An Introduction to Holistic 3D Reconstruction

Yi Ma

EECS Department, UC Berkeley
What is 3D Reconstruction?
Applications of 3D Reconstruction

Image Source: Internet
Traditional 3D Reconstruction Pipeline

Feature Extraction & Matching → Multiview Geometry → Point Cloud

Image Source: Internet
Are Point Clouds Universal Representation?

Very Difficult to Storing, Computing, Editing, Visualizing, Interact, and Interpret.

A Gear Wheel Scanned by eviXscan 3D Pro

Streets Scanned by Velodyne Lidar

Image source: Velodyne website
Image source: “Diagnostics of machine parts by means of reverse engineering procedures”
Limitation of Traditional 3D Reconstruction

- Textureless Scenes
- Reflection/Transparency
- Repetitive Patterns
- Medium/Large Baseline (Correspondence Fail)
- Multiple Moving Objects
Data-Driven Learning-Based Approaches

Depth Map Regression
Li, Z., & Snavely, N. (2018)

Pose Estimation
Kehl, Wadim., et al. (2017)

3D Instance Segmentation

Plane Detection

Layout Prediction

Voxel Generation
Song, S., et al. (2017)

Mesh Generation
Recently research [1] suggests deep encoder-decoder networks do not perform reconstruction but classification.

For data-driven based depth recovery, DNN is not better (or even worse) than nearest neighbors (NN).

Equivalent to Image Classification?

Our World is Full of Structures

• Man-made environments are rich of structural regularities
  • straight lines
  • smooth curves
  • parallelism
  • orthogonality
  • Symmetry
  • Building code & grammar…
Importance of Structure in Human 3D Perception

- Structure: **spatial relationships** among multiple points, lines, patches, etc.
- Human perceives 3D space by recognizing **many types of structure** in the scene

Exploring Structures for 3D Reconstruction

**LOCAL:** face-edge-vertex graph, smooth curves & surfaces

**SEMI-GLOBAL:** symmetry, parallelism & orthogonality

**GLOBAL:** shape grammar

Image Source: Chen et al., 2007; Pauly et al., 2008
Image Source: https://stuckeman.psu.edu/adapting-modern-architecture-local-context
Exploring Local Structures -- What Does a 2D Line Drawing Tell Us about the 3D Geometry?

“Given a single picture ... we usually have definite ideas about the 3-D shapes of objects. To do this we need to use **assumptions about the world and the image formation process**, since there exist a large number of shapes which can produce the same picture.”

-- Takeo Kanade, 1981

[Sinha and Adelson 1993]
Exploring Local Structures – Some History of Single Line Drawing Interpretation

- **Prototype-based interpretation:** Roberts (1965), Falk (1972), Grape (1973)
- **Curved objects:** Turner (1974), Shapira and Freeman (1979), Lee et al. (1985), Malik (1987) …
- **Dynamic scenes:** Asada et al. (1984)
- ……
Exploring Local Structures – Line Labeling [Huffman-Clowes, 1971]

• Every line in natural pictures of **polyhedron objects** should have exactly one of the four labels
  - Convex (+), concave (−), or occluding (→, ←)
Exploring Local Structures – Junction Dictionary and Consistent Labeling [Huffman-Clowes, 1971]

- **12 valid configurations** for trihedral vertex
  - L-, Y-, W-types
  - Represents just 11.5% of all possible configurations

- **T-junction** occurs when an edge occludes another partially.
  - Does not correspond to a three-dimensional vertex.
Exploring Local Structures – Junction Dictionary and Consistent Labeling [Huffman-Clowes, 1971]

- **12 valid configurations** for trihedral vertex
  - L-, Y-, W-types
  - Represents just 11.5% of all possible configurations

- **T-junction** occurs when an edge occludes another partially.
  - Does not correspond to a three-dimensional vertex.
Exploring Local Structures – A linear Algebra Approach to 3D Reconstruction [Sugihara, 1982]

• Consider a picture which is obtained as the orthographic projection of the object
  • i-th vertex: \((x_i, y_i, z_i)\)
  • j-th face: \((a_j, b_j, c_j)\)

• 3D reconstruction can be formulated as estimating the unknowns \(z_i, a_j, b_j, c_j\)
Exploring Local Structures – A linear Algebra Approach to 3D Reconstruction [Sugihara, 1982]

Line label assignments to the picture provide two forms of constraints:

1. Vertex $i$ should be on the $j$-th face:
   
   $$a_j x_i + b_j y_i + z_i + c_j = 0$$

2. Vertex $t$ should be nearer than the $k$-th face
   
   $$a_k x_t + b_k y_t + z_t + c_k > 0$$
Exploring Local Structures – A linear Algebra Approach to 3D Reconstruction [Sugihara, 1982]

\[
Aw = 0
\]

\[
Bw > 0
\]

