On Symmetry and Multiple-View Geometry:
Structure, Pose, and Calibration from a Single Image*

WEI HONG, ALLEN YANG YANG, KUN HUANG AND YI MA
*Department of Electrical & Computer Engineering, University of Illinois at Urbana-Champaign,
1308 West Main St., Urbana, IL 61801, USA*
weihong@uiuc.edu
yangyang@uiuc.edu
kunhuang@uiuc.edu
yima@uiuc.edu

Abstract. In this paper, we provide a principled explanation of how knowledge in global 3-D structural invariants, typically captured by a group action on a symmetric structure, can dramatically facilitate the task of reconstructing a 3-D scene from one or more images. More importantly, since every symmetric structure admits a "canonical" coordinate frame with respect to which the group action can be naturally represented, the canonical pose between the viewer and this canonical frame can be recovered too, which explains why symmetric objects (e.g., buildings) provide us overwhelming clues to their orientation and position. We give the necessary and sufficient conditions in terms of the symmetry (group) admitted by a structure under which this pose can be uniquely determined. We also characterize, when such conditions are not satisfied, to what extent this pose can be recovered. We show how algorithms from conventional multiple-view geometry, after properly modified and extended, can be directly applied to perform such recovery, from all "hidden images" of one image of the symmetric structure. We also apply our results to a wide range of applications in computer vision and image processing such as camera self-calibration, image segmentation and global orientation, large baseline feature matching, image rendering and photo editing, as well as visual illusions (caused by symmetry if incorrectly assumed).

Keywords: structure from symmetry, multiple-view geometry, symmetry group, reflective symmetry, rotational symmetry, and translational symmetry

1. Introduction

One of the main goals of computer vision is the study of how to infer three-dimensional (3-D) information (e.g., shape, layout and motion) of a scene from its two-dimensional (2-D) image(s). A particular thrust of effort is to extract 3-D geometric information from 2-D images by exploiting geometric relationships among multiple images of the same set of features on a 3-D object. This gives rise to the subject of *multiple-view geometry*, a primary focus of study in the computer vision community for the past two decades or so. Unfortunately, certain relationships among features *themselves* have been, to a large extent, ignored or at least under-studied. Some of those relationships, as we will see from this paper, have significant impact on the way that 3-D information can be (and should be) inferred from images.

Before we proceed further, let us pause and exam the images given in Fig. 1 below. What do they have in common? Notice that these images are just a few
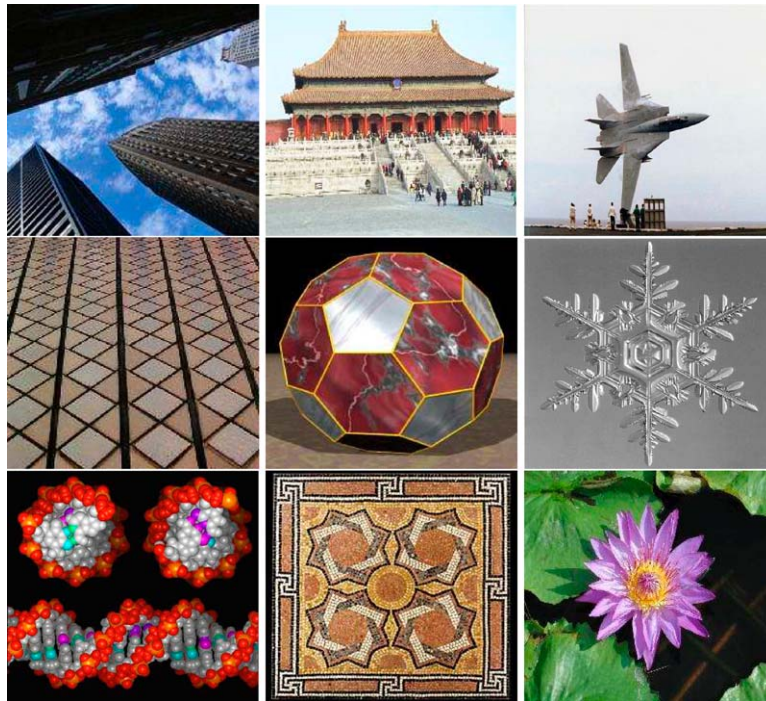
Figure 1.   Symmetry is in: architecture, machines, textures, crystals, molecules, ornaments, and nature, etc.

representatives of a common phenomenon exhibited in nature or man-made environment: *symmetry*. It is not so hard to convince ourselves that even from only a *single* image, we are able to perceive clearly the 3-D structure and relative pose (orientation and location) of the object being seen, even though in the image the shape of the object is distorted by the perspective projection. The reason is, simply put, there is symmetry at play.[1]

The goals of this paper are to provide a principled explanation why symmetry could encode 3-D information within a single perspective image and to develop algorithms based on multiple-view geometry that efficiently extract the 3-D information from single images. There are two things which we want to point out already:

1. Symmetry is *not the only* cue which encodes 3-D information through relationships among a set of features (in one image or more images). For instance, incidence relations among points, lines, and planes may as well provide 3-D information to the viewer;
2. The concept of symmetry that we consider here is *not* just the (bilateral) reflective symmetry, or the (statistical) isotopic symmetry which has been stud-

ied in a certain extent in the computer vision literature. Instead it is a more general notion describing global structural invariants of an object under the action of any group of transformations. To clarify this notion is one of the goals of this paper.

Symmetry, as a useful geometric cue to 3-D information, has been extensively discussed in psychological vision literature (Marr, 1982; Plamer, 1999). Nevertheless, its contribution to computational vision so far has been explored often through statistical methods, such as the study of isotropic textures (e.g., for the 4th image of Fig. 1) (Gibson, 1950; Witkin, 1988; Zabrodsky et al., 1995; Mukherjee et al., 1995; Malik and Rosenholtz, 1997; Rosenholtz and Malik, 1997; Leung and Malik, 1997). It is the works of Garding (1992, 1993) and Malik and Rosenholtz (1997) that have provided people a wide range of efficient algorithms for recovering the shape (i.e. the slant and tilt) of a textured plane based on the assumption of isotropy (or weak isotropy). These methods are mainly based on collecting statistical characteristics (e.g., the distribution of edge directions) from sample patches of the texture and comparing them with those of adjacent patches against the isotropic hypothesis. Information about the

surface shape is then often conveniently encoded in the discrepancy or variation of these characteristics.

But symmetry is by nature a geometric property! Although in many cases the result of symmetry indeed causes certain statistical homogeneity (like the 4th image of Fig. 1), there are reasons to believe that more accurate and reliable 3-D geometric information can be retrieved if we can directly exploit this property through geometric means. For example, for the texture shown in the 4th image of Fig. 1, shouldn't we directly exploit the fact that the tiling is invariant under certain proper translations parallel to the plane? To a large extent, such a geometric approach is complementary to extant statistical approaches: if statistical homogeneity can be exploited for shape recovery, so can geometric homogeneity, especially in cases where symmetry is the underlying cause for such homogeneity. Of course, for cases where statistical methods no longer apply (e.g., the 5th image of Fig. 1), geometric methods remain as the only option. One may call this approach as *structure from symmetry*.

We are by no means the first to notice that symmetry, especially reflective symmetry, can be exploited by geometric means for retrieving 3-D geometric information. Mitsumoto et al. (1992) studied how to reconstruct a 3-D object using the mirror image based on planar symmetry, Vetter and Poggio (1994) proved that for any reflective symmetric 3-D object one non-accidental 2-D model view is sufficient for recognition, Zabrodsky and Weinshall (1997) used bilateral symmetry assumption to improve 3-D reconstruction from image sequences, and Zabrodsky et al. (1995) provided a good survey on studies of reflective symmetry and rotational symmetry in computer vision at the time.

In 3-D object and pose recognition, Rothwell et al. (1993) pointed out that the assumption of reflective symmetry can be used in the construction of projective invariants and is able to eliminate certain restrictions on the corresponding points. Cham and Cipolla (1996) built the correspondences of contours from reflective symmetry. For translational symmetry, Schaffalitzky and Zisserman (2000) used it to detect the vanishing lines and points. Liu et al. (1995) analyzed the error of obtaining 3-D invariants derived from translational symmetry. In addition to isometric symmetry, Liebowitz and Zisserman (1998), Criminisi and Zisserman (1999) and Criminisi and Zisserman (2000) showed that other knowledge (e.g., length ratio, vanishing line, etc.) in 3-D also allows accurate reconstruction of structural metric and camera pose.

For the detection of symmetry from images, Marola (1989), Kiryati and Gofman (1998) and Mukherjee et al. (1995) presented efficient algorithms to find axes of reflective symmetry in 2-D images, Sun and Sherrah (1997) discussed reflective symmetry detection in 3-D space, and Zabrodsky et al. (1995) introduced a *symmetry distance* to classify reflective and rotational symmetry in 2-D and 3-D spaces (with some related comments given in Kanatani (1997)). Carlsson (1998) and Gool et al. (1996) derived methods to find 3-D symmetry from invariants in the 2-D projection. Liu and Colline (1998) proposed a method to classify any images with translational symmetry into the 7 Frieze groups and 17 wallpaper groups.

However, there is still a lack of formal and unified analysis as well as efficient algorithms which would allow people to easily make use of numerous and different types of symmetry that nature offers. Is there a unified approach to study 3-D information encoded in a 2-D perspective image of an object that exhibits certain symmetry? This paper will try to provide a definite answer to this question. Our work differs from previous results in at least the following three aspects:

1. We study symmetry under *perspective projection* based on existing theory of multiple-view geometry.[2] We claim that in order to fully understand such 3-D information encoded in a single image, one must understand geometry among multiple images.

2. In addition to recover 3-D structure of a symmetric object from its image, we show that any type of symmetry is naturally equipped with a *canonical (world) coordinate frame*, from which the viewer's relative pose to the object can be recovered.

3. We give the *necessary and sufficient conditions* in terms of the symmetry group of the object under which the canonical pose can be uniquely recovered, and we characterize the *inherent ambiguity* for each fundamental type of symmetry. Thus, for the first time, geometric group theory and (perspective) multiple-view geometry are elegantly and tightly integrated.

During the development, an important principle associated with images of symmetric objects will be examined with care: *One image of a symmetric object is equivalent to multiple images.* This principle is however not entirely correct since, as we will see, often relationships among such "images" will not be the

same as those among conventional images. It in fact requires careful modifications to existing theories and algorithms in multiple-view geometry if they are to be correctly applied to images of symmetric objects.

2.   Problem Formulation

Before we formulate the problem in a more abstract form, let us take a look at a simple example: a planar board with a symmetric pattern as shown in Fig. 2. It is easy to see that, from any generic viewpoint, there are at least four equivalent vantage points (with only the rotational symmetry considered, for now) which give rise to an identical image. The only question is which corners in the image correspond to the ones on the board. In this sense, these images are in fact *different* from the original one. We may call these images as "hidden."[3] For instance, in Fig. 2, we labeled in bracket corresponding corner numbers for such a hidden image.

In addition to the rotational symmetry, another kind of symmetry, the reflective symmetry, can give rise to a not so conventional type of hidden images, as shown in Fig. 3. Notice that, in the figure, the two "hidden images" with the four corners labeled by numbers in bracket cannot be an image of the same board from any (physically viable) vantage point![4] Nevertheless, as we will see below, just like the rotational symmetry, this type of hidden images also encodes rich 3-D geometric information about the object.

There is yet another type of symmetry "hidden" in a pattern like a checker board. As shown in Fig. 4 below, for a pattern that repeats a fundamental region indefinitely along one or more directions, the so-called "infinite rapport," one would obtain exactly "the same" image had the images been taken at vantage points that differ from each other by multiples nT of one basic translation T. Although all images would appear to be the same, features (e.g., points, lines) in these images correspond to different physical features in the world.
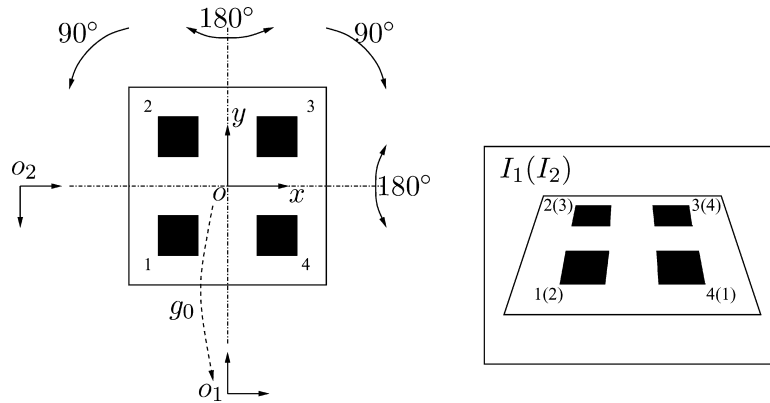

Figure 2.   Left: a checker board whose symmetry includes reflection along the x and y axes and rotation about o by $90°$. Right: an image taken at location o_1. Notice that the image would appear to be exactly the same if it were taken at o_2 instead. g_0 is the relative pose of the board we perceive from the image on the right.
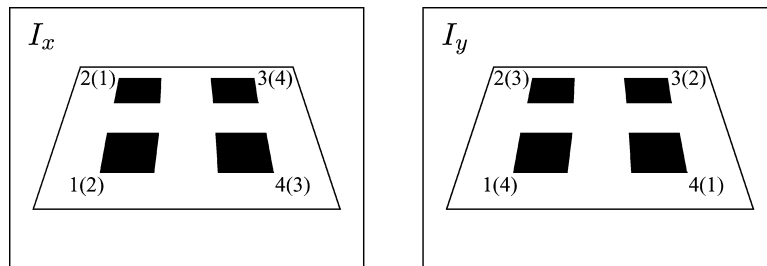

Figure 3. I_x: Corner correspondence between the original image of the board and an "image" with the board reflected in the x-axis by $180°$; I_y: Corner correspondence between the original image of the board and an "image" with the board reflected in the y-axis by $180°$.
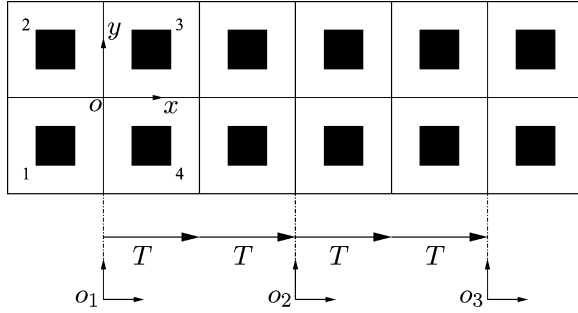
Figure 4. The checker pattern is repeated indefinitely along the *x*-axis. Images taken at o_1, o_2, and o_3 will be the same.

