

# Matching Photometric Observation Vectors with Shadows and Variable Albedo

Helena Cristina G. Leitão  
Inst. of Computing, Fluminense Fed. Univ.  
R. Passo da Pátria, 156, Niterói, RJ, Brazil  
hcgl@ic.uff.br

Rafael Felipe V. Saracchini  
Jorge Stolfi  
Inst. of Computing, State Univ. of Campinas  
Cx. P. 6176, 13083-970, Campinas, SP, Brazil  
{ra069320, stolfi}@ic.unicamp.br

## Abstract

*We describe a procedure to solve the basic problem of Variable Lighting Photometric Stereo – namely, recovering the normal directions and intrinsic albedos at all visible points of an opaque object, by analyzing three or more photos of the same taken with different illuminations. We follow the gauge-based approach, where the lighting conditions and light scattering properties of the surface are given indirectly by photographing a gauge object with known shape and albedo, under the same lighting conditions. Unlike previous solutions, our method yields reliable results even when some of the images contain cast shadows, penumbras, highlights, or inter-object lighting, at a cost. The cost of inner loop grows quadratically, (rather than exponentially) with the number  $m$  of input images. Usable approximations can be obtained in  $m \log m$  time.*

## 1. Introduction

### 1.1. Variable-lighting photometric stereo

In *variable-lighting photometric stereo* (VLPS), the goal is to determine the 3D geometry of a scene from a list of  $m \geq 3$  monochromatic digital photos  $S_1, \dots, S_m$ , all taken with different lighting conditions but with the same pose and viewpoint — such as those shown in figure 1. As R. J. Woodham pointed out in 1980 [14], by analyzing the  $m$  pixel intensities  $S_i[p]$  at any image point  $p$ , one can recover the unit vector  $\vec{s}[p]$  that is perpendicular to surface element which is visible at  $p$ . The third dimension (depth) of the scene can then be recovered by integration of this data.

### 1.2. Shadows, highlights and other anomalies

A major nuisance in photometric stereo are variations in the lighting conditions (such as cast shadows, penumbras,

and secondary scene-scattered light), and variations in the optical properties of the scene (such as highlights due to glossy finish). These effects obscure the similarity between the scene and gauge pixels with same normal, and therefore make the normal recovery problem substantially more difficult. Moreover, they also break all known acceleration methods for this problem, such as bucketing [5].

### 1.3. Previous work

The principles of photometric stereo must have been known at least since the establishment of photography and the development of photometry over a hundred years ago. Early articles assumed very simple settings: a smooth object with uniform color and Lambertian finish, illuminated by a single light source, distant and point-like (i.e. by a uniform and unidirectional light field). Their emphasis was on methods that could extract depth information from a single monochromatic image [9, 10, 15]

In 1980, Woodham [14] introduced the principles of VLPS proper — namely, the analysis of multiple images under different light fields — and the use of light gauges to obtain information on the lighting conditions. Several authors later extended the gauge method to more general BDRFs, including colored objects with glossy finishes [11, 7, 8]

Analysis of the gauge images provide the shading function that maps a surface normal to its the apparent intensity under each light. The most time-consuming part of all those algorithms is the inversion of these shading function, to recover the normal from the intensities. Woodham himself [15] introduced the use of a look-up table for this task. There are many papers concerned with accelerating the table look-up step [16, 7, 1, 17, 12, 5, 6, 17]. Other authors, such as Basri, Jacobs and Kemelmacher [3], replace the lookup table by an analytic model of the shading functions (a spherical harmonics series) that gets fitted to the gauge images.

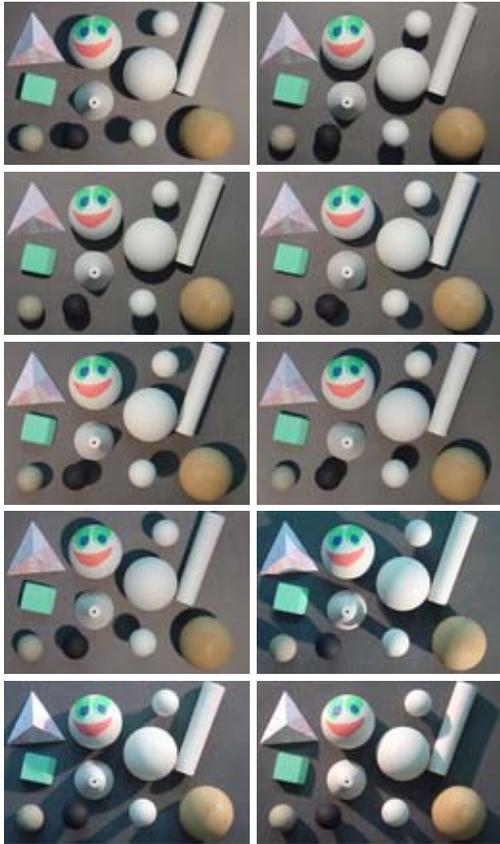


Figure 1. Images of a scene under various lighting conditions.

