Belief Propagation with Directional Statistics for solving the Shape-from-Shading problem

Tom S. F. Haines and Richard C. Wilson

The University of York,
Heslington, YO10 5DD, U.K.

Abstract. The Shape-from-Shading [SfS] problem infers shape from reflected light, collected using a camera at a single point in space only. Reflected light alone does not provide sufficient constraint and extra information is required; typically a smoothness assumption is made. A surface with Lambertian reflectance lit by a single infinitely distant light source is also typical.
We solve this typical SfS problem using belief propagation to marginalise a probabilistic model. The key novel step is in using a directional probability distribution, the Fisher-Bingham distribution. This produces a fast and relatively simple algorithm that does an effective job of both extracting details and being robust to noise. Quantitative comparisons with past algorithms are provided using both synthetic and real data.

1 Introduction

The classical problem of Shape-from-Shading [SfS] uses irradiance captured by a photo to calculate the shape of a scene. A known or inferred reflectance function provides the relationship between irradiance and surface orientation. Surface orientation may then be integrated to obtain a depth map. Horn[1] introduced this problem with the assumptions of Lambertian reflectance, orthographic projection, constant known albedo, a smooth surface, no surface inter-reflectance and a single infinitely distant light source in a known relation with the photo. This constrained scenario has been tackled many times since[2–7, to cite a few], and will again be the focus of this work.

Zhang et al.[8] surveyed the area in 1999, concluding that Lee and Kuo[4] was the then state of the art. Lee and Kuo iteratively linearised the reflectance map and solved the resulting linear equation using the multigrid method. More recent methods include Worthington and Hancock[5], which iterated between smoothing a normal map and correcting it to satisfy the reflectance information; Prados et al[6], which solved the problem with viscosity solutions; and Potetz[7] which used belief propagation. This last work by Potetz is particularly relevant due to it also using belief propagation, though in all further details it differs. Belief propagation estimates the marginals of a multivariate probability distribution, often represented by a graphical model. Potetz makes use of two variables per pixel, $\delta x/\delta z$ and $\delta y/\delta z$, and uses various factor nodes to provide the reflectance information, smoothness assumption and integrability constraint.

Whilst this model can be implemented simply with discrete belief propagation it would never converge and require a large number of labels, instead advanced continuous methods are used.

The following three sections, 2 through to 4, cover the component details, starting with the formulation, then belief propagation and finally directional statistics. Section 5 brings it all together into a cohesive whole, and is followed by section 6 which solves a specific problem. Following these sections we give results in section 7 and conclusions in the final section.

2 Formulation

Using previously given assumptions, of Lambertian reflectance, constant known albedo, orthographic projection, an infinitely distant light source and no inter-reflection the irradiance at each pixel in the input image is given by

$$I_{x,y} = A(\hat{\mathbf{l}} \cdot \hat{\mathbf{n}}_{x,y}) \tag{1}$$

where $I_{x,y}$ is the irradiance provided by the input image. A is the albedo and $\hat{\mathbf{l}} \in \mathbb{R}^3, |\hat{\mathbf{l}}| = 1$ is the direction to the infinitely distant light source; these are both provided by the user. $\hat{\mathbf{n}}_{x,y} \in \mathbb{R}^3, |\hat{\mathbf{n}}_{x,y}| = 1$ is the normal map to be inferred as the algorithm's output. The normal map can be integrated to obtain a depth map, a step with which we are not concerned. By substituting the dot product with the cosine of the angle between the two vectors you get

$$\frac{I_{x,y}}{A} = \cos\theta_{x,y} \tag{2}$$

where θ is therefore the angle of a cone around $\hat{\mathbf{l}}$ which the normal is constrained to[5]. This leaves one degree of freedom per pixel that is not constrained by the available information. A smoothness assumption provides the extra constraint.

Directional statistics is the field of statistics on directions, such as surface normals. Using a directional distribution allows the representation of surface orientation with a single variable, rather than the two used in Potetz[7] and many others. We propose a new SfS algorithm using such distributions within a belief propagation framework. This leads to a belief propagation formulation not dissimilar to Gaussian belief propagation[9] in its simplicity and speed.

