

Carlos Akcelrud Cony

**Simulação de Agentes Virtuais Dotados de
Percepção Sintética em Situações de
Emergência**

São Leopoldo

2007

Carlos Akcelrud Cony

Simulação de Agentes Virtuais Dotados de Percepção Sintética em Situações de Emergência

Dissertação submetida a avaliação como re-
quisito parcial para a obtenção do grau de
Mestre em Computação Aplicada.

Orientador:
Cláudio Rosito Jung

UNIVERSIDADE DO VALE DO RIO DOS SINOS
CIÊNCIAS EXATAS E TECNOLÓGICAS
PROGRAMA INTERDISCIPLINAR DE PÓS-GRADUAÇÃO EM
COMPUTAÇÃO APLICADA

São Leopoldo

2007

Dedicatória

Dedico esta dissertação a todas as pessoas que me ajudaram e acreditaram no meu esforço e capacidade.

Agradecimentos

Para começar gostaria de agradecer aos meus orientadores, em primeiro lugar a Profa. Soraia, a qual confiou em mim desde o primeiro dia, me oferecendo as melhores condições de trabalho que um bolsista e orientado de mestrado pode ter, tanto no suporte intelectual, como profissional. Não muito distante disso, o Prof. Cláudio que me “adotou” ao longo do mestrado como seu orientado (forças dos destino), mas também de fundamental importância para o desenvolvimento deste trabalho, dando idéias e soluções que somente uma pessoa com intelecto diferenciado poderia dar.

Um pouco mais abaixo dos meus queridos orientadores está um ex-professor, e atualmente amigo e colega, companheiro de jogos do Grêmio, estudos, dicas, soluções e até mesmo de viradas de noites escrevendo artigos. Valeu Bicho pela força e pelo apoio nestes 2 anos de mestrado.

Dentro do PIPCA, existem algumas pessoas que merecem destaque. Uma delas é o Prof. Osório, que esteve presente em alguns momentos decisivos (artigos e bancas). Outras são meus queridos colegas de batalha, que ao longo do curso deixaram de ser apenas colegas e se tornaram amigos importantes. Valeu JR, Paty e Roberto pelas horas de estudos, parcerias de lanches e cervejas no Rapach. Não posso deixar de agradecer a todos os professores e colegas que participaram ou contribuíram de alguma forma com este trabalho, que não foram citados aqui.

Uma pessoa que não participou diretamente da construção desta dissertação, mas que agüentou muita barra, aceitou esse voto de pobreza comigo e foi tão ou mais im-

portante que as já citadas neste singelo agradecimento, a minha esposa amada. Beijos Camila, te amo muito.

Tenho que mandar um obrigado para os meus queridos pais, beijão pai e mãe!! Vocês são a razão da minha existência e do meu sucesso.

Para não me alongar mais, agradeço a todos que conviveram comigo ao longo destes dois anos! Obrigado também a HP Brasil pelo o apoio e suporte financeiro para a realização deste trabalho.

Resumo

Este trabalho apresenta um modelo para prover habilidades perceptivas (visão, olfato e audição) aos agentes virtuais, focando em aplicações de simulação de multidões em situações de emergência. Os modelos visual e olfativo foram baseados em trabalhos já publicados por autores da literatura, enquanto que a audição virtual foi modelada com maiores detalhes.

A modelagem do sistema auditivo baseou-se em características reais, tais como as atenuações do som em relação às perdas energéticas causadas pelas paredes (refrações) e distância percorrida pelo som. Além destas, foram estudados alguns comportamentos fisiológicos do ouvido humano, tais como a dificuldade em que um ser humano tem em ouvir determinadas bandas de frequência, perdas auditivas decorrentes do envelhecimento, e ainda possíveis doenças auditivas. Ainda, aplicou-se um modelo de mascaramento sonoro, no qual alguns sons podem ser mascarados por outros com frequências similares.

O modelo perceptivo virtual foi avaliado em aplicações de simulação de multidões em situações de emergência, nas quais um alarme era posicionado em diferentes posições do ambiente, e os agentes tomavam decisões com base nas percepções virtuais. Um modelo de pseudo-comunicação entre os agentes também foi implementado e testado, e seu impacto no total de pessoas salvas foi avaliado.

Abstract

This work presents a model to provide perceptual abilities (vision, smelling and hearing) to virtual agents, focusing on applications of crowd simulation for emergency situations. The models for synthetic vision and hearing were based on papers published by known authors, while the virtual hearing was modeled in details.

The synthetic hearing model was based on real characteristics, such as sound attenuation caused by walls and distance from the sound source. Also, some physiological behaviors of the human hear were considered, such as the natural difficulty of a human being to hear some frequency bands, and hearing losses due to aging. Finally, the hearing model also included masking effects, in which some sounds may be masked by others with similar frequencies.

The proposed perceptual model was evaluated in applications of crowd simulation for emergency situations, in which an alarm was placed into different positions of the virtual environment, and the agents had to decide what to do based on their perceptions. A model for pseudo-communication was also implemented and tested, to evaluate its impact on the number of saved agents.

Lista de Figuras

1	A Mente Humana (RIO, 1996)	19
2	Caminho dos feixes de nervos ópticos em direção ao cérebro (FERRAZ, 1998).	21
3	Mucosas Nasais (GUYTON, 1981)	22
4	Cavidade Nasal Humana (PONTES, 2006)	23
5	Cadeia Ossicular (SANTOS; RUSSO, 1986)	24
6	Passos da Audição (SANTOS; RUSSO, 1986)	25
7	Principais Tipos de Mascaramento. (LEITE, 2003)	31
8	Efeito do mascaramento simultâneo (LEITE, 2003)	31
9	Gráfico de sensibilidade auditiva do ser humano em função da frequência.	42
10	(a) Ambos agentes se percebem (b) usando o mapa de percepção para aumentar a performance da abordagem por zona.	45
11	Zona de percepção olfativa para os agentes.	48
12	Componentes do Simulador.	50
13	Captura da tela do simulador.	53
14	Audição sem comunicação.	55
15	Audição com comunicação.	55

16	Influência das paredes no modelo auditivo.	56
17	Visão <i>versus</i> Olfato.	57
18	Audição <i>versus</i> Olfato <i>versus</i> Visão.	58
19	Impacto do mascarador sobre o alarme de frequência 6 kHz	58
20	Impacto do mascarador sobre o alarme de frequência 12 kHz	59

Lista de Tabelas

1	Nível de pressão sonora para exemplos do cotidiano.(PAINTER; SPANIAS, 2000)	28
2	Bandas Críticas (CAVE, 2002)	29
3	Valores atribuídos aos coeficientes α , β e γ , para diferentes frequências chaves.	42
4	Bandas de Frequência.	43
5	Decreto Número 3.048 da Previdência Social de 06 de maio de 1999	43
6	Tabela de materiais <i>versus</i> atenuação em dBs por faixa de frequência . . .	44
7	Populações Utilizadas - Perdas Auditivas	54

Anexos

Anexo A - Resumo do artigo publicado no SBGames 2006 - Simpósio Brasileiro de Jogos 2006

Anexo B - Resumo do artigo publicado no CGI 2007 - Computer Graphics International

Anexo C - Resumo do artigo publicado no ENIA 2007 - Encontro Nacional de Inteligência Artificial

Lista de Abreviaturas

NPC – *Non-Player Characters*

IA – Inteligência Artificial

XML – *eXtensible Markup Language*

SNC – *Sistema Nervoso Central*

FFT – *Fast Fourier Transform*

IEC – *International Electrotechnical Commission*

Sumário

1	Introdução	14
1.1	Motivação	15
1.2	Objetivo	16
1.2.1	Objetivo Específico	17
2	Percepções Humanas: Visão, Olfato e Audição	18
2.1	Visão Humana	19
2.2	Olfato Humano	21
2.3	Audição Humana	23
2.3.1	Ouvido Humano	24
2.3.2	Fisiologia da Audição	25
2.3.3	Percepção de Volume do Som	27
2.3.4	Resposta em frequência do sistema auditivo	28
2.3.5	Bandas Críticas	28
2.3.6	Mascaramento	30
2.3.6.1	Mascaramento Simultâneo	30

3	Percepções Virtuais: Visão, Olfato e Audição	34
3.1	Visão Virtual	34
3.2	Olfato Virtual	36
3.3	Audição Virtual	36
4	Modelo	39
4.1	Modelo Auditivo	40
4.2	Modelo Visual e Pseudo-Comunicação	45
4.2.1	Pseudo-comunicação	46
4.3	Modelo Olfativo	47
4.4	Modelo Comportamental	48
4.5	O Protótipo	49
5	Resultados	51
5.1	Condições Iniciais	52
5.2	Análise dos Resultados	54
5.2.1	Impacto da Frequência do Som	54
5.2.2	Impacto da atenuação causada pelas paredes	54
5.2.3	Comparações entre as percepções	56
5.2.4	Impacto do Som Mascarador	57
6	Considerações Finais	60
6.1	Trabalhos Futuros	

Referências

Capítulo 1

Introdução

A partir do crescimento das pesquisas na área de animação comportamental, as grandes empresas de entretenimento mundial, tais como as companhias de desenvolvimento de jogos e das artes cenográficas, adotaram a utilização de agentes virtuais. Este artifício pode ser observado claramente na grande maioria dos jogos atuais (ex.: *World of Warcraft*), nos filmes que requerem a interação entre um grande número agentes (ex. cenas de batalhas do filme *Senhor dos Anéis* e *Star Wars*) e ainda em cenas muito perigosas onde seria impossível a utilização de um ator humano (ex.: *Titanic*).

A área de simulação, em tempo real ou não, também tem encontrado na animação comportamental uma possibilidade de simular computacionalmente comportamentos complexos, tais como reação das pessoas em situações de emergência. Para isso, é necessário que sejam desenvolvidos bons modelos para que os resultados apresentem boas estimativas em relação ao mundo real. Com o uso de simulações, é possível evitar situações de riscos para o homem, tornando possível a extração de informações pertinentes ao seu comportamento. Assim, quanto mais as entidades simuladas forem providas de habilidades perceptivas, tais como visão, audição e olfato sintéticos, mais realistas poderão ser os resultados de suas aplicações.

Entretanto, Reynolds (REYNOLDS, 1987) sugere que é inadequado um agente pos-

suir completa informação do mundo virtual, já que é mais realista possuir imprecisão nos dados, pela inerente imperfeição dos sentidos humanos, fatores associados as distâncias, oclusões, etc. Portanto, o desenvolvimento de modelos perceptivos e do conseqüente comportamento baseado nas informações adquiridas pelas percepções não é uma tarefa trivial.

Vários autores na literatura têm investigado a área de humanos virtuais (MUSSE; THALMANN, 2001; BRAUN; BODMANN; MUSSE, 2005; CHENNEY, 2004), principalmente em aplicações envolvendo simulações comportamentais, *serious games*, ferramentas de treinamento, entre outras. Assim, humanos virtuais perceptivos podem ser de grande interesse para a indústria de jogos, bem como para a área de simulação. Quanto mais perceptivos forem os NPCs (*Non-Player Characters*), ou as entidades virtuais na simulação, maior é a possibilidade destas tomarem decisões acertadas, através da modelagem conjunta com técnicas da Inteligência Artificial.

1.1 Motivação

Visto que a área de modelagem comportamental continua a crescer, faz-se necessário o desenvolvimento de modelos que tornem os agentes virtuais mais inteligentes. Para isso, diversos autores da comunidade tem estudado as mais variadas formas de tornar isto realidade.

Seguindo a linha de modelagem comportamental, pretende-se com este trabalho criar um modelo para prover habilidade auditiva a humanos virtuais. Para isso, desenvolveu-se um modelo onde os agentes virtuais são capazes de “ouvir” sons reais (arquivo no formato *wave*, por exemplo). Além dessa habilidade, o modelo ainda visa tornar os agentes individualizados, ou seja, o modelo leva em consideração parâmetros como atenuação do som em relação a distância do agente até a fonte sonora, a frequência do som, mascaramento sonoro, a presença de paredes e até mesmo a idade ou doenças auditivas do agente em questão.

Para tornar o trabalho mais completo, serão providas outras percepções aos agentes virtuais (visão e olfato), porém estas são desenvolvidas utilizando técnicas mais simples, que necessitam menos recursos computacionais e, além disso, possam ser, futuramente, substituídas por novos modelos mais robustos.

1.2 Objetivo

Este trabalho visou desenvolver um modelo perceptivo, focado principalmente na audição sintética, a qual foi implementada de maneira a prover individualidades similares as dos seres humanos. Ainda, outras percepções como visão, olfato e ainda uma pseudo-comunicação, também foram incluídos, aproveitando modelos simplificados já propostos por diversos autores na literatura. Para este trabalho as informações obtidas pelos agentes do ambiente através de seus sensores (sentidos) são baseadas nas mais diversas maneiras de obtenção de dados, tais como:

- arquivos de som para audição (arquivos *Wave*);
- refração do som nas paredes;
- mascaramento sonoro;
- idade individualizada para cada agente;
- tipo de audição (doente ou não);
- campos sensoriais para visão e olfato;
- interação entre os agentes (pseudo-comunicação).

Este trabalho pode ser aplicado em diversas áreas de conhecimento como será mostrado posteriormente na Seção 1.2.1. Especificamente, o interesse deste trabalho está em avaliar o comportamento dos agentes nas mais diversas situações, partindo do pressuposto

que se pode especificar habilidades perceptivas para os humanos virtuais e avaliar como isto afeta o resultado da simulação.

1.2.1 Objetivo Específico

O objetivo é a criação de um modelo auditivo, com o qual seja possível desenvolver um modelo comportamental de agentes virtuais em situações de emergência. Para isso, desenvolveu-se um simulador de multidões em situações de emergência, onde os humanoides utilizam a sua percepção para definir qual será a sua reação. Este modelo é de grande importância para simulações na área de segurança, onde os resultados numéricos podem indicar problemas de planejamento ou na arquitetura do ambiente. Por exemplo, um alarme mal posicionado pode resultar no aumento de pessoas que não evacuam o ambiente por não perceberem o evento. Além disso, pode ter grande aplicabilidade em projetos arquitetônicos e de entretenimento.

Capítulo 2

Percepções Humanas: Visão, Olfato e Audição

Através de seus órgãos sensitivos os homens percebem a realidade que está a sua volta. Isto é constituído a partir do que está sendo percebido pelos diversos sentidos (visão, olfato, audição, tato e gustação), passando pela cognição (ato de conhecer, que envolve atenção, percepção, memória, raciocínio, juízo, imaginação, pensamento e linguagem), até a ação a ser realizada.

