

Digital Characters in the Real World

A Review of Embodied Agents in Augmented Reality

Christian Graf



Hello and welcome!

My name is ... This presentation is about ...

- I haven't ever done a presentation in English before, so it's a challenge for me as it is for you, my audience.
- I decided to speak in English instead of in German because my work was all written in English, and because I want to practise the language in an academic setting.
- If I you don't catch some words or I don't explain some concepts clearly, feel free to ask, probably after I have finished the slide.
- As you know there will be enough time for discussion in the end.

Contents

1. Introduction
2. Embodied Agents
3. Augmented Reality
4. Evaluation of Existing Agents
5. Design Requirements
6. Prototyp Implementation
7. Research Directions & Conclusion



30.06.2004

Christian Graf - „Digital Characters in the Real World“

2 / 31

I have structured my presentation in 7 sections

1. Introduction
2. Embodied Agents
3. Augmented Reality
4. Evaluation of Existing Agents
5. Design Requirements
6. Prototyp Implementation
7. Research Directions & Conclusion

Basic Terms

1. ■ Human Computer Interaction - HCI
2. ■ User Interface - UI
3. ■ User Centric Design
4. ■ Usability
5. ■ Virtual Reality - VR
6. ■ Augmented Reality – AR
- 7.

30.06.2004 Christian Graf - „Digital Characters in the Real World“ 3 / 31

Not all of us are familiar with particular terms, therefore some short definition:

HCI: „Human-Computer Interaction is the study of the relationship that exist between human users and the computer systems they use in the performance of their various tasks” (Faulkner 1998)

UI: The in-between of technology and human user

What you actually see and handle

User Centric Design: Design Approach that takes the user as criteria

Troughout the whole process user is asked and concerned

Usability: Criteria to measure how good a product is for the end- user

• „good“ = Easy to use:

- easy to learn
- effective
- easy to remember
- error- free
- pleasant (*Nielsen 1993*)

good usability adds productivity and satisfaction

VR: synthetic place where the user is immersed

senses are controlled by the system, at least vision and audio

AR: mixture of reality and virtuality

virtual object exist in the real world

Setting

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.


■ HITLab NZ

- High-Technology research lab
- Orientation to real-world problems
- Exploration of potentials

■ Current focus

- Augmented Reality as User Interface
- Distance collaboration





Motivation

- Objective: Ease presentation of complex information through computers
- Background: Mismatch between humans and computers
- Question: Are human like interfaces favourable in Human-Computer Interaction?

⇒ Feasibility study

30.06.2004 Christian Graf - „Digital Characters in the Real World“ 5 / 31

Background:

- User Interface of Computers does not match humans' primary abilities:
 - Analogue vs. Digital
 - Sense, Thoughts, Emotions vs. Logic, Complexity, Computation
 - Interaction via speech, not alphanumerical sign!

Feasibility study:

- Research centered work which aims on ...
- => ... Identifying the potentials and future research directions

1.

2.

3.

4.


5.

6.

7.

Research Task

„What is the relation between embodied agents, augmented reality and collaboration? Review the literature and compile a comprehensive report with appropriate references about your findings! Implement and possibly test a prototype! What are your recommended directions for further research?“



30.06.2004

Christian Graf - „Digital Characters in the Real World“

6 / 31

Background:

Nobody in the HITLab had done anything like that before.

No expertise in embodied agents.

Big project on remote collaboration was due to be given to the Lab

Application needed some kind of ‚edgy‘ but useful gimmick

Idea: Virtual human as remote presentator?!

Agents

- 1.
2.
 - Definition
 - Why agents?
 - Application Areas
 - Taxonomy
- 3.
- 4.
- 5.
- 6.
- 7.

