Talking Head and User Interfaces in VR

Miroslav Macík

CTU in Prague
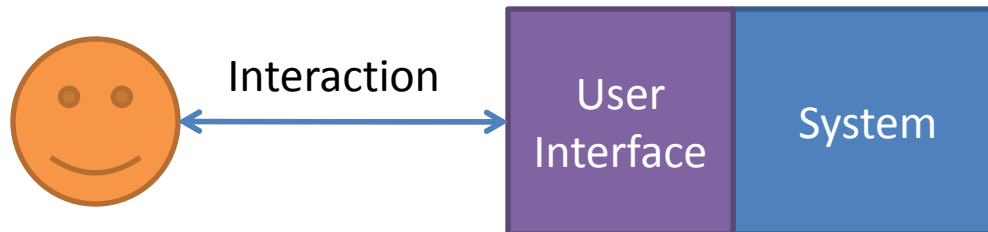
Department of Computer Graphics and Interaction

Outline

- User Interfaces
- User Interfaces in Virtual and Augmented Reality
- Interaction with Avatars and Conversational agents
- Speech interaction
- Talking heads
- Talking heads in i2home project
- Conclusion

User Interfaces

- Not only GUIs (Graphical UIs)
- UI is set of means by which people interact with a system
  - Real buttons, switches etc.
  - Speech interaction, haptic (touch) interaction
  - Conversational agents

UIs in virtual reality - Taxonomy

- User interfaces in virtual reality
  - Navigation and control of the virtual world
  - Embedded equivalents of real world UIs
    - Touch sensors, avatars, conversational agents
- (G)UIs in augmented reality
  - Interaction with real world
- Virtual reality in user interfaces
  - Avatars, Talking heads

User Interfaces in Virtual Reality Control and navigation: examples

- VRML – Walk, fly, examine
- Games – genre dependent
  - Simulators: imitation of real controls of a plane, car, …
  - 3D action games: focus on fast and effective navigation.
  - Massive multi player games: complex interaction, communication with other players

User Interfaces in Virtual Reality Embedded in Virtual Words: examples

- VRML - sensors
  - Touch, proximity, …
  - Feedback
- 3D Adventure games
  - Interaction by clicking on objects
  - Avatars
  - Dialogs
- Massive multiplayer games
  - + Communication with other players and NPCs (non player characters)

(G)UIs in augmented reality: examples

- HUDs [Head Up Displays]
  - Airplanes, cars
- GUI in augmented reality
  - Accessing real world using conventional GUI
  - VIDEO

Virtual Reality in User Interfaces

- Interaction like with real world objects or persons
- Invisible ubiquitous computing
- Intelligent user interfaces
- Examples
  - 3D animated avatars
  - Conversational agents
  - Talking heads
  - Visualization

Avatars, conversational agents

- Avatar – representation of a user in virtual word
  - 2D, 3D or just special symbol
- Conversational agents
  - Representation of the system
  - Provides realistic look and feel of conversation
  - Talking heads
  - Body animation

Human to human speech interaction

- Multiple modalities
  - Speech
    - Sound domain
    - Visual domain
  - Mimic motions, Eye gaze
  - Hand gestures
  - Overall motion, position of body parts …
- Everything synchronized
- Primary information – sound domain (usually)
- Other modalities – both added and redundant information

Talking heads

- Artificial model of a human head
- Mostly used to represent visual domain of speech
- Need synchronization with sound domain
- Why to use them?
  - More realistic speech interaction
  - Representation of the system as an intelligent being
  - Some users has problem to accept notifications in voice only
- How
  - 3D approach – face muscle simulation, key frame interpolation
  - 2D approach – key frame interpolation
    - feature points (MPEG4 FA)
    - alpha blending

Speech

- Visual and sound domain
  - Phoneme – basic units in sound domain
  - Visemes – basic units in visual domain
- There is not direct relation between phenomes and visemes
- In general visemes are context dependent
  - Appearance of viseme = current phonemes+ previous phoneme (s) + next phoneme(s)

