



Aalto University
School of Electrical
Engineering

Multi-view stereo reconstruction for computer graphics content capture

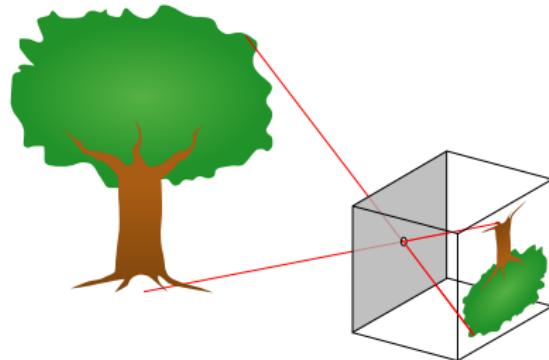
Konsta Hölttä

Master's thesis presentation

konsta.holttta@aalto.fi

Nov 7, 2014

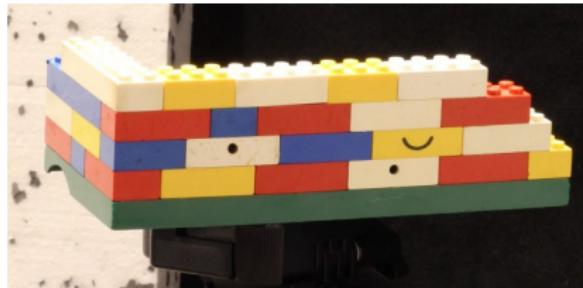
Multi-view stereo reconstruction for computer graphics content capture



- ▶ Interest in realistic content in computer graphics
- ▶ Computer graphics synthesizes images from geometric data
- ▶ Computer vision analyzes geometry from images
- ▶ This work: Scanning real life objects using several digital cameras simultaneously

Contents

- ▶ Intro, motivation
- ▶ Imaging fundamentals
- ▶ 3D reconstruction
- ▶ Scanning rig implementation
- ▶ Results
- ▶ Conclusion



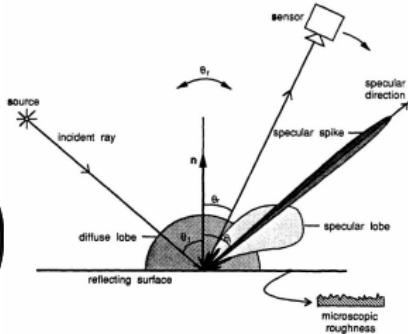
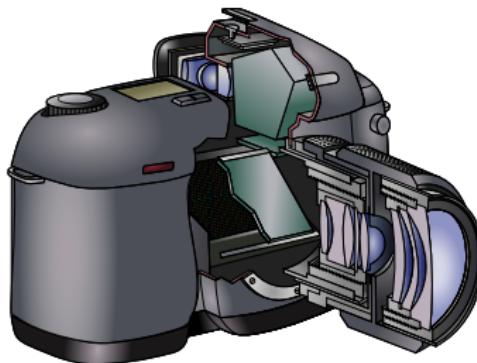
Motivation



Digital Ira (NVIDIA), CGI

- ▶ The processing power of GPUs big and increases
- ▶ Able to process massive level of detail
- ▶ Creating content manually tedious, difficult to get right
- ▶ Digital cameras are really good today
- ▶ ⇒ Scan the geometry and color using digital cameras

