Shape from Planar Curves: A Linear Escape from Flatland

Ady Ecker   Kiriakos N. Kutulakos   Allan D. Jepson
University of Toronto

Abstract

We revisit the problem of recovering 3D shape from the projection of planar curves on a surface. This problem is strongly motivated by perception studies. Applications include single-view modeling and fully uncalibrated structured light. When the curves intersect, the problem leads to a linear system for which a direct least-squares method is sensitive to noise. We derive a more stable solution and show examples where the same method produces plausible surfaces from the projection of parallel (non-intersecting) planar cross sections.

1. Introduction

Extracting depth from a single image of curves is a long standing problem. Since projection discards depth information, the problem is severely under-constrained. Clearly, for any curve in the image, there is always an infinite set of 3D curves whose projections coincide with that image curve. Therefore additional assumptions have to be made. One assumption that has been studied extensively is that the curves are planar, either lying on planar faces of polyhedra or planar cross sections of curved surfaces. Perception studies [16, 19] have demonstrated cases where humans clearly make the planarity assumption. Pizlo *et al*. [10] showed that planarity can even override stereo depth cues. Although the extraction of salient curves from real images is a difficult problem by itself, it is hard to believe the sophisticated ability to interpret curves as 3D surfaces had been evolved with no relation to depth perception of real scenes.

Figure 1 illustrates fully uncalibrated structured light. The curves in the image are projections of planar cross sections of the surface. The goal is to compute the 3D surface from the observed network of curves without any information about the planes. Notice that all planar curves could be placed on the same plane. A key here is to escape from flat and nearly-flat solutions.

This paper makes several contributions to the theory of shape from planar curves. We present a unified framework for interpreting projections of intersecting planar faces or cross sections, and demonstrate cases where the same principles can be used to deal with parallel planar cross sections. These important connections have not been pointed out in the literature before. We extend the method of Bouguet, Weber and Perona [2] to deal with arbitrary planes and derive a more stable algorithm. Specifically, we look for a solution that minimizes the algebraic error of a linear system, under the constraint that a geometrically meaningful measure of non-planarity is held constant. In addition, we provide statistical analysis for the component of the solution vector which lies in the trivial subspace, and explore solution ambiguities.

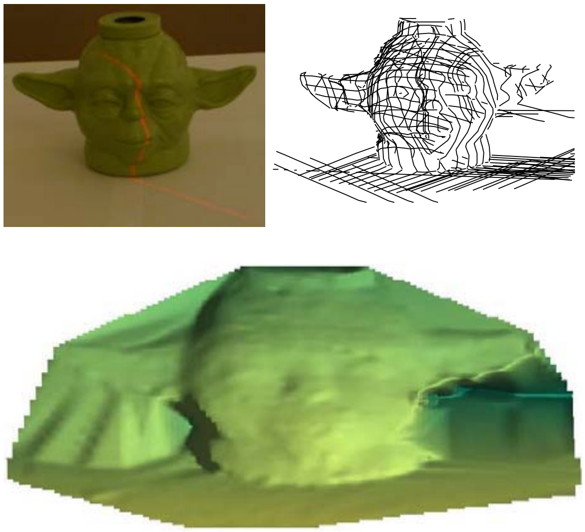

Figure 1. 3D reconstruction from uncalibrated structured light. Top left: a planar laser strip projected on the object. Top right: curves extracted from a video sequence of projected strips. Bottom: rendering of the computed surface.

2. Algebraic structure

In this section we review the algebraic structure of the problem and characterize the space of solutions in the ideal noiseless case. The theory for interpreting line drawings of polyhedra was formulated algebraically as a linear system by Sugihara [18]. Ulupinar and Nevatia [21] found similar structure in Straight Homogeneous Generalized Cylinders (SHGC). Our presentation is closer to

1-4244-1180-7/07/$25.00 ©2007 IEEE

Rothwell *et al.* [12, 13] and Bouguet *et al.* [2].

Consider a surface on which N planar curves $\Gamma_1, \ldots, \Gamma_N$, lying on N planes $\pi_1, \ldots, \pi_N$, are marked. Each curve Γ_i is just a set of points on a common plane π_i, and does not need to be continuous. For now assume an image is taken under orthographic projection, and the intersection points (x_{ij}, y_{ij}) between the projections of the curves Γ_i and Γ_j onto the image are identified. The setting is illustrated in figure 2. Our input is the intersection points, together with their association to the intersecting curves. The goal is to recover the planes and compute the depth along the curves. Note that two curves may not intersect in the image or they may intersect several times. For simplicity we omit the intersection multiplicity index. If two curves share a straight line segment, we pick its two endpoints as intersection points. When more than two curves intersect at a point, we consider them as pairs in a cyclic order.

Assuming the planes do not contain the projection direction (an edge-on plane provides no information and should be ignored), we can parameterize plane π_i as

$$z_i(x, y) = a_i x + b_i y + d_i \; . \tag{1}$$

A 2D intersection point in the image corresponds to a 3D intersection on the surface, which allows us to eliminate the unknown depths:

$$\begin{aligned} &z_i(x_{ij}, y_{ij}) - z_j(x_{ij}, y_{ij}) = \\ &\quad (a_i - a_j)x_{ij} + (b_i - b_j)y_{ij} + (d_i - d_j) = 0 \; . \end{aligned} \tag{2}$$

The last equation has another geometric interpretation. Planes π_i and π_j intersect at a 3D line whose projection onto the image is the image line L_{ij} defined by

$$(a_i - a_j)x + (b_i - b_j)y + (d_i - d_j) = 0 \; . \tag{3}$$

Equation (2) simply means that (x_{ij}, y_{ij}) is on L_{ij}. Collecting equations (2) for all visible intersection points results in a homogeneous linear system:

$$\mathbf{A}\mathbf{v} = 0 \; , \tag{4}$$

where $\mathbf{v} = (a_1, \ldots, a_N, b_1, \ldots, b_N, d_1, \ldots, d_N)^T$ is a vector collecting the planes' parameters and $\mathbf{A}$ is a sparse matrix whose rows contain $x_{ij}, -x_{ij}, y_{ij}, -y_{ij}, 1, -1$ at the appropriate columns. Any vector $\mathbf{v}$ defines surface curves $\{\Gamma_i\}$ by the back-projection of the image curves onto their planes. Sugihara already noted that the problem has similar structure under perspective projection (see appendix A.1).