Example of visemes

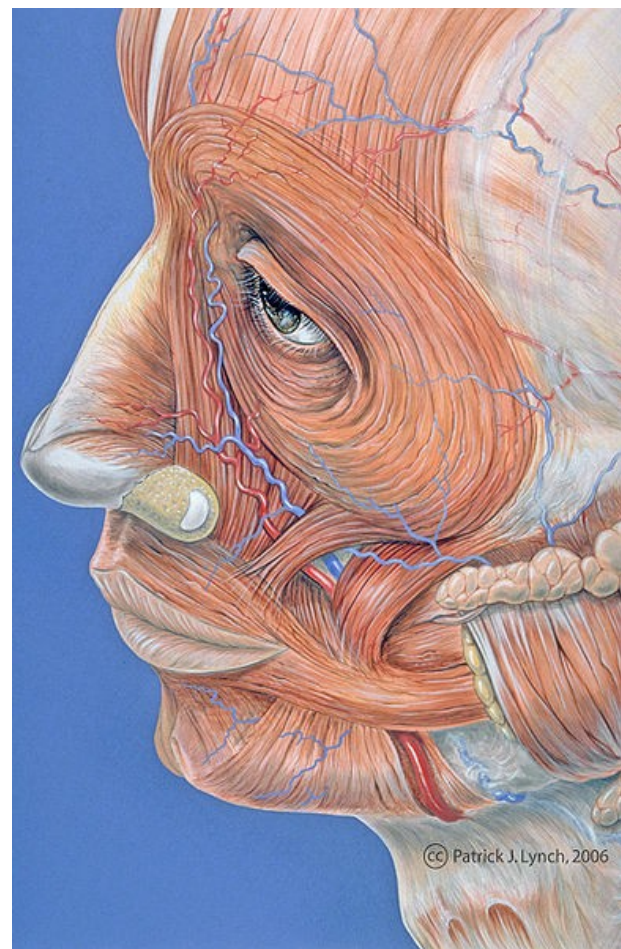eh   th   D,S,T

Speech synthesis

- Text to speech
- Input: text string
- Additional input information
  - Context, emotions, cultural origin
- Output:
  - Voice
  - Stream of phenomes

Talking head – 3D approach

- Representation
  - Volume
  - Surface (NURBS, Polygonal)
- Methods
  - 3D key frame interpolation
  - Parametric models
    - Walters pseudo muscles
  - Realistic models
    - Simulation of real muscles and skin
    - Computational time demanding

CC Patrick J. Lynch, 2006

Talking head – 2D approach

- Based on morphing of key frame images
  - Key frames – usually visemes
  - Feature points (MPEG4 FA)
  - Optical flow estimation
  - Linear or non linear alpha blending
- Features
  - Less computational power needed
    - Can be used on wide range of devices
  - Less fidelity
  - Easy to adapt new face
    - Just replace key fames with new set (and recompute optical flow, feature points)
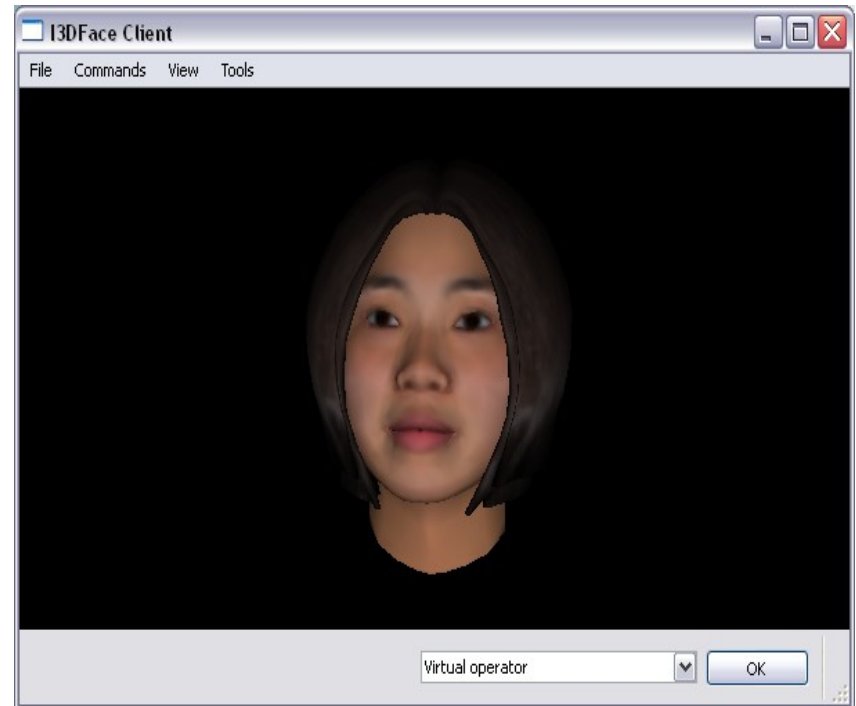
EXAMPLES OF TALKING HEADS

