



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Computer Vision and Image Understanding 99 (2005) 210–240

Computer Vision
and Image
Understanding

www.elsevier.com/locate/cviu

Symmetry-based 3-D reconstruction from perspective images[☆]

Allen Y. Yang^{*}, Kun Huang, Shankar Rao, Wei Hong, Yi Ma

*Coordinated Science Laboratory, University of Illinois at Urbana-Champaign,
1308 W. Main, Urbana, IL 61801, USA*

Received 14 April 2004; accepted 31 January 2005

Available online 10 March 2005

Abstract

Symmetry is an important geometric cue for 3-D reconstruction from perspective images. In this paper, we introduce a unified theoretical framework for extracting poses and structures of 2-D symmetric patterns in space from calibrated images. The framework uniformly encompasses all three fundamental types of symmetry, i.e., reflective, rotational, and translational, based on a systematic study of the homography groups in image induced by the symmetry groups in space. We claim that if a planar object admits rich enough symmetry, no 3-D geometric information is lost through perspective imaging. Based on two fundamental principles that utilize common spatial relations among symmetric objects, we have developed a prototype system which can automatically extract and segment multiple 2-D symmetric patterns present in a single perspective image. The result of such a segmentation is a hierarchy of new geometric primitives, called symmetry cells and complexes, whose 3-D structures and poses are fully recovered. Finally, we propose a new symmetry-based matching technique, which can effectively establish correspondences among the extracted symmetry cells across multiple images. We demonstrate the application of the proposed algorithms on image segmentation, matching, and 3-D reconstruction with extensive experimental results. The algorithms and systems are more accurate and easier to implement than existing point- or line-based methods.

Published by Elsevier Inc.

[☆] This work is supported by UIUC ECE/CSL startup fund and NSF Career Award IIS-0347456.

^{*} Corresponding author. Fax: +1 217 2442352.

E-mail addresses: yangyang@uiuc.edu (A.Y. Yang), kunhuang@uiuc.edu (K. Huang), srrao@uiuc.edu (S. Rao), weihong@uiuc.edu (W. Hong), yima@uiuc.edu (Y. Ma).

Keywords: Structure from symmetry; Symmetry group; Homography group; Symmetry-based matching; 3-D reconstruction

1. Introduction

Reconstruction of scene structures and camera poses from one or multiple images is a fundamental problem in computer vision literature. In the most extant methods, scene structures and camera poses are reconstructed based on structure from motion techniques (e.g., [1]) by exploiting geometric relations of point and line features among images, such as the epipolar constraint. However, we observe that humans have remarkable ability in inferring 3-D information from even single 2-D perspective images, which suggests the existence of other, probably more effective, cues for retrieving geometric information from perspective images. For example, in Fig. 1, people can effortlessly find the most probable 3-D interpretations for the first two images, and the elaborate effort to generate a “regular” 2-D perspective image with irregular 3-D structures creates the famous Ames room illusion in Fig. 1C.

This observation poses the following interesting questions: What kind of assumptions make human beings so capable and willing to derive 3-D information from single perspective images, despite the potential for failure? To what extent can we use these assumptions to reconstruct the most important 3-D information, such as scene structures and camera poses, only from single images. In this paper, we will examine the computational basis for one such assumption.

Much of the “regularity” in 3-D space can be characterized mathematically by the notion of *symmetry*. For example, two parallel lines in space are translationally symmetric, a human face is reflectively symmetric, and a cube is both reflectively and rotationally symmetric. If points and lines are viewed as the basic primitives that one uses to model an arbitrary scene, different types of symmetry can be considered as high-level primitives that can be used to model a variety of regular objects and scenes, especially in man-made environment. Psychologically symmetry serves as a strong cue for human visual perception [2,3], and its computational advantages have also been addressed in the computer vision community [4–7]. However, there is still a lack of a unified effort to demonstrate how all types of symmetric primitives can be

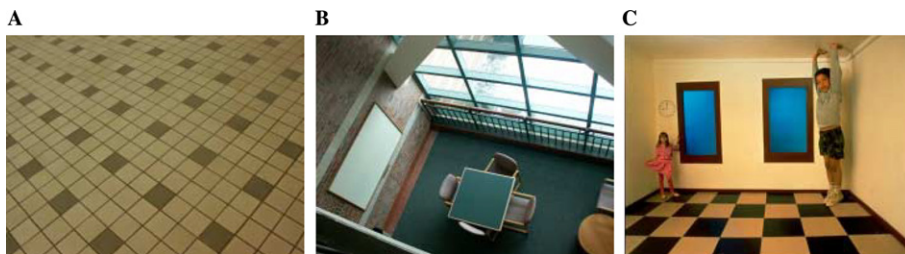


Fig. 1. (A,B) In man-made environments, symmetric patterns and structures are abundant. (C) The Ames room illusion.

detected, segmented, and matched in perspective images. In this paper, we propose a novel approach to automatically detect and segment all types of planar symmetry from single calibrated images, and match them across multiple images.¹

1.1. Structure from symmetry

One fundamental difficulty for inferring 3-D information from a single 2-D image is that infinitely many structures in space may give rise to exactly the same perspective image. Therefore, to make the problem well-conditioned, additional assumptions on the scene need to be imposed.

As we have learned from a general study of the relations between symmetry and multiple-view geometry [8], if an object admits rich enough symmetry, no 3-D geometric information (including structure and pose) is lost through perspective imaging except a scaling factor. This serves as the very basis for the entire symmetry-based reconstruction process. However, in this paper we are more interested in the inverse problem for planar symmetry: how to automatically detect and extract symmetric objects from each image, and further match them across multiple images. As we will show, a successful extraction of symmetric structures from single images relies on the interaction of the following two principles:

1. *Symmetry hypothesis testing.* Given a region of interest, to what extent can we claim that it could be the image of an object with certain types of symmetry in the space? An answer to this question requires us to understand how symmetry is *locally* encoded through perspective imaging so that we can verify whether *all* geometric relations for a valid image of a symmetric object are satisfied and consistent. From this we deduce the principle of “symmetry hypothesis testing.”

2. *Global symmetry testing.* Nevertheless, a region that passes certain types of local symmetry testing does not automatically imply that it must be an image of a symmetric object. Although each individual tile in Fig. 1 will pass any local testing as a square, what really makes the square interpretation “unquestionable” is the fact that this interpretation is also overwhelmingly *consistent* among all the tiles. This leads to the principle of “global symmetry testing” that we rely on in order to *robustly* derive 3-D structure from images.

In an image, the local symmetry hypothesis testing eliminates all regions which cannot be the perspective projection of any symmetric structures, hence gives correct focus to the regions which are the candidates of symmetric structures and thus contain richer 3-D information. The global symmetry testing carries the global constraints of the symmetric structures, which will resolve the inconsistency between multiple candidate regions, and also group the regions that contain the same symmetry type and spatial orientation.

¹ In this paper, we assume the intrinsic parameters of the camera have been calibrated. However, our research also shows the parameters can be calculated through a self-calibration process using symmetry [8].

1.2. Relation to the literature

A lot of research has been done on extracting general features or regions of interest from images. For the general problem of image primitive detection, there are feature-based methods, such as points [9] and lines [10], and pixel-based methods, such as active contours [11,12] and region segmentation [13–15]. If the segments sought can be parameterized, several techniques have been introduced, such as the well-known Hough transform [16] and nonlinear parameter estimation [17]. On the other hand, [18–20] addressed the topic of global object structure extraction from single images. Furthermore, there are methods that deal with other special kinds of invariant structures from perspective images [21–23].

Various instances of symmetry have been studied for the purpose of recognizing and recovering structures under the perspective, orthogonal and affine projections. This paper is by no means the first to notice that symmetry, especially reflective (or bilateral) symmetry, can be used to retrieve 3-D information. Mitsumoto et al. [24] were among the first who studied how to reconstruct a 3-D object using mirror image-based planar symmetry; Francois et al. [5] provided a more complete study of the reflective symmetry; Vetter and Poggio [25] proved that for any reflective symmetric 3-D object, one nonaccidental 2-D model view is sufficient for recognition; Zabrodsky and Weinshall [6] used bilateral symmetry assumptions to improve 3-D reconstruction from image sequences; and Zabrodsky et al. [26] provided a good survey on studies of reflective symmetry and rotational symmetry in computer vision. In 3-D object and pose recognition, Rothwell et al. [19] pointed out that the assumption of reflective symmetry can also be used in the construction of projective invariants and is able to eliminate certain restrictions on the corresponding points. For symmetry detection, [4,27,28] presented efficient algorithms to find axes of reflective symmetry in 2-D images, Sun and Sherrah [29] discussed reflective symmetry detection in 3-D space, and Zabrodsky et al. [26] introduced a so-called symmetry distance to classify reflective and rotational symmetry in 2-D and 3-D spaces (with some insightful comments given in [30]). Glachet et al. [31] are on locating general 2-D reflective contours under perspective imaging. For more about 2-D and 3-D symmetric curve deformation and reconstruction, please refer to [32].

On the role of symmetry groups in computer vision, Liu et al. [33] proposed a computational model to detect and represent all 17 Frieze patterns under the affine projection assumption. The limitation of their work is that the image can only be taken from the frontal view, and the repeated patterns must cover the majority of the image, although certain level of noise can be tolerated. On the study of symmetry in multiple-view geometry, Mundy et al. [34] presented a theoretical framework which is close to our work. It develops the epipolar geometry for repeated similar structures under the translation, rotation, affine, and bilateral relations. The key difference between [34] and this paper is: we show that an image of one symmetric structure itself is equivalent to multiple images by

investigating the geometric relations between the original image and its “hidden” counterparts. The relation between repeated similar structures in one image is presented in this paper under the concept of translational symmetry under the perspective projection. And most importantly, we propose a theoretical framework to extract 2-D symmetric structures from images and discuss the ambiguity in matching one or multiple symmetric patterns across a set of images, which are not covered in [34].

1.3. Contribution of this paper

In the literature, reflective, rotational, and translational symmetries have been primarily studied and used independently, often not under the exact perspective projection. However, our research shows that the key to consistent detection and segmentation of symmetric structures from their 2-D perspective images is to examine the relations among all types of symmetry as an *algebraic group*. By introducing symmetry group as a geometric cue into conventional image segmentation techniques, we are able to, for the first time, segment an image based on the precise and consistent 3-D geometry of the segmented regions. The output of such a segmentation is a hierarchy of geometric primitives (called *symmetry cells* and *complexes*) whose 3-D geometric information is fully recovered.

