Swimming in a Virtual Aquarium

Tapio Takala, Lauri Savioja, Tapio Lokki

Helsinki University of Technology,
P.O. box 5400, 02015 HUT
Espoo, Finland
tapio.takala@hut.fi, lauri.savioja@hut.fi, tapio.lokki@hut.fi
http://www.tml.hut.fi/, http://eve.hut.fi/

1 Introduction

We present a simulated underwater world to be experienced in a cave-like virtual environment [Cruz-Neira 1992]. The user, wearing stereo glasses and data gloves, moves among three-dimensionally modeled plants and schools of fish in the virtual water with hand gestures that simulate swimming. In addition to this physically based motion, the user may interact with the virtual fish, attracting them by gentle feeding gestures or scaring them by abrupt motion such as clapping hands. Besides visual feedback, the system includes 3D audio effects simulating water flow at the moving hands.

Figure 1. A view into the virtual aquarium.

Virtual environments have been utilized in many ways, ranging from simulator training to scientific visualization to entertainment. Having worked with computer animation, our purpose here was to make a compelling demonstration of simulated underwater life, similar to the Virtual Fishtank [Resnick 1999]. In order to make it more fun, and to enhance feeling of immersion, we added gestural motion control based on simplified hydrodynamics of swimming.

Intuitive navigation in virtual environments has been approached with different human motion interfaces. Pointing the direction is probably the most common, but other more embodied techniques from walking to bicycling have been experimented. Swimming would be the most natural in underwater environment. However, the only reference we know about virtual swimming happens on the surface [Chen et al. 2004].

2 Visual aquatic environment

We filled the underwater world with gently moving plants, and various types of fish. Schooling was simulated through flocking behavior loosely based on ideas presented by Reynolds [1987].

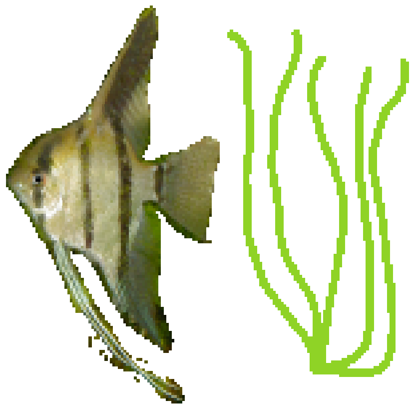
Fish and plant modeling and movement

The basic setup of the virtual aquatic world is described in a simple text file listing all the flocks and their attributes. Overall, there are three types of objects that can form flocks: fishes, crawlers, and plants.

All the modeling is performed with minimal effort such that one fish is presented with only eight triangles. They form a closed volume with a given thickness at the center of the fish. The actual fishlike outlook is achieved with alpha-blended texturing. The textures were obtained by an ordinary digital camera and post processing with *gimp* image manipulation software. An example of a fish texture is presented in Figure 2a). The movement of a fish is simplistic as well. Realistic animation of movement would require shape deformations, but in our case rigid models were enough. The only swim movement the fish has, consists of small rotations around its center. The obtained result is far from high quality fish animations [Terzopoulos 1994], but it is sufficient for our purposes. In addition to schools of fish, we have crawlers such as starfishes moving on the bottom. Their basic structure and movement are very similar to those of the fish. The only differences are that the models have only four triangles and in the movement each outer corner of the model is forced to touch the bottom or even dig into the bottom a bit. By this means we achieve surprisingly good-looking crawling in a sand bottom.

The modeling of plants is a bit more complicated since realistic movement of leafs requires more polygons. One plant consists of two orthogonal quad strips with the same textures. The level of tessellation is user definable and it affects directly the smoothness of achieved swaying. The movement itself is implemented with a digital waveguide model [Smith 1992] that propagates random sine waves back and forth in the leaves. The applied plant textures are currently hand drawn, and an example is shown in Figure 2b).

The bottom of the environment is modeled with a random height map and a repeating texture. User has control over the smoothness of the map. The given factor states how steep height variations there might be in the bottom.



a) b)

Figure 2. a) An example of employed texture presenting an angel fish, b) a simple hand drawn tape grass like texture.

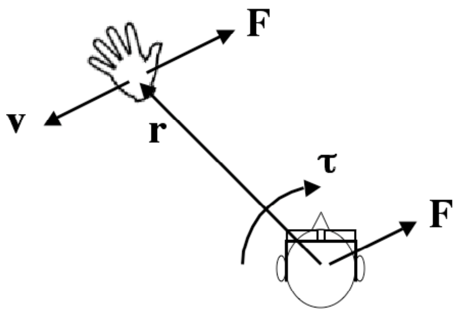
Schooling behavior

In our system, there is some variation in the sizes of fish, and the largest fish in a school is selected to be its leader, although its role is quite weak. All the fish have a random target, and whenever the leader reaches its goal a new goal is selected for all the fish in the same school. Each goal has a predefined optimal density, and it determines the proximity of the targets inside a school. In addition, each fish has a characteristic restlessness. It defines the acceleration in all the movements and describes the probability that the fish unexpectedly changes its target.

