Graphics, 3D and virtual reality interfaces

Job Zwiers
University of Twente,
HMI

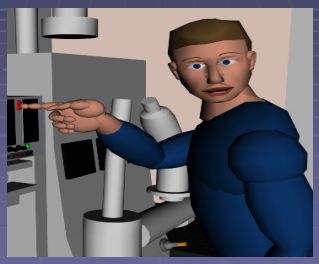
Overview

- Why Graphics &VR
- Modeling, 3D Graphics
- Avatars and Body Animation
- Where does it fit in the "big picture" of interfaces for Human Media Interaction

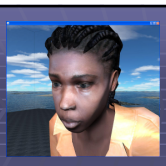
Why Graphics & VR

- User interfaces
- Training
- Simulation
- Games
- Serious gaming, decision support.
- Virtual presence
- Tutors
- Sales assistant

Instruction, explanation (Steve, USC/ISI)



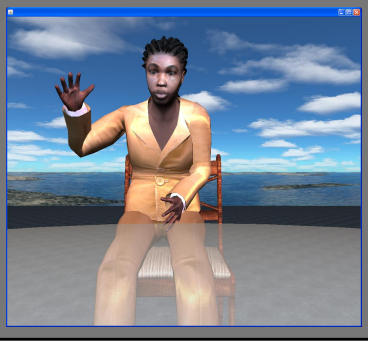
Simulation, tutoring



Remote Presence

Multimodal meeting support
for remote participants
AMIDA project

Remote meeting participant "Requests the floor"



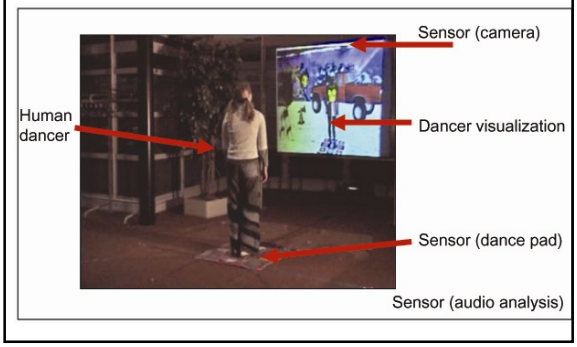
Serious gaming

Mission Rehearsal Exercise (Rickel et al, USC/ISI)

Tutoring, training: Virtual Conductor

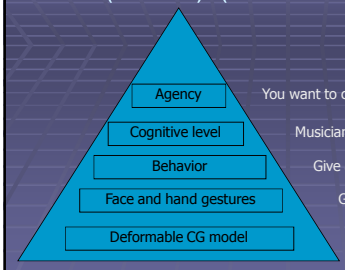
Virtual dancer

Virtual Dancer interacting with real dancer



Modeling, 3D Graphics

Layered control of E(mbodied)C(onversational)A(gent)s



You want to conduct a piece of music

Musician X starts lagging

Give a sign towards X

Generate a pointing gesture

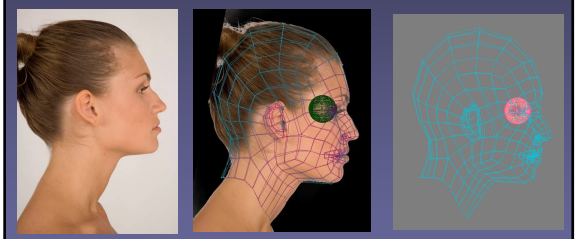
Generate deformation of the 3D mesh for the conductor avatar

Example of some event for the virtual conductor.

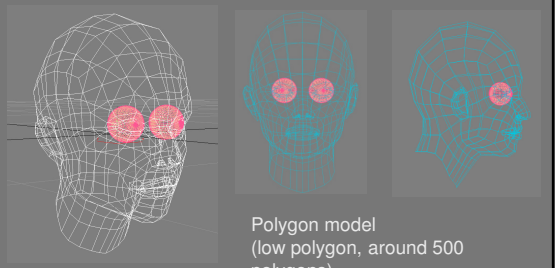
"Deformable CG models"

An impression of the techniques involved

Modeling low polygon model



Face geometry



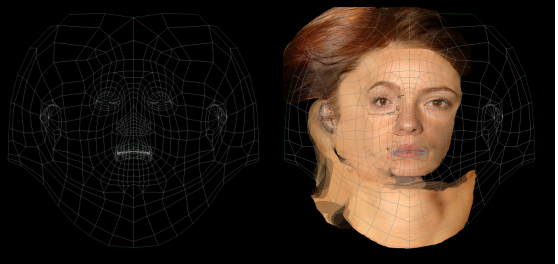
Polygon model
(low polygon, around 500 polygons)

Texturing

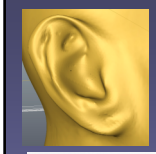
The photographs of the face must be combined into a 2D picture (the texture).
This texture is to be "wrapped around" the face geometry.

