
Microsoft
Microsoft Research
Faculty
Summit
2016

SPATIAL AUDIO FOR AUGMENTED REALITY

Mark Billinghurst

mark.billinghurst@unisa.edu.au

July 14th 2016

Augmented Reality

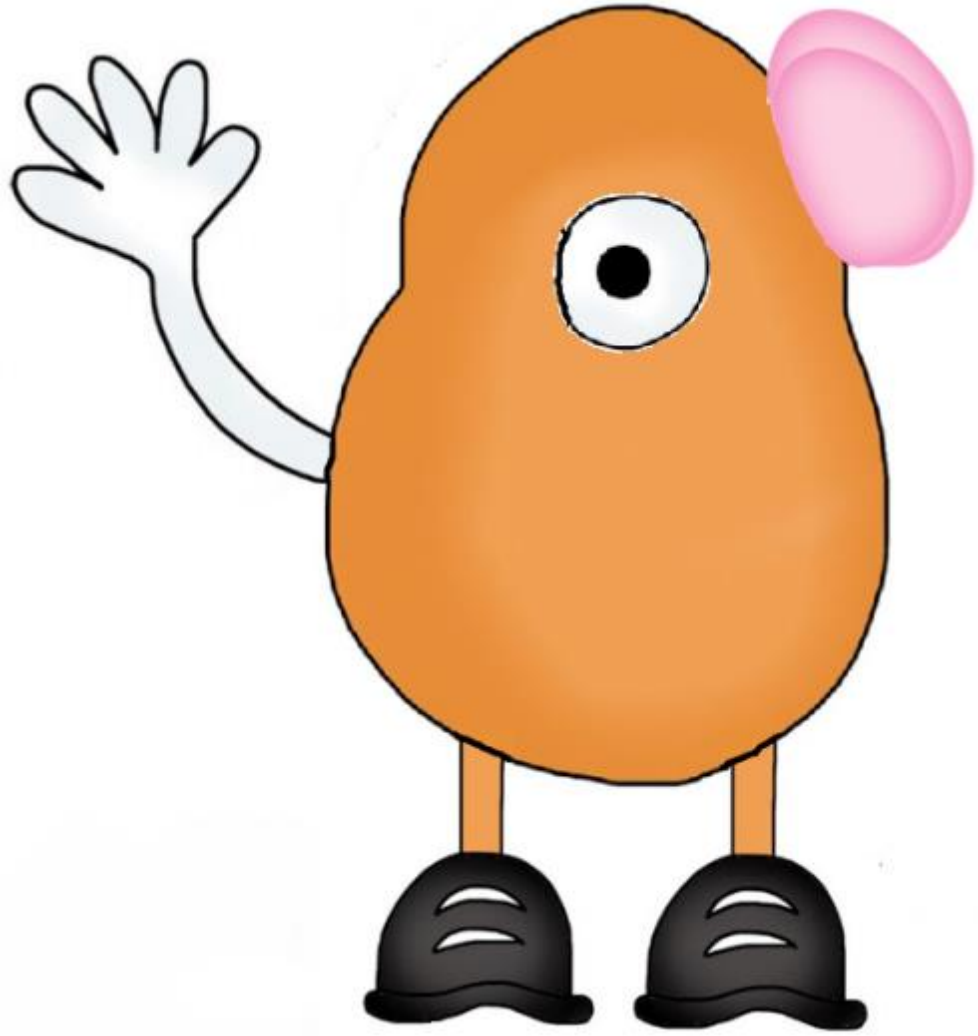
1. Combines Real and Virtual Images
   - Both can be seen at the same time
2. Interactive in real-time
   - The virtual content can be interacted with
3. Registered in 3D
   - Virtual objects appear fixed in space

Azuma, R. T. (1997). A survey of augmented reality. Presence, 6(4), 355-385.

Pokemon GO ..

