A MAXIMUM LIKELIHOOD SURFACE NORMAL ESTIMATION ALGORITHM FOR HELMHOLTZ STEREOPSIS

Jean-Yves Guillemaut[1], Ondřej Drbohlav[2], John Illingworth[1] and Radim Šára[3]

[1]*Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, Surrey, GU2 7XH, UK*

[2]*Visual Cognitive Systems Laboratory, University of Ljubljana, Tržaška 25, SI-1001 Ljubljana, Slovenia*

[3]*Center for Machine Perception, Czech Technical University, Karlovo náměstí 13, Prague 2, Czech Republic*

Keywords: Helmholtz Stereopsis, Photometry, Image-Based Object Reconstruction.

Abstract: Helmholtz stereopsis is a relatively recent reconstruction technique which is able to reconstruct scenes with arbitrary and unknown surface reflectance properties. Conventional implementations of the method estimate surface normal direction at each surface point via an eigenanalysis, thereby optimising an algebraic distance. We develop a more physically meaningful radiometric distance whose minimisation is shown to yield a Maximum Likelihood surface normal estimate. The proposed method produces more accurate results than algebraic methods on synthetic imagery and yields excellent reconstruction results on real data. Our analysis explains why, for some imaging configurations, a sub-optimal algebraic distance can yield good results.

1 INTRODUCTION

Most image-based reconstruction techniques assume that the surface reflectance of the scene is known or follows a parametric model. Conventional binocular stereo relies on the assumption that the scene is Lambertian, i.e. the appearance of a surface point is independent of the viewing direction. Photometric stereo can cope with non Lambertian reflectance models, but still relies on a known parametric surface reflectance model. Shape-from-silhouette does not make any assumption on surface reflectance but the obtainable reconstruction is only approximate (limited to the visual hull) and often requires capture of a large number of images. In contrast, Helmholtz stereopsis (HS), introduced in (Magda et al., 2001; Zickler et al., 2002a), has the unique property of being able to reconstruct objects with arbitrary and unknown surface reflectance properties.

HS is based on the Helmholtz reciprocity principle, which states that at any surface point the surface reflectance, represented by its bidirectional reflectance distribution function (BRDF) (Nicodemus et al., 1977), remains unchanged when the incident and reflected directions are interchanged. The method requires a camera and a point light source whose positions can be interchanged in order to produce recip-

rocal pairs of images. The method was first implemented in the case of a general multi-camera setup (Magda et al., 2001; Zickler et al., 2002a; Zickler et al., 2002b) and then in a binocular configuration (Tu and Mendonça, 2003; Zickler et al., 2003b). In (Zickler et al., 2003a; Zickler, 2006), HS was later extended to uncalibrated cameras. Other research concentrated on improving the accuracy of HS by radiometric calibration (Jankó et al., 2004), as well as generalising the method to structured and strongly textured surfaces (Guillemaut et al., 2004). Applications of the Helmholtz reciprocity principle have also been considered in the contexts of surface registration (Tu et al., 2003) and computer graphics (Sen et al., 2005).

HS produces a scene reconstruction consisting of a depth estimate and surface normal estimate at each pixel in a reference view. This yields a 2.5D representation of the scene from which the surface can be extracted by integration of the normal field. Accurate surface normal estimation at each pixel is therefore critical. In previous implementations, reciprocal pairs of images were used to define a set of linear constraints, from which the normal is estimated by finding the vector best satisfying all constraints. Singular value decomposition (SVD) (Press et al., 1992) has been used for such task, because it allows computation of a global solution at a low computational cost.

A drawback however is that the algebraic error function thus defined often lacks a physical meaning and is not optimum in a statistical sense.

In contrast, in this paper we propose to formulate the surface normal estimation problem in terms of minimising a distance that we call the radiometric distance. This distance is physics-based and measures the intensity modification to be applied in order to satisfy exactly the Helmholtz reciprocity constraint at a surface point. Under standard Gaussian image noise conditions, its minimisation yields a maximum likelihood (ML) surface normal estimate. We demonstrate that the computation of this optimum solution can be done efficiently by minimising a non-linear cost function with two variables only. The proposed algorithm also incorporates treatment of image saturation, a phenomenon which has often been ignored in previous work.

The paper is structured as follows. We start by giving an overview of HS, where we describe the conventional surface normal estimation algorithm. In the following section, we propose a novel radiometric distance and an efficient minimisation algorithm for deriving the surface normal. Finally, we present results on synthetic and real imagery, and conclude.
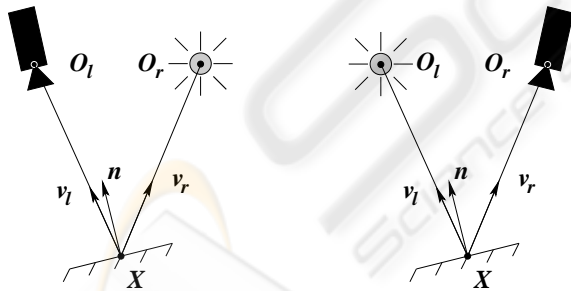
2 OVERVIEW OF HELMHOLTZ STEREOPSIS

Figure 1: A reciprocal configuration.

