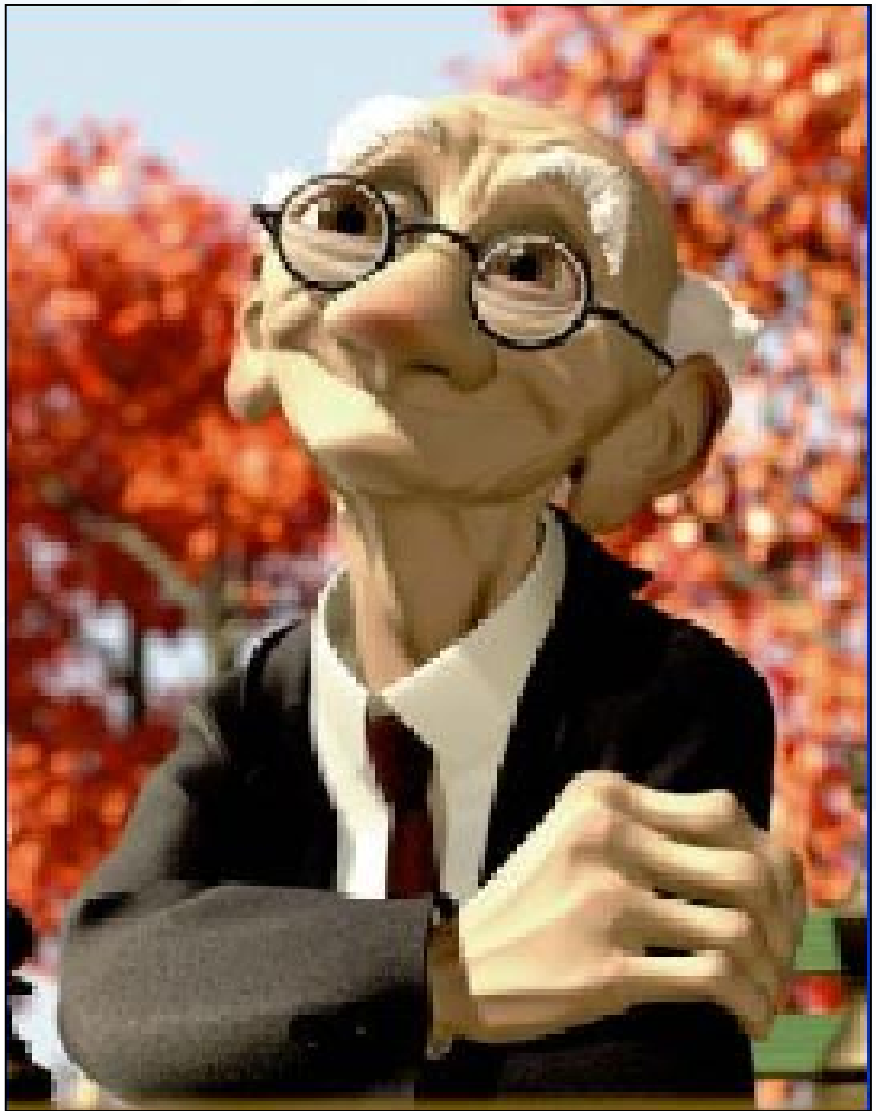
Sound/Audio
or
Listening to Geri‘s Game

Andrej Ferko

Comenius University, Bratislava, ferko@fmph.uniba.sk

www.sccg.sk/~wega

Geri's Game/Sound



Creating Web Graphics

You and Andrej Ferko

Class and Individual Work

Agenda

- Project Specification: Your Project
- Web Page Life Cycle – revisited
- Human Perception of Sound
- Listening to Geri's Game

Project Specification

- IS 9001
- Quality management
- Economic, efficient, safe, precise solutions
- Quality measures – formal, informal
- Number of copies sold, downloads, Google count by Dusan Hamar, page hits…

Project Questions

- Goal, requirements, project decision
- Functionality specification
- L. Weinmann recommends…
- www.web-redesign.com
- Qs on client, users, project, activas, content, maintenance, budget, scheduling

