3D Curve Sketch: Flexible Curve-Based Stereo Reconstruction and Calibration

Ricardo Fabbri and Benjamin Kimia
Laboratory for Engineering Man/Machine Systems (LEMS), Brown University

THE PROBLEM

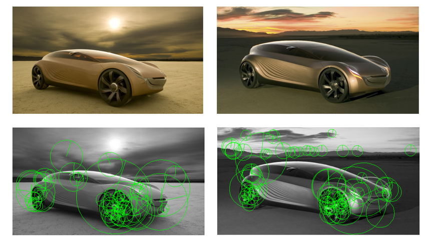
- Given a **large** sequence of 2D images
  - A video from a handheld camera
  - A set of unorganized snapshots
- Reconstruct the 3D scene

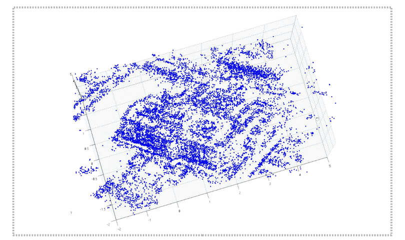
MOTIVATION

Current methods are either

a) Based on **isolated point features** (SIFT)

- Pro: no controlled acquisition is required
- Con: give only a point cloud 3D reconstruction; require textured scenes



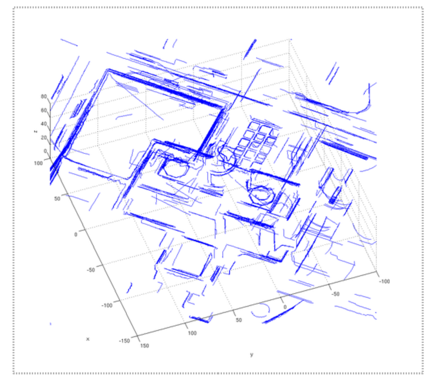


b) Based on **photometric consistency**

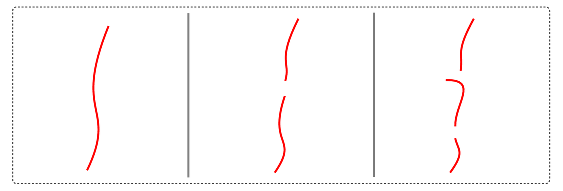
- Pro: produce detailed, texturized 3D reconstructions
- Con: require controlled acquisition and precise calibration, and use a large amount of resources. Unscalable.

We propose a middle ground approach based on curves

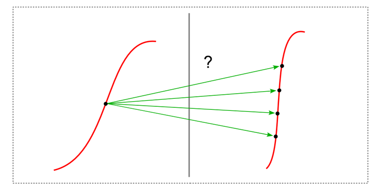
- **More distinctive** features than points, allowing for applications such as modeling and object matching
- **More efficient** in space and time when compared to volumetric approaches
- **More flexible** than isolated keypoints: works for scenes where there aren't enough feature points



