

Geometric Segmentation of Perspective Images Based on Symmetry Groups *

Allen Y. Yang Shankar Rao Kun Huang Wei Hong Yi Ma

Electrical & Computer Engineering Department

University of Illinois at Urbana-Champaign

{yangyang, sr Rao, kunhuang, weihong, yima@uiuc.edu}

Abstract

Symmetry is an effective geometric cue to facilitate conventional segmentation techniques on images of man-made environment. Based on three fundamental principles that summarize the relations between symmetry and perspective imaging, namely, structure from symmetry, symmetry hypothesis testing, and global symmetry testing, we develop a prototype system which is able to automatically segment symmetric objects in space from single 2-D perspective images. The result of such a segmentation is a hierarchy of geometric primitives, called symmetry cells and complexes, whose 3-D structure and pose are fully recovered. Such a geometrically meaningful segmentation may greatly facilitate applications such as feature matching and robot navigation.

1 Introduction

Let us first examine the images shown in Figure 1. For

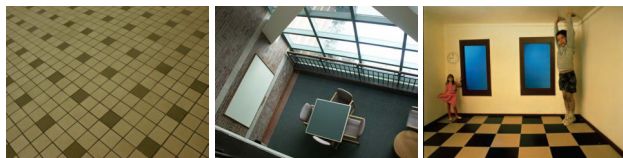


Figure 1: Left and middle: In man-made environment, symmetric patterns are abundant. Right: Ames room illusion.

the first two images, people have little difficulty finding the most probable 3-D interpretation for each image. We can easily describe the geometric relations among objects in the images, especially for objects which appear to be “regular enough.” Of course, our will and ability to perform such tasks will fail if our underlying assumptions are not valid, as indicated by the third image (the Ames room). So 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? In this paper, we summarize some of the key assumptions into three principles which essentially allow a computer to perform a similar task.

*This work is supported by UIUC ECE startup fund. The authors would like to thank anonymous reviewers for their valuable comments.

Structure from symmetry. One of the fundamental difficulties for deriving 3-D information from a single 2-D image is: Without knowing *anything* about 3-D geometry of a scene, there are infinitely many structures in space that may give rise to exactly the same image. To narrow down the infinite possible solutions to a unique one, we must impose extra assumptions. As one answer to this question, we derive the principle of “structure from symmetry”: If an object admits rich enough symmetry, no 3-D geometric information (including structure and pose) is lost through perspective imaging.

Symmetry hypothesis testing. However, this immediately leads to another fundamental difficulty: Given a region of interest, to what extent can we claim that it could be the image of an object with a certain kind of symmetry? One answer to this question requires us to understand how symmetry is precisely encoded through perspective imaging so that we can verify whether *all* geometric relations for a valid image of a symmetric object are satisfied in the region. Hence we deduce the principle of “symmetry hypothesis testing.”

Global symmetry testing. Nevertheless, a region that passes certain types of symmetry testing does not automatically imply that it must be the image of a symmetric object in space. Although each individual tile in the first image of Figure 1 passes any image based 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 deduce *3-D structure* from images. In fact, it is exactly this third principle that makes the Ames room illusion so compelling.

The goal of this paper is to study a *computational means* for the implementation of the above principles and develop a fully automatic system that is able to detect, extract and segment symmetric objects in terms of both symmetry types and geometric relations in space. Without loss of generality, in this paper, we introduce our algorithms and system only for planar symmetric objects. This simplification is based on the observation that planar symmetric objects are ubiquitous in man-made environment.

Relation to the literature. A lot of research has been done to recognize geometric structures with or without symmetry from images. The problem of local geometric cell ex-

traction is typically separated into two steps: image primitive detection and primitive parameter fitting. For the general problem of image primitive detection, there are feature-based methods, such as corner points [17] and lines [12], and pixel-based methods, such as active contours [8, 10] and region segmentation [16, 1, 20]. If the segments sought can be parameterized, several techniques have been introduced, such as the well-known Hough transform [3] or non-linear parameter estimation [24].

