Dense mirroring surface recovery from 1D homographies and sparse correspondences

Stas Rozenfeld
Faculty of Industrial Engineering and Management, Technion, Israel Institute of Technology
Haifa, Israel, 32000, rozefeld@tx.technion.ac.il

Ilan Shimshoni
Department of Management Information Systems, University of Haifa
Haifa, Israel, 31905, ishimshoni@mis.hevra.haifa.ac.il

Michael Lindenbaum
Computer Science Department, Technion, Israel Institute of Technology
Haifa, Israel, 32000, mic@cs.technion.ac.il

Abstract

In this work we recover the 3D shape of mirroring objects such as mirrors, sunglasses, and stainless steel objects. A computer monitor displays several images of parallel stripes, each image at a different angle. Reflections of these stripes in a mirroring surface are captured by the camera. For every image point, the directions of the displayed stripes and their reflections in the image are related by a 1D homography which can be computed robustly and using the statistically accurate heteroscedastic model, without monitor–image correspondence, which is generally required by other techniques. Focusing on a small set of image points for which monitor–image correspondence is computed, the depth and the local shape may be calculated relying on this homography. This is done by an optimization process which is related to the one proposed by Savarese, Chen and Perona[10], but is different and more stable. Then dense surface recovery is performed using constrained interpolation, which does not simply interpolate the surface depth values, but rather solves for the depth, the correspondence, and the local surface shape, simultaneously at each interpolated point. Consistency with the 1D homography is thus required. The proposed method as well as the method described in [10] are inherently unstable on a small part of the surface. We propose a method to detect these instabilities and correct them. The method was implemented and the shapes of a mirror, sunglasses, and a stainless steel ashtray were recovered at sub-millimeter accuracy.

1. Introduction

The recovery of specular and particularly mirroring surfaces has been considered for the last few years. It seems however, that the current algorithms are not practical yet for commercial use.

The problem was first addressed by Blake and Brelstaff [2], who proposed a system that can recover the depth map and the orientation of a specular surface, when enough reference Lambertian points exist on it. Several approaches focus on the reconstruction of surface curves. Zisserman, Giblin and Blake [15] track the motion of specularities to obtain information on the surface. Oren and Nayar [8] show how to discriminate between Lambertian and specular (virtual) features. They track the specular features and, using known camera motion, recover the surface at their traces (but not the whole surface). The known camera motion has also been used by Solem et al. in [12], where the proposed framework uses level-set based energy minimization. The minimization provides a smooth surface, that fulfills a relatively small number of constraints that can be motivated by both secular and lambertian properties of the surface. The calibration of the system can be challenging. Also, increasing the density of the depth map can be time consuming, and will mainly base on a smoothness of the surface. In [9] Park and Cho proposed a system which contains a retroreflective hemisphere and a specular object placed in it. The first point is recovered using two laser beams. All other points are recovered through propagation. Zheng and Murata [14] recover the whole shape of a rotating specular object by tracking the specularities created by a toroidal light source.

1-4244-1180-7/07/$25.00 ©2007 IEEE

Several approaches focus on particular object classes. Ripsman and Jenkin [10] recover planar specular objects using a three camera system. Halstead et al. [4] recover roughly symmetric surfaces from the reflection of a light pattern containing concentric circles. A medical application for the shape recovery of the human cornea was also demonstrated.

Other approaches attempt to recover a general object without moving it. In [1] Baba et al. proposed a laser scanning system which is able to scan specular surfaces as well as Lambertian ones. Unfortunately the system works very slowly (it takes about 20 minutes to acquire 100 measurements). Bonfort and Sturm [3] introduce a local and effective process which recovers the shape of the surface using two cameras which observe the distorted reflections of images displayed on a monitor. Similar experimental setup was used by Knauer et al. in [5], where normals to the surface are recovered first and then they are used to recover the depth map. The calibration necessary for this technique is challenging. A method proposed by Tarini et al. [13] uses a single camera which observes the reflection of several images displayed on a monitor. It establishes dense correspondence between the image and the monitor using a color based process which relies on the uniformity of certain photometric properties. This correspondence provides local constraints on depths and the surface normals. They are then integrated into a full surface estimate by global optimization using smoothness assumptions.

Like [13], we use a single camera and a monitor displaying several known patterns. The algorithm we propose, however, is inspired by the work of Savarese, Chen and Perona [11]. Relying on a single pattern containing intersecting lines, they recover the surface depth as well as its higher order properties at a sparse set of points where at least three lines intersect and for which correspondence is available. Their process is fully local: analyzing the differential curve distortion leads to the construction of a matrix parameterized by the unknown depth. This matrix becomes degenerate for the correct depth. See Figure 3 for a plot of their cost function as a function of depth.

In our system a computer monitor displays several images of parallel stripes, each image at a different angle. Distorted reflections of these stripes in a mirroring surface are captured by the camera. For every image point, the directions of the displayed stripes and their reflections in the image are related by a 1D homography. This homography is estimated using robust and statistically valid (heteroscedastic) methods (see [6] for related work), without any knowledge about correspondence. Focusing on a small set of points for which correspondence is extracted using a pseudo-random color pattern proposed by Morano [7], the depth and the local shape may be calculated relying on this homography. This is done by an optimization process which is related to the one proposed in [11], but is different and more stable.

Dense surface recovery is performed as an iterative process which constructs a Delaunay triangulation at every step and adds a new point at the center of each triangle. For every new point the depth, the correspondence, and the local surface shape are estimated locally, using the pre-computed homography and the vertices of the triangle.