Project Profile

- L. Weinmann recommends…
- www.secretsites.com/profiler/set-partnering.html
- INVESTIGATION…
- … FBI, detective story metaphor

Agenda

- Project Specification: Your Project
- Web Page Life Cycle – revisited
- Human Perception of Sound
- Listening to Geri's Game

Popular Audio Example

- Ladovska zima – a song by J. Nohavica
- A hit in a few days at CZ/SK web

Sound Spaces Motivation

- Why Sound within Computer Graphics?
- Sound vs. Image
- Human Understanding Sounds
- An Example - Data Sonification
- Sound Spaces Construction
- Sounds On-line and Authoring Tools

Why Sound within CG ?

- Historical Reasons: BELL, ASCII 7
- GUI: eye, ear, other senses
- Visually Impaired & Blind People
    - Braille line instead of display...

HUMANISATION of IT

- Motivation: What sound can do?

Sound Uses

- more information
- extending the scene
- emotional content
- real sounds from model and reality
- visually impaired or poor display
- sound as iconic, indexed, symbolic representation of meaning, plus signal,
- using speech, music and non-speech audio cues (beeps, noises)

The Graphics Pipeline Steps

input devices

modeler

Modeling

transformations

projection

clipping

visibility

Rendering

shading

rasterization

Output

device drivers

output devices

The Sound Rendering Steps

input devices

modeler

Modeling

A/D (1D)

requantization

filtering

compress

Rendering

segmentation

(recognition)

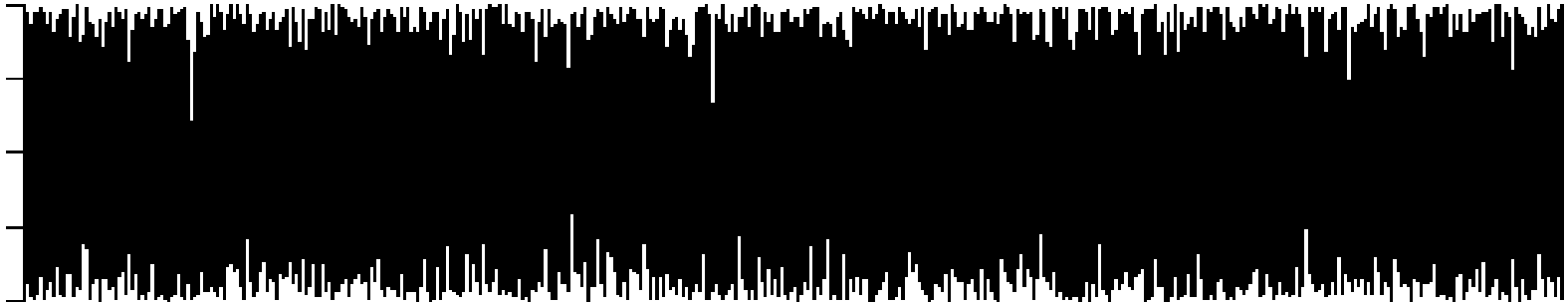
Output

device drivers

output devices

Examples by P. Bourke

- White noise, quite black



Many tools for each phase

- Sound processing tools:
- Transformations, D/A, A/D conversions
- Fourier transform, convolution, filters
- Wavelet transform
- Compression, databases, archives, MP3...

Sound vs. Image

- Sound is temporal, image is static
- Good eyes: hawk: predators, good ears: bunny
- Sound from parent layer, image from child one
- Digital sound and analogue one
- Music understanding differs with culture
- Curve and melody… tones and colors...
- … Computer Aided Musicology

User Interaction Model etc.