3 Belief Propagation

Loopy sum-product belief propagation is a message passing algorithm for marginalising an equation of the form

$$P(\mathbf{x}) = \prod_{v \in V} \psi_v(\mathbf{y}_v) \tag{3}$$

where $\mathbf{x}$ is a set of random variables and $\forall v, \mathbf{y}_v \subset \mathbf{x}$. Such an equation can be represented by a graphical model where each variable is a node and nodes that

interact via ψ functions are linked. In this case the random variables are directions, represented by normalised vectors. Message passing then occurs within this model, with messages passed along the links between the nodes. As the variables are directions the messages are probability distributions on directions. The method uses belief propagation to obtain the maximum a posteriori estimate of a pairwise Markov random field where each node represents the orientation of the surface at a pixel in the image. The message passed from node p to node q at iteration t is

$$m_{p\to q}^{t}(\hat{\mathbf{x}}_q) = \int_{\hat{\mathbf{x}}_p} \psi_{pq}(\hat{\mathbf{x}}_p, \hat{\mathbf{x}}_q)\psi_p(\hat{\mathbf{x}}_p) \prod_{u\in(N\setminus q)} m_{u\to p}^{t-1}(\hat{\mathbf{x}}_p)\delta\hat{\mathbf{x}}_p \tag{4}$$

where $\psi_{pq}(\hat{\mathbf{x}}_p, \hat{\mathbf{x}}_q)$ is the compatibility between adjacent nodes, $\psi_p(\hat{\mathbf{x}}_p)$ is the prior on each node's orientation and N is the 4-way neighbourhood of each node. Once message passing has iterated sufficiently for convergence to occur the belief at each node is

$$b_p(\hat{\mathbf{x}}_p) = \psi_p(\hat{\mathbf{x}}_p) \prod_{u\in N} m_{u\to p}^{t-1}(\hat{\mathbf{x}}_p) \tag{5}$$

From $b_p(\hat{\mathbf{x}}_p)$ the most probable direction is selected as output.

4 Directional Statistics

The Fisher distribution, using proportionality rather than a normalising constant, is given by

$$P_F(\hat{\mathbf{x}}; \mathbf{u}) \propto \exp(\mathbf{u}^T\hat{\mathbf{x}}) \tag{6}$$

where $\hat{\mathbf{x}}, \mathbf{u} \in \mathbb{R}^3$ and $|\hat{\mathbf{x}}| = 1$. Similarly, the Bingham distribution may be defined as

$$P_B(\hat{\mathbf{x}}; \mathbf{A}) \propto \exp(\hat{\mathbf{x}}^T\mathbf{A}\hat{\mathbf{x}}) \tag{7}$$

where $\mathbf{A} = \mathbf{A}^T$. By multiplying the above we get the 8 parameter Fisher-Bingham[10] [FB_8] distribution

$$P_{\mathrm{FB}_8}(\hat{\mathbf{x}}; \mathbf{u}, \mathbf{A}) \propto \exp(\mathbf{u}^T\hat{\mathbf{x}} + \hat{\mathbf{x}}^T\mathbf{A}\hat{\mathbf{x}}) \tag{8}$$

All three of these distributions have the advantage that they can be multiplied together without introducing further variables, which is critical in a belief propagation framework. We may decompose the FB_8 distribution. As $\mathbf{A}$ is symmetric we may apply the eigen-decomposition to obtain $\mathbf{A} = \mathbf{B}\mathbf{D}\mathbf{B}^T$, where $\mathbf{B}$ is orthogonal and $\mathbf{D}$ diagonal. This allows us to write

$$P_{\mathrm{FB}_8}(\hat{\mathbf{x}}; \mathbf{u}, \mathbf{A}) \propto \exp(\mathbf{v}^T\hat{\mathbf{y}} + \hat{\mathbf{y}}^T\mathbf{D}\hat{\mathbf{y}}) \tag{9}$$

where $\mathbf{v} = \mathbf{B}^T\mathbf{u}$ and $\hat{\mathbf{y}} = \mathbf{B}^T\hat{\mathbf{x}}$. As $|\hat{\mathbf{y}}| = 1$ we may offset $\mathbf{D}$ by an arbitrary multiple of the identity matrix, this allows any given entry to be set to 0. We can therefore consider it the case that $\mathbf{D} = \mathrm{Diag}(\alpha, \beta, 0)$, with $\alpha > 0$ and $\beta > 0$ so that

$$P_{\mathrm{FB}_8}(\hat{\mathbf{x}}; \mathbf{u}, \mathbf{A}) \propto \exp(\mathbf{v}^T\hat{\mathbf{y}} + \alpha\hat{\mathbf{y}}_x^2 + \beta\hat{\mathbf{y}}_y^2) \tag{10}$$

For convenience we may represent the FB_8 distribution as

$$\exp(\mathbf{u}^T\hat{\mathbf{x}} + \hat{\mathbf{x}}^T\mathbf{A}\hat{\mathbf{x}}) = \Omega[\mathbf{u}, \mathbf{A}] \tag{11}$$

Using this notation multiplication is

$$\Omega[\mathbf{u}, \mathbf{A}]\Omega[\mathbf{v}, \mathbf{B}] = \Omega[\mathbf{u} + \mathbf{v}, \mathbf{A} + \mathbf{B}] \tag{12}$$