Para modelar a percepção humana em computador e levando em consideração a sua complexidade inerente, usam-se alguns conceitos e variáveis de maneira a simplificar este processo. Assim, é possível mostrar as relações existentes entre a percepção e a reação do ser humano em uma determinada situação.

Segundo (MACHADO, 1996), os cinco sentidos permitem aos seres humanos reagir aos mais diversos estímulos. Estas informações obtidas pelos sentidos, somadas às adquiridas de maneira indireta, através dos livros, estudos, contatos inter-pessoais, meios de comunicação, por palavras escritas ou verbais, etc., formam o conhecimento de cada indivíduo.

Cunha (CUNHA, 2004) diz que, as informações adquiridas pelas percepções táteis,

auditivas, visuais, olfativas ou gustativas são processadas pelo sistema nervoso central, e somadas àquelas previamente armazenadas no cérebro, formando uma imagem mental de um objeto, se for o caso. O papel da mente humana é fazer a organização e a representação das percepções realizadas pelos sentidos (RIO, 1996), conforme ilustrado na Figura 1.

Nas próximas seções apresentam-se informações referentes aos sentidos humanos que serão modelados computacionalmente neste trabalho (visão, olfato e audição).



Figura 1: A Mente Humana (RIO, 1996)

2.1 Visão Humana

A retina do olho humano é composta por camadas de células. Uma dessas camadas é formada pelas células responsáveis por captar a luz, chamadas de células foto-receptoras. As outras camadas são responsáveis pela condução da informação visual até as fibras que compõe os nervos (PADULA; SPUNGIN, 2000; FIGUEIRA, 2000).

Existem dois tipos de células foto-receptoras (CASTLEMAN, 1996):

- Bastonetes: espalhados pela retina, existem aproximadamente 120 milhões de células bastonetes em cada olho. Este tipo de célula é aproximadamente 100 vezes mais sensível à luz do que as células cones, apresentando sensibilidade mesmo com baixos níveis de luminosidade. Contudo, “enxergam” em tons de cinza;
- Cones: existem cerca de 7 milhões de cones em cada olho. Este tipo de célula é bem menos espessa do que os bastonetes e reagem a luz quatro vezes mais rapidamente.

Os cones se subdividem em três tipos, os que respondem a ondas:

- longas, responsáveis pelas cores avermelhadas;
- médias, responsáveis pelas cores amareladas e esverdeadas;
- curtas, responsáveis pelas cores azuladas.

Além de ser responsável pela visualização das cores, os cones dão nitidez às imagens.

Após os impulsos nervosos, gerados pelas células foto-receptoras, os sinais percorrem os nervos ópticos até o SNC (sistema nervoso central). Assim, as fibras do nervo óptico recebem a composição dos impulsos de um grupo de células receptoras (CASTLEMAN, 1996), então, a informação que segue pelo nervo óptico percorre um caminho até chegar ao córtex visual (região do cérebro diretamente relacionada à visão). Este caminho está ilustrado na Figura 2.

Os feixes de fibras dos dois olhos se encontram no quiasmo óptico. Na verdade os feixes de ambos os olhos são divididos em duas partes. Cada parte transporta informação obtida por uma parcela de retina. A seção reta vertical que corta o centro do olho humano divide a retina em duas porções: a nasal e a temporal. As informações provenientes da porção temporal de cada retina projetam-se no córtex visual no mesmo lado (esquerdo ou direito) da cabeça ocupado pela retina. As informações provenientes da porção nasal de cada retina, por outro lado, projetam-se em hemisférios opostos do córtex visual (FERRAZ, 1998).

Esta mistura de feixes de fibras ópticas permite ao córtex visual processar duas amostras de uma cena. A combinação dos estímulos provenientes das duas retinas é justamente umas das duas principais transformações de informação realizadas pelo córtex visual (HUBEL; WIESEL, 1979).

Após estas transformações no córtex os impulsos nervosos prosseguem em seu caminho através do sistema nervoso central. Os demais estágios deste caminho irão traduzir o conjunto de estímulos visuais na visão propriamente dita (FERRAZ, 1998).

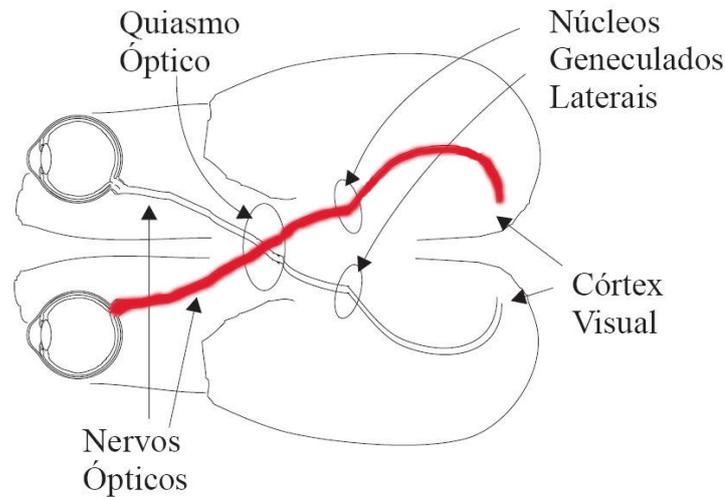


Figura 2: Caminho dos feixes de nervos ópticos em direção ao cérebro (FERRAZ, 1998).

2.2 Olfato Humano

O olfato humano é pouco desenvolvido se comparado ao de outros mamíferos. O epitélio olfativo humano contém cerca de 20 milhões de células sensoriais, cada qual com seis pêlos sensoriais (um cachorro tem mais de 100 milhões de células sensoriais, cada qual com pelo menos 100 pêlos sensoriais). Os receptores olfativos são neurônios genuínos, com receptores próprios que penetram no sistema nervoso central (GUYTON, 2002).

A cavidade nasal contém os órgãos do sentido do olfato e é forrada por um epitélio secretor de muco. Quando o ar respirado passa pelas cavidades nasais, o ar é purificado, umedecido e esquentado. O órgão olfativo é a mucosa que reveste a parte superior das fossas nasais (GUYTON, 1981).

As fossas nasais possuem dois tipos de mucosas (ver Figura 3):

- vermelha: contém as glândulas que secretam o muco e que mantém a região úmida. Possui esta coloração pelo fato de ser rica em vasos sanguíneos (GUYTON, 1981).
- amarela: é onde estão localizadas as terminações nervosas do nervo olfativo. Quando um odor é inspirado, a mucosa amarela é impregnada e então, os prolongamentos sensoriais (pêlos olfativos) são atingidos (GUYTON, 1981).

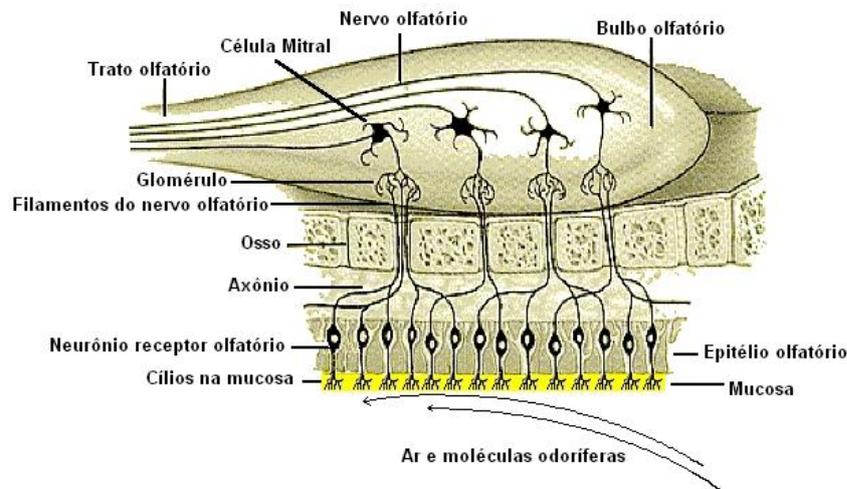


Figura 3: Mucosas Nasais (GUYTON, 1981)

Quando os pêlos olfativos são atingidos, geram impulsos nervosos, os quais são conduzidos até o corpo celular das células olfativas, passando para os axônios, que se comunicam com o bulbo olfativo. O agrupamento dos axônios penetra no osso etmóide para chegar ao bulbo olfatório, onde converge para formar estruturas sinápticas chamadas glomérulos. Estas se conectam em grupos que convergem para as células mitrais. Fisiologicamente essa convergência aumenta a sensibilidade olfatória que é enviada ao Sistema Nervoso Central (SNC), onde o processo de sinalização é interpretado e decodificado (GUYTON, 2002) (Figura 4).

A alta sensibilidade da mucosa olfativa produz a sensação de odor mesmo com poucas moléculas. A intensidade do odor está diretamente ligada a sensação percebida pelo indivíduo, ou seja, a sensação será tanto mais intensa quanto maior for a quantidade de receptores estimulados, o que depende da concentração da substância odorífera no ar (GUYTON, 2002).

O olfato ainda está intimamente ligado ao paladar, pois o gosto é uma combinação do olfato e paladar. Isto acontece para alguns tipos de sabores, onde o olfato ajuda na sua identificação, tal fato pode ser constatado quando se está resfriado onde perde-se um pouco da capacidade gustativa (GUYTON, 2002).

Segundo Guyton (GUYTON, 1981), o olfato tem grande capacidade adaptativa, ou

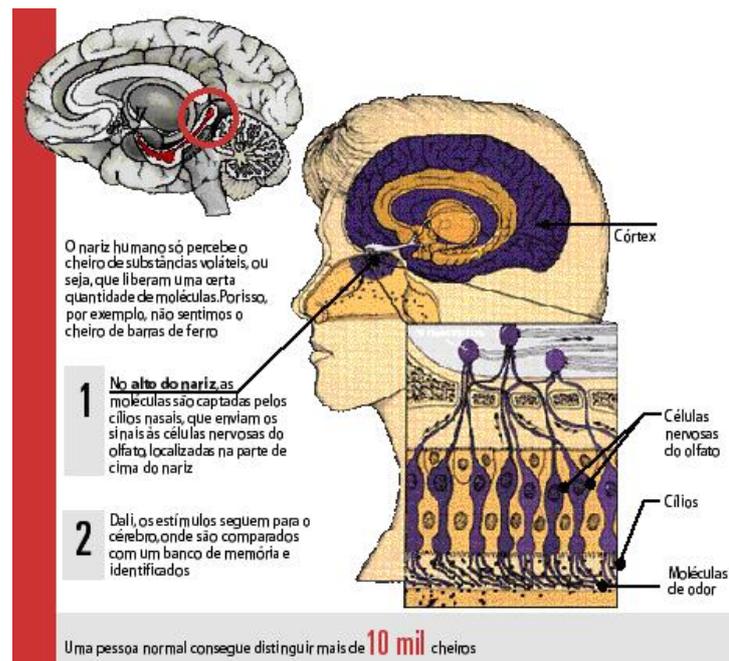


Figura 4: Cavidade Nasal Humana (PONTES, 2006)

seja, no início da exposição a um odor muito forte, a sensação olfativa pode ser bastante forte também, mas, após um certo tempo, o odor será quase imperceptível. Além disso, consegue detectar apenas uma sensação de odor por vez, se no ambiente possuir mais de um odor, o que contiver maior concentração prevalecerá.

2.3 Audição Humana

A onda sonora é gerada a partir das vibrações de um corpo, e esta, por ser uma onda mecânica, se propaga através de um meio físico. Normalmente esse meio é o ar e a onda sonora corresponde à variação da pressão atmosférica.

A primeira ação do sistema auditivo humano é a conversão das ondas sonoras em vibrações mecânicas. E, depois, em impulsos elétricos que serão enviados para o cérebro, onde são processados os dados recebidos, resultando na sensação da audição.

2.3.1 Ouvido Humano

O ouvido é o órgão reponsável por coletar os estímulos do mundo exterior, transformando as ondas sonoras em impulsos elétricos para o cérebro. O ouvido é uma estrutura mecânica muito complexa e sensível pois é capaz de detectar as menores quantidades de energia (LACERDA, 1988).

A seguir serão descritas as partes mais importantes do ouvido: ouvido externo, ouvido médio e ouvido interno (SANTOS; RUSSO, 1986).

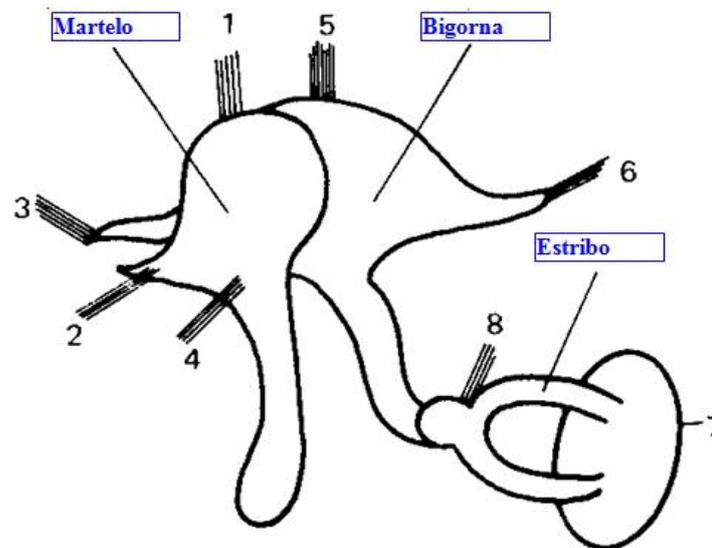


Figura 5: Cadeia Ossicular (SANTOS; RUSSO, 1986)

- Ouvido Externo: Composto do pavilhão auditivo, do canal auditivo e do tímpano. A função da orelha é de uma corneta acústica, capaz de ligar o espaço interior e o canal auditivo, otimizando a transferência de energia. A orelha ainda ajuda na localização da fonte emissora do som. As paredes do canal auditivo são compostas por ossos e cartilagens. O tímpano tem a forma de um cone e fecha o fundo do canal auditivo.
- Ouvido Médio: É uma cavidade preenchida por ar e contém três pequenos ossos: o martelo, a bigorna e o estribo. A função de tais ossículos é, através de uma

alavanca, acoplar mecanicamente o tímpano à cóclea (caracol), triplicando a pressão do tímpano (ver Figura 5).

