Bas-Relief Ambiguity Reduction in Shape from Shadowgrams

Jamil Draréni, Sébastien Roy
DIRO, Universite de Montreal , CP 6128 succ Centre-Ville, Montreal QC, Canada.
drarenij@iro.umontreal.ca, roys@iro.umontreal.ca

Peter Sturm
INRIA Rhone-Alpes , 655 Avenue de lEurope, 38330 Montbonnot St Martin, France.
Peter.Sturm@inrialpes.fr

Abstract

Coplanar shadowgrams provide an affordable mean to retrieve the 3d shape of an object especially when classical stereopsis fails (eg:textureless objects). Its principles are similar to the concepts used for Shape-From-Silhouettes with the only exception that here, light sources and cameras are interchanged. However, it is well known that any attempt to use the shadowgram to retrieve light sources positions is subject to a 4-parameter ambiguity. In this paper, we show how using the light spot visible in the camera reduces this ambiguity to a single parameter. We also suggest some practical solutions to gain a supplemental constraint on light sources positions and break the ambiguity. We demonstrate the effectiveness of our method using synthetic and real images.

1. Introduction

Many cues have been used in computer vision in order to infer and understand the 3d shape of objects and visual scenes in general. The resulting methods are quoted as *Shape-From-X*, where *X* refers to the main cue used for the reconstruction process. Stereoscopic disparity, apparent contour, shading, motion and shadows are some examples.

Shape from shadows received a great attention in the computer vision community, this is not surprising because valuable information on objects can be revealed from their cast shadows along with the corresponding light source. An other appealing aspect of the techniques based on shadows is that they do not rely on correspondences or a matching process like with classical stereopsis.

Early work on using shadows for structure retrieval dates back to Shafer and Kanade [8] who first established constraints on the orientations of the surfaces of interest in terms of observed cast shadows. Multiple light sources framework was proposed by [6] where a directional light moves in arc around the object of interest to acquire a sequence of planar cast shadows. The planar cast shadows are referred to as *planar shadowgrams*.

Later, Daum and Dudek [3] extended this framework to non single arc light trajectories. In [10], Yu *et al*. proposed a graph-based method to represent the constraints on the surfaces from light orientations. However, this method works on terrain-like surfaces lit by a directional source.

Following the success of space carving [4], Savarese *et al*. in the same vein proposed a shadow carving approach [7] where both silhouettes and shadows are used to infer the shape of an object. This is done by first building an occupancy volume from the silhouettes and carving away the regions that are not consistent with the observed shadows.

Recently, Shuntaro *et al*. [9] proposed a theory of shape from coplanar shadowgrams using a moving light source with no constraints on the trajectory of the light source. Once the positions of the light sources recovered, the convex hull of the object [1] is computed from the shadowgram using the classical shape from silhouette approach [5]. However, as proven by the authors themselves, when the shape of the object is unknown, the location of all the points sources can be recovered from the coplanar shadowgrams, only up to a four parameter perspective transformation related to the Generalized Perspective Bas-Relief ambiguity [2].

To break this ambiguity and to infer the location of the light sources, Shuntaro *et al*. use the shadowgrams of two additional spheres placed adjacent to the object of interest.

In this paper we propose a Shape-From-Silhouette (SFS) method based on shadowgrams. The proposed method exploits the projection of the light sources in the camera, visible through the translucent shadowgrams screen. We will show how the four parameter ambiguity described in [9] drops to a single parameter ambiguity by using the images of the light spots readily available in the image. In addition to this theoretical result, the proposed method en-

