

Visão Estereoscópica, Realidade Virtual, Realidade Aumentada e Colaboração

Alberto B. Raposo, Flávio Szenberg, Marcelo Gattass, Waldemar Celes

Tecgraf—Grupo de Tecnologias em Computação Gráfica
Departamento de Informática, PUC-Rio
R. Marquês de São Vicente, 225
Gávea, Rio de Janeiro – RJ, 22453-900

{abraposo, szenberg, mgattass, celes}@tecgraf.puc-rio.br

Abstract

This course is about new challenges in computer graphics and aims at making an overview of technologies related to the fields of virtual reality, augmented reality and collaborative graphic applications. Besides the presentation of these technologies, the course focuses on some applications developed at Tecgraf/PUC-Rio. The topic on generating stereographic images is further detailed, since it is one of the most important resources for virtual reality. The goal is to show that it is possible to develop applications in these areas with reasonable performance and realism without the need of very sophisticated computational resources.

Resumo

Este curso se insere no contexto de novos desafios em computação gráfica e tem como objetivo fazer um apanhado geral das tecnologias ligadas às áreas de realidade virtual, realidade aumentada e aplicações gráficas colaborativas. Além da apresentação e comparação de tecnologias, o curso enfocará algumas aplicações desenvolvidas pela equipe do Tecgraf/PUC-Rio. Haverá um aprofundamento maior no tópico de geração de imagens estereoscópicas, um dos mais importantes recursos para a realidade virtual. O objetivo é mostrar que possível desenvolver aplicações nessas áreas com desempenho satisfatório e bom grau de realismo sem o uso de recursos computacionais muito sofisticados.

7.1 Introdução

Nos últimos anos, a utilização da computação gráfica para visualização científica e modelagem geométrica tridimensional tem trazido mudanças fundamentais aos métodos de trabalho, tanto na indústria como em pesquisa e desenvolvimento. A possibilidade de visualização e manipulação interativa de modelos virtuais com auxílio do computador tem revolucionado muitas atividades, pois permite a compreensão e análise de enormes quantidades de informação de natureza espacial com eficiência sem precedentes, explorando a grande capacidade humana de raciocinar e se comunicar visualmente. Com a disponibilidade crescente de redes de comunicação de dados de alta velocidade, a interligação destes ambientes de visualização é o próximo passo na evolução do uso desta tecnologia.

É nesse contexto de novos desafios que se insere este curso, que analisa aspectos do espectro de Realidade Virtual (do inglês *Virtual Reality Continuum*). Este espectro passa pela Realidade Virtual (RV), Virtualidade Aumentada e Realidade Aumentada (RA) [Milgram e Kishino 1994]. Como se observa na Figura 7.1, o espectro de RV começa no mundo real e termina na RV. Entre os extremos, existe a Realidade Misturada (do inglês *Mixed Reality*), onde elementos do mundo real se misturam com elementos do mundo virtual. A Realidade Aumentada tem como base o mundo real, colocando sobre ele informações ou objetos virtuais. A Virtualidade Aumentada, por sua vez, já está inserida em ambiente virtual, no qual é colocada alguma visão do mundo real (por exemplo, uma imagem de videoconferência).

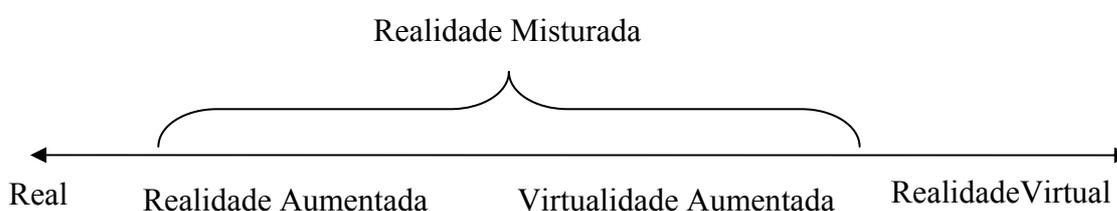


Figura 7.1 – Espectro de Realidade Virtual.

Dentro do espectro apresentado, este curso abordará a RV e a RA, que são as duas tecnologias onde os esforços de pesquisa e desenvolvimento estão mais concentrados atualmente. Também serão abordados tópicos relacionados à colaboração em ambientes virtuais.

Não existe um consenso sobre a definição de realidade virtual. Alguns autores tentam listar elementos-chave de RV, tais como imersão (sensação de estar em um ambiente), interatividade e *feedback* sensorial (visual, auditivo e tátil) [Sherman e Craig 2003]. Outros autores descrevem RV de uma maneira geral, como uma técnica avançada de interface, onde o usuário pode navegar e interagir em um ambiente sintético tridimensional gerado por computador, estando completa ou parcialmente presente ou imerso pela sensação gerada por canais multi-sensoriais, sendo o principal a visão [Krueger 1991], [Burdea and Coiffet 1994].

O objetivo principal do estímulo visual, tanto em RV quanto em RA, é fornecer ao usuário a sensação de tridimensionalidade. Para atingir esse objetivo, o sistema deve gerar, ao mesmo tempo, duas imagens diferentes, correspondendo às visões de cada um dos olhos. Esse tipo de geração de imagem é chamado de visão estereoscópica ou estereoscopia, e será detalhado a seguir.

7.2 Estereoscopia

Nesta seção é apresentada a estereoscopia visual. Devido à evolução das tecnologias de áudio, cada vez mais presente entre nós, seja via rádio de automóvel ou CD, estamos muito acostumados em relacionar estéreo ao som. Canais de áudio ligados a caixas de som independentes fazem chegar aos nossos ouvidos sons ligeiramente diferentes, e nos casos de fone de ouvido, exclusivamente diferentes. Estes sons são processados por nós no cérebro, dando a sensação de imersão ao ambiente onde originou os sons.

Esta tecnologia de estéreo de som pode ser estendida ao visual [Lipton 1982]. Os animais, em específico os seres humanos, são dotados de dois olhos que enxergam o mundo diferente, pois estão separados por uma distância. Isto por ser observado com um simples experimento: alinhe o polegar da mão esquerda com uma bandeirinha e seu nariz, e foque sua visão para o dedo. Você verá a bandeirinha como sendo duas, uma para cada olho (feche um olho e abra o outro e em seguida inverta), conforme Figura 7.2 [StereoGraphics 1997].



Figura 7.2 – Os olhos estão convergindo para o polegar e a bandeirinha é vista como dupla imagem.

Agora convergindo a visão para a bandeirinha, a visão que você terá com os dois olhos abertos é mostrada na Figura 7.3 – o polegar agora é visto como sendo dois.



Figura 7.3 – Os olhos agora convergindo para a bandeirinha e o polegar é visto como dupla imagem.

Estas diferenças entre imagens geradas pelo olho direito e pelo olho esquerdo são processadas pelo cérebro dando a nós uma noção de profundidade e com isto tem-se a idéia de imersão em um ambiente com objetos com diferentes distâncias.

Em computação gráfica, normalmente são geradas imagens mono (isto é, a partir de uma única câmera virtual) e estas imagens são visualizadas por um monitor ou projetadas em um plano. A estereoscopia visual, ao contrário, é feita a partir da geração de duas imagens a partir das localizações das câmeras virtuais separadas de uma determinada distância como será visto nas seções a seguir. Serão vistos nas próximas seções também os parâmetros e suas nomenclaturas e como eles afetam o resultado final que desejamos – imersão em um ambiente virtual.

Os equipamentos de visualização também afetam, e muito, na qualidade da visão estéreo. Serão abordados equipamentos para tal finalidade e seus funcionamentos. Muitos equipamentos necessitam formatos especiais de imagens para cada olho. Os formatos mais usuais também são listados, lembrando que não há um padrão universalmente aceito.

7.2.1. Noções Básicas

Quando tiramos uma foto através de uma câmera fotográfica usual, como por exemplo, a Figura 7.4, existem alguns efeitos que nos fornecem características tri-dimensionais da cena. Estes efeitos são chamados efeitos passivos. São efeitos inerentes à aparência do mundo externo e independem dos nossos olhos, isto é, giramos o papel fotográfico e a imagem continua estática.



Figura 7.4 – Foto tirada por uma máquina fotográfica usual.

Alguns efeitos passivos que podemos observar na foto acima são:

- Perspectiva: esse efeito pode ser notado pelo fato de existirem edifícios no “fundo” da imagem maiores que o prédio “mais à frente”, com tamanhos projetados menores. Vale ressaltar aqui que, “fundo” e “mais à frente” em uma foto são características impróprias, pois ao máximo que possa parecer ser 3D, foto é uma imagem planar. Na perspectiva, objetos de mesmo tamanho devem aparecer maiores na foto quando mais estiverem próximos da câmera. Na Figura 7.5, os dois paralelepípedos têm mesmo tamanho, mas pelo efeito de perspectiva gerado a partir de linhas existentes (definindo um ponto de fuga), concluímos que o paralelepípedo que aparece mais ao centro da imagem tem tamanho real maior.

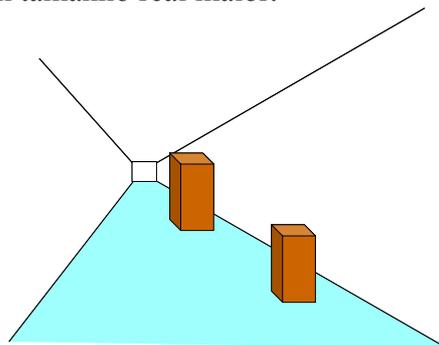


Figura 7.5 – Efeito de perspectiva.

