Specularities Reduce Ambiguity of Uncalibrated Photometric Stereo

Ondřej Drbohlav and Radim Šára

Center for Machine Perception, Department of Cybernetics
Faculty of Electrical Engineering, Czech Technical University in Prague
Technicka 2, 166 27 Prague 6, Czech Republic
{drbohlav,sara}@cmp.felk.cvut.cz

Abstract. Lambertian photometric stereo with uncalibrated light directions and intensities determines the surface normals only up to an invertible linear transformation. We show that if object reflectance is a sum of Lambertian and specular terms, the ambiguity reduces into a 2dof group of transformations (compositions of isotropic scaling, rotation around the viewing vector, and change in coordinate frame handedness).

Such ambiguity reduction is implied by the *consistent viewpoint constraint* which requires that all lights reflected around corresponding specular normals must give the same vector (the viewing direction). To employ the constraint, identification of specularities in images corresponding to four different point lights in general configuration suffices. When the consistent viewpoint constraint is combined with integrability constraint, binary convex/concave ambiguity composed with isotropic scaling results. The approach is verified experimentally.

We observe that an analogical result applies to the case of uncalibrated geometric stereo with four affine cameras in a general configuration observing specularities from a single distant point light source.

1 Introduction

Photometric stereo [13] is a method that recovers local surface geometry and reflectance properties from images of an object that are taken by a fixed camera under varying distant illumination. The principle of photometric stereo is in inverting a parametric model of surface reflectance. A usual version of photometric stereo uses a single distant point light source at a time to illuminate an object, and assumes Lambertian surface reflectance which implies that brightness value $I_{i,j}$ of i-th pixel in the image capturing the object appearance under j-th point light source is (see Fig. 1)

$$I_{i,j} = E_j \rho_i \cos\theta_{i,j} = (\rho_i \mathbf{n}_i)^\top (E_j \mathbf{l}_j)\,, \tag{1}$$

where E_j is the intensity of the light source, $\mathbf{l}_j$ is the light source direction, $\mathbf{n}_i$ is the normal vector of a surface patch that projects into the i-th pixel, $\theta_{i,j}$

A. Heyden et al. (Eds.): ECCV 2002, LNCS 2351, pp. 46–60, 2002.

is the angle between $\mathbf{n}_i$ and $\mathbf{l}_j$ (the angle of incidence), and ρ_i is a reflectance parameter of a small surface patch. This parameter is called *albedo* and describes what portion of incident light is re-emitted back into space in the form of diffuse reflection.

It is well understood that the reflectance model described by equation (1) is bilinear. To see that, it is convenient to denote $\rho_i \mathbf{n}_i$ and $E_j \mathbf{l}_j$ by $\mathbf{b}_i$ and $\mathbf{s}_j$, respectively; then, the above equation takes very simple, compact form

$$\mathbf{I} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N]^\top [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_M] = \mathbf{B}^\top \mathbf{S}, \tag{2}$$

where $\mathbf{I}$ is the matrix which has $I_{i,j}$ from (1) as its elements, matrix $\mathbf{B}$ collects the $\mathbf{b}_i$'s and matrix $\mathbf{S}$ analogically collects the $\mathbf{s}_j$'s. For the sake of brevity, we call $\mathbf{B}$ the *normals* and $\mathbf{S}$ the *lights*.

In the original version of photometric stereo [13], the light source directions $\mathbf{l}_j$ and intensities E_j are calibrated, thus the lights $\mathbf{S}$ are known. To evaluate normals and albedos, it suffices to right-multiply the matrix $\mathbf{I}$ in (2) by the inverse (or pseudo-inverse if the number of lights is greater than 3) of $\mathbf{S}$; by that the normals $\mathbf{B}$ are acquired. Normals $\mathbf{n}_i$'s are then $\mathbf{b}_i$'s scaled to unity, and albedos ρ_i's are the lengths of $\mathbf{b}_i$'s.

If, however, the light sources $\mathbf{S}$ are *not* known, then (2) represents a bilinear calibration-estimation problem [8] whose ambiguity can be phrased as follows:

> **Uncalibrated photometric stereo ambiguity.** Let there be images of an object of Lambertian reflectance observed from a fixed viewpoint, but illuminated sequentially from different unknown directions by a distant point light source. Then it is possible to factorize the input data matrix $\mathbf{I}$ from (2) into *pseudonormals* $\overline{\mathbf{B}}$ and *pseudolights* $\overline{\mathbf{S}}$ [6] that give the true normals $\mathbf{B}$ and the true lights $\mathbf{S}$ up to an unknown linear invertible transformation $\mathbf{A} \in GL(3)$: $\overline{\mathbf{B}} = \mathbf{A}\mathbf{B}$, $\overline{\mathbf{S}} = \mathbf{A}^{-\top}\mathbf{S}$.

