Estimating Illumination Direction from Textured Images

Manik Varma and Andrew Zisserman
Dept. of Engineering Science
University of Oxford, UK
{manik,az}@robots.ox.ac.uk

***Abstract*— We study the problem of estimating the illuminant's direction from images of textured surfaces. Given an isotropic, Gaussian random surface with constant albedo, Koenderink and Pont [JOSA 03] developed a theory for recovering the illuminant's azimuthal angle from a single image of the texture formed under a Lambertian model. In this paper, we extend the theory to deal with cases of spatially varying albedo.**

First, we generalise the theory to explain why their method should work even for certain types of spatially varying albedo. Our generalisation also predicts that the coherence of the structure tensor should lie below 0.8 in such non-constant albedo cases and accurately predicts the "deviation" from the true value observed by Koenderink and Pont on the Columbia-Utrecht (CUReT) texture database. Next, we extend the theory to account for arbitrarily varying albedo. We also investigate local, rather than global, estimates of the direction, and demonstrate our theory on the CUReT and the Heriot-Watt TextureLab databases where estimated directions are compared to ground truth.

I. Introduction

In this paper we address the problem of determining the illuminant's azimuthal angle from textured images. Traditionally, techniques from Shape from Shading have been used to estimate the direction of the light source [3], [15], [18]–[20], [22], [24]. Assuming a Lambertian [8] image formation model, most of these techniques try to simultaneously recover both shape, i.e. the surface height map or the surface normals, and the direction of the light source. However, this is an ill posed problem and so many constraints have to be imposed in order to find a reasonable solution. Some of the most common constraints are that the albedo is constant and that the surface is smooth or the normals integrable. Alternatively, other methods focus on local estimates or the occluding contour but, once again, have to impose very similar constraints to determine the illuminant's direction.

Recently, methods have been developed which specifically exploit the statistical nature of rough textures. Chantler *et al.* [4], [5] have shown that the variance of filter responses obtained from a textured image lie on Lissajous' ellipses as a function of the illuminant's azimuthal angle. Given three reference images of the texture, taken under fixed viewpoint and illuminant elevation, it is possible to determine the ellipse. This ellipse can then be used to read off the illuminant's azimuthal angle for any novel image of that texture. Koenderink and Pont [12] assume that the texture has constant albedo and that its surface has shallow relief, is isotropic and has a Gaussian random distribution. They develop a theory based on second order statistics to recover the illuminant's azimuth from a single image viewed orthographically under the Lambertian model.

Such statistical methods which recover the illuminant's direction from textured images are useful and have many applications. For instance, they can be used in Texture Analysis to provide information about the imaging conditions and thereby improve classification. Knowledge of the light source's direction can also help in Computer Graphics when introducing characters and objects with realistic shadows into an image. However, the applicability of most illumination estimation algorithms is restricted by the fact that they require the surface to have constant albedo. In this paper, we take a first step towards tackling this problem. In particular, we extend the method of Koenderink and Pont [12] to cases where the albedo is spatially varying.

The organisation of the rest of the paper is as follows: In section II we generalise the theory of Koenderink and Pont for the case where the albedo is isotropic and can be modelled as a random variable drawn from a log-normal distribution. In this case, the eigenvectors of the *structure tensor* $\mathbf{S} = < (\nabla \log I)(\nabla \log I)^T >$ turn out to be identical to those found by [12]. However, the coherence has a very different form and now, rather than being just a constant, becomes a function of the illuminant's elevation as well as the texture's albedo. We also examine how the coherence behaves in the presence of shadows and in section III verify our theory experimentally on the Columbia-Utrecht (CUReT) [6] database. Next, in section IV we explore how the theory can be further generalised to take into account arbitrarily varying albedo if extra information is present in the form of an additional reference image. The theory is then tested in section V on the Heriot-Watt TextureLab database, where additional reference images are available, and it is demonstrated that superior results are achieved with the new formulation. In section VI we investigate the advantages of using local regions, rather than the entire image, to form estimates of the azimuthal angle. Finally, we conclude in section VII with a discussion on the implications of our theory for resolving the Generalised Bas-Relief ambiguity.

II. Estimating the Light Source Azimuth

This section develops the basic theory for recovering the illuminant's azimuth from a single texture image. We consider the case where the underlying texture surface can be modelled as a Gaussian, random, rough surface. None of the parameters of the surface, the mean, variance or even the auto-correlation

function, need actually be known. Thus no knowledge of the geometry is required. Instead by making general assumptions about the surface height distribution, the second order statistics of the surface derivatives can be used to robustly recover the light source azimuth. The derivation will follow principally along the lines of [12].

Under the basic assumptions that the underlying model which produced the textured image has (a) an isotropic, Gaussian random rough surface with shallow relief viewed orthographically, (b) an albedo which is also isotropic but distributed log-normally, (c) an illuminant whose elevation ν is high as compared to the surface tangent plane, and (d) a perfect Lambertian image formation model without shadowing, specularities or inter-reflections, it will be shown that the illuminant's azimuthal angle ψ can be recovered from the largest eigenvector of the structure tensor $\mathbf{S}$.

As these assumptions might appear to be overly restrictive, it will be demonstrated that the theory holds even for cases when the textures deviate strongly from this model. For example, the results are empirically valid for elevations as small as $\nu = 5^\circ$ and when there are significant shadows. We will explain why this might be the case by considering the situation where the effects of shadowing, specularities, inter-reflections etc. can be incorporated into the albedo map.

If a textured surface is imaged under the Lambertian model [8], then the image intensities are independent of the viewing direction and depend on only the angle between the surface normal at each point and the light source direction. When there is a single, collimated, parallel light source, relatively high enough from the surface tangent plane so that shadows can be neglected, the image intensities are given by

$$I(x,y) = \frac{\rho(x,y) L_\lambda \sin\nu}{\sqrt{1+h_x^2+h_y^2}}[1-\cot\nu(h_x\cos\psi + h_y\sin\psi)] \quad (1)$$

where $\mathbf{L} = L_\lambda[\cos\nu\cos\psi, \cos\nu\sin\psi, \sin\nu]$ is the light source vector with elevation ν and azimuthal angle ψ, $h(x,y)$ is the Monge patch parameterisation of the surface height with partial derivatives $h_x(x,y)$ and $h_y(x,y)$, and $\rho(x,y)$ is the spatially varying surface albedo. Thus we are only considering a very simple image formation model and neglecting effects due to specularities, inter-reflections and shadows. Yet, as will be demonstrated, even this simple analysis can give very good results on real world datasets.

