6.801/866

Image-Based Rendering

T. Darrell

Vision for Graphics

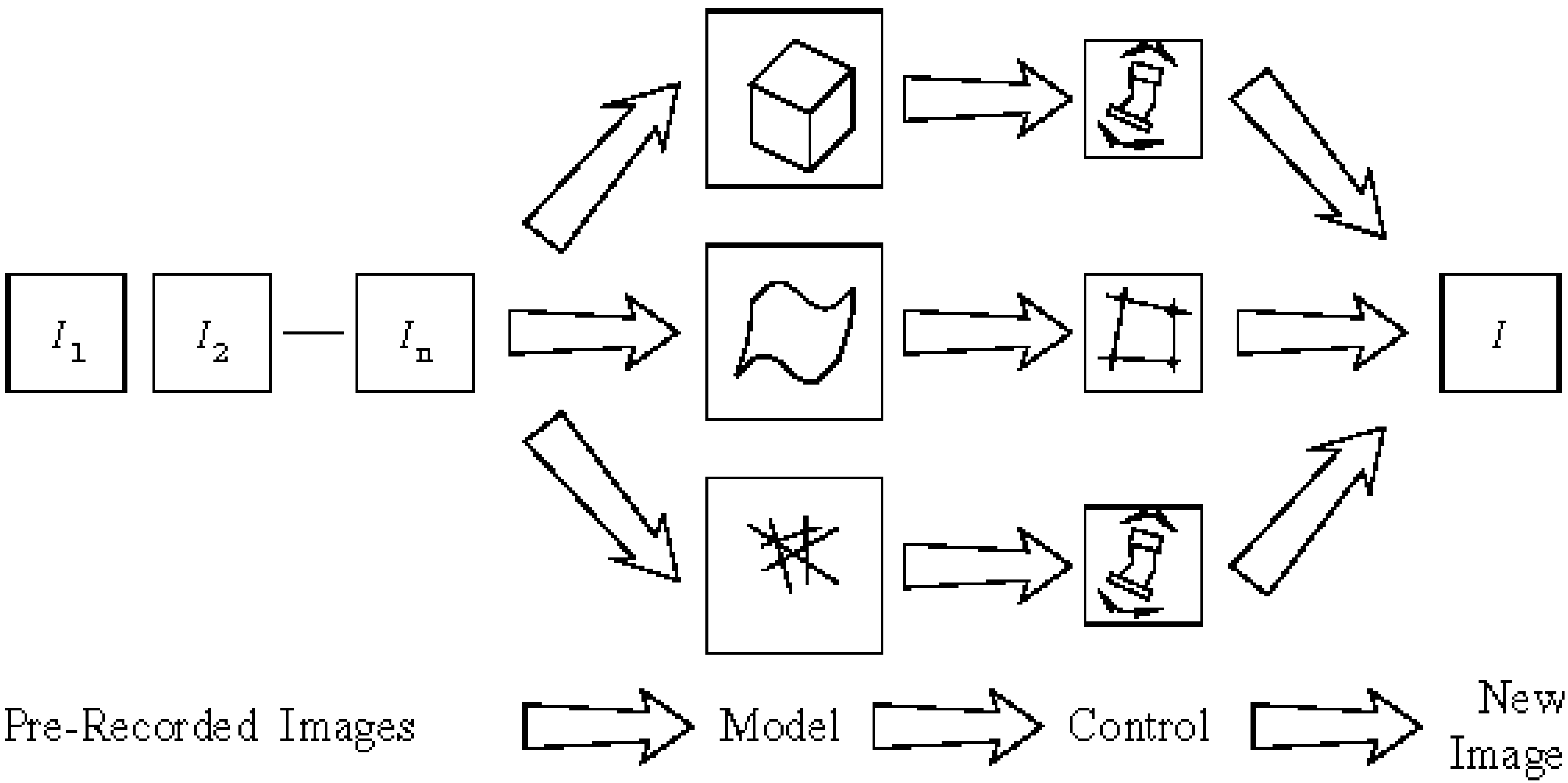
So far: stereo, motion, tracking, model-based recognition, most focusing on recovering 3-D models with accurate shape…

One of the main applications of vision is making new pictures!

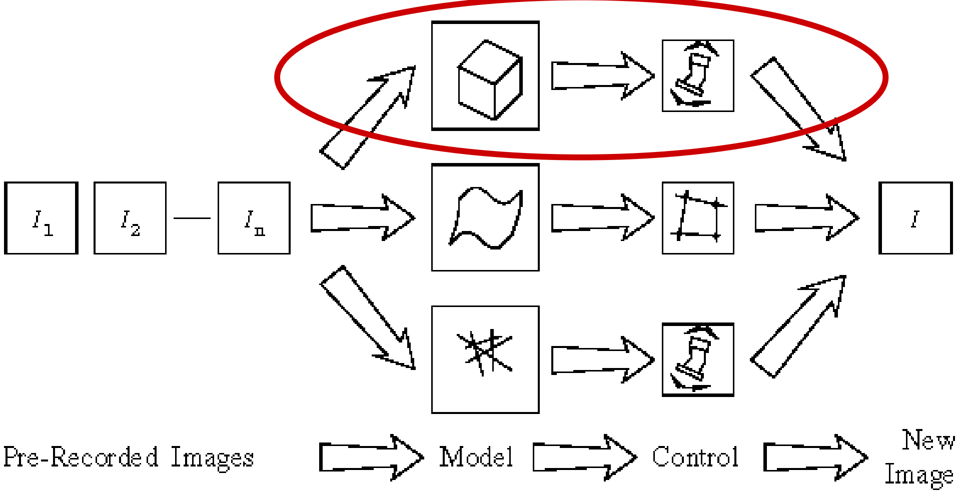
- Do we need detailed models?
- Do we need Euclidean 3-D shape?
- Are dense range images useful?

Image-based rendering

Synthesize new views from a set of pictures.

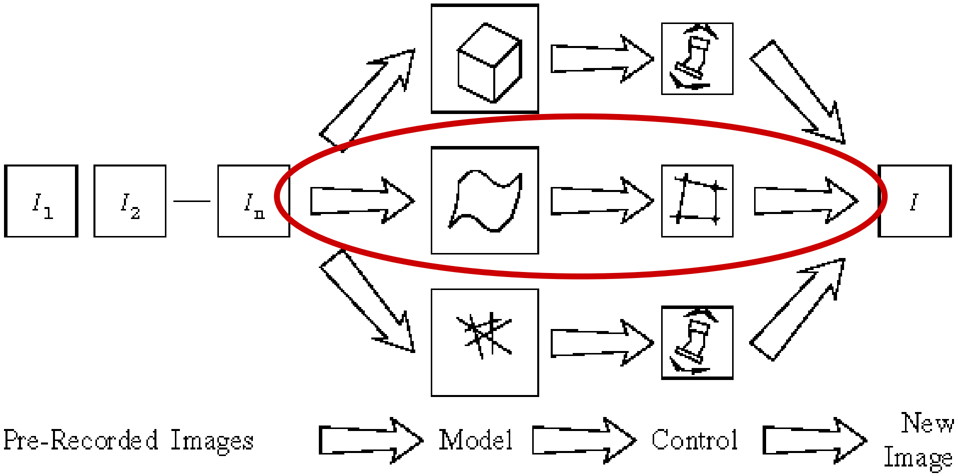


Taxonomy



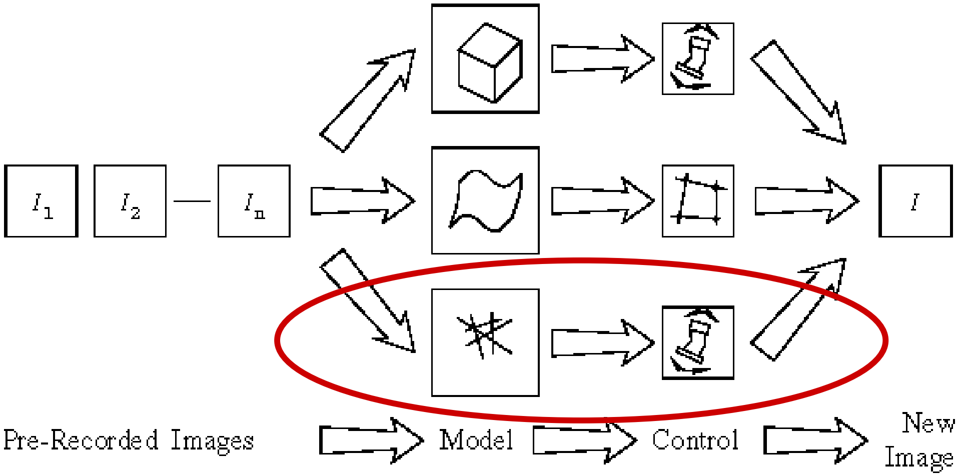
1. Build a 3-D model; re-render from new viewpoint
   - Multi-view stereo; “Virtualized Reality”
   - Visual Hulls
   - Model-based stereo
2. Establish correspondences; use view transfer
   - Affine view synthesis
3. Model sets of light rays
   - Lightfields, Lumigraphs, …

Taxonomy



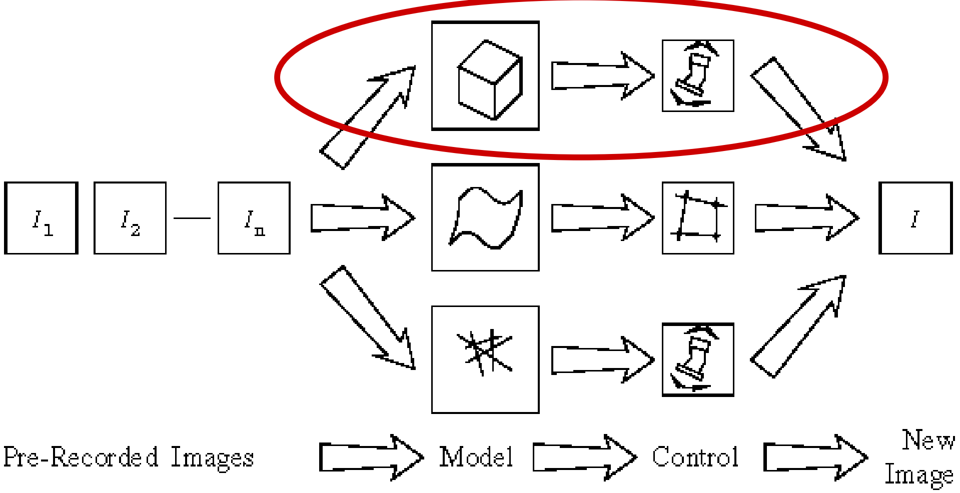
1. Build a 3-D model; re-render from new viewpoint
   - Multi-view stereo; "Virtualized Reality"
   - Visual Hulls
   - Model-based stereo
2. Establish correspondences; use view transfer
   - Affine view synthesis
3. Model sets of light rays
   - Lightfields, Lumigraphs, …

Taxonomy



1. Build a 3-D model; re-render from new viewpoint
   - Multi-view stereo; "Virtualized Reality"
   - Visual Hulls
   - Model-based stereo
2. Establish correspondences; use view transfer
   - Affine view synthesis
3. Model sets of light rays
   - Lightfields, Lumigraphs, …

Taxonomy



1. Build a 3-D model; re-render from new viewpoint
   – Multi-view stereo; “Virtualized Reality”
   – Visual Hulls
   – Model-based stereo
2. Establish correspondences; use view transfer
   – Affine view synthesis
3. Model sets of light rays
   – Lightfields, Lumigraphs, …

