Drbohlav, O. and Chantler, M.: How do joint image statistics change with illumination?
Technical Report HW-MACS-TR-0020, School of Mathematical and Computer Sciences,
Heriot-Watt University, Edinburgh, UK. August 11, 2004.

How do joint image statistics change with illumination?

Ondrej Drbohlav and Mike Chantler
Texture Lab, School of Mathematical and Computer Sciences
Heriot-Watt University
Edinburgh, UK
email: {O.Drbohlav,M.J.Chantler}@hw.ac.uk

Abstract
The dependence of statistical properties of an image as a function of illumination direction is an exciting topic which was so far investigated experimentally, and also addressed theoretically for a special case of low-order moments of image features. In this paper we observe the principal relationship between the *joint* (co-occurrence) distribution of surface properties and the corresponding joint distribution of local image features. We focus on two kinds of statistics computed from local image neighbourhoods: (a) the joint distribution of pixel intensities, and (b) the joint distribution of binary patterns obtained by taking the signs of intensity differences between a selected reference pixel and all other pixels. We work with a non-parametric histogram representation of the probability distribution functions, and show how the frequencies in a histogram bin depend on illumination. Finally, we focus on approximating the illumination dependence of image statistics by a harmonic series. Experimental results obtained using a real surface are presented.

Keywords
Illumination, 3D Texture, Illumination models, Object Recognition

1 Introduction

It is well known that the image statistics are strongly affected by the illumination conditions which are used for taking an image. In face recognition, for example, illumination dependence is the source of deep intra-class image variability which is often much higher than the inter-class variability implied by a change in identity of a person observed [6]. In the field of 3D texture analysis, the illumination direction was shown to play a crucial role as well [4]. As many recognition algorithms across computer vision are based on probabilistic approaches, it is extremely important to understand how the image statistics are dependent on illumination.

Such topic was approached both from empirical and theoretical directions: Ginneken et al [9] presented an empirical approach, providing the experimental evidence of dependence of image intensity histograms on illumination. Most interestingly, the paper also divides 3D textures into several categories and highlights the difference in histogram behaviour between them. A surface model-based approach of Chantler et al. [5] opens the way to understanding how texture surface statistics implies the statistics of an image under given illumination. The paper shows that the variance of a histogram is a first order harmonic of illuminant tilt, under assumptions that the surface is of gently varying height and of uniform reflectance. Barsky [2] relaxes the assumptions and shows that such dependence can be described by harmonic expansion of second order in a general situation.

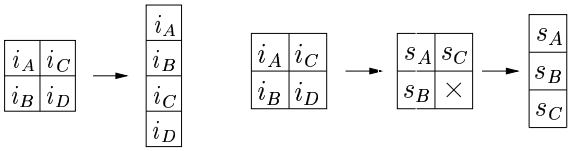
This paper delivers the following contributions to the topic: (i) unlike all previous work, we observe the formation of *joint* (co-occurrence) image statistics, i.e. not statistics of individual pixels, but statistics of local image neighbourhoods, and (ii) we explain the illumination dependence of individual multidimensional histogram bins, as opposed to predicting the behavior of low-order histogram moments (variance, skew, curtosis, etc.).

The paper is organised as follows. In Section 2 we present the local image features selected to be investigated, and explain why we consider the focus on them to be of importance. In Section 2.1 we show that given the illumination conditions, the process of image statistics formation is essentially related to integrating over portions of the surface statistics space. In Section 2.2 we show how a histogram bin population depends on illumination. In Section 3 we show that to capture the dependence of image statistics on illumination tilt *exactly* generally requires many parameters. However, we show that a lower-dimensional harmonic approximation of such dependence is in hand which reduces the number of parameters greatly, and we observe the feasibility of the approximation based on real experiments. In Section 4 we bring the discussion of the results and identify future directions of research.

2 Local image statistics

In this paper, we aim to gain insight into the dependence of image statistics on illumination. The first question is what features will be considered. For the purpose of this study, two types of local features were selected:

(a) **Intensities:** A 'trivial' feature obtained by reshaping the intensities in a local image neighbourhood into



(a) (b)

Figure 1. Local image features considered is this paper: (a) the raw image intensities in a local image neighbourhood, reshaped into a single feature vector; (b) the signs of differences $s_k = \text{sign}^{\pm}(i_\times - i_k)$ between a *reference* pixel intensity (here pixel D) and the other pixel intensities.

a single feature vector (see Fig. 1(a)). Besides that this feature is a 'natural' choice for this introductory study, the motivation also comes from that this feature has been recently used for texture recognition, with success rate outperforming the filter-based approaches [11].

