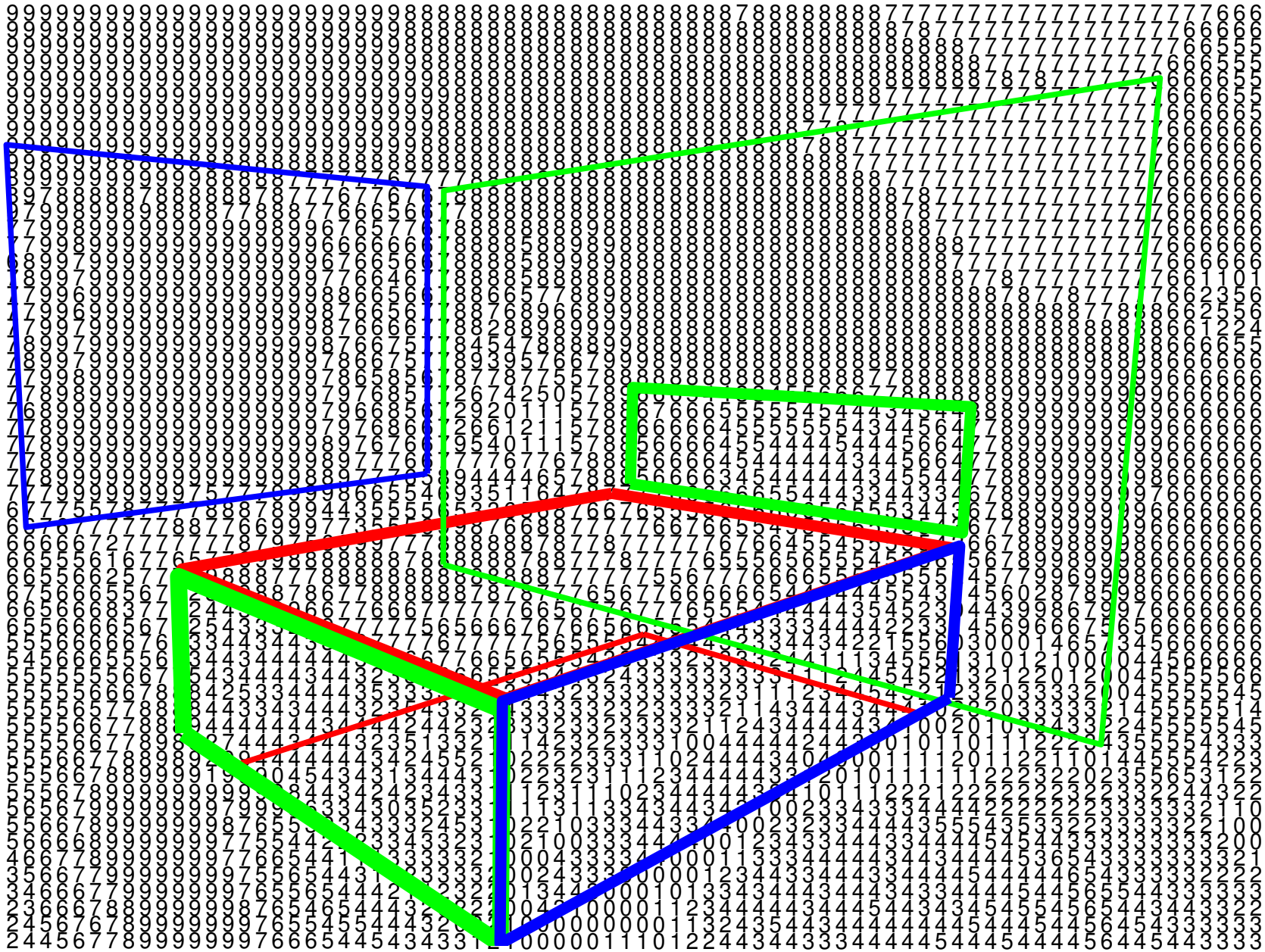
Inferring Spatial Layout from A Single Image via Depth-Ordered Grouping

Stella X. Yu, Computer Science, Boston College
Hao Zhang, Computer Science, UC Berkeley
Jitendra Malik, Computer Science, UC Berkeley

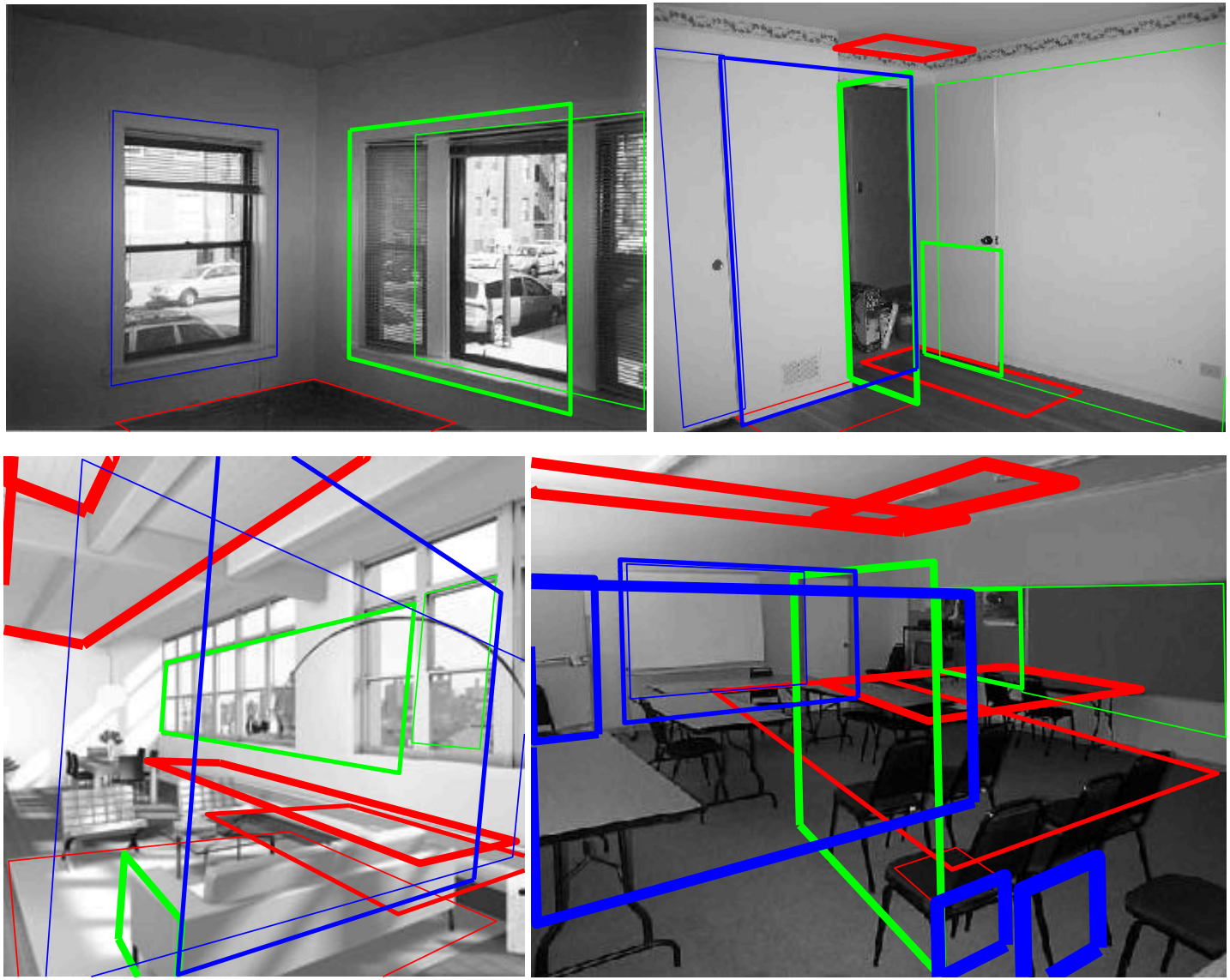
Input: A Single Image of An Indoor Scene