Models from stereo

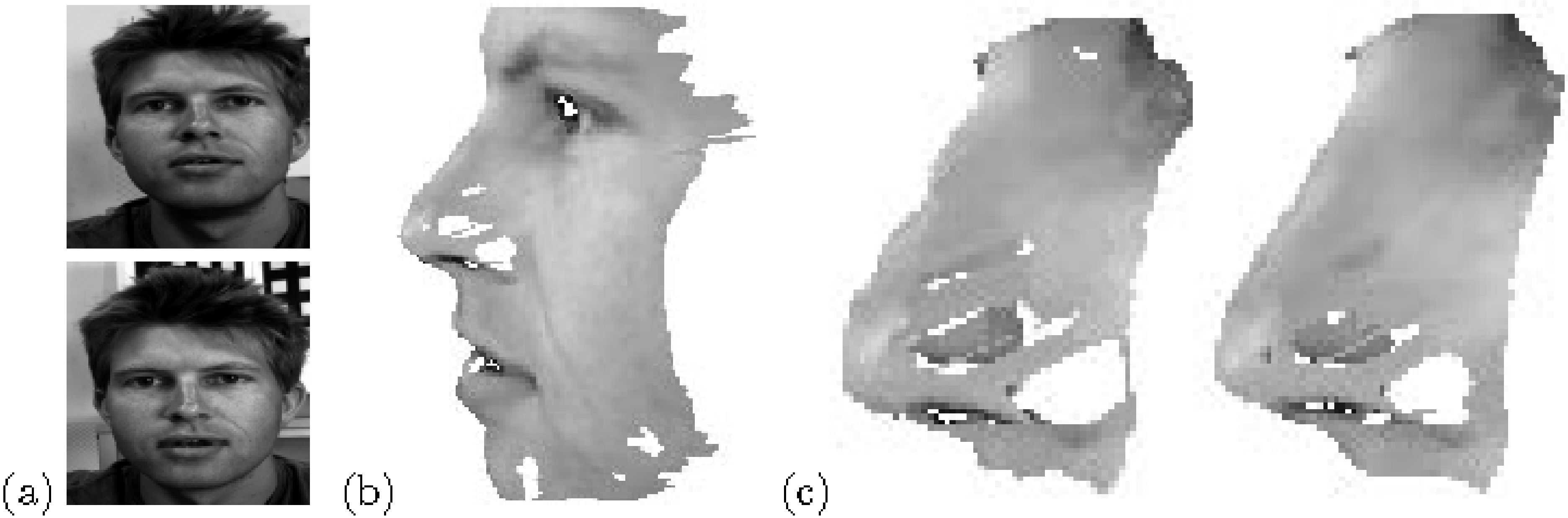
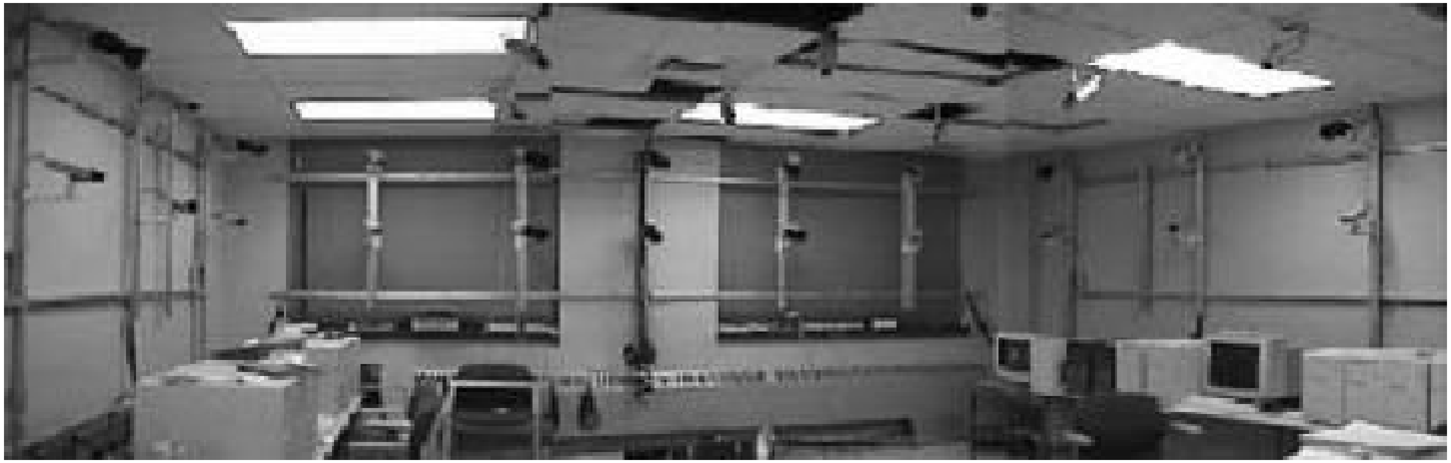


FIGURE 12.13: Correlation-based stereo matching: (a) a pair of stereo pictures; (b) a texture-mapped view of the reconstructed surface; (c) comparison of the regular (left) and refined (right) correlation methods in the nose region. Reprinted from [Devernay and Faugeras, 1994], Figures 5, 8 and 9.

CMU's 3-D Room

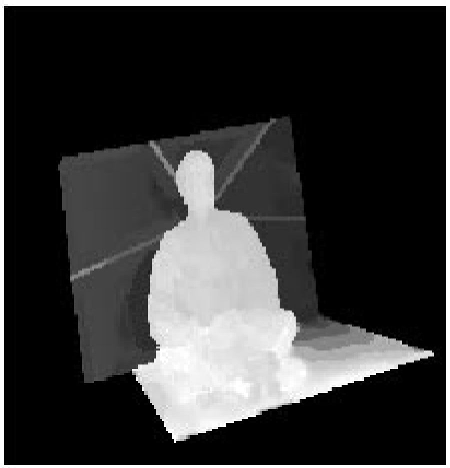
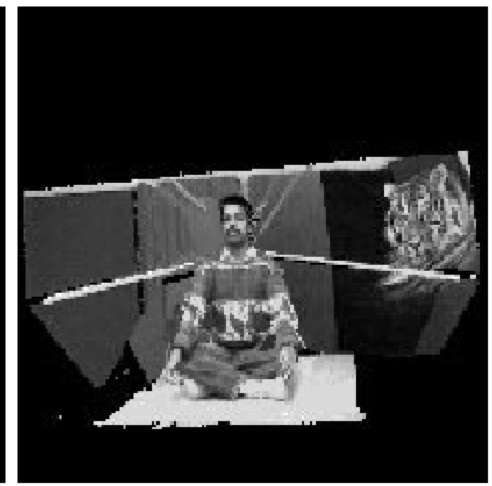
49 camera 3-D room:



[Kanade et al. 1998]

Multi-view stereo for VR

Compute dense range image from 3-6 nearby cameras:

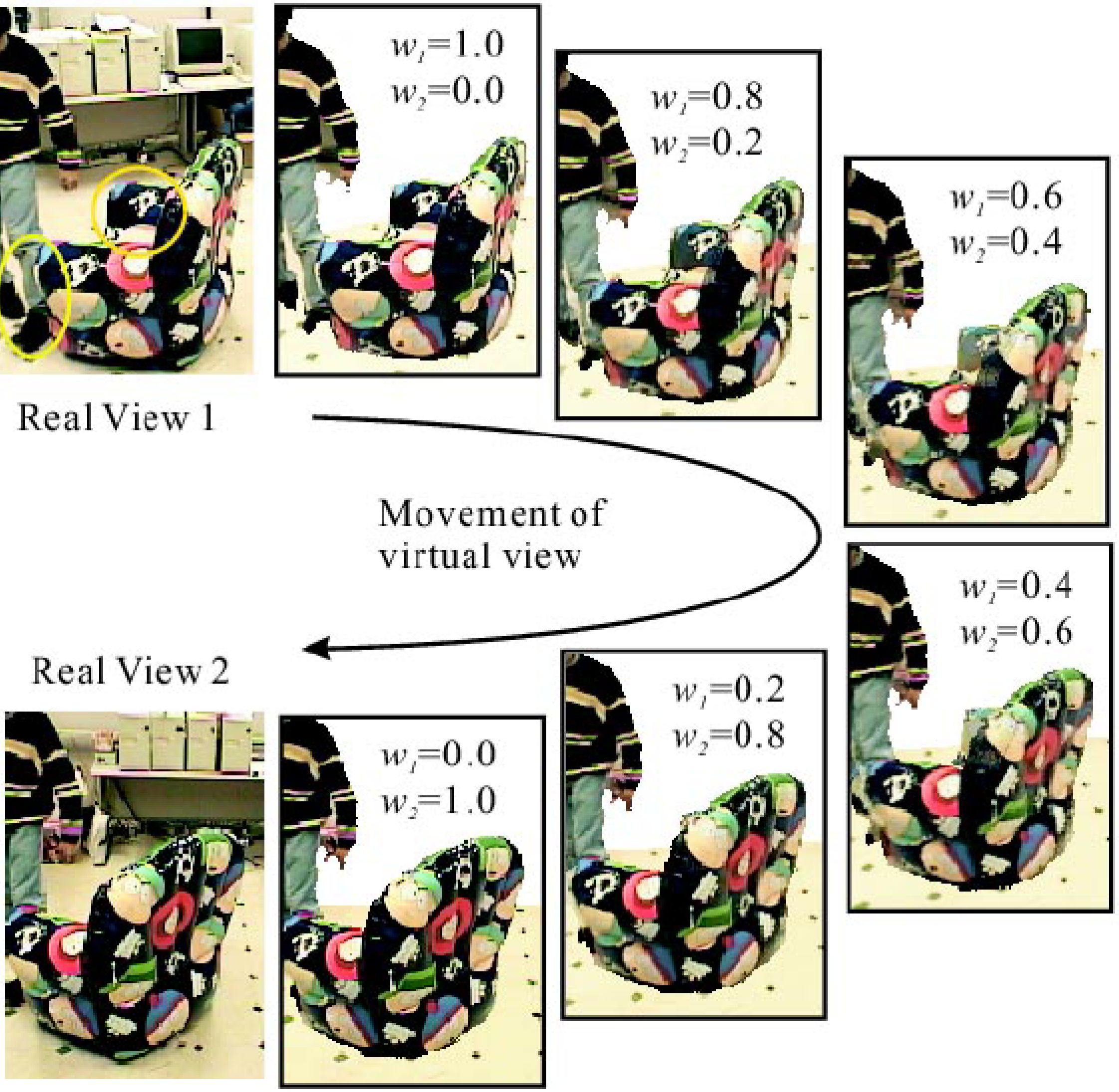
  

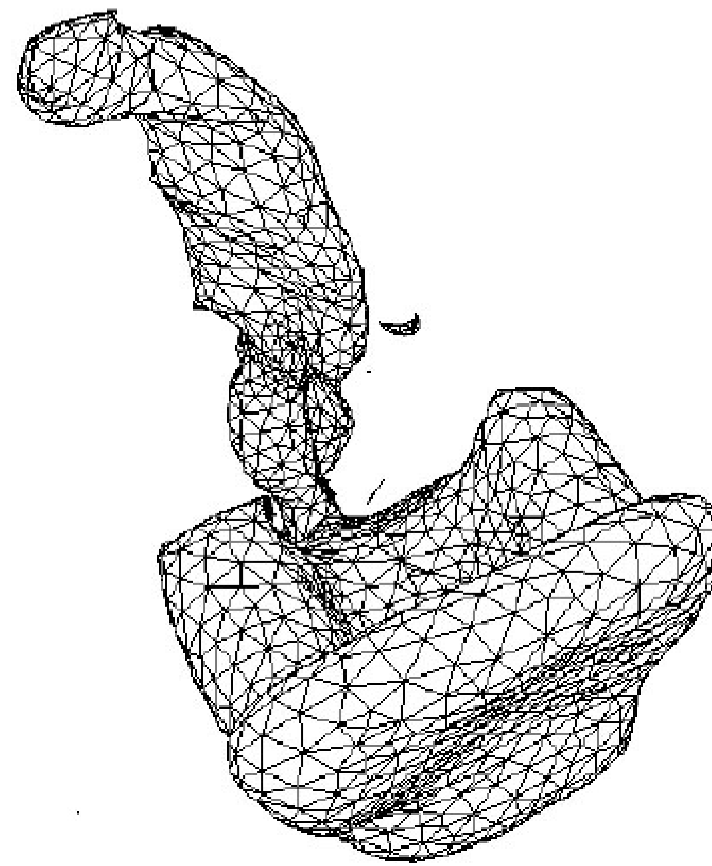
Merge into global mesh.