It is also well known that if multiple images of point and line features are provided, we can reconstruct general 3-D structures through classical structure-from-motion technique. But points and lines themselves contain little 3-D information in single views, and matching across images is ill-conditioned in many cases. Nevertheless, symmetry-based segmentation enables the use of symmetry cells as new geometric primitives for 3-D reconstruction from multiple views. From the system point of view, these symmetry cells can be treated as landmark features, which themselves contain rich 3-D information. If one common cell is detected in an image sequence, the reconstruction process will become more robust and much faster. Our study reveals the inherent ambiguities in matching symmetry cells as well as offers effective solutions to resolve the ambiguities. The outcome is a complete process for symmetry-based 3-D reconstruction and modeling: from extraction of symmetry cells in each image, matching them across multiple images, to a full reconstruction of 3-D camera poses and scene structures.

1.4. Paper organization

We apply the theory of multiple-view geometry to single views of symmetric structures in Section 2. In Section 3, a hierarchical symmetric pattern extraction algorithm is introduced. The algorithm will segment a perspective image based on its geometry information. Section 4 presents a novel method to achieve fast 3-D matching across multiple images using the symmetric patterns extracted. We show several other experimental results in Section 5.

2. Geometry for a single image of a planar symmetric structure

2.1. Planar symmetry and its group representation

We first define what symmetry is for a rigid body in 3-D space. Mathematically, the symmetry of an object is usually described by a symmetry group. Each element in the group represents a symmetric transformation under which the object is invariant.

Definition 1 (*Symmetric structure and its group action*). A set of 3-D features (points or lines) $S \subset \mathbb{R}^3$ is called a *symmetric structure* if there exists a nontrivial subgroup G of the Euclidean group $E(3)$ that acts on it. That is, for any element $g \in G$, g defines an automorphism on S (i.e., a one-to-one map from S to itself). In particular, we have $g(S) = g^{-1}(S) = S$ for any $g \in G$.

Generally, there are three types of symmetry for a rigid body: reflective symmetry, rotational symmetry, and translational symmetry. In this paper, we will only focus on 3-D recovery of symmetric planar structures. The reader is referred to [8] for the discussion on other symmetry properties of general rigid bodies (e.g., a cube) and [32] for deformable surfaces (e.g., a leaf).

When the structure S is on a plane P , the symmetry group G is instead a subgroup of the 2-D Euclidean group $E(2)$. If we choose a coordinate frame (x, y, z) attached to the structure S such that the z -axis is a normal vector of the plane P , then any point on the plane is determined by its (x, y) coordinates. Using homogeneous representation, any element $g \in G$ can be expressed as

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

where $R \in \mathbb{R}^{3 \times 3}$ is an orthogonal matrix representing the rotation or reflection on the xy -plane, and $T \in \mathbb{R}^3$ is a vector for translation on the xy -plane with $z = 0$.

Example 2 (*The symmetry group of a rectangle*). The symmetry group of a rectangle shown in Fig. 2 can be represented with respect to the canonical object frame by the following four matrices:

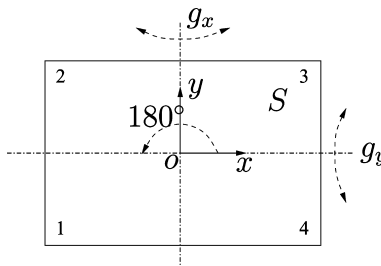


Fig. 2. A rectangle whose symmetry includes reflections along the x and y axes and a rotation about o by 180° . These transformations form a symmetry group of order 4.

$$g_e = I_{4 \times 4}, \quad g_x = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad g_y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$g_z = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where g_x and g_y denote the reflections along the x and y axes, respectively, and g_z the rotation about the z -axis by 180° . For example, the permutation result of $g_x((1,2,3,4))$ is $(4,3,2,1)$, which is exact the point correspondence of the structure S before and after action g_x .

Elements in the group $G = \{g_e, g_x, g_y, g_z\}$ satisfy the group relations

$$g_x^2 = g_y^2 = g_z^2 = g_e, \quad g_x g_y = g_y g_x = g_z, \quad g_x g_z = g_z g_x = g_y, \quad g_y g_z = g_z g_y = g_x.$$

These relations will be used later as the criteria for local symmetry testing. Only those regions that satisfy these group relations (under the homography representation to be precise, which will be introduced in the next section) can be considered as candidates of planar rectangle structures to the global symmetry testing.

2.2. Homography group in a perspective image

Now we consider a perspective image of the planar object S that is taken by a camera at a vantage point (or pose) $g_0 = (R_0, T_0) \in SE(3)$ with respect to the object frame, where $R_0 \in SO(3)$ and $T_0 \in \mathbb{R}^3$. This section will show how to represent the symmetry group relations described in the previous section under perspective imaging.

Fig. 3 shows a rectangle S on the plane P and a vantage point at g_0 . We can denote the image as $I(S)$. Since the rectangle is reflectively symmetric and g_x is a reflection along the yz -plane, then there exists an equivalent vantage point o' from which we would get exactly the same image $I(S)$ after g_x acts on S , and the transformation of the camera frames from o to o' can be represented as g'_x .

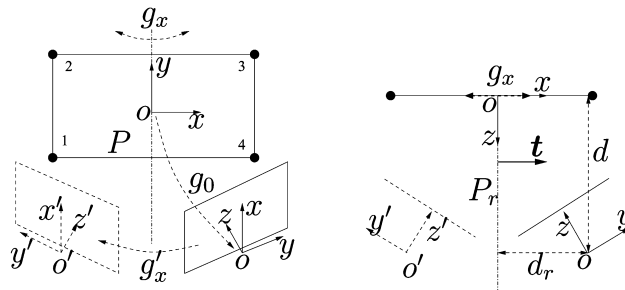


Fig. 3. Equivalent images of a rectangle, before and after a reflection g_x . (left) Frontal view; (right) top view. P_r is the plane of reflection and t is its (unit) normal vector.

In general, in a calibrated perspective image, we know under the homogeneous coordinates

$$\lambda \mathbf{x} = \Pi_0 g_0 \mathbf{X}, \quad (2)$$

where $\Pi_0 = [I_{3 \times 3} \ 0]$ is the projection matrix and g_0 is the vantage point. After applying a symmetry action $g \in G$, the new image of the new point $g\mathbf{X}$, denoted by $g(\mathbf{x})$, satisfies the following equation:

$$\lambda' g(\mathbf{x}) = \Pi_0 g_0 (g\mathbf{X}) = \Pi_0 (g_0 g g_0^{-1}) g_0 \mathbf{X}. \quad (3)$$

The second equality expresses the fact that the image of the structure S will remain the same, which is projected from the original point \mathbf{X} , if the image is taken from a vantage point different from the original one by $g_0 g g_0^{-1}$. The transformation $g_0 g g_0^{-1}$ is exactly the action of g on S expressed with respect to the camera coordinate frame. Then the symmetry group G on the plane P in space is naturally represented by its conjugate group $G' \doteq g_0 G g_0^{-1}$. We call G' the *homography group* (we will explain why it is called so later in this section).

Example 3 (*The homography group of a rectangle*). For the rectangle case studied in Example 2, the homography group $G' = g_0 G g_0^{-1}$ is given by $\{I, g_0 g_x g_0^{-1}, g_0 g_y g_0^{-1}, g_0 g_z g_0^{-1}\} \doteq \{I, g'_x, g'_y, g'_z\}$, and its elements satisfy the same set of relations as G in Example 2 since they are isomorphic:

$$(g'_x)^2 = (g'_y)^2 = (g'_z)^2 = I, \quad g'_x g'_y = g'_y g'_x = g'_z, \quad g'_x g'_z = g'_z g'_x = g'_y, \quad g'_y g'_z = g'_z g'_y = g'_x.$$

On the other hand, any element $g' \in G'$ itself is nothing but a linear transformation in $SE(3)$ with

$$g' = \begin{pmatrix} R' & T' \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{4 \times 4} \quad (4)$$

under the homogeneous coordinates, where $R' \in \mathbb{R}^{3 \times 3}$ is a rotation or reflection matrix, and $T' \in \mathbb{R}^3$ a translation, both in 3-D space. Hence, R' and T' can be represented by the following equality:

$$\begin{pmatrix} R' & T' \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} R_0 & T_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R_0 & T_0 \\ 0 & 1 \end{pmatrix}^{-1}, \quad (5)$$

which gives

$$\begin{cases} R' = R_0 R R_0^T, \\ T' = (I - R_0 R R_0^T) T_0 + R_0 T. \end{cases} \quad (6)$$

Our goal in recovering the pose of the symmetric pattern is to find R_0 and T_0 , the transformation between the structure and the camera. To use the relation described in Eq. (6), we first need to know R' and T' , and they can be solved from the homography relation between the two image planes with respect to the two camera frames at o and o' .

Since S is planar, there exists a homography H_0 mapping the plane $P \supseteq S$ in space to the image plane with H_0 given by

$$H_0 \doteq [R_0(1) \ R_0(2) \ T_0], \quad (7)$$

where $R_0(1)$ and $R_0(2)$ are the first two columns of R_0 . Given a point $X = [X, Y, Z]^T$ on the object plane, its coordinates on the image plane are given by $g_0(X) = H_0X$.

Furthermore, a symmetry action $g \in SE(3)$ is restricted to planar rotations and reflections, and it can be represented with respect to the object frame by a 3×3 matrix H , although H is not a homography matrix from the object to the camera in space. That is,

$$H \doteq \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{3 \times 3}, \quad (8)$$

where $R \in \mathbb{R}^{2 \times 2}$ is a rotation or reflection matrix on the xy -plane and $T \in \mathbb{R}^2$. Hence, a reflection or rotation of a point on the xy -plane can also be represented in matrix form as: $g(X) = HX$. It is easy to see that all the H 's form a group isomorphic to G . By abuse of language, we can denote $g = H$ (in the sense of isomorphism), and hence G consists of all the H matrices.