30.06.2004 Christian Graf - „Digital Characters in the Real World“ 7 / 31

Definition:

A software program, that is:

- reactive (sensing and acting), i.e. responds in a timely fashion to changes in the environment;
- autonomous, i.e. exercises control over its own actions;
- goal-oriented (pro-active, purposeful), i.e. does not simply act in response to the environment;
- temporally continuous, i.e. is a continuously running process.

Why agents:

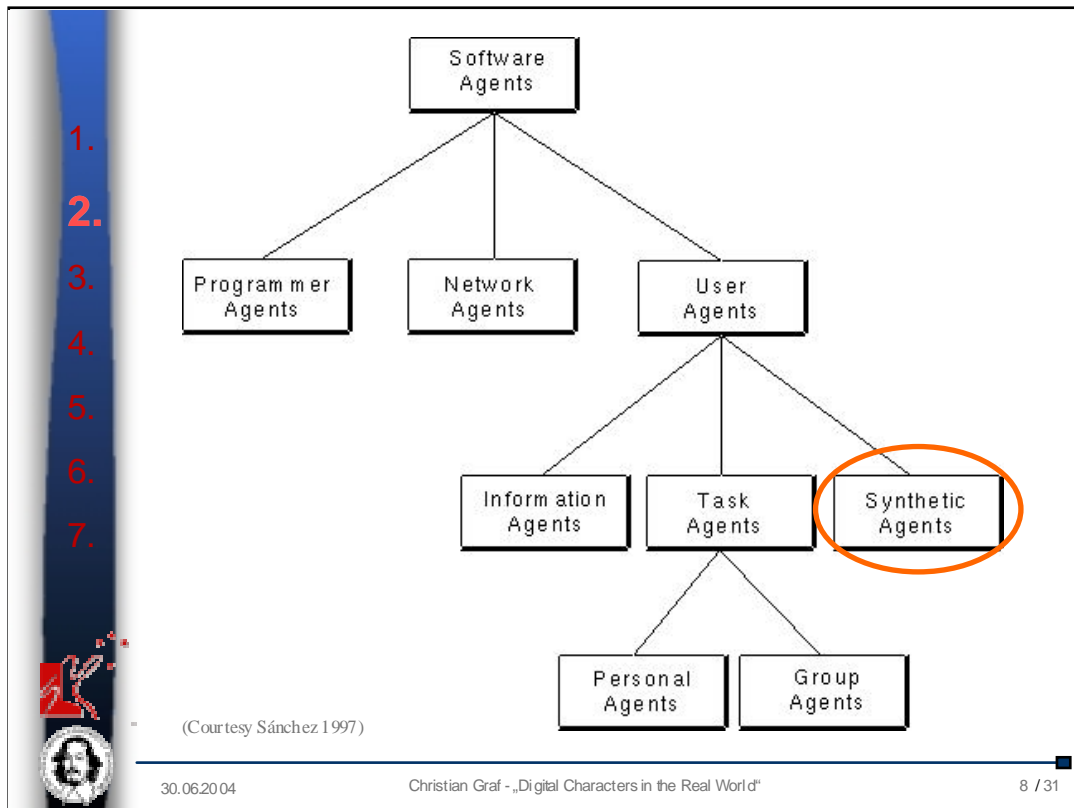
- they should take workload from the human!

Application Areas

agent = „software assistants“

in telecommunications, network management, electronic commerce and information gathering, and user interfaces

in UI: designed to be perceived directly by end users (contrasting to the other purposes)



There is no such thing as THE agent!

