SELF-CALIBRATED RECONSTRUCTION OF PARTIALLY VIEWED SYMMETRIC OBJECTS

Hassan Foroosh *Murat Balci* *Xiaochun Cao*

School of Computer Science, University of Central Florida
{balci, foroosh, xccao}@cs.ucf.edu

ABSTRACT

The traditional stereo reconstrction techniques based on point correspondences and the estimation of the cameras from the fundamental matrix introduce a four-fold ambiguity. Moreover, there is a projective ambiguity inherent in the fundamental matrix. We show that a symmetric object can be modeled even under partial occlusion with a pair of uncalibrated stereo images. This implies that unlike traditional stereo algorithms, we can extract 3D information from two arbitrary viewpoints, even when there is no left-to-right point correspondences. To demonstrate the effectiveness of the method, we present experimental results on both synthetic and real images.

1. INTRODUCTION

Three dimensional reconstruction recovers the geometry of the scene and the camera from multiple perspective views. Typically traditional methods for construction from two views [1] consist of three steps: compute the fundamental matrix from point correspondences, find the camera matrices, and then for each point correspondence compute the point in space that projects to these two image points using triangulation. Recently, Rother [2] has presented a new linear direct reference plane method for reconstructing simultaneously 3D features (points, lines and planes) and cameras from many perspective views by solving a single linear system. The method finds the linear projective relationship between lines or planes and cameras, and thus a minimum amount of image measurements is sufficient. However, it assumes that a reference plane is visible in all views, which might be difficult to obtain in practice. Starck and Hilton [3] presented a model-based framework for the reconstruction of shape and appearance (especially people) from multiple views. The technique generates improved results over many model-free approaches.

In this paper, we exploit the properties of mirror symmetry in order to self-calibrate and reconstruct the 3D model for partially viewed mirror symmetric objects. Symmetry, which is a property shared by many natural and man-made objects, is a rich source of information in images. Methods that exploit symmetry in order to impose constraints on the 3D structure of the scene have been occasionally explored in the past [4, 5]. Francois et al. [6] have presented a method for reconstructing mirror symmetric scenes from a single view by synthesizing a second camera based on the first one. However, they assume that their cameras are positioned in a restricted mirror symmetric setup. As a result the calibration is fixed by the restricted geometric configuration of the cameras, and hence is assumed known. We describe a more general framework, where unknown cameras can view a symmetric object from any arbitray position and orientation.

This work was partially supported by ONR grant #N000140410512 and Sun Microsystem's grant #EDUD-7824-030482-US.

2. THE MIRROR SYMMETRY CONSTRAINT

Symmetry is a ubiquitous property of many natural and man-made objects. Symmetric objects impose strong geometric constraints that can be exploited in recovery of motion and structure. Symmetry is present in objects in the form of central symmetry, axial symmetry (surface of revolution), or the most frequently encountered mirror symmetry, which we investigate in this paper. Examples, include a face, a car, a chair, etc. In mirror symmetry, object symmetry is defined with respect to a plane π in the 3D space, so that for any point $\mathbf{M} = [X \;\; Y \;\; Z \;\; 1]^T$ on the object there exist a point $\mathbf{M}' = [X' \;\; Y' \;\; Z' \;\; 1]^T$ on the object such that $\mathbf{MM}'$ is orthogonal to π and $d(\mathbf{M}, \pi) = d(\mathbf{M}', \pi)$, where $d(\cdot, \cdot)$ stands for the Euclidean distance.

Given two pairs of such symmetric points $(\mathbf{M}_1, \mathbf{M}'_1)$ and $(\mathbf{M}_2\mathbf{M}'_2)$, it can be readily verified that they are coplanar, e.g. in the plane π'. Suppose now that these four points are imaged with a perspective camera whose 3×4 projection matrix is given by $\mathbf{P}$. Define the world coordinate frame as follows: the x-axis along $\mathbf{M}_1\mathbf{M}'_1$, y-axis along the line of intersection of π and π', and z-axis according to the right-hand rule (see Figure 1). We can then show that in the 3D space the symmetric pairs of points are related via

$$\mathbf{M}'_i = \left(\mathbf{I}_4 - 2\frac{\mathbf{va}^T}{\mathbf{v}^T\mathbf{a}}\right)\mathbf{M}_i \quad i = 1, 2 \tag{1}$$

where $\mathbf{v} = [-1 \;\; 0 \;\; 0 \;\; 0]^T$, $\mathbf{a} = [1 \;\; 0 \;\; 0 \;\; 0]^T$, $\mathbf{I}_4$ is a 4×4 identity matrix, and T denotes the transpose.