Due to the symmetry of S , for any $g \in G$, we have $g(S) = H(S) = S$, and therefore $H_0(H(S)) = H_0(S)$. For a particular point $X \in S$ with coordinates $[X, Y, Z]^T$ in the object frame, we have

$$H_0(HX) = H_0HH_0^{-1}(H_0(X)). \quad (9)$$

The quantity on the left can be viewed as the image of X taken at H_0 after a symmetry transformation H has been applied to S . This is equivalent to that the image is taken at a “new” vantage point $H_0HH_0^{-1} \cdot H_0 = H_0H$. Hence, the homography transformation between the two image planes with respect to the two camera frames can be represented as

$$H' = H_0HH_0^{-1} \in \mathbb{R}^{3 \times 3}. \quad (10)$$

We say H' is the homography transformation of the two images induced by g' , the 3-D transformation of the two vantage points. Notice that H' plays the role of the relative transformation between the real image and the equivalent “hidden” view. This relation is illustrated in the following diagram:

$$\begin{array}{ccc} S & \xrightarrow{g=H} & S \\ H_0 \downarrow & & \downarrow H_0 \\ I(S) & \xrightarrow{H'=H_0HH_0^{-1}} & I(S) \end{array} \quad (11)$$

By abuse of language, we can denote $g' = H'$, and G' is the group consists of all the H' matrices. Hence G' is a conjugate group of G , that is, $G' = H_0GH_0^{-1}$. Now we are clear that G' as a conjugacy class is actually describing the homography relations between the real image and its hidden counterparts supposing the planar structure

undergoing symmetry actions on the object plane. That is the reason G is called *homography group* acting on an image of a symmetric structure.

Suppose the object plane P is defined by the equation $N^T X = d$ for any point $X \in \mathbb{R}^3$ on P , where $N \in \mathbb{R}^3$ is the unit normal vector of the plane, and $d \in \mathbb{R}_+$ is its distance to the center of the first camera frame (e.g., see Fig. 3). We know that a 3-D transformation in $O(3)$, $g' = (R', T')$, and its associated homography matrix between two image planes, H' , satisfy the following relation:

$$H' = R' + \frac{1}{d} T' N^T \in \mathbb{R}^{3 \times 3}. \quad (12)$$

Computationally, with more than four corresponding points between any pair of equivalent views, such as the four corner features of the rectangle in Fig. 3, $\{R', T', N\}$ can be linearly recovered from the following equation with two possible solutions up to scale [35]:

$$\mathbf{x}' \sim H' \mathbf{x} \iff \hat{\mathbf{x}}' H' \mathbf{x} = 0, \quad (13)$$

where \mathbf{x} and \mathbf{x}' are the images of the corresponding symmetric points on X and $g(X)$, respectively, and $\hat{\mathbf{x}}' \in \mathbb{R}^{3 \times 3}$ is the *skew symmetric matrix* of the vector \mathbf{x}' . The homography relation in Eq. (13) can also be expressed as $\mathbf{x}' \times (H' \mathbf{x}) = 0$.

We have shown that R' and T' can be calculated by decomposing the homography H' with more than 4 pairs of image points. In the next section, we will discuss how to get the camera pose (R_0, T_0) under the three different symmetry assumptions. we will also show that by introducing extra symmetry constraints, the solution to the 3-D pose recovery is unique up to a scale indicating the unknown distance between the camera and the plane P .

2.3. 3-D pose recovery of planar symmetric structures

In this section, we study the ambiguities in recovering g_0 from a single reflective, rotational, or translational symmetry.

2.3.1. Reflective symmetry

A reflection on a plane can always be described by a motion $g = (R, 0)$, where $\det(R) = -1$. Hence in its homography representation, $R' = R_0 R R_0^T$ and $T' = (I - R_0 R R_0^T) T_0$.

Lemma 4 (Relative camera pose from reflective symmetry). *Let $\mathbf{n} = T' / \|T'\|$, then $R' = I - 2\mathbf{n}\mathbf{n}^T$. The induced epipolar constraint becomes*

$$(\mathbf{x}')^T \hat{T}' \mathbf{x} = 0. \quad (14)$$

The proofs of the lemmas and propositions in this section are given in Appendix A. We conclude that T' can be recovered from two pairs of symmetric points using Lemma 4, and R' is uniquely determined by T' . Accordingly, by decomposing the associated reflective homography matrix H' , only one solution

$$H' \mapsto \left\{ R', \frac{1}{d} T', N \right\}$$

preserves symmetry of the pattern, that is, the reconstruction of image points from only one of the two possible solutions results in a symmetric pattern that is consistent with our assumption.

Then the camera pose $g_0 = (R_0, T_0)$ can be solved by the following proposition:

Proposition 5 (Canonical pose from reflective homography). *Given an image of a planar structure S with a reflective symmetry R in 3-D space, $R' = R_0 R R_0^T$, and N all known, if the origin of the object frame is chosen to be on the object plane, the initial pose g_0 can be determined up to an arbitrary translation of the frame along the intersection line of the object plane and the plane of reflection. If we set the x -axis to be the normal of the reflection plane and the z -axis to be the normal N of the planar structure, as shown in Fig. 3, we get a solution for R_0 :*

$$R_0 = [\pm v_1, \pm \widehat{N} v_1, N] \in SO(3), \quad (15)$$

where v_1 is the eigenvector of R' corresponding to the eigenvalue $\lambda_1 = -1$.

The vector T_0 can always be set to be the average of four vertices in the space, which are recovered by a reprojection from the image using R' and T' .

2.3.2. Rotational symmetry

Given a rotation $g = (R, 0)$, where $\det(R) = 1$, its homography representation is $R' = R_0 R R_0^T$ and $T' = (I - R_0 R R_0^T) T_0$.

Lemma 6 (Relative camera pose from rotational symmetry). *If $\omega \in \mathbb{R}^3$ is the scaled rotational axis of R' , i.e., $R' = e^{\widehat{\omega}}$, we have $\omega \perp T'$.*

Since $\omega \perp T'$, the motion between equivalent views is an orbital motion. Hence, only four points are needed to recover the degenerated essential matrix $E' = \widehat{T}' R'$ [35]. By decomposing the associated rotational homography matrix H' , only one solution

$$H' \mapsto \left\{ R', \frac{1}{d} T', N \right\}$$

is consistent with our symmetry assumption in reconstruction. Similar to the previous section, we have the following proposition:

Proposition 7 (Canonical pose from rotational homography). *Given an image of a planar structure S with a rotational symmetry with respect to an axis $\omega \in \mathbb{R}^3$, the initial pose g_0 is determined up to an arbitrary rotation around the axis (i.e., $SO(2)$). If we assume v_1, v_2 , and v_3 are three eigenvectors of R' that correspond to the eigenvalues $+1$, $e^{+j\theta}$, and $e^{-j\theta}$, R_0 is of the form*

$$[\text{Im}(v_2) \cos(\alpha) + \text{Re}(v_2) \sin(\alpha), \quad \text{Re}(v_2) \cos(\alpha) - \text{Im}(v_2) \sin(\alpha), \pm v_1], \quad (16)$$

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.

Since rotating a solution of R_0 by an arbitrary angle is still a valid solution, we may choose one with the simplest form: $R_0 = [\text{Im}(v_2), \text{Re}(v_2), \pm v_1]$, and the sign of the third column is chosen such that $\det(R_0) = 1$. The vector T_0 can be set to be the average of four vertices in the space, which are recovered by a reprojection from the image using R' and T' .

2.3.3. Translational symmetry

In this case, $R = I$ and $T \neq 0$. We have $R' = R_0 I R_0^T = I$ and $T' = R_0 T$. Furthermore, if the origin of the world frame is chosen in the object plane, the plane normal as the z -axis, and T' as the x -axis, we have

$$R_0 = [T' \quad \hat{N}T' \quad N] \in SO(3), \tag{17}$$

where both T' and N can be uniquely recovered by decomposing the translational homography equation assuming $d = 1$:

$$H - I = T'N^T. \tag{18}$$

We end up with a two-parameter family of ambiguities in determining g_0 , that is, translating o arbitrarily inside the plane. An extra translational symmetry along a different direction does not help reducing the ambiguities.

2.3.4. Summary

To summarize the above results, we have the following theorem:

Theorem 8 (Structure from symmetry). *For any (planar) symmetric object, its 3-D structure and pose are already encoded in a single 2-D perspective image. As long as the assumption of symmetry is correct, no 3-D information is lost through perspective imaging except a scaling factor.*

The next example will demonstrate how to recover the pose of a regular polygon from its image using one rotational symmetry assumption.

Example 9 (3-D pose recovery of a regular pentagon). Suppose a pentagon shown in Fig. 4 is on the x, y -plane with $z = 0$, and its vertices have the coordinates in space:

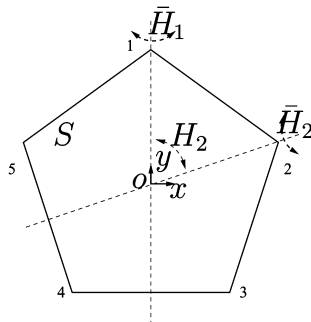


Fig. 4. A pentagon whose symmetry transformations form a dihedral group of order 10, i.e., D_{10} .

$\mathbf{X}_1 = [0, 1, 0]^T$, $\mathbf{X}_2 = R_{xy}(\frac{2\pi}{5})\mathbf{X}_1$, $\mathbf{X}_3 = R_{xy}^2(\frac{2\pi}{5})\mathbf{X}_1$, $\mathbf{X}_4 = R_{xy}^3(\frac{2\pi}{5})\mathbf{X}_1$, $\mathbf{X}_5 = R_{xy}^4(\frac{2\pi}{5})\mathbf{X}_1$, where $R_{xy}(\theta)$ is the rotation matrix on the x, y -plane by θ . We set

$$R_0 = \begin{bmatrix} \cos(\frac{\pi}{10}) & 0 & -\sin(\frac{\pi}{10}) \\ 0 & 1 & 0 \\ \sin(\frac{\pi}{10}) & 0 & \cos(\frac{\pi}{10}) \end{bmatrix}, \quad T_0 = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$$

as the unknown camera pose. After the perspective projection $\lambda\mathbf{x} = R_0\mathbf{X} + T_0$, the homogeneous coordinates of the five points in the image become $\mathbf{x}_1 = [2, 4, 1]^T$, $\mathbf{x}_2 = [1.0955, 3.3090, 0.7061]^T$, $\mathbf{x}_3 = [1.4410, 2.1910, 0.8184]^T$, $\mathbf{x}_4 = [2.5590, 2.1910, 1.1816]^T$, and $\mathbf{x}_5 = [2.9045, 3.3090, 1.2939]^T$.