- Handheld AR, touch input, GPS/compass sensors

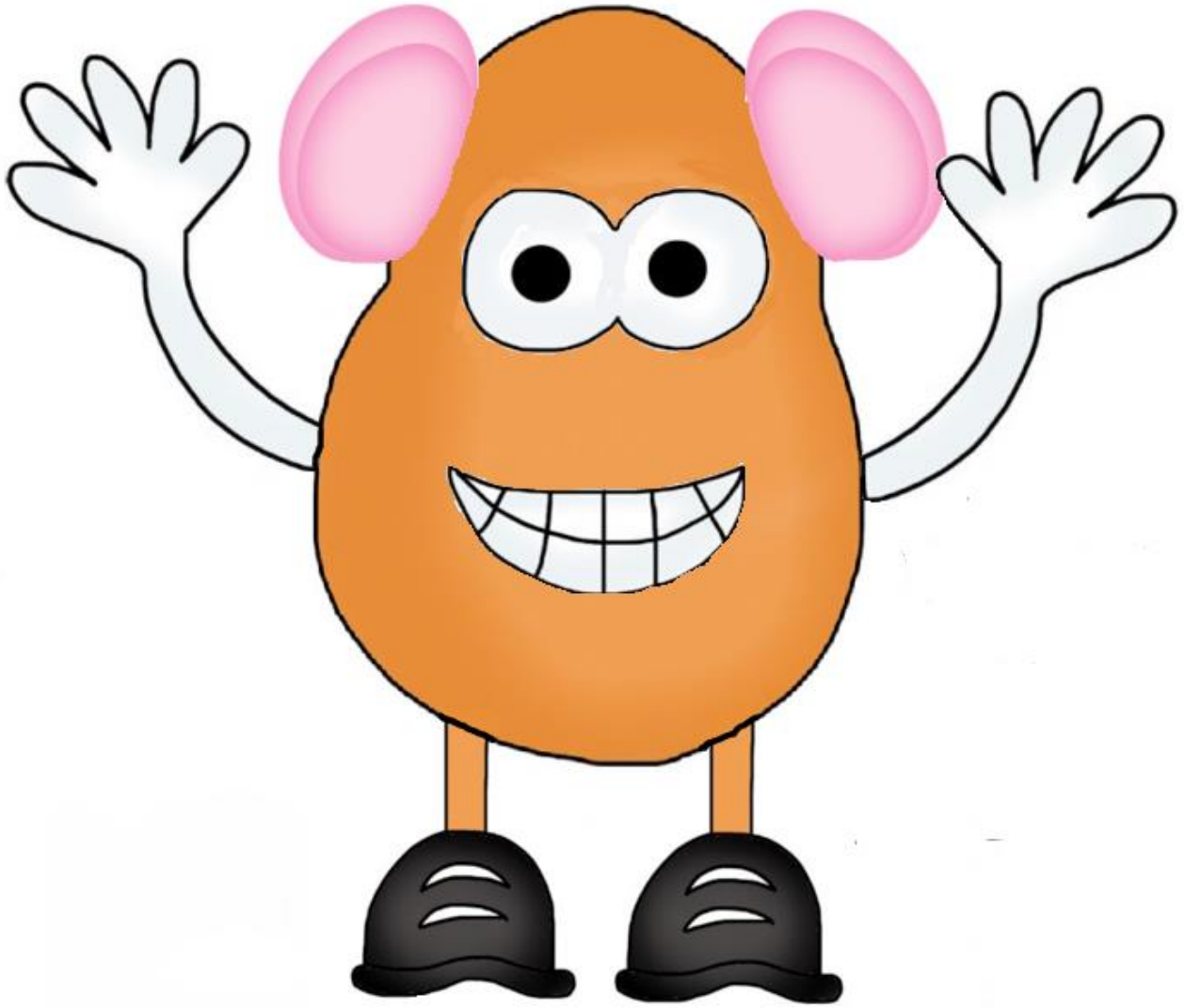
How We Look to Pokemon GO …



Hololens

- Head Mounted Augmented Reality
  - Speech, gesture input, stereo view

How We Look to Hololens



2 Eyes + 2 Ears = AR Spatial Interface

- Visual interface
  - See through HMD has ~ 30^{o} – 90^{o} Field of View
- Audio interface
  - Binaural headphone has 360^{o} Field of Hearing

Wearable Spatial Audio Interfaces

- Previous research
  - Audio only interfaces
    - Navigation, visually disabled, gaming. mobile UI
- Little work in Hybrid Interfaces
  - Small wearable display + spatial AR

