

Tópicos Especiais em Computação Gráfica

Realidade Virtual

Som Tridimensional

Márcio Serolli Pinho

Faculdade de Informática – PUCRS

5. O Som em ambientes virtuais

A adição de sons, em qualquer processo de exibição de imagens que pretende ser de alguma forma interativo, torna a exibição muito mais realista. Quando em uma tela, aparece no centro uma bola quicando, a agregação de um som mono, com a simples reprodução do som real sincronizado com o movimento da bola, já produz uma sensação de realismo muito maior para quem observa.

Historicamente, as discussões sobre imersão e realismo em ambientes virtuais têm dado quase que nenhuma ênfase aos aspectos de produção sonora. O ser humano, entretanto, é constantemente bombardeado de estímulos sonoros. Em diversos casos o estímulo auditivo não é somente uma forma de aumentar o realismo de uma simulação. Sua ausência pode, de fato, tornar inviável uma aplicação. Begault [Begault, 1994], por exemplo, discute algumas experiências feitas pelo *NASA Ames Research Center* onde pilotos de avião tinham extremas dificuldades em controlar um painel de botões virtuais nos quais não havia a correspondente geração de um som quando um deles era pressionado.

5.4 Características do som

Todos os sons, independentemente de onde vem, ou de como parecem para quem os ouvem, são resultado da vibração de um objeto, causada pela aplicação de uma força sobre este objeto [Aukstaklanis, 1992]. Estas vibrações empurram as moléculas de ar ao redor do objeto, estas moléculas por sua vez colidem com moléculas vizinhas e assim sucessivamente, formando ondas sonoras que se propagam pelo ar. Nossos ouvidos, quando percebem estas ondas, as transformam em sons.

Na realidade o som pode se propagar não só pelo ar, mas por qualquer meio que seja suficientemente denso para que as moléculas “se empurrem umas as outras”. Quanto mais denso for o meio, mais rápido o som irá se propagar. No ar a velocidade de propagação é de 340 m/s, na água de 1500 m/s, no ferro de 5000 m/s.

Uma onda sonora é uma curva senoidal que possui duas características relevantes: **amplitude** e **frequência**. A amplitude, expressa em **decibéis (dB)**, define o volume do

som. Uma conversação normal atinge 60 dB. A partir de 120 dB o ouvido humano começa a sentir dor. A amplitude (ou o volume) de um som é inversamente proporcional à distância entre a fonte sonora e o ouvinte. A frequência, expressa em **hertz (Hz)**, representa o número de vezes que a curva se repete em um segundo e tem uma relação direta com tipo de som percebido: sons graves são de baixa frequência, sons agudos são de alta frequência.

A sensibilidade auditiva de um ser humano normal é extremamente grande. Um ouvido sadio é capaz de captar mudanças na pressão do ar causadas por movimentos bastante pequenos, chegando a extremos de que algumas pessoas percebem sons causados por deslocamentos de um décimo do diâmetro de uma molécula de ar.

5.5 Localização de uma fonte sonora

Para tornar uma simulação realista, a produção (ou reprodução) de um som deve buscar a emissão de um som que seja o mais parecido possível com o som real. Esta “semelhança”, entretanto, não é a única característica relevante. A localização deste som é de extrema importância. Um apito, por exemplo, sendo acionado a partir de uma posição à direita do usuário, soa diferente do mesmo apito sendo acionado à esquerda deste mesmo usuário. Se o sistema que controla o ambiente virtual for capaz de prover esta característica, muitas vezes conhecida como “**som tridimensional**”, um elevado grau de imersão pode ser atingido.

A identificação da posição de uma fonte sonora é feita pelo ser humano a partir de vários mecanismos. Os principais são a **ILD (Interaural Level Difference)** e as **HRTFs (Head Related Transfer Functions)**.

Conforme foi mencionado anteriormente, a distância entre a fonte sonora e o ouvinte afeta o nível de percepção do som. A partir disto é fácil de concluir que haverá diferença entre os sons captados pelos dois ouvidos. Esta diferença é chamada **ILD (Interaural Level Difference)**. Na Figura 5.1, por exemplo, o ouvido direito irá captar um som mais alto do que ouvido esquerdo. Esta informação enviada ao cérebro, juntamente com a diferença no tempo de captação entre os dois ouvidos, permite a

determinação da posição da fonte sonora. Gulich [Gulich, 1989] apresenta uma extensa discussão sobre ao assunto.