- Prompt, measure, trigger, input data record, echo, acknowledgement, processing, prompt...
- Request, sample and event mode
- 6-7 logical input devices
- GUI: 1D, 2D (WIMP), 3D (noimmersive and immersive solutions)
- menu-choice tree (acc.) or hypertext

Architecture of Multimedia System

- A. Data bus for structured pictures
- B. Data bus for unstructured images
- C. Data bus for structured sounds
- D. Data bus for unstructured, natural sounds

- Input for A & C: model, data and functionality
- Input for B & D: scanner and microphone

- [Stuc91] STUCKI, P.: Graphics and Multimedia, tutorial at Eurographics Conference, Vienna 1991

Architecture of Multimedia System

- **A.** Data bus for structured pictures VRML, CGM, SVG
- **B.** Data bus for unstructured images JPG, GIF, PNG
- **C.** Data bus for structured sounds MIDI
- **D.** Data bus for unstructured, natural sounds MP3

- Input for A & C: model, data and functionality
- Input for B & D: scanner and microphone

- [Stuc91] STUCKI, P.: Graphics and Multimedia, tutorial at Eurographics Conference, Vienna 1991

Architecture of MM System II

- A to B functional unit is Computer Graphics.
- B to A functional unit is Image Analysis.
- C to D functional unit is Sound Synthesis.
- D to C functional unit is Sound Analysis.

- No model cases:
- B to B is Image Processing: image to image.
- D to D is Sound Processing: sound to sound.

- [Stuc91] STUCKI, P.: Graphics and Multimedia, tutorial at Eurographics Conference, Vienna 1991

