

3

Percepção Visual

Wann e Mon-Williams (1996) definem ambiente tridimensional como um ambiente que se beneficia dos aspectos naturais da percepção humana ao estender a informação visual para três dimensões espaciais, complementar com outros estímulos e mudanças temporais e ainda permitir que o usuário interaja com a informação exibida.

Desta forma, conhecer a percepção visual humana nos ajuda a entender fenômenos relacionados aos ambientes virtuais 3D como animação, perspectiva, iluminação e nível de detalhe.

Este capítulo irá estudar a percepção visual que faz parte dos sentidos humanos e consiste na habilidade de detectar e interpretar a luz que atinge os objetos.

3.1.

Introdução à Visão

Estamos tão acostumados ao nosso sistema visual que raramente nos damos conta da complexidade envolvida em simples tarefas como abrir os olhos ao acordar ou mover nosso olhar para ver objetos ao nosso redor. Apesar de vermos coisas todos os dias, a visão é na realidade um fenômeno bastante complicado e o olho é provavelmente um dos órgãos mais complexos do corpo humano.

Quando criança, nós usamos os olhos para interagir com o mundo a nossa volta: observamos objetos coloridos que chamam nossa atenção, reconhecemos rostos familiares e acompanhamos o movimento das pessoas. Ao ver dois objetos em locais diferentes, as crianças sabem discriminar entre eles e alcançar com a mão aquele que está mais perto. Este tipo de comportamento indica que existe

mais processamento de informação acontecendo do que simplesmente uma resposta a estímulos apresentados.

Neste momento é interessante fazer uma distinção entre ver e perceber. Ver é um processo passivo de obter informação sensorial visual, enquanto que perceber possui um componente de reconhecimento e compreensão. Muitas vezes podemos ver, porém não perceber.

O olho humano, nossa ferramenta primária de visão, é também um instrumento biológico de captura que assim como uma câmera fotográfica coleta, foca e processa luz. Ele possui lentes para focalizar e concentrar luz, um obturador que controla a quantidade de luz entrando, uma área sensível a luz localizada no fundo e que funciona como um filme e finalmente um cérebro que processa e revela a imagem. A analogia com uma câmera fotográfica é bastante simplista considerando o sofisticado sistema visual que os olhos representam, portanto para melhor entender este instrumento, faremos uma rápida revisão de sua anatomia e fisiologia.

3.2. Sistema Visual

O sistema visual pode ser dividido em três locais principais de processamento conforme a ilustração a seguir. Primeiramente destacamos os olhos, que agem como a entrada para nosso sistema visual, em seguida os caminhos visuais que transportam e organizam a informação visual e o córtex visual que é a parte do cérebro que gera a percepção (figura 1).

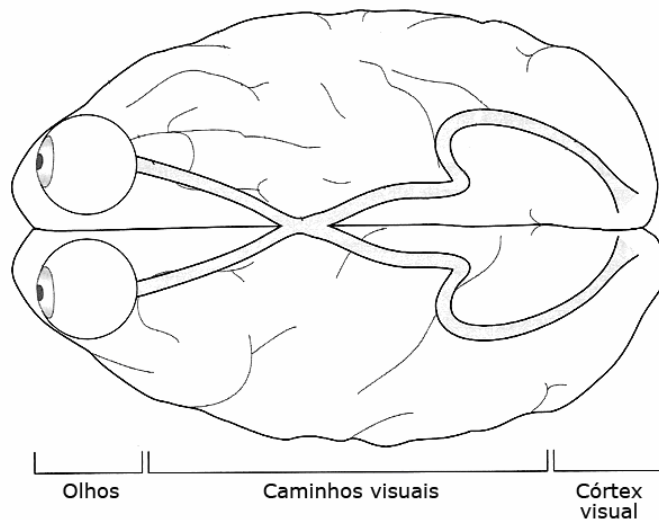


Figura 1 – Ilustração do cérebro demonstrando o sistema visual humano (LUEBKE, 2003)

Para fins de nossa análise da percepção visual, os caminhos visuais têm pouca contribuição, portanto vamos focar nossa descrição nos olhos e no córtex visual.

3.2.1. Olhos

Grandjean (1973) compara o olho humano a uma câmera fotográfica. A córnea, as lentes e a íris constituem a parte ótica e a retina funciona como um filme sensível à luz. A córnea e as lentes formam um mecanismo de refração que focaliza a luz entrante para atingir a retina. A coróide, um novelo vascular intermediário, serve para absorver a luz dispersa de forma similar ao interior escuro de uma câmera fotográfica.