Task: Depth-Ordered Pixel Grouping

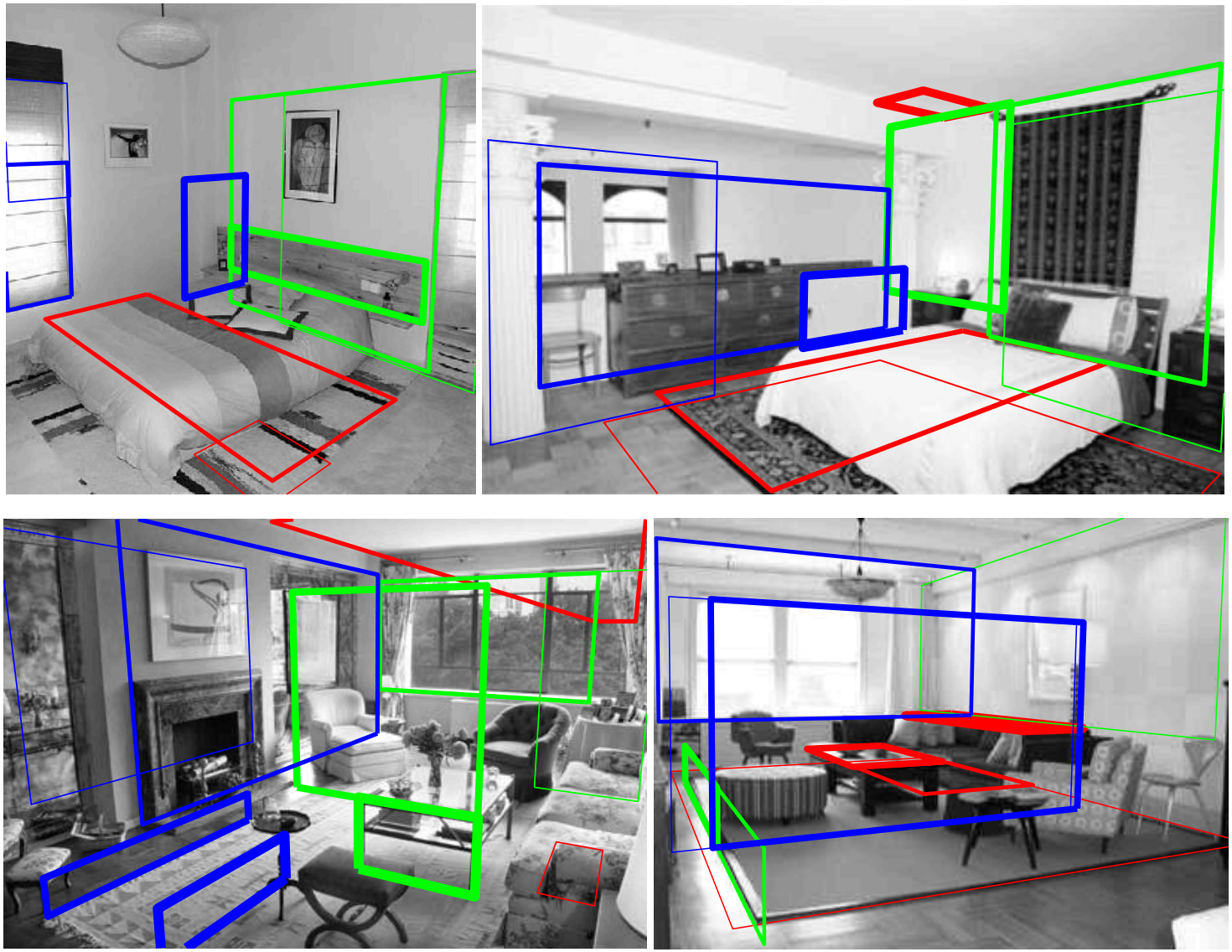


Visualization: 3D Scene from A New Viewpoint

Results: Scene Partition w. Depth-Ordered Planes



Results: Scene Partition w. Depth-Ordered Planes



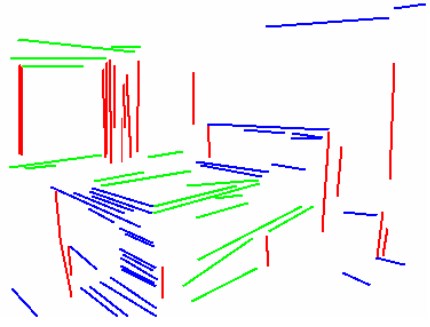
Line-Based Depth-Ordered Grouping Model

detection
localization

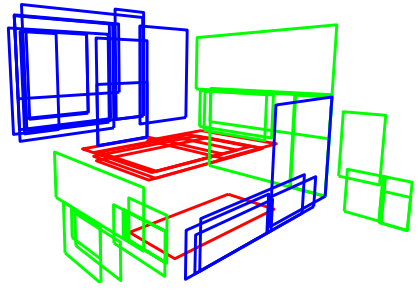
proximity
curvilinearity

Spectral Clustering with
Attraction and Repulsion
and Depth Ordering
Yu & Shi, 2001

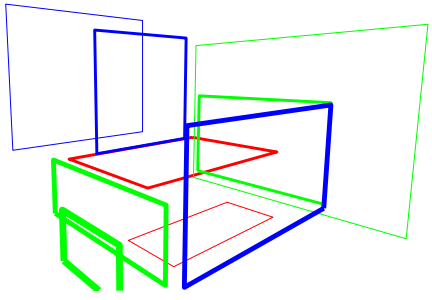
curvilinearity
parallelism
orthogonality
perspectivity