The proposed method as well as the method described in [11] are inherently unstable on a small part of the surface where both cost functions are insensitive to the depth. We quantify the instability by measuring the derivative of the cost function and use it as a weight for a smoothing process.

The method was implemented and the shapes of a mirror, sunglasses, and a stainless steel ashtray were recovered. For the planar mirror the estimated depths were consistent with a plane with sub-millimeter accuracy.

2. Notations and mathematical derivations

We start by presenting some notations and results described in [11]. Consider a monitor that displays a known image and a camera that captures its reflection in a mirroring surface; see Figures 1 (left) and 11 (left).

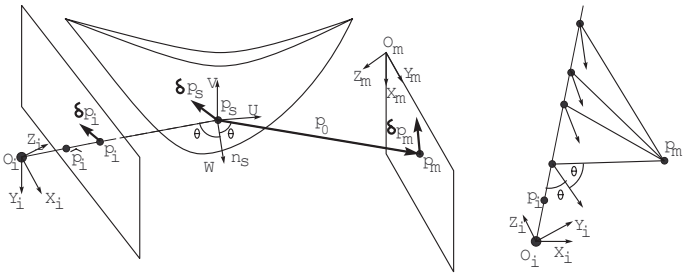

Figure 1. Optical reflection geometry (left), and Surface normal and depth correspondence (right).

Let $\boldsymbol{p_i}$ be an image point corresponding to the point $\boldsymbol{p_m}$ of the reference plane (monitor) and to its reflectance point $\boldsymbol{p_s}$. The depth recovery goal is to estimate the depth s, such that $s\hat{\boldsymbol{p}}_{\boldsymbol{i}} = \boldsymbol{p_s}$, where $\hat{\boldsymbol{p}}_{\boldsymbol{i}} = \boldsymbol{p_i}/\|\boldsymbol{p_i}\|$.

A basic rule of optics is that $\boldsymbol{n_s}$, the normal at $\boldsymbol{p_s}$, lies in the plane specified by the camera's optical center $\boldsymbol{O_i}$, $\boldsymbol{p_i}$ and $\boldsymbol{p_m}$. Let $\boldsymbol{V}$ be the normal to this plane. Moreover, $\boldsymbol{n_s}$ is the bisector of the angle $\angle \boldsymbol{O_i p_s p_m}$. Denote half of this angle as θ. Let $\{\boldsymbol{O_i}, [\boldsymbol{X_i}, \boldsymbol{Y_i}, \boldsymbol{Z_i}]\}$ be the camera coordinate system specified relative to the camera's optical center, $\{\boldsymbol{p_m}, [\boldsymbol{X_m}, \boldsymbol{Y_m}, \boldsymbol{Z_m}]\}$ the monitor coordinate system, and $\{\boldsymbol{p_s}, [\mathbf{U} = \mathbf{V} \times \boldsymbol{n_s}, \mathbf{V}, \mathbf{W} = \boldsymbol{n_s}]$ the (local) reflective surface coordinate system. The first coordinate system is known. The second coordinate system is estimated using an external calibration process. The point $\boldsymbol{p_m}$ is specified by a correspondence process, which defines $\boldsymbol{V}$ and a one-to-one relationship between s and $\boldsymbol{n_s}$ as illustrated in Fig-

ure 1 (right). For a known s, the third coordinate system together with $\boldsymbol{p_0} = \boldsymbol{p_m} - \boldsymbol{p_s}$ are specified as well. Most of our calculations are done in the surface coordinate system.

Consider a line $\boldsymbol{p_m}(t) = \boldsymbol{p_m} + t\boldsymbol{\delta p_m}$, $t \in R$ where $\boldsymbol{\delta p_m} = \cos\alpha_m \boldsymbol{X_m} + \sin\alpha_m \boldsymbol{Y_m}$. Clearly, $\boldsymbol{p_m}(t)$ lies in the monitor plane and passes through $\boldsymbol{p_m} = \boldsymbol{p_m}(0)$. Let $[\boldsymbol{p_u}, \boldsymbol{p_v}, \boldsymbol{p_w}]^T$ be the coordinates of vector $\boldsymbol{\delta p_m}$ in the $\{\boldsymbol{p_m}, [\boldsymbol{X_m}, \boldsymbol{Y_m}, \boldsymbol{Z_m}]\}$ coordinate system. The line $\boldsymbol{p_m}(t)$ induces, in a natural way, a curve $\boldsymbol{p_s}(t)$ on the mirroring surface and another curve $\boldsymbol{p_i}(t)$ in the image plane. Note that if the mirroring surface is a plane, then $\boldsymbol{p_s}(t)$ and $\boldsymbol{p_i}(t)$ are straight lines and not curves. Clearly, $\boldsymbol{p_s}(0) = \boldsymbol{p_s}$ and $\boldsymbol{p_i}(0) = \boldsymbol{p_i}$. Let $\boldsymbol{\delta p_s} = \dot{\boldsymbol{p_s}}(0)$ and $\boldsymbol{\delta p_i} = \dot{\boldsymbol{p_i}}(0)$ be the derivatives of $\boldsymbol{p_s}(t)$ at $\boldsymbol{p_s}$ and of $\boldsymbol{p_i}(t)$ at $\boldsymbol{p_i}$, respectively. Let α_i be the angle for which $\boldsymbol{\delta p_i} = \|\boldsymbol{\delta p_i}\|(\cos\alpha_i \boldsymbol{X_i} + \sin\alpha_i \boldsymbol{Y_i})$. The vector $\boldsymbol{\delta p_s}$ lies in the tangent plane of the surface at $\boldsymbol{p_s}$. Therefore, there is an angle α_s such that $\boldsymbol{\delta p_s} = \|\boldsymbol{\delta p_s}\|(\cos\alpha_s \boldsymbol{U} + \sin\alpha_s \boldsymbol{V})$. Finally, let a, b and c be the parameters describing the second order approximation $w = \frac{1}{2}au^2 + cuv + \frac{1}{2}bv^2$ of the mirroring surface close to $\boldsymbol{p_s}$ in its local coordinate frame.