2.1. GBR ambiguity and trivial subspace

The vector of true planes is a solution to (4) and hence in the null space of $\mathbf{A}$, which we denote $Null(\mathbf{A})$. However, $Null(\mathbf{A})$ also contains *trivial solutions*, which place all curves in the same arbitrary plane. Observe that (2) has a *basic trivial subspace* of solutions spanned by

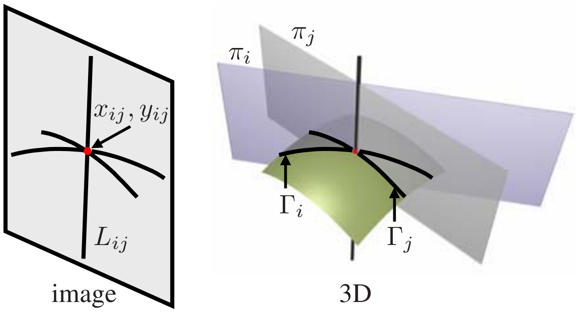

Figure 2. Intersecting planar curves in 3D and their projection.

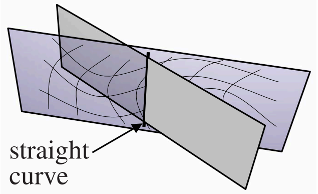

Figure 3. A trivial solution can embed all curves into the same geometric plane while assigning the straight line a different plane.

$$\mathbf{v}_1 = \frac{(1_N, 0_{2N})^T}{\sqrt{N}}, \mathbf{v}_2 = \frac{(0_N, 1_N, 0_N)^T}{\sqrt{N}}, \mathbf{v}_3 = \frac{(0_{2N}, 1_N)^T}{\sqrt{N}}, \tag{5}$$

where we use the notation c_k to denote a k-vector whose entries are all c. For any non-flat solution $\mathbf{v}$ to (4) there is a 4D subspace of ambiguous solutions, since any linear combination of $\mathbf{v}$,$\mathbf{v}_1$,$\mathbf{v}_2$,$\mathbf{v}_3$ is also a solution to (4). This is known as Generalized Bas-Relief (GBR) ambiguity under orthographic projection and perspective GBR (GPBR [8]) under perspective projection. The GBR ambiguity was already known to Sugihara, although the term GBR became common later [1]. Interestingly, in human perception studies of pictorial relief, Koenderink *et al.* [7] found that variations in 3D perception of the same picture by different observers can be attributed to the GBR ambiguity.

It might be though that $\mathbf{v}_1$,$\mathbf{v}_2$,$\mathbf{v}_3$ do not span the full algebraic subspace of flat solutions. This happens when there is a curve whose intersection points are along a straight line (in particular a curve with less than three intersections), as illustrated in figure 3 (assuming the planes are not edge-on implies that a straight curve in the image is straight in 3D). Note that Zero Gaussian Curvature (ZGC) surfaces [20] are a special case of surfaces that contain lines.

To characterize the complete space of flat solutions we need a measure of surface flatness. A natural approach is to pick a set of points on the surface, fit a plane with linear regression, and measure the residual error. The error will be zero whenever all points are coplanar. Although we don't know the depths at the points, we can express these depths as a linear function of $\mathbf{v}$. Let (x_i, y_i) be a set of k representative points on the curves (*e.g.* the intersection points).

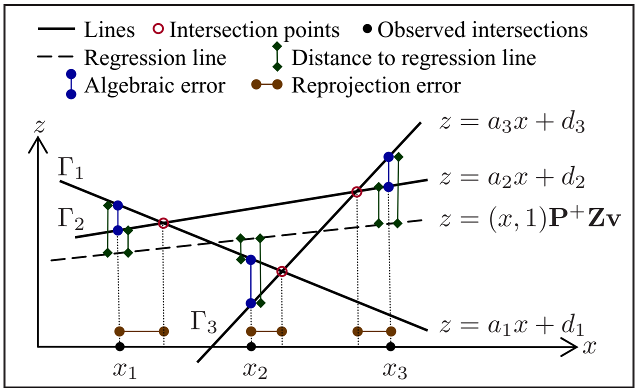

Figure 4. System of three intersecting lines $(\Gamma_1, \Gamma_2, \Gamma_3)$ in 2D. The image is formed by projection on the x-axis. Here

$$\mathbf{Z} = \begin{bmatrix} x_1 & 0 & 0 & 1 & 0 & 0 \\ 0 & x_1 & 0 & 0 & 1 & 0 \\ x_2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & x_2 & 0 & 0 & 1 \\ 0 & 0 & x_3 & 0 & 0 & 1 \\ 0 & x_3 & 0 & 0 & 1 & 0 \end{bmatrix}, \ \mathbf{v} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ d_1 \\ d_2 \\ d_3 \end{bmatrix}, \ \mathbf{P} = \begin{bmatrix} x_1 & 1 \\ x_1 & 1 \\ x_2 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_3 & 1 \end{bmatrix}.$$

Let $\mathbf{Z}$ be a matrix that reverse-projects the points onto their planes. The rows of $\mathbf{Z}$ contain $x_i, y_i, 1$ at the appropriate columns, so that $\mathbf{Zv} = (z_1, \ldots, z_k)^T$, where $(z_1, \ldots, z_k)^T$ are the depths of the planes at the points (x_i, y_i). The geometry is illustrated in figure 4, where for simplicity it shows the intersections of lines in 2D instead of planes in 3D. Define the matrix $\mathbf{C}$ by

$$\mathbf{P} = \begin{bmatrix} x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots \\ x_k & y_k & 1 \end{bmatrix}, \quad \begin{aligned} \mathbf{C} &= \left(\mathbf{I} - \mathbf{PP}^+\right)\mathbf{Z}/\sqrt{k} \\ &= \left(\mathbf{Z} - \mathbf{P}\left(\mathbf{P}^+\mathbf{Z}\right)\right)/\sqrt{k} \end{aligned}, \tag{6}$$

where $\mathbf{P}^+$ is the pseudo-inverse of $\mathbf{P}$. Then $\|\mathbf{Cv}\|^2$ is the averaged squared norm of the deviation in the z direction of the points (x_i, y_i, z_i) from the best fitting plane obtained by plane-fitting linear regression.

