Shading Primitives: Finding Folds and Shallow Grooves

John Haddon and David Forsyth

Computer Science Division
University of California, Berkeley CA 94720
{haddon,daf}@cs.berkeley.edu

Abstract

Diffuse interreflections cause effects that make current theories of shape from shading unsatisfactory. We show that distant radiating surfaces produce radiosity effects at low spatial frequencies. This means that, if a shading pattern has a small region of support, unseen surfaces in the environment can only produce effects that vary slowly over the support region. It is therefore relatively easy to construct matching processes for such patterns that are robust to interreflections. We call regions with these patterns "shading primitives."

Folds and grooves on surfaces provide two examples of shading primitives; the shading pattern is relatively independent of surface shape at a fold or a groove, and the pattern is localised. We show that the pattern of shading can be predicted accurately by a simple model, and derive a matching process from this model. Both groove and fold matchers are shown to work well on images of real scenes.

1 Introduction

There is a long history of study of the relationship between surface shading and shape (reviewed in greater detail in, for example, [9]). Horn's seminal thesis [8] established the image irradiance equation as the fundamental relationship between surface geometry and image brightness. Throughout the extensive literature on solutions to this equation (see bibliography in [9]) the emphasis has been on obtaining a dense depth map from shading, assuming that the object has been segmented, that shadow, reflectance and illumination boundaries either do not occur or have been identified, and that sufficient information is available to determine the reflectance map that applies.

This model of shading omits the effects of diffuse interreflections, a source of various substantial effects that are noticeable to human observers [17] and represent an important component of shading in most real scenes [4]. Diffuse interreflections create significant computational difficulties, because brightness at a patch is not simply a function of the surface normal, but a function of all radiating surfaces, modelled by the *radiosity equation* [18]. The relationship between an object's shape and its radiosity—the main matter to study—is exceptionally complex.

2 Shape and shading

Very few techniques for extracting shape information from shading fields are robust to the effects of diffuse interreflections. Nayar's method [16] for extracting a dense depth map uses estimates of interreflections, but does not take into account the effects of shadows, and of distant surfaces. Belhumeur *et al*. [1] showed that different surfaces under different illumination conditions can yield the same image. To overcome the difficulty of finding dense depth maps, one may either change the type of shape representation sought in the shading field, or the model of shading. Langer *et al*. adopted a model of shading where surface brightness at a point was proportional to the visible solid angle subtended by the sky at that point [15, 14]. Koenderink pointed out that deep pits and grooves had characteristic photometric properties, and considered shading properties characteristic of particular shapes [12]. The other approach, as advocated by Forsyth and Zisserman, is to attempt to build a dictionary of corresponding stylised shape and shading events, so that shading properties that are stable under interreflections can be used to infer a shape representation [5].

In our opinion, the best prospect for extracting shape information from shading is to construct programs that observe stylised properties of shading and associate those properties with shape primitives or their properties. In what follows, we show that, for two particular shading primitives, the shading field is robust to interreflections, its appearance can be predicted with a simple model, and these predictions can be used to find folds and grooves in images of real scenes.

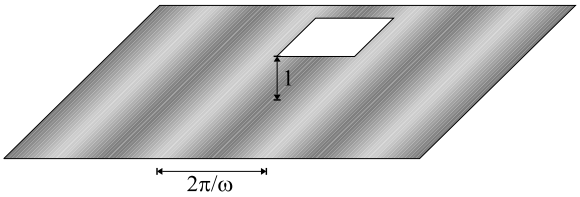


Figure 1: *The effects of distant surfaces can be estimated by considering a small surface patch illuminated by an infinite plane carrying a radiosity of* $a \sin \omega x$*; the patch is sufficiently small that its contribution to the plane's radiosity can be ignored. The radiosity on the patch in this configuration gives a guide as to where the shape information is in the patch's shading field.*

