Advances and Trends in Augmented Reality Systems

Dieter Schmalstieg

Graz University of Technology

Let’s Start With A Quiz

- Who knows the name of this device?

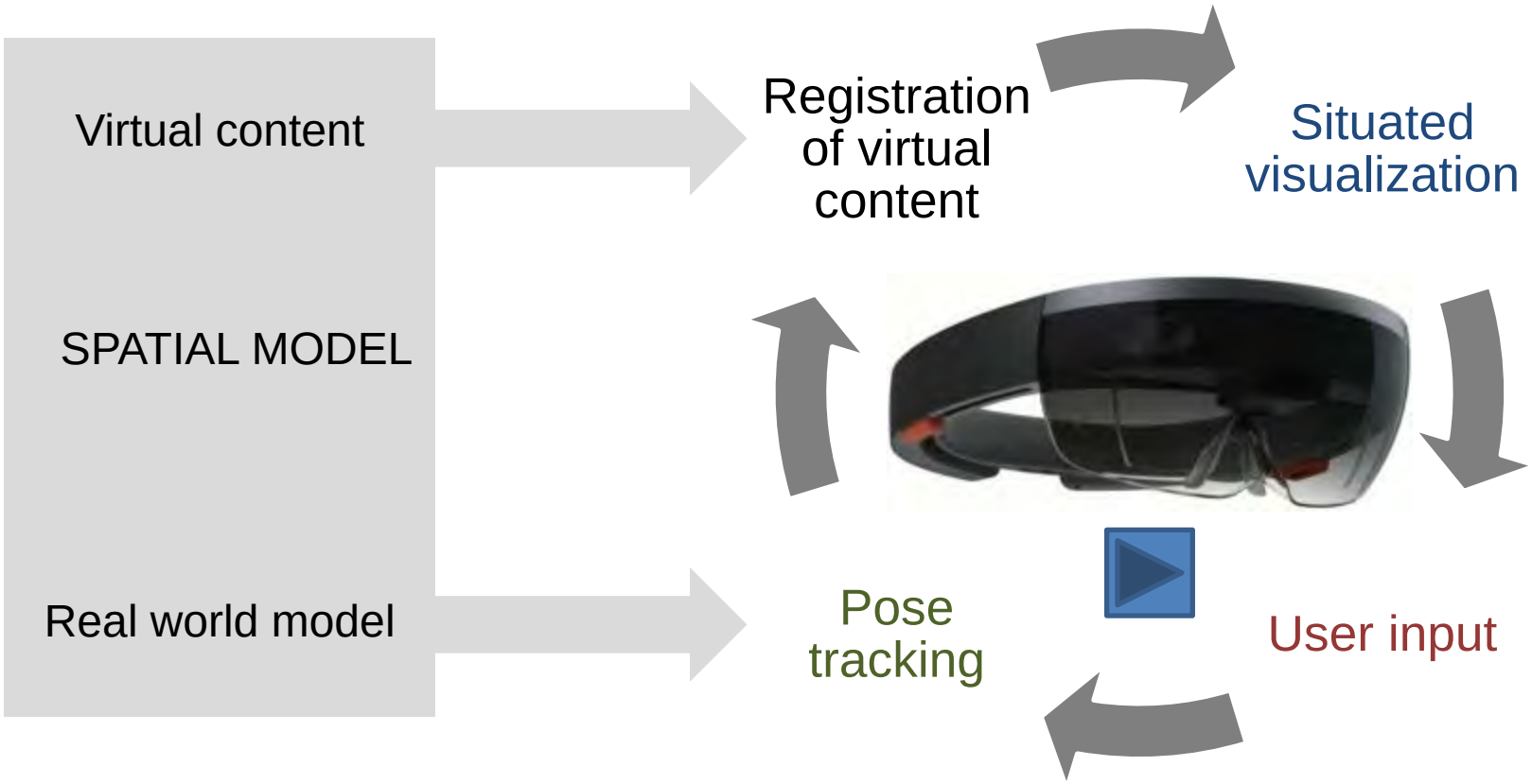
Microsoft Hololens,
a head-worn device for Augmented Reality

A Strong Disturbance in the Force

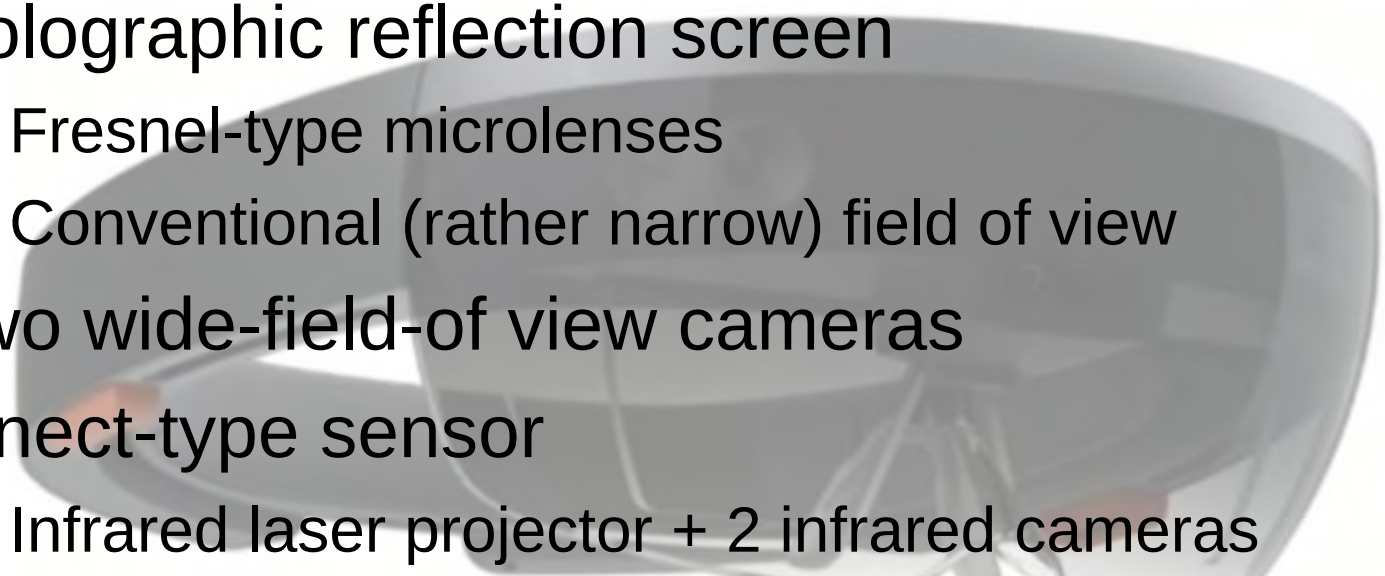
- 1990s: 1st wave of excitement about Virtual Reality (VR) and Augmented Reality (AR)
  - But remains a niche market
- 2010s: Massive investments by industrial players
  - Microsoft releases HoloLens
  - Facebook acquires Oculus Rift
  - Apple acquires PrimeSense, Metaio etc.
  - Valve, Sony, Samsung launch **VR gaming** platforms

- What is the application area of AR?

