_{ij;kl}^{mn}, q_{ij;kl}^{mn}, 1)^T$ and a reflectance function $\rho_{ij;kl}^{mn}(\lambda)$. Our problem is to predict I_{ij}^u, for a given direction of illumination u, given $\rho_{ij;kl}^{mn}(\lambda)$, $p_{ij;kl}^{mn}$ and $q_{ij;kl}^{mn}$ for all values of i, j, k and l. The values of $\rho_{ij;kl}^{mn}(\lambda)$, $p_{ij;kl}^{mn}$ and $q_{ij;kl}^{mn}$ have been computed from four

images by Colour Photometric Stereo. We shall not go into details of this now as they have been published elsewhere [2, 3]. Although the CPS scheme we use can deal with non-Lambertian surfaces, we assume here that the surface we are dealing with is Lambertian.

If the sensitivity of the sensor is $\mathcal{S}(\lambda)$, the spectral distribution of the incident light is $\mathcal{L}(\lambda)$, and the reflectance funcion of the surface patch projected on superpixel ij is $\mathcal{R}(\lambda, N_{ij}, u)$, where N_{ij} is the normal to the surface vector and u the direction of illumination vector, then the intensity value registered by the sensor at superpixel ij is:

$$I_{ij}^u = \int \mathcal{S}(\lambda)\mathcal{L}(\lambda)\mathcal{R}(\lambda, N_{ij}, u)d\lambda \tag{1}$$

$\mathcal{R}(\lambda, N_{ij}, u)$ could be separated in a geometric component $\mathcal{G}(N_{ij}, u)$ and a surface material component, $\rho_{ij}(\lambda)$. Assuming we are dealing with a Lambertian surface,

$$\mathcal{G}(N_{ij}, u) = \frac{N_{ij} \cdot u}{|N_{ij}||u|} \tag{2}$$

where N_{ij} is $(p_{ij}, q_{ij}, 1)^T$, so:

$$I_{ij}^u = \int \mathcal{S}(\lambda)\mathcal{L}(\lambda)\mathcal{G}(N_{ij}, u)\rho_{ij}(\lambda)d\lambda \tag{3}$$

Superpixel ij is made up from several pixels each of which may have its own reflectance function, and its own orientation. So, $\mathcal{G}(N_{ij}, u)\rho_{ij}(\lambda)$ is really the sum of several such factors, one for each pixel that contributes to the superpixel. Then

$$I_{ij}^u = \int \mathcal{S}(\lambda)\mathcal{L}(\lambda) \sum_{k,l=1}^{K,L} \mathcal{G}(N_{ij;kl}^{mn}, u)\rho_{ij;kl}^{mn}(\lambda)d\lambda \tag{4}$$

By exchanging the order of integration and summation, we obtain:

$$I_{ij}^u = \sum_{k,l=1}^{K,L} \int \mathcal{S}(\lambda)\mathcal{L}(\lambda)\mathcal{G}(N_{ij;kl}^{mn}, u)\rho_{ij;kl}^{mn}(\lambda)d\lambda \tag{5}$$

Note that this formula is quite general: it allows us to predict the value of superpixel ij from the information we have on its constituent pixels, even when seen by a different sensor, under illumination with different spectral distribution and different orientation than those under which the original images, from which the surface information was extracted, were captured.

We shall restrict ourselves here to the case where the sensor and light source are the same. If we assume that the unknown illumination direction is the same as one of the illumination directions in the original data used by the photometric stereo, then this equation collapses to a trivial one:

$$I_{ij}^u = \sum_{k,l=1}^{K,L} I_{ij;kl}^{u;mn} \tag{6}$$

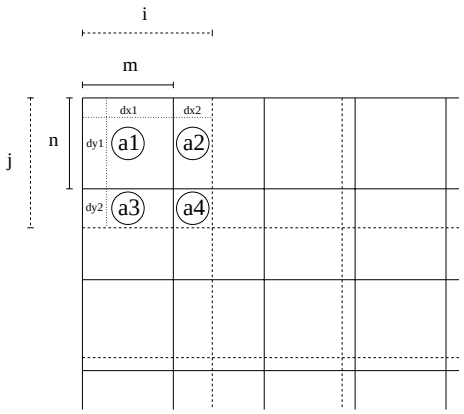


Fig. 1. General case of image resolution change.

This approach allows us to go straight to predicting how an image of the surface would look like when imaged under different conditions, from those under which the images in the database were imaged.