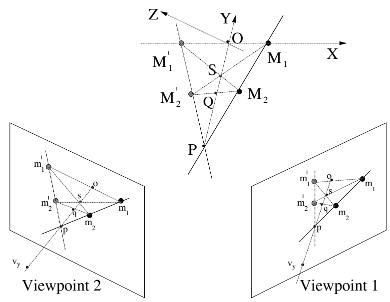


Fig. 1. Configuration of two pairs of symmetric points.

This equation equally applies to any symmetric pairs of points on the object. The corresponding image points are given by $\mathbf{m}_i = \mathbf{P}\mathbf{M}_i$ and $\mathbf{m}'_i = \mathbf{P}\mathbf{M}'_i$, with

$$\mathbf{P} = \mathbf{K}[\mathbf{r}_1 \;\; \mathbf{r}_2 \;\; \mathbf{r}_3 \;\; \mathbf{t}] \tag{2}$$

where $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ are the columns of the 3×3 rotation matrix $\mathbf{R}$, $\mathbf{t}$ is the translation vector, and $\mathbf{K}$ is the 3×3 upper triangular camera intrinsic matrix. Due to mirror symmetry, the two pairs of symmetric points are coplanar, and for our choice of the world coordinate frame thay are in the plane $Z = 0$. Therefore, the image coordinates are related to the world coordinates of the 3D points by a 3×3 homography, i.e. $\mathbf{m}_i = \mathbf{H}_w\bar{\mathbf{M}}_i$ and $\mathbf{m}'_i = \mathbf{H}_w\bar{\mathbf{M}}'_i$, where $\bar{\mathbf{M}}_i = [X_i \;\; Y_i \;\; 1]^T$, $\bar{\mathbf{M}}'_i = [X'_i \;\; Y'_i \;\; 1]^T$, and

$$\mathbf{H}_w = \mathbf{K}[\mathbf{r}_1 \;\; \mathbf{r}_2 \;\; \mathbf{t}] \tag{3}$$

Also, clearly the line $\mathbf{a}$ and the point $\mathbf{v}$ lie in the plane $Z = 0$. It therefore follows from (1) that the images of the symmetric points are related via

$$\mathbf{H}_w^{-1}\mathbf{m}'_i = \left(\mathbf{I}_3 - 2\frac{\bar{\mathbf{v}}\bar{\mathbf{a}}^T}{\bar{\mathbf{v}}^T\bar{\mathbf{a}}}\right)\mathbf{H}_w^{-1}\mathbf{m}_i \tag{4}$$

where $\bar{\mathbf{a}} = [1 \;\; 0 \;\; 0]^T$ and $\bar{\mathbf{v}} = [-1 \;\; 0 \;\; 0]^T$.
Since $\mathbf{a}$ and $\mathbf{v}$ are the world y-axis and the vanishing point along the x-axis, upon rearranging, we get

$$\begin{aligned} \mathbf{m}'_i &= \mathbf{H}_w\left(\mathbf{I}_3 - 2\frac{\bar{\mathbf{v}}\bar{\mathbf{a}}^T}{\bar{\mathbf{v}}^T\bar{\mathbf{a}}}\right)\mathbf{H}_w^{-1}\mathbf{m}_i & (5)\\ &= \left(\mathbf{I}_3 - 2\frac{\mathbf{v}_x\mathbf{l}^T}{\mathbf{v}_x^T\mathbf{l}}\right)\mathbf{m}_i & (6)\\ &= \mathbf{H}_h\mathbf{m}_i & (7) \end{aligned}$$