Most photometric stereo algorithms assume that the images are free from cast shadows, meaning that the scene consists of low-relief terrain or well-separated convex objects that are approximately Lambertian. Only recently have researchers considered the more general case when there are cast shadows, highlights, and other deviations from that ideal situation.

Barsky and Petrou [2] describe heuristics to detect shadows and highlights in a set of four or more images. The basic heuristic uses global intensity thresholds.

Chandraker, Agarwal, and Kriegman [4] recently proposed a VLPS algorithm for the case with  $m$  images of scenes lighted by  $n \geq k$  sources. They use combinatorial optimization techniques to identify pixels which are illuminated by each source in each image. Their method is limited to a relatively small number of point sources, so that the shadowing of each pixel can be represented by a string of  $m$  bits.

#### 1.4. Lighting and light scattering model

The most direct approach to VLPS requires detailed knowledge of the *bidirectional radiance distribution function* (BRDF) of the surface as well as of the *lighting conditions* of each image  $S_i$ . The BRDF of the scene's surface at point  $p$  of the image is a function  $\sigma[p](\vec{n}, \vec{u}, \vec{v})$  that gives the apparent brightness of the surface when oriented with normal  $\vec{n}$ , viewed from the direction  $\vec{v}$ , and illuminated with unidirectional light of unit intensity flowing in the direction  $\vec{u}$ . (Note that we include the geometric light spread factor  $\max\{0, -\vec{u} \cdot \vec{n}\}$  in the BRDF itself.)

The lighting conditions of each image can be modeled by a *light field*  $\Phi_i[p](\vec{u})$ , which gives the intensity of light flowing in the direction  $\vec{u}$  at the point  $p$  (which is meaningful only if  $p$  is on the scene's surface or in the empty regions outside all objects).

#### 1.5. Gauge-based VLPS

In the *gauge-based*(GB) variant of VLPS, the BRDF information is indirectly given by  $m$  images  $G_1, \dots, G_m$  of a *light gauge*, a sample object of known shape and color; where each gauge image  $G_i$  is taken under the same lighting conditions as the corresponding scene image  $S_i$ . See figure 2. (Depending on the application, it may be convenient to include the gauge object as part of the scene itself. In that case, each  $G_i$  will be just a sub-image of  $S_i$ .) In this pa-

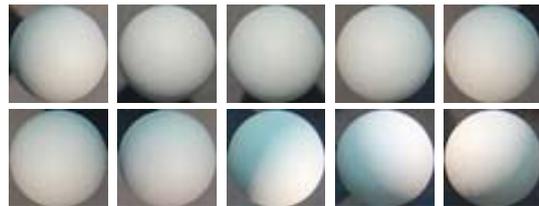


Figure 2. Gauge images for figure 1.

per, we assume that all images  $S_1, \dots, S_m$  have been geometrically corrected, trimmed, and aligned, so that each point  $p$  on their common domain  $\mathcal{S}$  corresponds to the same point on the scene's visible surface. The same condition is assumed for the gauge images  $G_1, \dots, G_m$ , whose common domain will be denoted by  $\mathcal{G}$ . We also assume a linear pixel value scale, so that sample values are directly proportional to physical light intensity. Finally, we assume that the surface normal vector  $\vec{g}[q]$  is known for every point  $q$  on the gauge's images.

In its basic form, GB VLPS is viable only if all visible scene and gauge surfaces have the same finish everywhere, except for variations in intrinsic color. That is, the BRDF  $\sigma[p]$  of the scene at each image point  $p$ , and the BRDF  $\gamma[q]$

of the gauge at any point  $q$ , must be multiples of some fixed BRDF  $\bar{\beta}$ :

$$\begin{aligned}\sigma[p](\vec{n}, \vec{u}, \vec{v}) &= \check{s}[p] \bar{\beta}(\vec{n}, \vec{u}, \vec{v}) \\ \gamma[p](\vec{n}, \vec{u}, \vec{v}) &= \check{g}[p] \bar{\beta}(\vec{n}, \vec{u}, \vec{v})\end{aligned}\quad (1)$$

The constant factors  $\check{s}[p]$  and  $\check{g}[q]$  in these formulas are the *intrinsic lightness* or *albedo* of the scene and gauge surface, respectively, at those points. Observe that the gauge's albedo  $\check{g}[q]$  and normal direction  $\vec{g}[q]$  must be known for all  $q \in \mathcal{G}$ . Typically, one uses a spherical gauge with uniform albedo, preferably white ( $\check{g}[q] = 1$  everywhere).