If the surface has shallow relief then the factor in the denominator can be ignored as $h_x, h_y \ll 1$. Following [12], [14], we work with the log intensity distribution given by

$$\log I(x,y) = \log(\rho L_\lambda \sin\nu) - \cot\nu(h_x\cos\psi + h_y\sin\psi) \quad (2)$$

where we have used the fact that $\cot\nu$ is small to form the truncated Taylor series expansion $\log(1-x) = -x$. Denoting $LI = \log I$, $s = \sin\psi$, $c = \cos\psi$ and taking partial derivatives gives

$$LI_x(x,y) = \frac{\rho_x}{\rho} - \cot\nu(ch_{xx} + sh_{xy}) \quad (3)$$

$$LI_y(x,y) = \frac{\rho_y}{\rho} - \cot\nu(ch_{xy} + sh_{yy}) \quad (4)$$

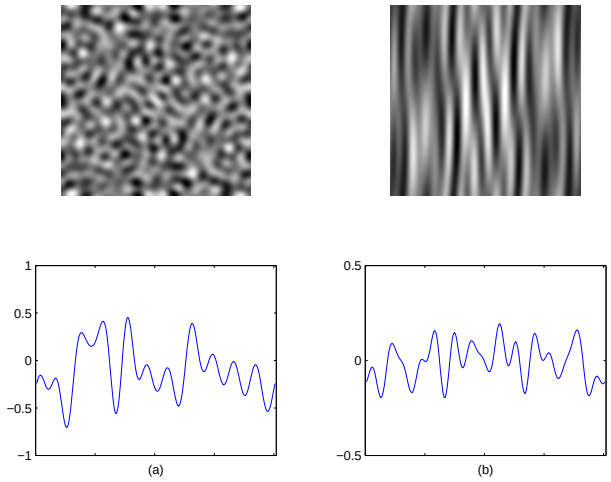

Fig. 1. Gaussian random rough surfaces and a corresponding 1D horizontal slice as generated by (5): (a) isotropic and (b) anisotropic.

Generic information about the surface height and albedo distributions is now needed in order to proceed further with the analysis. Many Lambertian rough surfaces in the real world can be described by a Gaussian height distribution with a given auto-correlation function. In [2], [17], it is shown that a Gaussian random rough surface can be generated by the interaction of a number of waves at different frequencies and orientations. Thus,

$$h(x,y) = \sum_n \sum_m h_{nm}\cos(nx + my) \quad (5)$$

where $n, m \in \mathbb{Z}$ and h_{nm} are random variables which determine the auto-correlation function. Figure 1 shows some sample Gaussian random rough surfaces which can be expressed as (5). Since a Gaussian rough surface must have an equal number of protrusions and indentations, the first order statistics will not reveal any information about the illuminant's azimuth (as the bright image regions will cancel out the dark image regions). Mathematically, $< LI_x >$ and $< LI_y >$ should vanish as the expected values of all partial derivatives of h must equal zero. Hence we turn to the square terms $< LI_x^2 >$, $< LI_y^2 >$ and $< LI_x LI_y >$ which become

$$\begin{aligned} < LI_x^2 > \;=\; & < (\rho_x/\rho)^2 > \\ & + \cot^2\nu < (ch_{xx} + sh_{xy})^2 > \\ & - 2\cot\nu(c < \rho_x h_{xx}/\rho > + s < \rho_x h_{xy}/\rho >) \end{aligned} \quad (6)$$

To account for the albedo, a similar kind of assumption is made about its distribution. If the albedo can be modelled as a random variable with a log-normal distribution [7], then $\log\rho$ should also be of the form (5) and therefore the third term in (6) must vanish as the product of any odd and even numbered derivatives has zero expected value. Denoting, $A_x = < (\rho_x/\rho)^2 >$ we then have,

$$< LI_x^2 > \;=\; A_x + \cot^2\nu < (ch_{xx} + sh_{xy})^2 > \quad (7)$$

$$< LI_y^2 > \;=\; A_y + \cot^2\nu < (ch_{xy} + sh_{yy})^2 > \quad (8)$$

where $< LI_y^2 >$ has been obtained by a similar treatment. The expression for $< LI_x LI_y >$ is also very similar

$$< LI_x LI_y >= A_{xy} + \cot^2\nu < (ch_{xx} + sh_{xy})(ch_{xy} + sh_{yy}) >$$

where $A_{xy} =< \rho_x \rho_y / \rho^2 >$.

We now need to evaluate the expectations of the height derivatives. Some straight forward trigonometry and integration yields

$$\begin{aligned}
< h_{xx}^2 > &= (1/2)\sum_n \sum_m n^4 h_{nm}^2 \\
< h_{yy}^2 > &= (1/2)\sum_n \sum_m m^4 h_{nm}^2 \\
< h_{xy}^2 > &= (1/2)\sum_n \sum_m n^2 m^2 h_{nm}^2 =< h_{xx} h_{yy} > \\
< h_{xx} h_{xy} > &= (1/2)\sum_n \sum_m n^3 m h_{nm}^2 \\
< h_{yy} h_{xy} > &= (1/2)\sum_n \sum_m n m^3 h_{nm}^2 \qquad (9)
\end{aligned}$$

