A Rough Guide to Shape from Shading

Tom S. F. Haines
thaines@eecs.qmul.ac.uk

Queen Mary, University of London

14$^{\text{th}}$ June 2012

Roadmap

1. The Original Model
2. A Little Bit Of Theoretical Stuff
3. Alternate Assumptions
4. SfS with...

Shape from Shading (SfS)

- Image formulation rules tell you how to go from a 3D model and its materials to a 2D image.
- Shape from shading is the inverse problem.
- It can be seen as a constraint on the set of possible realities.
- Justifying working with it can take several arguments - the simplest is that multiple species of animals, ourselves included, use it.

Roadmap

1. The Original Model
2. A Little Bit Of Theoretical Stuff
3. Alternate Assumptions
4. SfS with...

Pre-Horn

- In other fields the term *photoclinometry* is used instead of *Shape from Shading*.
- Earliest reference I know of: Diggelen, *A photometric investigation of the slopes of the heights of the ranges of hills in the maria of the moon*, 1951. Observed that the moon was a Lambertian surface and inferred the relative depths of 1D slices. By hand.
- First use of a computer that I know of: Rindfleisch, *Photometric method for lunar topography*, 1966.

Horn

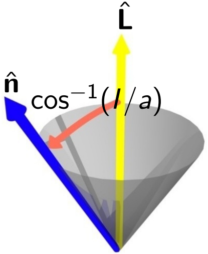
- *Shape From Shading: A Method For Obtaining The Shape Of A Smooth Opaque Object From One View*, 1970.
- First to work with entire surfaces, rather than 1D slices.
- Used the assumptions: Unknown object, Lambertian reflectance, orthographic projection, constant known albedo, a smooth surface, valid surface not required, no surface inter-reflectance and a single infinitely distant and known light source, with no falloff.
- Still ill defined - 2 degrees of freedom for each surface normal when only 1 degree is constrained.
- Will explain the model in geometric terms later.

Lee & Kuo

- *Shape from shading with perspective projection*, 1994.
- Best approach in 1999 according to a review paper: Zhang, Tsai, Cryer, and Shah, *Shape from shading: A survey*, 1999.
- This and all other approaches in the survey used Horn's assumptions - this was a comparison of who had the best optimiser.
- Solved by repeatedly linearising the equations and solving the constraints for each linearisation - closely related to Newtons method.
- Also used something called a 'multigrid method'. This is basically solving at various resolutions, to avoid the really bad local minima.
- Actually contains an integration constraint, which we will come to later...

Worthington & Hancock

- *New constraints on data-closeness and needle map consistency for shape-from-shading*, 1999.
- Used a hard constraint - that re-rendering the model must produce the original image. This leads to a simple geometric interpretation of the model:



- $\hat{\mathbf{n}} \cdot \hat{\mathbf{L}} = I/a$
- For each pixel - $\hat{\mathbf{L}}$ is the direction to the light source; I is the image brightness; a is the albedo and $\hat{\mathbf{n}}$ is the (unknown) surface normal.

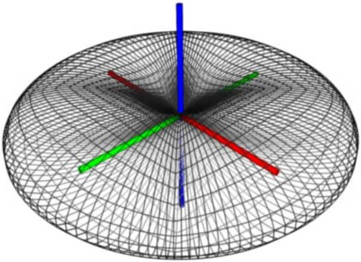
Worthington & Hancock, continued...

- Solved the model by iterating between smoothing the field of surface normals and then projecting the normals back onto the closest point on the cone.
- Initialised the field using image gradients, pointing away from the brightest direction. This is an implicit assumption that the light source is probably near the camera.
- Whole bunch of different smoothing approaches tried, but the initialisation is has the greatest impact on the results.