Let us consider the configuration illustrated in Figure 1. O_l and O_r are two points in space and X is a surface point. We denote by $d_l = \|O_l - X\|$ and $d_r = \|O_r - X\|$ the distances from O_l and O_r respectively to X, and define $v_l = \frac{1}{d_l}(O_l - X)$ and $v_r = \frac{1}{d_r}(O_r - X)$, which represent the unit vectors pointing from X to O_l and O_r respectively. The surface normal at X is represented by the unit vector n. The BRDF $f(X, u, v)$ at X is by definition the ratio of the outgoing radiance along the direction v to the incident irradiance along

the direction u. If we position an isotropic light source of intensity κ at O_l and a camera at O_r, the pixel intensity[1] i_r observed by the camera is:

$$i_r = f(X, v_l, v_r) \frac{v_l \cdot n}{d_l^2} \kappa. \tag{1}$$

If the positions of the light source and the camera are now swapped, an analogous formula is obtained for the radiance i_l observed by the camera at position O_l:

$$i_l = f(X, v_r, v_l) \frac{v_r \cdot n}{d_r^2} \kappa. \tag{2}$$

The two images captured by such cameras form a reciprocal pair. The Helmholtz reciprocity principle states that $f(X, v_l, v_r) = f(X, v_r, v_l)$. Combining Eq. (1) and Eq. (2), and denoting $s_l = \frac{1}{d_l^2} v_l$ and $s_r = \frac{1}{d_r^2} v_r$, we obtain (Zickler et al., 2002b):

$$(i_l s_l - i_r s_r) \cdot n = 0. \tag{3}$$

The vector $(i_l s_l - i_r s_r)$ must lie in the tangent plane at the surface point and the normal vector n can, in principle, be estimated from two or more such vectors (derived from two or more reciprocal pairs).

2.1 Surface Points Identification

Surface points are represented by a depth value at each pixel in a reference view. If at least $N \geq 3$ constraints defined by Eq. (3) are available (one per reciprocal pair of images), it is possible to define a measure for identifying such points. At each candidate surface point, we can form the following system of $N \geq 3$ equations (Zickler et al., 2002b):

$$Wn = 0 \quad \text{with} \tag{4}$$
$$W = \begin{bmatrix} i_{l_1} s_{l_1} - i_{r_1} s_{r_1}, \ldots, i_{l_n} s_{l_n} - i_{r_n} s_{r_n} \end{bmatrix}^\top.$$

The distribution of row vectors in W is then used to identify surface points as follows. If the intensities used for constructing the matrix W come from a surface point, these vectors must be coplanar and, therefore, the matrix W must be of rank 2. If the point is not a surface point, the rows of the matrix W are likely not to conform to this constraint. In practice, the rank 2 constraint is never exactly satisfied, even at surface points, because of image noise. For this reason, an alternative measure of rank, which we call the support measure, is defined by:

$$s = 1 - \frac{\sigma_3}{\sigma_2}, \tag{5}$$

---

[1]We adopt the convention that the pixel intensity is equal to the scene radiance. This is usually a reasonable assumption for high quality cameras. If it is not the case, radiometric calibration can be performed in order to meet this requirement (Jankó et al., 2004).

where σ_2 and σ_3 denote the second and third singular values of W. It is assumed here that the SVD of W is written in the form:

$$W = UDV^\top \quad \text{with} \tag{6}$$
$$D = \text{diag}(\sigma_1, \sigma_2, \sigma_3) \text{ and } \sigma_1 \geq \sigma_2 \geq \sigma_3 \geq 0.$$

The measure defined in Eq. (5) is strictly equivalent to the measure defined in (Zickler et al., 2002b), but has the advantage of normalising the value of the support measure between 0 and 1, a value close to 1 corresponding to a high chance of the point being located on a surface.

The conditions given in Eq. (4) are necessary, but not sufficient, for a point to be located at the surface. It can, therefore, happen that the support measure defined in Eq. (5) is large although the point does not belong to the surface. A possibility to resolve this ambiguity is to optimise the measure over a small window instead of a single pixel, assuming that the surface can be locally represented by a fronto-parallel planar patch as in (Zickler et al., 2002b). Surface points are then identified by maximising the support measure along each ray in the reference image sampled with regular depth increments. This yields a 2.5D surface reconstruction represented by a depth map. Typically the depth map obtained may be noisy, but a more accurate estimate of the object geometry can be obtained by computing the surface normal at each point and then integrating the obtained normal field - this is a standard procedure used in photometric stereo.

