

Specular Surface Recovery from Reflections of a Planar Pattern Undergoing an Unknown Pure Translation

Miaomiao Liu¹, Kwan-Yee K. Wong¹, Zhenwen Dai², and Zhihu Chen¹

¹ Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong

² Frankfurt Institute for Advanced Studies, Goethe-University Frankfurt

Abstract. This paper addresses the problem of specular surface recovery, and proposes a novel solution based on observing the reflections of a translating planar pattern. Previous works have demonstrated that a specular surface can be recovered from the reflections of two calibrated planar patterns. In this paper, however, only one reference planar pattern is assumed to have been calibrated against a fixed camera observing the specular surface. Instead of introducing and calibrating a second pattern, the reference pattern is allowed to undergo an unknown pure translation, and a closed form solution is derived for recovering such a motion. Unlike previous methods which estimate the shape by directly triangulating the visual rays and reflection rays, a novel method based on computing the projections of the visual rays on the translating pattern is introduced. This produces a depth range for each pixel which also provides a measure of the accuracy of the estimation. The proposed approach enables a simple auto-calibration of the translating pattern, and data redundancy resulting from the translating pattern can improve both the robustness and accuracy of the shape estimation. Experimental results on both synthetic and real data are presented to demonstrate the effectiveness of the proposed approach.

1 Introduction

Shape recovery has always been a hot topic in computer vision over the past few decades. Traditional shape recovery methods can be classified into different categories according to the information they employed, such as structure from motion [1], shape from silhouettes [2], and shape from shadings [3]. Most of the existing methods, however, can only handle diffuse objects due to the fact that the information they exploited is derived under the assumption of a diffuse surface. For instance, structure-from-motion methods cannot be applied to a specular (mirror) object as the features on the object surface are the reflections of its surrounding environment and are therefore not viewpoint independent. Similarly, shape-from-silhouettes methods have difficulties in handling specular objects as it is not a trivial task to extract the silhouette of a mirror object. The underlying principle of shape-from-shadings methods states that the shadings

of the object follow the Lambert’s cosine law, and therefore it is obvious that specular surfaces cannot be recovered using shape-from-shading methods. In addition, the shape of a specular object also cannot be obtained using a laser scanner since the surface may often reflect the laser ray away from the sensor of the scanner.

In order to allow shape recovery of specular objects, special algorithms have to be developed under the assumption of a specular surface whose appearance follows the law of reflection. In [4], Zisserman et al. pointed out that a moving observer can tell the physical properties, such as convexity or concavity, of a specular surface. Fleming et al. [5] also showed that human beings have the ability to distinguish the shape of a specular surface even when no knowledge about the environment is available. In [6], Oren and Nayar explored the relation between the image trajectories of the virtue features induced by specular reflection and the 3D geometry of the specular object they travelled on. Roth and Black [7] proposed to use both optical flow and specular flow to recover an object which is a mixture of specular and diffuse materials. In [8], Adato et al. analyzed the valued information of the specular flow induced by the movement of a camera, and derived a non-linear partial differential equation (PDE) to describe the relation between the 3D shape and motion of the specular object. In [9], Canas et al. linearized this non-linear PDE in order to make it easier for practical use. The aforementioned works are regarded as shape from specular flow (SFSF) methods. SFSF methods generally require a dense observation of specular flow, but dense specular flows often merge together in an undetermined way. Besides, SFSF methods recover the specular surface by solving the PDE, and the solution is greatly affected by the initial boundary condition. Furthermore, the motion of the specular object must be small and continuous in order to derive the PDE. All these limitations make SFSF methods not very useful in practice.

