Mirror Symmetry $\Rightarrow$ 2-View Stereo Geometry

Alexandre R.J. François[+], Gérard G. Medioni[+] and Roman Waupotitsch[*]

[+]Institute for Robotics and Intelligent Systems
University of Southern California,
Los Angeles, CA 90089-0273
{afrancoi,medioni}@usc.edu

[*]Geometrix Inc.
1590 the Alameda #200
San Jose, CA 95126
romanw@geometrix.com

Abstract

We address the problem of 3-D reconstruction from a single perspective view of a mirror symmetric scene. We establish the fundamental result that it is geometrically equivalent to observing the scene with two cameras, the cameras being symmetrical with respect to the unknown 3-D symmetry plane. All traditional tools of classical 2-view stereo can then be applied, and the concepts of fundamental/essential matrix, epipolar geometry, rectification and disparity hold. However, the problems are greatly simplified here, as the rectification process and the computation of epipolar geometry can be easily performed from the original view only. If the camera is calibrated, we show how to synthesize the symmetric image generated by the same physical camera. An Euclidean reconstruction of the scene can then be computed from the resulting stereo pair. To validate this novel formulation, we have processed many real images, and show examples of 3-D reconstruction.

Keywords: Mirror symmetry; Stereo; 3-D modeling.

1 Introduction

Symmetry is a rich source of information in images. *Mirror symmetry* (also called *bilateral symmetry*) is a property shared by many natural and man-made objects. A 3-D object exhibits mirror symmetry if there exists a plane that separates the object into two (identical) chiral parts. The plane is called *symmetry plane*. For each point of such an object, there exists an identical symmetric point that also belongs to the object. A line joining two symmetric points is called a *symmetry line*. All symmetry lines are perpendicular to the symmetry plane, which by construction contains all symmetry pair midpoints.

Methods to exploit symmetry and infer constraints on the 3D reconstructed shape of the object/scene have a long history in Computer Vision. Based on the assumption that "a skew symmetry depicts a real symmetry viewed from some unknown angle" (Kanade [6]), a large number of approaches rely on the analysis of skewed symmetries in images to infer constraints on the 3-D geometry of depicted objects. In most cases, (scaled) orthographic projection is assumed (see e.g. [8][3]). Recently, François and Medioni proposed an interactive bilateral-symmetric object reconstruction from a single view under orthographic projection [2]. A more limited number of studies deal with symmetry observed under perspective projection. Some, such as Ulupinar and Nevatia [13], and Mitsumoto *et al.* [7] use the vanishing point of symmetry lines, whose computation is unstable unless the perspective distortion is severe. Tan's reconstruction method [12] is instead based on a trapezium reconstruction algorithm and variants aimed at increasing the robustness of the method. The point ranges are recovered up to a scale factor. Different studies make use of the epipolar geometry arising from the bilateral symmetry of a scene in a single perspective view. For example, Rothwell *et al.* [10] use it to recover the projective 3-D structure of a set of points, in the context of object recognition. Zhang and Tsui [14] use a planar mirror to allow (partial) reconstruction of arbitrary objects. Shimshoni *et al.* [11] use a hybrid photo-geometric approach to perform projective reconstruction of faces. Most recently, Huynh [5] proposed an affine reconstruction algorithm from a monocular view in the presence of symmetry plane.

We establish more general epipolar projective properties of single views of mirror symmetric scenes, and apply them in a different context: we perform Euclidean 3-D reconstruction of mirror symmetric objects from a single perspective view *using traditional 2-view stereo geometry*. This is allowed by the *Mirror-Stereo Theorem*.

Theorem: *One perspective view of a mirror symmetric scene, taken with an arbitrary projective camera, is geometrically equivalent to two perspective views of the scene from two projective cameras symmetrical with respect to the unknown 3-D symmetry plane.*

We show how to extract the epipolar geometry relating the two views from the single input image. Classical 2-view stereo results can then be applied (see e.g. [1] or [4]), and the concepts of fundamental/essential matrix, epipolar geometry, rectification and disparity hold. If the camera is calibrated, we show how to synthesize the image generated by the original camera placed symmetrically, in order to be able to use traditional 2-view stereo tools to obtain an *Euclidean reconstruction* of the scene.

