Korea - New Zealand Joint Research Workshop

Recent works on Wearable Interaction and Usability

A Way to Enrich Physical Immersion?

Gerard J. Kim (김정현)

Digital eXPerience Laboratory

Korea University

Wearable computers (Witt, 07)

- Limited Capabilities: The wearable computer system is often very limited or constrained in terms of available computation power, energy consumption, and available I/O modalities (Smart phones?)
- Portable / Mobile / (Self contained?)
- Physical constraints: Enables hands-free or hand-limited use
- Adapted Interaction:
  The wearable computer system may automatically adapt the interaction style and/or interface rendering of a application, contents, environment in order to make interaction easier and more efficient while minimizing mental effort

Some previous and on-going works

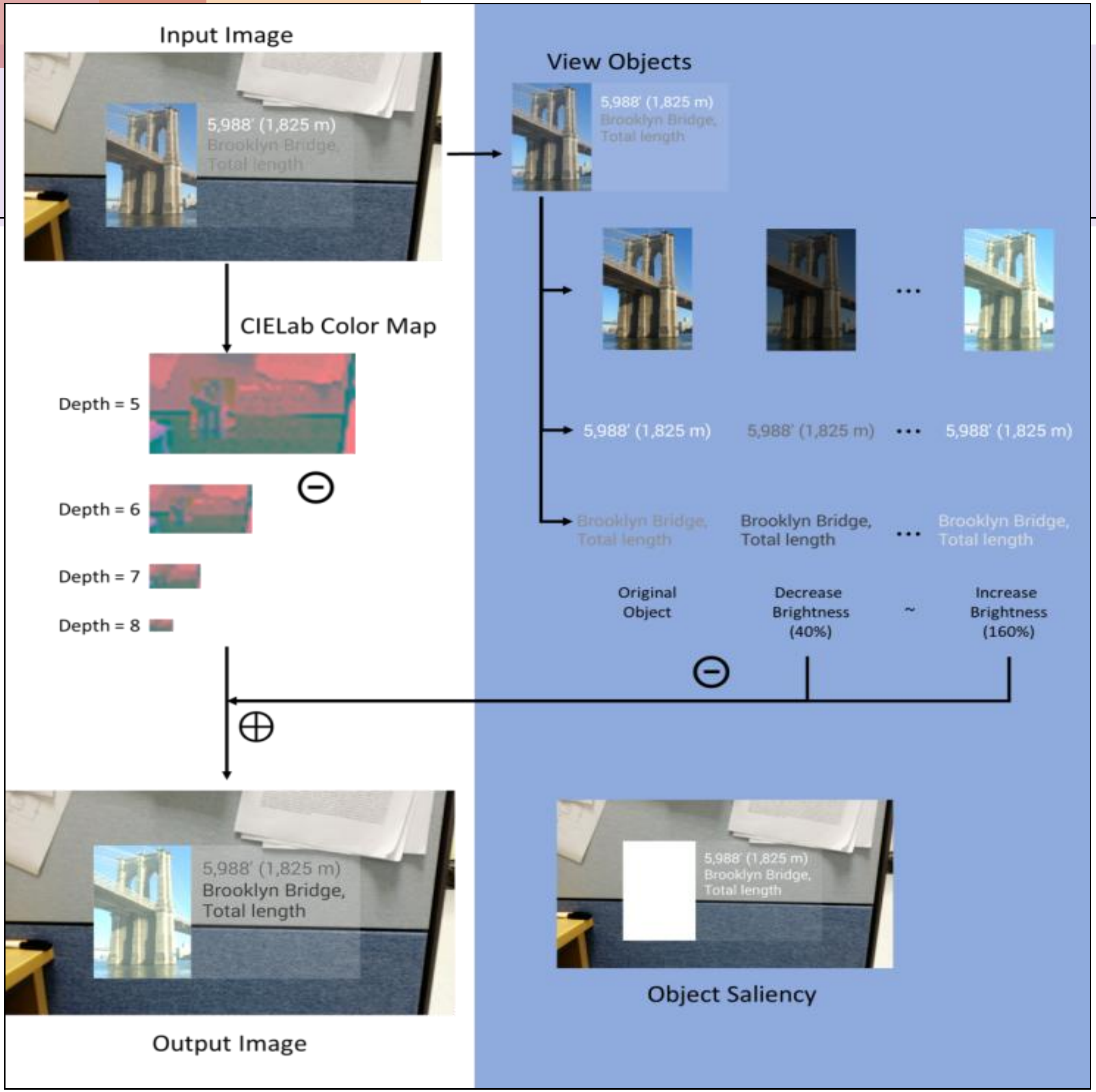
- Display
  - Readability of augmented information
  - Google glass study: Dual focus problem
  - (Richer tactile feedback)

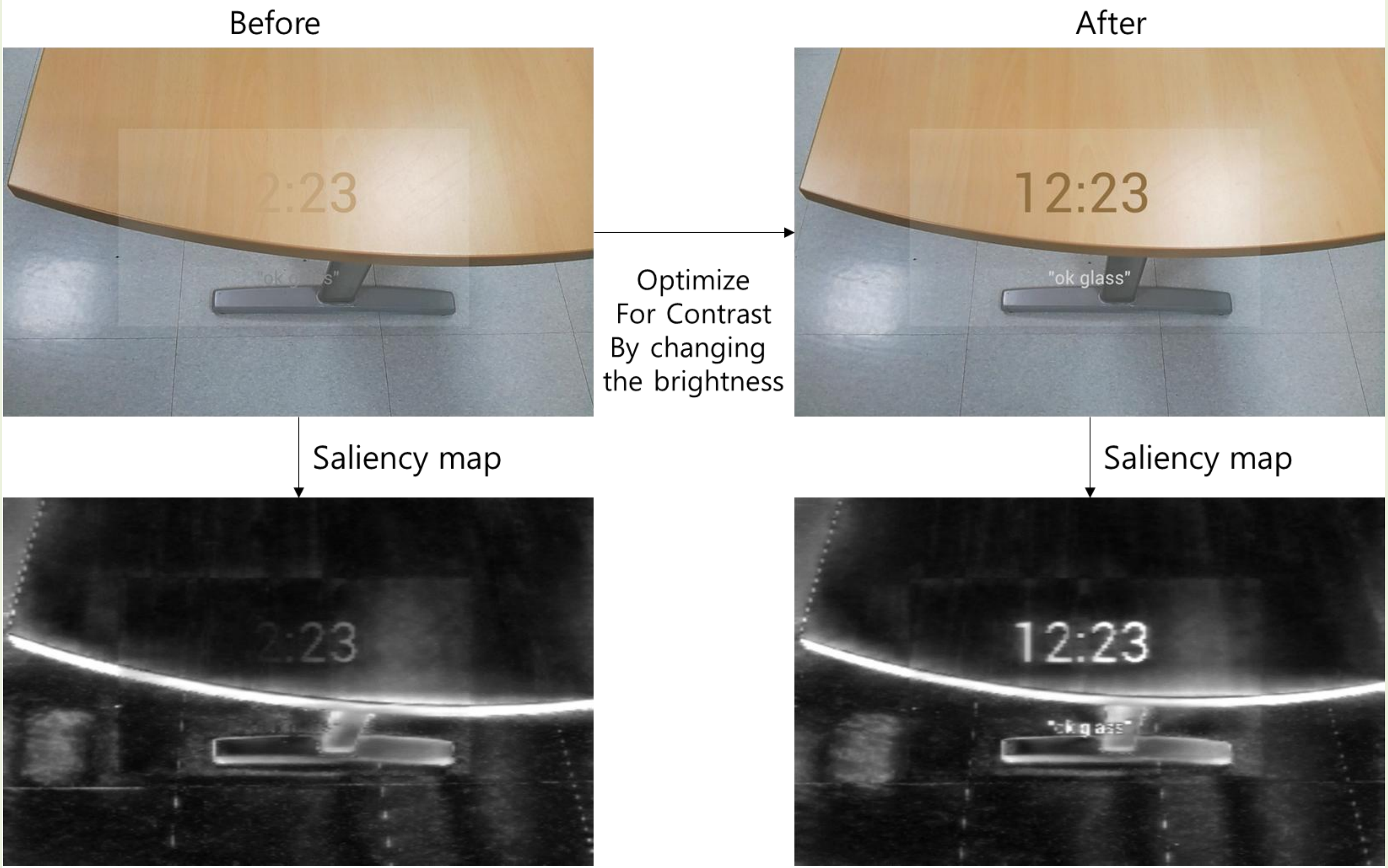
- Arm/Hand input
  - Mid-air gestures: Performance model
  - Hand-augmented interface

- Body-based input
  - (Body based metaphors)
  - EMMI: Extreme body movements
  - Pressure based whole body interface: G-bar
  - QUI: Quadruped user interface
  - Body based robot control Sprint