Now, we wish to analyse the general case, in which superpixel ij is made up of several pixels, some of which contribute only partially (see figure 1). In this situation, it is necessary to know by how much each pixel of the fine resolution (area a_{kl}) contributes to superpixel ij of the coarse resolution. In this sense we calculate the point spread function of the imaging device. Then it is not necessary to compute these areas a_{kl} which after all are only a crude approximation of the point spread function, but the weights arising from the point spread function itself can be used instead.

3 Point Spread Function

Practically every imaging system introduces some kind of degradation into the data it captures. A common phenomenon when deling with very fine textures is the spatial blurring that the imaging system introduces into the data. This phenomenon can be quantified in terms of how spread a point source appears to be when its image is captured. This is expressed by the *point spread function* (PSF) of the system.

As we explained at the end of section 2, in the more general case of our prediction method, it is not necessary to compute the areas a_{kl} which after all are only a crude approximation of the point spread function. Thus, the objective now is to obtain the PSF of the imaging system and use it in the prediction method.

In order to derive the PSF [6], an image with many edges of various known orientations is needed. However, it may not be necessary to use more than one orientation if the PSF of the sensor is circularly symmetric. To establish whether

Fig. 2. Test chart for the derivation of the point spread function.

the PSF of our sensor is circularly symmetric or not, a simple chart like the one in figure 2 is used. This chart can be used to measure the PSF at 0^o, 45^o, 90^o and 135^o degrees. The test chart is captured using the camera-grabber system used throughout our experiments, using the viewing and illumination conditions used for all the experiments reported.

To check the symmetry of the sensor, we compute the derivative of the image at 0^o, 45^o, 90^o and 135^o degrees using the Robinson operators [7].

The profiles of the resultant images along several lines orthogonal to the original edges are computed and averaged to produce the four profiles for 0^o, 45^o, 90^o and 135^o plotted in figure 3. These are the profiles of the point spread function. Note that first column shows the derivative images obtained. Two of the four profiles of the point spread function plotted there are narrower than the other two. This is because they correspond to orientations 45^o and 135^o and the distance of the pixels along these orientations is $\sqrt{2}$ longer than the distance of pixels along 0^o and 90^o.

In figures 4.a and 4.b we plot separately the two pairs of profiles and see that the system has the same behaviour along the 0^o, 90^o and 45^o, 135^o orientations. Taking into account the $\sqrt{2}$ correction for the 45^o and 135^o, we conclude that the point spread function of this imaging system is to a high degree circularly symmetric. Figure 4.c shows the plot of the four profiles while in figure 4.d we zoom into the central part of the plot of figure 4.c. Then, in our practical application these four profiles ara averaged to produce a single cross-section of a circularly symmetric point spread function.

When the profile of the PSF is known next step is to use this information in order to infer the weights of the filter which will be used in the prediction method instead of the areas a_{kl}. To do that, the obtained profile has been fitted to a gaussian distribution. The obtained gaussian distribution has $\mu = -0.038889$ and $\sigma = 1.157407$. Afterwards, the weights of the PSF filter which we shall refer as g are obtained as follows:

$$\begin{aligned} g_0 &= 1 \\ g_x &= e^{-\frac{x^2}{2\sigma^2}} \\ g_N &= e^{-\frac{N^2}{2\sigma^2}} \end{aligned} \tag{7}$$

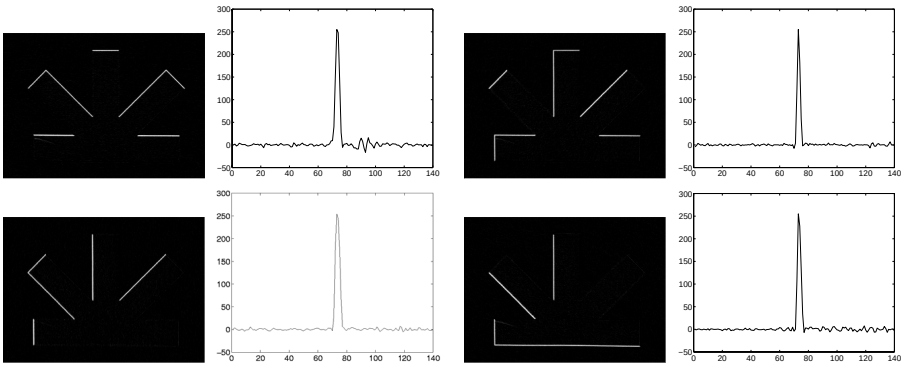
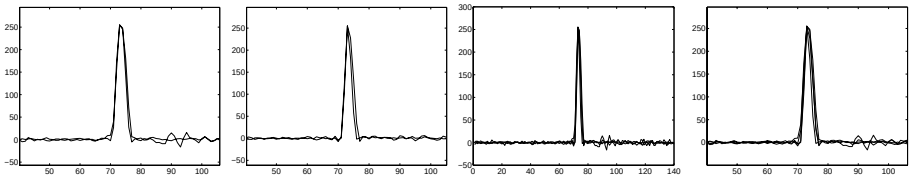