Ladislav Kunc – Jukebox UI

- DCGI, IBM
- 3D, speech synthesis, speech analysis, head motions
- Dialogs
- (Video demo)

Roman Hák – Mobile 3D talking head

- DCGI, master's thesis

- Mobile environment

- Aim: design realistic male and female TH, show the concept on example of call center

- Using SIP [Session Initialization Protocol]

Vicomtech – interaction with users with special needs

- Part of i2home project

- 3D, speech synthesis, no speech analysis

- Simple notifications and dialogs

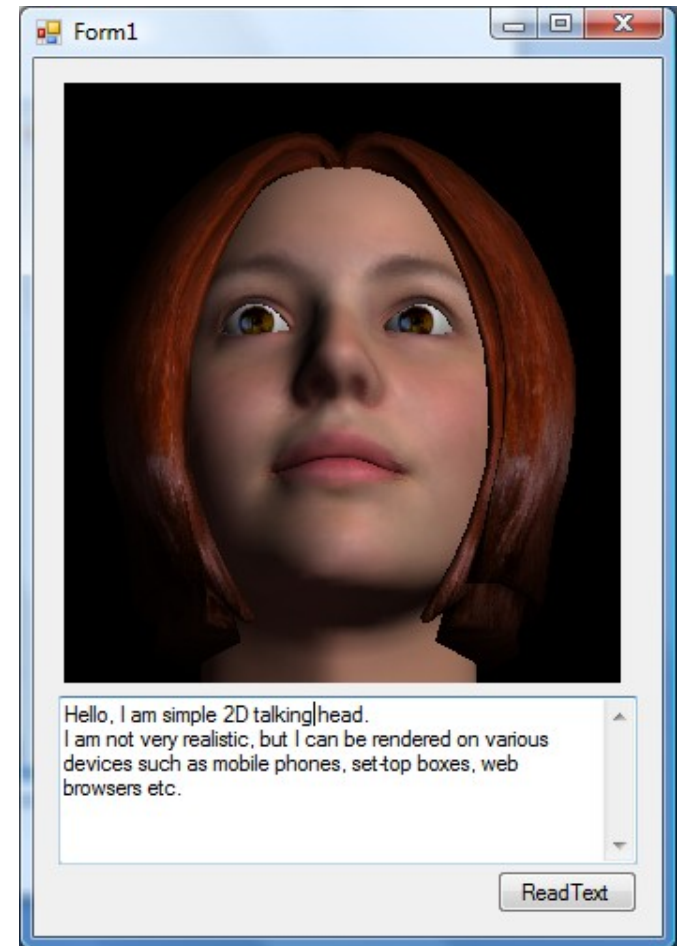- Target group: patients with Alzheimer's disease

Ezat, Poggio - Mike talk – Talking head based on morphing visemes

- MIT, Massachusetts, 1998
- 2D, videorealistic
- key frames morphing
  - Optical flow estimation
- No head motion
- Problem: holes
  - Hole filling

Miroslav Macík – simple 2D talking head

- CTU, CGG, 2008

- 2D

- Key frames morphing
  - non linear alpha blending

- Win Forms 2.0 component

- DEMO

Form1

Hello, I am simple 2D talking head.
I am not very realistic, but I can be rendered on various devices such as mobile phones, set-top boxes, web browsers etc.
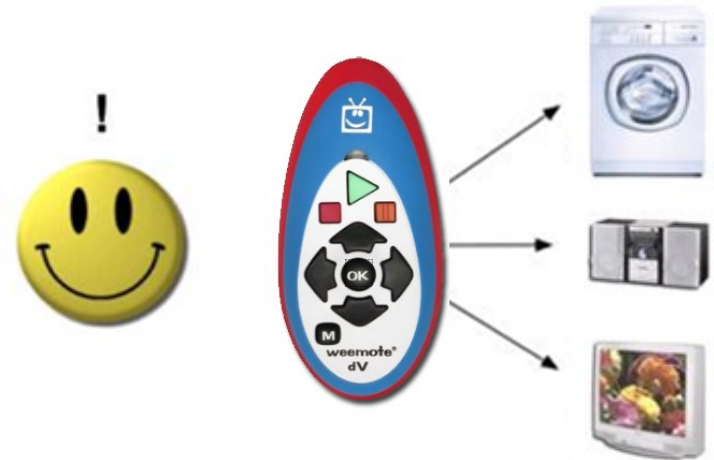
ReadText