AR as Multimedia System Loop



Microsoft HoloLens



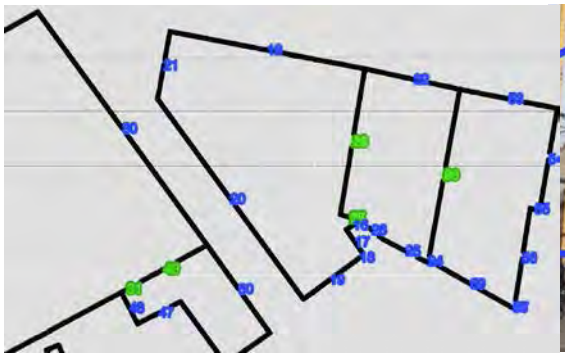
- Per eye-microprojectors
- Holographic reflection screen
  - Fresnel-type microlenses
  - Conventional (rather narrow) field of view
- Two wide-field-of view cameras
- Kinect-type sensor
  - Infrared laser projector + 2 infrared cameras
- Dedicated acceleration chip, probably for
  - Simultaneous localization and mapping (SLAM)
  - Sensor fusion

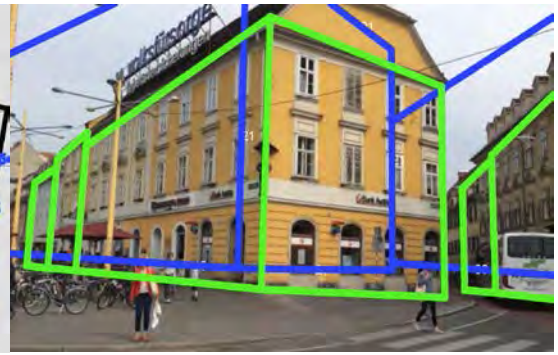
What the HoloLens Will Bring

- Tracking is solved now?
- Many years of 3D computer vision research
  (including my own)
  now obsolete?
- Not quite…
  - Infrared does not work outdoors
  - SLAM does not solve *global* localization problem

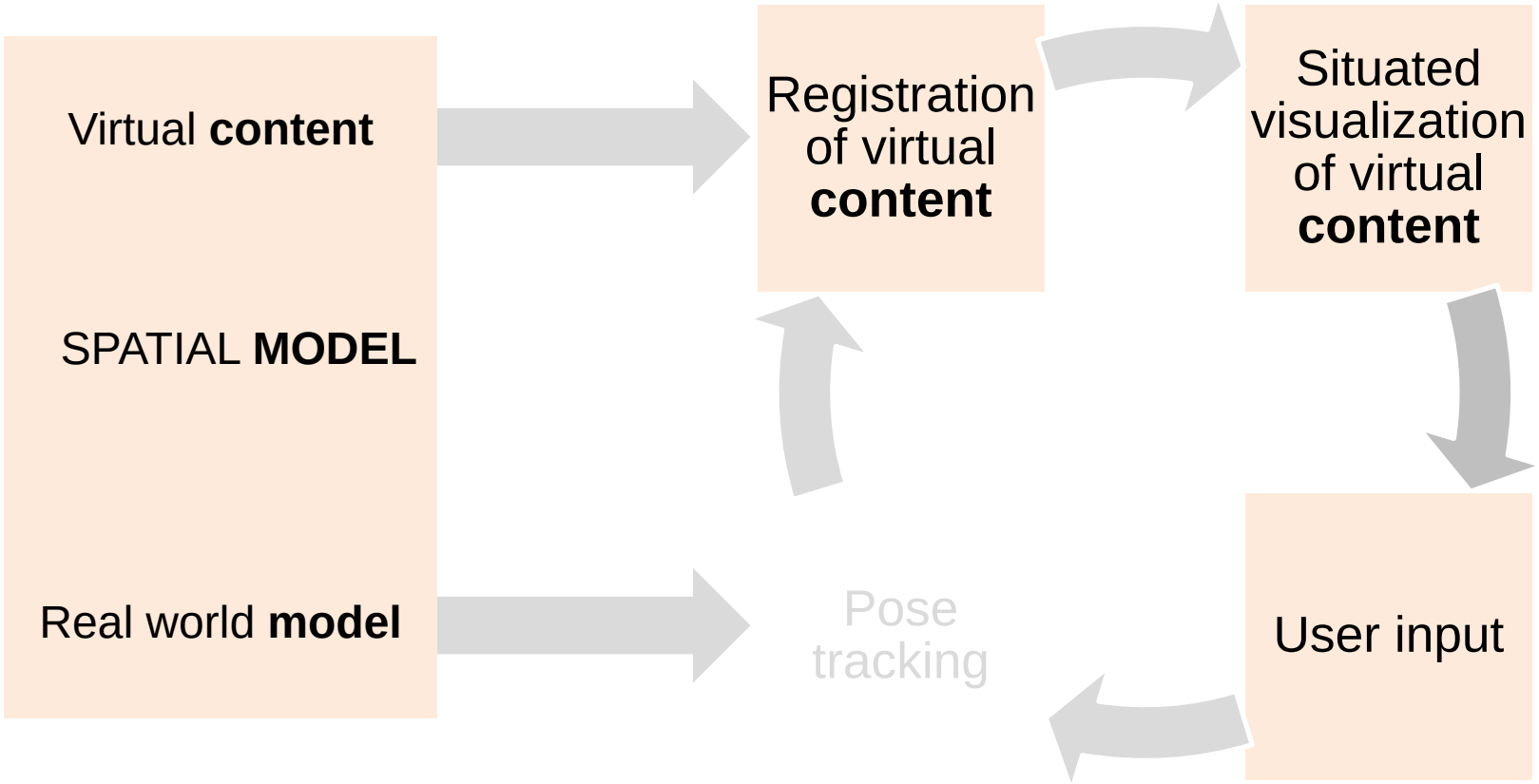
Global Localization with 2D Maps

- Problem: City-scale 3D scans are costly
- Localization with **existing** data
  - 2D map with building footprints and roof height
  - 3D objects extruded from 2D map
- Sensors (GPS, IMU, compass) give prior
- Detect facade outline in camera image





But Enough About Tracking…



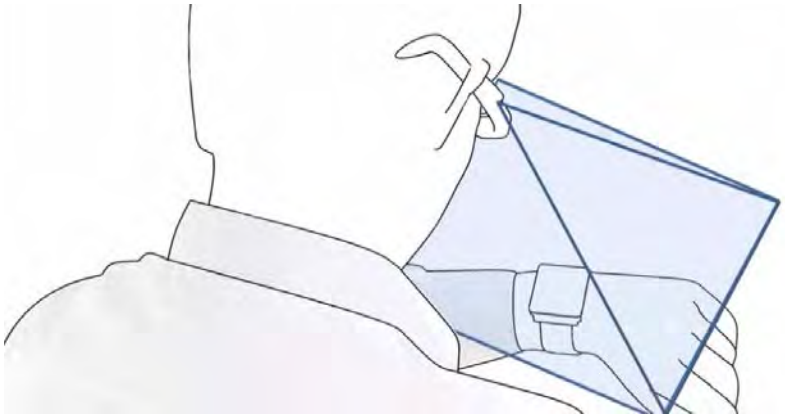
….what about the **content**?

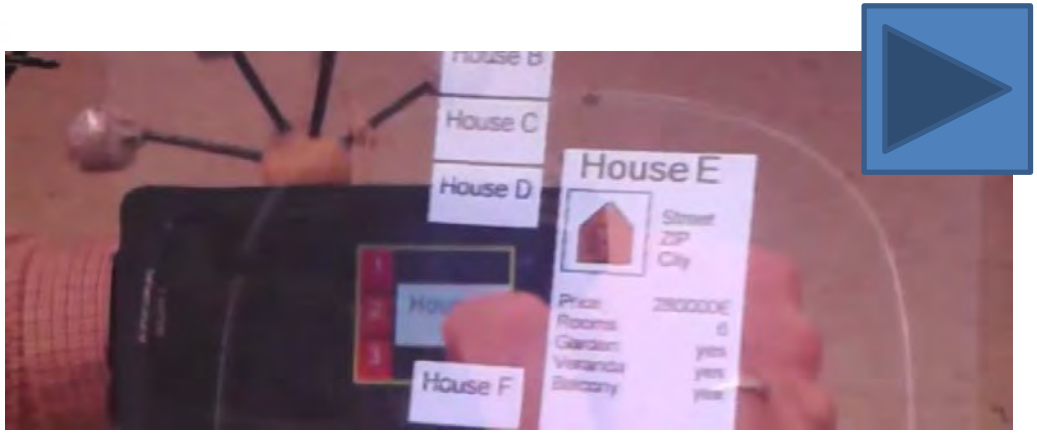
Where Does Digital Content Originate?

