

A cue to shading: elongations near intensity maxima

Daria Gipsman*
School of Computer Science
McGill University
Montreal, Quebec H3A 2A7 Canada

Michael S. Langer†
School of Computer Science
McGill University
Montreal, Quebec H3A 2A7 Canada

Abstract

The human visual system is often able to recognize shading patterns and to discriminate them from surface reflectance patterns. To understand how this ability is possible, we investigate what makes shading patterns special. We study a statistical property of shading patterns, namely that they tend to be more elongated near intensity maxima. Second-order derivatives of shading and of surface height are compared, and it is shown that intensities typically have an elongated structure relative to surface heights, and this elongation is more extreme near intensity maxima. This elongation property is formalized in terms of a skewness statistic on the aspect ratios of iso-intensity vs. iso-height curves. Keywords: shading, natural images, texture.

1 Introduction

What causes an object to appear darker or lighter? There are two main causes: the albedo, and the illumination. The former is the physical property of an object, and the latter - commonly called shading - is related to the combination of object geometry and light source properties.

Shading is one of the fundamental cues to 3D shape of visible surface, along with stereo, texture, and motion. Most previous shading analysis has concerned shape perception - both in computer vision and in human vision [2, 12, 14, 15]. In this paper we don't address the shape from shading problem. Rather, we address a problem that is arguably more fundamental. How does a visual system infer that it is looking at shading in the first place, as opposed to variations in reflectance? To be concrete, if a smooth matte terrain surface such as a wrinkled shirt is illuminated by a collimated light source, the luminance or shading is a smooth function. But why does this smooth luminance function appear to be the result of an illuminated 3D shape? Why does it not instead appear to be a flat surface with smooth reflectance? What properties of shading are used to distinguish smooth shading patterns from smooth reflectance patterns?

To illustrate, Figure 1 (a) depicts shading image of a bumpy terrain surface illuminated from slightly above the line of sight.

Fig. 1 (b) displays an image having the same power spectrum as the one in (a), but with randomized phase. The image in (b) does not give rise to percepts of 3D surface shape, but rather appears as blurred texture.

Since the two images have the same power spectrum, they must have the same autocorrelation function [8], which captures the second order image statistics. Thus, the statistical difference between the image in Fig. 1 (a) and (b) must lie in the higher-order statistics.

Intuitively, one prominent visual difference between the images in Fig. 1 (a) and (b) is that (a) is composed of elongated structures. In other words, the iso-intensity curves in (a) tend to be long and relatively straight, as opposed to the spotting structure in (b). This statistical elongation property has not, to our knowledge, been considered before as a cue to shading. Our goal in the research we present here is to explain and quantify this elongation property.

2 Background

2.1 Shape from shading

Many studies of shading have addressed the shape-from-shading problem: given a normalized image $I(x,y)$ and a model

$$I(x,y) = \frac{pL_x + qL_y + L_z}{\sqrt{p^2 + q^2 + 1}}, \quad (1)$$

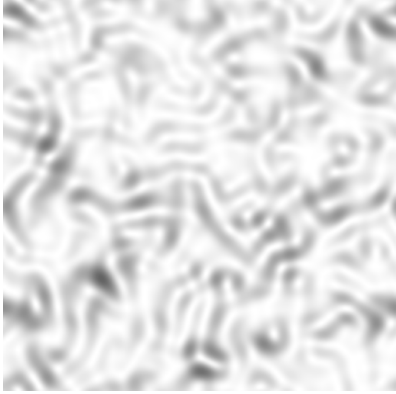
where $p = \frac{\partial}{\partial x}Z$, $q = \frac{\partial}{\partial y}Z$, compute a function $Z(x,y)$ satisfying the model and given image. This problem assumes that the surface reflectance is constant and that the intensity at (x,y) is entirely determined by the surface normal at that point, relative to the collimated light source direction (L_x, L_y, L_z) .

One idea for discriminating shading from non-shading patterns might be to attempt to explain a pattern with a shading model and, if this fails, attribute the model to non-shading. This raises the question of the existence of a solution $Z(x,y)$ for a given intensity image $I(x,y)$.