i2home project

Intuitive Interaction for
Everyone with
Home Appliances based
on Industry Standards

- Intuitive Interaction for Everyone with Home Appliances based on Industry Standards.

- EU project, 9 organization from 5 EU states
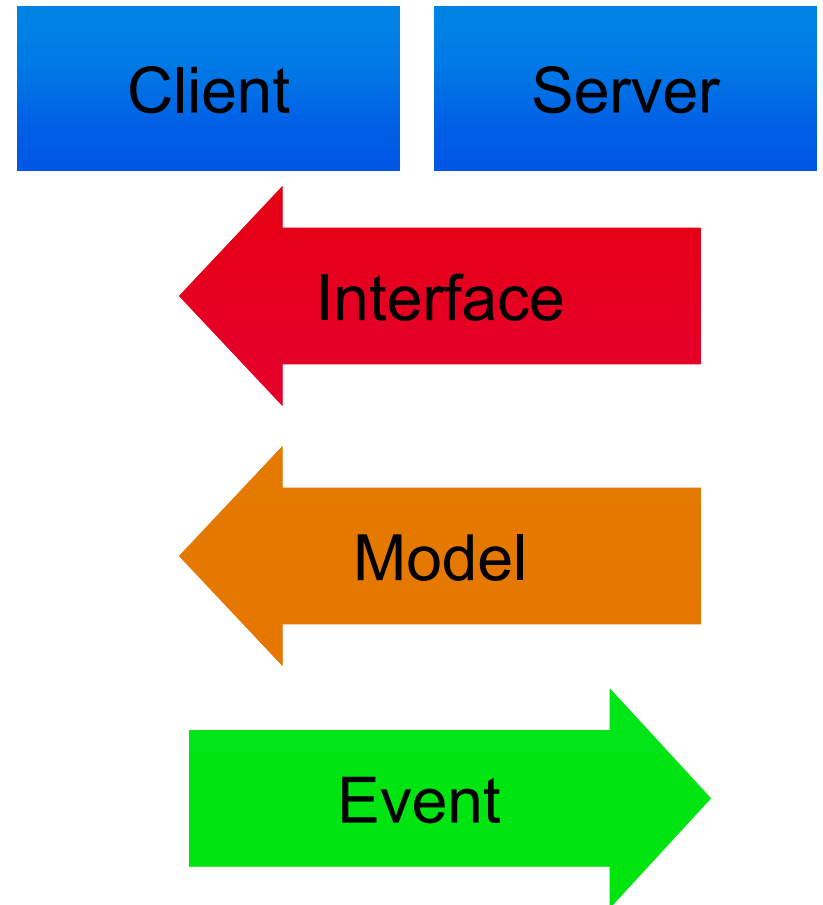
- Motivation:

Talking heads in i2home

- VICOMTECH – 3D TH for people with Alzheimer's disease
- Our role: User interfaces for aged people (65+)
  - Design + Underlying implementation
  - UIProtocol approach
    - Novel protocol for UI description and communication
    - User interface components in hierarchy
  - Nice to have talking head as UIProtocol component