- Digital artefacts are free from physical constraints
- Moore's law applies to many kinds of phenomena
  - 1980's: Digital office work
  - 1990's: Digital private life (email, photography, music)
  - 2000's: Social and mobile computing
  - 2010's: Cloud (hardware irrelevant), IoT, Ubicomp
- But: not *calm* computing
  - App for every aspect of life?
  - AR can provide *situated*, less obstrusive interface

AR as an Interface

- Internet of things allows control over physical environment
- But physical objects have no input or output
- AR can provide direct manipulation of the parameters





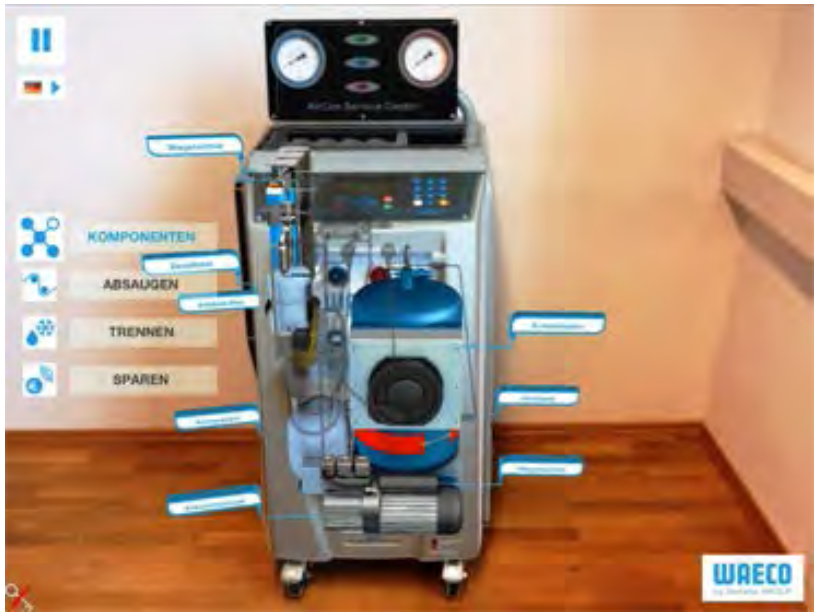
AR as a Dramatic Medium

- AR progresses from technology to dramatic medium
- New medium has new characteristics
  - Before “Citizen Kane”, movies were like stage recordings
- Characteristics of AR as a medium
  - Combines real+virtual → can display content anywhere
  - Spatially registered → free chooce of viewpoint
  - Interactive in real time → always interacting with physical space
- Requires new conventions
  - E.g., narrative focus vs free camera control
  - No cut scenes as in games
  - AR is more like a theater stage

AR as a Communication Medium

- AR as a medium for *communication*
- Content provided by
  - Professionals (entertainment, journalists etc.)
  - Authorities (traffic etc.)
  - Individuals → *social*
- Not geo-location, but precise spatial annotation (part of an object)
- *Linking* (like web links) between virtual and *real*
- *Channels* (like blogs) to organize content
- Always-on, context-driven, non-linear streams

Example: AR Instructions



Process data visualization and control

Maintenance

What Content is Required for AR Instructions?

- A sensor (Kinect) for tracking
- 3D model of the real object
- 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?

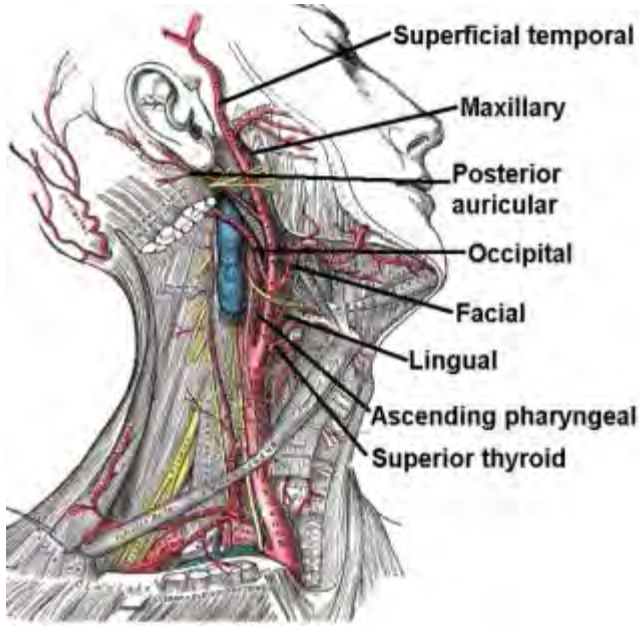
- Manually
  - Use 3D modeling tools + (maybe) scripting
  - Tedious, requires expert modeling knowledge

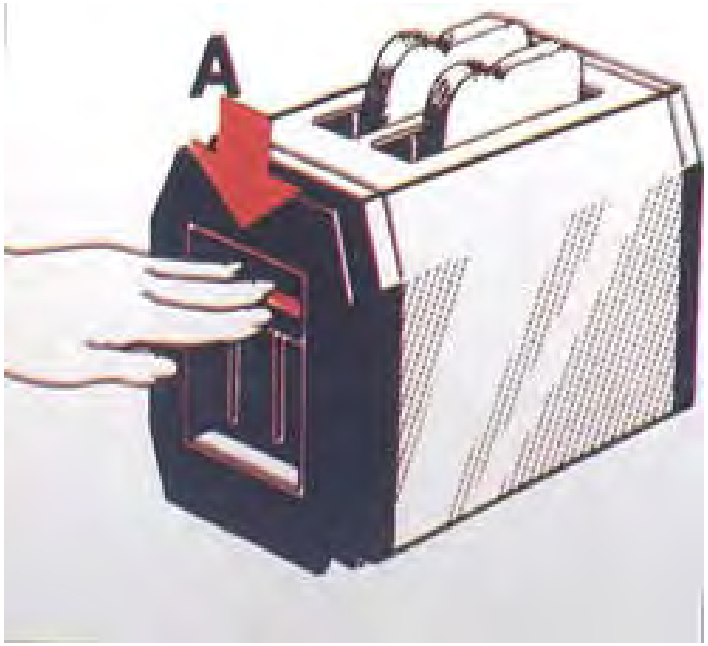
- **From existing printed instruction manuals**

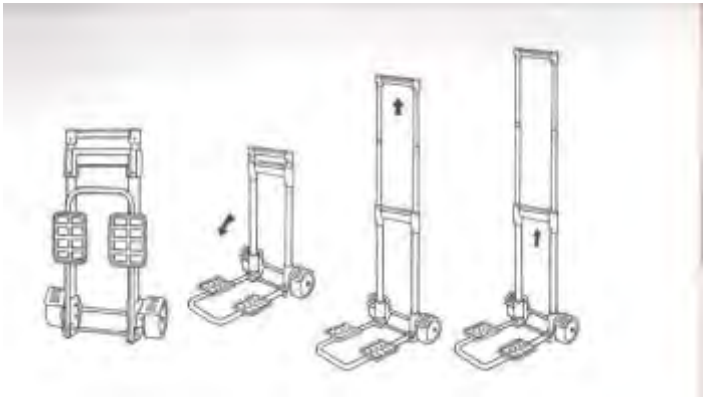
- From existing videos

What are the Elements of a Manual?