proximity
orthogonality



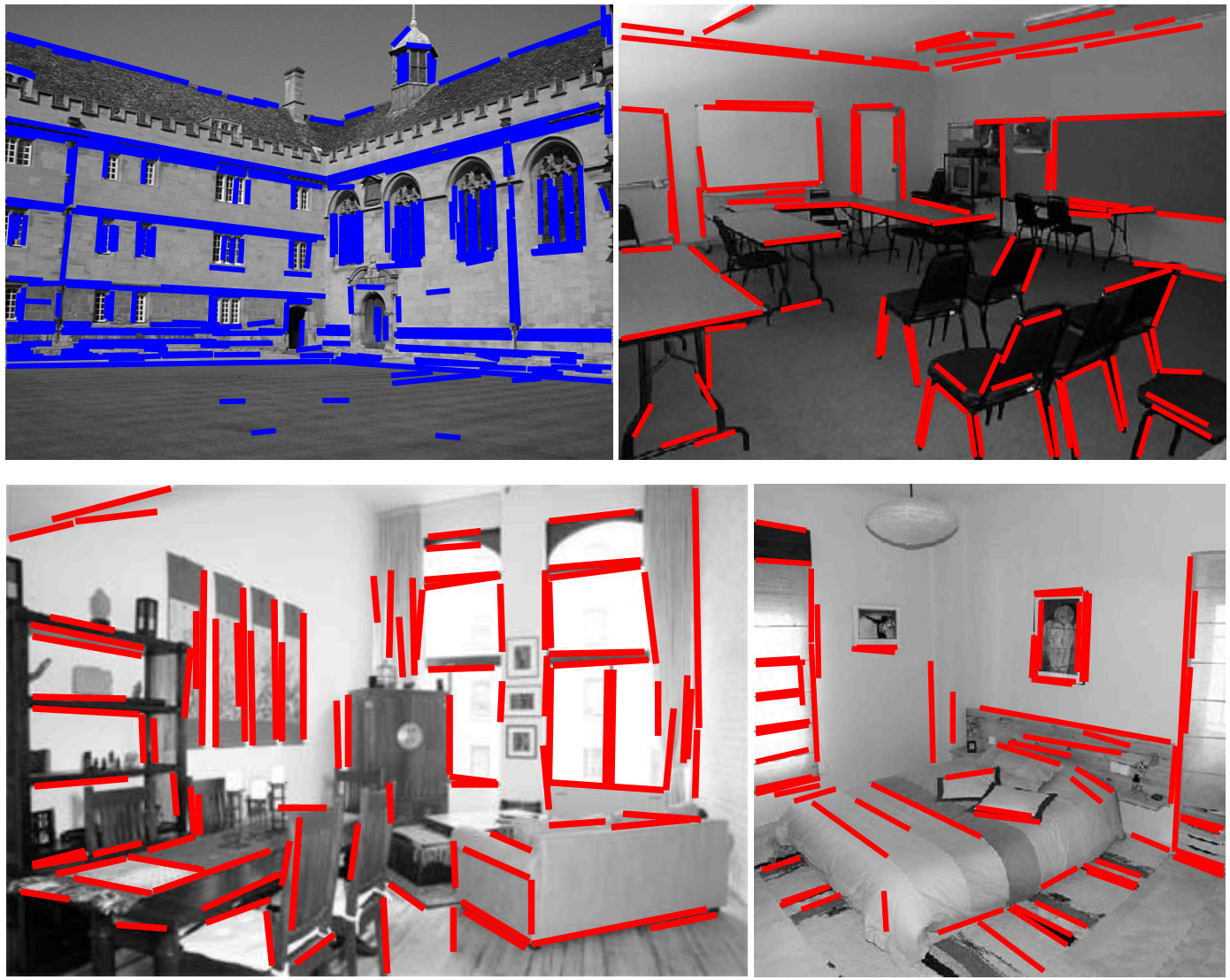
coplanarity
depth ordering



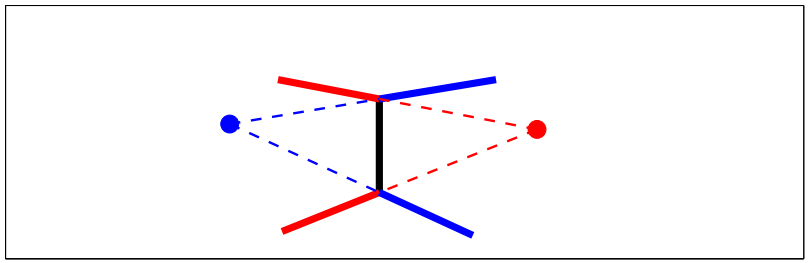
Talk Outline

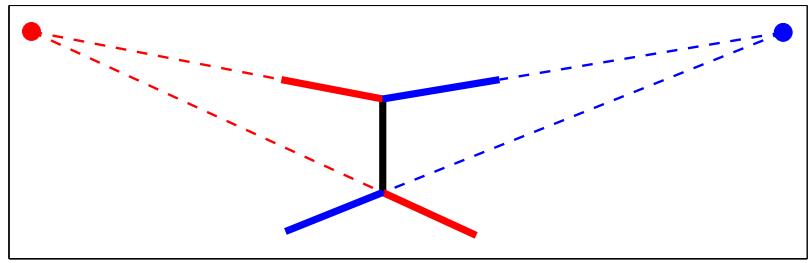
- Grouping Cues from 2D Lines to 3D Line Clusters
- Grouping from 3D Lines to 3D Quadrilaterals
- Grouping Cues from 3D Quadrilaterals to 3D Planes
- Spectral Clustering As the Core Grouping Engine
- Literature Review
- Further Research

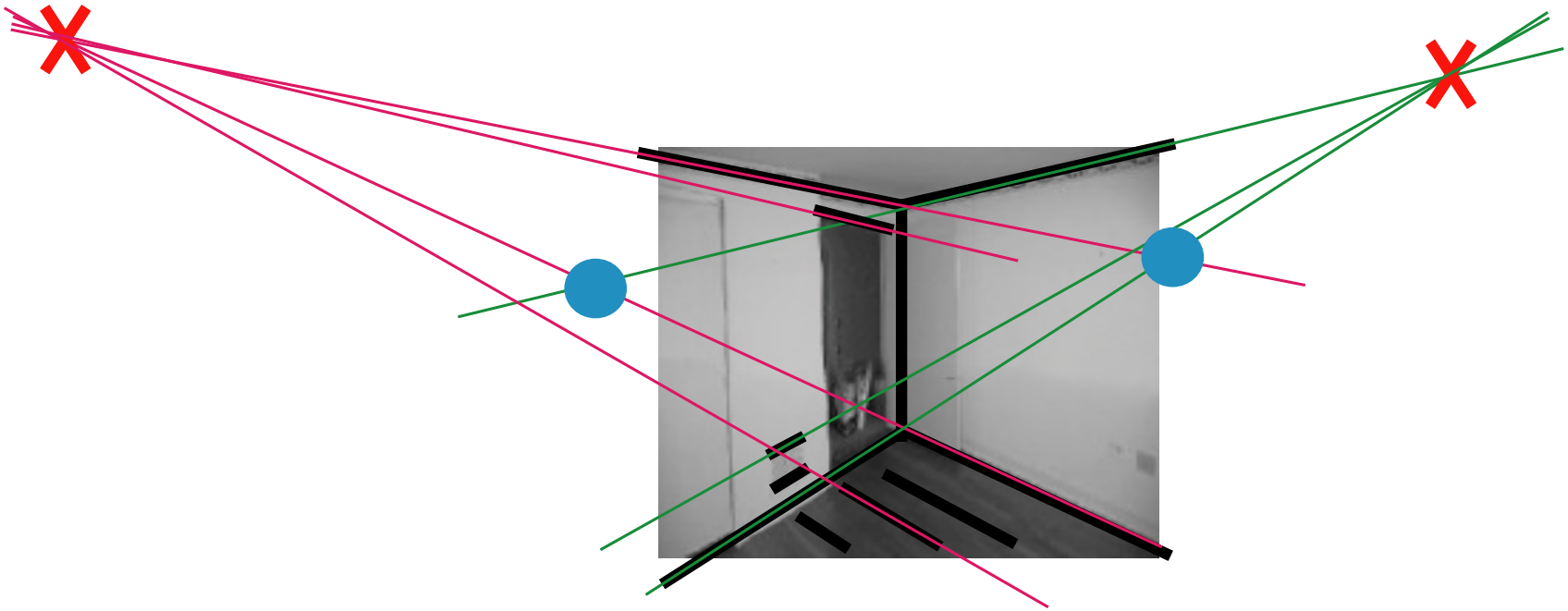
Challenge: Occlusion, Clutter, Weak Perspectivity



