

# Vision & Graphics

Frédo Durand  
MIT CSAIL

# Interaction is already strong

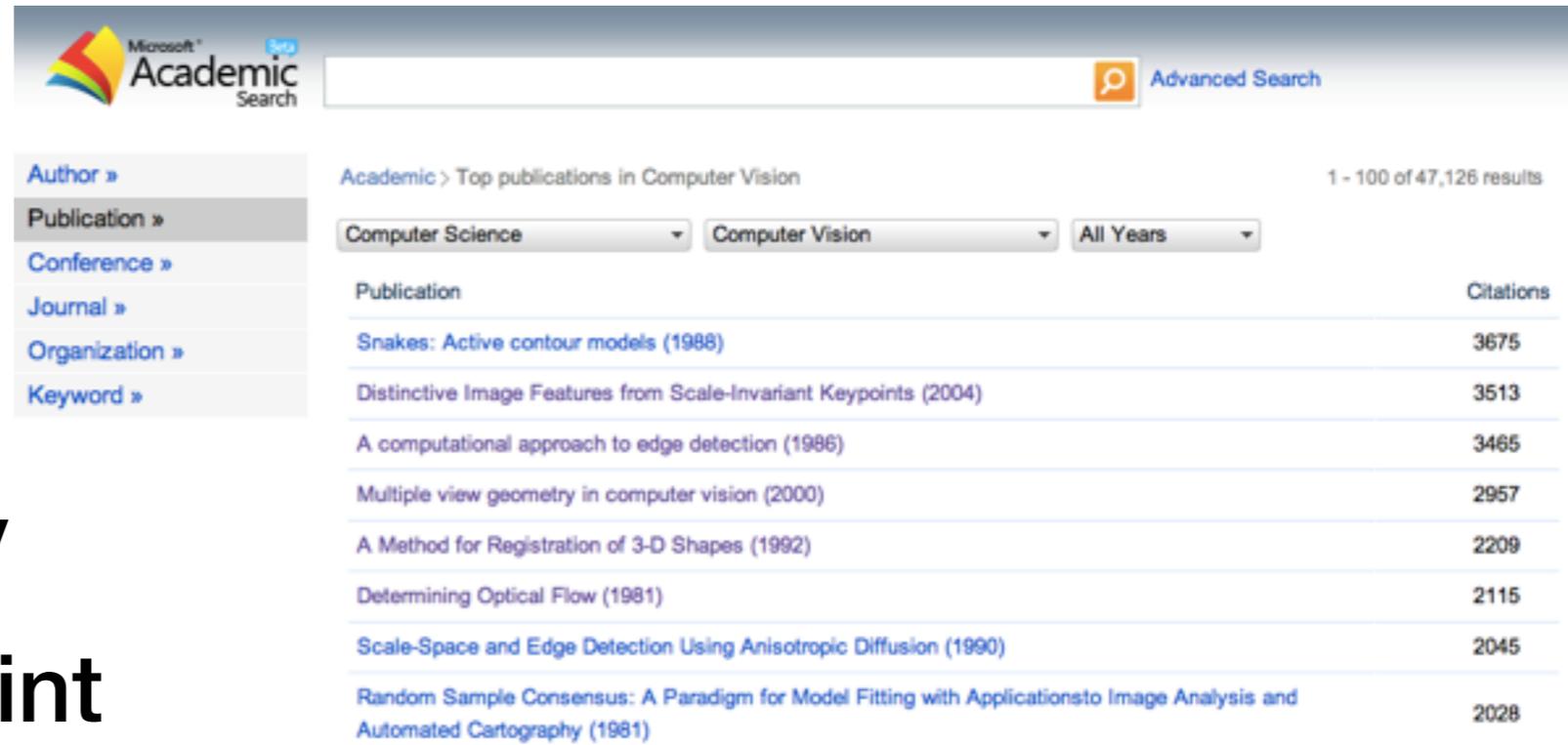
---

- **Rick Szeliski received the Siggraph Achievement Award 2011**
- **Anat Levin received the Eurographics Young Researcher Award 2010**
- **Vision is used in special effects**
- **Computational photography**



# Bibliometrics: top vision papers

- Snakes
- SIFT
- Edge detection
- Multiview geometry
- Iterated Closest Point
- Optical flow
- Anisotropic diffusion
- Ransac
- Stereo
- Harris corners
- ... are used in graphics



Microsoft Academic Search

Advanced Search

Academic > Top publications in Computer Vision

1 - 100 of 47,126 results

Computer Science Computer Vision All Years

Publication	Citations
Snakes: Active contour models (1988)	3675
Distinctive Image Features from Scale-Invariant Keypoints (2004)	3513
A computational approach to edge detection (1986)	3465
Multiple view geometry in computer vision (2000)	2957
A Method for Registration of 3-D Shapes (1992)	2209
Determining Optical Flow (1981)	2115
Scale-Space and Edge Detection Using Anisotropic Diffusion (1990)	2045
Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography (1981)	2028