(b) **Signs of intensity differences:** Intensity $i_\times$ of a selected reference pixel is compared with intensities i_k of all other pixels in a neighbourhood. If the reference pixel intensity is higher or equal to k-th pixel intensity, '$+1$' is recorded for the k-th pixel; and '-1' otherwise. For example, if the pixel neighbourhood is 2×2 then the number of recorded signs of differences is 3 (see Fig. 1(b)). The k-th feature component s_k can thus be written as

$$s_k = \text{sign}^{\pm}(i_\times - i_k)\,, \tag{1}$$

where the sign '$\pm$' indicates the modification of a standard signum operator: the case $i_\times = i_k$ is merged with the case $i_\times > i_k$, making the operator $\text{sign}^{\pm}$ binary.

Such feature is invariant, in particular, to any monotonic transformations in the pixel neighbourhood. It makes the recognition robust to variations in local surface illumination, indeed at the price of loosing discriminability. In texture recognition, it was proposed and applied with great success by Ojala et al. [8]. It is closely related to ordinal measures successfully applied for establishing image correspondences [3].

Having the features selected, the second question is what feature statistics are to be studied. The *joint* image statistics, as opposed to *marginal* image statistics, are considered to be the key to successful texture recognition [10]. This is why we study the full (joint) distributions of feature vectors. For representation of feature distributions, multi-dimensional histograms are used.

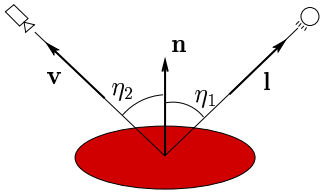


Figure 2. The illumination geometry

2.1 Statistics and illumination

The goal of this section is to analyse the way how the object surface statistics, together with given illumination conditions, imply the image statistics. To do so, a link between the surface and the image (the model of image formation) is needed. To keep the subsequent analysis simple, we adopt the following assumptions: (i) the model of reflectance is Lambertian, and (ii) inter-reflections and shadows are neglected. While all illustrations are done using these assumptions, it is not difficult to generalise for a different model of image formation. We make remarks on this topic at relevant places of the paper.

Let the light source be of direction $\mathbf{l}$ and of intensity σ, and let the pixel whose intensity is recorded be of normal orientation $\mathbf{n}$ and of albedo[1] ρ. Let the angle between the light vector and the normal vector be η_1 (see Fig. 2). Then the intensity i observed is [7]

$$i = \rho\sigma \cos\eta_1 = (\rho\mathbf{n})^\top(\sigma\mathbf{l}) = \mathbf{b}^\top\mathbf{s}\,, \tag{2}$$

where $\top$ mark stands for vector transposition. The right-hand side of this equation combines the surface albedo and normal direction into a single vector $\mathbf{b}$ termed the *scaled normal vector*, and the light intensity and direction into vector $\mathbf{s}$ termed the *scaled light vector*. These two entities enable to express the Lambertian model of reflectance in especially simple form.

Let us now start with an elementary example showing how single pixel intensity distribution can be computed. Let the scaled light vector $\mathbf{s}$ as well as the distribution of scaled normal vectors $p(\mathbf{b})$ be given. Then, due to (2), image intensity distribution $p(i)$ can be computed easily: As the scaled normals producing a given intensity i are those whose dot product with the light vector $\mathbf{s}$ is the same, one just needs to integrate over the a plane perpendicular to the light direction. More precisely, there holds

$$p(i) = \int\limits_{\{\mathbf{b}:\,\mathbf{b}^\top\mathbf{s}=i\}} \frac{p(\mathbf{b})}{\|\mathbf{s}\|}\,\mathrm{d}A_{\mathbf{s}}\,, \tag{3}$$

where $A_{\mathbf{s}}$ denotes the surface element of an integration plane perpendicular to $\mathbf{s}$. This can be shown as follows. First, both distributions must sum to unity, meaning $\int_{\mathrm{R}} p(i)\mathrm{d}i = \int_{\mathrm{R}^3} p(\mathbf{b})\mathrm{d}\mathbf{b}$. Choosing the scaled normals

[1] Albedo is a reflectance parameter which states what portion of incident irradiance is emited back into space in the form of diffuse reflection.