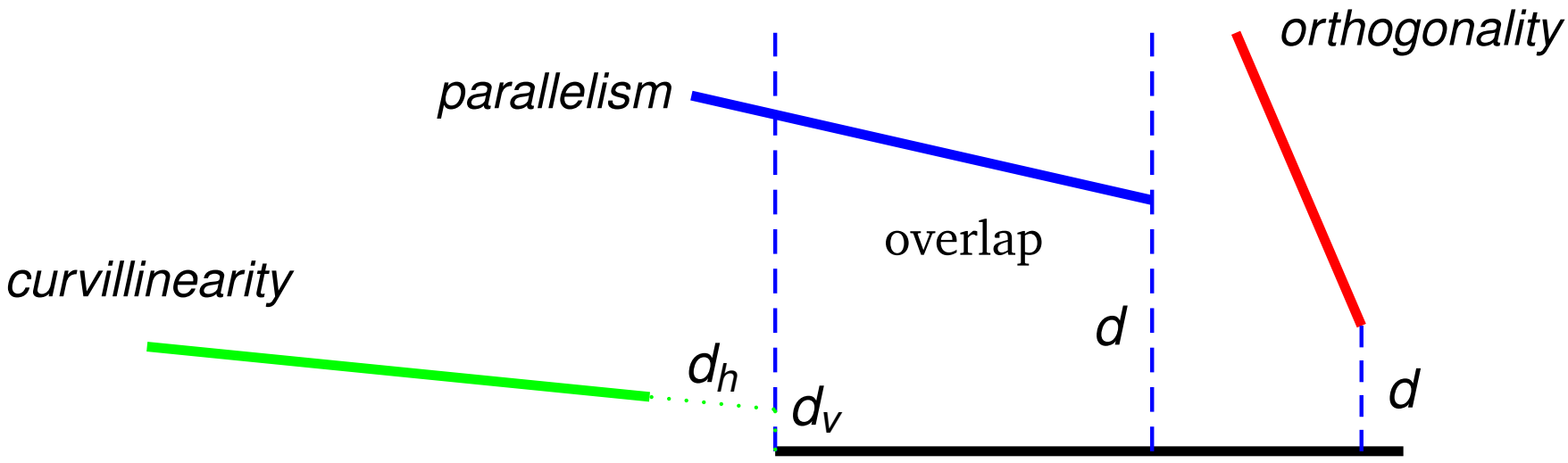
Ambiguity of Vanishing Points







Grouping Cues from 2D Lines to 3D Line Clusters

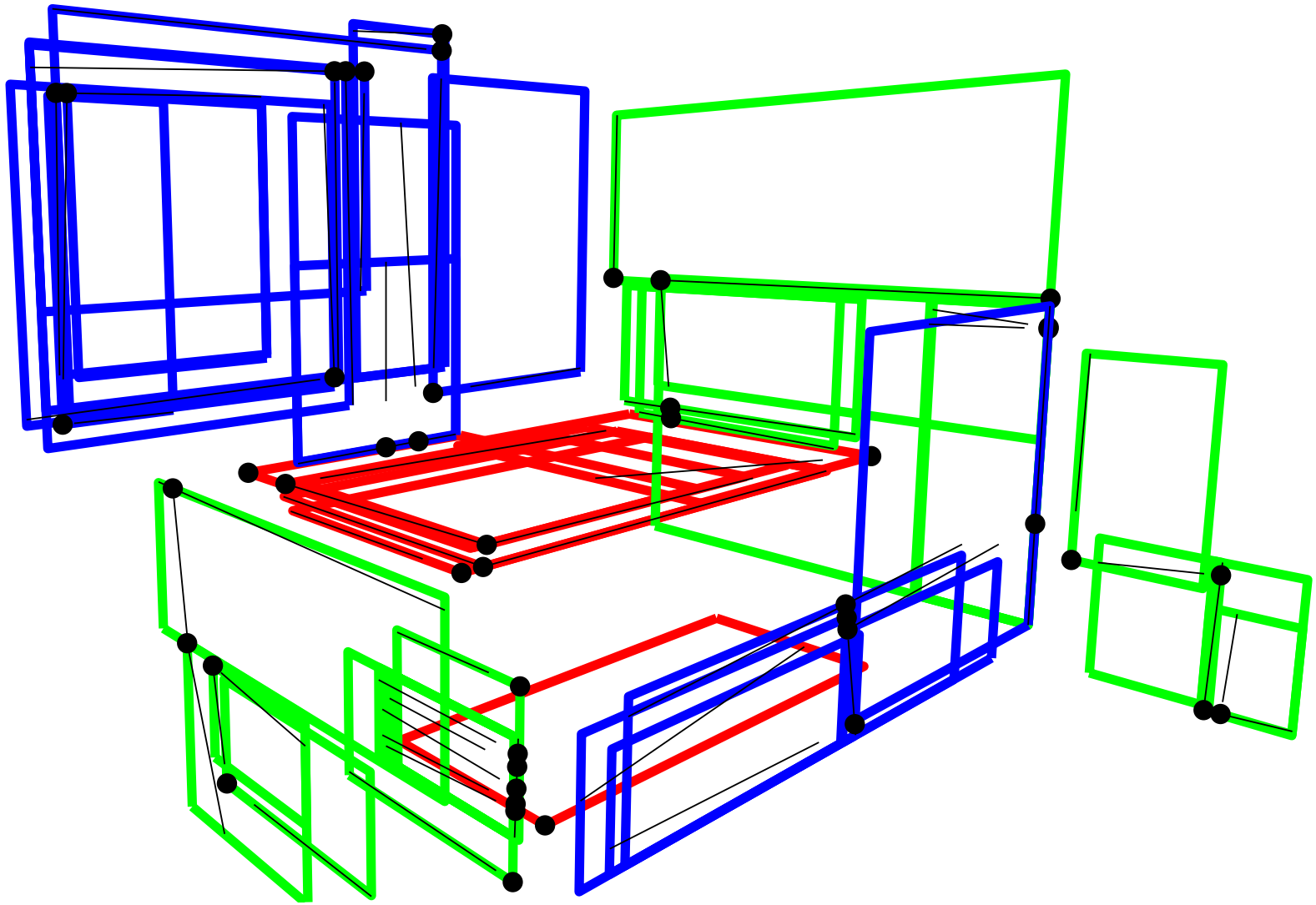


$$A_{-} = \exp\left(-\frac{d_h^2}{2\sigma_{c1}^2} - \frac{d_v^2}{2\sigma_{c2}^2} - \frac{1-\cos^2\theta}{2\sigma_{c3}^2}\right)$$

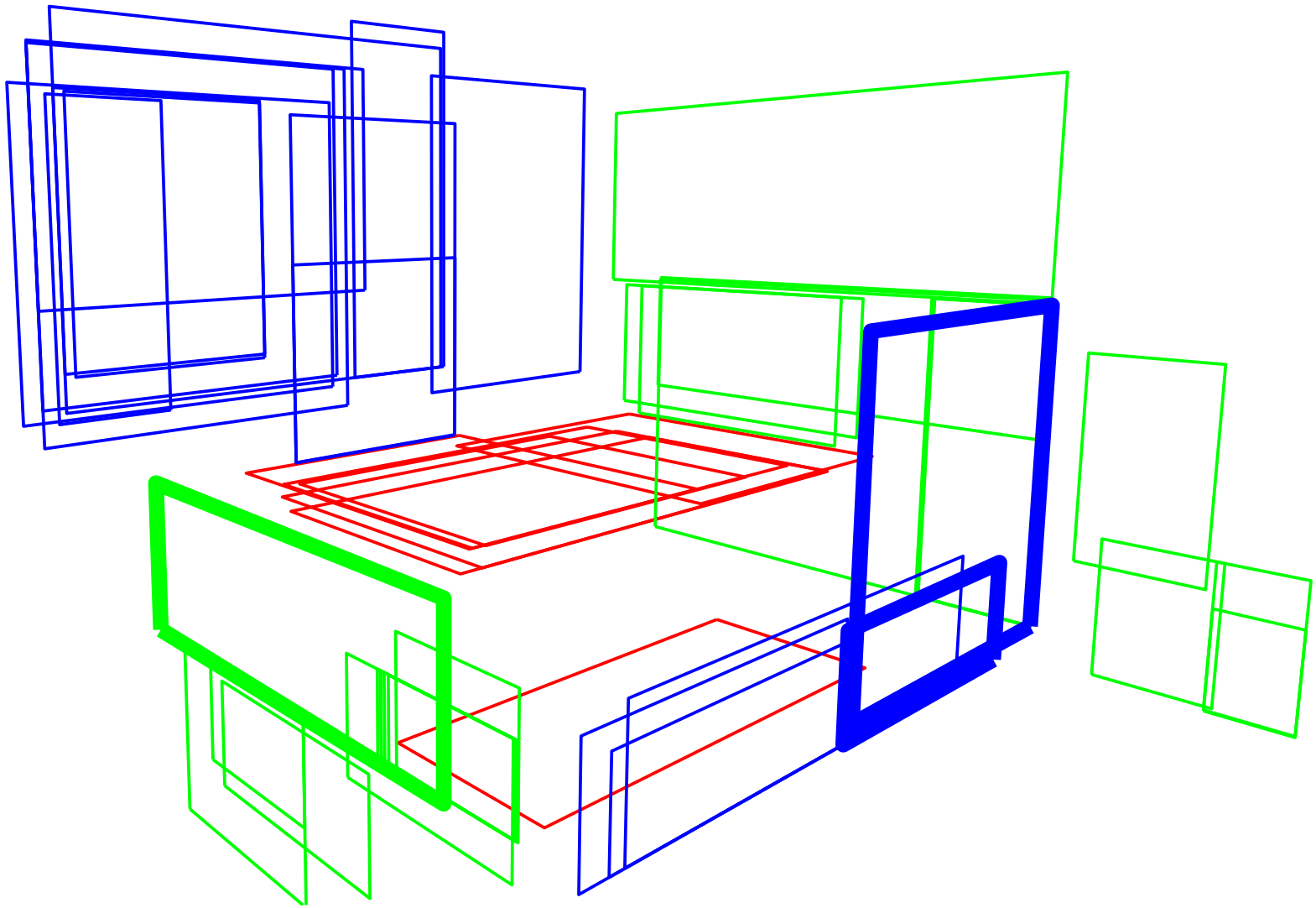
$$A_{\parallel} = \exp\left(-\frac{d^2}{2\sigma_{p1}^2} - \frac{(1-\text{overlap})^2}{2\sigma_{p2}^2} - \frac{1-\cos^2\theta}{2\sigma_{p3}^2}\right)$$

$$R_{\perp} = \exp\left(-\frac{d^2}{2\sigma_{o1}^2} - \frac{\cos^2\theta}{2\sigma_{o2}^2}\right)$$