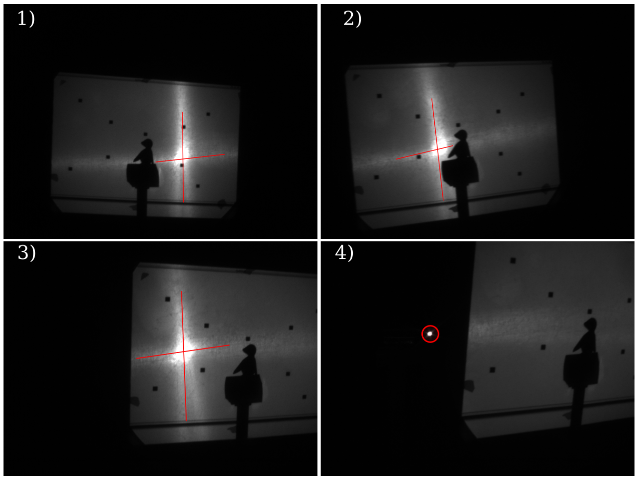

Figure 1. The projection of the light source is seen as a white spot through the screen. As the camera moves to the left (1,2,3), so the spot until the light source is directly visible (4).

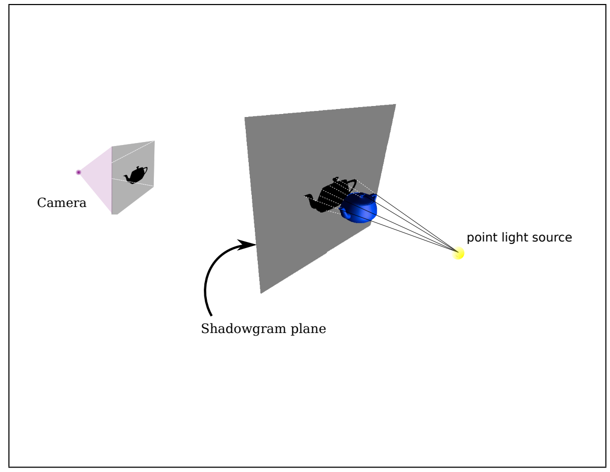

Figure 2. Setup to implement SFS. A point light source lit an object that in turn cast a shadow on a screen. A camera , placed on the other side of the screen, captures the shadowgram.

joys some interesting practical aspect by not relying on additional calibration objects, as their shadows may interact with the shadow of the object of interest. In deed, distinguishing a spot from a shadowgram silhouette can be done by a simple image thresholding even under sever camera distortions. We will also suggest some simple procedures that can be done at the acquisition time to constrain the remaining ambiguity and to fully solve the problem. We should point out that our method does not require the visibility of every spot corresponding to the light source we wish to estimate. Only two spots are needed and the rest are estimated using epipolar geomtry.

The rest of the paper is organized as follow. We outline the concept of epipolar geometry of shadowgrams in Section 2. Our main framework based on light triplets is presented in Section 3. We propose two simple methods to break the final ambiguity in Section 4. Experiments and conclusion are the subject of sections 5 and 6 respectively.

2. Shadowgrams and Epipolar Geometry

We assume the world's coordinate system attached to the shadowgram plane $\mathbf{\Pi}$ and the latter is located at $Z = 0$.

We represent the location of a light source $\mathbf{L}_i \in \mathbb{RP}^3$ in the global coordinate system by $\mathbf{L}_i = (X_i, Y_i, Z_i, 1)^\mathsf{T}$ and it's projection in the camera by $\mathbf{l}_i = (u_i, v_i, 1)$. The shadowgrams acquired by the camera are related to the shadowgrams on $\mathbf{\Pi}$ by a homography that remains fixed and can be estimated independently using standard computer vision techniques. Thus, we may assume the camera aligned with $\mathbf{\Pi}$ and express the projection of the light source $\mathbf{L}_i$ as:

$$\mathbf{l}_i \sim [\mathtt{I}_{3\times3}|-\mathbf{t}]\,\mathbf{L}_i$$

Where $\mathbf{t}$ is the position of the camera's center of projection.