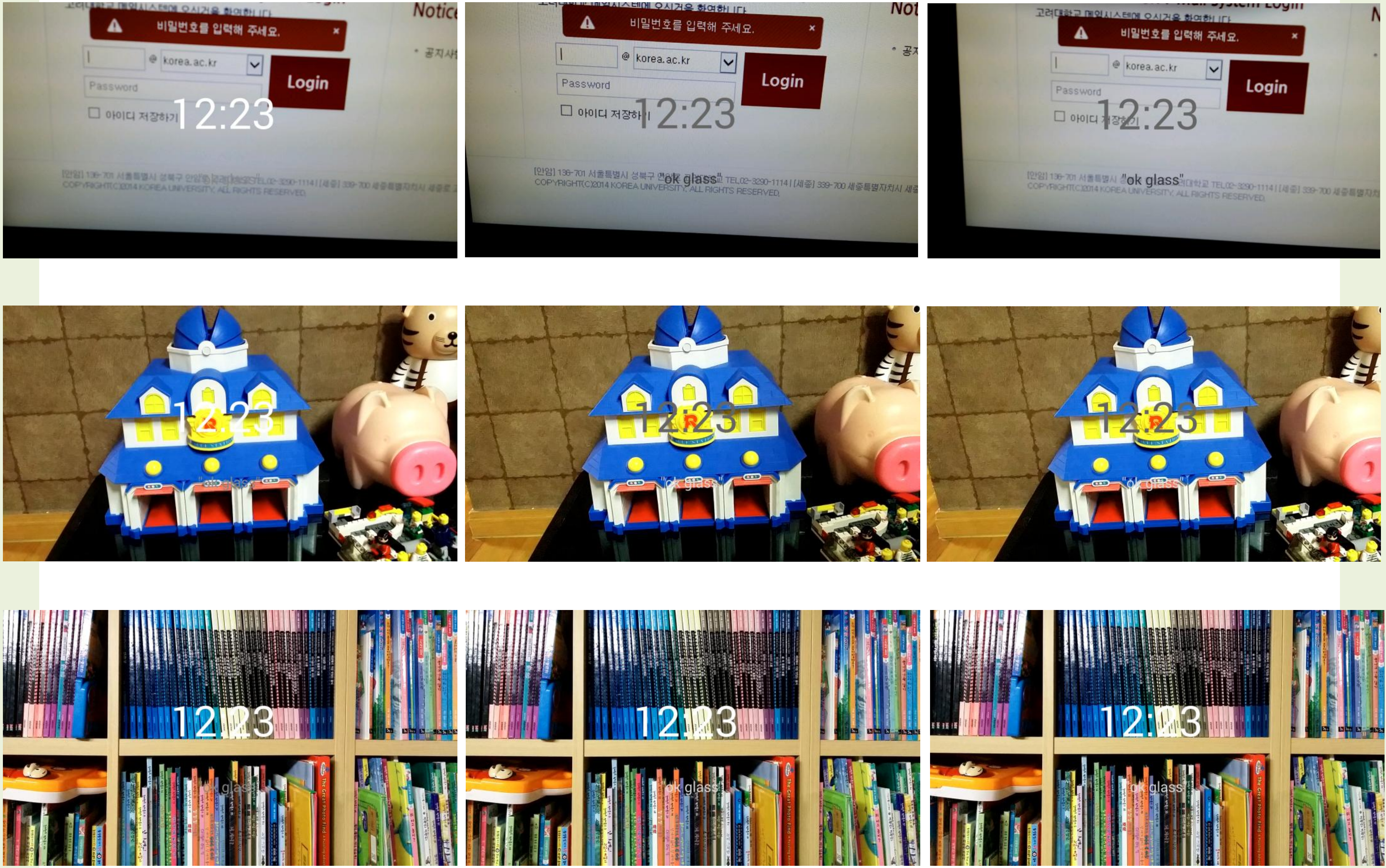
Readability for see-through contents

- Motivation
  - Dynamic background
  - Eye fatigue:
    - Lack of sufficient contrast (luminance/chromatic)
    - Dual focus
    - Visual acuity (center vs. periphery)
- Approach
  - Analyze "conspicuity (or saliency)" of augmented region
  - Adjust brightness of augmentation to enhance conspicuity
  - Mobile implementation
    - Object level (vs. pixel level) analysis and adjustment

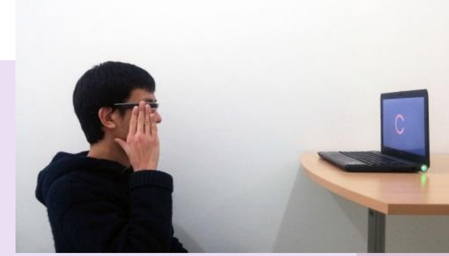

Input Image
5,988' (1,825 m)
Brooklyn Bridge,
Total length
CIELab Color Map
Depth = 5
Depth = 6
Depth = 7
Depth = 8
View Objects
Original Object
Decrease Brightness (40%)
~
Increase Brightness (160%)
Output Image
Object Saliency


Before
After
2:23
12:23
"ok glass"
Optimize
For Contrast
By changing
the brightness
Saliency map
Saliency map

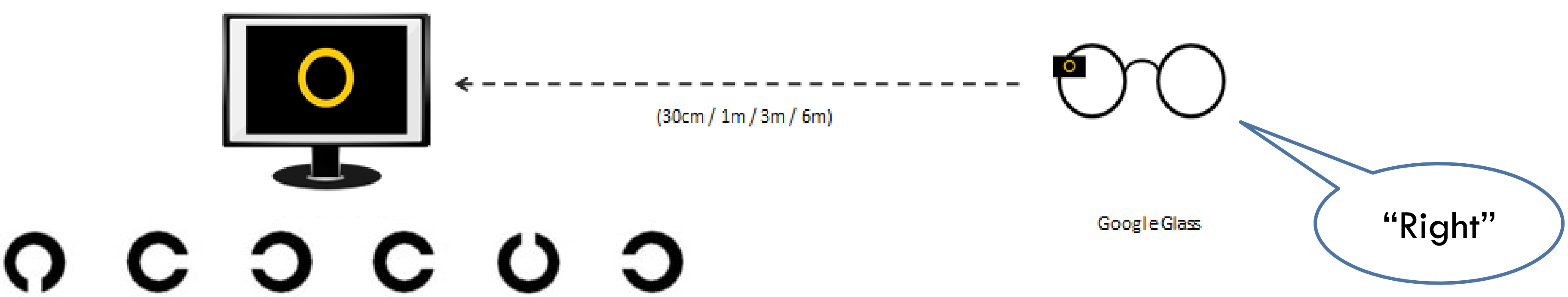
Without adjustment | Our approach | Ground truth



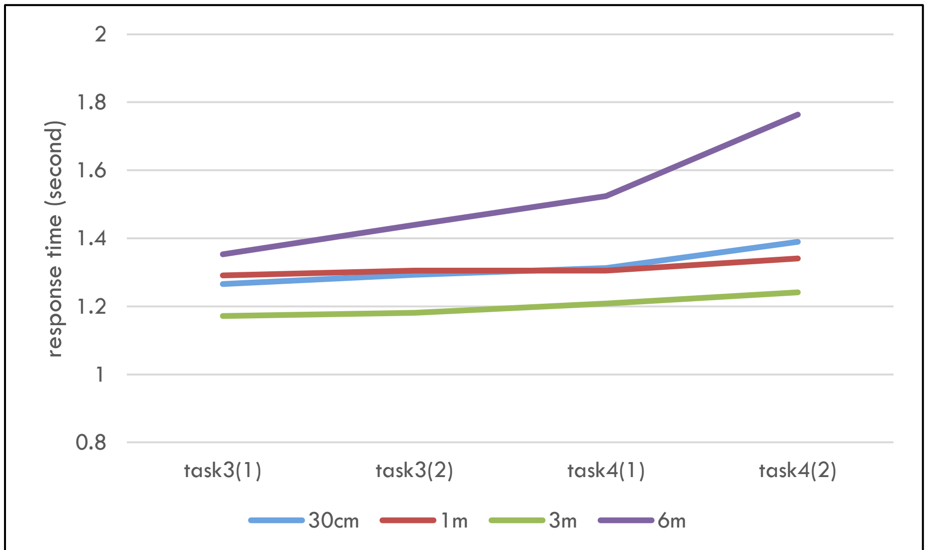
Eye fatigue assessment for see-through glasses

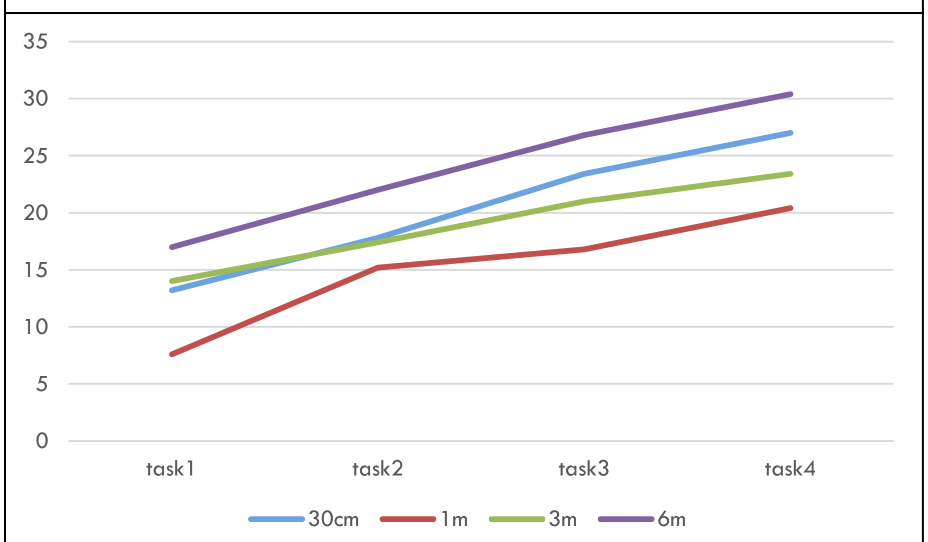


- Motivation
  - Quantify and establish an eye fatigue/stress (due to refocusing)
  - Later use it as a guideline in glass-based application design (or even glass design itself)
    - E.g. Object recognition and overlay
- Basic experiment
  - "Refocus and identify" task at different lengths using google glass
  - Measure fatigue indirectly by response time and survey

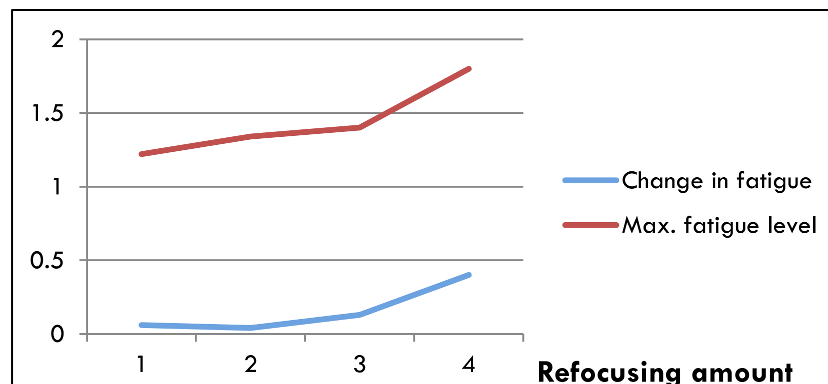


- Notes
  - Google glass designed to focus on a 25 inch virtual screen 2.4 m away
  - Obviously, longer the refocusing distance, more tired it will be
    - Interested in form of the function (linear, exponential, saturating, step, …?)