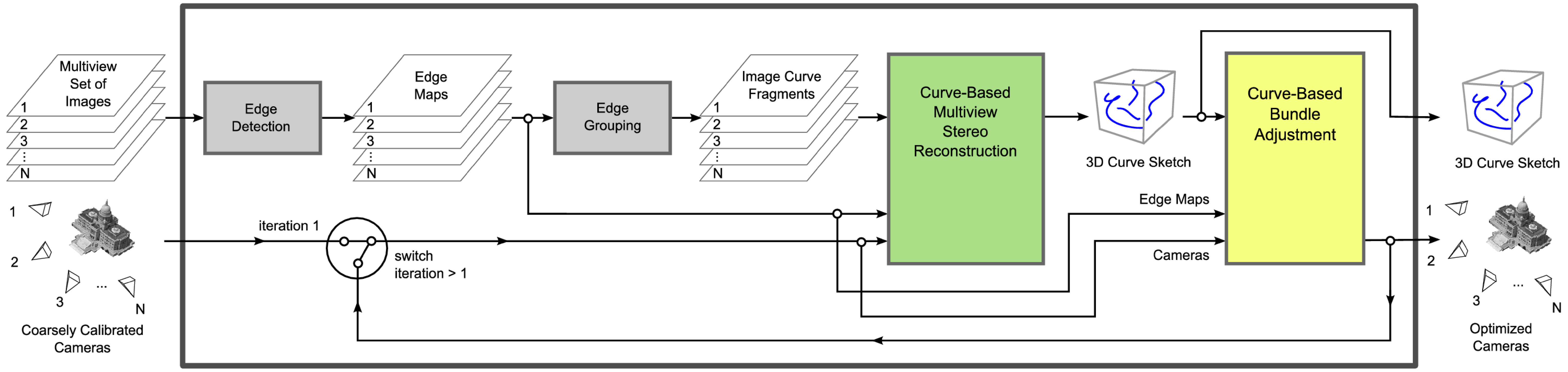
CHALLENGES OF USING CURVES



Edge grouping is unstable



Pointwise correspondence is ambiguous



OVERVIEW OF THE APPROACH

- Given a **large** sequence of views (at least 6)
- Rough Intrinsic and extrinsic camera calibration
  - Given by traditional approaches based on point features
- Goal: to produce the 3D Curve Sketch
  - A dense collection of 3D curve fragments
  - View-stationary curves such as reflectance contours, ridges, shadow and shade curves
- First, reconstruct a **core 3D curve sketch** of reliable curves
- The core 3D curve sketch model can be used to refine the cameras, which are then used to generate a final 3D Curve Sketch

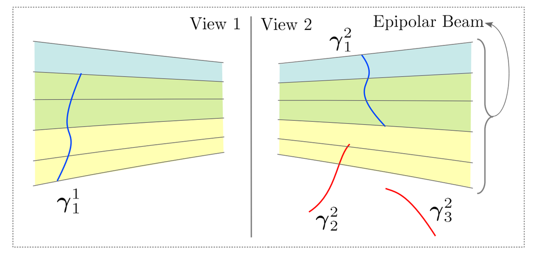
OVERVIEW OF CURVE-BASED MULTIVIEW STEREO

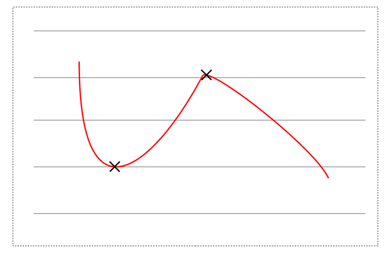
HYPOTHESIZE-AND-TEST FRAMEWORK

- Pick two views to be called **hypothesis views**
  - Form a 3D curve hypothesis by pairing curve fragments with sufficient epipolar overlap
- Test each 3D curve hypothesis by reprojecting onto other views, the **confirmation views**
  - Measure consistency to the subpixel edge map using differential geometry (tangent orientation)
  - Use distance transform for quick lookup
- Repeat for many different hypothesis views