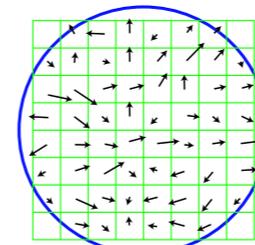
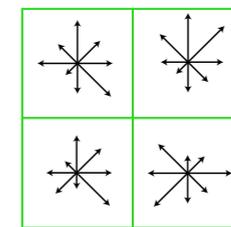
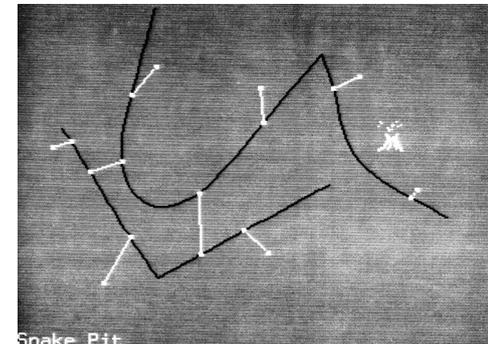


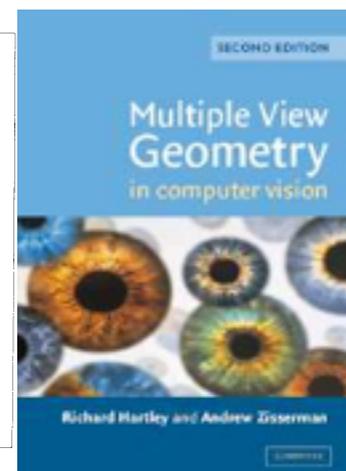
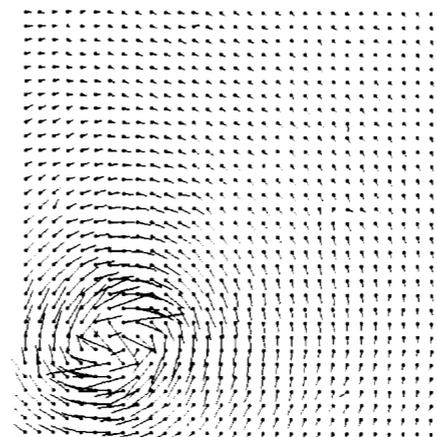
Image gradients



Keypoint descriptor



Snake Fit



# Computer vision

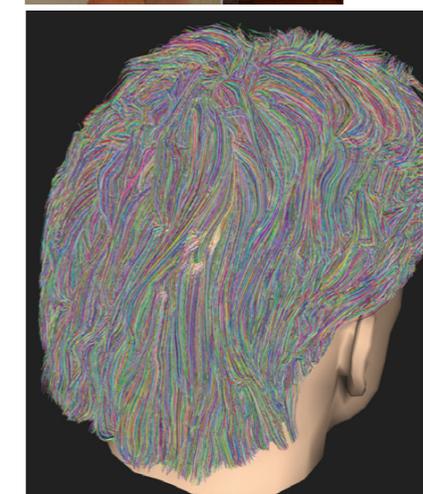
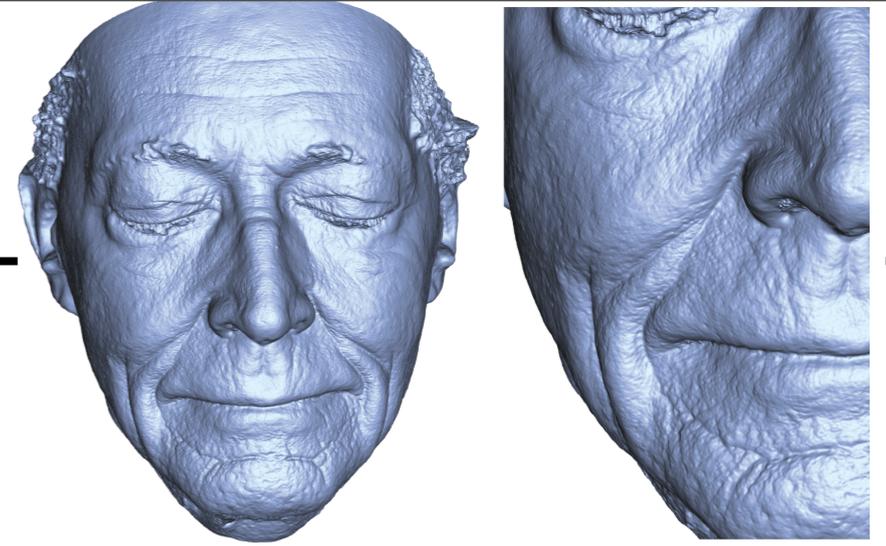
---

- Add value to visual inputs
- Be open minded about
  - input
    - active, passive, 2D, 3D, 4D, etc.
  - added value
    - depth, recognition, flow, etc.
  - way to add
    - math, hack, stat, engineering, etc.