We define the *trivial subspace* to be the *null space of* $\mathbf{C}$. All vectors in the trivial subspace correspond to coplanar points (x_i, y_i, z_i), including cases with straight curves as in figure 3. There are three degrees of freedom in choosing this plane, but the dimension of $Null(\mathbf{C})$ could be larger.

The trivial subspace has degrees of freedom which could be set with additional information (*e.g.* depth of some known points). In absence of additional information, under orthography one may get a convincing qualitative shape by picking a solution which is orthogonal to the trivial subspace (under perspective this might place observed points behind the camera). This choice is natural since while the planar solutions are common to any set of curves, the non-trivial component is specific to the observed curves. Related ideas appear in [2, 12]. Statistical support to this choice is provided in appendix A.2.

2.2. Dimension of non-trivial solutions

Ideally our system would have a 4D solution space: three dimensions are due to the GBR ambiguity and the fourth comes from the true shape. One may wonder whether systems with more independent non-trivial solutions exist. The dimension depends on the observed pattern of intersections. For instance, disjoint components of curves leave some degrees of freedom. If the image is of a triangular mesh, we can place all its vertices freely in depth. Rothwell and Stern [12] demonstrated polyhedra with more than four degrees of freedom but relatively few intersections. However, for highly-connected networks of curves the system will typically be over-determined. For example, in a structured light scenario we may have $O(N^2)$ equations in $3N$ unknowns. Although extremely rare, highly-connected systems with $O(N^2)$ intersections and a 2D non-trivial subspace do exist, as shown in figure 5. We explore this question further in appendix A.3.

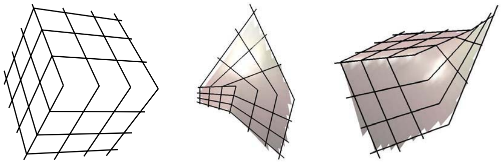

Figure 5. Multiple interpretations. The left drawing can be interpreted as a cube or as having doubly-ruled faces.

3. Proposed linear method

We now turn into solving the system in practice. Sugihara already noticed what he called the superstrictness problem: the linear system is typically over-determined and has no exact non-trivial solution in presence of noise. Various approaches have been taken to address this problem. Sugihara [18] explored finding a maximal generically reconstructible subset of equations. This approach does not spread the error evenly [11]. Shimshoni and Ponce [15] developed a linear programming method for a specified level of uncertainty in the vertices' positions. Ros and Thomas [11] proposed a nonlinear optimization algorithm based on a special ordering of vertices and faces (resolvable sequence). Grimstead and Martin [6, 22] used an iterative reweighting least-squares method. Sturm and Maybank [17] used SVD to solve for the displacements of the planes assuming the normals of the planes are known. Our approach extends the method of Bouguet *et al.* [2], which will be discussed in section 4.1.

In the presence of noise the trivial (planar) solutions are still exact, but in general there are no additional exact solutions. Geometrically this means there may be depth gaps between the planes at (x_{ij}, y_{ij}). A standard approach is to minimize the norm of the residuals of (4), *i.e.* looking for a vector $\mathbf{v}$ such that $\|\mathbf{Av}\|$ is small. However, when the points are nearly planar $\|\mathbf{A}\mathbf{v}\|$ is small for any set of curves. Thus, we want at the same time to keep the points (x_i, y_i, z_i) away from a common plane. The condition $\|\mathbf{C}\mathbf{v}\| = 1$ holds the points away from their best fitting plane. Recall that $\mathit{Null}(\mathbf{C})$ is made of the planar solutions and hence contained in $\mathit{Null}(\mathbf{A})$, even when $\mathbf{A}$ is noisy. We may assume for now that $\mathbf{v}$ is orthogonal to $\mathit{Null}(\mathbf{C})$, since any component of $\mathbf{v}$ in the trivial subspace does not affect $\|\mathbf{A}\mathbf{v}\|$. The trivial component can be added later if additional information is provided. We reformulate (4) as

$$\operatorname*{argmin}_{\mathbf{v}} \|\mathbf{A}\mathbf{v}\| \text{ s.t. } \|\mathbf{C}\mathbf{v}\| = 1 \,,\ \mathbf{v} \perp \mathit{Null}(\mathbf{C}) \,. \quad (7)$$

To solve (7), let $\mathbf{C} = \mathbf{U}\tilde{\mathbf{D}}\tilde{\mathbf{V}}^T$ be the SVD of $\mathbf{C}$. Define $\mathbf{V}$ and $\mathbf{D}$ by removing the columns of $\tilde{\mathbf{V}}$ and rows of $\tilde{\mathbf{D}}$ that correspond to singular values smaller than ε. The removed columns of $\tilde{\mathbf{V}}$ span the trivial subspace, which is at least 3D. The columns of $\mathbf{V}$ form a basis for the orthogonal complement of the trivial subspace. Writing $\mathbf{v} = \mathbf{V}\mathbf{D}^{-1}\mathbf{w}$, and using the fact that $\mathbf{U}$ is orthogonal, the problem becomes

$$\operatorname*{argmin}_{\mathbf{w}} \|\mathbf{A}\mathbf{V}\mathbf{D}^{-1}\mathbf{w}\| \text{ s.t. } \|\mathbf{U}\mathbf{D}\mathbf{V}^T\mathbf{v}\| = \|\mathbf{w}\| = 1 \,. \quad (8)$$