Texture and render new views….

[Kanade et al. 1998]

“Virtualized Reality”





[Kanade et al. 1998]

Models

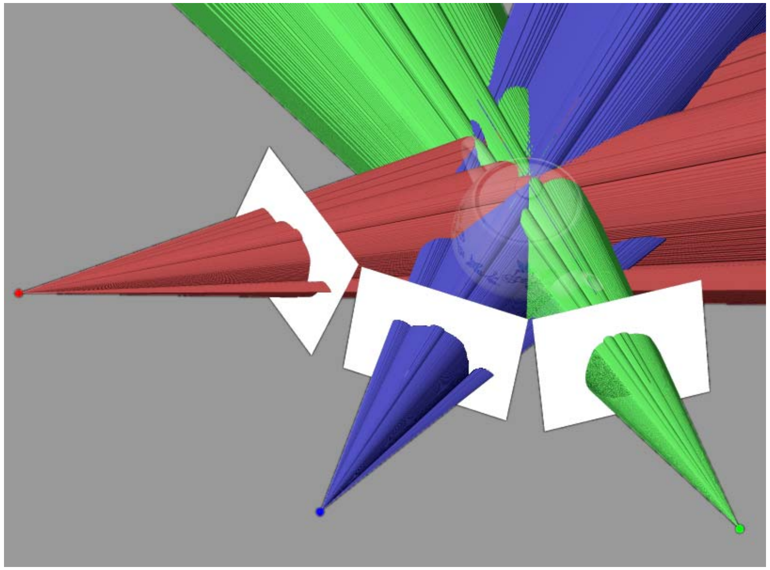
- Virtualized reality
  - very accurate
  - many correspondences
  - many cameras

- What can you do with a few cameras, and just silhouettes?

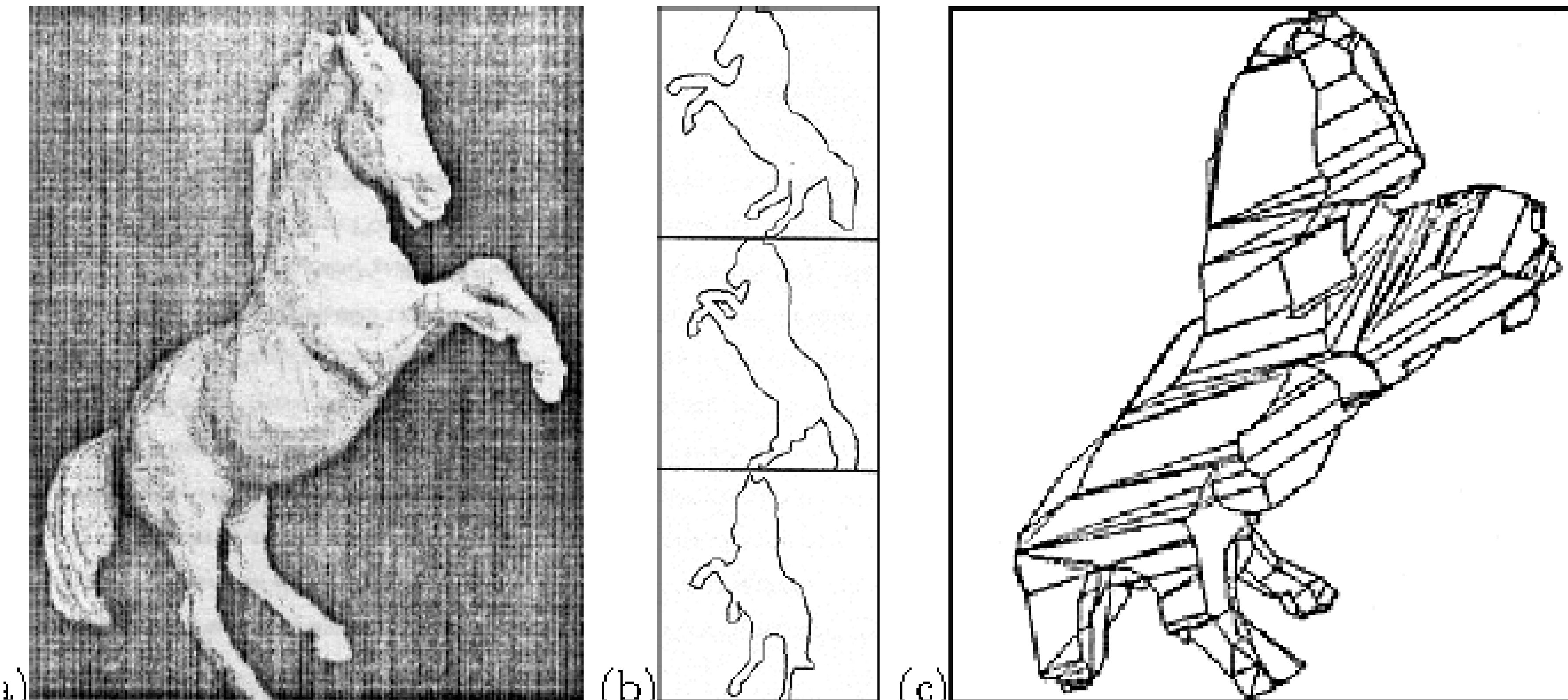
Visual Hulls

Visual Hull [Laurentini, 91]: the minimal object that produces the given silhouettes

- 3-D model contains the true object
- visual cone intersection
- texture mapped for a desired viewpoint

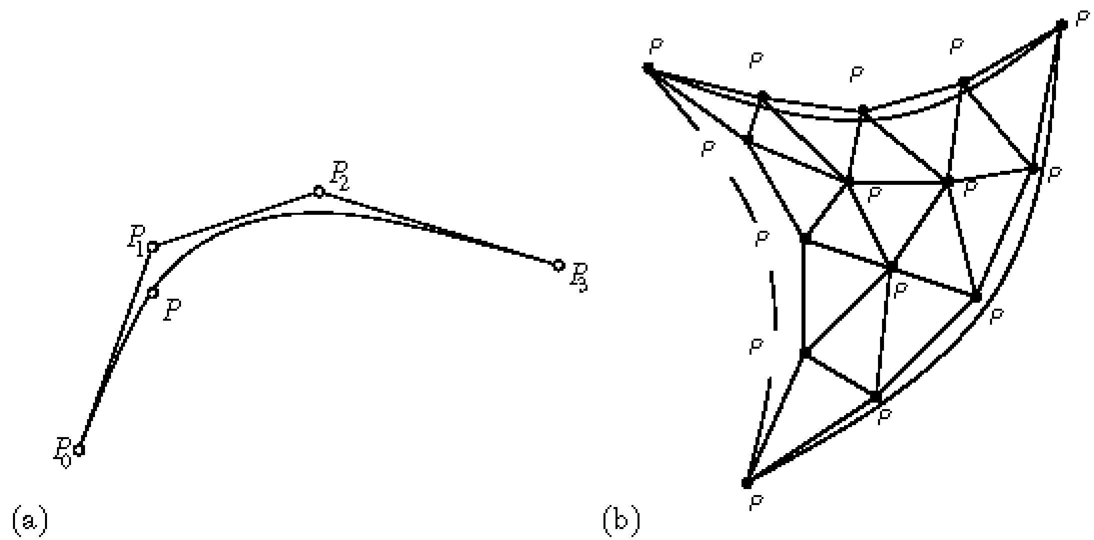


[Matusik]

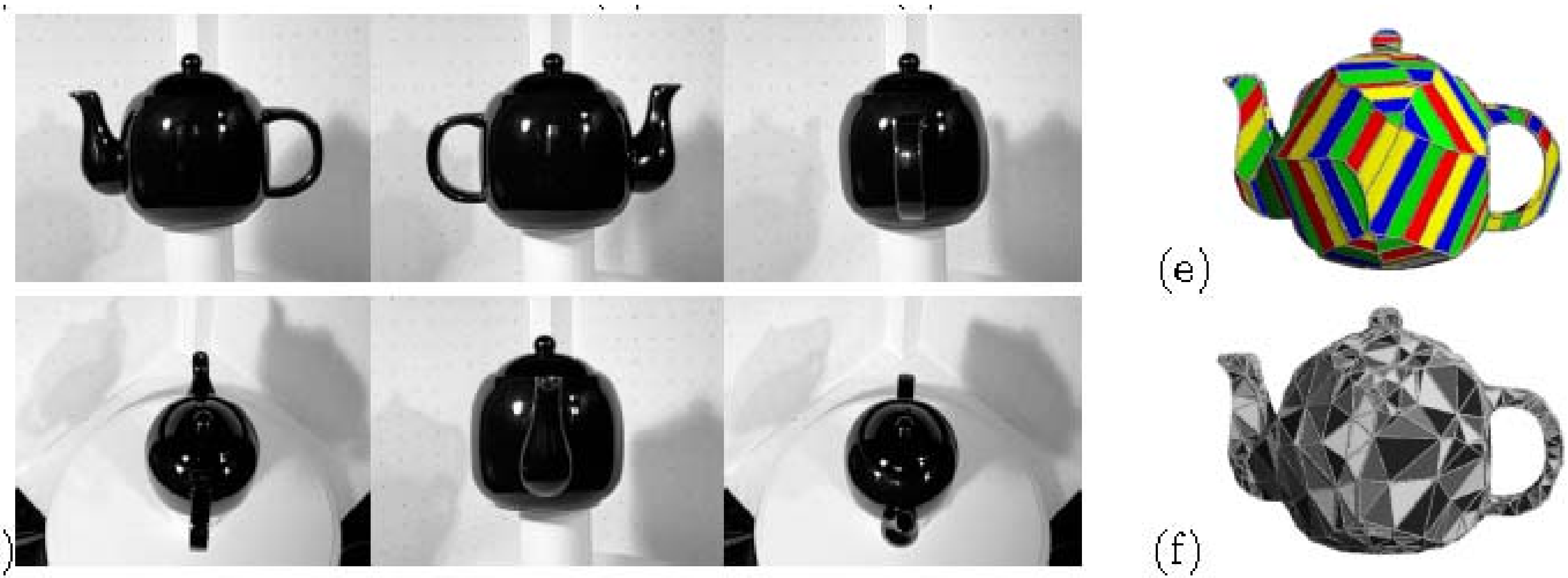

a)
(b)
(c)

