Image Processing for Personalized Reality

Kiyoshi Kiyokawa[1]
1 Graduate School of Science and Technology, Nara Institute of Science and Technology
8916-5, Takayama, Ikoma, 630-0192 Japan
E-mail:kiyo@is.naist.jp

Abstract There is a growing interest in augmented human (AH) that extends human capabilities with technology. AH is closely related to virtual and augmented reality (VR/AR) technologies which manipulates the way we perceive and interact with the real world. By flexibly manipulating the feedback cycle, we can *personalize* the reality to better support one's life. Flexible manipulation of the sense of sight is an important example of personalized reality (PA). Head mounted displays (HMDs) are continuously evolving, and researchers envision that one's sight will become fully reprogrammable in the near future. In this talk, introduced are a variety of PA applications and image processing approaches to realize them.

Keywords: virtual reality, augmented reality, personalized reality, redesigning vision

1. Introduction

Humans have acquired new capabilities by inventing various tools long before computers came up and mastering them as if they were part of the body. There is a growing interest in augmented human (AH) that extends human capabilities with technology. AH is closely related to virtual and augmented reality (VR/AR) technologies which manipulates the way we perceive and interact with the real world. We can now *personalize* the reality for each individual, by flexibly manipulating the feedback cycle by making full use of human and environmental sensing, sensory representation, wearable computing, context awareness, machine learning, biological information processing and other technologies. Through such personalized reality (PA) systems, we aim to contribute to the realization of an inclusive society where all people can maximize their abilities and help each other. In this talk, we take the sense of sight as an important example of personalized reality (PA) and introduce a variety of PA applications.

2. Head Mounted Displays

Head mounted displays (HMDs) are continuously evolving, and researchers envision that one's sight will become fully reprogrammable in the near future. Ultimately, visual stimuli should be presented in a field of view (FOV) of 200°(H) x 125°(V), at an angular resolution of 0.5 minute of arc, with a dynamic range of 80db, at a temporal resolution of 120Hz, and the device should look aesthetically good and socially acceptable. User should not notice any visual, temporal, and spatial inconsistency when presented augmentations. This goal is still far away, but we are gradually and steadily approaching.

3. Personalized Reality in the Case of Sense of Sight

A variety of PA applications are conceivable assuming that an arbitrary visual stimuli can be shown by an advanced HMD. For example, we can have a telescopic vision by using zoom cameras [1], a super wide field of view by fisheye lenses with a video see-through HMD [2][3]. We can see future trajectories of moving objects by tracking them [4], virtually have a bird's eye view of the real environment with real-time 3D reconstruction [5], expand the visible wavelength by using a thermal camera [6], replay quick 3D motion in-situ at a slower speed [7], and remove unwanted real objects by diminished reality [8]. To realize such flexible visual manipulation, a large variety of real-time image processing techniques are necessary.

A new opportunity for more advanced personalization of one's sight is emerging thanks to the evolution of CNN-based image conversion algorithms such as GAN and its variations. For example, one can see how a sightseeing place would look like from her viewpoint in real-time on a sunny day when it actually is raining. We currently have several research projects on real-time video conversion based on GAN for video see-through PA.

4. Conclusion

Opportunities for personalizing reality are rapidly increasing. There are many problems that we need to solve before PA is actually be disseminated, such as, a risk of misunderstanding, collision, addiction etc. Long-term impacts on our health are completely unknown either. Nevertheless, we believe that the potential of PA is huge which can completely change our life and the society. What is more important is to think about how human welfare and happiness can be achieved with this new and powerful technology.

References

[1] J. Orlosky et al.: "ModulAR: Eye-controlled Vision Augmentations for Head Mounted Displays," IEEE TVCG, Vol. 21, No. 11, pp. 1259-1268 (Nov. 2015).

[2] J. Orlosky et al.: "Fisheye Vision: Peripheral Spatial Compression for Improved Field of View in Head Mounted Displays," Proc. ACM SUI, pp. 54-61 (Oct. 2014).

[3] Y. Yano et al.: "Investigation of Dynamic View Expansion for Head-Mounted Displays with Head Tracking in Virtual Environments," Proc. ICAT-EGVE, pp. 37-44 (Dec. 2014).

