Trends and Challenges of Augmented Reality

Dieter Schmalstieg
Graz University of Technology
Let’s Start With A Quiz

• Who knows the name of this device?

Microsoft Hololens, a head-worn device for Augmented Reality
How Does Augmented Reality Work?

Overlay the real world with computer graphics

Virtual content → Registration of virtual content → Situated visualization

SPATIAL MODEL

Real world model → Pose tracking → User input

Virtual content

Real world model

Qualcomm Vuforia
We Have Come A Long Way

2000
6-8kg
€12,000

2006
1.5kg
€5,000

2008
0.1kg
€500
2016: Microsoft HoloLens

What we know:
- Optical see-thru display
- Wearable computer
- RGBD camera(s)
- Hardware-accelerated tracking and mapping

What we don’t know:
- A lot
- (Despite marketing, nothing holographic about it)
A Strong Disturbance in the Force

• 1990s: 1\textsuperscript{st} wave of excitement about Virtual and Augmented Reality
  – But remains a niche market

• 2010s: Massive investments by industrial players
  – Microsoft releases HoloLens
  – Facebook acquires Oculus Rift
  – Apple acquires PrimeSense, Metaio etc.
  – Valve, Sony, Samsung launch VR gaming platforms

• What is the goal this time?
AR Business Models

• Business models not yet clear
  – Increase **consumer** adoption, games, advertising
  – Consumer (currently) drives hardware development
  – Increase **industrial** use
  – Industrial may allow higher cost, more learning?

• Example
  – Vuforia, a leading AR SDK
  – Marketed by Qualcomm for consumer
  – Sold to PTC (industrial solution provider) in 2015

• **My speculation:** Industrial use will be big first
What are Industrial Use Cases?

Discrepancy checking

Construction progress monitoring

Hidden infrastructure visualization

[Schönfelder, Schmalstieg, ISMAR2008]

[Schall. Mendez, Schmalstieg, ISMAR2008]
More Industrial Use Cases

Maintenance instructions

Process data visualization and control

Tele-assistance

Image courtesy of Steffen Gaugglitz
Recap: Augmented Reality Systems

Virtual content → Registration of virtual content → Situated visualization

Real world model → Pose tracking → User input

SPATIAL MODEL

Graz University of Technology
Topic Today: Authoring of Instructions

Virtual content

Registration of virtual content

Situated visualization

Pose tracking

User input

SPATIAL MODEL

Real world model
What is Required for AR Instructions?

• *(A Kinect for tracking; won’t talk about it)*

• 3D model of the real object
  – Scanned with Kinect or existing CAD model

• Decomposition of model into parts

• Sequence of parts
  – Disassembling: remove parts
  – Assembling: add parts
  – Maintenance: remove, manipulate, add

• Representation of the necessary motions

• Visualizations that convey the actions well
How Can We Generate AR Instructions?

• Manually
  – Use 3D modeling tools + (maybe) scripting
  – Tedious, requires expert modeling knowledge

• From existing printed instruction manuals

• From existing videos
How Can We Generate AR Instructions?

• Manually
  – Use 3D modeling tools + (maybe) scripting
  – Tedious, requires expert modeling knowledge

• From existing printed instruction manuals

• From existing videos
What are the Elements of a Manual?

- Labels
- Directional arrows
- Before-after sequences
- Explosion diagrams
Retargeting from 2D to 3D

• Problem 1
  – Where is every part located in real world?

• Approach
  – Must be able to find parts (semi-)automatically

• Problem 2
  – What intent does the illustration have?

• Approach
  – Synthesize animation of the parts to communicate the intent
Preparations

• Scan 2D manual (or download PDF)
• Obtain 3D model of the machine
  – Get CAD data from vendor
  – Alternatively, use 3D scanner (Kinect again)
• Register 3D model with illustration
  – Same problems as 3D tracking-by-detection
  – For just a few camera poses, this is an easy task
Labels

• Read labels with optical character recognition

• Generate ID buffer
  – Every pixel refers to the part underneath

• Search line
  – Look up endpoint of line in ID buffer
  – Points to the part
Motion Arrows

• Detect arrow shapes (they are *pointy!* 😊)
• Interpret direction of 2D arrow in 3D

• Search referred part
  – Near arrow shaft
  – Must be removeable in direction of the arrow
  – Uses automatic CAD disassembly planer

[Kerbl, Kalkofen, Schmalstieg, CGF2015]
Before-After Sequences

- CAD model registered to “Before”
- Difference to “After” applied to ID buffer generates candidates
- Disassembly planner determines possible motions of candidates
- Compare difference image of “After” to possible motions
Explosions Diagrams

- Register CAD model to one part in the image
- Disassembly planner searches candidates for explosions
- Move candidate
- Compare to image
- Robust comparison with oriented chamfer distance
Multiple Moving Parts

Input 1

Processing

Output animation

oriented chamfer distance

difference mask (Input 1 / 2)

calculate best fit

Input 2
Results
How Can We Generate AR Instructions?

• Manually
  – Use 3D modeling tools + (maybe) scripting
  – Tedious, requires expert modeling knowledge

• From existing printed instruction manuals

• From existing videos
Cooking as a Video Game

“Sight” (short film by Eran May-raz and Daniel Lazo, Israel, 2012)

Note: These images are created offline by an animation artist!
Here is our version 😊

Knife skills video

Knife skills AR tutorial
Your Whole Life Is Already On Youtube

We can use these videos!
Overview of the Approach

Edit motion
Temporal segmentation
3D registration

Extract from input video
Track objects
Reconstruct 3D motion

Visualize
3D glyph synthesis
Ghosted rendering
Motion Extraction of Unknown Rigid Objects

- Unknown object in video $\rightarrow$ no 3D model
- Input video material usually not good enough for structure from motion
- Scan a similar object with a Kinect
- Create a simple rigging
- Automatically deform by skinning
- Deformed object can be tracked
Motion Extraction of Tools on Surfaces

- Track the tooltip and map trajectory to atlas
- Can retarget motion to any surface with same atlas
Motion Segmentation

• Segment the motion by combining
  – Path: only unique motions
  – Curvature: separate orientation change from jitter
  – Velocity: cut, if no motion for a certain time

• Can be extended to skeleton tracking
  – Greedy segmentation based on all bones

• Can be used as input to synthesize arrow glyphs
Motion Registration to Target Object

- Attach to source of motion
  - Guide user to source object

- Attach to target motion
  - Guide user to destination object

- Attach to both source and target
  - Rarely needed
  - Must scale the motion
3 Visualization

Ghosted objects

Arrow glyphs
Summary

• Authoring is overlooked as an important element of AR experiences

• Re-use of existing media sources
  – Printed manuals, video tutorials

• Future alternative: authoring by demonstration inside AR
What Comes Next?

• Better Input
  – Programmable cameras, camera arrays, better optics
• Output
  – Wide FOV, lightfield displays, adjustable focus, eyetracker
• AR meets the Internet of Things
• Storytelling: AR as a dramatic medium
• Social Computing: Situated Facebook?
• *Industrial it great, but consumer success would be even better*…
Questions?

Ask Now!

Spring 2016

www.augmentedrealitybook.org