Therefore, for an image like the 4th one in Fig. 1, it in fact may give rise to many (in theory, possibly infinitely many) "hidden images." There is clearly a reason to believe that it is these (many) hidden images that give away the geometry of the plane (e.g., tilt, slant) to the viewer's eyes.

It is then not hard to imagine that the combination of the rotational, reflective and translational symmetries will give rise to all sorts of symmetric objects in 2-D or 3-D space, many of which could be rather complicated. In our man-made world, symmetric objects are ubiquitous, under the names of "ornament," "mosaic," "pattern," or "tiling," etc. Fascination about symmetric objects can be traced back to ancient Egyptians and Greeks.[5] Nevertheless, a formal mathematical inquiry to symmetry is known as Hilbert's 18th problem, and a complete answer to it was not found till 1910 by Bieberbach (1910). While in the appendix we briefly review results of a complete list for 2-D and 3-D symmetric structures and groups, this paper will focus on how to combine this knowledge about symmetry with multiple-view geometry so as to infer 3-D information of a symmetric object from its image(s).

In order to explain why symmetry gives away accurate information about structure and location of a symmetric 3-D object from a single 2-D perspective image, we will need a mathematical framework within which all types of symmetries (that we have mentioned or not mentioned in the above examples) can be uniformly taken into account. Only if we can do that, will the introduction of symmetry into multiple-view geometry become natural and convenient.

Definition 1 (Symmetric structure and its group action). A set of points $S \subset \mathbb{R}^3$ is called a *symmetric structure* if there exists a non-trivial subgroup G of the Euclidean group $E(3)$ that acts on it. That is, for any element $g \in G$, it defines a bijection (i.e. a one-to-one, onto) map from S to itself:

$$g \in G : S \to S.$$

Sometimes we say that S has a symmetry group G. Or G is a group of symmetries of S.

In particular, we have $g(S) = g^{-1}(S) = S$ for any $g \in G$. Mathematically, symmetric structures and groups are equivalent ways to capture symmetry: any symmetric structure is invariant under the action of its symmetry group; and any group (here as a subgroup of $E(3)$) defines a class of (3-D) structures that are invariant under this group action (see Appendix A). Here we emphasize that G is in general a subgroup of the Euclidean group $E(3)$ but not the special one $SE(3)$. This is because many symmetric structures that we are going to consider are invariant under reflection which is an element in $O(3)$ but not $SO(3)$.[6] For simplicity, in this paper we consider G to be a discontinuous (or discrete) group.[7]

Using the homogeneous representation of $E(3)$, any element $g = (R, T)$ in G can be represented as a 4×4 matrix of the form

$$g = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 4}, \tag{1}$$

where $R \in \mathbb{R}^{3 \times 3}$ is an orthogonal matrix ("R" for both rotation and reflection) and $T \in \mathbb{R}^3$ is a vector ("T" for translation). Note that in order to represent G in this way, a world coordinate frame must have been chosen. It is conventional to choose the origin of the world coordinate frame to be the center of rotation and its axes to line up with the axes of rotation and reflection or direction of translation. Often the *canonical* world coordinate frame results in the simplest representation of the symmetry (Ma et al., 2003).

Now suppose that an image of a symmetric structure S is taken at a vantage point $g_0 = (R_0, T_0) \in SE(3)$—denoting the pose of the structure relative to the viewer or the camera. Here g_0 is assumed to be represented with respect to the canonical world coordinate frame for the symmetry. If so, we call g_0 the *canonical pose*. As we will see shortly, the canonical pose g_0 from the viewer to the object can be uniquely determined from a single image as long as symmetry admitted by the object (or the scene) is "rich" enough.

A (calibrated) image of S is a set of image points $\mathcal{I}_0 \subset \mathbb{R}^3$ and in homogeneous coordinates each image point $x \in \mathcal{I}_0$ satisfies

$$\lambda x = \Pi_0 g_0 X = [R_0, T_0] X, \qquad (2)$$

where $\Pi_0 \doteq [I, 0] \in \mathbb{R}^{3 \times 4}$, $X \in \mathbb{R}^4$ is homogeneous coordinates of a point $p \in S$, and $\lambda \in \mathbb{R}$ is the depth scale. Now since $g(S) = S$ for all $g \in G$, we have $g(\mathcal{I}_0) = \mathcal{I}_0$.[8] That is, if we change the vantage point from g_0 to $g_0 g$, due to the symmetry of S we would get apparently the same image from S. Although at a different vantage point $g_0 g$ an image x of a point $p \in S$ will become a different point on the image, say x', x' must coincide with one of the image points in $\mathcal{I}_0$ (taken from the original vantage point). That is $x' \in \mathcal{I}_0$ and we call $x' = g(x)$. Thus, the group G does nothing but permutes image points in $\mathcal{I}_0$, which is an action induced from its action on S in 3-D. For simplicity, if we ignore (for now) the effect of self-occlusion and limited field of view, we may always assume $g(x) \in \mathcal{I}_0$ for all $x \in \mathcal{I}_0$ and all $g \in G$.

From the symmetry G, the Eq. (2) yields a system of equations

$$\lambda' g(x) = \Pi_0 g_0 g X, \quad \forall g \in G. \qquad (3)$$

Therefore, given only one image $\mathcal{I}_0$, if we know the type of symmetry G in advance and how its elements act on points in $\mathcal{I}_0$, every image point in the set $\{g(x) : g \in G\}$ can be interpreted as a different image of the same point X seen from a different vantage point. That is, we effectively have as many as $|G|$ images of the same 3-D structure S.[9]

The problem *we are interested here is that given one image of a symmetric structure S (with a known symmetry group G) taken at a vantage point g_0, to what extent we may recover the pose g_0 and the structure S, and how?*

Notice that here g_0 is not the relative motion between different vantage points but the relative pose from the object to the viewer! As we will soon see, symmetry in general encodes strong 3-D information which often allows us to determine g_0. Of course, due to symmetry, there is no unique solution to the canonical pose g_0 either. That is, because $x \sim \Pi_0 g_0 g(g^{-1} X)$,[10] the image point x might as well be the image of the point $g^{-1} X$ seen from the vantage point $g_0 g$. Hence the above question should be restated as "to what extent we may identify the set $\{g_0 g\}$." Algebraically, this is to identify

the left coset $g_0 G$ among the set of all cosets of G (in $E(3)$). The set of all cosets of G in $E(3)$ is typically denoted as $E(3)/G$. Since G is in general not a normal subgroup of $E(3)$, $E(3)/G$ is not necessarily a group itself and is called the left coset space. Since in most cases we will be dealing with a discrete and finite group G, determining g_0 up to one of its cosets will then give sufficient information about the relative orientation of the object.

3.   Symmetry and Multiple-View Geometry

As we have suggested before, although symmetry is a phenomenon associated with a single image, a full understanding of its effect on 3-D reconstruction depends on the theory of multiple-view geometry. In this section, we examine analytically how the geometry of a single image is closely related to that of multiple ones whenever symmetry is present.

3.1.   Symmetric Multiple-View Rank Condition

Let $\{g_i = (R_i, T_i)\}_{i=1}^m$ be m different elements in G. Then one image $x \sim \Pi_0(g_0 X)$ of a symmetric structure with the symmetry G is equivalent to at least m different images that satisfy the following equations

$$g_1(x) \sim \Pi_0 g_0 g_1 g_0^{-1}(g_0 X),$$
$$g_2(x) \sim \Pi_0 g_0 g_2 g_0^{-1}(g_0 X),$$
$$\vdots$$
$$g_m(x) \sim \Pi_0 g_0 g_m g_0^{-1}(g_0 X).$$

These images must be related by the so-called multiple-view rank condition (Ma et al., 2002). That is, the following multiple-view matrix[11]

$$M_g(x) \doteq \begin{bmatrix} \widehat{g_1(x)} R(g_1) x & \widehat{g_1(x)} T(g_1) \\ \widehat{g_2(x)} R(g_2) x & \widehat{g_2(x)} T(g_2) \\ \vdots & \vdots \\ \widehat{g_m(x)} R(g_m) x & \widehat{g_m(x)} T(g_m) \end{bmatrix} \in \mathbb{R}^{3(m-1) \times 2},$$

$$(4)$$

with

$$\begin{cases} R(g_i) \doteq R_0 R_i R_0^T \in O(3), \\ \qquad\qquad\qquad\qquad\qquad i = 1, 2, \ldots, m \\ T(g_i) \doteq (I - R_0 R_i R_0^T) T_0 + R_0 T_i, \in \mathbb{R}^3, \end{cases} \qquad (5)$$

satisfies the rank condition

$$\boxed{\operatorname{rank}(M_g(\boldsymbol{x})) \leq 1, \quad \forall \boldsymbol{x} \in \mathcal{I}_0.} \tag{6}$$

Note that this rank condition is independent of any particular order of the group elements $g_1, g_2, \dots, g_m$ and it captures the *only* fundamental invariant that a perspective image of a symmetric structure admits.[12] Thus, we call it the *symmetric multiple-view rank condition*. Note that if $G \subseteq O(3)$ (i.e. $T_i = 0$ for all i), the expression for $T(g_i)$ can be simplified to

$$T(g_i) = \left(I - R_0 R_i R_0^T\right) T_0, \quad i = 1, 2, \dots, m. \tag{7}$$

Therefore, one image of a symmetric structure S with its symmetry group G is equivalent to $m = |G|$ images of $n = |S|$ feature points.[13] The reconstruction of $(R(g_i), T(g_i))$ and the 3-D structure of S can be easily solved by a factorization algorithm based on the above rank condition (see Ma et al. 2002). Nevertheless, in order to solve for the "canonical pose" $g_0 = (R_0, T_0)$, we need to further solve them from Eq. (5) once $R(g_i)$ and $T(g_i)$ are recovered. To solve for R_0, it requires us to solve a system of *Lyapunov* type equations:[14]

$$R(g_i) R_0 - R_0 R_i = 0, \quad i = 1, 2, \dots, m, \tag{8}$$

with $R(g_i)$ and R_i known. The uniqueness of the solution R_0 depends on the relation between $R(g_i)$ and R_i. Before we can reach a general statement on the solvability of such equations, let us first study individual types of symmetry, i.e. , the reflective, rotational, and translational symmetry.

3.2. Ambiguity Induced from Three Types of Symmetry

3.2.1. Reflective Symmetry.
Many man-made objects, for example a building or a car, are symmetric with respect to a central plane (the plane or mirror of reflection). That is, the structure concerned is invariant under a reflection with respect to this plane. Without loss of generality, suppose this plane is the yz-plane of a pre-chosen world coordinate frame. For instance, in Fig. 2, the board is obviously symmetric with respect to the yz-plane if the z-axis is the normal to the board.

Then a reflection in this plane can be described by the motion $g = (R, 0)$ where

$$R = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \in O(3) \subset \mathbb{R}^{3 \times 3}, \tag{9}$$

is an element in $O(3)$ and it has $\det(R) = -1$. Notice that a reflection always fixes the plane of reflection. If a reflection is the only symmetry that a structure has, then the group G consists of only two elements $\{e, g\}$ where $e = g^2$ is the identity map. In other words, G is isomorphic to the group $\mathbb{Z}_2$.

If one image of such a symmetric object is taken at $g_1 = (R_0, T_0)$, then we have the following two equations for each image point on this structure:

$$\lambda \boldsymbol{x} = \Pi_0 g_0 X, \quad \lambda' g(\boldsymbol{x}) = \Pi_0 g_0 g X. \tag{10}$$

To simplify the notation, define $R' \doteq R_0 R R_0^T$ and $T' \doteq (I - R_0 R R_0^T) T_0$. Then the symmetric multiple-view rank condition, in the two-view case, is reduced to the following well-known epipolar constraint:

$$g(\boldsymbol{x})^T \widehat{T}' R' \boldsymbol{x} = 0. \tag{11}$$

Note that here $\det(R') = \det(R) = -1$ is not a rotation matrix. Therefore the matrix $E = \widehat{T}' R'$ is not a conventional Essential matrix and the classical 8-point (or 7-point) algorithm needs to be modified accordingly, if it is to be used to recover (R', T')![15]

Once the $R' = R_0 R R_0^T$ is obtained, we need to use R' and R to solve for R_0. The associated Lyapunov equation can be rewritten as

$$R' R_0 - R_0 R = 0, \tag{12}$$

with R' and R known.