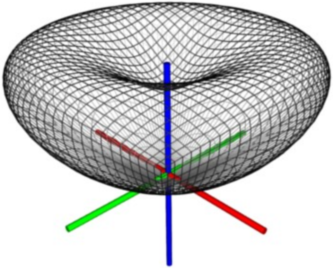
Haines & Wilson

- *Belief propagation with directional statistics for solving the shape-from-shading problem*, 2008.
- A probabilistic interpretation of Worthington & Hancock - solved the same model but used belief propagation with directional statistics to solve a Markov random field.
- Directional statistics are probability distributions over directions - to make it work an approximation of convolving a Fisher-Bingham-8 distribution by a Fisher distribution has to be developed.

Directional Statistics Visualisation



Bingham distribution, $\alpha = \beta = 5$. Can represent gradient information.

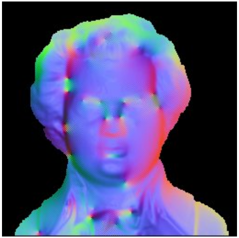


Bingham-Mardia distribution, $k = 8, \texttt{angle} = 45^\circ$. Can represent the cone constraint.

Results: Mozart at 90°

Mozart, lit head on.

Ground truth



Worthington & Hancock

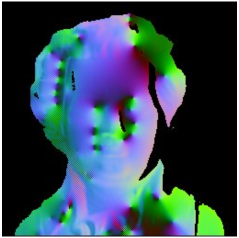
Lee & Kuo

Haines & Wilson

Results: Mozart at 45°

Mozart, lit at 45° from head on.

Ground truth

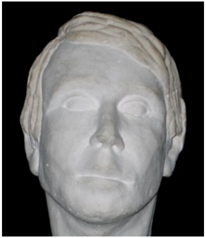


Worthington & Hancock

Lee & Kuo

Haines & Wilson

Results: Head

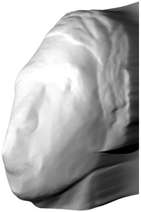


Head, lit head on.

Ground truth

Worthington & Hancock

Lee & Kuo

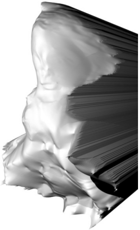


Haines & Wilson

Results: Venus

Venus, lit head on.

Ground truth



Worthington & Hancock

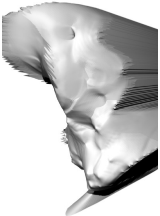
Lee & Kuo

Haines & Wilson

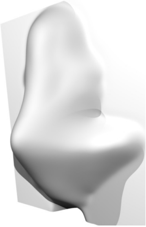
Results: Bard

Bard, lit head on.

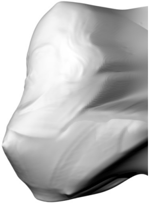
Ground truth



Worthington & Hancock



Lee & Kuo

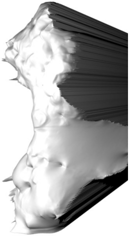


Haines & Wilson

Results: Sunev

Sunev, lit head on.

Ground truth



Worthington & Hancock

Haines & Wilson

Roadmap

1. The Original Model
2. A Little Bit Of Theoretical Stuff
3. Alternate Assumptions
4. SfS with. . .

Error Measurement

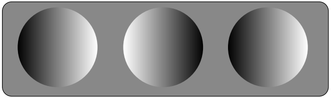
- Shading provides directional information, not depth.
- If you integrate to get depth the errors in surface orientation will add up - the further you go the larger they will be. This is called *curl*.
- Consequentially it makes more sense to measure SfS algorithms in terms of their angular error.
- Most papers use some function of depth error however, including both of the SfS review papers (Though the 2008 review is pointless anyway.).
- Funnily enough some of the 'best' algorithms have strong fronto-planar assumptions, that push the answer towards being a flat plane.

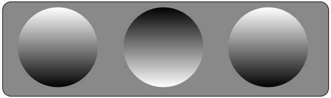
Calibration

- Cameras do not output a linear measure of irradiance - the function between pixel values and actual irradiance often forms an 'S' shaped curve.
- SfS algorithms need to know it - if you are doing SfS you must calibrate your camera, as otherwise it will destroy performance.
- In the event that the camera does not provide it or there is no way to infer it from the available data offline calibration can be done in the same way that a multiple exposure HDR photo is taken.