- Ouvido Interno: Começa pela janela oval, segue um canal semicircular até o caracol (cóclea) e é dividido em duas galerias pela membrana basilar. O caracol é constituído por um canal duplo, enrolado em torno de um eixo ósseo. A janela oval fecha o compartimento superior e transmite suas vibrações para a membrana basilar através da endolinfa, líquido viscoso que preenche esse conduto. A janela redonda é uma membrana circular, elástica, que fecha a parte superior do canal e, suas contrações compensam as variações de pressão produzidas pelas oscilações da membrana basilar.

2.3.2 Fisiologia da Audição

A audição pode ocorrer de duas maneiras: via aérea e via óssea.

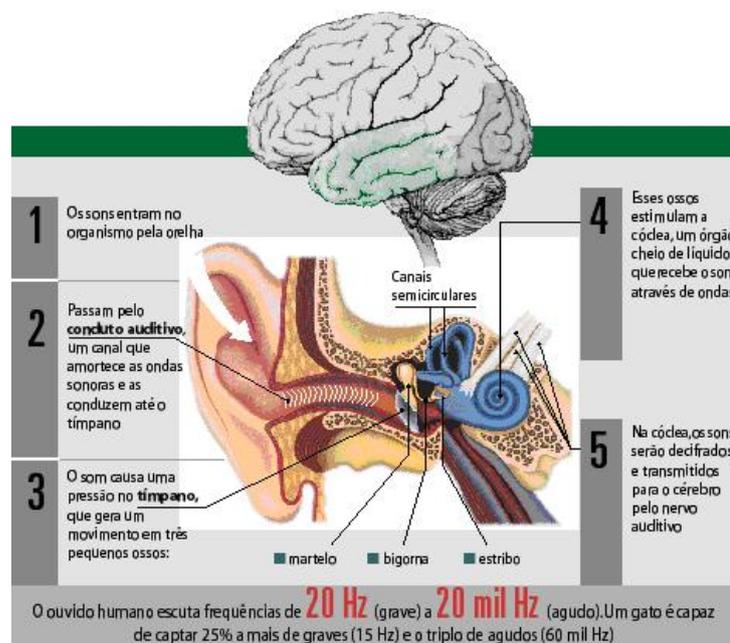


Figura 6: Passos da Audição (SANTOS; RUSSO, 1986)

- Via Aérea: A transformação do som em impulsos elétricos ao cérebro é o processo fundamental de audição. Este processo pode ser esquematizado da seguinte maneira

(Figura 6) (NEPOMUCENO, 1994; SANTOS; RUSSO, 1986):

- As ondas entram pelo pavilhão auditivo e percorrem o canal auditivo. A colisão da onda sonora com o tímpano, faz com que a membrana timpânica vibre na mesma frequência e amplitude do som de entrada. As ondas sonoras são convertidas de pressão para vibrações.
 - A vibração do tímpano é transmitida para o martelo que faz movimentar toda a cadeia ossicular. A vibração do martelo é transmitida para a bigorna e para o estribo, através de um sistema de alavancas.
 - A vibração do estribo é transmitida sobre a janela oval, que está em contato com o líquido do ouvido interno. A vibração é transformada em ondas de pressão no líquido.
 - As ondas sonoras se propagando nos líquidos do ouvido interno provocam a vibração da membrana basilar e do Órgão de Corti. A vibração chega até as células ciliadas, fazendo com que seus cílios oscilem saindo de sua posição de repouso.
 - A oscilação dos cílios (na mesma frequência da onda sonora original) causa uma mudança na carga elétrica endocelular, provocando um disparo de um impulso elétrico para as fibras nervosas que é conduzido para o nervo acústico e para o cérebro.
 - A indicação de qual célula ciliada irá responder ao estímulo vibratório depende da frequência do som: para sons agudos o deslocamento da membrana basilar é maior na região basal (próxima à janela oval) estimulando as células desta região; se o som é grave, o movimento maior da membrana basilar será na região apical.
- Via Óssea: Outra forma de um ser humano captar as energias sonoras é através das vibrações dos ossos, principalmente pela caixa crâniana. Um exemplo disso é quando se mastiga os alimentos ou no coçar a cabeça. A audição via óssea acontece

quando as ondas sonoras chegam até aos ossos da cabeça, fazendo-os vibrar, assim os ossos do ouvido médio também vibram e propagam a vibração até a cóclea, provocando ondas nos líquidos internos e fazendo que ocorra o processo auditivo (NEPOMUCENO, 1994; SANTOS; RUSSO, 1986).

2.3.3 Percepção de Volume do Som

O volume percebido pelo ser humano não é diretamente correspondente aos níveis de pressão do ar. O ouvido humano suporta variações de mais de 1.000.000 de vezes na pressão do ar da onda sonora, mas isto não corresponde a um aumento significativo na percepção do volume.

O ser humano é mais suscetível as variações de pressão da onda sonora para as pressões mais baixas do que as mais altas. Por exemplo, uma variação de $1.000\mu Pa$ para $2.000\mu Pa$ é mais perceptível que de $20.000\mu Pa$ para $21.000\mu Pa$. Isso mostra que não existe uma linearidade entre a pressão do ar e o volume percebido, por isso é que se usa uma escala logarítmica para indicar a percepção em volume.

A unidade de medida mais comum para o nível de pressão sonora é a SPL (*Sound Pressure Level*), a qual expressa o nível de pressão sonora em relação a um nível de referência. Segundo (PAINTER; SPANIAS, 2000), a pressão sonora no limiar de audibilidade em 1kHz

$$L = 20 \log_{10}(p/p_0)(dB SPL), \quad (2.1)$$

Na Tabela 1, pode-se verificar os volumes em dB SPL para alguns exemplos do cotidiano.

Situação	Pressão Sonora (dB SPL)
Limiar de Audibilidade	0
Murmúrio	30
Conversação Normal	60
Trânsito Pesado	80
Indústria Mecânica	100
Britadeira	120
Limiar da Dor	130
Avião à Jato	150

Tabela 1: Nível de pressão sonora para exemplos do cotidiano. (PAINTER; SPANIAS, 2000)

2.3.4 Resposta em frequência do sistema auditivo

A percepção do sistema auditivo humano também não ocorre de maneira linear, já que o ouvido humano é mais suscetível a frequências mais baixas do que as mais altas. Isso acontece pela estrutura fisiológica do ouvido humano. A membrana basilar, é responsável pela análise da frequência do som, e a maior parte desta membrana responde a sons com frequência menor que 3000 Hz, onde é localizada a maior parte do espectro de frequência da fala humana (LEITE, 2003).

2.3.5 Bandas Críticas

Moore (MOORE, 2003) caracteriza uma banda crítica como uma faixa de frequência ao redor de uma frequência central, na qual o sistema auditivo humano modifica-se abruptamente. Ou seja, o comportamento do ouvido humano varia de acordo com cada uma das bandas críticas.

É importante notar que cada pessoa tem uma frequência central para cada uma das bandas críticas, porém, para a maioria dos ouvintes, é possível fazer a seguinte generalização proposta por (PAINTER; SPANIAS, 2000):

$$BW_c(f) = 25 + 75[1 + 1.4(f/1000)^2]^{0.69} (\text{Largura da Banda Crítica em Hz}), \quad (2.2)$$

É comum discretizar as frequências nas bandas críticas. Para este trabalho foi utilizado o conjunto discreto mais referenciado na literatura, conforme pode ser visto na Tabela 2, a qual utiliza uma escala discreta chamada de escala Bark (CAVE, 2002).

Banda Crítica	Frequência (Hz)			Largura de Banda (Hz)
	Inferior	Central	Superior	
1	0	50	100	100
2	100	150	200	100
3	200	250	300	100
4	300	350	400	100
5	400	450	510	110
6	510	570	630	120
7	630	700	770	140
8	770	840	920	150
9	920	1000	1080	160
10	1080	1170	1270	190
11	1270	1370	1480	210
12	1480	1600	1720	240
13	1720	1850	2000	280
14	2000	2150	2320	320
15	2320	2500	2700	380
16	2700	2900	3150	450
17	3150	3400	3700	550
18	3700	4000	4400	700
19	4400	4800	5300	900
20	5300	5800	6400	1100
21	6400	7000	7700	1300
22	7700	8500	9500	1800
23	9500	10500	12000	2500
24	12000	13500	15500	3500
25	15500	19500		

Tabela 2: Bandas Críticas (CAVE, 2002)

As características das bandas críticas estão intimamente ligadas às da membrana basilar. Pohlmann (POHLMANN, 2000) diz que cada banda crítica corresponde a 1.3mm de espaçamento na membrana basilar. A associação posição-frequência não é linear. Isso explica o porquê das larguras das bandas críticas não serem lineares. A distância de uma banda crítica é conhecida como um Bark. A função a seguir permite converter frequências em Hertz para a escala Bark:(PAINTER; SPANIAS, 2000)

$$z(f) = 13 \arctan(0.00076f) + 3.5 \arctan[(f/7500)^2] \text{ (em Barks)} \quad (2.3)$$

2.3.6 Mascaramento

Quando um som se torna imperceptível para um ser humano devido a presença de outro, dizemos que este som foi mascarado. Som que provoca o mascaramento é chamado de mascarador. Este fenômeno é comum no dia-a-dia das pessoas; um exemplo simples é o toque de um celular que pode ser facilmente percebido quando se está em um ambiente em silêncio, porém, quando se está em um *shopping center*, por exemplo, esse som torna-se muito difícil de ser percebido. Ou seja, o limiar de audibilidade depende do ambiente e do seu nível de ruído.

O uso do mascaramento em computação é muito usado na área de codificação de sinais de áudio. Tanto que, hoje em dia, os mais importantes algoritmos de compressão de áudio utilizam este tipo de técnica, mascarando frequências que o ouvido não perceberia. Para este trabalho, utilizou-se o mascaramento para tornar ainda mais realístico o modelo auditivo implementado.

Existem duas categorias principais de mascaramento: simultâneo e o não simultâneo (ou temporal). A Figura 7 mostra esses tipos de mascaramento. Nela pode-se observar o limiar de mascaramento (linha pontilhada) em função do tempo, na qual o mascarador (linha sólida) está presente por 200ms, enquanto que seus efeitos estão presentes por cerca de 450ms. A seguir o mascaramento simultâneo é apresentado em mais detalhes, pois foi a categoria utilizada como inspiração neste trabalho.

2.3.6.1 Mascaramento Simultâneo

O mascaramento simultâneo ocorre quando os sinais relativos ao efeito do mascaramento estão presentes simultaneamente no ouvido. Ele reflete as limitações do sistema

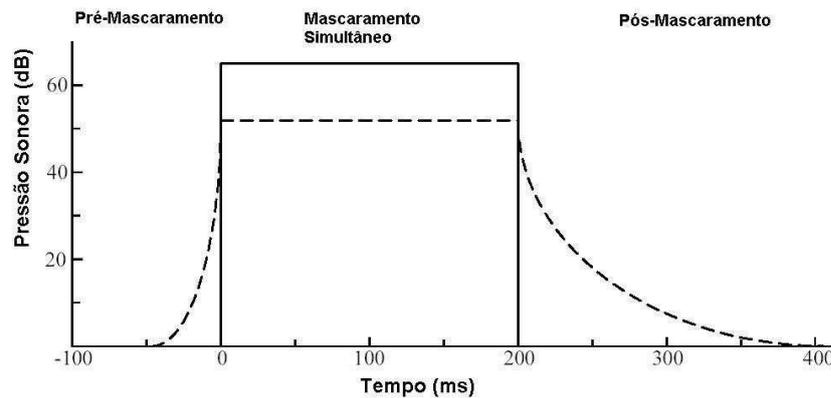


Figura 7: Principais Tipos de Mascaramento. (LEITE, 2003)

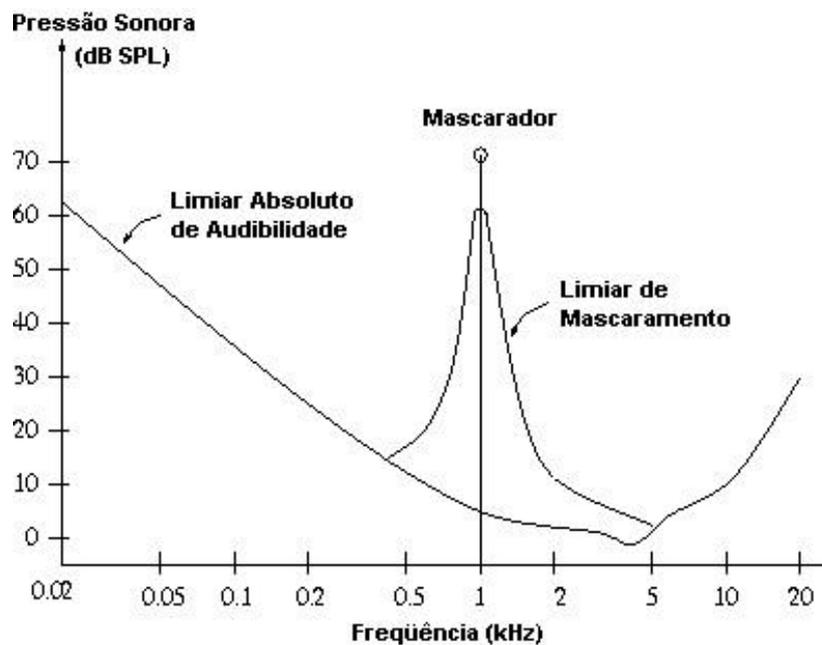


Figura 8: Efeito do mascaramento simultâneo (LEITE, 2003)

auditivo humano com relação à sua resolução em frequência.

Segundo Leite (LEITE, 2003), a Figura 7 mostra o mascaramento simultâneo, considerado o mais importante dos tipos de mascaramento, pois atinge os melhores níveis e ainda pode ser facilmente observado. Na Figura 8, é possível verificar a alteração do limiar de audibilidade na presença de um sinal tonal com nível de 70 dBs e com frequência de 1KHz. Qualquer sinal com intensidade inferior ao limiar de mascaramento será mascarado.

Existem alguns tipos de mascaramento simultâneo, tais como:

- Ruído mascarando tom
- Tom mascarando ruído
- Ruído mascarando ruído
- Tom mascarando tom

Neste trabalho foram considerados somente dois tipos de mascaramento simultâneos, ruído mascarando tom e tom mascarando tom.