At this point, there are more unknowns than equations and therefore the system must be constrained further for the light source azimuth to be recovered. One way of reducing the number of free variables is by constraining the underlying surface and albedo. In the case that both are isotropic, the expectations in (9) can be greatly simplified [2] to

$$\begin{pmatrix} < h_{xx}^2 > & < h_{xx} h_{yy} > & < h_{xx} h_{xy} > \\ < h_{yy} h_{xx} > & < h_{yy}^2 > & < h_{yy} h_{xy} > \\ < h_{xy} h_{xx} > & < h_{xy} h_{yy} > & < h_{xy}^2 > \end{pmatrix} = H \begin{pmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

while the albedo expectations simplify to

$$\begin{pmatrix} A_x & A_{xy} \\ A_{yx} & A_y \end{pmatrix} = A \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad (10)$$

where H and A are constants which depend on the surface height and albedo of the textured material (for instance, $A = 0$ for constant albedo textures). Substituting these values back into the expressions for $< LI_x^2 >$, $< LI_x^2 >$ and $< LI_x LI_y >$ gives

$$\begin{aligned}
< LI_x^2 > &= A + H\cot^2\nu(3\cos^2\psi + \sin^2\psi) \\
< LI_y^2 > &= A + H\cot^2\nu(\cos^2\psi + 3\sin^2\psi) \\
< LI_x LI_y > &= H\cot^2\nu(1+1)\sin\psi\cos\psi
\end{aligned}$$

There are now exactly three equations in three unknowns and therefore it is possible to recover the illuminant azimuth ψ from the eigenvectors of the structure tensor [12] defined as

$$\mathbf{S} = \begin{pmatrix} < LI_x^2 > & < LI_x LI_y > \\ < LI_x LI_y > & < LI_y^2 > \end{pmatrix} \qquad (11)$$

In the present case, the structure tensor turns out to have a very simple form

$$\mathbf{S} = A\mathbf{I} + H\cot^2\nu \begin{pmatrix} 2 + \cot 2\psi & \sin 2\psi \\ \sin 2\psi & 2 - \cot 2\psi \end{pmatrix} \qquad (12)$$

where $\mathbf{I}$ is the 2×2 identity matrix. The larger eigenvalue and corresponding eigenvector of the structure tensor are given by

$$\lambda_1 = A + 3H\cot^2\nu \Rightarrow \mathbf{v}_1 = \begin{bmatrix} \cos\psi \\ \sin\psi \end{bmatrix} \qquad (13)$$

while the smaller eigenvalue and eigenvector are given by

$$\lambda_2 = A + H\cot^2\nu \Rightarrow \mathbf{v}_2 = \begin{bmatrix} \cos(\psi + \pi/2) \\ \sin(\psi + \pi/2) \end{bmatrix} \qquad (14)$$

Thus, $\mathbf{v}_1$ points in the direction of the illuminant's azimuthal component and represents the desired solution. However, note that there is an ambiguity of 180 degrees in the recovered angle as $\mathbf{S}$ depends on 2ψ rather than ψ.

The *coherence* of the structure tensor $\mathbf{S}$ is defined to be

$$\begin{aligned}
\text{coh} &= \frac{\lambda_1^2 - \lambda_2^2}{\lambda_1^2 + \lambda_2^2} \qquad (15) \\
&= \frac{H\cot^2\nu(2A + 4H\cot^2\nu)}{A^2 + H\cot^2\nu(4A + 5H\cot^2\nu)} \qquad (16)
\end{aligned}$$

which must be less than or equal to 0.8. Thus, the coherence depends upon both ν and A. For example, when A^2 is negligible as compared to the second term in the denominator, the expression for the coherence simplifies to

$$\text{coh} = \frac{2A + 4H\cot^2\nu}{4A + 5H\cot^2\nu} \qquad (17)$$

which varies between 0.5 and 0.8 depending on the elevation ν. Of course, if A^2 is not negligible then the coherence can be lower still.

Deviations from the perfect Lambertian model: The model up till now has been derived under the assumption of perfect Lambertian reflectance without any shadowing (see figure 2), specularities, inter-reflections etc. However, in general, it is not possible to distinguish these effects from albedo variations given just a single image (unless there is prior information available) [10], [13]. For example, it is not possible to tell apart dark regions due to shadows (either cast or attached) from dark regions due to low albedo from only one image. Therefore, it might be possible to model these effects as albedo variations, as long as the distribution remains roughly log-normal (which can accommodate a large number of low intensity shadow regions in the bulk of the distribution with the specularities fitting into the long tail). In such a situation, (13), (14) and (16) will still hold and the largest eigenvector will point in the direction of the azimuth. However, A will now become a function of both ν and ψ (as well as the camera position) and therefore, for a given textured material,

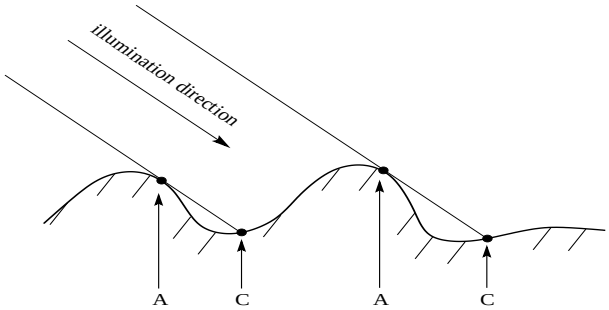

Fig. 2. A indicates an attached (self) shadow boundary, and C a cast shadow boundary. Significant cast shadows can suddenly appear below a certain elevation for rough surfaces.

the coherence can no longer be expected to be a monotonic function of the elevation. If we were to focus on shadowing as the major source of deviation from the the model, then depending on how quickly A increases with decreasing ν (i.e. the rate of shadowing), as compared to $H\cot^2\nu$, the coherence curve can either increase or decrease. It can also do both if the shadowing pattern changes after a certain elevation and one can expect kinks in the graph. Figure 3 plots some sample scenarios.