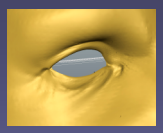
UV unwrapping:

mapping from 2D to the 3D face geometry



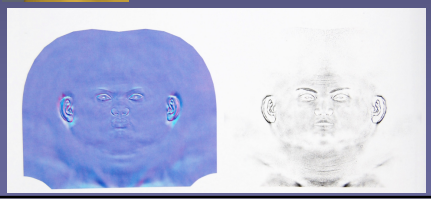
Normal maps, bump maps





Added detail…

exported in form of *normal map* or *displacement map*



Normal maps, Bump maps

Bump maps can create small scale detail, for example, skin pores.

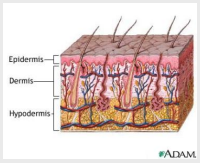
Technique: perturbation of surface normal vectors (rather than using very detailed geometry)

Diffuse and specular lighting maps

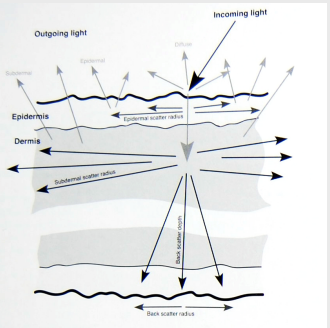
Diffuse light map
"Diffuse" here means: equal reflection in all directions.

Diffuse + Specular + Bump map
Specular reflection is stronger in direction of reflected rays. (Use normal vectors, direction of light and direction of viewer)

More refined skin modeling



"Real" skin consists of layers each with different material properties.
For instance, the epidermal layer contains no blood, filters mostly blue light.



Various subsurface layers

Epidermal reflection

Subdermal reflection.
Blue light is missing

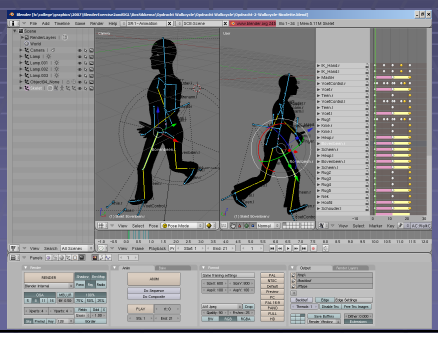
Combined reflection

Various combinations of epidermal and subdermal reflection, also including "backscatter"

Subsurface scattering included…

"Face and hand gestures" body animation in general

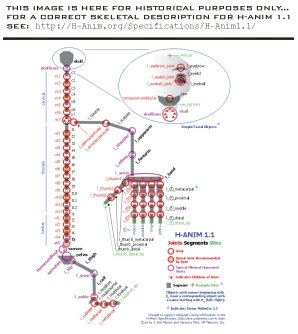
Body animation: based upon skeletons or "bone" structures



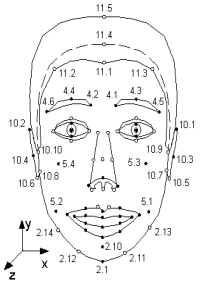
H-Anim and MPeg4

- HAnim is a standard for "humanoids" See http://h-anim.org
- It determines the names of human joints and bones, and the connection between them.
- MPeg4 builds upon HAnim, but also includes Facial expressions
- Aim is standardization, industrial acceptance (of VR aspects) fairly low

H-Anim skeleton



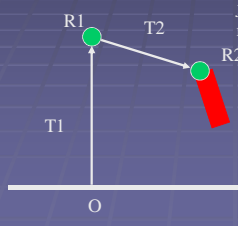
MPeg4 FDP/FAP



- FDP: Face Definition Parameters
- FAP: Face Animation Parameters
- There are similar BDP/BAP parameters, for *body* animation

Bone structures

Sketch of a bone structure with two joints (R1, R2), and three bones (the last one is shown in red)



Red: Segment attached to the R2 joint. Its coordinates are defined relative to the R2 joint

Transformation matrix for the "Red" segment:
T1 o R1 o T2 o R2

Skinning

- "**Skinning**" means the process where a mesh is associated with a bone structure.
- This can be done in a **simple way**: each limb defines its own mesh, and all vertices of one mesh are associated with exactly one bone.
- A **more sophisticated way** of skinning associates one ***or more*** bones with every vertex.
- The vertex position after joint rotations have been applied is a **weighted sum** of the rotated displacements M d, where the weights w_i determine the influence of each bone on that vertex.

$$\sum_{i=0}^{n} M_i d_i w_i$$

Rotation math

If the axis of rotation is **n**, and the angle is θ then the corresponding matrix is as follows:

$$M_R = \begin{pmatrix} 1-2(y^2+z^2) & 2xy-2sz & 2sy+2xz & 0 \\ 2xy+2sz & 1-2(x^2+z^2) & -2sx+2yz & 0 \\ -2sy+2xz & 2sx+2yz & 1-2(x^2+y^2) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$where:$