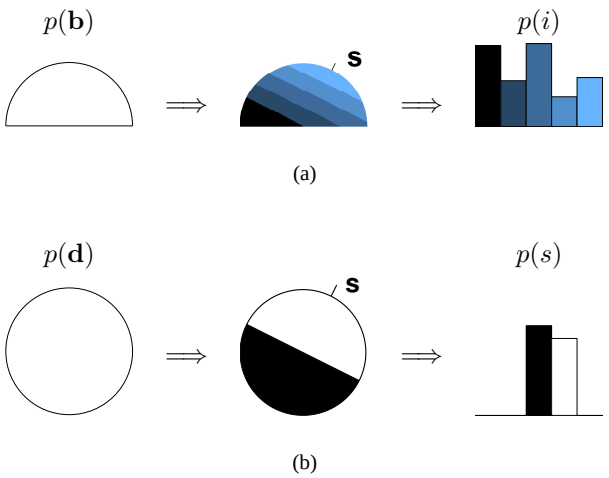


Figure 3. Given the scaled light vector $\mathbf{s}$ and local surface statistics, the feature histograms can be computed: (a) histogram of image intensities, given by integrating respective slabs of distribution in scaled normals space. The slabs are perpendicular to $\mathbf{s}$; (b) histogram of signs of co-occurring intensities, computed by integrating the distribution $p(\mathbf{d})$ of co-occurring scaled normals differences over respective half-spaces.

space coordinates such that the third one coincides with the light source direction, there holds $\mathrm{d}i = \|\mathbf{s}\|\mathrm{d}b_3$. The result (3) then follows immediately.

Fig. 3(a) summarises the mechanism of intensity distribution formation. In the figure, intensity distribution is parametrised using a histogram. Given $\mathbf{s}$, each histogram bin range implies a slab in the scaled normals space over which the distribution is integrated. The bin shown in black in the figure corresponds to zero-intensity bin collecting normals which are shadowed. Note also that the distribution is depicted as a half-space because in real imaging just the *visible* normals (those not inclined away from the viewing direction) contribute to the intensity distribution.

Given a local neighbourhood larger than one pixel, the *joint* (co-occurrence) intensity distribution can be constructed from joint distribution of scaled surface normals in a very analogous way. For example, a two-pixel co-occurrence $p(i_A, i_B)$ can be computed using the joint distribution of scaled surface normals as

$$p(i_A, i_B) = \int\int_{\substack{\{\mathbf{b}_A \,:\, \mathbf{b}_A{}^\top \mathbf{s} = i_A\} \\ \{\mathbf{b}_B \,:\, \mathbf{b}_B{}^\top \mathbf{s} = i_B\}}} \frac{p(\mathbf{b}_A, \mathbf{b}_B)}{\|s\|^2} \mathrm{d}A^A_{\mathbf{s}} \mathrm{d}A^B_{\mathbf{s}} \,. \tag{4}$$

Having identified the intensity distribution formation, let us now turn our attention to the other local feature: the signs of intensity differences. Thus, let there be two pixel neighbourhood with pixels A and B as just above, and let pixel B be the reference one. To evaluate s_A let us insert

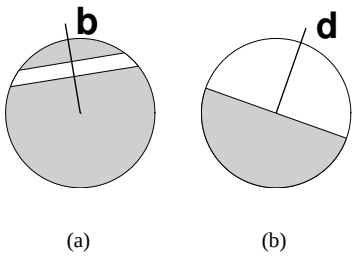


Figure 4. For a given histogram bin, each feature sample votes into the space of scaled light vectors: (a) scaled normal $\mathbf{b}$ votes for light sources giving intensity within the bin range; (b) the difference vector $\mathbf{d}$ votes for light sources giving signum '$+1$' in the case shown. For both cases, favourable scaled light vectors are indicated by white.

(2) into (1) to get

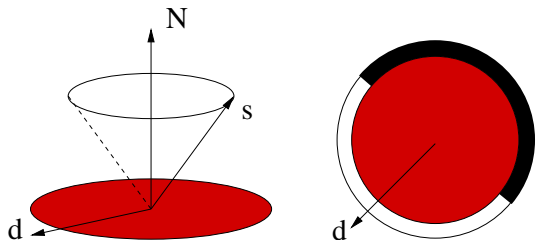
$$\begin{aligned} s_A &= \mathrm{sign}^{\pm}(i_B - i_A) = \mathrm{sign}^{\pm}[(\mathbf{b}_B - \mathbf{b}_A)^\top \mathbf{s}] = \\ &= \mathrm{sign}^{\pm}(\mathbf{d}_A{}^\top \mathbf{s}) \,, \end{aligned} \tag{5}$$

where $\mathbf{d}_A = \mathbf{b}_B - \mathbf{b}_A$ is termed *scaled normals difference vector*, or shortly a *difference vector*. Obviously, the mechanism of signs distribution formation is as illustrated in Fig. 3(b): The $+1$ bin (shown in white) integrates the scaled normal difference vector distribution over the half-space whose projection to $\mathbf{s}$ is positive; the -1 bin (shown in black) integrates the other half-space. Note that the figure reflects the fact that the domain of difference vectors distribution is the whole three-dimensional space. The formation of joint distributions of signs of intensity differences is done in a way analogous to the case of joint intensities distribution.