Para o primeiro, um ruído mascara um sinal tonal, o mascaramento só ocorre quando a intensidade do tom a ser mascarado é menor que um determinado limiar, o qual é diretamente relacionado com a intensidade do ruído. Na maioria dos estudos, o limiar de mascaramento varia entre -5 e +5 dB. Portanto, em alguns casos, um ruído de menor intensidade pode mascarar um tom de maior intensidade.

O segundo cenário, tom mascarando tom, é pouco útil na área de codificação de áudio, já que os mascaramentos para sinais de áudio são mais complexos do que sinais puramente tonais. Porém, em nosso estudo é importante, pois podemos simular o impacto de dois alarmes ocorrendo ao mesmo tempo.

O mascaramento ocorre com mais intensidade para sons dentro das mesmas bandas críticas, porém, existe uma certa propagação para as outras bandas críticas. Este efeito é conhecido como espalhamento do mascaramento (LEITE, 2003).

Para os casos de ruídos mascarando tons, é possível observar que o mascaramento máximo ocorre na banda crítica que coincide com a frequência do sinal tonal. Porém, como já foi dito, ocorre o espalhamento das outras bandas críticas pela banda do sinal tonal, e para este trabalho usou-se uma das técnicas mais aceita de espalhamentos, a de Schroeder (PAINTER; SPANIAS, 2000), a qual possui um decaimento conforme:

$$SF_{dB}(z) = 15.81 + 7.5(z + 0.474) - 17.5\sqrt{1 + (z + 0.474)^2}(\text{em dB}), \quad (2.4)$$

onde z é a frequência em Bark.

Assim, para modelar o comportamento de humanos virtuais e torná-lo mais semelhante aos modelos perceptivos reais, é importante para os modelar com habilidades que os permitam adquirir informações do ambiente virtual onde estão inseridos. Como a quantidade e variedade de dados disponíveis para os agentes aumenta, seu processo de decisão tende a ser mais preciso. As próximas seções apresentam trabalhos relacionados no sentido de aumentar a informação recebida pelos agentes virtuais, através dos sensores (sentidos) virtuais.

Capítulo 3

Percepções Virtuais: Visão, Olfato e Audição

Para modelar o comportamento de humanos virtuais, é importante provê-los com habilidades, adquirir informações do ambiente virtual onde estão inseridos. Como a quantidade e variedade de dados disponíveis para os agentes aumenta, seu processo de decisão tende a ser mais preciso. Esta seção apresenta trabalhos relacionados no sentido de aumentar a informação recebida pelos agentes virtuais, através dos sensores (sentidos) virtuais.

3.1 Visão Virtual

Um aspecto importante para aumentar o realismo de simulações é prover os agentes virtuais com visão sintética. Existem diversos trabalhos na área, e por isso, existem muitos modelos propostos na literatura. Aqui serão demonstrados somente alguns trabalhos relacionados, pois este tema não foi abordado em maior nível de detalhe, já que se decidiu pela implementação de um modelo simplificado para este tópico.

Pan (PAN; HAN; LAW, 2005) implementou um protótipo que inclui um sensor visual

usando o método de *ray tracing*. Assim, os agentes podem analisar o ambiente, pelo uso de raios que partem da posição dos seus olhos, em um ângulo de visão determinado. Com isso, o humanóide pode computar as intersecções dos raios com os objetos que estejam próximos a ele. Outra proposta foi desenvolvida por Tu e Terzopoulos (TU; TERZOPOULOS, 1994), neste, um sensor local de temperatura e um sensor visual tem acesso a um banco de dados geométrico e físico da simulação.

Na área de visão baseada em *rendering*, Noser et al. (NOSER et al., 1995) propõem um método que associa cada objeto da cena a uma cor diferente, então os objetos podem continuar sendo identificados na imagem 2D projetada. Com isso, é fácil saber qual objeto está dentro do campo de visão, mantendo uma tabela de correspondência entre as cores e os objetos. Noser e Thalmann (NOSER; THALMANN, 1998) aprimoraram o método anterior adicionando um *buffer* de profundidade e um *buffer* de cor, de maneira a aumentar a precisão da localização dos objetos. Terzopoulos e Rabie (TERZOPOULOS; RABIE, 1995) propuseram um paradigma para visão ativa, usando algoritmos de visão computacional para analisar continuamente o fluxo da imagem da retina captadas pelos olhos de um agente animado, o qual tem o intuito de se locomover através de seu ambiente.

No trabalho de Reynolds (REYNOLDS, 1987) para concentrações de bandos de pássaros e cardumes de peixes, assim como no trabalho de Musse para humanos virtuais em multidões (MUSSE; THALMANN, 1997), cada agente sabe a posição de seus vizinhos através da percepção geométrica. Shao (SHAO; TERZOPOULOS, 2005) propôs uma melhoria na performance desta técnica utilizando mapas de percepção. Tal método inclui mapa (*grid maps*) para representar os objetos estáticos do ambiente, em cada região, e um mapa global que contém o trajeto dos objetos móveis (pedestres). Estes mapas armazenam informações dentro de cada uma de suas células, as quais identificam todos que estão ocupando aquela área celular.

3.2 Olfato Virtual

Diferentemente das outras percepções, o olfato tem sido pouco explorado na literatura, principalmente devido a complexidade inerente a esta percepção. E tal qual a visão virtual, um modelo simples foi implementado neste trabalho.

Em um dos artigos encontrados, Mamlouk (MAMLOUK et al., 2003) usa um método numérico para apresentar uma infraestrutura robusta para a análise e interpretação dos experimentos psico-físico e neuro-fisiológicos correntes e futuros, em termos do “espaço da percepção olfativa”. Um banco de dados, de percepção olfativa, foi projetado no espaço dimensional Euclidiano mais próximo, usando uma escala aproximada multidimensional. Isto produz uma interpretação Euclidiana independente da percepção do odor, não importando se este espaço é métrico ou não. Mapas auto-organizáveis são aplicados para produzir um mapa bidimensional desta aproximação Euclidiana do espaço da percepção olfativa.

Zhang e Wyvill (ZHANG; WYVILL, 1997) apresentaram um grupo de borboletas que navegam utilizando seu sistema olfativo. Os agentes são capazes de direcionar seus movimentos de acordo com um odor químico no espaço do *voxel* baseado na teoria da transferência de massa. Objetos no ambiente são escaneados e convertidos para representações em *voxels* para facilitar a detecção de colisão.

3.3 Audição Virtual

A inclusão de sons em ambientes virtuais é um aspecto importante para melhorar o sentido de presença e imersão (LARSSON; KLEINE, 2002). Existem vários trabalhos na literatura, relacionados com a síntese, propagação, atenuação e percepção de áudio em ambientes virtuais.

Vários autores tratam especificamente da renderização de sons com múltiplas fon-

tes geradoras de áudio. Em particular, é de especial interesse, a renderização de áudio espacial baseado em exemplos (arquivos de áudio), o que necessita de complexos modelos de processamento de sinais, mesmo para um pequeno número de fontes sonoras (TSINGOS; GALLO; DRETTAKIS, 2004). Este processamento pode incluir padrões direcionais de fontes de rendering (LOKKI et al., 2002), posição do som 3D (BEGAULT, 1994) e reverberação artificial (GARDNER, 1997; LOKKI et al., 2002).

Na maioria dos casos, a modelagem da propagação do som em ambientes virtuais, é tratada apenas geometricamente, sem levar em consideração sua natureza ondulatória (MAFFRA, 2003). Essa abordagem é conhecida como acústica geométrica e, apesar de não ser fisicamente correta, fornece uma simplificação razoavelmente aproximada à realidade, de maneira que análises do ambiente simulado, possam ser realizadas (KUTTRUFF, 2000).

Uma outra linha de pesquisa relacionada com a geração e síntese de áudio diz respeito à comunicação verbal de agentes autônomos, onde eles devem ter a habilidade de “falar” e “perceber” o que os outros agentes “dizem”. Monzani e Thalmann (MONZANI; THALMANN, 2000) descrevem simples métodos para propagação do som durante a conversação de agentes autônomos. A aproximação é feita através de distribuições radiais e angulares, respectivamente utilizando a distância e variação angular entre os agentes. Desta maneira o som é atenuado se os agentes estão distantes, ou ainda se não estão de frente um para o outro.

Noser (NOSER et al., 1995) trata agentes autônomos perceptivos, de maneira homogênea e que sintetizam vozes. Neste caso, as vozes são entrada para um sofisticado modelo de animação comportamental onde reações são disparadas. Mais recentemente, Conde e Thalmann (CONDE; THALMANN, 2004) apresentaram um modelo de percepção multi-sensorial para agentes autônomos, incluindo audição, além de outros sentidos. No caso da audição, esta foi implementada através de cones de percepção que filtram informação e propagam o som através de técnicas de acústica geométrica. A principal contribuição do trabalho de Conde é integrar em uma única máquina de decisões, os

canais multisensoriais dos humanos virtuais.

Ainda existem estudos referentes ao mascaramento do som, tais estudos na maioria das vezes são aplicados para compactação de áudio. Um trabalho recente foi proposto com esta finalidade, onde Leite (LEITE, 2003) propõe um método de compactação de áudio baseado no modelo perceptual auditivo de Johnston (JOHNSTON, 1998). Este modelo, que será melhor detalhado na seção 4.1, utiliza o mascaramento sonoro com a finalidade de excluir bandas de frequência que não serão percebidas pelo ouvinte devido ao mascaramento do som.

Capítulo 4

Modelo

Neste capítulo será descrito o modelo utilizado para prover os agentes virtuais de habilidades de percepção artificial. No modelo proposto, as percepções de cada agente levam em consideração as características do sistema auditivo humano, tais como: sensibilidades distintas em diferentes bandas de frequências e a perda natural da audição em relação ao envelhecimento (seção 4.1). É importante salientar que a perda auditiva devido ao envelhecimento não é constante no espectro audível. Então, é importante analisar a distribuição espectral do som emitido, isso pode ser obtido de maneira trivial com o uso da Transformada Rápida de Fourier (FFT).

A pseudo-comunicação e os sensores visuais e olfativos usam abordagens sensorial por zona, respeitando a distância e o ângulo descritos nas áreas sensoriais. Estas percepções são mais detalhadas nas seções 4.2 e 4.3. Para o modelo comportamental adotou-se uma técnica conhecida de IA chamada de árvores de decisão (seção 4.4). É importante salientar que tais métodos foram escolhidos de maneira a facilitar a sua implementação, já que o foco maior deste trabalho era prover aos agentes virtuais um modelo auditivo robusto, enquanto os outros modelos podem ser facilmente substituídos no futuro.

4.1 Modelo Auditivo

No modelo de audição virtual, os seguintes fatores são usados para medir a percepção dos agentes, os quais estão relacionados ao som emitido, ambiente virtual, bem como aos atributos individuais dos agentes:

- Som emitido e ambiente virtual:

1. Extração de propriedades do som a partir de um arquivo de som:

- intensidade relativa do som em cada banda de frequência;

- mascaramento sonoro causado por múltiplas fontes emissoras de som;

2. Propagação Sonora:

- perda energética devido a refração do som nas paredes;

- perda energética devido a distância até a fonte sonora.

- Individualidades dos Agentes:

1. Curvas de capacidade auditiva normal das pessoas em diversas faixas etárias;

2. Perda auditiva em relação as diferentes bandas de frequência;

3. Perda auditiva devido a problemas auditivos.

Inicialmente, dois arquivos de som (por exemplo, arquivo *wave*) são fornecidos ao sistema, contendo um estímulo auditivo que será inserido no ambiente (alarme) e um ruído existente no ambiente (barulhos de um escritório, por exemplo). Também é fornecida a localização da fonte sonora, dos ruídos e a intensidade do som (em dB).

A sensação auditiva é caracterizada por contrações e dilatações de volumes de ar, que causam variações de pressão. À medida que as ondas sonoras se dissipam, a energia da onda é atenuada, de modo que a energia decai quando a distância à fonte aumenta. De fato, a atenuação da energia é uma função inversamente proporcional ao quadrado da distância, de modo que a intensidade sonora a uma distância x da fonte é dada por:

$$I(x) = A_d(x)I_e, \quad A_d(x) = \frac{1}{x^2}, \quad (4.1)$$

onde $A_d(x)$ representa a atenuação pela distância e I_e a intensidade emitida na fonte.

Uma pessoa situada a uma distância x da fonte sonora recebe como entrada em seu sistema auditivo um som com intensidade $I(x)$. Na verdade, essa intensidade está distribuída ao longo do espectro auditivo (aproximadamente entre 125Hz e 20kHz), e a sensibilidade do ouvido humano é diferente para cada faixa de frequência, já que existem faixas mais audíveis do que outras. Por exemplo, a faixa da voz humana varia entre 300 e 3000 Hz, e é mais facilmente ouvida. Para levar em consideração essa variação de sensibilidade em função da frequência, foi utilizada a curva C de atenuação de frequência (A_f), definida pelo padrão IEC 60179 da Comissão Internacional Eletrotécnica (International Electrotechnical Commission). Esta função de atenuação é dada por:

$$A_f(f) = \frac{k_c(2\pi f)^2}{((2\pi f)^2 + 129.42^2)((2\pi f)^2 + 76655^2)} \quad (4.2)$$

onde $k_c \approx 5.91797 * 10^9$, e f é a frequência do alarme ou ruído em Hz. O gráfico da sensibilidade A_f em função da frequência é ilustrado na Figura 9 (o eixo das frequências é mostrado em escala logarítmica, para melhor visualização).