The algorithm is summarized below:

```
Step 1:  Form the matrices $\mathbf{A}, \mathbf{Z}, \mathbf{P}, \mathbf{C}$.
Step 2:  Compute the SVD of $\mathbf{C}$. Form $\mathbf{V}$ and $\mathbf{D}$.
Step 3:  Compute the SVD of $\mathbf{A}\mathbf{V}\mathbf{D}^{-1}$. $\mathbf{w}$ is the last
         right singular vector. Return $\mathbf{v} = \mathbf{V}\mathbf{D}^{-1}\mathbf{w}$.
```

$\mathbf{w}$ essentially picks a solution in the $\mathbf{V}$ basis. The matrix $\mathbf{D}^{-1}$ weights the columns of $\mathbf{V}$ so that vectors closer to a flat solution get higher cost.

4. Results

Demonstrations of the method to uncalibrated structured light are shown in figures 1 and 6. The curves were extracted from a video sequence of projected laser strips controlled by hand. Curves whose intersection points were nearly linear were pruned. These surfaces were computed using the orthographic model. Surface interpolation between the curves was done by Matlab, which interpolates over a triangulation.

We also applied the method for single-view modeling [17]. The user creates a 3D model from a photograph by drawing line segments and linking them to planar faces. Results are shown in figures 7 and 8. Though many faces had to be defined to fix all degrees of freedom in these examples, we used nothing but grouping of segments to planes to infer the 3D shape.

4.1. Comparison

Several previous works have been attempting to do uncalibrated structured light. Chen, Gao and Chen [4] assume the projected planes form an orthogonal grid. Caspi and Werman [3] assume the planes belong to two pencils. In the setting of Bouguet *et al.* [2] it was assumed that all planes intersect at a known point (the light source) so each plane had only two degrees of freedom. Our formulation does not restrict the planes and thereby allows scanning any visible part of the surface.

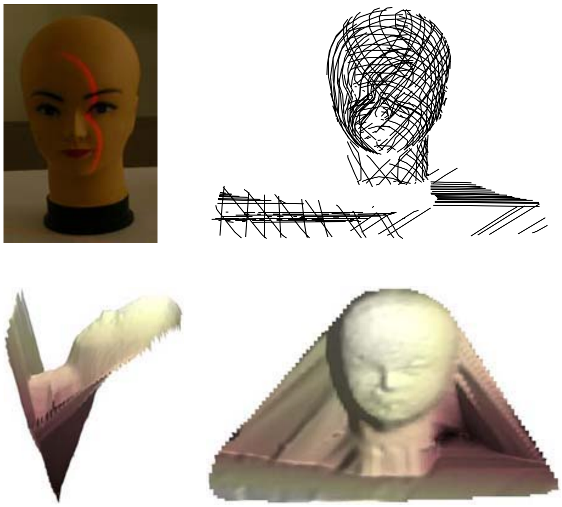

Figure 6. Surface reconstruction from 78 curves and 1053 intersection points. The slant is due to non-zero trivial component in the collection of true planes.

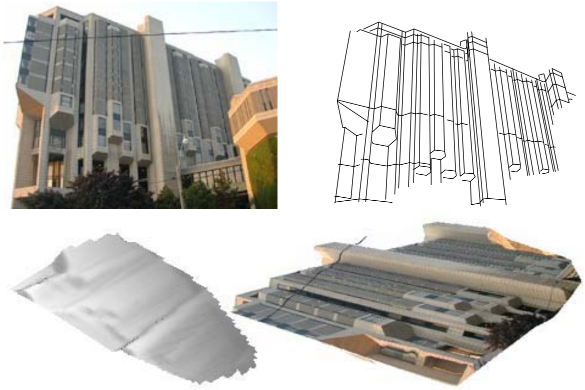

Figure 7. Top row: input image and manually drawn curves (54 planar faces). Bottom row: shaded and texture-mapped computed surface from different viewpoints.

In this subsection we derive the algorithm of Bouguet *et al.* [2] in a different way (their formulation makes the solution exactly orthogonal to the basic trivial subspace only when an exact solution exists), and compare it to ours. To avoid the zero solution ($\mathbf{v}_0 = 0_{3N}$) to (4), write

$$\operatorname*{argmin}_{\mathbf{v}} \|\mathbf{A}\mathbf{v}\| \text{ s.t. } \|\mathbf{v}\| = 1,\ \mathbf{v} \perp \mathit{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} \,. \quad (9)$$

Essentially one is looking for the fourth smallest right singular vector of $\mathbf{A}$. To enforce the solution to be orthogonal

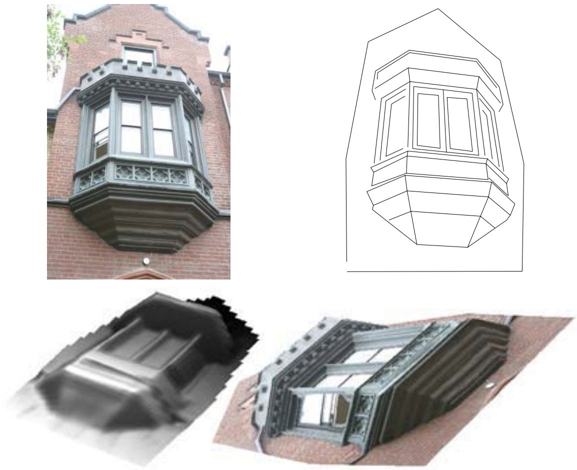

Figure 8. Image, curves, shaded and texture-mapped surface.

