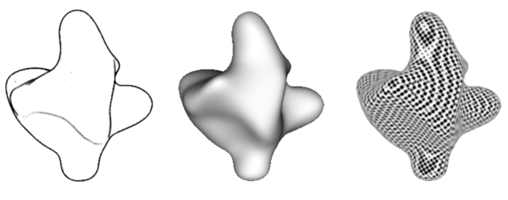
Shapecollage: occlusion-aware, example-based shape interpretation

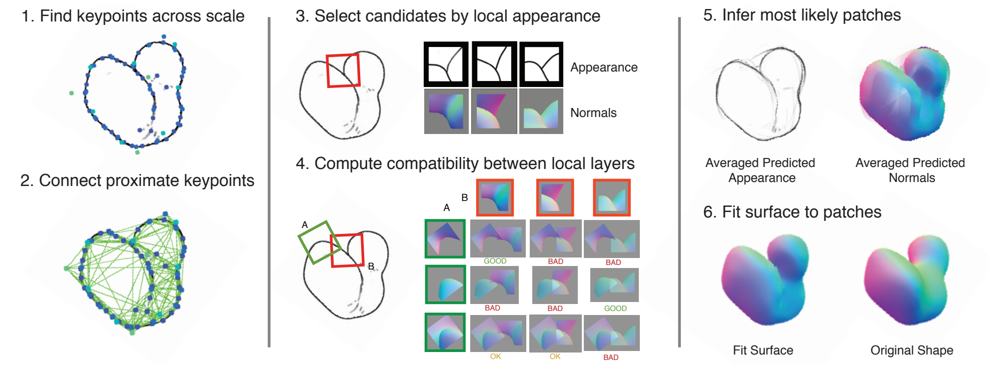
Forrester Cole, Phillip Isola, William T. Freeman, Frédo Durand, Edward H. Adelson
Massachusetts Institute of Technology

Goal: 3D shape interpretation from multiple styles of depiction

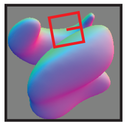


Approach: we explore an **example-based** approach that can interpret different styles of depiction within one system. **Occlusion boundaries** are a key challenge for the example-based approach, because they produce a combinatorial explosion of shape possibilities inside an image region. Our method avoids the combinatorial explosion by allowing patches to explain only the parts of the image they best fit, using an occlusion-aware patch representation and compatibility term.

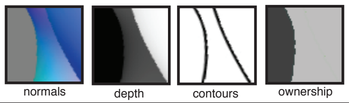
Method



Occlusion-Aware Patch Representation

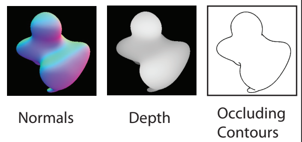


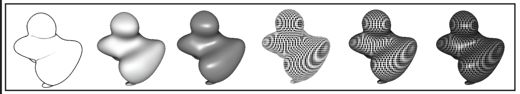
Each patch is extracted from an oriented square region. Besides appearance, the patch stores the normals oriented with the region, the depth and occluding contours, and an ownership mask encoding what part of the square region the patch explains.