Bas-relief Ambiguity

- If the light source position is unknown then there is a 1D family of valid surfaces.
- If the light source position is known then two members of that family are valid, as illustrated by the optical illusion below. This is the concave/convex ambiguity. (Does not exist if the light source is at the camera.)





- Not a major problem once you move beyond just SfS, as other sources of information can resolve it.
- One trick often used is to observe that most objects are approximately convex, and initialise as such, so you at least get stuck in a convex local minima.

Roadmap

Integration

- So far *needle maps* (Surface orientation at every pixel.) have been generated.
- Depth is usually desired, and as orientation is depths differential we have to integrate to get it.
- If you start travelling and end up at the same place as where you started the sum of how far you travelled in every dimension must be zero. An arbitrary needle map does not necessarily enforce this condition, and most of the approaches so far do not enforce it.
- There are actually papers on obtaining surfaces from dodgy needle maps (I have always used Gaussian belief propagation.).
- Note that the $+c$ of integration exists - depth output can only be relative, unless you have further information.

Frankot & Chellappa

- *A method for enforcing integrability in shape from shading algorithms*, 1988.
- Added an integrability constraint.
- Represented by $\frac{\delta^2 z}{\delta x \delta y} = \frac{\delta^2 z}{\delta y \delta x}$.
- They formulate a method of projecting a needle map onto the nearest needle map that satisfies this constraint.
- Can therefore work with any SfS algorithm.

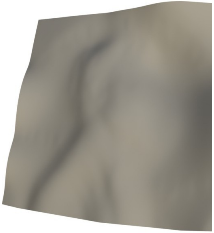
Potetz

- *Efficient belief propagation for vision using linear constraint nodes*, 2007.
- Easiest way to add an integration constraint is to use a surface representation that enforces it, such as a depth map.
- Potetz did this (Not the first), and solved it using an extremely hardcore BP method, with large cliques and arbitrary continuous multidimensional probability distributions.
- Insanely demanding in terms of both memory and computation, but the results are probably the best possible for this kind of model.

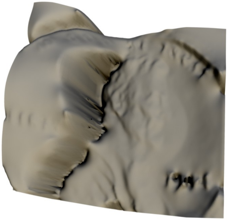
Potetz Results: Penny

Coin, lit at 45°.

Ground truth



Lee & Kuo



Haines & Wilson

Potetz

Smooth Surface

- Assuming a smooth surface causes issues, particularly at discontinuities, and explicitly handling them has a similar effect to the same in stereopsis.
- Zheng & Chellappa, *Estimation of illuminant direction, albedo, and shape from shading*, 1991: Used the basic idea that large changes in irradiance probably indicate a discontinuity in the surface.

Non-Lambertian Shading

- The Lambertian Bidirectional Reflectance Distribution Function (BRDF) is obviously not exhibited by many objects.
- Some simpler BRDF models can be converted into Lambertian as a pre-processing step - as long as it forms a cone constraint the irradiance can be converted so it makes the correct angle.
- The biggest issue is knowing the model - for Lambertian reflectance it is possible to estimate its albedo (Using points that face the light source), but as the model gets harder so does estimating it.

Non-Lambertian Shading Examples

- Healey & Binford, *Local shape from specularity*, 1988: Used the Torrence-Sparrow model, which is a model of specularities. Assumes that there is no range issues and fits the parameters after segmenting small bright regions.
- Ahmed & Farag, *A new formulation for shape from shading for non-lambertian surfaces*, 2006: Oren-Nayar, learns the parameters by expressing the entire problem as a partial differential equation and using a sophisticated solver.
- Approaches such as Vega & Yang, *Shading logic: A heuristic approach to recover shape from shading*, 1997; Lee & Kuo, *Shape from shading with a generalized reflectance map model*, 1997; only ever need to evaluate the BRDF, and calculate its differential, which can be done using finite differences. Any BRDF can be used with them.