to the basic trivial subspace (both in the noisy and noiseless cases), write $\mathbf{v} = \mathbf{B}\mathbf{u}$, where $\mathbf{B}$ is a matrix whose columns form an orthonormal basis to the orthogonal complement of $\mathbf{v}_1$,$\mathbf{v}_2$,$\mathbf{v}_3$ defined by (5). A possible closed-form choice is

$$\mathbf{B} = \begin{bmatrix} \mathbf{E} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{E} \end{bmatrix}_{3N\times(3N-3)} , \ \alpha = \frac{\frac{1}{\sqrt{N}} - 1}{N-1} ,$$
$$\mathbf{E} = \begin{bmatrix} -\frac{1}{\sqrt{N}} & -\frac{1}{\sqrt{N}} & \cdots & -\frac{1}{\sqrt{N}} \\ 1+\alpha & \alpha & \cdots & \alpha \\ \alpha & 1+\alpha & \cdots & \alpha \\ \vdots & \vdots & \ddots & \vdots \\ \alpha & \alpha & \cdots & 1+\alpha \end{bmatrix}_{N\times(N-1)} \tag{10}$$

(a less symmetric choice for $\mathbf{E}$ is the Helmert matrix). Since $\mathbf{B}$ has orthonormal columns the problem becomes

$$\operatorname*{argmin}_{\mathbf{u}} \|\mathbf{A}\mathbf{B}\mathbf{u}\| \ \text{ s.t. } \|\mathbf{B}\mathbf{u}\| = \|\mathbf{u}\| = 1 \,. \tag{11}$$

Let $\mathbf{A}\mathbf{B} = \mathbf{U}\mathbf{D}\mathbf{V}^T$ be the SVD of $\mathbf{A}\mathbf{B}$. $\mathbf{u}$ is the last column of $\mathbf{V}$ and $\mathbf{v} = \mathbf{B}\mathbf{u}$ is the desired solution. It can be verified that except for the columns associated with the intersection points of the first curve, the other columns of $\mathbf{A}\mathbf{B}$ are sparse and contain the same elements as $\mathbf{A}$. Since only the last singular vector is used, there might be iterative methods that are faster than the computation of the SVD (we didn't experiment with numerical methods).

Comparing (8) to (11), the matrix $\mathbf{V}\mathbf{D}^{-1}$ is replaced with $\mathbf{B}$. The methods differ in two aspects. First, excluding only the basic trivial subspace yields a flat solution when a straight line is present. Secondly, omitting $\mathbf{D}^{-1}$ is prone to nearly-flat solutions. These solutions place all curves, except a small number, near a single plane. Almost all intersection points contribute nearly zero to the total error, and the error at the small number of intersection points of curves far from this plane might be relatively negligible. Note also that unlike the condition $\|\mathbf{C}\mathbf{v}\| = 1$, the condition $\|\mathbf{v}\| = 1$ has no geometric meaning. In fact, it mixes units of slope (a_i, b_i) and depth (d_i).

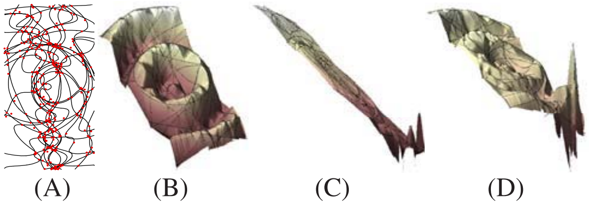

Figure 9. (A) 25 synthetic cross sections of a radial sine. The 351 intersection points are perturbed (red dots). (B) Our solution. (C) Last singular vector of the simple SVD method. This is a nearly-flat solution (note the bottom-right curve). (D) Second-last singular vector of the simple method. Although resembling the solution, it must be perpendicular to the last singular vector (C).

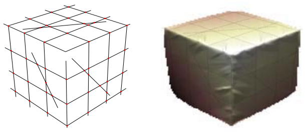

Figure 10. Left: adding four straight segments to figure 5 resolved the ambiguity. Right: computed surface by our method. The Last four singular vectors of the simple SVD method are flat, and the shape appears only at the fifth (note that our method thresholds by ϵ the singular vectors of $\mathbf{C}$, not $\mathbf{A}$).

Figure 9 compares both methods on synthetic randomly oriented cross sections of a radial sine. The positions of the intersection points were perturbed and shown by red dots. Both methods succeed solving this example without the perturbation. However, this perturbation was enough to break down the simple SVD method while our method succeeded. Figure 10 demonstrates our method when straight curves are present.

4.2. Examples with parallel curves

In previous sections we discussed systems of planar curves with many intersections. Next we demonstrate the method on examples with parallel cross sections where standard shape from texture algorithms would fail. In what may seem counterintuitive at first, Todd and Reichel [19] showed an example where humans perceive parallel cross sections better than randomly oriented cross sections with plenty of intersection points. Consider the topographic maps in figure 11. In this problem one has to set the order and distances (in z) between the planes. Notice the strong depth perception in absence of elevation values. Human perception doesn't rely completely on the density of the curves, since erasing a small number of curves will not change the perception as would be expected if they were perceived equally spaced. We are not particularly interested here in Euclidian reconstruction of topographic maps with fixed depth gaps between adjacent cross sections. Our aim is to demonstrate a 3D interpretation mechanism similar to the case with intersecting curves. The parallel case is obviously a harder under-constrained problem. Even for humans many topographic maps will not pop-out naturally.