A retina é a camada mais interna e pode ser considerada como sua estrutura fundamental. Ela é composta por um fino tecido composto por células receptoras fotossensíveis que detectam a entrada de luz, e por uma rede de neurônios que processam os impulsos luminosos. A saída destes neurônios forma o nervo ótico, que transmite as informações óticas para o cérebro, conforme ilustrado na figura 2.

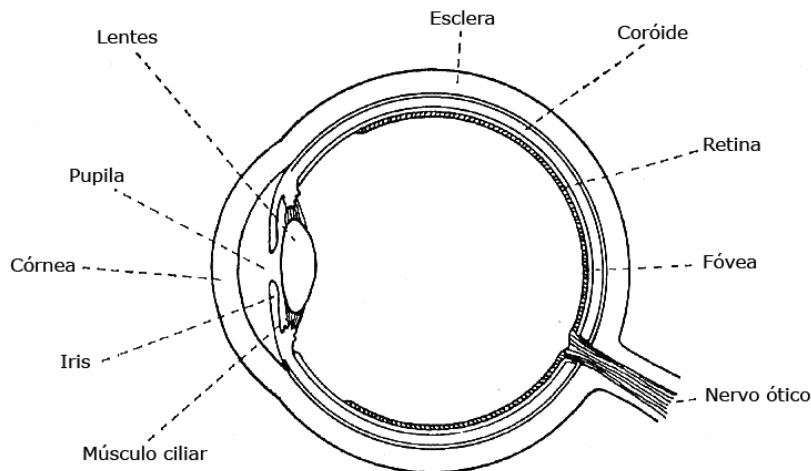


Figura 2 – Diagrama do olho em corte longitudinal (GRANDJEAN, 1973)

3.2.2. Células receptoras fotossensíveis

As células receptoras fotossensíveis são responsáveis por traduzir a imagem recebida em uma representação mais concisa e gerenciável. Existem dois tipos de células, os bastonetes e cones. Os bastonetes provêm percepção de contraste, atuam em baixa iluminação e fazem detecção de movimento. Os cones funcionam apenas quando há luz suficientemente disponível e são responsáveis pela visão detalhada e pela percepção apurada das cores.

Cones e bastonetes não são uniformemente distribuídos, de acordo com Grandjean (1973) existem no olho aproximadamente 130 milhões de bastonetes e 7 milhões de cones. Os cones ficam mais concentrados centralmente em uma região conhecida como fóvea e diminuem conforme aumenta a excentricidade retinal que é distância angular medida a partir do centro da retina. Os bastonetes por outro lado, quase não existem na região da fóvea e tem sua máxima concentração a cerca de 20 graus de excentricidade. Ambos os cones e bastonetes atingem sua menor concentração a aproximadamente 75 a 80 graus de excentricidade conforme indica Ferwerda (2001).

Estes dois tipos de células fotossensíveis trabalham juntos, os cones proporcionando a visão central detalhada e colorida e os bastonetes provendo

contraste, detecção de movimento e visão com baixa luminosidade. Quando o nível de iluminação cai, os bastonetes assumem e dominam o processamento visual de forma gradual e imperceptível.

Gallardo (2003) ilustra este processo descrevendo o que ocorre quando entramos em uma sala escura de cinema. Nos primeiros momentos podemos apenas discernir as áreas claras, tais como os sinais vermelhos de saída. Aos poucos nossos olhos se ajustam para que possamos enxergar os assentos e as pessoas na sala de cinema. Isto acontece porque logo que entramos na sala, são os cones que estão primariamente funcionando, porém aos poucos os bastonetes começam a assumir e se adaptar a escuridão.

3.2.3. Variações através da retina

Em função desta fisiologia do olho humano, a sensibilidade ao tamanho de estímulos não é uniforme através de toda a retina. Verificamos que a fóvea, onde existe a maior concentração de cones, tem a habilidade de tratar os menores detalhes. Entretanto, esta habilidade reduz de acordo com a excentricidade retinal, de tal forma que no campo de visão periférica, a retina tem muito pouca capacidade de discriminar pequenos detalhes.

Conforme indica Grandjean (1973), na prática esta formação fisiológica significa que sempre que queremos focar nossa atenção em um determinado objeto, devemos realizar uma combinação de movimentos dos olhos e da cabeça de forma que a luz projetada por este objeto seja projetada da região da fóvea em nossa retina. Isto garante que poderemos ver o objeto com o maior nível de detalhes possível.