Lemma 2 (Reflective Lyapunov equation). *Let $L: \mathbb{R}^{3 \times 3} \to \mathbb{R}^{3 \times 3}; R_0 \to R' R_0 - R_0 R$ be the Lyapunov map associated to the above equation, with R a reflection and $R' = R_0 R R_0^T$ both known. The kernel $ker(L)$ of L is in general 5-dimensional. Nevertheless, for orthogonal solutions of R_0, the intersection $ker(L) \cap SO(3)$ is only a 1-parameter family which corresponds to an arbitrary rotation in the plane of reflection.*

Proof: Let $\{\lambda_1, \lambda_2, \lambda_3\}$ and $\{\mu_1, \mu_2, \mu_3\}$ be the eigenvalues of R' and R respectively. Let $\{v_1, v_2, v_3\}$ be the

right eigenvectors of R' and $\{w_1, w_2, w_3\}$ be the left eigenvectors of R.

$$R'v_i = \lambda_i v_i, \quad w_j^T R = \mu_j w_j^T.$$

The eigenvalues of L are $\{\lambda_i - \mu_j\}$, $i, j = 1, 2, 3$ with $v_i w_j^T$ as corresponding eigenvectors because

$$L(v_i w_j^T) = R'v_i w_j^T - v_i w_j^T R = (\lambda_i - \mu_j)v_i w_j^T.$$

Then the kernel of L is the span$\{v_i w_j^T \mid \lambda_i = \mu_j\}$.

Without loss of generality, we assume that the reflection R is with respect to the yz-plane,

$$R = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{3 \times 3}. \qquad (13)$$

Its eigenvalues are

$$\mu_1 = -1, \quad \mu_2 = \mu_3 = 1.$$

The left eigenvector associated with $\mu_1 = -1$ is $w_1^T = [1\ 0\ 0]$. Two real left eigenvectors associated with the two repeated eigenvalue $\mu_2 = \mu_3 = 1$ are $w_2^T = [0\ 1\ 0]$, $w_3^T = [0\ 0\ 1]$.

For the R', we have

$$R'v_i = R_0 R R_0^T v_i = \lambda_i v_i \Leftrightarrow R(R_0^T v_i) = \lambda_i (R_0^T v_i).$$

From the equation above, the eigenvalues of R' are the same as the eigenvalue of R which are

$$\lambda_1 = -1, \quad \lambda_2 = \lambda_3 = 1.$$

Since w_1, w_2, w_3 are obviously also the right eigenvectors of R, the eigenvector of R' associated with $\lambda_1 = -1$ is $v_1 = R_0 w_1$; and two real eigenvectors of R' associated with the two repeated eigenvalues $\lambda_2 = \lambda_3 = 1$ are $v_2 = R_0 w_2$, $v_3 = R_0 w_3$, respectively. v_2 and v_3 are real and orthogonal to each other.

We have a total of five combinations of i, j such that $\lambda_i = \mu_j$. Thus, the kernel of L is a 5-dimensional space which has a basis consisting of the elements

$$v_1 w_1^T = [v_1,\ 0,\ 0], \quad v_2 w_2^T = [0,\ v_2,\ 0],$$
$$v_2 w_3^T = [0,\ v_3,\ 0], \quad v_3 w_2^T = [0,\ 0,\ v_2],$$
$$v_3 w_3^T = [0,\ 0,\ v_3] \in \mathbb{R}^{3 \times 3}.$$

Any linear combination of the basis elements

$$R_0 = \sum_{i=1}^{5} a_i v_i w_i^T \in \mathbb{R}^{3 \times 3}, \quad a_i \in \mathbb{R},$$

will be a solution of the equation $R'R_0 - R_0 R = 0$. Because $R_0 \in SO(3)$, we can impose the constraint of rotation matrix. This gives us

$$R_0 = [\pm v_1,\ v_2 \cos(\alpha) + v_3 \sin(\alpha),$$
$$- v_2 \sin(\alpha) + v_3 \cos(\alpha)] \in SO(3),$$

where $\alpha \in \mathbb{R}$ is an arbitrary angle. So R_0 has infinitely many solutions. Geometrically, the three column of R' can be interpreted as the three axes of the world coordinate frame that we attached to the structure. The ambiguity in R_0 then corresponds to an arbitrary rotation of the yz-plane around the x-axis. $\qquad \square$

If the structure also has symmetry with respect to another plane, say the xz-plane as in the case of the checker board (Fig. 2), this 1-parameter family ambiguity can therefore be eliminated.

In practice, it is very common that the object, like the checker board in Fig. 2, is flat and all the feature points on S lie in a 2-D plane P in 3-D space, i.e. $S \subset P$. In this case, the above ambiguity can also be reduced even if an extra symmetry is not available. Without loss of generality, the plane P can be described by an equation

$$N^T X = d, \qquad (14)$$

with $N \in \mathbb{R}^3$ as its normal vector and $d \in \mathbb{R}_+$ as the distance from the center of the camera to the plane P. Using this notation, the image point $\boldsymbol{x}$ and its symmetry point $g(\boldsymbol{x})$ satisfy the planar homography[16]

$$\widehat{g(\boldsymbol{x})} H \boldsymbol{x} = 0, \qquad (15)$$

where the matrix H is the homography matrix of the form

$$H = R' + \frac{1}{d} T' N^T \quad \in \mathbb{R}^{3 \times 3}. \qquad (16)$$

With more than 4 points, the matrix H can be recovered from Eq. (15). Decomposing the homography matrix H yields[17]

$$H \rightarrow \left\{ R',\ \frac{1}{d} T',\ N \right\}. \qquad (17)$$

If we assume the normal of the planar structure is in the direction of the z-axis, we may get a unique solution for the rotation,

$$R_0 = [\pm v_1, \pm \hat{N} v_1, N] \in SO(3),$$

where v_1, v_2, v_3 are eigenvectors of R' as in the proof of Lemma 2.

After R_0 is recovered, T_0 is recovered up to the following form

$$T_0 \in \left(I - R_0 R R_0^T\right)^\dagger T' + \text{null}\left(I - R_0 R R_0^T\right), \quad (18)$$

where $\left(I - R_0 R R_0^T\right)^\dagger$ is the pseudo inverse of $I - R_0 R R_0^T$ and $\text{null}(I - R_0 R R_0^T) = \text{span}\{v_2, v_3\}$ since both v_2 and v_3 (as in the proof of Lemma 2) are in the null space of the matrix $I - R_0 R R_0^T$. Such ambiguity in the recovered $g_0 = (R_0, T_0)$ is exactly what we should have expected: With a reflection with respect to the yz-plane, we in principle can only determine the y-axis and z-axis (including the origin) of the world coordinate frame up to any orthonormal frame within the yz-plane, which obviously has three degrees of freedom, parameterized by (α, β, γ) (where α is as in the proof of Lemma 2). If S itself is in a plane, we may choose the z-axis of the world frame to be the nor-

mal to the plane and the origin of the frame to be in the plane. Thus, we can reduce this ambiguity to a 1-parameter family: only the origin o now may translate freely along the y-axis, the intersection of the plane where S resides and the plane of reflection.

To conclude our discussion on the reflective symmetry, we have

Proposition 3 (Canonical pose from reflective symmetry). *Given an image of a structure S with a reflective symmetry with respect to a plane in* **3**-*D, the canonical pose g_0 can be determined up to an arbitrary choice of an orthonormal frame in this plane, which is a* **3**-*parameter family of ambiguity (i.e. $SE(2)$). However, if S itself is in a (different) plane, g_0 is determined up to an arbitrary translation of the frame along the intersection line of the two planes (i.e. $\mathbb{R}$).*

Figures 5 and 6 demonstrate an experiment with the reflective symmetry. The checker board is a planar structure which is symmetric with respect to the central line of itself (in fact there are many more local reflective symmetry on parts of the board).

3.2.2. Rotational Symmetry. Now suppose we replace the reflection R above by a proper rotation. For instance, in Fig. 2, the pattern is symmetric with
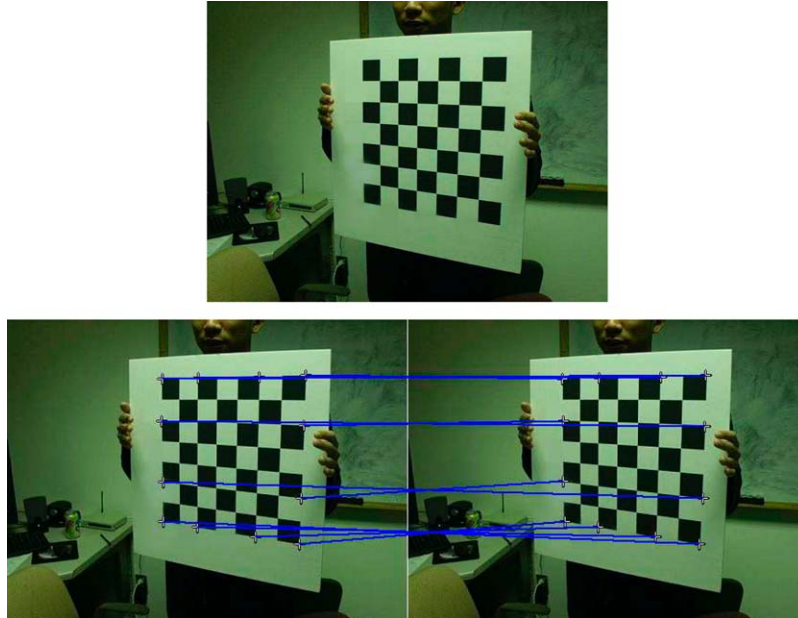

Figure 5. Top: An image of a reflectively symmetric checker board. Bottom: The symmetry is represented by some corresponding points. We draw two identical images here to illustrate the correspondence more clearly: Points in the left image are corresponding to points in the right image by a reflective symmetry.
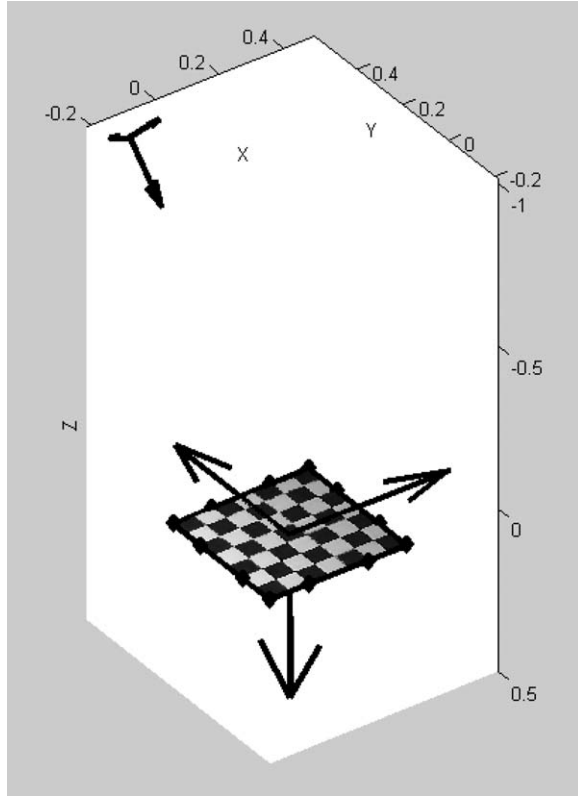
Figure 6. The reconstruction result from the reflective symmetry. The recovered structure is represented in the canonical world coordinate frame. From our discussion above, the origin *o* of the world coordinate frame may translate freely along the *y*-axis. The smaller coordinate frame is the camera coordinate frame. The longest axis is the *z*-axis of the camera frame which represents the optical axis of the camera.

respect to any rotation by $n \cdot 90°$ radians around *o* in the xy-plane. Now the question becomes, knowing the rotation R and its conjugation $R' = R_0 R R_0^T$, to what extent we can determine R_0 from the equation $R' R_0 - R_0 R = 0$. Without loss of generality, we assume R is of the form $R = e^{\hat{\omega}\theta}$ with $\|\omega\| = 1$ and $0 < \theta < \pi$, hence it has three distinct eigenvalues $\{1, e^{+j\theta}, e^{-j\theta}\}$.

Lemma 4 (Rotational Lyapunov equation). *Let L: $\mathbb{R}^{3 \times 3} \to \mathbb{R}^{3 \times 3}$; $R_0 \to R' R_0 - R_0 R$ be the Lyapunov map associated to the above equation, with R a rotation and $R' = R_0 R R_0^T$ both known. The kernel $ker(L)$ of this Lyapunov map is in general 3-dimensional. Nevertheless, for orthogonal solutions of R_0, the intersection $ker(L) \cap SO(3)$ is a 1-parameter family cor-*

responding to an arbitrary rotation (of α radians) about the rotation axis ω of R.

