Shape from X

• One image:
  • Shading
  • **Texture**

• Two images or more:
  • Stereo
  • Contours
  • Motion
Shape From X

- One image:
  - Shading
  - Texture

- Two images or more:
  - Stereo
  - Contours
  - Motion
Shape From Texture

Recover surface orientation or surface shape from image texture:

• Assume texture ‘looks the same’ at different points on the surface.

• This means that the deformation of the texture is due to the surface curvature.
Structural Shape Recovery

Basic hypothesis: Texture resides on the surface and has no thickness.

→ Computation under:
  - Perspective projection
  - Paraperspective projection
  - Orthographic projection
Reminder: Perspective Projection

\[ u = f \frac{x}{z} \]
\[ v = f \frac{y}{z} \]
The perspective projection distortion of the texture
• depends on both depth and surface orientation,
• is anisotropic.
Depth vs Orientation:

• Infinitesimal vector \([Δx, Δy, Δz]\) at location \([x, y, z]\). The image of this vector is

\[
\frac{f}{z} [Δx - \frac{x}{z} Δz, Δy - \frac{y}{z} Δz]
\]

• Two special cases:
  • \(Δz=0\) : The object is scaled
  • \(Δx=Δy=0\) : The object is foreshortened
Reminder: Orthographic Projection

Special case of perspective projection:
- Large f
- Objects close to the optical axis
  $\rightarrow$ Parallel lines mapped into parallel lines.

$u = sx$
$v = sy$
Orthographic Projection

Center of Projection

Image

Scaling

Object

Orthographic Projection
Tilt And Slant
Orthographic Projection

- **Tilt:** Derived from the image direction in which the surface element undergoes maximum compression.

- **Slant:** Derived from the extent of this compression.
Cheetah

Perpendicular Lines

Orthographic projections of squares that are rotated with respect to each other in a plane inclined at $\omega=60^\circ$ to the image plane.

$$\frac{\left\| (p_1/l_1) \times (p_2/l_2) \right\|}{\left\| p_1/l_1 \right\|^2 + \left\| p_2/l_2 \right\|^2} = \frac{\cos(\omega)}{1 + \cos^2(\omega)}$$
Paraperspective Projection

Generalization of the orthographic projection:

- Object dimensions small wrt distance to the center of projection.

→ Parallel projection followed by scaling
For planar texels:

\[ A' = \frac{f^2}{3z_0} \mathbf{n} \cdot [x_0 y_0 z_0] A \]
Paraperspective Projection

Texels:

- Image regions being brighter or darker than their surroundings.
- Assumed to have the same area in space.

→ Given enough texels, it becomes possible to estimate the normal.
Texture Gradient
Mesure texture density as opposed to texel area, that is, the number of textural primitives per unit surface.

Assuming the texture to be homogeneous, we have:

\[
\psi \mathbf{n} \propto \mathbf{b}
\]

\[
\psi = \begin{bmatrix}
u_1 & v_1 & 1 \\
\vdots & \vdots & \vdots \\
u_n & v_n & 1
\end{bmatrix}^t
\]

\[
\mathbf{b} = [b_1, \ldots, b_n]^t
\]

\[
\Rightarrow \mathbf{n} = \frac{\psi \mathbf{n}}{\|\psi \mathbf{n}\|}
\]
Machine Learning

Input Image

Superpixels

Train a regressor to predict depth —> Noisy predictions
Markov Random Field (MRF)

Graph with vertices and edges

Assign values to the nodes to minimize

\[ E(Y) = \sum_i \varphi(y_i) + \sum_{(i,j)} \psi(y_i, y_j) \]

Unary \hspace{1cm} Pairwise

\[ \rightarrow \text{Enforces consistency} \]
Deep Learning with MRF

Liu et al., PAMI 2016
The network can be designed to enforce normal consistency.
But only for the class of scenes it has been trained for.
Normals from a Single Image

Input  Ground Truth  Output

Wang et al., CVPR 2015
Forcing the deep net to be consistent across tasks increases robustness.

Zamir et al., CVPR’20
A Very Diverse Training Database Helps

Eftekhar et al., ICCV’21 vs Chen et al., CVPR’20
.. and so does a Transformer Architecture
People seem to be sensitive to orientation fields in the cases of both texture and shading.
Optional: Shape from Smear

**Hypothesis:** If orientation and scale fields are the key source of information for 3D shape perception, it should be possible to induce a vivid sense of 3D shape by creating 2D patterns with appropriate scale and orientation fields.

**Test:** Use a technique known as Line Integral Convolution to smear the texture along specific orientations and scale appropriately.
Optional: Scaling and Smearing

Scaling:

Smearing:
Optional: Inconsistent Stimulus

The orientation field cannot be integrated

➢ No depth perception.
➢ Do we integrate in our heads?
➢ Is this what the deep nets learn to do?
Strengths and Limitations

Strengths:
• Emulates an important human ability.

Limitations:
• Requires regular texture.
• Involves very strong assumptions.
• Deep learning can be used to weaken them.