When two light sources $\mathbf{L}_i$ and $\mathbf{L}_j$ are considered, an epipolar geometry is defined akin to the classical binocular stereo configuration. In this case, the light sources are analogous to the centers of projection of the cameras and the line that joins them is the baseline. The intersection of the baseline with $\mathbf{\Pi}$ is the epipole $\mathbf{e}_{ij}$. The homogeneous coordinates of $\mathbf{e}_{ij}$ can be expressed in terms of $\mathbf{L}_{i,j}$ coordinates as:

$$\mathbf{e}_{ij} = \begin{pmatrix} X_j Z_i - X_i Z_j \\ Y_j Z_i - Y_i Z_j \\ 0 \\ Z_i - Z_j \end{pmatrix} \tag{1}$$

This result stems form the intersection of the line joining $\mathbf{L}_{i,j}$ and $\mathbf{\Pi}$. It is worth noting that here, we only have one epipole as opposed to the classic stereo setup.

The same epipole can be estimated using the observed shadowgrams casted by an object lit by $\mathbf{L}_i$ and $\mathbf{L}_j$. In fact, the epipole is estimated from the intersection of two bitangent lines to the shadowgrams as depicted in Fig.2.

In case the light sources are located at the same distance from $\mathbf{\Pi}$, the epipole is located at infinity, and so the bitagent intersection.

Using only the information from the shadowgrams, any attempt to extract light positions is subject to a 4-parameters Bas-Relief ambiguity as shown in [9]. In the next section, we will show that considering the relation between three light sources reduces the ambiguity up to a single parameter.

3. Three Light Source Relation

Because three light sources are always coplanar and for the sake of simplicity, let us consider light sources $\widetilde{\mathbf{L}}_\mathbf{i}$ =

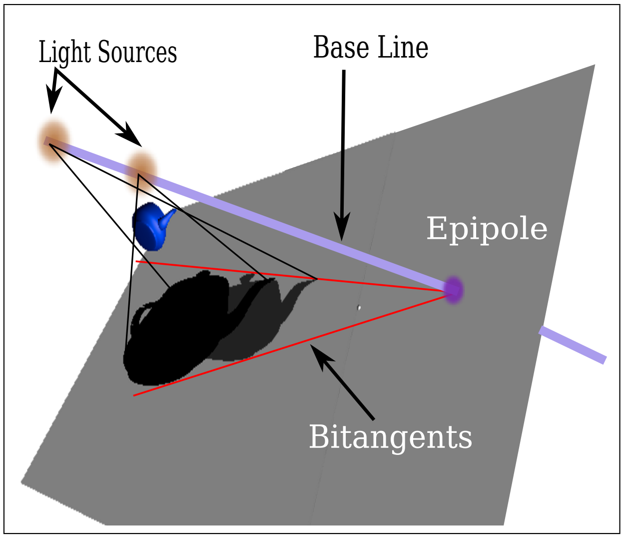

Figure 3. Shadowgrams from different light sources. The white spot is the projection of the spot.

$(0, Y_i, Z_i, 1)^\mathsf{T}$ that falls in the YZ-plane. From (1), the resulting epipoles $\tilde{\mathbf{e}}_{\mathbf{ij}}$ read off:

$$\tilde{\mathbf{e}}_{\mathbf{ij}} = \begin{pmatrix} 0 \\ Y_j Z_i - Y_i Z_j \\ 0 \\ Z_i - Z_j \end{pmatrix} \tag{2}$$

It's easy to see that the three $\tilde{\mathbf{e}}_{\mathbf{ij}}$ are colinear on $\mathbf{\Pi}$ along an epipolar line $\phi(Y)$ defined as:

$$\phi(Y) : Y = Y_j Z_i - Y_i Z_j / (Z_i - Z_j)$$

Let $\mathbf{\Gamma}$ represents a 3D plane that passes by $\phi(Y)$ and whose normal makes an angle of α with the XZ-plane:

$$\mathbf{\Gamma} \sim \begin{pmatrix} \cos\alpha \\ 0 \\ \sin\alpha \\ 0 \end{pmatrix}$$