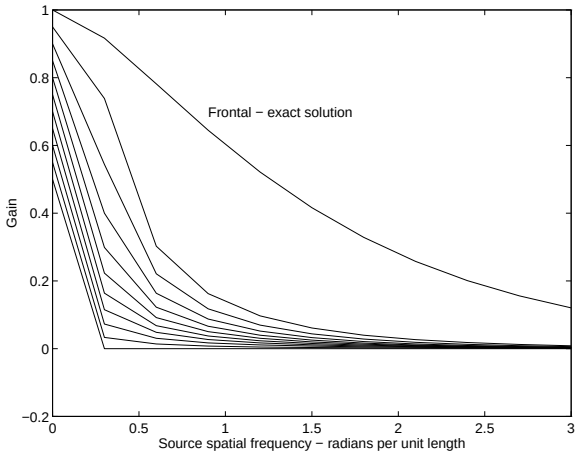


Figure 2: *A patch with a frontal view of an infinite plane which is a unit distance away and carries a radiosity* $\sin \omega x$ *can be shown to carry a radiosity of* $\rho \omega K_{-1}(\omega) \sin \omega x$*; we refer to* $\omega K_{-1}(\omega)$ *as the* gain *of the spatial frequency* ω*. This graph shows the gain for a frontal patch, and numerical estimates of the gain for patches at ten equal steps in slant angle with respect to the plane, from 0 to* $\pi/2$*, as a function of spatial frequency* on the plane*. Note that the gain falls extremely fast as a function of spatial frequency, meaning that large terms at high spatial frequencies must be regional effects, rather than the result of distant radiators.*

3 The effect of distant radiators

To be useful, any analysis of a shading field must be robust to the effects of surfaces that passively radiate to the object being studied, but may not be visible. Mutual illumination has a characteristic smoothing effect; shading effects that have a high spatial frequency (referred to a frame on the surface in question) and a high amplitude generally cannot come from distant surfaces. The effect is not well described in the literature, but is exploited in algorithms for determining the distribution of illumination given a set of surfaces, particularly [7, 6].

To illustrate this effect, consider a small surface patch illuminated by a distant, infinite plane as in Figure 1. The surface patch *is small enough that energy transfer from the patch to the plane is insignificant in determining shading distributions*. This is a good model for many important natural situations, such as objects in a room at some distance from the walls. Shading is scale-invariant, so we can fix the distance between the patch and the plane at one unit. Assume the plane carries a radiosity of $a \sin \omega x$ (there is no reason to ensure this is non-negative, since the radiosity of the patch is linear in the radiosity of the plane). Now assume that the patch carries a resulting radiosity $b\rho \sin \omega_r x$, and call b/a the gain at frequency ω, which we shall write as $G(\omega)$. By studying this gain, we can determine whether a signal is likely to have come from a surface region or from a distant radiator.

If the patch is parallel to the plane, the radiosity on the patch is exactly $\rho \omega K_{-1}(\omega) \sin \omega x$. Here K_{-1} is a modified Bessel function of the second kind and ρ is the reflectance of the patch. In this case, $G(\omega) = \omega K_{-1}(\omega)$, which falls very quickly with ω (as figure 2 shows) and approaches zero exponentially as $\omega \to \infty$.

If the patch is slanted with respect to the plane by σ, the effect is even more pronounced. Figure 2 also shows numerical estimates of the component of radiosity at spatial frequency $\omega_r = \omega \cos \sigma$ for various slant angles. Again, the gain function dies off to zero very quickly with increasing ω or σ. Thus, if one observes a high amplitude term at a high spatial frequency, *it is unlikely to have resulted from the effects of distant, passive radiators.* While very high frequency shading effects ("edges") are often identified as being due to changes in reflectance, [13, 10, 2] and low frequency shading effects may be due to distant radiators, there is a mid range of spatial frequencies that are largely unaffected by mutual illumination from distant surfaces.

The distinction between reflectance and shading effects raises a difficulty when one considers perspective—as an object recedes into the distance, the spatial frequencies of the corresponding image region rise. This means that effects that are classified as illumination changes in a nearby object become reflectance changes when the object is distant.

The scaling property of the gain function for distant radiators overcomes this difficulty. Because shading is scale invariant, the gain that occurs when the patch is d units away from the plane is the same as the gain when the

geometry is scaled by d:

$$G(d\omega, \sigma, 1) = G(\omega, \sigma, d)$$

The gain from a passive radiator at the focal point to the image of an object in the scene may then be shown to be a function of the product of the frequency of the illumination changes and the distance to that object. This means that image spatial frequencies can be used to assess whether the shading field contains local information or global information (from invisible, distant surfaces) if the focal length is known and some estimate of the slant is available. In particular, the support of filters designed to extract shape information from the shading field can be determined in an image frame, rather than a 3D frame.

This leads us to a meaning of the term "regional"—image scales in which no surfaces can be omitted in determining the effects of interreflections. Clearly, interreflections must be considered over larger scales than just a handful of pixels, but we have also shown that we do not need to consider the whole image.