The symmetry group of the pentagon isomorphic to the dihedral group D_{10} can be represented by the following matrices:

$$H_1 = I, \quad H_2 = \alpha, \quad H_3 = \alpha^2, \quad H_4 = \alpha^3, \quad H_5 = \alpha^4, \\ \bar{H}_1 = \beta, \quad \bar{H}_2 = \beta\alpha, \quad \bar{H}_3 = \beta\alpha^2, \quad \bar{H}_4 = \beta\alpha^3, \quad \bar{H}_5 = \beta\alpha^4,$$

where

$$\alpha = \begin{bmatrix} \cos(\frac{2\pi}{5}) & -\sin(\frac{2\pi}{5}) & 0 \\ \sin(\frac{2\pi}{5}) & \cos(\frac{2\pi}{5}) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \beta = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We assume the object to be rotationally symmetric, so the point correspondence in the image with respect to H_2 is

$$\mathbf{x}'_1 = \mathbf{x}_2, \quad \mathbf{x}'_2 = \mathbf{x}_3, \quad \mathbf{x}'_3 = \mathbf{x}_4, \quad \mathbf{x}'_4 = \mathbf{x}_5, \quad \mathbf{x}'_5 = \mathbf{x}_1.$$

And the homography matrix of this set of points is:

$$H' = \begin{bmatrix} -3.4913 & -0.9045 & 11.6960 \\ 0.9323 & 0.3090 & 0.2083 \\ -1.4593 & -0.2939 & 4.8003 \end{bmatrix}.$$

Next we decompose H' into four solutions by the standard method[35], and two of them give positive depths of the normal vector N of the pentagon under the perspective projection. However, we use these two sets of (R', T') to reconstruct the points $\{\bar{\mathbf{X}}_i\}$ in 3-D space from the image by assuming $d = 1$, only one set gives symmetric points with the edges having the same length:

$$R' = \begin{bmatrix} 0.3750 & -0.9045 & -0.2031 \\ 0.9045 & 0.3090 & 0.2939 \\ -0.2031 & -0.2939 & 0.9340 \end{bmatrix}, \quad T' = \begin{bmatrix} 12.5115 \\ -0.0900 \\ 4.0652 \end{bmatrix}, \quad N = \begin{bmatrix} -0.3090 \\ -0.0000 \\ 0.9511 \end{bmatrix}.$$

Finally, we solve for R_0 and \bar{T}_0 using Proposition 7

$$R_0 = \begin{bmatrix} 0.9511 & 0.0000 & -0.3090 \\ 0 & 1.0000 & 0.0000 \\ 0.3090 & -0.0000 & 0.9511 \end{bmatrix}, \quad \bar{T}_0 = \begin{bmatrix} 6.0056 \\ 9.0084 \\ 3.0028 \end{bmatrix}.$$

T_0 may be fully recovered by normalizing \bar{T}_0 using the true distance scale $d = N^T x_1 = 0.3330$. In summary, no 3-D information but one scaling factor is lost by the perspective projection of symmetric patterns in single views.

People may wonder why the other solution did not pass the symmetry test. The four-point algorithm computes (R', T') by decomposing the matrix product $H'^T H'$. If there exists another homography matrix H'' , such that $H''^T H'' = H'^T H'$, its corresponding relative position (R'', T'') is another valid solution based only on the positive depth constraint [35]. However, if the planar structure is symmetric, as in this case a pentagon, there are five 3-D camera positions that are qualified to have a correct “hidden” view. In general, (R'', T'') does not coincide with these positions, hence the reconstructed pattern is not symmetric. Theoretically, the other solution can be another valid “hidden” view. But if one more symmetric pattern is available in the scene, this ambiguity can be easily resolved. We will discuss this issue in Section 4.

In this example, we use one rotational symmetry assumption to recover the relative pose of the pentagon (R_0, \bar{T}_0) up to an unknown distance scale. It shows that if the symmetry assumption of a planar structure is correct, the pose of the structure with respect to the camera can be fully recovered. However, in this paper, our goal is to automatically detect symmetric structures only from their perspective images. Without knowing the 3-D information, applying reflective and rotational symmetries to the image of an irregular polygon will give inconsistent pose recoveries of the same object. The consistency is the key to the local symmetry testing we will propose in the next section. We claim a planar structure to be a symmetry cell candidate only when the poses recovered by all the elements of its homography group are consistent, especially the normal vectors recovered shall be the same.

Without loss of generality, in the following sections, we will use rectangle/square examples to demonstrate 3-D pose recovery of arbitrary planar symmetric structures. This simplification is valid in our discussion because the pose recovery method only relies on the correspondence of corner points of any polygon from its real view to its “hidden” views. It is trivial to extend the rectangle example to any polygons with $2n$ ($n \geq 2$) vertices, and extend the pentagon example to polygons with $2n + 1$ ($n \geq 2$) vertices.

3. Symmetry-based hierarchical image segmentation

Based on the theory we have in the previous sections, we are now ready to show how it may help us to segment a 2-D image of a scene in terms of 3-D geometric information encoded in it. Our goal is to segment the image into regions, identify the regions which potentially have a consistent interpretation as symmetric objects in space, and recover relative 3-D geometric relations between these regions.

Fig. 5 outlines the architecture of a prototype system that we have developed to achieve these tasks. The hierarchical architecture consists of three main levels. In this section, we will demonstrate how to implement the low-level symmetry cell

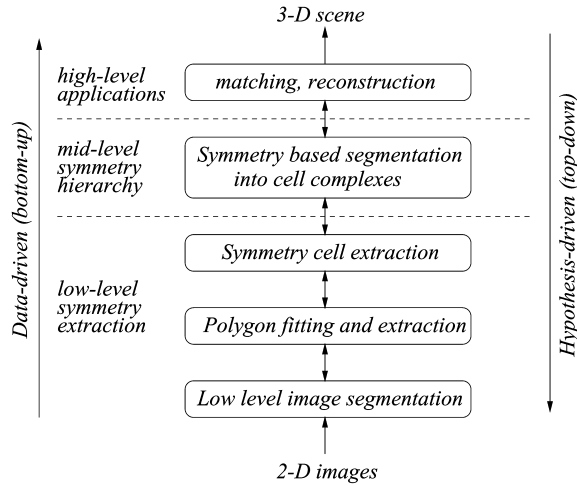


Fig. 5. The system diagram of the symmetry-based geometric segmentation algorithm.

extraction and middle-level symmetry hierarchy construction. We will also introduce a novel method to match symmetric cells across multiple images as an example of high-level applications in the next section.

3.1. Low-level symmetry cell extraction

As we know from the previous chapter, the homography matrices associated to different symmetry elements in G must preserve the same group relations. Furthermore, each homography should lead to the same 3-D reconstruction of the structure S . This is a very important condition and it leads to the following principle.

Principle 1 (Symmetry hypothesis testing). To verify if (part of) an image can be interpreted as that of an object with a symmetry group G , we need to verify for all the elements in the group G whether the resulting homography matrices preserve the same group structure and all lead to a consistent 3-D structure and pose (in particular, surface normal) recovery for the object.

To “recognize” or “detect” symmetry in an image, we first select some candidate image regions, which we call *cells*, and determine whether each cell passes the testing of certain types of local symmetry criteria. For instance, we can choose any set of four connected line segments that enclose a homogeneous region in the image as a cell. It may correspond to a 3-D quadrilateral that falls into one of the following cases: (1) nonsymmetric quadrilateral, (2) one reflective symmetry ($G = \mathbb{Z}_2$), (3) a rectangle ($G = D_4$), or (4) a square ($G = D_8$). Using the algorithms provided in the previous section, we can test the cell against all four cases and determine which types of symmetry this cell admits. If it falls into the first two categories, we can simply discard it, since we cannot distinguish the solution with the correct reflection

assumption from the ones with the incorrect assumptions when the reconstruction results are all inconsistent. That is to say, the first two cases are not distinguishable. A cell that belongs to any of the other two categories has consistent pose recovery solutions, and is called a *symmetry cell* for a quadrilateral region, with the type of symmetry labeled on it. Similar definitions may apply to other polygon regions with more than 4 edges.

To make our explanation intuitive, we will use the image shown in Fig. 6 to demonstrate the process of the overall system. In Section 5, other examples will be given for more generic scenes.

Polygon extraction. To obtain candidate regions for symmetry cells, we use a color-based mean shift algorithm [14] to obtain homogeneous color pattern regions from the image. The two primary parameters for the mean shift algorithm are set to be $\sigma_S = 7$, $\sigma_R = 9$. For every region that is larger than a minimal size, we compute the convex hull of the region in order to guarantee its convexity and reduce the noise on its boundary.

Since the contour of a polygon consists of piecewise line segments as special curves with zero constant curvature and corners with local curvature peaks, we developed a polygon fitting technique based on the constant curvature criterion proposed in [36]. The result for Fig. 6 is shown in Fig. 7. Compared to the original image, some of the “cells” are missing due to noise. Here we want to point out again that the overall algorithm is not restricted to quadrilaterals, and in fact from the constant curvature criterion, we directly know how many edges a convex region contains. The convex hull plus constant curvature technique is not essential to the proposed algorithm, it can be replaced by any other boundary smoothing and polygon fitting methods.

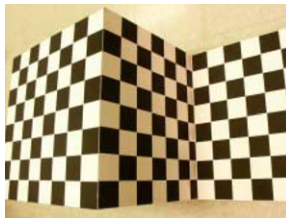


Fig. 6. An image of a cube and a checkerboard.

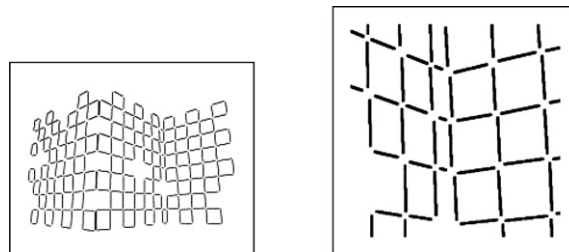


Fig. 7. (Left) The output of polygon extraction. (Right) Details.