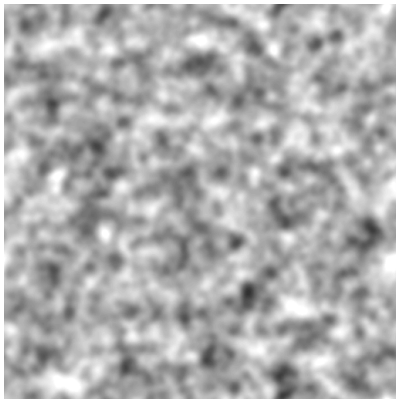
First, one can always find a shading model that explains a given intensity pattern. As Pentland observed [20], if the surface is near planar and the light source is near grazing incidence, then a linear shading model holds. For linear shape from shading, a solution always exists.

* email: dgipsm@sympatico.ca

† email: langer@cim.mcgill.ca



(a)



(b)

Figure 1. (a) Synthetic shaded surface. (b) Non-shading image having the same power spectrum as the image on (a).

Many shading patterns do not satisfy the linear models, however. For example, when a surface is illuminated by collimated light along the line of sight ("light source coordinates"), the shading model reduces to

$$I(x, y) = \frac{1}{\sqrt{p^2 + q^2 + 1}}. \quad (2)$$

Here the linear model does not apply - indeed the linear term of shading vanishes in this case and only second order and higher terms are present. This case has been studied extensively in the shading analysis literature. It is convenient because there are no shadows, and also the intensity gradient is in the same direction as the depth gradient. For example, Bruss addressed this case and presented the first uniqueness results in SFS [1], followed by Oliensis in [18, 19].

The issue of existence was addressed by Horn, Szelinski and Yuille [10]. They analyzed patterns that cannot result from shading on a smooth surface. For instance, an isolated dark region on a background of maximal intensity is proved to be an unfeasible shading pattern.

Extremal, or singular, points of shading play crucial role in many shape from shading techniques. In [1, 19], surface solutions are developed in the neighborhood of singular points of Z . Koenderink and van Doorn [13] studied intensity critical points while playing special attention to parabolic points. (Parabolic points lie on curves that separate elliptic from hyperbolic points.) Yuille [24], following an earlier work of Koenderink and van Doorn [11], demonstrated that extremal points of intensity often lie on parabolic curves. Weinshall [22] used geometrical properties of isophotes near shading maxima to locally classify shape.

2.2 Shading versus paint

The problem of telling "shading from paint" has been posed by Freeman and Viola in [3]. In the context of their work, the authors regard "paint" as any kind of marking made on a flat surface, while shape is regarded as a 3D object of constant reflectance. They employ Bayesian approach to find the most probable shape explanation while assuming Pentland's linear shading model [20].

In our research, we concentrate on the case where the light source is overhead. In this case, the linear model cannot be used since there is no first order component. The difference between the two models is significant. If a linear shading model is used, then the shape from shading problem always has a solution: one simply integrates intensities along parallel lines in the direction of a conjectured light source direction [38]. The linear model only holds in certain situations, though, namely when the surface is near planar and the light source is near grazing incidence.

In other work, Freeman et al drop the linear shading model assumption, and instead rely on probabilistic learning. Freeman and Bell in [7] attempt to recognize the cause of intensity variations by a learning-based technique. The problem addressed in [5] is extended by allowing an image to contain both shading and reflectance changes. The cause of intensity variation is assumed to be unique for every set of image location, scale and orientation. The training process involves building a steerable pyramid, every coefficient in which corresponds to the above-mentioned set of position, scale, and orientation. Every coefficient is then assigned a probability density function for both classes (i.e., shading and reflectance). This probability function is learned.

In a later research, Freeman, Pasztor and Carmichael [6] use a Markov network together with the shading/paint component patches to be recovered. The patches are estimated by maximizing their posterior probability given the image patches.