This hierarchy shows the diversity of the area

Programmer Agents:

an abstraction to conceptualize, design and implement complex systems

Network Agents:

autonomously migrating entities that act on behalf of network nodes in a distributed environment

User Agents:

an abstraction for end users to interact with computer systems

Information Agents: help users deal with information spaces that are typically unorganized and highly dynamic

Task Agents: help users perform computer-supported task

Synthetic Agents:

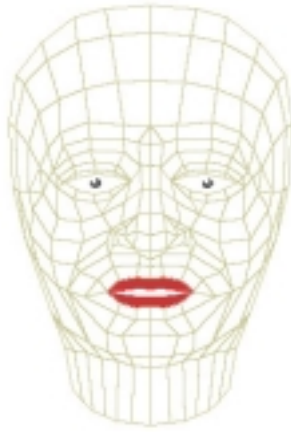
- create engaging environments for users
- lifelike characters into the computer interface
- promote a *suspension of disbelief* from the user
- illusion of autonomous, animistic entities

My focus: ,synthetic agents‘ aka ,conversational agents‘ aka ,embodied agents‘

DECFace

Gandalf

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.

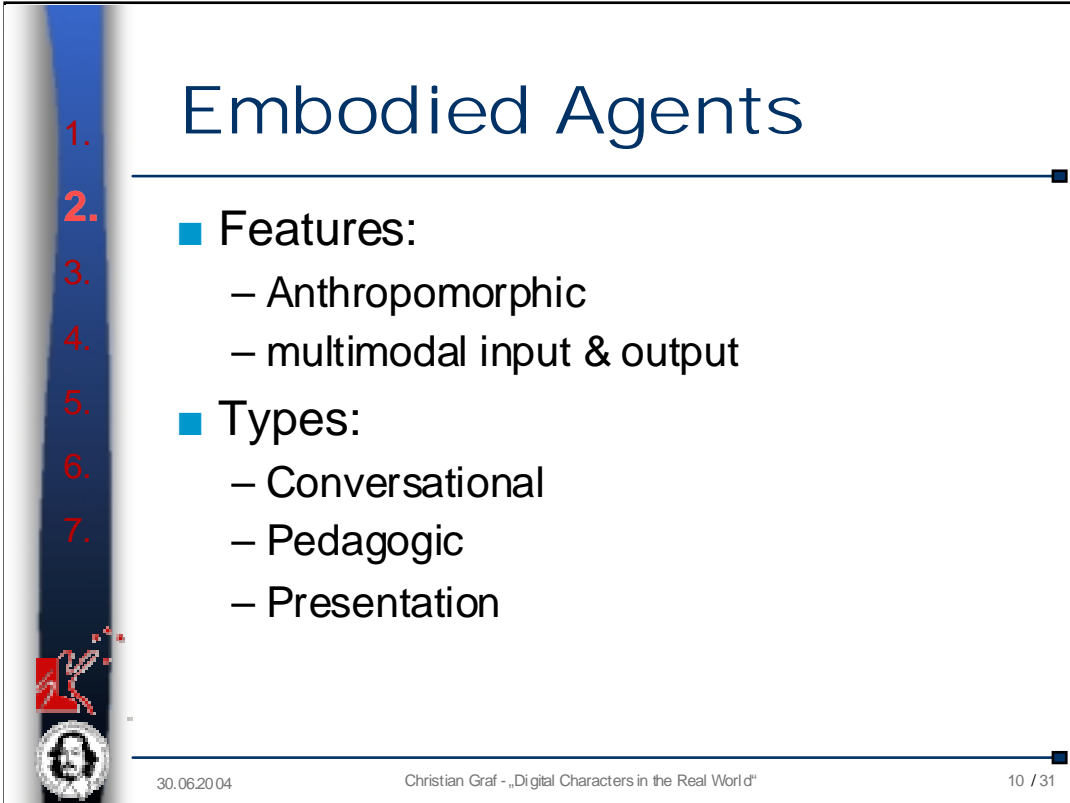


Waters & Levergood 1993



Thórisson 1996





Embodied Agents

- Features:
 - Anthropomorphic
 - multimodal input & output
- Types:
 - Conversational
 - Pedagogic
 - Presentation

30.06.2004 Christian Graf - „Digital Characters in the Real World“ 10 / 31

Features:

Anthropomorphic = „human-like“

not necessarily photo-realistic

comic-likeness levels expectation of humans

multimodal = uses of more than one sense

e.g. speech with gestures and emotional display

Types:

Conversational: language focused

aim: engage humans in conversation

Pedagogic: planning focused

aim: teach humans about a topic

Presentation: mixture of both - needs language to achieve a planned goal

aim: present new information to user and engage him with that content

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.