Local symmetry testing. For each extracted quadrilateral, using Principle 1, we can test whether it satisfies the symmetry of, for instance, a rectangle or a square. Fig. 8 shows cells extracted from the image in Fig. 7 that pass the testing as squares with their object frames attached. However, in generic images, the normal vectors obtained from segmentation at this level can be noisy, as experiments in Section 5 will show.

3.2. Middle-level symmetry hierarchy: geometric segmentation

A cell that passes the local symmetry testing of a rectangle or a square does not necessarily correspond to the image of a rectangle or square in space. We should further verify its validity in the context of the whole scene. For this purpose, we propose the following principle:

Principle 2 (*Global symmetry testing*). Symmetry cells which have consistent 3-D geometric properties (e.g., orientation) with other adjacent cells more likely correspond to symmetric 3-D structures in space.

For example, if a number of neighboring symmetry cells have mutually consistent normals, it is likely that these cells come from the same 3-D plane. Furthermore, a set of cells together may correspond to a 3-D structure which admits a new symmetry at a higher level. For example, two neighboring cells can be translational copies of each other in space, which is often the case for window patterns on the side of buildings. This leads to the notion of *symmetry complex*: a group of (preferably neighboring) cells with consistent 3-D geometric properties. Different geometric properties may lead to different types of symmetry complex segmentation. In this paper, we consider three properties as examples: orientation, topology (connectivity), and coplanarity.

In our prototype system, we first use the ISODATA algorithm [37] to cluster the spherical coordinates of the normal vectors of the obtained planar symmetry cells. The outliers are discarded. We further relate the symmetry cells by their connectivity relations. For two cells in the same orientation group, if their edges or corners are

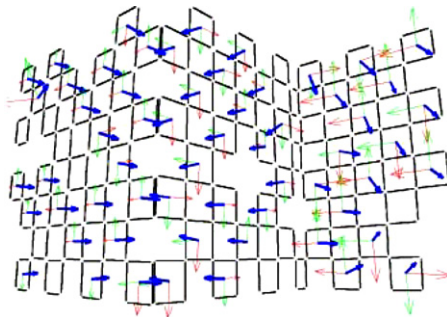


Fig. 8. Cells that pass the local symmetry testing are marked with their 3-D object frame (x, y, z)-axes drawn in red, green, and blue arrows, respectively.

adjacent to each other, they can be considered as “connected.” Finally, for those connected cells with the same local reflective and rotational symmetries, we can verify if they are coplanar by considering the translational symmetry among them. If two cells are coplanar and congruent, the plane normal recovered from the translational homography between the two cells needs to be the same as the normals of both cells.

Collective adjustment. Coplanar cells within the same group must share two things: a common normal vector N and the same distance d from the camera center to their supporting plane. We take \bar{N} to be the average of all the normal vectors within the group. For each cell in the group, the rotation R of its object frame is adjusted such that the z -component is \bar{N} . The translation T of its object frame is then scaled such that $T^T \bar{N} = d$. Such an adjustment may significantly improve the 3-D reconstruction of cells in each group. Fig. 9 shows the result that symmetry cells extracted from the example image become collectively adjusted based on their coplanarity.

Through the above segmentation process, every image gives rise to a hierarchy of symmetry cells and complexes that resembles a tree-like structure, as shown in Fig. 10. One can view such a tree-like hierarchy of symmetry cells and complexes as a “ $(2\frac{1}{2})$ -D” representation of the image: Although the hierarchy is not yet a full

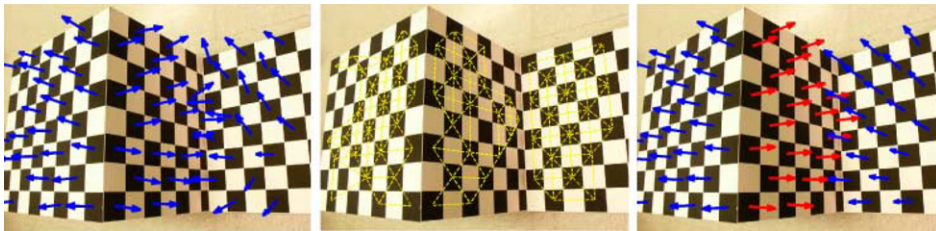


Fig. 9. (Left) Cells that pass the local symmetry testing are marked with their 3-D normal vectors drawn in blue arrows. (Middle) The output of the geometry-based grouping. Coplanar cells are grouped together, with yellow lines connecting cells in the same group. (Right) The final output of collective adjustment. The normal vectors of the cells within the same group are unified. The normal vectors of the cells in both the left and right figures were reversed for better visualization result.

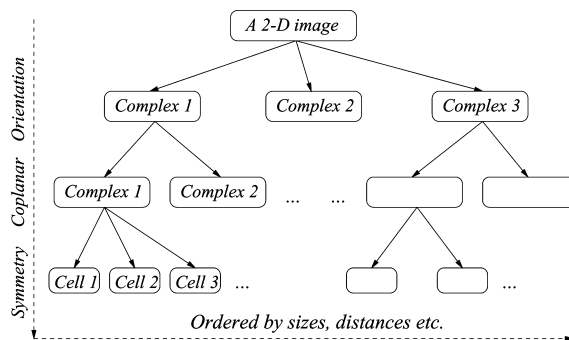


Fig. 10. A hierarchical segmentation of an image by 3-D geometric properties.

reconstruction of the entire scene, 3-D geometric information on regions of the scene is already available, and relations between these regions are also encoded within the hierarchy. This kind of hierarchy can also be used as landmarks from single images to allocate the possible position of the camera (and hence the robot) in the Simultaneous Localization and Mapping (*SLAM*) problem. We will elaborate this idea in the next section.

4. 3-D matching from images of symmetric structures

In this section, we deal with the situation when multiple views of (parts of) an object are given, and multiple symmetry cells on the structure are detected in the images. We will derive the conditions for matching multiple cells across the images.

4.1. Two-view reconstruction from one matched cell

We start with the simplest case. Suppose two images of the same cell are given, as shown in Fig. 11. For $i = 1, 2$, the pose (R_i, T_i) of the i th camera frame with respect to the object frame of the rectangle can be determined separately using the method given above. However, the reconstructed T_i 's have adopted different scales. The cause of this difference is that for each image, the distance d_i from the i th camera center to the supporting plane of the rectangle was set to be $d_i = 1$ to fix the unknown scale in translation.

Suppose that the four corners of the cell recovered from the first image have coordinates $(x_j, y_j)^T$ and those from the second image have coordinates $(u_j, v_j)^T$, $j = 1, 2, 3, 4$. Due to the difference in the adopted scales, the recovered coordinates of corresponding points differ by the same ratio. Since the corresponding points are physically the same points, the ratio $\alpha = \frac{d_2}{d_1}$ can be found in

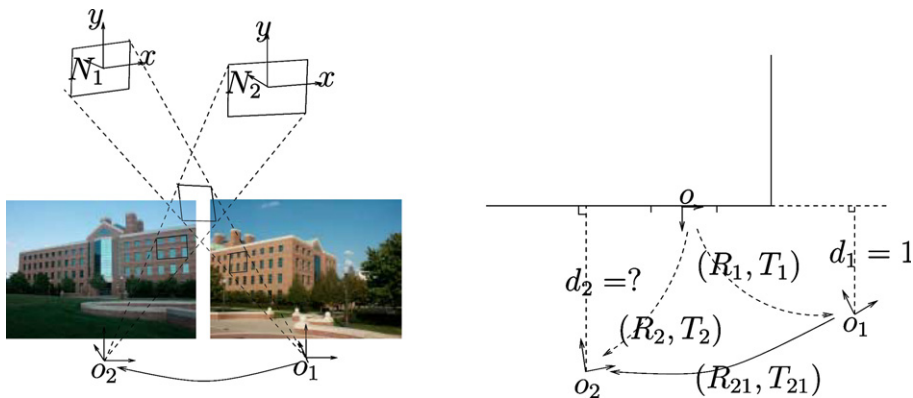


Fig. 11. The case for one rectangle in a two images. Usually the distance d_1 from the first camera center o_1 to the plane is set to be unit 1. Then the key is to recover d_2 , the distance from o_2 to the plane with respect to the reference d_1 .

$$\begin{bmatrix} x_j - \bar{x} \\ y_j - \bar{y} \end{bmatrix} = \alpha \begin{bmatrix} u_j - \bar{u} \\ v_j - \bar{v} \end{bmatrix}, \quad j = 1, 2, 3, 4, \quad (19)$$

where $\bar{x} = \frac{1}{4} \sum_{j=1}^4 x_j$ and similarly for \bar{y} , \bar{u} , and \bar{v} . In the presence of noise, α can be easily calculated using the least-square method. Therefore, if the first image is chosen as the reference image, then the relative pose of the second camera frame with respect to the canonical object frame becomes $(R_2, T_2) \leftarrow (R_2, \alpha T_2)$, and the distance from the second camera frame to the supporting plane of the rectangle is $d_2 = \alpha$.

4.2. Multiple views of multiple symmetry cells

First let us consider m perspective images $I_i(S)$ of a symmetry cell S taken at vantage points $(R_i, T_i) \in SE(3)$, $i = 1, 2, \dots, m$. For each (R_i, T_i) , there is a homography $H_i = [R_i(1) \ R_i(2) \ T_i] \in \mathbb{R}^{3 \times 3}$ which maps the plane $P \supseteq S$ in space to the i th image plane. The homography between the i th and j th images is denoted by $H_{ij} \doteq H_j H_i^{-1}$. We also denote $H \in \mathbb{R}^{3 \times 3}$ as a matrix representation of the symmetry action g on the xy -plane, although it is not a homography transformation in space. That is,

$$H \doteq \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{3 \times 3}, \quad (20)$$

where $R \in \mathbb{R}^{2 \times 2}$ is a rotation or reflection matrix on the xy -plane and $T \in \mathbb{R}^2$.