Além dessa variação de sensibilidade por frequência, há também uma perda natural de sensibilidade auditiva à medida que envelhecemos conhecida como presbiacusia. Tal perda de sensibilidade não é uniforme para todas as frequências, e aumenta progressivamente com a idade. Utilizou-se uma extensão do Critério de Fowler (ALMEIDA; CAMPOS; GÂNDARA, 2004), que fornece a atenuação da sensibilidade em função da idade, em algumas frequências chave:

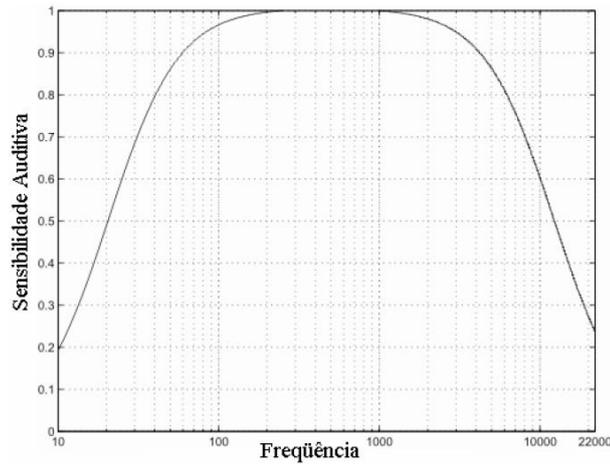


Figura 9: Gráfico de sensibilidade auditiva do ser humano em função da frequência.

$$A_a(\text{age}; f) = \alpha(f) + \beta(f)(\text{age}) + \gamma(f)(\text{age})^2, \quad (4.3)$$

onde os coeficientes α , β , γ , são dados na Tabela 3 de acordo com a frequência em teste¹, age é a idade do indivíduo.

Frequência (f)(Hz)	Coeficientes		
	$\alpha(f)$	$\beta(f)$	$\gamma(f)$
125	3,31	-0,262	0,0052
250	7,21	-0,483	0,0078
500	8,85	-0,594	0,0096
1000	12,36	-0,794	0,0120
2000	14,06	-0,925	0,0145
3000	12,16	-0,879	0,0157
4000	9,1	-0,747	0,0153
6000	9,11	-0,794	0,0172
8000	9,62	-0,87	0,0194

Tabela 3: Valores atribuídos aos coeficientes α , β e γ , para diferentes frequências chaves.

Como a Tabela 3 fornece a atenuação apenas para algumas frequências, o espectro de potência de um dado som foi dividido em bandas de frequência, conforme ilustrado na Tabela 4. Todas as frequências em cada banda são atenuadas de acordo com o coeficiente de atenuação da frequência chave da banda, conforme a Tabela 3.

¹Como pode ser observado, não se possui dados referentes a pesquisas com frequências mais altas do que 8 kHz, portanto, para estas frequências foram utilizados os mesmos coeficientes de 8 kHz.

Banda de Frequência (Hz)
0 - 125
125 - 250
250 - 500
500 - 1000
1000 - 2000
2000 - 4000
4000 - 6000
6000 - 8000
8000 - 20000

Tabela 4: Bandas de Frequência.

Além disso, podem-se considerar as perdas adicionais de audição, devido a fontes externas (como exposição exagerada a ambientes ruidosos). Cada indivíduo pode sofrer uma perda adicional de acordo com a Tabela 5, que é utilizada pela Previdência Social no Brasil, para classificar a redução da audição da pessoa estudada. Ainda, existe a atenuação provocada pela refração do som nas paredes (Tabela 6) (BECKETT; ARIBA, 1972).

Problema auditivo	Extensão da perda auditiva
Audição Normal	até 25 dB
Perda Leve	entre 26 dB - 40 dB
Perda Média	entre 41 dB - 60 dB
Perda Severa	acima de 60 dB

Tabela 5: Decreto Número 3.048 da Previdência Social de 06 de maio de 1999

Finalmente, quando o som chega ao ouvido do agente, ele sofre atenuações devido à distância entre a fonte e o agente, à variação na sensibilidade auditiva do ouvido humano, e às perdas esperadas de audição devido ao envelhecimento. Se a intensidade percebida pelo agente for maior do que um limiar T e maior do que o ruído existente², o agente reage ao estímulo. Embora o limiar da audição corresponda a 0 dB, sons muito baixos geralmente não são percebidos pelos agentes, Assim, neste trabalho foi utilizado um valor de T correspondendo a sua condição auditiva (Tabela 5).

²Para o cálculo do ruído, todas as atenuações (causadas pela distância, paredes, frequência e idade) foram aplicadas de maneira análoga ao som emitido pelo alarme.

	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz
alvenaria ou concreto 15cm	35	36	40	46	52	58
bloco concreto vazado 15cm	33	34	35	39	45	51
gesso 12mm	21	26	27	28	29	30
madeira 6mm	5	11	15	17	19	20
chapa de aço 1.5mm	25	31	37	42	46	49
vidro monolítico 3mm	12	17	22	26	27	28
vidro monolítico 6mm	17	23	25	26	27	28

Tabela 6: Tabela de materiais *versus* atenuação em dBs por faixa de frequência

O algoritmo proposto para percepção auditiva de um certo agente virtual pode ser resumido nos seguintes passos:

1. Ler a posição do alarme, o arquivo de áudio *wave*, e a intensidade I_e do som emitido (em dB).
2. Calcular a FFT do som, obtendo o vetor $F_0(i)$ com reações diferentes para cada frequência.
3. Compactação do vetor $F_0(i)$, explicado na seção 2.3.5, gerando novo vetor $F_0(j)$
4. Com base na distância x do agente à fonte, e as perdas de audição, multiplicar o vetor $F_0(j)$, pelas atenuações e reduções $A_d(j)$, $A_a(j)$ e $A_f(j)$ (aqui, j denota as bandas de frequências em bandas críticas, ver seção 2.3.5). O vetor resultante é chamado de $F(j)$.
5. Calcula a perda L (em dB) entre o sinal emitido e o sinal percebido, através da razão das energias.
$$L = 10 \log_{10} \left(\frac{\sum F_0(i)^2}{\sum F(i)^2} \right)$$
6. Acrescentar ao valor L as perdas adicionais para o agente em questão, conforme a Tabela 5 e perdas causadas pelas paredes, obtendo a perda total L_T .
7. Todos os itens anteriores, são computados para calcular o ruído existente (mascaramento, ver seção 2.3.6) no ambiente, gerando R_e .

8. Se $(I_e - L_T > T)$ e $(I_e - L_T > R_e)$, a *flag* de audição é marcada.

4.2 Modelo Visual e Pseudo-Comunicação

Em uma simulação com muitos agentes virtuais, se torna essencial o uso de uma técnica de percepção visual que capture informações do ambiente com um *frame rate* razoável, já que quanto mais rápido o agente detectar os objetos que estão no ambiente, mais rápido será a sua decisão, resultando em uma animação mais suave e realista. Este argumento é reforçado quando outros processamentos ainda devem ser realizados, como é o caso da percepção auditiva.

Neste trabalho, se utiliza uma abordagem por zona geométrica para descrever os modelos visual e de pseudo-comunicação. Esta região é a combinação da percepção da distância d_i e o ângulo de visão θ_i , os quais são atributos com valores determinados no início da simulação. A percepção de outro agente e/ou um evento é realizada quando o elemento entra na área sensorial do agente (Figura 10.a).

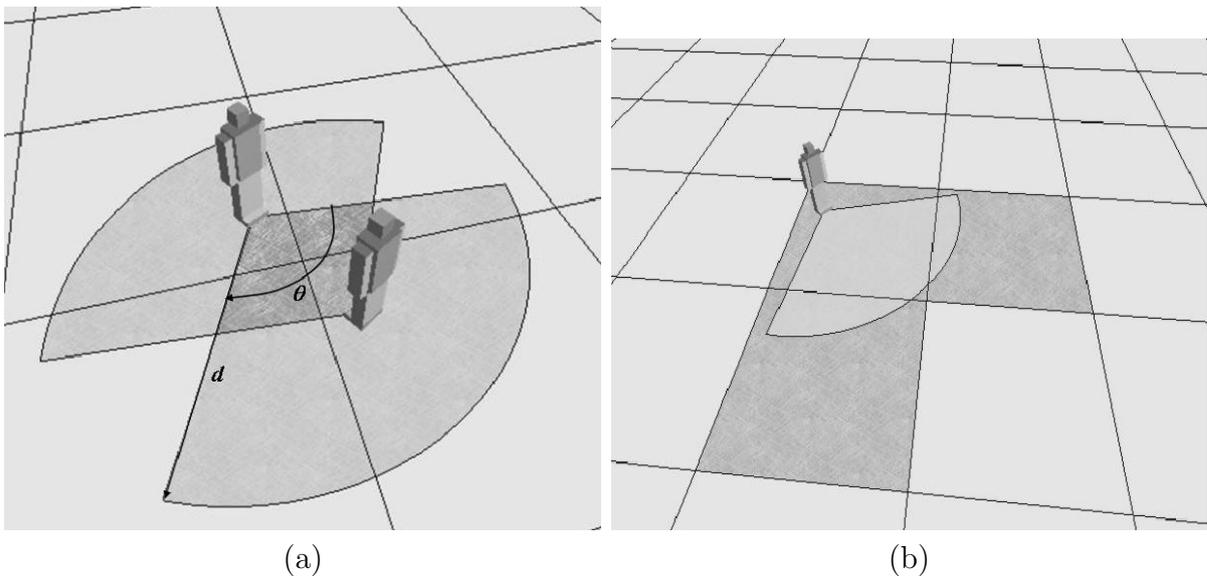


Figura 10: (a) Ambos agentes se percebem (b) usando o mapa de percepção para aumentar a performance da abordagem por zona.

4.2.1 Pseudo-comunicação

Cada agente tem um atributo individual (chamado C) o qual caracteriza seu conhecimento e associado as outras informações providas pelos modelos perceptivos. Como o modelo perceptivo contém três sentidos (visão, olfato e audição), o valor do atributo é uma representação decimal da codificação binária dos sentidos. Então, quando existe uma comunicação entre agentes, o valor do atributo para ambos agentes será uma aplicação do operador binário “ou”. A pseudo-comunicação (PC_{ij}) entre o agente i e o agente j é definida como:

$$PC_{ij} = C_i \vee C_j \quad (4.4)$$

Após a pseudo-comunicação, o valor do atributo C para ambos agentes será o resultado da equação 4.4.

Normalmente, as técnicas que utilizam a abordagem por zona possuem um tempo de processamento computacional alto. Para otimizar o desempenho, representa-se o chão virtual por um *grid map* (chamado de mapa de percepção), adaptado do Shao (SHAO; TERZOPOULOS, 2005). Este mapa representa objetos estacionários em uma região, tais como eventos ou obstáculos, bem como elementos dinâmicos, tais como os agentes virtuais. Neste trabalho, esse *grid* armazena informações que identificam todos os elementos (eventos ou agentes), que estão ocupando cada área celular. Cada agente transmite ao chão sua nova posição e, se necessário, a área celular é atualizada. Entretanto, somente as células as quais pertencem a intersecção entre chão virtual e a respectiva zona de um agente são consideradas para achar os vizinhos (Figura 10.b).

4.3 Modelo Olfativo

Existem certas aplicações onde a percepção olfativa pode apresentar relevância na coerência dos resultados adquiridos. Particularmente no caso de simulação de multidões, com uso de agentes que possuem a habilidade de detectar odores, pode resultar em simulações mais realísticas, provendo diagnósticos e comportamentos mais precisos durante um evento de perigo.

Entretanto, a precisão de um modelo de percepção olfativa não é uma tarefa simples. Na realidade, estímulos olfativos no nariz, produzido por moléculas, são transmitidas pelos neurônios até o córtex cerebral, mas ainda sim é desconhecido como o cérebro gera o reconhecimento do padrão deste odor. O sistema olfativo humano é capaz de detectar odores em pequenas concentrações (uma parte por milhão até uma parte por bilhão) dependendo do odor em questão. Além disso, é mais fácil detectar incrementos do que decrementos na concentração e magnitude percebida, e essas mudanças não são lineares, na realidade, estão mais próximas de uma escala logarítmica (ZELLNER; BARTOLI; ECKARD, 1991).

Devido a inerente complexidade do sistema olfativo humano, neste trabalho utilizou-se um modelo simplificado com o intuito de implementar um sistema olfativo para agentes virtuais. O modelo proposto usa abordagem sensorial e por zona do agente. Após a detecção do odor, a zona de percepção do olfato é subdividida em sensores que irão conter valores normalizados de acordo com a equação:

$$v_s = \begin{cases} 1 - \frac{d_s}{d_m}, & \text{se } d_s \leq d_m \\ 0, & \text{caso contrário.} \end{cases} \quad (4.5)$$

onde d_s é a distância do sensor (agente virtual) ao foco do evento olfativo, d_m é a distância máxima de propagação do odor e v_s indica a proximidade do sensor em relação ao foco (posição) do evento olfativo. Uma vez calculados os valores dos sensores, é realizada uma busca pela região (uma matriz quadrada de 3x3 para cada sensor) que

conterá a maior média sensorial. Uma vez conhecida esta região, a posição do seu centro será utilizada para direcionar o agente para a localização do evento. Cabe salientar que esta orientação, por meio da percepção olfativa, somente é utilizada quando a reação do agente ao detectar o evento seja “investiga”, dada pelo modelo comportamental (Seção 4.4). A Figura 11 ilustra a zona da percepção olfativa para o agente.

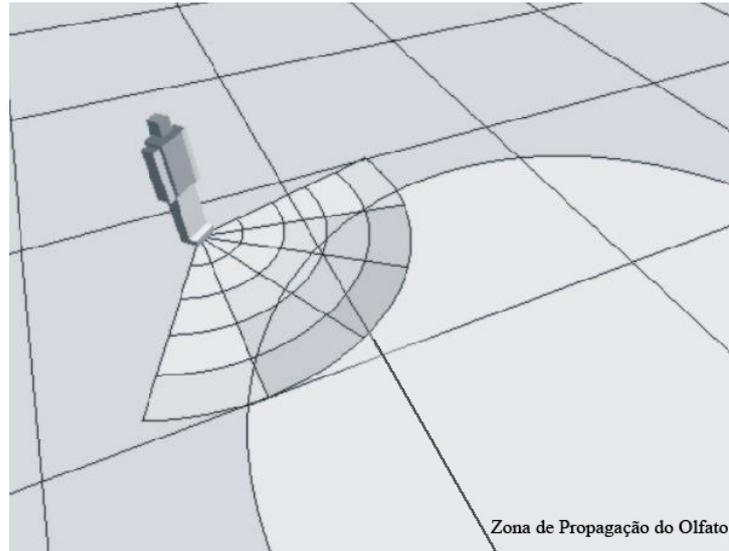


Figura 11: Zona de percepção olfativa para os agentes.

4.4 Modelo Comportamental

A modelagem do comportamento humano não é uma tarefa trivial, existem muitas variáveis inerentes para os seres humanos. Isso motivou a criação de um modelo comportamental aos agentes virtuais que utilizassem os sensores perceptivos (sentidos) e com base nos dados percebidos os agentes decidam coerentemente o próximo passo de simulação.