Various distributions may be represented by the Fisher-Bingham distribution, of particular use is the Bingham-Mardia distribution[11]

$$\exp(-k(\hat{\mathbf{u}}^T\hat{\mathbf{x}} - \cos\theta)^2) = \Omega[2k\cos(\theta)\hat{\mathbf{u}}, -k\hat{\mathbf{u}}\hat{\mathbf{u}}^T] \tag{13}$$

where $\hat{\mathbf{u}}$ is the direction of the axis of a cone and θ the angle of that cone. This distribution has a small circle as its maximum, which allows the irradiance information (Eq. 2) to be expressed as a FB_8 distribution.

5 Method

We construct a graphical model, specifically a pairwise Markov random field. Each node of the model is a random variable that represents an unknown normal on the surface. Belief propagation, as described in section 3, is then used to determine the marginal distribution for each node. To define the distribution to be marginalised two sources are used: the irradiance information (Eq. 2) and a smoothness assumption.

We model the smoothing assumption on the premise that adjacent points on the surface will be more likely to have a small angular difference than a large angular difference. We can express this idea by setting

$$\psi_{pq}(\hat{\mathbf{x}}_p, \hat{\mathbf{x}}_q) = \exp(k(\hat{\mathbf{x}}_p^T\hat{\mathbf{x}}_q)) \tag{14}$$

where $\psi_{pq}(\hat{\mathbf{x}}_p, \hat{\mathbf{x}}_q)$ is from the message passing equation (Eq. 4). This is a Fisher distribution with concentration k. Using FB_8 for the messages and dropping equation 14 into equation 4 we have

$$m_{p\to q}^t(\hat{\mathbf{x}}_q) = \int_{\mathbb{S}^2} \exp(k(\hat{\mathbf{x}}_p^T\hat{\mathbf{x}}_q))t(\hat{\mathbf{x}}_p)\delta\hat{\mathbf{x}}_p \tag{15}$$

$$t(\hat{\mathbf{x}}_p) = \psi_p(\hat{\mathbf{x}}_p) \prod_{u\in(N\setminus q)} m_{u\to p}^{t-1}(\hat{\mathbf{x}}_p) \tag{16}$$

Message passing therefore consists of two steps: calculating $t(\hat{\mathbf{x}}_p)$ by multiplying FB_8 distributions together using equation 12, followed by convolution of the resulting FB_8 distribution by a Fisher distribution to get $m_{p\to q}^t(\hat{\mathbf{x}}_q)$. The next section documents a method for doing the convolution.

For each node we have an irradiance value. Using equations 2 and 13 we can define a distribution

$$\Omega[2k\frac{I_{x,y}}{A}\hat{\mathbf{l}}, -k\hat{\mathbf{l}}\hat{\mathbf{l}}^T] \tag{17}$$

In principle $\psi_p(\hat{\mathbf{x}}_p)$, from equation 16, can be set to this Bingham-Mardia distribution to complete the model to be marginalised. This fails however due to the concave/convex ambiguity[12]. The formulation presented so far will converge to a bi-modal distribution at each node, with the modes corresponding to the concave and convex interpretations. A bias towards one of the two interpretations is required, to avoid arbitrarily selecting between them on a per-pixel basis. Taking the gradient vector at each node and rotating it onto the irradiance defined cone to get $\hat{\mathbf{g}}$ provides a suitable bias direction. This is identical to the initialisation used by Worthington & Hancock[5]. We then multiply equation 17 by a Fisher distribution using this direction vector with concentration h to get

$$\psi_p(\hat{\mathbf{x}}_p) = \exp((h\hat{\mathbf{g}} + 2kI_{x,y}A^{-1}\hat{\mathbf{l}})^T\hat{\mathbf{x}}_p + \hat{\mathbf{x}}_p^T(-k\hat{\mathbf{l}}\hat{\mathbf{l}}^T)\hat{\mathbf{x}}_p) \tag{18}$$

Using the gradient vector unmodified will produce a concave bias, whilst negating it will produce a convex bias. The pseudo-gradient defined in appendix A is used.

Once belief propagation has converged equation 5 can be used to extract a final FB_8 distribution for each node. For output we require directions rather than distributions. A method for finding the maximal mode of the FB_8 distribution is given in appendix B. To optimise the method a hierarchy is constructed and belief propagation is applied at each level. Each level's messages are initialised with the previous, lower resolution, levels messages. This results in less message passes being required for overall convergence[13].