Different symmetry assumptions have also been studied to recognize and recover structures under perspective, orthogonal or affine projections. This paper is not the first to notice that symmetry, especially (bilateral) reflective symmetry, can be used to retrieve 3-D information. [13] first studied how to reconstruct a 3-D object using mirror image based planar symmetry, [2] provided a more complete study of the reflective symmetry, [19] proved that for any reflective symmetric 3-D object one non-accidental 2-D model view is sufficient for recognition, [23] used bilateral symmetry assumptions to improve 3-D reconstruction from image sequences, and [22] provided a good survey on studies of reflective symmetry and rotational symmetry in computer vision. In 3-D object and pose recognition, [15] pointed out that the assumption of reflective symmetry can also be used in the construction of projective invariants and is able to eliminate certain restriction on the corresponding points. For symmetry detection, [11, 9, 14] presented efficient algorithms to find axes of reflective symmetry in 2-D images, [18] discussed reflective symmetry detection in 3-D space, and [22] 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 [7]).

Contributions of this paper. Reflective, rotational, translational symmetries have been primarily studied independently. However, this paper shows that the key to consistent detection and segmentation of symmetric structures from their 2-D perspective images is analysis of the relations among all symmetries 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. These new types of geometric primitives can be used to replace corner or edge features and significantly simplify feature matching across multiple images. We have developed a fully automatic prototype system that is able to accomplish the above tasks.

2 Geometry for a single image of a planar symmetric structure

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 non-trivial subgroup G of the Euclidean group $E(3)$ that acts on it. That is, for any element $g \in G$, g defines an isomorphism (i.e. a one-to-one map) from S to itself: $g \in G : S \rightarrow S$. In particular, we have $g(S) = g^{-1}(S) = S$ for any $g \in G$.

[4] studied multiple-view geometry for general symmetric structures. However, it treated the planar case as a special case, and the results were not sufficient for our purposes here. In this paper, we provide some new characterization of the planar case that is more pertinent to symmetry detection and segmentation.

When the structure S is in a plane P , the symmetry group G is also a subgroup of the 2-D Euclidean group $E(2)$ (for the plane). With a choice of a coordinate frame¹ (x, y, z) attached to the structure S , any element $g = (R, T)$ in G can be represented (in the homogeneous representation) as a 3×3 matrix of the form

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

where $R \in \mathbb{R}^{2 \times 2}$ is an orthogonal matrix in $O(2)$ (“ R ” for rotation and reflection) and $T \in \mathbb{R}^2$ is a vector (“ T ” for translation). For any symmetric structure, there is a natural choice for the coordinate frame that makes this representation the simplest. Such a frame is called the *object centered canonical frame*, or simply the *object frame*.

Example 1 (The symmetry group of a rectangle) The symmetry group of a rectangle shown in Figure 2, the dihedral

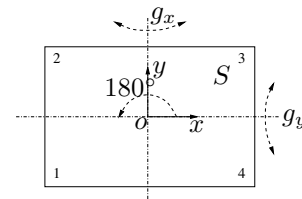


Figure 2: A rectangle whose symmetry includes reflections along the x and y axes and a rotation about o by 180° . These transformations form a dihedral group of order 2, i.e. D_2 .

group D_2 of order 2, can be represented with respect to the canonical object frame (x, y, z) by the following four matrices: $g_e = I$,

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

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

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

Note that the symmetry of the rectangle has no translational part (with respect to the object frame) and therefore in all g 's, $T = 0$.

¹We here choose the z -axis of the frame to be the normal to the plane such that any point on the plane is determined by its (x, y) coordinates.

2.1 Homography groups in an image

Now consider a perspective image of S that is taken by a camera at a vantage point $g_0 \in SE(3)$. Typically, g_0 is represented with respect to the object frame. Since S is planar, it is known from multiple-view geometry that there exists a homography $H_0 \in GL(3)$ depending on the vantage point g_0 which maps the plane $P \supseteq S$ in space to the image plane. Suppose that the vantage point g_0 is $(R_0, T_0) \in SE(3)$ and the plane P is defined by the equation $N^T \mathbf{X} = d$ for any point $\mathbf{X} \in \mathbb{R}^3$ on P . $N \in \mathbb{R}^3$ is the unit normal vector of the plane and $d \in \mathbb{R}_+$ is its distance to the center of the camera frame (e.g., see Figure 3). With a homography H_0