Great efforts have also been devoted to specular shape recovery techniques under known surroundings. In [10, 11], Savarese and Perona demonstrated that local surface geometry properties of a mirror object can be determined by analyzing the local differential properties of the reflection of two calibrated lines. In [12], Bonfort and Sturm introduced a voxel carving method to recover a specular surface using a normal consistency criterion. Seitz et al. [13] reduced the specular shape recovery problem to a light path reconstruction problem. They showed that surface points can be obtained if the positions of two reference points are known in space for one single view, or the position of one reference point is known and is visible in two different views. In [14], a phase shift method was used to determine the correspondence between a reference point in space and its reflection on the image plane. Nehab et al. [15] reduced the shape recovery problem to an image matching problem by minimizing a cost function based on normal consistency. Note that calibration plays an important role in specular shape recovery methods under known surroundings. Bonfort et al. [16] used a mirror to calibrate a reference plane after changing its position every time. In [17], Rozenfeld et al. suggested an approach to recover the surface by using

an empirically calculated 1D homography. Their proposed method only needs sparse plane-image correspondences, but depends on optimization greatly.

In this paper, we adopt the approach of specular shape recovery under known surroundings. Unlike previous methods which assume a fully calibrated environment, we propose a novel solution based on observing the reflections of a translating planar pattern. We assume a reference planar pattern is initially calibrated against a fixed camera observing a specular surface. This pattern is then allowed to undergo an *unknown pure translation*. The motion of the pattern as well as the shape of the specular object can then be estimated from the reflections of the translating pattern. The main contributions of this paper are 1) a closed form solution for recovering the motion of the translating pattern, and 2) a novel method for shape recovery based on computing the projections of the visual rays on the translating pattern. The proposed auto-calibration method allows an easy calibration of the translating pattern, and data redundancy resulting from the translating pattern can improve both the robustness and accuracy of the shape estimation. The proposed shape recovery method produces a depth range for each pixel in which the true surface point exists. This in turn provides a measure of the accuracy and confidence of the estimation.

The rest of the paper is organized as follows. Section 2 describes the physical configuration of our proposed translating pattern approach. Based on geometric constraints, a closed form solution is derived for recovering the unknown motion of the translating pattern in Section 3. A degenerate case is also studied and analyzed. Section 4 introduces a projection based method for specular shape recovery using the estimated motion. Experimental results on both synthetic and real data are presented in Section 5, followed by conclusions in Section 6.

2 Translating Planar Pattern

Consider a pinhole camera, with camera center O , looking at a specular object S which reflects a nearby planar pattern Π to the image I (see Fig.1). Consider now a point P on S . Let A be a point on Π which is reflected by S at P to a point x on I . Suppose now the planar pattern undergoes a pure translation given by an *unknown* translation vector \mathbf{T} , and let Ω denotes the planar pattern at the new location after translation. Let B be a point on Ω such that it is collinear with A and P . According to the law of reflection, B will also be reflected by S at P to the same point x on I .

As shown in [13, 16], the 3D position of P can be obtained by intersecting the visual ray (defined by O and x) with the reflection ray (defined by A and B). In order to construct the reflection ray, both the 3D positions of A and B are needed, and their accuracies would directly affect the accuracy of the estimated 3D position of P . The relative positions of A and B on the planar pattern can be resolved by encoding the pattern using gray code and observing the corresponding intensities at x . Besides, in order to obtain the 3D positions of A and B , the relative positions and orientations of Π and Ω with respect to O must also be known. In this paper, Π is assumed to have been calibrated