Olga © KTH Centre for Speech Technology, Stockholm





1.
2.
3.
4.
5.
6.
7.

STEVE © Information Science Institute, University of Southern California

30.06.2004 Christian Graf - „Digital Characters in the Real World“ 12 / 31

Contents

1. Introduction
2. Embodied Agents
3. Augmented Reality
4. Evaluation of Existing Agents
5. Design Requirements
6. Prototyp Implementation
7. Research Directions & Conclusion



30.06.2004

Christian Graf - „Digital Characters in the Real World“

13 / 31

We have learned about:

What an agents is and can do.

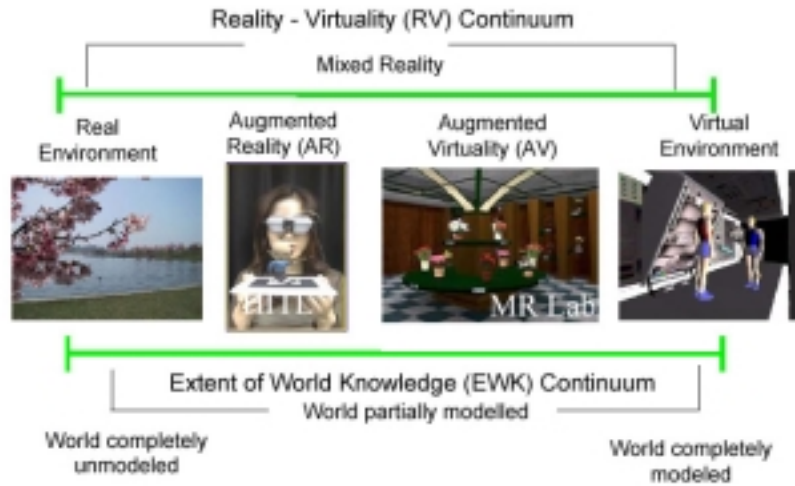
Agents as UI

- embodiment

Agents in VR

RV - Continuum

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.



Adopted after Milgram et al. 1994



1.
2.
3.
4.
5.
6.
7.

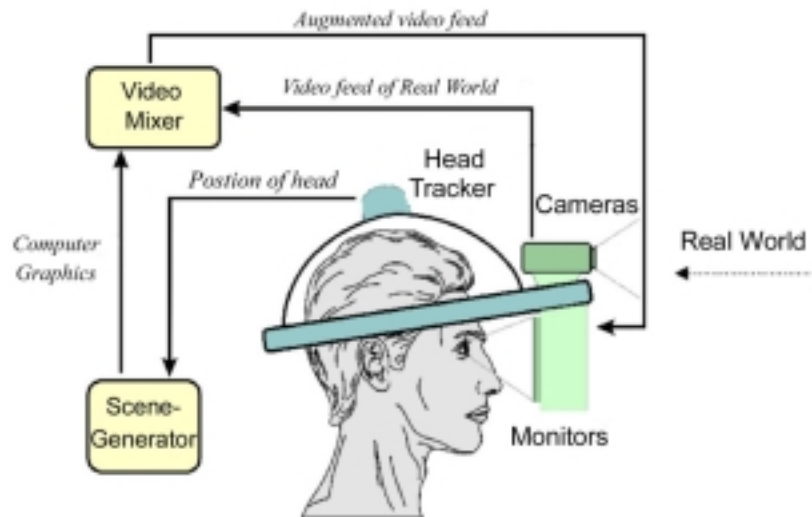


© V. Paelke, Institut für Kartographie und Geoinformatik, University of Hannover

30.06.2004 Christian Graf - „Digital Characters in the Real World“ 15 / 31

Augmented Reality I

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.



