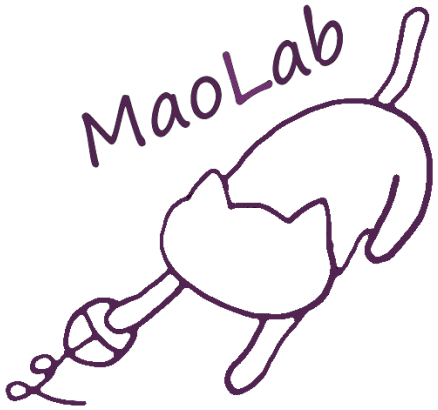
Video Visualization with Computer Vision

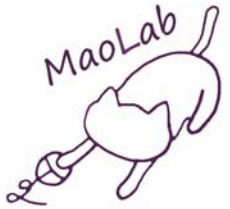
Masahiro Toyoura
University of Yamanashi

03/11/2014

NII Shonan Meeting on
Computer Visualization:
Concepts and Challenges



Vision & Visualization



Visualization for Vision

M. Toyoura, K. Kashiwagi, A. Sugiura, X. Mao, “Mono-glass for Providing Distance Information for People Losing Sight in One Eye,” ACM SIGGRAPH VRCAI, pp.39-42, 2012.

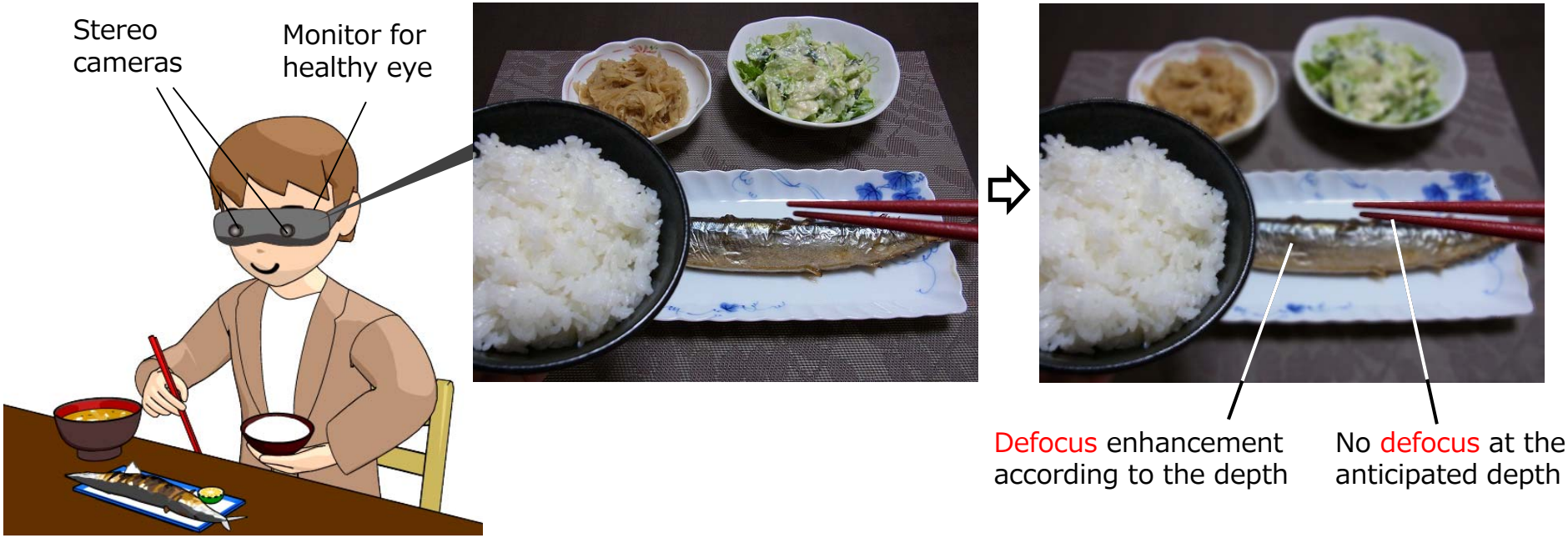
See on behalf of Humans

Vision for Visualization

M. Toyoura, S. Nishiguchi, X. Mao, M. Murakami, "Activity Visualization for Multiple Targets in a Video," Pacific Visualization (Poster), 2014.