[4] Y. Itoh et al.: "Laplacian Vision: Augmenting Motion Prediction via Optical See-Through Head-Mounted Displays," Proc. ACM Augmented Human (Feb. 2016).

[5] M. Mori et al.: "A Transitional AR Furniture Arrangement System with Automatic View Recommendation," Proc. ISMAR 2016 (Sep. 2016).

[6] J. Orlosky et al.: "VisMerge: Light Adaptive Vision Augmentation via Spectral and Temporal Fusion of Non-visible Light," Proc. ISMAR 2017 (Oct. 2017).

[7] T. Tao et al.: "An Interactive 4D Vision Augmentation of Rapid Motion," Proc. ACM Augmented Human 2018 (Feb. 2018).

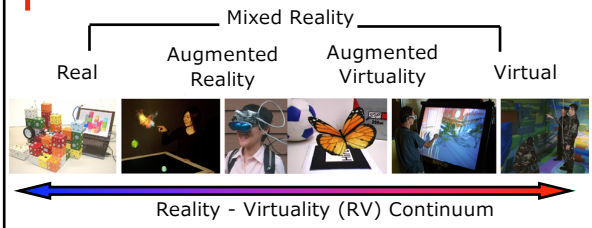
[8] N. Kawai et al.: "Diminished Reality based on Image Inpainting Considering Background Geometry," IEEE TVCG, Vol. 22, No. 3, pp. 1236-1247 (Mar. 2016).

Image Processing for Personalized Reality

Kiyoshi Kiyokawa

Cybernetics & Reality Engineering Laboratory
Nara Institute of Science and Technology

Reality-Virtuality Continuum

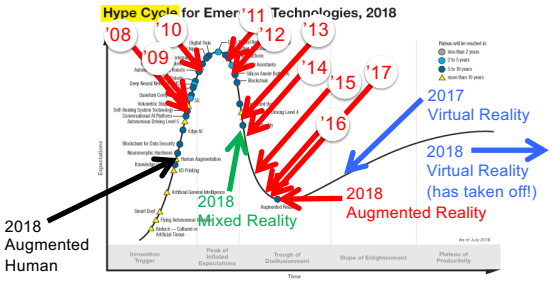


- (in a narrow sense) **Real is primary, virtual is secondary in AR**
- (in a narrow sense) **Virtual is primary, real is secondary in AV**
- (in a narrow sense) **MR is everything that combines both**
- (in a narrow sense) **100% is virtual in VR**

P. Milgram and F. Kishino, "A Taxonomy of Mixed Reality Visual Displays," IEICE Trans. Information Systems, Vol.E77-D, No.12, 1994, pp.1321-1329.

The Gartner Hype Cycle 2018

- AR getting thru disillusionment, VR has gone thru it



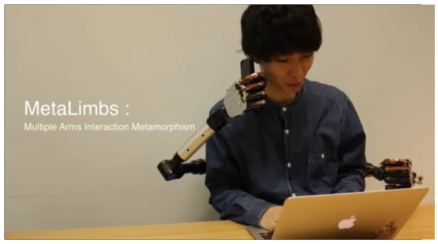
Potential of XR: Becoming the one you wished

- 7 Finger Robot (MIT, 2014)
  - Add two more fingers on your hand

https://www.youtube.com/watch?v=FTJW5YSRZhw

Potential of XR: Becoming the one you wished

- MetaLimbs (The Univ. of Tokyo, 2017)
  - Add two more arms on your body

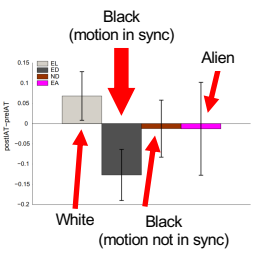


https://www.youtube.com/watch?v=sKjAp0iZ7dc

Potential of XR: Becoming a different self

- Changing your skin color reduces **racial bias**

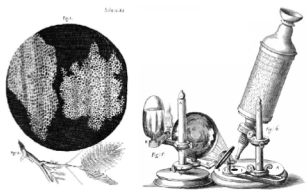
https://www.youtube.com/watch?v=HUN3iQX090



Peck et al., "Putting yourself in the skin of a black avatar reduces implicit racial bias," Cons. & Cognit., 22, 2013.