4 Finding grooves

Although grooves and folds may be geometrically small, they may have large photometric consequences, so that any attempt to discover the large-scale shape of an object must first identify and discount these shading primitives. The shading across the groove or fold is a regional property under the assumption that the underlying surface is convex and has low curvature. While there are an infinite number of possible groove shapes, the similarities in images of grooves are more striking that the differences. Figures 3(b,c) show two images of grooves. The reader may wish to attempt to guess which of the groove shapes in Figure 3(a) these two grooves represent.[1] We present a simple model which we use to predict the radiosity across a groove. We then show that the model agrees extremely well with experiment, and can therefore be used to identify the presence of grooves in images.

4.1 Modelling grooves

We model the radiosity at any point on the groove as the sum of two components: the first due to distant radiators, which is uniform (because any spatial frequency high enough to be non-uniform over the support of the groove was suppressed by the low gain at high spatial frequencies); and the second due to a single point source, modeled as a source at infinity. This is a version of a model suggested by Koenderink [11] and also used by Langer *et al*.[15].

[1] Figure 3(b) has the triangular cross-section, while 3(c) has semicircular cross-section.

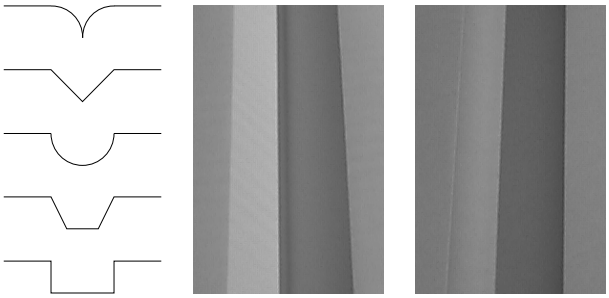


Figure 3: *(a) Cross-sections across the groove axis of paper models of grooves. Pictures of such models made from black, grey and white paper were taken under different lighting conditions in order to validate the predictions of the theory. (b,c) Typical images of experimental grooves. Note that although it is clear that these are images of grooves, it is difficult to determine the cross-section shape of the groove purely from the shading. The reader may wish to guess which of the grooves in (a) these images represent. The answer appears in the footnote.*

Because the groove has translational symmetry, we can model the "sky" (distant radiators) as an infinitely long half cylinder above the groove with its axis at the centre of the groove. We can then write the brightness at a point $\mathbf{u}$ in the groove as:

$$B(\mathbf{u}) = \frac{E_a}{2}[\sin(\theta_1 - \theta_u) - \sin(\theta_0 - \theta_u)] + E_p \cos(\theta_p - \theta_u)$$

where θ_1 and θ_0 are the polar angles of the edges of the unobscured sky (measured from the zenith), θ_u is the polar angle of the the normal at $\mathbf{u}$, θ_p is the polar angle of the point light source, and E_a and E_p are the brightnesses of the ambient and point light sources. This simple model allows us to predict the radiosity given a particular groove shape.

4.2 Confirming the model

In order to confirm our shading model, we wish to compare the theoretical predictions of radiosity profiles to those obtained from real images. A support vector machine [3] trained on theoretical data predicted from the model gives a low false negative rate on real images, indicating that the model accurately reflects reality.

Procedures. For the experimental data, we constructed a set of models of canonical grooves, in white, grey and black, with cross-section shapes shown in Figure 3(a). Photographs of these models were taken at different positions

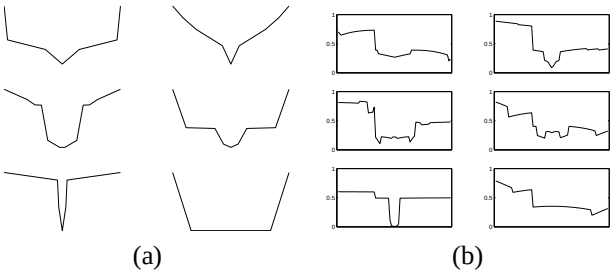


Figure 4: *(a) Typical random grooves used to predict groove radiosities. Each of the grooves are symmetric, twice as wide as they are deep, and have polygonal cross section. Shown here is the width of the groove with none of the surrounding plane—these cross sections may be considered as "chisel shapes". We use such a wide range of shapes because we are uninterested in the differences between different groove shapes. (b) The intensity cross section of the grooves in (a), as predicted by the model.*

in the room (giving variations in the ambient light) with different point light source directions (see Figures 3(b,c)). In order to control the direction of the point light source, we may either fix the position of the light source, and allow the shadow position to vary, or fix the shadow position, and move the point light source. Since we are more interested in localising the groove than the details of the illumination, we choose the latter.