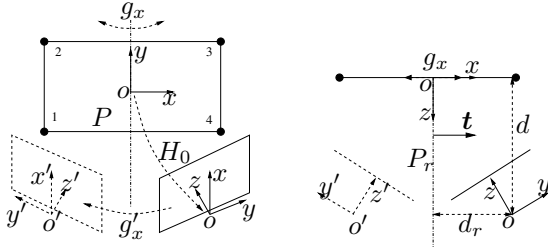


Figure 3: Equivalent images of a rectangle, before and after a reflection g_x . Left: frontal view; right: top view. P_r is the plane of reflection and t is its (unit) normal vector.

given by

$$H_0 \doteq R_0 + \frac{1}{d}T_0N^T \in \mathbb{R}^{3 \times 3}, \quad (2)$$

the coordinates of a point $\mathbf{X} = [x, y, z]^T$ on the plane are given by $g_0(\mathbf{X}) = H_0\mathbf{X}$. In essence, H_0 gives a matrix representation for taking an image of S at g_0 .

Due to the symmetry of S , we have $g(S) = S$ and therefore $H_0(g(S)) = H_0(S)$. For a particular point $\mathbf{X} \in S$, we have

$$H_0(g(\mathbf{X})) = H_0gH_0^{-1}(H_0(\mathbf{X})). \quad (3)$$

The quantity on the left can be viewed as the image of S taken at g_0 after a symmetry transformation g has been applied to S . This is equivalent to that the image is taken at a “new” vantage point g_0g . One image of S is then equivalent to $|G|$ such “new” images. These images will be referred to as *equivalent images* of the original one.

Between the structure S with a symmetry group G in space and its image $I(S)$ taken at g_0 , we can use the following commutative diagram to describe their relations

$$\begin{array}{ccc} S & \xrightarrow{H_0} & I(S) \\ g \downarrow & & \downarrow H_0gH_0^{-1} \\ S & \xrightarrow{H_0} & I(S) \end{array} \quad (4)$$

The group action of G on the plane P in space is then naturally represented by its conjugate group $G' \doteq H_0GH_0^{-1}$

acting on the image plane. We call it the *homography group*. Any element $g' = H_0gH_0^{-1} \in G'$ represents the homography transformation between two equivalent images, as shown in Figure 3 for a reflective symmetry.

Computationally, with more than 4 corresponding points between any pair of equivalent views, such as the 4 corner features of the rectangle in Figure 3, the associated homography matrix H' of g' can be linearly recovered from the following equation up to scale, where \mathbf{x}' is the image of its symmetric point $g(\mathbf{X})$

$$\mathbf{x}' \sim H'\mathbf{x} \Leftrightarrow \mathbf{x}' \times (H'\mathbf{x}) = 0. \quad (5)$$

Example 2 (The homography group of a rectangle) For the rectangle case studied in Example 1, the homography group $G' = H_0GH_0^{-1}$ is given by $\{I, H_0g_xH_0^{-1}, H_0g_yH_0^{-1}, H_0g_zH_0^{-1}\} \doteq \{I, g'_x, g'_y, g'_z\}$ and its elements satisfy the same set of relations as G in Example 1, since they are isomorphic:

$$\begin{aligned} (g'_x)^2 &= (g'_y)^2 = (g'_z)^2 = I, & g'_xg'_y &= g'_z, \\ g'_xg'_z &= g'_yg'_x = g'_y, & g'_yg'_z &= g'_zg'_y = g'_x. \end{aligned}$$

Although the above examples only show the reflective symmetry of a rectangle, the framework encompasses all three types of symmetry in a plane: *reflection, rotation, and translation*.

2.2 Consistent 3-D recovery from symmetry

Once the homography matrix $H' = H_0gH_0^{-1}$ is obtained from equivalent views, we can decompose it into

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

to obtain the relative pose (R', T') between the equivalent views. The 3-D structure of S can then be uniquely determined by triangulation. Furthermore, since H' and g are known, we may further use $H' = H_0gH_0^{-1}$ to recover information about the homography matrix $H_0 = R_0 + \frac{1}{d}T_0N^T$. H_0 obviously satisfies the following *Lypunov* type linear equation

$$H'H_0 - H_0g = 0, \quad \forall g \in G \quad (6)$$

with both H' and g now known. Once H_0 is solved, we can further decompose it into

$$H_0 \rightarrow \left\{ R_0, \frac{1}{d}T_0, N \right\}$$

to obtain the initial camera pose $g_0 = (R_0, T_0)$.