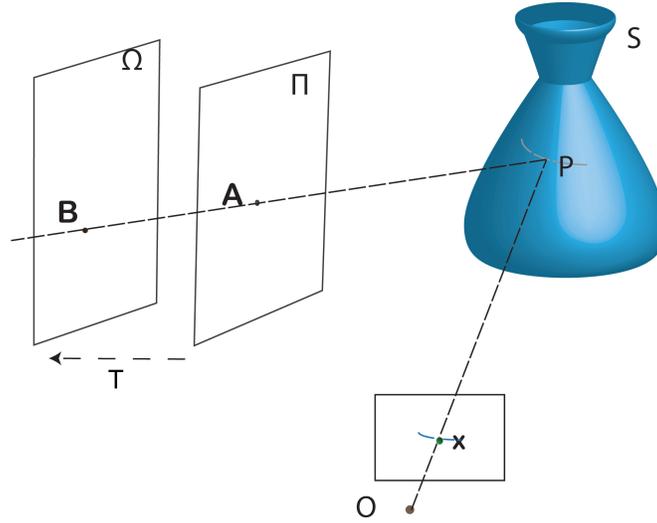


Fig. 1. Configuration of a translating planar pattern. A pinhole camera with camera center O is viewing a specular object S which reflects a nearby planar pattern Π to the image I . Ω is a planar pattern obtained by translating Π by an unknown translation vector \mathbf{T} . A and B are points on Π and Ω respectively which are reflected by S at P to the same image point x on I .

against O initially. The problem of calibrating Ω is thus reduced to the problem of estimating the unknown translation vector \mathbf{T} . In the next section, a closed form solution for \mathbf{T} will be derived based on some geometric constraints.

3 Recovering the Unknown Translation

Referring to the configuration described in the previous section, Ω and Π are related by a pure translation given by the translation vector \mathbf{T} . Consider a point X on Π . Let X' be a point on Ω such that it has the same relative position on the planar pattern as X (i.e., both X and X' are encoded by the same gray code). X and X' are referred to as a pair of corresponding points, and are related by $\mathbf{X}' = \mathbf{X} + \mathbf{T}$, where \mathbf{X} and \mathbf{X}' denote the position vectors for X and X' respectively. Now referring back to Fig. 1, and recalling that A and B are points on Π and Ω , respectively, which are reflected by S at P to the same image point x on I . Let Q be a point on Π which is encoded by the same gray code as B (i.e., B and Q form a pair of corresponding points). Hence, we have

$$\mathbf{B} = \mathbf{Q} + \mathbf{T}, \quad (1)$$

and the vector \mathbf{AB} in the direction of the reflection ray can be written as

$$\mathbf{AB} = \mathbf{AQ} + \mathbf{T}. \quad (2)$$

Consider the plane Γ formed by the visual ray and the reflection ray at P , and let \mathbf{N} be its unit normal vector. It is obvious that \mathbf{AB} and \mathbf{N} are mutually orthogonal and therefore

$$\mathbf{AB} \cdot \mathbf{N} = 0. \quad (3)$$

Substituting (2) into (3) gives

$$(\mathbf{AQ} + \mathbf{T}) \cdot \mathbf{N} = 0. \quad (4)$$

Note that Q can be inferred from the gray code of B , and the vector \mathbf{AQ} can be constructed from the 3D positions of A and Q on the calibrated planar pattern Π . The unit normal vector \mathbf{N} can be computed as the vector cross product of \mathbf{Ox} and \mathbf{OA} . Hence, the only unknown in (4) is the translation vector \mathbf{T} , which has three degrees of freedom. Since a pair of reflections observed at one pixel provides one linear constraint on \mathbf{T} , \mathbf{T} can therefore be estimated using a linear least squares method from the reflections observed at three or more pixels:

$$\mathbf{T} = - \begin{pmatrix} \mathbf{N}_1^T \\ \mathbf{N}_2^T \\ \vdots \\ \mathbf{N}_n^T \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{A}_1 \mathbf{Q}_1 \cdot \mathbf{N}_1 \\ \mathbf{A}_2 \mathbf{Q}_2 \cdot \mathbf{N}_2 \\ \vdots \\ \mathbf{A}_n \mathbf{Q}_n \cdot \mathbf{N}_n \end{pmatrix} \quad (5)$$

If $n > 3$, matrix composed of normal vectors will be a rectangle matrix. The symbol of inverse used in the (5) will be denoted as Moore-Penrose pseudo-inverse. The pseudo-inverse gives the solution in a least-squared sense.