# Acquisition

---

- **Geometry**
  - Also non-Lambertian materials
  - Also messy materials (hair, etc.)
- **Complex deforming geometry**



# Acquisition: Geometry++

- A depth map is often not the final answer

- not even always necessary

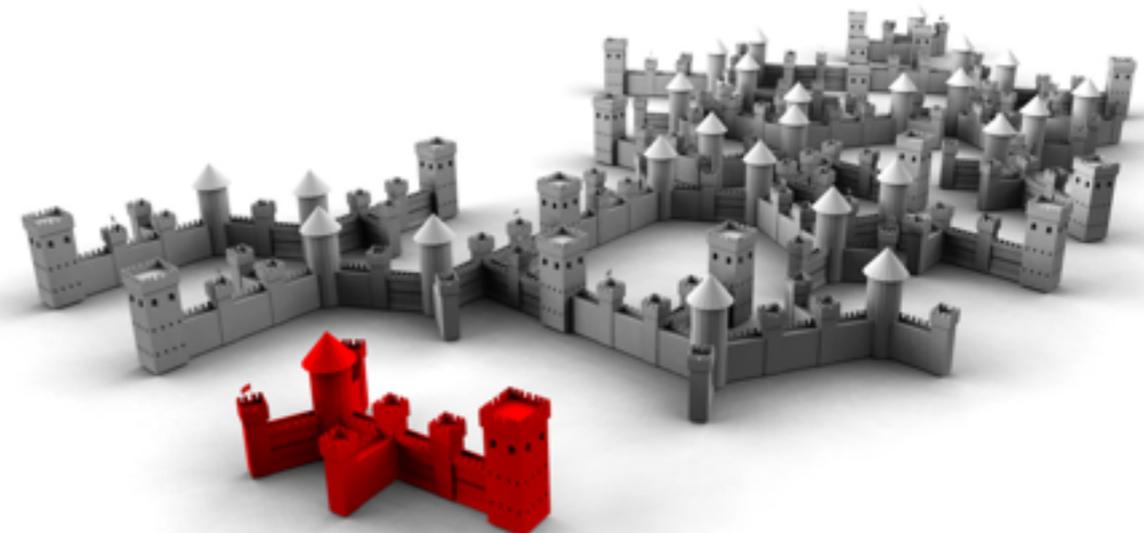
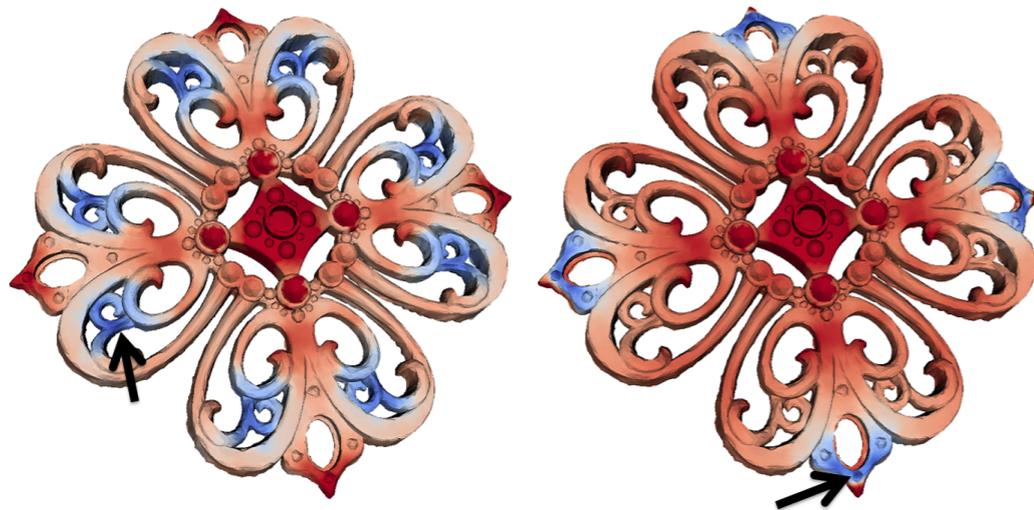
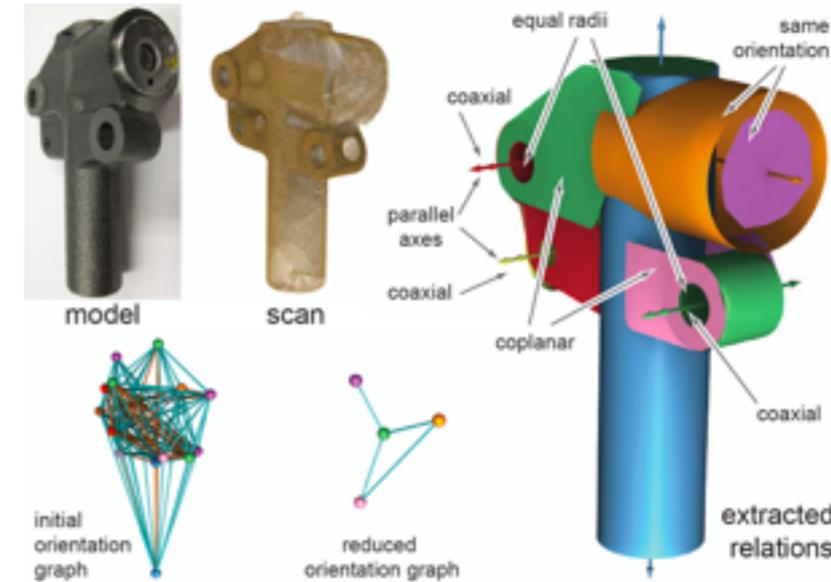
- Segments, parametric patches

