

Scene Depth Profiling Using Helmholtz Stereopsis

Hironori Mori^(✉), Roderick Köhle, and Markus Kamm

Sony European Technology Center, Stuttgart, Germany
{Hironori.Mori,Roderick.Koehle,Markus.Kamm}@eu.sony.com

Abstract. Helmholtz stereopsis is a 3D reconstruction technique, capturing surface depth independent of the reflection properties of the material by using Helmholtz reciprocity. In this paper we are interested in studying the applicability of Helmholtz stereopsis for surface and depth profiling of objects and general scenes in the context of perspective stereo imaging. Helmholtz stereopsis captures a pair of reciprocal images by exchanging the position of light source and camera. The resulting image pair relates the image intensities and scene depth profile by a partial differential equation. The solution of this differential equation depends on the boundary conditions provided by the scene. We propose to limit the illumination angle of the light source, such that only mutually visible parts are imaged, resulting in stable boundary conditions. By simulation and experiment we show that a unique depth profile can be recovered for a large class of scenes including multiple occluding objects.

Keywords: 3D reconstruction · Depth profiling · Stereopsis · Helmholtz reciprocity

1 Introduction

Three-dimensional optical depth profiling and sensing is one of the most actively pursued research areas in computational vision. Some of the key applications are industrial metrology, robotic vision, automotive driver assistance systems, 3D object scanning and media user interfaces. Optical depth sensing may be categorized in two groups, passive depth sensing such as stereo-vision or depth-from-defocus and active sensing technologies requiring an auxiliary light source, such as laser scanning or structured light.

Passive methods such as stereo-vision rely on the availability of scene features and require the measurement object to have diffuse reflection properties. Fringe projection methods [1] extend the applicability of stereoscopy to surfaces with homogeneous texture. The surface to be measured is illuminated by using a light projector to project a line or dot grid pattern. Since usually a repetitive pattern is used, one requirement of the fringe projection is the surface to be sufficiently smooth. For discontinuous surfaces with large height steps, phase ambiguities arise and a unique disparity assignment may not be possible. Structured light [2]

generalizes the concept of fringe projection and use time multiplexing to project an encoding pattern to uniquely identify surface patches.

In a recent series of publications, a novel approach to surface reconstruction was proposed by [3–5]. The idea is to use Helmholtz reciprocity to establish a relation between image intensity and the surface orientation of an observed object. Helmholtz reciprocity is a promising approach for 3D depth profiling, as it is largely independent of material reflectance and does not rely on surface texture. Our goal is to study the limitations of using Helmholtz reciprocity for general scene depth acquisition.

There have been a variety of proposals for solving the depth from Helmholtz reciprocal images. In the publication by [3,6], a multiview recovery approach is taken. It requires an initial depth estimate based on multiview-stereo correspondences and uses the surface normal estimate to refine the initial depth estimate. The approach in [7] uses a sensor fusion approach, refining the depth information obtained from a structured light scanner by utilizing the surface normal information obtained from Helmholtz reciprocity. In the publications of [4,5,8] the concept of Helmholtz stereopsis was introduced, it uses PDE integration methods to recover the depth profile along epipolar lines from a single pair of Helmholtz reciprocal images. Another interesting approach is proposed by [9], it uses variational techniques to iteratively refine an initial triangular mesh, Delaunoy showed that it is possible to recover the surface of complex scenes with multiple occluding objects.

In Helmholtz stereopsis, the scene is illuminated by a light source whereas a pair of reciprocal images is captured by exchanging the position of camera and light source. In its original conception, a point light source is used for object illumination. In our setup we use a LED pattern projector, which is commonly used for industrial laser scanning and structured light acquisition. The use of a projector for illumination allows limiting the angle of illumination. By matching the illumination aperture with the field aperture of the camera, only mutually visible parts of the scene are illuminated.

This report is organized as follows, first we will review the Helmholtz condition and formulate it in the context of perspective imaging. As we will see, this leads to an integral representation which allows solving for the surface profile in the case of multiple occluded objects. By using a light projector with a matched aperture we ensure stable boundary conditions for arbitrary scenes. For validation, we will provide simulation and experimental results of object and scene measurements.

2 Preliminaries

2.1 Perspective Surface Representation

Assuming a pinhole camera model, the image position (x, y) is related to the angular direction of an incident ray (u, v) by the focal length f . Given a discrete representation of the sensor image, the relation between the pixel index (i, j) and the angular image coordinates (u, v) is $u = \frac{\Delta x}{f} (i - i_0)$, $v = -\frac{\Delta y}{f} (j - j_0)$,

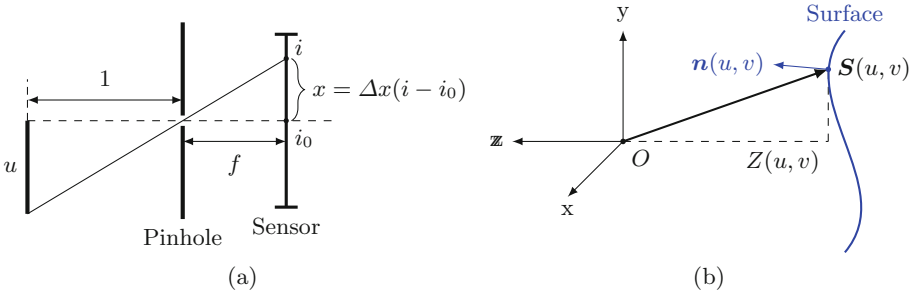


Fig. 1. (a) Illustration of the relation between the pixel index (i, j) and the angular image coordinates (u, v) . (b) Illustration of surface vector description by a given scalar depth map $Z(u, v)$ at a given angular image coordinate (u, v) .

whereas $\Delta x, \Delta y$ is the pixel resolution of the image sensor and (i_0, j_0) is the pixel position of the principle point (See Fig. 1a).

