The Future of Mixed Reality Interactions

Hrvoje Benko

IEEE AIVR 2019
A Vision of All-day MR

- Sensory and social **superpowers**
- Communicate and collaborate **at a distance**
- Next computing platform

Facebook F8 2017
A Vision of All-day MR

- Sensory and social superpowers
- Communicate and collaborate at a distance
- Next computing platform

Facebook F8 2017
What is taking so long?
New Computing Era =
New Display Form Factor + New Input Method + New Interactions
Not convinced?
<table>
<thead>
<tr>
<th>Product X</th>
<th>EMEA</th>
<th>US</th>
<th>Americas</th>
<th>APAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Units</td>
<td>Units</td>
<td>Units</td>
<td>Units</td>
<td>Units</td>
</tr>
<tr>
<td>15.00</td>
<td>17.00</td>
<td>11.25</td>
<td>12.75</td>
<td>2.00</td>
</tr>
<tr>
<td>22.00</td>
<td>20.00</td>
<td>16.5</td>
<td>4.00</td>
<td>1.00</td>
</tr>
<tr>
<td>20.00</td>
<td></td>
<td>1.00</td>
<td>2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>11.25</td>
<td></td>
<td>1.00</td>
<td>2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>12.75</td>
<td></td>
<td>2.00</td>
<td>4.00</td>
<td>2.00</td>
</tr>
</tbody>
</table>

Shipments:
- 15
- 4.00
- 2.00

Sum = 2.00
<table>
<thead>
<tr>
<th>1960s</th>
<th>1980s</th>
<th>2000s</th>
<th>2020s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Command Line Interfaces (keyboard)</td>
<td>Graphical User Interfaces (mouse)</td>
<td>Natural User Interfaces (touch/gestures, tablets, smartphones)</td>
<td>Mixed Reality Interfaces</td>
</tr>
</tbody>
</table>

New Computing Era =
New Display Form Factor + New Input Method + New Interactions
1960s  
Command Line Interfaces (keyboard)

1980s  
Graphical User Interfaces (mouse)

2000s  
Natural User Interfaces (touch/gestures, tablets, smartphones)

2020s  
Mixed Reality Interfaces

1968  
Engelbart & English MOAD

1983  
Microsoft Mouse / Apple Mouse
1960s

Command Line Interfaces
(keyboard)

1980s

Graphical User Interfaces
(mouse)

2000s

Natural User Interfaces
(touch/gestures, tablets, smartphones)

2020s

Mixed Reality Interfaces

1985

MultiTouch Tablet

2007

Apple iPhone

A multi-touch three dimensional touch-sensitive tablet.
In Proc. of the ACM CHI '85.
“The future is already here – it’s just not evenly distributed yet.”

William Gibson
What makes for compelling Mixed Reality interfaces?
What makes for compelling Mixed Reality interfaces?
Compelling MR interfaces are **adaptive**
IllumiRoom: Peripheral Projected Illusions for Interactive Experiences.
IllumiRoom

Jones, Benko, Ofek and Wilson, CES Las Vegas and ACM SIGCHI 2013
Segmented Focus
Appearance
Lighting
Magic of MR interactions happens when they are tightly coupled to the user’s environment.
Context

**environment**
(space geometry, object semantics, people around, ...)

**task**
(communication, navigation, calendar, ...)

**user actions**
(gestures, body pose, bio-signals, ...)

**user’s mental state**
(emotional, mental load, cognitive focus, ...)

Context not known at design time.
Person

Computer

Reality

World

HCI

MR

Sensing
Compelling MR interfaces are adaptive.
Compelling MR interfaces are adaptive, believable.
In MR, we are obsessed with creating a rich sense of reality!
Deep Appearance Models for Facial Rendering

STEPHEN LOMBARDI, JASON SARAGIH, TOMAS SIMON, YASER SHEIKH
Facebook Reality Labs
For interactions, realistic is not always better
Faces or vase?
Straight or crooked?
Moving or static?
Think about MR interfaces as perceptual illusions that give the user a believable experience!
Believable ≠ Realistic
Example...
But, passive haptics don’t scale!
Haptic Retargeting
Dynamic Repurposing of Passive Haptics for Enhanced Virtual Reality Experiences