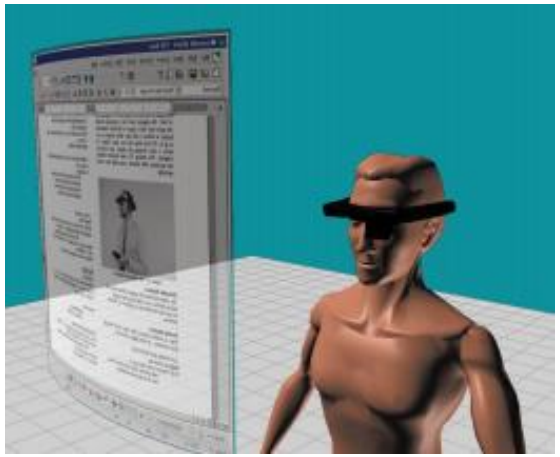
Benefits of Adding Spatial Audio to AR

- Cognitive
  - More information display without additional cognitive load
    - Different visual/auditory systems
- Information
  - Simultaneous information display using multiple modalities
  - Use appropriate modality for information
- Interface
  - Overcome limitations of limited visual display
    - Small screen size, Divided attention
  - Increase interface design options

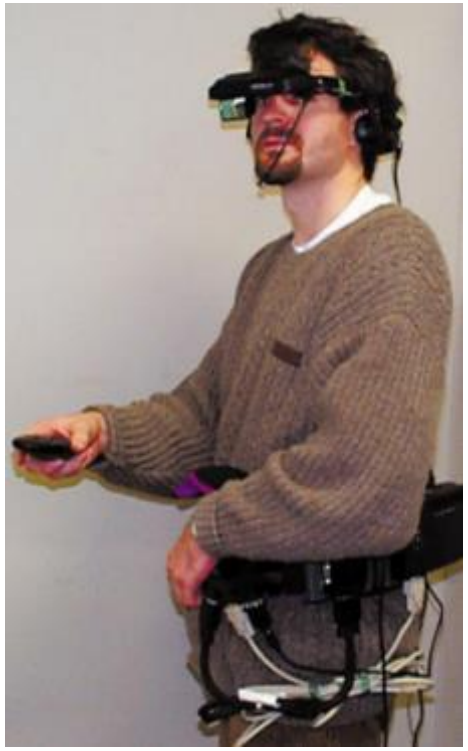
Example AR Applications of Spatial Audio

- Information Presentation
  - Wearable information space (Billinghurst 1999)
  - Attention Redirection (Barde 2016)
- Remote Collaboration
  - Wearable AR conferencing (Billinghurst 1998)
  - Hyrid conferencing spaces (Bleeker 2013)
- Location Based Audio
  - High Street Stories (Lee 2013)
- Authoring/Annotation
  - Audio Stickies (Langlotz 2013)
  - Augmented Sound Reality (Dobler 2002)

Wearable Information Spaces (1998)

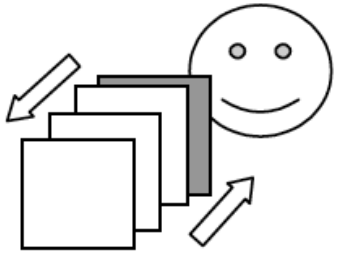


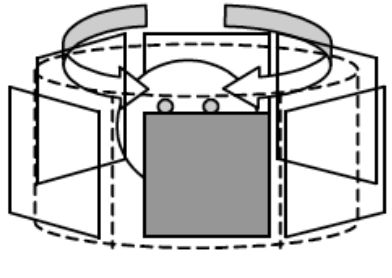
- Exocentric wearable information space
  - See through HMD
  - Wearable computer
  - Spatial audio/visual cues
  - Body stabilized information displays

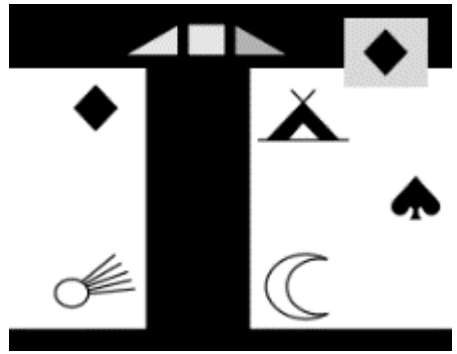


Billinghurst, M., Bowskill, J., Dyer, N., & Morphett, J. (1998). Spatial information displays on a wearable computer. *IEEE Computer Graphics and Applications*, *18*(6), 24-31.