Others have used multiple images either under different lighting [23] or multiple color channels [7] under a single lighting condition. These approaches are peripheral to our problem, since we are trying to address the single grey level condition only.

3 Critical point analysis

Let's now turn to the original results of this paper. Recall Fig. 1 and the observation that a surface illuminated along the line of sight tends to produce elongated intensity structures.

When decomposing an image into intensity levels, we notice that the elongation effect is more prominent near the global maxima of intensity. To illustrate this, in Figure 2(a) the points of 20%

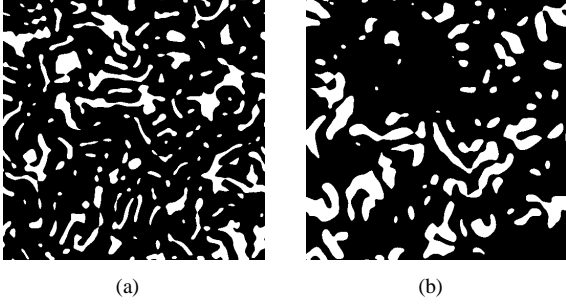


Figure 2. (a) 20% highest intensity and (b) 20% lowest intensity pixels of the image in Figure 1(a).

highest intensity are depicted in white, and the 20% lowest intensity points are displayed in Figure 2(b). The regions around global intensity maxima and minima bear significant differences. First, there are more islands of high intensity than low intensity. Since the number of pixels in the two sets is the same, the average areas in (a) are smaller than in (b). Second, the high intensity islands are more elongated compared to the low intensity ones. We will concentrate on the areas near intensity global maxima, based on the observed effect of elongated structures being most prominent there. (Properties of maxima are also important as previous work has shown maxima can be used in shape reconstruction.)

From the definition of I , it's easy to show that in case of the light source in the viewer direction, an intensity global maximum corresponds to a critical point of the surface height function Z , namely a local maximum, a local minimum, or a saddle point. The regions near critical points of Z are approximately elliptic (local maximum or local minimum of Z) or hyperbolic (saddle point of Z). The regions around global maxima of intensity I are approximately elliptic. In a small neighborhood of a critical point of Z , elliptic or hyperbolic regions of the surfaces Z or I yield elliptic or hyperbolic iso- Z and iso- I contours.

Up to a rotation factor, an iso- Z curve can be approximately represented by an equation $ax^2 + by^2 = c$. Note that the equation is valid for both the elliptic and the hyperbolic surface types. We'll address the constants a and b as second-order, or quadratic, coefficients of a contour.

We first show that an iso- Z curve with the quadratic coefficients of a and b yields an elliptic iso- I contour with the coefficients of $2a^2$ and $2b^2$. *The squaring of the coefficients squares their ratio, which in turn causes an elongation of the iso- I contours in the neighborhood of the maxima.*

As mentioned above, we study the case of the shading in light source coordinates, that is, the collimated source is along the line of sight (Eq. 2). Let's consider a second-order Taylor expansion of the height function around a critical point. Without loss of generality, the mixed second-order term can be annulled by a suitable rotation. Therefore, a surface Z can be locally approximated by a polynomial

$$Z(x, y) \approx ax^2 + by^2. \quad (3)$$

Then its partial derivatives would be given by

$$p = Z_x(x, y) = 2ax, \quad q = Z_y(x, y) = 2by.$$

Therefore, from Eq. 2 we obtain the following expression for the image:

$$I(x, y) = \frac{1}{\sqrt{1 + 4a^2x^2 + 4b^2y^2}}.$$

If we now compute its partial derivatives, and substitute into the Taylor series expansion at $(x, y) = (0, 0)$, then a second order approximation is:

$$I(x, y) \approx 1 - 2a^2x^2 - 2b^2y^2. \quad (4)$$

Observe that the a , b coefficients are squared in Equation 4, whereas they are not squared in Equation 3. *This difference between I and Z is the main reason why shading patterns are elongation near intensity maxima*, as we now elaborate.