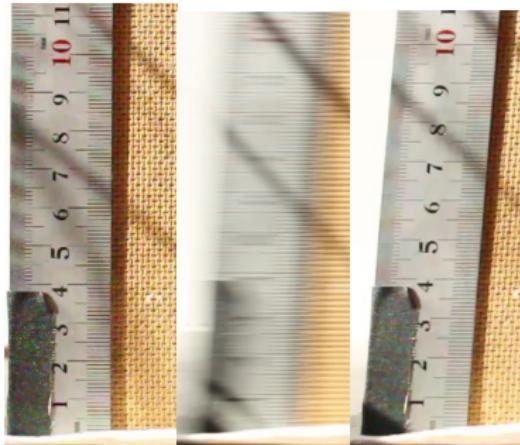
Imaging



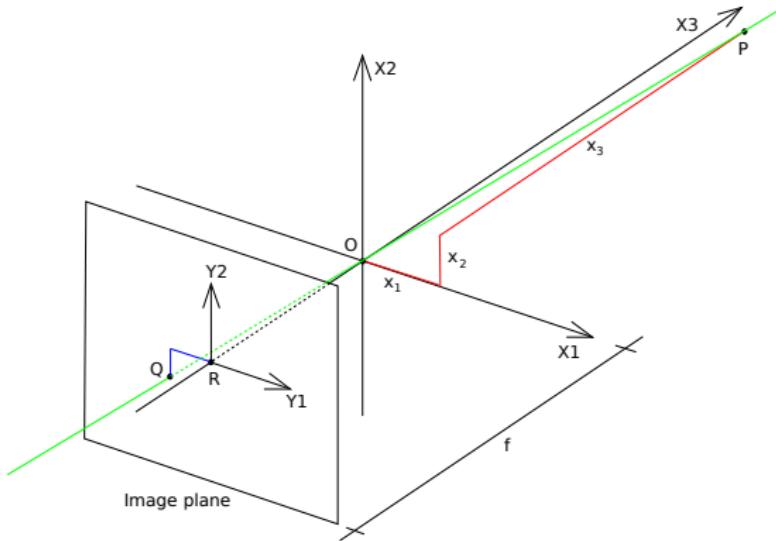
- ▶ Lens systems, geometric distortion, color distortion, depth of field, image resolution, frame rate, motion blur, compression quality, ...
- ▶ Complete imaging pipeline is complex
- ▶ Light travels from subject to sensor, to digital file
- ▶ Assumed uniform lighting; still, non-shiny subject
- ▶ Video = consecutive image frames

Video

- ▶ Frame after frame
- ▶ Moving subject
- ▶ Non-flickering lighting
- ▶ Motion blur
- ▶ Rolling shutter
- ▶ Line skipping
- ▶ Synchronization



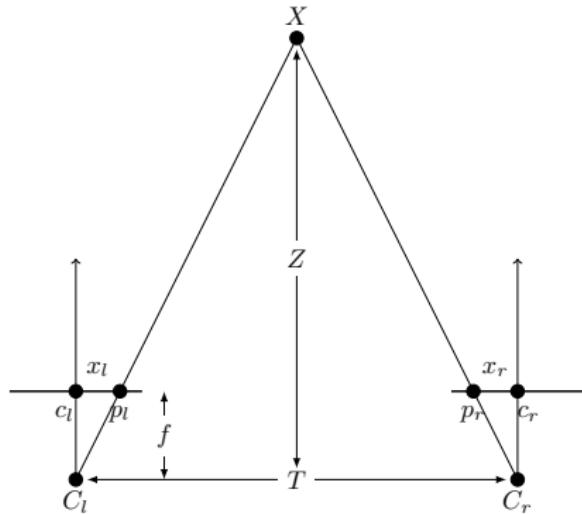
Simplified methods



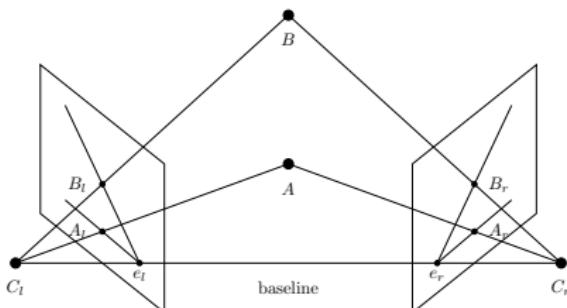
- ▶ Linear camera model, simple optical distortion models, orthogonal problems as separate steps
- ▶ Assumptions, corner detection, correspondence matching, triangulation, surface reconstruction, ...

2D to 3D: depth from disparity

- ▶ Just simple triangulation, in principle
- ▶ Two views, two positions in images \Rightarrow 3D position with depth
- ▶ In practice, many more sophisticated dense algos
- ▶ Complete pipeline needs to consider everything



(Multi-)Stereo system



- ▶ Manual or self-calibration for cameras
- ▶ Reconstruct 3D geometry \Leftrightarrow apply projection in inverse
- ▶ Need camera poses
- ▶ Reconstruct 3D view
- ▶ Determine scene from multiple viewpoints