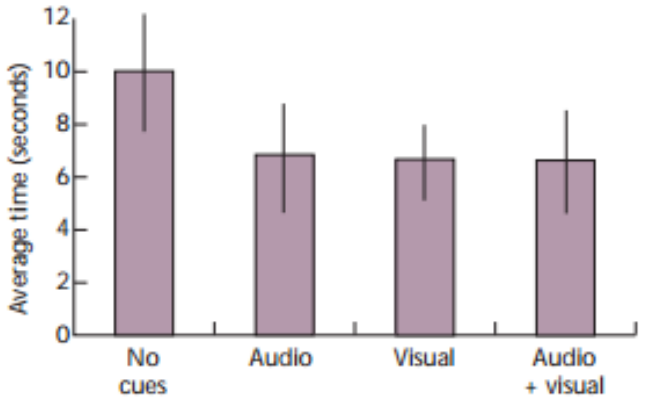
User Evaluation



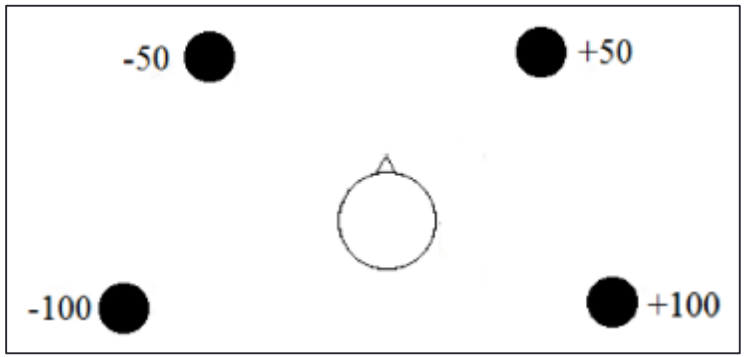




- Task
  - Finding target icon on pages of icons
- Conditions
  - Head stabilized vs. body stabilized
  - Additional spatial audio/visual cues for guidance
- Results
  - Body stabilized 30% faster performance
  - Spatial audio reduces search time by further 35%
  - No difference between spatial audio/visual cues



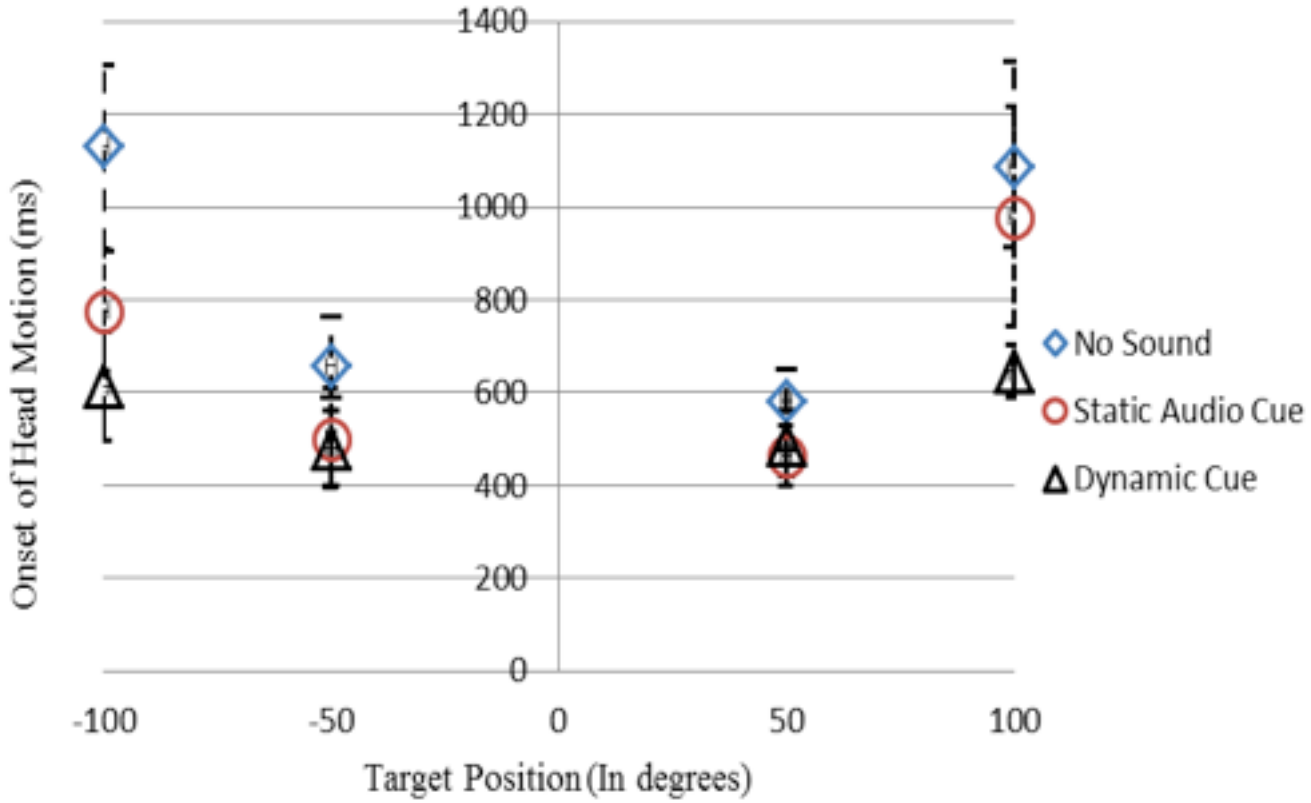
Attention Redirection (2016)