2.2 Surface Normal Estimation

Given some identified surface points, previous approaches (Magda et al., 2001; Zickler et al., 2002b) have estimated the normal n at each point as the third singular vector, i.e. the last column vector of V, from the SVD of W expressed in Eq. (6). It can be shown (see e.g. (Hartley, 1997)) that the solution obtained is the vector n which minimises

$$\|W \cdot n\|^2 = \sum_j [(i_{l_j} s_{l_j} - i_{r_j} s_{r_j}) \cdot n]^2 \text{ subject to } \|n\| = 1. \tag{7}$$

This cost function can be written in the form

$$\sum_j d_{\text{alg}}(n, i_{l_j}, i_{r_j}, s_{l_j}, s_{r_j})^2, \tag{8}$$

where d_{alg} denotes the algebraic distance

$$d_{\text{alg}}(n, i_{l_j}, i_{r_j}, s_{l_j}, s_{r_j})^2 = [(i_{l_j} s_{l_j} - i_{r_j} s_{r_j}) \cdot n]^2 \tag{9}$$
$$= \|i_{l_j} s_{l_j} - i_{r_j} s_{r_j}\|^2 \cos^2 \alpha_j.$$

In this expression, α_j denotes the angle between the vector $(i_{l_j} s_{l_j} - i_{r_j} s_{r_j})$ and the surface normal n. The term $\cos^2 \alpha_j$ is small when the constraint Eq. (3) is

close to be satisfied and, conversely, large when the constraint is far from being satisfied; this term therefore intuitively corresponds to an entity to be minimised. However, the physical interpretation of the scaling factor $\|i_{l_j} s_{l_j} - i_{r_j} s_{r_j}\|^2$ is not obvious. The influence of this term could be eliminated by normalising the rows of W to unit values, however it is possible that these coefficients play an important role by attenuating the contribution of measurements corresponding to low intensities or cameras/light sources located far away from the scene point. In any case, the SVD-based solutions, considered in previous works mainly for simplicity reasons, did not explicitly address the nature of errors affecting the measurements, and therefore, have no guarantee of optimality. In the next section, we propose a novel measure, which we prove to be optimum under the assumption that errors affecting the normal computation are due to Gaussian noise in intensity measurements.

3 RADIOMETRIC CONSTRAINT

3.1 Definition

Since the fundamental entities observed are image intensities (or equivalently, radiances), it is a natural choice to perform the minimisation directly in the space of radiances. We define an optimum solution by searching for the surface normal n and the intensity pairs $\{\hat{i}_{l_j}, \hat{i}_{r_j}\}_j$ which minimise the cost function[2]:

$$\sum_j [(\hat{i}_{l_j} - i_{l_j})^2 + (\hat{i}_{r_j} - i_{r_j})^2] \tag{10}$$
$$\text{subject to} \quad \forall j \quad (\hat{i}_{l_j} s_{l_j} - \hat{i}_{r_j} s_{r_j}) \cdot n = 0.$$

This cost function measures the correction to be made in the intensities observed in each reciprocal pair of images in order to fulfil exactly the constraints defined in Eq. (4). If we assume that the error in geometric calibration of the camera and light source is negligible compared to the intensity measurement error, and that the intensity measurement errors are independent and follow Gaussian distributions, the resulting cost function defines a sum of squared errors affected only by Gaussian noise, and its minimisation leads to an ML estimate.

After elimination of the constraints by substitution

---

[2]Note that for conciseness, we use the notation $\{\hat{i}_{l_j}\}_j$ instead of $\{\hat{i}_{l_j}\}_{j \in \{1,\ldots,N\}}$. Whenever the range has been omitted in this paper, it should be understood that the range is $\{1,\ldots,N\}$.

into the cost function, we obtain:

$$\sum_j [(\hat{i}_{l_j} - i_{l_j})^2 + (\frac{s_{l_j} \cdot n}{s_{r_j} \cdot n}\hat{i}_{l_j} - i_{r_j})^2], \qquad (11)$$

where the variables to optimise are n and $\{\hat{i}_{l_j}\}_j$. This defines a priori a complex minimisation problem with $3 + N$ unknowns, N being the number of reciprocal image pairs. We demonstrate in the appendix that this minimisation problem can be simplified into an equivalent problem where n is the only unknown. The obtained cost function to minimise is

$$\sum_j d_{\text{rad}}(n, i_{l_j}, i_{r_j}, s_{l_j}, s_{r_j})^2, \qquad (12)$$

where d_{rad} denotes the novel radiometric distance introduced and defined by:

$$d_{\text{rad}}(n, i_{l_j}, i_{r_j}, s_{l_j}, s_{r_j})^2 = \frac{[(i_{l_j}s_{l_j} - i_{r_j}s_{r_j}) \cdot n]^2}{(s_{l_j} \cdot n)^2 + (s_{r_j} \cdot n)^2}. \quad (13)$$