6 Message Passing

As indicated by equation 15 when passing messages we have to convolve a FB_8 distribution by a Fisher distribution. Doing this directly is not tractable, so we propose a novel three step procedure to solve this problem:

1. Convert the FB_8 distribution into a sum of Fisher distributions.
2. Convolve the individual Fisher distributions.
3. Refit a FB_8 distribution to the resulting mixture of Fisher distributions.

All three steps involve approximation, in practise this proves not to be a problem.

Step 1 We approximate the Fisher-Bingham distribution as a sum of Fisher distributions. Starting with equation 10 and rewriting the right-hand side

$$\exp(\mathbf{v}^T\hat{\mathbf{y}})\exp(\alpha\hat{\mathbf{y}}_x^2 + \beta\hat{\mathbf{y}}_y^2) \tag{19}$$

we may substitute an approximation of the right-hand multiple to get

$$\exp(\mathbf{v}^T\hat{\mathbf{y}})\int_0^{2\pi}\exp(m\hat{\mathbf{y}}_x\cos(\theta) + n\hat{\mathbf{y}}_y\sin(\theta))\delta\theta \tag{20}$$

In practise a small number of Fisher distributions will be sampled, to get

$$\exp(\mathbf{v}^T\hat{\mathbf{y}})\sum_i\exp([m\cos(\theta_i), n\sin(\theta_i), 0]\hat{\mathbf{y}}) \tag{21}$$

which may be re-written as a sum of Fisher distributions[1]

$$\sum_i \exp((\mathbf{v} + [m\cos(\theta_i), n\sin(\theta_i), 0]^T)^T \hat{\mathbf{y}}) \tag{22}$$

m and n need to be determined. To explicitly write the approximation

$$\exp(\alpha\hat{\mathbf{y}}_x^2 + \beta\hat{\mathbf{y}}_y^2) \propto \int_0^{2\pi} \exp(m\hat{\mathbf{y}}_x\cos(\theta) + n\hat{\mathbf{y}}_y\sin(\theta))\delta\theta \tag{23}$$

$$\exp(\alpha\hat{\mathbf{y}}_x^2 + \beta\hat{\mathbf{y}}_y^2) \propto 2\pi\mathbf{I}_0(\sqrt{m^2\hat{\mathbf{y}}_x^2 + n^2\hat{\mathbf{y}}_y^2}) \tag{24}$$

where $\mathbf{I}_0$ is the modified Bessel function of the first kind, order 0. Whilst similar[2] the two sides of equation 24 are different, and so a match is not possible, however, we may consider six values of $\hat{\mathbf{y}}$; $[\pm1, 0, 0]^T$, $[0, \pm1, 0]^T$ and $[0, 0, \pm1]^T$. These vectors are the minimas and maximas of the Bingham distribution. Using $[0, 0, \pm1]^T$ we get

$$\exp(0) \propto 2\pi\mathbf{I}_0(0) \equiv 1 \propto 2\pi \tag{25}$$

and, because of normalisation, can write

$$\exp(\alpha) = \mathbf{I}_0(\sqrt{m^2}) \qquad \exp(\beta) = \mathbf{I}_0(\sqrt{n^2}) \tag{26}$$

which can be rearranged to get suitable values of m and n

$$m = \mathbf{I}_0^{-1}(\exp(\alpha)) \qquad n = \mathbf{I}_0^{-1}(\exp(\beta)) \tag{27}$$

This approximation leaves the minimas and maximas in the same locations with the same relative values.

Step 2 Mardia and Jupp [14, pp. 44] give an approximation of the convolution of two Von-Mises distributions. (i.e. distributions on the circle.) If we represent the n-dimensional von-Mises-Fisher distribution as

$$P_{vMF}(\hat{\mathbf{x}}; \hat{\mathbf{w}}, k) \propto \exp(k\hat{\mathbf{w}}^T\hat{\mathbf{x}}) = \psi_n[\hat{\mathbf{w}}, k] \tag{28}$$

where $\hat{\mathbf{x}}, \hat{\mathbf{w}} \in \mathbb{R}^n$ and $|\hat{\mathbf{x}}| = |\hat{\mathbf{w}}| = 1$ then the approximation given is

$$\psi_2[\hat{\mathbf{w}}_1, k_1] * \psi_2[\hat{\mathbf{w}}_2, k_2] \approx \psi_2[\hat{\mathbf{w}}_1 + \hat{\mathbf{w}}_2, A_2^{-1}(A_2(k_1)A_2(k_2))] \tag{29}$$

where $A_p(k) = \frac{\mathbf{I}_{p/2}(k)}{\mathbf{I}_{p/2-1}(k)}$. This may easily be extended to the Fisher distribution with no angular offset between the distributions

$$\psi_3[\hat{\mathbf{w}}, k_1] * \psi_3[\hat{\mathbf{w}}, k_2] \approx \psi_3[\hat{\mathbf{w}}, A_3^{-1}(A_3(k_1)A_3(k_2))] \tag{30}$$

As a computational bonus, $A_3(k)$ may be simplified

$$A_3(k) = \frac{\mathbf{I}_{1.5}(k)}{\mathbf{I}_{0.5}(k)} = \coth(k) - \frac{1}{k} \tag{31}$$

[1] Note that they are written here without normalisation terms; to maintain this under the usual mixture model each Fisher distribution has to be weighted by its inverse normalisation term.