This ambiguity exists because it holds that $\mathbf{I} = \overline{\mathbf{B}}^\top \overline{\mathbf{S}} = \mathbf{B}^\top \mathbf{A}^\top \mathbf{A}^{-\top} \mathbf{S} = \mathbf{B}^\top \mathbf{S}$. The uncalibrated photometric stereo ambiguity can be reduced and/or removed only if additional information about lights or normals is available. This information may have different form. First possibility is to estimate normal vectors and albedos in several points by an independent method and use them to disambiguate the photometric stereo (note that due to the symmetry of (2), the value of this information is the same as if light directions and intensities are known). Another possibility is to assume that at least six light sources are of equal (or known relative) intensity, or that albedo is uniform (or known up to a global scaler) for at least six normals at a curved surface. Such possibilities were employed and/or discussed in [6,14,1], and it was shown that such knowledge reduces the ambiguity from the $GL(3)$ group into the group of scaled orthogonal transformations $\mathbf{A} = \lambda \mathbf{O}$ ($\mathbf{O} \in O(3), \lambda \neq 0$). Yet another important possibility is given by the integrability constraint that requires the normals recovered by photometric stereo to correspond to a continuous surface [1,4]. As shown by

Belhumeur et al. [1], in this case the original ambiguity is reduced into ambiguity represented by the group of four-parametric (or three-parametric, if overall scaling is not counted) generalized bas-relief transformations. And, integrability constraint together with the knowledge of six albedos (or six light intensities as described above) reduces the original ambiguity into binary convex/concave ambiguity composed with isotropic scaling.

In our recent work on uncalibrated photometric stereo [3] we showed that inherent symmetries of reflectance models that are separable with respect to the viewing and illumination directions can be exploited to construct two new geometrical constraints. The constraints are represented by projections of normals onto planes perpendicular to the viewing and illumination directions, respectively. We constructed the constraints using polarization measurement under the assumption of separable reflectance model for smooth dielectrics and showed that the two constraints alone combined together reduce the ambiguity to convex/concave ambiguity composed with isotropic scaling.

In this paper we show that if object reflectance is a sum of Lambertian reflectance and a mirror-like reflectance, then the original ambiguity represented by a group $GL(3)$ reduces into a two-parametric group of transformations. These transformations are compositions of isotropic scaling (1dof), rotation around the viewing vector (1dof), and change in the global coordinate frame handedness (binary ambiguity). This ambiguity reduction is implied by a condition that all lights reflected around corresponding specular normals must give the same vector (the viewing direction). We call this condition the *consistent viewpoint constraint*. We show that specularities in as few as *four* images corresponding to four different distant point lights in general configuration are sufficient to utilize the consistent viewpoint constraint.

By this result, we make a step towards *uncalibrated* photometric stereo for objects whose reflectance includes not only body (diffuse) component, but also interface (specular) component. Such composite reflectance models are certainly not new to photometric stereo applications, see e.g. [2,11,10,9], but in those methods, in contrast to the ours, the light sources are supposed to be known.

The specific representative of composite reflectance model (the superposition of Lambertian and mirror-like reflectance) is selected in this work because as specularities are sparse in the images, they can be treated as *outliers* to the *Lambertian* reflectance model. This gives us a valuable possibility to study the problem as Lambertian photometric stereo with additional information represented by the consistent viewpoint constraint.

2 Consistent Viewpoint Constraint

The problem we will analyze is photometric stereo with uncalibrated lights $\mathbf{S}$ for objects whose reflectance is given by superposition of Lambertian and specular terms. As discussed in Section 1, we treat this problem as uncalibrated Lambertian photometric stereo with additional geometrical information provided by

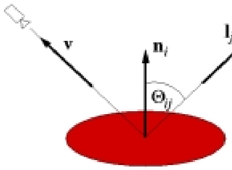


Fig. 1. Reflectance geometry. For Lambertian reflectance, brightness seen by a camera is dependent on cosine of the angle of incidence and independent on the viewing direction.

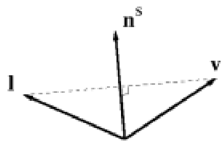


Fig. 2. Specular geometry configuration. Surface normal $\mathbf{n}^S$ is a bisector between the viewing direction $\mathbf{v}$ and the illumination direction $\mathbf{l}$.

specularities. In this section we review the geometry of mirror-like reflection, and formulate the constraint implied by the observation of specularities in images.

It is well understood that specularities occur at loci where light reflects on a smooth surface in a mirror-like manner towards the observing sensor. Hence, mirror-like reflection carries valuable information about geometrical configuration of the viewing vector, the illumination vector, and the surface normal: if a specularity is observed in an image, then at the corresponding surface point, surface normal is a bisector of the viewing and illumination vectors (see Fig. 2). Therefore for the viewing direction it holds that

$$\mathbf{v} = 2(\mathbf{l} \cdot \mathbf{n}^S)\mathbf{n}^S - \mathbf{l} = 2\left(\widehat{\mathbf{s}} \cdot \widehat{\mathbf{b}}^S\right)\widehat{\mathbf{b}}^S - \widehat{\mathbf{s}} \tag{3}$$

where $\widehat{}$ denotes normalization to unity, and $\mathbf{n}^S$ is a normal that is observed as specular under illumination of direction $\mathbf{l}$. The right-most part of the equation essentially states the same fact in "natural" photometric stereo variables. We call $\mathbf{l}$ and $\mathbf{n}^S$ (as well as $\mathbf{s}$ and $\mathbf{b}^S$) a *specular pair*.

The equation may be viewed as a formula for computing viewpoint direction from known light $\mathbf{s}$ and specular normal $\mathbf{b}^S$. The key fact to be observed is that this relation states: no matter which specular pair is used for viewing direction evaluation, all give the same result.

Consistent viewpoint constraint. A collection of specular pairs follows the *consistent viewpoint constraint* if they all, by (3), give the same viewing direction $\mathbf{v}$.