A surface iso-height contour with the second-order coefficients of a and b around a critical point of Z is projected to an elliptic iso-intensity contour with the coefficients equal to $2a^2$ and $2b^2$. The surface iso-height contour is elliptic around minimum or maximum (i.e. a and b have the same sign) and is hyperbolic around a saddle point (a , b have different signs). An iso-intensity contour is always elliptical, though, since a critical point of Z always produces a maximum in intensity I .

We define an aspect ratio α of either elliptic or hyperbolic contour with the coefficients a , b as follows:

$$\alpha = \max\left(\frac{|b|}{|a|}, \frac{|a|}{|b|}\right).$$

Obviously, it always holds that $\alpha \geq 1$. In case the aspect ratio of iso- Z contour equals 1, α of the corresponding iso-intensity curve is also 1. However, when an iso- Z aspect ratio α_z is greater than 1, the corresponding iso- I aspect ratio $\alpha_i = \alpha_z^2 > \alpha_z$. This aspect ratio squaring causes the elongation of iso-intensity contours relative to iso-height curves. This complies with the observations of Fig. 1.

4 Experiments

In this section, we present experiments that explore properties of image statistics near intensity maxima. Our results are reminiscent of the recent work of Motoyoshi et al [17], who observed and explained correlation between third-order statistics (skew) of image intensities and surface glossiness. In our paper, we do not investigate glossiness, but rather we consider matte (Lambertian) surfaces only. We consider various aspects of skew that arises from shading near intensity maxima.

Recall the shading image shown in Fig. 1a. In Figure 3, we present histograms of the surface heights (Fig. 3(a)) and of the image intensities (Fig. 3(b)). We can see that the image intensities of a shading image are quite skewed, even though the heights are not skewed.

We next examine histograms of the second order coefficients mentioned earlier, namely the distribution of second-order coefficients of images and surfaces in the neighborhood of intensity maxima.

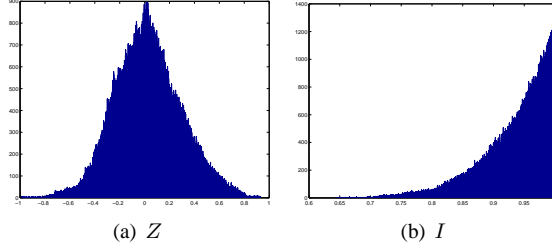


Figure 3. Histogram of a surface function Z (a) and of a shading image I (b).

We generate a large number of synthetic surfaces $Z(x, y)$ and images $I(x, y)$ of size 512×512 . Each surface Z is built using spectral synthesis [33], namely by adding up a set of 2D sinusoids. The amplitudes $\tilde{Z}(f_x, f_y)$ are taken from the two-dimensional Gaussian density function

$$\rho(f_x, f_y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}\|(f_x, f_y)\|^2}$$

where f_x, f_y are spatial frequencies and σ^2 is the variance. Phases are picked at random from the interval $[-\pi, \pi]$, subject to conjugacy constraint for Fourier transforms of real images. A surface is then constructed by applying the inverse Fourier transform. The only parameter characterizing the surface is therefore the standard deviation of the Gaussian. A surface generated by σ -Gaussian is denoted by Z_σ . Intensity functions are then rendered from the surfaces using Eq. 2. An image rendered from Z_σ surface will be denoted by I_σ .

In order to investigate the statistical differences between the images and the surfaces, we compare second-order coefficients. Namely, we estimate the partial derivatives numerically, subsequently solving

$$\begin{aligned} a(x \cos \theta + y \sin \theta)^2 + b(y \cos \theta - x \sin \theta)^2 \\ = Ax^2 + By^2 + Cxy, \end{aligned}$$

to compute the second-order coefficients a, b . In the formula above, θ represents the direction of elongation.