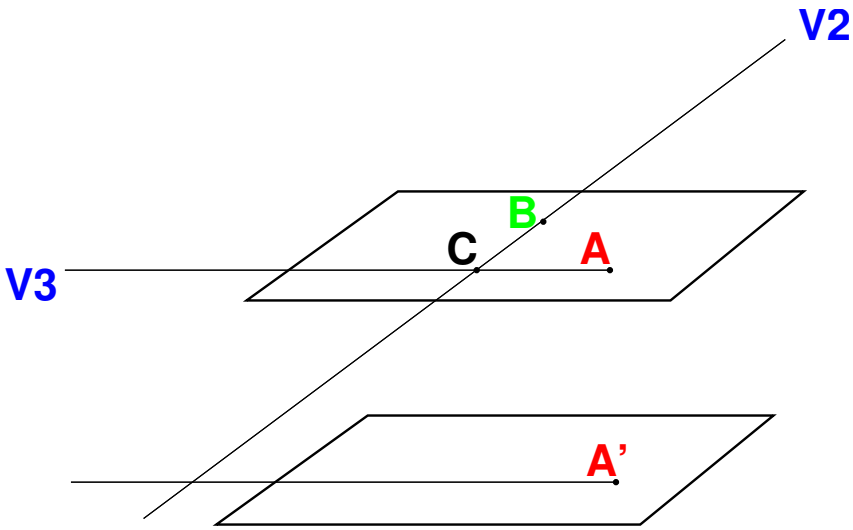
Grouping from 3D Lines to 3D Quadrilaterals