- Response time change in time
  - 6m > 30cm > 3m ~ 1m
- Survey on "eye/refocus" fatigue
  - 6m > 30cm > 3m ~ 1m
  - Same as above

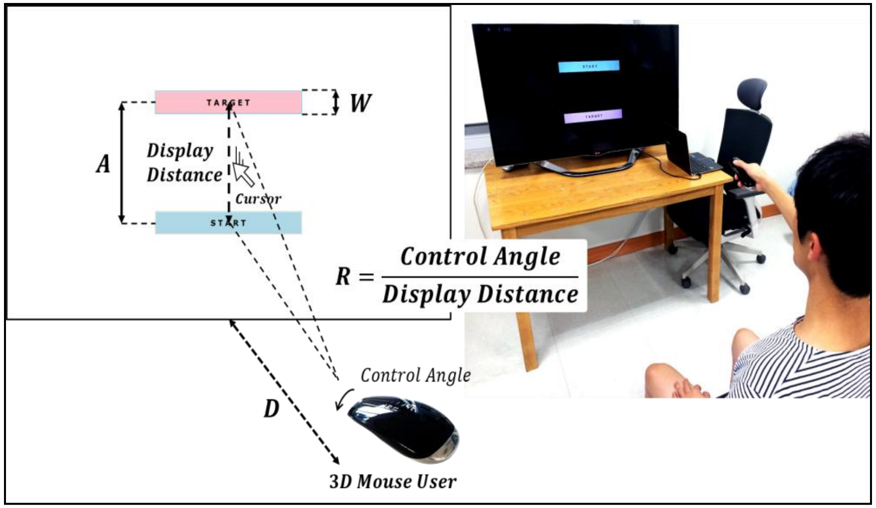


Extended Fitt's Law for Mid-air Gestures

- Motivation
  - Wearable/Hand-held devices will use gestures heavily in the future
  - Validate if traditional Fitt's Law applies, or extend it if necessary
  - Use as guideline in wearable/remote interaction design

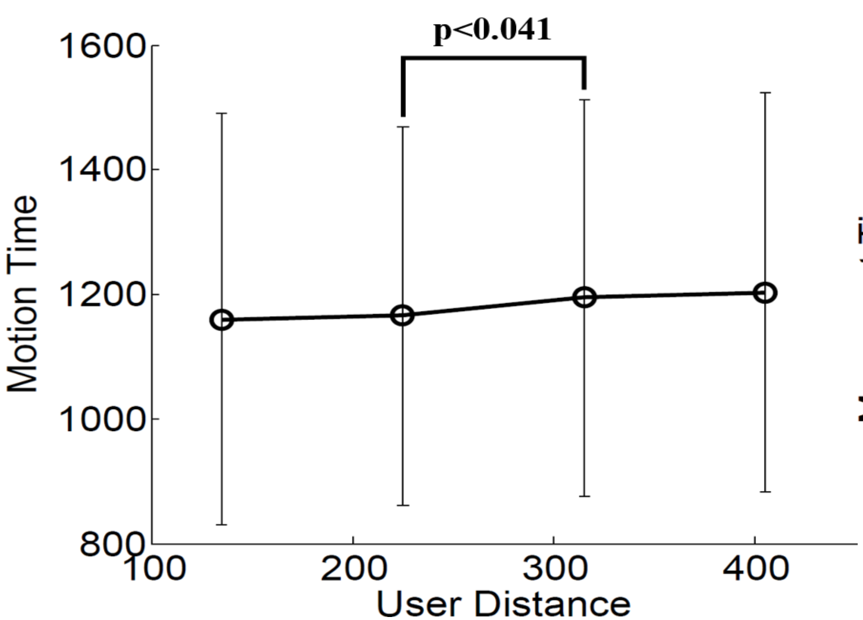
- Basic experiment
  - Hypothesis: Performance will depend on:
    - Object width (W) and Movement distance (A) +
    - Viewing/distance and C/D ratio
  - Logitech MX Air Mouse

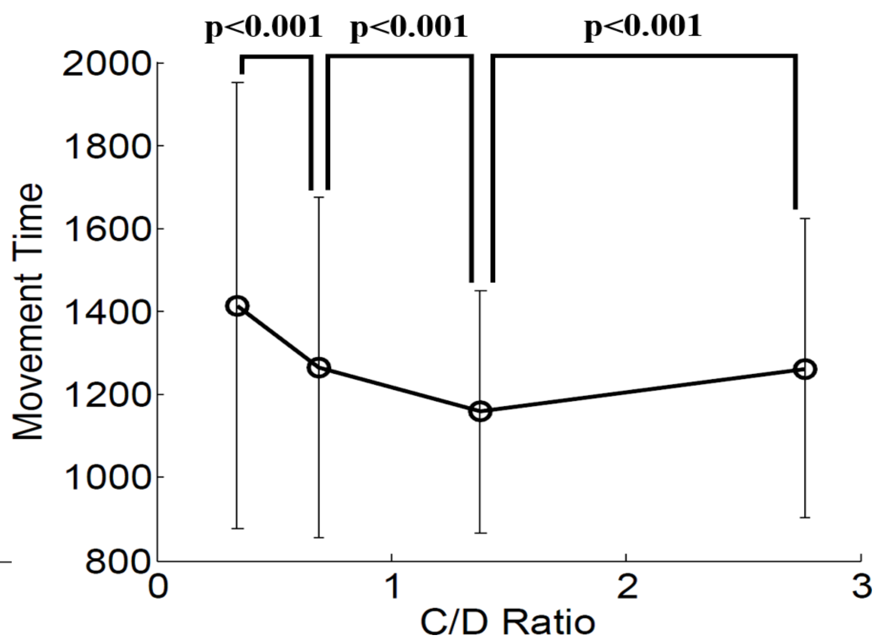
Factor	Levels
Movement Distance (A)	10.41, 15.62, 23.43 cm
Object Width (W)	4.63, 6.94, 10.41 cm
User Distance (D)	135, 225, 315, 405 cm
C/D Ratio (R)	2.76, 1.38, 0.69, 0.34 °/cm



Collect performance data for different combinations and fit a function

$$MT = a + b\log_2\left(\frac{A}{W}+1\right) + c\tan^{-1}\left(\frac{(D-D_{ref})^2}{D_{ref}^2}\right) + d\left(R-R_{ref}\right)^2$$

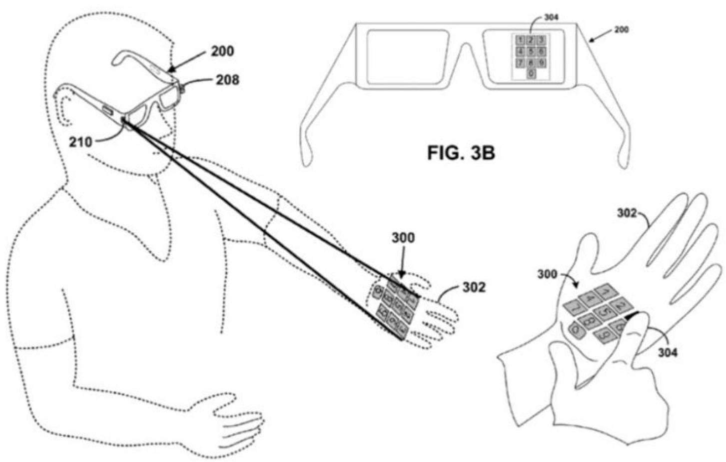




Model	a	b	c	D	RMS	R^2	$R^2_{(adj)}$
Fitts [5]	710	299	N/A	N/A	88.34	0.718	0.714
Welford [15]	399	232	N/A	N/A	102.3	0.621	0.616
Bi [3]	688	276	N/A	N/A	99.72	0.64	0.635
Kopper [8]	710	298	N/A	N/A	88.41	0.717	0.713
Ours	578	299	35.9	139	34.50	0.958	0.956

Hand augmented AR interface

- Depth sensor on the glass (near future?)
- Use hand/palm as a tangible surface on which interaction occurs
  - Tracking hands/palms/fingers/finger parts
  - Low end mobile platform - no learning / no IK