The following expression, specifying a linear transformation between tangent vectors, follows from the results described in [11]:

$$\|\boldsymbol{\delta p_s}\| \begin{bmatrix} \cos\alpha_s \\ \sin\alpha_s \end{bmatrix} = \mathcal{A}\mathcal{B} \begin{bmatrix} \cos\alpha_m \\ \sin\alpha_m \end{bmatrix}, \tag{1}$$

where

$$\Delta = (J_u - 2a\cos\theta)(J_v - 2b\cos\theta) - (2c\cos\theta)^2,$$
$$J_u = \cos^2\theta \frac{s+\|\mathbf{p}_0\|}{s\|\mathbf{p}_0\|},$$
$$J_v = \frac{s+\|\mathbf{p}_0\|}{s\|\mathbf{p}_0\|},$$
$$\mathcal{A} = \frac{1}{\Delta}\begin{bmatrix} J_v - 2b\cos\theta & 2c\cos\theta \\ 2c\cos\theta & J_u - 2a\cos\theta \end{bmatrix},$$
$$\mathcal{B} = \frac{1}{\|\mathbf{p}_0\|}\begin{bmatrix} -\cos^2\theta & 0 & \cos\theta\sin\theta \\ 0 & -1 & 0 \end{bmatrix}[(\boldsymbol{X_m})^{\boldsymbol{s}}, (\boldsymbol{Y_m})^{\boldsymbol{s}}],$$

and $(\boldsymbol{X_m})^{\boldsymbol{s}}$ and $(\boldsymbol{Y_m})^{\boldsymbol{s}}$ are the column vectors $\mathbf{X}_m, \mathbf{Y}_m$ expressed in the surface coordinate system.

We observe that this linear transformation, denoted T_{ms}, also induces a 1D homography between the monitor angles and the corresponding surface angles. We denote this homography H_{ms}. The camera observes the surface tangent plane, implying that there is another 1D homography between the surface tangent plane angles and the image plane angles. This homography, H_{si}, can be computed from $\boldsymbol{U}$ and $\boldsymbol{V}$ as follows:

$$H_{si} \cong \begin{bmatrix} \boldsymbol{U}_x - (\boldsymbol{p_i})_x \boldsymbol{U}_z & \boldsymbol{V}_x - (\boldsymbol{p_i})_x \boldsymbol{V}_z \\ \boldsymbol{U}_y - (\boldsymbol{p_i})_y \boldsymbol{U}_z & \boldsymbol{V}_y - (\boldsymbol{p_i})_y \boldsymbol{V}_z. \end{bmatrix}. \tag{2}$$

The homography $H_{mi} = H_{si}H_{ms}$ relates the monitor angles and the image plane angles. Therefore the combined homography, as well as its components, can be computed from the depth s, the image-monitor correspondences between $\boldsymbol{p_i} = [(\boldsymbol{p_i})_x, (\boldsymbol{p_i})_y]^T$ and $\boldsymbol{p_m}$, and the surface local shape characterization parameters a, b and c.

On the other hand, the same homography H_{mi} may be computed empirically, without knowing the depth, the corresponding point, and the surface parameters. Comparing the two versions of the same homography provides a constraint that can be used to estimate the depth.

To estimate H_{mi}, monitor–image angle correspondences are needed. For that we use a sequence of displayed images, each associated with a constant tangent direction, defining the monitor angles $\{\alpha_m^k\}_{k=1}^K$ for all the points in the displayed image. Let $\{\alpha_i^k\}_{k=1}^K$ be the sequence of tangent angles associated with a particular image point. See examples of such images in Figure 2. With this sequence of angle pairs, H_{mi} can be estimated combining a robust method with the heteroscedastic technique that we developed for 1D homography estimation. This method provides results superior to standard linear regression techniques such as SVD. Let H_{mi}^{emp} denote the estimated homography. Note that it may be calculated independently for every point in the image.

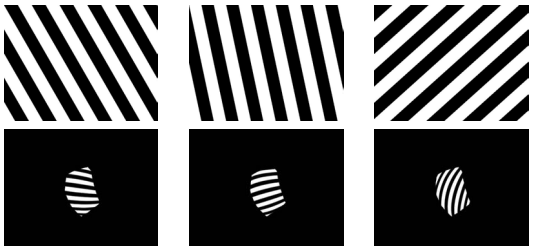

Figure 2. Typical displayed images (top), and their corresponding captured images (bottom). Such pairs are used for the 1D homography estimation.

The next subsection describes, following [11], how the various components of the analytical homography H_{mi} are calculated.