The surface height is described by a positive valued profile function $Z(u, v)$. As illustrated in Fig. 1b, we assume a fixed coordinate system, where the entrance pupil of the camera is located at the origin and the viewing direction of the camera is oriented along the negative z -axis. The spatial coordinates of the surface $S(u, v)$ are then described by the perspective projection relation,

$$S(u, v) = \begin{bmatrix} u \\ v \\ -1 \end{bmatrix} Z(u, v). \tag{1}$$

With the surface parameterized as a vector function of the image coordinates (u, v) , the normal field is obtained by the cross product of the partial derivatives of the surface,

$$N(u, v) = \frac{\partial S}{\partial u} \times \frac{\partial S}{\partial v} = \begin{vmatrix} i & j & k \\ Z + uZ_u & vZ_u & -Z_u \\ uZ_v & Z + vZ_v & -Z_v \end{vmatrix} = Z^2 \begin{bmatrix} Z_u/Z \\ Z_v/Z \\ 1 + uZ_u/Z + vZ_v/Z \end{bmatrix},$$

where Z_u, Z_v are partial derivatives of $Z(u, v)$ with respect to the variable u, v respectively.

2.2 Helmholtz Image Formation

For the measurement of a Helmholtz reciprocal image pair, the surface reflection is modeled by the radiance equations. Given a point on the surface of observation, the projected image coordinates are given by (u, v) and (u_2, v_2) for both views. The distance between the light source and the point on the surface is measured by the euclidean distances r_1, r_2 . The angle between the incident ray from the light source and the surface normal is given by θ_{in} as illustrated in Fig. 2a.

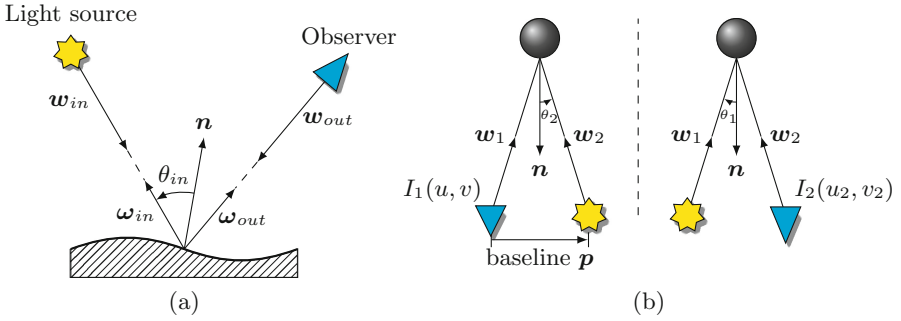


Fig. 2. (a) Reflection of light of an incident light ray at w_{in} to the reflected ray w_{out} by bidirectional reflection distribution function $f_r(w_{in}, w_{out})$. A pair of Helmholtz reciprocal images is taken, by exchanging the position of light source and observer. (b) Capturing of a pair of Helmholtz reciprocal images $I_1(u, v)$ and $I_2(u_2, v_2)$ by exchanging the position of light source and observer. The angular image coordinates of the left view are given by (u, v) and of the right view by (u_2, v_2) .

The reflection of an object point at a surface is given by the radiance,

$$I_1(u, v) = \alpha f_r(w_2, w_1) \cos\theta_2 / r_2^2, \tag{2}$$

by exchanging the position of camera and light source as shown in Fig. 2b, the radiance of the reciprocal image is,

$$I_2(u_2, v_2) = \alpha f_r(w_1, w_2) \cos\theta_1 / r_1^2, \tag{3}$$

where $f_r(w_i, w_o)$ is the bidirectional reflectance distribution at the observed surface point with w_i and w_o being the unit directions of the incident and reflected ray. The factor α is a constant scaling factor incorporating the illumination intensity and the camera gain.

For physical materials the BRDF satisfies Helmholtz reciprocity. Helmholtz reciprocity states that the reflection coefficient is the same if the path of light travel is reversed $f_r(w_1, w_2) = f_r(w_2, w_1)$ with $w_1 = -w_1/r_1$ and $w_2 = -w_2/r_2$ being the normalized directions of the observation and illumination vector. By measuring a pair of Helmholtz reciprocal images, the reflection function can be eliminated resulting in the Helmholtz reciprocity condition,

$$I_1(u, v) \cos\theta_1 / r_1^2 = I_2(u_2, v_2) \cos\theta_2 / r_2^2. \tag{4}$$

Using this notation, the Helmholtz reciprocity condition is given by the scalar product as stated in [3–6],

$$n^T \tau = 0 \tag{5}$$

whereas the vector τ is the tangent vector to the observed surface,

$$\tau = \frac{I_1(u, v)}{r_1^3} w_1 - \frac{I_2(u_2, v_2)}{r_2^3} w_2. \tag{6}$$

We may geometrically interpret the Helmholtz condition as a vector tangent field spanned over the space of possible depth values. If in this field, an initial starting depth is given, the surface depth can be reconstructed by path integration over the tangent field. In the next section, we restate the Helmholtz partial differential equation for the perspective imaging case.