There are two specific methods that can get triggered by behavior of the diver. First one is scaring. If the user gets too close to a school and is too loud, the fish get scared and the school diverges. It is implemented by moving the target locations of every fish further away from the target of the leader. The opposite of scaring is attracting the fish. If the diver makes gentle gestures towards the fish they get interested and swim closer to the user. This again is achieved by changing the goal positions of the fish.

2 Gestures used in swimming

The swimmer's motion is captured with a three-dimensional magnetic motion tracker with one sensor at each hand and one on the head. In order to make the experience as natural as possible, we made the user to move in the aquatic world by virtual swimming, based on physical calculations instead of symbolic gestures. As a sensor system for the whole body was not available, we reduced the control to the hands only. This also allows one to stand safely in balance without affecting the fluidity of virtual motion.



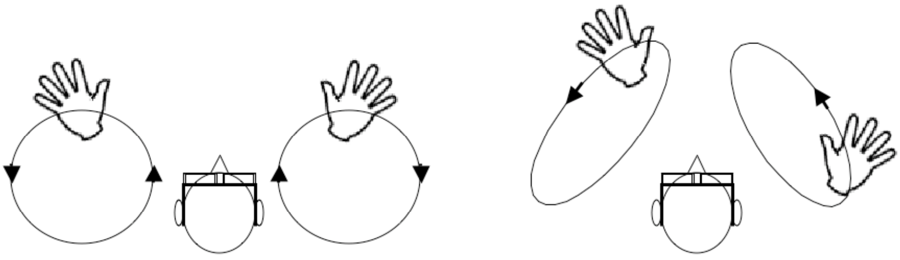
$$\mathbf{F} = -\eta\,\mathbf{v}\cdot|\,\mathbf{r}\,|\cdot\sin(\mathbf{v},\mathbf{r})$$
$$= -\eta\,\mathbf{v}\cdot|\,\mathbf{r}\times\mathbf{v}\,|\,/\,|\mathbf{v}|$$

$$\boldsymbol{\tau} = \mathbf{r}\times\mathbf{F}$$

Figure 3. Simplified fluid physics used in the simulation.

Attaching one position sensor to the user's head and one to each hand we get the configuration shown in figure 3. We assume the swimmer's center of mass to be at the head. Thus all forces and torques are applied to it, and the viewing position is moved accordingly. All forces affecting the motion are supposed to be due to the viscosity, thus proportional to velocity and cross-sectional area of a moving body. Thus the force **F** virtually acting on a moving hand is proportional to water viscosity η and to the hand's velocity **v** relative to the environment. Direction of the force is opposite to the velocity. The effective hand area is not exactly modeled, but simply assumed to be proportional to length of the extended hand, i.e. to distance **r** from the head, and to sine of the angle between **v** and **r**, as shown in the formulas. This can be calculated by taking the cross product of these two vectors. Orientation of the hand also matters when we calculate the torque $\boldsymbol{\tau}$ applied to the mass center due to the force. As shown, $\boldsymbol{\tau}$ is the vector cross product of hand position **r** and the force **F**.

Summing up forces and torques due to each hand, we get the totals affecting the user's mass center, causing linear and angular acceleration, respectively. These are continuously integrated in time, resulting in linear and angular velocities and corresponding changes in position and orientation of the virtual world. As viscosity damps all motion, we add before integration also a force proportional and opposite to the instantaneous velocity vector of the swimmer and, respectively, a torque against the current angular velocity.



a) b)

Figure 4. Gestures for (a) moving forward, and (b) turning right.

These simple and general calculations of physics allow one to move in very different ways in the environment. Symmetrical circular motion of both hands, similar to breaststroke, is very efficient for swimming forwards or backwards (figure 4a). The same motion with one hand only (or with both hands in the same direction) makes one to turn continuously (4b), and a short instantaneous turn can be made by just changing the orientation of hands relative to the head. Perceptionally these gestures are very natural, similar to real swimming. The swimmer's motion after a gesture continues a while, but gradually stops due to viscosity, if no further movements are done.

As the forces and torques are modeled in three dimensions, the vertical motion of hands is also taken into account, allowing one to dive deeper or raise towards surface. Three-dimensional rotation, however, appeared to difficult to manage by users as they accidentally turned upside down and couldn't return to a normal position. Thus we turned these calculations off and allowed rotation around the vertical axis only.