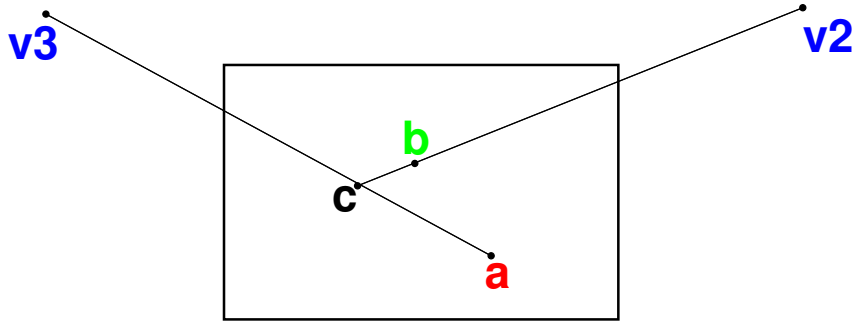
Coplanarity between 3D Quadrilaterals



Depth Ordering of 3D Points Along A Direction

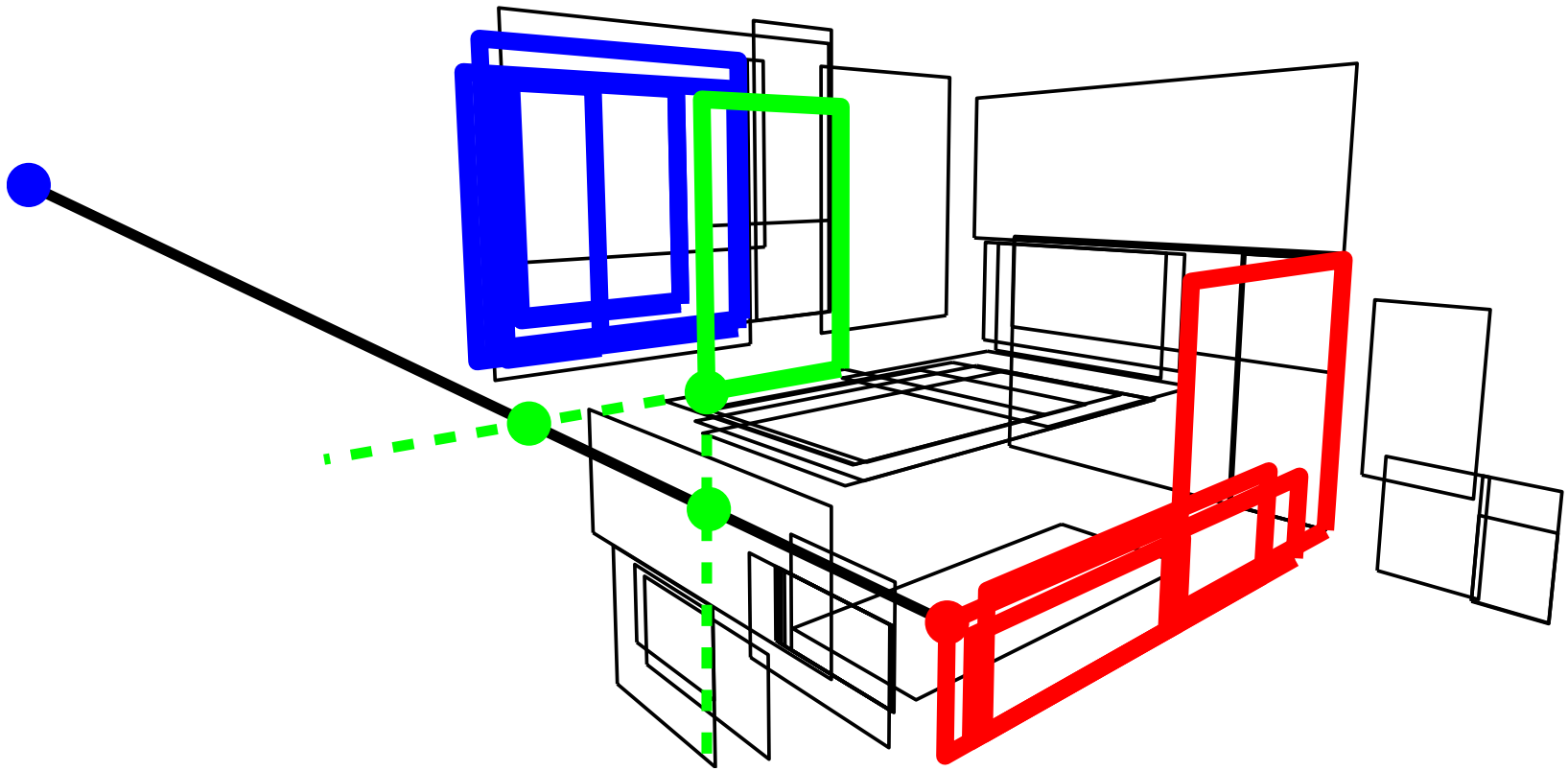


in 3D space

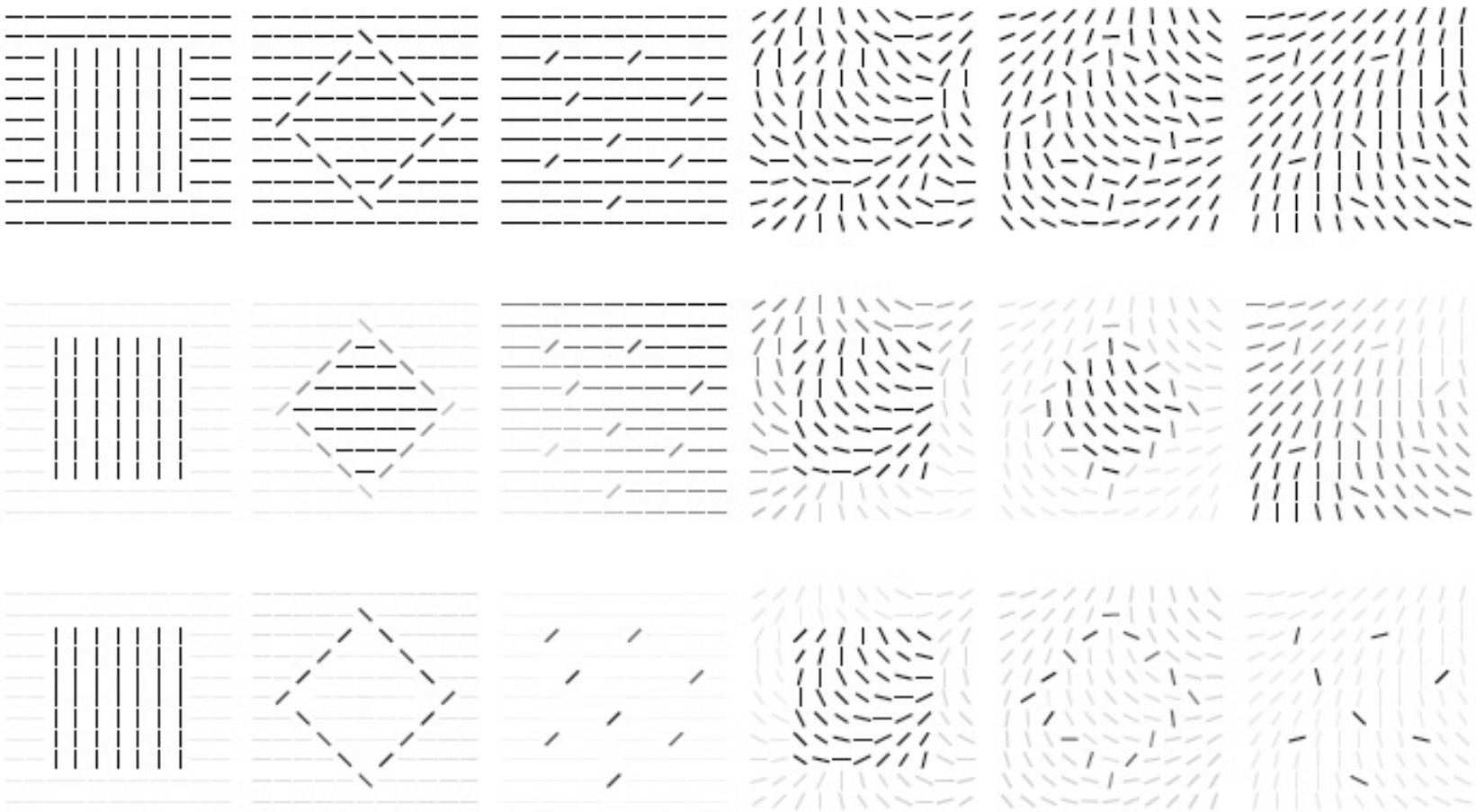


in 2D image

Depth Ordering between 3D Quadrilaterals

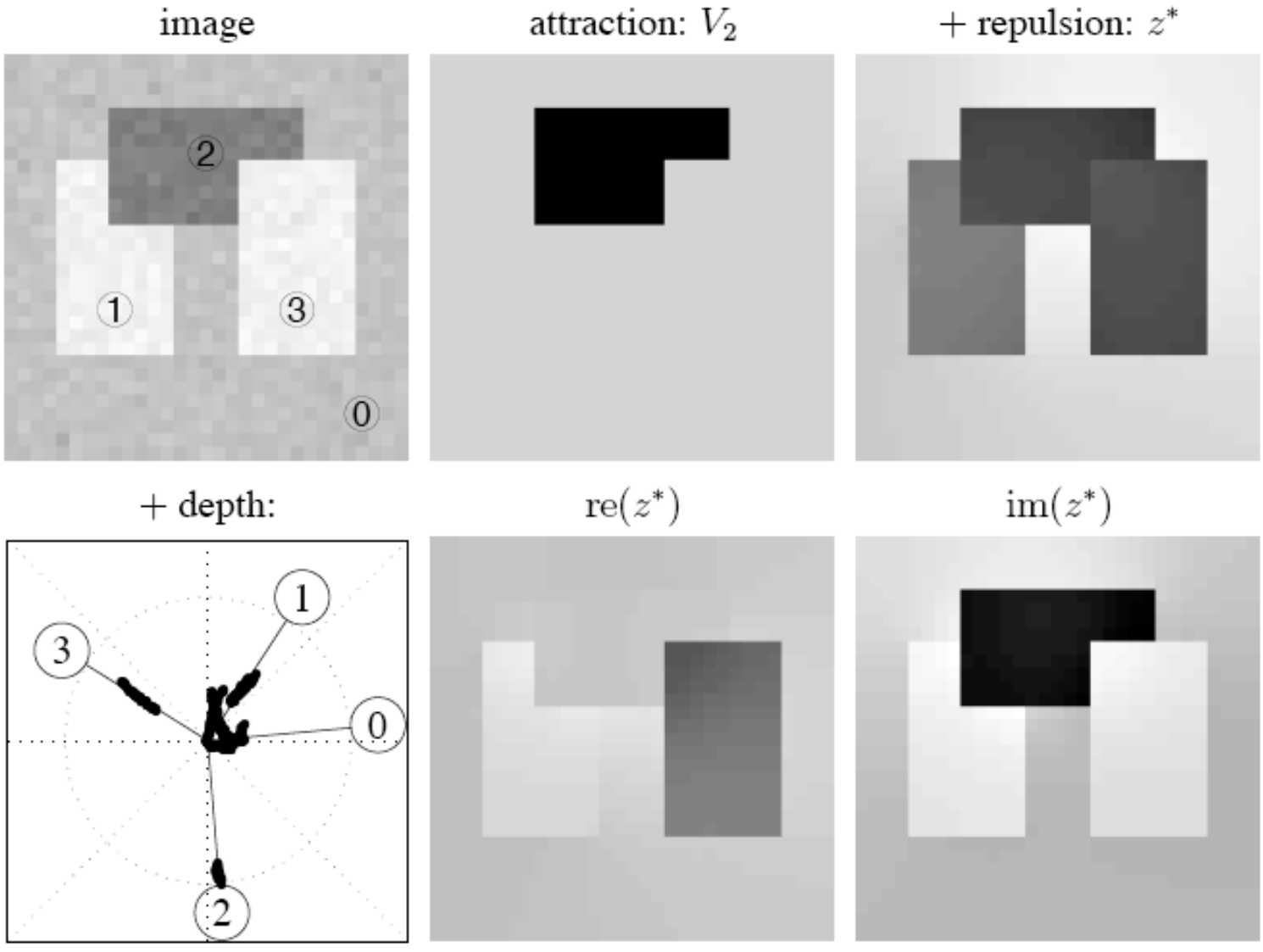


Spectral Clustering w. Attraction & Repulsion



Eigensolution of $(A - R + D_R, D_A + D_R)$, *Yu & Shi, CVPR 2001*

Spectral Clustering w. Attraction & Depth Ordering



Eigensolution of $(A + \sqrt{-1} \cdot R, D_A + D_R)$, *Yu & Shi, ICCV 2001*

Review: Different Approaches from 2D to 3D

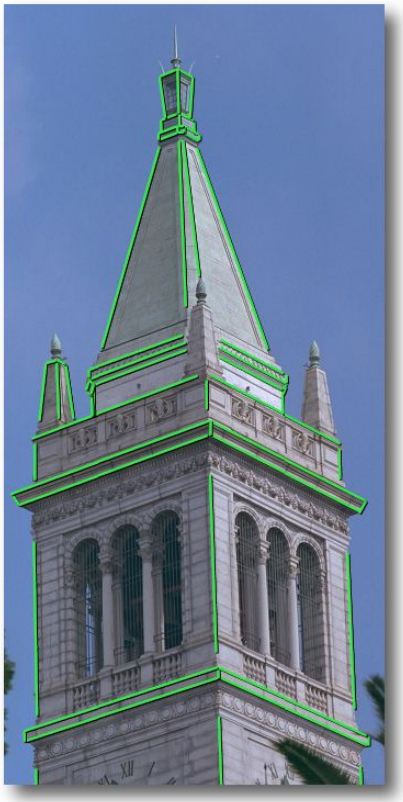
- Geometrical Reconstruction
- Shape-from-X
- Rule-based Generative Approach
- Statistical Learning