The observed spots from the camera image constrain the sought light source $\mathbf{Q}_i$ on a line:

$$\mathbf{Q}_i(\lambda) = \begin{pmatrix} \mathbf{t} \\ 1 \end{pmatrix} + \lambda \left[\begin{pmatrix} \tilde{\mathbf{l}}_{\mathbf{i}} \\ 1 \end{pmatrix} - \begin{pmatrix} \mathbf{t} \\ 1 \end{pmatrix} \right]$$

The light source Q_i must project into $\tilde{\mathbf{l}}_{\mathbf{i}}$ and also lies on the plane $\mathbf{\Gamma}$, which leads to the following constraint in terms of λ:

$$\mathbf{\Gamma}^\mathsf{T} \mathbf{Q}_i(\lambda) = 0 \tag{3}$$

Which yield the following value of λ:

$$\lambda = \frac{(\mathbf{t}_1 \cos\alpha + \mathbf{t}_3 \sin\alpha)}{(\mathbf{t}_1 \cos\alpha + (\mathbf{t}_3 - \mathbf{Z}_i)\sin\alpha)}$$

The line that joins two light sources $\mathbf{Q}_i$ and Q_j is expressed as $\gamma\mathbf{Q}_i + \rho\mathbf{Q}_j$ and the coordinate of the related epipole $\mathbf{e}'_{ij}$ must satisfy:

$$\mathbf{e}'_{ij} \sim \gamma\mathbf{Q}_i + \rho\mathbf{Q}_j = \begin{pmatrix} 0 \\ \dots \\ 0 \\ \dots \end{pmatrix} \tag{4}$$

By Zeroing the first and the third component, we force the epipole to be the projection of a point on the YZ-plane according to the assumptions we made on the light sources locations. In a general context, one must fix the appropriate constraints to ensure that the epipoles lie into the observed epipolar line.

Further, to satisfy the equation (4), the scalars γ and ρ read off:

$$\begin{pmatrix} \gamma \\ \rho \end{pmatrix} \sim \begin{pmatrix} Z_j \\ -Z_i \end{pmatrix}$$

Substituting the values of γ and ρ in (4) gives:

$$\begin{aligned} \mathbf{e}'_{ij} &\sim \begin{pmatrix} 0 \\ (\mathbf{Y}_i\mathbf{Z}_j - \mathbf{Y}_j\mathbf{Z}_i)(\mathbf{t}_3\sin\alpha + \mathbf{t}_1\cos\alpha) \\ 0 \\ (\mathbf{Z}_j - \mathbf{Z}_i)(\mathbf{t}_3\sin\alpha + \mathbf{t}_1\cos\alpha) \end{pmatrix} \\ &\sim \begin{pmatrix} 0 \\ \mathbf{Y}_i\mathbf{Z}_j - \mathbf{Y}_j\mathbf{Z}_i \\ 0 \\ \mathbf{Z}_j - \mathbf{Z}_i \end{pmatrix} \\ &\sim \mathbf{e}_{ij} \end{aligned} \tag{5}$$

Thus, for each angle α we can compute an epipole e'_{ij} that verifies consistent with the epipolar geometry and the observed spots.

This ambiguity is materialized by a single parameter, namely the swiveling angle α as illustrated in Fig.4.

4. Solving the Ambiguity

In the previous section we showed that when the light spots are identified in the shadowgram images, the relationship between three light sources via their mutual epipolar geometry constrain the light positions up to a 1 parameter ambiguity. In order to break this ambiguity, an extra "information" on the light sources configuration is mandatory. Bare in mind, like it will be shown, that we do not need to observe every spot to infer the locations of all light sources; two spots are sufficient.

We propose two simple procedures that a user can perform while moving the light source in order to estimate the location of three light sources and alleviate the ambiguity of the whole system.