3 Three-Dimensional Auditory Environment

Besides visual feedback, the system includes underwater 3D auditory environment. Auditory environment is composed of static ambient ocean wave sounds, dynamic bubbling sound and heart rate sound controlled by hand movements, and individual sound events. Spatial audio is reproduced with 14 channel loudspeaker system which creates fully immersive spatial sound environment.

Spatial sound reproduction

To create an immersive auditory environment a spatial audio reproduction system have to be applied. In principle, spatial audio reproduction can be handled with headphones [Begault 1994] or multiple loudspeakers [Rumsey 2001]. In our Virtual Aquarium, we decided to apply loudspeaker reproduction, since additional wires and wearable devices disturb the swimmer.

Our spatial sound reproduction system consists of 14 loudspeakers and a sub-woofer. The applied spatial rendering method is the vector base amplitude panning (VBAP) [Pulkki 1997], which is a practical method in immersive cave-like environments. VBAP allows almost arbitrary positioning of the loudspeakers and still good directional sound can be reproduced. In our case we have four loudspeakers on the floor, six on ear level and four in the ceiling. With such a configuration accurate enough directional panning can be achieved so that loudspeakers are not disturbing the visual images. The more detailed presentation of the applied audio system is presented by Hiipakka et al. [2001].

Implemented auditory environment

An immersive auditory environment is built by using both static ambient sounds and dynamic interactive sounds. Static underwater soundscape was mixed from low pass filtered ocean wave and waterfall samples. Low pass filtering muffles sound so that it gives a feeling of underwater listening. To create an immersive ambience this static stereo mix is reproduced from all loudspeakers so that left channel is routed to eight loudspeakers and right channel to eight loudspeakers.

Dynamic interaction sound consists of several elements. For hand movements a bubbling sound stream is continuously running and the volume is controlled by the amount of movements. The more hands are moving the louder the bubbling sound is reproduced from the direction defined by the vector from the swimmer's head to the hand. If hands are standing still no bubbling sound is heard. In addition, we implemented a heart rate sound which is reproduced on top of the swimmer. This sound is a pulse-like synthetic sound and the rate of these pulses is controlled. When swimmer move his/her hands the rate is raised and when movement stops the rate is slowed down with a leaky integrator. Third interaction sound was a hand clap sample which is triggered if swimmer tries to scare fishes.

4 Conclusions

We have created a virtual environment that simulates underwater diving among plants and fish. It has an embodied user interface based on hand motion similar to swimming. Having demonstrated the system to various visitors, we have got positive reactions. The experience is considered entertaining and very natural. Because the movement is physical, no specific symbolic gestures are needed to learn. A formal user study, however, concerning the effect of gestural movement on the depth of immersion or sense of presense, remains to be future work.

References

1. Carolina Cruz-Neira (1992). The CAVE: audio visual experience automatic virtual environment. *Comm. of the ACM*, **35**(6), pp. 64 – 72.

2. Mitchel Resnick (1999). Virtual Fishtank. http://llk.media.mit.edu/projects/fishtank/

3. Tzu-Pei Chen, Yuichiro Kinoshita, Yasufumi Takama, Sidney Fels, Kenji Funahashi, Ashley Gadd (2004). Swimming Across the Pacific: A Virtual Swimming Interface. SIGGRAPH'2004 Emerging Technologies.

4. Craig Reynolds (1987). Flocks, Herds, and Schools: A Distributed Behavioral Model. *Computer Graphics*, **21**(4), pp. 25-34.

5. Dimitri Terzopoulos, X. Tu, R. Grzeszczuk, Artificial fishes: Autonomous locomotion, perception, behavior, and learning in a simulated physical world, Artificial Life, **1**(4), December, 1994, 327-351.

6. Julius O. Smith. Physical modeling using digital waveguides. Computer Music J., **16**(4):74-87, 1992 Winter.

7. Durand R. Begault. 3-D Sound for Virtual Reality and Multimedia. Academic Press, Cambridge, MA, 1994.

8. F. Rumsey. Spatial Sound. Focal Press, Oxford and Boston, 2001.

9. V. Pulkki. Virtual sound source positioning using vector base amplitude panning. J. Audio Eng. Soc., **45**(6):456-466, June 1997.

10. J. Hiipakka, T. Ilmonen, T. Lokki, M. Gröhn, and L. Savioja. Implementation issues of 3D audio in a virtual room. In Proc. 13th Symposium of IS&T/SPIE, Electronic Imaging 2001, volume 4297B, San Jose, California, USA, Jan. 21-26 2001.