Mahdi Azmandian, Mark Hancock
Hrvoje Benko, Eyal Ofek, Andy Wilson
Microsoft Research

SIGCHI 2016

Leverage the dominance of vision to retarget people's hand motions.
Body Warping

The Rendered Body Shifts to The Right
World Warping

physical cube

30°

to

virtual cube

The World Also Rotates
(At Different Rate)

90°
Putting it all together...
We can even estimate the mass effect!

\[ m_{\text{per}} = \frac{0.08 \text{kg}}{0.82 + 0.18\lambda} \]

Focusing on “as real as possible” designs can lead to sub-optimal MR experience. Design for BELIEVABILITY, not REALISM.
Compelling MR interfaces are adaptive, believable.
Command Line Interfaces (keyboard)       Graphical User Interfaces (mouse)       Natural User Interfaces (touch/gestures, tablets, smartphones)       Mixed Reality Interfaces

1960s                                        1980s                                        2000s                                        2020s

- Location fixed
- Precise and accurate inputs
- Sensing poor
- Explicit (command driven)

- Mobile
- Imprecise and noisy inputs
- Sensing rich
- Implicit (context assisted)
How to deal with imprecise, noisy, but sensing-rich inputs?
Compelling MR interfaces are adaptive, believable, and computational.
Can you type on a phone keyboard?
Probabilistic Phone Touch Keyboard

Keyboard geometry model
+ Touch precision model
+ Dictionary model
+ Language model

+ N-best list UI for error correction

+ Gesture model
Smart virtual keyboard can be better than a physical keyboard

Entry Rate Results

Plot shows participant $q_1$, median and $q_3$ (sorted by median) entry rates as well as lumped condition $q_1$, median and $q_3$ entry rates. Only entries where error rate < 10%.

A. 2 fingers, mid-air
B. 2 fingers, on-surface
C. 10 fingers, mid-air
D. 10 fingers, on-surface
2 Finger VR Typing at >100 WPM

Stimulus: How are things with you?
The role of MR Interaction Designer

- Iterative Design
- Sensor Fusion
- Machine Learning
MR Interaction Pipeline

Understand Context

- Eye
- Controller
- Microphone
- Brain
MR Interaction Pipeline

Understand Context → Infer Goals

[Images of sensors and brain]
MR Interaction Pipeline

1. Understand Context
2. Infer Goals
3. Adapt Interaction
MR Interaction Pipeline

1. Understand Context
2. Infer Goals
3. Adapt Interaction
4. User Input
MR Interaction Pipeline

1. Understand Context
2. Infer Goals
3. Adapt Interaction
4. User Input
Learning MR UI Policies from Gaze Data

Trained RL agents to predict when an MR label is meaningful to the user.

**Context:** User’s gaze behavior + task + environment

**Output:** Inferring task-specific goals + reduced clutter of MR labels

https://ait.ethz.ch/label-agent/

Gebhardt et al. “Learning Cooperative Personalized Policies from Gaze Data”
ACM UIST 2019
Key Challenge = Missing Data
Replica: 3D Spaces and Object Semantics

https://github.com/facebookresearch/Replica-Dataset

Habitat: A Platform for Embodied AI Research

facebook Artificial Intelligence

https://www.aihabitat.org/
EPIC-Kitchens Dataset

Damen et al. Scaling Egocentric Vision: The EPIC-KITCHENS DATASET. ECCV2018
https://epic-kitchens.github.io/2019
Call to action!

Collect and release datasets of MR interactions with different:

- Environments
- Objects (both real and virtual)
- Tasks
- People

Including eye-tracking, bio-signals, hand interactions, body movements, etc.
Summary
New Computing Era =
New Display Form Factor
+ New Input Method
+ New Interactions
Compelling MR interactions are adaptive, believable, and computational.
Design interactions that adapt to the user’s actions, the world around them, and the context of use.
Focus on interaction believability. Reality is overrated!
Harness the computational methods to overcome uncertainty, scale, noise, and enable personalization.
MR Interaction Pipeline

1. Understand Context
2. Infer Goals
3. Adapt Interaction
4. User Input
Thanks to all my collaborators!

Hrvoje Benko
benko@fb.com
Facebook Reality Labs
We are hiring!