Training Set

100 random shapes, 2000 total viewpoints

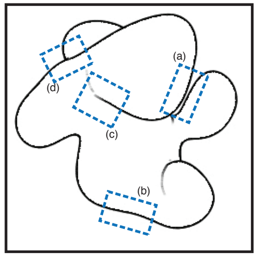




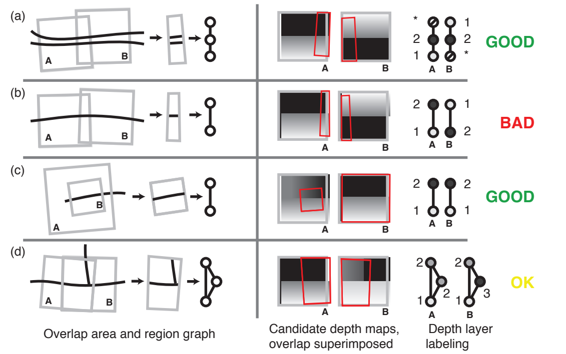
Six rendering styles

Local Depth Compatibility

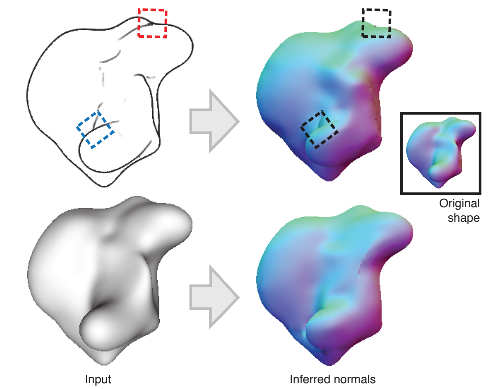
Our local depth order compatibility metric handles challenging areas by segmenting the overlap area into a graph (below left) and using the depth maps to assign an ordinal depth value to each segmented region (below right). Compatibility is determined by the similarity of the graph labelings, weighted by the segment area.



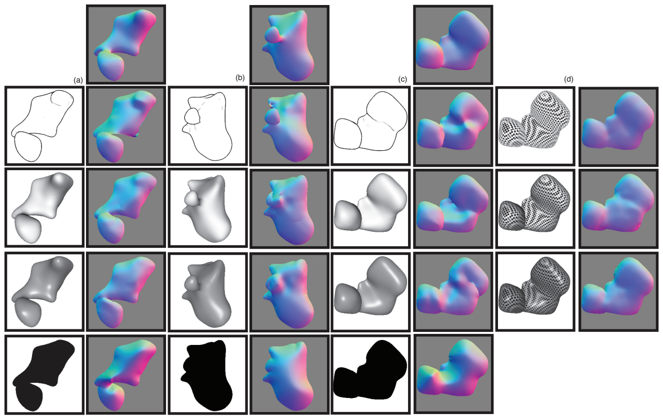
(a) the two shapes are compatible because the overlap graph allows for a gap.
(b) the two shapes are incompatible because they compete to label both sides of the single contour.
(c) the large scale patch has a single depth layer in the entire patch, but two layers inside the overlap region.
(d) at the T-junction the two labelings are slightly incompatible but the incompatible labeling has small weight.



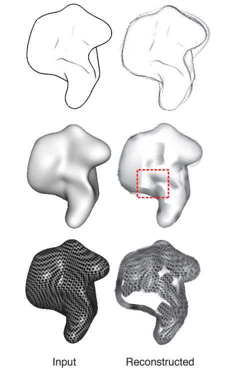
Results



A line drawing and diffuse rendering (left) are interpreted as normal maps by our system (right). Each rendering produces a different plausible interpretation. Local depth ordering is used to interpret self-occlusion (blue box) and near-occlusion (red box).



Examples of normal map estimation for varying rendering styles. Top: ground truth normal map. (a,b,c): line drawing, diffuse and glossy shading with solid material, with interpreted normal map. (d): same as (c) but with texture only, diffuse and glossy shading with texture. Bottom: normal interpolation from shape boundaries.

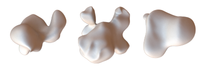


Reconstructing the stimulus from patch appearance. Reconstructed patches are usually faithful to the original rendering, though are sometimes confused (boxed). The styles used for reconstruction gives an estimate of the stimulus style.

Photographic input with complex lighting

With a training set constructed by rendering a realistic material model with a complex lighting environment, our system can extend to photographic input. In this example, the same sculpture was cast in different materials, providing ideal test images.

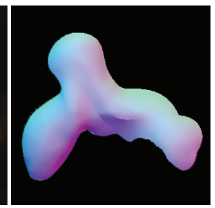
In future work, we hope to expand the training set to many other types of shapes, and to extend the appearance matching step to handle noise or clutter in the background.



Training images

Bronze version (photograph)

Jean Arp, S'accroupissant (photograph)



Interpreted normal map

Bronze re-rendering