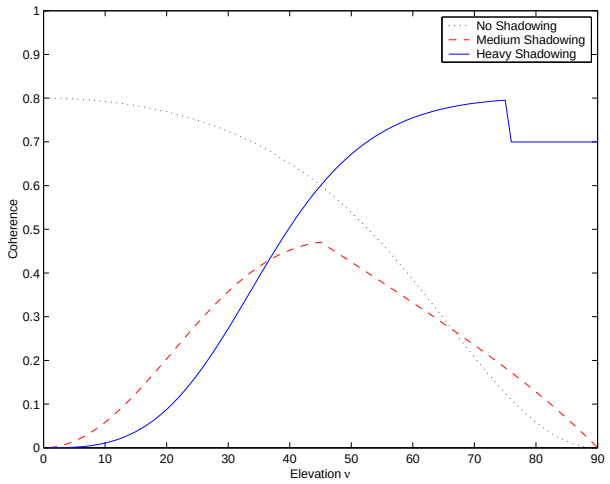

Fig. 3. The variation in coherence with elevation in the presence of shadows. The curve can change dramatically if the shadowing pattern changes after a certain elevation. This can also cause a jump in the curve, for example, when significant cast shadows suddenly appear below a certain elevation. It should be noted that these are just a few sample curves from the set of all such possible. Each can vary considerably depending on how shadowing influences the albedo parameter A.

III. Single Image Experiments and Comparisons

It is interesting to note that the eigenvectors recovered in (13) and (14) are identical to those found by Koenderink and Pont [12]. Thus, even though their theory was derived for a constant albedo map, their method should hold for a much wider range of textures. However, since a constant albedo map implies $A = 0$, their model expects that the measured coherence should always equal 0.8 and should not change with varying elevation, azimuth or texture sample. The theory derived here predicts otherwise. For almost constant albedos (i.e. small A), (17) expects the coherence to lie between 0.5 and 0.8 with lower values being expected for larger variations in albedo. These predictions match very well with the "deviations" from the ideal as measured in [12].

In the case of a true Gaussian random rough surface with painted white albedo, [12] report that the azimuth is estimated correctly within a few degrees but "the coherences are significantly lower" and vary between 0.4 and 0.7 with changing elevation. Similarly, on a sample texture from the Columbia-Utrecht (CUReT) database [6], the illuminants azimuth is detected to within a degree of the ground truth ($\psi = 0$) but the coherences are again found to be slightly lower with the 25 to 75 percentiles being 0.53 to 0.78.

It should be noted that while the current model has been derived by assuming Gaussian and log-normal distributions, it may also hold to some degree for other distributions for which the appropriate expected values cancel out. To determine how well the model copes with various materials with differing albedo and height distributions, we apply it to all the textures in the CUReT database. There are a total of 61 materials present in this database and each texture has been imaged under 205 different viewing and illumination conditions. Out of all the images available, we selected 92 images per material for which a big enough texture region could be extracted from the image. While the most extreme viewpoints are excluded, there are still many images for which the viewing direction is far from head on. Again, in order to verify the robustness of the model, we do not photometrically or geometrically calibrate the images but instead use the raw pixel intensities after they have been converted to grey scale.

Figure 4 shows the results of the algorithm on some CUReT textures. Only a few samples are shown for lack of space. For each texture, 7 images are chosen for which the viewing angle is almost in the direction of the surface plane normal (within $15°$). The value of the illuminant's azimuth is estimated using (13) and the estimation error in degrees is plotted as a function of ν in the middle row. The error is less than a few degrees even though the view is not perfectly normal, the albedo not constant and the surface not necessarily isotropic Gaussian. The results are valid even in the presence of shadows for the smaller values of elevation. In the bottom row, the associated coherences have also been plotted as a function of ν. As can be seen, they are not always equal to the constant value 0.8 but vary with ν and albedo as predicted by the theory.

Next, we apply the method to all $92 \times 61 = 5612$ images selected from the CUReT database and estimate the light source azimuth. As can be expected some results will not be very good due to the oblique viewpoint and the strong deviation of the textures from our assumptions. Nevertheless, in a majority of the cases, the azimuth is recovered to within a few degrees. Figure 5 is a plot of the estimation error versus the number of images having that error. Thus, for 1475 images the azimuth is estimated to within an accuracy of $1°$ while 3255 images (roughly 58% of those selected) have an error less than $5°$.

However, the algorithm does have a source of error which could be biasing these results. When a texture is strongly anisotropic, the perpendicular partial derivative dominates the structure tensor and forces the estimated illuminant to lie in its direction irrespective of the true azimuthal angle. For example, for a texture with translational symmetry, the iso-illumination contours are straight lines parallel to the translation direction and hence the derivatives in this direction will be negligible as compared to the perpendicular derivatives. So, for images which are vertically oriented (see figure 6), the x derivative becomes very large and forces the structure tensor to assume the form

$$\mathbf{S} = H \begin{pmatrix} 1 & \epsilon \\ \epsilon & \epsilon \end{pmatrix} \Rightarrow \lambda_1 = 1, \lambda_2 = 0, \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