Does the consistent viewpoint constraint reduce the uncalibrated photometric stereo ambiguity? We will analyze what transformations may be applied to the true normals and the true lights, such that the transformed specular pairs, inserted into (3), all give the same vector. Let us denote this vector $\mathbf{u}$ and write the equivalent of (3) for the transformed lights and normals:

$$\mathbf{u} = \frac{2\left[(\mathbf{A}^{-\top}\mathbf{s}) \cdot (\mathbf{A}\mathbf{b}^S)\right]\mathbf{A}\mathbf{b}^S}{||\mathbf{A}^{-\top}\mathbf{s}||||\mathbf{A}\mathbf{b}^S||^2} - \frac{\mathbf{A}^{-\top}\mathbf{s}}{||\mathbf{A}^{-\top}\mathbf{s}||}, \tag{4}$$

where $||\cdot||$ are explicitly written normalization factors. Multiplying both sides of the equation by $||\mathbf{A}^{-\top}\mathbf{s}||||\mathbf{A}\mathbf{b}^S||^2\mathbf{A}^\top$, we get

$$\alpha(\mathbf{s},\mathbf{b}^S)\mathbf{w} = 2(\mathbf{s}\cdot\mathbf{b}^S)\mathbf{P}\mathbf{b}^S - (\mathbf{b}^S\cdot\mathbf{P}\mathbf{b}^S)\mathbf{s}\,, \tag{5}$$

where $\alpha(\mathbf{s},\mathbf{b}^S) = ||\mathbf{A}^{-\top}\mathbf{s}||||\mathbf{A}\mathbf{b}^S||^2$ absorbs (unknown) scaling factors, $\mathbf{P}$ denotes $\mathbf{A}^\top\mathbf{A}$ and $\mathbf{w}$ denotes $\mathbf{A}^\top\mathbf{u}$; and we applied the fact that $(\mathbf{A}^{-\top}\mathbf{s})\cdot(\mathbf{A}\mathbf{b}^S) = \mathbf{s}\cdot\mathbf{b}^S$. Note that in this equation, vector $\mathbf{w}$ may be treated as fully independent on $\mathbf{P}$ because $\mathbf{P} = \mathbf{A}^\top\mathbf{A}$ gives $\mathbf{A}$ only up to arbitrary orthogonal transformation. We show in Appendix that for a convex smooth specular object illuminated from all directions it must hold that $\mathbf{P} = \lambda^2\mathbf{I}$, $\lambda \neq 0$ ($\mathbf{P}$ is a scaled identity). From that it follows that the only transformations under which the consistent viewpoint constraint is preserved are $\mathbf{A} = \lambda\mathbf{O}$, $\mathbf{O} \in O(3)$. Fixing the coordinate frame by a usual choice (image plane spans plane $x - y$, viewing direction coincides with axis z), the allowable transformations $\mathbf{A}$ are those that preserve the viewing direction. Writing them explicitly,

$$\mathbf{A} = \lambda \begin{bmatrix} \pm1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \pm1 \end{bmatrix} \mathbf{R}_z(\xi) \qquad \xi \in \langle 0, 2\pi)\,, \lambda > 0\,. \tag{6}$$

In this equation, $\mathbf{R}_z(\xi)$ stands for rotation around the z-axis (the viewing direction) by angle ξ. The ambiguity in sign of the third coordinate was included in (6) only for the sake of completeness, but naturally the correct sign is easily set by orienting the normals towards the viewing direction (normals that are inclined from the viewpoint are invisible). The ambiguity in sign of the first coordinate stays unresolved until some additional constraint is applied (or until it is resolved manually).

In this paper we resolve the remaining ambiguity using the integrability constraint. Integrability constraint fixes both the sign of the first coordinate (thus the handedness of the coordinate system) *and* the rotation angle ξ modulo π. This means that the final ambiguity is convex/concave ambiguity composed with isotropic scaling. This result follows from the fact that the intersection of the $O(3)$ group (of which transformations (6) with $\lambda = 1$ are a sub-group) with the generalized bas-relief group is a two-element set of the identity transformation and the transformation that reflects the first two coordinates [1].

Finally, let us observe how many specular pairs do we need to establish the consistent viewpoint constraint. Equation (5) represents three scalar equations for each specular pair. After eliminating the unknown constant $\alpha(\mathbf{s},\mathbf{b}^S)$, there are two independent equations per specular pair. The unknowns $\mathbf{w}$ and $\mathbf{P}$ are both up to scale, so the number of degrees of freedom to fix is 2 (from the vector $\mathbf{w}$) plus 5 (from the symmetric matrix $\mathbf{P}$). We thus observe that at least *four* specular pairs in general configuration are needed to apply the consistent viewpoint constraint.

So far, we have not analyzed which configurations of four specular pairs are singular, nor the problem of (possible) finite solution multiplicity for non-singular configurations. However, in experiments we observed unique solution in all cases. Analysis of which sets of four specular pairs give well-conditioned solution is a topic for future research.

3 Experiment

In this experiment we show normal and albedo reconstruction for two objects:

1. *WhiteBall* which is a highly polished billiard ball of uniform albedo,
2. *ChinaPot* which is a glazed china tea pot with painted motif.

Images were acquired by 12 bit cooled camera (COOL-1300 by Vosskühler, standard Computar 75mm lens) under tungsten illumination (150W, stabilized direct current). The light was moved by hand around the object. The distance between object and light was not kept constant. No information about lights has been measured nor recorded. Input images for the WhiteBall object are shown in Fig. 3.