This defines a simple minimisation problem with only two degrees of freedom (the norm of n must be one). The solution can be computed using a non-linear optimisation algorithm such as the Levenberg-Marquardt algorithm (Press et al., 1992). The algorithm can be initialised with the result of the SVD solution and usually converges rapidly after only a few iterations. To be physically valid, the solution obtained must satisfy the visibility constraints stating that $s_l \cdot n > 0$ and $s_r \cdot n > 0$. Although these constraints can be enforced during minimisation of the cost function, we found it simpler and sufficient to verify that the visibility constraints are satisfied after convergence of the search algorithm. Note that these constraints also guarantee that the denominators of the equations considered in this section and in the appendix are non-zero.

3.2 Comparison with the Algebraic Constraint

From Eq. (9) and Eq. (13), we have:

$$d_{\text{rad}}(n, i_{l_j}, i_{r_j}, s_{l_j}, s_{r_j})^2 = \qquad (14)$$
$$\frac{1}{(s_{l_j} \cdot n)^2 + (s_{r_j} \cdot n)^2}d_{\text{alg}}(n, i_{l_j}, i_{r_j}, s_{l_j}, s_{r_j})^2.$$

Although the interpretation of the multiplicative factor relating the two measures does not seem immediately obvious, it is apparent that the difference between the two constraints is due only to the positions of the camera and light source relatively to the surface point, and does not involve the measured intensities. The two measures will produce identical normal estimates whenever the term $(s_{l_j} \cdot n)^2 + (s_{r_j} \cdot n)^2$ is constant for all j. One particular situation when this occurs in practice is for horizontal surfaces located at the centre of the turntable used in our experimental set up (see Section 4.2).

3.3 Treatment of Image Saturation

Specularities can be difficult to capture due to the limited range of the camera sensor and often result in saturated pixels. This affects the constraints previously defined because the measured intensities are no longer proportional to the true radiance. So far, very little attention has been paid to this problem apart from (Tu and Mendonça, 2003) where a solution is reported in the case of binocular HS. We propose a similar treatment of image saturation in the multiocular case, and adapt the radiometric distance accordingly.

As in (Tu and Mendonça, 2003), we observe that when a saturation is present in a reciprocal pair of images (usually it is observed simultaneously in both images at reciprocal positions), the normal bisects the incident and emerging rays, and we have

$$(\frac{1}{\|s_{l_j}\|}s_{l_j} - \frac{1}{\|s_{r_j}\|}s_{r_j}) \cdot n = 0. \qquad (15)$$

This constraint must replace Eq. (3) whenever a saturation is observed. In this case, the appropriate row of W in Eq. (4) is replaced by $(\frac{i_{\text{sat}}}{\|s_{l_j}\|}s_{l_j} - \frac{i_{\text{sat}}}{\|s_{r_j}\|}s_{r_j})^\top$ before computing the SVD, and for the radiometric constraint, Eq. (13) is replaced by:

$$d_{\text{rad}}(n, i_{l_j}, i_{r_j}, s_{l_j}, s_{r_j})^2 = [(\frac{i_{\text{sat}}}{\|s_{l_j}\|}s_{l_j} - \frac{i_{\text{sat}}}{\|s_{r_j}\|}s_{r_j}) \cdot n]^2. \qquad (16)$$

In these expressions, the value i_{sat}, representing the pixel saturation intensity value, is necessary to guarantee that the dimension is the same as in Eq. (3) or Eq. (13). This distance is similar to the cost function defined by Tu and Mendonça in (Tu and Mendonça, 2003) in the case of binocular HS.

4 RESULTS

4.1 Synthetic Data

We conduct a simple experiment to compare the performances of the proposed radiometric distance with the algebraic distance. Two implementations are considered for the algebraic distance described in Section 2.2. The first one, which was adopted in (Zickler et al., 2002b), uses the matrix W as defined in Eq. (4) for computation of the SVD and is called *unnormalised algebraic*; the second implementation proceeds similarly but with rows normalised to unit values and is called *normalised algebraic*. The issue of normalisation has been discussed earlier in Section 2.2. We consider the reconstruction of a surface patch with ground truth normal n whose BRDF

is modelled by the modified Phong reflectance model described in (Lewis, 1994; Lafortune and Willems, 1994). The model represents the BRDF as the sum of a diffuse component and a specular component:

$$f_{n,k_d,k_s}(X,v_l,v_r) = k_d \frac{1}{\pi} + k_s \frac{n+2}{2\pi} \cos^n \alpha, \qquad (17)$$

