Light Control and 3D Video

Markus Gross
Computer Graphics Laboratory
ETH Zürich
&
Disney Research Zurich (DRZ)

graphics.ethz.ch
Example - LiberoVision
Purpose of the Talk

- High level overview of research that led to the development of such technology
- Mostly focused on controlled environments
- Detail: graphics.ethz.ch/3Dvideo
Overview

- Introduction and Motivation
- #1 blue-c: light control
- Towards modular camera/projector systems
- #2 3D video processing
- Outlook
#1 blue-c

ETH Computer Center

ETH Hönggerberg
Idea

ETH Hönggerberg

ETH Center
System Overview

Gross et al. Siggraph 2003
Camera Layout

• 16 Firewire cameras
“Active” Projection Panels

- Shuttered glass panels (based on LC-technology)
- Screen is opaque for back projection
- Screen is transparent for video acquisition

Shutter panels (62.5 Hz)
Lighting and Projection

- Active lighting using pulsed LED arrays
- Active stereo projection using conventional LCD projectors with additional LC-shutters
Active Lighting 10’000 LEDs
Timing - Animation
3D Mirror
Software Components

- 3D acquisition and reconstruction
  - Calibration
  - Segmentation
  - 3D reconstruction
  - 3D video streaming, compositing, and rendering
- Networking and communication (BCL)
  - Event synchronization
  - Multimedia streaming (TAO/ACE)
- Application programming interface (API)
  - Distributed scene graph
  - Integration of 3D video inlays and audio
3D Video

- Extension of 2D video into the third dimension
- Appearance & geometry of real-world scenes
- Spatio-temporal representation
  - Depth compositing
  - Object replacement
  - Animation
- Novel display features
3D Video Pipeline

Application

3D scene

Rendering

image

3D Video Object Rendering

Network Streaming

point stream

3D Reconstruction

points with color and normal

Silhouette Extraction

silhouette

Background Subtraction

foreground image

color
3D Reconstruction

- Shape-from-Silhouettes method
  - Based on visual hull
  - Laurentini, Matusik
  - Real-time algorithm
  - Precise camera calibration
  - Photometric calibration

- Point-based representation of the user
  - 3D points with color and normal
  - Exploit spatio-temporal coherence
Dynamic Point Samples

- 3D Video primitive
- Generalizes pixels to 3D points
3D Video Stream

Waschbüsch et al. PG 2005
Differential 3D Video Stream

green: new
red: expired
blue: color changed
white: color unchanged
black: background

Insert
Delete
UpdateCol
UpdatePos possibly
3D Video
Close-Up
Video Recording Pipeline

Lamboray et al. ICIP 2004

Würmlin et al. VMV 2005
Results
Findings and Limitations

• blue-c was a major engineering milestone
  – Expertise from many different areas
• Experience in realtime 3D video
• Design monolithic, not scalable
• Proprietary hardware needed
• Environment makes it difficult to acquire high-quality video
• Limitations in processing speed and power
Concept of „Presence Brick“

- Modular setup
- Combination of DLP projector, cameras, PC, and MC
- Off-the-shelf components
- Collaboration with Henry Fuchs
- Highly scalable
Concept: “Presence Brick”

- Stereo cameras
- Texture camera
- Structured light projector
System Overview

3D Video Brick
- Acquisition
  - Projector
  - Stereo Cams
  - Texture Cam
- Depth Extraction
  - Stereo Matching
  - Discontinuity Optimization

View-independent Scene Representation
- Merging
  - Photo Consistency
  - Outlier Removal
  - Union
- Editing / Rendering
  - Compositing
  - Spatio-Temp. Effects
  - EWA Volume Splatting

Multiple Additional 3D Video Bricks...
Prototype Installation

• 3 Bricks
  – Scene coverage: 70° horizontally, 30° vertically
  – Synchronized via micro controller unit (MCU)
Texture & Depth

- Structured light invisible for texture cameras
- Alternating projection of inverse patterns
- Synchronize cameras shutters
Example – One Brick

Grayscale camera  Color camera  Grayscale camera

Textures

Stereo
Results – All Bricks
Enforcing Photo Consistency

Merging **without** photo consistency enforcement

Merging **with** photo consistency enforcement
Input: Three 3D Video Bricks

Baseline between bricks:
- approx. 2.5 m
- approx. 35°
Results – Final 3D Video
Interactive 3D Video Editing

- Video Hypervolume
  - Similar to video cube
  - Irregular 4D data

Waschbüsch et al. PG 2005
Editing Pipeline

• Graph cut algorithm for selection
Graph Construction
Results
3D Video Billboard Cloud

- One billboard from each input viewpoint
- Planar geometric proxy
- Displacement map
- Blending of multiple maps
- Bilateral disparity filtering

Waschbüsch et al. EG 2007
Concept

multi-view video

depth extraction

3D video billboard cloud

3D video

filtering

Computer Graphics Laboratory ETH Zurich
Input

Color maps

Depth maps

Alpha mattes
Filtering Geometry

- Input displacements are noisy
- Regular sampling: efficient signal processing tools
- Spatio-temporal bilateral filter for smoothing the geometry
Bilateral Disparity Filter

\[ \tilde{d}(x) = \frac{1}{s} \cdot \iiint d(\xi) \cdot c(\xi, x) \cdot s(d(\xi), d(x)) \cdot \delta \xi \]

Normalization term

input disparities

range filter kernel

output disparities

over displacement map & over time

domain filter kernel
Filter Kernels

• Domain filter kernel
  – Cubic B-spline weighted by alpha matte
    \[ c(\xi, x) = \alpha(\xi) \cdot B(\xi - x) \]
  – Low-pass filter
  – Lower weight at more inaccurate boundaries

• Range filter kernel
  – Step function
  – Preserves depth discontinuities
Filtering Over Multiple Views

• Problem
  – Billboards from multiple views are filtered independently
  – Overall geometry may diverge

• Solution
  – Filter over all billboards at the same time
Implementation

• For filtering view \( i \)
  – Warp domain filter kernel to all views \( j \)
  – Convolve all views \( j \) with warped kernel
  – Accumulate in view \( i \), using range filter

Splatting \( \rightarrow \) Graphics Hardware
Filtering - Comparison

- no disparities, only planes
- raw disparities
- views filtered independently
- views filtered together
View-Dependent Rendering

• Unstructured Lumigraph (UL) framework [Buehler01]
• Render consistent depths
  – Render views into separate depth buffers
  – Resolve occlusions via fuzzy z-buffer
  – Blend remaining depths using UL
• Render consistent colors
  – Projective texturing
  – Blend colors using UL
Rendering Comparison

color blending only  color & depth blending
Results
Results
Conclusions

• You will have to live with the imperfections of your input data!
• Much of this can be corrected by clever rendering and display
• Major Research Challenges:
  – Theory of plenoptic sampling
  – Spatio-temporal representations (MLS)
  – TOF sensors for controlled environments/studio
  – Compression
  – Editing and art direction
Questions?
Example - LiberoVision
Scene Sampling

- Pixel coordinates, $u_L, u_R$
- Disparity, $d = u_L - u_R$
- Depth, $z \approx \frac{1}{d}$
- Camera coordinates:

\[
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix} = P^{-1} \cdot z \cdot \begin{pmatrix}
u \\
v \\
1
\end{pmatrix}
\]
Sampling Spaces

disparity space  ray space  camera space
Error Model

- Uniform error model
- Uncertainty of disparity due to pixel sampling
Error Spaces

disparity space
ray space
camera space