Hence, we may summarize the results so far as the following principle:

Principle 1 (Structure from symmetry) For any (planar) symmetric object, its 3-D structure and pose are already

encoded in a single 2-D perspective image. No 3-D information is lost through the perspective imaging, as long as the symmetric assumption is correct.

As we know from the previous section, 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 2 (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.*

Example 3 (Symmetry of a tiled floor) In practice, an object often admits all three types of symmetry, such as a tiled floor shown in Figure 4 left. In the figure, the three regions with corners marked are obviously translational copies of one another. Each region itself is a square whose symmetry group is the dihedral group of order 4, i.e. D_4 , which contains a rotation subgroup (the cyclic group C_4). The inter-relations between the homography are in-

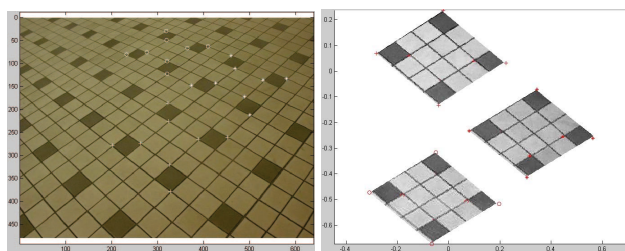


Figure 4: Left: An image of a tiled floor with all types of planar symmetry encoded. Right: Three regions on the floor recovered from the homography group among themselves.

duced from all three elements of the symmetry group, which allow a unique recovery of not only the 3-D structure and pose of each region but the relative 3-D positions between the three different regions, as the reconstruction result shows in Figure 4 right.

Lastly, we need to point out that there are several degenerated cases that only partial 3-D information of their structures can be retrieved from their images, and the interested reader is referred to [4].

3 Symmetry-based hierarchical image segmentation

Equipped with basic knowledge about symmetry and perspective imaging and two geometric principles, we are now ready to show how 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 sym-

metric objects in space, and recover relative 3-D geometric relations between these identified regions (e.g., as congruent tiles on the same 3-D plane etc.). Our remarkable ability to perceive such information from a single painting or photograph suggests that this is a viable task, at least for scenes consisting overwhelmingly of objects with “regular structures.”

Figure 5 outlines the architecture of a prototype system that we have developed to achieve these tasks. The hier-

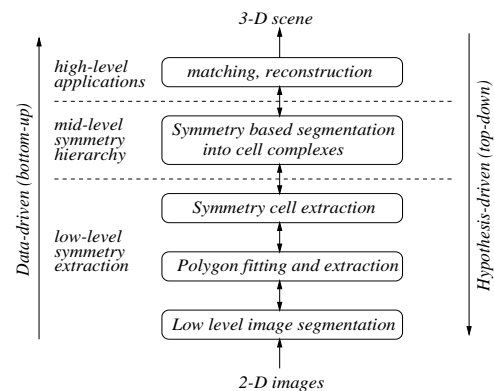


Figure 5: The system diagram of the symmetry based geometric segmentation algorithm.

archical architecture consists of three main levels. In this paper, we show how to realize the low-level symmetry cell extraction and middle-level symmetry hierarchy construction on images of quadrilateral structures. But again the given system also works on general symmetric polygons. The output of the system provides “geometric primitives” which encodes much richer 3-D geometric information than corner or edge features and facilitates high-level tasks such as matching features in different images. In Section 4, we will show one such example with details given in [5].

3.1 Low-level symmetry cell extraction

In order to “recognize” symmetry in an image, we first select some candidate sets of 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 the following cases: 1. one reflection symmetry ($G = \mathbb{Z}_2$); 2. a rectangle ($G = D_2$); 3. a square ($G = D_4$). Using the algorithm provided in the previous section, we can test the cell against all three hypotheses and determine which type of symmetry this cell admits. If it falls into the first category, we can simply discard it. A cell that belongs to any of the other categories is called a *symmetry cell*, with the type of symmetry labeled on it.