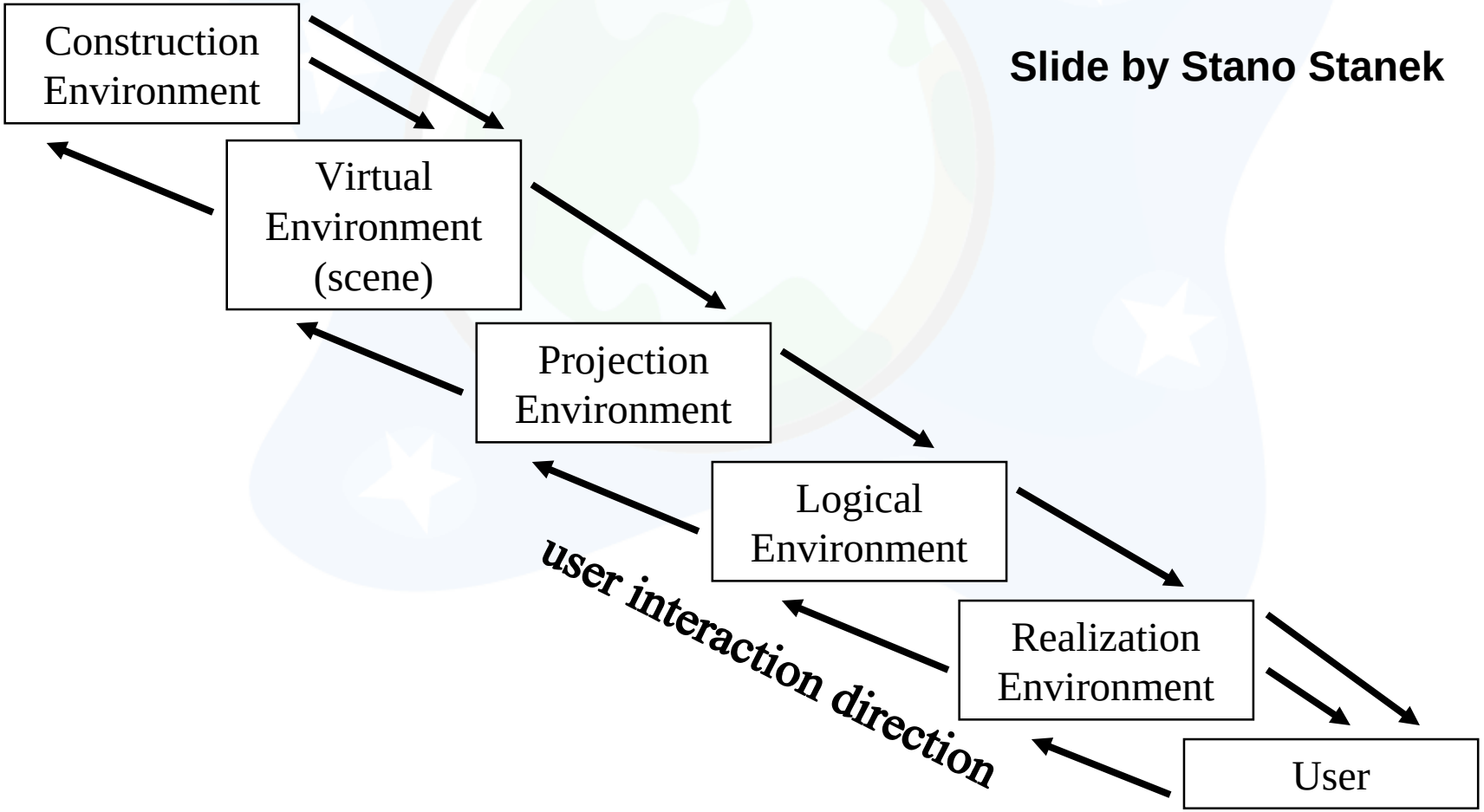
Sonification

- B to D images to sounds
- D to B sounds to images

- If we resign on error/model, we can use
- any of the functional units for VISUALIZATION or SONIFICATION

- Their goal is to enhance UNDERSTANDING.

PREMO Environments



Slide by Stano Stanek

Areas of Information Processing

input \ output	description	image	sound	smell	taste	touch
description	Symbolic manipulation	Computer graphics	Sound synthesis	Smell synthesis	Taste synthesis	Touch synthesis
image	Image recognition	Image processing	Sonification			
sound	Sound recognition		Sound processing			
smell	Smell recognition			Smell processing		
taste	Taste recognition				Taste processing	
touch	Touch recognition					Touch processing

Slide by Stano Stanek

Four General Criteria

- **ISO/IEC JTC1 Strategic Policy Statement formulated in 1994 the strategic objective as follows: to promote world-wide economies and efficiencies and global trade by creating an international standards environment that will enable IT suppliers to provide IT users with timely means to manage information efficiently, economically, accurately and securely. The strategic criteria give four "external" limitations to each medium: from efficiency to security.**

PREMO Example = Your room

- “Pilot” Example of Multimedia Object
- PREMO international standard
- MPEG-4 and MPEG-7 projects

- Object is OOP/CORBA and has data containers

Web Page Perception

- Document
- Painting
- Radio
- Theater
- Movie
- Human (audio)visual system
- VR
- Interactive & Adaptive Hypermedia
- IT product
- Legal entity
- Future avatar

WWW & XML >> WWD

- Million User Interface
- VRML 1.0 and VRML 2.0 and X3D
- VRML ECMA Script, VRML EAI...
- Data Mining
- Collaborative Hypermedia, Virtual Sculpting, MUDVR
- MPEG-4 & MPEG-7 Goals
- Content Age, Semantic Web

Communication

- Original <---> Recipient
- Original … Two recipients
- Incomplete original
- Two parts of original, two recipients
- No original => Model, representation
- No model => Darstellung, Ostension
- Knowledge direct or indirect
- Metacommunication, semiotics

– Optimize the download time

Directing the eye trajectory

- What happens the second 10 seconds, after the first click?

- head tracking
- eye tracking - entry point & trajectory
- vision
- cognitive processes
- and memory, etc. etc. etc.

Directing the E. Trajectory

- Static techniques used in painting
- Dynamic techniques used in theatre
- Web page as the fourth wall
- Goal oriented using curtain, actors, story and scenography, lighting and sound space - and interaction
- Web page is not 2D: structuring

Directing the E. T.

- The simplest case: plain text
- The page is presented as a book
- Documents and DTP rules, TEX
- Web page is ~2D: structuring texts
- Directing of reading, index, links

Directing the ear trajectory

- The simplest case: radio
- The page is presented as a 1D sound stream
- Listening to the read document

- Directing of listening, index, links,
- search, rewind, repeat...