- Use dynamic spatial audio cues to direct attention
  - Audio moving in direction of target position
- Experimental Test
  - Divided attention task (wearable screen, projection screen)
  - Use no cue, static audio, dynamic moving spatial cue
  - Directing user attention to one of four target positions

Barde, A., Ward, M., Helton, W., & Billinghurst, M. (2016). Attention Redirection Using Binaurally Spatialised Cues Delivered Over a Bone Conduction Headset. *HFES 2016*

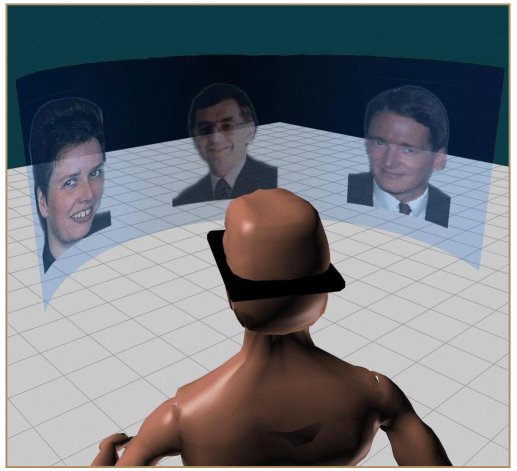
Experimental Results

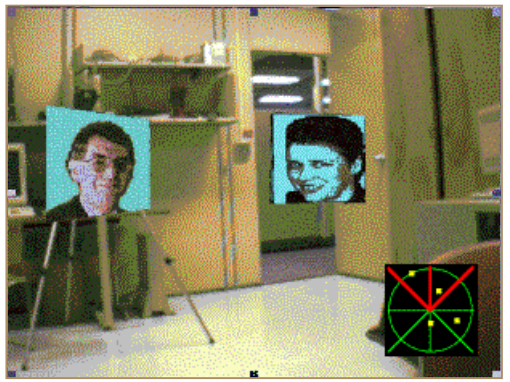


- Dynamically moving audio cue significantly reduces onset time
  - 30-40% faster than static audio cue for targets out of view
  - Up to 50% faster than no audio cue

Wearable AR Conferencing (1998)