UIProtocol features

- single interface on any platform
  - desktop, web, mobile, TV
  - Flash, Java, JavaFX, Silverlight, Web
- designed for client-server applications
  - separation of application logic and interface
  - rich clients
  - animations, media, maps, charts
- communication protocol
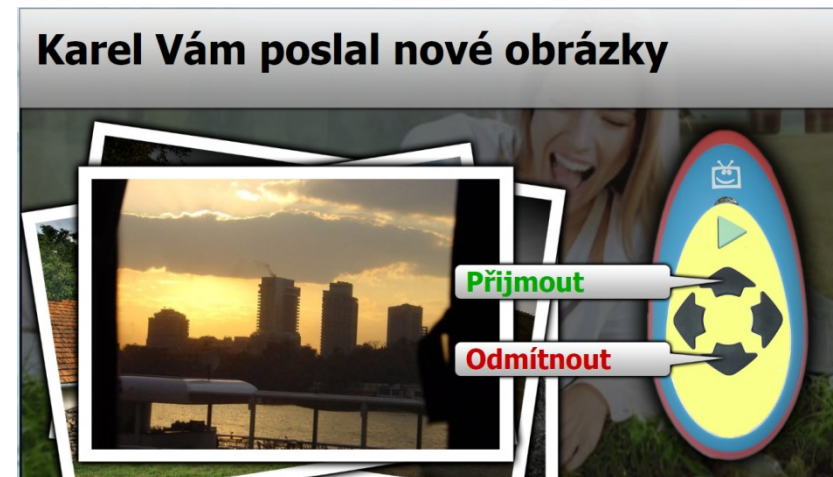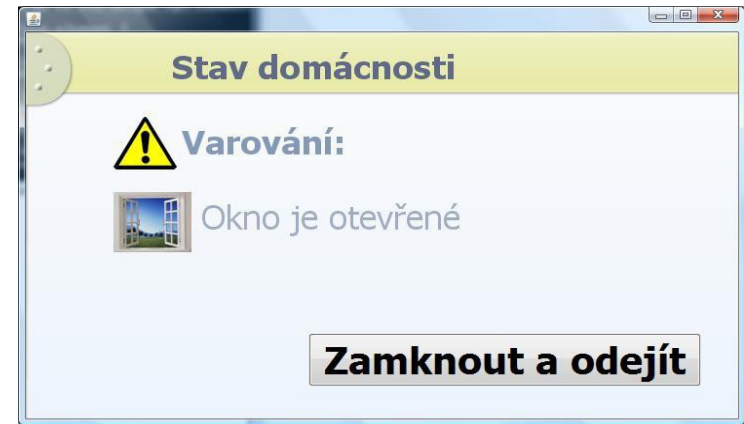 - thin, but powerful client

Client-server communication

- interfaces
  - hierarchical structure of the UI
  - provided by UIGenerator
- models
  - dynamic data used in a interface
  - updated by application logic
- events
  - description of a user action sent to server

Client   Server

Interface

Model

Event

UIProtocol client implementations

- Java reference client (Václav Slováček)

- UiTv - .NET 3.5 Virtual Television (Miroslav Macík)

- Under development
  - Silverlight client (Tomáš Vichta)
  - ASP.NET based web client (Jiří Vaško)
  - PHP based web client (Vojtěch Sedláček)

Talking head UIProtocol component

- We presume that talking head interaction can be good for our target group
  - Notifications (pills reminder, doorbell, …)
- Only way to proof it -> testing
- Implement 2D talking head as UIProtocol component
  - Simple -> we can deliver it on many platforms
  - Graceful fallback (degradation)

Demands

- Deliver the same TH to more platforms
  - Presumption: a user is familiar with particular face
- Easy to implement
  - Can ask clients developers to do it
- Can use real faces
  - Good topic to test the difference
- Can be seamlessly replaced by more sophisticated 3D variant
- On simple clients can be represented by plain text

public.notification.talkingHead

Conclusion

- Virtual reality and user interfaces are tightly connected
  - UIs in virtual and augmented reality
  - Virtual reality in user interfaces
  - Avatars, Conversational agents
- Speech interaction is demanding (sound and visual domain, gestures, mimic)
- Talking heads can do interaction more natural

THANK YOU FOR YOUR ATTENTION

QUESTIONS AND ANSWERS

References

[1] Cassell J. – Embodied Conversational Agents, AI Magazine Volume 22, no. 2, 2001

[2] Carrasco et al. – Natural Interaction between Virtual Characters and Persons with Alzheimer's Disease, International Conference on Computers Helping People]with Special Needs, 2008

[3] Kunc L. – Facial Animation, presentation in PhD. subject virtual reality, 2008?

[4] Ezzat T., Poogio T. - Videorealistic talking faces: A morphing approach, Cambridge, AVSP 1997

[5] Hák R. - Mobile 3D Talking Head, project web, http://maverick.webz.cz/3DHead/index.php