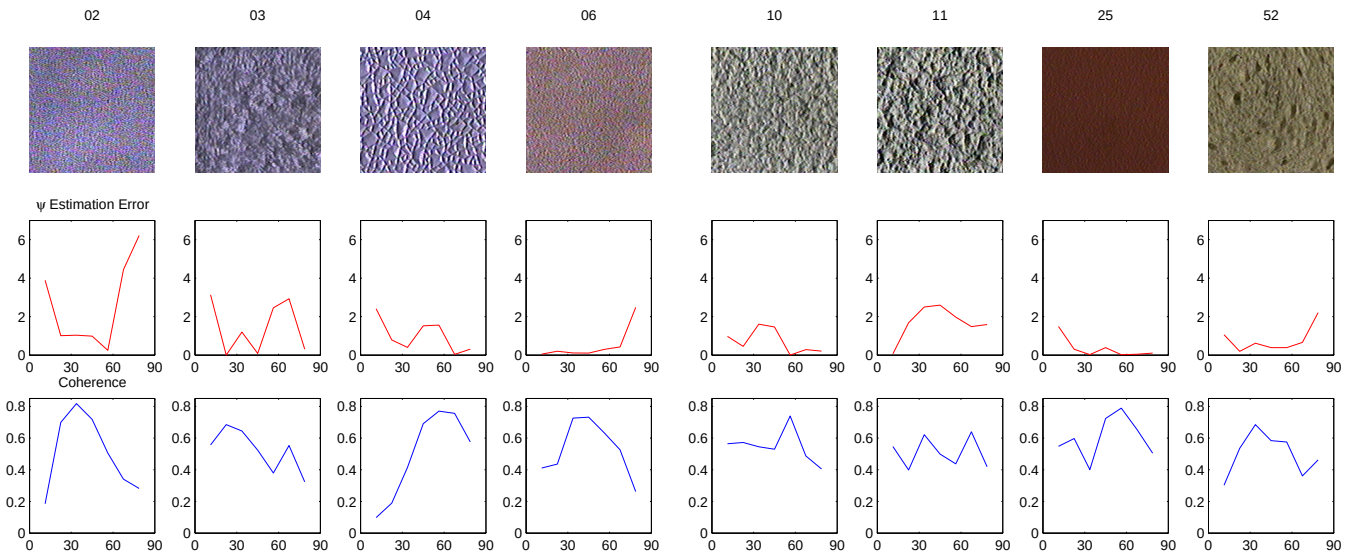

Fig. 4. The top row shows some sample CUReT textures from a fronto-parallel view (only a few samples are shown because of space limitations). The middle row plots, for each material, the error in estimating ψ (in degrees) as the illuminant's elevation varies from 11.25° to 78.75° (the viewing angle also varies but is always within 15° of the surface plane normal). Clearly, ψ is estimated to an accuracy of within a few degrees even when the albedo is not constant, the texture not necessarily isotropic Gaussian and even in the presence of strong shadows for smaller values of the elevation. The associated coherence values are plotted in the bottom row. As can be seen, the measured coherence varies with elevation and texture sample as predicted by the theory. It is not a constant equal to 0.8 in all cases as predicted by [12]. The jumps in the curves are most probably due to shadowing effects primarily with change in elevation but could also be due to the other effects with change in viewpoint (as the camera's azimuthal angle fluctuates between 0° and 180° from image to image). The samples are: Polyester (texture number 02), Terrycloth (03), Rough Plastic (04), Sandpaper (06), Plaster A (10), Plaster B (11), Quarry Tile (25), and White Bread (52). Note that for each sample, derivatives are computed at various scales and the best result reported. No photometric or geometric calibration has been done and all images are converted to grey scale.

and thus the estimated azimuth is 0° irrespective of the actual direction of the illuminant. A similar problem exists for horizontal textures and $\psi = 90^\circ$. And since most illuminant directions in the CUReT database are either $\psi = 0^\circ$, $\psi = 90^\circ$ or $\psi = 180^\circ$ it is difficult to tell whether the algorithm is working properly or giving erroneous results because of the dominance of oriented edges. However, in these cases the coherence will be *greater* than 0.8 and in fact will approach 1 and can therefore be used to flag errors. Figure 6 illustrates this effect. The algorithm seems to be working well as the estimated azimuth appears to lie very close to ground truth for $\psi = 0^\circ$. However, in reality, it is the orientation effects

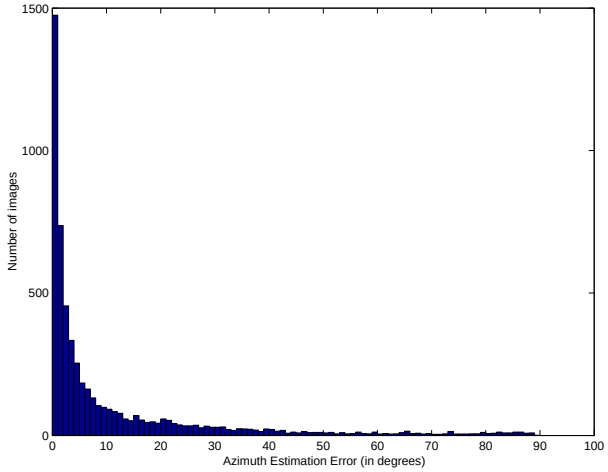

Fig. 5. A count of the azimuth estimation errors (in degrees) for all 5612 images in the CUReT database. Results are given for the best scale for computing derivatives.

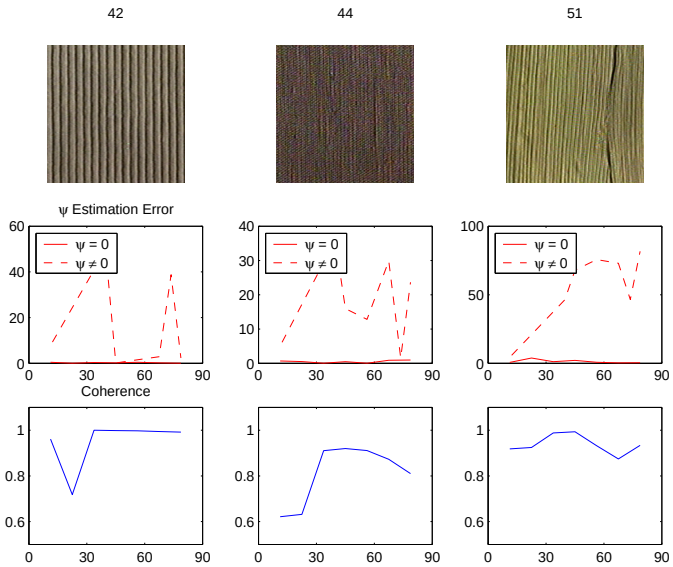

Fig. 6. The model can appear to be working well even though it is being fooled by orientation effects. As long as the illuminant's true azimuth is around 0° the algorithm returns good results (solid red curve in the graphs in the middle row) for Corduroy (42), Linen (44) and Corn Husk (51). However, the estimates for all other illuminant directions are very poor as can be seen by the dashed curve in the same graphs. The fact that the coherence is near 1 for the $\psi = 0^\circ$ curves can be used to flag this.

which are causing this and once the true illuminant direction moves away from 0° the errors become very large. The fact that the coherences are greater than 0.8 can be used to flag this occurrence.