A região periférica da retina, apesar de menos sensível a detalhes visuais, tem maior sensibilidade para movimentos. Desta forma a visão humana percebe menos detalhes em objetos se movendo se comparado a objetos estáticos. O resultado é a sensação de objetos borrados à medida que estes se movem e passam por nosso ponto fixo de visão.

3.2.4.

Fatores que influenciam a percepção visual

A partir de Preece (1994) e Grandjean (1973) podemos considerar que existem fatores ambientais e individuais relevantes no resultado final de nossa percepção visual. A seguir destacamos os mesmos:

Fatores ambientais

- **Iluminação de fundo:** A intensidade da luz de fundo usada para iluminar um estímulo pode afetar substancialmente sua visibilidade.
- **Adaptação à intensidade de luz:** O olho humano é sensível a intensidade de luz devido ao nível de fotoreceptores na retina e também a dilatação da pupila. Este é o efeito que ocorre, por exemplo, quando entramos em uma sala escura de cinema.
- **Dispositivo de exibição da imagem:** dado que vemos imagens geradas por computador através de monitores de vídeo, as configurações de brilho, contraste, cor e gama irão afetar a percepção de qualquer estímulo visual. Claramente, nosso sistema visual mais sofisticado, fica limitado pela fidelidade do dispositivo de exibição sendo utilizado.
- **Interação com outros sentidos:** Estudos recentes (SHAMS, 2000; ASTHEIMER, 1993) demonstram que existe uma complexa interação entre os sentidos, mais especificamente descrevendo impactos do som na percepção visual.

Fatores individuais

- **Idade:** A sensibilidade a contraste varia como uma função da idade, degradando conforme seu aumento.

- **Percepção de cor:** as cores não são percebidas igualmente por todas as pessoas. Algumas pessoas possuem daltonismo, fenômeno mais comum entre os homens e representado pela deficiência de visualizar as cores vermelho e verde.
- **Visão estereoscópica:** Uma grande parte da população, estimado em uma pessoa a cada dez, não consegue perceber a profundidade estereoscópica, ou seja, não consegue perceber profundidade como resultada da disparidade entre as imagens formadas por cada olho.
- **Formações da lente:** Uma visão incorreta pode resultar da falta de capacidade do olho de ajustar corretamente a distância focal de sua lente para projetar o objeto fixado exatamente na retina. A miopia é um exemplo de formação da lente que leva a este fenômeno.
- **Estado emocional:** O estado emocional do observador afeta a dilatação de sua pupila, consequentemente afetando a quantidade de luz que pode atingir a retina.
- **Experiência:** Nossa percepção de objetos pode ser influenciada por conhecimento prévio e experiências anteriores. Estudos especulam, por exemplo, que laranjas são percebidas, em telas de computador, com cores mais ricas e naturais após serem percebidas e identificadas como tal pelo observador.

3.3. Nível de Detalhe

Objetos ou ambientes representados tridimensionalmente e virtualmente na tela de um computador podem assumir diferentes níveis de detalhe. Seja um personagem em um jogo ou uma maquete eletrônica de arquitetura, todos são apresentados visualmente para o usuário e, portanto faz sentido analisarmos estes à luz do sistema visual humano.

Segundo Gallardo (2003), baseado no comportamento e nas limitações do nosso sistema de visão, os critérios mais comuns para estabelecer o nível de detalhe de um objeto são:

- **Distância:** menor nível de detalhe para objetos mais afastados do ponto de visão;
- **Tamanho:** quanto menor o tamanho projetado na tela do computador, menor o nível de detalhe;
- **Excentricidade:** menor nível de detalhe para objetos localizados na visão periférica;
- **Velocidade:** objetos que se movem rapidamente pelo campo de visão podem ter seu nível de detalhe diminuído em relação aos que se movem mais lentamente;
- **Profundidade de campo:** menor nível de detalhe para os objetos que estão visualmente embaçados em função do foco da visão estar em uma profundidade diferente;

Vários pesquisadores na área de computação gráfica têm feito propostas de sistemas que se aproveitam da noção de que nós percebemos menos detalhes em objetos na visão periférica ou em rápido movimento. Destacamos estudos de Clark (apud LUEBKE, 2003) que sugeriu simplificar o nível de detalhe de objetos quanto mais próximos da periferia do campo de visão e quanto maior suas velocidades e Blake (apud LUEBKE, 2003) que desenvolveu métricas para prever o nível de detalhe apropriado em circunstâncias como os movimentos relativos dos objetos em relação ao observador.