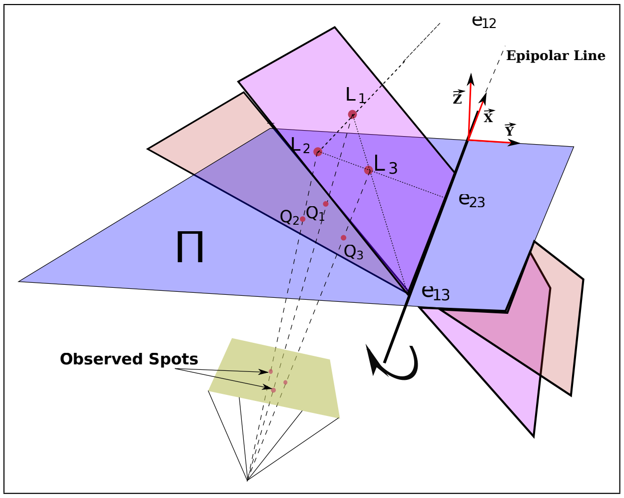

Figure 4. The 1-parameter ambiguity. When the deprojection plane swivel around the epipolar line, new light sources Q_1, Q_2 and Q_3 can be infered with the same properties as the real one.

Moving the Camera. After acquiring the desired sequence, a user can move the camera while the light source remain static. It is then possible to triangulate the last light position since the camera has a direct observation of the spot. Here, moving the camera can be avoided if one can afford using a second camera. Let us denote the reconstructed light source by L_1, we can see from Fig.3 that a light source L_3 can be reconstructed by intersecting the line $L_1 - e_{13}$ and the line that joins e_{23} and the spot of the light L_2. The remaining lights can be reconstructed the same way we did for L_3.

Baseline Normal to the Screen. If the user identifies a pair of light with a baseline orthogonal to the screen (pure front/back motion w.r.t the screen), an additional constraint on light locations is gained. In deed, a plane defined by these light and any other light is fully constrained.

5. Experiments

In this section we present our experiments involving synthetic and real images. Here we asses the presented method using the first solution to alleviate the ambiguity ; by moving the camera we fully reconstruct a light source and deduce the others using the epipolar geometry.

5.1. Synthetic Tests

Our synthetic test consist of 10 lights positioned randomly. We computed the 3D location of one light source by triangulating its spots observed in 2 virtual cameras. Along with another spot, we estimated the location of the remaining light sources using epipolar geometry. The sensitivity of the process was measured by adding zero mean gaussian noise to the virtual spots. The result of this experiment with different noise level is depicted in Fig.5 .

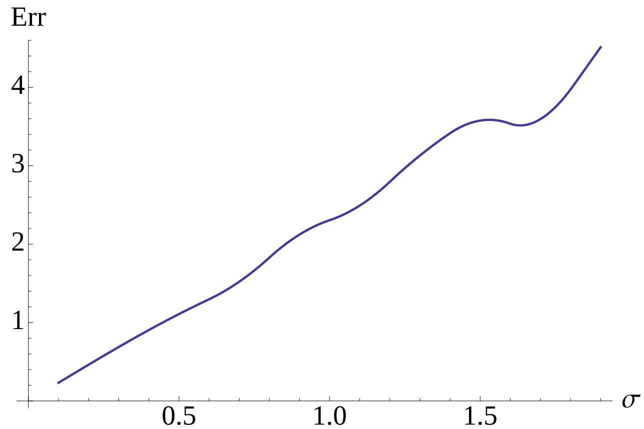

Figure 5. Sensitivity of the method in terms of noise level.

The reported error is the average distance of the reconstructed lights with the original one. We can see that the error grows linearly in terms of noise level.

5.2. Real Tests

We tested our method on real shadowgrams of a penguin figurine. The figurine was lit by a moving projector and the resulting shadows were recorded by a video camera through a screen. The resolution of the camera is 640×480 pixels. At the end of the sequence the camera underwent a motion while the projector remained fixed. Hence, the 3D position of the last projector could be reconstructed by triangulating the observed spots using conventional structure from motion techniques.

Once this done, the remaining light sources were estimated one by one using the 3-light sources geometry presented in this paper combined with the epipolar geometry. The triplets consist of the fully reconstructed light source (the last one), a light source with a visible spot and the light source of interest.