[2] Written as power series they are identical except for the denominators of the terms, for which the Bessel functions are the square of the exponentials.

Step 3 To derive a Fisher-Bingham distribution from the convolved sum of Fisher distributions we first need the rotational component of the Bingham distribution, which we calculate with principal component analysis.

$$\bar{\mathbf{m}} = \frac{\sum_i W_i \mathbf{u}_i}{\sum_i W_i} \tag{32}$$

W_i is the normalisation constant of the indexed Fisher distribution, $\bar{u}_i$ is its direction vector multiplied by its concentration parameter.

$$\mathbf{X} = \begin{bmatrix} W_0(\mathbf{u}_0 - \bar{\mathbf{m}}) \\ W_1(\mathbf{u}_1 - \bar{\mathbf{m}}) \\ \vdots \end{bmatrix} \tag{33}$$

$$\mathbf{X}^T\mathbf{X} = \mathbf{R}\mathbf{E}\mathbf{R}^T \tag{34}$$

$\mathbf{E}$ is the diagonal matrix of eigenvalues. $\mathbf{R}$ is then the rotational component of the Bingham distribution.

Given six directions and their associated density function values we may fit the rest of the parameters to get a distribution with matching ratios between the selected directions. Given six instances of[3]

$$\exp(\mathbf{v}^T\hat{\mathbf{x}} + \hat{\mathbf{x}}^T\mathbf{D}\hat{\mathbf{x}}) = p \tag{35}$$

for a known p and $\hat{\mathbf{x}}$ where $\mathbf{D}$ is diagonal, we can apply the natural logarithm to both sides to get

$$\mathbf{v}^T\hat{\mathbf{x}} + \hat{\mathbf{x}}^T\mathbf{D}\hat{\mathbf{x}} = \ln(p) \tag{36}$$

This is a linear set of equations, which can be solved using standard techniques to get $\mathbf{v}$ and $\mathbf{D}$. The final FB_8 distribution is then proportional to

$$\exp((\mathbf{R}\mathbf{v})^T\hat{\mathbf{x}} + \hat{\mathbf{x}}^T\mathbf{R}\mathbf{D}\mathbf{R}^T\hat{\mathbf{x}}) \tag{37}$$

The six directions have to be carefully selected to produce a reasonable approximation, as only these sampled directions will be fitted and the convolved distribution can differ greatly from a Fisher-Bingham distribution. The selection strategy used is based on the observation that with no Fisher component the optimal selection is $[\pm 1, 0, 0]^T$, $[0, \pm 1, 0]^T$ and $[0, 0, \pm 1]^T$ (There is also a computational advantage of this selection as they are linearly separable.). Given a Fisher component we may divide through the mixture of Fisher distributions to leave only a (supposed) Bingham component; the estimation procedure will then estimate another Fisher component as well as the Bingham component. This leads to an iterative scheme, where the Fisher component is initialised with the mean of the mixture of Fisher distributions and updated after each iteration. In practise convergence happens after only two iterations. It should be noted that this approach is the inverse operation of the initial conversion to a mixture of Fisher distributions, i.e. it has error precisely opposite the error introduced by step 1, ignoring the use of a finite number of Fisher distributions.

[3] It should be noted that equality rather than proportionality is used here. This is irrelevant as multiplicative constants have no effect.

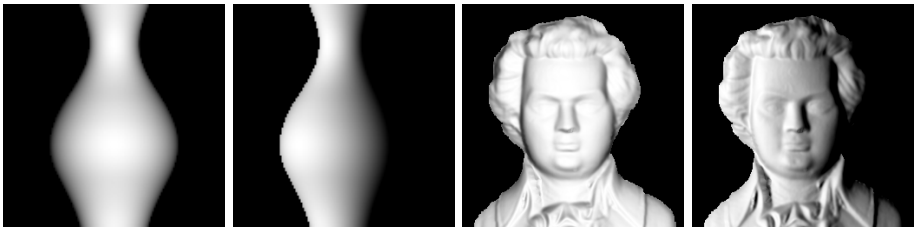

Fig. 1. Synthetic inputs, derived from the set used by Zhang et al[8]. From left to right they are referred to as Vase 90°, Vase 45°, Mozart 90° and Mozart 45°. The light source direction vector for the 90° images is $[0, 0, 1]^T$, whilst for the 45° images it is $[-\sqrt{2}, 0, \sqrt{2}]^T$.