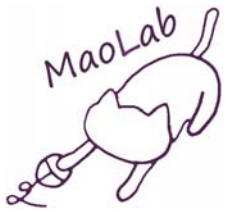
Mono-glass

Provide the function of binocular disparity on behalf of user's eyes.



⇨ Informative + Perceptual/Intuitive Vis of depth

Motivation



Cannot do binocular disparity processing with **one eye**.

⇨ It makes grasping an object difficult, and reduces QoL (Quality of Life).

The number of people losing sight in one eye:

(Not published, but) is considered to be **the same order to the visually impaired patients**.

(Poor vision: **246** million, blindness: **39** million (2010))

Design of Mono-glsss

- Depth estimation with stereo cameras
  - Relatively light, real-time process
- Defocused image synthesis with software
  - An arbitrary degree of defocusing

Original image


VISUAL
PERCEPTION
FROM A
COMPUTER
GRAPHICS
PERSPECTIVE
GEORGIA
EUROPEAN
Synthesized image

Why defocusing?

L

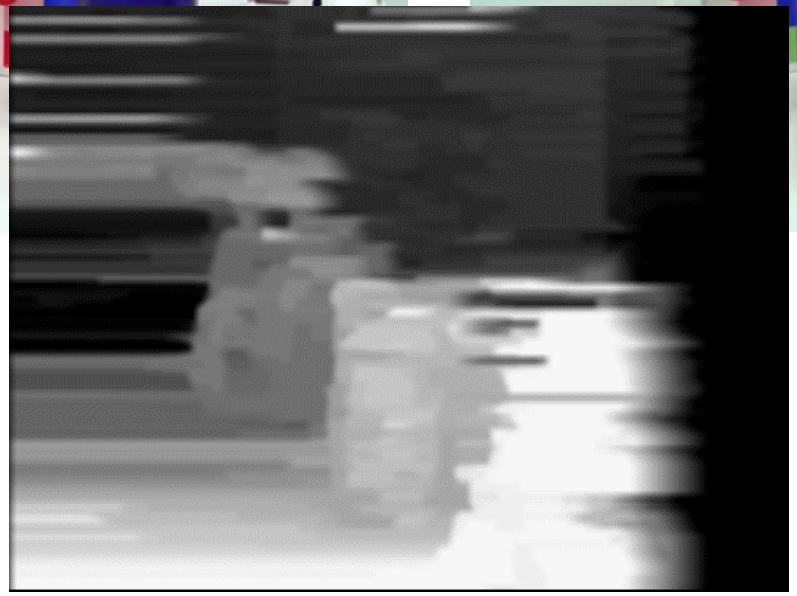
R

Why defocusing?

Why defocusing?

Why defocusing?

L



Depth

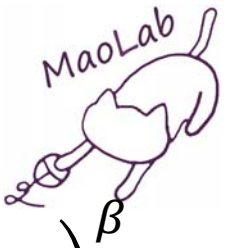
R

Why defocusing?