We generated two experimental data sets, each of fifteen images. The first consisted of all combinations of shape and reflectance, taken with the same shadow position and same ambient light. The second set had varying shadow positions and ambient light.

For our theoretical data, we generated a large number of random symmetric polygonal grooves (see Figure 4(a)). After finding the position of the point light source that will cast a shadow boundary at the desired position, we used our model to generate the radiosity across the groove (see Figure 4(b)). Since the angle of incidence of the point light source varies, the average radiosity of each groove varies. To correct for this, each intensity cross-section is normalized to have zero mean and unit variance.

Negative examples were taken from randomly chosen lines in random images.

Comparison. A support vector machine is a learning machine that determines the hyperplane in the feature vector space which best separates the positive and negative data [3]. We trained an SVM using theoretical data for positive examples. None of the real groove cross-sections tested negative, no matter what their shadow position was, indicating that the test data (from experimental images) are consistent with the model data (theoretically generated). Thus, our simple groove model reflects reality accurately.

Since the theoretical predictions match the experimental data, we can use the theory to construct filters to find grooves in real images. Furthermore, since the shading model works for grooves, we expect it to work for folds.

5 Implementation and results

Grooves appear in images in a range of orientations, aspect ratios, scales and with different shadow positions and shapes. To find grooves of a particular orientation, the image is blurred along the desired direction, and the image intensity perpendicular to this direction, (across the groove) is tested using the SVM. Different aspect ratios can be searched by varying the width of the blur. Size differences are easily dealt with by scaling the image. We have already seen that the SVM is relatively insensitive to shadow position. Finally, we have seen that the shape of the groove has very little effect on the image.

Running the groove finder for one groove size and various orientations on the images in the top row of Figure 5 yields the images in the bottom row. Black pixels indicate the presence of a groove. In all images, the spine is clearly labeled as a groove. In (a) and (b), the shoulder blades are also marked as grooves. The detectors are selective as to size and orientation—the figure shows examples at a range of orientations tuned to one size and orientation, and therefore do not pick up all the grooves in an image.

The groove detector also marks some things as grooves which are not (for example, the left side of the back in (b)). It is unreasonable to expect template matching to yield a zero false positive rate. At a higher level, when constructing a representation of the object, we can take into account other evidence (such as edge boundaries) and reject many of the false positives.

Folds can be detected using the same strategy. We generate random symmetric polygonal folds, and then use the shading model to predict the radiosity over a fold. An SVM based on the predicted images is generated, and is then run on real images, giving extremely good results, as may be seen in Figure 6.

6 Discussion

Folds and grooves are small deformations on surfaces that produce large photometric effects. The pattern of shading on a fold or a groove is predicted well by a simple vignetting model, and shows significant consistency in the shading over different shapes. This means that folds and grooves can be marked effectively by filters predicted