Nevertheless, the model appears to be quite robust when its basic conditions are met. For example, for Plaster A (texture number 10) which appears to be isotropic, the azimuth was estimated to within 5° nearly 90% of the times, irrespective of viewpoint and shadowing. Thus, even though there is room for improvement, the simple model derived without taking into account many physical phenomenon still appears to work quite well.

IV. Estimation from Two Images

There are often cases when multiple images are available of a texture taken from the same viewpoint but with varying illumination. Photometric Stereo techniques rely on such data for example. In these cases, it is possible to use the extra information available to lift some of the restrictions imposed on the model in section II. In particular, it is possible to have freely varying albedo and, in this section, we develop a theory for estimating the illuminant's azimuth under such circumstances.

Suppose we have available two registered images I_1 and I_2 imaged by varying the illuminant's azimuth. Then, under the Lambertian model, the image intensities are given by

$$I_i(x,y) = \frac{\rho(x,y)L_\lambda \sin\nu}{\sqrt{1+h_x^2+h_y^2}}[1-\cot\nu(h_x\cos\psi_i + h_y\sin\psi_i)]$$

Note that by taking the ratio of the two images, it is possible to immediately get rid of both the albedo variation as well as the normalising constant in the denominator. Thus, we no longer have to make explicit the assumption that the surface has shallow relief in order to remove the $\sqrt{1+h_x^2+h_y^2}$ factor. Furthermore, the albedo can be allowed to vary arbitrarily as it has no influence on the ratio. Taking logarithms and again making use of the truncated Taylor series expansion gives

$$\begin{aligned} LR &= \log(\frac{I_1}{I_2}) \\ &= \cot\nu[h_x(\cos\psi_2-\cos\psi_1)+h_y(\sin\psi_2-\sin\psi_1)] \end{aligned}$$

Denote $C=\cos\psi_2-\cos\psi_1$ and $S=\sin\psi_2-\sin\psi_1$. Then

$$\begin{aligned} LR &= \cot\nu(Ch_x+Sh_y) \\ \Rightarrow\ LR_x &= \cot\nu(Ch_{xx}+Sh_{xy}) \\ \Rightarrow\ LR_y &= \cot\nu(Ch_{xy}+Sh_{yy}) \end{aligned} \tag{18}$$

Again, $< LR_x >$ and $< LR_y >$ are not expected to contain any information, so one must look at the second order statistics contained in $< LR_x^2 >$, $< LR_y^2 >$ and $< LR_xLR_y >$. If the surface is isotropic and Gaussian, then $< h_{xx}^2 >=< h_{yy}^2 >= 3H, < h_{xy}^2 >=< h_{xx}h_{yy} >= H$ while all other expectations are zero. Therefore,

$$\begin{aligned} < LR_x^2 > &= H\cot^2\nu(3C^2+S^2) \\ < LR_y^2 > &= H\cot^2\nu(C^2+3S^2) \\ < LR_xLI_y > &= H\cot^2\nu 2CS \end{aligned} \tag{19}$$

and the structure tensor is given by

$$\mathbf{S} = H\cot^2\nu \begin{pmatrix} 3C^2+S^2 & 2CS \\ 2CS & C^2+3S^2 \end{pmatrix} \tag{20}$$

Making use of the trigonometric identities $\cos(\psi_2 \pm \psi_1) = \cos\psi_2\cos\psi_1 \mp \sin\psi_2\sin\psi_1$, $\sin(\psi_2\pm\psi_1) = \sin\psi_2\cos\psi_1 \pm \cos\psi_2\sin\psi_1$ and performing some careful, but straight forward, algebra yields

$$\mathbf{S} = \alpha \begin{pmatrix} 2-\cos(\psi_1+\psi_2) & -\sin(\psi_1+\psi_2) \\ -\sin(\psi_1+\psi_2) & 2+\cos(\psi_1+\psi_2) \end{pmatrix} \tag{21}$$

where

$$\alpha = 4H\cot^2\nu\sin^2\left(\frac{\psi_2-\psi_1}{2}\right) \tag{22}$$

The eigenvalues of the structure tensor are now $\lambda_1 = 3\alpha$ and $\lambda_2 = \alpha$ while the larger eigenvector is

$$\mathbf{v}_1 = \begin{bmatrix} -\sin(\psi_1+\psi_2) \\ 1+\cos(\psi_1+\psi_2) \end{bmatrix} \tag{23}$$

from which it is possible to recover the joint angle $\psi_1+\psi_2$.

The coherence of the structure tensor now becomes

$$\text{coh} = \frac{\lambda_1^2-\lambda_2^2}{\lambda_1^2+\lambda_2^2} = 0.8 \tag{24}$$

V. Experimental Results for Two Images

We now assess the validity of the theory developed in section IV on the Heriot-Watt TextureLab database [21]. The database has 30 textures representing various kinds of materials: isotropic, oriented (in both surface and albedo), rough, etc. Figure 7 shows one image of each sample present in the database. Each material has been imaged under a fixed viewpoint. The illuminant's elevation is also fixed at $\nu = 45^\circ$ but the azimuth varies between $\psi = 0^\circ$ and $\psi = 315^\circ$.

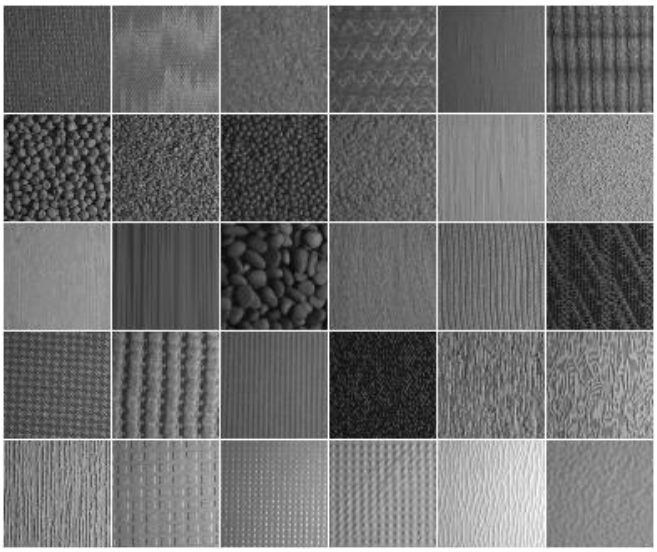

Fig. 7. Materials present in the Heriot-Watt TextureLab database. There are 30 textures and each has been imaged from a fixed viewpoint. The illuminant elevation is also fixed at $\nu = 45^\circ$ but the azimuth varies between $\psi = 0^\circ$ and $\psi = 315^\circ$.