© V. Paelke, Institut für Kartographie und Geoinformatik, University of Hannover

30.06.2004

Christian Graf - „Digital Characters in the Real World“

16 / 31

User interacts with real and virtual objects

Augmented Reality II

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.

■ Properties

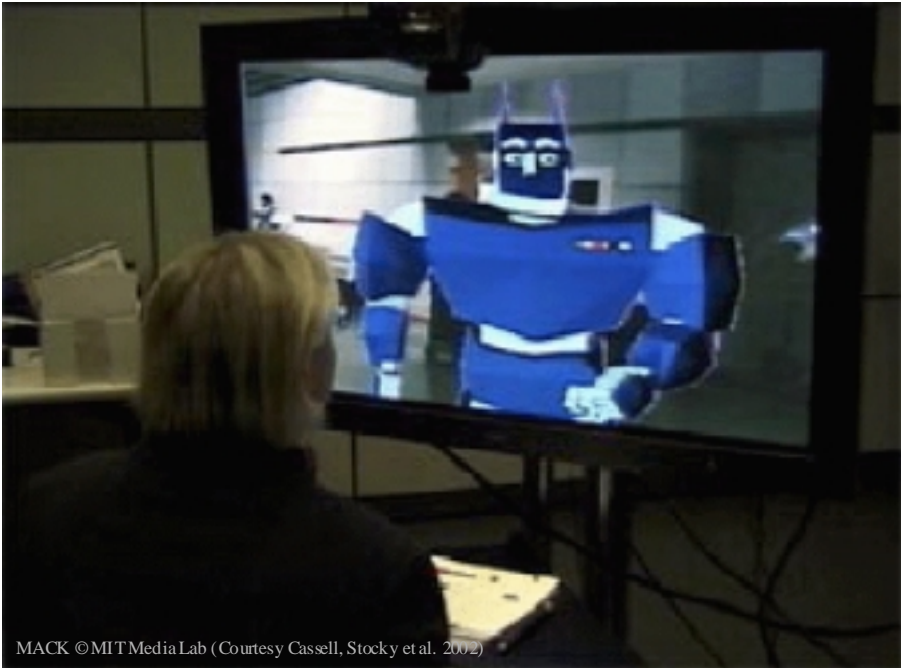
- Combination of reality & virtual objects
- 3D real environment & real time
- Enhanced interactivity with all senses
- Virtual immersion: partly

■ Realisation:

- Head Mounted Displays or screen based



1.
2.
3.
4.
5.
6.
7.



MACK ©MIT Media Lab (Courtesy Cassell, Stocky et al. 2002)

30.06.2004 Christian Graf - „Digital Characters in the Real World“ 18 / 31

The image shows a person with short blonde hair sitting at a desk, looking at a large computer monitor. On the monitor, a blue, stylized robot character with a white face and large eyes is displayed. The robot has a boxy, friendly appearance. The background of the screen shows a simple, brightly lit room. The person is seen from the back, and their hands are resting on the desk. The overall scene is set in what appears to be a laboratory or office environment.

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.



Courtesy R. Torre, EPFL, Suisse

30.06.2004

Christian Graf - „Digital Characters in the Real World“

19 / 31

1.
2.
3.
4.
5.
6.
7.