This paper is organized as follows. In section 2, we establish a lemma introducing a virtual camera, symmetric to the camera that produced the input image with respect to the object's symmetry plane. In section 3, we establish a lemma linking the symmetry lines and epipolar lines in images of a symmetric scene taken by cameras in symmetrical locations. The proof for the mirror-stereo theorem is given in section 4.

In section 5, we show, in the case of a calibrated pinhole camera, how to build a virtual view taken by the same camera placed symmetrically in location and orientation. Examples of Euclidean 3-D reconstructions obtained from such pairs with an off-the-shelf stereo package are presented. A summary of the contribution concludes the paper in section 6.

Throughout the paper, the approach and notation used are consistent with that of [4].

2 Virtual view

Lemma 1: *The image of a scene that is symmetric with respect to an unknown plane, formed by an arbitrary projective camera, is identical to the image of the scene formed by the (virtual) projective camera symmetric of the first one with respect to the scene's 3-D (unknown) symmetry plane.*

Without loss of generality, we place the origin Ω of the world Euclidean, orthogonal, right-handed coordinate system (Ω, x, y, z) on the symmetry plane, which we also define as the (Ω, y, z) plane. The setting is illustrated in figure 1.

Let X denote a world point represented by the homogeneous 4-vector $\begin{bmatrix} x & y & z & 1 \end{bmatrix}^T$. Let x denote the corresponding image point represented by a homogeneous 3-vector $\begin{bmatrix} U & V & W \end{bmatrix}^T$. A general projective camera is defined by a 3x4 projection matrix $P = M[I| - \tilde{C}]$, where M is a 3x3 matrix, and $\tilde{C}$ denotes the inhomogeneous 3-vector of the camera center coordinates in the world coordinate system.

X is mapped to x according to the relation:

$$x = M[I| - \tilde{C}]X$$

By construction, the world point symmetric to X with respect to the (Ω, y, z) plane is $\bar{X} = ZX$, where:

$$Z = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ (we note } \tilde{Z} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{)}$$