Smoothed Visual Hull

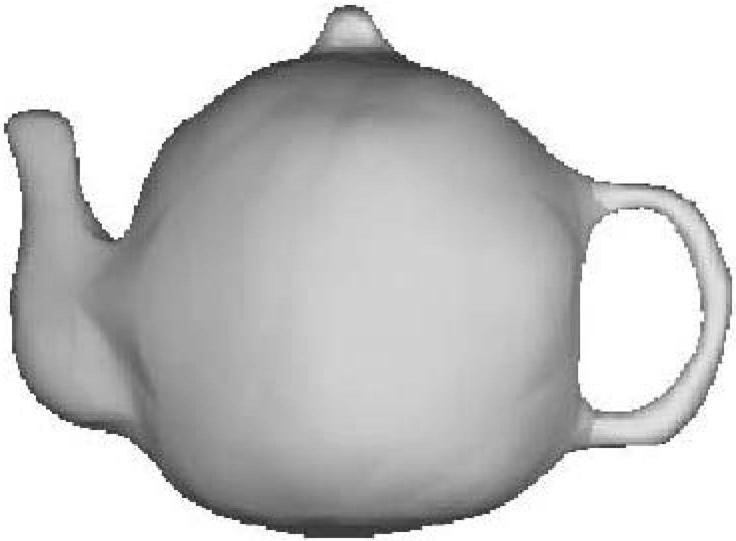
Fit surface spline to mesh; relax model according to smoothness assumption. [Sullivan and Ponce]



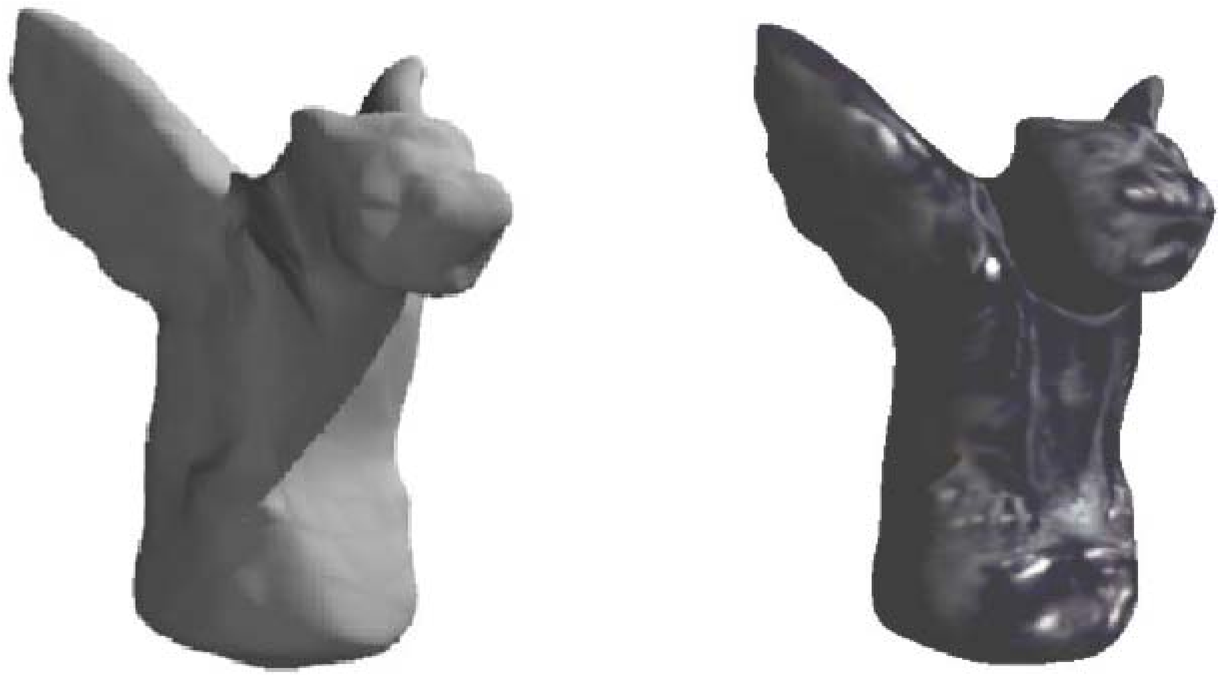
Smoothed Visual Hull



Smoothed Visual Hull Result



Smoothed Visual Hull Result



Smoothed Visual Hull Result

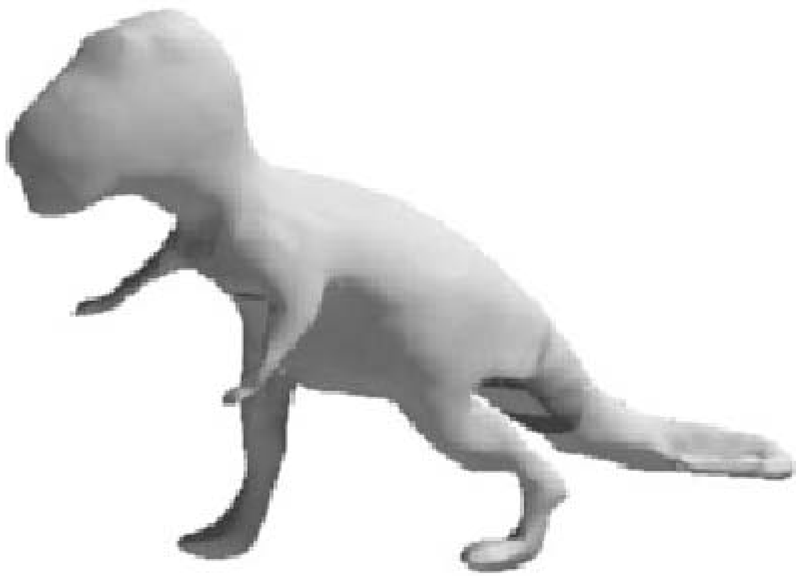
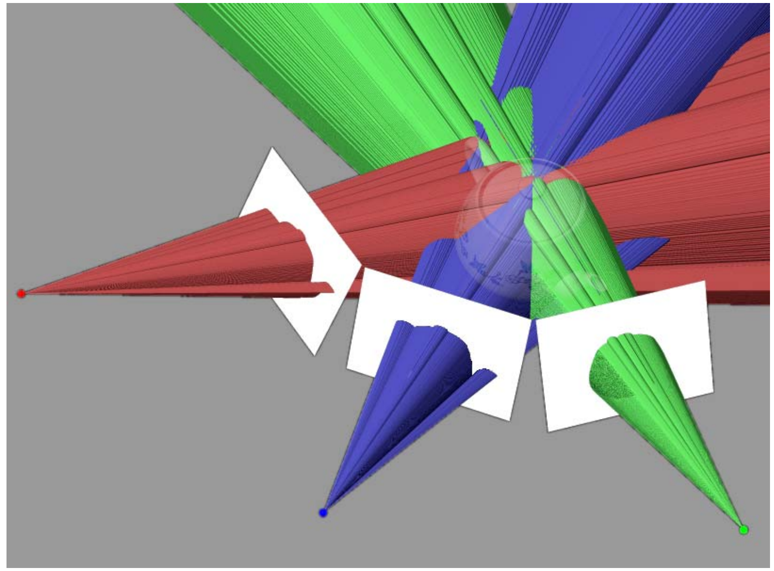


Image-based Visual Hulls



Visual Hull can be computed in $O\left(Kn^2\right)$ from K images with $n \times n$ pixels, without computing any explicit 3-D geometery

(Matusik *et al,* 2001)

Exploit view-dependent texture mapping (more later…)

Image-Based Visual Hulls

Model-based SFM

- Assume parametric shape model
  - boxes
  - prisms
  - solids of revolution
  - unknown height, width, etc…
  - constraints between unknowns
- Given marked features, fit model to image using (relatively simple) non-linear search.

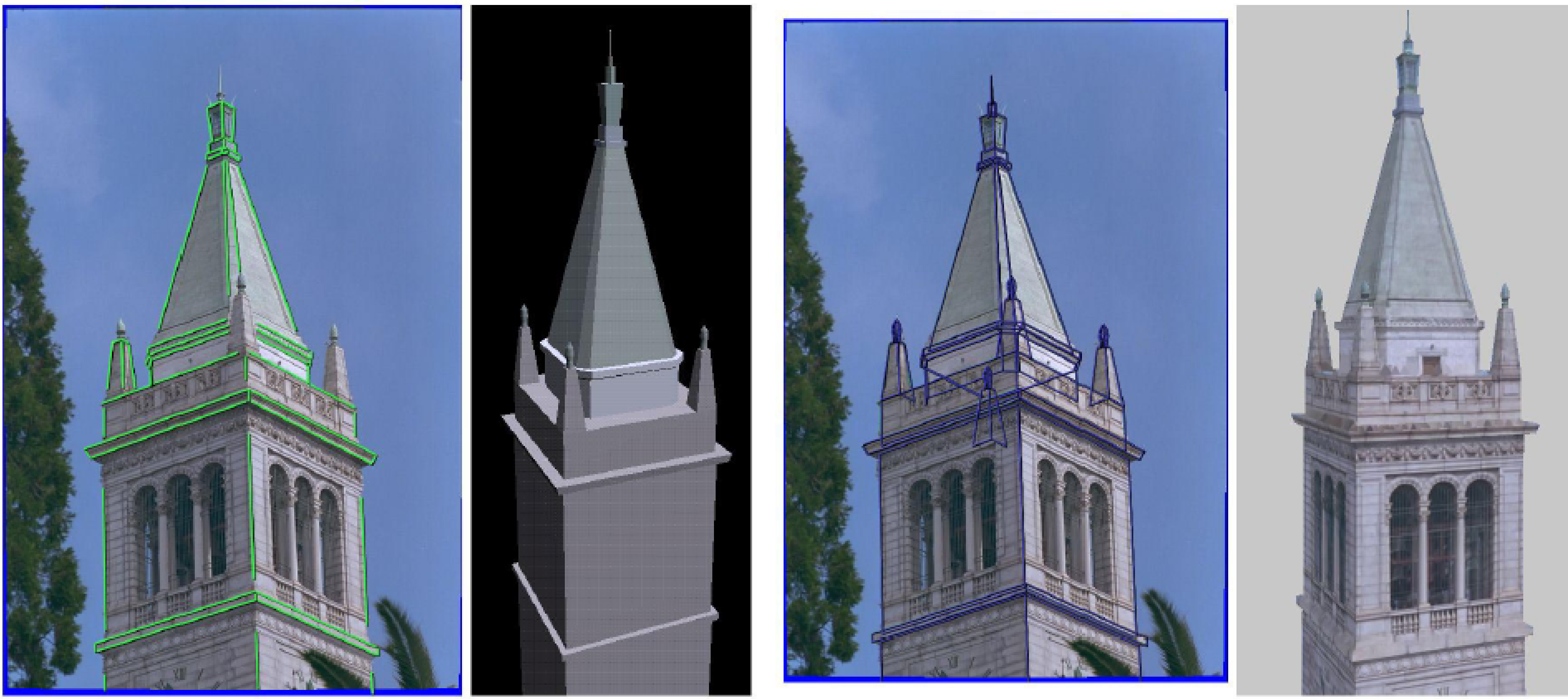
Façade

Visually compelling model from just a few photographs!