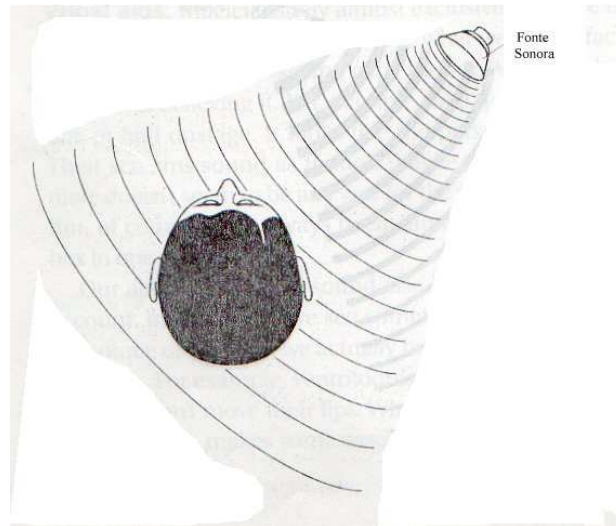


Figura 5.1 - Captação diferenciada de sons pelos ouvidos

O segundo aspecto relevante na localização da fonte sonora é o formato e o tamanho das estruturas anatômicas que circundam o canal auditivo, como a cabeça, as orelhas e os ombros do ouvinte. Segundo [Gardner, 1998 e Gilkey, 1997] as várias dobras existentes na orelha modificam as frequências que compõem um som, reforçando algumas, enfraquecendo outras, dependentemente da posição de onde provem o som. Estas transformações sofridas pelo som, até chegar ao nervo auditivo são chamadas de **Head Related Transfer Functions (HRTFs)**. Este fenômeno foi descoberto a partir da constatação de mesmo com um dos ouvidos totalmente surdo, algumas pessoas ainda podem detectar a posição de uma fonte sonora [Batteau, 1967; Wightman, 1995].

Estas funções podem ser medidas através da inserção de pequenos microfones no conduto auditivo de um ouvinte exposto a fontes sonoras (Figura 5.2). Na medição, uma fonte sonora é colocada em algum ponto onde o ouvinte é capaz de percebê-la e um som é emitido. Os sons captados pelos microfones são então processados e comparados com o som emitido. O resultado destas comparações é tomado juntamente com a posição conhecida da fonte sonora a fim de criar as funções que transformam o som emitido no som captado [Wenzel, 1994; Wightman, 1993].

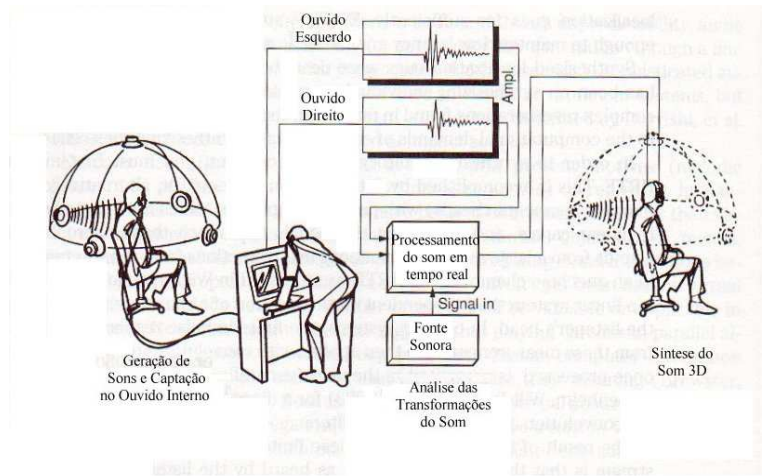


Figura 5.2 – Captação de sons para determinação das HRTFs [Wenzel, 1994]

5.6 Geração de som tridimensional

A forma mais simples de geração de som tridimensional em um ambiente virtual é a disposição de alto-falantes ao redor do usuário e o correto acionamento destes conforme a necessidade de emissão do som. Esta alternativa, entretanto, pode necessitar de um demorado processo de calibração e ajuste para o correto posicionamento dos alto-falantes. Por este motivo esta tecnologia é usada preferencialmente em ambiente com as CAVES [Bargar, 1993], onde, por suas próprias características, o processo de montagem da estrutura também é demorado, ou em ambientes onde não são necessárias mudanças muito freqüentes.

