Virtual Concert Performance – Synthetic Animated Musicians Playing in an Acoustically Simulated Room

Rami Hänninen, Lauri Savioja, and Tapio Takala

Department of Computer Science
Helsinki University of Technology
Otakaari 1, FI-02150 Espoo, Finland

{Rami.Hanninen,Lauri.Savioja,Tapio.Takala}@hut.fi
URL: http://www.hut.fi/HUT/Acoustics/DIVA/

Abstract
This paper describes a system that implements a virtual visual and acoustic model of a closed space, such as a concert hall. The system features an animated musician playing a flute, physically modeled flute sound synthesis, visual and acoustical room models, auralization, and real time graphical user interface. The user can move in the modelled space and experience the visual and aural environment.

Keywords: Virtual Reality, Animation, Physical Modeling, Room Acoustics, Auralization

![Diagram of DIVA system overview and information flowgraph.](image)

Hänninen et al. 402 ICMC Proceedings 1996
1 Introduction

We have developed a soft- and hardware system called DIVA (Digital Virtual Acoustics) for producing virtual audiovisual performances in real-time. The whole processing chain from sound synthesis through sound propagation in a room to auralization at the listener's ears is implemented, and combined with synchronized animation of music players and lighting of the virtual scene. The result is a fully synthetic concert where an animated musician plays a physically modelled instrument in an acoustically modelled hall. Instrument fingering is automated and, by inverse kinematics calculations, the finger motions are synchronized to exactly perform each note on an instrument model.

Acoustics of the enclosing room are simulated based on geometric model and material information of the room. Sound binaural auralization of the instrument and concert hall model is auralized to the listener, who can freely move in the virtual concert hall.

This paper describes the overall structure of the system, and the MIDI-based animation control of virtual musicians. Detailed discussion of physical instrument modeling, as well as auralization and room acoustics simulation can be found in other papers [Huopaniemi et al. 1994, Valimäki et al. 1996, Savioja et al. 1996].

2 Overview and Related Work

Modern computer graphics and signal processing systems can create realistic audiovisual representations for Virtual Reality [see e.g. Kallaway 1993, Begault 1994]. For animations, motion has been generated from MIDI music by procedural control on one hand [Lähte 1990], and sound effects have been generated by physically-based simulation on the other [Takala and Fahn 1992]. However, only few systems [Cook 1995] have combined these with musician modeling — closest to our work is the virtual performer by [Katajou et al. 1993].

The system components and the overall information flow are shown in figure 1. The upper half of the figure shows the audio stream while the lower half shows the visual stream.

Our sound sources are physical synthesis models of instruments. Several approaches to physical modeling can be found in [Smith 1996, Valimäki and Takala 1996]. We use the waveguide method, currently implemented for flute and piccolo aing instruments such as acoustic guitar, banjo and mandolin [Valimäki et al. 1995].

Concert hall geometry, including visual and acoustic material properties, defines the 3D model of the virtual space. We use a hybrid method to calculate the room's response. Early reflections are computed binaurally with image-source method. Reverberant part of the impulse response is obtained by the ray-tracing method. Diffraction effects at low frequencies could be taken into account with another model based on signal propagation in a three-dimensional waveguide mesh. The latter technique, however, are not applicable to real-time performance, but only to off-line prepared animations. Real-time performance calls for coarse simplifications, taking only the direct sound and first reflections as echoes into account, for example, and modeling late reverberation by digital filter structure [Scherroeder 1962].

Modeling and synthesis of binaural hearing (auralization) is an important part of immersive audio environments. Simple means for giving a directional sensation of sound are the interaural amplitude and time differences (IAD and ITD), but they cannot resolve the front-back confusion. The head-related transfer function (HRTF) gives better performance but is computationally more demanding. Our approach is to process only the direct sound with a filter approximation of the HRTF at the respective direction, and to use IAD and ITD for other reflections (image sources). For late-reverberation we use a recursive filter structure consisting of comb and allpass filters [Huopaniemi et al. 1994].

3 Animation Control from MIDI

The MIDI-to-movement mapper unit acts as musician and instrument models to convert MIDI events into movements. As an inverse kinematics system then transforms the goals to arm and finger movements.

The instrument definition list the possible fingering combinations that an instrument supports for each note. As there may be alternative fingerings for the same note, the mapper first makes a sequence of compatible grips to play the music. Then each grip is replaced with the set of associated elementary movements. They are passed to an inverse kinematics system that generates a path through all given position and temporal movement goals in 3D space.

Our human instrumentalist model (figure 2) consists of a static body with movable arms and fingers. Each arm has 27 degrees of freedom in total. The inverse kinematics, i.e. deciding intermediate arm joint positions from the finger positions, is solved iteratively by using the Jacobian matrix method [Wat and Warr 1992]. As the inverse kinematics is computationally very demanding, the finger positions are precomputed, and only their synchronization and rendering is done in real time.

Finally, the interactive animation control unit provides a graphical user interface to the DIVA system. Through this interface, the user moves a virtual viewer and listener around in a concert hall.
4 Experiments and Future Work

Our implementation is distributed to run in parallel processors, connected by Ethernet. Currently we use one Silicon Graphics Power Challenge workstation for real-time visualization and user interface and another for image-source calculations. A TMS320C40-based signal processor system performs the auralization. We are investigating the possibility of using high-speed asynchronous transfer mode (ATM) networks for virtual audio quality applications.

We have an animated video presentation of our system (figure 2 is a snapshot from the video). An animated flutist plays in the model of a real concert hall, while the camera moves around, demonstrating room acoustics at different listening positions.

For more immersive experience, we are planning to build a virtual studio environment, similar to the CAVe system. [Cruz-Neira et al. 1992], where the user may freely move and control the virtual world with a gesture sensor – for example to conduct a virtual orchestra.

Hänninen et al.

References