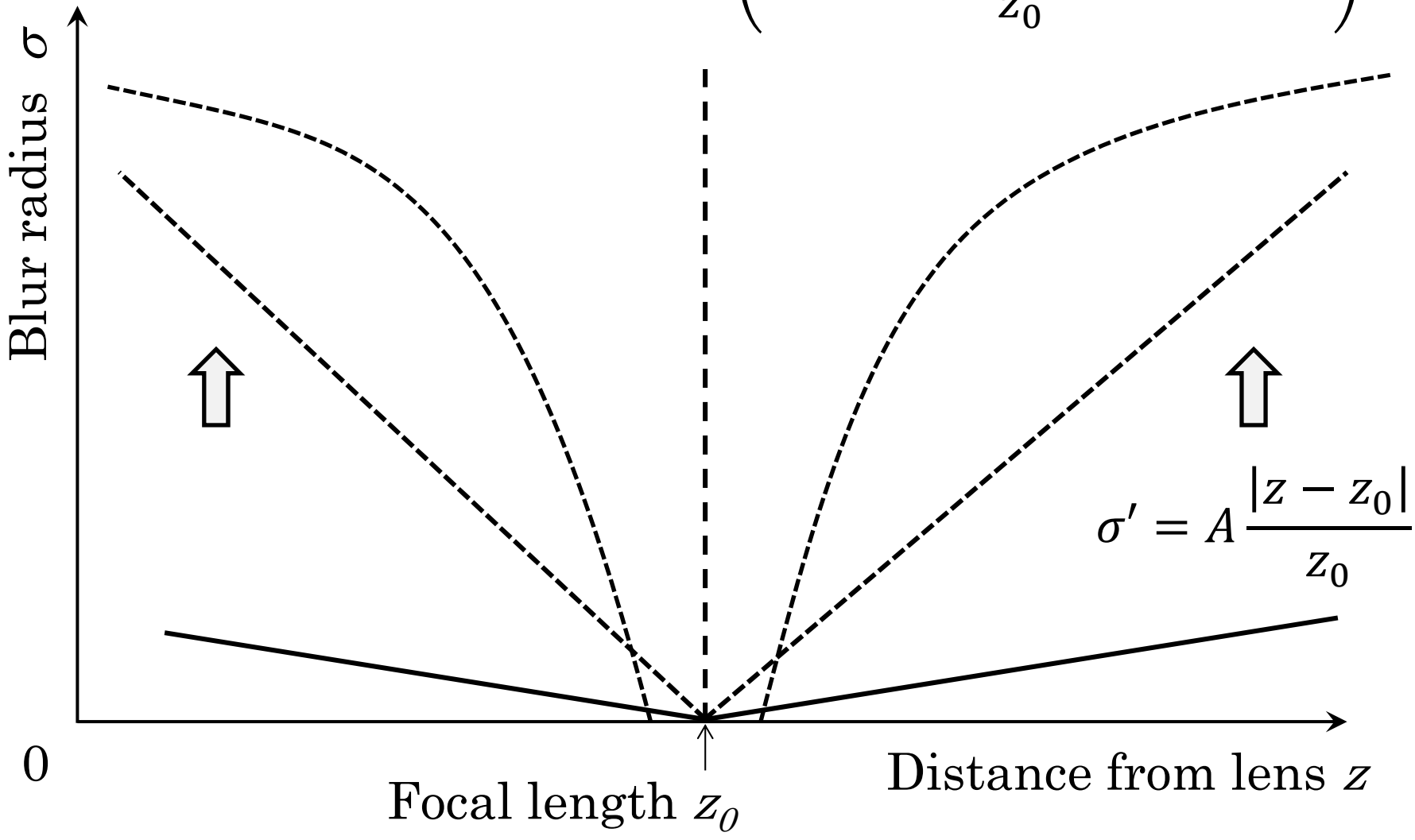
Only **informativeness** is not enough.

Perceptionality and **intuitiveness** are also important.