To make our explanation intuitive, we will use the image shown in Figure 6 to demonstrate the process of the overall

system. In Section 4, other examples will be given for more

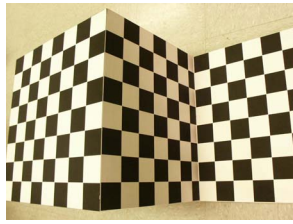


Figure 6: An image of a cube and a checkerboard.

generic scenes.

Image segmentation. We first use the color based mean shift algorithm [1] to obtain candidate regions for symmetry cells from the image.² Second, for every region that is larger than a minimal size, we compute the convex hull of the region in order to guarantee its external contour to be convex and reduce the noise on its boundary.

Polygon fitting and extraction. Next, we fit the extracted convex hulls from image segmentation with a polygon. 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 by [21]. From the constant curvature criterion, we directly know how many edges each convex region contains. In this paper, our system extracts only quadrilaterals.

The output from these two steps is the image coordinates of line segments of polygonal regions. The result for Figure 6 is shown in Figure 7. Compared to the original image,

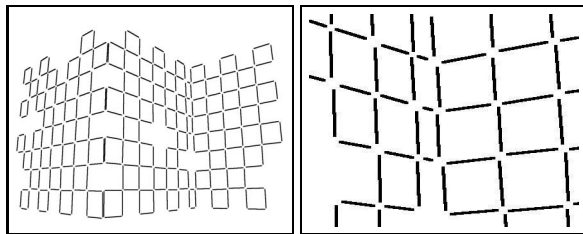


Figure 7: Left: The output of polygon extraction. Right: Details.

some of the “cells” are missing due to noise.

Local symmetry testing. For each quadrilateral extracted from above, using the symmetry hypothesis testing principle given in the previous section, we can test whether it satisfies the symmetry of a rectangle or a square. However, in real images, the boundaries obtained from segmentation can be extremely noisy sometimes, as examples in Section 4 show. As a more practical testing criterion, say for a rectangle, we may compute the three non-identity homography matrices H'_x, H'_y, H'_z and if the reconstructed plane normals

$$\max\{\cos^{-1}(N_x^T N_y), \cos^{-1}(N_x^T N_z), \cos^{-1}(N_y^T N_z)\} \leq \tau$$

²The two primary parameters for the mean shift algorithm are set to be $\sigma_S = 7, \sigma_R = 9$.

for some small threshold τ (in our experiments $\tau = 15^\circ$), we claim that the polygon passes the test for symmetry as a rectangle. Otherwise, we discard this polygon as non-symmetric.

If the cell size is small (as in this example, each cell is only about 1/20 the size of the image), due to a low SNR, some of the cells will fail the above symmetry test. To improve accuracy, we may use nonlinear programming to optimize the homography matrices subject to *all* the group relations that they are supposed to satisfy. The result of the symmetry test for the extracted polygons of Figure 7 is shown below in Figure 9 left.

3.2 Middle-level symmetry hierarchy: geometric segmentation

A cell that passes the symmetry test of a rectangle or a square does not necessarily corresponds 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:

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

For example, if a number of neighboring symmetry cells have mutually consistent normals, then 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 on a larger scale. For example, two neighboring cells can be translational copies of one another in space, which is often the case for window patterns on the side of a building. This leads to the notion of a *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.

Orientation clustering. To identify the orientations of the obtained planar symmetry cells, we can classify the spherical coordinates of their normal vectors using any standard clustering algorithm. In our current implementation, we use the ISODATA algorithm [6], known for its simplicity and efficiency³. For the above image, the clustering result of the space angles (θ, ϕ) of the normal vectors of all symmetry cells is shown in Figure 8. The two primary groups correspond to symmetric cells of the two main orientations in space. The remaining cells are outliers due to noise and polygon fitting errors.

Topology. With the orientations classified, symmetry cells can be further related or separated by their neighborhood relations (within each orientation group). For example, for two cells in the same orientation group, if their edges or corners are adjacent to each other (say ≤ 50 pixels), they can be

³The inter-cluster angle distance is chosen to be $\pi/12$.

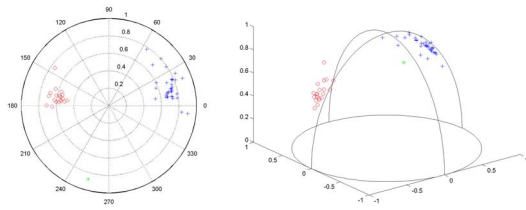


Figure 8: Clusters of normals of symmetry cells.

considered as “connected.” This will segment the symmetry cells into different groups in terms of their connectivity.