2.3 Differential Helmholtz Condition

To formulate the Helmholtz condition for the perspective imaging case, consider a camera and light source being aligned in a common focal plane at $z=0$. The orientation of the camera is towards the negative z -axis. The camera and light source are located in a common focal plane and separated by a baseline vector $p = [p_x \ p_y \ 0]^T$. The object surface is parameterized by a positive depth profile $Z(u, v)$ in terms of the image coordinates (u, v) of the left view. The Helmholtz reciprocal image pair is given by the image functions $I_1(u, v)$ for the left and $I_2(u_2, v_2)$ for the reciprocal right view. The perspective projection of surface point $S(u, v)$ in both reciprocal views is given by the relation,

$$S(u, v) = \begin{bmatrix} u \\ v \\ -1 \end{bmatrix} Z(u, v) = \begin{bmatrix} p_x \\ p_y \\ 0 \end{bmatrix} + \begin{bmatrix} u_2 \\ v_2 \\ -1 \end{bmatrix} Z(u, v). \tag{7}$$

From above perspective projection, the angular disparity of a point between two views is given by,

$$\begin{bmatrix} u - u_2 \\ v - v_2 \end{bmatrix} = \begin{bmatrix} p_x \\ p_y \end{bmatrix} \frac{1}{Z(u, v)}. \tag{8}$$

The observation vector for both reciprocal views is,

$$w_1 = \begin{bmatrix} u \\ v \\ -1 \end{bmatrix} Z(u, v), \quad w_2 = \begin{bmatrix} u_2 \\ v_2 \\ -1 \end{bmatrix} Z(u, v). \tag{9}$$

When capturing the image of the left view $I_1(u, v)$, the vector w_1 is the observation vector starting from the focal position of the camera located at the left position pointing to a surface point and the vector w_2 is the illumination vector starting from the position of the light source and pointing to the observed surface point. When capturing the reciprocal image, the vector w_1 becomes the illumination vector and w_2 is the observation vector for the right image.

We may now simplify the Helmholtz reciprocal condition, by substituting the scalar product between the normal and observation vector. The scalar product of illumination vector and surface normal is,

$$|n| r_1 \cos\theta_1 = -n^T w_1 = Z^3, \tag{10}$$

$$|n| r_2 \cos\theta_2 = -n^T w_2 = Z^3(1 + (u - u_2)Z_u/Z + (v - v_2)Z_v/Z). \tag{11}$$

By using the disparity relation 8, the scalar product 11 simplifies to,

$$-n^T w_2 = Z^3(1 + p_x Z_u/Z^2 + p_y Z_v/Z^2). \tag{12}$$

The distance between focal points and object point is given by $r_1 = |w_1| = Z\sqrt{u^2 + v^2 + 1}$, $r_2 = |w_2| = Z\sqrt{u_2^2 + v_2^2 + 1}$. Inserting the scalar products 10, 11 for perspective projection into the Helmholtz condition 4, the perspective Helmholtz condition becomes,

$$I_1(u, v) \frac{Z^3}{r_1^3} = I_2(u_2, v_2) \frac{Z^3}{r_2^3} (1 + p_x Z_u / Z^2 + p_y Z_v / Z^2). \quad (13)$$

By introducing the following changes in notation, we eliminate the explicit dependency on the depth $Z(u, v)$ and above expression becomes more concise. For a Helmholtz reciprocal pair of images, the measured intensities are normalized by a radial weight function,

$$J_1(u, v) = \frac{I_1(u, v)}{\hat{r}_1^3}, \quad \hat{r}_1 = \sqrt{u^2 + v^2 + 1}, \quad (14)$$

$$J_2(u_2, v_2) = \frac{I_2(u_2, v_2)}{\hat{r}_2^3}, \quad \hat{r}_2 = \sqrt{u_2^2 + v_2^2 + 1}. \quad (15)$$

The explicit dependency on $Z(u, v)$ is eliminated by introducing the reciprocal depth $D(u, v) = 1/Z(u, v)$. With the partial gradients of $D_u(u, v)$, which are $D_u(u, v) = -Z_u/Z^2$ and $D_v(u, v) = -Z_v/Z^2$, the Helmholtz condition for rectified, perspective imaging becomes

$$J_1(u, v) = J_2(u_2, v_2) (1 - p_x D_u - p_y D_v). \quad (16)$$

For a point on the surface being mutually visible in both views, the cosine term $\cos\theta_2$ must be positive. For a continuous surface the weighting term in 16 satisfies the visibility constraint $(1 - p_x D_u - p_y D_v) > 0$.

