Vision Sensing
Multi-View Stereo for Community Photo Collections

Venus de Milo
The Digital Michelangelo Project, Stanford
How to sense 3D very accurately?

Range image
How to sense 3D very accurately?

- Contact
- Mechanical (CMM, jointed arm)

Range acquisition:
- Transmissive
  - Industrial CT
  - MRI
- Reflective
  - Non-optical
    - Radar
    - Ultrasound
    - Sonar
- Optical
optical methods

passive

shape from X:
stereo
motion
shading
texture
focus
defocus

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

classical active

active

time of flight

triangulation
Triangulation

- Depth from ray-plane triangulation:
  - Intersect camera ray with light plane
Example: Laser scanner

Cyberware® face and head scanner

- very accurate < 0.01 mm
- more than 10sec per scan
Example: Laser scanner

Digital Michelangelo Project
http://graphics.stanford.edu/projects/mich/
Shadow scanning

http://www.vision.caltech.edu/bouguetj/ICCV98/
Basic idea

- Calibration issues:
  - where’s the camera wrt. ground plane?
  - where’s the shadow plane?
    - depends on light source position, shadow edge
Two Plane Version

- **Advantages**
  - don’t need to pre-calibrate the light source
  - shadow plane determined from two shadow edges
Estimating shadow lines

$x_{c} = (104, 128) = (x_{c}, y_{c})$

$t = t_0 = 0134$

$x_{top}(t_0)$

ref. top row: $y_{top} = 10$

$x_{bot}(t_0)$

ref. bot. row: $y_{bot} = 230$

$\Delta I(x_c, y_c, t)$

$t_{s}(x_c, y_c) = 133.27$

$\Delta I(x, y_{top}, t_0)$

$x_{top}(t_0) = 118.42$
Shadow scanning in action
Results

accuracy: 0.1mm over 10cm ~ 0.1% error
Textured objects
Scanning with the sun

accuracy: 1mm over 50cm ~ 0.5% error
Scanning with the sun

accuracy: 1cm over 2m
~ 0.5% error
Faster Acquisition?

- Project multiple stripes simultaneously
- Correspondence problem: which stripe is which?

- Common types of patterns:
  - Binary coded light striping
  - Gray/color coded light striping
Binary Coding

Faster:

\[ 2^n - 1 \] stripes in \( n \) images.

Example:

3 binary-encoded patterns which allows the measuring surface to be divided in 8 sub-regions
Binary Coding

- Assign each stripe a unique illumination code over time [Posdamer 82]
Binary Coding

Example: 7 binary patterns proposed by Posdamer & Altschuler

Codeword of this pixel: 1010010 identifies the corresponding pattern stripe

Projected over time
More complex patterns

Works despite complex appearances

Works in real-time and on dynamic scenes

- Need very few images (one or two).
- But needs a more complex correspondence algorithm
Continuum of Triangulation Methods

Single-stripe

Multi-stripe
Multi-frame

Single-frame

Slow, robust

Fast, fragile
**Time-of-flight**

- No baseline, no parallax shadows
- Mechanical alignment is not as critical
- Low depth accuracy
- Single viewpoint capture


Working Volume: 1500mm - Accuracy: 7%
Spatial Resolution: 1x32 - Speed: ??
Comercial products

Canesta
64x64@30hz
Accuracy 1-2cm

Not accurate enough for face modeling, but good enough for layer extraction.
Depth from Defocus
Depth from Defocus
Depth from Defocus

+ Hi resolution and accuracy, real-time
− Customized hardware
− Single view capture?

Working Volume: 300mm - Accuracy: 0.2%
Spatial Resolution: 512x480 - Speed: 30Hz
Capturing and Modeling Appearance

- Mies Courtyard House with Curved Elements
- Close-up of green leaves
- Close-up of underwater scene with fish
Capture Face Appearance
Image-Based Rendering / Recognition

Schechner et. al. *Multiplexed Illum*
Light Stage Data

Original Resolution: 64\times32

Lighting through image recombination: Haeberli ‘92, Nimeroff ‘94, Wong ‘97
Shape Recovery
BRDF
Material Recognition
Human Vision
Rendering
Object / Face Recognition