- Full volumetric model

- Procedural models

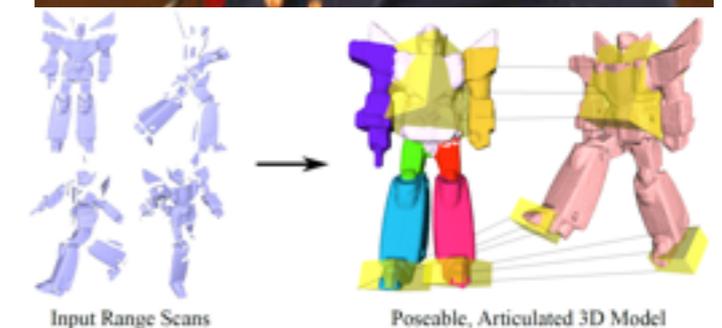
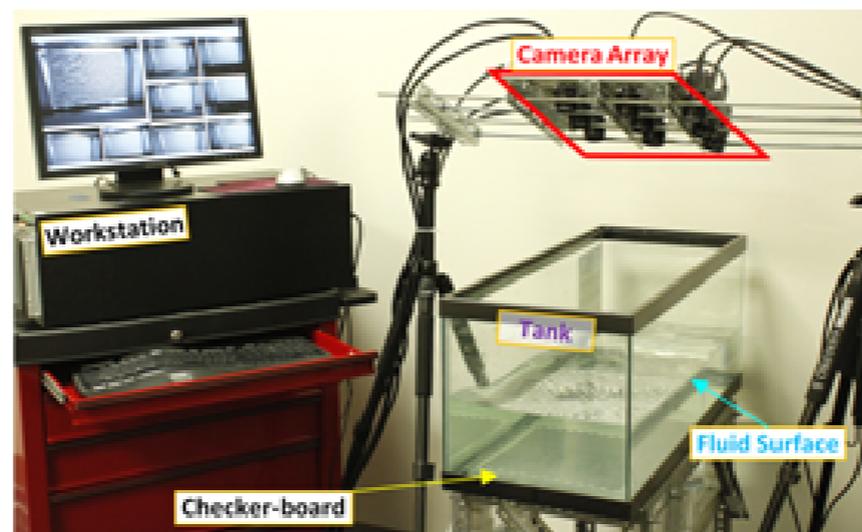
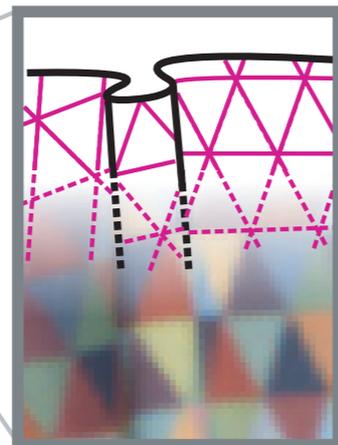
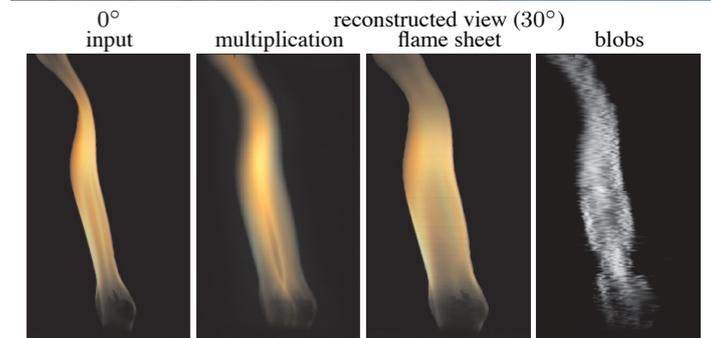
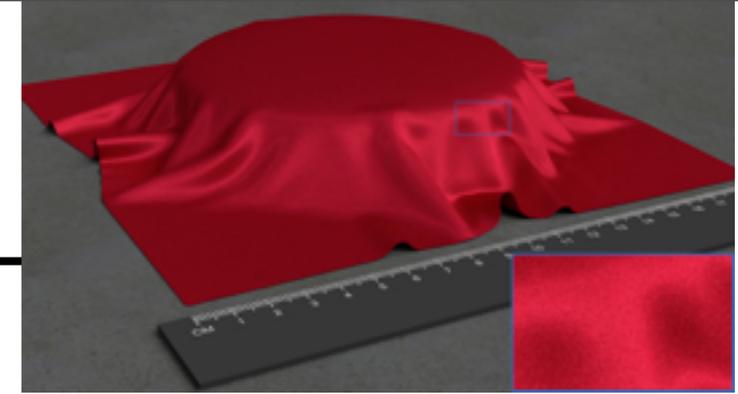
- buildings, trees