2.4 Integral Helmholtz Condition

The Helmholtz condition for perspective imaging suggests the image pairs being related by a perspective image transform by a perspective coordinate transform. Suppose, the object under observation is an arbitrarily oriented plane placed in front of camera and light source. A rectangular element in the coordinate system (u, v) of the left view, is mapped into a parallelogram in the reciprocal image coordinate system (u_2, v_2) of the right view. The change of area due to a perspective view transformation is given by the Jacobian determinant,

$$\det \frac{\partial(u_2, v_2)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial u_2}{\partial u} & \frac{\partial u_2}{\partial v} \\ \frac{\partial v_2}{\partial u} & \frac{\partial v_2}{\partial v} \end{vmatrix} = 1 - p_x D_u - p_y D_v. \quad (17)$$

We may express the Helmholtz reciprocity condition using the more general notation,

$$J_1(u, v) \partial(u, v) = J_2(u_2, v_2) \partial(u_2, v_2). \quad (18)$$

The intensities, normalized according to Eq. 14, of a pair of reciprocal images are related by a perspective view transform. The differential terms $\partial(u, v)$ are the size of an area patch projected into the respective coordinate system.

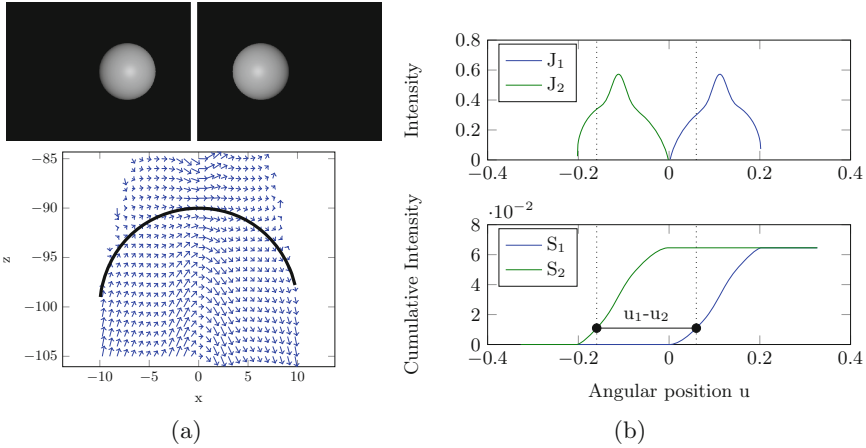


Fig. 3. (a) Raytracing of a Helmholtz image pair of a spherical object and geometric interpretation of the Helmholtz condition in terms of a surface tangent field from the figure above. (b) Intensity graph and cumulative intensity of the Helmholtz reciprocal images of the spherical object in Fig. 3a. The cumulative intensity forms a hysteresis curve. The disparity is determined by measuring the lag between both curves.

By integration, we obtain another interpretation of the Helmholtz condition. Consider an surface patch Ω mutually visible in both reciprocal images and let its projected areas be Ω_1 and Ω_2 , then the area integral of the normalized intensities is identical.

$$\iint_{\Omega_1} J_1(u, v) \, d\Omega_1 = \iint_{\Omega_2} J_2(u_2, v_2) \, d\Omega_2 \quad (19)$$

If we further assume a rectified stereo configuration with a baseline vector of $p_x > 0$ and $p_y = 0$, the epipolar lines of the image pairs are parallel to the x-axis. If the object is of finite size and the background is black or sufficiently far away such that the object is mutually visible in both images, then the intensity sum along each epipolar line is expected to be equal.

In the case of rectified stereo, the identity in Eq. 19 suggests a reconstruction rule which measures the disparity from the sum of intensities along epipolar lines. For the simulated example of the spherical object in Fig. 3a, the measurement of the disparity is illustrated in Fig. 3b. At a given intensity sum level, the measured disparity is the difference of the abscissa at which the intensity sum is equal for both image pairs. By measuring the disparity for each level of the intensity sum, we may recover a disparity map, which is a valid but not necessarily unique solution to the differential equation formulated by the Helmholtz condition.

2.5 Simulation Examples

For the existence of a solution, the object to be measured is required to have sufficient reflectance and to be close to the light source such that the image signal has a sufficiently high signal noise ratio. Occluded and non-visible surface areas in one image will be located in the shadow regions of the corresponding reciprocal image. For complex geometrical scenes, the software Blender in combination with LuxRender is a suitable modeling and rendering environment. The LuxRender implementation is based on the physically based rendering engine by Pharr and Humphreys [10], it correctly models point illumination and provides a physically correct specular model (Cook-Torrance model [11]).

Boundary Conditions: For uniqueness, we need suitable boundary conditions for the Helmholtz condition. In case the object is of finite size and has a continuous surface, we may recover a unique disparity map by capturing the object such that it is mutually visible in both images. If the object is extended beyond the viewing cone of the camera, the intensity sum will have an offset compared to the intensity sum of the reciprocal image. In this case, it will be necessary to use other boundary constraints, e.g. by using stereo correspondences to determine this offset.