3.3.1. Percepção de Movimento

Nosso campo visual pode ser pensado como um sistema de processamento visual dirigido por importância, ou seja, primeiro ele detecta um estímulo e depois responde ao mesmo. O contraponto deste funcionamento é que quando o estímulo torna-se estático, a habilidade de percepção do sistema se deteriora rapidamente,

pois as células receptoras fotossensíveis ficam acostumadas ao mesmo sinal e param de responder.

Existem basicamente dois tipos de detectores independentes e coordenados de movimento no sistema visual: o sistema visual na retina e o sistema *eye-head*. O primeiro não requer movimento dos olhos para responder e seguir uma imagem através do campo visual, ele responde a mudanças no contraste e na iluminação. Através dele percebemos que um objeto se move, porém não percebemos que ele se move de forma independente em relação ao espaço.

Os olhos têm uma melhor percepção de movimento quando existem objetos reconhecíveis no fundo, ao invés de apenas uma cor ou fundo escuro. Em outras palavras, nós somos mais capazes de perceber o movimento relativo do que o movimento absoluto. Esta preferência pode causar ilusões visuais, como por exemplo, quando um objeto maior se move contra um objeto menor parado e nós percebemos que o movimento ocorre com o menor.

3.4. Estímulos de Profundidade

Idealmente parece ser desejável apresentar na tela de um computador informações visuais que possuam características similares aos objetos que percebemos ao nosso redor. Desta forma o sistema visual poderia usar o mesmo processo aplicado na percepção dos objetos no mundo real.

Preece (1994) afirma que um ambiente tridimensional gerado por computador precisa necessariamente ser representado através de uma tela que é uma superfície bidimensional. Portanto para fazer com que objetos sejam percebidos como tridimensionais precisamos utilizar o que chamamos de estímulos de profundidade monoculares.

Dalgarno (2003) descreve os estímulos monoculares como aqueles disponíveis a partir da visão de apenas um único olho e sugere alguns tipos de estímulos de profundidade que iremos resumir a seguir e ilustramos na figura 3.

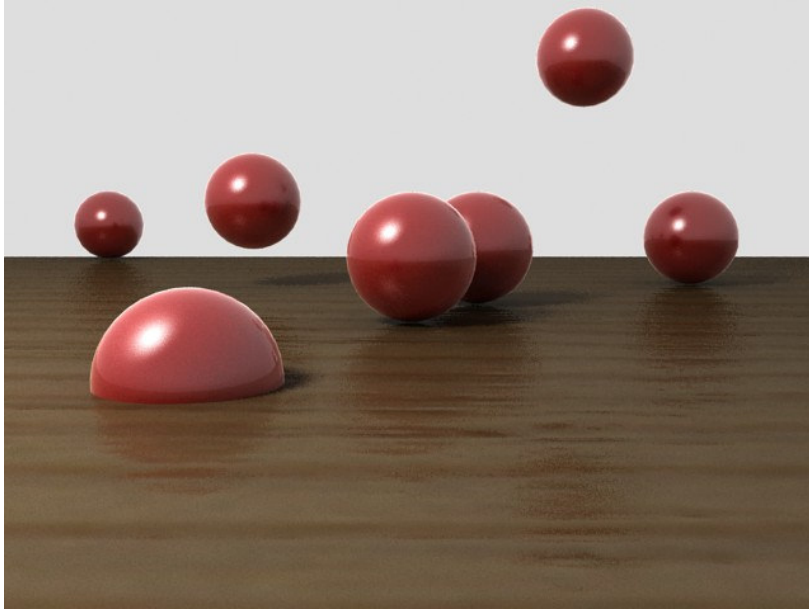


Figura 3 – Imagem gerada através do programa 3ds Max para ilustrar estímulos de profundidade

- Perspectiva é convergência de linhas paralelas ao se afastar do observador;
- Tamanho relativo: um objeto pequeno ao lado de um objeto idêntico grande parece estar mais distante do observador;
- Gradiente de textura: também chamado efeito de impressionismo, é a percepção de profundidade baseada na mistura gradual dos detalhes repetidos em uma textura que avança no para o horizonte de visão;
- Interposição ou oclusão é a percepção de profundidade quando objetos se sobrepõem;
- Perspectiva aérea é a percepção de profundidade devida a dispersão da luz causada pela atmosfera e que torna os objetos mais distantes coloridos de acordo com o fundo;
- Altura relativa é a percepção de que objetos acima do horizonte estão mais longe e mais distantes do que os objetos abaixo;
- Sombras nos ajudam a entender o tamanho, forma, textura e densidade dos objetos. A percepção acontece em função da posição da luz e orientação da sombra gerada;
- Paralaxe de movimento é a mudança na oclusão de objetos quando o observador se move na mesma distância;