- Iluminação: o efeito de iluminação aplicado à cena adiciona à imagem uma descrição melhor da forma dos objetos presentes, contribuindo com o aumento do grau de realismo. Este efeito é exemplificado na Figura 7.6. Note que na figura (a), o que parece é que existem um círculo e um hexágono, preenchidos com determinadas cores de forma uniforme. Já na figura (b), podem-se observar as verdadeiras formas dos objetos tridimensionais, esfera e cubo.



Figura 7.6 – Efeito de iluminação.

- Oclusão: a oclusão é responsável em esconder partes de objetos, ou por inteiro, que estejam por de trás de outros. Com isto, observa-se a ordem dos objetos em termos de distância à câmera. Este efeito é ilustrado na Figura 7.7. Na imagem da esquerda a esfera está por trás do cubo na cena e na imagem da direita o cubo é que esta por trás.

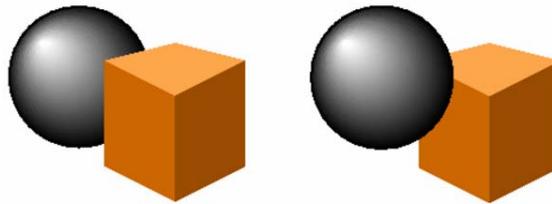


Figura 7.7 – Efeito de oclusão.

- Sombra: quando podemos afirmar que um objeto está colado a um plano de apoio? Esta pergunta, apesar de simples, é de difícil resposta em uma imagem resultante de uma projeção 3D para 2D. Isto porque o inverso dessa projeção, dado um ponto na imagem, é uma reta no espaço 3D, isto é, diversos pontos na cena 3D alinhados com a câmera projetam-se em um mesmo ponto. Isto pode levar à conclusão que um objeto pode estar fixo a um plano ou então mais próximo ao observador e solto no ar. A sombra auxilia na resposta de uma pergunta parecida com a proposta: quando podemos afirmar que um objeto não está colado ao plano? Uma sombra de um objeto está sempre no plano de apoio e se o objeto está a uma distância na imagem da sombra, podemos concluir que o objeto não está no plano de apoio, como podemos ver na Figura 7.8 com a esfera. Na mesma figura abaixo, parece que o cubo está fixado no plano de apoio, mas esta afirmação não pode ser comprovada baseada apenas em uma imagem projetada.

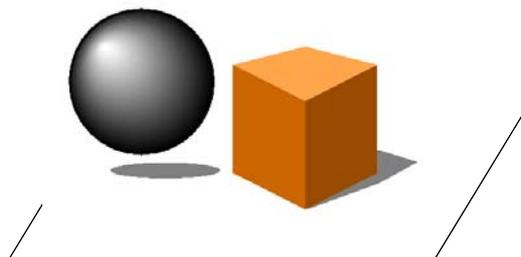


Figura 7.8 – Efeito de sombra.

- Gradiente da textura: segundo especialistas em obras de arte, este é o único efeito passivo proposto no século XX. Este efeito é ilustrado na Figura 7.9. Pode-se observar nesta figura, que os quadrados brancos e pretos indicam um decrescimento nos tamanhos indicando pontos onde se projetam algumas direções, definidos como ponto de fuga. Com isto, temos uma noção de profundidade.

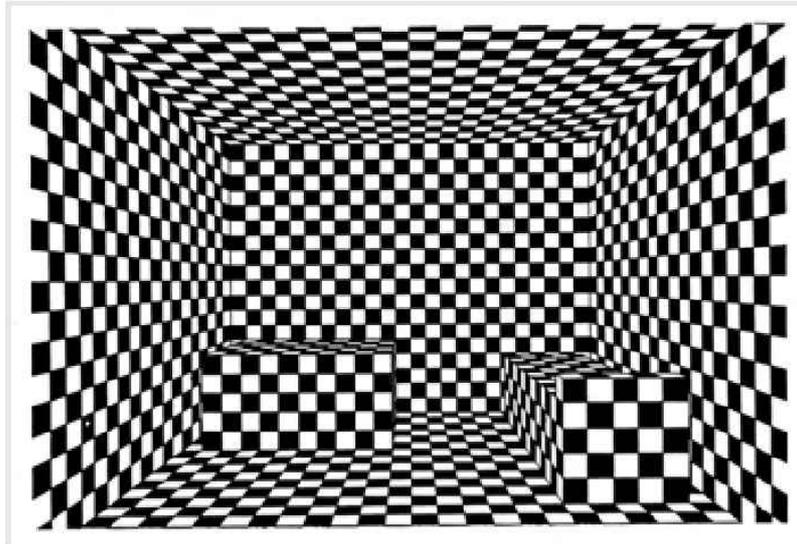


Figura 7.9 – Efeito do gradiente de textura.

Todos estes efeitos também podem ser visto na Figura 7.10, com alguns sendo indicados.



Figura 7.10 – Exemplo de alguns efeitos passivos.

A estereoscopia visual é classificada como efeito ativo, pois não está presente diretamente em uma imagem. O efeito estereoscópico é produzido através de um conjunto elementos, como será visto a seguir.

Podemos dizer que “um visor estereoscópico é um sistema óptico cujo componente final é o cérebro humano” [StereoGraphics 1997]. Isto porque, como foi dito na seção anterior, é o cérebro que faz a fusão das duas imagens (uma para cada olho) resultando em noções de profundidade, como é ilustrado na Figura 7.11. São capturadas duas imagens

ligeiramente diferentes relativas à mesma cena, da mesma maneira que vemos o mundo real. Não entraremos em detalhes no funcionamento da captação de imagens no sistema ocular humano nem tão pouco no funcionamento do cérebro no que diz respeito à visão.

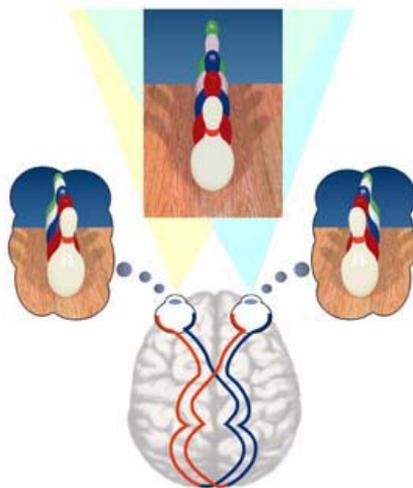


Figura 7.11 – Visões de cada olho da mesma cena.

Uma consequência imediata da diferença das imagens capturadas pelos dois olhos é o espaçamento entre o mesmo ponto projetado nas duas retinas, chamado de disparidade na retina. Por exemplo, veja a Figura 7.12: o olho da esquerda vê a árvore marrom à esquerda da árvore verde, enquanto o olho direito a vê à direita.

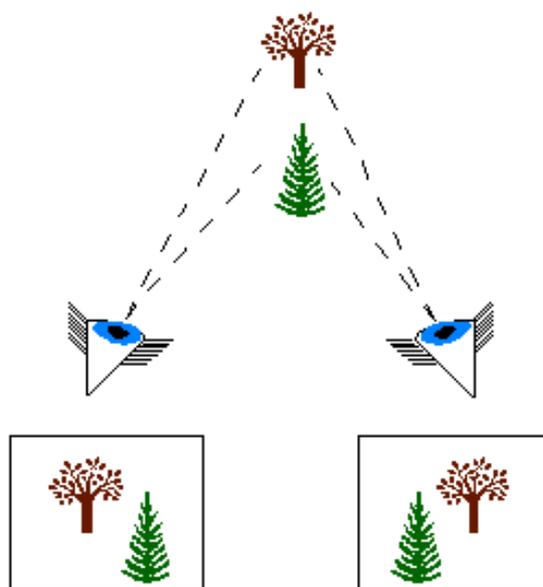
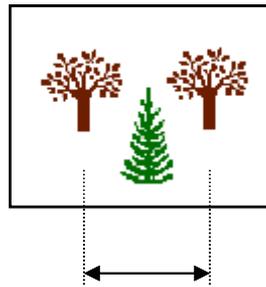


Figura 7.12 – Visões da mesma cena pelos dois olhos.

Estas duas imagens chegam ao cérebro onde é feita uma superposição, resultando na Figura 7.13. Será abordado mais adiante o relacionamento da disparidade com a noção de profundidade na cena 3D.



disparidade na retina

Figura 7.13 – Superposição das imagens e a disparidade na retina.

Na seção a seguir, veremos como é realizado a estereoscopia no computador, os equipamentos necessários para exibição e como as imagens projetadas chegam corretamente em cada olho.

7.2.2. Estereoscopia no Computador

Na seção anterior, foram abordadas noções básicas de computação gráfica, mais especificamente, os efeitos passivos e o efeito ativo, que é a estereoscopia visual propriamente dita. Nesta seção será visto como se faz para que imagens distintas cheguem a cada olho humano.

Em todo os sistemas usuais de estereoscopia visual, as imagens para cada olho são exibidas ou projetadas em um mesmo espaço. Então como se faz para cada olho enxergar o que cada um deve ver? Existem duas classificações de equipamento para estereoscopia – estéreo passivo e estéreo ativo. Tanto uma categoria quanto a outra, utilizam óculos para separar as imagens (olho esquerdo e olho direito). Dependendo do tipo de óculos, o estéreo é classificado na categoria.