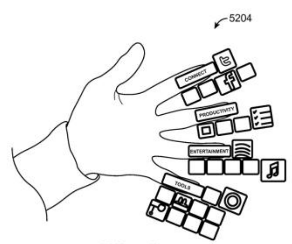
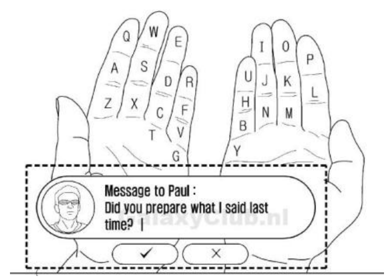
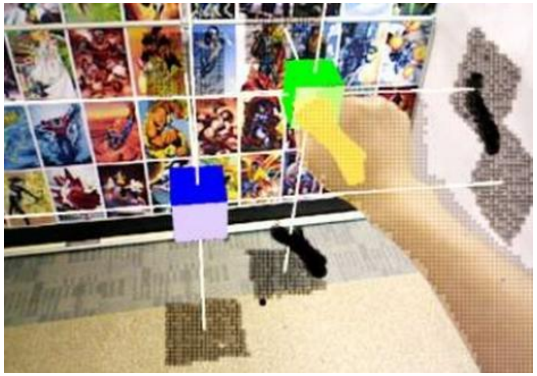


FIG. 52b





Skin Color Detection

HSV Color Space Conversion

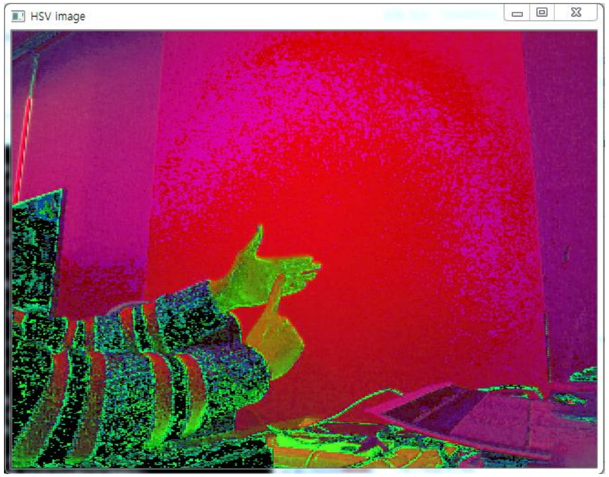
HSV 색상 공간으로 변환 후 피부색에 해당하는 영역의 마스크를 생성함

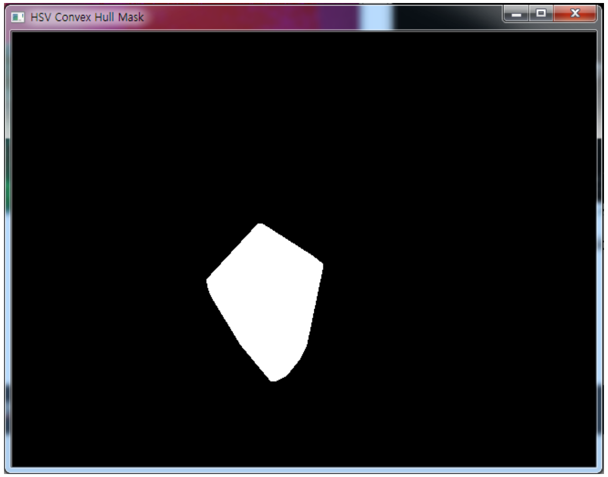
Contour

HSV 색상 공간에서 생성한 마스크에서 나타나는 영역의 Contour를 구하고 각 Contour의 크기 및 면적 등을 이용하여 필터링 함

Convex Hull

선택된 Contour를 이용하여 각 Contour를 둘러싸는 Convex Hull을 구하고 이를 이용하여 마스크를 생성함





Depth Map processing

Range Filtering

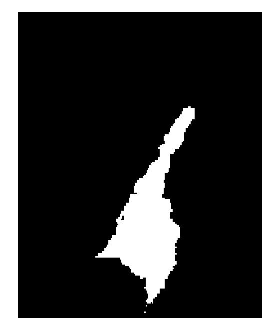
Depth 센서의
거리값을 이용하여
50cm~95cm 범위 내의
값만 사용하도록 함

Masked Depth Map

RGB 영상을 이용한
Convex Hull 마스크와
필터링된 Depth Map을
이용하여 손 영역으로
추정되는 영역 추출

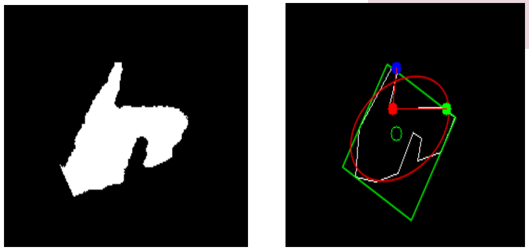
Classify Hands

RANSAC 알고리즘을
이용하여 스크린이
되는 손과 입력하는
손을 구분함

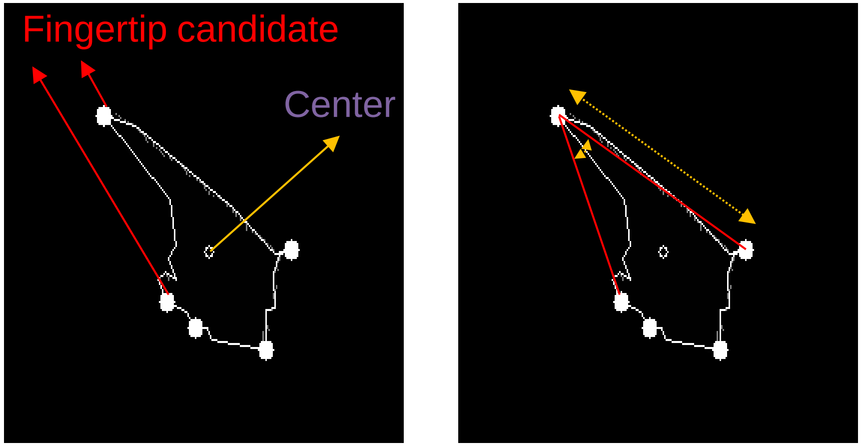


Palm Orientation and Interaction Fingertip

- Palm
  - Z: normal to the palm plane
  - Estimate thumb and index finger from the convex hull
  - Calculate relative rotation to the reference (palm in right angle to the sensor)

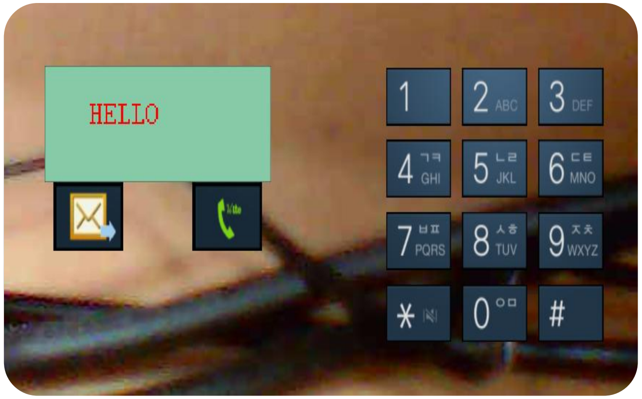


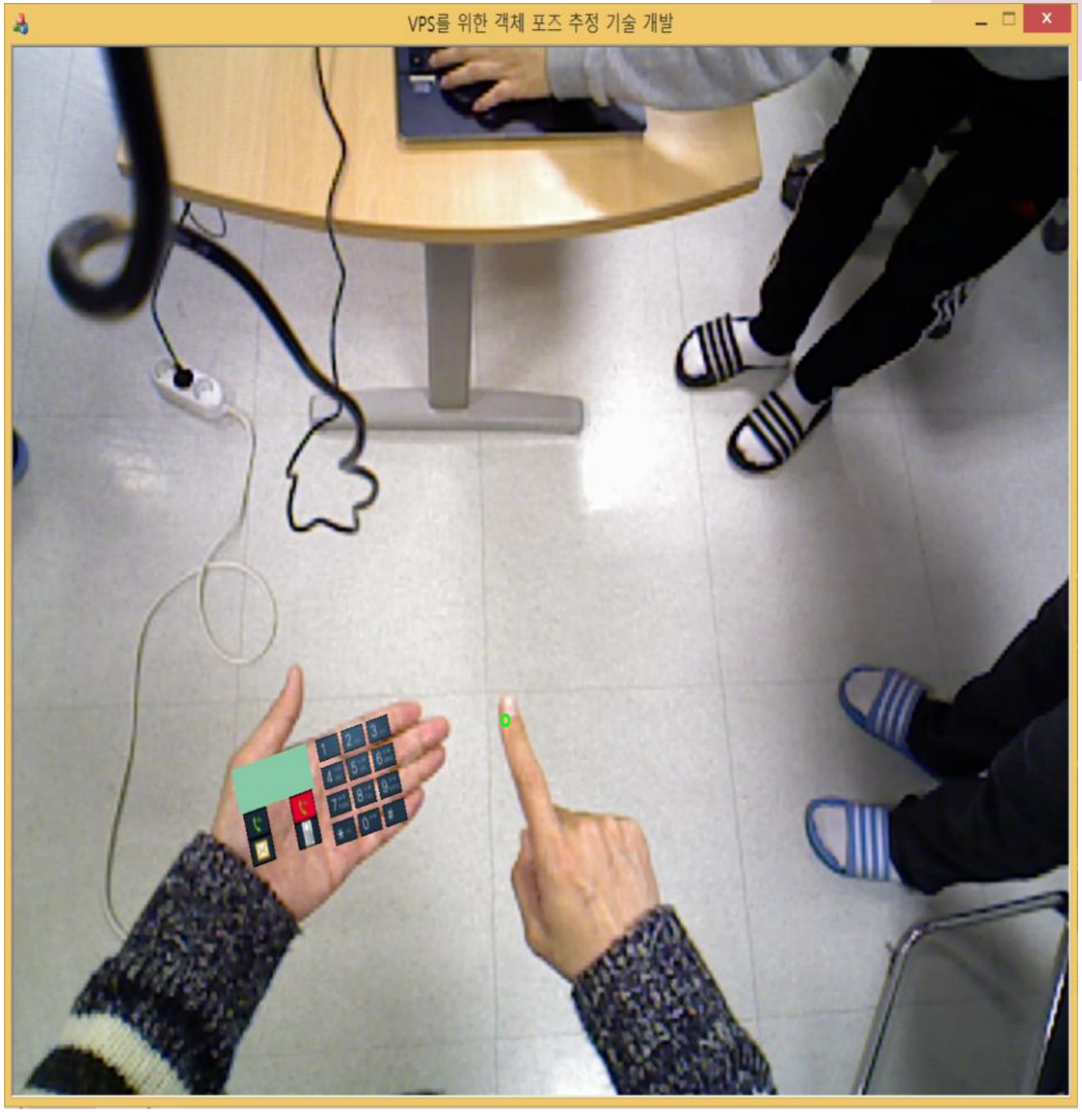
- Interaction fingertip
  - Noise filtering + convex hull and consider relative angles among the vertices
  - Event press detection
    - Threshold distance between fingertip and palm (< 3cm)



UI Application

movie





Body based interfaces (circa 2004)

Q: How can we interact using our body parts (other than just hands) ?