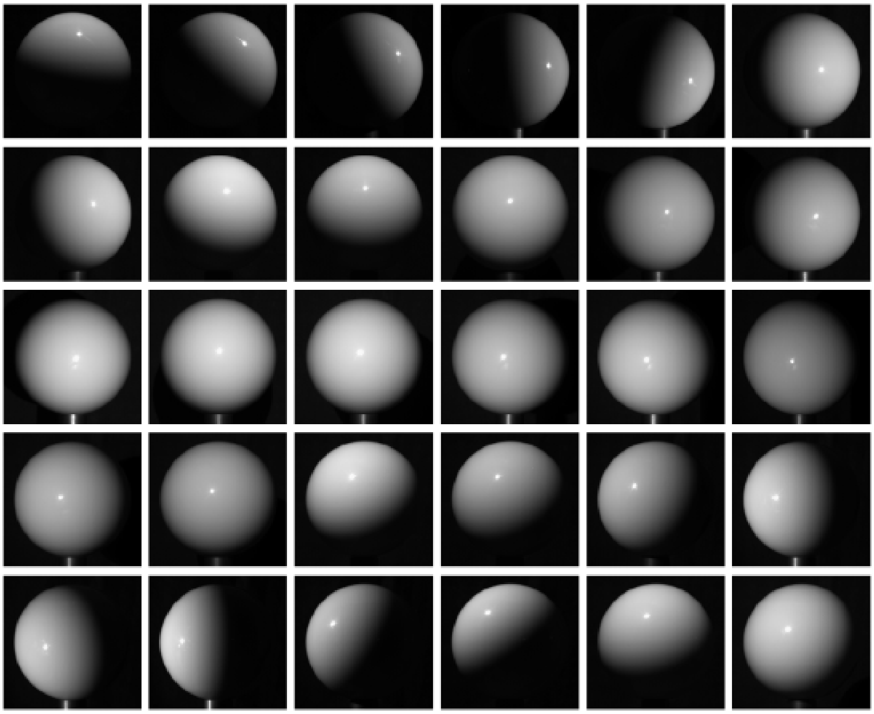


Fig. 3. Input data for the WhiteBall object.

Data was processed in 9 consecutive steps:

1. The mean of 10 dark frames was subtracted from each of the input images.

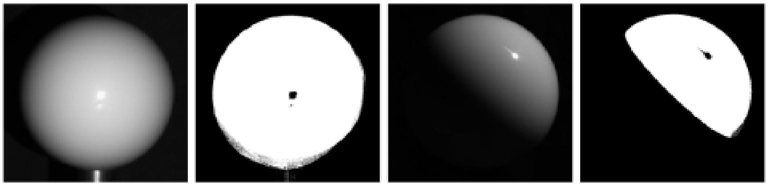


Fig. 4. Two examples of Lambertian behaviour masks for the WhiteBall object.

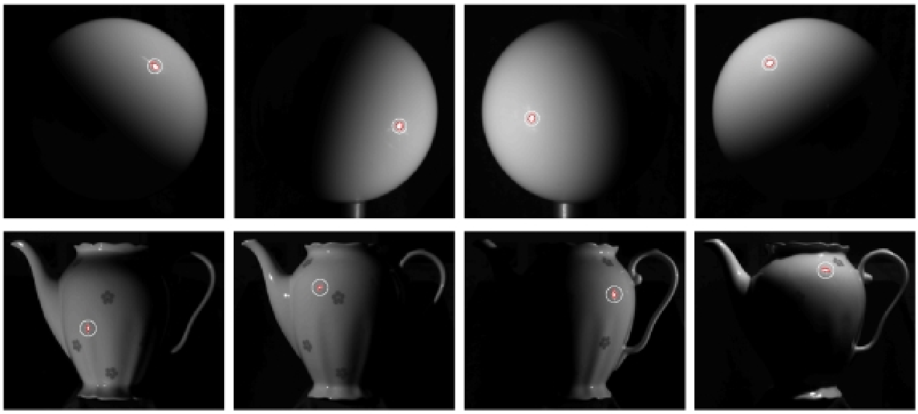
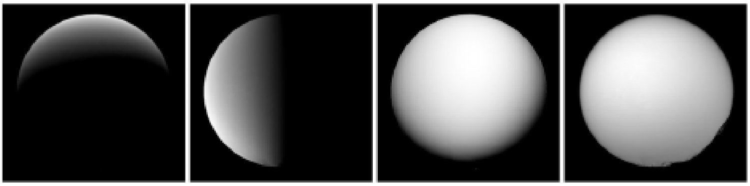
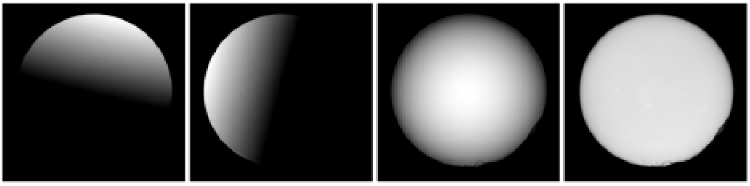


Fig. 5. Selected specular regions (marked with red contour and rounded by small white circle for better reading) used to apply the consistent viewpoint constraint, for both objects.