where α denotes the angle between the perfect specular reflective direction and the emerging direction. The three parameters n, k_d and k_s represent respectively the specular exponent (large values result in sharper specular reflections), the diffuse reflectivity and the specular reflectivity. In our experiment, we considered the values $n = 40$, $k_d = 0.4$ and $k_s = 0.05$. There exist more sophisticated models which could have been considered, such as the one described in (Cook and Torrance, 1982), but it is important to note that there is no "perfect" model, all existing models being restricted to a particular class of surfaces. Also it was observed in Section 3.2 that the relative performance of algebraic and radiometric algorithms should be independent of the surface model, therefore the choice of the reflectance model is not critical as long as it is physically plausible. We chose the modified Phong model because it is simple to implement and physically plausible (it satisfies energy conservation and Helmholtz reciprocity).

4.1.1 Simple Camera/light Configuration

We consider a simple case of eight pairs of reciprocal images captured with cameras and light sources placed regularly on a circle (and thus separated by $22.5°$ increments). The plane of the circle is assumed to be horizontal, and the surface reference point is placed on the axis of rotation of the circle such that the incident and emerging rays all make an angle of $30°$ with the vertical direction. This configuration is important in practice because it occurs e.g. when an object placed on a turntable is captured with fixed light source and camera. A similar experimental set-up has been considered for the acquisition of real data in this paper and in (Zickler et al., 2002b). For a given surface normal orientation, represented by the inclination angle with respect to the vertical direction, we generate radiance values for each reciprocal pair, and perturb them by adding zero-mean Gaussian noise with standard deviation σ. The strength of the light source is constant and equal to $\kappa = 1\,000$. The inclination angle of the surface normal is varied from 0 to $45°$, by $1°$ increments, and at each setting the experiment is repeated 10 000 times. For evaluation, we compute the root mean squared (RMS) angular error between the set of estimated normals $\hat{n}_{ij}$ and true normals $\bar{n}_{ij}$
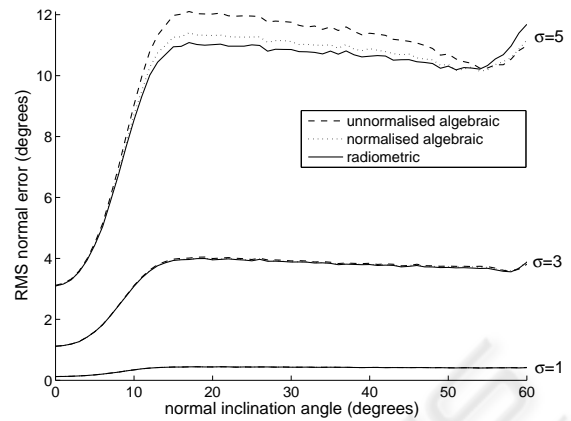

Figure 2: Influence of the inclination angle of the surface normal on the accuracy of the normal reconstruction.

defined by $\sqrt{\frac{1}{N}\sum_i \sum_j \theta(\bar{n}_{ij},\hat{n}_{ij})^2}$, where $\theta(\bar{n}_{ij},\hat{n}_{ij})$ denotes the angles between the vectors $\bar{n}_{ij}$ and $\hat{n}_{ij}$.

The results are shown in Figure 2, for three different image noise levels corresponding to $\sigma = 1$, 3 and 5. We can observe that the radiometric and algebraic methods perform very closely. In the case of $\sigma = 1$ or 3, the curves seem to overlap and are difficult to distinguish. Magnification would show that the radiometric distance is the most accurate, however the improvement remains very small. With more severe noise conditions ($\sigma = 5$), the difference in error increases to the order of one degree. This shows that the algebraic solution is a very good estimate (almost as good as the ML estimate) in spite of possibly large noise conditions for this particular camera/light configuration. In this case, the distance separating the camera/light source to the surface point is constant, and there is little incident and emerging angle variations. As a consequence, the multiplicative factors appearing in Eq. (14) do not exhibit large variations which would otherwise penalise the algebraic distance. In the case of a surface normal oriented vertically, these factors are exactly equal, and identical results are expected from both methods. The discrepancy between the different methods is expected to increase with the inclination angle of the normal. This can be verified in Figure 2.