A utilização dos óculos é ilustrada na Figura 7.14. Duas imagens são geradas por duas câmeras como mostrado na Figura 7.12. Os óculos separam as imagens para cada olho e o cérebro realiza a superposição (ou fusão).

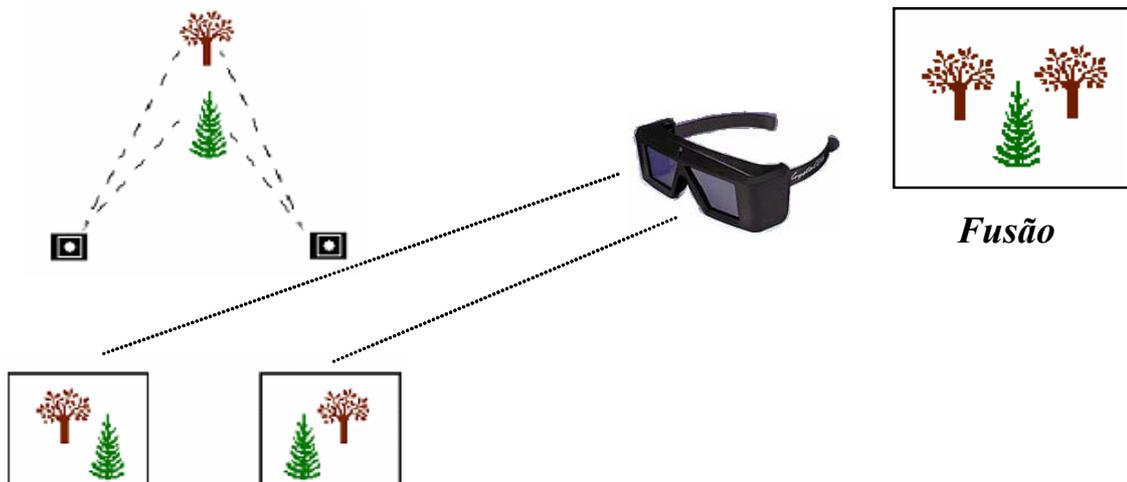


Figura 7.14 – Esquema da utilização de óculos na estereoscopia visual.

Estéreo Passivo

A primeira categoria abordada é o estéreo passivo. Nesta categoria, as duas imagens são exibidas simultaneamente e os óculos atuam como filtros.

O estéreo anáglifo (do inglês *anaglyph*) é o exemplo mais comum e muito usado nos cinemas antigamente. Neste estéreo, a filtragem é feita por cores. No início da utilização desse estéreo, eram utilizadas duas cores apenas, vermelho e azul. Isto é, a imagem para o olho esquerdo tinha apenas componente de cor vermelha e a imagem para o olho direito componente azul, como ilustrado na Figura 7.15.

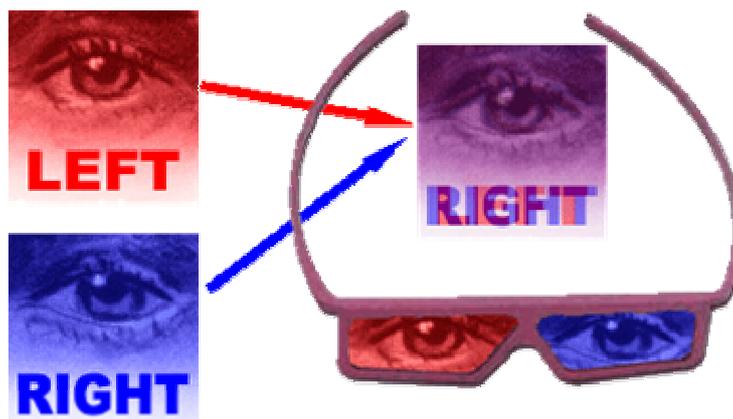


Figura 7.15 – Estéreo anáglifo.

Cada lente dos óculos é feita de papel celofane da mesma cor da componente para o respectivo olho (vermelho para o olho esquerdo e azul para o olho da direita).

Atualmente adicionou-se a componente verde para um dos olhos para melhorar a qualidade da imagem, pois era restrita apenas a dois canais de cores. Exemplo de imagem com a componente verde adicionada é vista na Figura 7.16.

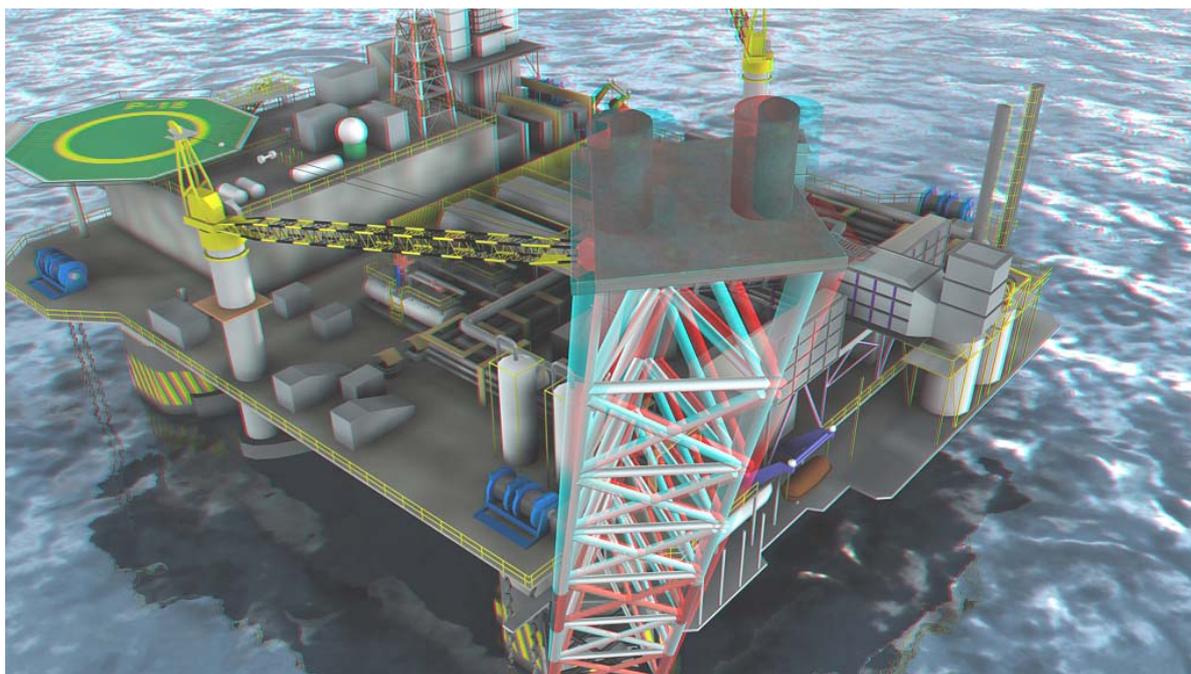


Figura 7.16 – Estéreo anáglifo com a componente verde adicionada.

As vantagens desse tipo de estéreo são:

1. Necessita apenas um projetor ou monitor;
2. Pode ser impressa;
3. Baixo custo – os óculos são facilmente confeccionados.

A desvantagem principal é a perda de qualidade que a coloração impõe.

Uma solução alternativa é a filtragem por espectro de cores. Exemplo dessa alternativa é o ChromaDepth™ [Chromatek 2003]. Nessa solução, as lentes dos óculos refratam a luz conforme o espectro de cores. O gráfico “distância versus espectro” é mostrado na Figura 7.17.

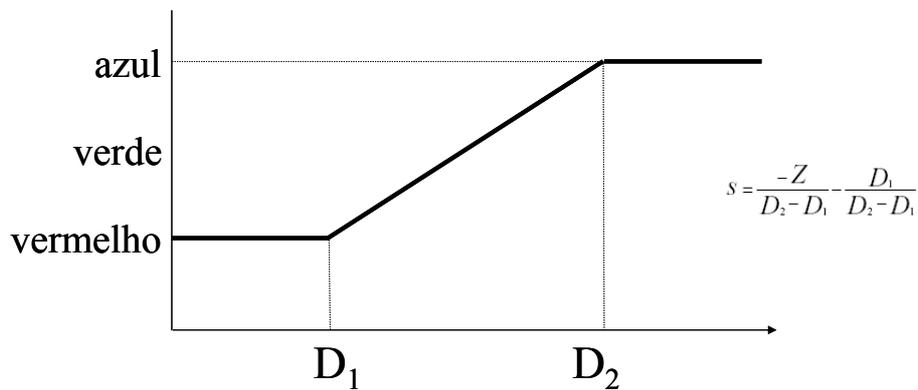


Figura 7.17 – Gráfico distância versus espectro

Isto significa que, objetos ou partes que tenham cores quentes (próximas ao vermelho) estarão mais próximas do observador do que as cores frias (próximas ao azul). Exemplo de imagem gerada para este tipo de estéreo é visto na Figura 7.18.