Body inspired metaphors
(easy to remember)

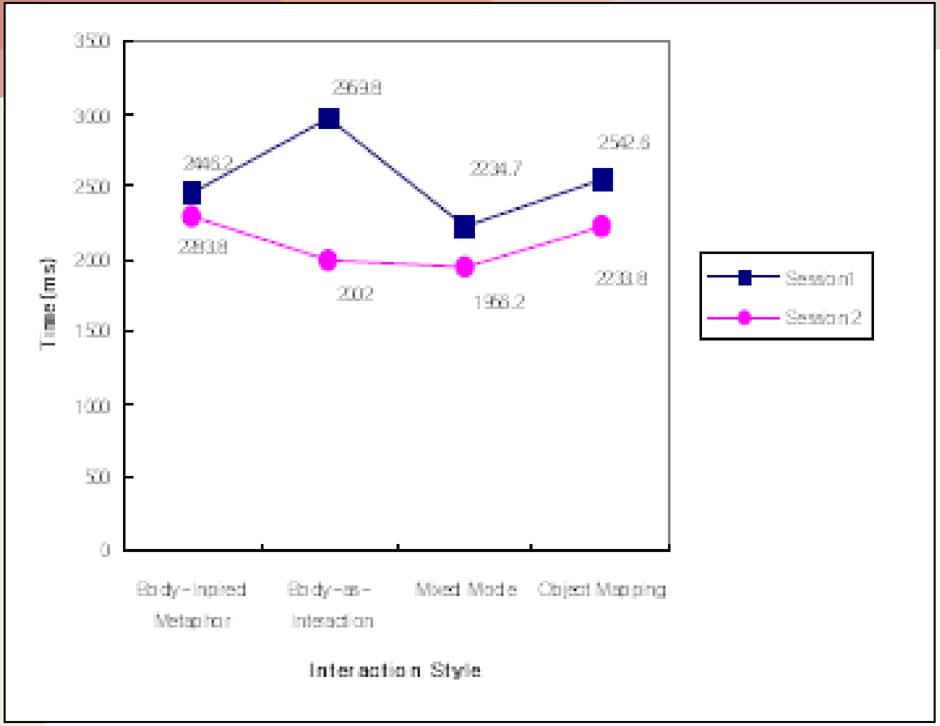
Body as interaction surface
(ergonomic)

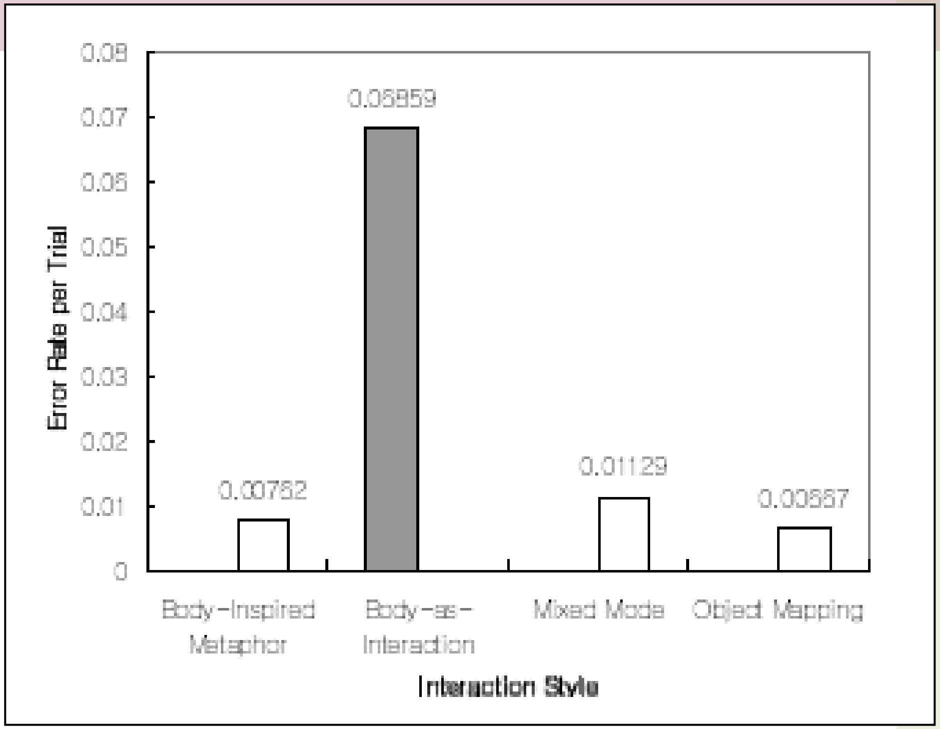
Proprioceptive

Body based metaphor (circa 2004)

Body Parts / Clothing	Original Semantics	Applicable Tasks
Eye(s) / Eye Glass	Seeing Window	presentation, check emails, viewers, turn on TV, activate window,
Mouth	Eating Speaking Blowing	beep, take items (in games), coloring (by blowing)
Ear(s) / Ear Rings	Hearing Collect Info	play sound, increase volume, join in mail list
Head	Importance Thought	homing, bring out notepad
Hand(s) / Fingers	Hold	temporary storage,
Leg(s)	Locomotion Support	navigation, constraining
Skeleton	Hierarchy Structure	avatar control, data traversal
Buttons	Switch / Key Open/Close	activation
Pockets	Containment Safe	folders, trash can, file transfer

Application	Tasks	Body-Inspired Metaphor
Powerpoint	Start/End	Touch the Eye
	Next Slide	Right Shoulder
	Prev. Slide	Left Shoulder
MP3 Player	Start/End	Touch the Ear
	Play	Touch the Mouth
	Stop	Touch the Chest
	Next Song	Right Shoulder
	Prev. Song	Left Shoulder
Desktop Manager	Open Folder	Touch a Pocket
	Log in	Touch the Back Pocket

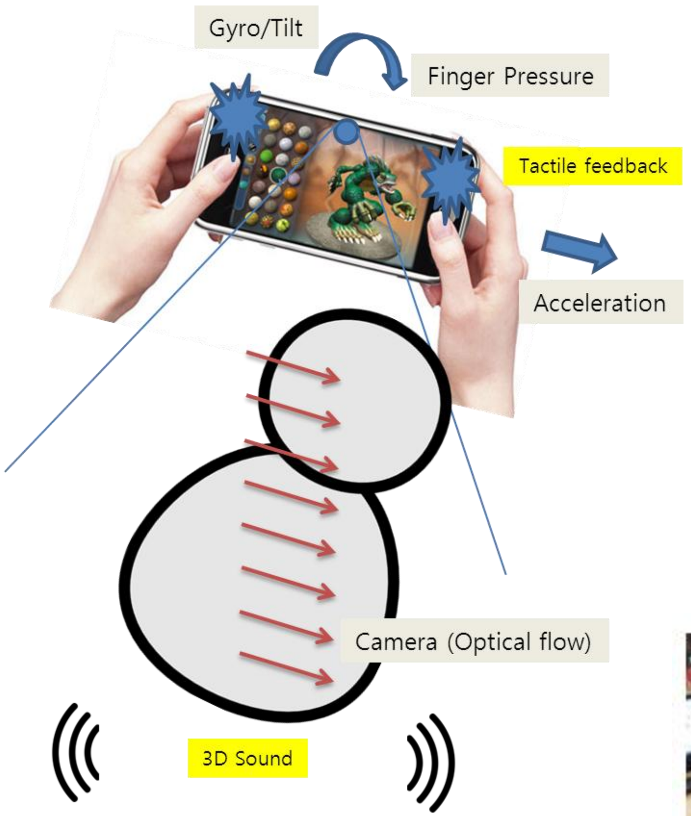




	Folder	Control Panel	Power point	MP3 Player	Wallet
Eye	1		13 (87%)	1	
Ear				12 (80%)	
Mouth			1	2	
Head	1	14 (93%)			
Pocket	13 (87%)			1	1
Back Pocket	1				14 (93%)

