Space-Time Stereo

James Davis – Honda Research Institute
Ravi Ramamoorthi – Columbia University
Szymon Rusinkiewicz – Princeton University
Diego Nehab – Honda & Princeton

BAVM - UCSC
June 2003
Goal: Acquire range images

- **Applications**
  - Object reconstruction
  - Robotic navigation
  - Face tracking
  - Motion capture
Active vs. passive classification

Active

Passive

Optical

Triangulation

Active depth from defocus

Active stereo

Interferometry

Moiré

Holography

Shape from shading

Shape from silhouette

Depth from focus

Stereo
Spatial vs. temporal classification

- Triangulation
  - Spatial
    - Stereo
    - Laser scanning
  - Temporal
    - Structured light
    - …
Triangulation

1. Find corresponding pixels
2. Triangulate to get depth

Ambiguity in determining correspondence is the main challenge
Traditional stereo
Works poorly when scene lacks spatial texture
How can we use the temporal information?
Temporal matching vector
Coded structured light

Compare temporal pattern to known set of projected patterns
Coded structured light

Compare temporal matching vector to other camera
Laser scanning

Matching vector
Spatial vs. temporal classification

Spacetime Stereo

Temporal only processing
- Temporal laserscanner
  - ASP87
  - KGC91
  - CL95
  - BP98
  - PADSS98
- Structured light
  - ISM84
  - BMS98

Combined processing
- RHL02
- ZCS02

Spatial only processing
- Spatial laserscanner
  - JKC85
  - MJ87
  - SJM91
  - BFBBPS98
  - DCC01
- Traditional stereo
  - KWZK95
  - KRS96
  - CHCW97
  - SS02

Temporal extent of processing
- Many frames
- One frame

Spatial extent of processing
- One pixel
  - Many pixels
Static scene, variable illumination
Spatial vs. temporal stereo
Errors due to spatial occlusion
Errors due to temporal occlusion
Which space-time window size should we use?

Error as a function of spatio-temporal window size

- 1x1 spatial window
- 5x5 spatial window
- 11x11 spatial window
- 19x19 spatial window
Temporal laser scanning
Benefit of two real cameras

Real camera

\[
\text{Camera}_{\text{real}} = (\text{Light}+\text{Ambient}) \cdot \text{Surface\_reflectance}
\]

Virtual camera viewpoint

\[
\text{Camera}_{\text{virtual}} = (\text{Light})
\]

\[
\text{Camera}_{\text{second\_real}} = (\text{Light}+\text{Ambient}) \cdot \text{Surface\_reflectance}
\]
Errors due to spatial distortion
Errors due to temporal distortion