However, when the specular object is a plane or a surface of revolution, there exist some degenerate cases which are described below in Proposition 1 and Proposition 2.

Proposition 1. *If the specular object is a plane, the translation vector \mathbf{T} cannot be determined by using the proposed theory.*

Proof. Let S be a specular plane. Without loss of generality, let P be an arbitrary point on S and N_P be the surface normal at P . Consider now the visual ray OP . By the law of reflection, the visual ray OP , the surface normal N_P and the reflection ray must be all lie on the same plane. Let Γ_P denotes such a reflection plane. Note all points on S have the same surface normal N_P . It follows that all the reflection planes will intersect along a line passing through the camera center O and with an orientation same as N_P , and hence their normal vectors are all orthogonal to N_P . This makes the matrix composed of the normal vectors of the reflection planes in (5) to have a rank of at most 2, and therefore \mathbf{T} cannot be recovered using (5).

Proposition 2. *If the specular object is a surface of revolution and its revolution axis passes through the camera center, the translation vector \mathbf{T} cannot be determined by using the proposed theory.*

Proof. Let S be a specular surface of revolution with its revolution axis l passing through the camera center O . Without loss of generality, let P be an arbitrary point on S and N_P be the surface normal at P . Consider now the visual ray OP . By the law of reflection, the visual ray OP , the surface normal N_P and the reflection ray must be all lie on the same plane. Let Γ_P denotes such a reflection plane. By the properties of a surface of revolution, N_P lies on the plane defined by P and l . Since l passes through O , it is easy to see that l also lies on Γ_P . It follows that all the reflection planes will intersect at l and hence their normal vectors are all orthogonal to l . This makes the matrix composed of the normal vectors of the reflection planes in (5) to have a rank of at most 2, and therefore \mathbf{T} cannot be recovered using (5).

Since a sphere is a special kind of surface of revolution with any line passing through its center being a revolution axis, Corollary 1 follows immediately.

Corollary 1. *A single sphere cannot be recovered by using the proposed theory.*

4 Specular Surface Recovery

With the estimated translation vector \mathbf{T} , the position of the reference planar pattern after translation can be determined. If accurate correspondences between points on the reference planar pattern and pixels on the image can be obtained, the surface of the specular object can be recovered by triangulating the corresponding visual ray and reflection ray at each pixel (see Section 2). However, most of the encoding strategies cannot achieve a real point-to-point correspondences. Usually, one pixel corresponds to one encoded area (e.g., a square on the reference planar pattern). In [16], a dense matching of the pixels on the image and positions on the reference plane was implemented by minimizing an energy function which penalized the mis-encoding problem, and bilinear interpolation was used to achieve a sub-pixel accuracy in the matching. Bilinear interpolation, however, is indeed not a good approximation as it is well-known that ratio of lengths is not preserved under perspective projection. Besides, the reflection of the pattern by a non-planar specular surface would also introduce distortion in the image of the pattern, and such a distortion depends on both the shape and distances of the object relative to the camera and the planar pattern. In order to obtain good shape recovery results, the resolution of the encoding pattern and the positions of the plane need to be chosen carefully.

In this paper, the specular surface is assumed to be smooth and locally planar. After associating a gray code (and hence an encoded area on the reference plane) to each pixel, the algebraic centers c_i for pixels with the same gray code is computed and associated with the center C_i of the corresponding encoded area on the reference plane. This is the approximation to the point-to-point correspondences in the underlying encoding strategy. This approximation has been verified with experiments on synthetic data, and experimental results show that the error distances introduced by approximating the ground truth reflection

