What's wrong with video? – Frédo Durand, MIT CSAIL

Disclaimer

• This is a non-technical talk
• I have no credential or authority to talk about the following subjects
• A lot of it is common-sense/well-known
• My goal is to point out research challenges to create compelling videos
• And chat about what makes it compelling

Traditional Photo & Video

• Scene preparation
  – Make up, lighting, viewpoint
• Capture
  – Optics + film
• Post production
  – Dodge & burn, movie editing
  – some tedious, some creative
• Viewing
  – Flip through photos, passive in front of TV

Computational Photo & Video

• Scene preparation
  – Get rid of it!
• Capture
  – Data-rich, active
• Post production
  – Reconstruct image
  – Automate tedious post-processing
  – Facilitate creative choices
• Viewing
  – Additional dimensions (autostereoscopic, motion, HDR)
  – Interaction

Recipe for computational photography

• If you think photorealistic graphics is difficult
• If you think vision is too hard

Recipe for computational photography

• If you think photorealistic graphics is difficult
• If you think vision is too hard

• Do computational photography:
  – Do a bad job at solving an ill-posed vision problem
  – And cheat: record more information, have user in the loop
  – Modify things a little bit
  – Re-synthesize
  – And voila, a photorealistic picture
Image-based editing

• With B. Mok Oh, Max Chen and Julie Dorsey, [Siggraph 2001]
• See www.mok3.com
• 3D model from single photograph
• 3D not accurate, but looks great
• Lots of user input

Input image New viewpoint Relighting

Defocus Matting

• With Morgan McGuire, Wojciech Matusik, Hanspeter Pfister, John “Spike” Hughes
• Data-rich: use 3 streams with different focus

Tone mapping

With Julie Dorsey
Three wrongs make one right
• Analyze image
  – Intrinsic image: albedo & illumination
  – Simple bilateral filter
• Modify
  – In our case, reduce contrast of large-scale (illumination)
• Recombine
  – Get final image

Flash Photography

• Elmar Eisemann and Fredo Durand, [Siggraph 2004]
• Available light is too weak, image is blurry/noisy

Flash Photography

• Available light is too weak, image is blurry/noisy
• Flash photos look harsh, ambiance is not nice
Flash Photography

- Available light is too weak, image is blurry/noisy
- Flash photos look harsh, ambiance is not nice
- Our work combines the two to get the best of both

Recipe for computational photography

- Do a bad job at solving an ill-posed vision problem
  - But be non-committal about it, no hard decision
- Modify things a little bit
  - But not too much, be conservative
- Re-synthesize
  - And voila, you get a photorealistic picture

Other example: colorization

- Colorization Using Optimization
  Anat Levin, Dani Lischinski, Yair Weiss [Siggraph 2004]
- Optimization with similarity on pixel intensity

The “film look”

- What makes Hollywood production different from a home video
  - The story is better!
  - Hum, wait, there must be a better explanation!
- It is not only the medium:
  a film transferred to DVD still has the “film look”
- What makes a still picture different from video?

Why is home video usually bad?

People we should fire:
- Director of Photography (image quality)
- Director, script author (story)
- Editor
**Response curve, dynamic range**

- Video sensors have poor dynamic range

![Response Curve](image)

**Traditional craft: dodging and burning**

- *Clearing Winter Storm* by Ansel Adams

![Clearing Winter Storm](image)

**Tone mapping in comp. photo**

- Much success
- No one question is control (work with Philip Guo)
  - Use bilateral for selection
- Open problem: dodging/burning for video

![Tone Mapping](image)

**Low contrast is also an advantage**

- W. Eugene Smith photo of Albert Schweitzer
- 5 days to print!
- Things can be related because the intensity is more similar
- Balance, composition

![Low Contrast](image)

**Lighting**

- E.g. 3-point lighting
  - Reduce dynamic range
  - Emphasize silhouettes → 3D cues
- Goals of lighting:
  - Manage dynamic range
  - Reveal shape, layout, material
  - Tell story

![Lighting](image)

**Lighting research challenges**

- Get rid of scene preparation:
  - Lighting as a post-process
- Dynamic range can be handled by tone mapping
- Reveal shape, layout, material
- Tell story
- Note that cinematic lighting is often “non-realistic” → relighting does not need to be physical

![Lighting Research Challenges](image)
**Make Up**

- Can we do post-processing make up?
  - Color response, wrinkle removal
  - Face tracking, template-based processing

**Depth of field**

- Two types of photographers
  - Hate/love depth of field
- Computational imaging
  - Reduce dof (George)
  - Increase it (Berthold)

**Depth of field**

- It’s all about the size of the lens aperture

**Sensor size**

- [http://www.mediachance.com/dvdlab/dof/index.htm](http://www.mediachance.com/dvdlab/dof/index.htm)
- Use two optical systems

**The coolest depth of field solution**

- [http://www.mediachance.com/dvdlab/dof/index.htm](http://www.mediachance.com/dvdlab/dof/index.htm)
General issue: (de)-emphasis

- Crucial pictorial tools
  - E.g. Lighting, depth of field, Gaussian blur
- New tool: Emphasis and texture variation (with Sara Su & Maneesh Agrawala)
  - Modify texture variation to “hide” distracting background

General pictorial issues & techniques

- Lighting, dodge/burn, filtering, make up, touch up can be used for the same effects
- They solve the same depiction challenge
- What are the general (medium-independent) depiction challenges
- What are the general pictorial techniques

Holistic notion of “look”

With Soonmin Bae & Sylvain Paris
- Can we characterize “look”?
- Can we transfer look?
- Coarse-grain characterization
  - Intensity & color histogram
  - Frequency content
  - Etc.

Before

After transfer of Adams “look”
Evil fluctuations in video

- Zoom
- Auto-exposure
- Idea for the zoom issue:
  - Zooming is like cropping
  - Why not capture gigapixel video and post-crop?
    - Well, maybe file size and processing time
  - Do only when user zooms
    - And as a post-process, go directly to final framing
  - Or use superresolution

Story & story telling in video

- Hard to improvise, anticipate
  - Go back in time!
  - At least have a buffer.
- Editing is usually missing
  - Discrete, pre-digested editing helps
  - Virtual cinematographer?
- Pictorial techniques are needed for selection & (de)emphasis
  - And often there is just no story, just a sense of place/moment
    - Free exploration

UI

- Current video editing requires much decision
- Deal with continuous variable (time)

Video equivalent of photo browser?

- Discrete interface is key:
  - Keep or delete
- Sort vs. search

Less is more

- Suspension of disbelief
  - The more realistic, the more you notice the flaws
- Beholder’s imagination
- Free exploration vs. guided medium

- Computational imaging often reveals more
- My suggestion: use it to hide more

System & performance issues

- Video & computational imaging raise system challenges
  - Large-data management
  - High computation cost
- It is important to appreciate & tackle these issues
- See the example of other fields
  - Real-time rendering
  - Out of core mesh processing
- Side note: writing a good system paper is hard
  - To learn how to write a bad one, see my web page
Summary: How can we help?

People whose job we should make easier:
• Director of Photography (image quality)
• Editor

Use automation to replace tedious component and complex scene preparation
Provide creative or expressive knobs
Data-rich imaging, mid-level vision