3D vision

Jason Lawrence (some slides by Szymon Rusinkiewicz)
CS 651, Spring 2007: Computer Vision
3D perception: illusions

[Block and Yuker]
3D perception: illusions

[Block and Yuker]
3D perception: illusions

[Block and Yuker]
3D perception: illusions
3D perception: illusions
3D perception: conclusions

- perspective is assumed
- relative depth ordering
- occlusion is important
- local consistency
3D perception: stereo
3D perception: stereo

- experiments show that absolute depth estimation not very accurate
  - low “relief” judged to be deeper than it is

- relative depth estimation very accurate
  - can judge which object is closer for stereo disparities of a few seconds of arc
3D computer vision

- recover representation of 3D shapes from (primarily) images:

[Debevec et al.]
3D data types

- point data
- volumetric data
- surface data
3D data types: point data

• “point clouds”
• advantage: simplest data type
• disadvantage: no information on adjacency/connectivity
3D data types: volumetric data

- regularly-spaced grid in \((x,y,z)\): “voxels”
- for each grid cell, store
  - occupancy (binary: occupied/empty)
  - density
  - other properties
- popular in medical imaging
  - CAT scans
  - MRI
3D data types: volumetric data

The National Library of Medicine's
Visible Human Project (TM)

Human-Computer Interaction Lab
Univ. of Maryland at College Park

3D data types: volumetric data

- **advantages:**
  - can “see inside” objects
  - uniform sampling: simpler algorithms

- **disadvantages:**
  - lots of data
  - wasteful if only storing surface
  - most “vision” sensors / algorithms return point or surface data
3D data types: surface data

- polyhedral
  - piecewise planar
  - polygons connected together
  - most popular: “triangle meshes”
- smooth
  - higher-order (quadratic, cubic, etc.) curves
  - Bezier patches, splines, NURBS, subdivision surfaces, etc.
3D data types: surface data

• advantages:
  - preserves adjacency information
  - usually returned by vision sensors / algorithms

• disadvantages:
  - difficult to define for translucent objects
  - non-uniform parameterization
  - non-topology-preserving algorithms difficult
3D data types: surface data

- implicit surfaces (cf. parametric)
  - zero set of a 3D function
  - usually regularly sampled (voxel grid)
- advantage: easy to write algorithms that change topology
- disadvantage: wasted space / time
3D data type: implicit surfaces
2.5-D data

color along each ray
depth along each ray

http://www.ixbt.com
2.5-D data

- image: stores an intensity / color along each set of regularly-spaced rays in space
- range images: stores a depth along each of a set of regularly-spaced rays in space
- not a complete 3D description: does not store objects occluded (from some view)
- view-dependent scene description
2.5-D data

- this is what most sensors / algorithms actually return

- advantages:
  - uniform parameterization
  - adjacency / connectivity information

- disadvantages:
  - does not represent entire object
  - view dependent
2.5-D data

- range images
- range surfaces
- depth images
- depth maps
- height fields
- 2.5-D images
- surface profiles
- xyz maps
- ...

...
range acquisition taxonomy

Range acquisition
  - Contact
    - Mechanical (CMM, jointed arm)
    - Inertial (gyroscope, accelerometer)
    - Ultrasonic trackers
    - Magnetic trackers
  - Transmissive
    - Industrial CT
    - Ultrasound
    - MRI
  - Reflective
    - Non-optical
      - Radar
      - Sonar
  - Optical
touch probes

- jointed arms with angular encoders
- return position, orientation of tip

Faro Arm - Faro Technologies, Inc.
range acquisition taxonomy

Optical methods

Passive

- Shape from X:
  - stereo
  - motion
  - shading
  - texture
  - focus
  - defocus

Active

- Active variants of passive methods
  - Stereo w. projected texture
  - Active depth from defocus
  - Photometric stereo

- Time of flight
- Triangulation
optical range acquisition methods

✦ advantages:
  - non-contact
  - safe
  - usually inexpensive
  - usually fast

✦ disadvantages:
  - sensitive to transparency
  - confused by specularity and interreflection
  - texture (helps some methods, hurts others)
stereo

- find feature in one image, search along epipolar line in other image for correspondence:
epipolar geometry
stereo example
stereo

- advantages:
  - passive
  - cheap hardware (2 cameras)
  - intuitive analogue to human vision
- disadvantages:
  - only acquire good data at “features”
  - sparse, relatively noisy data (correspondence is hard)
  - bad around silhouettes
  - confused by non-diffuse surfaces
- variant: incorporate multiple cameras into multibaseline stereo to reduce ambiguity
shape from motion

- "limiting case" of multibaseline stereo
- track a feature in a video sequence
- for $N$ frames and $F$ features
  have $2NF$ knowns and $6N+3F$ unknowns

[Tomasi and Kanade]
shape from motion

✦ advantages:
  - feature tracking easier than correspondence in far-away views
  - mathematically more stable (large baseline)

✦ disadvantages:
  - does not accommodate object motion
  - still problems in areas of low texture, non-diffuse regions, and around silhouettes
shape from shading

\[ \vec{n}(x, y) \quad \text{and} \quad h(x, y) \]

[For syth and Ponce]
shape from shading

- problem: ambiguity

[Pentland]
shape from shading

✦ advantages:
  - single image
  - no correspondences
  - analogue in human vision
✦ disadvantages:
  - mathematically unstable
  - can’t have texture
✦ “photometric stereo” (active method) more practical than passive version
shape from texture

- mathematically similar to shape from shading, but uses distortion of texture
shape from texture

- analogue to human vision
- same disadvantage as shape from shading
shape from focus

- shape from focus: at which focus setting is a given image region sharpest?
- shape from defocus: how out-of-focus is each image region?
- passive versions rarely used
- active depth from defocus can be made practical
active optical methods

✦ advantages:
  - can usually obtain dense data
  - usually much more robust and accurate than passive techniques

✦ disadvantages:
  - introduces light into the scene
  - not motivated by human vision
terminology

- range acquisition, shape acquisition, rangefinding, range scanning, 3D scanning
- alignment, registration
- surface reconstruction, 3D scan merging, scan integration, surface extraction
- 3D model acquisition
related fields

- computer vision
  - passive range sensing
  - rarely construct complete, accurate models
  - application: recognition
- metrology
  - main goal: absolute accuracy
  - high precision, provable errors more important than scanning speed, complete coverage
  - industrial inspection, quality control, etc.
related fields

- humanities
  - want complete and accurate model
  - field acquisition
  - archival, visualization, education, etc.

- computer graphics
  - often want complete model
  - low noise, geometrically consistent model more important than absolute accuracy
  - application: animated CG characters
vision in geometric modeling

Greg Dykstra