Another necessary condition for gauge-based VLPS is that the BRDF  $\bar{\beta}$  must be dominated by wide-angle scattering, with little mirror-like reflection or sharp glossy scattering. The standard example is the Lambertian BRDF,

$$\bar{\beta}(\vec{n}, \vec{u}, \vec{v}) = \max\{0, -\vec{u} \cdot \vec{n}\} \quad (2)$$

However, almost any BRDF  $\bar{\beta}$  will do, as long as it is not dominated by impulse-like components (sharp peaks or ridges). For simplicity, we will also assume that the images are taken under nearly parallel projection and illuminated by distant light sources; so that the viewing direction  $\vec{v}$  and the primary light flow is the same at every point of  $\mathcal{S}$  or  $\mathcal{G}$ .

Gauge-based VLPS can be extended to multichannel (e.g. RGB trichromatic) images, yielding a single normal map  $\vec{s}[p]$  but a different albedo map  $\check{s}_\lambda[p]$  for each spectral band  $\lambda$ . The latter provide the illumination-independent *intrinsic color* of the scene at each pixel. See figure 3.



Figure 3. Albedo map from figure 1.

## 1.6. Fundamental equations

The key idea of gauge-based VLPS is that the intensity of each point on a scene photo  $S_i$  or a gauge photo  $G_i$  can be analyzed into the product of two factors: the *intrinsic albedo*,  $\check{s}$  or  $\check{g}$  (that depends only on the surface's material and finish) and the *lighting factor*, (that depends only

on the index  $i$  and on the surface's slope,  $\vec{s}$  or  $\vec{g}$ ). Specifically,

$$\begin{aligned}S_i[p] &= \check{s}[p] L_i[p](\vec{s}[p]) \\ G_i[q] &= \check{g}[q] L_i[q](\vec{g}[q])\end{aligned}\quad (3)$$

Here, each  $L_i$  is the *shading function* implied by the lighting conditions at the point  $p$  and the basic BRDF  $\bar{\beta}$ . It maps each unit vector  $\vec{n}$  to the apparent lightness of a white surface perpendicular to  $\vec{n}$ , and is given by the formula

$$L_i[p](\vec{n}) = \int_{\mathbb{S}^2} \Phi_i[p](\vec{u}) \bar{\beta}(\vec{n}, \vec{u}, \vec{v}) d\vec{u} \quad (4)$$

Note that, in this model, the intrinsic albedos  $\check{g}$  and  $\check{s}$  and the normals  $\vec{s}$  and  $\vec{g}$  are distinct from each other and vary from point to point, but are same for all  $i$ .

## 1.7. GB VLPS with uniform lighting

If the lighting field  $\Phi_i[p](\vec{u})$  does not depend explicitly on the position on the surface  $p$ , but only on the light source's direction  $\vec{u}$ , then the shading function  $L_i[p](\vec{n})$  depends on the normal  $\vec{n}$  only. In that case, the recovery of the normal map  $\vec{s}[p]$  is essentially a solved problem [13]. Namely, it suffices to find a point  $q$  in the gauge images that reacts in the same way as  $p$  to changes in lighting directions — except possibly for a constant factor that depends on the albedos  $\check{s}[p]$  and  $\check{g}[q]$ . More precisely, we must find  $q \in \mathcal{G}$  such that the  $m$ -vectors

$$\begin{aligned}\mathbf{S}[p] &= (S_1[p], S_2[p], \dots, S_m[p]) \\ \mathbf{G}[q] &= (G_1[q], G_2[q], \dots, G_m[q])\end{aligned}\quad (5)$$

are multiples of each other. That is,

$$S_i[p] = \tau G_i[q] \quad \text{for all } i, \quad (6)$$

where  $\tau = \check{s}[p]/\check{g}[q]$ . The vectors  $\mathbf{S}[p]$  and  $\mathbf{G}[q]$  are called the *observation vectors* (OVs) of the points  $p$  and  $q$ .

Having located the matching gauge point  $q$ , we can recover the normal vector  $\vec{s}[p]$  and albedo  $\check{s}[p]$  of the scene at  $p$  by the formulas

$$\vec{s}[p] = \vec{g}[q] \quad \check{s}[p] = \frac{|\mathbf{S}(p)|}{|\mathbf{G}(q)|} \check{g}(q) \quad (7)$$

This method will fail if there are two points  $q', q''$  on the gauge images which have different normals (that is,  $\vec{g}[q'] \neq \vec{g}[q'']$ ) but collinear OVs (that is,  $\mathbf{G}[q'] = \alpha \mathbf{G}[q'']$  for some scalar  $\alpha$ ). To avoid this problem, the number of images  $m$  must be at least 3, and the light fields  $\Phi_1, \dots, \Phi_m$  must be sufficiently varied to break any such ambiguities. We will assume that this condition is satisfied in what follows.