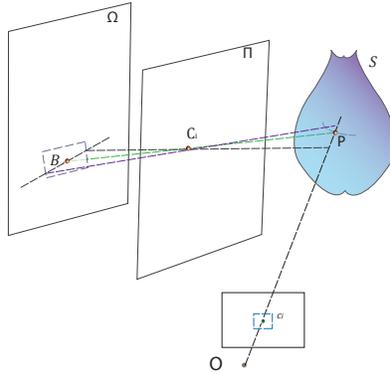


Fig. 2. Shape recovery of specular object. The visual ray at c_i is projected onto the translated reference plane Ω using its corresponding point C_i on the initial reference plane Π as the projection center. The projection of the visual ray on Ω will result in a line which will intersect the corresponding encoded area as observed at c_i . This intersection line segment can be back projected onto the visual ray and this gives a depth range in which the surface point should exist.

points on the plane with the centers of the encoded areas are negligible. Nevertheless, this approximation can only be applied to obtain a matching between the image pixels and points on one reference plane, as the algebraic centers c_i computed for one reference plane position will not in general coincide with those computed for a different reference plane position. Instead of trying to locate a point match on the translated reference plane Ω for each c_i , the visual ray for c_i is projected onto the translated reference plane Ω using its corresponding point C_i on the initial reference plane Π as the projection center (see Fig. 2). The projection of the visual ray on Ω will result in a line which will intersect the corresponding encoded area (which is a square) as observed at c_i . This intersection line segment can be back projected onto the visual ray and this gives a depth range in which the surface point should exist. If Ω is far from Π , the depth range will be small and the midpoint can be taken as the estimated surface depth. Alternatively, the depth range can be made smaller by increasing the resolution of the encoding pattern (i.e., decreasing the size of the encoded area). This, however, is not a practical strategy since the encoding pattern will be blurred in the imaging process and may result in an aliasing effect.

5 Experimental Analysis

The proposed approach was tested using both synthetic and real image data. The synthetic images were generated using the *pov-ray* software. For the real experiments, images were taken by using a simple setting composed of a monitor, a camera and a drawer. The monitor was used for displaying gray patterns. The

camera used was a Cannon 450D with a focal length of 50mm. The drawer was used for carrying the monitor and making it move in pure translation motions.

5.1 Synthetic Experiments

In the synthetic experiments, the scene was composed of two spheres or two planes with different positions and orientations. Note that the cases with multiple spheres or multiple planes are not degenerate, and the shape of the objects can be recovered using the proposed method. In order to analyze the effect of the translation scale on the result, the reference plane was translated with multiple scales.

The reconstruction for a part of a sphere is shown in Fig. 3, which also gives a comparison with the ground truth data. Fig. 4 shows the relation between the translation scale of the reference plane and the shape recovery error. It can be seen that, within the visible reflection range for the specular object, larger translation scale would result in smaller reconstruction error. After experimenting with a range of translation scales, 80 is chosen as the translation scale for the second position of the reference plane. The reconstruction error is smaller than 0.5%. The reconstruction for the plane is shown in Fig.5, which again also gives a comparison with the ground truth data. 70 is chosen as the translation scale and the average reconstruction error is less than 0.45%.

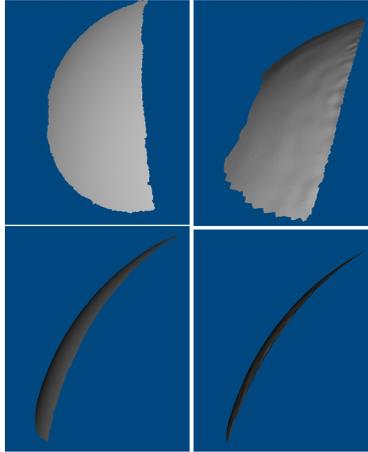


Fig. 3. Shape recovery result for a sphere. Left column: Ground truth observed in two different views. Right column: Recovered shape in two different views.