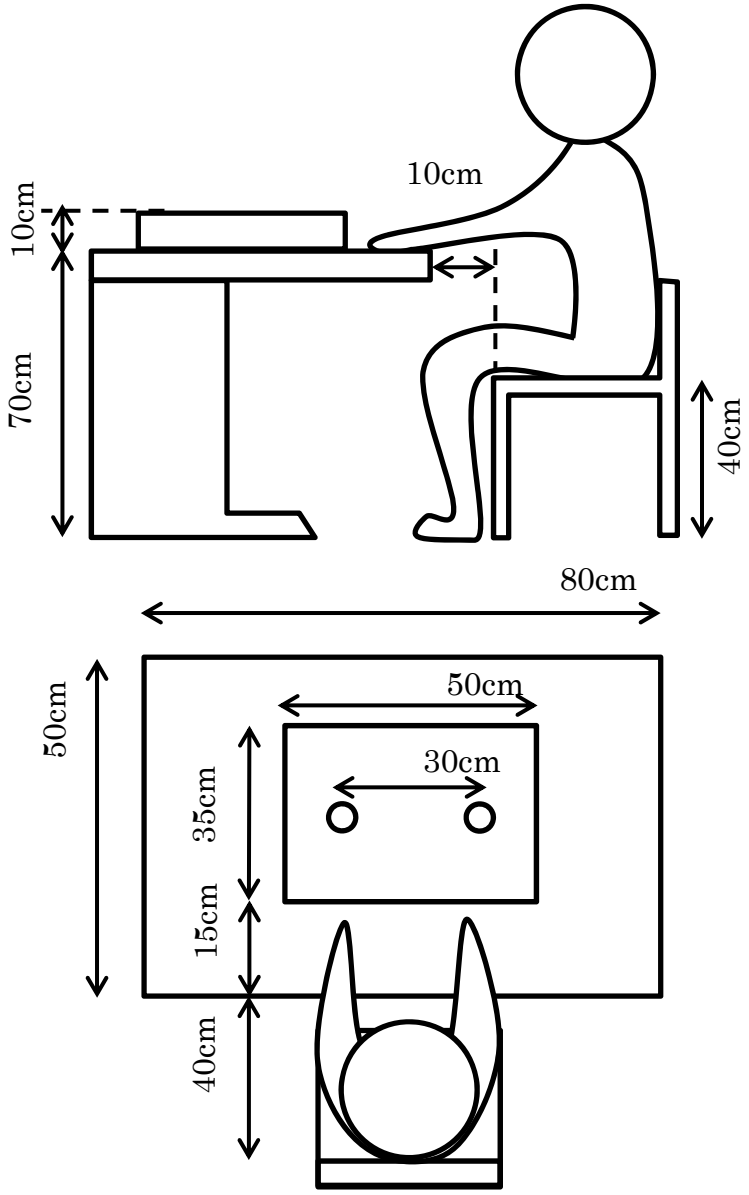
Enhancement of Defocusing

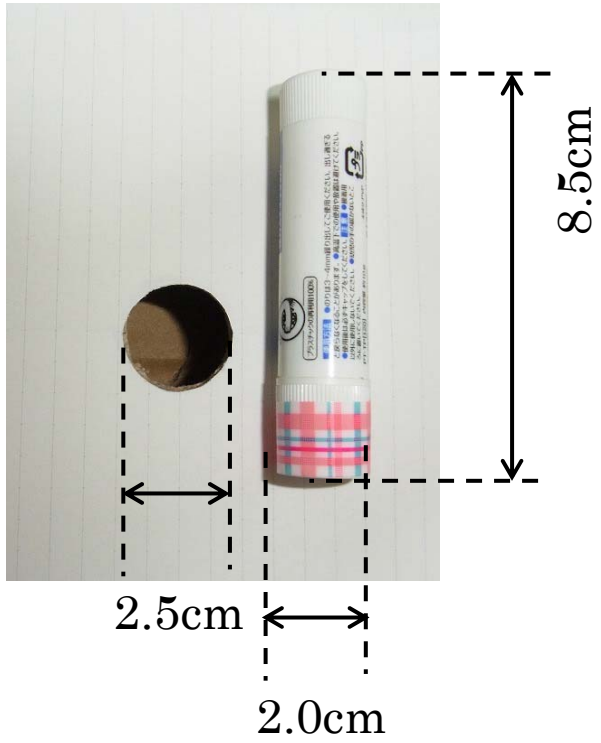


$$\sigma'' = \alpha\left(\gamma\frac{\max(0,|z-z_0|-\phi)}{z_0}+\delta\right)^{\beta}$$



Subject Study

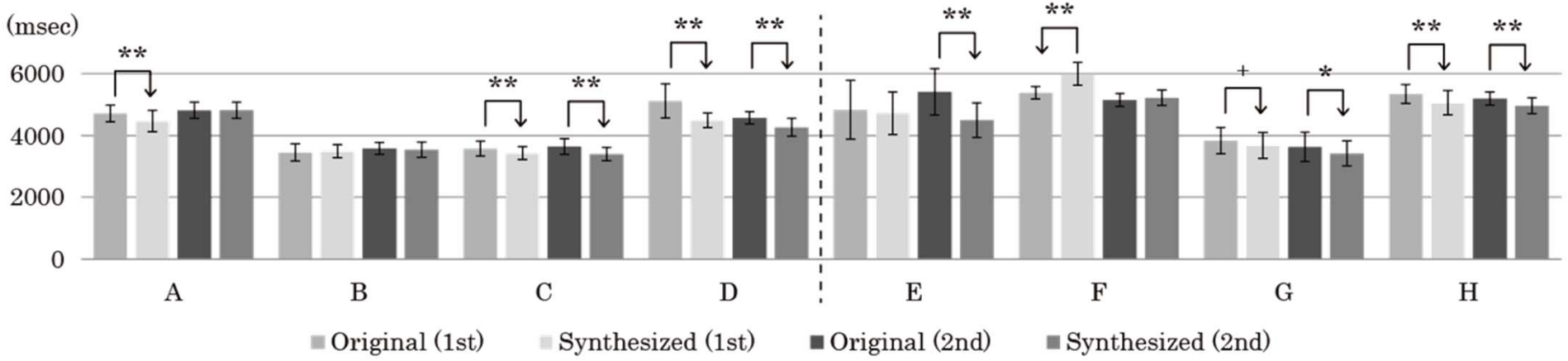




Task (Repeat 30 times)