Proof: Let $\{\lambda_1, \lambda_2, \lambda_3\}$ and $\{\mu_1, \mu_2, \mu_3\}$ be the eigenvalues of R' and R respectively. Let $\{v_1, v_2, v_3\}$ be the right eigenvectors of R' and $\{w_1, w_2, w_3\}$ be the left eigenvectors of R.

$$R' v_i = \lambda_i v_i, \quad w_j^T R = \mu_j w_j^T.$$

The eigenvalues of L are $\{\lambda_i - \mu_j\}$, $i, j = 1, 2, 3$ with $v_i w_j^T$ as corresponding eigenvectors because

$$L(v_i w_j^T) = R' v_i w_j^T - v_i w_j^T R = (\lambda_i - \mu_j) v_i w_j^T.$$

Then the kernel of L is the span$\{v_i w_j^T \mid \lambda_i = \mu_j\}$.

Without loss of generality, we assume that the rotation is with respect to the z-axis. So R is

$$R = e^{\hat{\omega}\theta} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

with $\omega = [0, 0, 1]^T$ and $0 < \theta < 2\pi$. Any such a rotation matrix has three distinct eigenvalues,

$$\mu_1 = 1, \quad \mu_2 = e^{+j\theta}, \quad \mu_3 = e^{-j\theta}.$$

The left eigenvector associated with $\mu_1 = 1$ is $w_1^T = [0, 0, 1]$ which has a real value. The other two left eigenvectors associated with the two complex eigenvalues $\mu_2 = \bar{\mu}_3$ are $w_2^T = [-i, 1, 0]$, $w_3^T = [i, 1, 0]$ with $w_2^T = \bar{w}_3^T$.

For the R', we have

$$R' v_i = R_0 R R_0^T v_i = \lambda_i v_i \Leftrightarrow R(R_0^T v_i) = \lambda_i (R_0^T v_i).$$

From the equation above, the eigenvalues of R' are the same as the eigenvalue of R which are

$$\lambda_1 = 1, \quad \lambda_2 = e^{+j\theta}, \quad \lambda_3 = e^{-j\theta}.$$

The eigenvector associated with $\lambda_1 = 1$ is v_1^T which has a real value. The other two eigenvectors associated with the two conjugate complex eigenvalues $\lambda_2 = \bar{\lambda}_3$ are v_2, v_3 where $v_2 = \bar{v}_3$.

We have three pairs of (λ, μ) such that $\lambda_i = \mu_i$, $i = 1, 2, 3$. Thus, we get the basis for the kernel of L

$$v_1 w_1^T = [0, 0, v_1] \in \mathbb{R}^{3 \times 3},$$
$$v_2 w_2^T = [-i v_2, v_2, 0],$$
$$v_3 w_3^T = [i v_3, v_3, 0] = [i v_2^T, v_2^T, 0] \in \mathbb{C}^{3 \times 3}.$$

The real and imaginary parts of the complex basis

$$\text{Re}[v_2 w_2^T] = [\text{Im}(v_2), \text{Re}(v_2), 0],$$
$$\text{Im}[v_2 w_2^T] = [-\text{Re}(v_2), \text{Im}(v_2), 0],$$
$$\text{Re}[v_3 w_3^T] = [\text{Im}(v_2), \text{Re}(v_2), 0] = \text{Re}[v_2 w_2^T],$$
$$\text{Im}[v_3 w_3^T] = [\text{Re}(v_2), -\text{Im}(v_2), 0] = -\text{Im}[v_2 w_2^T],$$

are also in the kernel of L. Then the real kernel of L is a 3-dimensional space which has the basis

$$\{[0, 0, v_1], [\text{Im}(v_2), \text{Re}(v_2), 0],$$
$$[-\text{Re}(v_2), \text{Im}(v_2), 0]\} \in \mathbb{R}^{3 \times 3}.$$

The solution of R_0 will be the linear combination of the basis,

$$R_0 = a_1[0, 0, v_1] + a_2[\text{Im}(v_2), \text{Re}(v_2), 0]$$
$$+ a_3[-\text{Re}(v_2), \text{Im}(v_2), 0],$$

where coefficients $a_1, a_2, a_3 \in \mathbb{R}$. Because $R_0 \in SO(3)$, we can impose the constraint of it being a rotation matrix. Then R_0 is of the form

$$[\text{Im}(v_2)\cos(\alpha) + \text{Re}(v_2)\sin(\alpha),$$
$$- \text{Im}(v_2)\sin(\alpha) + \text{Re}(v_2)\cos(\alpha), \pm v_1],$$

where $\alpha \in \mathbb{R}$ is an arbitrary angle. So R_0 has infinitely many solutions. Geometrically, the three columns of R' can be interpreted as the three axes of the world coordinate frame that we attached to the structure. The ambiguity in R_0 then corresponds to an arbitrary rotation of the xy-plane around the z-axis. $\quad\square$

This lemma assumes that $0 < \theta < \pi$. If $\theta = \pi$, R has two repeated -1 eigenvalues and the proof above no longer applies. Nevertheless, we notice that $-R$ is exactly a reflection with two $+1$ eigenvalues, with a reflection plane orthogonal to the rotation axis of R. Thus, this case is essentially the same as the reflective case stated in Lemma 2. Although the associated Lyapunov map has a 5-dimensional kernel, its intersection with $SO(3)$ is the same as any other rotation.

In addition, it can be verified directly that the null space of the matrix $I - R_0 R R_0^T$ is always 1-dimensional (for $0 < \theta \leq \pi$) and $(I - R_0 R R_0^T)v_1 = 0$ (where v_1 was defined in the proof of the above lemma). Thus, the translation T_0 is recovered up to the form:

$$T_0 \in (I - R_0 R R_0^T)^\dagger T' + \text{null}(I - R_0 R R_0^T), \quad (19)$$

where $\text{null}(I - R_0 R R_0^T) = \text{span}\{v_1\}$. Together with the ambiguity in R_0, g_0 is determined up to a so-called *screw motion* about the rotation axis ω. Similar to the reflective case, we have

Proposition 5 (Canonical pose from rotational symmetry). *Given an image of a structure S with a rotational symmetry with respect to an axis $\omega \in \mathbb{R}^3$, the canonical pose g_0 is determined up to an arbitrary choice of a screw motion along this axis, which is a 2-parameter family of ambiguity (i.e. the screw group $SO(2) \times \mathbb{R}$). However, if S itself is in a (different) plane, g_0 is determined up to an arbitrary rotation around the axis (i.e. $SO(2)$).*

Figures 7 and 8 demonstrate an experiment with the rotational symmetry. Each face of the cube is a planar structure which is symmetric to another face by a rotation about the longest diagonal of the cube by 120 degree.

3.2.3. Translational Symmetry. In this case, since $R = I$ and $T \neq 0$, Eq. (5) is reduced to the following equations

$$R' = R_0 I R_0^T = I, \quad T' = R_0 T. \quad (20)$$

Obviously, the first equation does not give any information on R_0 (and the associated Lyapunov map is trivial), nor on T_0. From the second equation however, since both T and T' are known (up to a scale), R_0 can be determined up to a 1-parameter family of rotations (i.e. $SO(2)$). Thus, the choice of the world frame (including T_0) is up to a 4-parameter family (i.e. $SO(2) \times \mathbb{R}^3$).

Furthermore, if S is planar, which often is the case for translational symmetry, the origin o of the world frame can be chosen in the supporting plane, the plane normal as the z-axis, and T as the x-axis. Thus
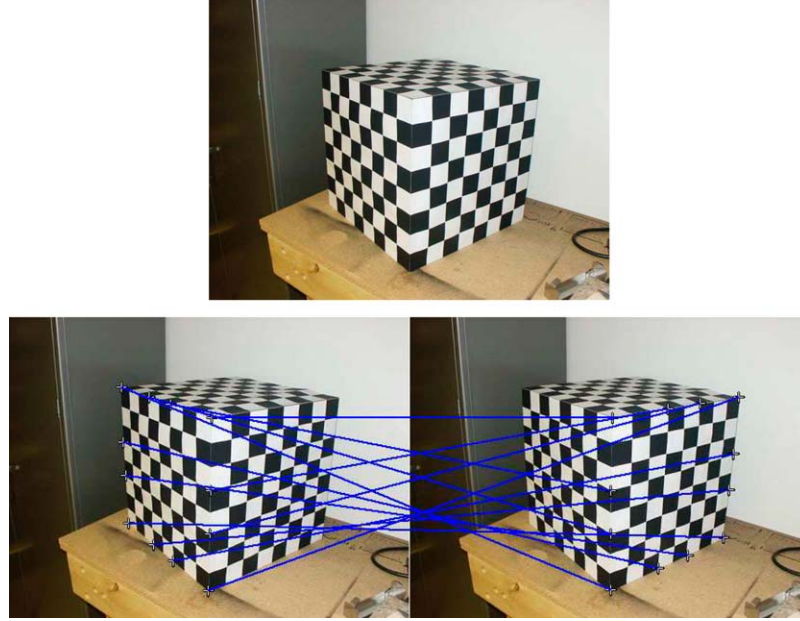
$$R_0 = [T', \hat{N}T', N] \in SO(3),$$

Figure 7. Top: An image of a cube which is rotationally symmetric about its longest diagonal axis. Bottom: The symmetry is represented by some corresponding points. Points in the left images correspond to points in the right image by a rotational symmetry.
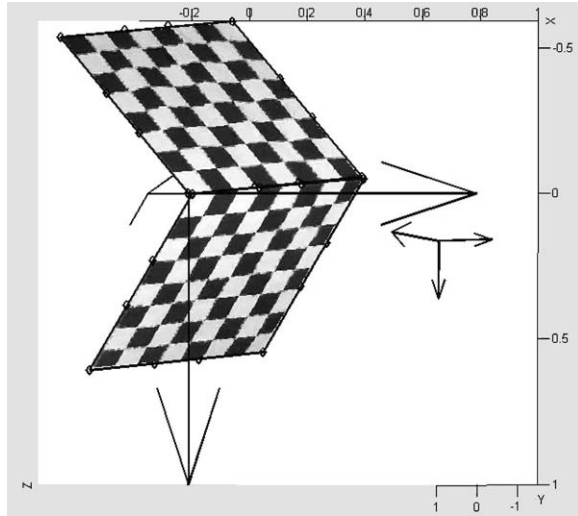

Figure 8. Reconstruction result from the rotational symmetry. The recovered structure is represented in the canonical world coordinate frame. From our discussion above, the origin *o* of the world coordinate may translate freely along the *z*-axis, and the *x*, *y*-axis can be rotated within the *xy* plane freely. The smaller coordinate frame is the camera coordinate frame. The longest axis is the *z*-axis of the camera frame which represents the optical axis of the camera.

where both T' and N can be recovered from decomposing the homography $H = I + \frac{1}{d}T'N^T$. We end up with a 2-parameter family of ambiguity in determining g_0 – translating *o* arbitrarily inside the plane (i.e.

$\mathbb{R}^2$). Even with an extra translational symmetry along a different direction, the origin *o* can be any place on the plane. Figures 9 and 10 demonstrate an experiment with reflective symmetry. A mosaic floor is a planar structure that is invariant with respect to the translation along proper directions.

We summarize in Table 1 the ambiguities in determining the pose g_0 from each of the three types of symmetry, for both generic and planar scenes.

3.3. Necessary and Sufficient Condition for a Unique Pose Recovery from Symmetry

As we have seen from above sections, there is always some ambiguity in determining the relative pose (g_0) from the vantage point to the canonical world coordinate frame (where the symmetry group G was represented in the first place) if only one type of symmetry is considered. In reality, most symmetric structures (or objects) have more than one types of symmetry. For instance, the board in Fig. 2 has all three types of symmetry present. It is the joint work of all symmetries as a group, not its individual elements, that conveys to the viewer a strong perception of the pose of the object.