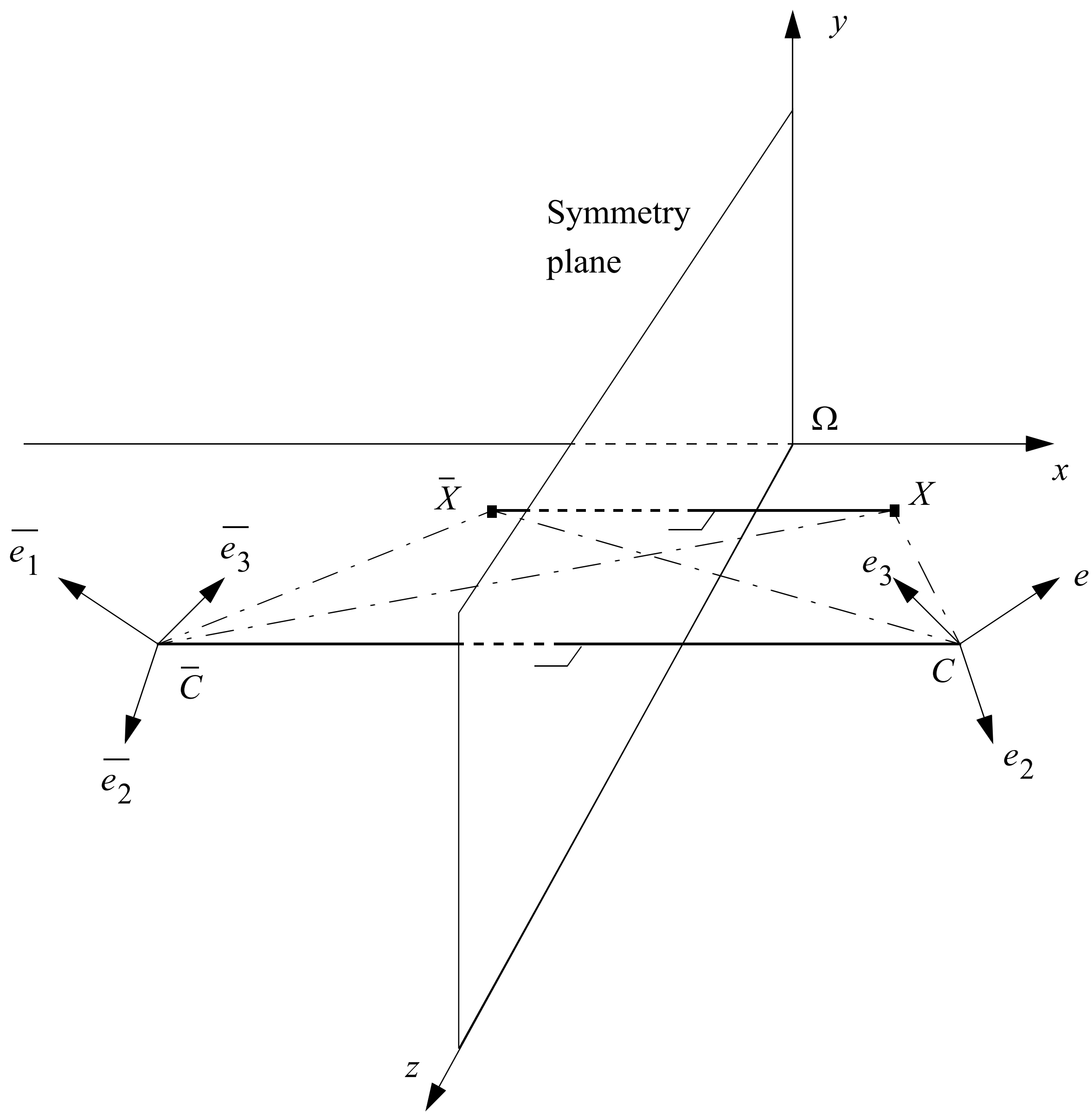


Fig. 1. Symmetric points seen from symmetric cameras

Thus the image point $\bar{x}$ of the world point $\bar{X}$ seen by camera C is:

$$\bar{x} = M[I| - \tilde{C}]\bar{X} = M[I| - \tilde{C}]ZX$$

Consider now the virtual camera $\bar{C}$, symmetrical of camera C with respect to the object's symmetry plane. By construction, its center is $\bar{C} = ZC$, and it projects a world point X into the image point x' such that:

$$x' = \bar{M}[I| - \tilde{\bar{C}}]X, \text{ where } \bar{M} = M\tilde{Z}$$

Replacing symmetric elements by their expression, we get:

$$x' = M\tilde{Z}[I| - \tilde{Z}\tilde{C}]X = M[I| - \tilde{C}]ZX = \bar{x}$$

Similarly:

$$\bar{x}' = M\tilde{Z}[I| - \tilde{Z}\tilde{C}]ZX = M[I| - \tilde{C}]ZZX = x$$

It follows that the image of a pair of symmetric points viewed by the real camera is identical to the image of the same symmetric pair viewed by the virtual symmetric camera, the actual image points of the symmetric points being reversed in the real and virtual view. Q.E.D.