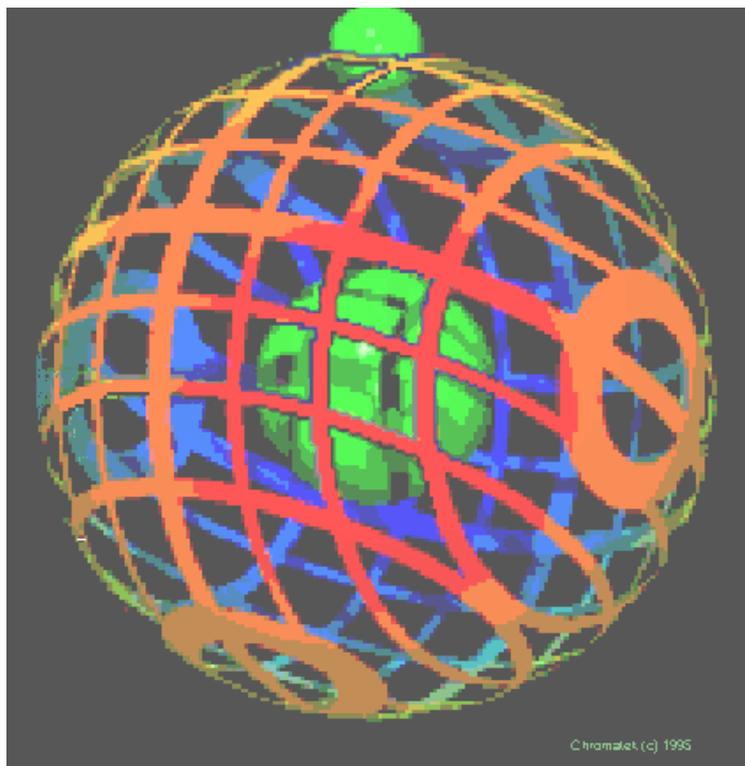


Figura 7.18 – Exemplo de imagem para estéreo por espectro de cor.

Uma grande desvantagem é a restrição à tabela de cores imposta no gráfico da Figura 7.17. Isto impossibilita seu uso, por exemplo, em visualização científica, pois a tabela de cores é em função de propriedades existentes ao longo dos objetos da cena e não em função à distância do usuário. Esta técnica funciona apenas para imagens estáticas – cenas e câmera sem alterações.

A vantagem é possibilidade de impressão. Para isto, existem dois tipos de óculos:

- ChromaDepth® 3D High Definition (HD): para visualização em monitores;
- ChromaDepth® 3D Standard Glasses: para visualização em impressos.

Uma alternativa ao estéreo baseado em cor é a utilização de polarizadores lineares. Neste caso são usados dois projetores para exibir as imagens e cada um contém uma lente polarizadora na frente. Estas lentes tem a propriedade de polarizar vertical ou horizontalmente a luz emitida pelo projetor, como ilustrado na Figura 7.19.

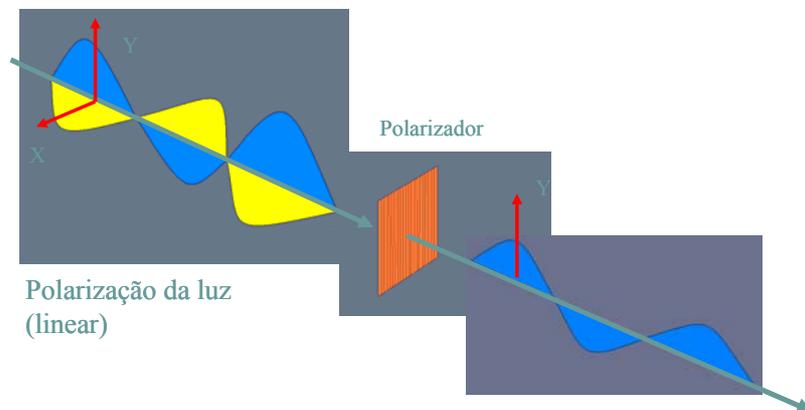


Figura 7.19 – Polarização vertical da luz.

O esquema a ser montado para este tipo de estéreo é mostrado na Figura 7.20.

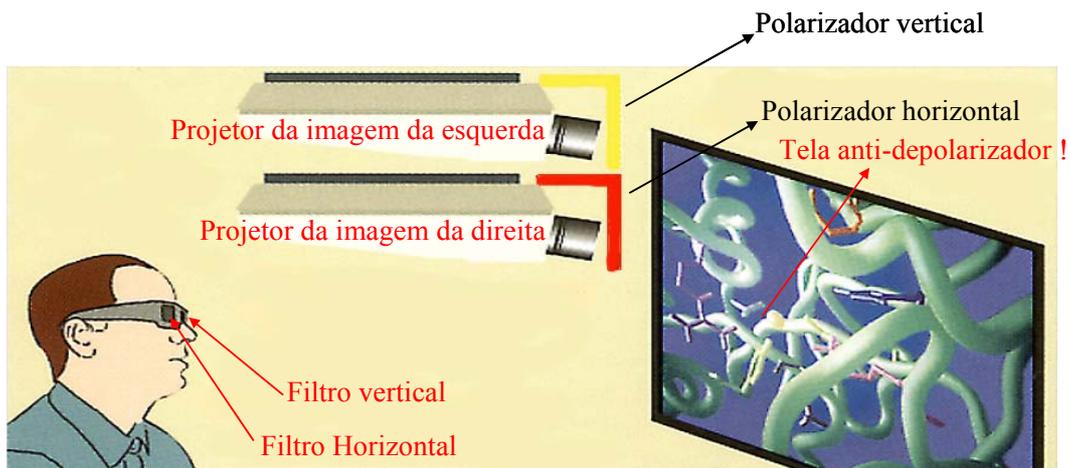


Figura 7.20 – Esquema do sistema para estéreo passivo por polarizador.

A desvantagem desse sistema é que a inclinação da cabeça do observador leva à perda do estéreo. A solução para isto é mudar o tipo de polarizador – ao invés de linear, utilizar polarizadores circulares. O esquema a ser montado é idêntico, mudando apenas os polarizadores.

Com o avanço da tecnologia, alguns monitores com display de cristal líquido estão permitindo a visualização estéreo sem uso de óculos, separando colunas pares para um olho e ímpares para o outro. O problema maior é a característica de ser mono usuário. Exemplo desses monitores é o fabricado pela DTI [DTI 2004], [About.com 2004] ilustrado na Figura 7.21.

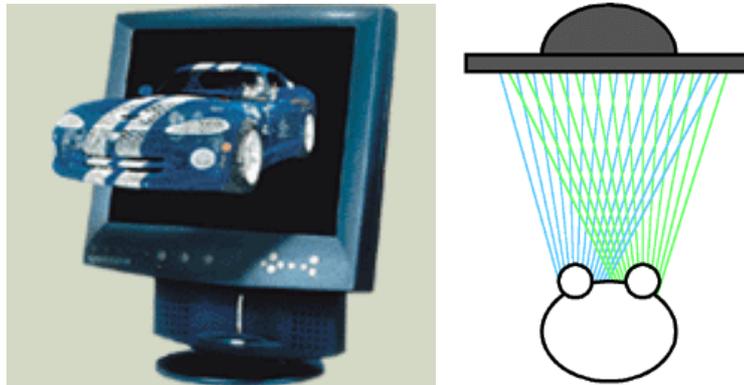


Figura 7.21 – Monitor DTI – estéreo dispensando o uso de óculos.

Estéreo Ativo

Nesse tipo de estéreo, os óculos são sincronizados com o projetor ou monitor, conforme ilustrado na Figura 7.22.

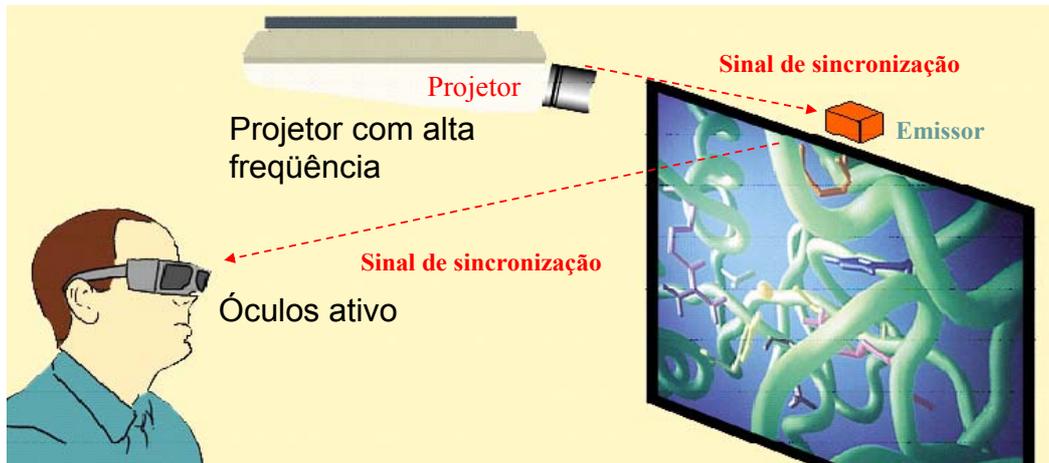


Figura 7.22 – Esquema do sistema para estéreo ativo.