Practical reconstruction system



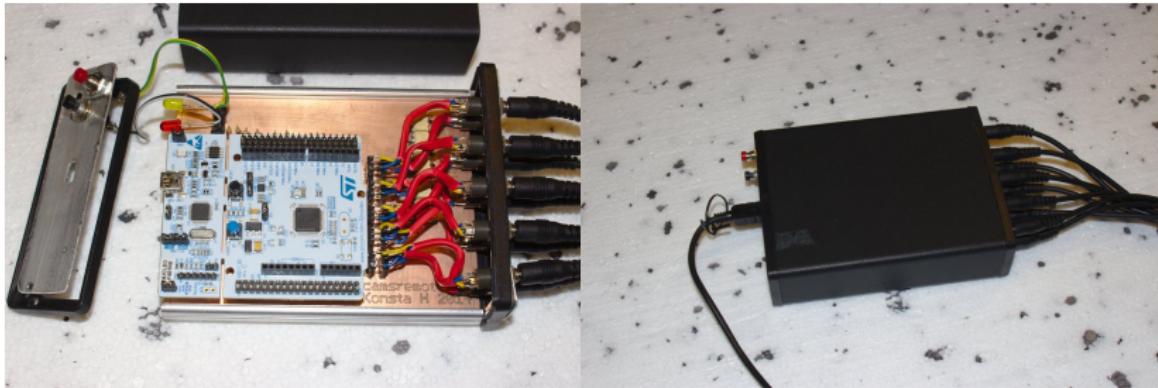
- ▶ Background theory ⇒ construct a reconstruction rig
- ▶ Acquire a geometric model of a subject
- ▶ Flexible and general purpose, easy to use
- ▶ Human head special interest
- ▶ Heavy study on available hardware and software

Rig implementation

- ▶ 9x Canon EOS 700D DSLR, 50 mm lens
- ▶ Wired remote trigger tool
- ▶ Magic Lantern
- ▶ auxiliary HW, support SW
- ▶ Not restricted to any specific reconstruction SW



Remote trigger tool

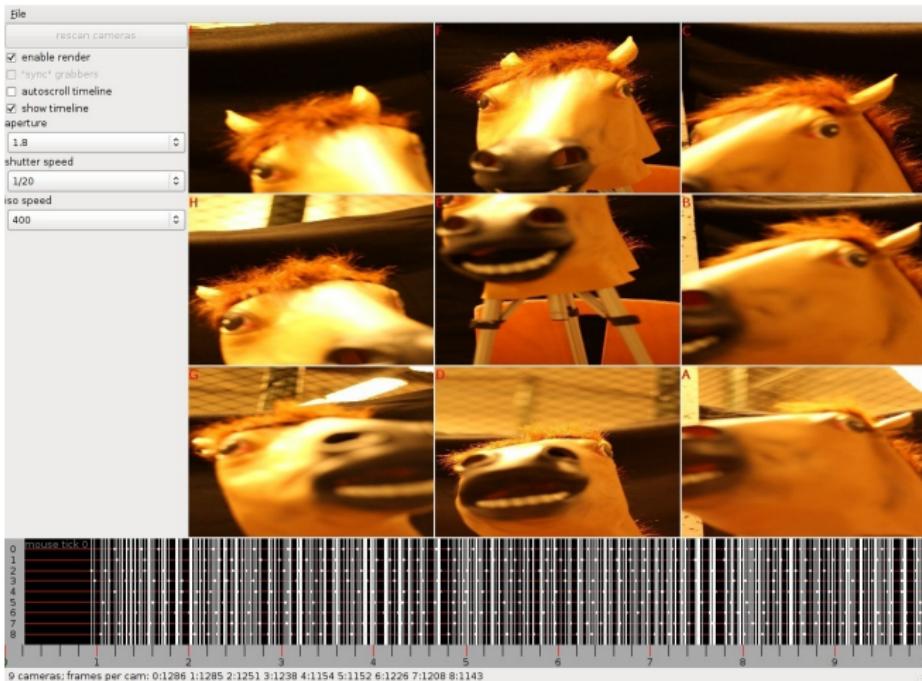


- ▶ Simultaneous or arbitrary focus/shutter release
- ▶ Custom built on a microcontroller using standard wired remote connections and a USB port
- ▶ Commercial tools mostly wireless, expensive

Software

- ▶ Few options available
- ▶ Just camera controls, based on *gphoto2*
- ▶ Generic open source library for cameras in C
- ▶ Preview tool, image download, camera configuration, remote shutter
- ▶ 3D reconstruction development out of scope

Previewer



- ▶ A great aid in aiming the rig as a whole
- ▶ Individual orientations with the viewfinders though