- Perspectiva de movimento é a mudança no tamanho dos objetos quando o observador se aproxima ou se afasta;

Modelos tridimensionais também podem ser construídos utilizando imagens estereoscópicas ou estímulos de profundidade binoculares, nos quais duas imagens do mesmo objeto em ângulos ligeiramente diferentes são apresentadas separadamente para cada olho. Este processo é equivalente ao que ocorre no mundo físico, pois devido à posição que nossos olhos estão, cada um recebe uma imagem ligeiramente diferente. Cabe ao sistema visual no cérebro combinar estas imagens para formar a representação tridimensional que vemos.

3.5.

Navegação e orientação tridimensional

Alem do aspecto visual, a interação tridimensional também envolve fundamentalmente duas habilidades dos usuários: navegação (chegar de um ponto a outro) e orientação. Parush e Berman (2004) descrevem em seu artigo o conceito de mapa cognitivo, uma representação mental do ambiente incluindo noções de distância e direção.

Este mapa cognitivo pode ser decomposto em três entidades de conhecimento ou representação: marcos, rotas e mapas. Os marcos são representações visuais de objetos inseridos no ambiente tridimensional. As rotas são descrições procedimentais e sequenciais do caminho entre dois pontos no ambiente, trazendo identificação dos locais e ações que devem ser tomadas. Os mapas são representações gráficas da área geográfica envolvida, incluindo todos os elementos que a compõem e suas relações espaciais.

Parush e Berman (2004) destacam a importância de mapas e marcos como auxílios de navegação e orientação tridimensional, que apesar de terem um impacto inicial no aprendizado da interação, tornam-se muito eficientes posteriormente no uso para execução de tarefas.

3.6. Conclusões Parciais

Conhecer a fisiologia do sistema visual nos permite examinar mais detalhadamente suas implicações na ergonomia de ambientes com interação tridimensional envolvendo nível de detalhe, movimento e estímulos de profundidade.

Utilizando os conhecimentos da visão humana, pesquisadores desenvolveram complexos modelos de nível de detalhe que são utilizados por programas geradores de imagens tridimensionais realistas, conhecidos como *renderizadores*. Estes modelos utilizam os fatores excentricidade retinal, velocidade e tamanho para calcular o nível de detalhe necessário para um objeto ou figura.

Um exemplo simples e intuitivo de aplicação destes conceitos pode ser observado através do *Percept* (REDDY, 2006), um programa de computador que simula as limitações da percepção visual humana borrando áreas de uma imagem a partir de um dado campo de visão e velocidade de movimento do observador. Na figura 4 exibimos duas imagens, a original na esquerda e na direita a imagem com áreas borradas justamente onde nossa visão percebe menor nível de detalhe. A simulação do *Percept* considerou um campo de visão de 120 por 135 graus e uma velocidade constante de 50 graus por segundo.



Figura 4 – O programa *Percept* aplicado a uma imagem da pintura de Rembrandt, Retorno do Filho Pródigo

Jogos eletrônicos também utilizam os conceitos de nível de detalhe para solucionar o desafio de exibir ambientes tridimensionais realistas sem exceder o poder de processamento dos computadores atuais. O *Unreal Engine* é uma plataforma para desenvolvimento de jogos tridimensionais da empresa americana *Epic Games*, e provê algoritmos adaptativos para eliminação de polígonos e simplificação de superfícies. Esta funcionalidade torna objetos orgânicos no jogo, como pessoas e animais, extremamente realistas e rápidos em tempo de resposta para o jogador. O *Unreal Engine* é utilizado para jogos em computadores pessoais e consoles *Xbox* e *PlayStation 2*, e pode ser visto em ação nos jogos *Unreal Tournament 2004* (figura 5), *Splinter Cell*, *Rainbow Six* e *Lineage II*.