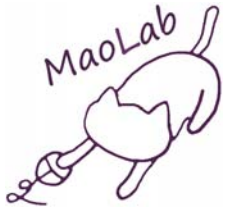
1. Put the cylinder on the front of the box.
2. Insert the cylinder into the right hole.
3. Look ahead.
4. Put the cylinder on the front of the box.
5. Insert the cylinder into the left hole.
6. Look ahead.

Result of Subject Study



- In case that synthesize image displayed…
  - 6 subjects significantly shorter
  - Only F significantly longer
- Discussions
  - Our synthesized images could provide distance information for some subjects
  - There were differences

Vision & Visualization



Visualization for Vision

M. Toyoura, K. Kashiwagi, A. Sugiura, X. Mao, “Mono-glass for Providing Distance Information for People Losing Sight in One Eye,” ACM SIGGRAPH VRCAI, pp.39-42, 2012.

See on behalf of Humans

Vision for Visualization

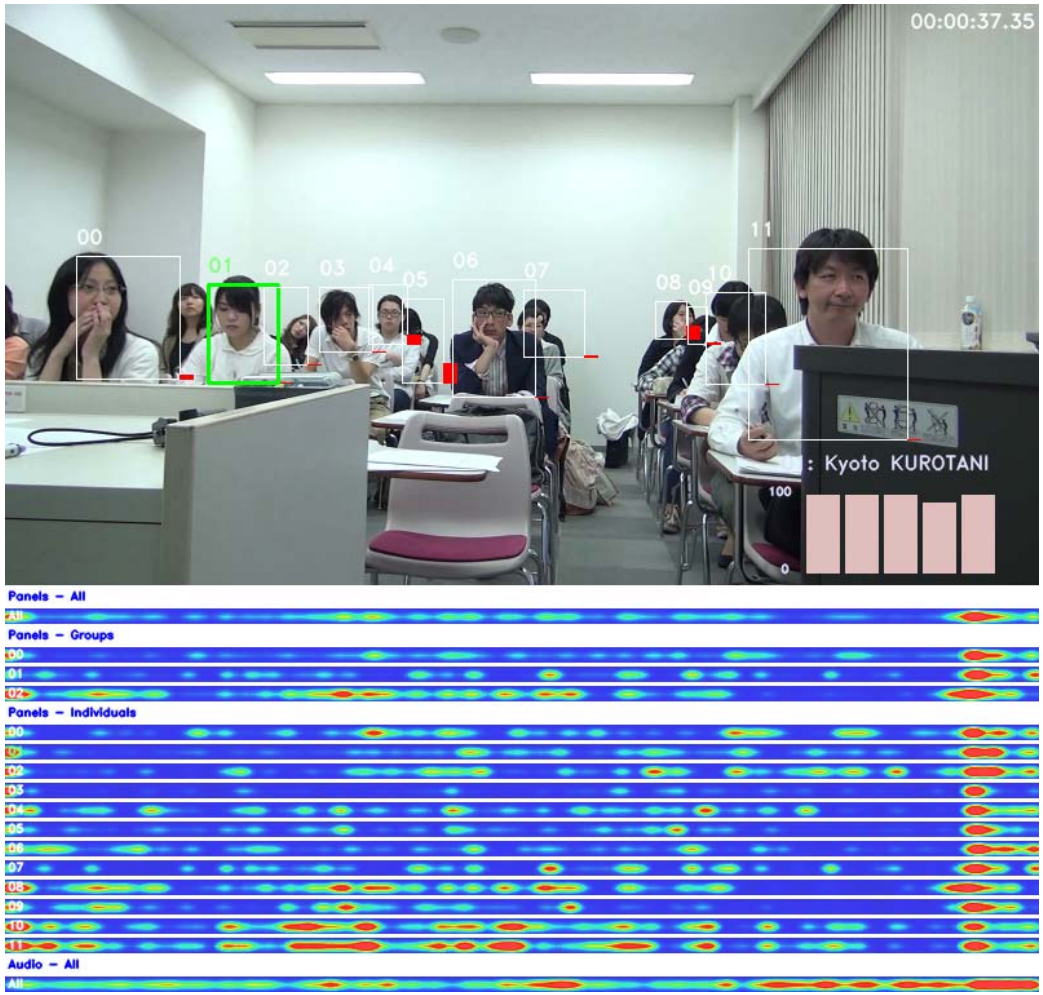
M. Toyoura, S. Nishiguchi, X. Mao, M. Murakami, "Activity Visualization for Multiple Targets in a Video," Pacific Visualization (Poster), 2014.