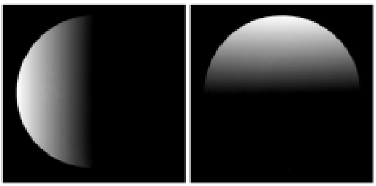
2. Image points whose intensity fell below or over respective thresholds were labeled as non-Lambertian, the other ones as *candidates* for Lambertian points.
3. Candidates for Lambertian points were confirmed to be Lambertian in the case that in four randomly selected images, they belonged to sufficiently large image pixel sets with Lambertian behavior (i.e., if intensities of the pixel set in any of four images could be sufficiently well expressed as a linear combination of intensities of the pixel set in the other three images). Only such quadruples of images were involved whose any three corresponding light directions were sufficiently far from being coplanar (automatic check of this condition was done using simple conditioning number tests). Two selected Lambertian-consistent masks resulting from this step are shown in Fig. 4.
4. Lambertian portion of data was factorized by Jacobs algorithm [7]. From the factorization pseudolights $\overline{\mathbf{S}}$ were obtained.
5. Pseudonormals $\overline{\mathbf{B}}$ were computed using Lambertian image regions and the pseudolights obtained in the previous step. Each normal was fit individually by using least-square fit. After that, pseudolights were re-evaluated by an analogous procedure, and this iterative process (alternating between re-



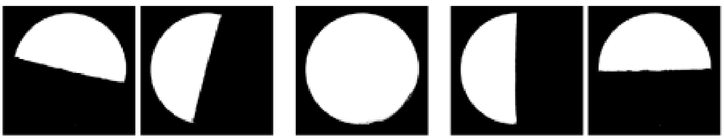
(a) Normals illuminated from the coordinate axis directions (left, top, viewpoint) and albedo (last image) reconstructed by Jacobs method from segmented input data assuming Lambertian reflectance. Note that normals look as illuminated from directions other then specified. The albedo map does not correspond to expectations either, since in reality it is uniform over the object.



(b) Normals (illuminated as above) and albedo conforming to the consistent viewpoint constraint. Note that images already appear as illuminated from three mutually perpendicular directions and that the albedo map is uniform, as expected.

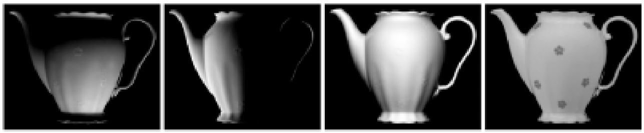


(c) Normals illuminated from the left and top directions conforming to both the consistent viewpoint and integrability constraints. The viewpoint-illuminated image and albedo are identical to those in Fig. 6(b). Note that the original ambiguity (cf. Fig. 6(a)) is already canceled.

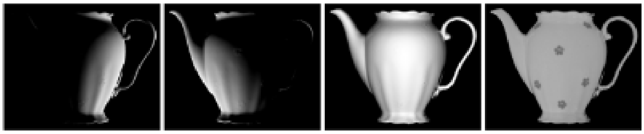


(d) Shadow boundaries for illuminations as above after applying the consistent viewpoint constraint (first three) and both consistent viewpoint and integrability constraints (last three).

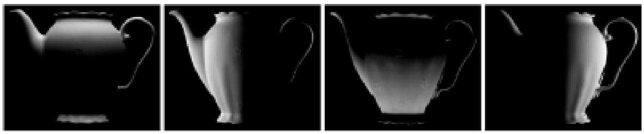
Fig. 6. Results on the WhiteBall object.



(a) Normals and albedo as in Fig. 6(a) (output from factorization).



(b) Normals and albedo as in Fig. 6(b) (consistent viewpoint constraint applied).



(c) Normals illuminated from four directions (top, left, below, right) after applying both consistent viewpoint and integrability constraints. The viewpoint illuminated image and albedo are identical to those in Fig. 7(b).

(d) Shadow boundaries for illuminations as above after applying the consistent viewpoint constraint (first three) and both consistent viewpoint and integrability constraints (last five).

Fig. 7. Results on the ChinaPot object.

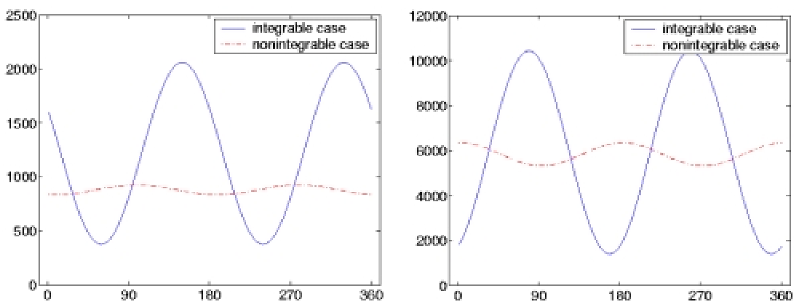


Fig. 8. Integrability violation measure as a function of rotation of normals around the viewing direction ($\xi \in \langle 0, 360)$ [deg]). The WhiteBall object (left) and the ChinaPot object (right). The two plots in each graph (shown in red and blue) correspond to two coordinate frame handednesses. Normals are integrable in only one of them.

computation of pseudonormals and pseudolights as in [5]) was repeated 10 times. The residual (the sum of squared differences between the predicted and observed intensities over the valid image regions) converged to about 1/3 of its initial value. The result of this step is shown in Fig. 6(a) for the WhiteBall object and in Fig. 7(a) for the ChinaPot object. Note neither the illuminated normals nor albedo do correspond to our expectations.