3 Through the looking glass

Lemma 2: *In stereo pair of images of a mirror symmetric scene, taken by cameras whose centers are symmetrical with respect to the scene's symmetry plane, the epipolar lines are images of symmetry lines, and reversely symmetry lines project into epipolar lines.*

Symmetry lines are 3-D lines perpendicular to the symmetry plane. With the notations defined above (see figure 1), for any world point X, the lines $X\bar{X}$ and $C\bar{C}$ are parallel by construction. It follows that the epipolar plane containing a point X also contains the symmetric point $\bar{X}$, and consequently the corresponding symmetry line $X\bar{X}$. In other words, all symmetry lines lie in an epipolar plane. It results that the epipolar line defined by the epipolar plane in each image is also the projection of all symmetry lines contained in this plane. Reversely, any symmetry line projects into the epipolar line defined by the epipolar plane containing the symmetry line. Q.E.D.

4 Mirror-stereo theorem

We can now give a proof for the Mirror-Stereo Theorem, stated in the introduction. Per Lemma 2, the symmetry lines in a single perspective view of a mirror symmetric scene, taken by an arbitrary projective camera, are also the epipolar lines of the stereo pair consisting of the given image and the image formed by any projective camera located symmetrically, and thus completely characterize the epipolar geometry between such a view and the original view. Per Lemma 1, one such view, specifically the one taken by the virtual camera symmetrical to the real one with respect to the scene's symmetry plane, is actually the same image as the original. In these two (identical) images, the image points of a world point and its (world) symmetrical point are reversed. It results that the original image is geometrically equivalent to two views from two different positions, related by a computable epipolar geometry. Q.E.D.

Given the epipolar pencil and a set of point correspondences in the one real image, it is possible, using classical epipolar geometry results, to compute:

- the fundamental matrix relating the two views,
- the camera matrices (up to a projective transformation)
- for each point correspondence, the 3-D point that projects to those points (up to a projective transformation).

Note that additional constraints derived from the mirror symmetry property can be used to make those computations simpler than in the general case, as computations can be done in a single image.

5 Experimental validation

We developed our novel formulation to use existing 2-view stereo tools to perform 3-D reconstruction of mirror symmetric objects from a single, real perspective image. Such systems take as input two images, taken by the *same* physical camera from different positions, and sometimes require the images to be rectified. If the rectification poses no theoretical problem given the availability of epipolar pencils, the synthesis of a virtual symmetric view taken by a physical camera requires more work. We show how to build such an image in the case of a calibrated pinhole camera, in order to allow Euclidean reconstruction of the scene. We then present the reconstruction results obtained by processing image pairs made of one real perspective image of a mirror symmetric object, and one synthetic symmetric view, in an off-the-shelf commercial package, Eos Systems' PhotoModeler Pro 4.0 [9].

5.1 Calibrated pinhole camera case

Suppose the camera used to take the available picture is a pinhole camera. It projects the world point into the image point according to the formula:

$$x = KR[I|(-\tilde{C})]X$$

The 3x3 matrix K is the camera calibration matrix. If C is a pinhole camera, its calibration matrix is of the form:

$$K = \begin{bmatrix} \alpha_x & 0 & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}$$

where α_x and α_y represent the focal length of the camera in the x and y directions respectively, and (x_0, y_0) is the principal point, all expressed in terms of pixel dimensions.

R is the 3x3 transform matrix from the world coordinate system to the camera coordinate system, in this case a rotation, defined by:

$$R = \begin{bmatrix} e_1^T & e_2^T & e_3^T \end{bmatrix}^T$$

where (e_1, e_2, e_3) is the orthonormal basis of the Euclidean coordinate system associated with the camera.

Let (e_1', e_2', e_3') be the orthonormal basis of the Euclidean coordinate system associated with the virtual symmetrical camera. By construction, $e_j' = \overline{e_j} = \tilde{Z}e_j$. Because of the symmetry, this coordinate system's orientation is reversed compared to that of (e_1, e_2, e_3) (if one is right-handed, the other is left-handed).