Proposition 6 (Rotational and reflective symmetry).
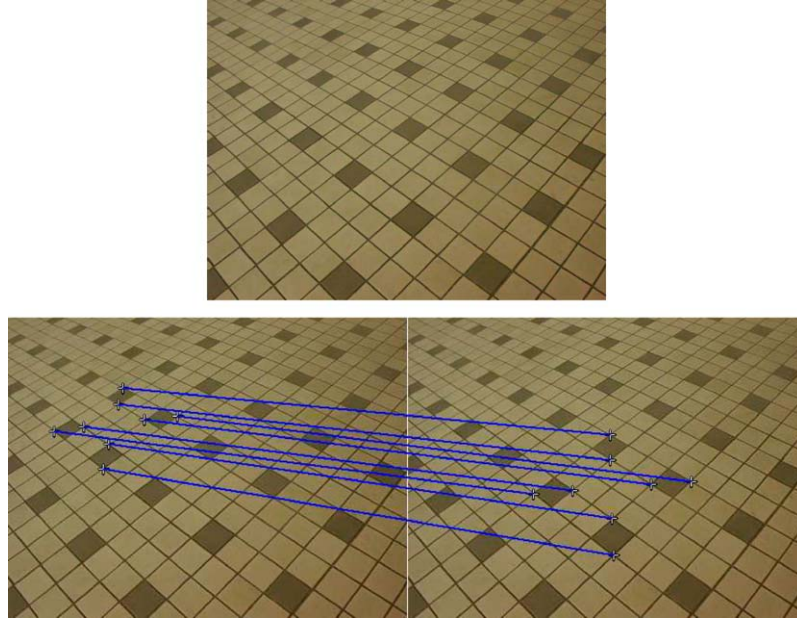*Given a* (*discrete*) *subgroup G of O*(3), *a rotation R_0 is*

Figure 9. Top: An image of a mosaic floor which admits translational symmetry. Bottom: The symmetry is represented by some corresponding points. We draw two identical images here to represent the correspondence more clearly: Points shown in the left images correspond to points shown in the right image by a translational symmetry.
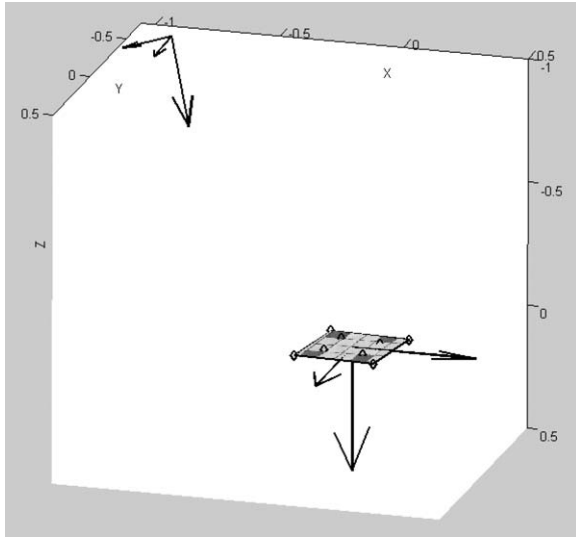

Figure 10. Reconstruction result for the translational symmetry. The structure is represented in the canonical world coordinate frame. From our discussion above, the origin o of the world coordinate may translate freely within the xy-plane. The smaller coordinate frame is the camera coordinate frame. The longest axis is the z-axis of the camera frame which represents the optical axis of the camera.

uniquely determined from the pair of sets $(R_0 G R_0^T, G)$ *if and only if the only fixed point of G acting on* $\mathbb{R}^3$ *is the origin.*

Proof: If $(R_0 G R_0^T, G)$ are not sufficient to determine R_0, then there exists at least one other $R_1 \in SO(3)$ such that $R_0 R R_0^T = R_1 R R_1^T$ for all $R \in G$. Let $R_2 = R_1^T R_0$. Then, $R_2 R = R R_2$ for all $R \in G$. Hence R_2 commutes with all elements in G. If R_2 is a rotation, all R in G must have the same rotation axis as R_2; if R_2 is a reflection, R must have its axis normal to the plane that R_2 fixes. This is impossible for a group G that only fixes the origin. On the other hand, if $(R_0 G R_0^T, G)$ is sufficient to determine R_0, then the group G cannot fix any axis (or a plane). Otherwise, simply choose R_2 to be a rotation with the same axis (or an axis normal to the plane), then it commutes with G. The solution for R_0 cannot be unique. □

Once R_0 is determined, it is then not difficult to show that, with respect to the same group G, T_0 can be uniquely determined from the second equation in

Table 1. Ambiguity in determining the canonical pose from three types of symmetry. (∗: "$(a + b)$-parameter" means there are an a-parameter family of ambiguity in R_0 of g_0 and b-parameter family of ambiguity in T_0 of g_0.)

Ambiguity	Lyapunov equation	g_0 (general scene)	g_0 (planar scene)
Reflective	5-dimensional	$(1 + 2)$-parameter*	$(0 + 1)$-parameter
Rotational	3-dimensional	$(1 + 1)$-parameter	$(1 + 0)$-parameter
Translational	9-dimensional	$(1 + 3)$-parameter	$(0 + 2)$-parameter

(5). Thus, as a corollary to the above proposition, we have the following theorem:

Theorem 7 (Unique canonical pose from a symmetry group). *Suppose that a symmetric structure S admits a symmetry group G which contains a rotational or reflective subgroup that fixes only the origin of $\mathbb{R}^3$. Then the canonical pose g_0 can always be uniquely determined from one image of S.*

Note that the group G does not have to be the only symmetry that S allows—as long as such a G exists as a subgroup of the total symmetry group of S, one may claim uniqueness for the recovery of g_0.

Now one may wonder why the third column in Table 1 is necessary? Why do we have to treat a planar structure separately and why not treat more "uniformly" the planar restriction as an extra reflective symmetry with respect to the plane itself? The problem with that is, even if we could add this reflection, say R, into the symmetry group G, it is not possible to recover its corresponding element $R_0 R R_0^T$ in $R_0 G R_0^T$ via multiple-view geometric means because features on the plane correspond to themselves under this reflection and no other feature points outside of the plane is available (by our own planar assumption). Thus, we can no longer apply the above theorem to this case.

In order to give a correct statement for the planar case, for a reflection R with respect to a plane, we call the normal vector to its plane of reflection as the *axis of the reflection*.[18] Using this notion, we derive the following corollary:

Corollary 8 (Canonical pose from a symmetry group of a planar scene). *If a planar symmetric structure S allows a rotational or reflective symmetry subgroup G (without the reflection with respect to the supporting plane of S itself) with two independent rotation or reflection axes, the canonical pose g_0 can always be uniquely determined from one image of S (with the world frame origin o restricted in the plane and the z-axis chosen as the plane normal).*

As a consequence, to have a unique solution for g_0, a planar symmetric structure S must allow at least two reflections with independent axes, or one reflection and one rotation (automatically with independent axes for a planar structure). This is consistent with the ambiguity given in Table 1.

3.4. Camera Self-Calibration from Symmetry

In addition to structure and pose, another important knowledge for 3-D reconstruction is in fact readily available from images of symmetric objects: the *camera intrinsic parameters*, or calibration, which we will study in this subsection.

Results given in the preceding subsections, such as Theorem 7, are based on the assumption that the camera has been calibrated. If the camera is uncalibrated and its intrinsic parameter matrix, say $K \in \mathbb{R}^{3 \times 3}$, is unknown, then the Eq. (2) becomes

$$\lambda \boldsymbol{x} = K \Pi_0 g_0 \boldsymbol{X}.$$

From the epipolar constraints between pairs of hidden images, instead of the Essential matrix $E = \widehat{T}' R'$, we can only recover the Fundamental matrix

$$F \sim K^{-T} \widehat{T}' R' K^{-1}, \qquad (21)$$

where as before $R' = R_0 R R_0^T \in O(3)$ and $T' = (I - R_0 R R_0^T) T_0 + R_0 T \in \mathbb{R}^3$. Notice that here K is automatically the same for the original image and all the hidden ones.

3.4.1. Calibration from Translational Symmetry. In the translational symmetry case, we have $R' = I$ and $T' = R_0 T$. If given three mutually orthogonal

translations $T_1, T_2, T_3 \in \mathbb{R}^3$ under which the structure is invariant, then from the fundamental matrix $F \sim K^{-T}\widehat{T'}K^{-1} = \widehat{KT'} = \widehat{KR_0T}$ we get vectors

$$v_i \sim KR_0T_i, \quad i = 1, 2, 3. \tag{22}$$

That is, v_i is equal to KR_0T_i up to an (unknown) scale.[19] Since T_1, T_2, T_3 are assumed to be mutually orthogonal, we have

$$v_i^T K^{-T}K^{-1}v_j = 0, \quad \forall i \neq j. \tag{23}$$

We get three linear equations on entries of the matrix $K^{-T}K^{-1}$. If there are less than three parameters in K unknown,[20] the calibration K can be uniquely determined from these three linear equations (all from a single image). The reader should be aware that these three orthogonal translations correspond precisely to the notion of "vanishing point" or "vanishing line" used in existing literature (Criminisi and Zisserman, 1999, 2000).[21]

3.4.2. Calibration from Reflective Symmetry.
In the reflective symmetry case, if R is a reflection, we have $R^2 = I$ and $R'T' = -T'$. Thus, $\widehat{T'}R' = \widehat{T'}$. Then $F \sim \widehat{KT'}$ is of the same form as a fundamental matrix for the translational case. Thus, if we have reflective symmetry along three mutually orthogonal directions, the camera calibration K can be recovered similarly to the translational case.[22] So having a reflective symmetry in certain direction is as good as having a translational symmetry along that direction, as far as camera calibration is concerned.

3.4.3. Calibration from Rotational Symmetry.
A more interesting case where symmetry may help with self-calibration is the rotational symmetry. In this case, it is easy to show that the axis of the rotation R' is always perpendicular to the translation T'. According to (Ma et al., 2000), the Fundamental matrix F must be of the form

$$F = \lambda \hat{e}KR'K^{-1}, \tag{24}$$

where $e \in \mathbb{R}^3$ of unit length is the (left) epipole of F and the scalar λ is one of the two non-zero eigenvalues of the matrix $F^T\hat{e}$. Then the calibration matrix K satisfies the so-called normalized Kruppa's equation

$$FKK^TF^T = \lambda^2\hat{e}KK^T\hat{e}^T, \tag{25}$$

with F, e, λ known and only KK^T unknown. This equation, as shown in Ma et al. (2000), gives two linearly independent constraints on KK^T. For instance, if only the camera focal length f is unknown, we may re-write the above equation as

$$F\begin{bmatrix} f^2 & 0 & 0 \\ 0 & f^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}F^T = \widehat{\lambda e}\begin{bmatrix} f^2 & 0 & 0 \\ 0 & f^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}\widehat{\lambda e}^T, \tag{26}$$

which is a linear equation on f^2. It is therefore possible to recover focal length from a single image (of some object with rotational symmetry).

Example 9 (A numerical example).   For a rotational symmetry, let

$$K = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$R = \begin{bmatrix} \cos(2\pi/3) & 0 & -\sin(2\pi/3) \\ 0 & 1 & 0 \\ \sin(2\pi/3) & 0 & \cos(2\pi/3) \end{bmatrix},$$

$$R_0 = \begin{bmatrix} \cos(\pi/6) & \sin(\pi/6) & 0 \\ -\sin(\pi/6) & \cos(\pi/6) & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$T_0 = \begin{bmatrix} 2 \\ 10 \\ 1 \end{bmatrix}.$$

Note that the rotation R by $2\pi/3$ corresponds to the rotational symmetry that a cube admits. Then we have

$$F = \begin{bmatrix} -0.3248 & -0.8950 & -1.4420 \\ 0.5200 & 0.3248 & -2.4976 \\ -1.0090 & -1.7476 & -0.0000 \end{bmatrix},$$

$$\hat{e} = \begin{bmatrix} 0 & -0.4727 & 0.4406 \\ 0.4727 & 0 & 0.7631 \\ -0.4406 & -0.7631 & 0 \end{bmatrix},$$

and $\lambda = 2.2900$ (with the other eigenvalue -2.3842 of $F^T\hat{e}$ rejected). Then the equation (26) gives a linear equation in f^2

$$-0.2653 \cdot f^2 + 1.0613 = 0.$$

This gives $f \approx 2$ which is the focal length given by the matrix K in the first place.

Therefore, using symmetry, one can not only recover a 3-D symmetric structure, pose to the structure, but also obtain the camera intrinsic parameters such as the focal length of the camera (from only one image if one wishes). Furthermore, symmetry based calibration techniques happen to be always linear!

4.   Applications and Experiments

Several distinctive features allow symmetry-based techniques to be extremely useful and effective in a wide range of applications in computer vision and image processing. Firstly, by imposing *global* structural invariants such as symmetry, multiple-view geometric algorithms become much more accurate and robust. The reason is that the "baseline" of the relative poses between the "hidden" images is large most of the time. Secondly, since reconstruction can now be done within a single image, no more tracking or matching features across multiple images is needed.[23] This dramatically reduces the computational complexity, which is important in time-critical applications such as robot vision. Last but not the least, a typical man-made scene consists of numerous symmetric structures. These structures can be used as a new type of "primitives" when we try to establish correspondences across multiple views. These complex primitives, serving a similar role as landmarks, may simplify the matching task which is otherwise a more difficult problem for simple features such as points and lines.

Here we do not claim to be the first to suggest the applications given in this section—many of them have been studied separately before by researchers in the computer vision, image processing, or graphics community. We here mostly want to demonstrate that, now under our unified framework, these previously isolated case studies can all be put coherently together.

4.1.   Symmetry-Based Segmentation and Global Orientation

When we tour around a man-made environment, we typically have little problem to orient and locate ourselves since objects of regular shape, especially symmetric shapes, easily stand out and provide us overwhelming geometric information. In order for a machine vision system to emulate this functionality and to utilize the rich geometric information encoded in symmetric objects, we first need to know how to segment such objects from the image. Using the techniques introduced earlier in this paper, we can test whether certain image segments, obtained by other low-level segmentation algorithms such as mean shift (Comanicu and Meer, 2002), can be the perspective projection of symmetric objects in 3-D. Such objects can already be used as landmarks by the machine vision system when navigating in the man-made environment. Moreover, the segmented symmetric objects may serve as primitives for other higher-level applications such as large baseline matching (See Section 4.2).

In the symmetry-based segmentation method, the input image will be first segmented into polygonal regions which may or may not be a perspective projection of symmetric objects in 3-D. Then the regions which potentially can be interpreted as images of symmetric objects in space will be tested against each fundamental type of symmetries described in Section 3.2. Here we show via an example. In Fig. 11, each region in the image enclosed by four connected line

Figure 11.   Examples of symmetry-based segmentation of rectangles. Rectangular regions are segmented and attached with local coordinate frames which represent their orientations. The segmentation does not need any human intervention.

segments may be interpreted as a 3-D quadrilateral that falls into the following cases: 1. quadrilateral (no symmetry) 2. trapezia (one reflective symmetry). 3. rectangle (two reflective symmetries). 4. square (two reflective symmetries + three rotational symmetries). We can test this region against the four hypotheses by using the reconstruction methods in Section 3.2 and determine to which type of symmetry it belongs. Most commonly, symmetric regions in man-made world are rectangles and squares. A region is simply discarded if it can only be interpreted as a quadrilateral or trapezium in 3-D. The symmetry-based testing will eliminate most of the useless segments and keep only symmetric ones. The resulting segments will be labeled by their type of symmetry, position, and orientation (the normal vector of the region in space). These output segments encode much richer 3-D geometric information than simple features such as corners and edges, as shown in Fig. 11.