No estéreo ativo, os óculos são compostos por duas lentes de cristais capazes de fechar a visão dos olhos. Quando uma visão é fechada, a outra é aberta, isto é, quando é projetada a imagem destinada ao olho direito, o projetor emite um sinal de sincronização para o emissor infravermelho, que é repassado para os óculos, fechando a visão do olho esquerdo e mantendo aberta a visão do olho direito. Similar para a visão destinada ao olho esquerdo, onde os óculos só permitem a visão para o olho esquerdo, bloqueando o direito. Levando em consideração que a frequência mínima para exibição em computador é 60 Hertz e que existem 2 imagens a serem projetadas (uma para cada olho), para termos um sistema estéreo ativo considerado bom, o projetor deve trabalhar em uma frequência mínima de 120 Hertz. Esta é uma das desvantagens desse sistema quando é usado projetor, pois seu preço é bem elevado. O preço do kit contendo um emissor e óculos sem fios também é alto. Apesar de ser possível montar o sistema com óculos ligados ao emissor através de fios, o não uso de fios permite que diversas pessoas utilizem o sistema simultaneamente, podendo então montar

salas de visualização estéreo. A maior vantagem é a qualidade, normalmente superior à projeção passiva.

7.2.3. Fundamentos Matemáticos

Nesta seção serão abordados os fundamentos matemáticos, nos pontos de vista para criação e visualização de imagens. Neste texto, as câmeras não são rotacionadas, isto é, são paralelas.

Para começar, como já foi citado ao longo do texto até aqui, existem diferenças entre imagens formadas nas retinas de cada olho quando sobrepostas. Estas diferenças são na direção horizontal. A **disparidade** é zero para objetos onde os olhos convergem. Como já foi mencionado, na Figura 7.2 os olhos estão convergindo para o polegar, enquanto na Figura 7.3 os olhos estão convergindo para a bandeirinha [StereoGraphics 1997].

Já a **paralaxe** é a distância entre os pontos correspondentes das imagens do olho direito e do esquerdo na imagem projetada na tela. Em outras palavras, disparidade e paralaxe são duas entidades similares, com a diferença que paralaxe é medida na tela do computador e disparidade, na retina. É a paralaxe que produz a disparidade, que por sua vez, produz o estéreo. Isto é, na criação de imagens em estéreo, temos que nos preocupar com a paralaxe.

Os três tipos básicos de paralaxe são:

- Paralaxe zero: conhecido como ZPS (do inglês *Zero Parallax Setting*). Este tipo é ilustrado na Figura 7.23(b). Um ponto com paralaxe zero se encontra no plano de projeção, tendo a mesma projeção para os dois olhos.
- Paralaxe negativo: significa que o cruzamento dos raios de projeção para cada olho encontra-se entre os olhos e a tela de projeção, como é visto na Figura 7.23(a), dando a sensação de o objeto estar saindo da tela.
- Paralaxe positivo: indica que o ponto está atrás do plano de projeção, como ilustrado na Figura 7.23(c), dando a sensação de que o objeto está atrás da tela de projeção.

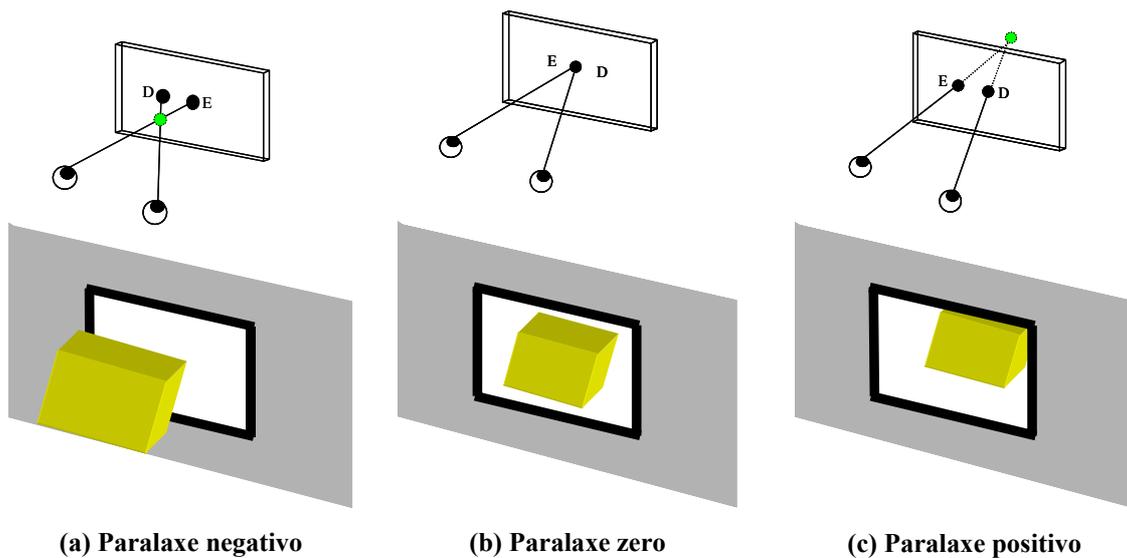


Figura 7.23 – Tipos de paralaxe.

A Figura 7.24 mostra como a paralaxe é gerada. O plano de projeção foi colocado entre os pontos A e B. Com isto, a paralaxe referente ao ponto A é positiva, enquanto para o ponto B, a paralaxe é negativa.

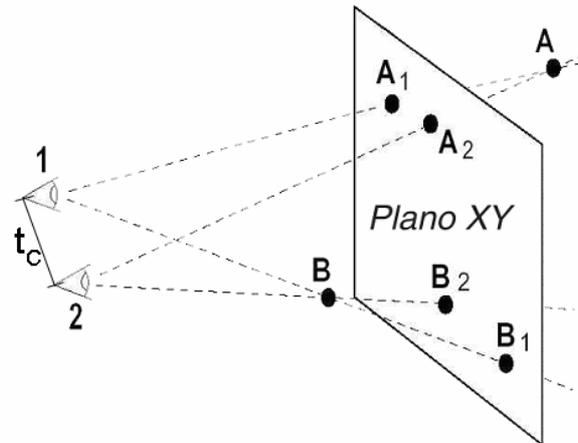


Figura 7.24 – Paralaxe a partir da projeção.

A paralaxe positiva P pode apresentar problema quando comparado à distância interaxial (t_c) – distância entre os olhos. Quando P tem valor menor, mas próximo a t_c o resultado é ruim, a menos que se queira posicionar o objeto no infinito. Se P for maior t_c , significa que há um erro, pois é um caso degenerado. Estes casos estão ilustrados na Figura 7.25.

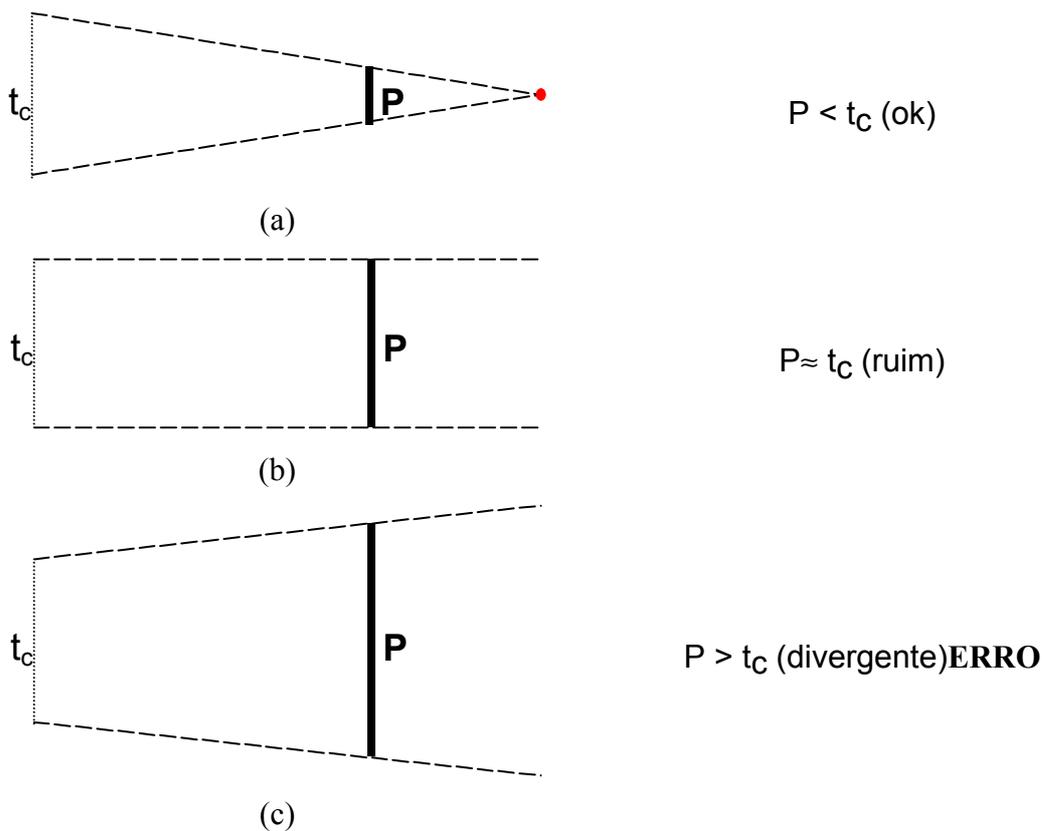


Figura 7.25 – Problemas com paralaxe positivo.

Deve-se tomar cuidado para que as projeções sempre caiam no retângulo que define o campo de visão no plano de projeção. Caso isto não ocorra, significando que apenas um dos olhos esteja vendo o ponto, a noção de estereoscopia é perdida. Isto apenas é aceitável para pontos que se movem rapidamente.