2.1. Calculating H_{mi} analytically

Consider a sequence of corresponding angle pairs $\{\alpha_i^k, \alpha_m^k\}_{k=1}^K$, a correspondence $\{\boldsymbol{p_i}, \boldsymbol{p_m}\}$, and a suggested depth s. The surface coordinate system, as well as the angle θ, may be calculated, leading to $\mathcal{B}$ and an estimate of H_{si}. Now, angles α_s^k are estimated as follows:

$$\begin{bmatrix} \cos\alpha_s^k \\ \sin\alpha_s^k \end{bmatrix} \cong H_{si}^{-1} \begin{bmatrix} \cos\alpha_i^k \\ \sin\alpha_i^k \end{bmatrix}. \tag{3}$$

The $K \times 3$ matrix

$$M = M(s, \boldsymbol{p_i}, \boldsymbol{p_m}, \{\alpha_i^k, \alpha_m^k\}_{k=1}^K) = \begin{bmatrix} B_v^1 & -B_u^1 \tan\alpha_s^1 & B_u^1 - B_v^1 \tan\alpha_s^1 \\ \dots & \dots & \dots \\ B_v^k & -B_u^k \tan\alpha_s^k & B_u^k - B_v^k \tan\alpha_s^k \\ \dots & \dots & \dots \\ B_v^K & -B_u^K \tan\alpha_s^K & B_u^K - B_v^K \tan\alpha_s^K \end{bmatrix}, \quad (4)$$

where

$$\begin{bmatrix} B_u^k \\ B_v^k \end{bmatrix} = \mathcal{AB} \begin{bmatrix} \cos\alpha_m^k \\ \sin\alpha_m^k \end{bmatrix}, \quad (5)$$

has rank 2. Its kernel is $\boldsymbol{\nu}/\|\boldsymbol{\nu}\|$, where $\boldsymbol{\nu}^T = [J_u - 2a\cos\theta\ ,\ J_v - 2b\cos\theta\ ,\ 2c\cos\theta]$. This kernel specifies $\mathcal{A}$ up to a scale factor.

The product $H_{mi}^{an} = H_{si}\mathcal{AB}$ is an estimate of H_{mi}. Therefore, the depth s can be estimated by comparing H_{mi}^{an} with H_{mi}^{emp}. This method is an alternative to the process described in [11], which builds only on the rank 2 property of M and does not use the alternative empirical estimate H_{mi}. For stable points the depth estimates were roughly the same. One advantage of our optimization criterion is its geometric interpretation. Another advantage is the larger basin of attraction as can be seen in Figure 3.

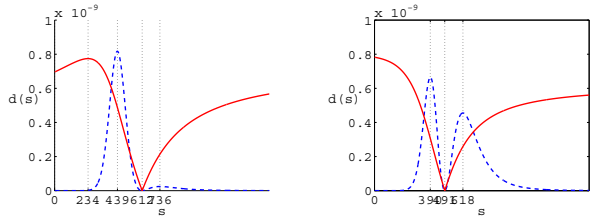

Figure 3. The cost function used in [11] (dashed line) vs. the proposed cost function (solid line). The two plots correspond to two different image points.

We would now like to turn our attention to the estimation of the second order surface parameters a, b and c. This only requires that we estimate the missing scale parameter to transform H_{ms} into the linear transformation T_{ms}. In our implementation this is done using the scale based method of [11]. The relationship between the length of a line on the monitor and its projection on the image is used to recover the missing scale parameter. From T_{ms}, $\mathcal{A}$ can be recovered, leading to the estimation of a, b and c.

2.2. The planar case

In the final part of this section we would like to describe a simple method that allows planar mirroring surfaces to be recovered. We would like to emphasize that this method is not a component of the general algorithm and is not even a special case of it. It just provides a simpler, intuitive, recovery method for a limited case.

For planar surfaces the reflection of the planar monitor is a planar virtual object as illustrated in Figure 4. The superscript r denotes the properties of the reflection. In addition, the 1D homography between the monitor and the image, H_{mi}, is also a 1D homography between the monitor's reflection and the image. This is because planar mirrors preserve angles and distances. In particular, $\alpha_m^k = \alpha_m^{k\,r}$. The directions associated with the coordinate frame, $\{\boldsymbol{p_m^r}, [\boldsymbol{X_m^r}, \boldsymbol{Y_m^r}, \boldsymbol{Z_m^r}]\}$, of the reflected (virtual) monitor are specified by three parameters. The 2×2 1D homography is specified up to a scale and thus provides three constraints on these parameters (eq. 2) and enables their recovery. For planar mirrors the difference between the normal to the monitor and the normal to its reflection is the normal to the mirror W:

$$\boldsymbol{W} = \frac{\boldsymbol{Z_m^r} - \boldsymbol{Z_m}}{\|\boldsymbol{Z_m^r} - \boldsymbol{Z_m}\|}.$$

Note that the normal can be recovered at any image point using H_{mi} only, as described above. Thus, an accurate estimation of W can be computed by averaging over several points in the image. There is however, more than one solution satisfying eq. 2, but only one of the solutions consistently yields the same normal for every point on the plane and thus can be recognized. In order to recover the plane displacement it is sufficient to establish the correspondence $\{\boldsymbol{p_m}, \boldsymbol{p_i}\}$ at a single point.