If  $\vec{s}[p] = \vec{g}[q]$ , then the observation vector at  $p$  will be a multiple of the observation vector at  $q$ ; that is, where  $\tau$  is the relative albedo  $\check{s}[p]/\check{g}[q]$  of the scene at  $p$ .

## 2. VLPS with non-uniform lighting fields

### 2.1. Causes of non-uniform lighting

Unfortunately, uniform lighting is difficult to obtain, even for carefully chosen scenes and lighting positions, and is unlikely to occur in practice. One major source of non-uniformities are the *cast shadows*, which occur when some other part of the scene blocks some or all of the light that reaches  $p$  from some directions  $\vec{u}$ . Other sources are the *secondary light* that reaches  $p$  after being scattered by some other part of the scene, and highlights — sharp reflections due to polished surfaces or transparent overcoats (wax, varnish, paint, binder, etc.). These *anomalous lighting* effects may cause the apparent intensity  $S_i[p]$  to be lower or higher, respectively, than predicted by equation (6).

To achieve uniform lighting, the *primary light field*  $\Phi_i^*(r, \vec{u})$  — the component of the light flow  $\Phi_i[r](\vec{u})$  that is due to sources external to the scene — must depend only on the direction  $\vec{u}$ , and must be zero if there is a line with direction  $\vec{u}$  that hits the visible part of the scene’s surface twice. This constraint severely limits the geometry of the scene (which is basically restricted to terrains with bounded slope) and the placement of light sources (which must be confined to the cone of directions which make an acute angle with all surface normals).

Moreover, in order to reduce the effects of secondary lighting, one must constrain the slope of the surface (so that most of the scattered light leaves the scene before hitting some other part of the surface), and/or constrain the albedo to very low values (so that the scattered light flow is small compared to the external flow).

Strictly speaking, the uniform lighting condition also excludes *attached shadows* — regions on the scene surface where the normal makes an acute angle with the direction of the external light flow, so that the light is blocked by the object at  $p$  itself. However, the value of  $\Phi_i[p](\vec{u})$  is irrelevant for directions  $\vec{u}$  that point outwards from the local surface, because the light flowing in these directions gets excluded anyway from the integral (4) by the factor  $\max\{0, \vec{u} \cdot \vec{n}\}$  included in the BRDF. Therefore, at any point  $p$  that lies on the attached shadow of some object, one may still assume that  $\Phi_i[p]$  is the light field  $\Phi_i$  that would be observed at  $p$  if that object did not exist. It follows that attached (but not cast!) shadows can be handled under the uniform lighting model.

### 2.2. The sufficient data hypothesis

In any case, we must assume that the gauge images are free from such defects; that is,  $\Phi_i[q]$  is equivalent to a fixed flow  $\Phi_i^*$  for any point  $q$  on the gauge. We will say that the

light flow  $\Phi_i[p]$  at a point  $p$  of the scene is *gauge-like* if it is equivalent to the gauge flow  $\Phi_i^*$ , apart from the effect of attached shadows and the usual level of quantization and measurement errors.

If some of the scene images are affected by anomalous lighting, the observation vectors  $S[p]$  and  $G[q]$  of two points with same surface normal may not be collinear. Equation (6) does not hold, and the method outlined in section 1.7 does not work.

Gauge-based VLPS is still possible in the presence of anomalous lighting, as long as, for each point  $p$ , there is a subset  $K[p]$  of the indices  $\{1, \dots, m\}$  with  $|K[p]| \geq 3$  such that  $\Phi_i[p]$  is gauge-like for all  $i \in K[p]$ . Note that the subset may (and, usually, will) be different for each point  $p$ . We will refer to this condition as the *sufficient data hypothesis*.

To model this situation, we introduce the *clarity factor*  $\kappa_i[p]$ , defined as the ratio between the observed intensity  $S_i[p]$  of a point  $p$  in the scene image, and the albedo-corrected intensity  $G_i[q]\tilde{s}[p]/\tilde{g}[q]$  of the gauge point  $q$  with same normal. That is, we replace equations (3) by

$$\begin{aligned} S_i[p] &= \tilde{s}[p] L_i^*(\vec{s}[p]) \kappa_i[p] \\ G_i[q] &= \tilde{g}[q] L_i^*(\vec{g}[q]) \end{aligned} \quad (8)$$

where  $L_i^*$  is the gauge’s shading function, assumed to be independent of the position  $q$ .