Um fator importante que deve ser levado em consideração é que a distancia do observador à tela afeta o efeito da estereoscopia. Quanto maior a distancia à tela, maior o efeito estereoscópico (tanto positivo quanto negativo). Na Figura 7.26 é ilustrado o caso para paralaxe positiva. Note que para o observador 2, que está mais distante da tela, o ponto também se encontra mais distante e para dentro, comparado ao observador 1.

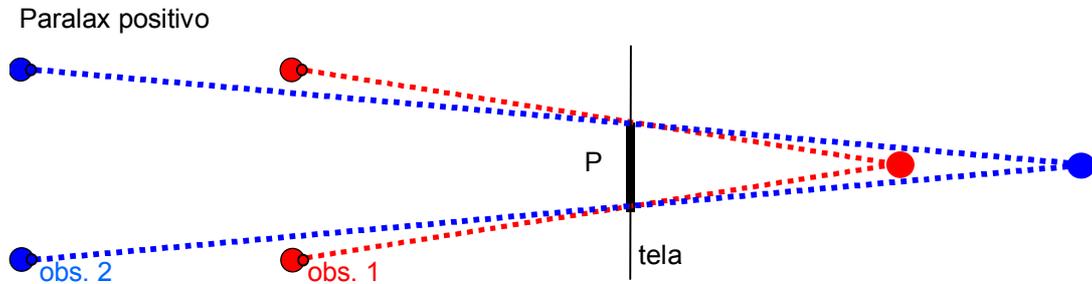


Figura 7.26 – Efeito estereoscópico com paralaxe positiva com distâncias diferentes do observador à tela.

Na Figura 7.27 é ilustrado o efeito para o caso da paralaxe negativa.

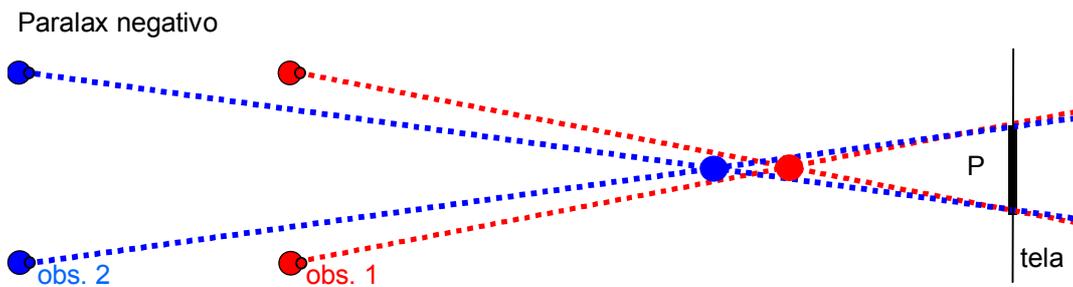


Figura 7.27 – Efeito estereoscópico com paralaxe negativa com distâncias diferentes do observador à tela.

Um grande desafio da estereoscopia é gerar maior efeito de profundidade com menor valor de paralaxe devido ao espaço físico limitado da tela e distância máxima que um ambiente comporta para os observadores. Em regra geral, o ângulo de paralaxe (β) deve estar no intervalo $[-1,5^{\circ}, 1,5^{\circ}]$, definindo paralaxes mínimos e máximos. O esquema de controle da paralaxe é ilustrado na Figura 7.28, onde d é a distância entre o observador à tela.

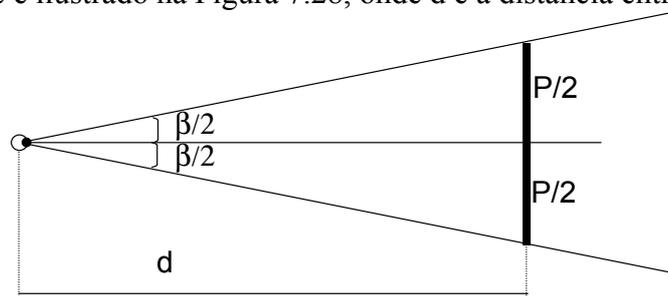


Figura 7.28 – Problemas com paralaxe positivo.

$$\text{Portanto } P = 2 * d * \tan\left(\frac{\beta}{2}\right).$$

Em uma situação de desktop, normalmente $d = 60\text{cm}$. Portanto o valor máximo de paralaxe P_{max} é $1,57\text{ cm}$. Em uma sala de visualização com distância média de 3m , o valor máximo de paralaxe P_{max} é $7,85\text{ cm}$.

Esta paralaxe é relativa ao mundo físico, onde são feitas as projeções. Para transformar para o mundo virtual temos que dividir esta grandeza pela largura da tela de projeção e multiplicar pelo número de pontos horizontais.

Quando temos câmeras rotacionadas, existe paralaxe na vertical, o que pode gerar desconforto, como mostrado na Figura 7.29.

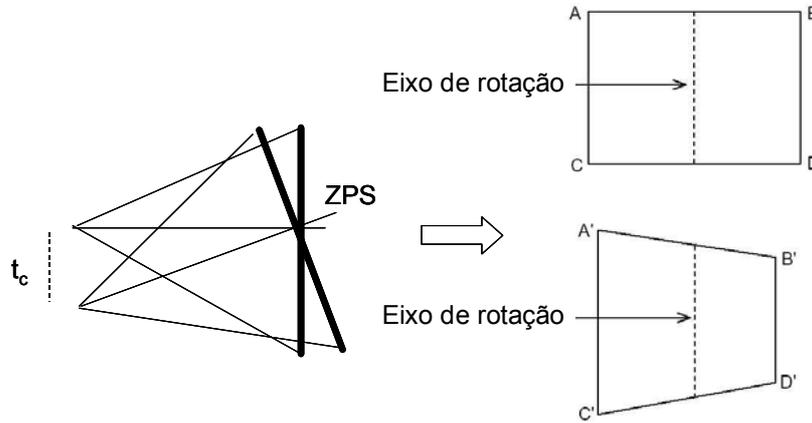


Figura 7.29 – Paralaxe vertical devido a câmeras rotacionadas.

A distância interaxial também influencia a paralaxe. Quanto maior a distância interaxial, maior é a paralaxe e, conseqüentemente, maior a sensação de estéreo. Lógico que se deve obedecer os limites para a paralaxe positiva.

Como foi dito, este texto trata apenas de câmeras paralelas, com isto, duas imagens não são sobrepostas por causa da distância interaxial, exceto no ZPS. O esquema de câmeras paralelas é mostrado na Figura 7.30 [StereoGraphics 1997], onde d_0 é a distância das câmeras ao ZPS.

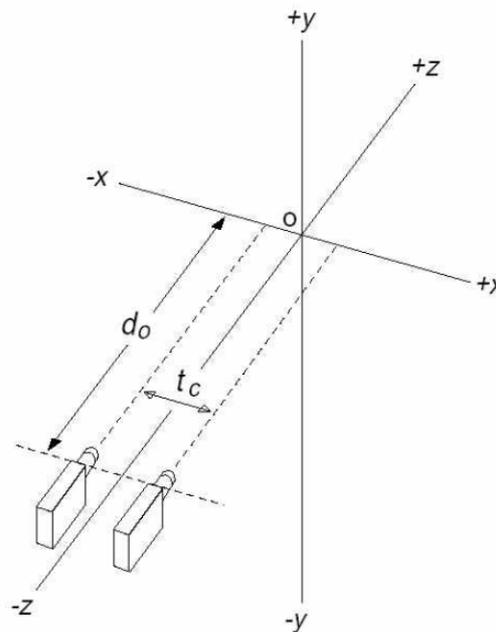


Figura 7.30 – Câmeras paralelas.

A relação da paralaxe em função da câmera é dada por:

$$P(d_m) = M \cdot f_c \cdot t_c \cdot (1/d_0 - 1/d_m)$$

onde

- M : magnificação (telas maiores, paralaxes maiores);
- d_m : objetos mais próximos da câmera produzem maior efeito de estereoscopia.

Em computação gráfica costuma-se definir o *frustum* de visão (i.e., o cone alcançável pela visão) sendo simétrico e a projeção é feita no plano Z_{near} . A simetria em estereoscopia é desfeita como será visto.

Em estereoscopia, em função de existência de duas câmeras, existem regiões de um *frustum* não alcançáveis pelo outro, e vice-versa, como é visto na Figura 7.31, observado de cima. Z_{zps} é a distância do ZPS às câmeras. Para solucionar o problema de diferença entre os *frustums*, eles são deformados, tornando-se não-simétricos. Esta deformação é dada na horizontal e é definida pelo HIT (do inglês *Horizontal Image Translation*).

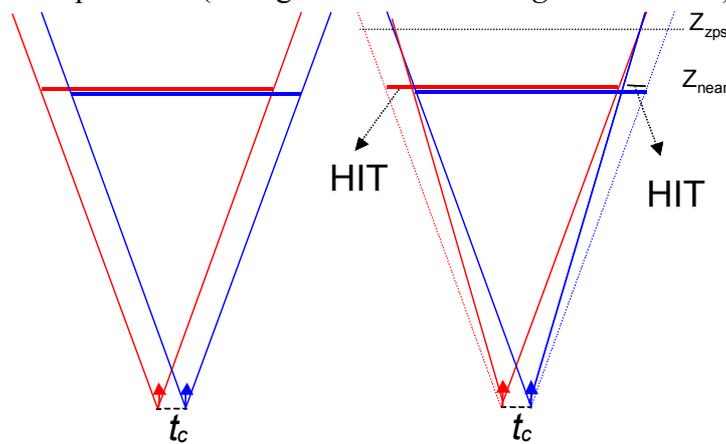


Figura 7.31 – Deformação do frustum.