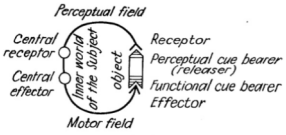
Hooke's Project on Human Augmentation

- Micrographia (1665)
  - "the adding of artificial organs" "to recover some degrees of those former perfections"
  - "Glasses have highly promoted our seeing"
  - "there may be found many Mechanical Inventions to improve our other Senses, of hearing, smelling, tasting, touching"



Jakob von Uexküll's Umwelt

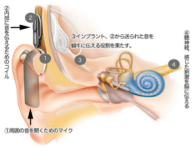
- The animal's umwelt or self-world is determined by the animal's sensory systems, the means by which sensory information is processed, perceived and its action systems
  - A hen responds to chicks' sound
  - A frog responds to object's motion
  - …



Functional cycle

Sensory Prosthesis, Substitution, & Augmentation

- Sensory Prosthesis
  - Restore a lost sensation
  - ex) Cochlear implant



- Sensory Substitution
  - Substitute a lost sensation by another
  - ex) The Optacon (OPtical to TActile CONverter)
- Sensory Augmentation
  - Enable to sense information beyond its original capabilities

Evolution of Hearing Aids

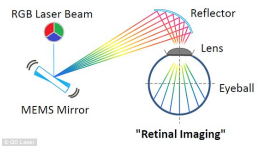
- Benefits by DSP
  - Auto volume adjustment
  - Noise reduction
  - Directional microphone
  - …

- More "smart" functions
  - Translation
  - Speaker-dependent processing
  - Sound synthesis
  - …

(Panasonic)

Will HMDs be the second hearing aids?

- Unstoppable downsizing / high performance trends
- One day, it will be more comfortable with an HMD?
- ex) VRD allows sharp views for those with low vision



QD Laser

As Personal Affairs

- What do **I** want?
- How do **I** want to live?

- **Personalized Reality**
  - For now, for here, for me
  - The level of augmentation does not matter
  - It is **not** Personal Reality
    Reality is personal by nature

Personalized Reality

- Smart Earphones
  - Adjust surrounding sound
  - Translate to your language
  - Emphasize voice of a person of interest
  - …
- Smart Eyeglasses
  - Change color
  - Change color temperature
  - Dim only periphery
  - Brighten a person of interest
  - …

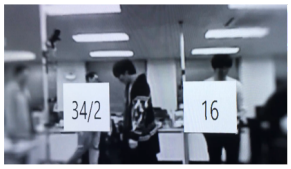
"Here One" AR earphones

"Pixel Buds" Translating earphones

"CTRL ONE" Ting changing glasses

Simple Example

- Blurred monochrome video is less distracting
- Simple math quiz can be solved 20% faster with HMD

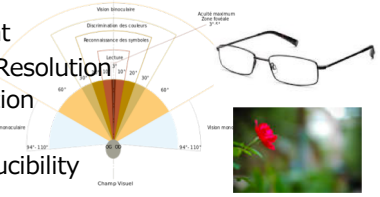


Personalized Reality in the Case of Sight

- Arbitrary operation of the retinal image
  1. "Perfect" HMD
     - Wide-view, high angular res., HDR, low latency…
     - Gradually realized with video see-through approaches
     - Slower, but steady progress is made with optical see-through approaches
     - One day, video and optical approaches may be fused
  2. Correction of visual perception processes
- Flexible editing / redesigning of vision
  3. What capabilities do we want?

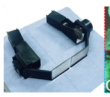
Challenges in HMD

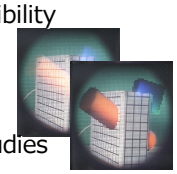
- Size & Weight
- FoV & Angl. Resolution
- Accommodation
- Occlusion
- Color reproducibility
- Calibration
- Latency
- …
- Many many studies

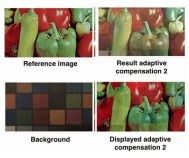


Challenges in HMD

- Size & Weight
- FoV & Angl. Resolution
- Accommodation
- Occlusion
- Color reproducibility
- Calibration
- Latency
- …
- Many many studies





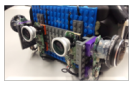


Personalized Reality in the Case of Sight

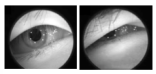
- Arbitrary operation of the retinal image
  1. "Perfect" HMD
     - Wide-view, high angular res., HDR, low latency…
     - Gradually realized with video see-through approaches
     - Slower, but steady progress is made with optical see-through approaches
     - One day, video and optical approaches may be fused
  2. Correction of visual perception processes
- Flexible editing / redesigning of vision
  3. What capabilities do we want?