Fig. 3. Derivative images and PSF profiles at 0°, 45°, 90° and 135° degrees.



(a) (B) (c) (d)

Fig. 4. The point spread function of the imaging system. (a) PSF profile for orientations 0° and 90°. (b) PSF profile for orientations 45° and 135°. (c) Four profiles of the PSF. (d) Zooming into (a).

In order to obtain the dimension of the filter we have imposed $g_N = 0.01$, consequently N can be obtained as

$$-\frac{N^2}{2\sigma^2} = \ln 0.01 \Rightarrow N = \lceil \sqrt{-2\sigma^2 \ln 0.01} \rceil \tag{8}$$

Furthermore, when the weights of the filter are obtained, a normalisation of them has done. We shall refer to the normalised filter as $\tilde{g}$ and it can be calculated as

$$\tilde{g}_{ij} = \frac{g_{ij}}{\sum_{i,j=1}^{N} g_{ij}} \tag{9}$$

As a result of this process our PSF filter $\tilde{g}$ is symmetric and of size 9×9.

Note that incorporating the use of PSF filter instead of areas a_{kl} the predicting method has replaced the exploration of a tile of size $K \times L$ pixels of the fine resolution by the exploration of a tile of size $N \times N$, where N is the dimension of the PSF filter. Hence, each pixel has a contibution in the superpixel ij equal to the weight of the PSF filter position which it corresponds.

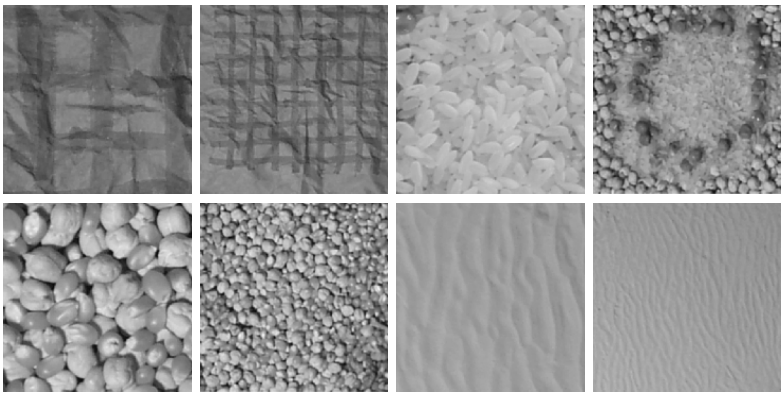


Fig. 5. First and third column: one image of the photometric stereo set. Second and fourth column: images captured from a longer distance.

4 Experimental results

The proposed method was tested with several photometric sets consisting of four images each, obtained by placing the camera at a distance along the zenith of the surface. Other images of the same surfaces were captured from a longer distance for 4 different illumination directions. These are the images we shall want to classify, using as reference images those of the photometric stereo sets.

The photometric database consists of seven surfaces. First and third column of figure 5 show one of the photometric set images for four of the surfaces. From the information in the original database, we predict how each surface looks like when seen from the longer distance, for 4 different directions of illumination. Hence we create a "virtual" database of 28 image textures against which we compare the unknown image texture in order to classify it.

We start our classification system by obtaining a representative feature vector for each texture image in the "virtual" database, using the co-ocurrence matrices. The co-ocurrence matrices were implemented in an isotropic way for fixed distance d, computing two of the most typical features, contrast and homogeneity, for 60 different values of d. Among all the computed features, those which could discriminate between the different classes best were chosen. These turned out to be contrast at distance 29, and homogeneity at distance 2. After that we build a classifier for this set of features, which calculates the feature vector for the unknown image texture, and assigns it to one of the classes of the "virtual" database. Second and fourth column of figure 5 show examples of the unknown images we want to classify.