To test the theory, we take samples from the database whose surface might be modelled as isotropic and Gaussian but for which the albedo varies considerably. For each sample, the image taken at $\psi = 0^\circ$ is retained as the reference image while (23) is then used to recover the azimuthal angle for all the rest. Figure 8 is a plot of the estimation error for four samples, AN4, TL2, TL3 and TL6, each of which has signification variation in its albedo. The middle row shows plots of the estimation error versus ψ for the remaining images. The solid blue curves represent the errors in the angle estimated using (23) and generally tend to be much lower than the dashed red curve representing the error in estimation due to (13). The bottom row is a plot of the associated coherences. Even though (24) predicts that the coherences should now equal 0.8 this is clearly not the case. The variation is most probably due to deviations from the model in terms of shadowing.

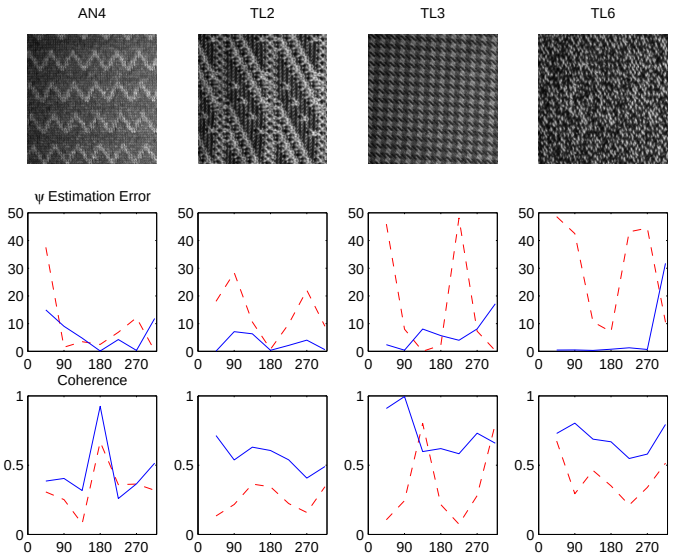

Fig. 8. Estimating the illuminant's azimuth for samples in the Heriot-Watt TextureLab database. For each material, the image at $\psi = 0^\circ$ is chosen as the reference image. The solid blue curves (middle row) then represent the error in estimating ψ in degrees for all the remaining images using (23). The dashed red curves represent the estimation error as measured using (13). The bottom row is a plot of the associated coherences. Note that for both methods, derivatives are computed at various scales and the best results reported.

VI. Local Estimation

Even though the methods developed in sections II and IV appear to cope fairly well with deviations in the model, there are often cases where a few bad measurements can adversely affect the recovery of the azimuthal angle. Therefore, it is desirable to estimate the illuminant's direction using local regions rather than the entire image.

As has been noted in section III, the presence of strong edges can bias the structure tensor and therefore these regions should be excluded while computing the expectations. Similarly, regions of constant intensity where the signal variation is very low should also be excluded.

There exist many such operators [9], [11], [16] to discard exactly such regions. Most of them are based around computing the second moment matrix which is extremely similar to the structure tensor $\mathbf{S}$. We use the Harris corner detector operator [11] to reject edge and constant intensity regions which might deviate from the assumed model and therefore give bad estimates. To estimate the statistics locally, we compute the most interesting Harris points and then for each point, use the region around it to calculate the expectations $< LI_x^2 >$, $< LI_y^2 >$ and $< LI_{xy} >$ and thereby the structure tensor. Thus at each chosen Harris point we compute the structure tensor and evaluate the local estimate of the illuminant direction. This can then be used to return the probability distribution of the azimuthal angle from which the mode can be chosen as the most likely estimate.

Preliminary experiments indicate favourable results. As discussed in section III the azimuth can be estimated to within an accuracy of a few degrees for most images of Plaster A in the CUReT database. This indicates that the texture satisfies the basic model. However, for a few images the estimation error is as high as 15° indicating that viewpoint and shadowing effects are causing deviations from the model and thereby contributing bad measurements. It is hoped that if these measurements can be excluded from the estimation process then we should be able to recover the azimuthal angle much more accurately. This is found to be exactly the case when the top 300 Harris points are used to choose the regions for computation. Figure 9 plots the probability distribution of the angles estimated using the Harris regions. The mode of the distribution is at 65° which is within 0.15° of the ground truth while using the entire image the recovered angle was $\psi = 49.61^\circ$ with an error of 15.49°.

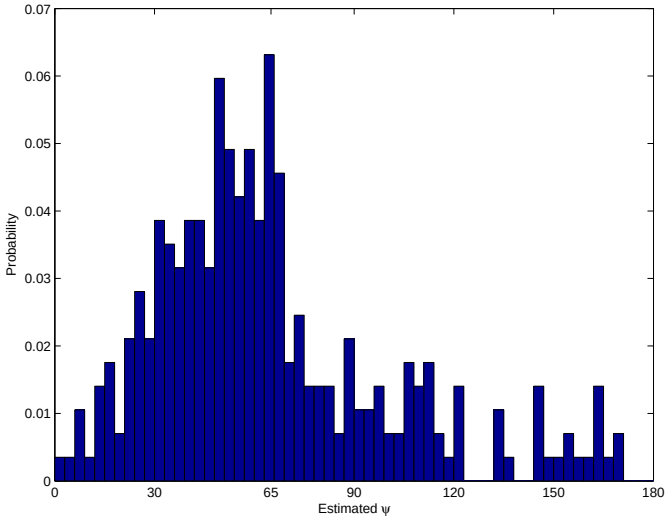

Fig. 9. Recovering the illuminant's azimuth using local estimates for an image of Plaster A from the CUReT database. The ground truth is $\psi = 65.10^\circ$ but the angle recovered using (13), which computes statistics over the entire image, is $\psi = 49.61^\circ$ due to shadowing and viewpoint deviations. By estimating the angle locally using Harris regions and rejecting others it is possible to improve the accuracy of the estimate as the mode of the distribution is 65°.