The resulting 3D model using 10 shadowgrams is depicted in Fig.6 .

6. Conclusion

In this paper we presented a method to reconstruct objects from their shadowgrams. As opposed to previous works, the presented method does not use a calibration rig or an object that may interfere with the object of interest. Instead, we exploited the images of the light sources visible in the camera as bright spots easily identified. We also showed that using these spots alone can only reduce the ambiguity

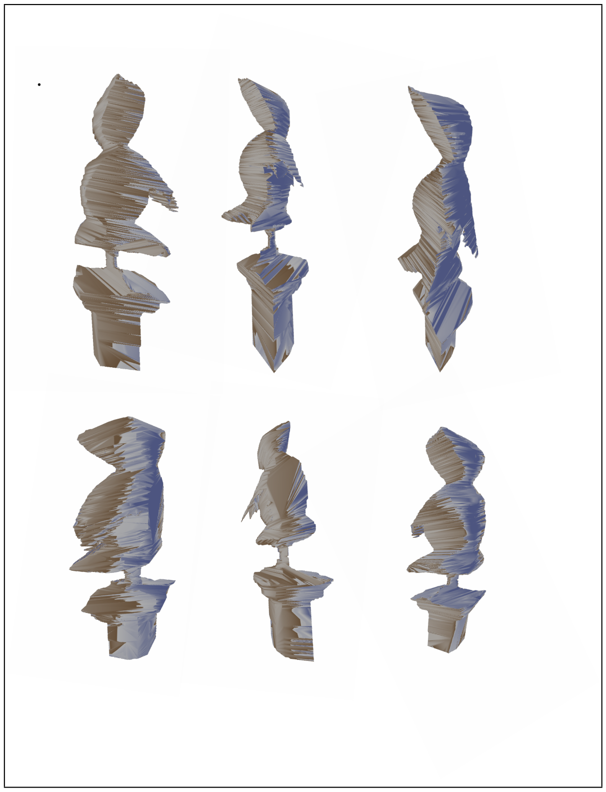

Figure 6. Snapshots of the reconstructed penguin using silhouettes from Fig.1

from 4 to 1 parameter. The later can be solved using simple procedures during the acquisition to yield a 3D model.

References

[1] B. Baumgard. *Geometric Modeling for Computer Vision*. PhD thesis, University of Stanford, 1974. 1

[2] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille. The bas-relief ambiguity. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 0:1060, 1997. 2

[3] M. Daum and G. Dudek. On 3-d surface reconstruction using shape from shadows. In *CVPR '98: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, page 461, Washington, DC, USA, 1998. IEEE Computer Society. 1

[4] K. N. Kutulakos and S. M. Seitz. A theory of shape by space carving. *Int. J. Comput. Vision*, 38(3):199–218, 2000. 1

[5] A. Laurentini. The visual hull concept for silhouette-based image understanding. *IEEE Trans. Pattern Anal. Mach. Intell.*, 16(2):150–162, 1994. 1

[6] D. Raviv, Y. Pao, and K. Loparo. Reconstruction of three-dimensional surfaces from two-dimensional binary images. 5(5):701–710, May 1989. 1

[7] S. Savarese, M. Andreetto, H. Rushmeier, F. Bernardini, and P. Perona. 3d reconstruction by shadow carving: Theory and practical evaluation. *Int. J. Comput. Vision*, 71(3):305–336, 2007. 1

[8] S. Shafer and T. Kanade. *Using Shadows in Finding Surface Orientations*, 22:145–176, 1983. 1

[9] S. Yamazaki, S. G. Narasimhan, S. Baker, and T. Kanade. Coplanar shadowgrams for acquiring visual hulls of intricate objects. In *Proc. International Conference of Computer Vision*, pages 1–8, 2007. 1, 2, 3

[10] Y. Yu and J. T. Chang. Shadow graphs and 3d texture reconstruction. *Int. J. Comput. Vision*, 62(1-2):35–60, 2005. 1