6. Specularly reflecting normals in four images were determined. Specular regions were selected by hand from the set of segmented non-Lambertian regions available from Step 3. The selections are shown in Fig. 5. Pseudonormals from the previous step were averaged over the whole extension of the respective specularity.
7. A transformation **A** was found that maps specular pseudonormals and corresponding pseudolights onto those which fulfill the consistent viewpoint constraint. The idea of the algorithm was to design 'ideal' specular pairs (which follow the consistent viewpoint constraint *exactly*), and look for transformation **A** that maps experimentally obtained specular pseudonormals and pseudolights *closest* (in a least-square sense) to these ideal specular pairs. The algorithm was essentially of the same type as the well-known bundle-adjustment method and we acknowledge the article of Triggs et al. [12] that helped us to design it. The transformation itself was parametrized as $\mathbf{A} = \mathrm{diag}\,[1, \lambda_2, \lambda_3]\,\mathbf{V}^\top$[1]. A unique solution existed in both objects.
8. This transformation was applied to pseudonormals and pseudolights output from Step 5. The consistent viewpoint direction resulting from the previous step was rotated to $[0,\ 0,\ 1]^\top$.

   The results of this step are shown in Figs. 6(b) and 7(b). Note that the resulting albedo is uniform, as expected. It is already the disambiguated albedo of the object. Note also the viewpoint-illuminated normal fields are already symmetric in both objects, as expected.
9. Integrability constraint was applied to resolve the rotation angle ξ and the sign of the first coordinate in transformation (6). The integrability constraint violation measure was constructed as a sum of squared height differences computed over elementary loops in the normal field. Note that for integrable surface the change in height over a closed loop vanishes. The measure was computed for $\xi \in \langle 0, 2\pi)$ on pseudonormals output from Step 8 as well as on these pseudonormals with the x-component reflected (blue and red plots in Fig. 8). The coordinate frame handedness was selected according

[1] This corresponds to the SVD decomposition of **A** (11) with **U** set to the identity and λ_1 set to 1. Specifically, the optimized parameters were: viewpoint direction **v**, four ideal specular normals $\mathbf{n}^S$, and the parameters of transformation **A**. Ideal specular lights were computed by reflecting the (optimized) viewing direction **v** around (optimized) ideal specular normals $\mathbf{n}^S$. Initial parameter values were: $\lambda_2 = \lambda_3 = 1$, **V** initiated randomly (random rotation, solution was observed not to be dependent on this initial choice), ideal specular normals were set to normalized specular pseudonormals, and **v** to normalized average of these. The iterative algorithm converged quickly (in about 15 iterations) into a unique solution.

to which of the two plots gave lower minima. The two minima in such plot correspond to the convex-concave ambiguity which was resolved by hand. The result of this step is shown in Figs. 6(c) (illuminated from left and top directions) and 7(c) (illuminated from four directions as indicated). Note the normals look as expected. This is confirmed by self-shadow boundaries as shown in Fig. 6(d) and 7(d).

4 Implications

In this section we show that there exist problems in computer vision that are formally very similar to the discussed problem of uncalibrated photometric stereo with consistent viewpoint constraint.

In Section 2, we analyzed the problem of uncalibrated photometric stereo, and the role of consistent viewpoint direction in constraining essentially affine ambiguity of uncalibrated photometric stereo. Here we show the validity of the following statement:

> Let affine structure of an object be evaluated by affine geometric stereo with uncalibrated cameras. If four cameras in a general configuration observe specularities being reflected by the object surface from one distant point light source, then the original affine ambiguity reduces into similarity (composition of rotation, isotropic scaling, and change in coordinate frame handedness).

This statement follows from the analysis given in Section 2, where all the results apply if we make substitution *light direction* $\leftrightarrow$ *viewing direction.*

To check the validity of this observation in detail, let there be a surface $\mathbf{X}$ that is parametrized by u and v, $\mathbf{X} = \mathbf{X}(u, v)$. Geometrical stereo with uncalibrated affine projection matrices evaluates the shape up to an affine transformation because the projections $\mathbf{x}^j(u, v)$ of point $\mathbf{X}(u, v)$ in the j-th affine camera $\mathbf{C}^j$ are[2]

$$\mathbf{x}^j(u,v) = \begin{bmatrix} C^j_{1,1} & C^j_{1,2} & C^j_{1,3} \\ C^j_{2,1} & C^j_{2,2} & C^j_{2,3} \end{bmatrix} \begin{pmatrix} X_1(u,v) \\ X_2(u,v) \\ X_3(u,v) \end{pmatrix} = \left[\mathbf{C}^j_1, \mathbf{C}^j_2\right]^\top \mathbf{X}(u,v) = \mathbf{C}^{jT}\mathbf{X}(u,v) \tag{7}$$

and thus $\mathbf{x}^j(u, v)$'s are invariant under transformation $\mathbf{C}^j \mapsto \mathbf{A}\mathbf{C}^j$, $\mathbf{X} \mapsto \mathbf{A}^{-\top}\mathbf{X}$, where $\mathbf{A} \in GL(3)$. It is known (see Yuille et al. [15]) that under these affine transformations the camera viewing vectors $\mathbf{v}_j \sim \mathbf{C}^j_1 \times \mathbf{C}^j_2$ are transformed co-variantly and the surface normals $\mathbf{n} \sim \frac{\partial \mathbf{X}}{\partial u} \times \frac{\partial \mathbf{X}}{\partial v}$ are transformed contravariantly: $\mathbf{v} \sim \mathbf{A}^{-\top}\mathbf{v}$, and $\mathbf{n} \sim \mathbf{A}\mathbf{n}$ ($\sim$ means "up to a scaling factor"). Thus the normals in affine geometrical stereo transform like in photometric stereo, and the camera viewing vectors behave just like the illumination directions. But, in addition,