2.2 Bin population

The previous section answered the question of how the local image statistics are generated given the surface statistics and the light source direction. In this section, we come closer to the very goal of this paper: Given again the surface properties distribution, how do the image statistics change with illumination? To proceed, we suppose that the image statistics is represented by (possibly multidimensional) histogram, and observe how the populations of individual bins depend on illumination direction.

The computation of bin population dependence on illumination is essentially a voting process. Let us first consider the case of single pixel intensity distribution. For a specific scaled normal sample $\mathbf{b}$, the set of scaled light vectors $\mathbf{s}$ producing the intensity within the bin range forms a slab (see Fig. 4(a)). The vote is represented by a membership function in the scaled lights space which is 1 where the scaled light is favourable to the bin and 0 where it is not. The votes are summed across all samples. Formally, if $p(\mathbf{b})$ is the scaled normals distribution and $v(\mathbf{b}, l, \mathbf{s})$ is the



(a) (b)

Figure 5. Illustrating the conditions of the experiment: (a) The light source moves along a cone whose axis is the macroscopic normal $\mathbf{N}$. As the surface is of gently varying depth and of uniform albedo, the difference vectors lie near the plane perpendicular to $\mathbf{N}$; (b) the voting space is a 1D illumination tilt space.

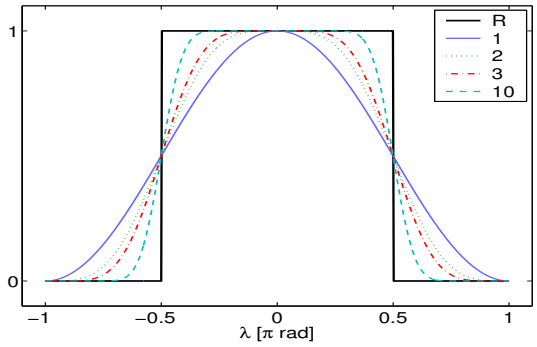


Figure 6. The rectangular voting function (R), and its harmonic approximations (first three, and tenth).

voting function for a scaled normal $\mathbf{b}$ and l-th histogram bin, the bin frequency p_l dependence on $\mathbf{s}$ is then

$$p_l(\mathbf{s}) = \int v(\mathbf{b}, l, \mathbf{s}) p(\mathbf{b}) \mathrm{d}\mathbf{b}\,. \qquad (6)$$

The voting process for the difference vectors is similar (see Fig. 4(b)). Observe that in this case, if light $\mathbf{s}$ is favourable to the bin, then $\alpha\mathbf{s}$, $\alpha > 0$ is favourable as well. Hence, in this case, the voting space can be represented by a suitable two-dimensional surface.

With joint statistics, the voting mechanism stays the same; briefly, for a two-pixel co-occurrence, for the bin frequency p_{lm} we can write $p_{lm}(\mathbf{s}) = \int v(\mathbf{b}_A, l, \mathbf{s}) v(\mathbf{b}_B, m, \mathbf{s}) p(\mathbf{b}_A, \mathbf{b}_B) \mathrm{d}\mathbf{b}_A \mathrm{d}\mathbf{b}_B$.

To give this section a summary, we have observed that evaluation of bin population as a function of illumination direction is a voting process and can be written in a form of (6). In the next section, we consider a special type of voting functions, and get a taste of how co-occurrence statistics change with illumination based on conducting a real experiment.

3 Voting as a convolution

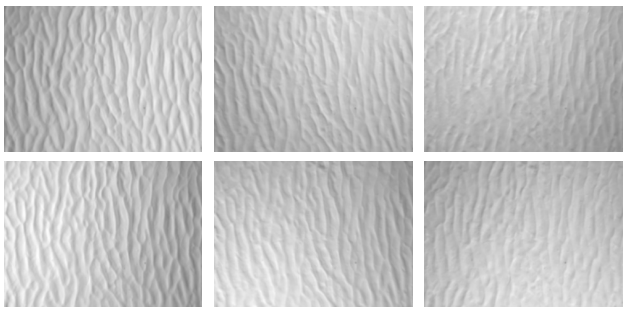
In this section, we focus on local features given by the signs of intensity differences. It has been noted in the previous section that the absolute magnitudes of difference vectors $\mathbf{d}$ and scaled light vectors $\mathbf{s}$ do not matter because the result of computing the sign by (1) is invariant under the transformation $\mathbf{s} \mapsto \alpha\mathbf{s}, \alpha > 0$ and $\mathbf{d} \mapsto \beta\mathbf{d}, \beta > 0$. This suggests to identify both the difference vectors and the scaled light vectors under these equivalence relations, and to use a suitable two-dimensional surface enclosing the origin for their representation; a sphere, for example. Under such identification, the distribution of difference vectors is a function on a sphere, and it is also the sphere what is the light directions voting space. Importantly enough, the voting functions for two non-equal difference vectors differ only by rotation because otherwise they both divide the space of light vectors into two complementary half-spaces. But that means that the voting process (6) is in this case a convolution of the voting kernel with the difference vectors distribution on a sphere.