Augmentation of Field of View + Visual Acuity

- Dynamically change the FOV by user behavior



Mix different types of stereo camera pairs



Intention estimation based on eye gestures

ex)

Looking around to expand the view

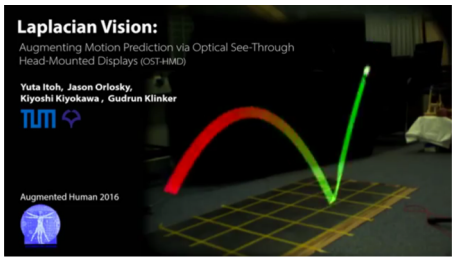
Normal video see-through view

Squinting to zoom the view

@ ISMAR 2015

Augmentation of Motion Prediction Capability

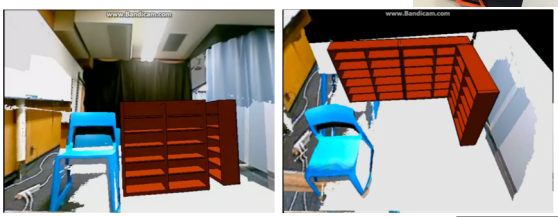
- Overlay a predicted trajectory of an moving object



@ Augmented Human 2016

Augmentation of Viewpoint

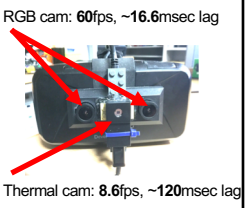
- Freely move the viewpoint while physically staying at the same position



@ ISMAR 2016

Adaptive Sensor Fusion (Augmentation of visible wavelength)

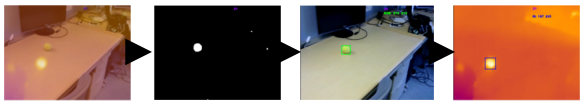
- Adaptively visualize temperature in the dark regions
- Temporal consistency is crucial
  - Fast RGB camera vs. Slow thermal camera



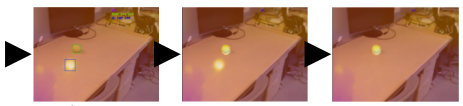
@ ISMAR 2017, AH2018

Synchronized Sensor Fusion

- Scene sync by timestamp-based rotation
- Object sync by per-object tracking and warping



Original | Background Subtraction (fastMCD, Yi 2013) | Object Tracking (KCF tracker, Henriques 2015) | Scene Sync



Object Sync | Image Warping | Image Inpainting

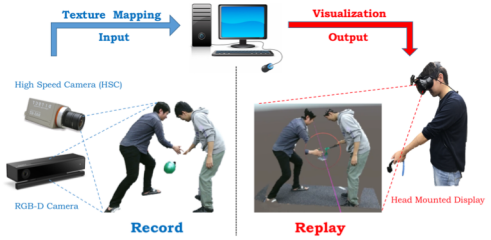
Synchronized Sensor Fusion

- Scene sync by timestamp-based rotation
- Object sync by per-object tracking and warping
- Registration error is reduced by 81.6%

Original | Synchronized

In-situ 4D Visualization
(Augmentation of dynamic visual acuity)

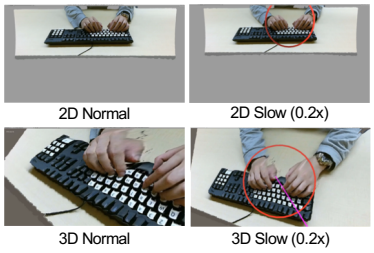
- Some events are too fast to observe
- Replay at a slower speed by HSC and RGBD camera