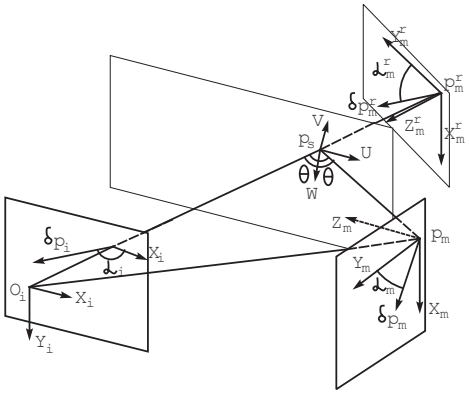

Figure 4. Surface reconstruction. Planar case.

3. Depth estimation for initial points

In this section we explain how to initialize the surface reconstruction process by estimating the depth of a few initial points. We select these points, denoted $\{\boldsymbol{p_i^l}\}_{l=1}^L$, as those for which corresponding monitor points $\{\boldsymbol{p_m^l}\}_{l=1}^L$ are available. For each one of these points, the depth s is estimated using the following optimization process:

$$\hat{s} = \arg\min_{s} \; d_{ang}(\quad H^{emp}_{mi}(\{\alpha^k_i, \alpha^k_m\}^K_{k=1}), \\ H^{an}_{mi}(s, \boldsymbol{p_i}, \boldsymbol{p_m}, \{\alpha^k_i, \alpha^k_m\}^K_{k=1})) \tag{6}$$

where $d_{ang}(\cdot,\cdot)$ is the distance between the two 1D homographies. The distance is defined by placing the entries of the two matrices into $4D$ vectors and computing the angle between them. Note that as H^{emp}_{mi} is known and H^{an}_{mi} is only a function of s, we can denote $d_{ang}(H^{emp}_{mi}, H^{an}_{mi})$ as $d(s)$.

Typical initial point reconstruction associated with a smooth surface and recovered using this 1D optimization is shown in Figure 5 (left). Clearly, the reconstruction is problematic around some curve. This problem, which was already observed in [11], occurs in locations where the homography H^{an}_{mi} is highly insensitive to depth. In such cases, the inevitable inaccuracy in the image data (image point locations $\boldsymbol{p_i}$ and their corresponding angles α^k_i) is significantly amplified. Fortunately, this insensitivity is simple to detect. Figure 6 shows the cost function associated with a stable and an unstable point near the correct depth. We refer to the above phenomenon as the *stability problem*.

To detect regions associated with instability, we calculate a *stability measure*, $\epsilon = \frac{\partial d(s)}{\partial s}$. Figure 5 (right) shows the stability measure computed for each initial point. Points resulting in a reconstruction discontinuity are exactly those corresponding to a low stability measure values.

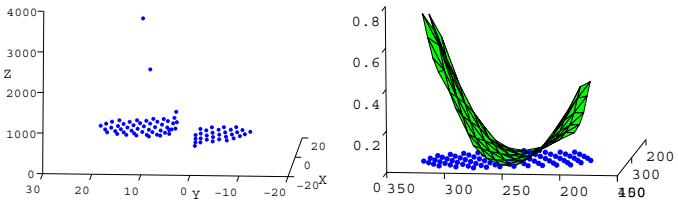

Figure 5. (Left) - depth estimates at the initial point set obtained by the local optimization. (Right) - The stability measure evaluated at the same points. Note that it is very low for the points where the calculation is inaccurate.

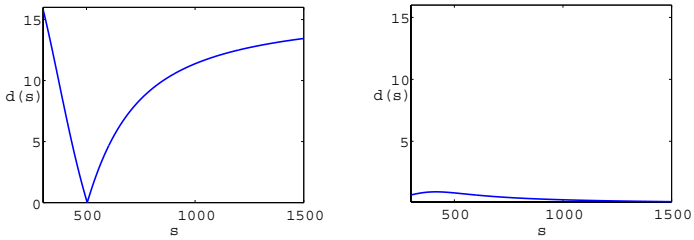

Figure 6. The cost function associated with a stable point (left) and a problematic, unstable point (right).

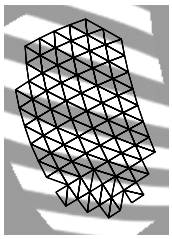

Figure 7. Image triangulation

3.1. Correcting instability

In contrast to [11], which only detected the unstable points, our goal is to fully reconstruct the shape. Thus, we correct the unstable depth estimates using a two stage smoothing process: First we perform a Delaunay triangulation on the set of initial image points, to specify each point's neighbors (Figure 7). The neighborhood $\mathcal{N}_l$ of $\boldsymbol{p^l_i}$ is the set of direct neighbors in the triangulation and $\boldsymbol{p^l_i}$ itself.

Let s_l be a depth variable, associated with the initial point $\boldsymbol{p^l_i}$. Initialize s_l to the initially estimated depth. Let $s_{l\leftarrow l'}$ be a depth estimate at $\boldsymbol{p^l_i}$ obtained by a first order surface approximation relying on a neighbor $\boldsymbol{p^{l'}_i}$. The estimate is calculated from the depth s_l and the corresponding normal to the surface at that point. Clearly, $s_{l\leftarrow l} = s_l$. Let ϵ_l be the stability measure associated with the point $\boldsymbol{p^l_i}$. We update the depth estimation at every point, using the depth estimates associated with its neighbors and their stability, by minimizing the MSE. That is

$$s_l := \arg\min_{s} \sum_{l' \in \mathcal{N}_l} \epsilon^{\gamma}_{l'} (s - s_{l\leftarrow l'})^2, \tag{7}$$

where γ is a constant ($\gamma = 10$ was used in most experiments). The optimal estimated depth is clearly the weighted average

$$s_l := \frac{\sum_{l' \in \mathcal{N}_l} \epsilon^{\gamma}_{l'} s_{l\leftarrow l'}}{\sum_{l' \in \mathcal{N}_l} \epsilon^{\gamma}_{l'}}. \tag{8}$$

This way, less stable neighbors have a lower contribution to the depth estimate. We repeated this weighted smoothing iteratively. In each iteration all the initial points are traversed in descending stability measure order. The depth is updated after each iteration, making the process similar to the Gauss-Seidel method. Note that updating the depth induces an updated surface normal as well. We repeat these iterations until the process converges. See Figure 8 (left) for typical results.