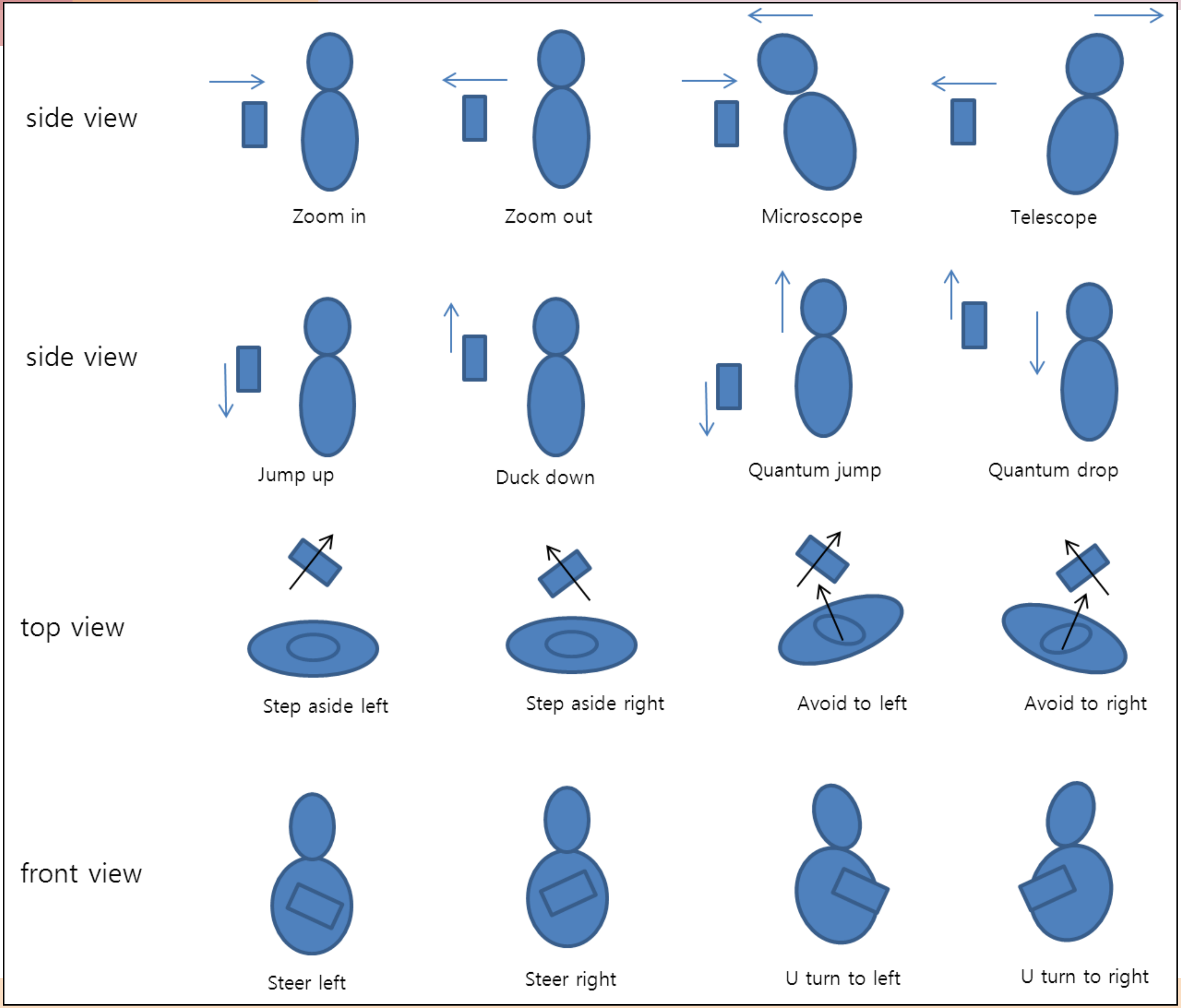
EMMI: Extreme Motion Mobile Interface

(a)

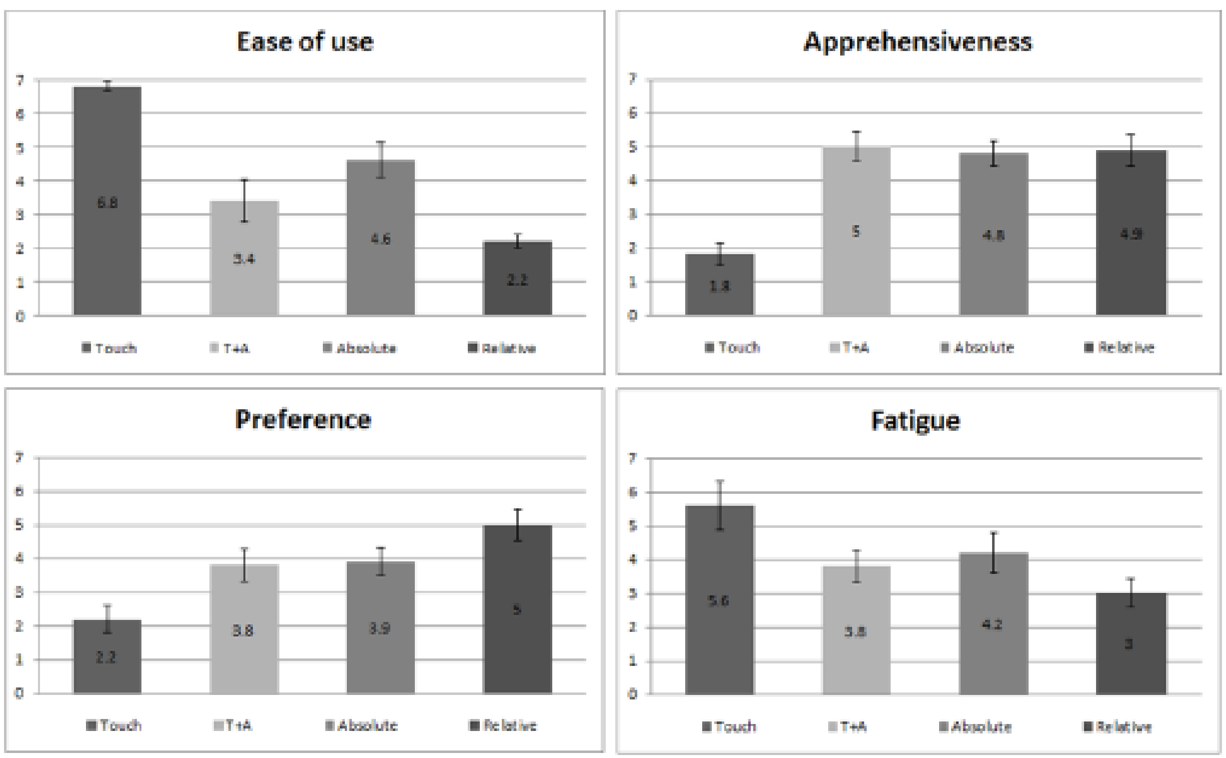


(b)

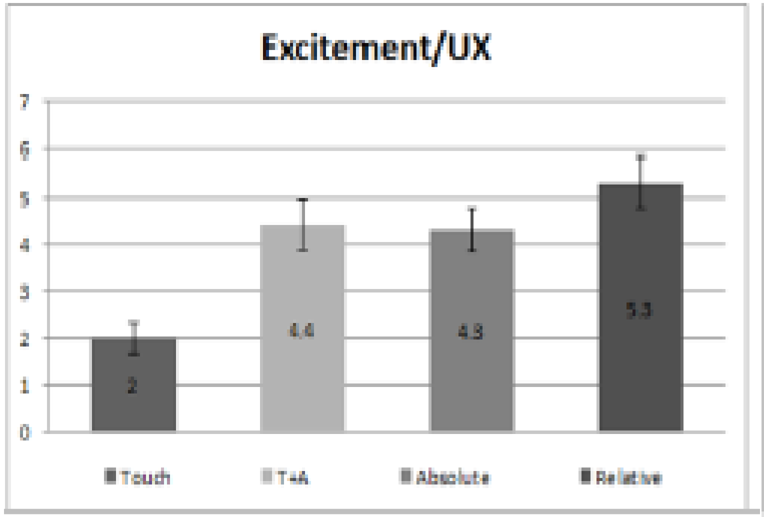
movie

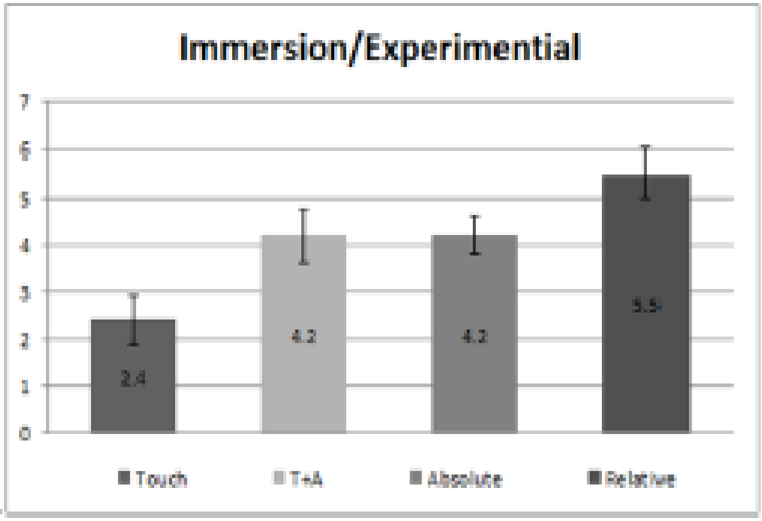

side view
Zoom in
Zoom out
Microscope
Telescope
side view
Jump up
Duck down
Quantum jump
Quantum drop
top view
Step aside left
Step aside right
Avoid to left
Avoid to right
front view
Steer left
Steer right
U turn to left
U turn to right