- Finding symmetries



# Acquisition

- Material Appearance
- Volumetric media
  - clouds, smoke, fire
- Articulated characters
  - active or markerless mocap
- Complex deforming geometry
  - physical parameters for simulation



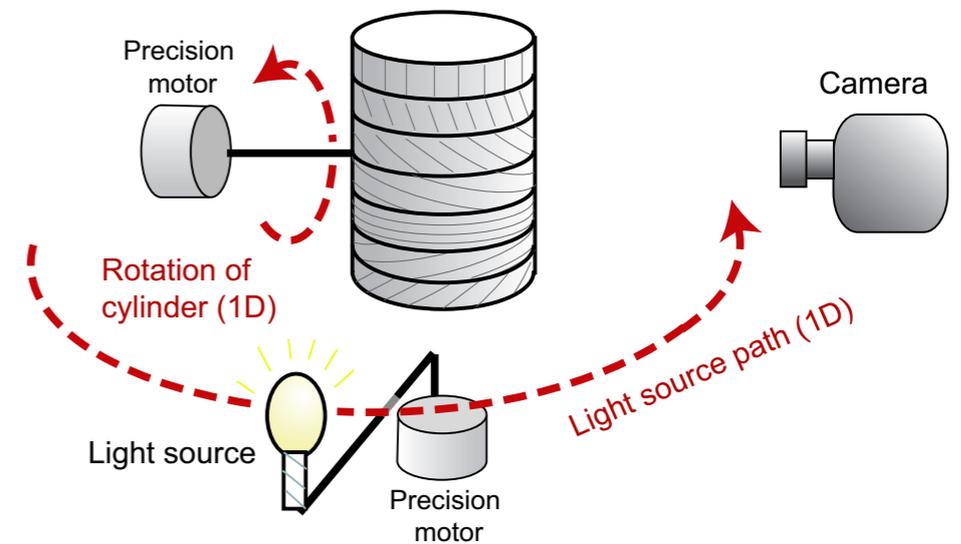
# Active techniques

- They're great, don't overlook them
- Still require a lot of computation



portable configuration

reconstruction



# Content creation

---

- Biggest challenge in graphics
- Can vision help?
- related to capture, 3D acquisition, etc.
- Sketch-based modeling
- Priors for modeling/editing