The virtual camera projects the world point X into the image point x' according to:

$$x' = KR\tilde{Z}[I|-\tilde{\tilde{C}}]X = K\overline{R}[I|-\tilde{\tilde{C}}]X$$

where:

$$\bar{R} = R\tilde{Z} = \begin{bmatrix} \bar{e}_1^T & \bar{e}_2^T & \bar{e}_3^T \end{bmatrix}^T$$

The image point of the same world point seen by the physical camera placed in $\bar{C}$ and with principal axis direction $\bar{e}_3^T$ is:

$$x'' = KR''[I|-\tilde{\bar{C}}]X$$

where

$$R'' = \begin{bmatrix} e_1''^T & e_2''^T & \bar{e}_3^T \end{bmatrix}^T$$

and (e_1'', e_2'') is an orthonormal basis of the principal plane such that $(e_1'', e_2'', \bar{e}_3)$ has the same orientation as (e_1, e_2, e_3) (both are either right or left-handed). Such a basis is obtained by applying a rotation of an arbitrary angle θ to the basis $(\bar{e}_1, \bar{e}_2)$ in the plane it defines, and inverting the direction of one of the axes. The rotation of the coordinate system corresponds to a rotation of the physical camera around its optical axis.

Using the notation

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \bar{R}[I|-\tilde{\bar{C}}]X$$

x'' can be written:

$$x'' = \begin{bmatrix} \alpha_x & 0 & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Since we are looking to relate x'' and

$$x' = \begin{bmatrix} \alpha_x & 0 & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

the simplest analytic form naturally corresponds to no rotation, i.e. to the cases where the real camera coordinate system axes are aligned with that of the virtual (symmetric) ones, the direction of one of the axes being reversed to preserve the coordinate system's orientation. The two corresponding cases are shown in figure 2. For sake of clarity and simplicity, without loss of generality,

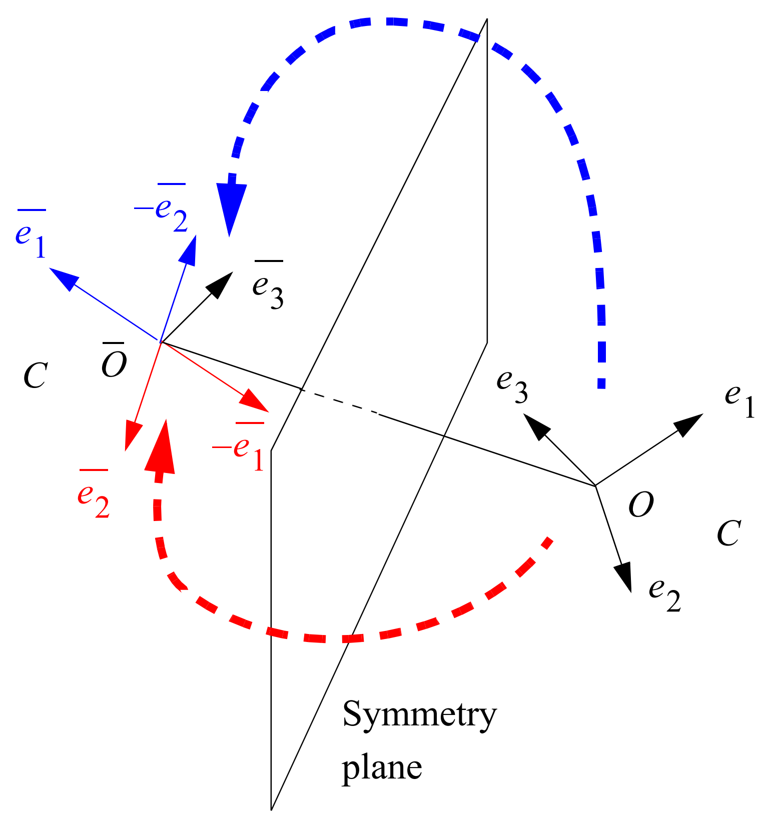


Fig. 2. How to relate the virtual view and the real camera's view from the virtual viewpoint.

we only consider an inversion of the x-axis.