To reduce the parallel case to the intersections case we use a simple shape prior, namely assume that the surface is roughly planar over local patches. We overlay a coarse rectangular grid over the image region, and group neighboring 2×2 cells into a set of W partially overlapping windows. The windows serve as virtual facets. We sample a set of points (x_i, y_i) along the curves. These points are treated as the image of the intersection points between the planar curves and the virtual planar facets. Although in general the planes will not intersect at the sampled points, we may still strive to minimize the distances. Rather than parameterizing the facets explicitly, for each window we measure the distances between the curves to the best fitting plane at the sampled points. This is done by constructing a matrix $\mathbf{C}_i$ as in (6) for each window (normalized by the number of points in each window). To enforce the curves to be parallel, we can either set $a_i = b_i = 0$ and solve for $\mathbf{v} = (d_1, \ldots, d_N)^T$, or alternatively add a penalty for the variance

$$\|\mathbf{R}\mathbf{v}\|^2 = \lambda\,(\,Var(a_i) + Var(b_i))\ , \tag{12}$$

$$\mathbf{R} = \sqrt{\frac{\lambda}{N}} \begin{bmatrix} \mathbf{I}_N - \frac{1}{N}\mathbf{1}_{N\times N} & \mathbf{0}_{N\times N} & \mathbf{0}_{N\times N} \\ \mathbf{0}_{N\times N} & \mathbf{I}_N - \frac{1}{N}\mathbf{1}_{N\times N} & \mathbf{0}_{N\times N} \end{bmatrix} .$$

The choice of λ depends on our confidence that the planes are parallel. The optimization problems becomes (semicolons denote vertical matrix concatenation):

$$\operatorname*{argmin}_{\mathbf{v}} \left\| \left[\mathbf{R}; \tfrac{1}{\sqrt{W}}\mathbf{C}_1; \cdots ; \tfrac{1}{\sqrt{W}}\mathbf{C}_W\right] \mathbf{v} \right\| \quad \text{s.t. } \|\mathbf{C}\mathbf{v}\| = 1 \ . \tag{13}$$

This problem is solved as in (8). The solution is up to GBR ambiguity. Examples are shown in figure 11.

5. Conclusions

We presented a unified analysis of shape from planar curves. We generalized the linear method of Bouguet *et al*. to deal with arbitrary planes, improved its robustness, and demonstrated its applicability to single view modeling and fully uncalibrated structured light. Then we showed examples with parallel curves where we assumed the curves intersect implicit planar facets. As the cases of polyhedra, SHGC's, ZGC's and uncalibrated structured light were treated independently in the literature, it is valuable that the same approach can deal with all of them.

A difficult issue we weren't dealing with is assigning curves to planar faces automatically, which is a perceptual grouping problem. One may try heuristics such as grouping parallel lines to a planar face, yet often multiple interpretations are possible. Similarly, the method proposed for dealing with parallel curves depends on the size of the windows that roughly approximate the surface. When they are too small the surface will be over-smoothed, as shown in figure 12.

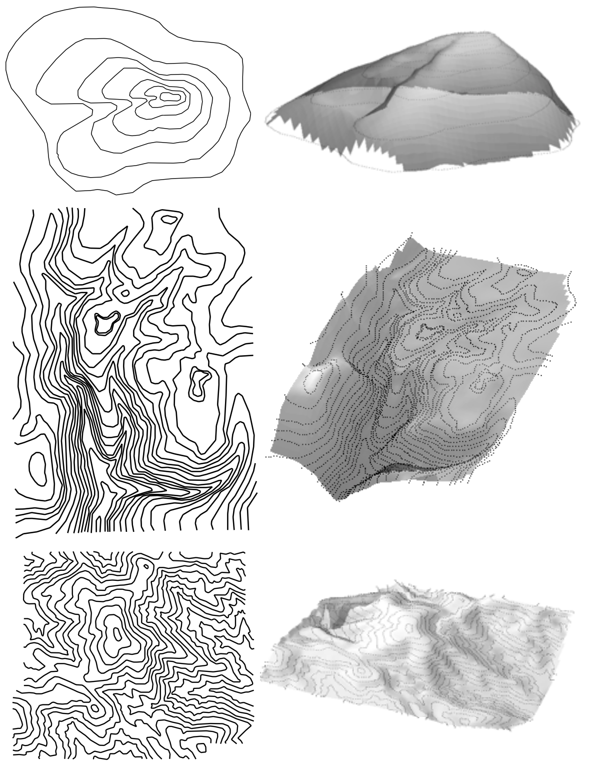

Figure 11. Reconstruction of topographic maps by piecewise planarity using a grid of 10×10 cells. Each window is made of 2×2 cells.

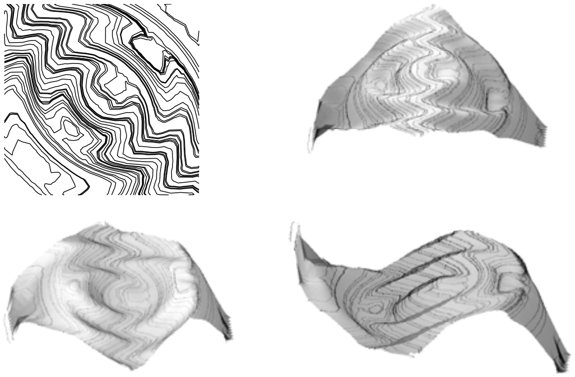

Figure 12. Clockwise: synthetic cross sections of a radial sine and last three singular vectors, solved by piecewise planarity on a grid of 10×10 cells. These solutions over-smooth the surface where a single plane fit (figure 13(C)) is appropriate for this shape.