Three steps:

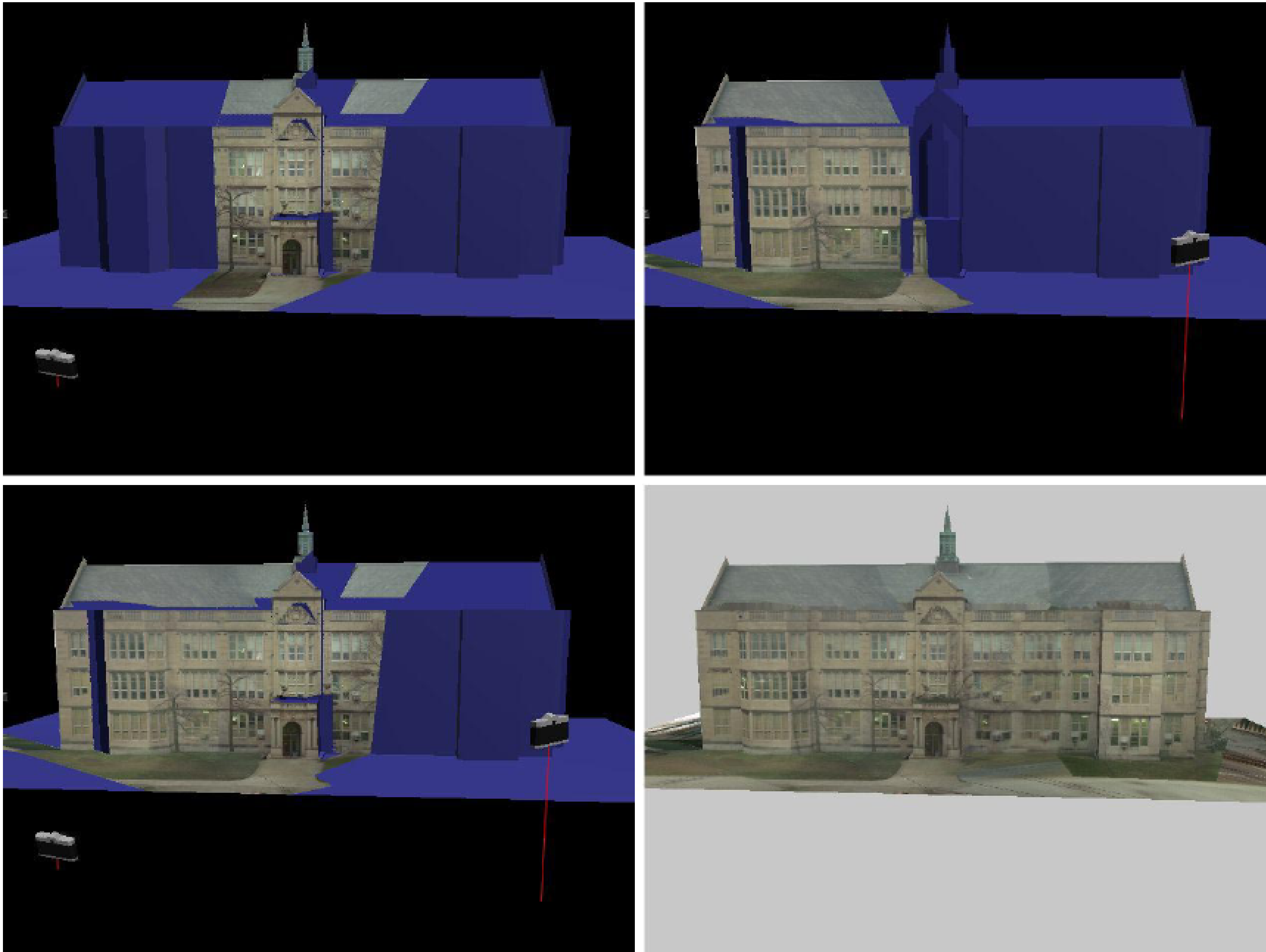
- Photogrammetry (Model-based SFM)
- View dependent Texture Mapping
- Model-based Stereopsis

Photogrammetry (Model-based SFM)

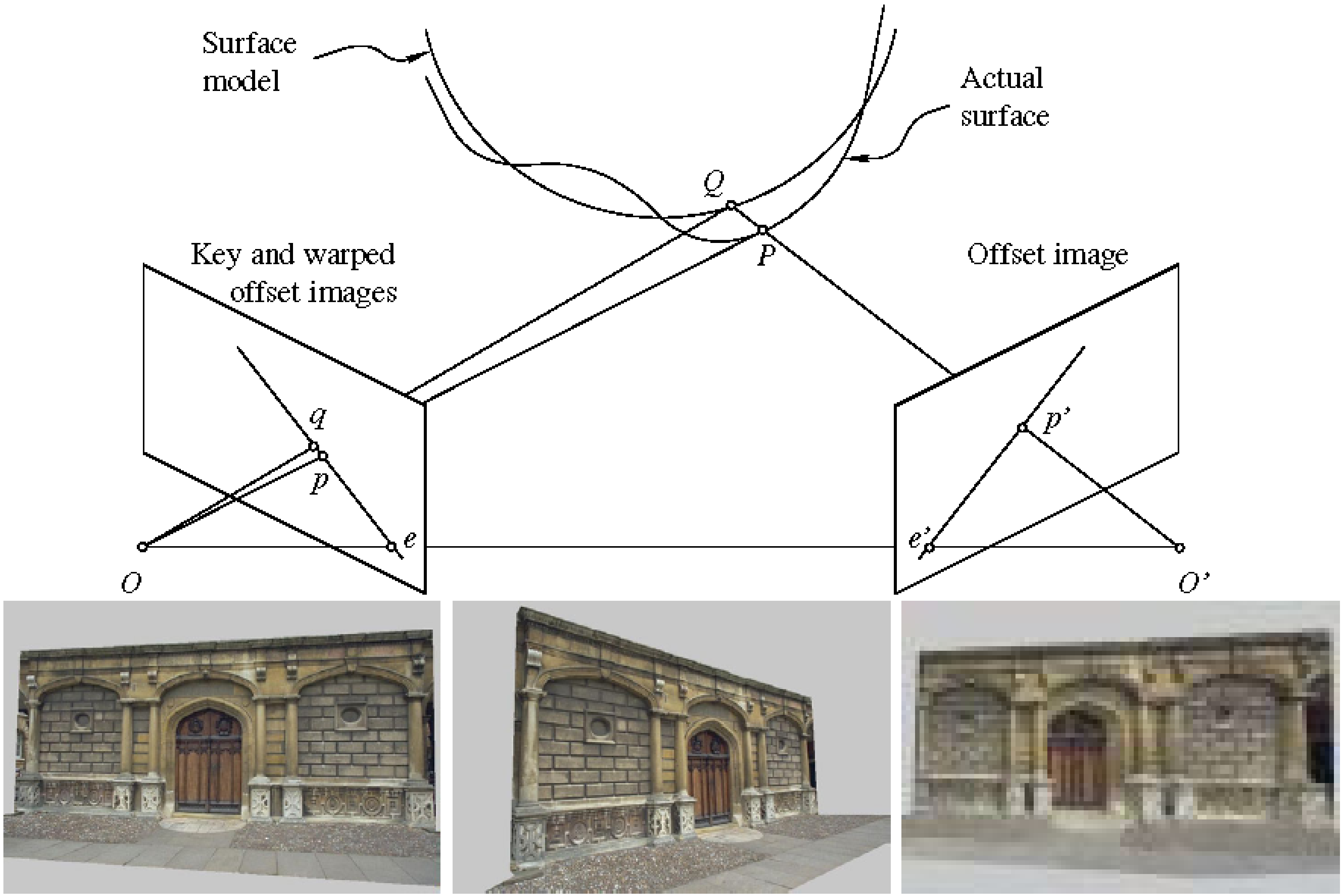


Line features recovered model model overlay recovered texture

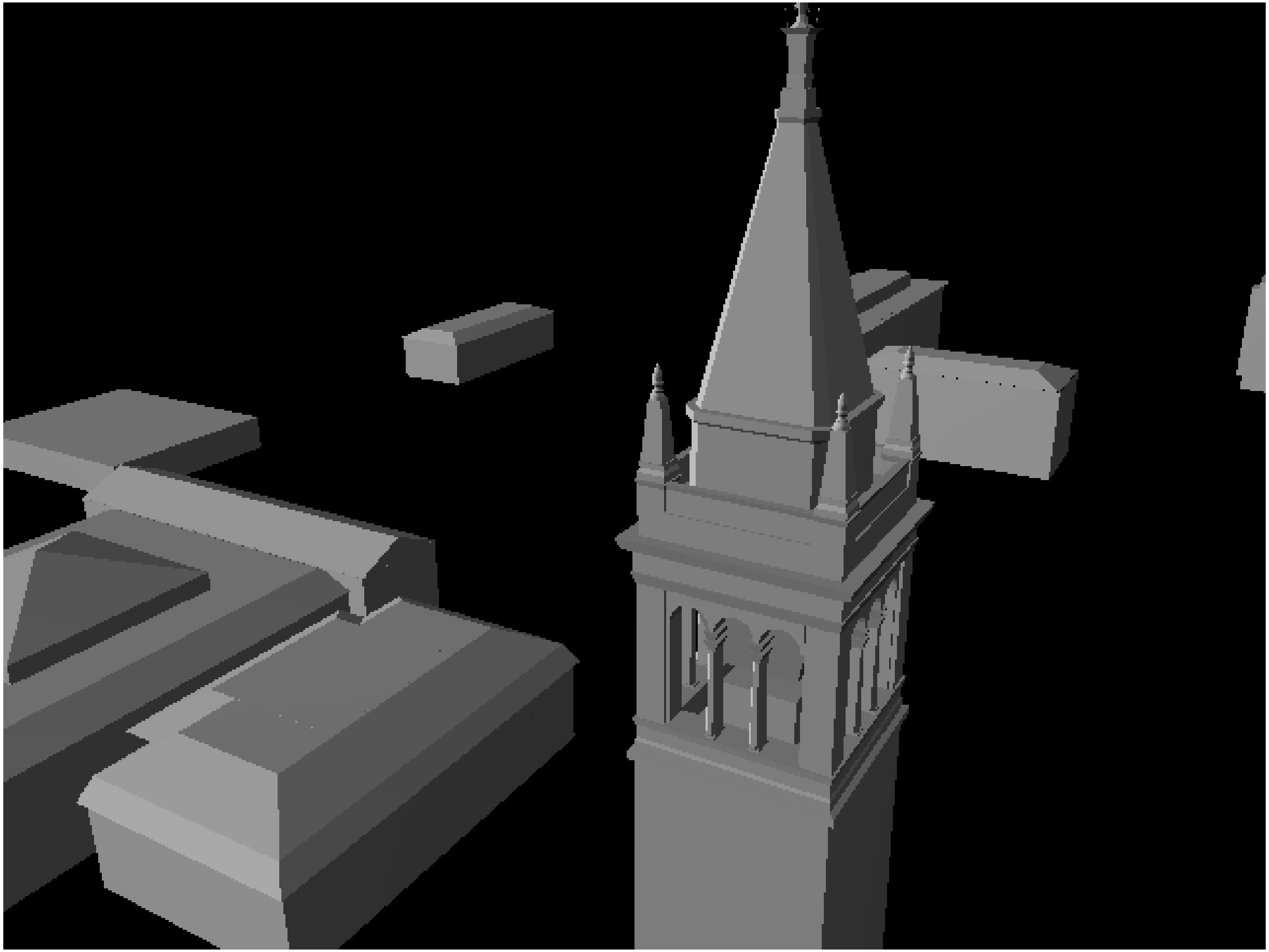
View-dependent texture



Model-based stereo for surface detail



Façade



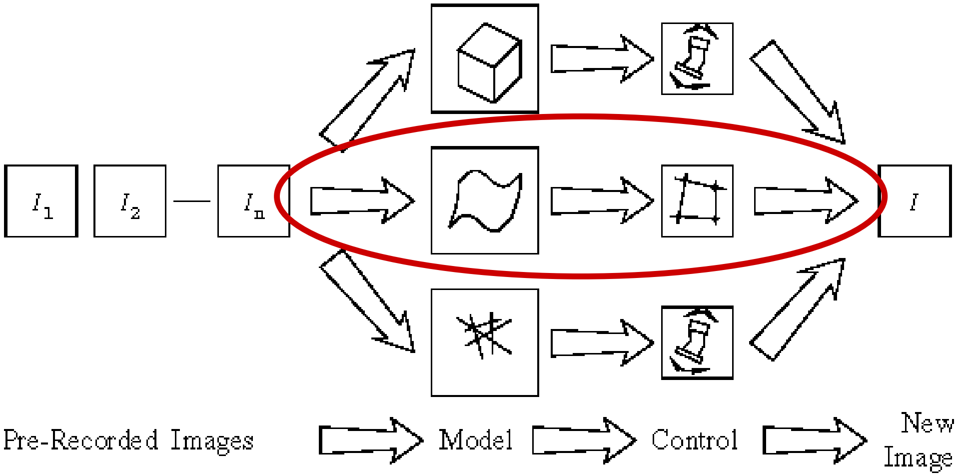
Façade

Façade Movie

Calibration/model free IBR?