Analyzing lecture room video for faculty development (FD)

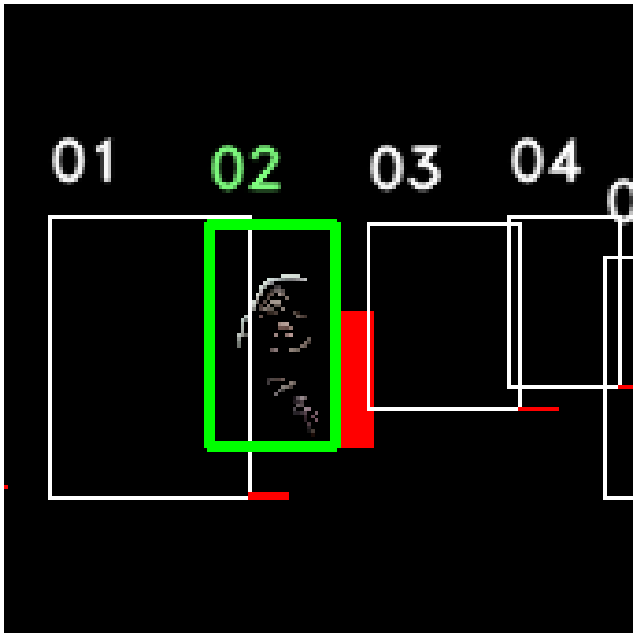
ActVis support analyzing videos
(**Act**ivity **Vis**ualizer)

Overview of ActVis

1. Capturing a video
   - By settled camera
   - Rough positions of target objects are not changed
2. Video processing
   - Place panels
   - Activation level is extracted
   - Generate seek bars
   - (Register meta-data)
3. Video analysis
   - Skim through with visualized data



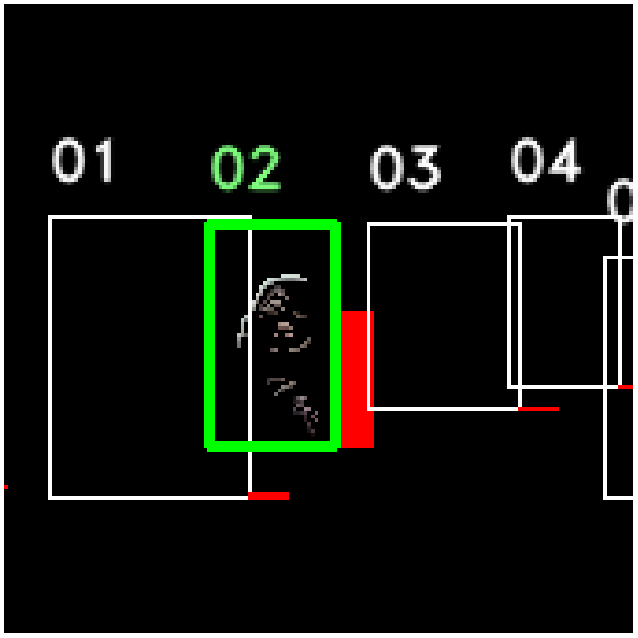
Panel Placement



Place panels at any positions on the screen

- Manual placement helps to avoid miss/error of target detection
- Allow to apply to distorted videos captured by omni-directional cameras
- Multiple panels examines multiple targets

Activation Level Extraction



Activation level is extracted from the rate of pixels with large temporal differential

- Temporal differential, as a low-level feature, can be robustly extracted
- Red linear gauge at the right of the panel also visualizes large movement within the panel