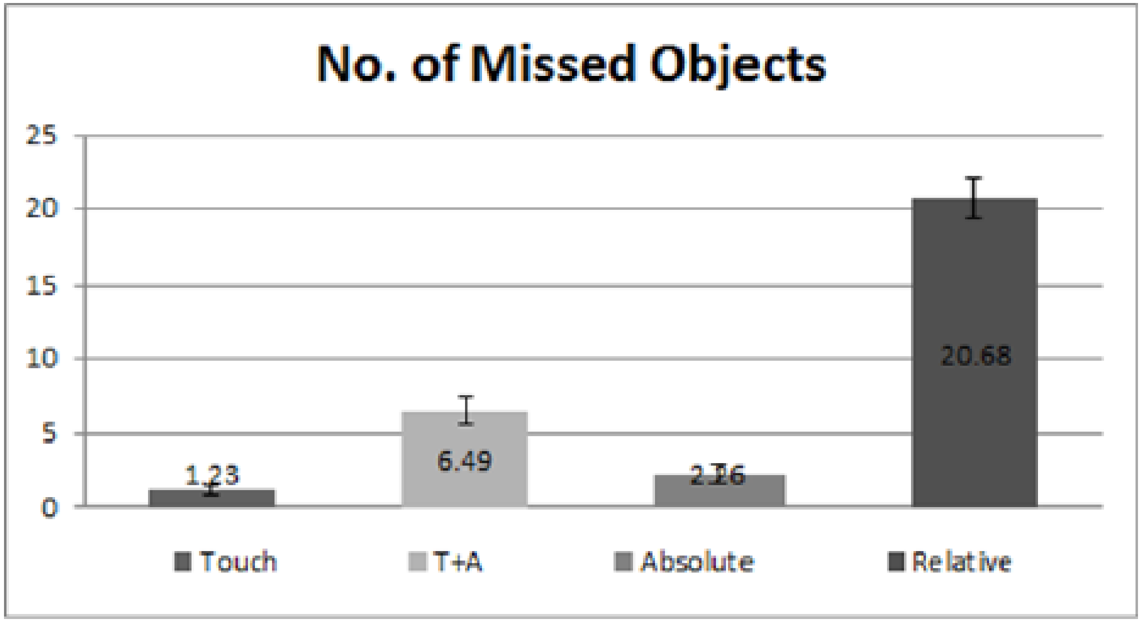
Interface type	Symbol	Explanation
Touch based	Touch	Touch the respective lane (e.g. lanes 1 and 5) to directly place the basket
Touch + Acceleration	T+M	Use direct touch and motion (detected by an acceleration threshold) simultaneously to move into lanes 1 and 5
Absolute motion by hand	Absolute	Use only hand motion (approximated by an acceleration threshold[1])
Relative motion by hand/head	Relative	Use hand motion (approximated by an acceleration[1] threshold) and head movement (approximated by the face detection / optical flow measurement from the phone camera)



Generally low usability
→
Q: Should mobile game be really experiential?


Excitement/UX
7
6
5
4
3
2
1
0
2
4.4
4.3
5.3
Touch
T+A
Absolute
Relative


Immersion/Experimential
7
6
5
4
3
2
1
0
2.4
4.2
4.2
5.5
Touch
T+A
Absolute
Relative

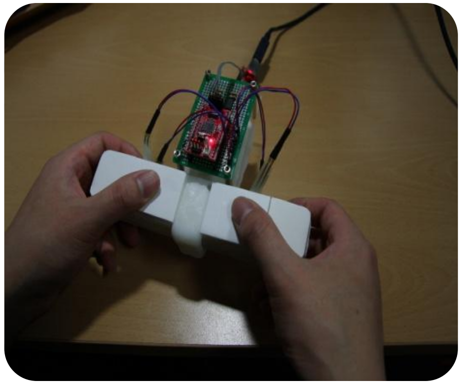

No. of Missed Objects
25
20
15
10
5
0
1.23
6.49
2.26
20.68
Touch
T+A
Absolute
Relative

G-Bar: Revisited

movie

Grounded \ Input Scheme	Isometric		Isotonic
YES	G-bar	-	-
NO	-	NG-bar	Wiimote

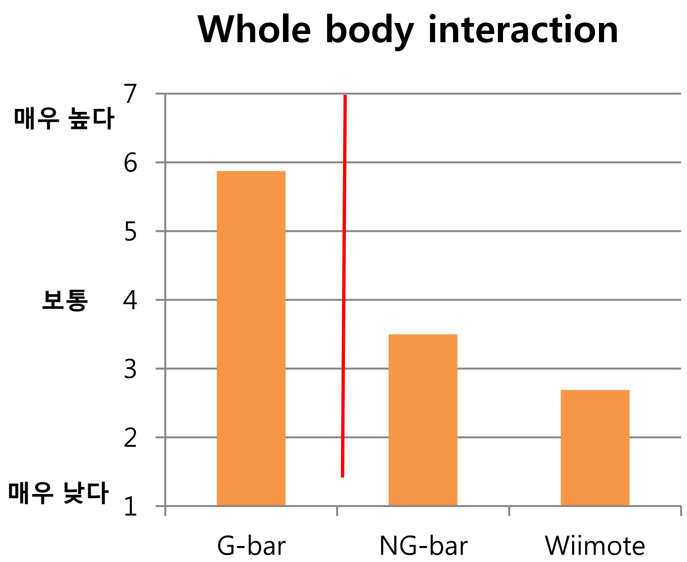
G-bar



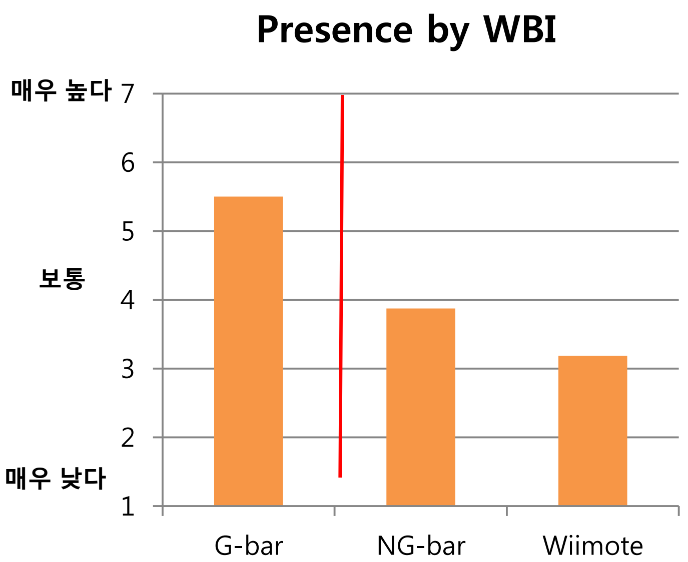
NG-bar

Wii-mote

Results (1)

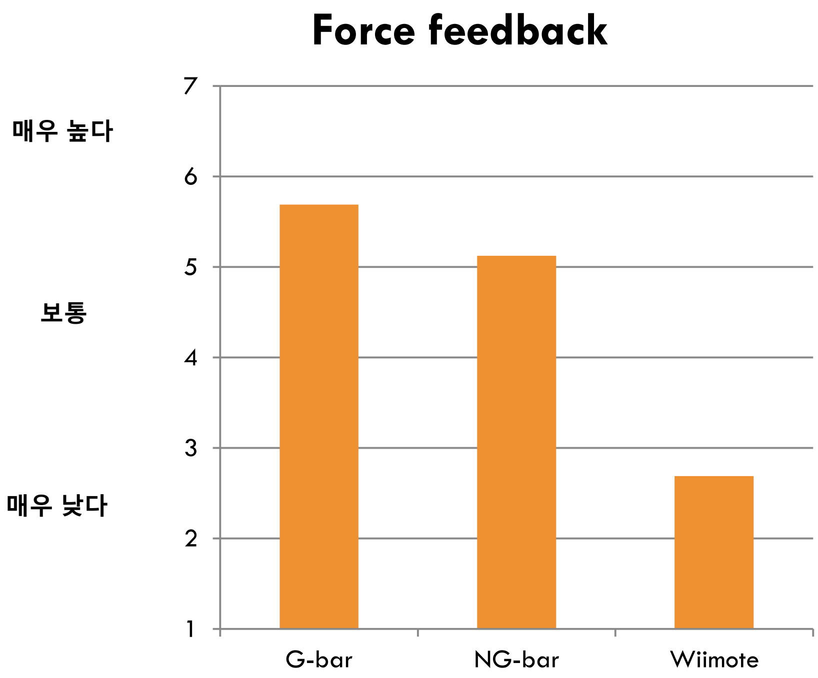


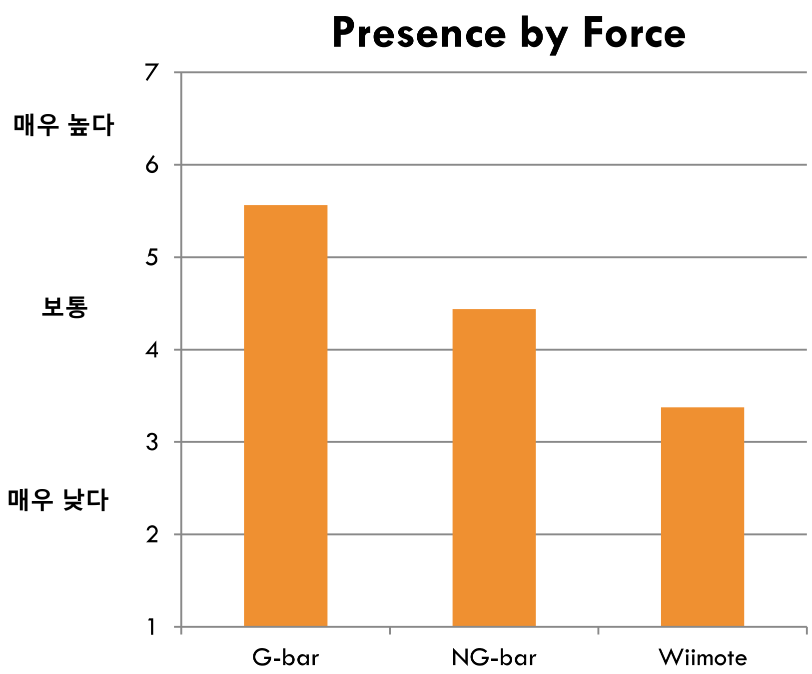
G-bar , NG-bar : $p = 0.000$
G-bar , Wiimote : $p = 0.000$
NG-bar , Wiimote : $p = 0.131$