[2] Origins of image frames in all cameras are aligned.

the specular geometry condition (see Fig. 2) is also symmetrical with respect to the change $\mathbf{l} \leftrightarrow \mathbf{v}$. After we formulate the equivalent of the consistent viewpoint constraint (in this case, it could be called *consistent specular illumination constraint*), the mathematics of the problem is the same. The affine ambiguity is therefore reduced into composition of scaling, rotation around (unknown) illumination direction, and change in coordinate frame handedness; but this is the similarity ambiguity.

5 Conclusions

As a basic result of this paper we have shown that if object reflectance is a sum of Lambertian and specular terms, the uncalibrated photometric stereo ambiguity is reduced into effectively 2dof group of transformations (compositions of rotation around the viewing vector, isotropic scaling and change in coordinate frame handedness). For that, identification of specularities in images corresponding to four different distant point lights in general configuration is sufficient. We expect a similar result will hold if the specular spike is blurred by isotropic surface roughness. This result brings us closer to the practical situation when 'one waves a torch in front of an object and Euclidean structure is revealed.' The good applicability of the approach was verified experimentally on two real objects made of different material.

Note that albedo is obtained without imposing the integrability constraint. The integrability is used to fix only 1dof of the normal field. Since integrability must be computed on normal derivatives, any reduction of the number of parameters to be found significantly improves the accuracy of the resulting normals.

As we noted, lights and cameras play a symmetric role in the consistent viewpoint constraint. Hence, by interchanging lights and cameras, the constraint may also be applied to the case of uncalibrated geometric stereo with four affine cameras in a general configuration observing specularities from a single distant point light source.

Acknowledgement. This work was supported by the Grant Agency of the Czech Republic under Project GACR 102/01/1371 and in part by the Czech Ministry of Education under Research Programme MSM 212300013.

References

1. P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille. The bas-relief ambiguity. *International Journal of Computer Vision*, 35(1):33–44, 1999.
2. E. N. Coleman and R. Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Computer Graphics and Image Processing*, 18(4):309–328, 1982.
3. O. Drbohlav and R. Šára. Unambiguous determination of shape from photometric stereo with unknown light sources. In *Proc. IEEE International Conference on Computer Vision*, volume 1, pages 581–586, 2001.

4. J. Fan and L. B. Wolff. Surface curvature and shape reconstruction from unknown multiple illumination and integrability. *Computer Vision and Image Understanding*, 65(2):347–359, 1997.
5. A. S. Georghiades, D. J. Kriegman, and P. N. Belhumeur. Illumination cones for recognition under variable lighting: Faces. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 52–58, 1998.
6. H. Hayakawa. Photometric stereo under a light source with arbitrary motion. *Journal of Optical Society of America A*, 11(11):3079–3089, 1994.
7. D. Jacobs. Linear fitting with missing data: Applications to structure from motion and to characterizing intensity images. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 206–212, 1997.
8. J. J. Koenderink and A. J. van Doorn. The generic bilinear calibration-estimation problem. *International Journal of Computer Vision*, 23(3):217–234, 1997.
9. S. Lin and S. W. Lee. Estimation of diffuse and specular appearance. In *Proc. IEEE International Conference on Computer Vision*, volume 2, pages 855–860, 1999.
10. S. K. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Journal of Robotics and Automation*, 6(4):418–431, 1990.
11. F. Solomon and K. Ikeuchi. Extracting the shape and roughness of specular lobe objects using four light photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4):449–454, 1996.
12. B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. F. Fitzgibbon. Bundle adjustment — a modern synthesis. In *Proc. Vision Algorithms: Theory and Practice. International Workshop on Vision Algorithms.*, volume 1883 of *Lecture Notes in Computer Science*, pages 298–372, 1999.
13. R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.
14. A. Yuille and D. Snow. Shape and albedo from multiple images using integrability. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 158–164, 1997.
15. A.L. Yuille, J.M. Coughlan, and S. Konishi. The KGBR viewpoint-lighting ambiguity and its resolution by generic constraints. In *Proc. IEEE International Conference on Computer Vision*, volume 2, pages 376–382, 2001.

Appendix

Let $\mathcal{S}$ be a unit sphere in $\mathbb{R}_3$. Let $\mathbf{v}$ be a vector from $\mathcal{S}$ and let $\mathcal{S}^{1/2}$ be a set of unit vectors $\mathbf{b}^S \in \mathcal{S}$ for which it holds that $\mathbf{v} \cdot \mathbf{b}^S \geq 0$ (thus $\mathcal{S}^{1/2}$ is a half-sphere, with $\mathbf{v}$ being its axis). Vector $\mathbf{v}$ represents viewing direction and vectors $\mathbf{b}^S$ represent visible normals. For given normal $\mathbf{b}^S$, let $\mathbf{s}$ denotes the light direction under which $\mathbf{b}^S$ is specular (so that $\mathbf{b}^S$ and $\mathbf{s}$ is a specular pair).

We are asking the following question: if the normals are transformed as $\mathbf{b}^S \mapsto \mathbf{A}\mathbf{b}^S$ and lights as $\mathbf{s} \mapsto \mathbf{A}^{-\top}\mathbf{s}$, what are the only transformations that preserve the consistent viewpoint condition (4)?