VII. Conclusions

In this paper, we have developed a theory for estimating the illuminant's azimuth for isotropic, Gaussian random textures with spatially variable albedo. When the albedo itself is isotropic and randomly distributed log-normally, then the solution for the illuminant's azimuth is identical to the one found by Koenderink and Pont [12]. However, the coherence of the structure tensor is no longer a constant but varies with both the elevation and the azimuth and is dependent on the texture's albedo and shadowing pattern. In the case that extra information is available in the form of a registered image with the same elevation, then it is possible to extend the theory to arbitrarily varying albedo as long as the surface itself is roughly isotropic Gaussian.

Being able to recover the illuminant's azimuth raises the interesting possibility of resolving parts of the Generalised Bas-Relief ambiguity (GBR) [1], [23]. Unfortunately, it turns out that once integrability has been enforced, the GBR does not affect the azimuthal angle of the light source but only its elevation and strength. However, the fact that we have imposed a Gaussian distribution on the height function does restrict the ambiguity. If the transformed surface is given by

$$\bar{h}(x,y) = \lambda h(x,y) + \mu x + \nu y + d \qquad (25)$$

then, in theory, both μ and ν must be zero and the ambiguity reduces to λ which affects the variance of the Gaussian, and the constant of integration in the surface reconstruction d which affects the mean. However, in practise, due to numerical reasons and because the Gaussian distribution is being approximated by a finite number of surface height points, it may well be the case that the ambiguity is not resolved to just λ and d but may also involve spurious values of μ and ν.

Acknowledgements

We are grateful to Mike Chantler for sharing the Heriot-Watt TextureLab database with us. Financial support was provided by a Rhodes Scholarship and the EC project CogViSys.

References

[1] P. Belhumeur, D. Kriegman, and A. L. Yuille. The bas-relief ambiguity. *IJCV*, 35(1):33–44, 1999.

[2] M. V. Berry and J. H. Hannay. Umbilic points on gaussian random surfaces. *J. Phys. A: Math. Gen.*, 10(11):1809–1821, 1977.

[3] M. J. Brooks and B. K. P. Horn. Shape and source from shading. In *Proc. Int. Joint Conf. Aritficial Intell.*, pages 932–936, 1985.

[4] M. J. Chantler, G. McGunnigle, A. Penirschke, and M. Petrou. Estimating lighting direction and classifying textures. In *Proc. BMVC.*, pages 737–746, 2002.

[5] M. J. Chantler, M. Schmidt, M. Petrou, and G. McGunnigle. The effect of illuminant rotation on texture filters: Lissajous's ellipses. In *Proc. ECCV*, pages 289–303, 2002.

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

[7] M Evans, N. Hastings, and B. Peacock. *Statistical Distributions*. Wiley-Interscience, 3rd edition, 2000.

[8] J. D. Foley, A. Van Dam, S. K. Feiner, and J. F. Hughes. *Computer Graphics: Principles and Practice*. Addison-Wesley, 1990.

[9] W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and center of circular features. In *Proc. of ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland*, pages 281–305, June 2-4 1987.

[10] D. Forsyth and A. Zisserman. Reflections on shading. *IEEE PAMI*, 13(7):671–679, 1991.

[11] C. J. Harris and M. Stephens. A combined corner and edge detector. In *Proc. Alvey Vision Conf.*, pages 147–151, 1988.

[12] J. J. Koenderink and S. C. Pont. Irradiation direction from texture. *J. of the Optical Society of America*, 20(10):1875–1882, 2003.

[13] J. J. Koenderink and A. J. Van Doorn. Geometrical modes as a general method to treat diffuse interreflections in radiometry. *J. of the Optical Society of America*, 73:843–850, 1983.

[14] J. J. Koenderink and A. J. van Doorn. Image processing done right. In *Proc. ECCV*, volume 1, pages 158–172. Springer-Verlag, 2002.

[15] C. H. Lee and A. Rosenfeld. Improved methods of estimating shape from shading using the light source coordinate system. *Artificial Intelligence*, 26:125–143, 1985.

[16] T. Lindeberg and J. Gårding. Shape-adapted smoothing in estimation of 3-d depth cues from affine distortions of local 2-d brightness structure. In *Proc. ECCV*, LNCS 800, pages 389–400, 1994.

[17] M. S. Longuet-Higgins. The statistical analysis of a random, moving surface. *Phil. Trans. R. Soc. Lond. A*, 249(966):321–387, 1957.

[18] P. Nillius and J-O. Eklundh. Automatic estimation of the projected light source direction. In *Proc. CVPR*, volume 1, pages 1076–1083, 2001.

[19] A. P. Pentland. Finding the illuminant direction. *J. of the Optical Society of America*, 72:448–455, 1982.

[20] E. V. Vega and Y-H. Yang. Default shape theory: With the application to the computation of the direction of the light source. *J. of the Optical Society of America*, 60:285–299, 1994.

[21] J. Wu and M. J. Chantler. Combining gradient and albedo data for rotation invariant classification of 3d surface texture. In *Proc. ICCV*, pages 848–855, 2003.

[22] Y. Yang and A Yuille. Sources from shading. In *Proc. CVPR*, pages 534–539, 1991.

[23] A. L. Yuille, D. Snow, R. Epstein, and P. Belhumeur. Determining generative models for objects under varying illumination: Shape and albedo from multiple images using SVD and integrability. *IJCV*, 35(3):203–222, 1999.

[24] Q. Zheng and R. Chellappa. Estimation of illuminant direction, albedo and shape from shading. *IEEE PAMI*, 13(7):680–702, 1991.