Então para a construção do modelo comportamental dos agentes, utilizou-se uma técnica bastante conhecida da Inteligência Artificial: as árvores de decisão. Certamente esta é uma das metodologias mais tradicionais de aprendizado de máquina, utilizada para a implementação da tomada de decisões por agentes (RUSSELL; NORVIG, 2002). Utilizando esta técnica, pretende-se prover agentes capazes de reagir de maneira distinta, de acordo com as informações adquiridas por meio de seus sensores (sentidos).

Para construir as árvores utilizadas neste trabalho, foi definido um questionário aplicado a um grupo de indivíduos que responderam perguntas relativas ao seu comportamento em situações de emergência. O principal objetivo do questionário era tornar possível a aquisição de um conjunto de regras, a partir das respostas dos entrevistados em determinadas situações. A seguir, é possível verificar um exemplo de pergunta utilizada no questionário:

- “O que você faria se visse um evento de perigo?”
 - (1) Nada
 - (2) Fugiria
 - (3) Investigaria

No caso citado, o sentido “visão” é a percepção relevante. Da mesma forma, outras questões considerando as demais percepções foram realizadas (inclusive considerando diferentes combinações perceptivas). As respostas dadas pelos entrevistados foram armazenadas em um banco de dados, sendo utilizado como entrada para o software Weka (WITTEN; FRANK, 2005). Este software foi utilizado para a geração automática das árvores (regras) utilizadas neste trabalho. Todas as árvores criadas são do tipo J48 e com *10-fold cross validation* (WITTEN; FRANK, 2005). Através das regras extraídas, foi possível recriar um comportamento generalizado das pessoas em função de um evento de perigo. Por consequência, essas regras foram integradas ao protótipo de simulação, com a intenção de guiar o comportamento dos agentes em situações de emergência.

4.5 O Protótipo

O protótipo foi desenvolvido inteiramente para este projeto, para isso utilizou-se a linguagem de programação C++, pois pudemos utilizar os conceitos de Orientação objetos para componentizar os modelos apresentados no capítulo anterior. Para uma melhor visualização do simulador, combinou-se com o C++, a ferramenta *OpenGL*.

A componentização tornou-se uma tarefa importante para este trabalho, pois assim, podemos, no futuro, aproveitar os conceitos de OO para acrescentarmos, substituímos e/ou atualizarmos o simulador já desenvolvido.

A figura a seguir mostra a relação entre os componentes dos sensores e a componente comportamental.

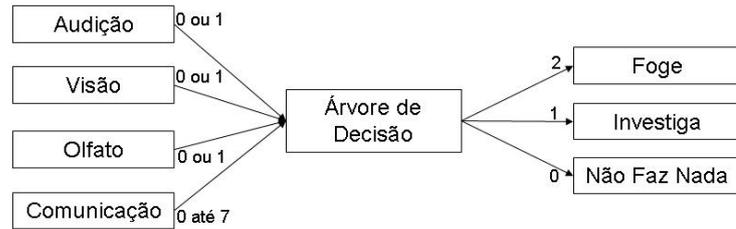


Figura 12: Componentes do Simulador.

Capítulo 5

Resultados

Neste estudo, desejou-se fazer com que a simulação das situações de emergência, envolvendo multidões, tivesse um comportamento mais próximo da realidade. Isto foi possível com a combinação dos modelos apresentados previamente na seção 4, já que possibilitou que os agentes reagissem a partir da combinação da percepção de sinais sonoros e do que foi percebido pelos sensores de olfato e de visão e ainda com o resultado da comunicação entre os agentes. Neste modelo, os agentes são representados por partículas e não possuem tratamento de colisão ou formação de grupos.

Para isso, a seguir, pode-se ver os dados de entrada do simulador desenvolvido. Para a audição de entrada são considerados os seguintes parâmetros.

- P: Propriedades do som do alarme (retirados dos dados de entrada do alarme);
- P': Propriedades do som do mascarador, ruído, retirados dos dados de entrada mascarador;
- D: Distância dos agentes a fonte emissora do alarme (calculado em tempo de execução);
- I: Idade dos agentes (definido no arquivo de configuração);
- A: Níveis de perda auditiva de cada agente (definido no arquivo de configuração).

Combinando com esses parâmetros, existem os dados da visão, olfato e comunicação, os quais são processados no mesmo modelo perceptivo. Estes são:

- Raio do campo sensorial;
- Ângulo do campo sensorial;
- Evento
 - Posição;
 - Raio de propagação do cheiro;
 - Dimensão do evento (no caso, raio).

Por fim, para avaliar os resultados deste modelo, calculou-se o número de agentes salvos na simulação. Em outras palavras, o agente que consegue perceber o evento de perigo e fugir do ambiente.

A ação de fugir está diretamente relacionada com o uso das árvores de decisão apresentada anteriormente.

5.1 Condições Iniciais

Para obter os resultados, as simulações foram realizadas em um ambiente contendo uma área de $600m^2$, subdivida em seis salas de $40m^2$, uma sala de $240m^2$ e um corredor de $120m^2$ (Figura 13). No maior ambiente, existem 50 agentes, enquanto cinco agentes foram posicionados em cada uma das salas restantes. As posições dos agentes são definidas aleatoriamente, totalizando 85 agentes na simulação.

Para validar este modelo, diversas situações com diferentes habilidades perceptivas foram simuladas. Para otimizar o desempenho computacional, as percepções visual e olfativa foram executadas somente quando o agente estivesse no mesmo recinto do evento visual e/ou olfativo. Os gráficos dos resultados e respectivas análises consideraram a

porcentagem de agentes salvos. Neste contexto, um agente foi considerado salvo quando estivesse há cerca de meio metro de uma das saídas de emergência do ambiente (foram definidas duas saídas de emergência, localizadas nos extremos do corredor). Além disso, a velocidade de deslocamento média de um agente utilizada foi de 0,9 m/s e um tempo de simulação de 180 segundos.

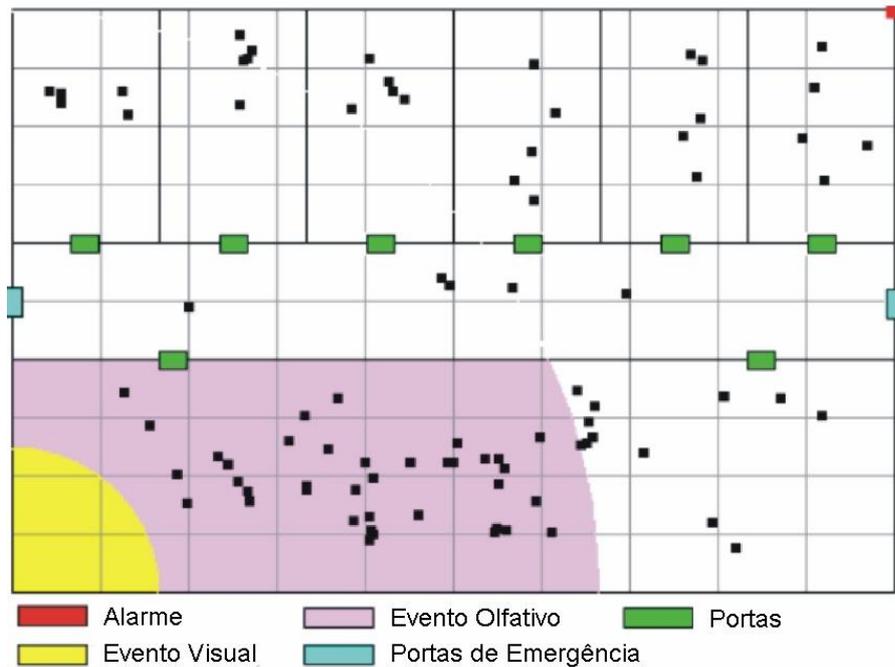


Figura 13: Captura da tela do simulador.

No caso da percepção auditiva, simulou-se o ambiente com um alarme, e este foi posicionado em três diferentes locais: canto inferior esquerdo, canto superior direito e no centro. Os arquivos de som contém informações de 6kHz e 12kHz, e ainda pode ou não haver ruído (mascaramento) na simulação. O ambiente foi povoado com nove possíveis combinações de população, considerando 3 grupos de acordo com suas idades (jovens, idosos e idade média de 55 anos) e 3 grupos de acordo com as suas condições auditivas (normais, médios e doentes - Tabela 5.1).

A fim de quantificar o impacto na quantidade de agentes salvos, quando estes estão presentes em um ambiente que possui algum tipo de ruído que pode mascarar o som, foi considerado o seguinte cenário: alarmes a 6 kHz e 12 kHz, agentes com condição auditiva normal e jovens. Em todas as simulações o alarme foi ajustado a 100dB.

Audição	População		
	Normal	Média	Grave
Normal	70%	25%	10%
Perda Leve	10%	25%	10%
Perda Média	10%	25%	10%
Perda Grave	10%	25%	70%

Tabela 7: Populações Utilizadas - Perdas Auditivas

5.2 Análise dos Resultados

Nesta seção serão mostrados e analisados os resultados obtidos.

5.2.1 Impacto da Frequência do Som

A Figura 14 apresenta o resultado das simulações considerando apenas a percepção auditiva com as devidas atenuações na propagação do som e desativando a pseudo-comunicação. É possível verificar que quanto maior a frequência do som, menor é o número de agentes que percebem o alarme, impactando o número de agentes salvos. Também é possível verificar que quanto mais doente e idosa a população, menor é o número de agentes salvos.

Na Figura 15 foram realizadas simulações ativando a pseudo-comunicação. É possível perceber que a disponibilidade de tal habilidade minimiza as perdas auditivas relacionadas a doenças e à idade dos agentes, como era esperado.

5.2.2 Impacto da atenuação causada pelas paredes

A Figura 16 apresenta o impacto da atenuação das paredes em relação a percepção auditiva de uma população de agentes jovens e saudáveis. É possível observar o decréscimo de agentes salvos para cada uma das posições testadas do alarme. Em contra-partida, desconsiderando a atenuação referente as paredes, praticamente a totalidade de agentes se salva, em decorrência da área do ambiente virtual ser relativamente pequena em relação

Impacto da frequência do alarme sem influência da comunicação

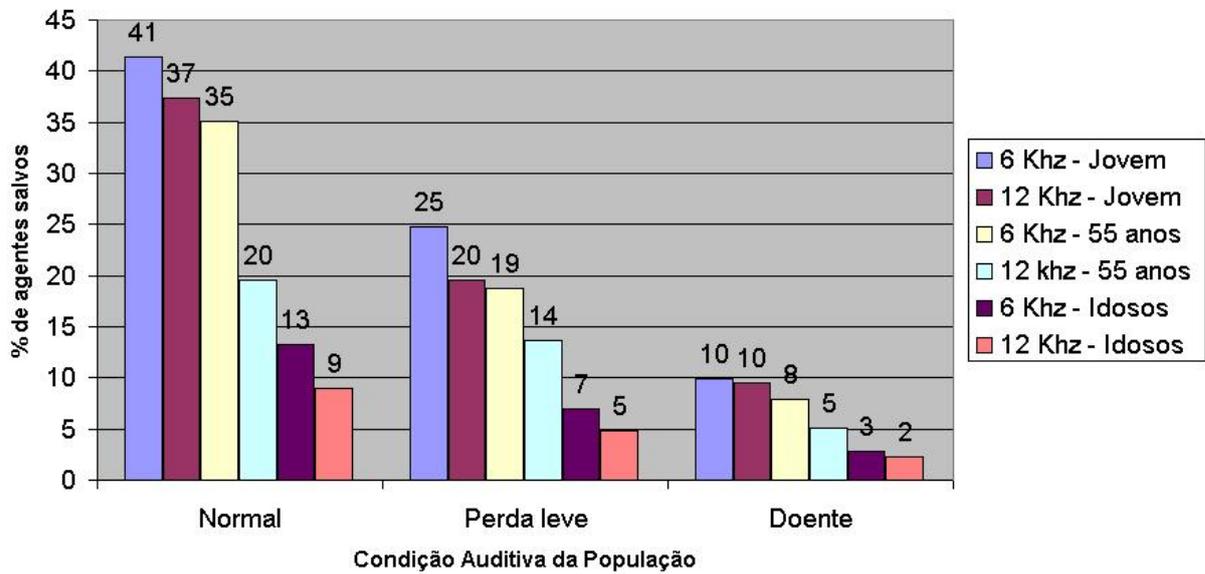


Figura 14: Audição sem comunicação.

Impacto da frequência do alarme com influência da comunicação

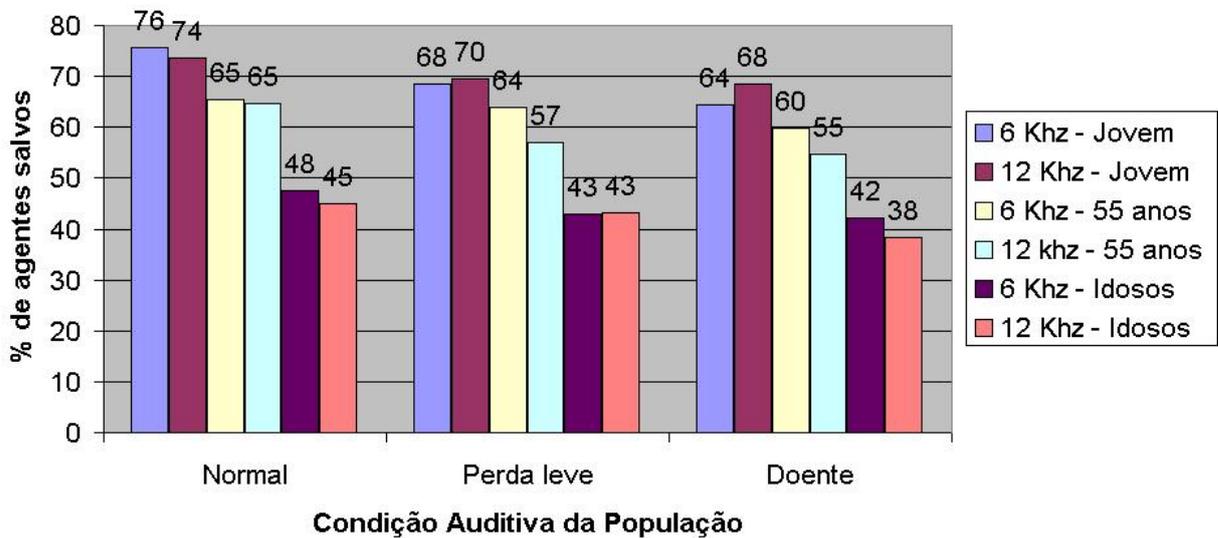


Figura 15: Audição com comunicação.

ao volume do alarme.

