Depth-Based Visual Signal Processing

C.-C. Jay Kuo
University of Southern California
Outline

- Introduction
- How to get the depth information
- How to use the depth information
- Future research directions
History of 3D Visual Experience

Earliest:

1915 → First 3D movie
1928 → First stereoscopic 3D television
1935 → First 3D color movie

Latest:

2009 → First 3D video on youtube
2010 → First 3D newspaper
2011 → Fastest selling consumer electronics device - Kinect
3D movies:
3D TV and movies:
Recent Developments: 2D-to-3D Conversion

The titanic:

Original 2D + Depth map

Step 1: Restoration to 4K resolution

Step 2: 3D conversion:
Time: 60 weeks
Cost: $18 millions
Outline

- Introduction
- How to get the depth information
- How to use the depth information
- Future research directions
Acquiring Depth Information

- Depth from stereo camera
  - Stereo matching
- Depth from depth camera
  - Depth acquisition directly
- Depth from single camera
  - Shape from X (X: Shading, Texture, Photometric Stereo, etc.)
  - Other shape inference methods -> focus of my talk
Example 2: Area Ratio
Example 3: Extended Vanish Point
Depth from In-Focus Detection (1)

Blur Estimation:

Requirement:
images taken from a fixed camera position and object position but using different focal settings

Depth from In-Focus Detection (2)

Which is in front?
Depth from In-Focus Detection (3)

Which is in front?

![Blue dragonfly on a branch](image-url)
Depth from In-Focus Detection (4)

Which is in front?
Focus:

By estimating the blurriness at edges, we recognize the in-focus region.
Focus:

In focus degree estimation

Requirement: images taken from a fixed camera position and object position but using different focal settings

**Focus:**

- Original
- Foreground extraction
- Initial In-focus Estimation
- After post-processing

**Depth from In-Focus Detection (7)**
Which is in front?
Depth from Haziness (2)

Which is in front?
**Haziness:**

- Relationship between atmosphere transmission \( t(x) \) and depth \( d(x) \)

\[
t(x) = e^{-\beta d(x)}
\]

- the radiance factor

- constant
Parallel lines appear to converge with distance, eventually reaching a vanishing point at the horizon.
Background Depth Modeling (2)

Which is in front?
Background Depth Modeling (3)

Which is in front?
Which is in front?
Background structures:

By modeling the scene to above structures, we can infer depth for background.
Introduction
- How to get the depth information
- How to use the depth information
- Future research directions
Depth-Assisted Navigation (1)

Walker alert
Cars without driver:

http://www.youtube.com/watch?v=JmpVhBFdKUg
Depth assisted tracking:

Solving occlusion problems!

Who is occluded? How much?
Depth assisted tracking:

Occlusion cases:

- Non-occlusion: there is no overlaid part
- Occlusion occurs: (partial and severe) there is overlaid part

Examples of tracking under non-occlusion

Examples of tracking under partial occlusion

Tracking under partial and severe occlusion

“Depth Assisted Visual Tracking”, Y. Ma, S. Worrall, A. M. Kondoz, Centre for communication systems research, University of Surrey, Guildford, Surrey, United Kingdom
The Kinect:

http://www.youtube.com/watch?v=T_QLguHvACs
Depth-assisted motion estimation:

http://www.youtube.com/watch?v=Mf44bWQr3jc&feature=results_video&playnext=1&list=PL0C6C641D376DDFF
Depth as Visual Cue: 3D and Virtual Reality (1)

Stanford University
Depth as Visual Cue: 3D and Virtual Reality (2)

Make3D: convert your still image into 3D model

“Learning 3D Scene Structure from a Single Still Image”, ICCV 2009
Color-based segmentation: Can color tell us everything about segments?

Example 1: mean shifts

Example of changing scores for different segmentation granularities: (a) Original image, (b)-(g) mean shift segmentations using scale bandwidth ($h_s$) 7 and color bandwidths ($h_r$) 3, 7, 11, 15, 19 and 23 respectively.