Under these conditions, x' and x'' differ only by their first component U' and U'' respectively, and thus only the image points abscissas differ:

$$u' = \frac{a}{c}\alpha_x + x_0$$

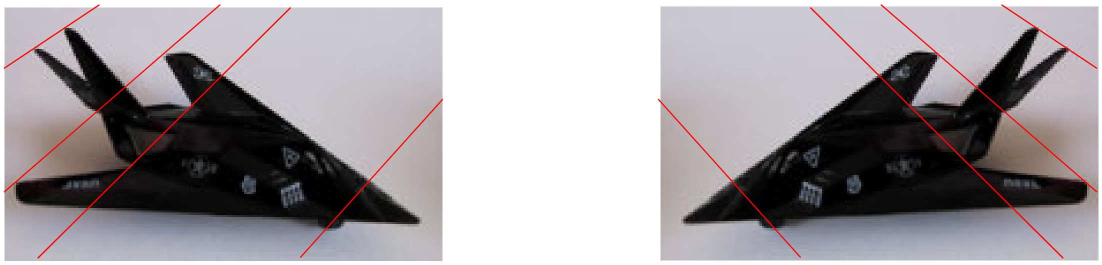
$$u'' = -\frac{a}{c}\alpha_x + x_0$$

It follows that if the camera is calibrated, then the symmetric view of the virtual stereo pair can be obtained by simply performing a horizontal mirror of the original image with respect to a vertical axis going through the principal point. Note that neither the camera-centered orthonormal bases nor the scene's symmetry plane need be explicitly computed in practice. The camera calibration process only concerns intrinsic camera parameters (principal point and focal length). Any deviations of the camera from the pinhole model (e.g. radial distortion) can be corrected following this initial calibration process, without invalidating the described reconstruction method. The relationship between the camera and the scene is *a priori* unknown (it will be later recovered by the stereo processing), and the only requirement for the above results to apply is for the camera center not to lie in the scene's symmetry plane (which is an obvious degenerate case).

The following reconstruction examples have been processed from image pairs made of a real image and the synthetic symmetric view produced according to the principle established above. Point correspondences were given by hand. If dense correspondences would allow a dense reconstruction, the usual correlation method used in 2-view stereo will generally not work since the two mirror views may correspond to viewing directions far apart.