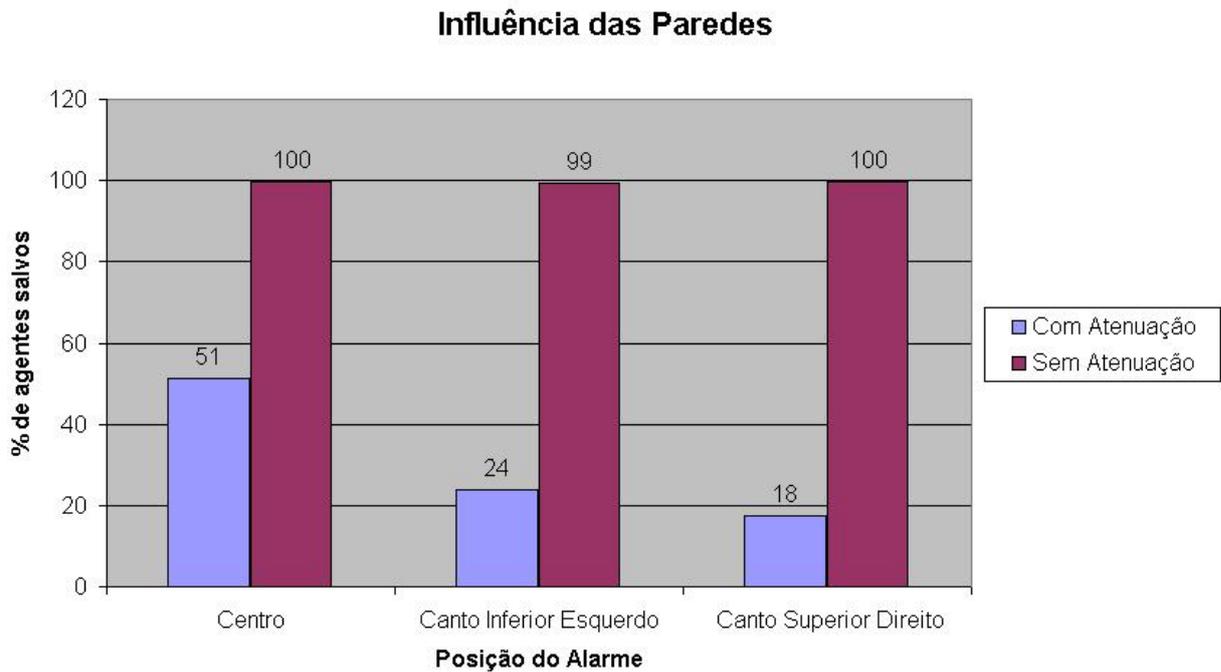


Figura 16: Influência das paredes no modelo auditivo.

5.2.3 Comparações entre as percepções

A Figura 17 apresenta uma análise do número de pessoas salvas considerando apenas a capacidade visual e olfativa para diferentes locais do evento de perigo. Conforme esperado, a percepção visual apresentou melhores resultados que a percepção olfativa, devido ao seu maior espaço de abrangência, além de apresentar uma maior precisão na informação adquirida. Com a ativação da pseudo-comunicação, os resultados encontrados foram proporcionalmente os mesmos, demonstrando que esta habilidade não influencia o predomínio da percepção visual sobre a olfativa.

Finalmente, a Figura 18 apresenta uma análise considerando todas as percepções (audição, visão e olfato) sem a pseudo-comunicação, para diferentes locais do evento do perigo e do alarme (posições coincidentes). Inicialmente, pode-se verificar que localizar o alarme no centro da edificação seria a melhor escolha para o ambiente virtual definido neste caso de teste. Além disso, devido ao número de agentes concentrados na maior sala, maior foi a porcentagem de agentes que perceberam, quer seja pela visão e/ou pelo olfato, o evento de perigo ocorrido neste recinto. Também é possível verificar um número

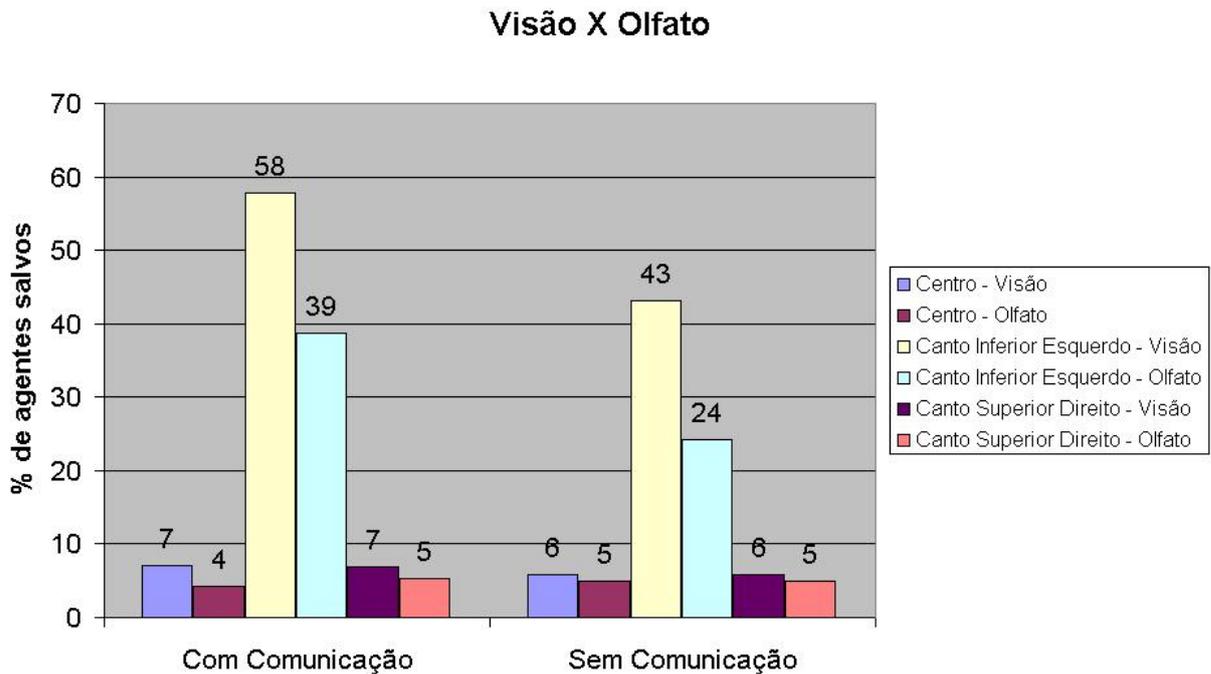


Figura 17: Visão *versus* Olfato.

expressivo de agentes que perceberam o evento de perigo, quando este ocorreu no centro do prédio. Esta situação se deve ao fato de que agentes situados na maior sala visualizaram o evento de perigo situado no corredor, ao evacuarem o prédio por terem escutado o alarme.

5.2.4 Impacto do Som Mascarador

Os resultados apresentados pelas figuras 19 e 20 mostram que, em ambos os casos, quando o som mascarador está presente na mesma banda de frequência do alarme, o número de agentes salvos é menor. Isso de fato deveria ocorrer, pois a interferência do mascarador é maior para faixas de frequências parecidas (bandas críticas próximas).

Outro fato importante a ser salientado é que quanto menor a distância entre as bandas críticas do mascarador e do alarme o efeito anterior também pode ser comprovado.

Por fim, nota-se também que o impacto da frequência continua presente, já que em todas as faixas analisadas o número de agentes salvos com alarme de 12 kHz é menor que o de 6 kHz.

Audição X Olfato X Visão

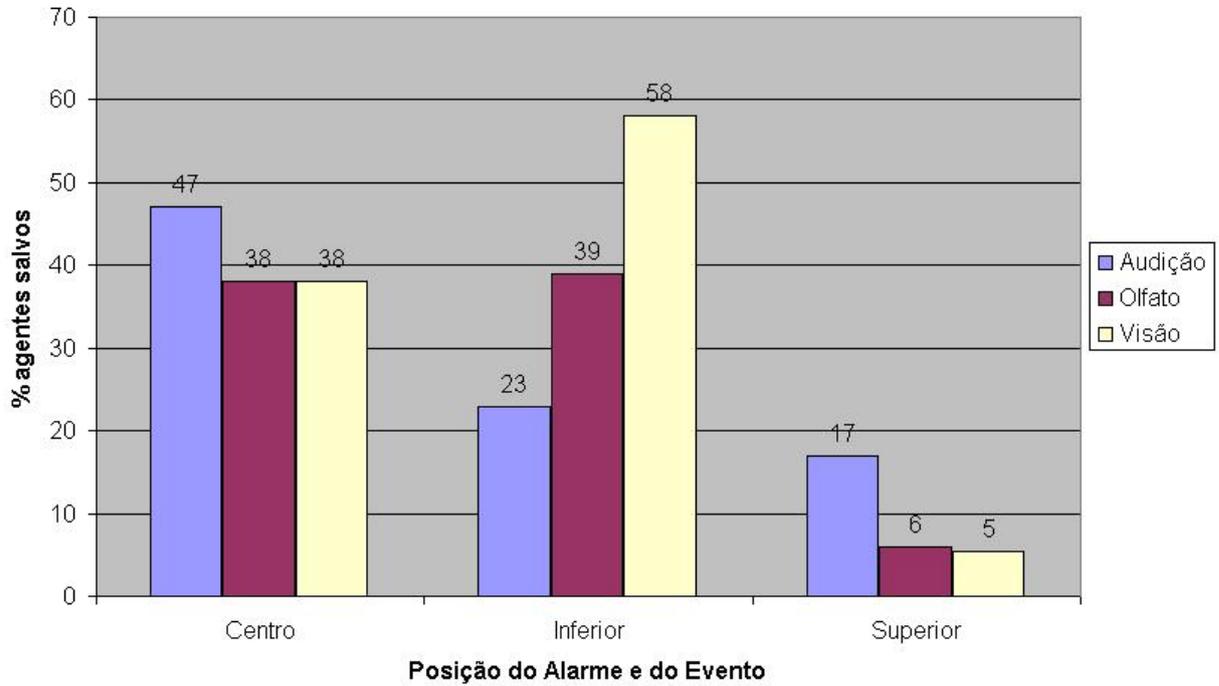


Figura 18: Audição *versus* Olfato *versus* Visão.

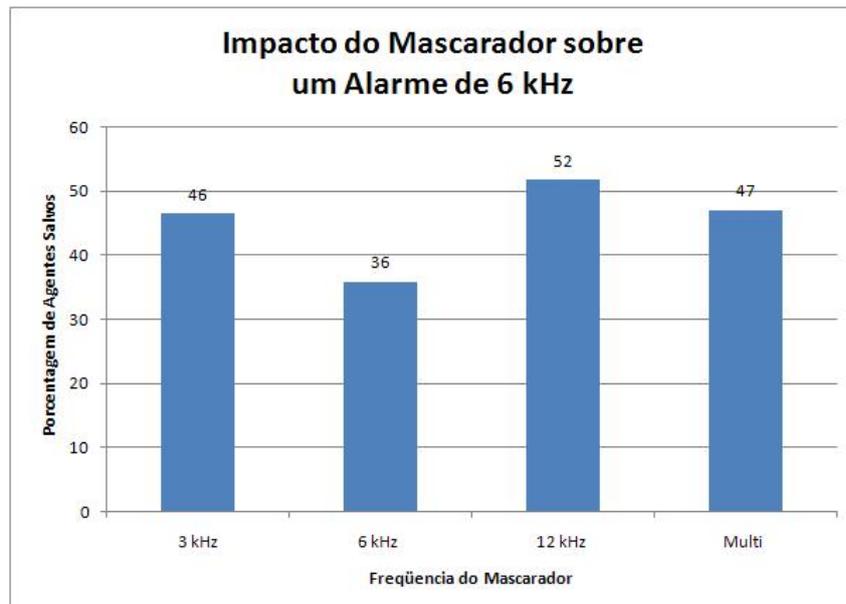


Figura 19: Impacto do mascarador sobre o alarme de frequência 6 kHz

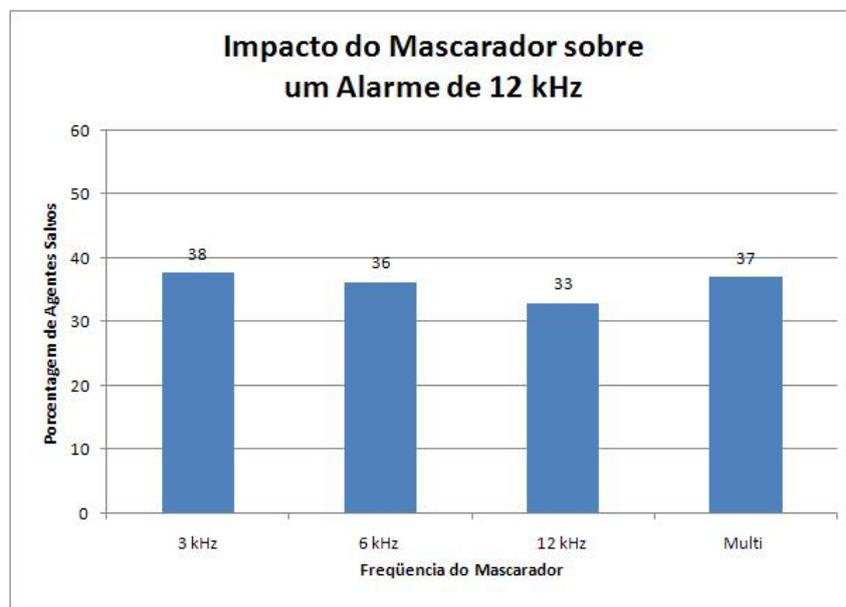


Figura 20: Impacto do mascarador sobre o alarme de frequência 12 kHz

Capítulo 6

Considerações Finais

O trabalho apresentado visou desenvolver um modelo multi-perceptivo de humanos virtuais, focado principalmente em desenvolver um trabalho mais realístico na percepção auditiva. As percepções visuais e olfativas foram desenvolvidas de maneira simplista, a fim de complementar o sistema auditivo desenvolvido.