4.1 Skewness of second order coefficients

The most natural comparison would be to analyze the distributions of the aspect ratio α , formally defined as

$$\alpha = \max\left(\frac{|b|}{|a|}, \frac{|a|}{|b|}\right).$$

However, its distribution appears to be not informative, since the ratio of the highest and the lowest value of α is too high. Instead, we introduce the modified second-order coefficients \tilde{a}, \tilde{b} , defined as the minimum and the maximum (in absolute value) between a and b , respectively, i.e.

$$\tilde{a} \equiv \begin{cases} a, & |a| < |b| \\ b, & \text{otherwise} \end{cases}, \quad \tilde{b} \equiv \begin{cases} b, & |b| > |a| \\ a, & \text{otherwise} \end{cases}.$$

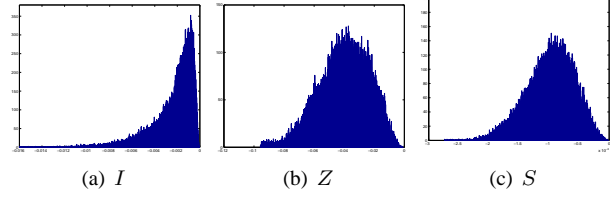


Figure 4. Histograms of the negative values of \tilde{b} for (a) rendered image I , (b) surface height function Z , (c) scrambled image S . Every figure displays the distribution measured of the highest-value 20% of the surface/image.

We then compare the distributions of \tilde{b} for the surfaces Z and the rendered images I with the hope that the differences between them will be representative enough.

Figures 4(a) and 4(b) depict the distribution of \tilde{b} for a rendered image I and a corresponding surface Z , both generated with same value of σ (namely $\sigma = 6$). *Only negative \tilde{b} values are displayed, since the regions around the intensity maxima are concave and hence have negative second-order coefficients.* Moreover, as our theoretical analysis deals with the second-order expansions around critical points of I , we only present the distributions of the highest 20% points (which would correspond to the areas around global intensity maxima).

The distributions of \tilde{b} coefficient for the surfaces and the images bear significant differences. The most salient observation is that the negative part of the \tilde{b} -distribution for the images is skewed comparing to that of the surfaces. We capture this difference quantitatively by the *skewness coefficient*, defined as

$$\gamma = \frac{E(X - \mu)^3}{st.d.^3}.$$

Table 1 displays the skewness coefficient of \tilde{b} of the surfaces Z and the images I . (A third skewness value labelled S will be explained below.) The skewness coefficient is estimated by running each experiment 30 times and taking the average. To verify the reliability of the obtained estimation, we also build the 95% confidence intervals; they are presented in the tables along with the average estimates.

Table 1. Average skewness γ and its 95% confidence interval of the negative part of \tilde{b} distribution for the surface heights Z , the shaded images I , and the scrambled images S .

	Average value	Confidence interval
Z	-0.39	[-0.44,-0.34]
I	-1.95	[-2.01,-1.88]
S	-0.45	[-0.47,-0.43]

The skewness coefficient of \tilde{b} is much higher (in absolute value) for the images I than for the surfaces Z . The confidence intervals are relatively small given the number of samples, which indicates the consistency of the results. These data suggest that the statistical differences in the second order order pattern near intensity maxima could indeed be used as indicator of shading.

As an aside, the reader may be surprised that \tilde{b} values for the surfaces Z have non-zero skewness. Recall however, that \tilde{b} is defined to be the maximum between the two quadratic coefficients, and that we are considering here only the negative \tilde{b} values. The key point here is not that I or Z is skewed, but rather that the skewness of I is much greater than that of Z .

To elaborate, let's consider how the shape of the distributions (i.e. longer tails in case of images) derives from the theoretical result of coefficients squaring of Section 3. Indeed, we note that the negative parts of \tilde{b} -distributions of the images I have longer "tails" comparing to those of the surfaces Z . Let X be a random variable distributed around a peak $-\mu$ in the negative part of the axis. If we assume that the number of samples in the intervals $[-\mu, 0]$ and $[-2\mu, -\mu]$ is the same (as is roughly the case for the surfaces Z), then the squared variable¹ $-X^2$ (corresponding to the \tilde{b} coefficient of I) will contain the same number of samples inside the intervals $[-\mu^2, 0]$ and $[-(2\mu^2), -\mu^2]$. The latter interval is three times larger, hence longer "tails".