Coplanarity. For a set of neighboring cells of the same orientation, we can further verify if they are also coplanar by considering the translational symmetry among them. From Principle 2 and Example 4 in the previous section, if two cells are coplanar, the plane normal recovered from the translational homography between the cells needs to be the same as the normal of the cells. This allows us to test if a set of cells of the same orientation are coplanar. Obviously, there are two planes in Figure 6 where cells have the same orientation but are not coplanar. The final segmentation result for the cells in this image is shown in Figure 9 right.

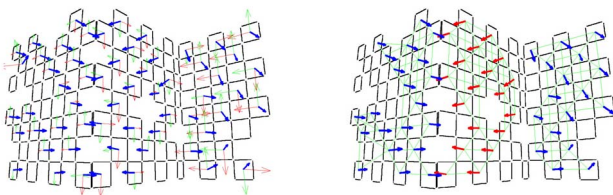


Figure 9: Left: Cells that pass the local symmetry test are marked with their 3-D object frame (x, y, z)-axes drawn in red, green and blue arrows, respectively. Right: The final output of the symmetry-based geometric segmentation. The normal vectors are drawn in different colors based on their grouping. Connectivity and coplanarity are illustrated as connected green lines between cells.

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 Figure 10. Cells are the leaves of the tree. At each level of the hierarchy, cells and complexes are ordered by their sizes, distances, or neighboring topology and labeled with their symmetry types. 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 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.

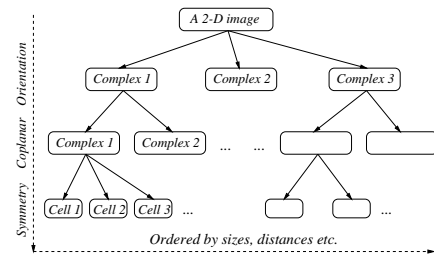


Figure 10: A hierarchical segmentation of a 2-D image by 3-D geometric properties.

4 Experiments and applications

In this section, we demonstrate the performance of the prototype system by applying it to a variety of images, shown in Figure 11. 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 one minute on a 1.4GHz PC in Matlab, which includes the mean shift segmentation and polygon fitting. The computation of symmetry hierarchy without nonlinear optimization takes less than five seconds, but the optimization version usually need several minutes depending on how many cells are extracted from the image.

In Figure 11, the first image is the side of a corridor in a library. Despite the noisy background, the side of the cubicle passes the symmetry test and eventually becomes the only cell extracted. The second and third images both return multiple cells and complexes which are obviously the most salient rectangular objects in the scene. Notice that rectangles with inhomogeneous internal color pattern are also correctly extracted. On the other hand, in the second image next to the TV stand, an illusory rectangle is extracted – which is created by occluded edges from surround objects – one of many manifestations associated with a single view. The last image is the most challenging one: cell boundaries are not so well-defined and 3-D geometric relations among cells are not precise. The system still does a decent job of extracting and segmenting most cells.

Applications in matching. If multiple images of one scene are given, it is much easier to establish correspondence between cells and complexes than using only pure point or line features. This is because the matching now can use both the 3-D geometric information and texture of the cells (or complexes). There is no need for any iterative robust statistic methods (such as RANSAC for point features) and the allowable motion between images can be very large. Once two cells in two different images are matched, the relative 3-D camera pose is automatically recovered. Figure 12 shows one such example. For details on matching, see [5].

5 Conclusions

This paper demonstrates that it is computationally feasible to segment and represent an image of a man-made environment based on accurate 3-D geometric information encoded in the image itself. The key to this is that symmetry groups admitted by (planar) symmetric objects in space are precisely encoded in the image through the so-called homography groups. This stipulates that no 3-D information is lost through 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 system is also hierarchical. This allows the development of algorithms for each layer to be relatively independent and easily replaceable. Computationally, a strong coupling (or feedback) is introduced between different layers through multiple hypothesis testings. 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. Tested on a variety of images, the system is robust and accurate.