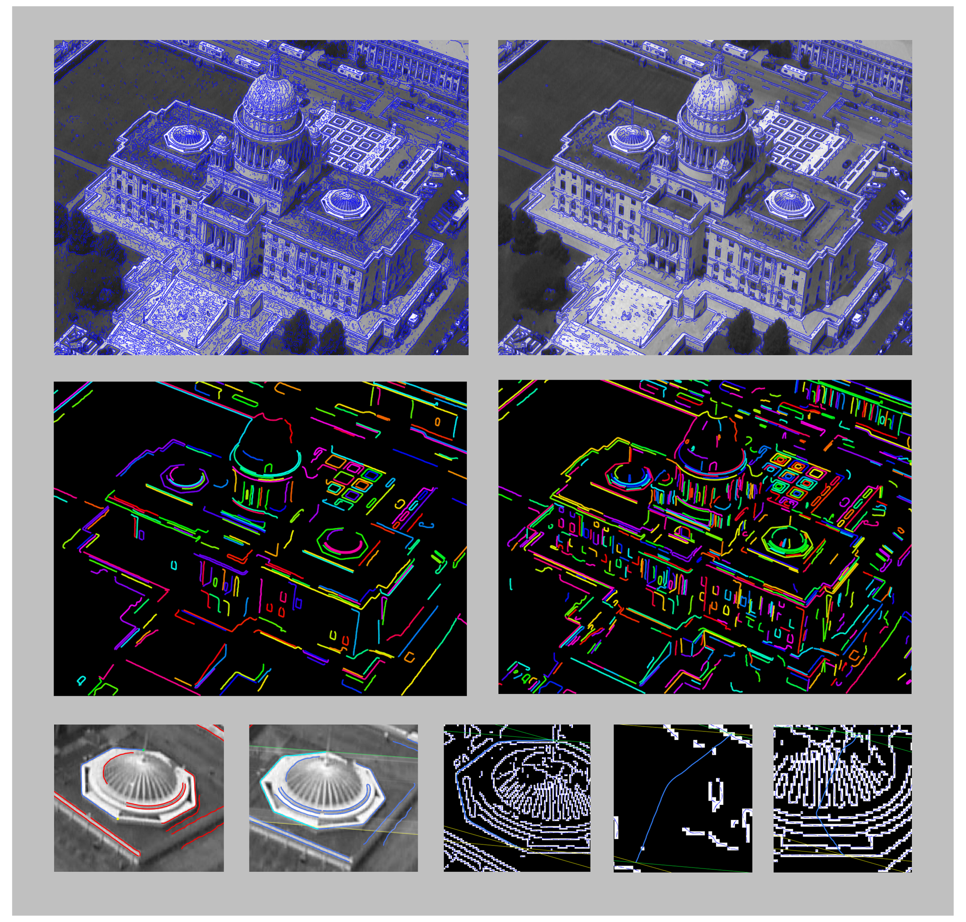
HYPOTHESIS FORMATION

- Curves in two views are paired if they have sufficient epipolar overlap



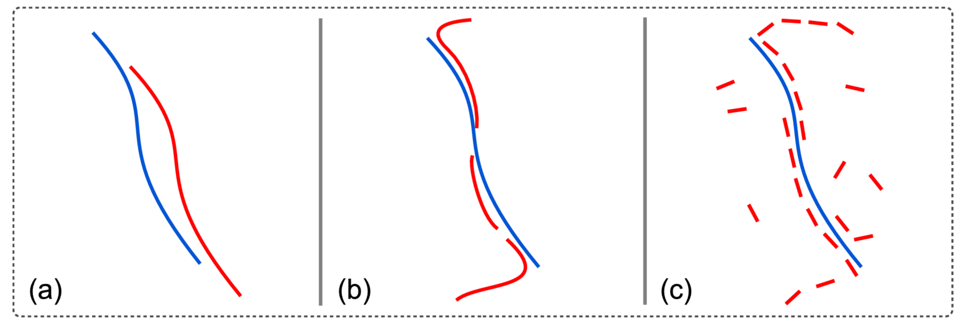


- Avoid multiple intersections with epipolar lines by breaking at epipolar tangencies



HYPOTHESIS TESTING

- Each 3D curve hypothesis is tested by reprojecting onto confirmation views
- We typically use at least 4 confirmation views for each pair of hypothesis views
- The reprojection is validated against the edge content in the confirmation view
  - Validating against grouped curve fragments (a) would be vulnerable to instabilities (b)
  - This motivates using smaller curve primitives. Here we use subpixel edgels attributed with differential geometry



HYPOTHESIS TESTING

- The support of a reprojected curve is the *number* of edgels that pass thresholds in distance and orientation difference
- For each point of the reprojected curve
  - Locate the nearest edgel in the edge map by distance transform lookup
  - Make sure it passes a distance threshold
  - Make sure it passes an orientation difference threshold
- Accumulate support over all confirmation views
  - Discard hypotheses with support less than a threshold
- Greedy matching: if an image curve participates in more than one hypothesis, keep the one with largest support
- Ratio test
  - Repeated structures can cause problems
  - Ambiguity happens if the second best hypothesis for a given image curve has support close to its best hypothesis
  - We do not reconstruct these hypotheses
  - Similar test to matching SIFT features

KEY IDEAS

- By using the *number* of 'inlier' edgels as the measure of support, we obtain very robust matching
- Use of orientation is essential to avoid clutter.

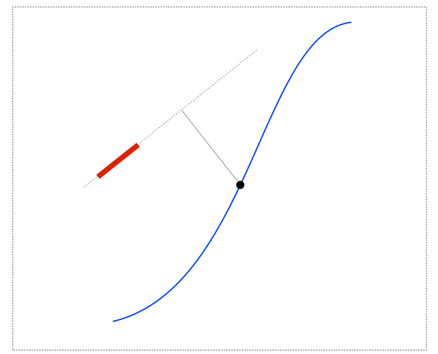
CURVE-BASED POSE OPTIMIZATION

OVERVIEW

- The 3D curve sketch is the set of all confirmed 3D curve hypotheses from curve-based multiview stereo.
- Basic idea: minimize reprojection error. Project the 3D curve sketch on all views and minimize discrepancy to the edge maps for those projected curves having sufficient support

Reprojection error

- For each point of each projected curve
  - Lookup the nearest edgel using distance transform
  - Measure the distance to the line containing the edgel