Semiotics

- BELL (sound) is a signal
- DRVD is a symbol of music
- ----------
- Sound categories: speech, music, non-speech audio cues (beeps and buzzes)

NLP about BELL & DRVD

- BELL is an audio input
- DRVD is both symbolic input and video/visual input

- INPUT means the type of human input channel here

NLP: Neuroliguistic Programming

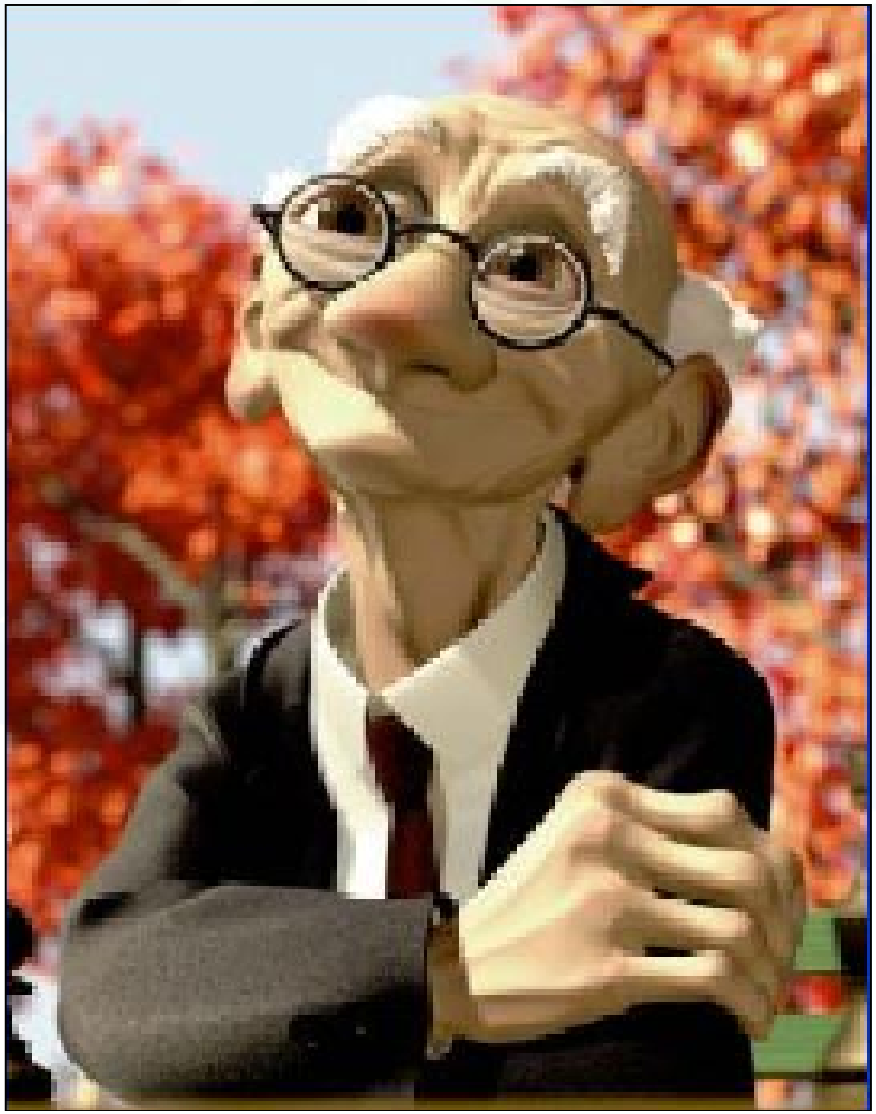
- [OCon89] O'CONNOR, J. - SEYMOUR, J.: Introducing Neuro-Linguistic Programming, Lambent Books 1989
- **The types of human input channels: VAKOG and S**
- video, audio, kinesthetic, ophtalactic and gustative - the first signal system
- symbolic - the second signal system