Note that the clarity factor  $\kappa_i[p]$  is 1 if the light flow  $\Phi_i$  is gauge-like at  $p$ . In images  $S_i$  where  $p$  lies in the cast shadow of some other object, the clarity  $\kappa_i[p]$  will be less than 1. Conversely, in images  $S_i$  where the light flow at  $p$  is incremented by highlights or secondary lighting, the clarity  $\kappa_i[p]$  will be greater than 1. The sufficient data assumption can be restated as follows: for any point  $p$  in  $\mathcal{S}$ , there are least three indices  $i$  such that  $\kappa_i[p]$  is 1.

### 2.3. The subsetting approach

The proportionality test (6) would still work if one could somehow identify the subset of  $K$  of “good” images. In principle, one could try to solve the problem for all subsets  $K$  with three or more images, as if the corresponding flows were shadow-like at  $p$ , and then select the answer that provides the best match between the scene and gauge observation vectors. However, the number of combinations  $\binom{m}{3}$  is approximately  $m^3/6$ , so this method is too slow when one has dozens of images.

Previous attempts to solve this problem have focused on heuristic criteria for detecting whether a pixel  $p$  is affected by shadows and/or highlights in some image  $S_i$ . The signature-matching algorithm would then be applied to the subset  $K$  of images that pass those criteria [2, 4].

## 2.4. Limitations of the subsetting approach

However, shadow- and highlight-detection heuristics are often unreliable, or require unrealistic constraints on the scene and lighting conditions.

In particular, one cannot merely exclude images  $S_i$  where  $S_i[p]$  is below a certain threshold (whether absolute, or relative to other values  $S_j[p]$ ), as proposed by Barski and Petrou [2], because this criterion would also exclude images with attached shadows — which do not break the signature-matching algorithm, provide useful information, and may even be necessary to achieve the sufficient-data condition.

Moreover, real light sources usually cast penumbras rather than sharp-edged shadows, so one can have shadowed points in that are substantially lighter than gauge-like points *in the same scene image*. For this same reason, one cannot expect that shadows will have sharp edges, and that “shadowed” is a boolean all-or-nothing attribute — as assumed by Chandraker, Agarwal and Kriegman [4]. The frequent presence of substantial secondary lighting is a further complication for this approach.

Indeed we conjecture that, under moderately general conditions, *it is impossible to reliably determine when a flow  $\Phi_i$  is gauge-like at any scene point  $p$  by analyzing the scene images  $S_1, \dots, S_m$  alone*. On the other hand, it seems that any analysis that also uses the gauge images  $G_1, \dots, G_m$  for this decision must necessarily identify the point  $q$  of the gauge that has the same normal as  $p$ .

## 3. Our approach

Given the practical and theoretical limitations of heuristic subsetting, we have opted for a different “holistic” approach. When looking for the gauge point  $q$  that best matches the point  $p$ , we use the whole image set  $S_1, \dots, S_m$ , but we compare the observation vectors  $\mathbf{S}[p]$  and  $\mathbf{G}[q]$  with a distance metric  $\Delta(\mathbf{S}, \mathbf{G})$  that is insensitive to anomalous lighting effects.

To understand the idea, consider the graphs shown in the left column of figure 4. Each graph refers to a single pair of pixels  $p \in \mathcal{S}$  and  $q \in \mathcal{G}$ , identified by red crosses in the image. Each mark in the graph represents a pair  $a_i = (\mathbf{S}_i, \mathbf{G}_i) = (S_i[p], G_i[q])$ , for  $i \in \{1, \dots, m\}$ , taken from the  $m$  scene and gauge images in figures 1 and 2.

Figure 4(a) is the ideal situation, where  $p$  and  $q$  have the same normal, and  $p$  is gauge-like in all images. As one can see, the  $m$  points  $a_i$  fall on a straight line through the origin, apart from small errors due to pixel quantization and measurement noise. The slope of the line is the relative albedo  $\tilde{s}[p]/\tilde{g}[q]$ .

The plot 4(b) shows the situation when  $p$  and  $q$  have the same normal, but  $p$  is gauge-like in only 4 of the images.

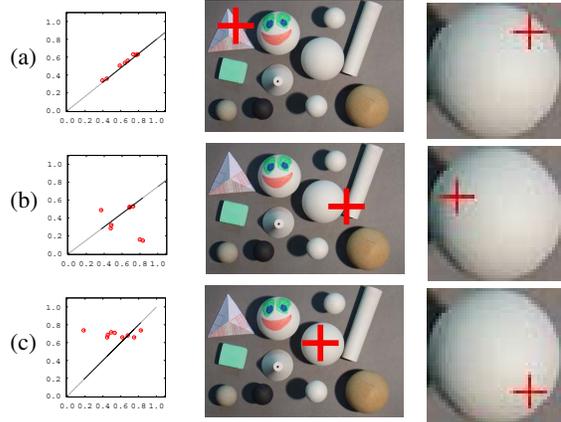


Figure 4. Plots of pairs  $(S_i, G_i)$ .