$s = \cos(\theta/2), \quad (x, y, z) = \sin(\theta/2)\,\mathbf{n}$

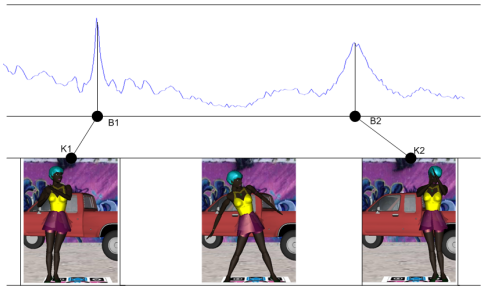
This is often represented by means of a quaternion q:
q = (s, x, y, z) = s + x**i** + y**j** + z**k**

Animating rotation

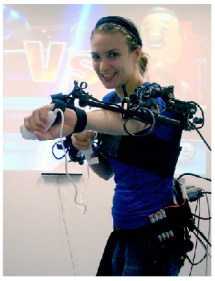
- Low level animation (OpenGL/Direct3D) is always based upon *matrices*
- Animation and interpolation is not done on the matrix level: a weighted average of two rotation matrices is not even a rotation matrix, but includes "skewing"
- Quaternion representation allows for good quality interpolation: "shortest path" and without unwanted acceleration/deceleration

Procedural Animation
Physical body simulation

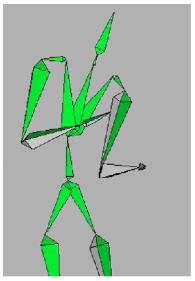
Motion analysis



Motion analysis by means of tracking devices

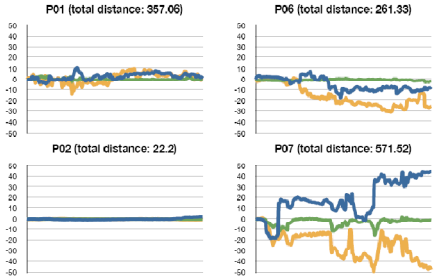


Participant in Gypsy 6 Motion Capture Suit

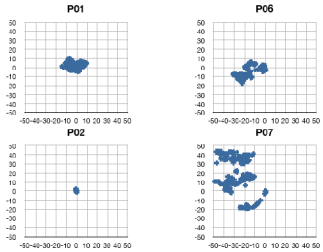


Output Motion Capture Data on Human Model

Movement of the body core in each dimension (results by Marco Pasch)



Movement of the body core top view



X-Axis:forward/backward movement
Y-Axis: lateral movement

Motion generation

Parameterized motion

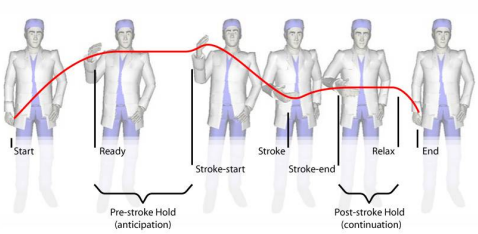
Predefined gestures need not be *fixed*: parameterization of procedural animation can be used to affect simple aspects like amplitude, or more complex aspects like "style" characteristics or even personality based aspects.

Physical Controllers

- Specifying animation can be time consuming and tedious, therefore, costly for industrial products like games
- *Physical simulation* can be used to calculate (part of) the body animation behaviour.
- **Idea**: the kinematically controlled parts (think: the arms of the conductor) result in forces and torques applied to the remainder of the body. Physical simulation can then be used to calculate the effects.
- Physical controllers allow for more "high level" goals to be specified, like "keep balanced", even when unexpected movements occur.

Body motion & Animation on a higher level

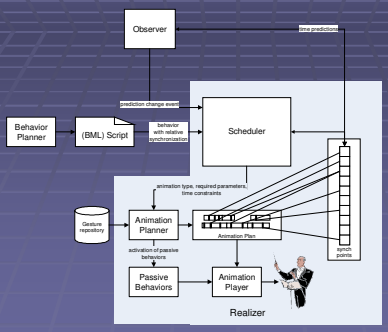
Phases in gestures



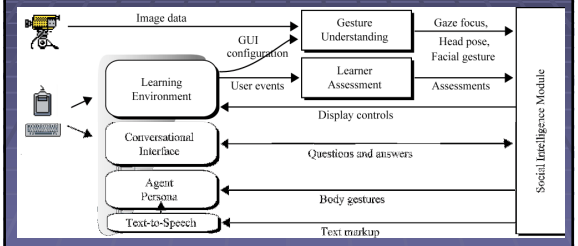
B(ody)M(otion)L(anguage)

```
<bml id="bml1">
  <gaze id="gaze1" target="AUDIENCE"/>
    <description level="1" type="RU.ACT">
      <target>PERSON1</target>
      <intensity>0.6</intensity>
      <lean>0.4</lean>
    </description>
  </gaze>
  <speech id="speech1" start="gaze1:ready">
    Welcome ladies and gentlemen!
  </speech>
</bml>
```

Architecture of the conductor



Architecture that incorporates some cognitive aspects and "agency"



Thank you