O cálculo do HIT (por semelhança de triângulos) é ilustrado na Figura 7.32.

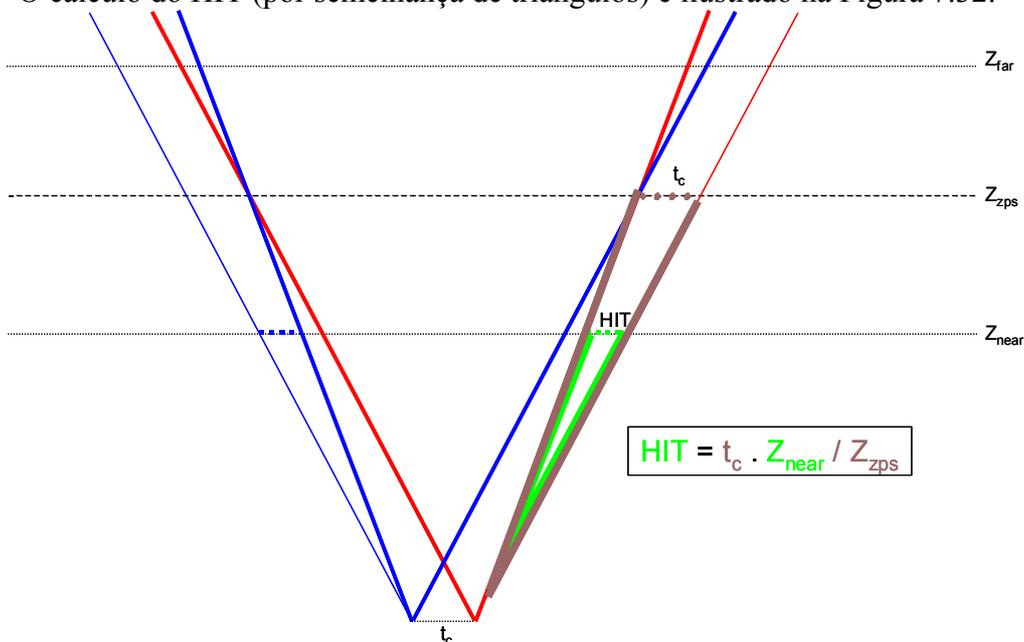


Figura 7.32 – Cálculo do HIT.

Existem diversas maneiras de se trabalhar com os parâmetros:

- Especifica-se Z_{zps} e calcula t_c a partir dele, conforme a Figura 7.33.

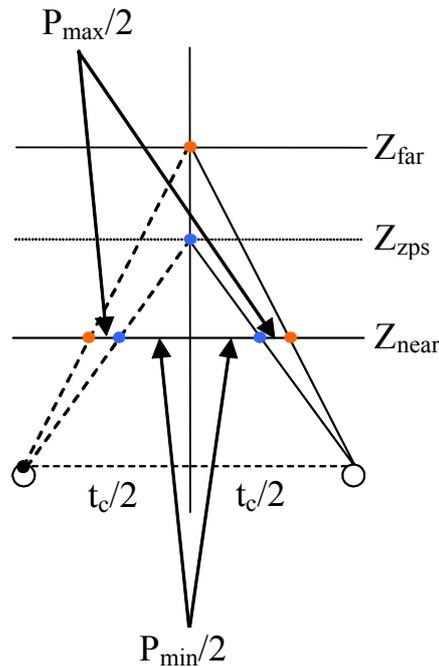


Figura 7.33 – Cálculo de t_c dado Z_{zps} .

Daí, conclui-se que:

- ✓ Para t_c em função de P_{min} apenas, temos:

$$t_{c_1} = \frac{P_{min}}{Z_{zps} - Z_{near}} Z_{zps}$$

- ✓ Para t_c em função de P_{min} e P_{max} apenas, temos:

$$t_{c_2} = \frac{P_{max}}{Z_{far} - Z_{near}} Z_{far}$$

Para que não ocorra um estouro nos limites mínimos e máximos das paralaxes, tomamos o menor valor entre os dois, isto é:

$$t_c = \min \{ t_{c_1}, t_{c_2} \}$$

- Calcula-se t_c e Z_{zps} que otimiza as paralaxes de cada cena entre um valor máximo e mínimo. Neste caso, t_c é dado em função de P_{min} e P_{max} conforme visto no caso anterior. Da Figura 7.33, podemos também encontrar que

$$Z_{zps} = \frac{t_c}{t_c - P_{min}} Z_{near}$$

No mundo real, a acomodação (mudança da forma da lente do olho) e a convergência (rotação dos olhos, corrigindo para o foco) estão associadas a um único ponto. Na estereoscopia visual através do computador, isto não ocorre. Nesse caso, as imagens são projetadas ou exibidas em uma tela e a sensação é que os objetos estão para fora ou para dentro dela. Isto é, a acomodação dos olhos é feita para a tela, enquanto que a convergência, baseada na paralaxe, não é para a tela. A acomodação e convergência são controladas por músculos e caminhos nervosos distintos. Esta quebra entre a acomodação e convergência pode gerar desconforto, causando inclusive mal-estar, devido ao hábito condicionado por toda a vida. Devido também a esta diferença, algumas pessoas podem ter dificuldade em visualizar o estéreo. Crianças e profissionais treinados em estereoscopia não costumam ter desconforto.

O único caso em que a quebra entre a acomodação e convergência não ocorre é no ZPS. Valores menores de paralaxe reduzem esta quebra, minimizando o desconforto.

Como podemos notar e comprova-se na prática, nem sempre é fácil gerar imagens estereográficas. Uma vez conhecidas as noções básicas, a composição de imagens estereográficas é uma arte, e não uma ciência.

7.2.4. TecStereoPlayer – Um Sistema para Visualização de AVIs Estéreos

TecStereoPlayer é um tocador de vídeos com estereoscopia visual. Todos os fundamentos e bases matemáticas citados na seção anterior já devem estar aplicados aos vídeos. O TecStereoPlayer apenas exibe.

O desenvolvimento do TecStereoPlayer objetivou-se em suportar diversos formatos de vídeos estéreos [Lipton 1997] e exibir também em diversos dispositivos estéreos, transformando as imagens do vídeo quando necessários para suas utilizações.

Um dos maiores problemas em sistemas de estereoscopia visual é a falta de padrão para imagens e, conseqüentemente, para vídeos. Os formatos suportados pelo TecStereoPlayer são:

- Lado a lado: as imagens estereoscópicas neste formato são formadas por duas imagens colocadas lado a lado. Exemplo desse formato é mostrado na Figura 7.34. Convencionalmente, a imagem colocada à esquerda é destinada ao olho esquerdo e a da direita, para o olho direito.

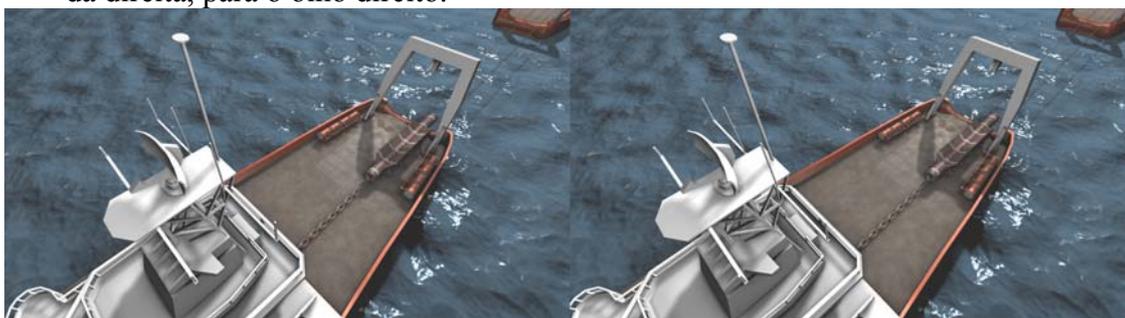


Figura 7.34 – Exemplo de formato lado a lado de imagem estéreo.

- Acima e abaixo: neste formato, as imagens estereoscópicas são formadas por duas imagens colocadas uma acima da outra. Exemplo desse formato é mostrado na Figura 7.35. Convencionalmente, a imagem colocada acima é destinada ao olho esquerdo, e a abaixo, ao direito.

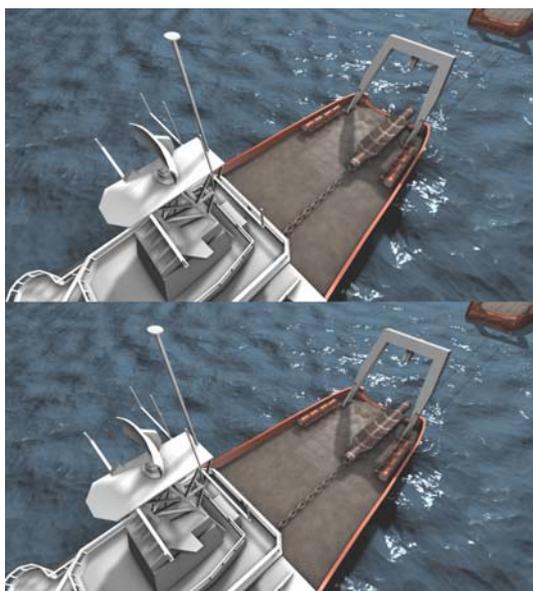


Figura 7.35 – Exemplo de formato acima e abaixo de imagem estéreo.