Control

- ▶ Camera exposes parameters via USB
- ▶ Shutter speed, aperture, ...
- ▶ Mostly implemented in gphoto2 CLI
- ▶ Image download as soon as captured
- ▶ USB connection to remote trigger box (Python)

```
./forallp.sh \
--set-config imageformat="RAW" \
--set-config imageformatsd="RAW" \
--set-config shutterspeed="1/100" \
--set-config iso="100" \
--set-config colorspace="sRGB" \
--set-config picturestyle="Standard" \
--set-config focusmode="One Shot" \
--set-config aperture="11" \
--set-config meteringmode="Evaluative" \
--set-config drivemode="Single" \
--set-config capturetarget="Memory card"
```

Experiments

- ▶ Synch tests
- ▶ Sample scans
- ▶ A few reconstruction programs tested
- ▶ Numerous libraries and commercial tools reviewed



Source data



18 Mpix per cam

Output geometry



1-5M points

Textures

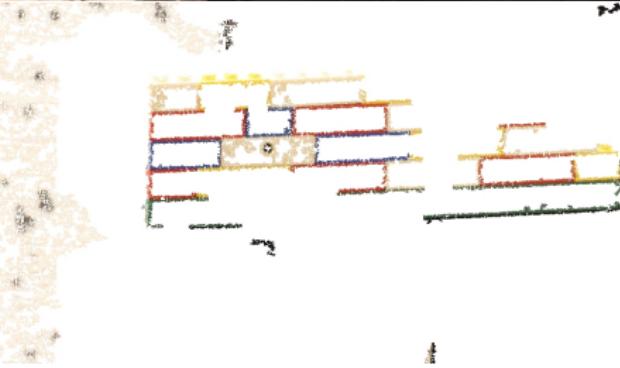
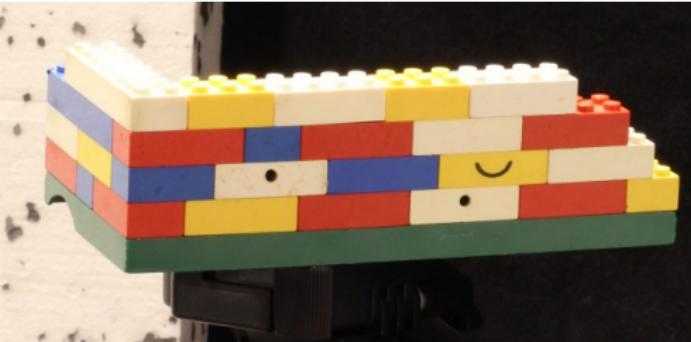


Hand



Occlusion issues

Legos



No texture

Rubber horse mask



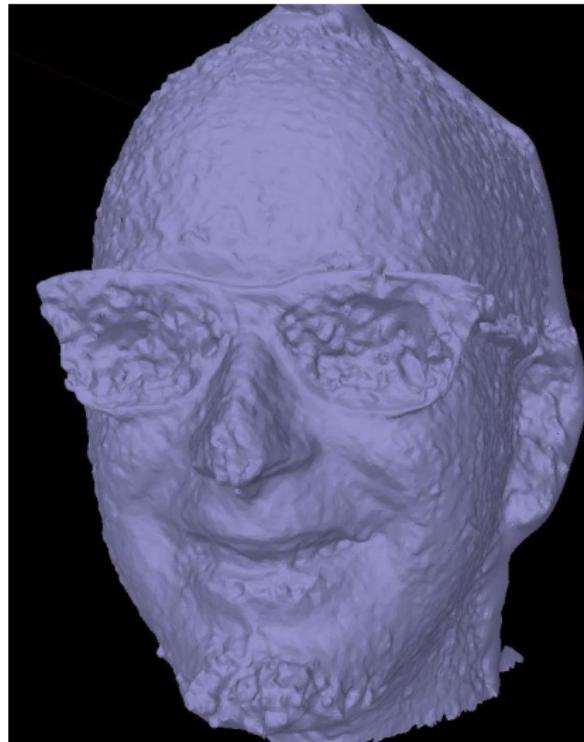
Uniform color areas

Level of detail



Issues

- ▶ Video synchronization
- ▶ Light requirements
- ▶ Post processing
- ▶ Camera quirks



Future work

- ▶ Actually use the rig
- ▶ Reproduce state-of-the-art
- ▶ Mesoscopic face geometry
- ▶ Research new ways



Conclusion



- ▶ Background and feasibility study ⇒ built reconstruction rig
- ▶ Sub-millimeter-scale content
- ▶ Data needs post-processing
- ▶ Future work: actually use it