To illustrate this, we lower the dimension of the problem by adopting the following assumption: The surface is of uniform albedo and gentle height variation. For such surface, the difference vectors computed on scaled normals in a local neighbourhood all lie near a plane perpendicular to the mean surface orientation (see Fig. 5(a)). Obviously, the light source is favourable to the event $s = +1$ provided that its projection onto the surface lies within the half-circle defined by the direction of $\mathbf{d}$ (see Fig. 5(b)), in other words, that the illumination tilt (the azimuth angle of the projection) lies within the nonzero part of the voting function. That means that the voting space can be collapsed into a circumference. The voting is a result of convolving the voting kernel (shown as a black thick solid line in Fig. 6) with the distribution of the difference vectors which now 'lives' on the circumference.

3.1 Harmonic expansion of the voting kernel

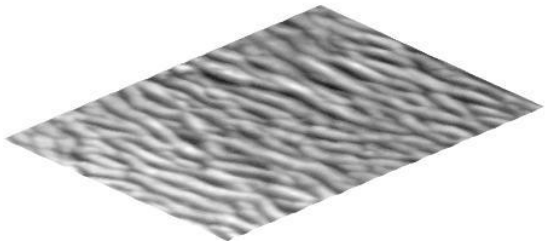
If the number of voting samples is q, in a general case the result of voting into the illumination space (the circumference) has $2q$ points of discontinuity, and therefore an exact representation of distribution illumination dependence is costly. The goal of this section is to demonstrate that the harmonic expansion of the voting kernel can reduce the number of parameters needed considerably. Approximating the voting kernel essentially corresponds to switching to 'fuzzy' voting when the voting result is not just 0 or 1 but there is a smooth transition in between.

To approximate the voting kernel (shown in black in Fig. 6) by harmonics, we avoid using the Fourier expansion. This is because the Fourier expansion solves the problem of least-squares optimal approximation of the kernel;



0°	60°	120°
180°	240°	300°

(a)



(b)

Figure 7. (a) A surface illuminated from six tilt directions specified. The illumination slant was kept at 45°; (b) the result of surface reconstruction by photometric stereo.

and such approximation has properties which are not desirable in this case: It has several local maxima and both overshoots and undershoots the values allowable for voting (1 and 0, respectively). Our aim is to make a harmonic approximation $f(\phi)$ which in addition to best least-square approximation of the kernel exhibits the following properties: (i) has a unique maximum 1 and a unique minimum 0, (ii) is symmetric around 0, (iii) f and $1-f$ are the same functions, only shifted by π. It is a short exercise to show that in that case the series is formed by odd orders of cosines. The first three approximations are:

$$
\begin{aligned}
v_1(\phi) &= \frac{1}{2}(1+\cos\phi) \qquad (7)\\
v_2(\phi) &= \frac{1}{2}\left(1+\frac{9}{8}\left(\cos\phi+\frac{1}{9}\cos 3\phi\right)\right)\\
v_3(\phi) &= \frac{1}{2}\left(1+\frac{75}{64}\left(\cos\phi+\frac{1}{6}\cos 3\phi+\frac{1}{50}\cos 5\phi\right)\right)
\end{aligned}
$$