Generates Seek Bars

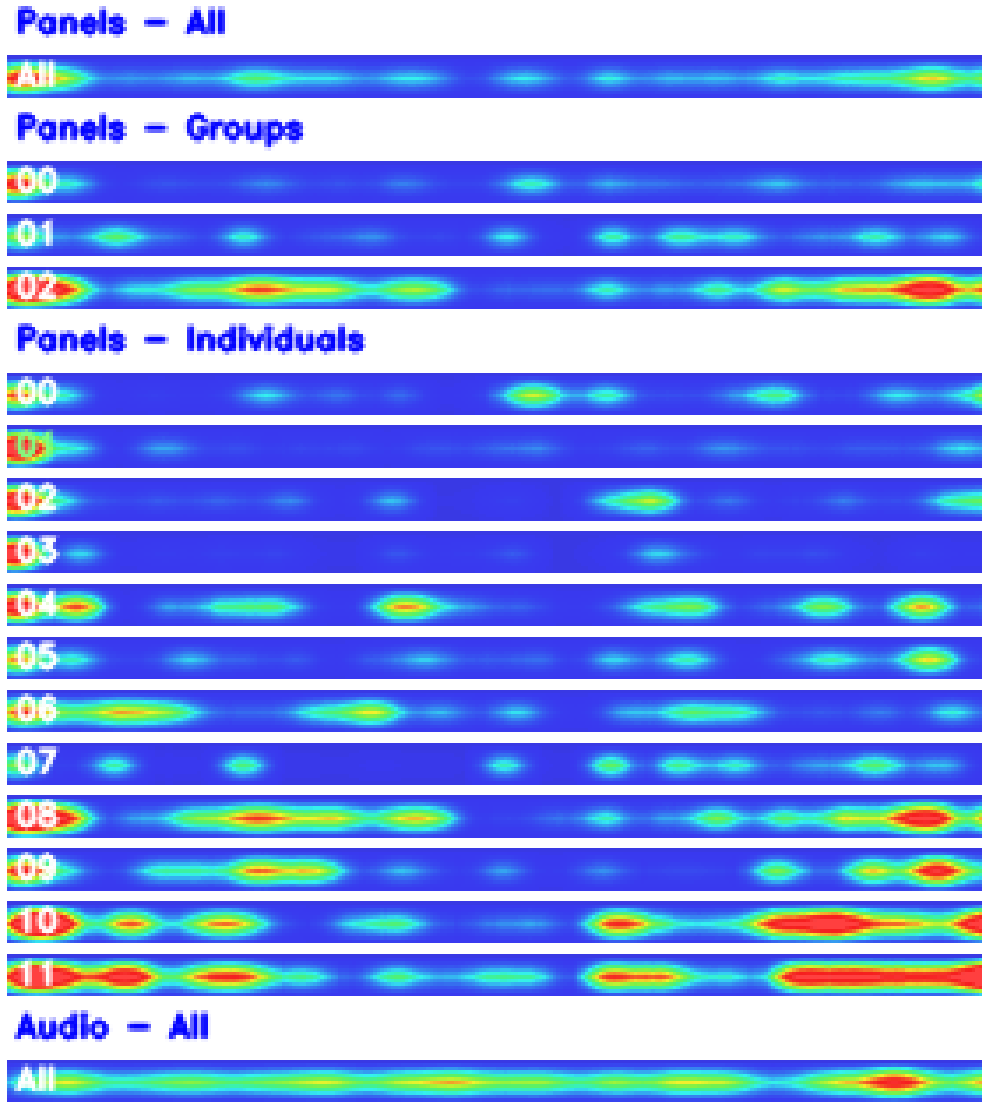
All ⇨

Groups ⇨

(Assign targets to a group manually)