In case of general scenes, the object may not be bounded and we may have multiple occluding objects. One approach to relief the uniqueness problem is to introduce an aperture to the light source. By limiting the angular extend of the light source, the illumination aperture will cast a shadow into the scene. If the aperture of the camera of the reciprocal image is matched to size and shape of the light source, only those parts of the scene are illuminated which are mutually visible in the reciprocal pair of images. Figure 4a shows a Helmholtz reciprocal image pair of a sphere placed in front of an extended, untextured plane. In the figure, the illumination cone and the camera viewing cone are bounded by a circular aperture. Figure 4b compares the resulting disparity map with the ground truth. Except of the tips of the spherical object, the disparity is accurately recovered including the occluding sphere.

Occlusion and Visibility Order: The recovery artifacts at the top of the sphere are caused by a change of the order of visibility for occluding objects which have a small width compared to the distance to the background object. In this case, the shadow cast on the background object becomes detached from the occluding front object. Figure 5 illustrates the occlusion artifacts, suppose a front object of width W is placed in front of a background plane at a distance H_1 and the distance between the common focal plane of the Helmholtz setup with baseline P and the background object is H_0 . If the ratio W/H_1 is less than P/H_0 , the shadow S cast by the frontal object is separated from its projected image F by a visible patch of background object B . By computing the intensity sum starting at the left image boundary, the sequence of visible object patches differs for the left and right view. Measuring disparity from the intensity sum results in a feature correspondence mismatch.

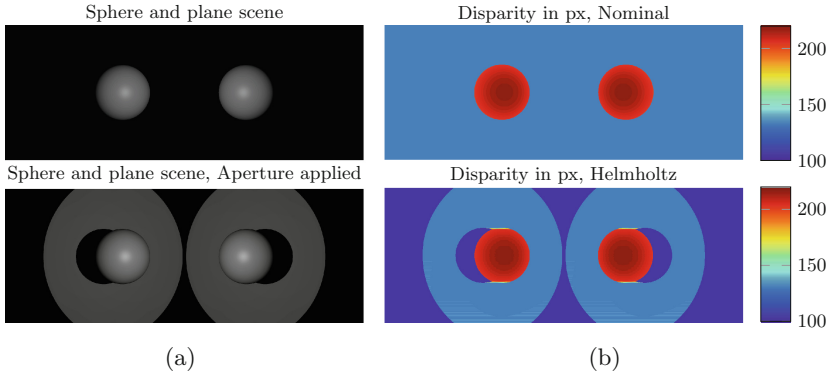


Fig. 4. (a) Top: Spherical object in front of an extended plane. Bottom: The scene is rendered with a circular illumination and camera aperture. By matching the aperture of camera and light source, boundary conditions are well defined for reconstruction. (b) Comparison of nominal versus estimated disparity.

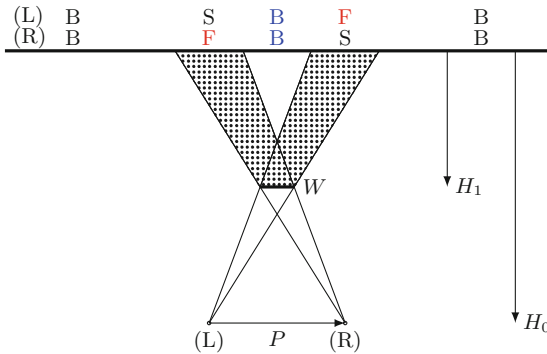


Fig. 5. Illustration of visibility order for a thin planar object in front of a background plane. The order of visibility of background and foreground object segments is reversed between left (L) and right view (R).

Noise Propagation: Since Helmholtz stereopsis estimates the disparity by integrating the intensity, an error in the input signal will not only affect the disparity estimation at the current pixel position, but also propagate to the neighboring pixel in the direction of integration. For Gaussian image noise, the error propagation of the disparity will result in a random walk. Close to the starting boundary condition, the disparity error will be small and grow continuously with the distance to the starting position. When repeating the procedure for all epipolar rows in a rectified image, the disparity will diverge randomly for each row. To reduce error propagation due to image noise, one option is to match the cumulative intensity at the left and right boundary. This is best illustrated with a planar rectangular patch, as we have in most cases two boundaries.

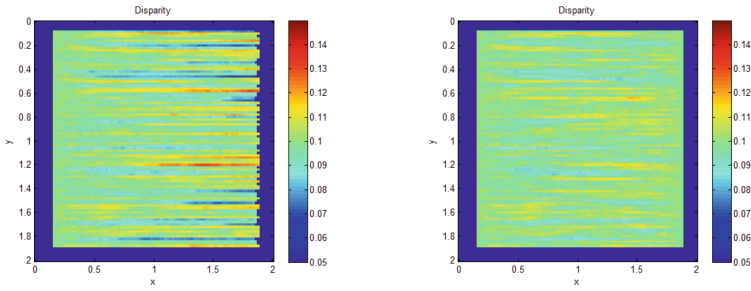


Fig. 6. In both test scenes, the baseline of the camera positions is 20 cm and the approximate working distance 1m. In the MATLAB test scene, a sphere of 10 cm radius is placed at the origin and viewed from a camera at a distance of 1m. A background plane is located 50 cm behind the sphere. Left: Disparity map initialized at the left boundary. Right: Disparity map of matched boundaries.