Semiotics

- **Imitating reality >> iconic representation.**
- **Index, indexed representation. Pars pro toto.**
- **Convention >> symbol.**
- **Signal, context dependent.**

- ***Examples:***
- ***icons, A for all letters, names, dialogue box.***

Geri's Game/Sound



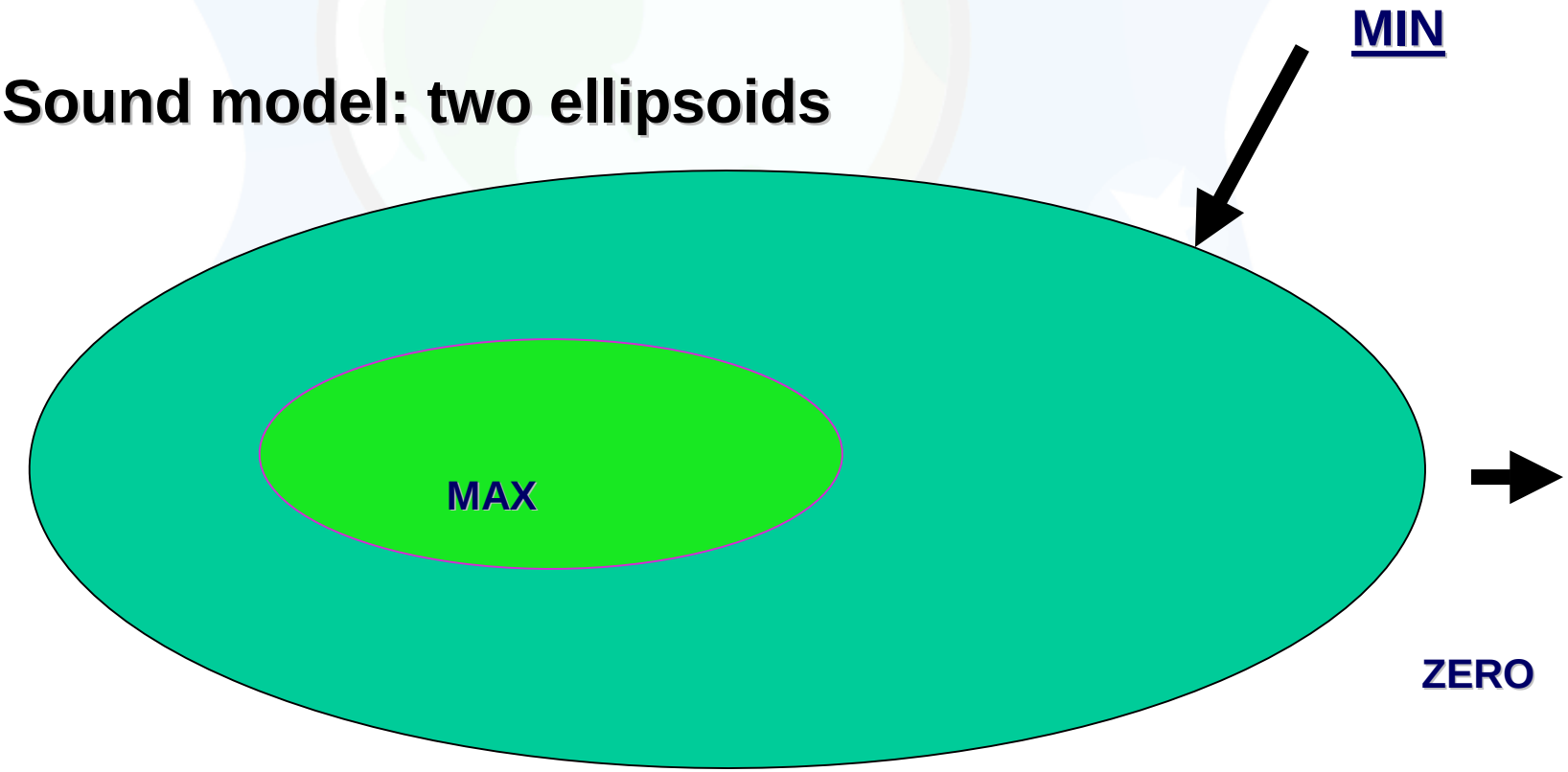
Geri's Game/Sound

- Music
- „Speech“
- Real Sounds
- Symbolic Sound

VRML 3D Sound

Sound source: any sound file (MIDI, MP3)