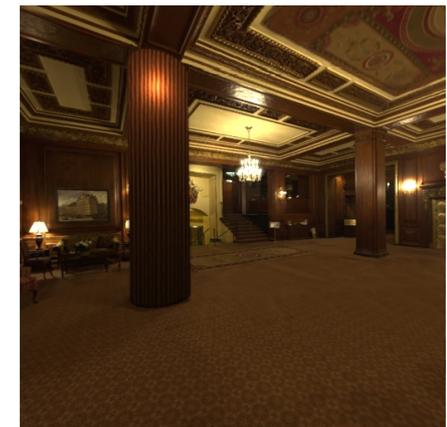


# Analysis-resynthesis for editing

---

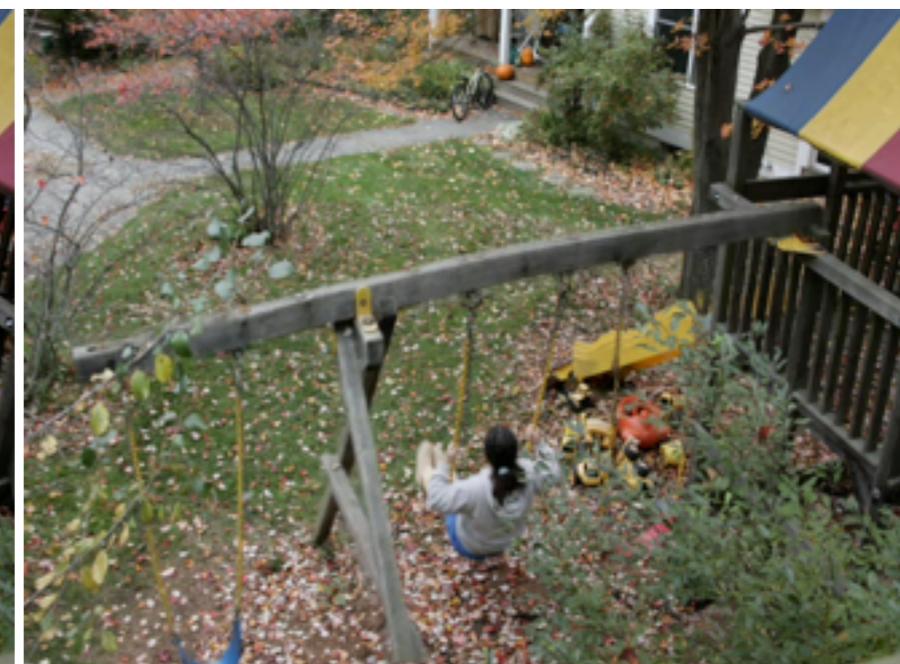
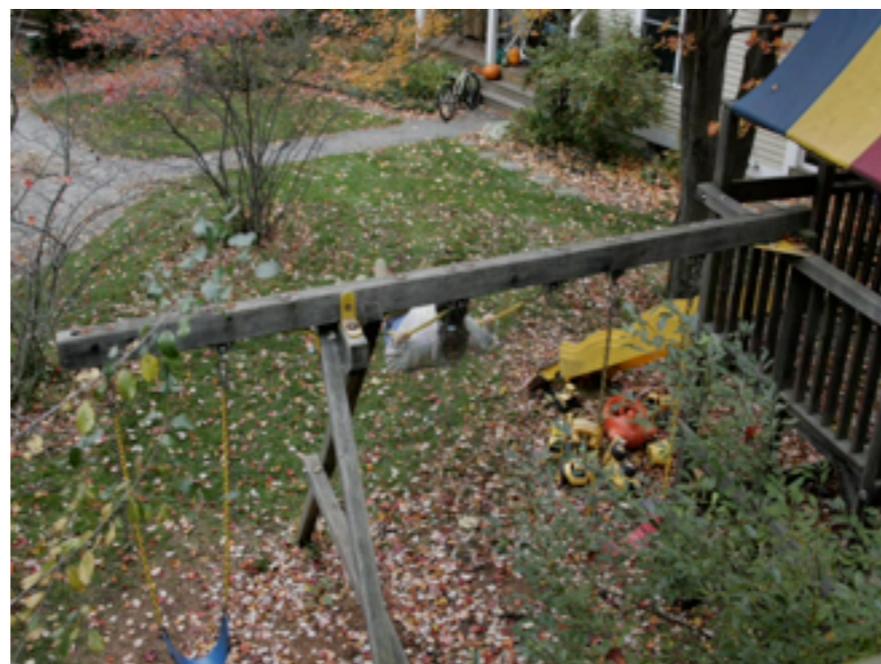
- **Recipe:**

- Analyze input/ decompose
- Edit layers/channels/components  
recombine / rerender



- **Images**

- HDR
- relighting



- **3D data**

- **Motion**

- => low to mid-level analysis is important

# Priors for synthesis

- Can we use natural priors to synthesize motion, images?

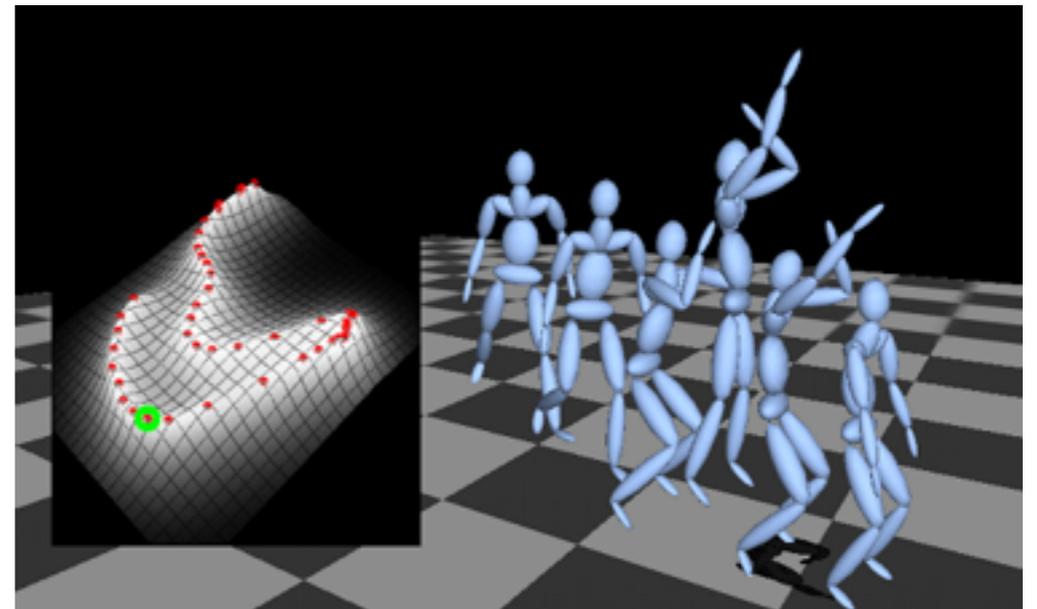
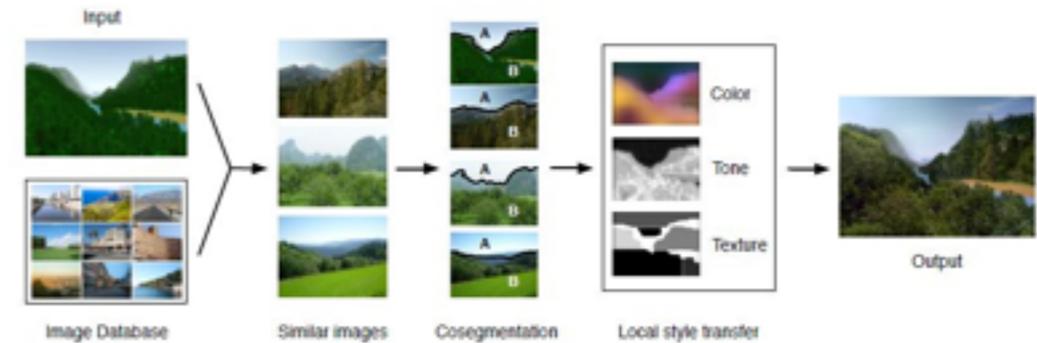
- **Texture synthesis**