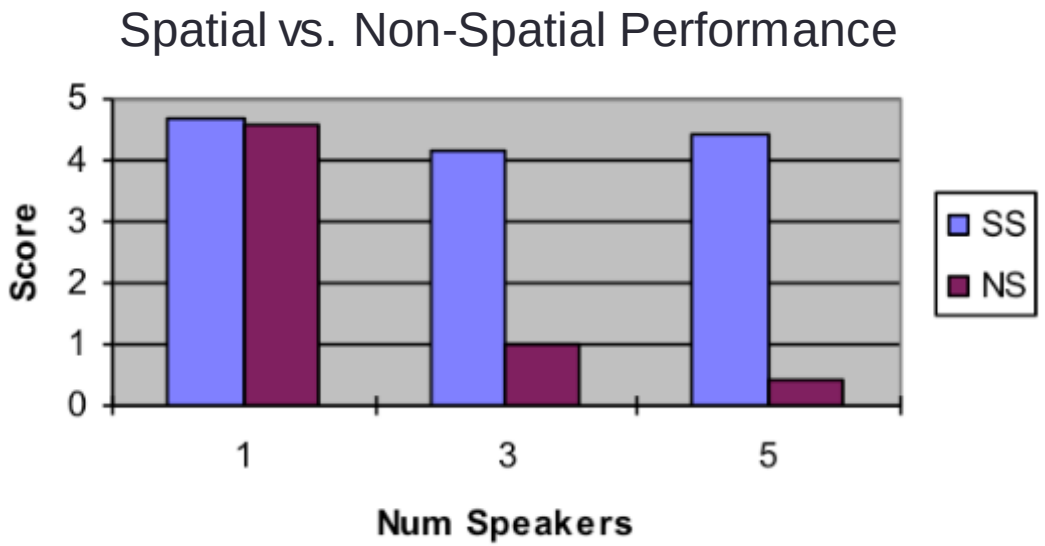
- Concept
  - mobile video conferencing
  - spatial audio/visual cues
  - body-stabilized data
- Implementation
  - see-through HMD
  - head tracking
  - static images, spatial audio

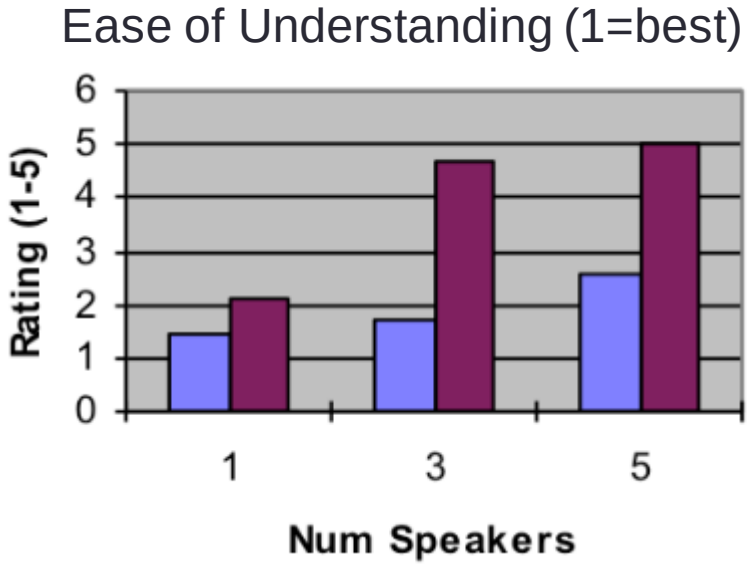




Billinghurst, M., Bowskill, J., Jessop, M., & Morphett, J. (1998). A wearable spatial conferencing space. In *Wearable Computers, 1998. Digest of Papers. Second International Symposium on (pp. 76-83). IEEE.*

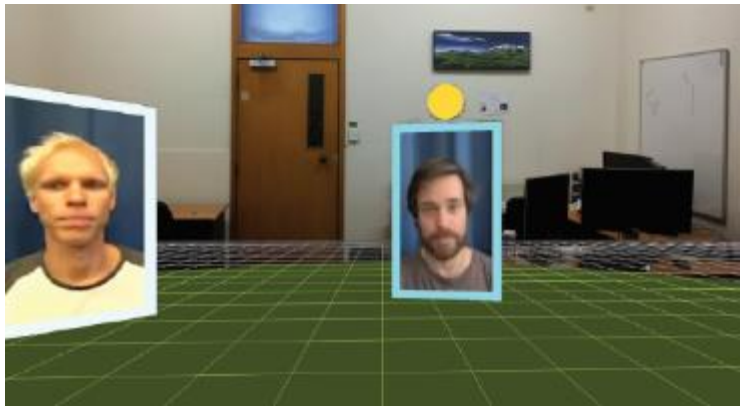
User Evaluation





- Speaker discrimination task
  - 1,3,5 speakers saying almost same phrase at same time
  - Spatial vs. non-spatial cues
- Results
  - Spatial performance significantly better, more higher rated
  - Even simple spatial visual cues (radar display) produced improvement

Using HHD and HMD (2013)



HMD AR View

Tablet View

- Use tablet to interact with AR conf. people
- Exo-centric view of conference space
- AE Spatial audio gives sense of direction

Bleeker, T., Lee, G., & Billinghurst, M. (2013). Ego-and Exocentric interaction for mobile AR conferencing. In *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on* (pp. 1-6). IEEE.