Figura 5 – Cena do jogo Unreal Tournament (EPIC GAMES)

Além das implicações em termos de nível de detalhe, as características do olho humano também devem ser observadas para aumentar a sensação de realismo e imersão do usuário. A iluminação em cenas geradas artificialmente por computador tem efeito importante por afetarem o nível de percepção, mais especificamente as áreas mais claras atraem mais a atenção em comparação as mais escuras. Somente após observar as áreas claras e coloridas, nosso foco irá passar a examinar as partes mais escuras.

Neste aspecto, os estímulos visuais de profundidade também desempenham papel fundamental na percepção da imagem apresentada, ou seja, os ambientes tridimensionais devem combinar a utilização de diversos estímulos monoculares como sombra, perspectiva, atmosfera e oclusão para auxiliar na percepção e consequentemente no realismo da interação. A figura 6 demonstra uma aplicação onde o usuário pode navegar livremente por um quarto e observar os objetos por diversos ângulos e distâncias. Percebemos a utilização de diversos estímulos visuais que auxiliam na experiência de imersão.



Figura 6 – Imagem de uma aplicação de interação tridimensional para arquitetura (QUEST3D)

Complementando os aspectos visuais e de nível de detalhe, a navegação e a orientação em ambientes tridimensionais é facilitada pela inserção de marcos ou mapas que auxiliam na formação do mapa cognitivo pelo usuário. O jogo *Spyro* da empresa americana *Insomniac Games* enfrentou o desafio de lidar com enormes ambientes tridimensionais de livre navegação que poderiam facilmente deixar o usuário perdido e desorientado. John Fiorito (2006) da *Insomniac* descreve que uma das soluções foi utilizar marcos sempre que possível. Marcos são estruturas ou características geográficas que distinguem diferentes áreas e oferecem um ponto de referência para o usuário. No jogo, estruturas arquitetônicas como castelos, fortes ou outras construções de impacto visual mostraram ser os marcos mais eficientes (figura 7).



Figura 7 – Imagem de um marco utilizado no jogo *Spyro* (FIORITO, 2006)

Em face aos exemplos descritos, verificamos facilmente como a percepção visual humana tem grande influência na interação tridimensional. Conhecer as implicações visuais da fisiologia do olho humano, entender os estímulos de profundidade e o efeito de marcos e mapas na orientação e navegação, provê caminhos para o projetista de aplicações, como jogos ou visualizações arquitetônicas, tornar a experiência do usuário mais satisfatória, eficiente e adequada.

3.7.

Referências Bibliográficas do Capítulo 3

ASTHEIMER, P. **What you see is what you hear-Acoustics applied in virtual world**. S, Fraunhofer-Inst. for Comput. Graphics, Darmstadt, 1993.

DALGARNO, B.; HEDBERG J.; HARPER, B. **The Contribution of 3D environments to conceptual understanding**. Proceedings of the 19th Annual Conference of the Australian Society for Computers in Tertiary Education., 2003.

FERWERDA, J.A. **Elements of Early Vision for Computer Graphics**. IEEE Computer Graphics and Applications, 2001.

FIORITO, J. **Spyro Production Design**. Gamasutra. Disponível em <http://www.gamasutra.com/features/20000502/spyro_03.htm>. Acesso em: 22 jun. 2006.

GALLARDO, A. **3D Lightning: History, Concepts and Techniques**. Rockland, MA: Charles River Media, 2003.

GRANDJEAN, E. **Ergonomics of the Home**. Taylor & Francis, 1973.

LUEBKE, D. **Level of Detail for 3D Graphics**, San Francisco: Morgan Kaufmann, 2003.

PARUSH, A.; BERMAN, D. **Navigation and orientation in 3D user interfaces: the impact of navigation aids and landmarks**. Int. Journal of Human-Computer Studies, 2004.

PREECE, J. **Human-Computer Interaction**. Harlow, England: Addison-Wesley, 1994.

QUEST3D Widescreen room demo. **Quest3D.** Disponível em <<http://www.quest3d.com/>>. Acesso em: 22 jun. 2006.

REDDY, M. **Percept Website.** 1997. Disponível em <<http://www.martinreddy.net/percept>>. Acesso em: 21 jun. 2006.

SHAMS, L.; KAMITANI, Y.; SHIMOJO, S. **What you see is what you hear.** London: Nature Publishing, 2000.

UNREAL Technology. **Epic Games.** Disponível em: <<http://www.unrealtechnology.com>> . Acesso em: 22 jun. 2006.

WANN, J.; MON-WILLIAMS, M. **What does virtual reality NEED?:** human factors issues in the design of three-dimensional computer environments. International Journal of Human-Computers Studies, 1996.