- inpainting
- making CG more real

- **Motion**

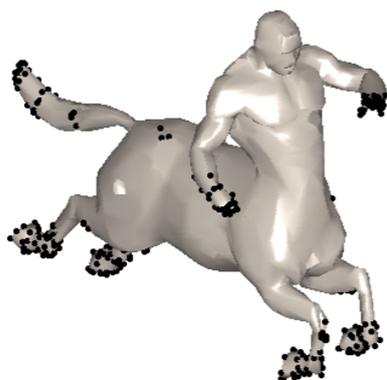
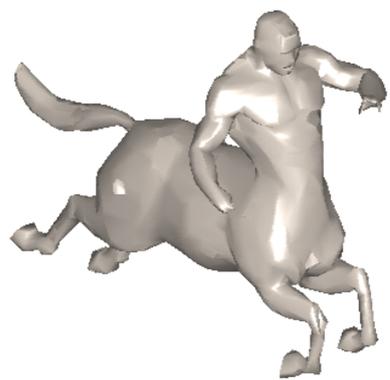
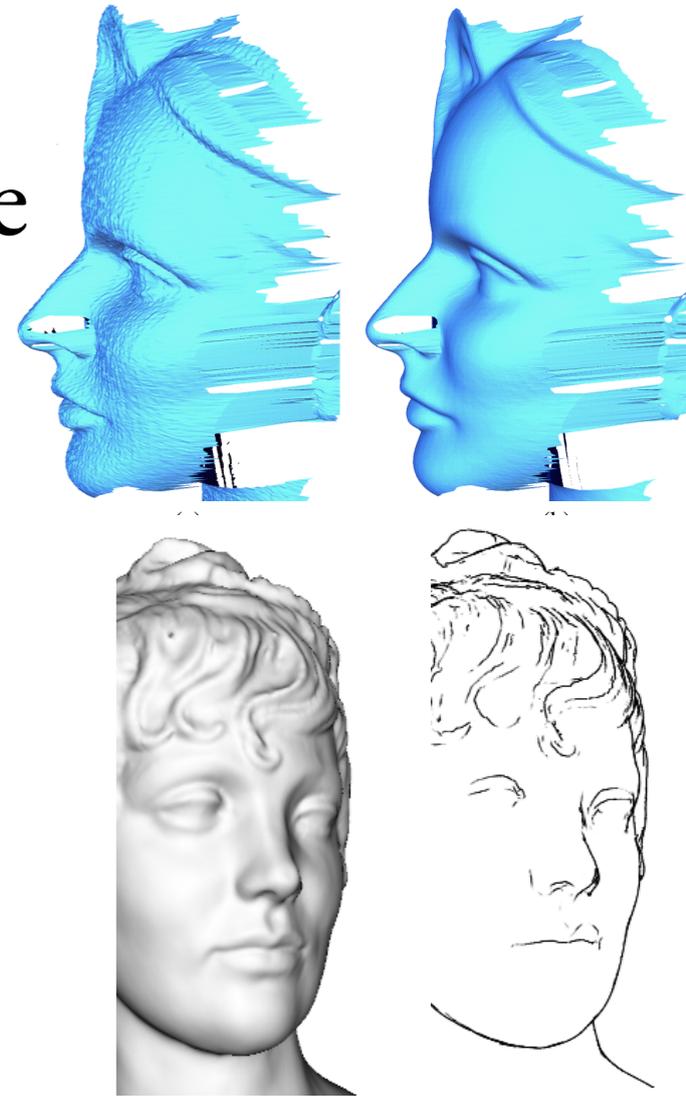
- Used a lot for humans (see Aaron)
- Cloth, hair, fluids?
- Directly from video?