With the relatively high γ value, the contribution of the initial large depth errors is negligible. However, with high γ, the number of values which are effectively averaged is low. Therefore, after the first smoothing stage the depth estimate is relatively rough, especially near the low stability

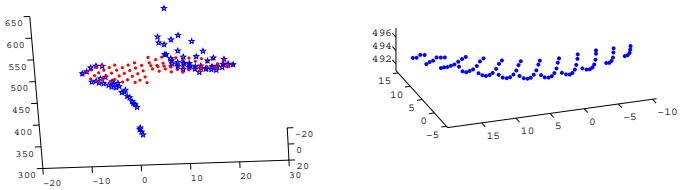

Figure 8. (Left) - the depth estimate before (blue stars) and after one smoothing iteration (red dots). (Right) - the final depth estimation of the initial points.

curve. In a second smoothing stage, the estimated depths are further uniformly averaged, creating a smooth, more accurate reconstruction. For results see Figure 8 (right).

Note that both smoothing processes are constrained by the correspondences between monitor points and image points, yielding a one-to-one correspondence between the depth and the surface normal. Therefore the process does not converge to a trivial constant solution.

4. Increasing the density of the depth Map

We now augment the initial points with additional points, leading to a dense surface reconstruction. Additional points are added by an iterative process that does not change the existing points.

In the beginning of each iteration, we triangulate the image plane using the already recovered points $\{\boldsymbol{p_i^l}\}_{l=1}^{L}$; see Figure 7. We add a new image point at the center of mass of each triangle. Every point will be calculated using a local process. For every such point we now aim to calculate the depth, the surface normal, the corresponding point on the display, and the second order parameters a, b and c. All these parameters are required for the next iterations.

Depth - A second order estimate of the depth at the new point is obtained from every vertex of its triangle. This estimate uses the depth and the second order shape parameters a, b and c, associated with them. The average of these three estimates serves as the estimate of the depth of the new point (Figure 9 (left)).

Correspondence - The point $\boldsymbol{p_m}$, corresponding to the new image point $\boldsymbol{p_i}$ is initially set ($\boldsymbol{p_m^{init}}$) to be the center of mass associated with the monitor points corresponding to the vertices of the triangle. A final estimate $\hat{\boldsymbol{p}}_{\boldsymbol{m}}$ is found by relating the 1D homography, which depends on this corresponding point, to the empirical 1D homography available at every image point.

$$\hat{\boldsymbol{p}}_{\boldsymbol{m}} = \underset{\boldsymbol{p_m} | d_{ang}(H_{mi}^{emp}, H_{mi}^{an}) < \lambda}{\arg\min} \|\boldsymbol{p_m^{init}} - \boldsymbol{p_m}\|^2 \qquad (9)$$

The limit λ on the distance between the homographies is set to account for discretization and homography estimation error. In our experiments λ was set to 2°. This process extracts the surface normal as well.

The second order surface parameters a, b and c - Given the correspondence, these parameters may be calculated up to an unknown scale parameter τ. Given an hypothesized value for τ, the depths at the three triangle vertices may be estimated from the implied second order surface estimate and compared with the known depths at these points (estimated in earlier iterations). The difference is used as the cost function for the 1D optimization process for finding τ (Figure 9 (right)). This method is much simpler than the one described in Section 2.1 for estimating the second order surface parameters for the initial points and does not require additional monitor–image correspondences.

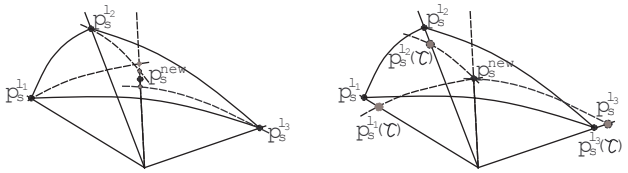

Figure 9. Adding a single point within a triangle: estimating the depth of the new point (left) and testing depth consistencies as part of estimating the second order parameters (right).

Now, after we have estimated the depth, the point correspondence and the second order surface parameters at each new point, we can start a new iteration and further increase the sampling density. This process continues until we reach the desired density.

5. Experimental results

We have implemented our technique and tested it on three mirroring objects, each made from a different material. The first one was a planar mirror, which allowed us to quantify the quality of the estimate. For the other objects, a pair of sunglasses and a stainless steel ashtray, which are general non-planar surfaces, no ground truth was available.

The images displayed on the monitor were designed so that the algorithm can extract the needed information:

1. A correspondence between the monitor and the camera for a sparse set of points, and

2. a correspondence between the orientation at every image point and the orientation at its corresponding point on the monitor, $\{\alpha_i^k, \alpha_m^k\}_{k=1}^{K}$, which is required to estimate H_{mi}^{emp} at every image point.