- Labels
- Directional arrows
- Before-after sequences
- Explosion diagrams







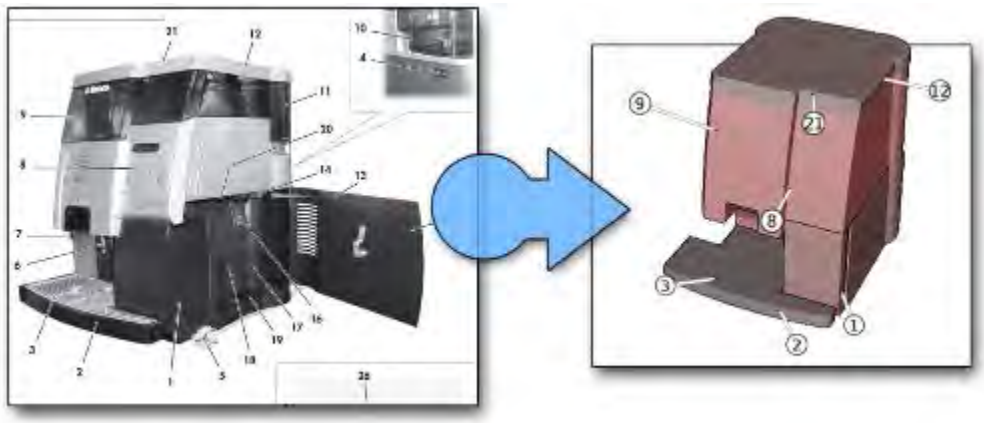
Retargeting from 2D to 3D

- 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

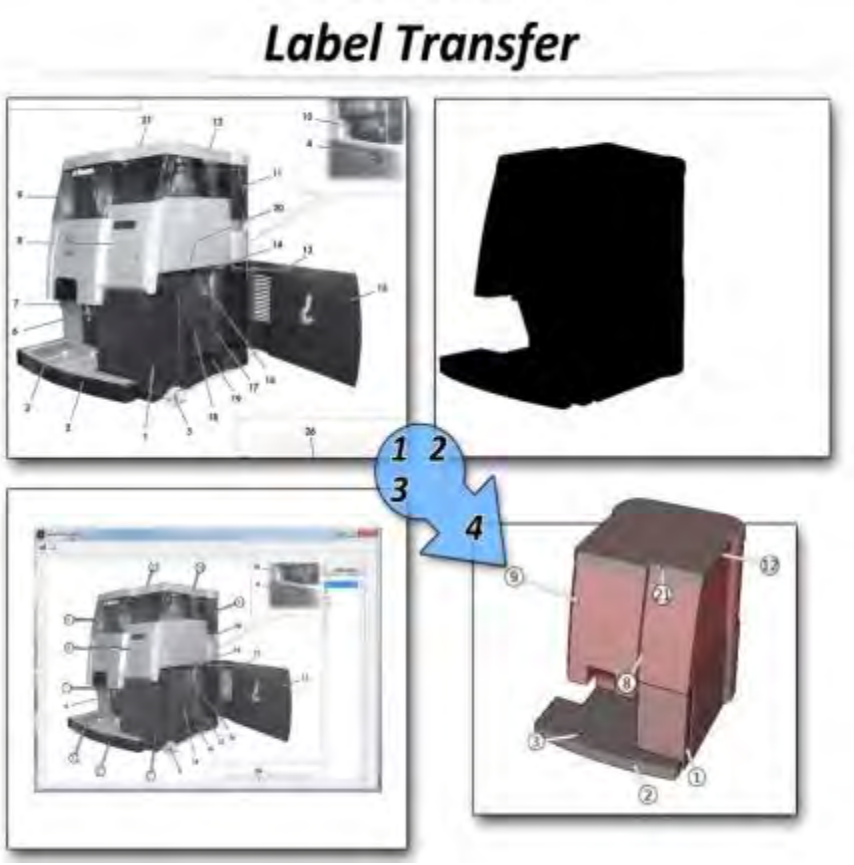
Preparations

- 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

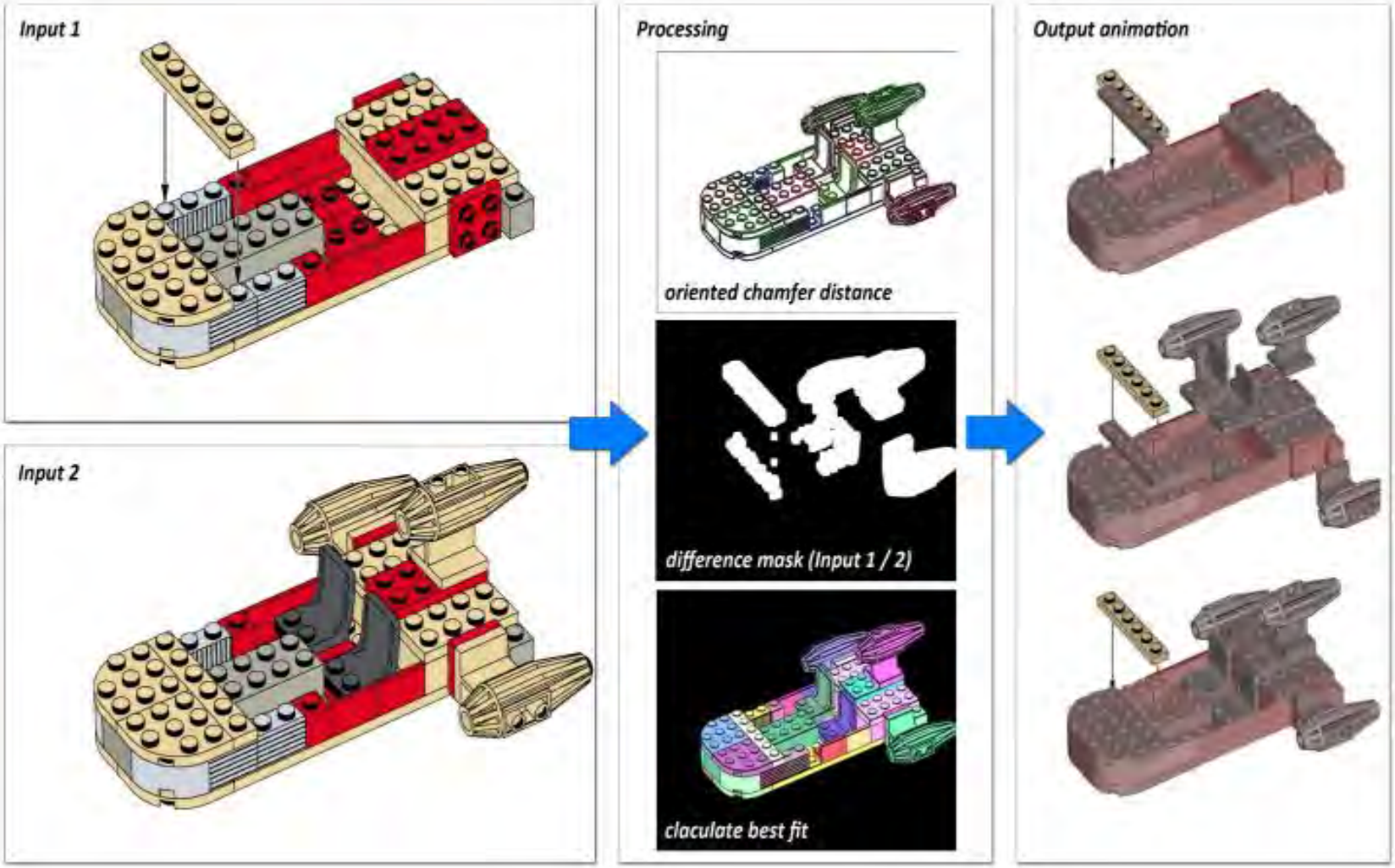


Labels

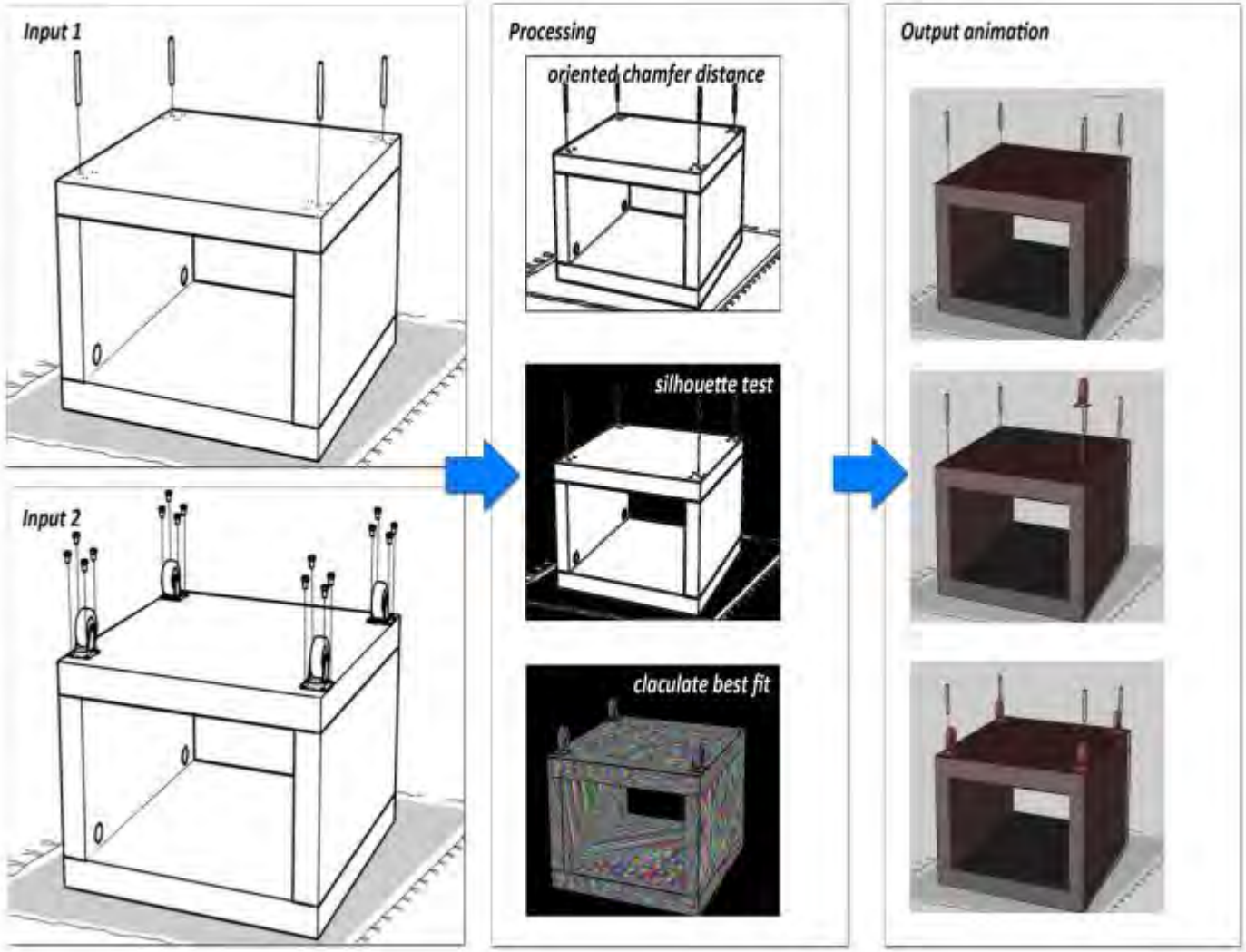
- Read labels with optical character recognition
- Generate ID buffer