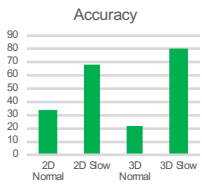


@ Augmented Human 2018

In-situ 4D Visualization

- 4D visualization helps to understand rapid events

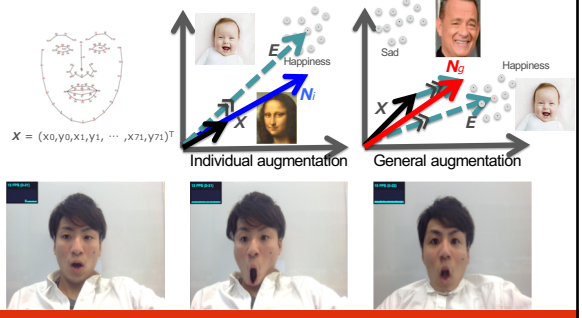




@ Augmented Human 2018

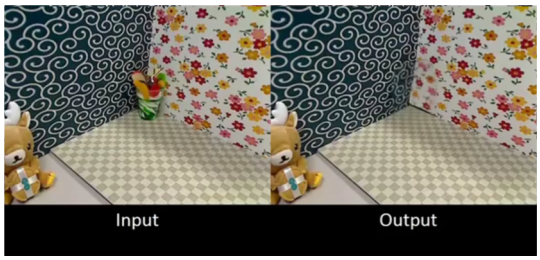
Augmentation of Facial Expression Reading

- Support conversation by easing reading expressions



Removal of Unwanted Objects

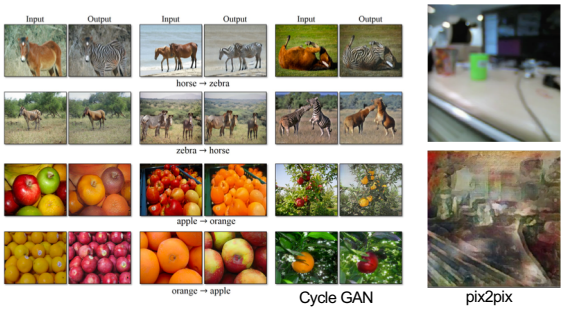
- Diminished Reality



@ ISMAR 2013

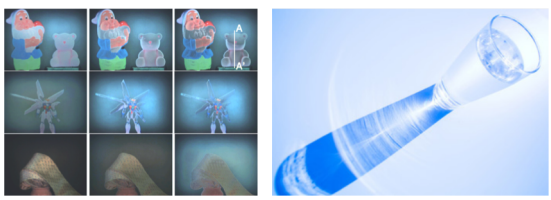
CNN-based Image Conversion

- Many possibilities for personalized reality

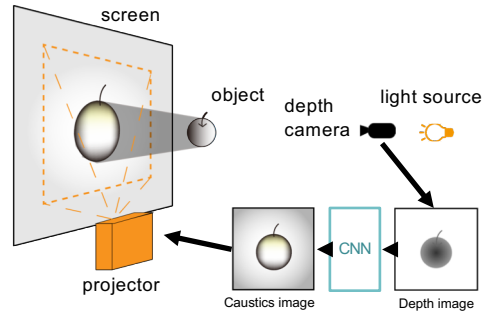


Realtime Caustics Rendering

- Interactive material manipulation getting popular
- Caustics rendering (by photon mapping) still time-consuming

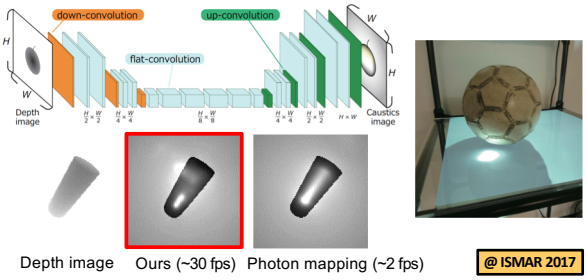


Realtime Caustics Rendering



Realtime Caustics Rendering

- Learning: (256x256) x 5K pairs x 8 shapes, 1300 min
- Estimation: 16.16 msec (NVIDIA GTX 1080)



Depth image Ours (~30 fps) Photon mapping (~2 fps)

@ ISMAR 2017

Summary

- Personalized Reality
  - Infinite possibilities!
  - More and more opportunities!
- Human Welfare
  - Standing by each individual is the key to success
  - What is “happiness” to human being?

Thank you very much!

Cybernetics & Reality Engineering