Since S has a symmetry group G , for each i , the matrices H'_i forms a homography group $G'_i = H_i G H_i^{-1}$ that is conjugate to G . The key observation is that the homographies $H'_i = H_i H H_i^{-1}$ with respect to the action g only depend on the i th image, but both H_i and H_{ij} can be determined from them by solving the following set of *Lyapunov* type linear equations:

$$H'_i H_i - H_i H = 0, \quad H'_j H_{ij} - H_{ij} H'_i = 0, \quad (21)$$

with H , H'_i , and H'_j known (see the proof of Proposition 5 in the [Appendix A](#) for detail). The benefit of symmetry then becomes clear: the relative poses between different views or the pose between each view and the object can be recovered from information that is encoded in individual images—no point matching across different images is needed. This confirms our observations in the previous sections.

However, some ambiguity remains in the above equations. It is easy to see that if H_i is a solution to the Eq. (21), so is $H_i \tilde{H}$ for any element \tilde{H} in the orientation-preserving² subgroup $\tilde{G} \doteq G \cap SE(2)$. This is because

$$H'_i(H_i \tilde{H}) - (H_i \tilde{H})(\tilde{H}^{-1} H H_i) = 0 \quad (22)$$

and $\tilde{H}^{-1} G \tilde{H} = G$. For a proper rectangular cell (i.e., not a square), there are only two possibilities:

² Because H_i needs to be orientation-preserving.

$$\tilde{G} = \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\} \simeq \mathbb{Z}_2. \quad (23)$$

This comes at no surprise since if the rectangle is rotated by 180° about its center, one obtains the same image of the cell from the same vantage point. Similarly, for a regular pentagon, $\tilde{G} = \{g_e = I_{3 \times 3}, g_1 = \alpha, g_2 = \alpha^2, g_3 = \alpha^3, g_4 = \alpha^4\}$, where α represents a rotation by $\frac{2\pi}{5}$ about its center

$$\alpha = \begin{bmatrix} \cos(\frac{2\pi}{5}) & -\sin(\frac{2\pi}{5}) & 0 \\ \sin(\frac{2\pi}{5}) & \cos(\frac{2\pi}{5}) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

From m images of a rectangular cell, we can get up to 2^m possible solutions to the camera poses (relative to the object frame) if we only rely on the homography group G' from each view. Nevertheless, this problem can be mostly resolved if the images contain more than one cell.

Theorem 10 (Cell-to-cell matching). *From $m \geq 2$ different views of $n \geq 2$ symmetry cells, given that each cell is matched correctly across different views, the solution of the camera poses is unique if and only if at least one cell has a different rotation axis.*

Proof. The proof is by examining every two views of two cells. To have a second solution, there must exist the same rigid-body transformation $g' \in SE(3)$ that is the homography conjugate $g' = g_i \tilde{g} g_i^{-1}$ of the same $\tilde{g} \in \tilde{G}$ on each cell. This situation is possible only when the two cells share the rotation axis, as we have illustrated in Fig. 12. \square

The above theorem requires that each cell is correctly matched across different views. If we relax this condition and allow a set of cells, i.e., a cell complex, to be matched among themselves, we have the following theorem.

Theorem 11 (Complex-to-complex matching). *Given a set of $n \geq 2$ symmetry cells matched among themselves across $m \geq 2$ views, the ambiguity in a valid solution for the*

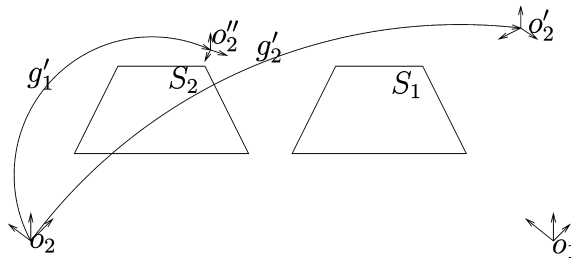


Fig. 12. Multiple views of multiple cells matching. o_1 and o_2 are two real vantage points. In solving for the translation between o_1 and o_2 , o_2 and o_2' are the two possible solutions for cell S_1 , while o_2 and o_2'' are the two possible solutions for cell S_2 . When the two cells do not share a rotation axis, the two ambiguous solutions o_2' and o_2'' can be eliminated since they are not consistent, i.e., $g_1' \neq g_2'$.

camera poses is only caused by the global rotational symmetry group of the set of cells as a single 3-D structure.

The proof is similar to Theorem 10. Obviously, Theorem 10 is a special case of Theorem 11. The latter stipulates that the only valid set of matching among a set of symmetry cells are caused by the global symmetry of all the cells. Notice that point features in general do not have the same nice properties as symmetry cells.

Finally, it is also possible that certain cells can be completely mismatched across different images, outside of a given set of cells. This is particularly the case when there are multiple repeated patterns in the scene. Since we know the 3-D pose and structure of each cell up to scale from each view, given the camera positions, there is essentially only one configuration in which a mismatch could occur. As shown in Fig. 13, the cell S_1^2 can be matched either to the cell S_2^2 or S_2^3 in the second image if the cells S^2 and S^3 in space are parallel to each other and their shape and texture differ by a similarity transformation. Such cells form exactly a 1-parameter family in the second image. This relation can be thought of as the “epipolar constraint” for cell matching between two views. Notice that a mismatch causes only a “depth” illusion in space. However, these types of mismatches can be easily eliminated once we know something about S^2 in relation with any other matched cells, for example, the relative size or depth between S^1 and S^2 .

4.3. Pictorial matching

The previous sections have covered the geometry for reconstruction of symmetry cells. However, several pictorial properties that the cells possess can further eliminate possible mismatches. For example, consider the shape similarity, a symmetry cell can only be matched to another one with the same number of edges and similar length ratios of the edges. For two cells which pass the shape matching, we can align their scales and compare the color texture inside the cells. This kind of 2-D information is not available for matching points, lines and certain local affine invariant regions, and it further highlights the advantages of aligning multiple images using symmetry cells.

Combining all the pictorial measurements, we can easily construct a matching matrix M for every two images, whose (i, j) component is the similarity of the i th and j th cells in two images, respectively. Nevertheless, the method searching for the optimal

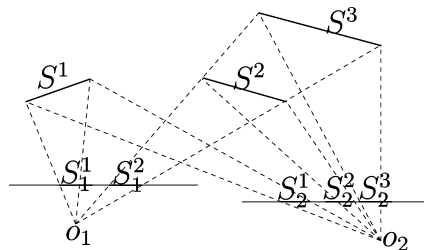


Fig. 13. “Epipolar constraint” for symmetry cells.

match in this matrix is not the focus of this paper, and the interested reader is referred to [38].

5. Experiments

5.1. Symmetry cell extraction

We first demonstrate the performance of the symmetry cell extraction algorithm by applying it to a variety of images of generic scenes. The images are shown in Fig. 14. For all the experiments, there is no manual interference at all. The system uses the same set of parameters for mean shift segmentation, polygon fitting, and symmetry testing. All images are 1280×960 pixels RGB color images. The speed for low-level cell extraction is within 1 min on a 1.4 GHz Pentium4 PC running Matlab, which includes the mean shift segmentation and polygon fitting. The computation of symmetry hierarchy without nonlinear optimization takes less than 5 s, but the optimization version usually needs several minutes, depending on how many cells are extracted from the image.

The order we select in our demonstration is from the easiest to the most difficult. The second and third images both contain multiple cells and complexes which are obviously the most salient rectangular objects in the scene. The system also successfully groups coplanar cells in the same groups. The last image taken from an aged book shelf is the most challenging one. Most of the geometric meaningful patterns contain characters inside the regions and they sit on the book contours which are not strictly planar surfaces. Some of the cell boundaries are not so well-defined and their 3-D geometric relations are not precise. The system still does a decent job of extracting and segmenting most cells.

5.2. Symmetry-based 3-D reconstruction

This section shows the experiments on matching symmetry cells across multiple images. Section 5.2.1 demonstrates the results of automatic cell matching and reconstruction with or without ambiguities across images. Section 5.2.2 demonstrates how to achieve fast 3-D image-based rendering by manually selecting several symmetry cells.

5.2.1. Automatic cell matching

Fig. 15 shows that two symmetry cells are automatically extracted and matched across three images of an indoor scene. Correspondence of the corners of the cells are derived from the only consistent solution for the camera poses found by our symmetry extraction algorithm. Notice that there is a very large translation between the first view and the last two views but a very large rotation between the second and third views. Classical feature tracking becomes noisy for such large motions. But the problem becomes very well-conditioned if the reconstruction of symmetry cells is first applied between the real images and their “hidden” counterparts, and then match them across the multiple views with the relative poses already extracted.

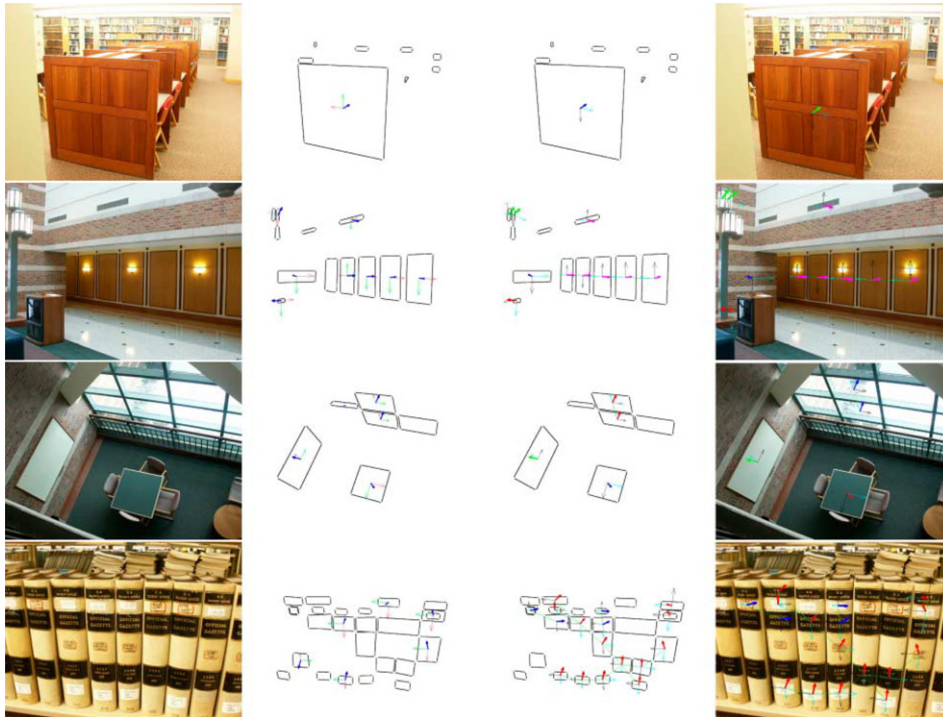


Fig. 14. Experiments. (Left) Original images; (Middle left) geometric cell extraction and local symmetry testing, cells without frame attached fail the test; (Middle right) local symmetry testing with nonlinear optimization. (Right) grouped symmetry cells, each with its object frame pose recovered and marked. Different groups of symmetry cells are identified by bold arrows with different colors for the normals.