Sound model: two ellipsoids



The VRML Sound Node

- **The syntax of the Sound Node:**

- **Sound {**
- **exposedField SFVec3f direction 0 0 1**
- **exposedField SFFloat intensity 1**
- **exposedField SFVec3f location 0 0 0**
- **exposedField SFFloat maxBack 1**
- **exposedField SFFloat maxFront 1**
- **exposedField SFFloat minBack 1**
- **exposedField SFFloat minFront 1**
- **exposedField SFFloat priority 0**
- **exposedField SFNode source NULL**
- **field SFBool spatialize TRUE**

- **}**

3D SOUND ON-LINE

- **http://www.dform.com/inquiry/tutorials/vrmlaudio/vrmlaudio1.html#Sound**
- **http://www.dform.com/inquiry/tutorials/vrmlaudio/vrmlaudio4.html#Websound**
- **http://www.wareing.dircon.co.uk/3daudio.htm**

Key Source for VRML

- **Introduction to VRML-97**
- by D. R. Nadeau et al., SIGGRAPH 98 Course Notes (CN), 511 slides+, vrml97_s98.zip ~ 7.5 Mega

- ***http://www.siggraph.org/education/materials/…***

- **Easy access -> WEGA page, AF page**

3D Sound Applications

- **Standard ones:**

e-shops, home theater…

- **Exceptionaly beautiful one:**

Interactive 3D Sound storytelling for blind children

- **e-accesibility**

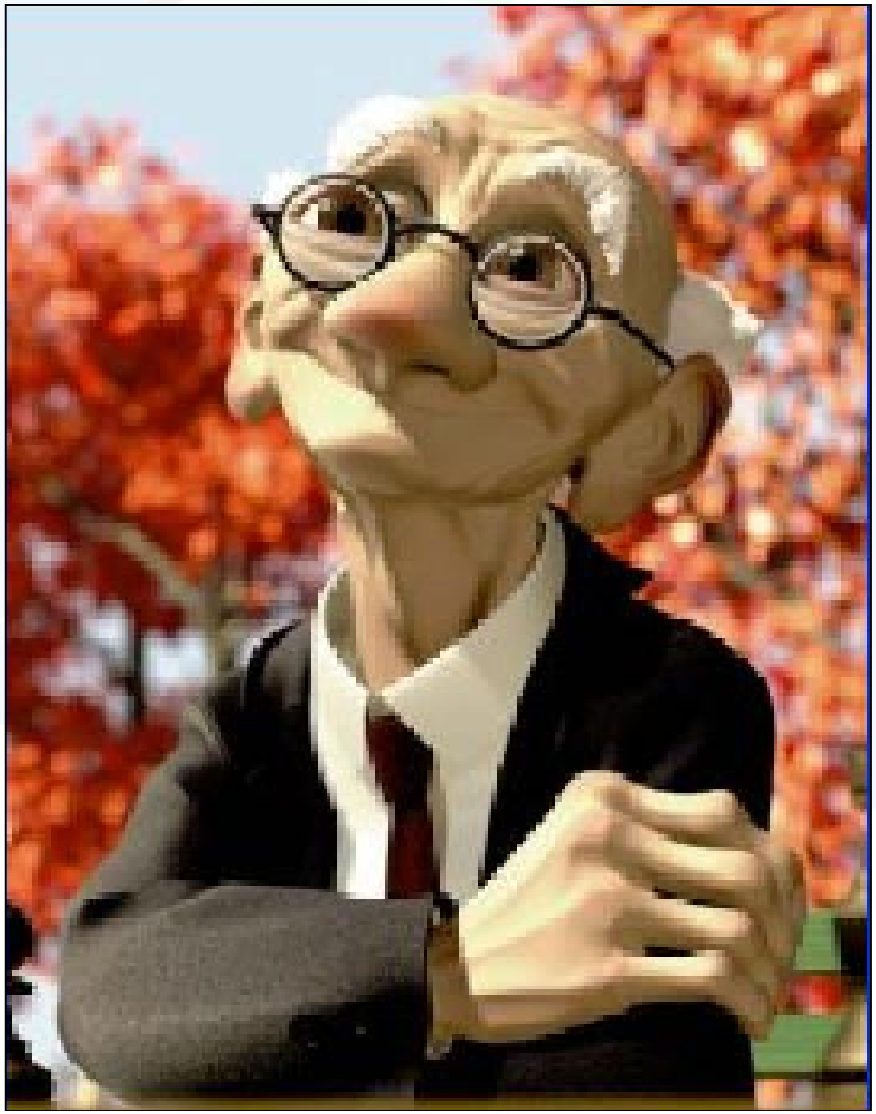
Sound Spaces

- GUI, application
- sound source tree
- attributes: volume, stereo field, frequency, rhythm, timbre...
- 2D or 3D grid sonification
- WIMP sonified (T. Fora), earcon
- keyboard modes: description, operating

Scientific Sonification

- DNA listening tool
- volume rendering surface properties: normals, curvatures, gradient sign, etc.
- KYMA language & SIGGRAPH CN
- pendulum example, ...

Geri's Game/Sound



More References (selection)

- [CGRM92] Information technology - Computer Graphics - Reference Model, (ISO/IEC 11072)