  - Every pixel refers to the part underneath
- Search line
  - Look up endpoint of line in ID buffer
  - Points to the part



Multiple Moving Parts



Results


Input 1
Input 2
Processing
oriented chamfer distance
silhouette test
claculate best fit
Output animation

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**

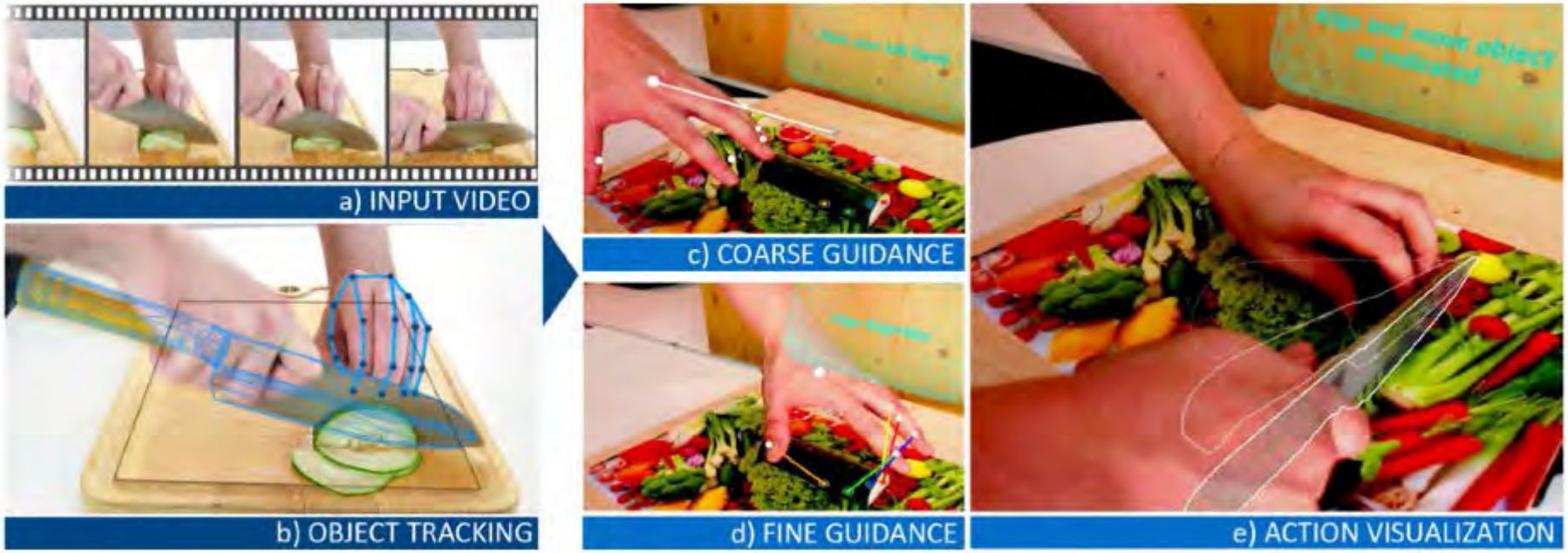
Cooking as a Video Game

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

Note: These images are created offline by an animation artist!