In the future, we will try to extend our system to work under varying lighting conditions, camera calibration, occlusion, and non-planar symmetric objects. Such a system will have a wide range of applications in image matching, object recognition, and robot navigation.

References

- [1] D. Comanicu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, pages 603–619, 2002.
- [2] A. Francois, G. Medioni, and R. Waupotitsch. Reconstructing mirror symmetric scenes from a single view using 2-view stereo geometry. In *Proceedings of Int. Conference on Pattern Recognition*, 2002.
- [3] Y. C. Hecker and R. M. Bolle. On geometric hashing and the generalized hough transform. *IEEE Transactions on systems, man, & cybernetics*, 24(9):1328–1338, 1994.
- [4] W. Hong, A. Y. Yang, and Y. Ma. On symmetry and multiple view geometry: Structure, pose and calibration from a single image. *Int. Journal on Computer Vision*, submitted 2002.
- [5] K. Huang, W. Hong, A. Y. Yang, and Y. Ma. Symmetry-based 3-D reconstruction from perspective images: matching and reconstruction. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, submitted 2003.
- [6] A.K. Jain and R.C. Dubes. *Algorithms for Clustering Data*. Prentice Hall, 1988.
- [7] K. Kanatani. Comments on symmetry as a continuous feature. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 19(2):246–247, 1997.
- [8] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi Jr. *Gradient flows and geometric active contour models*. Technique report, University of Minnesota, 1994.
- [9] N. Kiryati and Y. Gofman. Detecting symmetry in grey level images: The global optimization approach. *Int. Journal on Computer Vision*, 29(1):29–45, 1998.
- [10] R. Malladi, J. A. Sethian, and B. C. Vemuri. Shape modeling with front propagation: a level set approach. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 17(2):158–175, 1995.
- [11] G. Marola. 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, 1989.
- [12] P. Meer and B. Georgescu. Edge detection with embedded confidence. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 23(12):1351–1365, 2001.
- [13] H. Mitsumoto, S. Tamura, K. Okazaki, and Y. Fukui. 3-D reconstruction using mirror images based on a plane symmetry recovering method. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 14(9):941–946, 1992.
- [14] D. P. Mukherjee, A. Zisserman, and J. M. Brady. Shape from symmetry—detecting and exploiting symmetry in affine images. *Phil. Trans. Royal Soc. London A*, 351:77–106, 1995.
- [15] C. A. Rothwell, D. A. Forsyth, A. Zisserman, and J. L. Mundy. Extracting projective structure from single perspective views of 3D point sets. In *Proceedings of IEEE International Conference on Computer Vision*, pages 573–582, 1993.
- [16] J. Shi and J. Malik. Normalized cuts and image segmentation. In *Proceedings of Int. Conference on Computer Vision & Pattern Recognition*, pages 731–737, 1997.
- [17] J. Shi and C. Tomasi. Good features to track. In *Proceedings of Int. Conference on Computer Vision & Pattern Recognition*, pages 593–600, 1994.
- [18] C. Sun and J. Sherrah. 3D symmetry detection using the extended gaussian image. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 19(2):164–169, 1997.
- [19] T. Vetter and T. Poggio. Symmetric 3D objects are an easy case for 2d object recognition. *Spatial Vision*, 8:443–453, 1994.
- [20] Y. Weiss. Segmentation using eigenvectors: A unifying view. In *Proceedings of IEEE International Conference on Computer Vision*, pages 975–982, 1999.
- [21] D. M. Wuescher and K. L. Boyer. Robust contour decomposition using a constant curvature criterion. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 13(1):41–51, 1991.
- [22] H. Zabrodsky, S. Peleg, and D. Avnir. Symmetry as a continuous feature. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 17(12):1154–1166, 1995.
- [23] H. Zabrodsky and D. Weinshall. Using bilateral symmetry to improve 3D reconstruction from image sequences. *Comp. Vision and Image Understanding*, 67:48–57, 1997.
- [24] Z. Zhang. *Parameter estimation techniques: a tutorial with application to conic fitting*. INRIA, 1995.