- [Chia99] CHIARIGLIONE, L.: MPEG-4 - The fusion of natural and syntetic audio and video, SCCG99 Int. Conf. Proceedings, Budmerice 1999, pp. 45-64
- [Ferk92] FERKO, A.: Multimedia and their Limitations. SSCG92 Proceedings. Comenius University. Bratislava, May 1992.
- [Fole90] FOLEY, J. et al.: Fundamentals of Interactive Computer Graphics, Wiley, New York 1990
- [Koes64] KOESTLER, A.: The Act of Creation, New York, Macmillan, 1964
- [Herz94] HERZNER, W. (Ed.) Multimedia/Hypermedia in Open Distributed Environments, Proceedings of the Symposium in Graz, Austria, 1994, Springer-Verlag
- [Mart00] MARTTILA, O. - VUORIMAA, P. XML Based Mobile Services. pp. 282-289. In: WSCG'2000 Conf. Proceedings. Pilsen: Univ. of West Bohemia 2000.
- [OCon89] O'CONNOR, J. - SEYMOUR, J.: Introducing Neuro-Linguistic Programming, Lambent Books 1989
- [Orma97] ORMANDY, Roman: CALIGARI Manifesto, manuscript, California 1997.
- [Peng00] PENG, Ch. - VUORIMAA, P. Development of Java User Interface for Digital Television. pp. 259-265. In: WSCG'2000 Conf. Proceedings. Pilsen: Univ. of West Bohemia 2000.
- [PREM94] Information Processing Systems - Computer Graphics and Image Processing - Presentation Environments for Multimedia Objects (PREMO), (CD ISO/IEC 14782:1994)
- [Sten94] STENZEL H. et al.: PREMO, An Architecture for Presentation of Multimedia Objects in an Open Environment, in: [Herz94], pp. 77-96
- [Stuc91] STUCKI, P.: Graphics and Multimedia, tutorial at Eurographics Conference, Vienna 1991
- [Žára99] ŽÁRA, J.: VRML-97, laskavý průvodce virtuálními světy, ComputerPress, Brno 1999

Thank You

For Your Attention

Sound/Audio
or
Listening to Geri‘s Game

Andrej Ferko

Comenius University, Bratislava, ferko@fmph.uniba.sk

www.sccg.sk/~wega