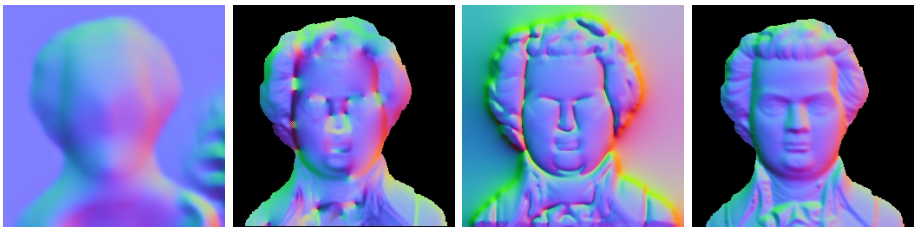

Fig. 2. Results for the synthetic Mozart 90° input. From left to right they are Lee & Kuo[4], Worthington & Hancock[5], the presented algorithm and then finally ground truth. They represent normal maps, with $x \rightarrow$ red, $y \rightarrow$ green and $z \rightarrow$ blue to represent the surface normal at each pixel. Red and Green are adjusted to cover the whole $[-1, 1]$ range, blue is left covering $[0, 1]$.

7 Results & Analysis

We compare the presented algorithm to two others, Lee & Kuo[4] and Worthington & Hancock[5], using both synthetic and real data. Figure 1 gives the four synthetic inputs used, figure 2 gives the results and ground truth for just one of the four inputs. Qualitatively, looking at figure 2, Lee & Kuo is simply too blurred to be competitive. Worthington & Hancock shows considerably more detail, but suffers from assorted artifacts and is still blurred. The presented algorithm has sharp details and less blurring compared to the others.

Figure 3 gives the results of a quantitative analysis of the synthetic results. Each table gives the results for one of the four inputs, with each row dedicated to an algorithm. The columns give the percentage of pixels in each image that are beneath an error threshold, the error being the angle between the ground truth and estimated normals. Sticking to the 90° images where the light is at $[0, 0, 1]^T$ Lee & Kuo consistently makes fewer large mistakes, this can be put down to its excessive blurring. Worthington & Hancock appears to have an advantage at the very lower ends of the scale, this is presumably because it perfectly matches

Vase 90°	$< 1°$	$< 2°$	$< 3°$	$< 4°$	$< 5°$	$< 10°$	$< 15°$	$< 20°$	$< 25°$
Lee & Kuo	0.8	3.3	7.0	12.0	21.6	75.7	97.7	100.0	100.0
Worthington & Hancock	6.8	13.3	17.8	22.2	26.7	46.5	59.1	67.9	75.3
Presented Algorithm	7.8	13.4	22.5	34.5	39.0	55.9	68.1	76.7	83.9

Vase 45°	$< 1°$	$< 2°$	$< 3°$	$< 4°$	$< 5°$	$< 10°$	$< 15°$	$< 20°$	$< 25°$
Lee & Kuo	0.9	3.9	7.4	11.4	15.7	47.0	73.8	85.1	88.8
Worthington & Hancock	6.6	13.4	17.4	20.4	24.1	37.3	49.1	57.9	65.1
Presented Algorithm	0.3	4.4	10.3	18.4	28.4	44.5	58.0	68.4	76.7

Mozart 90°	$< 1°$	$< 2°$	$< 3°$	$< 4°$	$< 5°$	$< 10°$	$< 15°$	$< 20°$	$< 25°$
Lee & Kuo	0.2	0.7	1.5	2.6	4.1	18.3	36.1	52.5	64.9
Worthington & Hancock	2.7	6.4	10.4	14.3	18.4	34.4	47.2	56.3	63.9
Presented Algorithm	0.9	3.7	8.5	15.4	21.7	42.2	53.5	61.9	68.5

Mozart 45°	$< 1°$	$< 2°$	$< 3°$	$< 4°$	$< 5°$	$< 10°$	$< 15°$	$< 20°$	$< 25°$
Lee & Kuo	0.2	0.7	1.5	2.5	3.8	16.1	35.0	54.7	67.2
Worthington & Hancock	2.4	5.4	8.0	10.4	13.4	25.0	33.4	40.5	46.8
Presented Algorithm	0.2	0.8	2.1	4.5	7.9	21.9	33.3	43.1	50.4

Fig. 3. Synthetic results. Each grid gives results for the input named in the top left. Each row gives results for a specific algorithm. Each column gives the percentage of pixels within a given error bound, i.e. the $< 1°$ column gives the percentage of pixels where the estimated surface orientation is within 1 degree of the ground truth. The percentage is only for pixels where ground truth is provided.

the irradiance information, unlike the others. The presented approach is always ahead for the Vase 90° input. For the Mozart 90° input our approach consistently exceeds Lee & Kuo but does not do so well at getting a high percentage of spot on estimates as Worthington & Hancock. However, for error thresholds of 4° and larger the presented algorithm is again better.