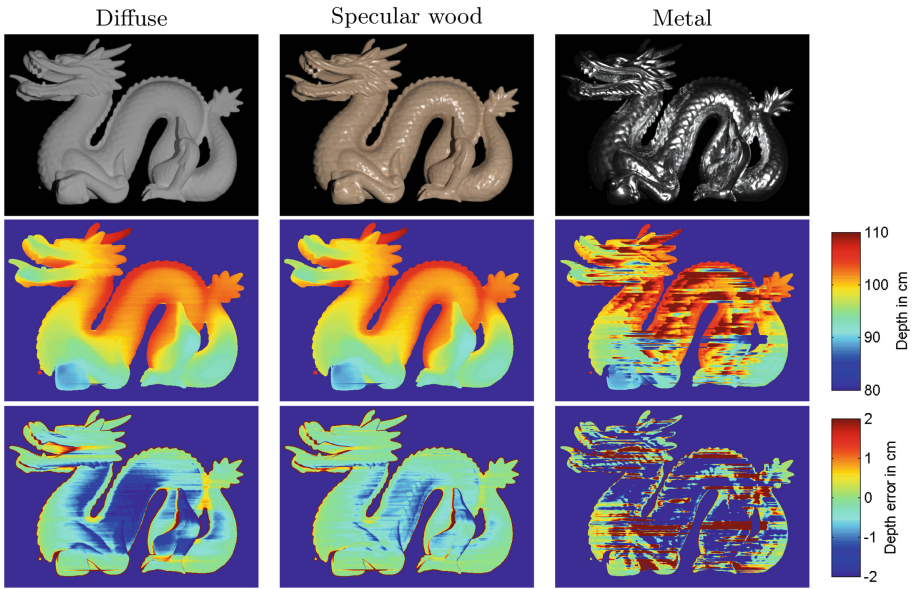


Fig. 7. Depth estimation for material simulations - (1) Diffuse, (2) Specular wood, and (3) Metal. The distance to the simulated object is approx. 100 cm at a baseline of 20 cm, a sensor width = 32 mm and $f = 35$ mm.

If we have multiple boundaries e.g. at the left/right aperture, we may scale the intensities in order to ensure the propagated depth matches at each boundary. The impact of error propagation in images is illustrated in Fig. 6. Since each row of the image has uncorrelated noise, each row of the disparity results in a different random walk. In the left image, the disparity is measured starting from the left of the scanline. Not only the error of the disparity will increase with the

distance to the initial boundary, but also the object contour will be scraggy at the unmatched boundary. In the right image, the boundaries are matched by rescaling the intensities. By matching the boundaries on both sides, the object contour will remain smooth and the error propagation is reduced.

Specular Reflections: The following simulation example in Fig. 7 demonstrates the result of Helmholtz stereopsis to a more complex shape and with different material reflections. The test scene is the Stanford Dragon model, available from the Stanford 3D Scanning Repository. The distance between the dragon and the camera is approximately 1m. The result of depth estimation and the difference from ground truth are shown in the middle row and in the bottom row, respectively. For diffuse and specular materials the estimated depth agree very well with the ground truth. The metallic surface has too strong reflections resulting in image saturation.

3 Evaluation

3.1 Experimental Setup

For the measurements we used an industrial USB camera (IDS UI-1480-C-SE) with 5 MP image resolution and 8 bit color depth. To reduce noise and increase dynamic range, an exposure series of 5 images are captured and an HDR image is reconstructed. Ambient light subtraction is done by capturing a pair of HDR images, and subtracting the images with the respective light source turned on and off. The light source used is an LED projector (LTPR, Opto Engineering). The objective lens used for both camera and projector has a focal length $f = 12$ mm and aperture $F/2.8$. The camera and two projectors are mounted with parallel orientation on a sliding stage whereas the camera is placed in the center between the two projectors at a baseline of approx. 20 cm. By sliding the stage, the position of light source and camera is exchanged.

3.2 Calibration

For geometric calibration we used standard photogrammetric techniques using a planar rectangular grid for the estimation of the camera model. The estimated camera model is used for stereo image rectification.

In a physical setup the pinhole/spotlight assumptions used in deriving Eqs. 14 and 15 are not valid, the apodization function depends on the vignetting of the devices. For symmetry, we use the same lens type for camera and projector and make the assumption that the apodization function is radially symmetric. The intensity apodization is calibrated using a target board, instead of measuring the vignetting of the individual optical components. A more detailed discussion about Helmholtz calibration is given in [12].

For radiometric calibration, we used a target with an inverted dot pattern which white dots locate on black background. Image normalization is done by

assuming radial symmetry and using a parabolic approximation. With the rectified image radiance pair given by $\hat{I}_{1,2}(u, v)$, the normalized Helmholtz image pair is given by $\hat{J}_{1,2}(u, v) = \mu(u, v)\hat{I}_{1,2}(u, v)$ with the apodization function $\mu(u, v) = 1 - \frac{1}{2}\alpha r^2$, $r^2 = u^2 + v^2$ and α being a radiometric correction parameter. By matching the cumulative intensity sum for each dot pattern of the target pattern, we may adjust the radiometric calibration parameter α .