5.2 Real Experiments

The real experiment was conducted to recover the shape of a small *spoon*. The experiment setting was quite simple. The monitor was carried by the drawer and

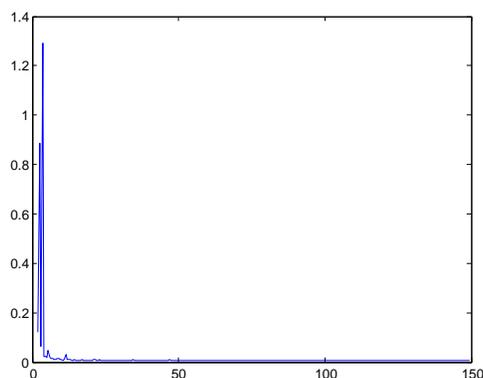


Fig. 4. Relationship between depth recovery error and translation scale. The scene is composed of two spheres which have the same radius of 19, and the translation scale ranges from 1 to 150. 12880 pixels are used for shape recovery. Y-axis shows the average depth recovery error for all pixels, and X-axis represents the translation scale.

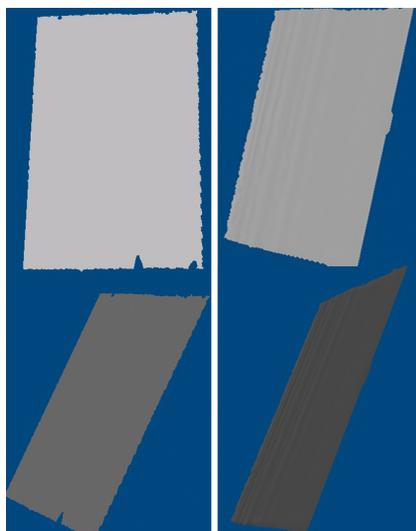


Fig. 5. Shape recovery result for a plane. Left column: Ground truth observed in two different views. Right column: Recovered shape in two different views.

translated on the table (see Fig. 6). The monitor displays gray pattern images and a set of images were taken when the monitor translated in different scales. A traditional structured light method [18] was used to get the correspondences between a coded area on the reference plane and a pixel on the image. The first reference planar pattern is initially calibrated. The reference plane can be translated along a certain direction in different scale and the reconstruction results can be obtained by choosing the best motion scale or combine the multiple motion information together to get a better result. Although the spoon is small and it only reflects a small area on the monitor, the shape can be well recovered (see Fig. 7).

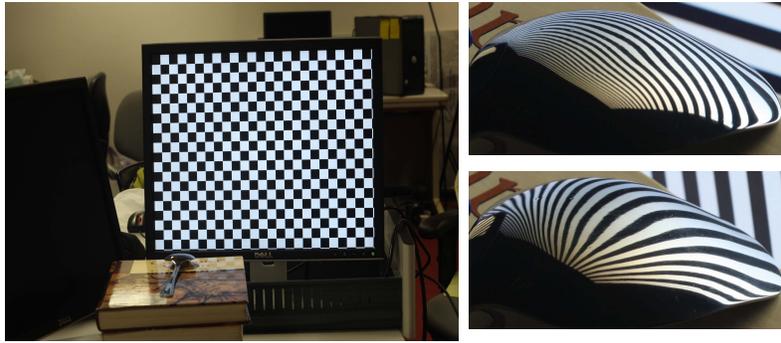


Fig. 6. Setting for the real experiment. Left Column: Setting used in the real experiment. Right Column: sample pattern images used for encoding. The proposed method is to recover the area reflecting the reference plane on the specular object.

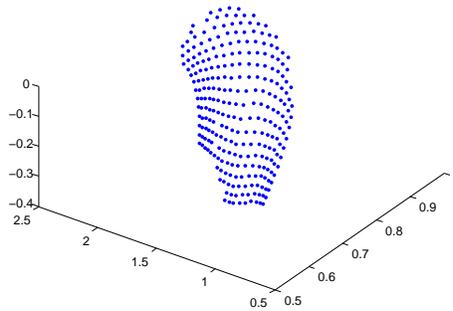


Fig. 7. Shape recovery result for a *spoon*. The recovered shape is the area with reflections on the *spoon*.