These pairs that correspond to those images still fall on a straight line, whereas some of the other pairs lie below the line (because of shadowing) or above it (because of highlights or secondary lighting). Finally, figure 4(c) is typical for the case when  $p$  and  $q$  have different normals. Note there is no significant subset of points that are collinear with the origin.

Our function  $\Delta(\mathbf{S}, \mathbf{G})$  is such that it is small in cases (a) and (b), and large in case (c). Specifically, the function tries to find a sufficiently large set  $K[p, q] \subseteq \{1, \dots, n\}$  of indices such that the corresponding pairs  $a_i$  lie sufficiently close to a straight line through the origin. If such set exists, it returns a small value (much less than 1), whose magnitude depends on how many pairs are in that set, and how close they are to the line. Otherwise it returns an arbitrary large value near 1.

The precise meanings of “sufficiently large” and “sufficiently close” obviously depends on  $m$ , on gauge intensities  $\mathbf{G}_i$ , and on the accuracy of the data. However, note that by varying the point  $q$  on the gauge we can cause the two abscissas  $G_i[q]$  and  $G_j[q]$  to vary independently of each other. Therefore, we can expect to have a whole one-dimensional manifold of points  $q \in \mathcal{G}$  for which  $a_i$  and  $a_j$  are aligned with the origin. It follows that the alignment of only two pairs  $a_i, a_j$  with the origin cannot be considered significant.

On the other hand, having three points  $a_i, a_j, a_k$  aligned with the origin is a situation that can arise only by accident, or because the two normals agree. To the extent that the first hypothesis is unlikely, the coincidence can be counted as evidence that  $p$  and  $q$  have the same normal.

The normal recovery problem can be framed as a Bayesian parameter estimation problem, where the observed output is the pixel’s observation vector  $\mathbf{S}$ , and the unknown parameters are the surface normal  $\vec{s}$  and the albedo  $\tilde{s}$ . Note that this approach differs from the subsetting approach of section 2.3. Instead of trying to choose the subset  $K$  a priori, as a function of the scene point  $p$  alone, our

distance looks for a different subset  $K$  for each pair  $p, q$ , determined from both OVs.

### 3.1. Computing the distance $\Delta$

Computing  $\vec{s}$  and  $\vec{s}^*$  from the Bayesian formula is prohibitively expensive, so we have developed the following practical approximation. If we knew the albedo  $\alpha$  of  $p$ , we could evaluate the likelihood of  $q$  being the correct gauge pixel by computing the *albedo score*

$$W(\alpha) = \sum_{j=1}^m \mathbf{G}_j \exp \left[ -\frac{1}{2} (\mathbf{S}_j - \alpha \mathbf{G}_j)^2 \right] \quad (9)$$

The score  $W(\alpha)$  is an indication of how many pairs  $a_i = (\mathbf{S}_i, \mathbf{G}_i)$  fall close to a straight line through the origin with slope  $\alpha$ , and how closely they do. In particular, if  $k$  pairs with  $\mathbf{G}_j$  close to 1 fall precisely over that line, while the rest are very far from it, then the score  $W(\alpha)$  will be about  $k$ . The factor  $\mathbf{G}_i$  compensates for the fact that pairs with low  $\mathbf{G}_i$  are more likely to lie near the line by mere chance.

We can take the maximum value of  $W(\alpha)$  as an indication of the likelihood that the point  $q$  has the same normal as  $p$ . For this purpose, we can use  $W_{\max} = \max \{W(\alpha_i) : 1 \leq i \leq m\}$ , where  $\alpha_i = \mathbf{S}_i/\mathbf{G}_i$ . Although the the maximum-score albedo need not be any  $\alpha_i$ , it is expected to very close to one of them.

Since these intensities can be contaminated by error, measured by the standard deviation  $\sigma$  of the quantization and measurement noise present in pixel intensities, we use the safer estimator  $\alpha_i = \sqrt{\mathbf{S}_i^2 + \sigma^2}/\sqrt{\mathbf{G}_i^2 + \sigma^2}$ , which tends to  $\sqrt{1/2}$  when both intensities tend to zero. Furthermore, we assume that the albedo cannot be greater than 1; so we can ignore (and skip the computation of)  $W(\alpha_i)$  when  $\mathbf{S}_i \geq \mathbf{G}_i + 3\sigma$ . Having found the maximum score  $W_{\max}$ , we compute the distance  $\Delta$  so that it is large when  $W_{\max}$  is 2 or less, and tends quickly to 0 as  $W_{\max}$  grows beyond 3. We use a Gaussian function

$$\Delta(\mathbf{S}, \mathbf{G}) = \exp \left[ -\frac{1}{2} \left( \frac{W_{\max}}{\omega} \right)^2 \right] \quad (10)$$

where  $\omega$  is arbitrarily set to 1.5.