Moving to the 45° inputs, where the light source direction vector is $[-\sqrt{2}, 0, \sqrt{2}]^T$, things do not go so well. For Vase 45° it gets the highest percentage of pixels with an error less than 5°, but above that is exceeded by Lee & Kuo and below that beaten by Worthington & Hancock. For Mozart 45° Worthington & Hancock is the clear victor. The presented algorithm doing poorly as the light source moves away from $[0, 0, 1]^T$ can be put down to the bias introduced to handle the concave/convex ambiguity[12]. The gradient information used for the bias is necessary to avoid a bi-modal result, but also pulls the solution away from the correct answer, this effect being more noticeable as the light deviates away from being at the camera.

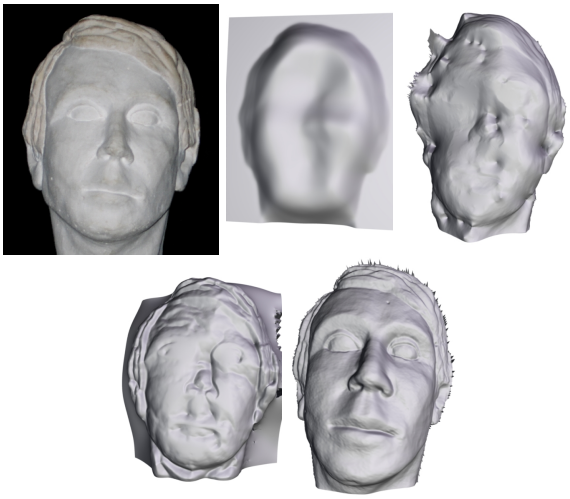

Fig. 4. Input and results for the head. From left to right they are input, Lee & Kuo[4] and Worthington & Hancock[5] on the first line and the presented algorithm and then ground truth on the second.

Head	$< 1°$	$< 2°$	$< 3°$	$< 4°$	$< 5°$	$< 10°$	$< 15°$	$< 20°$	$< 25°$
Lee & Kuo	0.1	0.4	0.8	1.4	2.2	8.6	19.7	32.1	43.5
Worthington & Hancock	0.1	0.6	1.4	2.6	4.0	13.7	23.4	32.6	41.0
Presented Algorithm	0.1	0.5	1.1	1.9	3.0	11.5	21.4	30.9	39.6

Fig. 5. Results for head input. See figure 3 for explanation.

Figure 4 gives a real world input and the results as 3D renders of the integrated output, figure 5 gives the same quantitative analysis used for the synthetic results. This input was captured in a dark room using a camera with a calibrated response curve and the shape determined with a head scanner, with the camera calibrated to the scanners coordinate system so that a ground truth normal map could be produced. Looking at figure 5 Worthington & Hancock is quantitatively ahead, but looking at the actual output it is more blob than face, though some features are recognisable. To use an analogy, an art restorer painting over a canvas with constant colour knowing that the original artist must have used that colour in some of the areas covered can get the most matches if the competition is terrible, despite producing a blurred result. Sticking to a qualitative analysis the presented algorithm is clearly not perfect, but it gives sharper results, with features such as mouth, eye sockets and hair that are superior to the competition.

For the head image the run time is over 12 hours for Lee & Kuo, 54 minutes for Worthington and Hancock and 9.5 minutes for the presented algorithm on a 2Ghz Athlon.

8 Conclusion

We have presented a new algorithm for solving the classical shape from shading algorithm, and demonstrated its competitiveness with previously published algorithms. The use of belief propagation with FB_8 distributions is in itself new, and a method for the convolution of a FB_8 distribution by a Fisher distribution has been devised. The algorithm does suffer a noticeable flaw in that overcoming the convex/concave problem biases the result, making the algorithm weak in the presence of oblique lighting. An alternative solution to the current bias is an obvious area for future research.

A Pseudo Gradient

A diffusion method is used to calculate an estimate of the gradient direction. This method is robust in the presence of noise and lacks the distortion of methods such as the Sobel operator. It is described here in terms of a random walk.