Color-based segmentation:

Example 2: graph-based segmentation

Example of changing scores for different parameters using efficient graph-based segmentation: (a) Original image, (b)-(d) efficient graph-based segmentations using scale bandwidth ($h_s$) 7, color bandwidth ($h_r$) 7 and $k$ values 5, 25, 125 respectively.

“A Comparison of Image Segmentation Algorithms”, Caroline Pantofaru Martial Hebert
CMU-RI-TR-05-40
Color-based segmentation:

Example 3: Hybrid solution (mean shift+efficient graph)

Example of changing scores for different parameters using a hybrid segmentation algorithm which first performs mean shift filtering and then efficient graph-based segmentation: (a) Original image, (b)-(g) segmentations using scale bandwidth ($h_s$) 7, and color bandwidth ($h_r$) and $k$ value combinations (3,5), (3,25), (3,125), (15,5), (15,25), (15,125) respectively.

Depth-assisted segmentation:

Color + depth = better segmentation

(a) Segmentation with color only, (b) depth only, (c) color + depth

“DEPTH ASSISTED OBJECT SEGMENTATION IN MULTI-VIEW VIDEO”, Cevahir Çığla and A.Aydın Alatan, Department of Electrical and Electronics Engineering, M.E.T.U, Turkey
Overview on 2D-to-3D Video Conversion

- Semi-automatic conversion
- Full-automatic conversion
Semi-automatic Conversion (1)

- **Typical 2D to 3D conversion pipeline**
  - Artists manually draw the depth map for each key frame in a commercial movie
  - Computer-aided software helps artists propagate the key frame depth map to all the other frames
  - Use the Depth Image based Rendering (DIBR) technique to create left and right views of one frame
  - Use the hole filling technique to create image patches in the hole area
An Example:

- Artist manually Draw
- DI BR + Hole Filling
Semi-automatic Conversion (3)

- **Challenges**
  - **Depth Map Propagation:**
    - Expansive to recruit artist to draw all depth maps
    - Temporal consistency issue
  - **DIBR + Hole Filling:**
    - Director requires different artistic effect, e.g., emphasize certain parts of an object to yield the “pop-out” effect
    - Fill in the missing patch in the rendered image, while preserving the texture & structure
3D image and video automatic conversion:

Depth map generation

Input source (image or video) → automatically → Depth map → What is the rules?

- Focuses;
- Haziness;
- Background structures;
Outline

- Introduction
- How to get the depth information
- How to use the depth information
- Future research directions
Future Research Directions (1)

- Automatic fine-grained depth map for foreground objects (e.g. human faces, translucent regions)

Future Research Directions (2)

- **Depth-assisted image/video content retrieval**
  
  Content-Based Image Retrieval (CBIR) – Color, BoW, Shape, Texture, etc.

  Original image + depth image

  Better segmentation, scene classification, 3D content recognition

  Better understanding of the scene content and category
Future Research Directions (3)

- Visual Attention and Quality Analysis

  - Evaluate depth map quality by stereo matching:
    - Synthesized view
    - Matching result

  - Learning based 3D video quality scoring system.

    Titanic 3D → Scoring system → 9.5!

Future Research Directions (4)

- Human-Centric Image/Video Segmentation and Analysis
Narrow Down Semantic Gap

- Human-centric versus pixel-based approaches

Multiple levels of segmentation hierarchy. Pixel-level over-segmentation on top-right. Larger region granularity in bottom row, with bottom-right having largest regions. Original frame from South Pacific, ©2009 BBC.

“Efficient Hierarchical Graph-Based Video Segmentation”, Matthias Grundmann, Vivek Kwatra, Mei Han, Irfan Essa, Georgia Institute of Technology, Atlanta, GA, USA, Google Research, Mountain View, CA, USA
Future Research Directions (5)

• **True 3D capturing, coding and reconstruction**
  
  • Example:
    - ‘Bullet Time’ – a visual effect that allows the audience’s point-of-view to move around the scene at a normal speed while the action unfolding is played out in slow motion

  • Capturing
    - Dense camera arrays, spatial-temporal sampling
    - RGB cameras + depth cameras
    - Optimal setting

  • Coding
    - Goes beyond MVC and 3DVC

  • Reconstruction
    - Rendering for virtual views