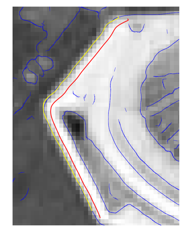


TIMING

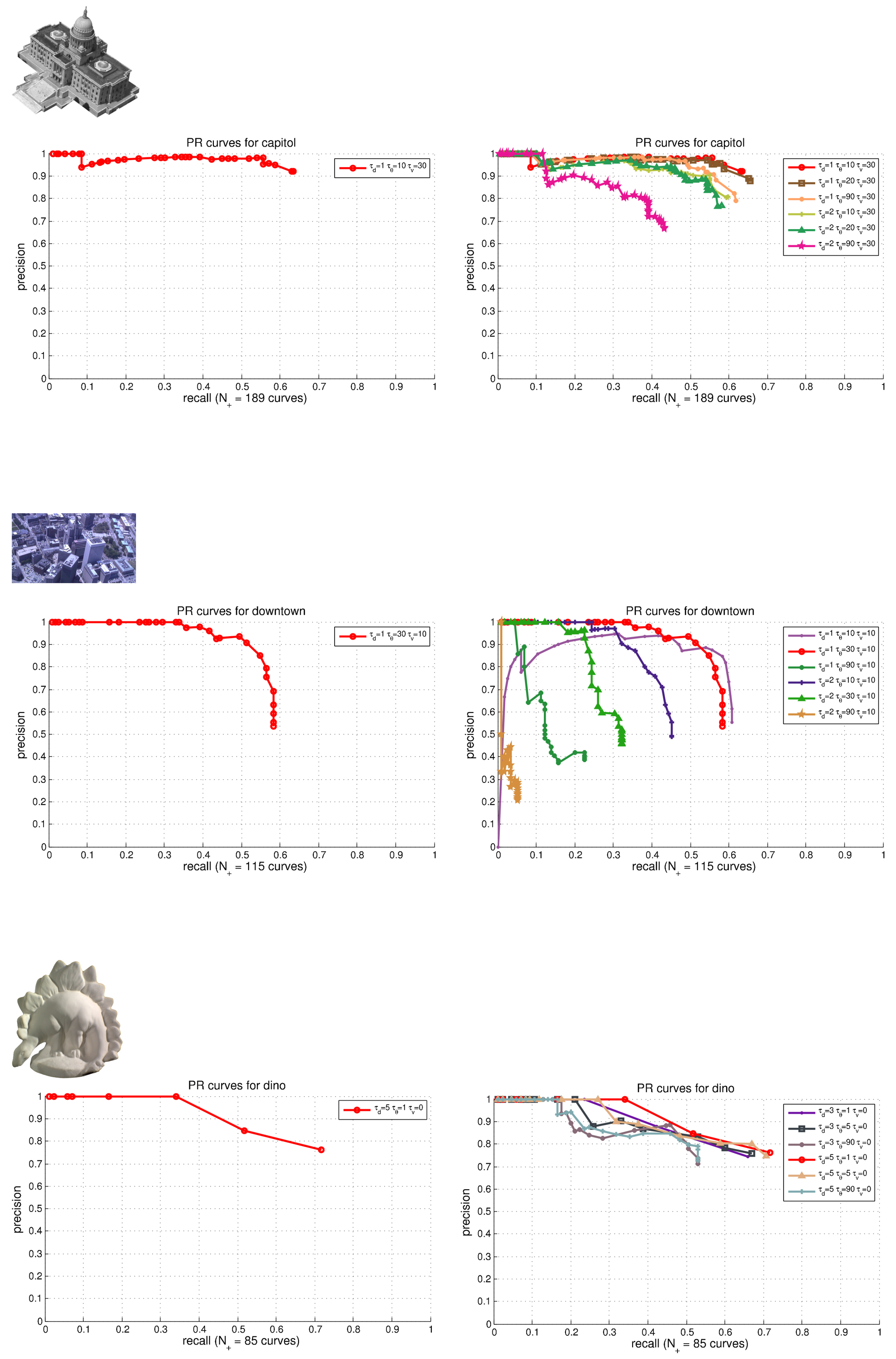
Dataset	# 3D Curves	# Hypot. Views	Min. Length	Time (s)
Dino	3712	100	20	466
Capitol	1231	30	40	567
Capitol High	1742	30	40	633
Downtown	2340	30	40	985

CURVE-BASED POSE OPTIMIZATION

- Stage 1: optimize the pose for each view
  - Only optimize the 6 degrees of freedom using Levenberg-Marquardt. Curve sketch stays fixed.
  - To fight clutter, can run this procedure many times, each time picking 3 curves and optimize. Output best.
- Stage 2: full bundle adjustment.
  - Optimize multiple cameras simultaneously
  - 3D curve samples are also being optimized
  - Sparse Levenberg-Marquardt is used
  - Strategies to scale it up to hundreds of views