After the segmentation, we can further apply clustering algorithms to classify the symmetric objects in terms of their orientations and positions. For instance, if a number of neighboring symmetric objects have mutually consistent normal vectors (e.g., the five panels in the left image of Fig. 11), it is likely that these objects come from the same plane. Such a plane provides a "global" reference of orientation for the agent that is navigating through the environment. More details about symmetry-based segmentation and classification can be found in Yang et al. (2003).

4.2. Large-Baseline Matching and Reconstruction

Matching features across multiple views is a key step for almost all reconstruction algorithms based on multiple-view geometry. Conventional methods for feature matching mostly rely on tracking features between images with small motions. However, the 3-D geometry recovered from images with small baseline is severely inaccurate due to a small signal-to-noise ratio (SNR). For images with large baseline, some methods such as RANSAC (Fishler and Bolles, 1981) and LMeDs (Zhang et al., 1995) use robust statistic techniques to iteratively match features and estimate the geometry. However, such techniques usually require good initialization. In these conventional methods, normally points and lines are used as features to match. As we have mentioned earlier, the symmetric regions that we can obtain from symmetry-based segmentation encode full geometric/textural information of their 3-D coun-

terparts. Hence they can be more easily identified and matched across different views.

Matching a symmetric object across two images induces less ambiguity than matching a points or a line because not only the position and orientation but also the shape and texture of the symmetric object need to be matched. Although there may be more than one possible matching for a single symmetric object, there is typically a unique and unambiguous matching that is consistent for all the symmetric objects in the scene. A systematic solution to matching symmetric objects in multiple images is given in a related paper (Huang et al., 2004), which converts the matching problem to one of identifying the maximal complete subgraphs, the cliques, of a matching graph.

As the experiment in Fig. 12 shows, the symmetry-based matching algorithm is effective even in the presence of both small and large baseline, in which case existing techniques normally fail: the baseline between the first and second images is large and the baseline between the second and the third image is almost zero. The camera pose and geometry of the symmetric structures are accurately reconstructed using the symmetry-based method given in the preceding section. The ground truth for the length ratios of the white board and table are 1.51 and 1.00, and the recovered length ratio are 1.506 and 1.003, respectively. Error in all the right angles is less than $1.5°$.[24]

In addition to rectangular shapes, other more complex symmetric shapes can also be reconstructed accurately using the method given in this paper. For example, Hong et al. (2004) has demonstrated that even smooth symmetric curves without any distinctive feature can be reconstructed accurately from their perspective images. A more detailed study of the numerical performance of the symmetry-based reconstruction algorithms under different levels of random noises and different view angles can also be found in that paper.

4.3. Symmetry-Based Photo Editing

The symmetry-based methods introduced earlier can also be used for editing photos, which is an important problem in digital photography and graphics. But conventional photo editing softwares such as Photoshop usually are lack of mechanisms that may preserve the perspective projection of geometric objects. It is then very desirable if we are able to edit the photos while automatically preserving the perspective geometric
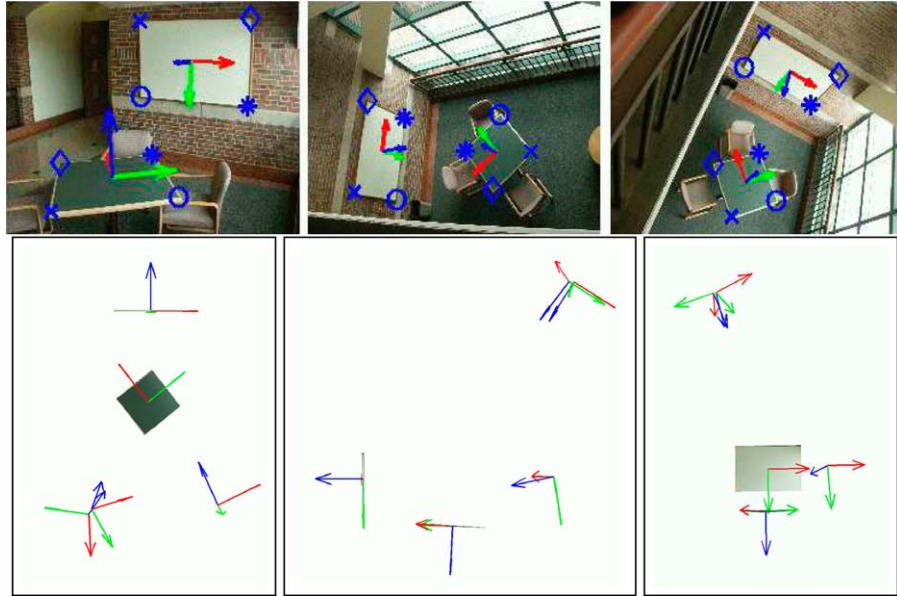
Figure 12. Top: Two symmetry objects matched in three images. From the raw images, symmetry objects segmentation and matching do not need any manual intervention. Bottom: Camera poses and structures of symmetric objects are recovered. From left to right: top, side, and frontal views of the matched and reconstructed objects and the camera poses.

relations of objects in the scene. With the knowledge in symmetry, it is possible for us to manipulate 2-D image regions based on their 3-D shapes and relationships. For example, many photo editing operations are based on the "copy-and-paste" function. The conventional "copy-and-paste" simply copies a region pixelwise onto a new location. Artifacts become easily noticeable if perspectiveness is distorted or destroyed by the editing. The symmetry-based methods allow us to manipulate a region realistically as if manipulating an object in 3-D without explicitly reconstructing the 3-D structure of the object.

With very little human intervention, we may first register symmetric objects of interest in one or multiple images, and recover their relative positions and orientations using the method given earlier. The transformation of such a region from one place or one image to another can therefore be easily performed. Figure 13 shows an comprehensive example of symmetry-based photo editing, which includes removing occlusion, copying and replacing objects in the scene, and adding new objects. More details and examples of symmetry-based photo editing can be found in a related paper (Huang et al., 2003).
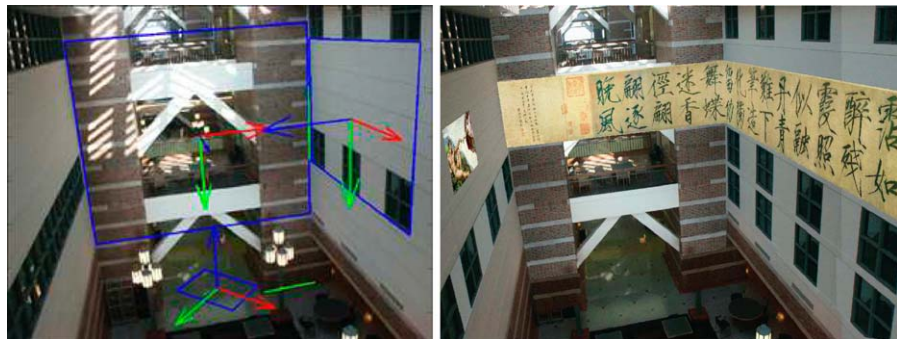

Figure 13. An example of photo editing. Left: The original picture with some symmetric regions registered. Right: The shadows of the roof on the frontal wall and the occlusions by the lamp poles are removed using symmetry-based "copy-and-paste." Some paintings are pasted on the side walls according to the correct perspectiveness. Additional windows are added on the side walls.
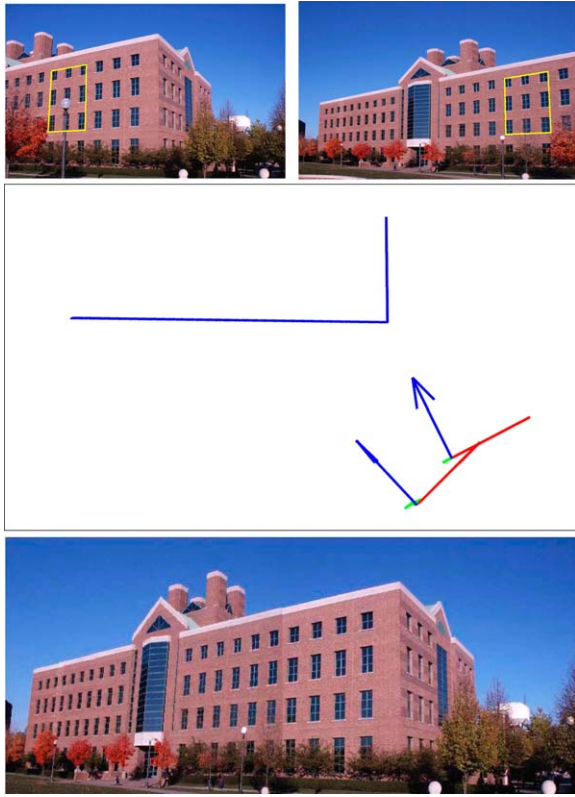
Figure 14. Mosaic of the two images on the top using corresponding symmetry objects in the scene (in this case windows on the front side of the building). The middle picture is a bird view of the recovered 3-D shape of the two sides of the building and the camera poses (the two coordinate frames). Notice that perspectiveness is precisely preserved for the entire building in the final result (bottom image).

Figure 14 shows another example of generating panorama or image mosaics from multiple images based on ideas similar to the symmetry-based photo editing. Conventional panorama approaches usually require the images are taken with a fixed camera center or the camera positions are known. Using our methods, the camera poses can be easily obtained as a "by-product" when we align the symmetric objects in different images.

4.4. *Visual Illusions and Symmetry*

Another interesting application of our computational theory of symmetry is to help understand certain psychological aspects of human vision. In previous sections, we have shown that if an 3-D object is indeed symmetric, multiple-view geometric techniques enable us to easily and accurately retrieve its 3-D geometry from even a single image. One might wonder if the same geometric assumption is strongly held up by our human vision system. More specifically, one might want to know: *Is symmetry a fundamental assumption adopted by human visual perception?*

Compelling examples from psychological studies of human vision give us sufficient confidence that this hypothesis is true. Ironically, the best way to demonstrate this is not by showing how useful symmetry is for human visual perception, but by showing certain negative "side effects" of the symmetry assumption on human perception. As we will see, this also serves as a ground for many famous visual illusions. The following two examples demonstrate two aspects of such side effects.

The first example is to show that human vision system often reconstructs 3-D geometry by imposing on the scene the assumption that "regular" objects are likely to be symmetric. Especially when parallax (or stereo) information is not available, this assumption becomes dominant. The Ames room, Fig. 15 (top), is a controlled experiment which deliberately excludes the parallax information from the viewer. Obviously, our human perception tends to assume that the frames of the windows and the tiles on the ground are rectangles or squares of the same size, only "properly" distorted by the perspective.

On the other hand, an illusion can also be created in such a way that it contains more than one set of symmetries which however are not compatible with each other. The consequence is to throw our visual perception into a dizzy dilemma. The Escher waterfall, Fig. 15 (bottom), is one good example for this type of illusion. If we only focus on the water tunnel, it has a horizontal symmetry along the y-axis, and the two towers also have their own vertical symmetry along the z-axis. A contradiction reaches at the intersections of the tunnel and the towers. Under these two different symmetries along different axes, it would be impossible for the tunnel and the towers to intersect on the locations as shown in the figure. Nevertheless, since the symmetry assumption is only applied locally to individual symmetric part, an illusion is created once the viewer trying to put everything back together globally.

Examples above show that symmetry is indeed one of many fundamental assumptions adopted by our human visual system. A good understanding of this assumption will help us improve the design of computer vision systems. Symmetric patterns or structures in one
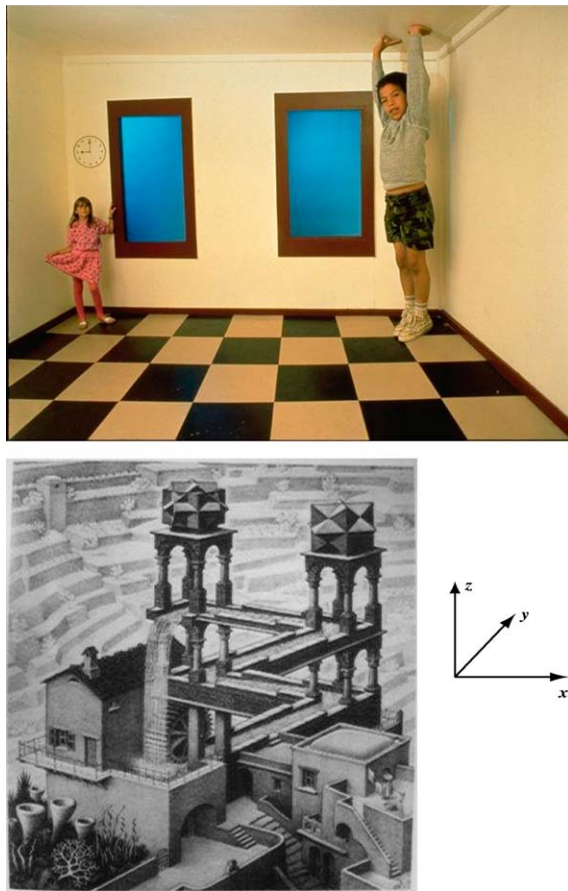
Figure 16. Necker's Cube. The illusion is caused by orthographic projection. The cube can be interpreted either concave or convex.