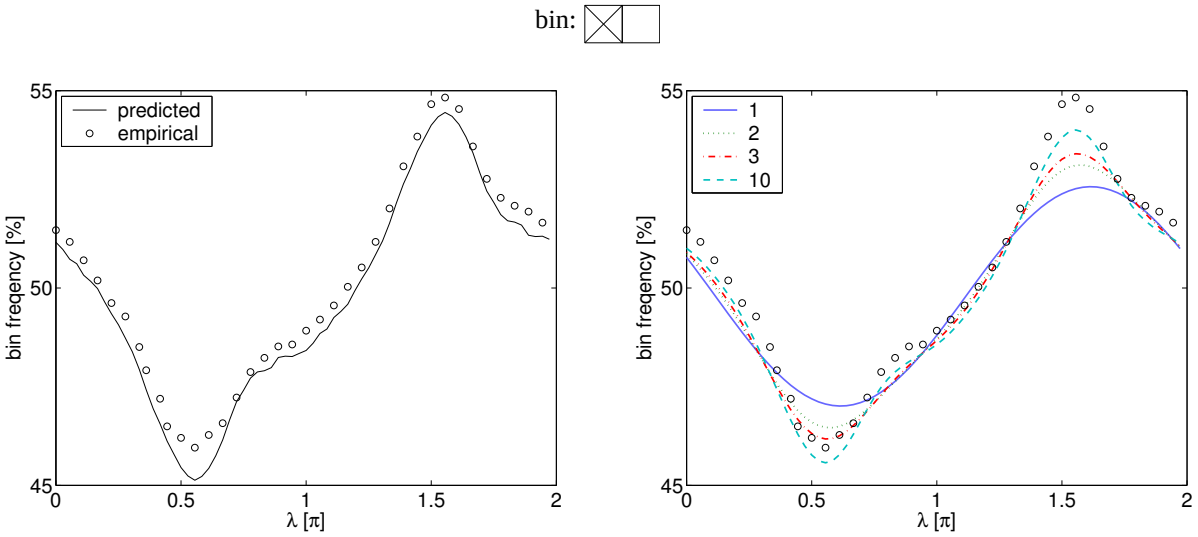
These three approximations, together with the tenth one (involving the cosine of 19th order) are shown in Fig. 6.

3.2 Experiment

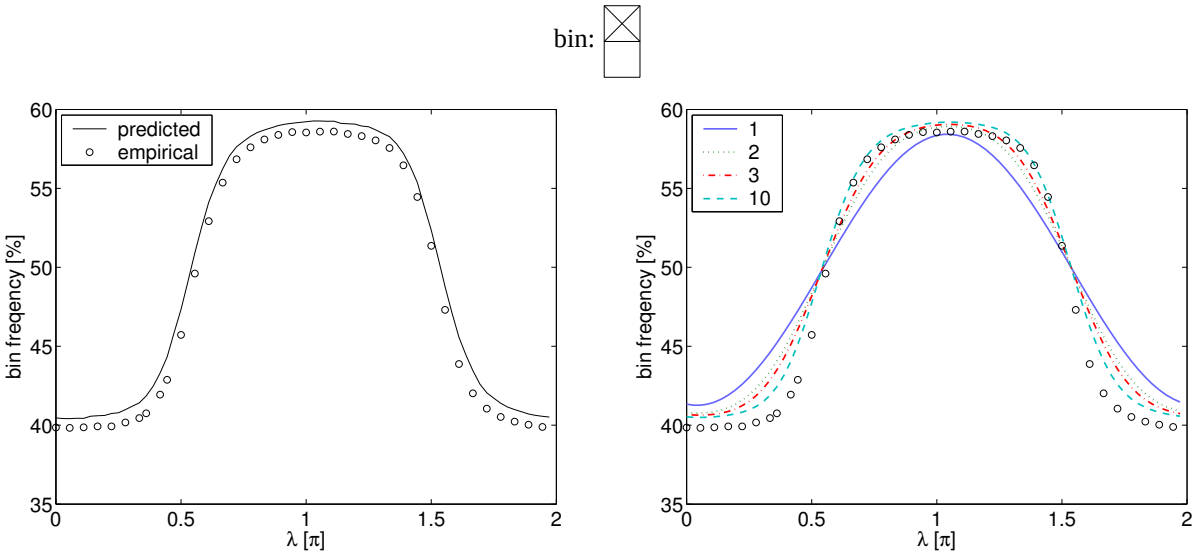
To test the analysis presented in the previous section we used a 3D texture ABA from the PhoTex database [1]. This is a plaster surface of uniform reflectance and of surface normal slant not exceeding 10°. This sample fulfills the working assumption described at the beginning of Section 3 with a sufficient degree of accuracy. The data used were as follows: 36 images of the surface taken with the light source of slant 45° and with tilt going from 0° to 350° in 10° steps. We used six of these images (see Fig. 7(a)) to get the surface normals at each pixel by photometric stereo method [12]. The result of the surface reconstruction is shown in Fig. 7(b).