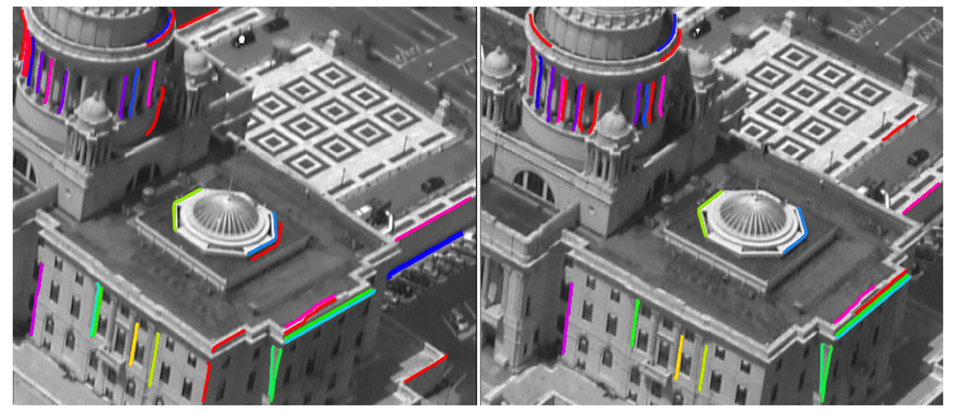


EVALUATION



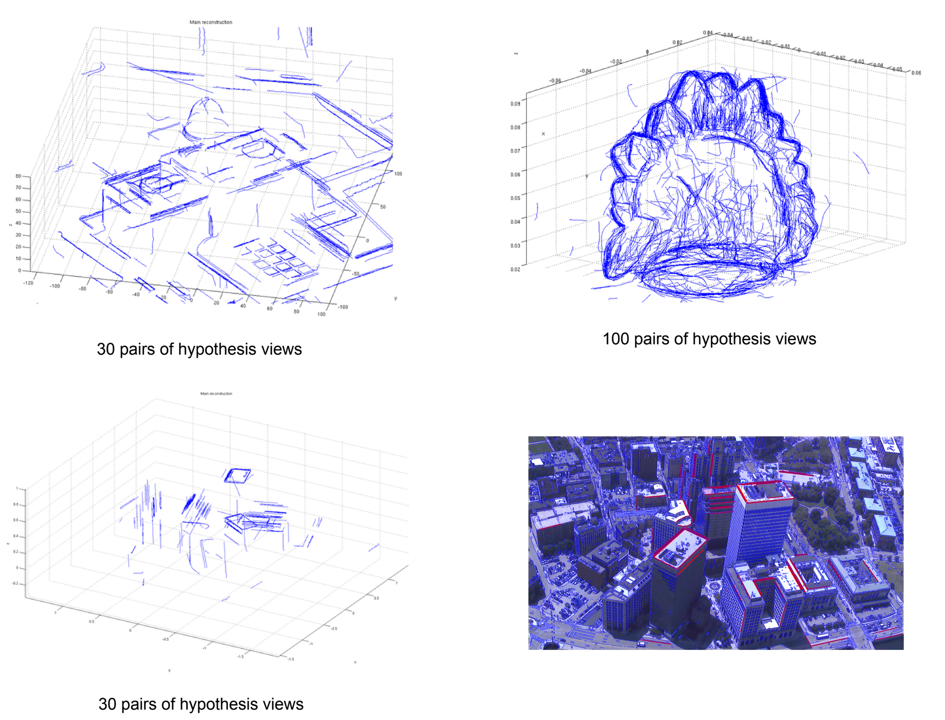
EVALUATION

- Ground truth is hard for real-world datasets
- We focus on the ability of the system to find the correct curve correspondence between two views
- We manually record the ground-truth correspondence between typical hypothesis views for each dataset



- We plot precision/recall curves comparing the core 3D curve sketch to ground truth
- All datasets produce 100% precision at 30% recall, enough for applications such as calibration/registration
- Tangential orientation plays a fundamental role

RECONSTRUCTION RESULTS



CONCLUSION

- Novel framework for multiview reconstruction and pose optimization based on image curve content
- Allows for applications where the assumptions of existing approaches fail but image curve content is present
- Integrates geometric information across many views
- Initial building block for complete reconstruction of general scenes