Figure 15. Top: Ames room. Ames room is located in San Francisco Exploratorium (www.exploratorium.edu). This room actually has no square corners, but carefully designed so as to exploit the nature of human vision, and viewer will get a wrong 3-D perception by looking into the room from this vantage point (only). Bottom: Escher's waterfall. It is a 2-D drawing of a bizzare waterfall by Maurits C. Escher. When we focus on the tunnel, it guides the water from low close to high far from the waterfall, but if we focus on the tower, it actually raises the tunnel directly from low to high along the waterfall.

single image may give us the same geometric information as multiple images only if our assumption about their symmetry is correct; otherwise, multiple images of the same object will be needed, in order to resolve ambiguities or illusions such as the ones shown above. The study in this paper provides a computational basis for the cause of such illusions.

Perspective Versus Orthographic Projection. Last but not the lest, despite the above illusions, our theory shows that in general, under perspective projection, reconstruction from symmetry is hardly ambigu-
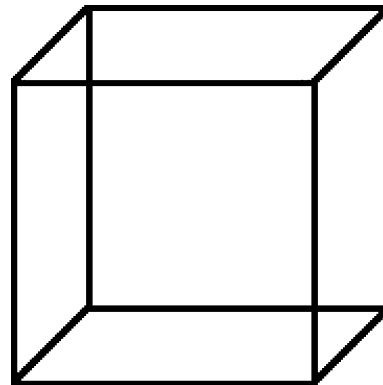
ous (at least locally) since perspective projection encodes depth information properly. However, this may no longer be true under other types of projection. Figure 16 shows an example of the famous Necker's Cube illusion, from which the viewer can easily identify two equally probable interpretations. The reason is that the drawing of the cube according to orthographic projection eliminates any cue for depth information. But this ambiguity hardly exists for a perspective image of a cube. This example shows that the effect of (assuming) symmetry under different projection models can be significantly different. But a more detailed comparative study is beyond the scope of this paper.

5. Discussions and Conclusions

Points, lines, and planes are special symmetric objects which have been extensively studied as primitive geometric features for reconstructing a 3-D scene from 2-D images. This paper begins to look at primitives at a higher level (hence with a larger scale): primitives which have certain "regular" structure, with the regularity characterized by certain symmetries that the structure admits. Obviously this is a more principled way to study assumptions about 3-D structure that people have exploited before in multiple-view geometry, such as orthogonality and parallelism (hence vanishing points) etc. Our experimentation demonstrates that, at this level of generalization, multiple-view geometric algorithms become extremely well-conditioned due to almost always large baseline between "hidden images." Because of this, nonlinear optimization (such as bundle adjustment) is not really needed to improve the results.

Since reconstruction can now be done reliably from even a single image, feature correspondence becomes less of a problem and a proper permutation of features from the same image is often all that an algorithm requires.

Probably the most important observation from this paper is that, in addition to the 3-D structure, the "canonical" pose between the canonical world coordinate frame of a symmetric object and the camera can also be recovered. This extra piece of information is extremely useful and allows us to achieve many new tasks which used to be difficult to do with only points, lines and planes. For example, in an environment rich with symmetric objects, a viewer can easily identify its location and orientation by referencing with these canonical frames. Furthermore, such information can be readily utilized to establish correspondence across images taken with a large baseline or change of view angle: As long as one common (local) symmetry can be recognized and aligned properly, the rest of the structures in the scene can then be correctly registered and reconstructed. The remaining problem is mostly an engineering one of how to efficiently register numerous symmetric parts and objects present in a scene and to obtain a consistent 3-D reconstruction. We believe that, together with conventional geometric constraints among multiple images, symmetry is indeed an important cue which eventually makes 3-D reconstruction a more well-conditioned problem.

In this paper, we have only shown how to perform reconstruction if we know the type of symmetry in advance. But a more challenging problem is how to identify what type of symmetry a structure admits from its (perspective) images when the symmetry is not known a priori. This leads to:

The inverse problem *that, given a* (*perspective*) *image of a symmetric structure S with an* unknown *symmetry group taken at some* (*unknown*) *vantage point, how to find a maximum subgroup G of the permutation group $\mathfrak{S}$ of S which can be represented as a subgroup of E(3) through a group isomorphism $\mathcal{R} : g \mapsto (R, T)$ and the representation is consistent with the symmetric multiple-view rank condition?*

Notice that this is a constrained group representation problem subject to the multiple-view rank condition. Although solutions to simple symmetric structures can be found already (e.g., the symmetry-based segmentation example described in Section 4.1), this problem, at this point, remains largely open for general symmetric structures, in both 2-D and 3-D.

Appendix A: Symmetric Structures and Associated Symmetry Groups

As we have mentioned in the problem formulation section, the quest to understand various types of symmetry and their relationships has a history of more than two thousand years. But not until the 20th century, with the aid of modern algebra, did mathematicians formulated it as a formal mathematical problem, which is ever since known as Hilbert's 18th problem:

> *Is there in n-dimensional Euclidean space also only a finite number of essentially different kinds of groups of motions with a fundamental region?*

A definite answer to this problem was given in Bieberbach (1910). In this appendix, we give a short survey of mathematical facts on the subject of group actions on symmetric structure. For applications in computer vision, we will primarily survey results related to symmetric structure and groups in 3-D Euclidean space. All facts and statements will be given without proofs, and interested readers may refer to Weyl (1952), Grünbaum and Shephard (1987) and Martin (1975).

A.1. Group Actions on Symmetric Structures

Definition 10 (Isometry). An isometry is a mapping from the set of points *S* into itself (i.e. an automorphism) which preserves distance.

Theorem 11. *An isometry is a distance preserving automorphism, and the set of all isometries forms a group $(G, \circ, e)$, where e is the identity mapping (or element) in G.*

Regarding isometry in any *n*-dimensional Euclidean space $\mathbb{E}^n$, we have the following facts:

Proposition 12. *Every isometry of the Euclidean space $\mathbb{E}^n$ is one of the following three fundamental types*[25] (*Fig.* 17):

1. Rotation *about a point o around an axis by a given angle θ. The point o is called the center of rotation.*
2. Reflection *in a given hyperplane P.*
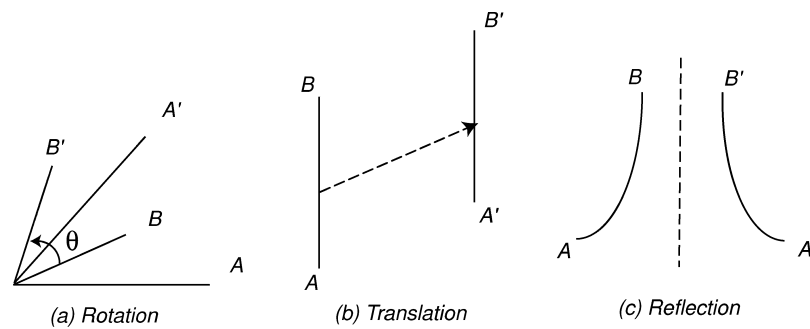3. Translation *in a given direction by a given distance.*

Figure 17.   The three basic types of isometric transformations: rotation, translation and reflection.

Figure 18.   A band ornament pattern (frieze pattern). It is an example of 1-D symmetric patterns.

A.2.   *Symmetry in a 1-D or 2-D Euclidean Space*

A.2.1.   *Groups Associated to 1-D and 2-D Symmetry.*
One special 1-D symmetric pattern is a band ornament pattern, also called a frieze pattern (Fig. 18). A real band ornament is not strictly 1-dimensional, but its symmetry often makes use of one of its dimensions. The only possible types of symmetry for band ornaments are reflection, translation and their combinations. The (bilateral) reflective symmetry is pervasive in both art and nature, which is also conspicuous in the structure of man-made objects including buildings and machines (Weyl, 1952). Any reflection symmetry is isomorphic to the group $\mathbb{Z}_2$. That is if we denote a reflection by r, we always have $r^2 =$ identity. A translation may be represented by a vector t. A pattern invariant under the translation t is also invariant under its iterations $t^1, t^2, t^3, \ldots$ and under the inverse t^{-1} of t and its iterations, which give rise to an infinite group. It is not hard to show that if a 1-D pattern admits both translation and reflection symmetry, the reflection may only occur at locations which are multiples of $\frac{1}{2}t$.

If we take a band ornament where we repeat its individual section again and again and sling it around a circular cylinder, we obtain a pattern which is in 3-D space but has a rotational symmetry in 2-D—we only have to pay attention to the cross-section of the cylinder. Any finite group of rotations contains a primitive rotation operation r around some axis by an angle which must be of the form $\theta = \frac{2\pi}{n}$ for some $n \in \mathbb{Z}_+$, and its iterations $r^1, r^2, \ldots, r^n =$ identity. We know this forms a cyclic group of order n, conventionally denoted a C_n. It is isomorphic to the Abelian group $\mathbb{Z}_n$ and the order of the group n completely characterizes the group. We may also generalize the cyclic symmetry to other 3-D objects. Then the cylinder may also be replaced by any surface or structure with cylindrical symmetry, namely by one that is carried into itself by a rotation around certain axis.

Rotational (or cyclic) symmetry does not change the orientation of an object. If reflections, also called *improper rotations*, are taken into consideration, we have the *dihedral group* of order n, denoted as D_n, and D_n is isomorphic to $\mathbb{Z}_n \times \mathbb{Z}_2$. One can think of this is a (cyclic) rotation group C_n combined with the reflections in n lines forming angles of $\frac{1}{2}\theta$. A reflection changes the orientation of the object it is applied to. Finally, we have *Leonardo's Theorem* (Martin, 1975).

Theorem 13 (Leonardo's Theorem). *The only finite groups of isometry in a* 2-*D plane are the cyclic groups C_n and the dihedral groups D_n.*

A.2.2. Groups Associated to Tiling of a 2-D Euclidean Space.

Although Leonardo's Theorem completely characterizes finite groups of 2-D symmetry, Hilbert's 18th problem, i.e. exhaust all possible realizations of the rotational and reflective symmetry together with the translational symmetry as 2-D lattice patterns (of a fundamental region), or the so-called tiling, turns out to be a much more difficult problem.[26]

The difference between a tiling pattern and general 2-D patterns is that a planar fundamental region (a "tile") fills the entire 2-D plane by its congruent pieces without gaps or overlaps (e.g., the 4th image of Fig. 1).

Definition 14. A plane tiling $\mathcal{T}$ is a countable family of closed congruent sets $\mathcal{T} = \{T_1, T_2, \dots\}$ which cover the plane $\mathbb{E}^2$ without gaps or overlaps.

The countability condition excludes families in which a tile is allowed to have zero area. From the definition, we see that the intersection of any finite set of tiles of $\mathcal{T}$ necessarily has zero area, such an intersection may be empty or may consist of a set of isolated points and arcs. In these cases the points will be called *vertices* of the tiling and the arcs will be called *edges*. A special case of tiling is that each tile is a polygon.

Let us start our discussion from the group of translations, because this operation is the most conspicuous in 2-D tilings. We have mentioned before the translations form a group. In general, the groups of reflections and rotations do not obey the commutative law, but the combination of translations is commutative, that is to say, the group of translations is *Abelian*.

If a group G contains translations, G will not be finite, but can still be discontinuous. For such a group there are three possibilities (Weyl, 1952): Either it only consists of identity; or all the translations in the group are iterations t^n of one basic translation $t \neq$ identity; or these translations form a 2-dimensional *lattice* $\{x_1 t_1 + x_2 t_2 :$ where $x_1, x_2 \in \mathbb{Z}\}$ and the independent vectors t_1, t_2 form a *lattice basis*.

At least 17 different ways of tiling were known to ancient Egyptians. However, it remained as a mystery for thousands of years that if these were the only possibilities. In 19th century (Fedorov, 1885, 1891a, 1891b, 1971) gave the first mathematical proof that this is indeed true which solved Hilbert's 18th problem for the 2-D and 3-D cases. [27] It was pointed out that for proper and improper rotation groups, the only possible choices in 2-D plane allowed by tiling are C_1, C_2, C_3, C_4, C_6 and D_1, D_2, D_3, D_4, D_6 since when $n = 5, 7,$ or $n \geq 8$, such an n-gon can not tile the whole plane without gaps (Weyl, 1952). These 10 groups can further be realized as 13 different types of 2-D lattices which are unimodularly inequivalent. Finally, by taking into account translations which have to be compatible with these lattices, there are essentially in total only 17 different ways to tile a 2-D plane. Interested readers can find a full description of all the 17 patterns in Grünbaum and Shephard (1987).