### 3.2. Analysis

The average computation cost per pixel is proportional to  $Nm^2$ , where  $N$  is the number of entries in the gauge OV table, and  $m$  is the number of lighting fields. The cost can be easily reduced to  $Nm \log m$  by eliminating pairs  $i, j$  that contribute little to the score. Therefore, our method remains viable even when  $m$  is a hundred or more.

## 4. Tests

To evaluate our method, we tested it on two sets of data: (1) synthetic images, (2) digital photos of a real scene.

### 4.1. Tests with synthetic image data

While synthetic images are not acceptable for validating complete VLPS algorithms, they are a useful check of whether the algorithm works when its basic assumptions are verified. Artificial images are free from spurious factors (such as non-linear camera sensor response, image alignment errors, non-uniform lighting, and image compression noise) that could affect the performance, but are not inherent to the algorithm.

The synthetic images were produced by ray-tracing a simple virtual scene, consisting of blocks of various shapes and colors lying on a flat checkered surface. The gauge object was a white sphere placed outside the reach of other object's cast shadows. We used seven images in the test reported here. Two of the images were rendered with near-normal lighting and were free of cast shadows; the other five had substantial cast shadows on the floor and on some of objects. See figures 5 and 6.

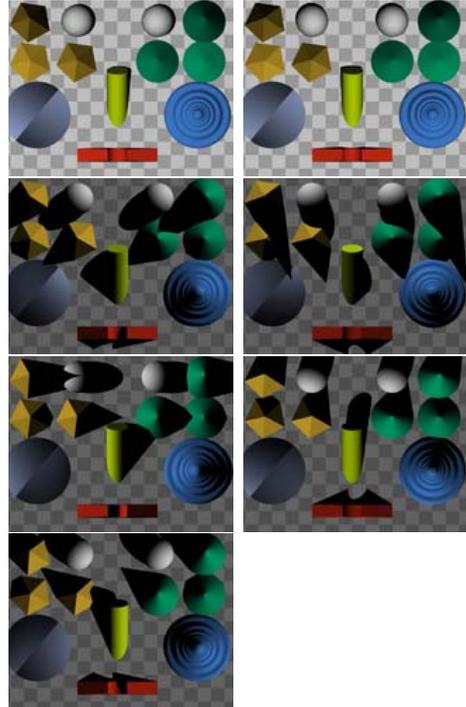
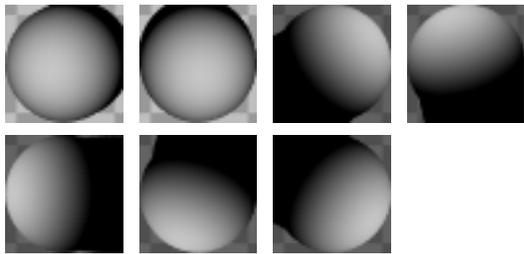


Figure 5. Synthetic scene images.

All objects had Lambertian finish and were illuminated by a point-like source. The source-to-scene distance was



**Figure 6. Gauge images for figure 5.**

---

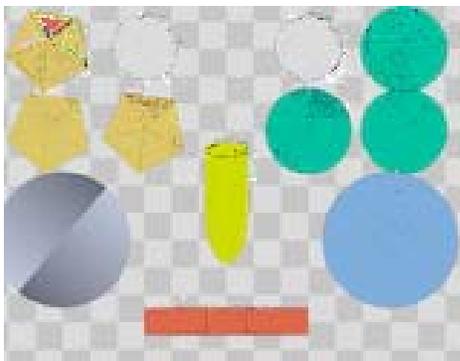
about 100 times the width of the visible part of the scene, and the camera field-of-view was narrowed to provide near-parallel image projection.

The gauge images were subsampled to provide 3771 gauge sample points  $q$ . The resulting normal and albedo maps are shown in figures 7 and 8.



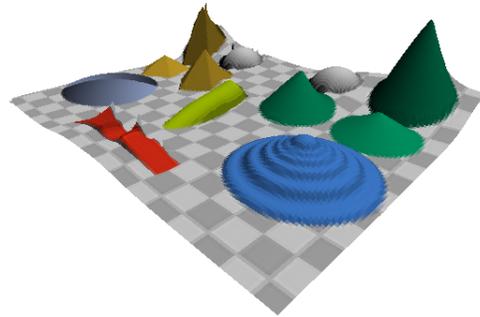
**Figure 7. Normal map from figure 5.**

---



**Figure 8. Albedo map from figure 5.**

---



**Figure 9. 3D view from figure 5.**

---

#### 4.2. Tests with real images

For the tests with real data, we took digital photographs of a group of simple objects: a truncated cone, a cylinder, a three-sided pyramid, half of a green rubber eraser, and seven spheres. See figure 1. The pyramid was made of heavy paper with printed red text. Five objects were left uncoated: the eraser, the pyramid, a wooden drawer knob, a gray rubber mouse-ball, and a black one. All the other objects were coated with spray-on matte white paint, and the happy face was then drawn on one of the spheres, with colored whiteboard markers. The white mouseball at bottom center, which was placed on pedestal slightly above the other objects, was used as the light gauge.