4.1.2 General Camera/light Configuration

We now consider a more general set-up which allows more freedom in camera and light source placement. This is the kind of experimental set-up that could be obtained if a camera and light source were mounted on a robot arm allowed to move freely around the scene. A surface patch with ground truth normal n is considered. N pairs of points O_l and O_r are generated
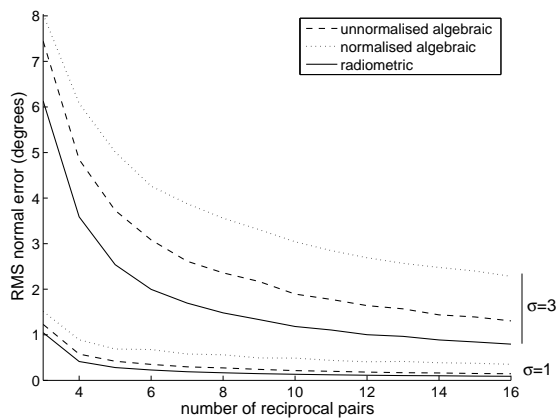
Figure 3: Influence of the number of reciprocal image pairs considered on the accuracy of the normal reconstruction.

randomly; they define N reciprocal pairs of images (here radiance values) of the surface patch. The points are constrained to be located on the same side of the patch such that the visibility constraint is satisfied. The distance from the points to the patch is selected randomly within the interval $[0.2, 1]$. The random positions are obtained by generating points with random spherical coordinates (r, θ, ϕ) in the respective intervals $[0.2, 1]$, $[10°, 80°]$ and $[0, 360°]$. The radiance values generated are perturbed by a zero-mean Gaussian noise with standard deviation σ. The strength of the light source is constant and equal to $\kappa = 1\,000$.

We considered two different noise levels $\sigma = 1$ and $\sigma = 3$, and the number of reciprocal pairs was allowed to vary between 3 (minimum number supported by the method) and 16. The experiment is repeated 10 000 times, and the RMS angular error between the estimated normal and the true normal is computed for each method. The results are shown in Figure 3. The method based on the radiometric distance is always more accurate than the methods based on the algebraic distances, with the normalised version slightly more accurate than the unnormalised version. The fact that the method based on the radiometric distance leads to the best results is not surprising since it is an ML estimator under the assumed noise model.

4.2 Real Data

The experimental setup consists of a camera, a light source and a turntable which performs the interchange of camera and light source positions. A 12 bit digital camera Vosskuhler CCD-1300 equipped with a 25 mm lens was used along with a halogen lamp equipped with a diaphragm and acting as a point light source. The distance between the camera and the centre of the table is approximately 80 cm and the dis-

tance between the camera and the light source 60 cm. The camera is geometrically calibrated. Experiments were carried out with four objects: a snooker ball, two teapots, and a terracotta doll (see first row of Figure 4). The first three object have specular surfaces, while the fourth one is Lambertian. The reconstruction pipeline is similar to the one described in (Zickler et al., 2002b), with the exception that we consider three different measures for surface normal estimation (unnormalised algebraic, normalised algebraic and radiometric). Eight reciprocal pairs of images are generated by rotating the turntable by regular 22.5° increments. In order to restrict the search for surface points, a bounding box is defined for each object. The bounding box is discretised into square voxels of resolution 1 mm × 1 mm × 1 mm in the case of the snooker ball and the doll, and 2 mm × 2 mm × 2 mm for the other objects. A window of size 5 × 5 pixels was used during depth search for all objects.

The results obtained for all objects are shown in Figure 4 in the case of the method based on the radiometric distance. The normal field is represented as a needle map shown in the second row of Figure 4. For each object, a 3D models is obtained by integration of the normal field weighted by the corresponding support measure using the method reported in (Terzopoulos, 1982). Views of the obtained models, rendered untextured with Phong shading, are shown in the third row of Figure 4. Although the objects are very challenging to reconstruct (especially the first three for which the surfaces are highly specular), the algorithm produces smooth and visually correct reconstructions in all cases. A few small artefacts are visible at the location of the specularities which are not as localised as expected in theory and result in larger areas of images saturation. The extended saturation observed are due to the fact that the point light source assumption is not exactly satisfied in practice, and also because the surface is not an ideal specular reflector (mirror surface). Other minor artefacts, due to occlusions essentially, are visible at the object boundaries. We would like to emphasise that our implementation is fully automatic and that no attempt was made to eliminate such artefacts by editing manually the data.

The results obtained by the algebraic methods are very close to the results obtained by the radiometric method and are not shown here due to space limitation. Unfortunately, there is no ground truth information available to allow for a formal evaluation of the performance of the different methods. Instead we compared the set of normals obtained for the algebraic and radiometric methods by computing the RMS, median and maximum angular differences. The results are reported in Table 1. The differences are