- Cameras are hard to calibrate…desirable to have IBR methods that work without external/scene knowledge
- Recover affine structure from motion
- Use to insert virtual objects that follow camera motion…

Taxonomy



1. Build a 3-D model; re-render from new viewpoint
   - Multi-view stereo; “Virtualized Reality”
   - Visual Hulls
   - Model-based stereo
2. Establish correspondences; use view transfer
   - Affine view synthesis
3. Model sets of light rays
   - Lightfields, Lumigraphs, …

Affine image transfer

Use affine model …Given P_0-P_3, and

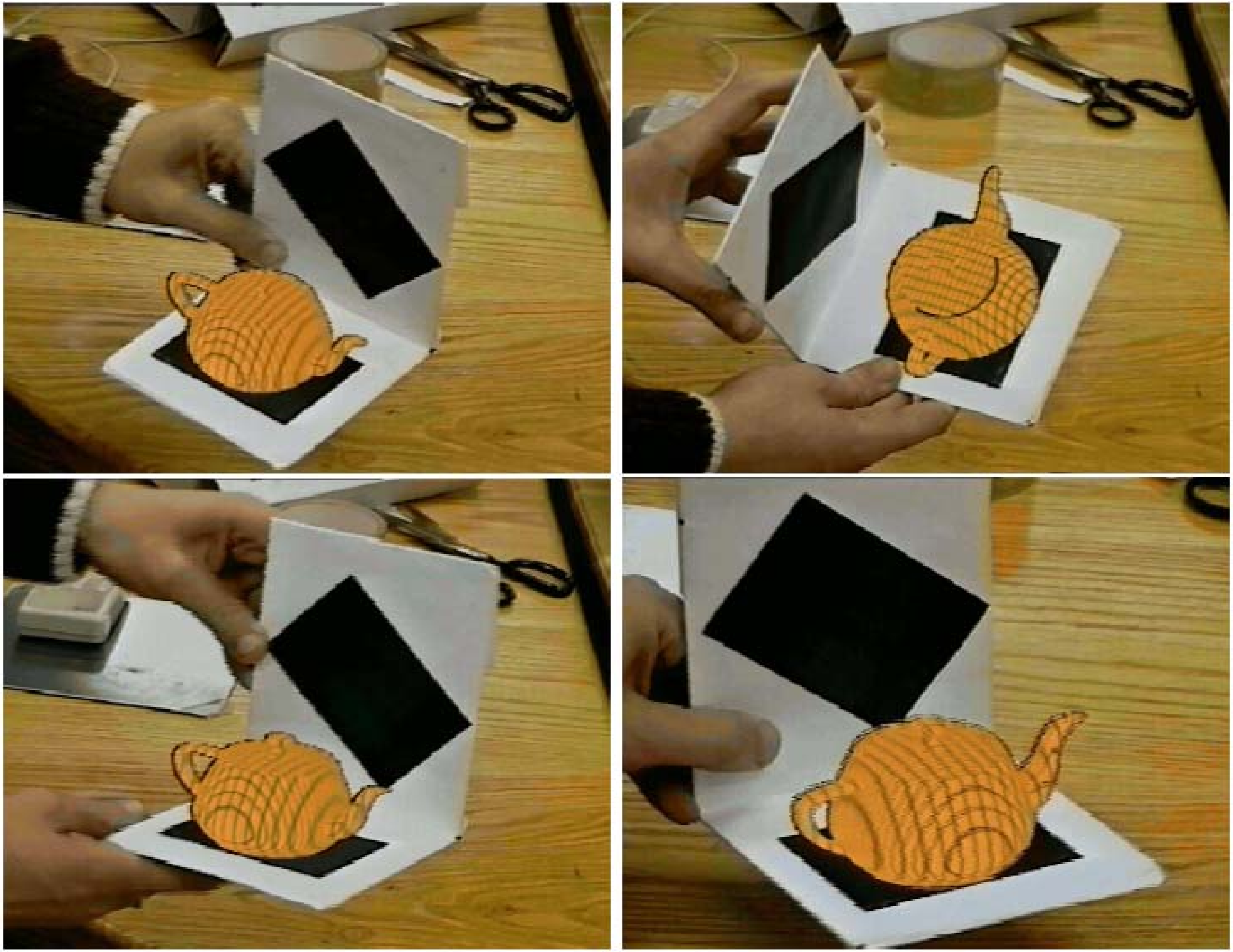
$$\boldsymbol{p} = \mathcal{A}\boldsymbol{P} + \boldsymbol{b}, \quad \text{where} \quad \mathcal{A} = \begin{pmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \end{pmatrix}$$

With appropriate choice of 4 bases we can express projected location of points as:

$$\boldsymbol{p} = (1 - x - y - z)\boldsymbol{p}_0 + x\boldsymbol{p}_1 + y\boldsymbol{p}_2 + z\boldsymbol{p}_3.$$

1. Given m>=2 images of p_0-p_3 and p solve using least-squares for x,y,z
2. Use x,y,z and positions of p_0-p_3 in new view to find p in new view.

Augmented reality

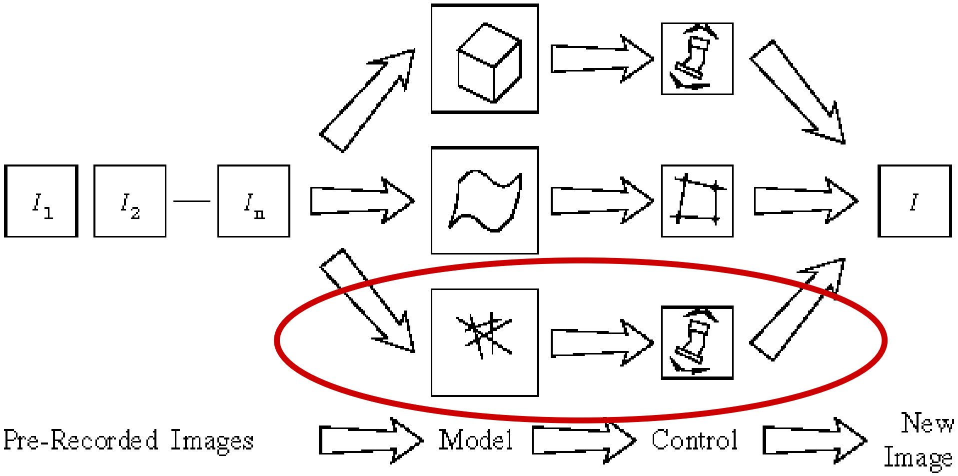


Find cameras with black squares; add virtual object to scene with correct camera motion.

Model recovery

- View transfer good for many special effects and augmented reality applications.
- For model recovery, dense correspondence is needed!
- But correspondence is hard! … (and/or models are approximate)
- What can we do without correspondence?
- Model visible rays, not shape….

Taxonomy

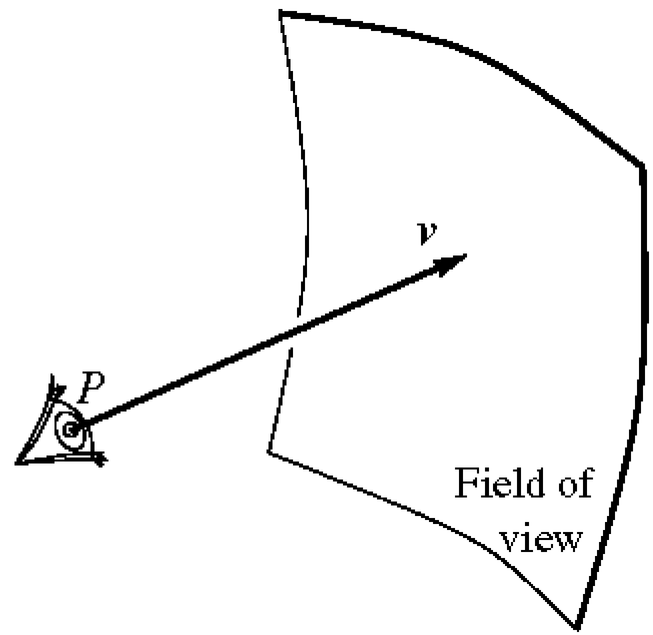


1. Build a 3-D model; re-render from new viewpoint
   - Multi-view stereo; “Virtualized Reality”
   - Visual Hulls
   - Model-based stereo
2. Establish correspondences; use view transfer
   - Affine view synthesis
3. Model sets of light rays
   - Lightfields, Lumigraphs, …

The Plenoptic function

- IBR → recover geometric and photometric models from photographs, bypass the modeling process.
- *Plenoptic function*: images that can be seen!

- What parameterizes visible rays?
  - Camera position
  - Viewing angle
  - Wavelength
  - Time



(In a non-dispersive medium…)

The Plenoptic function

- Adelson and Bergen's Plenoptic function
- 7D → 5D → 4D → 2D

7D: $(c_x, c_y, c_z, \theta, \varphi, \lambda, t)$

5D: $(c_x, c_y, c_z, \theta, \varphi)$

4D: (x_1, y_1, x_2, y_2)

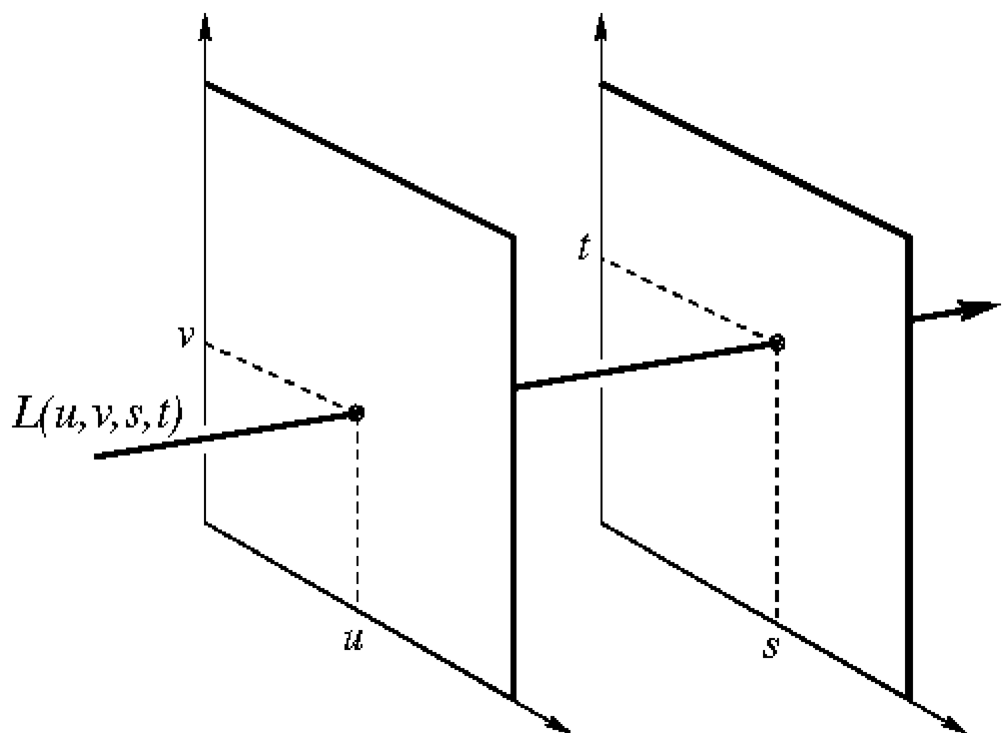
2D: (θ, φ)