We began by computing the distribution of signs of differences between a pixel and its eastern neighbour. As in this case the two bins' probabilities sum up to unity for each illumination, we evaluated just the dependence of event $s=+1$. This was done as follows. The normals estimated by photometric stereo were used for constructing the difference vectors statistics on a circle as suggested in the previous section. The difference vectors distribution was then convolved with the full (rectangular) voting kernel, such predicted dependence was compared with the $+1$ sign statistics computed directly on images acquired across varying tilt illumination. The result shown in Fig. 8(a)(left) presents a very good agreement of the predicted and empirical values of the histogram frequency. In Fig. 8(a)(right), the results of voting by kernel approximation is shown as well. The same procedure was applied to compute the dependence of frequency of $+1$ bin for a local neighbourhood formed by a pixel and its southern neighbour (see Fig. 8(b)).

Essentially analogous procedure has been applied to the eight-dimensional joint distribution of signs in a 3×3 local image neighbourhood (see Fig. 9). In this case, the total number of bins is $2^8=256$. To compute the voting result, the convolution could not be used any more, but a product of voting kernels was used as follows from the theory presented in Section 2.2. For presentation purposes, we sorted the bins according to their average frequency. Interestingly enough, around 50% of all data points is contained within the first four bins (observe the scales of frequency vertical axes). Note also that the bins come out from the sorting procedure in complementary couples (presented side by side in the figure): This is due to the symmetries in the texture surface properties. Predicting the bin population by voting with a full rectangular kernel again predicts the population quite well (see black solid lines in the figure). It is also notable that the harmonic approximations consistently underestimate the populations with decreasing order of approximation, which is caused by fuzzy voting. However, the *qualitative* shape of the first four bins is captured well even by the product of first harmonic approximations; this is no more true for the fifth and sixth bins which are, however, much less populated.



(a)



(b)

Figure 8. The dependence of bin populations on illumination tilt λ for signs of differences between pixels as shown above each of the graphs. The circular dots are empirical data and black solid lines is the result of predicting the bin population by voting with a full rectangular kernel. Other lines are approximations of the kernel as in Fig. 6.