Specularity Detection

- Specularities will trip up a SfS algorithm.
- Using a specularity detector as a pre-processing step is extremely helpful.
- A more integrated example is Ragheb & Hancock, *Separating lambertian and specular reflectance components using iterated conditional modes*, 2001; where belief propagation is used to learn the ratio of Lambertian and specular contribution to each pixel, in combination with a SfS algorithm.

Variable Surface Parameters

- SfS is usually posed for objects with a single material, e.g. a constant albedo Lambertian model.
- Variable albedo, or other parameters is usually trivial if known - you just have different parameters for each pixel. Knowing it is the problem.
- The area of intrinsic image research is dedicated to inferring albedo from an image, though Tappen, Freeman & Adelson, *Recovering intrinsic images from a single image*, 2005 is about as good as it gets.
- Integrated solving for albedo and SfS is demonstrated by Fua & Leclerc, *Object-centered surface reconstruction: combining multi-image stereo and shading*, 1995. It optimises the vertices of a 3D mesh, also making use of stereopsis. Has a cost term to prefer piecewise constant albedo.

Perspective

- As cameras tend to use perspective projection assuming orthographic will distort objects badly, particularly when close to the camera.
- Many of the more recent algorithms given over the following slides include perspective.
- Algorithms that consider depth can often get perspective for free.

Lighting falloff

- Prados & Faugeras. *Shape from shading: a well-posed problem*, 2005: Demonstrate that if you include perspective and model the falloff from the light source the problem is no longer ill-posed and there is only one answer.
- The only issue is it breaks if lighting falloff is not detectable - you will probably need to use 12bit per channel capture.
- There is an entire sequence of these papers, including ones by different authors. The model was initially limited to the light source being at the camera, which has been relaxed. Support for alternate shading models has also been added.
- They all formulate the problem as a PDE, and find 'viscosity solutions'.

Multiple lights

-
- Prados, Camilli & Faugeras, *A unifying and rigorous shape from shading method adapted to realistic data and applications*, 2006: Consider a light source at the camera, with perspective and lighting falloff, e.g. a flash on a camera.
- Tian, Tsui, Yeung & Ma, *Shape from shading for multiple light sources*, 1999: Consider multiple area lights. The approach requires the depth of singular points (The points that face a light source) however, which is hardly practical, and then propagate values from these points, which is error prone.

SfS outside the dark room

- Langer & Zucker, *Shape from shading on a cloudy day*, 1994: Considers uniform light coming from a hemisphere, e.g. the sky, consequentially pixel brightness is a function of occlusion. Uses a voxel grid in a space carving style algorithm - it effectively digs pits until they are deep enough to match the solid angle of visible sky implied by the pixels irradiance.
- A more recent instance of the above is Prados, Jindal & Soatto, *A Non-Local Approach to Shape From Ambient Shading*, 2009, which formulates the problem using (horrible) PDEs.
- Whilst a very cool idea these approaches just don't work - all of these papers give abstract results, i.e. runs on synthetic images that are not even recognisable as real objects.
- Note that Brooks & Horn, *Shape and source from shading*, 1985, introduced the idea of a uniform hemisphere to represent the sky. Whilst they gave an algorithm no results were presented, and I am doubtful it could have ever worked.

Inter-reflections

- Nayar, Ikeuchi & Kanade, *Shape from interreflections*, 1990: The only approach that considers interreflections.
- Reasons that interreflections lighten surfaces, and make them flatter/less concave.
- Therefore if you estimate the surface using any SfS algorithm and then use that to estimate the contribution from interreflections it will always be an underestimate.
- Therefore iteratively run SfS and interreflection estimation, using the interreflection estimate each time to reduce the interreflections in the image.

Shadows

- One approach is to use a shadow detection method and switch off SfS in such areas.
- Deformable model based approaches can model shadows explicitly, e.g. Samaras & Metaxas, *Incorporating illumination constraints in deformable models*, 1998.

Inpainting

- As strange as it sounds detecting areas where SfS is going to fail (e.g. texture), deleting them and applying an inpainting algorithm to fill in the gaps, before running SfS on the entire image, is a thing.
- It probably works because inpainting algorithms actually work quite well.
- A recent example is Zhang, Yip, Brown & Tan, *A Unified Framework for Document Restoration using Inpainting and Shape-from-Shading*, 2009, which uses this approach to infer the shape of pages to obtain a flat 'scan' of the page without damaging the book.