Tanto o modelo visual como o olfativo e ainda a pseudo-comunicação se valeram de técnicas bem conhecidas pela comunidade acadêmica, utilizando somente técnicas de aproximação geométrica e algumas otimizações afim de executar o simulador de maneira mais rápida. Já o modelo auditivo apresentado possuiu algumas características bem definidas, tais como:

- Características individuais dos agentes:
 - Perda auditiva em relação a idade;
 - Perda auditiva em relação a condição auditiva.

- Extração das propriedades de um arquivo de som no formato *wave*:
 - Atenuação inerente ao espectro de frequência da onda sonora (por exemplo, quanto mais agudo um som, mais difícil será a sua percepção).

- Atenuação da intensidade sonora devido a refração do som nas paredes;
- Atenuação referente a distância do agente à fonte sonora.
- Mascaramento do alarme causado por outros sons.

O simulador desenvolvido comprovou aquilo que era esperado, ou seja, à medida que as frequências ficam muito altas, a sensibilidade de percepção dos agentes diminui (Figura 9). Isso ocorre também com pessoas mais idosas em relação às jovens, e com problemas de deficiência auditiva em relação às pessoas com percepções na faixa de normalidade. A melhoria nos aspectos perceptivos dos humanos virtuais neste primeiro modelo é de extrema importância para a simulação de multidões em situações de emergência.

Sem o aspecto auditivo incluído, o modelo original, utilizando flags aleatórios ou apenas usando a distância em relação ao alarme, poderia simular as diferentes percepções dos indivíduos, através de métodos de acústica geométrica. No entanto, certamente o resultado não seria tão realístico e comparável à vida real. Mais ainda, simulações com áudio verdadeiro de alarmes reais podem ser simulados através deste modelo. Outro impacto esperado era a redução significativa causada pelas paredes (atenuação acentuada devido a refração).

Um ponto interessante de ressaltar é o fato que o sensor olfativo não causa tanto impacto nos resultados como o sensor visual. Isso era esperado também, pois o fator de certeza da visão é maior que o do olfato (para os seres humanos, ver o evento causa mais impacto do que sentir o cheiro), características comprovadas pela árvore de decisão criada no modelo comportamental. Outro ponto que foi demonstrado pelas simulações é que a pseudo-comunicação serviu de maneira bastante eficaz para propagar o conhecimento do evento de perigo.

O mais importante a se salientar em resultados de simulações em situação de emergência é a grande dificuldade em se validar este tipo de sistema. Foram executadas muitas simulações e procurou-se utilizar o sistema auditivo de maneira mais realista

possível para que os resultados fossem o mais fidedignos com a realidade. Os resultados experimentais foram coerentes com o esperado, mas é muito difícil realizar um experimento real em condições similares para uma validação mais abrangente.

6.1 Trabalhos Futuros

Como vimos, modelos perceptivos virtuais são de grande valia tanto na área de entretenimento como na de simulações. Para trabalhos futuros, deixam-se abertas algumas possibilidades, tais como:

1. Incrementar os modelos visual e olfativo,
2. Modificar o sistema de comunicação entre os agentes,
3. Incrementar o modelo de decisão dos agentes,
4. Explorar o comportamento investigativo dos agentes.

Anexo A - Simulação de Agentes com Capacidade Auditiva em Situações de Emergência

Este artigo propõe um novo modelo de percepção auditiva para estimular o movimento de agentes virtuais em situações de emergência. No modelo proposto, uma fonte sonora (alarme) é colocada em algum local do ambiente virtual, e o decaimento da intensidade do som se dá em função da distância do agente à fonte. Quando a onda sonora atinge o agente virtual, também são levadas em consideração as diferentes sensibilidades do ouvido humano por faixas de frequência, assim como a perda natural de audição em função da faixa etária. Se o estímulo final no ouvido virtual for suficientemente forte, ou seja, pode ser ouvida, o agente reage, evacuando o ambiente conforme modelo descrito em um simulador de multidões baseado em Física.

Anexo B - A Perceptive Model for Virtual Agents in Crowds

This paper proposes a multi-perceptive model to stimulate the movement of virtual agents in danger events. The sound source (alarm), visual event and smell event localizations, as well as their attributes, such as sound frequency, event dimension and odor's propagation area, are defined in the virtual environment. In hearing perception, the sound intensity is attenuated according to the distance from the source and collisions with the walls located between the agent and the alarm. Also, the different sensibilities of the human ear are taken in consideration, as well as the natural hearing loss because of the aging. Besides hearing, the model aggregates other perceptions, such as vision and smell, treated by zonal and sensory approaches regarding the distance and the agent view angle to the perceptive area. To provide realism to the model, a pseudo-communication among agents was defined to propagate the individual knowledge acquired by their perceptions. Using this knowledge, the behavior model will define which action is to be taken next in the environment through a decision tree.

Anexo C - Um Modelo Comportamental Baseado em Árvores de Decisão para Agentes Virtuais em Situações de Emergência

Este artigo propõe um modelo comportamental de humanos virtuais baseado em árvores de decisão. O contexto de simulações é focado nas reações de humanos virtuais em situações de emergência. No modelo proposto, os agentes virtuais são dotados de três percepções virtuais: audição, visão e olfato. Além destes sentidos, os agentes possuem a habilidade de comunicar, onde suas percepções são propagadas para outros agentes na multidão. Os dados adquiridos pelos sensores virtuais são utilizados como entradas na árvore de decisão, permitindo que o agente possa coerentemente reagir como uma função de sua percepção.

Referências

- ALMEIDA, E. R.; CAMPOS, V. A. R.; GÂNDARA, M. E. R. Implicações médico legais da pair. *Caderno de Debates da Revista Brasileira de Otorrinolaringologia*, p. 56–69, 2004.
- BECKETT, H. E.; ARIBA, J. A. G. *Windows Performance, Design and Installation*. London: RIBA Publications Ltd., 1972.
- BEGAULT, D. R. *3-D sound for virtual reality and multimedia*. San Diego, CA, USA: Academic Press Professional, Inc., 1994. ISBN 0-12-084735-3.
- BRAUN, A.; BODMANN, B. E. J.; MUSSE, S. R. Simulating virtual crowds in emergency situations. In: SINGH, G. et al. (Ed.). *VRST*. [S.l.]: ACM, 2005. p. 244–252. ISBN 1-59593-098-1.
- CASTLEMAN, K. R. *Digital Image Processing*. [S.l.]: Prentice-Hall, 1996.
- CAVE, C. R. *Perceptual modeling for low-rate audio coding*. Dissertação (Mestrado) — McGill University, Montreal, Quebec, Canada, 2002.
- CHENNEY, S. Flow tiles. In: *Proceedings of Symposium on Computer Animation*. Grenoble, França: ACM and Eurographics Association, 2004. p. 233–242.
- CONDE, T.; THALMANN, D. An artificial life environment for autonomous virtual agents with multi-sensorial and multi-perceptive features. *Journal of Visualization and Computer Animation*, v. 15, n. 3-4, p. 311–318, 2004.
- CUNHA, M. A expressão corporal e o deficiente visual. *Revista Benjamin Constant*, Rio de Janeiro, ano 10 n28, 2004.
- FERRAZ, M. C. *Codificação de Imagens*. Dissertação (Mestrado) — Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro, 1998.
- FIGUEIRA, M. Assistência fisioterápica à criança portadora de cegueira congênita. *Revista Benjamin Constant*, Rio de Janeiro, ano 6, n17, 2000.
- GARDNER, W. 3d sound for virtual reality and multimedia. In: . [S.l.]: Kluwer Academic Publishers, 1997.
- GUYTON, A. *Fisiologia Humana. 5ª ed.* Rio de Janeiro: Editora Interamericana, 1981.
- GUYTON, A. *Tratado de fisiologia médica*. Rio de Janeiro: Editora Guanabara, 2002.

- HUBEL, D. H.; WIESEL, T. N. Brain mechanisms of vision. *J-SCI-AMER*, v. 241, n. 3, p. 130–139, 41–44, sep 1979. ISSN 0036-8733.
- JOHNSTON, J. D. Transform coding of audio signals using perceptual noise criteria. *IEEE Journal on Selected Areas in Communications*, v. 6, n. 2, p. 314–323, February 1988.
- KUTTRUFF, H. *Room Acoustics*. [S.l.]: Spon Press, 2000.
- LACERDA, A. P. *Audiologia Clínica*. Rio de Janeiro: Editora Guanabara Koogan, 1988.
- LARSSON, V.; KLEINE, M. Better presence and performance in virtual environments by improved binaural sound rendering. In: *Conf. on virtual, synthetic and entertainment audio*. Espoo, Finland,: Proceedings of the AES 22nd Intl., 2002. p. 31–38.
- LEITE, S. B. *Melhoria do codificador de fala G.722.1 através do uso de um modelo perceptual*. Dissertação (Mestrado) — Universidade Estadual de Campinas, Campinas-SP, 2003.
- LOKKI, T. et al. Creating interactive virtual auditory environments. *IEEE Computer Graphics and Applications*, v. 22, n. 4, p. 49–57, 2002.
- MACHADO, L. M. Paisagem valorizada: a serra do mar como espaço e como lugar. In: *Percepção ambiental: a experiência brasileira*. São Paulo: Studio Nobel, 1996. p. 97–119.
- MAFFRA, S. A. R. S. *Propagação de Som em Ambientes Acústicos Virtuais Bidimensionais*. Dissertação (Mestrado) — PUC University, Rio de Janeiro, 2003.
- MAMLOUK, A. M. et al. Quantifying olfactory perception: mapping olfactory perception space by using multidimensional scaling and self-organizing maps. *Neurocomputing*, v. 52-54, p. 591–597, 2003.
- MONZANI, J.-S.; THALMANN, D. A sound propagation model for interagents communication. In: HEUDIN, J.-C. (Ed.). *Virtual Worlds*. [S.l.]: Springer, 2000. (Lecture Notes in Computer Science, v. 1834), p. 135–146. ISBN 3-540-67707-0.
- MOORE, B. C. *An Introduction to the Psychology of Hearing, Fifth Edition*. Academic Press, 2003. Paperback. ISBN 0125056281. Disponível em: <<http://www.amazon.ca/exec/obidos/redirect?tag=citeulike04-20&path=ASIN/0125056281>>.
- MUSSE, S. R.; THALMANN, D. A model of human crowd behavior. In: *Computer Animation and Simulation '97*. Heidelberg, Germany: Springer-Verlag, 1997. p. 39–51.
- MUSSE, S. R.; THALMANN, D. Hierarchical model for real time simulation of virtual human crowds. *IEEE Trans. Vis. Comput. Graph.*, v. 7, n. 2, p. 152–164, 2001.
- NEPOMUCENO, L. A. *Elementos de Acústica Física e Psicoacústica*. Porto Alegre: Editora Edgard Blucher Ltda, 1994.
- NOSER, H. et al. Navigation for digital actors based on synthetic vision, memory, and learning. *Computers & Graphics*, v. 19, n. 1, p. 7–19, 1995.

- NOSER, H.; THALMANN, D. Sensor-based synthetic actors in a tennis game simulation. In: . [S.l.: s.n.], 1998. v. 14, n. 4, p. 193–205.
- PADULA, W.; SPUNGIN, S. A criança visualmente incapacitada, do nascimento até a idade pré-escolar: a importância da estimulação visual. *Revista Benjamin Constant*, Rio de Janeiro, ano 6, n16, 2000.
- PAINTER, T.; SPANIAS, A. Perceptual coding of digital audio. *Proceedings of the IEEE*, v. 88, n. 4, p. 451–515, 2000.
- PAN, X.; HAN, C. S.; LAW, K. H. A multi-agent based simulation framework for the study of human and social behavior in egress analysis. In: *ASCE International Conference*. Cancun, Mexico: Soibelman, L., Peña-Mora, F., 2005. p. 135–146.
- POHLMANN, K. C. *Principles of Digital Audio*. [S.l.]: McGraw-Hill Professional, 2000. ISBN 0071348190.
- PONTES, R. *Enxergar através do corpo*. Dissertação (Mestrado) — Universidade Cândido Mendes, Rio de Janeiro, 2006.
- REYNOLDS, C. W. Flocks, herds, and schools: A distributed behavioral model. In: *Proc. ACM SIGGRAPH '87*. Anaheim, California: [s.n.], 1987.
- RIO, V. D. Cidade da mente cidade real: percepção ambiental e revitalização na área portuária do rio de janeiro. In: *Percepção ambiental: a experiência brasileira*. São Paulo: Studio Nobel, 1996. p. 3–22.
- RUSSELL, S. J.; NORVIG, P. *Artificial Intelligence: A Modern Approach (2nd Edition)*. Prentice Hall, 2002. Hardcover. ISBN 0137903952. Disponível em: <<http://www.amazon.ca/exec/obidos/redirect?tag=citeulike04-20&path=ASIN/0137903952>>.
- SANTOS, T. M. M.; RUSSO, I. C. P. *A Prática da Audiologia Clínica*. São Paulo: Cortez Editora, 1986.
- SHAO, W.; TERZOPOULOS, D. Autonomous pedestrians. In: *SCA '05: Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation*. New York, NY, USA: ACM Press, 2005. p. 19–28. ISBN 1-7695-2270-X.
- TERZOPOULOS, D.; RABIE, T. F. Animat vision: Active vision in artificial animals. In: *ICCV*. [S.l.: s.n.], 1995. p. 801–808.
- TSINGOS, N.; GALLO, E.; DRETTAKIS, G. Perceptual audio rendering of complex virtual environments. In: . [S.l.: s.n.], 2004. v. 23, n. 3, p. 249–258.
- TU, X.; TERZOPOULOS, D. Artificial fishes: physics, locomotion, perception, behavior. In: *SIGGRAPH*. [S.l.]: ACM, 1994. p. 43–50. ISBN 0-89791-667-0.
- WITTEN, I. H.; FRANK, E. *Data Mining: Practical machine learning tools and techniques, 2nd Edition*. San Francisco, CA, USA: Morgan Kaufmann, 2005.
- ZELLNER, D.; BARTOLI, A.; ECKARD, R. Influence of color on odor identification and liking ratings. In: *American Journal of Psychology* **104**(4), 547-561. [S.l.: s.n.], 1991.

ZHANG, H.; WYVILL, B. Behavioural simulation in voxel space. In: *CA*. [S.l.]: IEEE Computer Society, 1997. p. 119–126. ISBN 0-8186-7984-0.