3.3 Measurement Results

Black Background Scene: As a first measurement consider the rectified and normalized Helmholtz image of a Pythagoras statue, made of white plaster, in Fig. 8a. The statue is placed in front of a black background to have defined boundary conditions. The points which are mutually visible are identified by applying a hard threshold. The first valid point along an epipolar line can be starting point and the last valid point can be an end point. The mismatch between the intensity sum between both images is about 4%. Figure 8a shows the estimated angular disparity map and Fig. 8b the mesh reconstruction in the coordinate space of the left camera.

Combination with Stereo Matching: If the object has sufficient texture or a corner, a starting condition can be found by using conventional stereo matching. In that case, it can be used for the references to define corresponding initial points between left and right views. Once a reliable disparity is found for one view, the corresponding point in the other view can be easily derived from the disparity. Since reliable disparities tend to be on edges or textured areas, edge detection results are predisposed to be used as a reliability map. Regarding left view, scanning disparity and reliability map starting from the left boundary to up to the right boundary, the first pixel, where the pixel which fulfills a valid disparity and a certain texture level is located, is defined as a starting point. The last point of the same condition is defined as an end point (shown in left hand images of Fig. 8c). For right view, the starting-/end points are both derived by the disparity (shown in the higher right image of Fig. 8c). The estimated disparity map using this method is shown in the lower right of Fig. 8c.

Illumination Reticles: To support arbitrary scenes, one idea is to modify the Helmholtz stereoscopic setup and introduce an aperture to the light source. By matching the size and shape of sensor and illumination aperture, both camera and illumination have same field of view which can be a definition of boundary for generic scenes. Figure 8d shows the rectangular reticle pattern installed in between an object lens, and images with projected vertical lines. Subsidiary, a pseudo camera aperture pattern is overlying by a dotted line. Illumination reticle and camera aperture clarify the boundary condition for integration in each line. A projected light for left view defines a boundary condition on the left-hand, and corresponding boundary on the right-hand is defined by pseudo camera aperture pattern. For the right view, it is the other way around. The disparity of

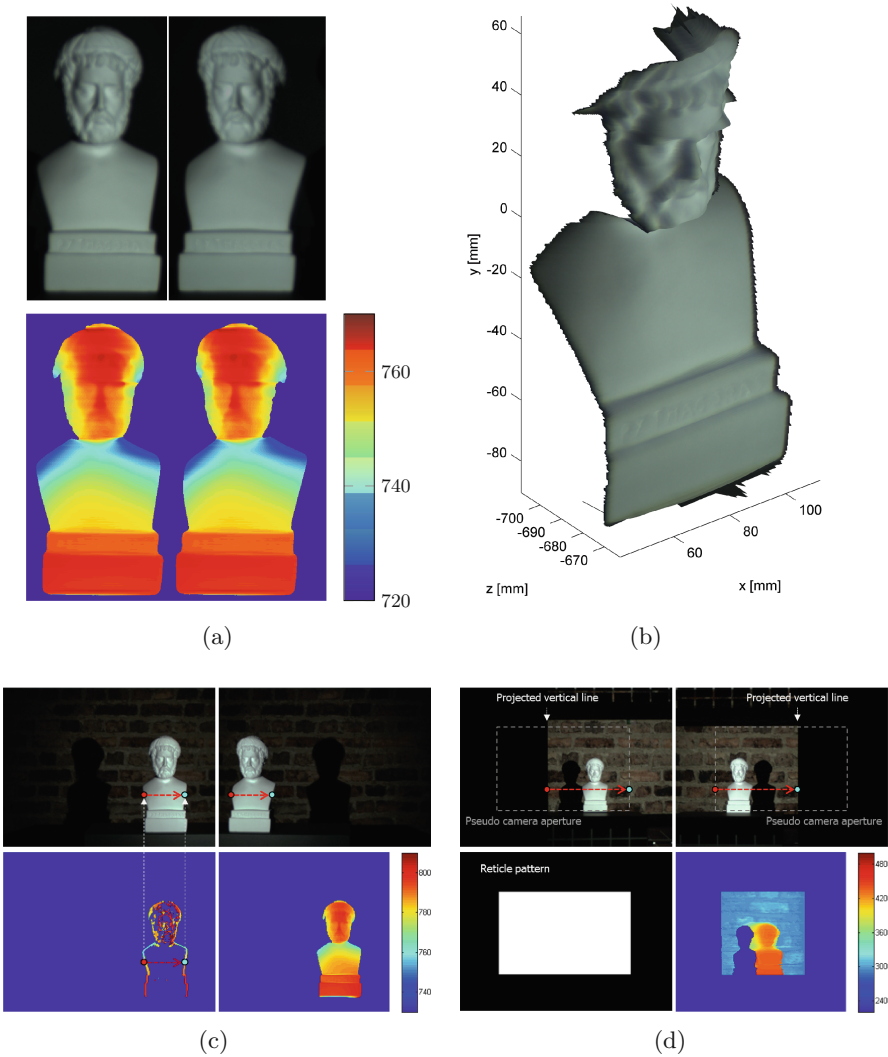


Fig. 8. (a) Normalized reciprocal image pair of Pythagoras plaster statue (above) and estimated pixel disparities of Pythagoras statue (below). The distance to the object is approx. 70 cm at a baseline of 19 cm, a pixel size of 4.4 μm and $f = 12$ mm. (b) Surface mesh reconstruction of Pythagoras statue. (c) Combination with stereo matching, initial condition is determined by disparity map of left view (lower left). Estimated disparity map is on the below right. The distance to the object is approx. 70 cm at a baseline of 21 cm. (d) Rectangular pattern projected by a reticle on light source. Initial condition is defined by pseudo camera aperture and projected reticle patterns. Estimated disparity map is on the below right. The distance to the object and that to the background at a baseline of 21 cm are approx. 130 cm and 190 cm respectively.

background and foreground can be estimated together, however, some artifacts are visible in the disparity map (shown in the lower right of Fig. 8d.). The reason for the occlusion artifacts is explained in Sect. 2.5.