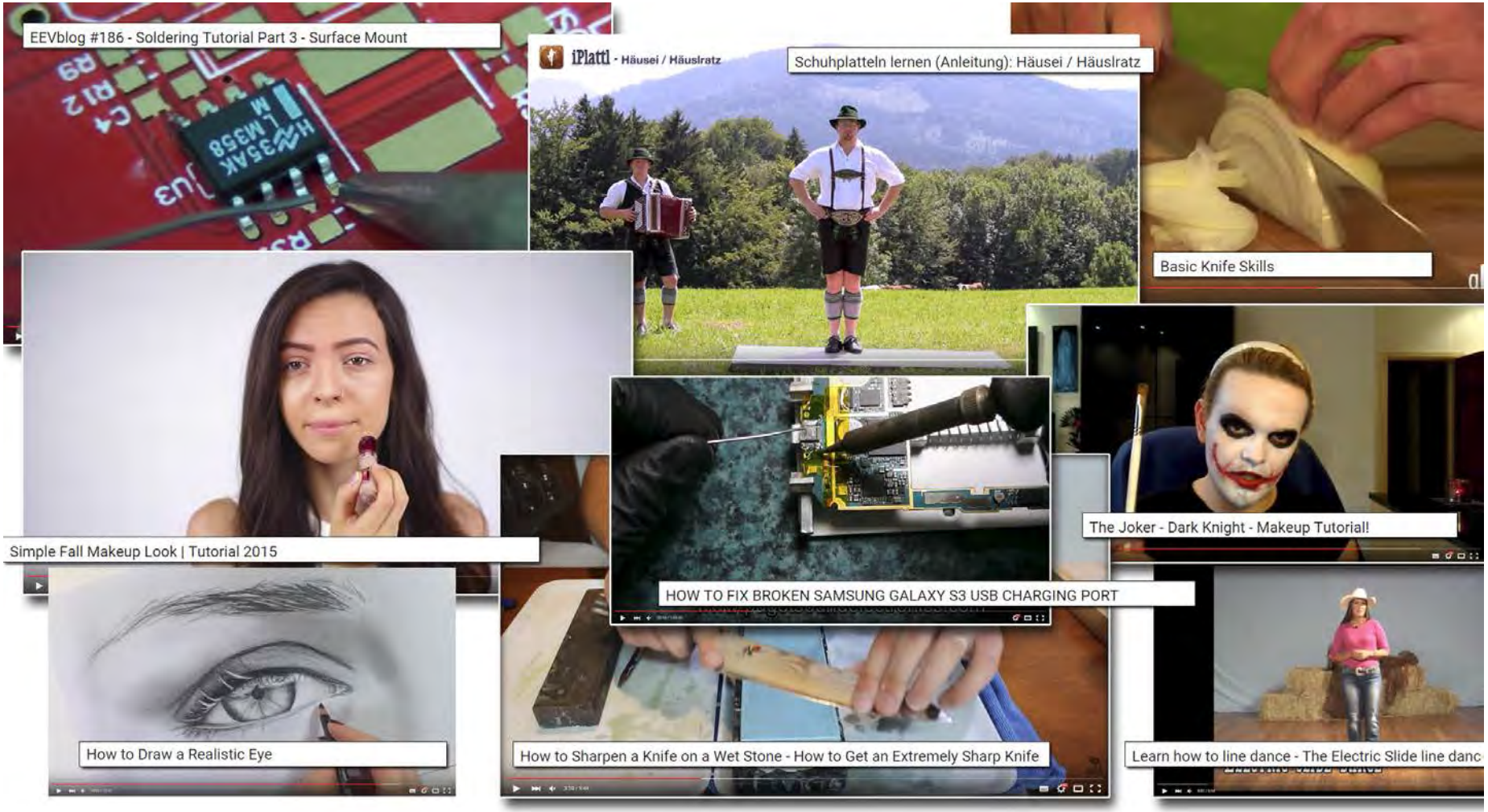
Here is our version ☺

Knife skills video



Knife skills AR tutorial

Your Whole Life Is Alreasy 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

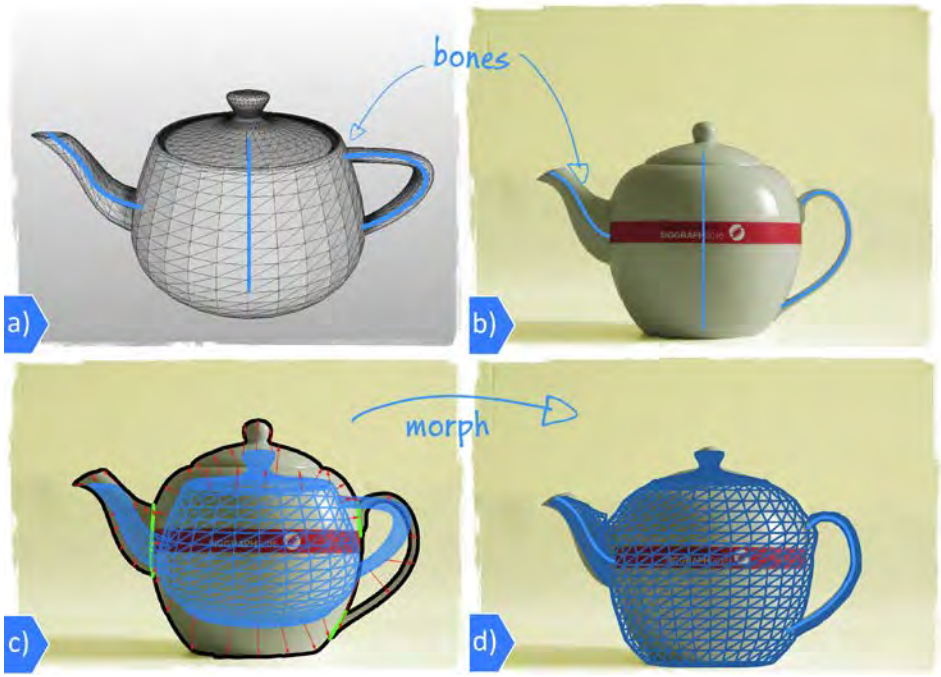
Visualize

3D glyph synthesis

Ghosted rendering

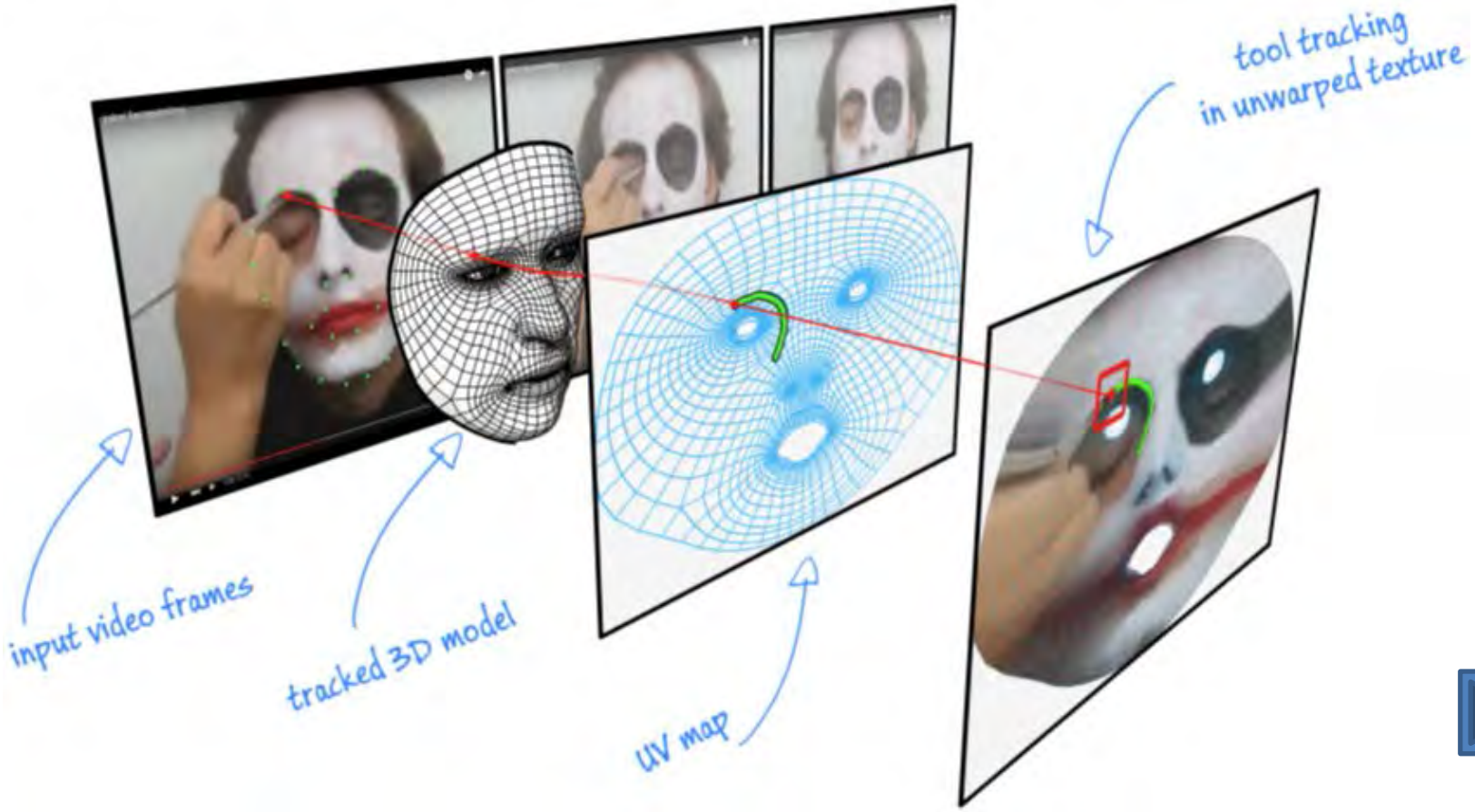
1 Motion Extraction of Unknown Rigid Objects