- Depth-Ordered Grouping

Reconstruction from Calibration & User Input

Modeling and Rendering Architecture from Photographs

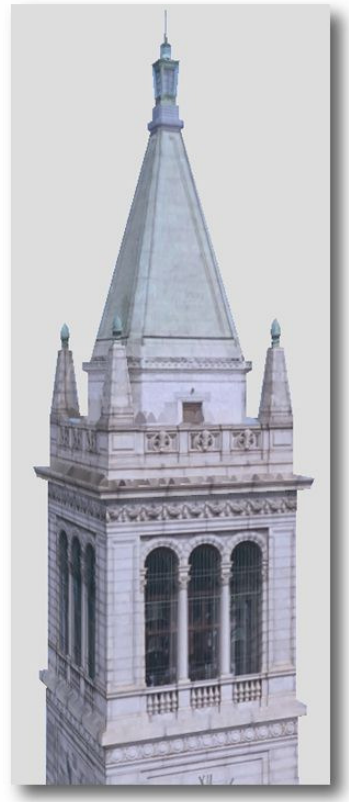
Debevec, Taylor, and Malik 1996



Original photograph with marked edges

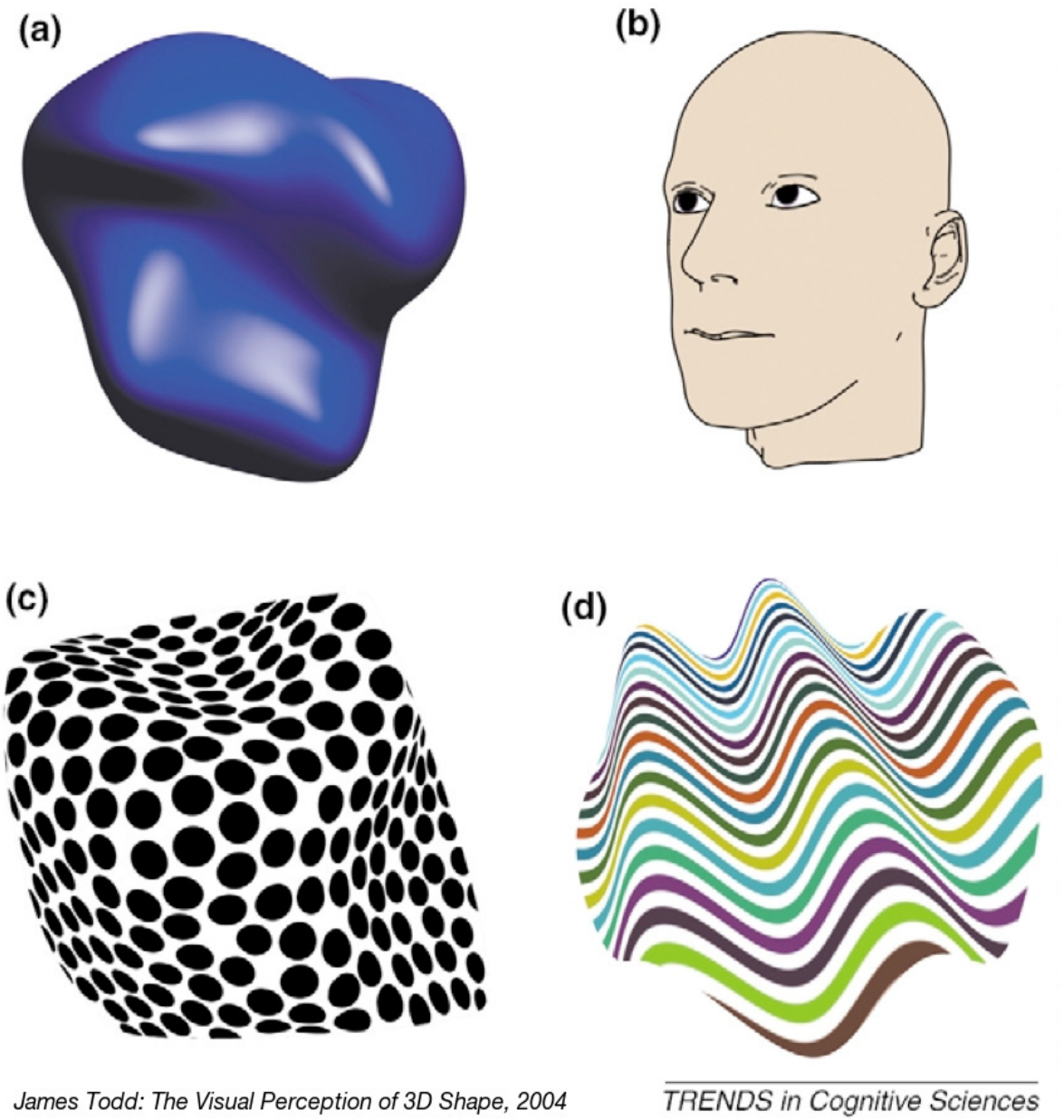
Recovered model

Model edges projected onto photograph



Synthetic rendering

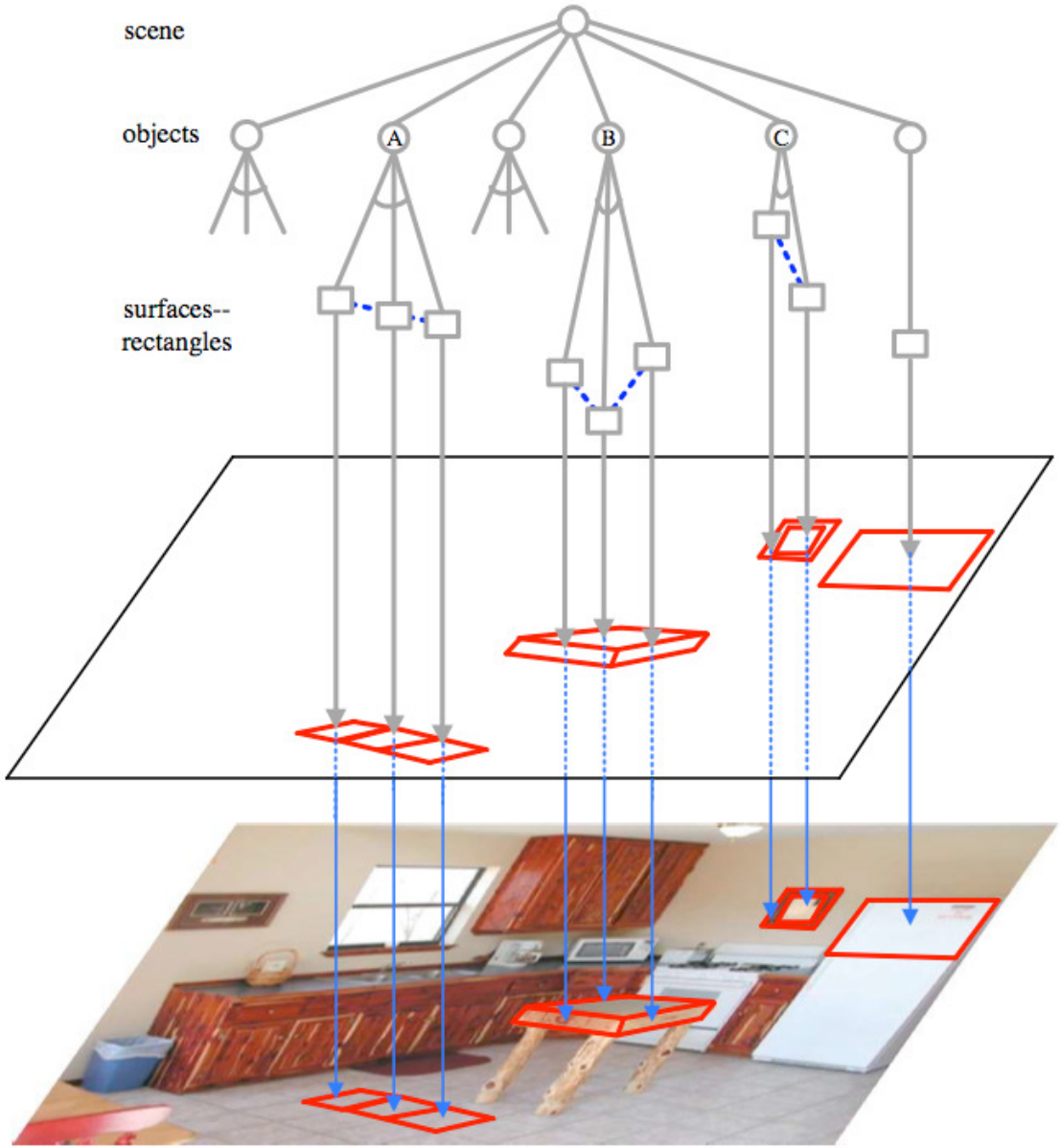
Shape-from-X with Applicable Cues



James Todd: The Visual Perception of 3D Shape, 2004

TRENDS in Cognitive Sciences

Generative Model on Limited Spatial Relations



Han & Zhu: Bottom-up / Top-down Image Parsing by Attribute Graph Grammar, ICCV, 2005