All walks start at the pixel for which the calculation is being applied and are of fixed length. Each walk contributes a vector going from the walks start point to the walks end point, the mean of these vectors is the output gradient direction. Every step the walk moves to one of the four adjacent pixels, the pixels are weighted by $\alpha + I^{\beta}_{(x,y)}$ where $I_{(x,y)}$ is the irradiance of the pixel and α and β are parameters. This creates a walk that tends towards brighter areas of the image, the mean being a robust gradient direction.

B Maximisation of FB_8

For visualisation and integration with non-probabilistic modules finding the direction with the greatest density is needed. This is the quadratic programming problem of maximising equation 9. It may be solved efficiently by observing that it is the same problem as finding the closest point on an ellipsoid to a given point. This latter problem can be expressed as an order 6 polynomial and solved with Newtons method. As a further convenience the initialisation can be done in such a way that it always converges directly to the maximal root[15].

The Fisher-Bingham distribution is a conditioned multivariate normal distribution[14, pp. 175]

$$\Sigma = \frac{-(\mathbf{D} + c\mathbf{I}_3)^{-1}}{2} \qquad \bar{\mu} = \Sigma\mathbf{v} \tag{38}$$

where $c\mathbf{I}_3$ is a scaled identity matrix selected to make $\mathbf{D} + c\mathbf{I}_3$ negative definite. The maximal point is therefore the closest point to $\bar{\mu}$ on the unit sphere using Mahalanobis distance. Mahalanobis distance is

$$\sqrt{(\hat{\mathbf{x}} - \bar{\mu})^T \Sigma^{-1} (\hat{\mathbf{y}} - \bar{\mu})} \tag{39}$$

To minimise the above equation we consider that Σ is diagonal and rewrite as

$$\sqrt{\sum_i [(\hat{\mathbf{y}}_i - \bar{\mu}_i)^2 \sigma_i^{-1}]} \tag{40}$$

where $\sigma_i, i \in \{1, 2, 3\}$ are the elements of Σ, which may be rearranged as

$$\sqrt{\sum_i [(\mathbf{z}_i - \sqrt{\sigma_i^{-1}}\bar{\mu}_i)^2]} \tag{41}$$

where $\mathbf{z}_i = \sqrt{\sigma_i^{-1}}\hat{\mathbf{y}}_i$. This is now Euclidean distance when solving for $\mathbf{z}_i$, and the constraint that $\hat{\mathbf{y}}$ be of unit length becomes the equation of an ellipsoid

$$\sum_i [(\frac{\mathbf{z}_i}{\sqrt{\sigma_i^{-1}}})^2] = 1 \tag{42}$$

References

1. Horn, B.K.P.: Shape From Shading: A Method For Obtaining The Shape Of A Smooth Opaque Object From One View. PhD thesis, Massachusetts Institute of Technology (1970)
2. Brooks, M.J., Horn, B.K.P.: Shape and source from shading. Artificial Intelligence (1985) 932–936
3. Zheng, Q., Chellappa, R.: Estimation of illuminant direction, albedo, and shape from shading. Pattern Analysis and Machine Intelligence **13(7)** (1991) 680–702
4. Lee, K.M., Kuo, C.C.J.: Shape from shading with perspective projection. CVGIP: Image Understanding **59(2)** (1994) 202–212
5. Worthington, P.L., Hancock, E.R.: New constraints on data-closeness and needle map consistency for shape-from-shading. Pattern Analysis and Machine Intelligence **21(12)** (1999) 1250–1267
6. Prados, E., Camilli, F., Faugeras, O.: A unifying and rigorous shape from shading method adapted to realistic data and applications. Mathematical Imaging and Vision **25(3)** (2006) 307–328
7. Potetz, B.: Efficient belief propagation for vision using linear constraint nodes. Computer Vision and Pattern Recognition (2007) 1–8
8. Zhang, R., Tsai, P.S., Cryer, J.E., Shah, M.: Shape from shading: A survey. Pattern Analysis and Machine Intelligence **21(8)** (1999) 690–706
9. Haines, T.S.F., Wilson, R.C.: Integrating stereo with shape-from-shading derived orientation information. British Machine Vision Conference **2** (2007) 910–919
10. Kent, J.T.: The Fisher-Bingham distribution on the sphere. Royal Statistical Society, Series B (Methodological) **44(1)** (1982) 71–80
11. Bingham, C., Mardia, K.V.: A small circle distribution on the sphere. Biometrika **65(2)** (1978) 379–389
12. Ramachandran, V.S.: Perception of shape from shading. Nature **331** (1988) 163–165
13. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient belief propagation for early vision. Computer Vision and Pattern Recognition **1** (2004) 261–268
14. Mardia, K.V., Jupp, P.E.: Directional Statistics. Wiley (2000)
15. Hart, J.C.: Distance to an ellipsoid. Graphics Gems IV (1994) 113–119