4 Conclusion

In this paper, our interest is to discuss the limits and principal problems of perspective Helmholtz stereopsis, in particular for challenging border cases such as no-texture, occlusion and stable boundary conditions for general scenes, e.g. a plain white unbounded wall or surface with specular reflections. As Helmholtz stereopsis is based on a PDE, boundary condition and noise propagation are fundamental key issues. We have extended Helmholtz stereopsis to the perspective imaging case and proposed a solution approach of the resulting differential equation based on energy conservation. By introducing an illumination aperture we are able to provide stable boundary conditions for unbounded surfaces to capture the depth profile of the scene.

The disparity estimation is robust with respect to depth discontinuities and object occlusion. However, an open problem is the visibility ordering problem in case of occlusion. The shadow cast by a occluding object may cause a change of object order between reciprocal image pairs. Without additional assumptions to match corresponding object segments, this leads to disparity mismatches. The fusion with stereo matching is one possible approach to obtain stable boundary conditions.

In the shown examples we only utilized the Helmholtz conditions and did not impose any regularizing constraints on the surface. When texture is present and smoothness assumptions can be imposed on the surface, fusion approaches such as proposed by Zickler [4] are preferable. Regularization may also help to solve issues with respect to occlusion order.

Helmholtz stereopsis is well suited for combination with structured light or fringe projection methods. Error propagation due to imaging noise and calibration errors may be reduced by using an illumination grating for scene segmentation and using partial integration for depth recovery. While structured light provides a high level of depth accuracy, its depth resolution is limited by the resolution capabilities of the pattern projector. Combining Helmholtz stereopsis with structured light allows to maintain a defined depth accuracy at a depth resolution close to the resolution limit of the imaging device.

References

1. Gorthi, S.S., Rastogi, P.: Fringe projection techniques: whither we are? *Optics Lasers Eng.* **48**(IMAC-REVIEW-2009-001), pp. 133–140 (2010)
2. Geng, J.: Structured-light 3D surface imaging: a tutorial. *Adv. Optics Photonics* **3**(2), 128–160 (2011)

3. Magda, S., Kriegman, D., Zickler, T., Belhumeur, P.N.: Beyond lambert: reconstructing surfaces with arbitrary BRDFS. In: Proceedings Eighth IEEE International Conference on Computer Vision, ICCV 2001, vol. 2, pp. 391–398. IEEE (2001)
4. Zickler, T.E., Ho, J., Kriegman, D., Ponce, J., Belhumeur, P.N.: Binocular Helmholtz stereopsis. In: Proceedings Ninth IEEE International Conference on Computer Vision, pp. 1411–1417. IEEE (2003)
5. Tu, P., Mendonça, P.R.: Surface reconstruction via Helmholtz reciprocity with a single image pair. In: Proceedings 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. I-541. IEEE (2003)
6. Guillemaut, J.Y., Drbohlav, O., Sara, R., Illingworth, J.: Helmholtz stereopsis on rough and strongly textured surfaces. In: Proceedings 2nd International Symposium on 3D Data Processing, Visualization and Transmission, 3DPVT 2004, pp. 10–17. IEEE (2004)
7. Weinmann, M., Ruiters, R., Osep, A., Schwartz, C., Klein, R.: Fusing structured light consistency and Helmholtz normals for 3D reconstruction. *ACM Trans. Graph.* **24**, 536–543 (2005)
8. Tu, P., Mendonça, P.R., Ross, J., Miller, J.: Surface registration with a Helmholtz reciprocity image pair. In: IEEE Workshop on Color and Photometric Methods in Computer Vision, vol. 7. Citeseer (2003)
9. Delaunoy, A., Prados, E., Belhumeur, P.N.: Towards full 3D Helmholtz stereovision algorithms. In: Kimmel, R., Klette, R., Sugimoto, A. (eds.) ACCV 2010. LNCS, vol. 6492, pp. 39–52. Springer, Heidelberg (2011). doi:[10.1007/978-3-642-19315-6_4](https://doi.org/10.1007/978-3-642-19315-6_4)
10. Pharr, M., Humphreys, G.: Physically Based Rendering: From Theory to Implementation. Morgan Kaufmann, San Francisco (2010)
11. Cook, R.L., Torrance, K.E.: A reflectance model for computer graphics. *ACM Trans. Graph.* **1**(1), 7–24 (1982)
12. Jankó, Z., Drbohlav, O., Sara, R.: Radiometric calibration of a Helmholtz stereo rig. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2004, vol. 1, pp. 166–171. IEEE (2004)