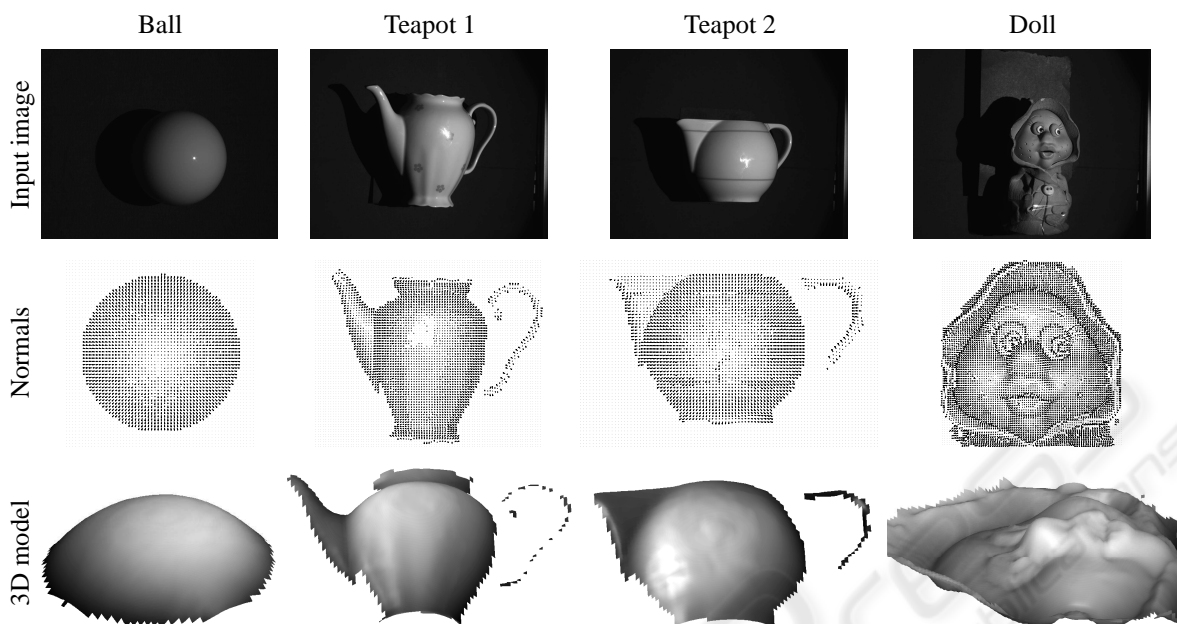|	Ball	Teapot 1	Teapot 2	Doll

Figure 4: Reconstruction results with some real objects.

very small, except for the maximum difference in which case the larger values are due mostly to occlusions around the object contour. The observation that the results obtained by the radiometric and algebraic methods are similar is compatible with our previous observation that for this particular camera/light placement, the algebraic constraint is a good approximation of the radiometric constraint.

Table 1: Angular differences between the normals computed with radiometric and unnormalised algebraic methods. The values are in degrees.

	Ball	Teapot 1	Teapot 2	Doll
RMS	0.0219	0.997	1.12	1.87
median	0.0104	0.0077	0.0086	0.0306
max	0.226	25.0	37.8	63.5

5 CONCLUSIONS

In this paper, we proposed a novel method for surface normal reconstruction using HS. The method is based on the minimisation of a cost function which consists of squared radiometric distances summed over all reciprocal pairs of images in which a surface point is visible. Physically, the cost function represents the modification to be applied to the intensities of the projected surface point in order to satisfy exactly the Helmholtz reciprocity principle. The normal computed by this method has been shown to be an ML estimate under standard Gaussian noise assumption. A solution can be computed at low computational cost due to the low number of optimisation variables.

In the case of synthetic data, it was verified experimentally that the method based on the radiometric distance results in more accurate surface normal estimates than methods based on SVD. Experiments carried out with real imagery showed that the proposed method produces similar results to SVD-based methods for the experimental set-up considered, and is able to generate realistic 3D models for a variety of objects with challenging surface properties.

In future work, we would like to build a more versatile experimental set-up which would allow more freedom in camera/light source placement, by e.g. mounting a pair of camera/light source on a robot arm. Such a set-up would allow reconstruction of a full 3D model instead of the current 2.5D representation. In this scenario, we would also expect larger differences in performance between the algebraic methods and the radiometric method, and we may be able to measure the optimality of the radiometric method under real noise conditions.

ACKNOWLEDGEMENTS

REFERENCES

Cook, R. L. and Torrance, K. E. (1982). A reflectance model for computer graphics. *ACM Transactions on Graphics*, 1(1):7–24.

Guillemaut, J.-Y., Drbohlav, O., Šára, R., and Illingworth, J. (2004). Helmholtz stereopsis on rough and strongly textured surfaces. In *Proc. International Symposium on 3D Data Processing, Visualization and Transmission*, pages 10–17.

Hartley, R. I. (1997). In defence of the eight-point algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(6):580–593.

Jankó, Z., Drbohlav, O., and Šára, R. (2004). Radiometric calibration of a Helmholtz stereo rig. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 166–171.