4.2 Control experiment

The \tilde{b} -distributions of the surfaces and the images are not only different in shape, but they also have different scales. To try to better control for this "apple" vs. "orange" comparison, we return to the example of Fig. 1. Rather than comparing the statistics of the shading image intensities to those of the depth map, we compare statistics of shading to those of "phase scrambled" images having the same power spectrum as the shading images. These images which we denote S are built in the Fourier domain, such that the Fourier amplitudes of S are averaged over those of several I , and its frequency phases are random values. An example of such a scrambled image was shown in Fig. 1b.

We compare the distributions of the negative \tilde{b} of the images I with those of the scrambled images S in Fig. 4 (a) and (c). The distributions are different in shape, even though the rendered and the scrambled images have similar power spectrum i.e. second order statistics.

Table 1 displays the skewness coefficient γ of the images I which is much greater (in absolute value) than that of the scrambled images S . This confirms the observation about the distributions of the negative values of \tilde{b} being different in shape. Also notice from Figure 4 and Table 1 that the \tilde{b} -distributions and skewness coefficients of S and Z are similar. Thus the scrambled images S appear to be *statistically* similar to the surfaces Z , but different from the rendered images I .

¹The factor of 2 is omitted in the explanation since it only affects the scale of the distribution.

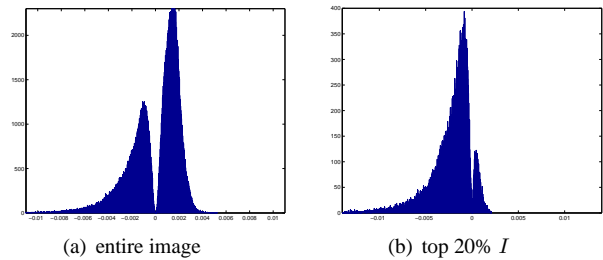


Figure 5. Histograms of the distributions of \tilde{b} for all pixels in image I and for the top 20% intensities.

5 Future Work

In this paper we presented statistical parameters that could serve as cues to shading. We also introduced the theoretical derivations supporting the statistical observations. As it was emphasized throughout the paper, our analysis was restricted to the areas of shading maxima, thus limiting the statistical observations to negative values of quadratic coefficients \tilde{b} . However, the positive part of b -distribution might be equally relevant. In Figure 5 we compare the distributions of \tilde{b} for an entire image (all the pixels) with the brightest 20% of the pixels. These data are not so meaningful, since our derivation earlier in the paper only considered the intensity maxima. Nonetheless, it is possibly interesting that one finds skewness in other statistics as well. Understanding this better is a subject of future research.

Another topic of future work is to consider an alternative illumination model, or other surface reflectances. In terms of illumination model, we assumed a distantly-illuminated Lambertian surface. Yet, extending our results to other models might be interesting. For instance, the analysis of intensity maxima in case of uniform hemispheric light source is straightforward. According to [9], intensity $I(x, y)$ is given by

$$I(x, y) = 1 + \frac{1}{\sqrt{p^2 + q^2 + 1}}$$

where p and q denote the partial derivatives of the surface height function $Z(x, y)$. Approximating Z by the second-order polynomial

$$Z(x, y) \approx ax^2 + by^2.$$

will yield the following second-order expansion of $I(x, y)$ around the origin:

$$I(x, y) \approx 2 - 2a^2x^2 - 2b^2y^2.$$

As this case is almost identical to the case of illumination in the viewer direction, which we addressed earlier in the paper, all the results and conclusions remain valid. (Note: the model of [9] is not a general model because it does not consider cast shadows, but this model may be sufficient for intensity maxima. For a more general model shading under diffuse lighting, see [16].) Other reflectance models, e.g. Phong model [21] or mirror reflectance [4] should be considered as well.