Evaluation

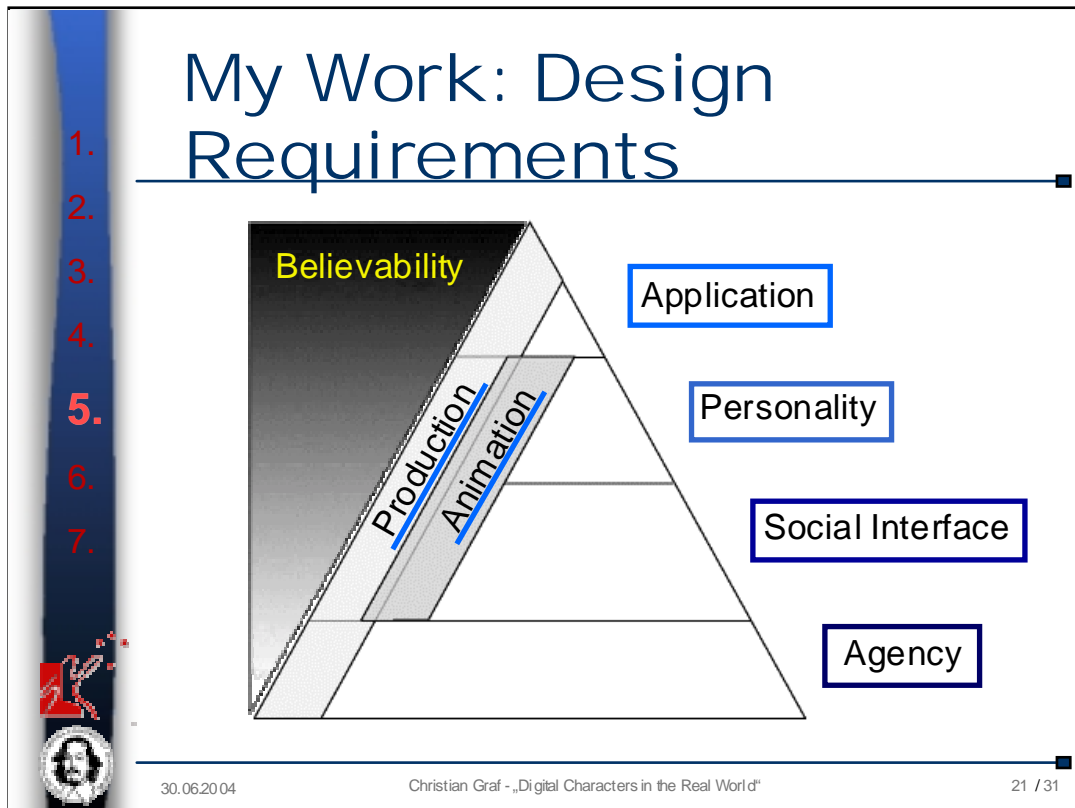
■ Talking Heads	Embodied Agents
– DECFace	- Cosmo
– Gandalf	- REA
	- MACK

Result: Anthropomorphic agents have positive effects on certain tasks

30.06.2004 Christian Graf - „Digital Characters in the Real World“ 20 / 31

- DECFace: „people’s responses could be predicted from known principles of human behaviour“ (Nass & Steuer 1993)
- GANDALF: new feature: turn-taking
users „prefered such system to another embodied character capable of only emotional expressions“ (Cassell, Bickmore et al. 1999)
- Cosmo: new feature: deictic expression
„most participants understood the agent’s advice most the time“ (Towns)
deixis framework produced clearer expressions
- REA: new feature: handles repairs gracefully
system is sensitive to the function of language und non-verbal behaviour
seemed to improve the interaction, no qualitative statements
- MACK: most modern system with turn-taking, deixis, non-verbal behaviour
„user trusted in information the system presented to them“ (Cassell et al. 2000)

Positive Effects: agent increases presence -> involvement -> commitment



Agency: basic techniques to ensure autonomy, reactivity, responsiveness, reliability, efficiency, goal directness ...

Social Interface: speech, non-verbal behaviour, human-likeness

Personality: individuality to avoid dullness
personal memory to recognize former customers

Application: training towards the domain of the application
- specific knowledge, behaviours, and language

Animation: human figure animation
- facial animation & body animation

Production: consistency
timing

Criteria & Goal: Believability - the suspension of disbelief

Contents

1. Introduction
2. Embodied Agents
3. Augmented Reality
4. Evaluation of Existing Agents
5. Design Requirements
6. Prototyp Implementation
7. Research Directions & Conclusion



30.06.2004

Christian Graf - „Digital Characters in the Real World“

22 / 31

We have seen:

- what AR means

- what VR and AR agents exist

- what requirements an AR agent has to meet

In the following sections

- we will introduce a prototyp implementation

1.
2.
3.
4.
5.
6.
7.