Dimension	Viewing space	Name	Year
7	free	plenoptic function	1991
5	free	plenoptic modeling	1995
4	inside a 3D box	Lightfield/Lumigraph	1996
3	inside a 2D circle	concentric mosaics	1999
2	at a fixed point	panorama	1994

[Shum and He]

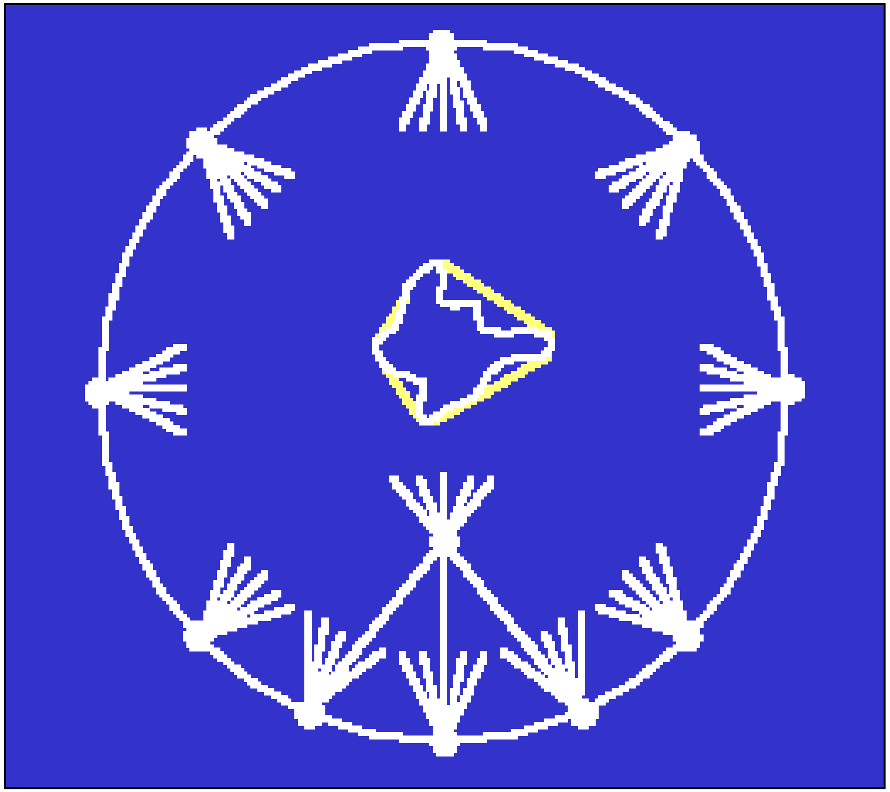
Lightfields

- Approximate Plenoptic function for fixed camera location, time, …
- Reparametrize rays based on planar intersection
- A "light slab":



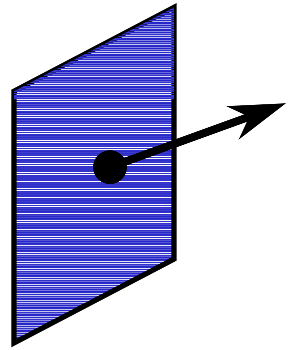
Lightfields

- Generally, 2D slices of 4D data set
- For a new views compute other 2D slices
- Challenges:
  - Capture
  - Parameterization
  - Compression
  - Rendering