6 Conclusions

In this paper, we have proposed a novel solution for specular surface recovery based on observing the reflections of a translating planar pattern. Unlike previous methods which assume a fully calibrated environment, only one reference planar pattern is assumed to have been calibrated against a fixed camera observing the specular surface, and this pattern is allowed to undergo an *unknown pure translation*. The motion of the pattern as well as the shape of the specular object can then be estimated from the reflections of the translating pattern. The main contributions of this paper are

- 1) A closed form solution for recovering the motion of the translating pattern. This allows an easy calibration of the translating pattern, and data redundancy resulting from the translating pattern (i.e., reflections of multiple planar patterns at each pixel location) can improve both the robustness and accuracy of the shape estimation.
- 2) A novel method for shape recovery based on computing the projections of the visual rays on the translating pattern. This produces a depth range for each pixel in which the true surface point exists. This in turn provides a measure of the accuracy and confidence of the estimation.

Experimental results on both synthetic and real data are presented, which demonstrate the effectiveness of our proposed approach. As for the future work, we are now trying to solve the more challenging problems in which 1) the initial reference planar pattern is also uncalibrated and 2) the motion of the reference planar pattern is not limited to pure translation.

References

1. Tomasi, C.: Shape and motion from image streams under orthography: a factorization method. *International Journal of Computer Vision* **9** (1992) 137–154
2. Wong, K.Y.K., Cipolla, R.: Structure and motion from silhouettes. In: *ICCV*. Volume II. (2001) 217–222
3. Woodham, R.: Photometric method for determining surface orientation from multiple images. In: *OptEng* 19(1). (Jan. 1980) 19(1):139–144
4. Zisserman, A., Giblin, P., Blake, A.: The information available to a moving observer from specularities. *Image Vision Comput.* **7** (1989) 38–42
5. Fleming, R. W., T.A., Adelson, E.H.: Specular reflections and the perception of shape. *Journal of Vision* **4** (2003)
6. Oren, M., Nayar, S.K.: A theory of specular surface geometry. *Int. J. Comput. Vision* **24** (1997) 105–124
7. Roth, S., Black, M.J.: Specular flow and the recovery of surface structure. In: *CVPR*. (2006) 1869–1876
8. Adato, Y., Vasilyev, Y., Ben-Shahar, O., Zickler, T.: Toward a theory of shape from specular flow. In: *ICCV*. (2007) 1–8
9. Canas, G.D., Vasilyev, Y., Adato, Y., Zickler, T., Gortler, S., Ben-Shahar, O.: A linear formulation of shape from specular flow. In: *ICCV 2009*. (2009)
10. Savarese, S., Perona, P.: Local analysis for 3d reconstruction of specular surfaces. In: *CVPR*. (2001) II:738–745

11. Savarese, S., Perona, P.: Local analysis for 3d reconstruction of specular surfaces - part ii. In: ECCV. (2002) II:759–774
12. Bonfort, T., Sturm, P.: Voxel carving for specular surfaces. In: ICCV. (2003) 691–696
13. Seitz, S., Matsushita, Y., Kutulakos, K.: A theory of inverse light transport. In: ICCV. (2005) II: 1440–1447
14. Yamazaki, M., Iwata, S., Xu, G.: Dense 3d reconstruction of specular and transparent objects using stereo cameras and phase-shift method. In: ACCV. (2007) II:570–579
15. Nehab, D., Weyrich, T., Rusinkiewicz, S.: Dense 3d reconstruction from specular consistency. In: CVPR. (2008) 1–8
16. Bonfort, T., Sturm, P., Gargallo, P.: General specular surface triangulation. In: ACCV. Volume II. (2006) 872–881
17. Rozenfeld, S., Shimshoni, I., Lindenbaum, M.: Dense mirroring surface recovery from 1d homographies and sparse correspondences. In: CVPR. (2007) 1–8
18. Scharstein, D., Szeliski, R.: High-accuracy stereo depth maps using structured light. Volume 1. (2003) 195–202