**Theorem:** A labeled line drawing represents a polyhedral scene if and only if the linear system has a solution.

In practice, there are usually infinite number of solutions ...
Exploring Local Structures – Shape Recovery via Optimization

• To resolve ambiguity, one option is to use additional cues such as shading [Sugihara, 1986]:

$$\text{min} \sum_{k} \gamma_k (d_k - \hat{d}_k(w))^2$$

subj. to $Aw = 0, \quad Bw > 0$

Lambertian surface with light source direction $l$:

$$d_k = L \cdot \cos \theta = \frac{L \cdot l \cdot n_k}{|l| \cdot |n_k|} = \frac{L \cdot l \cdot (a_k, b_k, 1)}{|l|\sqrt{(a_k)^2 + (b_k)^2 + 1}}.$$
Exploring Local Structures – Shape Recovery via Optimization

• Another option is to invoke additional structural priors, such as smoothness and regularity:

“To interpret a polygon in the image, we try to find a configuration of the vertices in space that makes the three-dimensional figure as regular as possible. Regularity might be measured in a variety of ways ... we prefer local features which are more likely to survive occlusion.”

-- Barrow and Tenenbaum, 1981

\[
\begin{align*}
\text{min } f(w) \\
\text{subj. to } Aw = 0, \quad Bw > 0
\end{align*}
\]

e.g., \( f(w) = \text{“sum of the squares of angles of faces”} \)
Exploring Local Structures – Some History of Additional Cues in Single Line Drawing Interpretation

- **Surface contour**: Stevens (1981), Barrow and Tenenbaum (1981), Marr (1982)
- **Texture**: Bajcsy and Lieberman (1976), Witkin (1981)
- **Vanishing points**: Nakatani and Kitahashi (1984)
- ......
Exploring Structures for 3D Reconstruction

**LOCAL:** face-edge-vertex graph, smooth curves & surfaces

**SEMI-GLOBAL:** symmetry, parallelism & orthogonality

**GLOBAL:** shape grammar

Image Source: Chen et al., 2007; Pauly et al., 2008
Image Source: https://stuckeman.psu.edu/adapting-modern-architecture-local-context
Symmetry Structures

Symmetry captures almost all “regularities”.

Symmetry Structures – Hidden Images from Rotation

Symmetry Structures

- Hidden Images from Rotation

**Diagram:**

- Diagram showing a transformation from one symmetry structure to another with labels and axes.

**Equation:**

- $g_0$

**Label:**

- $I_1$

- $I_2$

**Axes:**

- $x$, $y$, $z$

**Points:**

- $o$, $o_1$, $o_2$
Symmetry Structures – Hidden Images from Reflection
Symmetry Structures – Hidden Images from Translation
Symmetric Structure & Group

**Definition.** A set of 3-D features $S$ is called a symmetric structure if there exists a non-trivial subgroup $G$ of $E(3)$ that acts on it such that for every $g$ in $G$, the map

$$g \in G : S \rightarrow S$$

is an (isometric) automorphism of $S$. We say the structure $S$ has a group symmetry $G$.

- $X = [X, Y, Z, 1]^T \in \mathbb{R}^4$, $x = [x, y, z]^T \in \mathbb{R}^3$
- $g_0 = \begin{bmatrix} R_0 & T_0 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 4}$, $\Pi_0 = [I, 0] \in \mathbb{R}^{3 \times 4}$

$$x \sim \Pi_0 g_0 X \quad \Rightarrow \quad g(x) \sim \Pi_0 g_0 g X$$
Symmetry Structures – Hidden Multiple Views

\[ g_1(x) \sim \Pi_0 g_0 g_1 g_0^{-1}(g_0 X) \]
\[ g_2(x) \sim \Pi_0 g_0 g_2 g_0^{-1}(g_0 X) \]
\[ \vdots \]
\[ g_m(x) \sim \Pi_0 g_0 g_m g_0^{-1}(g_0 X) \]

\[ g = (R, T), g' = g_0 g g_0^{-1} \]

\[ g' : \begin{cases} 
R' \doteq R_0 R R_0^T \in O(3) \\
T' \doteq (I - R')T_0 + R_0 T \in \mathbb{R}^3
\end{cases} \]
Solving $g_0$ from Lyapunov equations:

$g_i'g_0 - g_0g_i = 0, \ i = 1, \ldots, m$

with $g_i'$ and $g_i$ known.