First, we showed in Section 2 that this question is equivalent to asking what symmetric positively definite matrices $\mathbf{P} = \mathbf{A}^\top\mathbf{A}$ exist such that

$$\alpha\mathbf{w} = 2(\mathbf{s} \cdot \mathbf{b}^S)\mathbf{P}\mathbf{b}^S - (\mathbf{b}^S \cdot \mathbf{P}\mathbf{b}^S)\mathbf{s}\,, \tag{8}$$

where $\mathbf{w} \in \mathcal{S}$ is some vector which is fixed for all $\mathbf{b}^S \in \mathcal{S}^{1/2}$, and α is a scaling constant (different for each specular pair).

Obviously, the necessary condition for the validity of (8) is that $\forall \mathbf{b}^S \in \mathcal{S}^{1/2}$:

$$\mathbf{w} \in span(\mathbf{P}\mathbf{b}^S, \mathbf{s}) \,, \tag{9}$$

or, equivalently,

$$\mathbf{P}\mathbf{b}^S \in span(\mathbf{w}, \mathbf{s}) \,. \tag{10}$$

But $\mathbf{P}$ is symmetric and positively definite and thus its effect on $\mathbf{b}^S$ represents anisotropic scaling in arbitrary three orthogonal directions. To see that, let us write the SVD decomposition of $\mathbf{A}$ and $\mathbf{P}$:

$$\mathbf{A} = \mathbf{U}\mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)\mathbf{V}^\top \quad \lambda_1, \lambda_2, \lambda_3 > 0; \quad \mathbf{U}, \mathbf{V} \in O(3) \,, \tag{11}$$

$$\mathbf{P} = \mathbf{V}\mathrm{diag}(\lambda_1^2, \lambda_2^2, \lambda_3^2)\mathbf{V}^\top \,. \tag{12}$$

Thus $\mathbf{P}$ scales along the direction of eigenvector $\mathbf{V}_i$ (i-th column of $\mathbf{V}$) by the respective λ_i.

Then normals $\mathbf{b}_i^S = \pm\mathbf{V}_i$ (where $\pm$ is properly selected according to whether $\mathbf{V}_i$ is or is not in $\mathcal{S}^{1/2}$) that are specular under corresponding lights $\mathbf{s}_i$ are mapped onto themselves (up to a scale), and consequently $\mathbf{w} \in span(\mathbf{b}_i^S, \mathbf{s}_i), i = 1 \ldots 3$. That implies that $\mathbf{w} = span(\mathbf{b}_1^S, \mathbf{s}_1) \cap span(\mathbf{b}_2^S, \mathbf{s}_2) \cap span(\mathbf{b}_3^S, \mathbf{s}_3) = \mathbf{v}$. But (10) must hold for all $\mathbf{b}^S$ and we must therefore require $\mathbf{P}\mathbf{b}^S \in span(\mathbf{v}, \mathbf{s}) = span(\mathbf{v}, \mathbf{b}^S)$. The only way to arrange it is to align one of the scaling directions (say, $\mathbf{V}_1$) with $\mathbf{v}$, and to set the scalings along the other two directions equal ($\lambda_2 = \lambda_3$). Next, we show that all λ_i's must be equal.

Let us complete (11) and (12) by writing decomposition of $\mathbf{A}^{-\top}$ as

$$\mathbf{A}^{-\top} = \mathbf{U}\mathrm{diag}(\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \frac{1}{\lambda_3})\mathbf{V}^\top \,. \tag{13}$$

Observe that a particular choice of the matrix $\mathbf{U}$ has no effect on the validity of the consistent viewpoint constraint, since it only transforms both $\mathbf{s}$ and $\mathbf{b}^S$

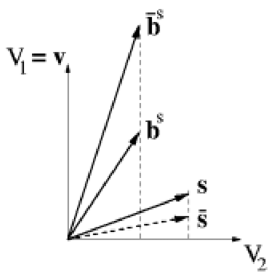


Fig. 9. Scaling the $\mathbf{V}_1$ component of normals by λ_1 while scaling the same component of lights by $1/\lambda_1$ results in that the new specular pair violates consistent viewpoint constraint if $\lambda_1 \neq 1$.

by a global orthogonal transformation (cf. (11)). It therefore suffices to consider the effect of transformations (11) and (13) on $\mathbf{b}^S$ and $\mathbf{s}$, respectively, with $\mathbf{U}$ set to identity.

Fig. 9 shows one pair of specular normal and corresponding light source before ($\mathbf{b}^S$,$\mathbf{s}$) and after ($\overline{\mathbf{b}}^S$,$\overline{\mathbf{s}}$) the photometric stereo transformation. Without the loss of generality we set $\lambda_2 (= \lambda_3) = 1$. The figure illustrates the fact that when, for example, $\lambda_1 > 1$, then the transformed normal makes a smaller angle with $\mathbf{v}$ than the original normal, while the transformed light makes a greater angle with $\mathbf{v}$ as compared with the original light. From that it follows that for transformed normals and lights $\mathbf{v}$ is not consistent with the specular geometry condition unless it holds that $\lambda_1 = \lambda_2 = \lambda_3$.

Thus we have the result that $\mathbf{P}$ may be only the scaled identity $\mathbf{P} = \lambda^2 \mathbf{I}$.