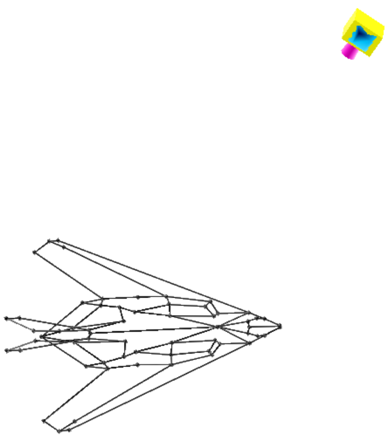
5.2 Stealth fighter model

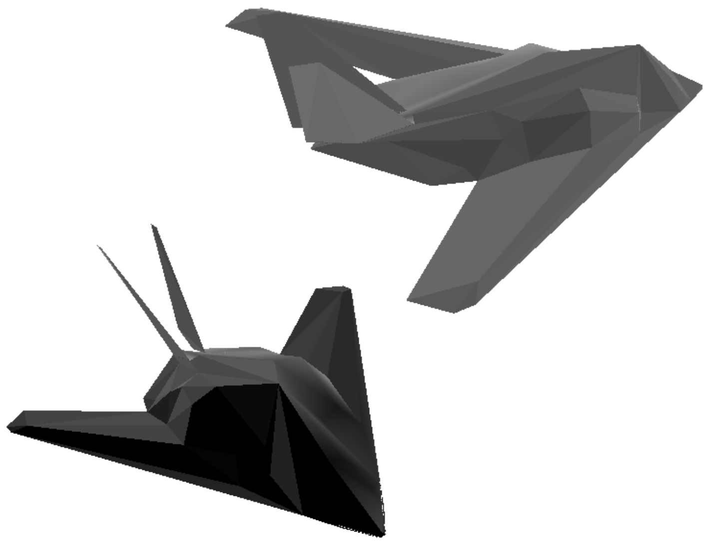
Our first example is a toy F-117 Stealth fighter. The input image and synthesized symmetric view are presented in figure 3(a) with their epipolar pencils. The internal parameters of the digital camera used were determined using Photomodeler's calibration tool. After specifying a minimum number of point correspondences between the two images, the software was able to perform a full calibration of both camera positions and to infer the 3-D position of the object points. Adding more point and line correspondences allowed to obtain a 3-D wireframe model of the object presented in figure 3(b). Note that the recovered object is mirror symmetric, and that the recovered cameras are actually symmetrical in location and orientation with respect to the object's symmetry plane, properties that are not used at any point in Photomodeler's reconstruction algorithms (the symmetry property is only used to synthesize the symmetric view). Once the 3-D inference is performed, the 3-D surface model shown in figure 3(c) is built and the texture can be extracted from the images. The final 3-D model with texture map is shown in figure 3(d). The construction of this 3-D textured model took only a few minutes to a novice Photomodeler user. Incorporating symmetry derived constraints in the workflow and in the interface would greatly facilitate the process and reduce the time needed to build a model, while also helping in improving the accuracy of the manual input and helping deal with self occlusions across the symmetry plane.



a. Input image (left) of a (toy) F117 and inverted image (right), used as a stereo pair. Epipolar pencils are overlaid on the images.

b. Wireframe model of the F117 and computed corresponding camera positions





c. 3-D surface model of the F117

d. Textured model of the F117

Fig. 3. F117 Stealth Fighter model.

5.3 Face model

The technique clearly applies beautifully to perfectly symmetric objects. We also wanted to try the more challenging problem posed by (quasi-)human faces. We used a bust of Darth Maul which is fairly symmetric. The input and synthetic symmetric views are presented in figure 4(a). The wireframe model and camera positions and orientations as recovered by Photomodeler are presented in figure 4(b). Finally, the resulting textured model is shown in figure 4(c). The polygonal model, although coarse, captures the geometry of the face fairly well, so that the resulting textured model is easily recognized as a face under various viewpoints.

6 Conclusion

We have proved the fundamental result that a single perspective view of a mirror symmetric scene, taken by an arbitrary projective camera, is geometrically equivalent to two perspective views of the scene with two cameras that are symmetric with respect to the unknown 3-D symmetry plane.

In the case of a calibrated pinhole camera, we have shown how to build a virtual stereo pair by synthesizing the view taken by the same physical camera located and oriented symmetrically. Such virtual stereo pairs allow to produce Euclidean reconstruction using traditional 2-view stereo tools, examples of which were presented.

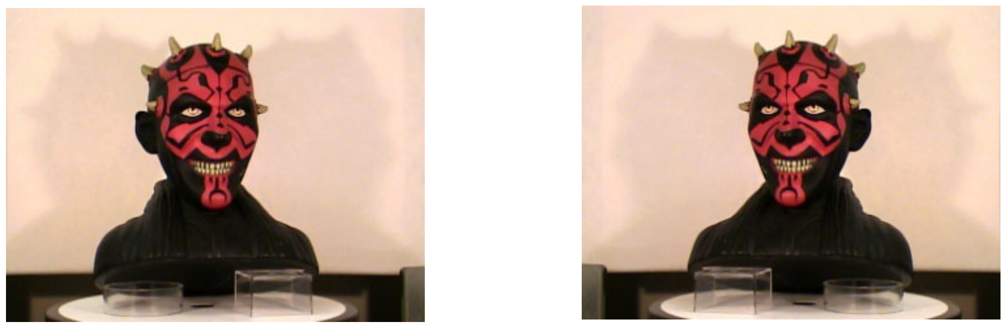
Additional constraints derived from the mirror symmetry property can be used to make the 2-view stereo computations simpler than in the general case. They can also be used to enhance reconstruction application workflow and interface to improve robustness and accuracy of manual interventions.