My Prototyp Implementation

- The Goal
- Possible approaches:
 - Wrap up from scratch: Pros & Cons
 - Use existing architecture: Pros & Cons
- The approach
- The solution
 - extending BEAT

30.06.2004 Christian Graf - „Digital Characters in the Real World“ 23 / 31

Goal: An AR agent for presentations

Task: Define architecture

Implement functioning demo system

Approach 1: Wrap up from scratch

Pros: application centered specialisation

Cons: too difficult (complete, sound architecture to be developed!)

too time consuming (concept, programming, tests etc.)

Approach 2: Use existing architecture

Pros: basic functionality already implemented

stable, mature system without bugs

Cons: details of architecture do not match needs of application

My approach:

find appropriate existing architecture

bring it to AR

adopt & train it on desired application

My solution:

extending BEAT

=> it took weeks to do research

=> and another three months to get the desired software



First Approach:

VRML figure (H-Anim) rendered through AR-Toolkit

ART: searches for markers
determines position in visual field
aligns geometry to that position
(for details come and see me!)

Pro: no integration problems

Contra: low resolution
combination with timed speech & body movements difficult

1.
2.
3.
4.
5.
6.
7.

BEAT

- „Behavior Expression Animation Toolkit“
- input: text
- analyses & extracts linguistic and contextual information
- adds information for animation
- output: appropriate non-verbal behaviour to accompany speech


30.06.2004 Christian Graf - „Digital Characters in the Real World“ 25 / 31

Technical details:

Part-of-Speech Tagger

Lexikal Dictionary - WordNET

Phantomime - Open Inventor based animation engine



Problems

- Implementing the agent:
 - Software licences
 - Other researchers latency
 - Appropriate format of body description
 - Architecture mismatch
- Using the agent:
 - Registration in the world
 - Placement in user's visual field
 - Determination of an appropriate size

30.06.2004 Christian Graf - „Digital Characters in the Real World“ 26 / 31

Implementation

Software Licences:

- no (free) compact SDK at that time
- huge bundle of different parts necessary
- commercial products cost money!

Researchers latency

- no policies about licenses of own research results
- personal reasons (PhD thesis etc.)

Appropriate format of body description

- VRML: H-Anim
- possible MPEG (includes H-Anim body description)

Architecture mismatch:

- OI works on a scenegraph where events trigger actions
- ART works continuously on video stream

Using the agent

Registration: size of markers & quality of video feed cannot afford world alignment -> agent is screen aligned!

Questions: how obstrusive is the agent?

- > where should it be placed?
- > How big should it be?

1.
2.
3.
4.
5.
6.
7.

Digital Character in the Real World



30.06.2004 Christian Graf - „Digital Characters in the Real World“ 27 / 31

Agent's properties:


- lip-synch speech
 - emotion display through eyes, eye brows, eye lids, arms, hands, and body posture
 - gestures:
 - beat (supporting)
 - iconic („as-if“)
 - simulation of lifelikeness in idle situation:
 - eye-lids close randomly
 - body with Perlin noise
 - fix position in visual field
 - size like a real human
 - knee-upward representation -> hands need to be seen
- => impression of a personal assistant!**

Results

1.
2.
3.
4.
5.
6.
7.