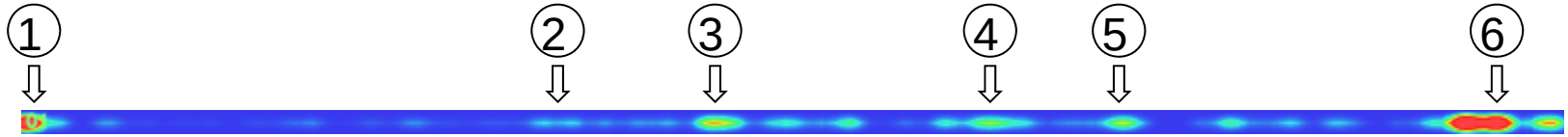
Individuals ⇨

Audio power ⇨



Click a seek bar to
- Jump to selected time
- Highlight the panel

Example of Video Analysis



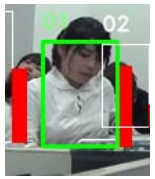
① Sneeze

② Push back her hair

③ Laugh with a colleague

④ Turn over handout

⑤ Face up

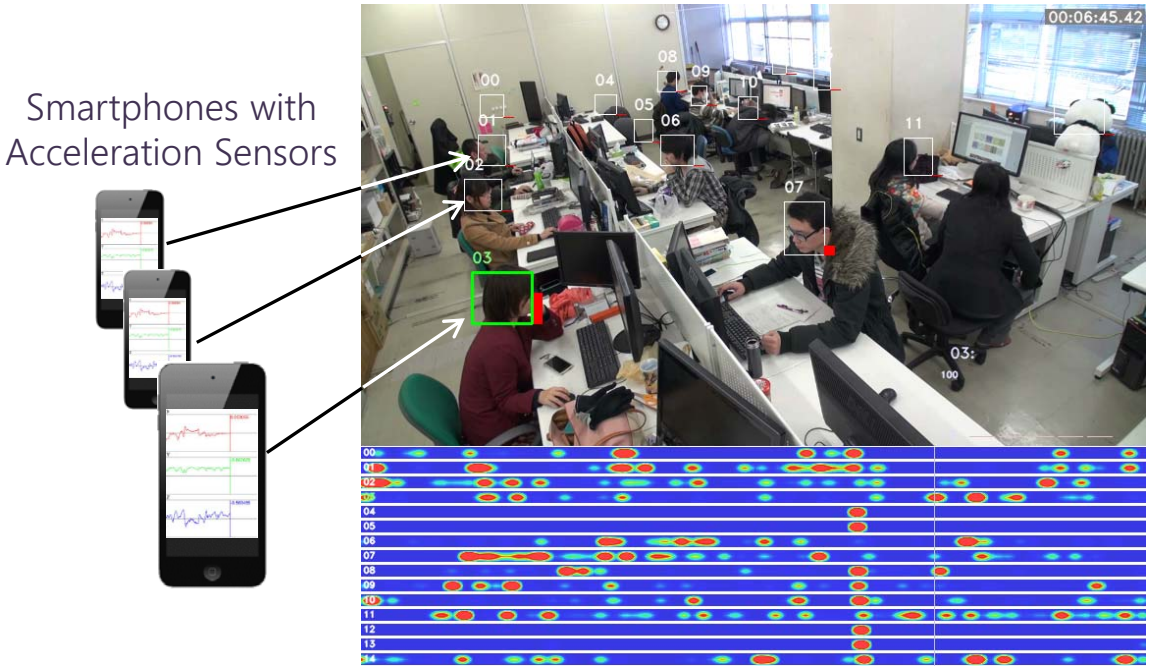


⑥ Bend down for accessing to her bag

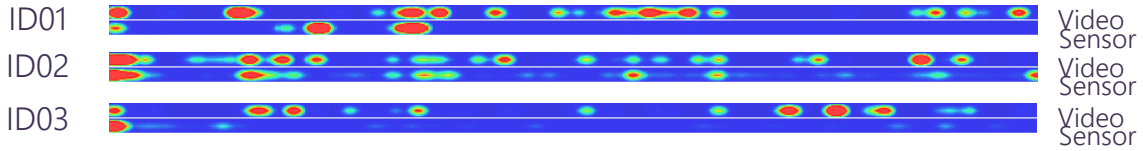
- ③ and ⑤ would be important cues of students' attention for the lectures
- ① and ② may be useless for FD
- Sleeping undetected, but important for FD

Ongoing ~ Sensor Fusion

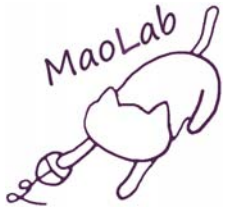
Multiple cameras, more sensors would solve the occlusion.



Corresponding bars (Video and acceleration sensor)



Vision & Visualization



Visualization for Vision

M. Toyoura, K. Kashiwagi, A. Sugiura, X. Mao, “Mono-glass for Providing Distance Information for People Losing Sight in One Eye,” ACM SIGGRAPH VRCAI, pp.39-42, 2012.

See on behalf of Humans

Vision for Visualization

M. Toyoura, S. Nishiguchi, X. Mao, M. Murakami, "Activity Visualization for Multiple Targets in a Video," Pacific Visualization (Poster), 2014.

Thank you for your attention.