As these images are large, we produce from each one of them 9 subimages to be used for testing. Thus, we had in all $7 \times 9 \times 4 = 252$ test images of size 94×94 each (7 different surfaces, 4 different directions of illumination, 9 subimages).

Table 1. Confusion matrix and illuminant classification rates for the four illuminant directions (O_1,O_2,O_3,O_4).

	T1	T2	T3	T4	T5	T6	T7
T1	0.89	0.11	0.0	0.0	0.0	0.0	0.0
T2	0.25	0.47	0.08	0.0	0.0	0.17	0.03
T3	0.0	0.0	1.0	0.0	0.0	0.0	0.0
T4	0.0	0.0	0.0	0.97	0.03	0.0	0.0
T5	0.0	0.0	0.0	0.0	0.86	0.14	0.0
T6	0.0	0.0	0.0	0.0	0.14	0.86	0.0
T7	0.0	0.0	0.0	0.0	0.0	0.0	1.0

	O_1	O_2	O_3	O_4
T1	0.67	0.25	1.0	0.50
T2	0.50	0.57	0.0	0.0
T3	0.55	0.89	0.11	0.78
T4	0.89	0.33	0.75	0.44
T5	0.44	1.0	1.0	0.89
T6	0.55	0.44	0.83	0.29
T7	1.0	0.44	0.67	0.67

The texture classification rate obtained classifying the unknown images into one of the 7 surfaces (without considering illumination direction) was 84.74%. While the illuminant classification rate obtained (considering 7 surfaces and 4 different direction of illumination) was 64.87%.

In order to illustrate these classification results we used the so called *confusion matrices* $A = \alpha_{ij}$, where α_{ij} is the frequency with which a test example from class i is classified as belonging to class j. Ideally a confusion matrix should be the unit matrix. Table 1 shows the confusion matrix when we classify the test images to one of the texture classes of the database. Table 1 also shows the classification rate for the four illuminant directions.

In two alternative experiments we tried to classify the same patterns without going through the prediction stage, but just using features constructed from the original high resolution images (original database). In this case we only classified correctly into the 7 surfaces 22.28% of patterns. In a third experiment, we used again features constructed from the original images but we scaled up the distance d used for the construction of the co-ocurrence matrix according to the change in image resolution. In this case the texture recognition rate of the classifier was 33.69%. Hence, we demonstrate it is necessary to go through the prediction stage in order to extract the features which allows the best recognition results.

5 Conclusions

We presented a general framework for recognising textures when seen from different distances. The 4-source CPS has been used in order to obtain the reflectance and the surface shape information of the surface from a close by distance. The proposed method allows one to predict how the texture will look like when seen by a different sensor and under different imaging geometry with an illuminant of different spectral properties. It is based on the assumption of Lambertian surfaces, but it can easily be generalised to other types of surface. The method has been validated using real sets of images in the context of texture recognition when the test data have been captured from different distances than those in the database.

Our futher studies will be focused on the inclusion of more features in order to characterise the textures. The idea is to introduce colour as well as additional textural features with the goal to improve the recognition rates of the classifier.

6 Acknowledgments

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

References

1. Chantler, M.: Why illuminant direction is fundamental to texture analysis. IEE Proceedings in Vision, Image and Signal Processing **142** (1995) 199–206
2. Barsky, S., Petrou, M.: Colour photometric stereo: Simultaneous reconstruction of local gradient and colour of rough textured surfaces. In Proceedings 8th International Conference on Computer Vision (2001) II: 600–605
3. Petrou, M., Barsky, S.: Shadows and highlights detection in 4-source colour photometric stereo. In Proceedings International Conference on Image Processing (2001) 967–970
4. Woodham, R.: Gradient and curvature from the photometric-stereo method, including local confidence estimation. Journal of the Optical Society of America, A **11** (1994) 3050–3068
5. Haralick, R., Shanmugan, K., Dunstein, I.: Textural features for image classification. IEEE Transactions on Systems, Man, and Cybertinetics **3** (1973) 610–621
6. Petrou, M., Bosdogianni, P.: Image Processing. The Fundamentals. John Wiley & Sons, LTD (1999)
7. Robinson, G.: Edge detection by compass gradient mask. Computer Graphics and Image Processing **6** (1977) 492–501