- Social Agents are advantageous
 - for education, presentation and others
 - generally: in soft skills focused areas
 - AR is more acceptable than VR by users
 - Good chance to be accepted as intuitive interface
 - Implementation is feasible
 - Communicative agent was brought to AR
 - ‚Natural‘ presenter possible
-



Future Directions

- **Research**
 - Culture invariant non-verbal behaviours
 - Presence in mixed realities
 - Recognition of gestures (deixis, iconic, ...)
 - Spatial audio for multi user settings
 - Development of standard user tests
- **Implementation**
 - Bring BEAT to open source
 - Migrate from VRML to MPEG-4

30.06.2004 Christian Graf - „Digital Characters in the Real World“ 29 / 31

Research:


Preconditions

Work on Agent & Modalities

Postconditions

Implementation:

Converting BEAT to open source



Lessons learned

- About the subject
 - Acquired sufficient expertise in the area
 - Prototyp needed lots of resources
 - First results demand testing
- Personal conclusion
 - Insight how research community works
 - Fostered research interest in AR & UI

⇒ **AR is great! Interaction is fun!**

30.06.2004 Christian Graf - „Digital Characters in the Real World“ 30 / 31

Subject:

Enough time to read through piles of papers

-> good overview of core technologies, problems and related subjects

Prototyp:

needs lots of resources to be assembled (many programs!)

Personal

Gained insight in academic world

Fostered wish to do research

Thank you for
your attention!

For further questions ask now or
come and see me later!



References



AR ToolKit Homepage: <http://www.hitl.washington.edu/artoolkit/>

BEAT Homepage: <http://gn.www.media.mit.edu/gn/project/beat/>

Azuma, R.T. "A survey of Augmented Reality". In *Presence: Teleoperators and Virtual Environments*, Vol 6, No. 4: 355-386

Azuma, Bailiot, Behringer, Feiner, Julier, MacIntyre: „Recent Advances in Augmented Reality“, In *IEEE Computer Graphics and Applications*, Vol. 21, No. 6: 34 - 47

Cassell, J., Vilhjálmsson, H.H., Bickmore, T. (2001). "BEAT: the Behavior Expression Animation Toolkit". In *Proceedings of SIGGRAPH 2001*

Cassell, J., Stocky, T., Bickmore, T., Gao, Y., Nakano, Y., Ryokai, K., Tversky, D., Vaucelle, C., Vilhjálmsson, H. (2002). "MACK: Media lab Autonomous Conversational Kiosk". In *Proceedings of IMAGINA 02*, Monte Carlo

Kato, H., Billinghurst, M. (1999). Marker Tracking and HMD Calibration for a video based Augmented Reality Conferencing System. In *Proceedings of the 2nd International Workshop of Augmented Reality (IWAR 99)*.

Milgram, P., Takemura, H., Utsumi, A. and Kishino, F. (1994): "Augmented reality: a class of displays on the reality- virtuality continuum". In *Telemanipulator and Telepresence Technologies*, SPIE Vol. 2351: 282- 292

Paelke, V. (2004). Augmented Reality - Einführung. Institut für Kartographie und Geoinformatik. University of Hannover. http://www.igk.uni-hannover.de/lehre/katalog/Augmented_Reality/pdf-files/VL-AR-04-01-b.pdf

Sanchez, J.A. (1997). A Taxonomy of Agents. Tech. Rep. No. ICT-97-1, Interactive and Cooperative Technologies Lab, Universidad de las Américas-Puebla, Cholula, Puebla, México. <http://ict.udlap.mx/people/alfredo/ihc-o99/clases/agentes.html>

Sundblad, O. and Sundblad, Y. (1998). OPGA - a Multimodal Interactive Information Assistant. CHI 1998

Thorisson, K.R. (1997). Gandalf: An Embodied Humanoid Capable of Real-Time Multimodal Dialogue with People. First ACM International Conference on Autonomous Agents. Marina del Rey, California, February 5-8, 1997: 536-537

Waters, K. and Levergood, T.M. (1993). DECface: An Automatic Lip-Synchronization Algorithm for Synthetic Faces. Technical Report CRL 93/4. Digital Equipment Corporation, Cambridge Research Lab, Cambridge