- **Priors used directly in 3D renderer?**



# Applying vision to other signals

- **Geometry**
  - challenge: domain is conflated with range
- **Motion capture data**
- **PDEs, signal processing, edge detection**
- **Features, retrieval, recognition**
- **Priors**



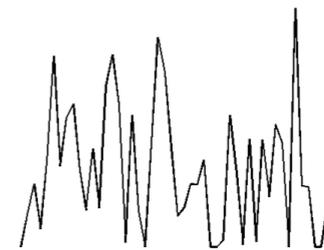
(Feature detection)



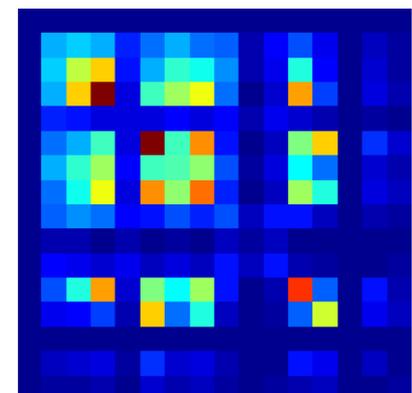
Feature description



Vector quantization



Bag of words

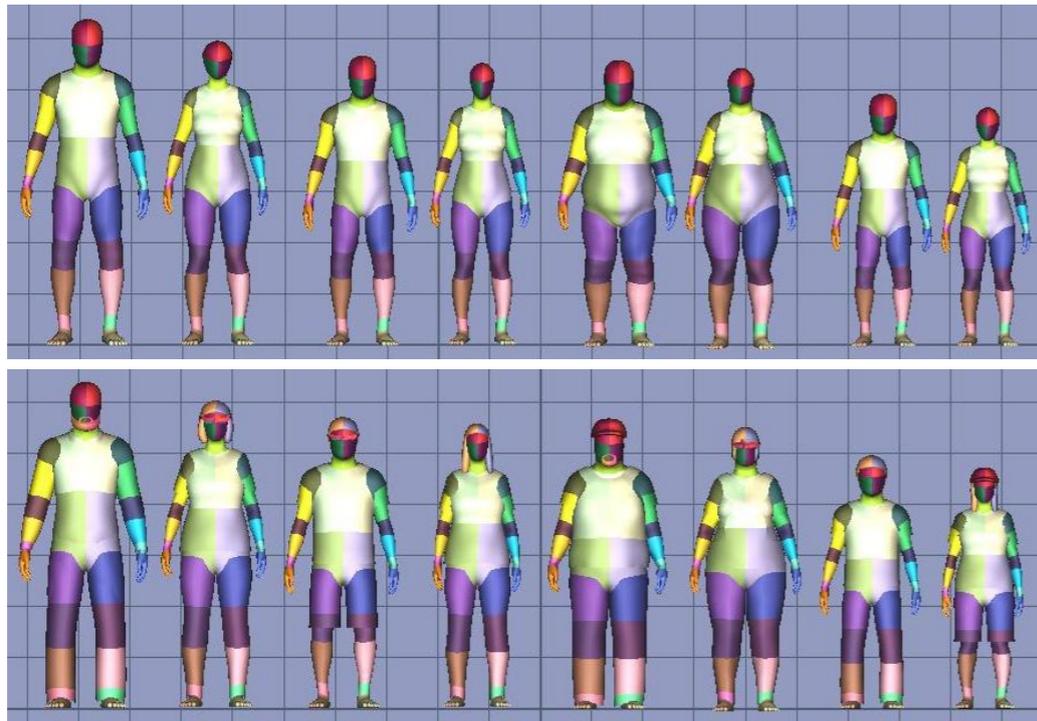


Bag of expressions

Fig. 2. Flow of the ShapeGoogle algorithm.

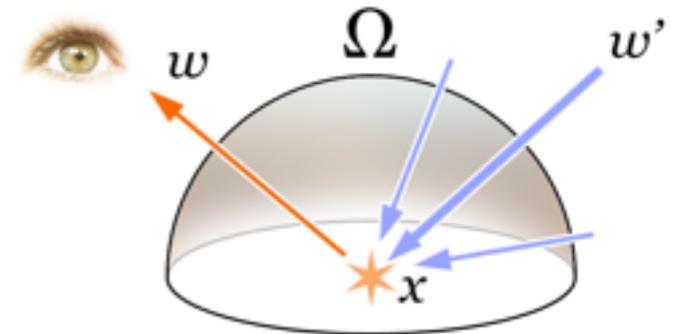
# Using graphics for Vision

- Ground truth, training data



- The image formation equations are the same

- Example: motion blur, depth of field
- Graphics knows more about light transport, material appearance



- Human motion modeling



# Where graphics culture is better

---

- Connection to the real world
- Speed, scalability
- Systems issues, languages, API
- Most important publications easy to identify
  - Fewer siggraph papers than ICCV+CVPR+ECCV
- ... Siggraph parties are better but vision conference locations are better



# Recap: connections

---

- Acquisition
  - Geometry, material appearance, motion
- Active techniques
- Content creation
- Analysis-resynthesis for editing
- Priors for synthesis
  - images, motion
- Applying vision to other signals
  - motion, geometry
- Using graphics for vision