Fig. 15. Two symmetry cells matched in three images. From the raw images, symmetry cell extraction, to cell matching, the process needs no manual intervention.

Our algorithm gives very accurate solutions. The ground truth for the length ratios of the white board and table are 1.51 and 1.00, and the recovered length ratios are 1.506 and 1.003, respectively. Error in all the right angles is less than 2.5° .

In Fig. 16, multiple solutions are found automatically by our algorithm for matching and reconstructing symmetric cells detected in two images of an indoor

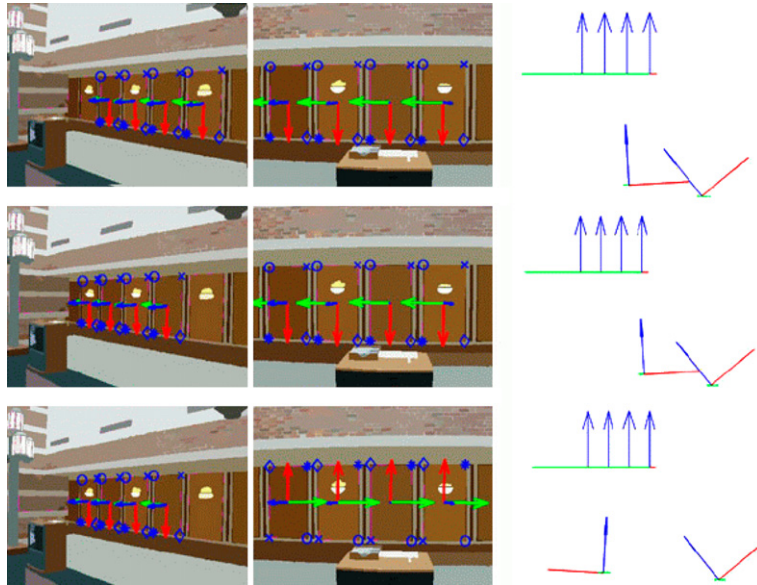


Fig. 16. From top to bottom: Different matches for four rectangles in each image with the first row showing the correct matchings. The right side shows the top view of the reconstructed 3-D scene with camera poses.



Fig. 17. Five images used for reconstruction of a building. The last image is used solely for obtaining roof information.

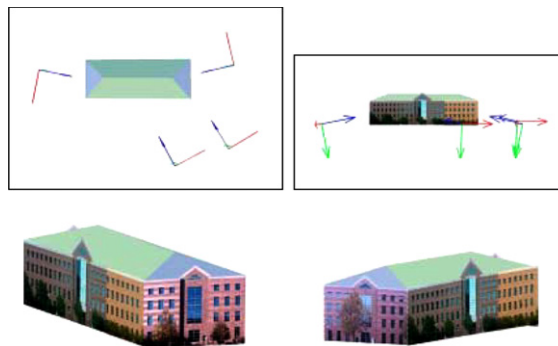


Fig. 18. (Top) The four coordinate frames show the recovered camera poses from the first four images in Fig. 17. The roof was substituted by a “virtual” one based on corners extracted from the fifth image with the symmetry-based algorithm. Blue arrows are the camera optical axes. (Bottom) Reconstructed 3-D model rendered with the original set of the four images.

scene. The top set of images shows the correct matching of four rectangles in the scene and the two camera poses are correctly recovered. The middle row is an ambiguous solution caused by the translational symmetry of the rectangular complex, in which each of the four rectangles in the left image is matched to the one to its right without using their color (texture) information. As a result, the recovered camera poses are shifted towards the right. The bottom row demonstrates another ambiguity that not only the matching is shifted due to the translational symmetry, but it is also mismatched due to the rotational symmetry of the four rectangles as a whole complex. Both the last two situations are examples of geometric ambiguities characterized by Theorem 11. Fortunately, both ambiguous solutions can be avoided when texture information is used in matching.

5.2.2. Semi-automatic scene reconstruction

The ambiguities displayed in Fig. 16 could become extremely complicated if many similar cells are present in a scene. For instance, the symmetry complex of the windows shown in Fig. 17 may lead to many ambiguous matches. In such cases, manual intervention is necessary if a realistic 3-D reconstruction is sought. Our algorithm still helps to minimize the necessary manual interaction.

For images in Fig. 17, the user only needs to point out which cells (not points) correspond to which in the images and our algorithm will then automatically find a consistent set of camera poses from the matched cells, as Fig. 18 top shows. Based on the recovered camera poses, the subsequent 3-D structure reconstruction is shown in Fig. 18 bottom. The 3-D model is rendered as piecewise planar with parts not on the building manually cut. The angles between the normal vectors of the two orthogonal walls differ from 90° by an average of 1° error even without any nonlinear optimization. The whole process (from taking the images) takes less than 20 min.

6. Conclusion

This paper demonstrates that it is computationally feasible to represent an image of man-made environment based on accurate 3-D geometric information. The key to this is that symmetry groups admitted by (planar) symmetric objects in space are precisely encoded in the image through the homography groups. This stipulates that no 3-D information is lost through the perspective projection for symmetric objects. The result of such geometry-based segmentation is a hierarchical representation of symmetric objects in terms of their spatial geometric properties and relations.

Correspondingly, the architecture of our symmetry detection method is also hierarchical. This allows the development of the algorithms for each layer to be relatively independent. Computationally, local and global symmetry testings are introduced between different layers to guarantee the consistency of the geometric hierarchy. Although this mechanism will probably reject some valid cells due to noise, it does significantly reduce the number of false positives in symmetry cell extraction, which is important for high-level applications such as matching.

If multiple images of the same scene are given, it is much easier to establish correspondence between cells and complexes than using only point or line features. This leads to the multiple-view geometry of symmetry cells. We have discussed the possible ambiguity in matching cells across images, and the extra geometric and pictorial information in symmetry cells allows us to recover them more easily.

A prototype reconstruction system based on the proposed algorithms is also developed. Experiments show that it gives very accurate 3-D recovery of the camera poses and scene structures. Although the proposed method is limited to scenes with symmetric objects, the type of symmetries can be very flexible and the number of symmetric objects needed is very small.

Appendix A

Proof of Lemma 4. First, it is easy to check that $T' = -R'T'$. Since $\det(-R') = 1$, we use a property of the skew-symmetric matrix,

$$\widehat{T}' = -\widehat{R}'\widehat{T}' = R'\widehat{T}'R'^T.$$

Then we have $R'\widehat{T}' = \widehat{T}'R' = \widehat{T}'$.

Next, we assume T' has unit length. Since $T'T'^T$ is an orthogonal projection matrix with respect to T' , and we have the constraints: $\det(R') = -1$ and $\widehat{T}'R' = \widehat{T}'$, $R' = I - 2T'T'^T$. The epipolar constraint induced becomes $\mathbf{x}^T\widehat{T}'\mathbf{x} = 0$. \square

Proof of Proposition 5. We first denote the *Lyapunov map* $L: \mathbb{R}^{3 \times 3} \rightarrow \mathbb{R}^{3 \times 3}$; $R_0 \mapsto R'R_0 - R_0R$ associated to the condition

$$R' = R_0RR_0^T \iff R'R_0 - R_0R = 0.$$

Without loss of generality, we assume the reflection matrix R is with respect to the yz -plane and has the form

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

Therefore its eigenvalues are

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

Let $\{u_1, u_2, u_3\}$ be the left eigenvector of R . Hence the left eigenvector associated with $\mu_1 = -1$ is $u_1 = [1 \ 0 \ 0]$. The two real left eigenvectors associated with the two repeated eigenvalue $\mu_2 = \mu_3 = 1$ are $u_2 = [0 \ 1 \ 0]$ and $u_3 = [0 \ 0 \ 1]$. If we denote $\{v_1, v_2, v_3\}$ to be the right eigenvectors of R' , and their corresponding eigenvalues are $\{\lambda_1, \lambda_2, \lambda_3\}$, we have

$$R'v_i = R_0RR_0^T v_i = \lambda_i v_i \iff R(R_0^T v_i) = \lambda_i (R_0^T v_i).$$

From the equation above, the eigenvalues of R' are the same as the eigenvalues of R , i.e.,

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

and their associated eigenvectors can be represented by

$$v_i = R_0 u_i, \quad i = 1, 2, 3.$$

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

$$\begin{aligned} e_1 &= v_1 u_1 = [v_1, 0, 0], & e_2 &= v_2 u_2 = [0, v_2, 0], & e_3 &= v_2 u_3 = [0, 0, v_2], \\ e_4 &= v_3 u_2 = [0, v_3, 0], & e_5 &= v_3 u_3 = [0, 0, v_3] \in \mathbb{R}^{3 \times 3}. \end{aligned}$$

Any linear combination of the basis elements

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

will be a solution of the equation $R'R_0 - R_0R = 0$. Since we assume the normal to be in the direction of the z -axis, we get a solution for R_0

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

and T_0 is recovered up to the following form:

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

where $(I - R_0 R R_0^T)^\dagger$ is the pseudo inverse of $I - R_0 R R_0^T$. Since both v_2 and v_3 are in the null space of the matrix $I - R_0 R R_0^T$ (since their associated eigenvalues are $\lambda_2 = \lambda_3 = 1$), we have $\text{Null}(I - R_0 R R_0^T) = \text{span}\{v_2, v_3\}$.