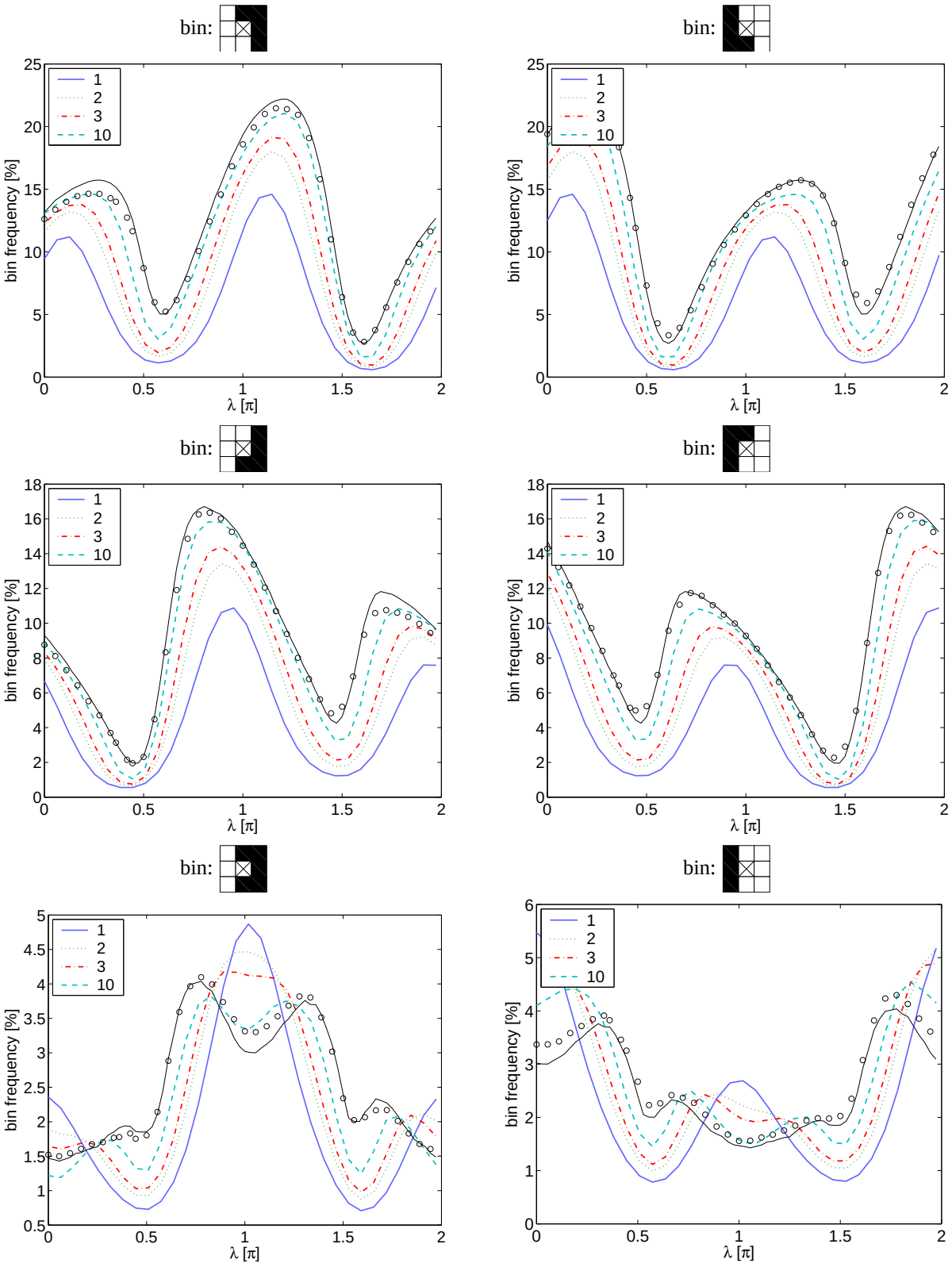


Figure 9. The dependence of bin populations on illumination tilt λ for signs of differences as shown above each of the graphs. The circular dots are empirical data and black solid lines is the result of predicting the bin population by voting with a full rectangular kernel. Other lines are approximations of the kernel as in Fig. 6.

4 Summary and Conclusion

The results obtained under assumptions used in Section 3 are quite encouraging and deserve several remarks. First, the assumption of uniform albedo, used quite often in physics-based analysis in computer vision, is usually quite restrictive. It is worthwhile to note that in this work, if the task is to characterise statistics of surfaces under different illumination then the uniform albedo surfaces are actually the hard case. This is because if the albedos in neighboring pixels differ greatly, the distribution of signs of intensity differences will tend to be invariant to illumination, as the ordering of pixels will be governed not by illumination conditions but by the albedo difference itself. Second, these results suggest exciting directions for future research. For the surface texture studied, the bin frequency variations showed up to be much smoother than they generally could be. This is of course because it was the underlying distribution of *difference vectors* which was smooth, and the voting result inherited this smoothness. While in this work we mainly outlined the principles governing the way from surface statistics to image statistics, in future work it will be interesting to identify *characteristic* surface statistics which often appear in reality, and to employ such priors in recognition tasks, for example. Likewise, a very important question is how the statistics behave in presence of strong shadowing, a phenomenon which was excluded in this introductory study.

References

[1] PhoTex database. Texture lab, Heriot-Watt University, Edinburgh, UK. Available on-line at `http://www.cee.hw.ac.uk/texturelab/database/photex/`.

[2] S. Barsky. *Surface Shape and Colour Reconstruction using Photometric Stereo*. PhD thesis, University of Surrey, October 2003.

[3] D. N. Bhat and S. K. Nayar. Ordinal measures for image correspondence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(4):415–423, 1998.

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

[5] M.J. Chantler, M. Schmidt, M. Petrou, and G. McGunnigle. The effect of illuminant rotation on texture filters: Lissajous's ellipses. In *Proc. European Conference on Computer Vision*, volume 3, pages 289–303, 2002.

[6] A. S. Georghiades, D. J. Kriegman, and P. N. Belhumeur. Illumination cones for recognition under variable lighting: Faces. In *Proc. of Computer Vision and Pattern Recognition*, pages 52–59, 1998.

[7] Johann H. Lambert. *Photometria sive de mensura et gradibus luminis, colorum et umbra*. Augustae Vindelicorum, Basel, 1760.

[8] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, July 2002.

[9] Bram van Ginneken, Jan J. Koenderink, and Kristin J. Dana. Texture histograms as a function of irradiation and viewing direction. *International Journal of Computer Vision*, 31(2/3):169–184, 1999.

[10] M. Varma and A. Zisserman. Classifying materials from images: to cluster or not to cluster? In *Texture2002 : The 2nd International Workshop on Texture Analysis and Synthesis*, pages 139–144, June 2002.

[11] M. Varma and A. Zisserman. Texture classification: are filter banks necessary? In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 691–8, 2003.

[12] R. J. Woodham, Y Iwahori, and R. A. Barman. Photometric stereo: Lambertian reflectance and light sources with unknown direction and strength. Technical Report 91-18, Laboratory for Computational Intelligence, Univ. of British Columbia, Vancouver, BC, Canada, August 1991.