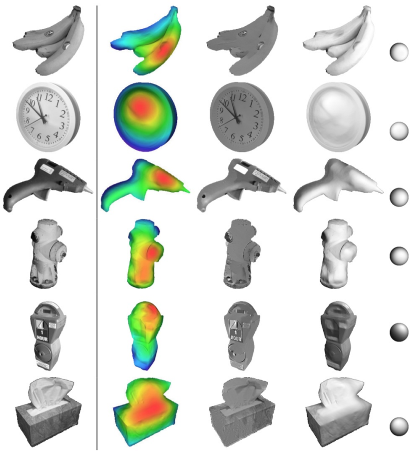
Roadmap

...Statistical models

- If the object in question belongs to a known class, for which a statistical model is avalible, then the model can be fitted to the SfS data.
- Atick,Griffin & Redlich, *Statistical approach to shape from shading: Reconstruction of three-dimensional face surfaces from single two-dimensional images*, 1996: First instance of this, using a model of human faces.
- Will Smith (Uni. of York) has done an awful lot of work on human faces with SfS, such as his thesis, *Statistical Methods For Facial Shape-from-shading and Recognition*, 2007.
- A further example is Dovgard & Basri, *Statistical symmetric shape from shading for 3d structure recovery of faces*, 2004, who make use of symmetry.

- Barron & Malik, *Shape, Albedo, and Illumination from a Single image of an Unknown Object*, 2012: Learns everything, by using the SfS constraint alongside various costs that provide it with clues as to which SfS solution to choose.
- For albedo estimation they have a smoothness prior and a minimal entropy - the first to encourage regions of contiguous or smoothly varying albedo, the second to encourage albedo reuse around the image.
- For shape they have a prior towards flat objects, towards occluding contours having surface orientation in the image plane, and a term to encourage constant changes in mean curvature (To encourage spherical objects.).

Figure 1 from the previous



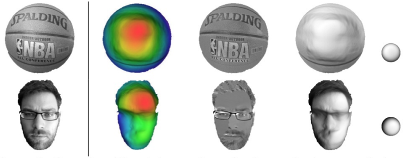


Figure 1. Our algorithm takes only a single masked grayscale image as input (shown on the left) and produces as output a depth map, albedo map, shading image, and spherical harmonic illumination model. These images were taken by the authors with a cellphone camera in uncontrolled indoor and outdoor illumination conditions. All images and results in this paper were produced using the same piece of code with the same parameter settings. The "shading" image is a rendering of the recovered depth under the recovered illumination. Depth is shown with a pseudo-color visualization (red is near, blue is far). Many more similar results can be found in the supplementary material.

. . . (Multi-view) Stereopsis

- Stereo is the obvious thing to combine SfS with:
  - SfS does well in smoothly shaded areas where stereo has no information.
  - Stereo does well in areas where there is texture to match, areas where SfS fails because we can't infer the albedo.
  - SfS tends to provide fine detail, but gets the large scale details wrong, whilst stereo gets the large details right but is not good with the fine details.
- Blake, Zisserman & Knowles, *Surface descriptions from stereo and shading*,1985 : Discussed the idea of using SfS to interpolate between sparse points found by stereo. No algorithm was given however.

Modular approaches

- Where you run them separately, without integrating them.
- Leclerc & Bobick, *The direct computation of height from shading*, 1991: An Sfs algorithm, but they initialise a depth map using stereo to avoid getting stuck in a nasty local minima.
- Mostafa, Yamany & Farag, *Integrating stereo and shape from shading*, 1999: Run both, using sparse Stereo, then subtract the SfS from the stereo and fit a smooth surface, before adding back the SfS.
- Cryer, Tsai & Shah, *Integration of shape from shading and stereo*, 1995: Take the stereo is low frequency information and SfS is high literally - use the fast Fourier transform to apply low and high pass filters respectively.