7 Acknowledgements

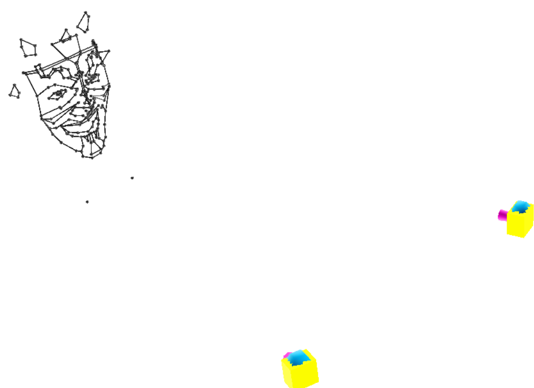
This research has been funded in part by the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, Cooperative Agreement No. EEC-9529152.

8 References

1. Faugeras O. D. *Three-Dimensional Computer Vision: a Geometric Viewpoint*. MIT Press, 1993.
2. François A. and Medioni G. Interactive 3-D Model Extraction From a Single Image. *Image*



a. Input image (left) of a Darth Maul bust and inverted image (right), used as a stereo pair



b. Wireframe model of the face and computed corresponding camera positions

c. Textured polygonal model of Darth Maul's face

Fig. 4. Darth Maul face model.

and Vision Computing, vol. 19, no. 6, April 2001, pp. 317-328.

3. Gross A.D. and Boult T. E. Analyzing Skewed Symmetries. Int. Jour. Computer Vision, vol. 13, no.1, September 1994, pp. 91-111.
4. Hartley R. and Zisserman A. *Multiple View Geometry*. Cambridge University Press, Cambridge, UK, 2000.
5. Huynh D. Q. Affine Reconstruction from Monocular Vision in the Presence of a Symmetry Plane. In *Proc. Int. Conf. Computer Vision*, vol. 1, pp. 476-482, Corfu, Greece, September 1999.
6. Kanade T. Recovery of the Three-Dimensional Shape of an Object from a Single View. *Artificial Intelligence*, vol. 17, 1981, pp. 409-460.
7. Mitsumoto H. *et al.* 3-D Reconstruction Using Mirror Images Based on a Plane Symmetry Recovering Method. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 14, no.9, September 1992, pp. 941-946
8. Nalwa V. S. Line-Drawing Interpretation: Bilateral Symmetry. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 11, no. 10, October 1989, pp. 1117-1120
9. Photomodeler Pro 4.0, Eos Systems. www.photomodeler.com
10. Rothwell C.A., Forsyth D.A., Zisserman A. and Mundy J.L. Extracting Projective Structure from Single Perspective Views of 3D Point Sets. In *Proc. Int. Conf. on Computer Vision*, Berlin, May 1993, pp. 573-582.
11. Shimshoni I., Moses Y. and Lindenbaum M. Shape Reconstruction of 3-D Bilateral Symmetric Surfaces. *International Journal of Computer Vision*, vol. 39, no. 2, September 2000, pp. 97-110.
12. Tan T. N. Monocular Reconstruction of 3-D Bilateral Symmetrical Objects. In *Proc. British Machine Vision Conf.*, Poster Session 1, September 1996.
13. Ulupinar F. and Nevatia R. Constraints for Interpretation of Line Drawings under Perspective Projection. *Computer Vision Graphics and Image Processing*, vol. 53, no. 1, January 1991, pp. 88-96.
14. Zhang Z.Y. and Tsui H.T. 3D Reconstruction from a Single View of an Object and Its Image in a Plane Mirror. In *Proc. Int. Conf. on Pattern Recognition*, Brisbane, Australia, August 1998, pp. 1174-1176.