G-bar , NG-bar : $p = 0.001$
G-bar , Wiimote : $p = 0.000$
NG-bar , Wiimote : $p = 0.223$

Results (2)



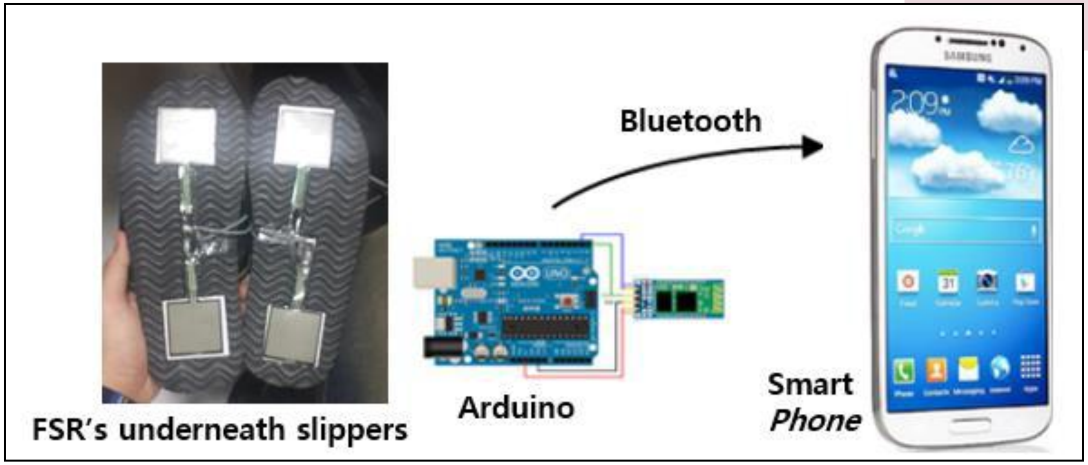


G-bar , NG-bar : p = 0.091

G-bar , NG-bar : p = 0.027

QUI: Quadruped interface

QUI: Quadruped interface



- Use both hands and feet
- Fast and simultaneous input
- Tasks
  - Typing
    - Mode change (Shift, Special characters)
    - Special keys (Backspace, Space bar)
  - Gaming
    - Shoot with hands/fingers and navigate with feet

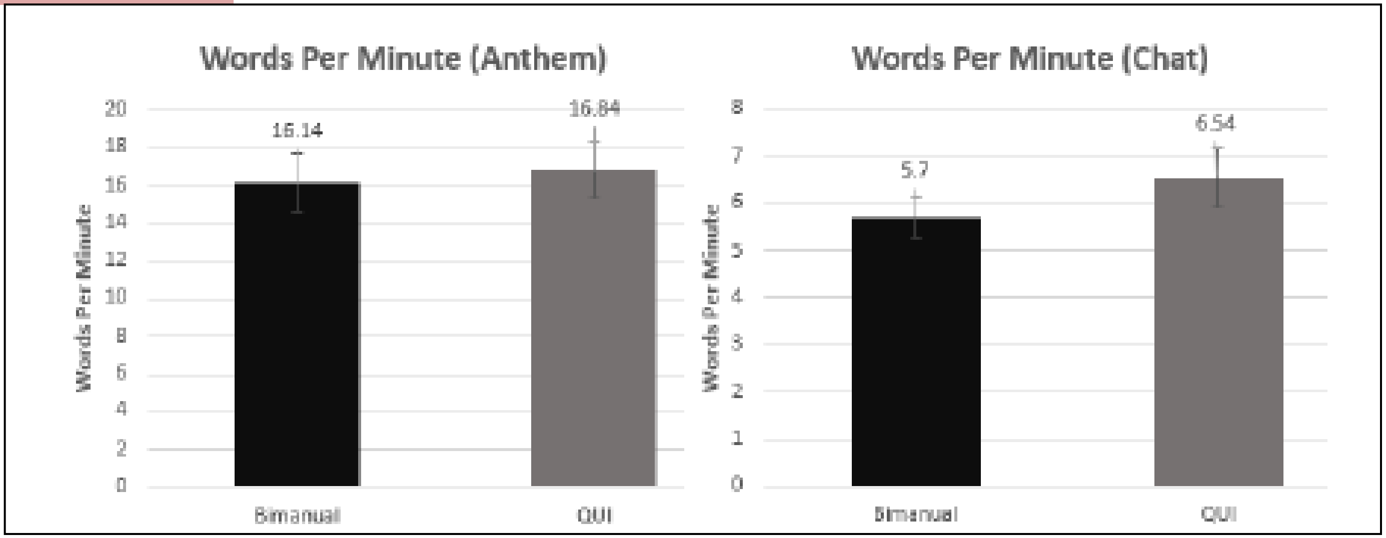
Typing/Chat

	Task 1 (Korean National Anthem)	Task 2 (Conversational Chat)
Task Characteristics	Memorized phrases (relatively easy)	Reactive / Involve special characters (more difficult)
Fingers (mapped to)	All keys in the virtual keyboard	All keys in the virtual keyboard
Left foot (mapped to)	Backspace	Switch to special character layout (like a function key)
Right foot (mapped to)	Space bar	Space bar

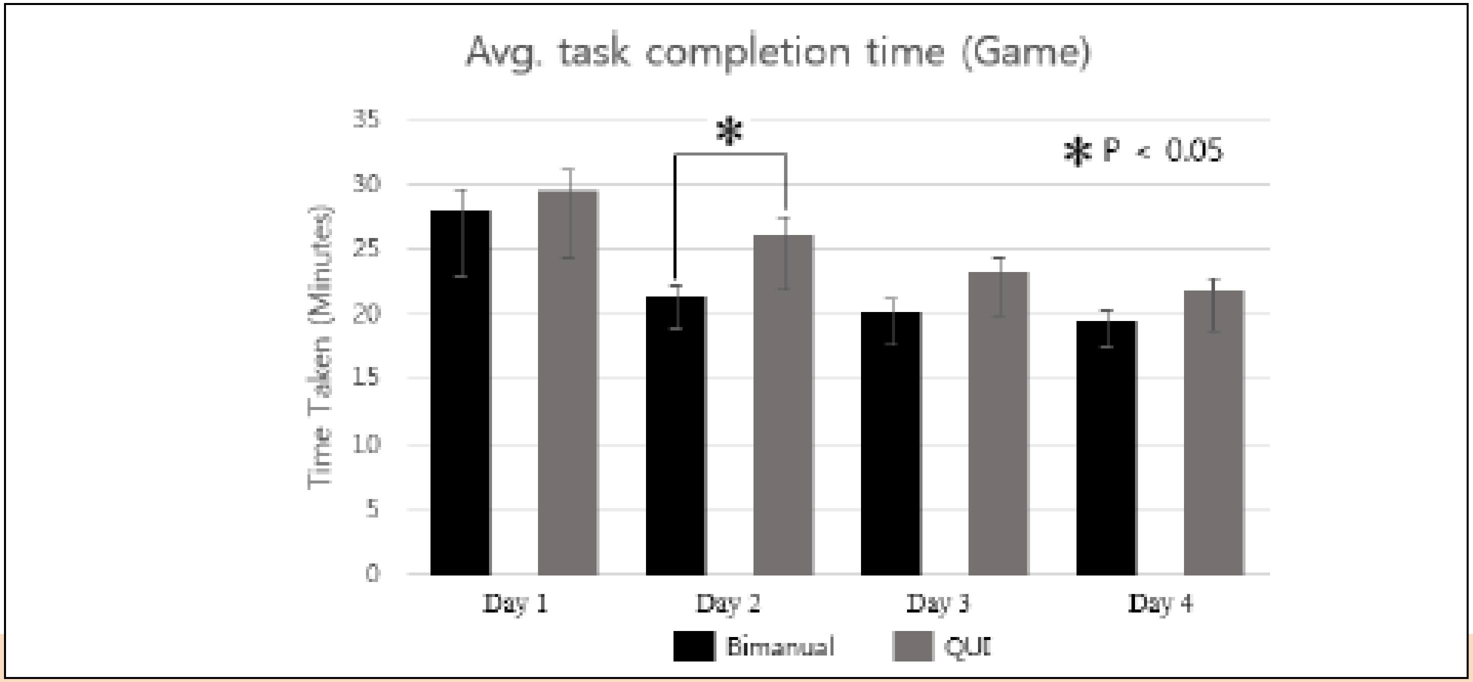
movie

Object manipulation

	Degrees of freedom / Control amount	Finger touch input	QUI
Position	2 / ± 800 pixels (left/right and up/down)	Fingers of left hand	Fingers of left hand
Scale	1 / ± 0.2% (bigger or smaller)	Fingers of right hand	Fingers of right hand
Orientation	1 / ± 30 degrees (rotate left/right)	Fingers of right hand	Left and right foot



Words Per Minute (Anthem)
Words Per Minute
16.14
16.84
Bimanual
QUI
Words Per Minute (Chat)
Words Per Minute
5.7
6.54
Bimanual
QUI



Avg. task completion time (Game)
Time Taken (Minutes)
*
* P < 0.05
Day 1
Day 2
Day 3
Day 4
Bimanual
QUI

Sprint: Body based robot control

movie

Summary and Collective Conclusion

gjkim@korea.ac.kr

Thank you