Lafortune, E. P. and Willems, Y. D. (1994). Using the modified phong reflectance model for physically based rendering. Technical Report CW 197, Department of Computing Science, K.U. Leuven, November, 1994.

Lewis, R. R. (1994). Making shaders more physically plausible. *Computer Graphics Forum*, 13(2):109–120.

Magda, S., Kriegman, D. J., Zickler, T., and Belhumeur, P. N. (2001). Beyond Lambert: Reconstructing surfaces with arbitrary BRDFs. In *Proc. IEEE International Conference on Computer Vision*, volume 2, pages 391–398.

Nicodemus, F. E., Richmond, J. C., Hsia, J. J., Ginsberg, I. W., and Limperis, T. (1977). Geometrical consideration and nomenclature for reflectance. NBS Monograph 160.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical recipes in C*. Cambridge University Press, second edition.

Sen, P., Chen, B., Garg, G., Marschner, S. R., Horowitz, M., Levoy, M., and Lensch, H. P. A. (2005). Dual photography. *ACM SIGGRAPH*, 24(3):745–755.

Terzopoulos, D. (1982). Multi-level reconstruction of visual surfaces: Variational principles and finite element representations. A.I. Memo 671, MIT.

Tu, P. and Mendonça, P. R. S. (2003). Surface reconstruction via Helmholtz reciprocity with a single image pair. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 541–547.

Tu, P., Mendonça, P. R. S., Ross, J., and Miller, J. (2003). Surface registration with a Helmholtz reciprocity image pair. In *IEEE Workshop on Color and Photometric Methods in Computer Vision*.

Zickler, T. (2006). Reciprocal image features for uncalibrated helmholtz stereopsis. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 1801–1808.

Zickler, T., Belhumeur, P. N., and Kriegman, D. J. (2002a). Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. In *Proc. European Conference on Computer Vision*, volume III, pages 869–884.

Zickler, T. E., Belhumeur, P. N., and Kriegman, D. J. (2002b). Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. *International Journal of Computer Vision*, 49(2/3):215–227.

Zickler, T. E., Belhumeur, P. N., and Kriegman, D. J. (2003a). Toward a stratification of Helmholtz stereopsis. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 548–555.

Zickler, T. E., Ho, J., Kriegman, D. J., Ponce, J., and Belhumeur, P. N. (2003b). Binocular Helmholtz stereopsis. In *Proc. IEEE International Conference on Computer Vision*, pages 1411–1417.

APPENDIX

Simplification of the Cost Function based on the Radiometric Distance

We would like to compute the values n and $\{\hat{i}_{l_j}\}_j$ which minimise the cost function F defined by:

$$F(n, \{\hat{i}_{l_j}\}_j) = \sum_j [(\hat{i}_{l_j} - i_{l_j})^2 + (\frac{s_{l_j} \cdot n}{s_{r_j} \cdot n}\hat{i}_{l_j} - i_{r_j})^2]. \quad (18)$$

We start by writing the partial derivatives of F with respect to $\hat{i}_{l_k}$; for any index k in $\{1, \ldots, N\}$, we have:

$$\frac{\partial F(n, \{\hat{i}_{l_j}\}_j)}{\partial \hat{i}_{l_k}} = 2(\hat{i}_{l_k} - i_{l_k}) + 2\frac{s_{l_k} \cdot n}{s_{r_k} \cdot n}(\frac{s_{l_k} \cdot n}{s_{r_k} \cdot n}\hat{i}_{l_k} - i_{r_k}). \quad (19)$$

At the optimum, the partial derivatives must be zero; this results in the following constraints:

$$\forall k, \quad \hat{i}_{l_k} + (\frac{s_{l_k} \cdot n}{s_{r_k} \cdot n})^2 \hat{i}_{l_k} = i_{l_k} + \frac{s_{l_k} \cdot n}{s_{r_k} \cdot n}i_{r_k}, \quad (20)$$

from which, we can deduce that, at the optimum, the values of the variables $\hat{i}_{l_k}$ are:

$$\forall k, \quad \hat{i}_{l_k} = \frac{i_{l_k}(s_{r_k} \cdot n)^2 + i_{r_k}(s_{l_k} \cdot n)(s_{r_k} \cdot n)}{(s_{l_k} \cdot n)^2 + (s_{r_k} \cdot n)^2}. \quad (21)$$

After substitution of these values into Eq. (18), we obtain the simplified cost function G defined by:

$$G(n) = \sum_j \frac{[(i_{l_j}s_{l_j} - i_{r_j}s_{r_j}) \cdot n]^2}{(s_{l_j} \cdot n)^2 + (s_{r_j} \cdot n)^2}. \quad (22)$$

G defines a simpler cost function with single unknown n (two degrees of freedom) which has the same global minimum as the original cost function F.