A technical limitation of our formulation for parallel planes is that we cannot group all sampled points to the same plane, since minimizing $\|\mathbf{C}\mathbf{v}\|$ s.t. $\|\mathbf{C}\mathbf{v}\| = 1$ is meaningless. It is interesting to note that several unresolved examples in the literature could be solved by arrang-

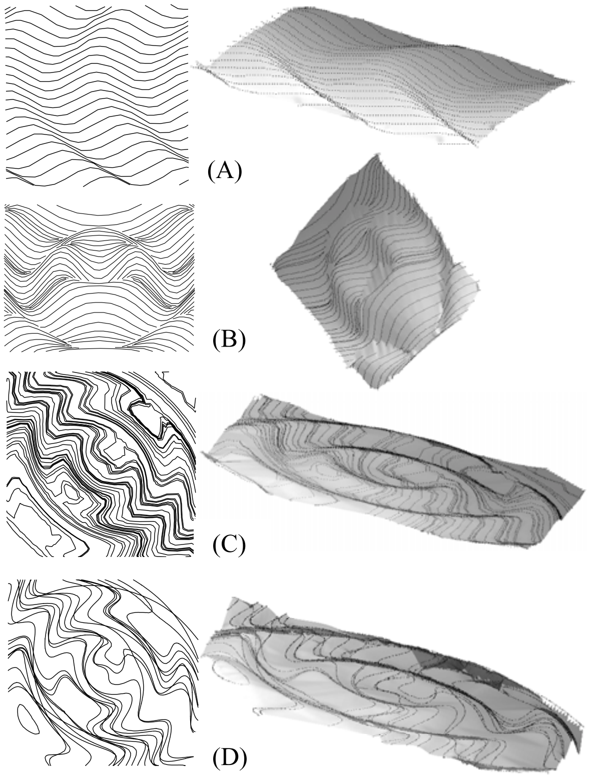

Figure 13. Shapes determined by a single plane fit. (A) Waves from Stevens [16]. (B) Radial sine from Todd and Reichel [19]. (C) Sine with non-uniform cuts (same as figure 12). (D) Sine with perturbed parallel cuts (intersection points were not used). $\lambda = 10^4$ for (A,B,C) and $\lambda = 10^3$ for (D).

ing the cross sections as close as possible to a single principle plane. Computationally this means finding the singular vector of the matrix $\left[\mathbf{R};\mathbf{C}\right]$ that corresponds to the smallest singular value greater than ε. Figure 13 demonstrates this singular vector on examples from Stevens [16], Todd and Reichel [19], and two synthetic examples with random gaps between parallel planes and nearly-parallel planes. While the case of intersecting planar curves is now well understood, the case of parallel cuts is more challenging and may require combination of additional depth cues.

It is possible to combine planarity with other depth cues to resolve the remaining degrees of freedom. For example, Lipson and Shpitalni [9, 22] maximized measures such as corner orthogonality and verticality of line segments. Another example is Surfaces Of Revolution (SOR [5]), which are SHGC's with two additional constraints. Shimshoni and Ponce [15] and Shimodaira [14] used also shading. These approaches require nonlinear optimization, potentially in high dimension. Note that Sugihara [18] already proposed to simplify the search by first solving the linear constraints and then searching for scale and trivial subspace coefficients (*e.g.* search in a 4D space) to optimize additional objectives. Our method can contribute in that direction as well.

A. Appendices

A.1. Formulation for perspective projection

Following standard conventions, the relation between image coordinates to world coordinates is $x = fX/Z$, $y = fY/Z$. A plane π_i which does not pass through the camera center is parameterized by

$$\begin{aligned} &a_iX + b_iY + c_iZ(X,Y) = \\ &(a_ix/f + b_iy/f + c_i)Z(X,Y) = 1\ , \end{aligned} \tag{14}$$

and the intersections between pairs of curves yield homogenous linear equations analogous to (2)

$$\begin{aligned} &1/Z_i(x_{ij}, y_{ij}) - 1/Z_j(x_{ij}, y_{ij}) = \\ &(a_i - a_j)x_{ij}/f + (b_i - b_j)y_{ij}/f + (c_i - c_j) = 0\ . \end{aligned} \tag{15}$$

In the perspective case we divide by f the entries x_{ij}, y_{ij}, x_i, y_i in the matrices $\mathbf{A}, \mathbf{Z}, \mathbf{P}$ of section 2, denoted as $\mathbf{A}_f, \mathbf{Z}_f, \mathbf{P}_f$, and minimize $\|\mathbf{A}_f\mathbf{v}_f\|$ s.t. $\|\mathbf{C}_f\mathbf{v}_f\| = 1$. Sugihara [18] observed that the orthographic system (2) has an exact non-trivial solution if and only if the perspective system (15) with the same intersection points has such a solution. Similarly, the singular values of our linear method in perspective are the same as the orthographic ones. By setting $\mathbf{v}_f = (fa_1, \ldots, fa_N, fb_1, \ldots, fb_N, d_1, \ldots, d_N)^T$ we get $\mathbf{A}_f\mathbf{v}_f = \mathbf{A}\mathbf{v}$, $\mathbf{Z}_f\mathbf{v}_f = \mathbf{Z}\mathbf{v}$, $\mathbf{P}_f\mathbf{P}_f^+ = \mathbf{P}\mathbf{P}^+$, and hence $\mathbf{C}_f\mathbf{v}_f = \mathbf{C}\mathbf{v}$. Note that in general, minimizing $\|\mathbf{A}_f\mathbf{v}_f\|$ s.t. $\|\mathbf{v}_f\| = 1$ is not invariant to f.

In perspective, one should pick a solution so that all visible points are in front of the camera (cheirality constraints).

A.2. Distribution of the basic trivial component

Assume a set of random planes is viewed orthographically from a random direction, or simply that the planes have random uniform orientations. What is then the distribution of the basic trivial component? The relative magnitude of the basic trivial component is

$$\left\|\left(\mathbf{v}^T\mathbf{v}_1, \mathbf{v}^T\mathbf{v}_2, \mathbf{v}^T\mathbf{v}_3\right)\right\| / \left\|\mathbf{v}\right\|\ . \tag{16}$$

Since absolute depth is lost in orthographic view, we may set $\mathbf{v}^T\mathbf{v}_3 = \frac{1}{\sqrt{N}}\sum d_i = 0$. An upper bound on (16) is

$$\frac{1}{\sqrt{N}}\left\|\left(\sum a_i, \sum b_i\right)\right\| / \left\|(a_1, \ldots, a_N, b_1, \ldots, b_N)\right\|\ . \tag{17}$$

When the planes have uniform orientations, the directions of the unscaled normals are $(n_i^x, n_i^y, -n_i^z)$, where $n_i^x, n_i^y, n_i^z \sim N(0,1)$. In the form of (1), $a_i \sim n_i^x/n_i^z$, $b_i \sim n_i^y/n_i^z$. The coefficients of the planes are Cauchy distributed. The sum of Cauchy variables is also distributed Cauchy, has a mode at zero but its mean is undefined (the CLT doesn't apply). The empiric histogram of (17) is shown in figure 14. We conclude that, under the assumption of uniform normals, it is probable that the relative magnitudes of $\mathbf{v}^T\mathbf{v}_1$ and $\mathbf{v}^T\mathbf{v}_2$ are small (figure 9 is an example).