Statistical Associations of Feature and Context

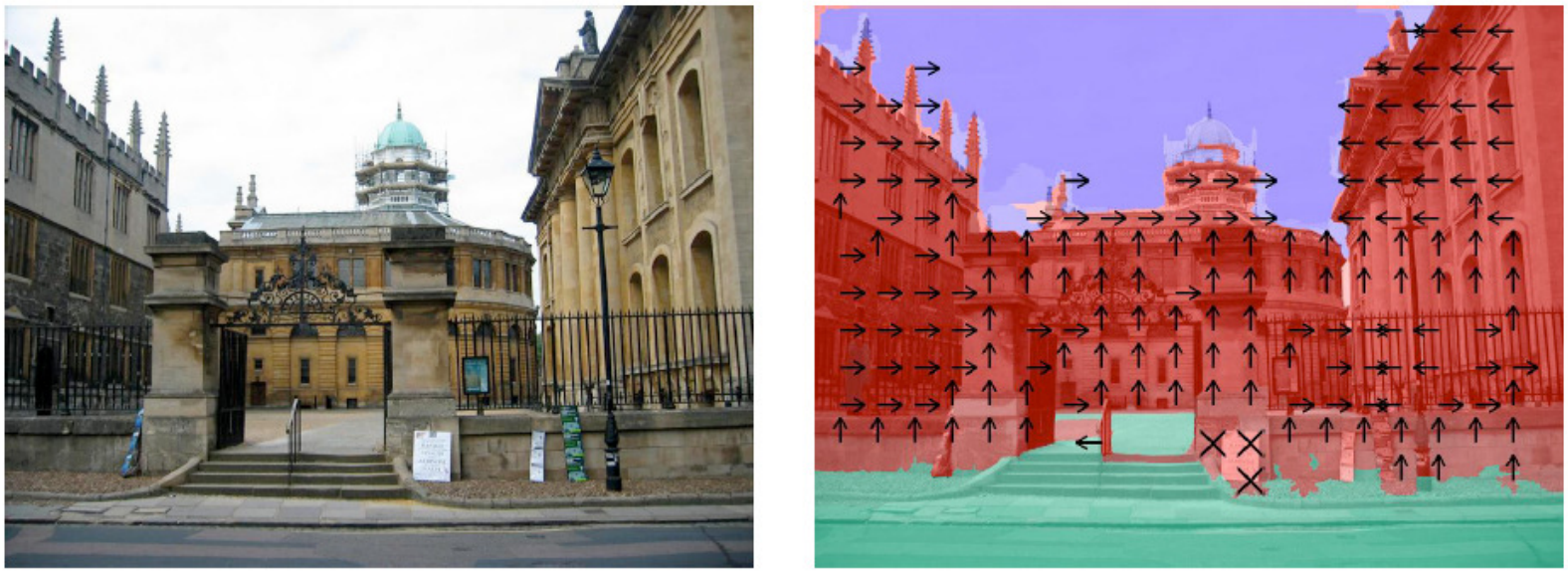


Figure 1: Geometric context from a single image: ground (green), sky (blue), vertical regions (red) subdivided into planar orientations (arrows) and non-planar solid ('x') and porous ('o').

Hoiem, Efros, & Hebert: Geometric Context from A Single Image, ICCV, 2005

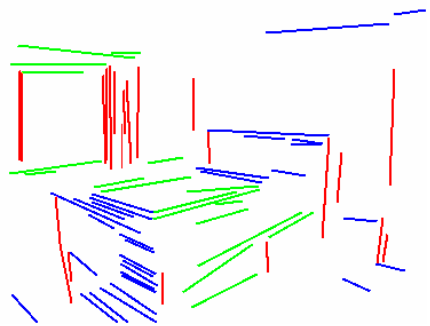
Line-Based Depth-Ordered Grouping Model

detection
localization

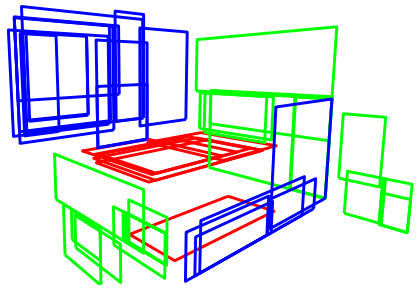
proximity
curvilinearity

Spectral Clustering with
Attraction and Repulsion
and Depth Ordering
Yu & Shi, 2001

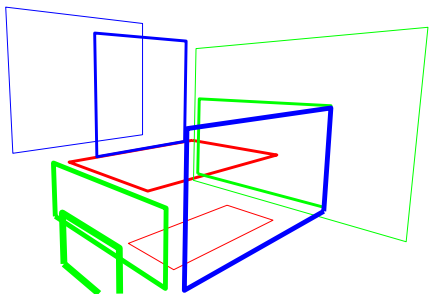
curvilinearity
parallelism
orthogonality
perspectivity



proximity
orthogonality



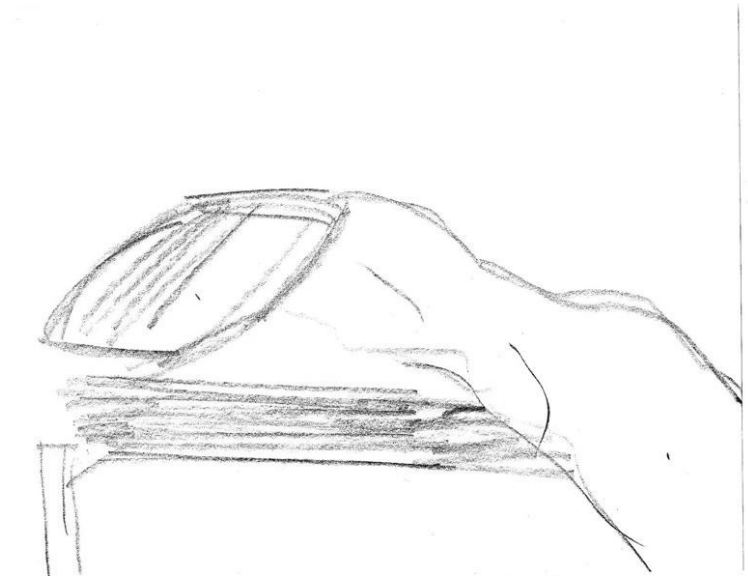
coplanarity
depth ordering

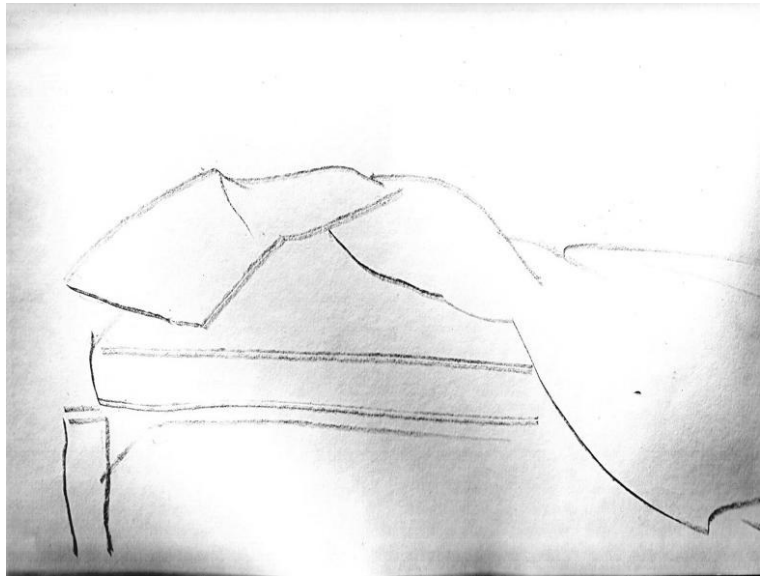


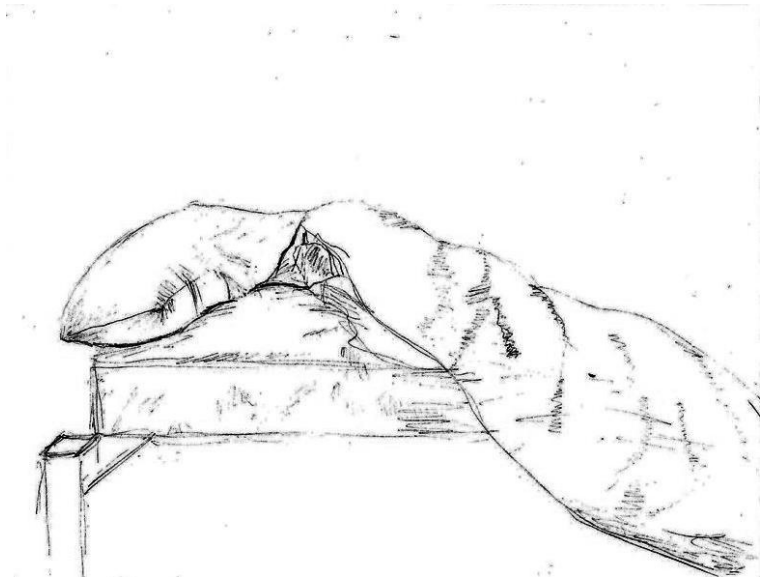
Further Research

- Group Lines into Quadrilaterals
- Relative Depth Ordering
- Cross Constraints on Depth and Extent
- Concurrent Grouping
- Vocabulary of Lines for Depth and Volume

Vocabulary of Lines For Depth and Volume







Drawings by Mandy Sakamoto, ©Art and Visual Perception Class, Yu & Mulhern, Boston College, Fall 2007