References

- [1] A.R.Bruss, *The eikonal equation: some results applicable to computer vision*, J. Math. Phys., Vol. 23, No. 5, pp. 890-896, 1982.
- [2] H.H.Buelthoff, H.A.Mallot, *Integration of depth modules: stereo and shading*, Journal of the Optical Society of America, A, Vol. 5, No. 10, pp. 1749-1758, 1988.
- [3] W.T.Freeman, P.A.Viola, *Bayesian model of surface perception*, Adv. in Neural Information Processing Systems, Vol. 10, 1998.
- [4] R. W. Fleming, A. Torralba, E. H. Adelson *Specular reflections and the perception of shape*, Journal of Vision, Vol. 4, No. 9, pp. 798-820, 2004.
- [5] W.T.Freeman, E.C.Pasztor, O.T.Carmichael, *Learning low-level vision*, International Journal of Computer Vision, Vol. 40, No. 1, pp. 25-47, 2000.
- [6] W.T.Freeman, M.Bell, *Learning local evidence for shading and reflectance*, Int. Conf. on Comp. Vision, pp. 670-677, 2001.
- [7] W.T.Freeman, M.F.Tappen, E.H.Adelson, *Recovering intrinsic images from a single image*, Advances in Neural Information Processing Systems 15, 2003.
- [8] R.C.Gonzalez, R.E.Woods, *Digital image processing*, 2nd edition, Prentice Hall, 2001.
- [9] B.K.P.Horn, R.W.Sjoberg, *Calculating the reflectance map*, Applied Optics, Vol. 18, No. 11, pp. 1770-1779, 1979.
- [10] B.K.P.Horn, R.S.Szeliski, A.L.Yuille, *Impossible shaded images*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, No. 2, pp. 166-170, 1993.
- [11] J.J.Koenderink, A.J. van Doorn, *Photometric invariants related to solid shape*, Optica Acta, Vol. 27, No. 7, pp. 981-996, 1980.
- [12] J.J.Koenderink, A.J. van Doorn, A.Kappers, *Surface perception in pictures*, Perception and Psychophysics, Vol. 52, No. 5, pp. 487-496, 1992.
- [13] J.J.Koenderink and A.J. van Doorn, *Illuminance critical points on generic smooth surfaces*, Journal of the Optical Society of America, A, Vol. 10, No. 5, pp. 844-854, 1993.
- [14] M.S.Langer, H.H.Buelthoff, *Depth discrimination from shading under diffuse lighting*, Perception, Vol. 29, No. 6, pp. 649-660, 2000.
- [15] M.S.Langer, H.H.Buelthoff, *A prior for global convexity in local shape from shading*, Perception, Vol. 30, No. 4, pp. 403-410, 2001.
- [16] M.S.Langer, S.W.Zucker, *Shape from shading on a cloudy day*, Journal of the Optical Society of America, Vol. 11, pp. 467-478, 1994.
- [17] I.Motoyoshi, S.Nishida, L.Sharan, E.H.Adelson, *Image statistics and the perception of surface qualities*, Nature, Vol. 447, pp. 206-209, 2007.
- [18] J.Oliensis, *Uniqueness in shape from shading*, International Journal of Computer Vision, Vol. 6, No. 2, pp. 75-104, 1991.
- [19] J.Oliensis, *Shape from shading as a partially well-constrained problem*, CVGIP: Image Understanding, Vol. 54, No. 2, September, pp. 163-183, 1991.
- [20] A.P.Pentland, *Shape information from shading: A theory about human perception*, Spatial Vision, Vol. 4, No. 2, pp. 165-182, 1989.
- [21] Bui Tuong-Phong, *Illumination for computer-generated images*, Communications of the ACM, Vol. 18, No. 6, pp. 311-317, 1975.
- [22] D.Weinshall, *Local shape approximation from shading*, Journal of Mathematical Imaging and Vision, Vol. 4, 119-138, 1994.
- [23] Y.Weiss, *Deriving intrinsic images from image sequences*, Proceedings of the International Conference on Computer Vision, Vancouver, Canada, 2001.
- [24] A.L.Yuille, *Zero crossings on lines of curvature*, Computer Vision, Graphics and Image Processing, Vol. 45, pp. 68-87, 1989.