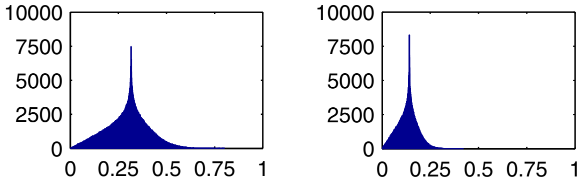

Figure 14. Histograms of (17) for random planes (10^6 trials). Left: $N = 10$ planes, $mean = 0.297$, $std = 0.108$. Right: $N = 50$, $mean = 0.132$, $std = 0.05$.

A.3. Ambiguous systems

The linear system (4) has a special structure: on each row we have $x_{ij}, -x_{ij}, y_{ij}, -y_{ij}, 1, -1$. We show how to construct systems with $O(N^2)$ intersections and a 2D non-trivial subspace artificially. Let $\{\pi_i^1\}$,$\{\pi_i^2\}$ be any two sets of N planes in general position, and $\{L_{ij}^1\}$,$\{L_{ij}^2\}$ the lines defined by (3). Set (x_{ij}, y_{ij}) to be the intersection point of L_{ij}^1 and L_{ij}^2. By construction, the two sets of planes $\{\pi_i^1\}$,$\{\pi_i^2\}$ are solutions to the system defined by this choice of (x_{ij}, y_{ij}). Having a 3D non-trivial subspace is much more accidental, since for each intersection the lines L_{ij}^1,L_{ij}^2,L_{ij}^3 must intersect at a point (or two lines coincide).

References

[1] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille. The bas-relief ambiguity. *IJCV*, 35(1):33–44, 1999.

[2] J.-Y. Bouguet, M. Weber, and P. Perona. What do planar shadows tell us about scene geometry? In *Proc. CVPR'99*, pages 514–520, 1999.

[3] Y. Caspi and M. Werman. Vertical parallax from moving shadows. In *Proc. CVPR'06*, pages 2309–2315, 2006.

[4] D. Chen, W. Gao, and X. Chen. A new approach of recovering 3-D shape from structure-lighting. In *Proc. ICSP'96*, pages 839–842, 1996.

[5] C. Colombo, A. D. Bimbo, and F. Pernici. Metric 3D reconstruction and texture acquisition of surfaces of revolution from a single uncalibrated view. *PAMI*, 27(1):99–114, 2005.

[6] I. J. Grimstead and R. R. Martin. Creating solid models from single 2D sketches. In *Proc. 3rd ACM Symposium on Solid Modeling and Applications*, pages 323–337, 1995.

[7] J. J. Koenderink, A. J. van Doorn, A. M. L. Kappers, and J. T. Todd. Ambiguity and the 'mental eye' in pictorial relief. *Perception*, 30(4):431–448, 2001.

[8] D. J. Kriegman and P. N. Belhumeur. What shadows reveal about object structure. *JOSA A*, 18(8):1804–1813, 2001.

[9] H. Lipson and M. Shpitalni. Optimization-based reconstruction of a 3D object from a single freehand line drawing. *Computer-Aided Design*, 28(8):651–663, 1996.

[10] Z. Pizlo, Y. Li, and G. Francis. A new look at binocular stereopsis. *Vision Research*, 45(17):2244–2255, 2005.

[11] L. Ros and F. Thomas. Overcoming superstrictness in line drawing interpretation. *PAMI*, 24(4):456–466, 2002.

[12] C. Rothwell and J. Stern. Understanding the shape properties of trihedral polyhedra. Technical Report 2661, INRIA, 1995.

[13] C. A. Rothwell, D. A. Forsyth, A. Zisserman, and J. Mundy. Extracting projective structure from single perspective views of 3D point sets. In *Proc. ICCV'93*, pages 573–582, 1993.

[14] H. Shimodaira. A shape-from-shading method of polyhedral objects using prior information. *PAMI*, 28(4):612–624, 2006.

[15] I. Shimshoni and J. Ponce. Recovering the shape of polyhedra using line-drawing analysis and complex reflectance models. *CVIU*, 65(2):296–310, 1997.

[16] K. A. Stevens. The visual interpretation of surface contours. *AI*, 17(1–3):47–73, 1981.

[17] P. F. Sturm and S. J. Maybank. A method for interactive 3D reconstruction of piecewise planar objects from single images. In *Proc. BMVC'06*, pages 265–274, 1999.

[18] K. Sugihara. *Machine Interpretation of Line Drawings*. MIT Press, 1986.

[19] J. T. Todd and F. D. Reichel. Visual perception of smoothly curved surfaces from double-projected contour patterns. *J. of Experimental Psychology: Human Perception and Performance*, 16(3):665–674, 1990.

[20] F. Ulupinar and R. Nevatia. Perception of 3-D surfaces from 2-D contours. *PAMI*, 15(1):3–18, 1993.

[21] F. Ulupinar and R. Nevatia. Shape from contour: Straight homogeneous generalized cylinders and constant cross section generalized cylinders. *PAMI*, 17(2):120–135, 1995.

[22] P. A. C. Varley and R. R. Martin. Estimating depth from line drawings. In *Proc. 7th ACM Symposium on Solid Modeling and Applications*, pages 180–191, 2002.