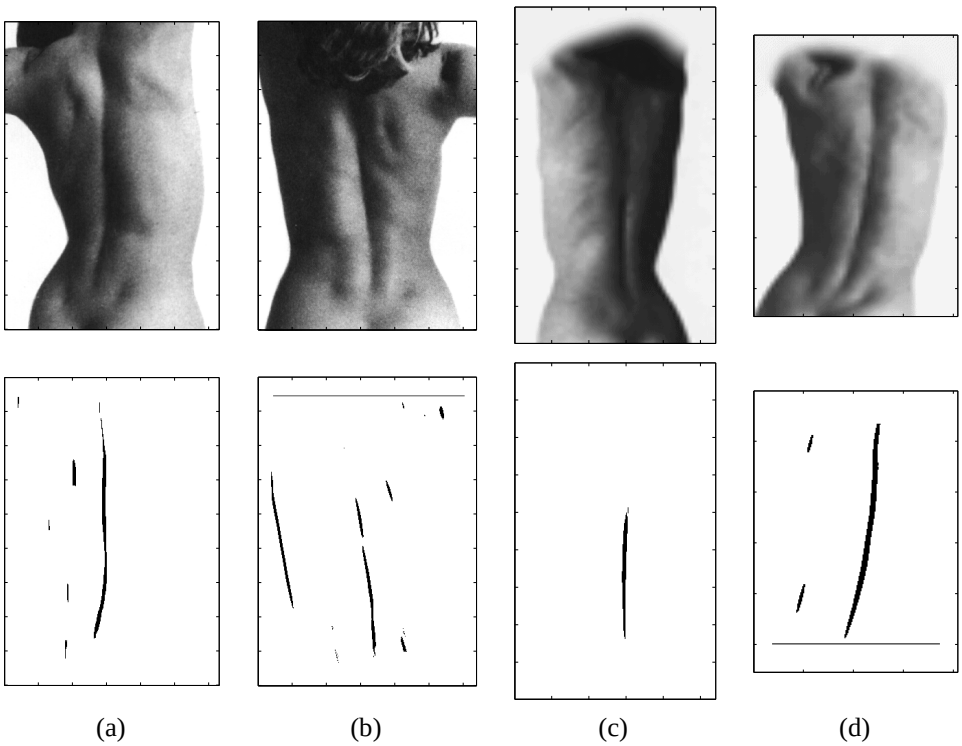


Figure 5: *Top row: Original images. While the groove caused by the spine is geometrically small, it has significant photometric consequences, affecting the shading for as much as a third of the width of the back. It is for this reason that we seek a representation which combines geometrical and photometric primitives.*
Bottom row: Groove detector output (black pixels indicate the presence of a groove). The detector has been trained on theoretical predictions of groove shape. It is highly selective in terms of size and orientation. In each case, the detector is looking for grooves that are 40 pixels wide. The ticks are every 50 pixels. (a) The predominant groove is vertical, so we look for grooves at $0°$*. Both the spine and the left should blade are identified, but the kidneys are not marked, due to their different orientation. (b) Groove angle of* $15°$ *marks the spine and right shoulder blade. The edge of the back is also marked—this false positive will need to be rejected, taking into account other evidence. (c) Because the illumination is from the left, we look for grooves at* $180°$*. (d) Groove angle of* $-15°$*. The horizontal lines in this image and in (b) are artifacts of the process of blurring along the groove.*

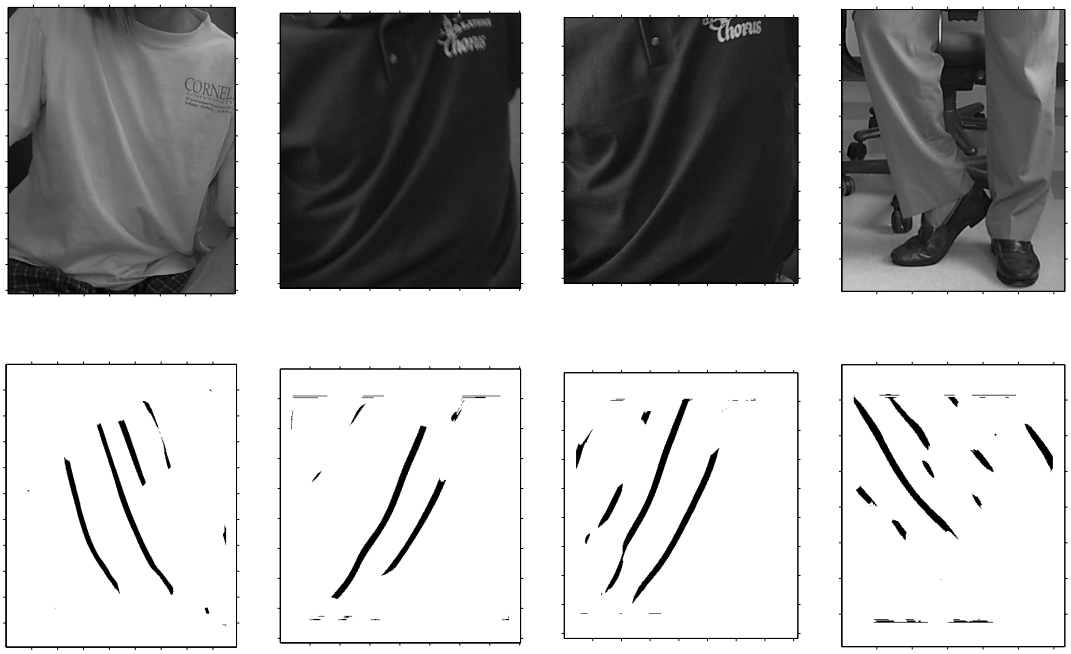
from a simple shading model without attempting to infer shape or the position of the illuminant. Furthermore, this approach is robust to interreflections, because it looks at regional properties and at spatial frequencies that are largely unaffected by distant radiators.