$g_1(x) \sim \Pi_0 g_0 g_1 g_0^{-1}(g_0 X)$

$g_2(x) \sim \Pi_0 g_0 g_2 g_0^{-1}(g_0 X)$

$\vdots$

$g_m(x) \sim \Pi_0 g_0 g_m g_0^{-1}(g_0 X)$
Symmetry Structures

3-D reconstruction with symmetry is simple, accurate and robust!
Symmetry Structures – Hidden Images in Each View

Symmetry on object

\[ g = (R, 0) \]
\[ R = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \]

Virtual camera-camera

\[ g' = (R', T') \]
\[ R' = R_0 R R_0' \]
\[ T' = (I - R') T_0 \]
Symmetry Structures – Reflective Homography

2 pairs of symmetric points

Reflective homography

\[ H = R' + \frac{1}{d} T' N^T \]

Decompose \( H \) to obtain \((R', T', N)\) and \(T_0\)

Solve Lyapunov equation

\[ R'R_0 - R_0 R = 0 \]

to obtain \( R_0 \).
Symmetry Structures – Alignment of Different Objects

\[ g_{21} = g_2^{-1} g_1 \]

\[ d_2 = 1 \]
Symmetry Structures – Scale Correction

For a point $p$ on the intersection line

$$\lambda_1 = \frac{d_1}{N_1^T \mathbf{x}} = \lambda_2 = \frac{d_2}{N_2^T \mathbf{x}}$$

$$\alpha = \frac{d_2}{d_1} = \frac{N_1^T \mathbf{x}}{N_2^T \mathbf{x}}$$

$$g_2 \leftarrow (R_2, \alpha T_2)$$

$$g_{21} = g_2^{-1} g_1$$
Symmetry Structures – Alignment of Different Views

\[ g_{21} = g_2^{-1} g_1 \]

\[ d_1 = 1 \]

\[ d_2 = ? \]
Symmetry Structures – Scale Correction

\[
\begin{bmatrix}
    x_i - \bar{x} \\
    y_i - \bar{y}
\end{bmatrix} = \alpha \begin{bmatrix}
    u_i - \bar{u} \\
    v_i - \bar{v}
\end{bmatrix}
\]

\[i = 1, 2, 3, 4\]

\[d_2 = \alpha\]

\[g_2 \leftarrow (R_2, \alpha T_2)\]

\[g_{21} = g_2^{-1} g_1\]

For any image \(x_1\) in the first view, its corresponding image in the second view is:

\[x_2 \sim R_2 R_1^T \left( \frac{1}{N_1^T} x_1 - T_1 \right) + T_2\]
Symmetry Structures – Alignment of Multiple Views

Method is **object-centered and baseline-independent**.
Symmetry Structures – Experiment Results

\[ \alpha = 0.7322 \quad \beta = 90.36^\circ \]
Symmetry Structures – Experiment Results

\[ \alpha = 0.7433 \]
Symmetry Structures – Image Transfer
Symmetry Structures – Camera Pose
Symmetry Structures – Full 3D Model
Reference on Multiview Geometry of Junctions, Lines, Planes, and Symmetries

An Invitation to 3D Vision, Yi Ma, S. Soatto, J. Kosecka, and S. Sastry
Exploring Structures for 3D Reconstruction

**LOCAL:** face-edge-vertex graph, smooth curves & surfaces

**SEMI-GLOBAL:** symmetry, parallelism & orthogonality

**GLOBAL:** shape grammar

Image Source: Chen et al., 2007; Pauly et al., 2008
Image Source: https://stuckeman.psu.edu/adapting-modern-architecture-local-context
Exploring Global Structures – Shape Grammar
Summary

• Here we outline the rest of the tutorial.

• A holistic 3D reconstruction pipeline consists of three main steps:
  1. Structure type identification, i.e., what types of structure are there in the scene?
  2. Structure instance identification, i.e., where are the instances of such structure in the image?
  3. Structure-based 3D reconstruction, i.e., how can we infer the 3D geometry from the detected structure instances?

We have focused on Step 3 so far. The rest of the tutorial will discuss Steps 1 and 2.
Example Problems in Structure Type Identification

• Local structures
  • Are the objects polyhedrons, smooth/curved surfaces, piece-wise planar, or some combination of those?

• Semi-global structures
  • What types of symmetry are there?
  • Manhattan world? Atlanta world? Something else?

• Global structures (shape grammar):
  • What rules are used (known as inverse procedural modeling)?
Example Problems in Structure Instance Detection

• Local structures:
  • Build the face-edge-vexter graph, i.e., via junction detection, line detection, face identification, etc.
  • Estimate the parameters of the geometric primitives involved

• Semi-global structures:
  • Symmetry detection
  • Vanishing point detection

• Global structures:
  • Procedural reconstruction