Location Based - High Street Stories (2013)

2009

2011

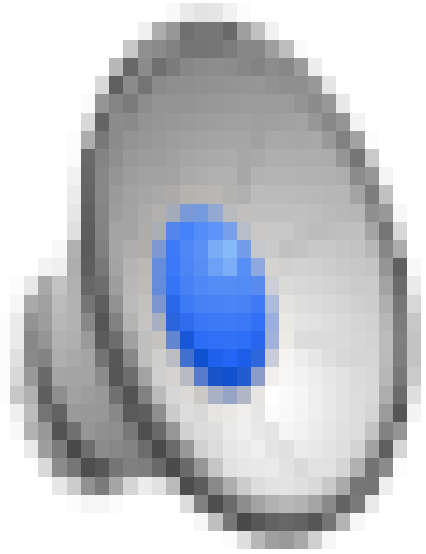
2016

- Christchurch 2011 earthquake
  - Destroyed High Street, historical heart of city
- High Street Stories
  - Mobile AR app with minimal visual cues
  - Geolocated spatial audio cues – stories from locals
  - See http://www.highstreetstories.co.nz/

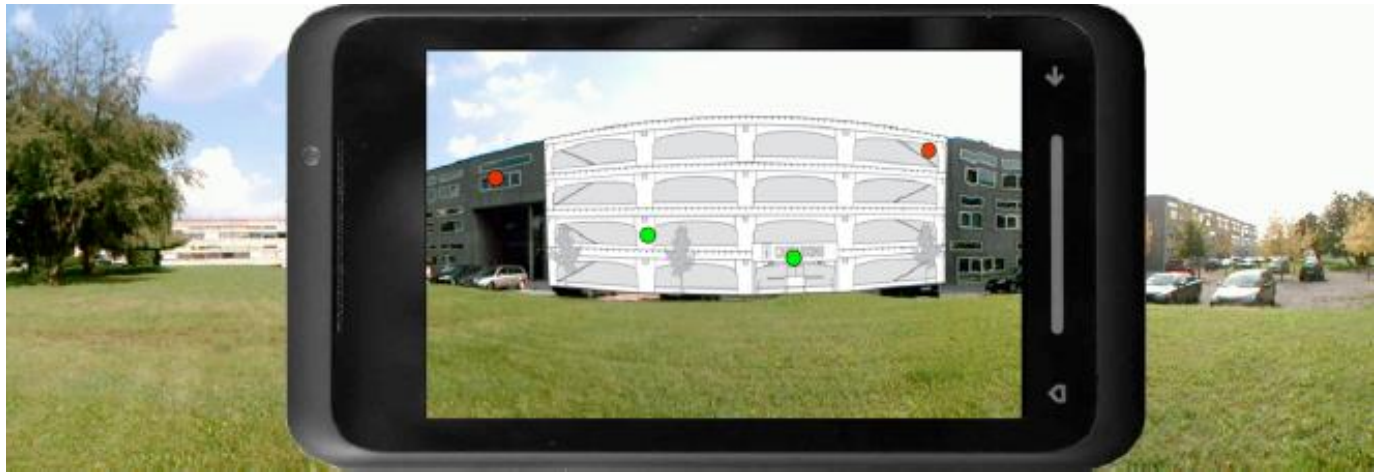
Location Based Information

- High Street Stories Interface
  - Map + AR View (GPS, compass interface)
  - Virtual tags showing geo-located stories
  - Spatial audio browsing based on viewpoint
  - Click to play complete story, view images

Demo Video



Authoring - Audio Stickies (2013)



- Mobile AR browser
  - Outdoor AR, GPS/compass tracking, panorama tracking
- User's can add spatial audio annotations
  - Precise placement of spatial audio notes

Langlotz, T., Regenbrecht, H., Zollmann, S., & Schmalstieg, D. (2013). Audio stickies: visually-guided spatial audio annotations on a mobile augmented reality platform. In *Proceedings of the 25th Australian computer-human interaction conference: augmentation, application, innovation, collaboration* (pp. 545-554). ACM.

