Introduction

People have no difficulty finding the most probable 3-D interpretation from single images. One image of a symmetric object is equivalent to multiple images!

Contributions of this paper

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. For the first time, we are able to segment images based on the precise and consistent 3-D geometry of the segmented regions. The output is a hierarchy of geometric primitives whose 3-D information is fully recovered.

From previous papers, we have the following definition:

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, \( g : S \to S \). In particular, we have \( g(S) = g^{-1}(S) = S \) for any \( g \in G \).

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

\[
H' = \begin{pmatrix} R' & t' \end{pmatrix} N
\]

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 \( H_0 = R_0 + \frac{1}{d}T_0N \). \( H_0 \) obviously satisfies the following Lyapunov type linear equation

\[
H' H_0 - H_0 g = 0, \quad \forall g \in G
\]

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

\[
H_0 = \begin{pmatrix} R_0 & t_0 \end{pmatrix} N
\]

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

Hierarchical Geometric Segmentation

Step 1 (Local region extraction): After the original color image feeds into the system, a mean-shift algorithm is applied to get homogeneous texture regions. We compute the convex hull of the texture regions to reduce the boundary noise.

Step 2 (Polygon fitting): Next we fit the convex hulls with polygons. Our technique is based on the constant curvature criterion and the polygon vertices are extracted to be used to decompose the homography matrices.

Setp 3 (Local symmetry testing): For each quadrilateral extracted, using the symmetry hypothesis testing principle, we can test whether it satisfies the symmetry of a rectangle by considering its three normal vectors:

\[
\max(\|\omega_{x} N_1^0\|, \|\omega_{y} N_2^0\|, \|\omega_{z} N_3^0\|) \leq \varepsilon
\]

Step 4 (Global symmetry testing): As an example, we here test the following three global constraints: Orientation clustering, Connectivity and Coplanarity.

The polygon fitting and global testing results are shown on the right.

Symmetry Group

The symmetry group of a rectangle \( S \) is shown on the right. The dihedral group \( D_4 \) can be represented with respect to the canonical object frame \((x,y,z)\) by the identity matrix and the following three matrices:

\[
g_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad g_y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad g_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\]

Now consider a perspective image of \( S \) that is taken by a camera at a vantage point \( g_0 \). Since \( S \) is planar, there exists a homography \( H_0 \). Due to the symmetry of \( S \), we have \( g(S) = S \) and \( H_0(g(S)) = H_0(S) \).

Hierarchical Geometric Segmentation

The homography group \( G' = H_0 G H_0^{-1} \) is given by

\[
\{(I, H_0gH_0^{-1}) : g \in G \}
\]

The polygon fitting and global testing results are shown on the right.

Experiments with Generic Images

Structure from symmetry (fully automatic)

Image alignment and scene reconstruction (semi-automatic)