We took a total of 13 photos of these objects with a consumer camera (Sony DSC-V1) fixed to a photographic table. Lighting was provided by a hand-held 500W cylindrical halogen lamp with a  $8\text{cm} \times 8\text{cm}$  reflector, about 50–100cm from the scene. There was also some diffuse fluorescent lighting (with a noticeably cooler tinge) from ceiling fixtures. The lamp was moved around so that half of the photos are mostly free of cast shadows, and half have substantial penumbras cast over several objects. All images were scaled down to scaled down to  $128 \times 85$  pixels. The resulting normal and albedo maps are shown in figure 10 and 3.

#### 5. Conclusions

The experiments we performed so far indicate that our method is quite robust and general. It can handle situations with shadows and secondary lighting, without the need for a prior identification of anomalous pixels. The method makes no assumption on continuity of normals across the scene; so it works independently on each scene pixel, and therefore can be trivially parallelized.

Although the cost of the normal identification step is relatively high, there is plenty of room for optimizations [5, 7].



Figure 10. Normal map from figure 1.



Figure 11. 3D view from figure 1.

## Acknowledgements

This project was partly supported by research grants from CNPq (304581/2004-6 and 301016/92-5), CAPES, FAPESP and FAPERJ. ★[Tematico, CT-Info...].

## References

- [1] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu. An optimal algorithm for approximate nearest neighbor searching in fixed dimensions. *Journal of the Association of Computing Machinery*, 45(6):891–923, 1998.
- [2] S. Barsky and M. Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in presence of highlights and shadows. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(10):1239–1252, Oct. 2003.
- [3] R. Basri, D. Jacobs, and I. Kemelmacher. Photometric stereo with general unknown lighting. *International Journal of Computer Vision*, 72(3):239–257, 2007.
- [4] M. Chandraker, S. Agarwal, and D. Kriegman. ShadowCuts: Photometric stereo with shadows. In *CVPR07*, pages 1–8, 2007.
- [5] H. C. da Gama Leitão, R. Saracchini, and J. Stolfi. A bucket grid structure to speed up table lookup in gauge-based photometric stereo. In *Proceedings of the 20th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAP 2007)*, pages 221–227. IEEE Computer Society Press, Oct. 2007.
- [6] A. Gionis, P. Indyk, and R. Motwani. Similarity search in high dimensions via hashing. *The VLDB Journal*, pages 518–529, 1999.
- [7] A. Hertzmann and S. M. Seitz. Shape and materials by example: A photometric stereo approach. In *Proceedings IEEE CVPR 2003*, volume 1, pages 533–540, June 2003.
- [8] A. Hertzmann and S. M. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying BRDFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1254–1264, Aug. 2005.
- [9] B. K. P. Horn. Understanding image intensities. *Artificial Intelligence*, 8(2):201–231, 1977.
- [10] B. K. P. Horn. Obtaining shape from shading information. In B. K. P. Horn and M. J. Brooks, editors, *Shape from Shading*, pages 121–171. MIT Press, Cambridge, Mass., 1989.
- [11] O. Ikeda. Shape reconstruction for color objects using segmentation and photometric stereo. In *Proceedings of ICIP 2004*, pages 1365–1368, 2004.
- [12] P. Indyk and R. Motwani. Approximate nearest neighbour: Towards removing the curse of dimensionality. In *Proceedings of the 13th Annual ACM Symposium on Theory of Computing (STOC'98)*, pages 604–613, 1998.
- [13] A. suppressed to avoid identification]. [title suppressed]. 2005.
- [14] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.
- [15] R. J. Woodham. Determining surface curvature with photometric stereo. In *Proceedings of the 1989 IEEE International Conference on Robotics and Automation*, volume 1, pages 36–42, May 1989.
- [16] R. J. Woodham. Gradient and curvature from the photometric stereo method, including local confidence estimation. *Journal of the Optical Society of America, Series A*, 11(11):3050–3068, 1994.
- [17] L. Zhong and J. J. Little. Photometric stereo via locality sensitive high-dimension hashing. In *Proceedings of the Second Canadian Conference on Computer and Robot Vision (CRV'05)*, pages 104–111, 2005.