where $\mathbf{v}_x$ is the vanishing point along the x-axis, and $\mathbf{l}$ is a line in the image plane corresponding to the projection of the world y-axis.
This shows that under mirror symmetry, the projections of symmetric points into the image plane are related via a harmonic homology $\mathbf{H}_h$ [1]. Given the image projection of any two pairs of symmetric points, this harmonic homology can be computed. As a result, given the image projection of any 3D point that is coplanar with the original four points, the projection of its symmetric counterpart can be readily determined even if it is not visible due to for instance occlusion. This is the key idea that is exploited herein in order to build 3D Euclidean models of symmetric objects, even when they are partially visible.
Before showing this, we extend the idea to an even more general situation that can be encountered in practice. Consider the case, where $\mathbf{v}_x$ is known, but $\mathbf{l}$ is unknown. Clearly, the homology $\mathbf{H}_h$ can not then be computed directly from this minimal information. It turns out that we still can find the homology using the mirror symmetry constraint. For this purpose note that any line through the vanishing point $\mathbf{v}_x$ that interects the image of the symmetric object has to be perpendicular to its plane of symmetry. Let $\mathbf{l}_1$ be one such line, and $\mathbf{m} \neq \mathbf{v}_x$ be a point on this line and on the image of the object. Then it can be shown that the projection of the world origin into the image plane is given in terms of $\mathbf{l}$ by

$$\mathbf{o} = \mathbf{l} \times (\mathbf{v}_x \times \mathbf{m}) \tag{8}$$

where $\times$ denotes the cross product of two vectors.
By exploiting the geometry of the mirror symmetry, we can verify that the point $\mathbf{m}'$ that is symmetrically situated with respect to $\mathbf{m}$ on the line $\mathbf{l}_1$ satisfies the following equality

$$\{\mathbf{v}_x, \mathbf{o}; \mathbf{m}, \mathbf{m}'\} = 1 \tag{9}$$

where $\{\cdot, \cdot; \cdot, \cdot\}$ is the cross-ratio of four points.
By adopting a sign convention, equation (9) can be used once with the x-coordinates and once with the y-coordinates. Upon substituting for $\mathbf{m}'$ and $\mathbf{o}$ from (6) and (8) respectively, we get two cubic polynomial equations in terms of the two unknown parameters of $\mathbf{l}$. The ambiguity in the solution can be resolved by repeating the above steps for a second line. Once $\mathbf{l}$ is known, the homology can be computed as before.

3. STEREO RECONSTRUCTION

Consider the case where we have a stereo pair of a mirror symmetric object. Assume also that some parts of the object are not visible due to for instance occlusion (including self-occlusion), or the object not being centered within boundaries of the image frames. Based on the results of the previous section, we can still fully reconstruct the object from these stereo views, using the symmetric counterparts of the occluded regions. The only challenge that we face is that a Euclidean reconstruction would require the world to image homography $\mathbf{H}_w$ to be known. This is equivalent to saying that we need the camera to self-calibrate using the given stereo pair.
For this purpose note that for a unit aspect ratio and zero camera skew, $\mathbf{K}$ is of the form

$$\mathbf{K} = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{10}$$

where f is the camera focal length, and $\tilde{\mathbf{c}} = [\mathbf{c}^T \quad 1]^T = [u_0 \;\; v_0 \;\; 1]^T$ is the principal point.
On the other hand [7, 8]

$$\mathbf{H}_w = [r_{31}\tilde{\mathbf{v}}_x \quad r_{32}\tilde{\mathbf{v}}_y \quad t_z\mathbf{o}] \tag{11}$$

where $\mathbf{v}_y$ is the vanishing point along the y-axis.
Since $\tilde{\mathbf{v}}_x$ and $\tilde{\mathbf{v}}_y$ are the vanishing points along two orthogonal directions, we have

$$\tilde{\mathbf{v}}_y^T \omega \tilde{\mathbf{v}}_x = 0 \tag{12}$$

where $\tilde{\mathbf{v}}_x = [\mathbf{v}_x \quad 1]^T$, $\tilde{\mathbf{v}}_y = [\mathbf{v}_y \quad 1]^T$, and ω is the image of the absolute conic (IAC) [1], whose knowledge would yield the intrinsic calibration parameters using the Cholesky factorization of its dual.
Now, we know that $\mathbf{v}_y$ must lie on the image of the world y-axis, i.e. $\mathbf{v}_y^T\mathbf{l} = 0$. Therefore

$$\tilde{\mathbf{v}}_y = [\mathbf{l}]_\times \omega \tilde{\mathbf{v}}_x \tag{13}$$

where $[\cdot]_\times$ is the usual notation for the skew symmetric matrix characterizing the cross product. Geometrically, $\omega\mathbf{v}_x$ is the vanishing line of the pencil of planes that are parallel to the plane of symmetry.