- Unknown object in video → 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



1 Motion Extraction of Tools on Surfaces

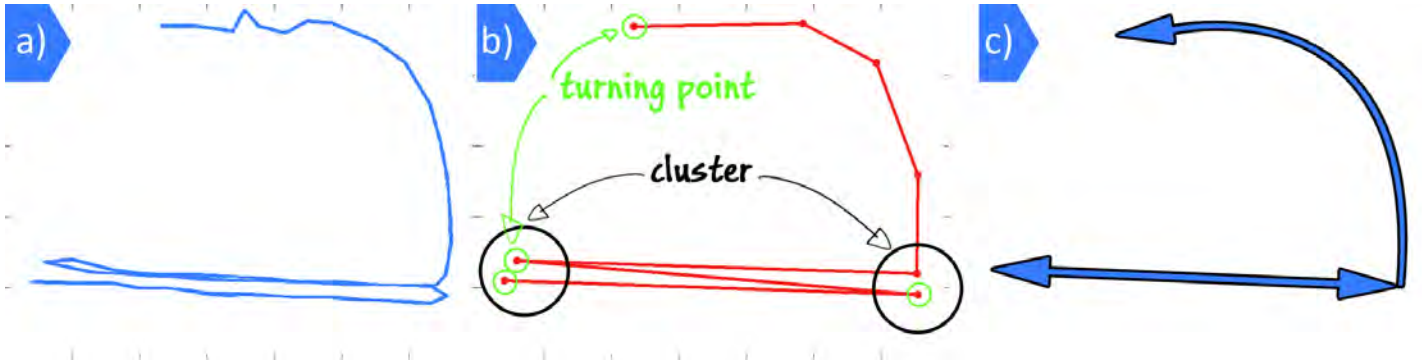
- Track the tooltip and map trajectory to atlas
- Can retarget motion to any surface with same atlas



2

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

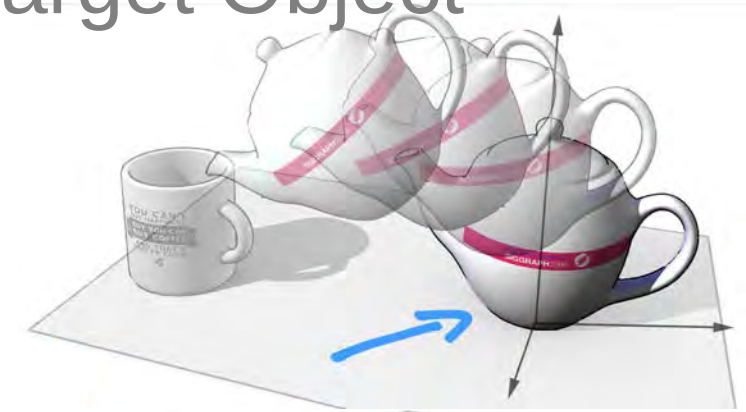


- Can be extended to skeleton tracking
  - Greedy segmentation based on all bones
- Can be used as input to synthesize arrow glyphs

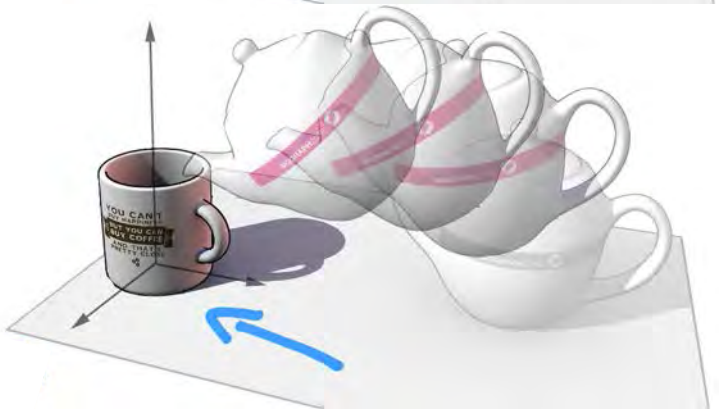
2

Motion Registration to Target Object

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



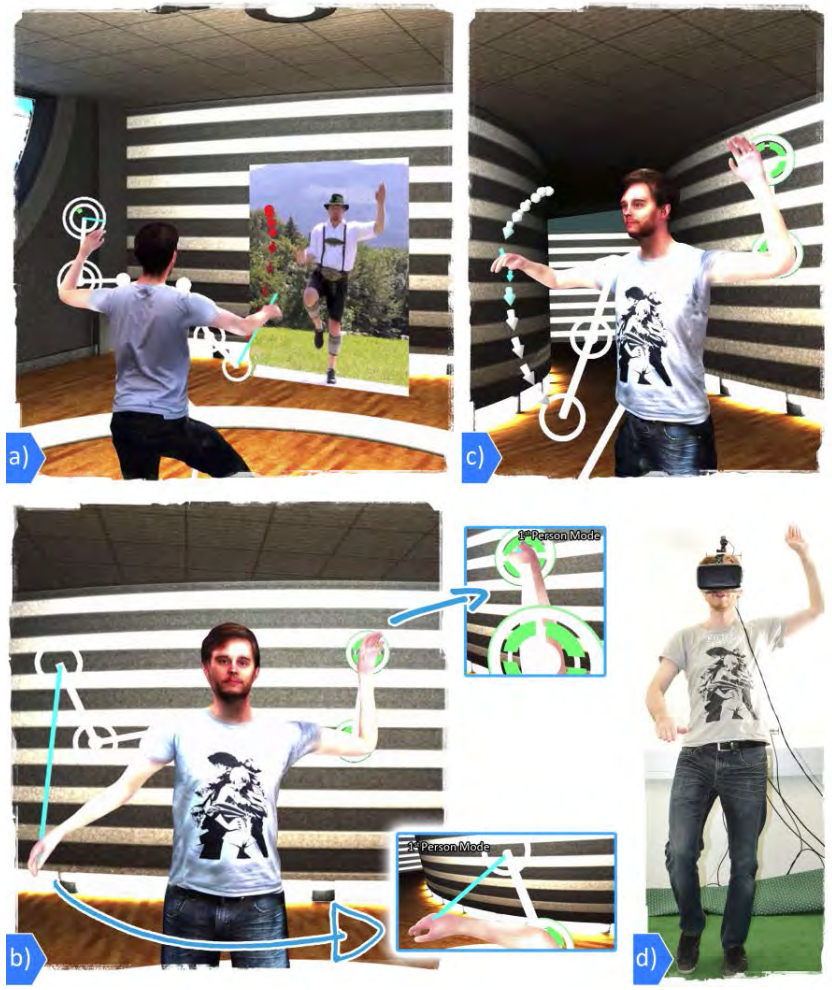
- Attach to target motion
  - Guide user to destination object