The first correspondence is obtained using the approach proposed by Morano *et al.* [7]. In this approach the monitor displays a color pattern such as the one shown in Figure 10 (left). This pattern is characterized by the fact that each 3×3 neighborhood is unique. Therefore, once we

have identified the neighborhood of a captured patch, shown in Figure 10 (right), the corresponding neighborhood in the displayed image shown in Figure 10 (left) can be found.

For each patch, a monitor point $\boldsymbol{p_m}$, in the initial set, is the center of a monitor patch. Its corresponding image point $\boldsymbol{p_i}$ is estimated as the patch's center of mass in the image. In addition, more images, containing several points placed at known distances and directions around each of the initial points are displayed as well. This enables the algorithm to estimate the true scale of $\mathcal{A}$, leading to the recovery of the local second order estimate of the shape around the point.

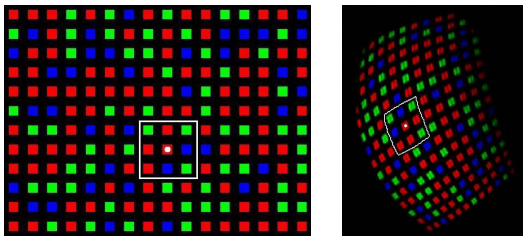

Figure 10. Finding the correspondence of the initial points using a pseudo-random color template. (left) - Displayed image; (right) - Captured image.

Achieving correspondence between the image orientations and the corresponding orientations at the monitor points generally requires dense correspondence of locations. To avoid this difficulty we display about 100 images of parallel black and white stripes, where the orientations are the same all over the image. See Figure 2 (top). In the corresponding images (Figure 2 (bottom)) the tangent directions are computed at the edges of the stripes. For points not lying on the edges, the image directions are computed by interpolation.

The recovered surfaces obtained in the experiments are shown in Figure 11. For each experiment, the first image in the row shows an image of the object. The plots in each row show how the sampling density increases as the algorithm progresses, resulting in smooth reconstructions.

The results of the planar mirror reconstruction allow us to quantify the accuracy of the results. First a plane is fit to the recovered points and the residuals of their depths relative to the plane are computed. Figures 12 (left and middle), show the residuals associated with the depths of the initial points, obtained by the local, homography comparison method, and after the improvement achieved by the smoothing process.

The third plot in Figure 12 shows a histogram of the residuals for the much larger dense set of points obtained by the full algorithm. These residuals are bounded by $[-0.1, 0.05]$ millimeters. We are aware of the fact that the planar mirror is a very simple example and thus is likely to allow accurate estimation. Still, the results are promising.

In the final experiment we compared the general algorithm to the special algorithm designed to deal with planar objects (Section 2.2) by applying them to the planar mirror. The algorithms yield quite accurate results, although the more general algorithm is somewhat better. On the other hand, the planar object specific method is faster. The algorithms were run on 114 points on the surface, recovering the surface normal at each point. In Figure 13 we plot the histogram of the angle differences between each pair of normals for both algorithms in degrees. Both algorithms estimate the normal with an average dispersion which less than $0.8°$.

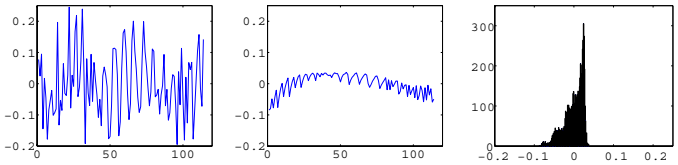

Figure 12. Residuals in millimeters for planar surface reconstruction: (left) - for the initial points after local depth estimation computation; (middle) - after smoothing. (right) - the residual histogram for the densely reconstructed surface.

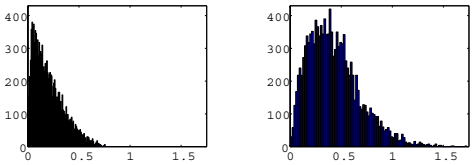

Figure 13. Histogram of differences between the normals estimated at the initial points of the general algorithm: (left) - using the general algorithm; (right) - using the simple algorithm.

6. Conclusions

This paper describes a new method for 3D shape recovery of mirroring surfaces. The proposed method builds on earlier approaches [11, 13] but differs from them in several important aspects.

Specifically, it builds on a 1D homography evaluated using robust and statistically valid (heteroscedastic) methods, at every point in the image. Therefore, it is eventually able to provide a densely recovered surface, while requiring only a few initialization points where correspondence is needed.

The depth calculation, at the initial points, is similar to that described in [11], but uses a different cost function, which does not suffer from "ghost solutions" [11], and usually has a single minimum. Unlike [13], these calculations use only local information, and are, therefore, both faster and less biased to smooth surfaces.

A special simple algorithm is also presented for planar surfaces. This algorithm is able to recover the plane accurately using many angle correspondences like in the general algorithm, but only one point correspondence.