Since t_z is a global scale factor, we can set it to one for one of the two images and compute it for the second one by forcing corresponding image points to be projected to the same 3D world point. The other two columnwise scale factors, i.e. r_{31} and r_{32} can be found in terms of the camera intrinsic parameters by using (11), (13) and the orthonormality property of the rotation matrix. As a result the world to image homography for both images can be expressed as a function of the intrinsic parameters. Therefore, in order to solve for the unknown intrinsic parameters, and hence all related extrinsic parameters, we formulate our problem in terms of the inter-image homography that minimizes the symmetric transfer error of geometric distances.

$$\begin{aligned}(f, \mathbf{c}) &= \arg\min_{\Gamma} \sum_i d(\mathbf{m}_i, \mathbf{H}_{f,\mathbf{c}}^{-1}\mathbf{m}_i'))^2 \\ &+ \quad d(\mathbf{m}_i', \mathbf{H}_{f,\mathbf{c}}\mathbf{m}_i)^2 \end{aligned} \tag{14}$$

where Γ is the 3D search space of the solution for f and (u_0, v_0), $\mathbf{H}_{f,\mathbf{c}} = \mathbf{H}'_w\mathbf{H}_w^{-1}$ is the inter-image homography (which is only a function of f and $\mathbf{c}$, and superscripts indicate the images in which the cross-ratios are taken.

In order to minimize this cost function, we take advantage of the fact that the principal points of recent CCD cameras are very close to the center of the image. Therefore, we first find an initial estimate for f from (14) by setting $\mathbf{c}$ as the center of the image. The search space for the intrinsic parameters $(f, \mathbf{c})$ that minimize the cost function in (14) is then narrowed down to a 3D window around this initial value of f and the image center. The solution is therefore found without resorting to non-linear minimization techniques, i.e. by sampling the solution space within the 3D search window, and performing an exhaustive search within this small window. Once the camera parameters are found, we can find the two world to image homographies, from which we can recover the 3D points by optimal triangulation as described in [1]. Again note that 3D reconstruction can be achieved for regions that are viewed by only one camera (e.g. due to occlusions or partial view), using the symmetric part of the object.

4. EXPERIMENTAL RESULTS

The proposed approach has been tested on an extensive set of simulated and real data. Due to lack of space only some are presented below.

4.1. Computer Simulation

Points	true X	Est. X	true Y	Est. Y	true Z	Est. Z
1^{st}	-85	-82.47	0	0.002	0	-0.127
2^{nd}	-125	-121.72	-150	-144.57	0	-0.037
3^{rd}	125	121.44	-150	-144.31	0	-0.455
4^{th}	85	82.13	0	0.133	0	-0.552
5^{th}	-85	-83.43	0	0.298	-100	-95.73
$6\&7^{th}$	±125	±124.37	-100	-96.30	-115.69	-109.70
8^{th}	85	81.98	0	0.577	-100	-95.93

Table 1. Estimated 3D coordinates at 1.5 pixel noise level

Noise	Avg.	Std. Dev.	Avg.	Std. Dev.	Avg.	Std. Dev.
0.2	1.60	0.78	1.48	1.77	1.80	1.88
0.4	1.44	0.51	1.24	1.40	1.43	1.60
0.6	1.76	0.37	1.37	1.60	1.46	1.76
0.8	2.14	0.66	2.51	2.82	2.86	2.92
1.0	3.30	0.83	2.95	3.48	3.10	3.14
1.2	1.97	0.92	1.94	2.33	2.88	2.55
1.5	2.70	0.65	2.87	2.83	2.58	2.77

Table 2. Performance vs noise (in pixels)

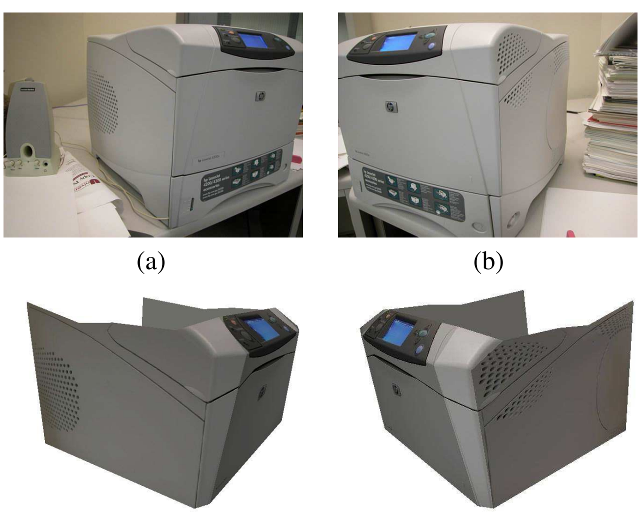
For the synthetic data shown herein the image resolution was 480×320. Eight points with 3D coordinates shown in Table 1 were reconstructed to recover four planes. Note that the point pair 6th and 7th in table 1 are not both visible in the images. The 7th point is not seen in the left image, while the 6th point can not be seen in the right image. Hence, they only differ in their x-coordinates by a sign. In the experiments presented herein, Gaussian noise with zero mean and a standard deviation of $\sigma \in [0, 0.5]$ was added to the projected image points. For each noise level, we ran our algorithm independently 100 times, and all the results shown are the averaged ones. The reconstructed 3D points were compared with the ground truth. In 1.5 pixel noise level, the results are shown in Table 1. To evaluate the performance vs noise, we also measured the absolute errors of X, Y and Z for every points. Mean and standard deviation of all coordinates are shown in Table 2 with different noise levels.