Building Annotation

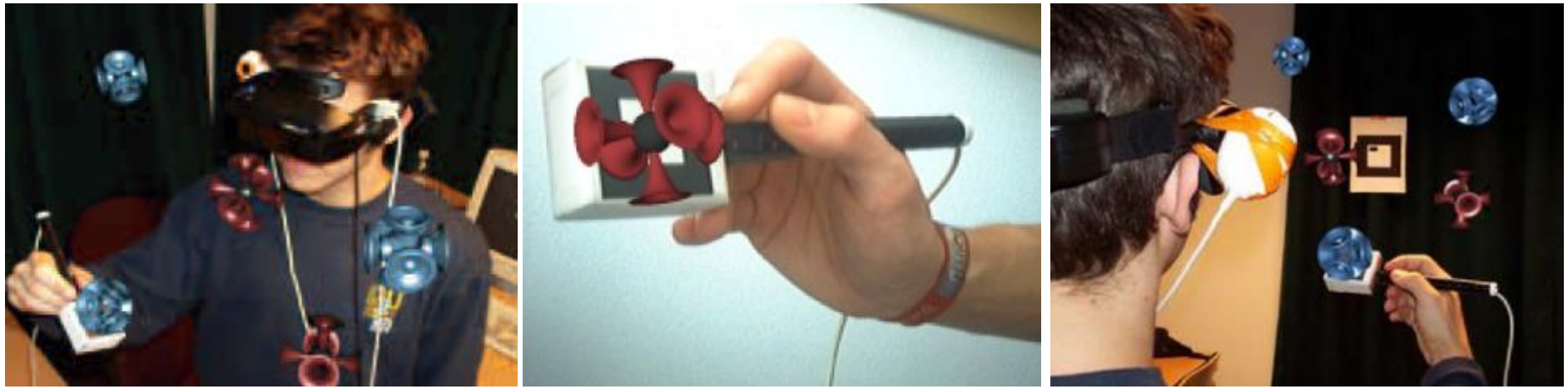
- Use mobile AR to view virtual buildings on site
  - View alternative AR designs
- Viewer can add audio comments
  - Simple tap and record interface
- Users can browse audio notes of others
  - Only play audio clips when in view

Demo Video

User Feedback

- 30 users tried system
  - 4 AR buildings viewed and commented on
  - Tested in two cities (Dunedin, Graz)
- Main feedback
  - Audio annotations seen as very useful
  - System easy to learn and use
  - Usable in noisy natural environment
  - Spatial audio supported discrimination between notes
  - Audio clutter an issue

Augmented Sound Reality (2002)



- Wearable interface for placing spatial audio cue
  - Virtual icons representing audio cues
  - 3D stylus for direct manipulation of sound sources
  - Viewing on stereo video see-through HMD
  - Spatial audio playback

Dobler, D., Haller, M., & Stampfl, P. (2002). ASR: augmented sound reality. In *ACM SIGGRAPH 2002 conference abstracts and applications* (pp. 148-148). ACM.

Lessons Learned

- Spatial audio helps with information presentation
  - Out of view information, multimodal presentation
- Spatial audio can direct user attention
  - Dynamic audio cues
- Spatial audio cues can improve AR conferencing
  - Speaker discrimination, localization, social presence
- Tools can be developed for spatial audio authoring
  - Recording, manipulation audio cues
- Spatial audio enables richer AR experiences
  - Engages more sensors, reduces cognitive load

Directions for Future Research

- User interface metaphors
  - How to interaction with hybrid interfaces?
  - How to present information between modalities?
- Collaborative Interfaces
  - Using spatial audio for sharing communication cues
  - Recording and sharing spatial audio
- Applications/Tools
  - Which AR applications should use spatial audio?
  - AR spatial audio development tools
- Technology
  - Using headphones vs. bone conducting transducers/other tech.
  - Spatial audio algorithms (individual HRTF vs. generic HRTF, etc)

Conclusions

- AR is becoming commonly available
  - Handheld, head mounted
- Spatial audio can significantly improve AR experience
  - User interface
  - Information presentation
  - Remote collaboration
- However there are still significant areas for research
  - User interface, algorithms, collaboration, applications, etc

www.empathiccomputing.org

mark.billinghurst@unisa.edu.au

@marknb00