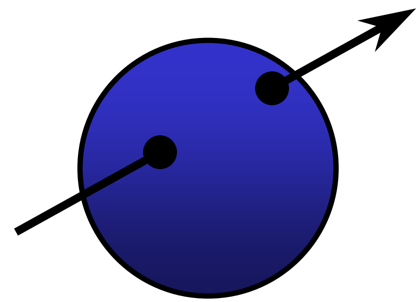


Alternate representations

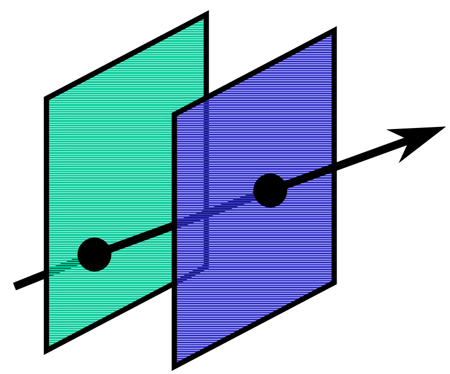
- Point / angle



- Two points on a sphere



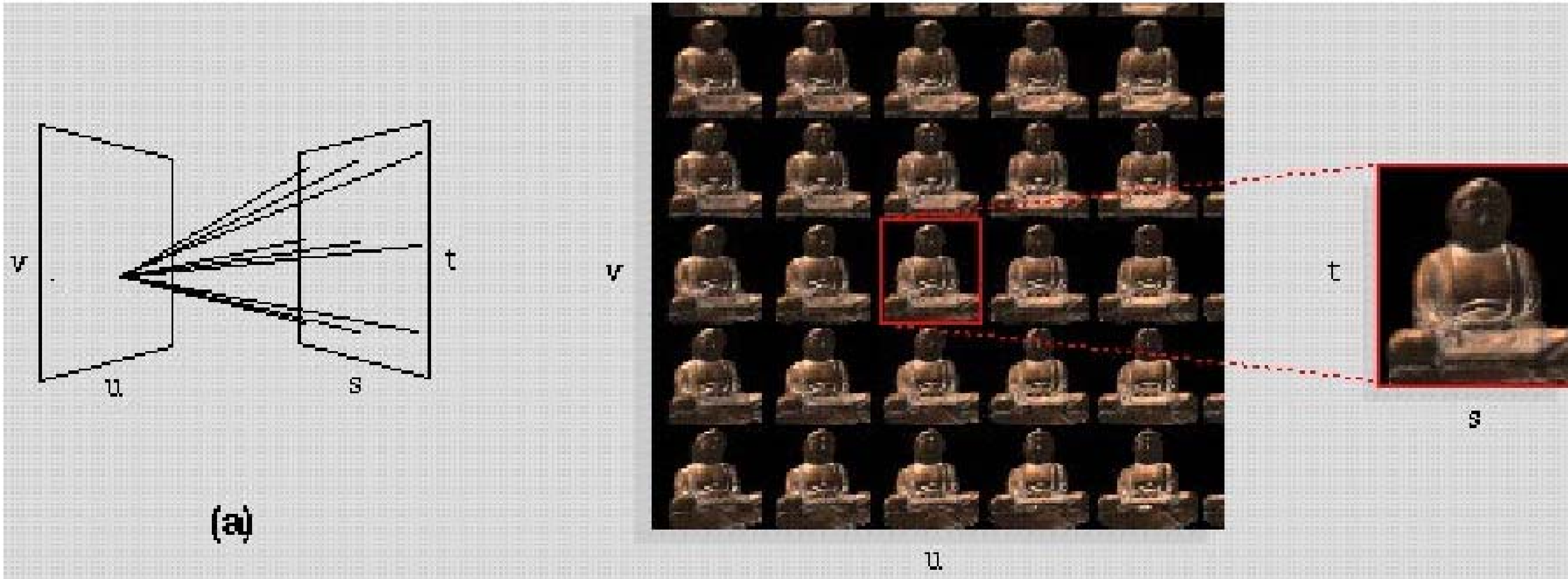
- Points on two planes

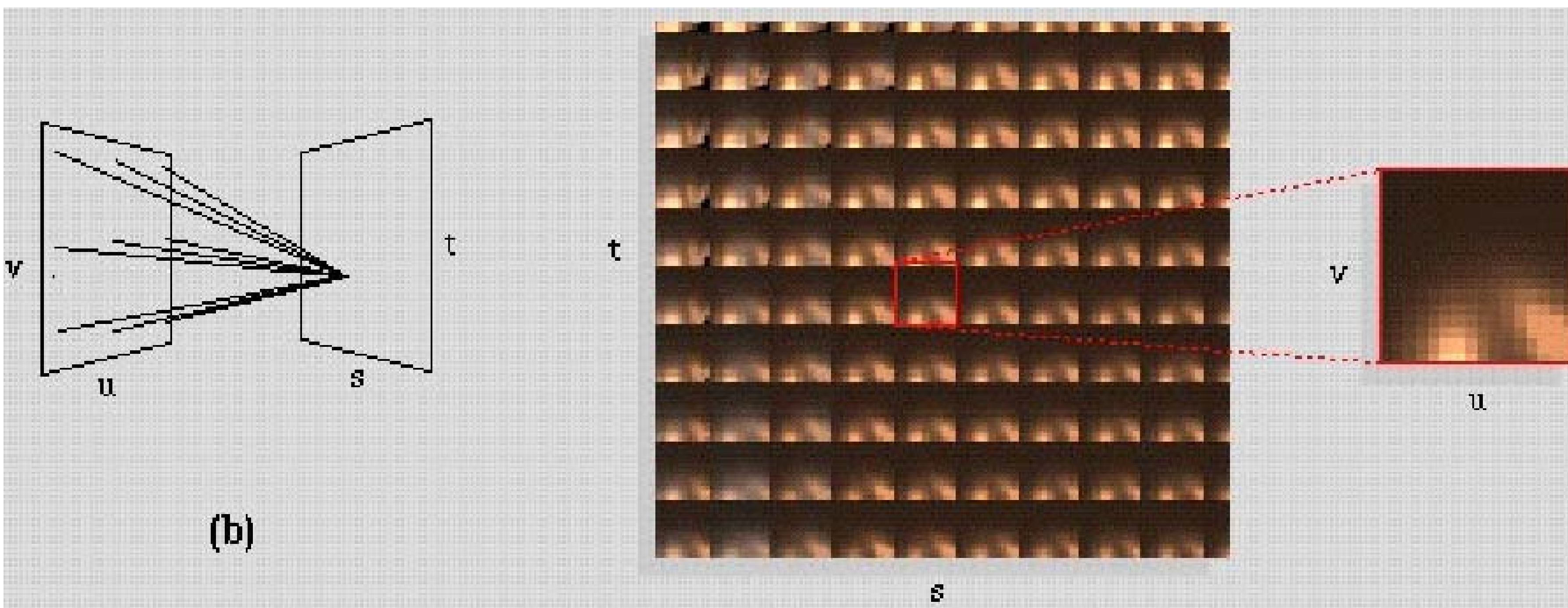


- Original images and camera positions…

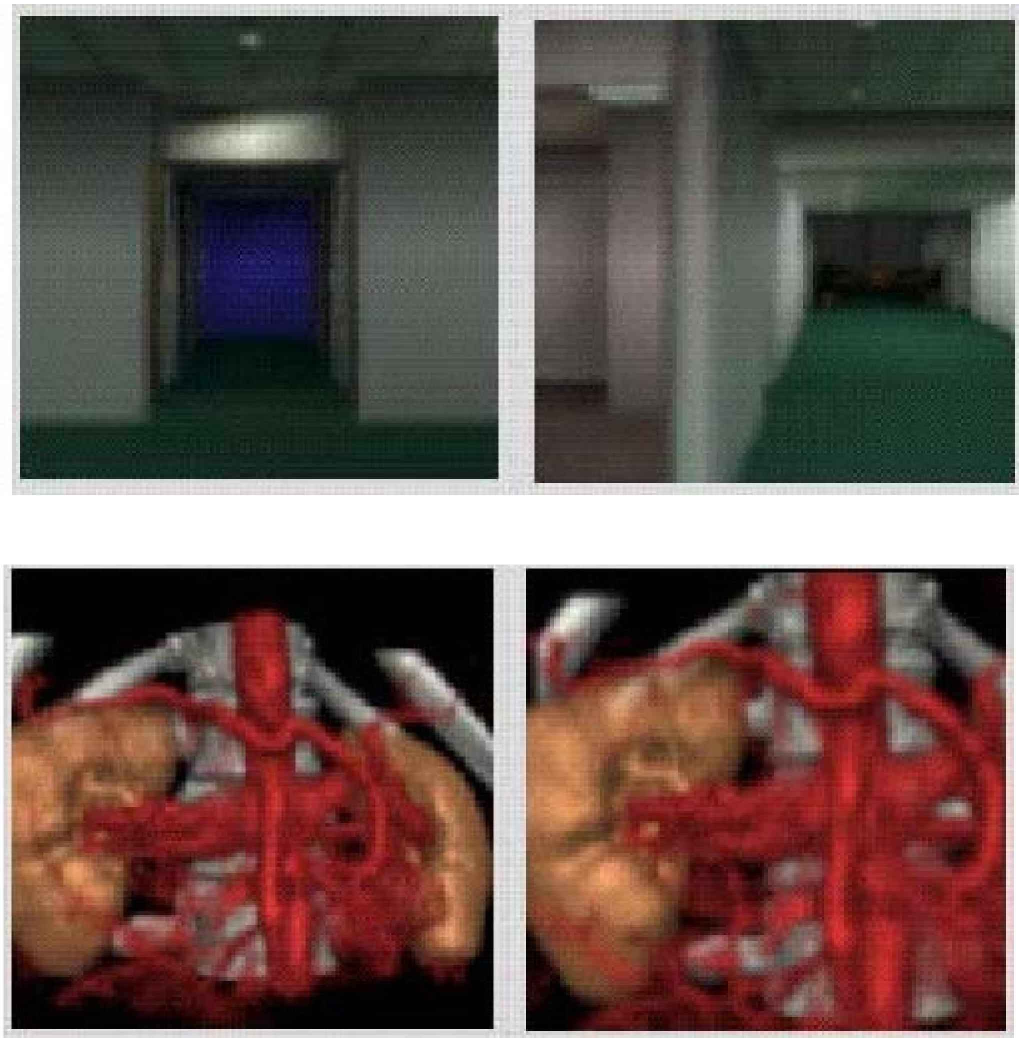
Light-field rendering

- Compute intersection with (u,v) and (s,t) planes, take closest ray
- Interpolation possibilities
  - Bilinear in (u,v) only
  - Bilinear in (s,t) only
  - Quadrilinear in (u,v,s,t)

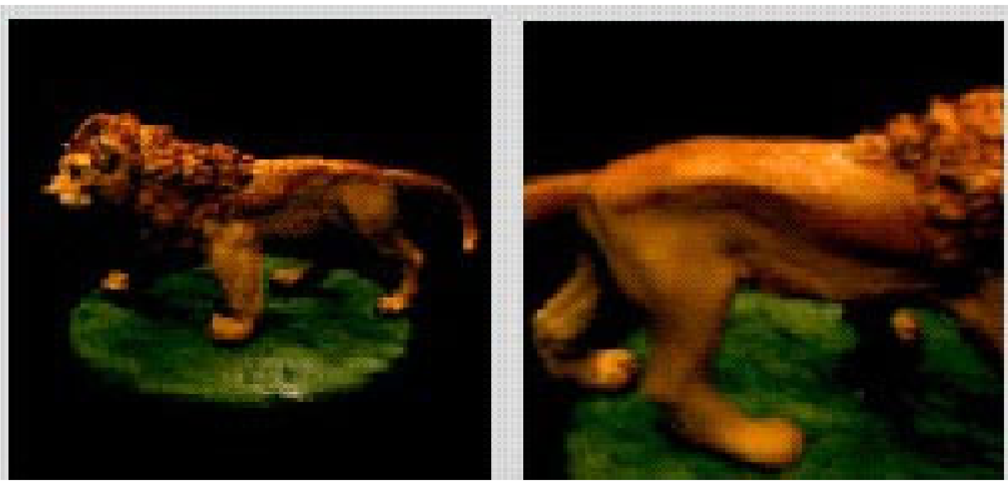

v
u
s
t
(a)
v
u
t
s


v
u
s
t
(b)
t
s
v
u

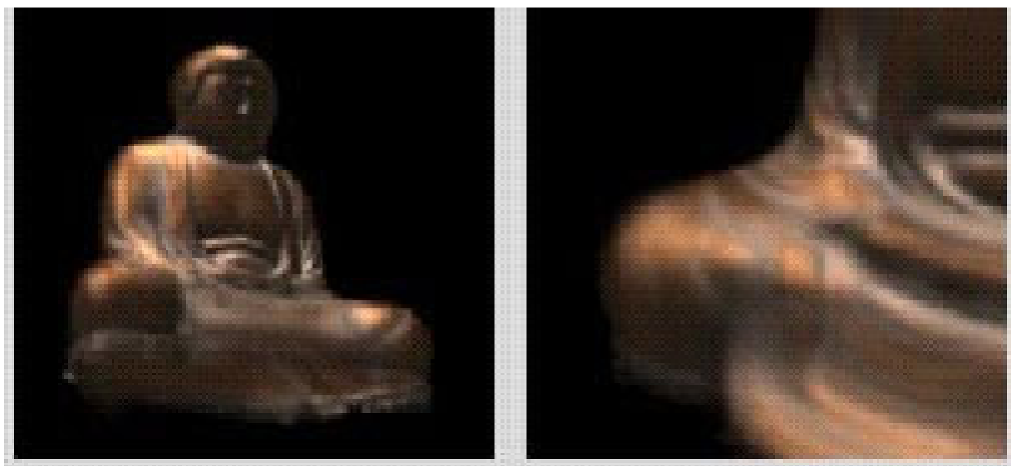
Example lightfields



Example lightfields



Example lightfields



Unstructured Lumigraph

Generalize model-based view-dependent texture mapping (e.g., Façade) and Lightfield

Both are methods for interpolating color values for a desired ray as some combination of input rays.

- VDTM: use geometric model as proxy
- LFR: planar light "slab"

The Unstructured Lumigraph [Buehler 2001] is an IBR algorithm that includes VDTM and LFR as special cases, and has nice properties of each.

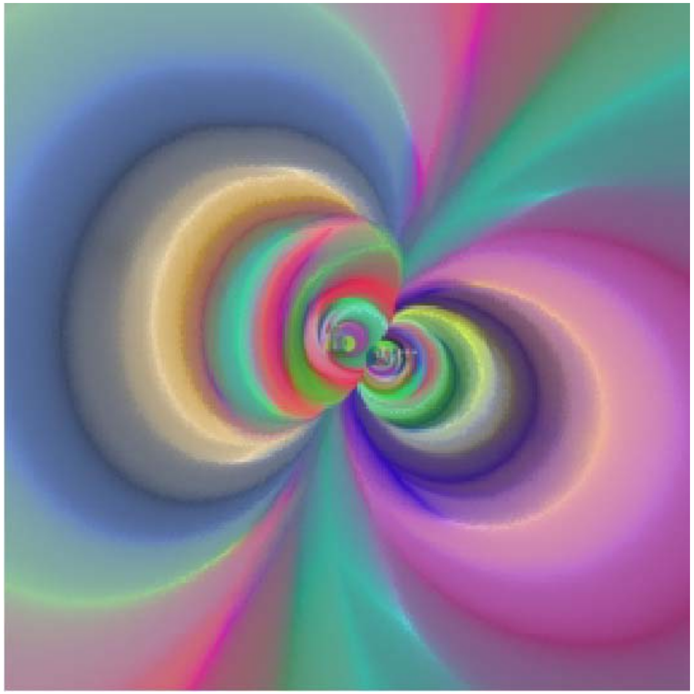
Unstructured Lumigraph

Desirable properties

- Geometric proxies
- Unstructured input
- Epipole consistency
- Minimal angular deviation
- Continuity
- Resolution Sensitivity
- Equivalent ray consistency
- Real-time

Unstructured Lumigraph Rendering

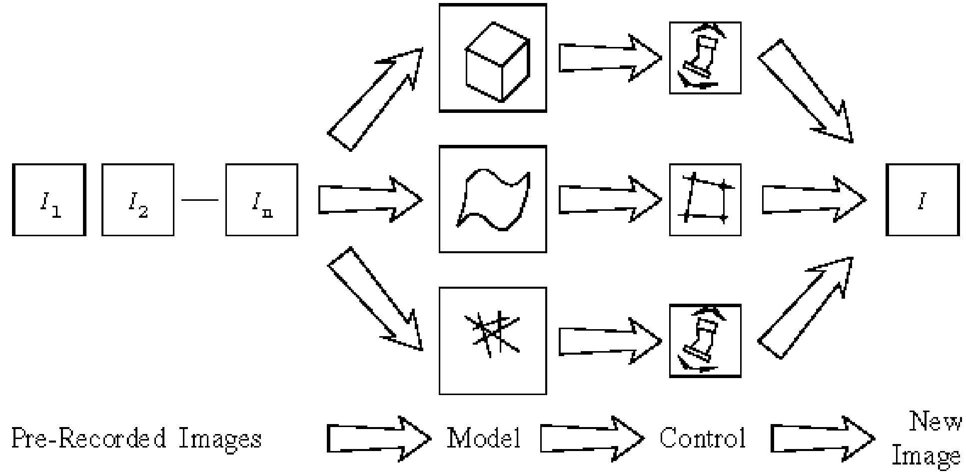
- Example: hallway with “tunnel” geometric proxy (inside of cube).
- Images gathered from translating robot.
- 3-D effect with no (local) 3-D structure….



Blending field

Rendering

Taxonomy



1. Build a 3-D model; re-render from new viewpoint
   - Multi-view stereo; “Virtualized Reality”
   - Visual Hulls
   - Model-based stereo
2. Establish correspondences; use view transfer
   - Affine view synthesis
3. Model sets of light rays
   - Lightfields, Lumigraphs, …

[Figures from Forsythe and Ponce unless Attributed]

Endgame

- Exams due today
- Project show and tell on Tuesday—3 slides for L.M.—stand up and give a 2-3 minute overview (videotaped)
- Projects by 12/10 (or electronically by 12/15 with extension)

23	11/26	Model-Based Vision	Req: FP 18	
	11/28	Thanksgiving (NO LECTURE)		
24	12/3	Image Databases	Req: FP 25	
25	12/5	Image-Based Rendering	Req: FP 26	Exam #2 Due
26	12/10	Project Show and Tell		Projects Due Submit 3 Slides

Recap

#	Date	Description
1	9/5	Course Introduction
2	9/10	Cameras, Lenses, and Sensors
3	9/12	Radiometry and Shading Models I
4	9/17	Radiometry and Shading Models II
5	9/19	Multiview Geometry
6	9/24	Stereo
7	9/26	Color
8	10/1	Shape from Shading
9	10/3	Image Filtering
10	10/8	Image Representations
11	10/10	Texture and Edges
12	10/17	Bayesian Analysis
13	10/22	Optic Flow and Direct SFM
14	10/24	Affine Reconstruction
15	10/29	Interactive Systems (Low-Level)
16	10/31	Face Detection and Recognition I
17	11/5	Face Detection and Recognition II
18	11/7	Projective Reconstruction
19	11/12	Segmentation I
20	11/14	Segmentation II
21	11/19	Tracking I
22	11/21	Tracking II
23	11/26	Model-Based Vision
	11/28	Thanksgiving (NO LECTURE)
24	12/3	Image Databases
25	12/5	Image-Based Rendering
26	12/10	Project Show and Tell

Thanks!

We learned a lot! (and we hope you did too!)

Comments are very welcome to refine this class in the future.

Thanks to:

- Erik
- Louis
- You