Outra técnica, mais utilizada, foi desenvolvida no *NASA AMES Auditory Display Project* [Wenzel, 1992; Wightman, 1993]. O objetivo era a geração de sons tridimensionais a partir de fones de ouvido comuns. A idéia foi realizar testes exaustivos sobre um grande conjunto de usuários a fim de obter um padrão para as funções **HRTF**.

Para o desenvolvimento do projeto cada usuário foi colocado sentado em uma cadeira, numa sala livre de eco, na qual havia 144 alto-falantes. No interior dos ouvidos do usuário eram colocados minúsculos microfones para captar os sons, depois que estes sofressem as transformações causadas pelas orelhas e pela cabeça. A partir disto, sons eram emitidos dos alto-falantes e captados pelos microfones.

Com estes sons armazenados e processados foi possível criar um conjunto de funções de transformação (as HRTFs) para “espacializar” um som emitido apenas em dois fones de ouvido.

Este projeto deu origem ao **Convolvotron** da empresa **Crystal River Engeneering** (<http://www.cre.com>). Este equipamento é, em essência, um **DSP** (processador digital de sinais) capaz de aplicar as transformações HRTFs a um sinal sonoro. Em 1995, o **MIT Media Lab** desenvolveu o projeto **KEMAR** (*Knowles Electronic Manikin for Acoustic Research*) [Gardner, 1995 e Gardner, 1997], repetindo a experiência da NASA, desta vez com um manequim no lugar de usuários humanos.

Neste projeto, financiado pela **Wave Arts 3D** (<http://www.wavearts.com>), foram feitas mais de 700 medições para diversas fontes sonoras dispostas em diversas posições. Atualmente a empresa está oferecendo no mercado dois produtos criados a partir deste projeto. O primeiro deles é o **WaveSurround DX Plugin**, um *software* que permite a criação de som tridimensional em qualquer aplicação que utilize a biblioteca **DirectX** da **Microsoft**. O segundo (Figura 5.3) é um *plugin* para o “tocador” de MP3, **WinAmp**. Ambos os *softwares* permite a simulação de som tridimensional proveniente de duas fontes sonoras.



Figura 5.3 – Plugin de som 3D para o WinAmp

Outro produto criado a partir da tecnologia desenvolvida no projeto KEMAR foi o **InMotion 3D Audio Producer**¹ da empresa americana **Human Machine Interfaces, Inc** (<http://www.humanmachine.com>). Este *software* possibilita a geração de sons a partir de seis fontes sonoras (Figura 5.4).

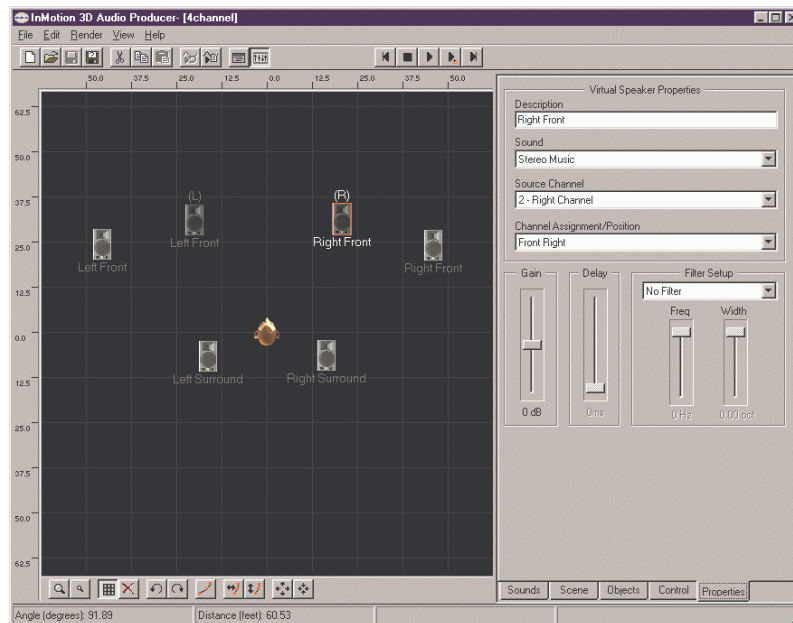


Figura 5.4 - InMotion 3D Audio Producer

¹ O custo deste *software* é de quinhentos dólares.