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

3 Visualization

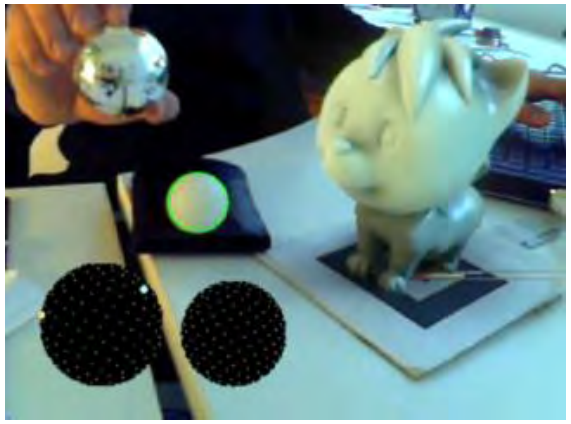
Ghosted objects



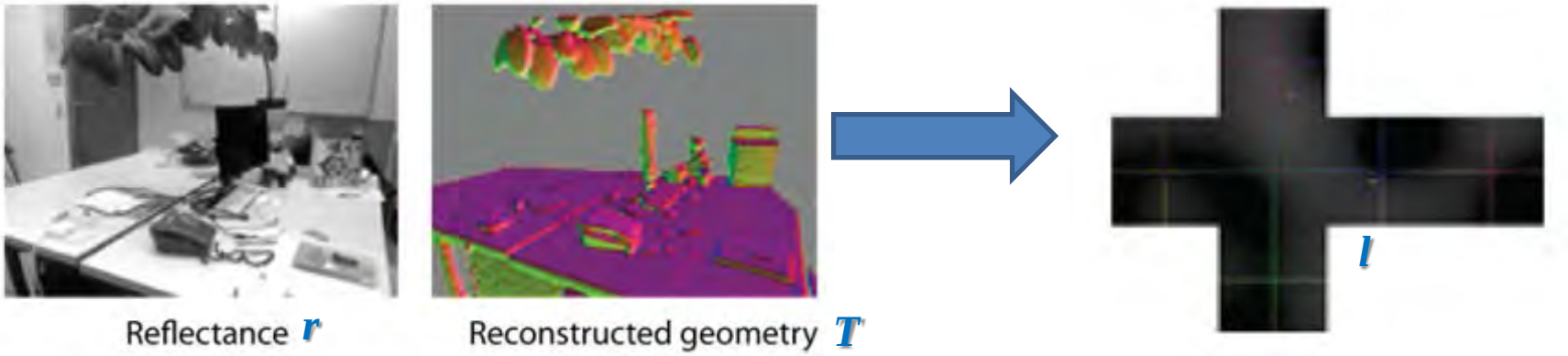
Arrow glyphs

More Realistic Visualization of AR Content

- Geometric scanning is easy now (depth sensors)
- Realistic lighting for AR also needs photometry
  - Surface reflectance (material properties)
  - Incident illumination (light sources)
- Photometric scanning not yet easy
  - Ambiguous problem
  - Mostly done in special studios: lightstage, lightprobe…
- **Can you do-it-yourself**?
  - Consumer hardware
  - Scan your **own real content** for AR

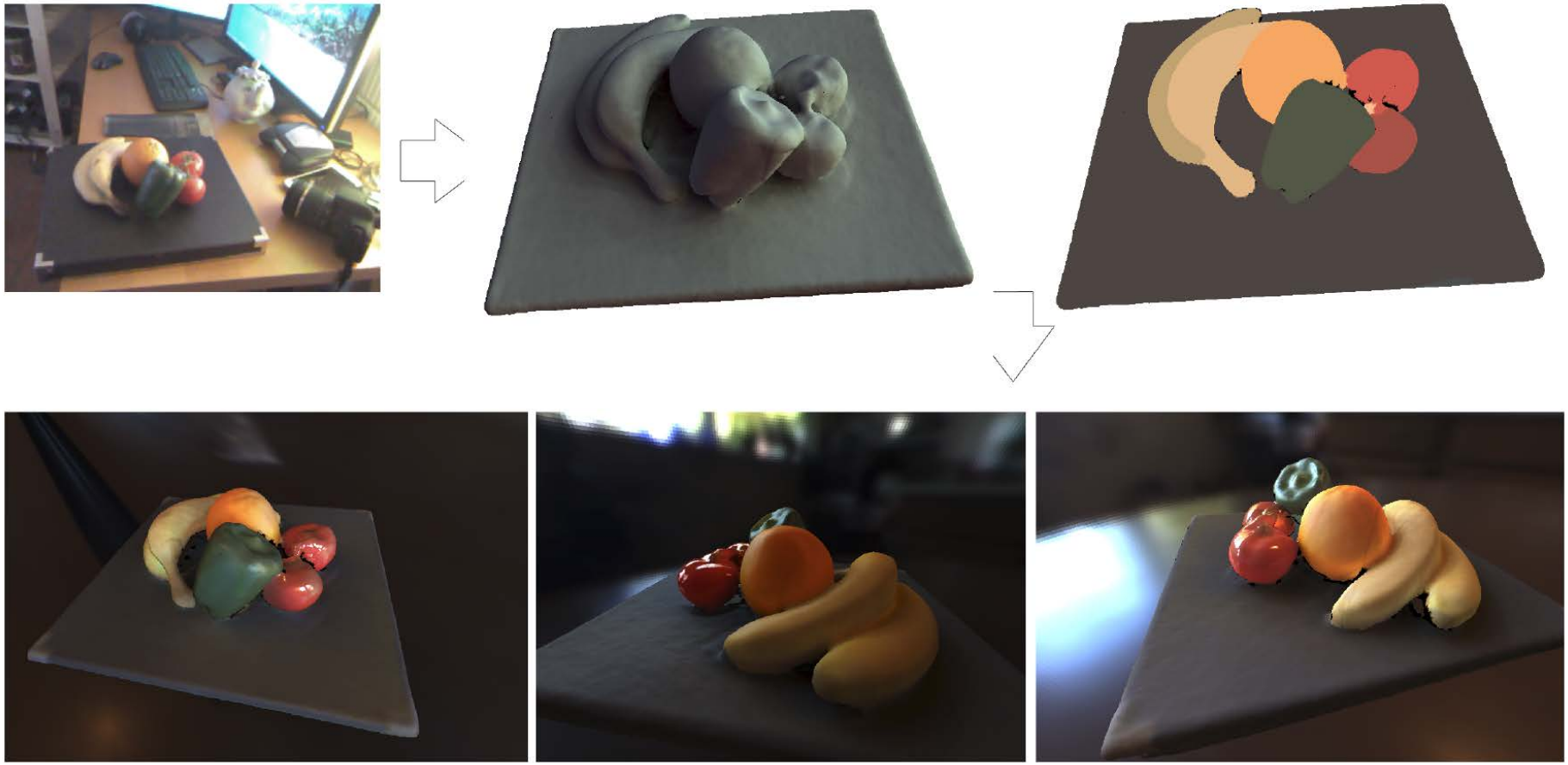


Photometric Reconstruction from Images



- Solve a linear system **T** * **l** = **r**
- Output:
  - **l** … incident light
- Input:
  - **r** … reflected light, as seen in video
  - **T** … radiance transfer, computed from scene geometry

Material Segmentation and Specularities

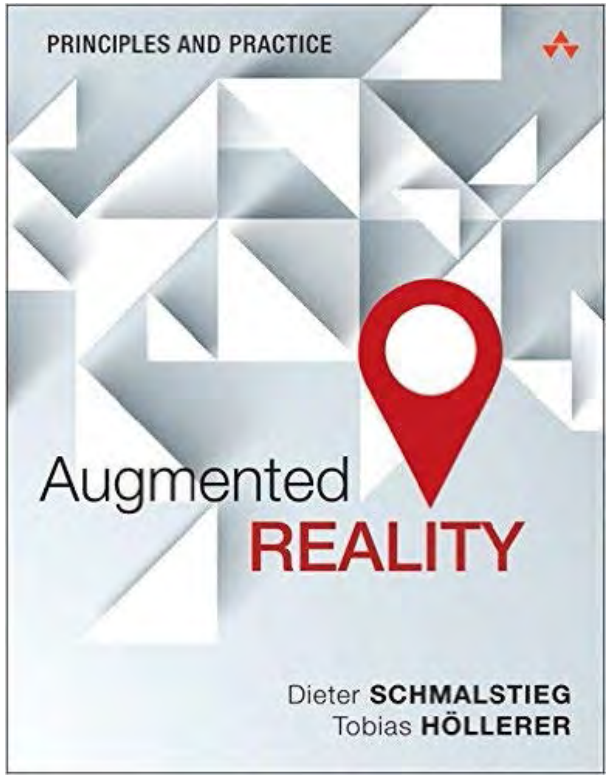


Summary and Outlook

- We will see the rise of new media: VR, AR
- New media will need content
  - For drama, narrative
  - For communication
- Re-use of existing sources
  - Printed manuals, video tutorials, 3D photometric scans
- The future
  - Authoring by demonstration
  - AR Telepresence
  - Augmented Humans?

Questions?

Find answers in this book! ☺



OUT 21 June 2016
www.augmentedrealitybook.org