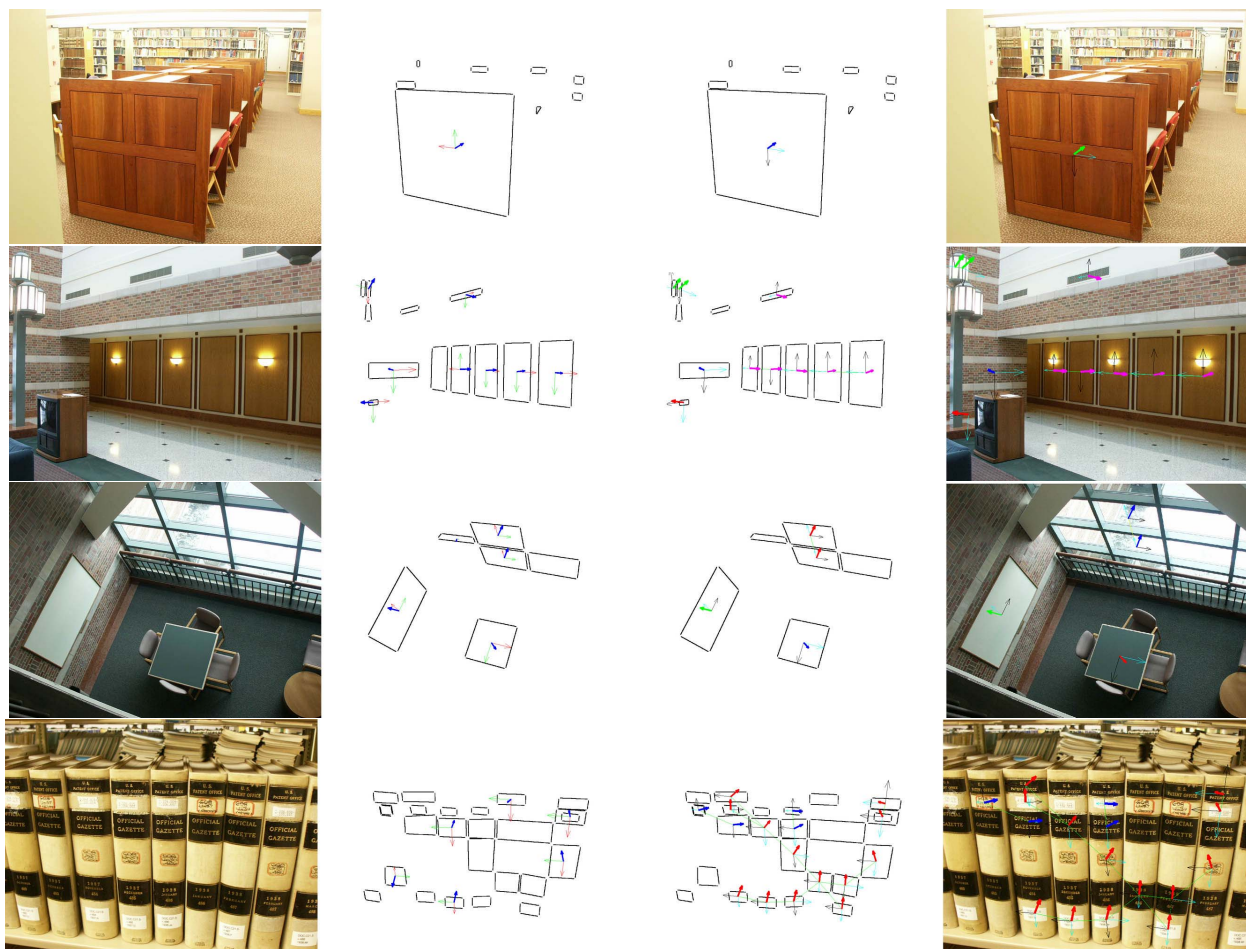


Figure 11: Experiments. Left: original images; Middle left: geometric cell extraction (1 minute) and local symmetry testing without nonlinear optimization (< 5 seconds), cells without frame attached fail the test; Middle right: local symmetry testing with nonlinear optimization. Right: grouped symmetry 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.



Figure 12: A unique matching and reconstruction of cells across images: from the raw images to symmetry cell extraction, matching, and 3-D cell structure and camera pose recovery, the whole process is fully automatic. Notice that conventional feature point matching or tracking techniques will fail with these images because the relative rotation is large but baseline is almost zero and symmetries in the scene would induce many ambiguous matches and generate outliers for the RANSAC type techniques.