A.3. Symmetry in a 3-D Euclidean Space

Similar analysis can be applied to study groups associated to symmetry or tiling in a 3-D Euclidean space. One surprising result in the 3-D space is that although in a 2-D plane, there is for every number $n > 2$ a regular polygon of n sides, in 3-D space there only exist five regular polyhedra, often called the Platonic solids (see Fig. 19). Based on this, one can show that any group of rotational and reflective symmetry around a center o in 3-D can be generated by C_n, D_n, and three additional isometric (rotational) groups associated to the five Platonic solids (Fedorov, 1971; Weyl, 1952; Bulatov, www).[28]

Among all these rotational and reflective groups, only 32 of them are allowed in any 3-D lattice, and they give rise to a total of 70 different realizations of
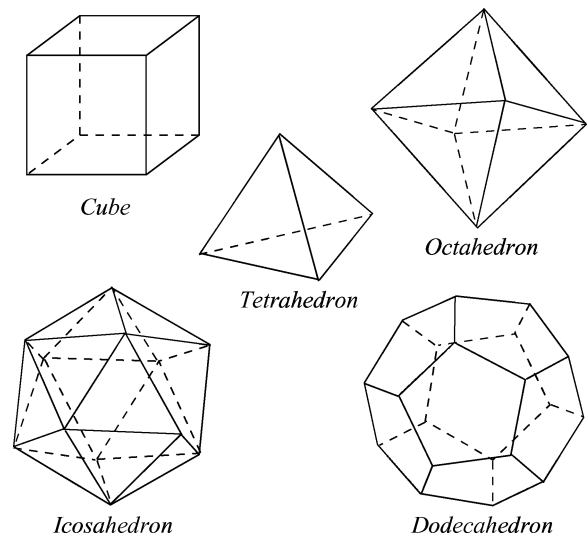

Figure 19. Five Platonic solids: cube, tetrahedron, octahedron, icosahedron and dodecahedron.

lattice. Together with proper translational components, there are a total of 230 different ways to tile a 3-D Euclidean space with a fundamental region (a tile).

Acknowledgments

Notes

1. In fact, there are strong psychological evidences to believe that symmetry is one of the most effective assumptions that humans adopt to process visual information. We will have a more thorough discussion on this topic at the end of this paper after we have understood better the geometric basis for it.
2. Whereas most existing studies of symmetry are based on orthogonal or affine camera models, as approximation to the perspective case.
3. The phenomenon with "multiple images" from a single one was already pointed out by Malik and Rosenholtz (1997). But we here will make a more geometric use of these images. For the lack of better words, we will call such images as "hidden." In fact, they are probably better be called as "ulterior" or "covert" images, or "equivalent" images.
4. One may argue that they are images taken from behind the board. This is true if the board is "transparent." If the symmetric object is a 3-D object rather than a 2-D plane, such a argument will nevertheless fall apart.
5. Egyptians certainly knew about all 17 possible ways of tiling the floor; and Pythagoras already knew about the five platonic solids which are the only non-trivial ones allowing 3-D rotational symmetry, that were only proven by mathematicians in the 19th century (Fedorov, 1885, 1891a, 1891b, 1971).
6. Here $O(3)$ denotes the group of 3×3 orthogonal matrices including both rotations ($SO(3)$) and reflections.
7. In Appendix A, we give explanation to why this assumption is valid. However, it will be interesting to consider in the future generalized notion of "symmetry" that is captured by a continuous or even non-rigid group.
8. Here we have effectively identified, through a group homomorphism, G as a subgroup of the permutation (or automorphism) group $\mathfrak{S}$ of the set $\mathcal{I}_0$ when $\mathcal{I}_0$ is a finite (or infinite) set.
9. Here we use $|G|$ to denote the cardinality of G. In particular, when G is finite, $|G|$ is the number of elements in G.
10. We use the symbol "$\sim$" to denote "equal up to a scale."
11. In the matrix, we use the notation $\hat{u}$ to denote the skew symmetric matrix associated to a vector $u \in \mathbb{R}^3$ such that $\hat{u}v = u \times v$ for all $v \in \mathbb{R}^3$.
12. Here "only" is in the sense of sufficiency: if a set of features satisfies the rank condition, it can always be interpreted as a valid image of an object with the symmetry G.
13. It is possible that both $|G$ and $|S|$ are infinite. In practice, one can conveniently choose only finitely many of them.
14. In some literature, such equations are also called Sylvester equations.
15. In fact, E is equal to $\widehat{T}'R' = \widehat{T}'$, which is a skew-symmetric matrix and the extraction of R' and T' from such an E is very much simplified. Furthermore, to recover such an E from the epipolar constraint, only 4, instead of 8, corresponding points are needed.
16. The planar homography is also a direct consequence of the (symmetric) multiple-view rank condition (Ma et al., 2002), applied to planar scene.
17. In general, one gets a total of four solutions from the decomposition, but only two of them are physically possible (Weng et al., 1993).
18. The role of the axis of a reflection is very similar to that of the axis of a rotation once we notice that, for any reflection R, $-R$ is a rotation of angle $\theta = \pi$ about the same axis.
19. The reader may check that v_i can also be obtained from the homography matrix if the scene is planar.
20. For example, pixels are square.
21. Here we see, symmetry is a more principled way to unify in computer vision notions like "orthogonality," "parallelism," or "vanishing point" without using points or plane at infinity.
22. Notice that in both the translational and reflective symmetry cases, for planar structures, one can obtain the vector KT' from now the "uncalibrated" homography matrix $H = K(I + T'N^T)K^{-1} = I + KT'(K^{-T}N)^T$.
23. Although matching, or more precisely permutation, of features is still required within a single image, since the techniques work very well with only a small number of features, matching often becomes relatively easier to do in a single image once the type of symmetry is specified.
24. Results from point- or line-based multiple-view algorithms give much larger errors.
25. In fact, all three types can be expressed in terms of product of (at most $n + 1$) reflections.
26. The problem of tiling in non-Euclidean (elliptic and hyperbolic) spaces turns out to be much easier mathematically, due to the work of Jordan, Fricke, Klein et al., as was surveyed by Hilbert when he proposed his 18th problem in 1901.
27. A general solution to Hilbert's 18th problem was not solved till 1910 by Bieberbach (1910).
28. The cube and the octahedron share the same isometry group, so do the icosahedron and the dodecahedron.

References

A. Criminisi, I. and Zisserman, A. 1999. Single view metrology. In *IEEE International Conference on Computer Vision*, pp. 434–441.

A. Criminisi, I. and Zisserman, A. 2000. Single view metrology. *Int. Journal on Computer Vision*.

Bieberbach, L. 1910. Uber die bewegungsgruppen des n-dimensionalen euklidischen raumes mit einem endlichen fundamentalbereich. *Nachr*, 75–84.

Bulatov (www). Polyhedra collection at http://www.physics.orst.edu/~bulatov/polyhedra/index.html.

Carlsson, S. 1998. Symmetry in perspective. In *European Conference on Computer Vision*.

Cham, T.J. and Cipolla, R. 1996. Geometric saliency of curve correspondances and grouping of symmetric comntours. In *European Conference on Computer Vision*, pp. 385–398.

Comanicu, D. and Meer, P. 2002. Mean shift: A robust approach toward feature space analysis. In *IEEE Transactions on Pattern Analysis & Machine Intelligence*, pp. 603–619.

Fedorov, E.S. 1885. The elements of the study of figures. *Zapiski Imperatorskogo S. Peterburgskogo Mineralogichesgo Obshchestva [Proc. S. Peterb. Mineral. Soc.]*, 2(21):1–289.

Fedorov, E.S. 1891a. Symmetry in the plane. *Zapiski Imperatorskogo S. Peterburgskogo Mineralogichesgo Obshchestva [Proc. S. Peterb. Mineral. Soc.]*, 2(28):345–390.

Fedorov, E.S. 1891b. Symmetry of finite figures. *Zapiski Imperatorskogo S. Peterburgskogo Mineralogichesgo Obshchestva [Proc. S. Peterb. Mineral. Soc.]*, 2(28):1–146.

Fedorov, E.S. 1971. *Symmetry of Crystals*. Translated from the 1949 Russian Edition. American Crystallographic Association.

Fishler, M. and Bolles, R. 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395.

Garding, J. 1992. Shape from texture for smooth curved surfaces in perspective projection. *Journal of Mathematical Imaging and Vision*, 2(4):327–350.

Garding, J. 1993. Shape from texture and contour by weak isotropy. *J. of Artificial Intelligence*, 64(2):243–297.

Gibson, J. 1950. *The Perception of the Visual World*. Houghton Mifflin.

Gool, L.V., Moons, T., and Proesmans, M. 1996. Mirror and point symmetry under perspective skewing. In *Int. Conference on Computer Vision & Pattern Recognition*, pp. 285–292.

Grünbaum, B. and Shephard, G.C. 1987. *Tilings and Patterns*. W. H. Freeman and Company.

Hong, W., Ma, Y., and Yu, Y. 2004. Reconstruction of 3-D deformed symmetric curves from perspective images without discrete features. In *European Conference on Computer Vision*.

Huang, K., Hong, W., and Ma, Y. 2003. Symmetry-based photoediting. In *9th International Conference on Computer Vision, Workshop on Higher Level Knowledge*, Nice, France.

Huang, K., Hong, W., Yang, A.Y., and Ma, Y. 2004. Large baseline matching and reconstruction from symmetry cells. In *IEEE Int. Conference on Robotics & Automation*.

Kanatani, K. 1997. Comments on symmetry as a continuous feature. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 19(2):246–247.

Kiryati, N. and Gofman, Y. 1998. Detecting symmetry in grey level images: The global optimization approach. *Int. Journal on Computer Vision*, 29(1):29–45.

Leung, T. and Malik, J. 1997. On perpendicular texture or: Why do we see more flowers in the distance? In *Int. Conference on Computer Vision & Pattern Recognition*, pp. 807–813.

Liebowitz, D. and Zisserman, A. 1998. Detecting rotational symmetries using normalized convolution. In *Int. Conference on Computer Vision & Pattern Recognition*, pp. 482–488.

Liu, J., Walker, E., and Mundy, J. 1995. Characterizing the stability of 3D invariants derived from 3D translational symmetry. In S.Z. Li, D.P. Mital, E.K. Teoh, and H. Wang (eds.), *Recent Developments in Computer Vision*, Springer, Berlin, pp. 415–424.

Liu, Y. and Colline, R. 1998. Frieze and wallpaper symmetry groups classification under affine and perspective distortion. Technical Report, CMU-RI-TR-98-37, The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA.

Ma, Y., Košecká, J., and Huang, K. 2002. Rank deficiency condition of the multiple-view matrix for mixed point and line features. In *Asian Conference on Computer Vision*.

Ma, Y., Košecká, J., Soatto, S., and Sastry, S. 2003. *An Invitation to 3-D Vision: From Images to Geometric Models*. Springer-Verlag.

Ma, Y., Vidal, R., Košecká, J., and Sastry, S. 2000. Kruppa's equations revisited: Its degeneracy, renormalization, and relations to chirality. In *European Conference on Computer Vision*.

Malik, J. and Rosenholtz, R. 1997. Computing local surface orientation and shape from texture for curved surfaces. *Int. Journal on Computer Vision*, 23:149–168.

Marola, G. 1989. On the detection of the axes of symmetry of symmetric and almost symmetric planar images. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 11(1):104–108.

Marr, D. 1982. *Vision*. W. H. Freeman and Company.

Martin, G.E. 1975. *The Foundations of Geometry and the Non-Euclidean Plane*. Springer-Verlag.

Mitsumoto, H., Tamura, S., Okazaki, K., and Fukui, Y. 1992. 3-D reconstruction using mirror images based on a plane symmetry recovering method. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 14(9):941–946.

Mukherjee, D.P., Zisserman, A., and Brady, J.M. 1995. Shape from symmetry—Detecting and exploiting symmetry in affine images. *Phil. Trans. Royal Soc. London A*, 351:77–106.

Plamer, S.E. 1999. *Vision Science: Photons to Phenomenology*. The MIT Press.

Rosenholtz, R. and Malik, J. 1997. Surface orientation from texture: Isotropy or homogeneity(or both)? *Vision Research*, 37(16):2283–2293.

Rothwell, C.A., Forsyth, D.A., Zisserman, A., and Mundy, J.L. 1993. Extracting projective structure from single perspective views of 3D point sets. In *IEEE International Conference on Computer Vision*, pp. 573–582.

Schaffalitzky, F. and Zisserman, A. 2000. Planar grouping for automatic detection of vanishing lines and points. *Image and Vision Computing*, 18:647–658.

Sun, C. and Sherrah, J. 1997. 3D symmetry detection using the extended gaussian image. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 19(2):164–169.

Vetter, T. and Poggio, T. 1994. Symmetric 3D objects are an easy case for 2D object recognition. *Spatial Vision*, 8:443–453.

Weng, J., Huang, T.S., and Ahuja, N. 1993. *Motion and Structure from Image Sequences*. Springer Verlag.

Weyl, H. 1952. *Symmetry*. Princeton Univ. Press.

Witkin, A.P. 1988. Recovering surface shape and orientation from texture. *J. of Artificial Intelligence*, 17:17–45.

Yang, A.Y., Rao, S., Huang, K., Hong, W., and Ma, Y. 2003. Geometric segmentation of perspective images based on symmetry groups. In *IEEE International Conference on Computer Vision*.

Zabrodsky, H., Peleg, S., and Avnir, D. 1995. Symmetry as a continuous feature. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 17(12):1154–1166.

Zabrodsky, H. and Weinshall, D. 1997. Using bilateral symmetry to improve 3D reconstruction from image sequences. *Comp. Vision and Image Understanding*, 67:48–57.

Zhang, Z., Deriche, R., Faugeras, O., and Luong, Q.-T. 1995. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78(1/2):87–119.