Haines & Wilson

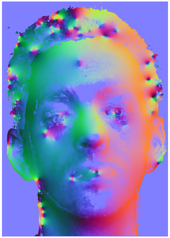
- Another modular approach, though it iterates between them so each estimate can inform the other.
- Uses Gaussian belief propagation on a Markov random field to combine them - Stereo provides the data term, SfS provides the smoothness term. Error estimates are used to set the Standard deviations of these sources correctly.
- Used Worthington & Hancock for SfS, and a BP based algorithm for stereo.
- Includes piece-wise constant albedo estimation, using segmentation to define the constant areas.

Results: Me, part 1

Left image.

Estimated Albedo

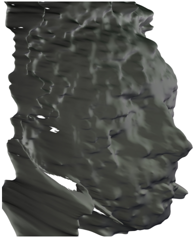
Stereo needle map



SfS needle map

Results: Me, part 2

Discrete stereo result



Smoothed stereo result

Ground truth

Output

Zombie

Object centred approaches

- Many examples of these - they all involve making some initial estimate of the model using stereo, then updating the model to obey SfS derived orientation data. The previous approach is technically one of them, though most of these use multiple cameras rather than just a stereo pair.
- Fua & Leclerc, *Object-centered surface reconstruction: combining multi-image stereo and shading*, 1995: Probably the first such algorithm - works exactly as described above. Includes albedo estimation. The model is a mesh, with per face albedo.
- There is a long chain of papers on the same theme - shadow handling, alternate shading models and specularity handling have all been explored.
- The biggest issue with them is initialisation and optimisation method - the early approaches did poorly because of weak optimisation. Initialisation remains an issue - a bad initial model will not get better due to SfS.

Wu, Wilburn, Matsushita & Theobalt

- *High-quality shape from multi-view stereo and shading under general illumination*, 2011.
- The best results I've seen with multi-view stereo and SfS. Its the deformable model approach with up to date techniques, polished to give good results.
- MVS is doing most of the work however - it provides the initial model, which has to be accurate.
- It estimates arbitrary lighting - they use spherical harmonics to represent the lighting environment, in the same form as a light probe - MVS gives surface orientation and the irradiance is a linear combination of the light emitted by each harmonic - a linear equation. It factors in shadows.
- Surface refinement uses the light model and irradiance to push the surface to comply. Adaptively decides if it should use the information or not, so stereo can remain king where there is texture.

Figure 1 from the previous

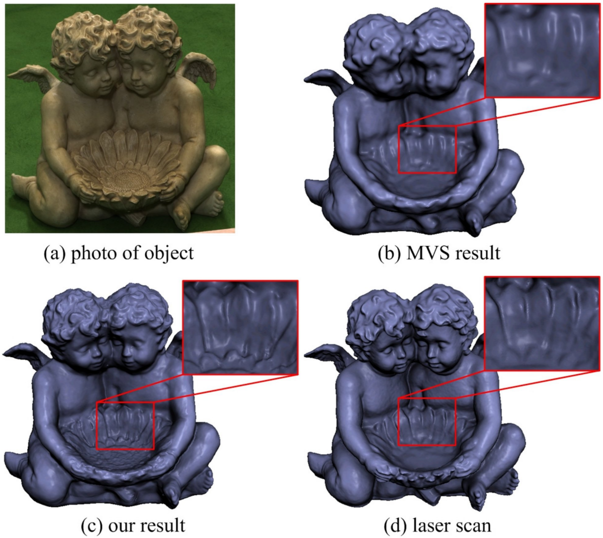


Figure 1. Our approach reconstructs models of much higher detail than state-of-the-art stereo approaches.

Conclusions

- SfS on its own is (mostly) a silly idea.
- There are valid applications - e.g. book page shape recovery.
- Combined with other cues is can work very well.
- It is an incredibly varied area, and people have worked on some really crazy formulations - this talk has only scratched the surface.
- Light source estimation, shadow and specularity detection and intrinsic images are all closely related areas that have only been touched on.