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 on the yz -plane. If S itself is on a plane, we may choose the z -axis of the world frame to be the normal to the plane and the origin of the frame to be on 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. \square

Proof of Lemma 6. Notice that $T' = T_0 - R'T_0$, so we denote $T_0 = T_1 + T_2$, where T_1 is perpendicular to ω and T_2 is parallel to ω . It is easy to check that the parallel components of T_0 and $R'T_0$ cancel each other, hence T' is perpendicular to ω . \square

Proof of Proposition 7. We first denote the *Lyapunov map* $L: \mathbb{R}^{3 \times 3} \rightarrow \mathbb{R}^{3 \times 3}; R_0 \mapsto R'R_0 - R_0R$ associated to the condition

$$R' = R_0 R R_0^T \iff R'R_0 - R_0R = 0.$$

Without loss of generality, we assume the rotation is with respect to the z -axis, so R is of the form

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

with $\omega = [0, 0, 1]^T$ and $0 < \theta < \pi$. Hence it has three distinct eigenvalues $\{1, e^{+j\theta}, e^{-j\theta}\}$. The left eigenvector associated with $\mu_1 = 1$ is $u_1 = [0, 0, 1]$ that is real. The other two left eigenvectors associated with the two complex eigenvalues $\mu_2 = \bar{\mu}_3$ are $u_2 = [-i, 1, 0]$, $u_3 = [i, 1, 0]$ with $u_2 = \bar{u}_3$. Notice that $\text{Im}(u_2)$ is orthogonal to $\text{Re}(u_2)$.

For the R' , we have

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

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 which is a real vector. The other two eigenvectors associated with the two conjugate complex eigenvalues $\lambda_2 = \bar{\lambda}_3$ are v_2, v_3 , where $v_2 = \bar{v}_3$. Furthermore, $\text{Im}(v_2)$ is orthogonal to $\text{Re}(v_2)$ since R_0^T is a rotation matrix.

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

$$v_1 u_1 = [0, 0, v_1] \in \mathbb{R}^{3 \times 3},$$

$$v_2 u_2 = [-iv_2, v_2, 0], \quad v_3 u_3 = [iv_3, v_3, 0] = [i\bar{v}_2, \bar{v}_2, 0] \in \mathbb{C}^{3 \times 3}.$$

The real and imaginary parts of the two complex bases

$$\text{Re}[v_2 u_2] = [\text{Im}(v_2), \text{Re}(v_2), 0],$$

$$\text{Im}[v_2 u_2] = [-\text{Re}(v_2), \text{Im}(v_2), 0],$$

$$\text{Re}[v_3 u_3] = [\text{Im}(v_2), \text{Re}(v_2), 0] = \text{Re}[v_2 u_2],$$

$$\text{Im}[v_3 u_3] = [\text{Re}(v_2), -\text{Im}(v_2), 0] = -\text{Im}[v_2 u_2]$$

are also in the kernel of L . Then the real kernel of L is a 3-D space which has the set of bases

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

The solution of R_0 will be the linear combination of the bases:

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

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. We have shown that $\text{Im}(v_2)$ is orthogonal to $\text{Re}(v_2)$, then R_0 is of the form

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

where $\alpha \in \mathbb{R}$ is an arbitrary angle, and the sign of the third column is chosen such that $\det(R_0) = 1$.

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. \square

References

- [1] J. Weng, T.S. Huang, N. Ahuja, *Motion and Structure from Image Sequences*, Springer Verlag, 1993.
- [2] N.F. Troje, H.H. Bulthoff, How is bilateral symmetry of human faces used for recognition of novel views, *Vis. Res.* 38 (1) (1998) 79–89.
- [3] T. Vetter, T. Poggio, H.H. Bulthoff, The importance of symmetry and virtual views in three-dimensional object recognition, *Curr. Biol.* 4 (1994) 18–23.
- [4] D.P. Mukherjee, A. Zisserman, J.M. Brady, Shape from symmetry—detecting and exploiting symmetry in affine images, *Phil. Tran. R. Soc. Lond., Series A* 351 (1995) 77–106.
- [5] A. Francois, G. Medioni, R. Waupotitsch, Reconstructing mirror symmetric scenes from a single view using 2-view stereo geometry, in: *Proc. Inter. Conf. on Pattern Recognition*, Québec City, Canada, 2002, pp. 12–16.
- [6] H. Zabrodsky, D. Weinshall, Using bilateral symmetry to improve 3D reconstruction from image sequences, *Comput. Vis. Image Understand.* 67 (1997) 48–57.
- [7] A.Y. Yang, W. Hong, Y. Ma, Structure and pose from single images of symmetric objects with applications to robot navigation, in: *Proc. IEEE Int. Conf. on Robotics and Automation*, 2003, pp. 1013–1020.
- [8] W. Hong, A.Y. Yang, Y. Ma, On symmetry and multiple view geometry: structure, pose and calibration from a single image, *Inter. J. Comput. Vis.* 60 (2004) 241–265.
- [9] J. Shi, C. Tomasi, Good features to track, in: *Proc. Inter. Conf. on Computer Vision and Pattern Recognition*, 1994, pp. 593–600.
- [10] P. Meer, B. Georgescu, Edge detection with embedded confidence, *IEEE Transact. Pattern Anal. Mach. Intell.* 23 (12) (2001) 1351–1365.
- [11] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, A. Yezzi, Gradient flows and geometric active contour models, in: *Proc. IEEE Inter. Conf. on Computer Vision*, 1995, pp. 810–815.
- [12] R. Malladi, J.A. Sethian, B.C. Vemuri, Shape modeling with front propagation: a level set approach, *IEEE Trans. Pattern Anal. Mach. Intell.* 17 (2) (1995) 158–175.
- [13] J. Shi, J. Malik, Normalized cuts and image segmentation, in: *Proc. Inter. Conf. on Computer Vision and Pattern Recognition*, 1997, pp. 731–737.
- [14] D. Comanicu, P. Meer, Mean shift: a robust approach toward feature space analysis, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (2002) 603–619.
- [15] Y. Weiss, Segmentation using eigenvectors: a unifying view, in: *Proc. IEEE Inter. Conf. on Computer Vision*, 1999, pp. 975–982.
- [16] Y.C. Hecker, R.M. Bolle, On geometric hashing and the generalized Hough transform, *IEEE Trans. Syst. Man Cybernet.* 24 (9) (1994) 1328–1338.
- [17] Z. Zhang, Parameter estimation techniques: a tutorial with application to conic fitting, Technical Report No. 2676, INRIA, France (October 1995).
- [18] T. Kanade, Recovery of the three-dimensional shape of an object from a single view, *Artif. Intell.* 17 (1981) 409–460.
- [19] C.A. Rothwell, D.A. Forsyth, A. Zisserman, J.L. Mundy, Extracting projective structure from single perspective views of 3D point sets, in: *Proc. IEEE Inter. Conf. on Computer Vision*, 1993, pp. 573–582.
- [20] J. Koščeká, W. Zhang, Video compass, in: *Proc. Eur. Conf. on Computer Vision*, 2002, pp. 476–491.
- [21] T. Leung, J. Malik, On perpendicular texture or: Why do we see more flowers in the distance? in: *Proc. Inter. Conf. on Computer Vision and Pattern Recognition*, 1997, pp. 807–813.
- [22] L. Van Gool, M. Proesmans, A. Zisserman, Grouping and invariants using planar homologies, in: *Workshop on Geometrical Modeling and Invariants for Computer Vision*, Xidian University Press, 1995.
- [23] J. Mundy, A. Zisserman, D. Forsyth, *Applications of Invariance in Computer Vision*, Springer-Verlag, 1994.
- [24] H. Mitsumoto, S. Tamura, K. Okazaki, Y. Fukui, 3-D reconstruction using mirror images based on a plane symmetry recovering method, *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (9) (1992) 941–946.
- [25] T. Vetter, T. Poggio, Symmetric 3D objects are an easy case for 2D object recognition, *Spatial Vis.* 8 (1994) 443–453.

- [26] H. Zabrodsky, S. Peleg, D. Avnir, Symmetry as a continuous feature, *IEEE Trans. Pattern Anal. Mach. Intell.* 17 (12) (1995) 1154–1166.
- [27] G. Marola, On the detection of the axes of symmetry of symmetric and almost symmetric planar images, *IEEE Trans. Pattern Anal. Mach. Intell.* 11 (1) (1989) 104–108.
- [28] N. Kiryati, Y. Gofman, Detecting symmetry in grey level images: the global optimization approach, *Inter. J. Comput. Vis.* 29 (1) (1998) 29–45.
- [29] C. Sun, J. Sherrah, 3D symmetry detection using the extended Gaussian image, *IEEE Trans. Pattern Anal. Mach. Intell.* 19 (2) (1997) 164–169.
- [30] K. Kanatani, Comments on symmetry as a continuous feature, *IEEE Trans. Pattern Anal. Mach. Intell.* 19 (3) (1997) 246–247.
- [31] R. Glachet, J.T. Lapreste, M. Dhome, Locating and modelling a flat symmetric object from a single perspective image, *CVGIP: Image Understand.* 57 (2) (1993) 219–226.
- [32] W. Hong, Y. Yu, Y. Ma, Reconstruction of 3-D symmetric curves from perspective images without discrete features, in: *Proc. Eur. Conf. on Computer Vision*, 2004.
- [33] Y. Liu, R.T. Collins, A computational model for repeated pattern perception using Frieze and wallpaper groups, in: *Proc. Inter. Conf. on Computer Vision and Pattern Recognition*, 2000.
- [34] J.L. Mundy, A. Zisserman, Repeated structures: image correspondence constraints and 3D structure recovery, in: *Applications of invariance in Computer Vision*, 1994, pp. 89–106.
- [35] Y. Ma, J. Kořecká, S. Soatto, S. Sastry, *An Invitation to 3-D Vision: from Images to Models*, Springer-Verlag, New York, 2004.
- [36] D.M. Wuescher, K.L. Boyer, Robust contour decomposition using a constant curvature criterion, *IEEE Trans. Pattern Anal. Mach. Intell.* 13 (1) (1991) 41–51.
- [37] A. Jain, R. Dubes, *Algorithms for Clustering Data*, Prentice Hall, New Jersey, 1988.
- [38] A.Y. Yang, K. Huang, S. Rao, W. Hong, Y. Ma, Symmetry-based 3-D reconstruction from perspective images, Technical Report UILU-ENG-03-2204, University of Illinois at Urbana-Champaign, 1997.