4.2. Real Data

We experimented with various real objects that contained mirror symmetry. Experimental results were verified against ground truth, which indicate an excellent performance for our approach, with the standard deviation of error in the recovered distance ratios under 3.5. Figures 2 and 3 show some real images of a symmetric objects. Note that some parts are only visible in one image. However, our technique accurately recovers all parts as shown in the snapshots of the reconstructed 3D models shown in in part (c) and (d) of these figures.

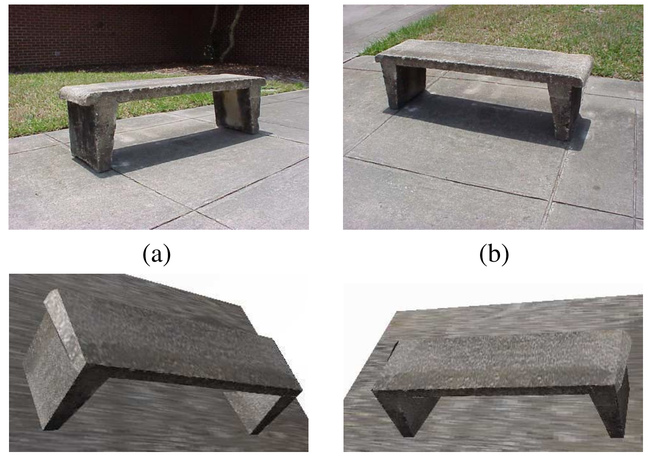
5. CONCLUSION

We have proposed a technique for reconstructing partially viewed symmetric objects given two views obtained by an unknown camera from unknown arbitrary positions and orientations. We use the symmetry of the scene to self-calibrate and recover both visible and occluded parts. The approach is shown to be very stable and provides excellent results.



(a) (b) (c) (d)

Fig. 2. (a) & (b) Two real images of a partially viewed symmetric object, (c) & (d) snapshots of the reconstructed 3D model including the occluded left and right portions



(a) (b) (c) (d)

Fig. 3. (a) & (b) Two real images of a partially viewed symmetric object, (c) & (d) snapshots of the reconstructed 3D model including the occluded left and right portions

6. REFERENCES

[1] R.I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2000.

[2] Rother C., "Linear multi-view reconstruction of points, lines, planes and cameras using a reference plane," in *Proc. of 9th International Conference on Computer Vision*, 2003, pp. 1210 – 1217.

[3] Starck J. and Hilton A., "Model-based multiple view reconstruction of people," in *Proc. of 9th International Conference on Computer Vision*, 2003, pp. 915 – 922.

[4] Mitsumoto H., Tamura S., Okazaki K., Kajimi N., and Fukui Y, "3-d reconstruction using mirror images based on a plane symmetry recovering method," *IEEE Trans. Pattern Analysis & Machine Intelligence*, vol. 14, no. 9, pp. 941–946, 1992.

[5] 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. of Fourteenth International Conference on Pattern Recognition*, 1998, vol. 2, pp. 1174 – 1176.

[6] Francois A.R.J., Medioni G.G., and Waupotitsch R., "Reconstructing mirror symmetric scenes from a single view using 2-view stereo geometry," in *Proc. of 16th International Conference on Pattern Recognition*, 2002, pp. 12 – 16.

[7] O.D. Faugeras, *Computer Vision: a Geometric Viewpoint*, MIT Press, 1993.

[8] A. Criminisi, I. Reid, and A. Zisserman, "Single view metrology," *International Journal of Computer Vision*, 2001.