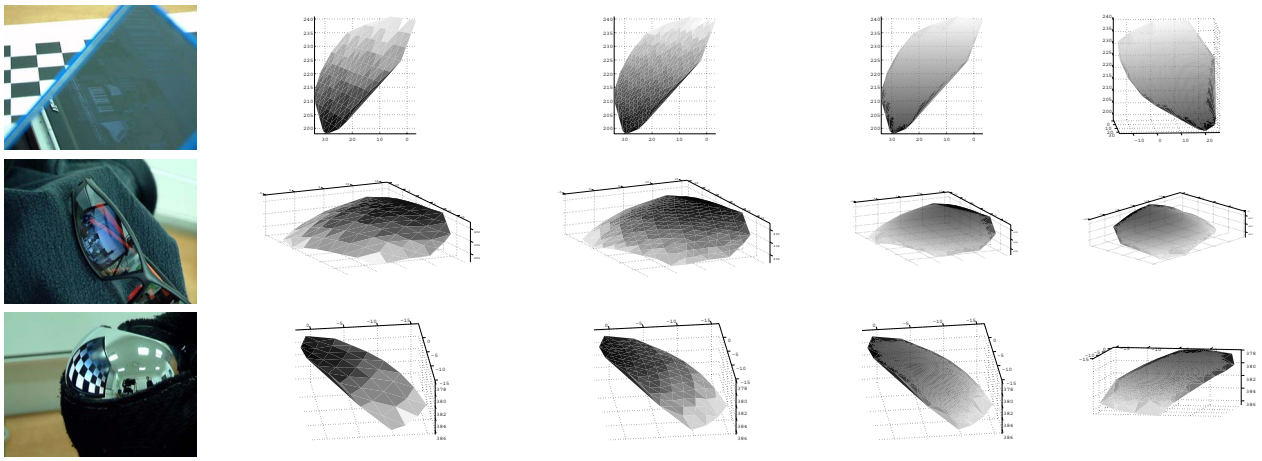

Figure 11. Experimental results: recovering the surfaces of a planar glass mirror (top), plastic sunglasses (middle), and a stainless steel ashtray (bottom). From left to right: the object, the final estimation of the initial points, after one density enhancement iteration, the final surface (four iterations), another view of the final surface.

The surface is recovered with high density using constrained interpolation, which does not simply interpolate the surface, but rather solves for the depth, the correspondence, and the local surface shape, simultaneously at each interpolated point, requiring consistency with the 1D homography. This process, is completely local and does not require a precalculated, dense, correspondence, which may be difficult to obtain.

The proposed method is essentially a structured light approach for recovering the shape of specular and mirror-like objects. Note that the method does not rely on photometric measurements. Therefore, no photometric calibration is needed. This characteristic, along with the dense, accurate surface recovery, make it suitable for real-world applications such as specular parts quality inspection.

References

[1] M. Baba, K. Ohtani, M. Imai, and T. Konishi. New laser rangefinder for three-dimensional shape measurement of specular objects. *Optical Engineering*, 40(1):53–60, 2001. 2

[2] A. Blake and G. Brelstaff. Geometry from specularities. In *Proc. Int. Conf. Comp. Vision*, pages 394–403, 1988. 1

[3] T. Bonfort and P. Sturm. Voxel carving for specular surfaces. In *Proc. Int. Conf. Comp. Vision*, pages 591–596, 2003. 2

[4] M. Halstead, B. Barsky, S. Klein, and R. Mandell. Reconstructing curved surfaces from specular reflection patterns using spline surfce fitting of normals. In *ACM Computer Graphics and Interactive Techniques*, pages 335–342, New Orleans, 1996. 2

[5] M. C. Knauer, J. Kaminski, and G. Häusler. Phase measuring deflectometry: a new approach to measure specular free-form surfaces. In *Optical Metrology in Production Engineering. Edited by Osten, Wolfgang; Takeda, Mitsuo. Proceedings of the SPIE, Volume 5457, pp. 366-376 (2004).* 2

[6] Y. Leedan and P. Meer. Heteroscedastic regression in computer vision: Problems with bilinear constraint. *Int. J. of Comp. Vision*, 37(2):127–150, June 2000. 2

[7] R. Morano, C. Ozturk, R. Conn, S. Dubin, S. Zietz, and J. Nissanov. Structured light using pseudorandom codes. *IEEE Trans. Patt. Anal. Mach. Intell.*, 20(3):322–327, March 1998. 2, 6

[8] M. Oren and S. Nayar. A theory of specular surface geometry. In *Proc. Int. Conf. Comp. Vision*, pages 740–747, 1995. 1

[9] W. Park and H. Cho. Measurement of the 3-dimensional shapes of specular objects using recursive triangulation. In *Pacific Conference on Manufacturing*, volume 1, pages 389–394, Seoul, Korea, 1996. 1

[10] A. Ripsman and M. Jenkin. Local surface reconstruction of objects in space. *IEEE Int. Symposium on Computational Intelligence*, 2001. 2

[11] S. Savarese, M. Chen, and P. Perona. Local shape from mirror reflections. *Int. J. of Comp. Vision*, 64(1):31–67, August 2005. 2, 3, 4, 5, 7

[12] J. Solem, H. Aanaes, and A. Heyden. Variational surface interpolation from sparse point and normal data. *IEEE Trans. Patt. Anal. Mach. Intell.*, 29(1):181–184, January 2007. 1

[13] M. Tarini, H. P. A. Lensch, M. Goesele, and H.-P. Seidel. 3D acquisition of mirroring objects using striped patterns. *Graph. Models*, 67(4):233–259, 2005. 2, 7

[14] J. Zheng and A. Murata. Acquiring a complete 3D model from specular motion under the illumination of circular-shaped light sources. *IEEE Trans. Patt. Anal. Mach. Intell.*, 22(8):913–920, 2000. 1

[15] A. Zisserman, P. J. Giblin, and A. Blake. The information available to a moving observer from specularities. *Image and Vision Computing*, 7(1):38–42, 1989. 1