Folds and grooves are best identified as "dictionary entries", where a characteristic pattern in shading evokes a stylised shape percept, rather than by a dense depth map, which may be difficult to obtain. This approach sees shading analysis as a collection of matching activities, where primitives are identified on top of each other. Consider the image of a clothed limb—once the effects of the folds in the cloth have been discounted, we can decide whether the remaining shading pattern is consistent with that of a cylinder. Furthermore, by examining the collection of folds, we may be able to extract some information on the configuration of the body underneath the clothes.

This formalism has a number of attractions. It does not attempt to extract redundant or ill-conditioned shape information; it can be made robust to interreflections; and shading models can be used to predict likely primitives. The central issue for further work is finding more shading primitives, ideally with the aid of simplified shading models, and using these primitives in recognition.

Acknowledgements

We are grateful for discussions on shading with Margaret Fleck, Jitendra Malik, Jean Ponce and Andrew Zisserman, and for Michael Langer's extensive comments on the first draft of this paper. This research was



(a) (b) (c) (d)

Figure 6: *Top row: Original images. While of a much smaller scale than the shape of the body, folds are very significant photometrically. Note the consistency in fold spacing, and the fact that the folds in the cloth contain valuable information about the configuration of the body underneath.*
Bottom row: Fold detector output (black pixels indicate the presence of a fold). The detector was trained on theoretical predictions of the shading pattern on folds. It is highly selective for size and orientation—in each case, it is looking for folds that are 46 pixels wide. The ticks are every 50 pixels. (a) The predominant folds at $20°$ *extracted. (b)* $150°$. *(c)* $160°$. *(d)* $40°$. *The false positives appearing in this image will need to be rejected at a higher level. The horizontal lines at the top and bottom of (b), (c) and (d) are artifacts of the process of blurring along the fold.*

supported by an NSERC PGS B fellowship, an NSF NYI grant and an NSF Digital Library award IRI-9411334.

References

[1] P. Belhumeur, D.J.Kriegman, and A.L.Yuille. The bas-relief ambiguity. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 1060–6, 1997.

[2] A. Blake. Boundary conditions for lightness computation in mondrian world. *Computer Vision, Graphics and Image Processing*, 32:314–327, 1985.

[3] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–97, 1995.

[4] D. Forsyth and A. Zisserman. Mutual illumination. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 466–473, 1989.

[5] D. Forsyth and A. Zisserman. Reflections on shading. *IEEE T. Pattern Analysis and Machine Intelligence*, 13(7):671–679, 1991.

[6] R. Gershbein, P. Schröder, and P. Hanrahan. Textures and radiosity: Controlling emission and reflection with texture maps. In *SIGGRAPH-94*, pages 51–58, 1994.

[7] P. Hanrahan, D. Salzman, and L. Aupperle. A rapid hierarchical radiosity algorithm. In *SIGGRAPH-91*, pages 197–206, 1991.

[8] B. Horn. *Shape from Shading: a Method for Obtaining the Shape of a Smooth Opaque Object from One View*. PhD thesis, MIT Department of Electrical Engineering, 1970.

[9] B. Horn and M. Brooks. *Shape from shading*. MIT Press, 1989.

[10] B. K. P. Horn. Determining lightness from an image. *Computer Vision, Graphics and Image Processing*, 3:277–299, 1974.

[11] J.J.Koenderink. Vignetting and reflexes. Shading Course Notes.

[12] J. Koenderink and A. V. Doorn. Geometrical modes as a method to treat diffuse interreflections in radiometry. *J. Opt. Soc. Am.*, 73(6):843–850, 1983.

[13] E. Land and J. McCann. Lightness and retinex theory. *J. Opt. Soc. Am.*, 61(1):1–11, 1971.

[14] M. Langer and S. Zucker. Diffuse shading, visibility fields and the geometry of ambient light. In *Proceedings Fourth International Conference on Computer Vision*, pages 138–147, 1993.

[15] M. Langer and S. Zucker. Shape-from-shading on a cloudy day. *J. Opt. Soc. America A*, 11(2):467–478, 1994.

[16] S. Nayar, K. Ikeuchi, and T. Kanade. Shape from interreflections. *Int. J. Computer Vision*, 6(3):173–195, August 1991.

[17] A. Parker and C. Christou. Virtual realism and virtual reality: a psychological perspective. In *Simulated and virtual realities: elements of perception*. Taylor and Francis, 1995.

[18] F. Sillion and C. Puech. *Radiosity and Global Illumination*. Morgan Kaufmann, 1994.