Para cada um desses formatos, ainda pode existir uma faixa preta separando as imagens. Uma justificativa para a existência dessa faixa, vertical no formato lado a lado e horizontal no formato acima e abaixo, é a imposição de alguns compressores de vídeo em relação à razão de aspecto das imagens. Por isto, não existe um padrão de altura (no caso vertical) ou largura (no caso horizontal) dessa faixa preta. Convencionalmente, busca-se adotar 5% da altura total para a faixa horizontal e 5% da largura total para a faixa vertical.

Outros formatos existentes, mas não suportados pelo TecStereoPlayer, são: entrelaçados por coluna e entrelaçados por linha. A existência desses formatos são devido ao fato de serem formatos específicos para determinados sistemas de equipamentos para visualização estéreo.

Os dispositivos suportados pelo TecStereoPlayer são:

- Placas gráficas com suporte ao OpenGL estéreo: neste caso, existem 4 *buffers*, chamados de QuadBuffer – dois para cada olho. Os óculos podem ser ligados diretamente à placa através de fios, eliminando a presença do emissor. Neste caso, geralmente, apenas um usuário pode utilizar o sistema por vez. Algumas placas suportam a utilização de emissores. O TecStereoPlayer divide as imagens de cada imagem do vídeo para cada *buffer* de trás de cada olho. Depois de preencher os dois *buffers* traseiros, é realizado o *SwapBuffers*.
- Monitores DTI: para estes monitores, como mencionado anteriormente, não é necessária a utilização de óculos. O TecStereoPlayer destina cada imagem de cada olho para um dos seus formatos internos e, através de uma comunicação via saída serial, é passado ao monitor o formato utilizado.
- VRex: são projetores capazes de realizar estéreo ativo e passivo, não simultaneamente. O formato exigido por este projetor é o entrelaçamento de colunas. Isto é, a imagem destinada para o olho da esquerda é distribuída pelas colunas pares da imagem enviada ao projetor e a imagem destinada para o olho da direita pelas colunas ímpares. Através de um equipamento ligado entre a placa de vídeo e o projetor, permite-se também a utilização do estéreo do OpenGL através de uma entrada de S-Video.

7.3 Realidade Virtual

Com aplicação na maioria das áreas do conhecimento, entre elas a medicina, mecânica, treinamento militar, ergonomia, jogos e entretenimento, e com um grande investimento das indústrias na produção de hardware, software e dispositivos de entrada e saída, a realidade virtual vem experimentando um desenvolvimento acelerado nos últimos anos e indicando perspectivas bastante promissoras para os diversos segmentos vinculados com a área [Stanney 2002] [Silva et al. 2004].

Como mencionado anteriormente, não existe uma definição universalmente aceita do que é RV. Porém, pelo menos duas palavras-chave estão sempre relacionadas: imersão e interatividade.

Imersão envolve a sensação de estar em um outro ambiente ou estar vendo o mundo sob outro ponto de vista. Não está necessariamente vinculada à RV, pois podemos falar em “imersão mental”, quando nos sentimos dentro de uma história narrada por um livro ou por um filme no cinema. A RV envolve a “imersão física”, que é o sentimento de estar com o corpo dentro desse outro ambiente [Sherman e Craig 2003]. Para isso, estímulos gerados pela tecnologia são enviados ao corpo que está imerso. Isso não quer dizer necessariamente que todo o corpo está imerso, ou que todos os sentidos são estimulados. Os melhores jogos hoje em dia já provêem ótima sensação de imersão em um simples PC. Dentre os estímulos gerados em simulações de RV, os mais importantes são os visuais, gerados através da estereoscopia e de configurações físicas de displays capazes de aumentar o sentimento de imersão (CAVEs e outros dispositivos que serão vistos mais adiante). Outros estímulos comuns são os sonoros (som 3D) e os tácteis (*force feedback*, para sentir colisões ou *haptics feedback*, para sentir a textura dos objetos).

Interatividade está relacionada à resposta do sistema de RV às ações do usuário. A experiência em RV envolve necessariamente a navegação (capacidade do usuário de alterar seu ponto de vista sobre um mundo virtual) e a capacidade de afetar objetos do mundo virtual. Para conseguir interatividade, é essencial a geração das imagens em tempo real, o que geralmente exige um sistema computacional robusto e uso de várias técnicas para otimização da renderização, especialmente quando o mundo virtual tem grande volume de dados [Akenine-Möller e Haines 2002]. Além disso, a interação envolve dispositivos de entrada não-convencionais, pois o mouse e teclado podem não estar disponíveis ou não serem adequados para uma situação de RV imersiva.

As seções seguintes apresentarão dispositivos de saída e de entrada usados em RV, mostrando como eles afetam, respectivamente, a imersão e a interatividade. Em seguida será apresentada a parte de software, mostrando alguns tipos de software utilizados em RV.

7.3.1. Dispositivos de Saída

Os dispositivos de saída são os elementos de um sistema de RV capazes de gerar os estímulos sensoriais para provocar o sentimento de imersão. Nesse texto, serão apresentados apenas os dispositivos visuais, embora dispositivos auditivos e tácteis também estejam disponíveis no mercado e sejam objetos de muitos trabalhos de pesquisa [Stone 2000], [Cook 2002].

Os dispositivos de saída visual podem ser classificados de acordo com o grau de imersão que provêem ao usuário: existem dispositivos não imersivos (baseados em monitor), parcialmente imersivos e (totalmente) imersivos [Silva et al. 2004], [McAllister 2002].

Os sistemas baseados em monitor fazem uso de monitores convencionais das estações de trabalho para a visualização dos ambientes virtuais. São considerados não imersivos porque o campo de visão do usuário (no ambiente virtual) é limitado às dimensões do monitor, que geralmente não passam de 21 polegadas. É possível obter a estereoscopia neste tipo de sistema através de óculos ativos ou dos monitores auto-estéreos, mencionados anteriormente. Apesar das limitações no que diz respeito ao sentimento de imersão, este tipo de sistema traz pelo menos duas vantagens: o custo reduzido e a facilidade de interação, pois o usuário não perde contato com o mundo real, podendo usar os dispositivos convencionais de interação, como teclado, mouse e *joystick*.

Os sistemas parcialmente imersivos provêm um sentimento de imersão maior que os sistemas baseados em monitor. Isso é alcançado por meio de sistemas de projeção compostos por uma ou mais telas de projeção ou vários monitores agrupados. Como a área de projeção é maior que nos sistemas não imersivos, a imersão do usuário é maior e é possível a participação de um número maior de pessoas. Por outro lado, a maior área de projeção leva a uma exigência maior de desempenho das máquinas, pois quanto mais imerso está o usuário, maior é o desconforto causado por atraso na resposta do sistema. Tanto por causa das várias projeções utilizadas quanto por causa da necessidade de máquinas mais sofisticadas, esse tipo de sistema já é bem mais caro que os não imersivos. Em sistemas parcialmente imersivos o usuário começa sair do mundo real em direção ao virtual. Assim, dispositivos de interação diferentes de teclado e mouse podem ser necessários. A Figura 7.36 mostra dois exemplos de sistemas de projeção parcialmente imersivos: o primeiro é um “canto”, composto de duas telas de projeção dispostas em ângulo de 90°, e o segundo é um *workbench*, uma espécie de mesa onde o ambiente virtual é projetado.

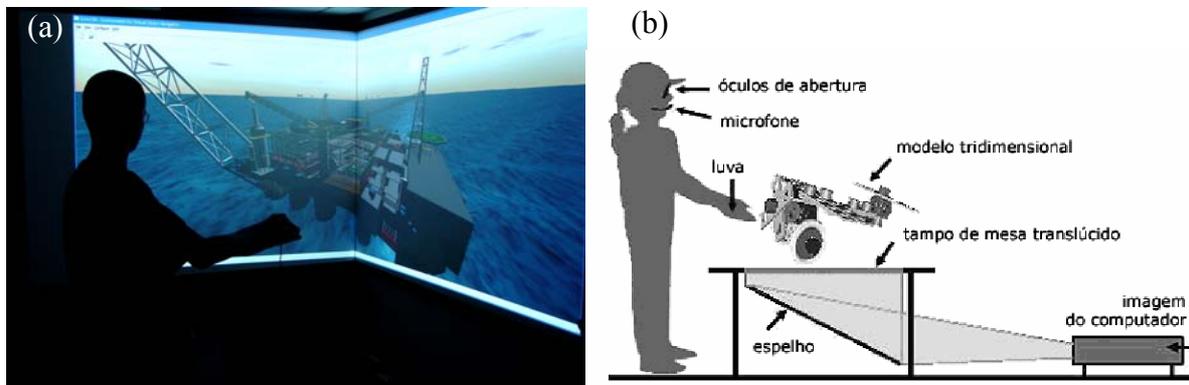


Figura 7.36 – Exemplos de sistemas parcialmente imersivos: (a) canto; (b) *workbench*.

Os sistemas de RV imersivos são aqueles que envolvem completamente o usuário, fazendo com que ele perca o contato visual com o mundo real. Capacetes e CAVEs [Cruz-Neira et al. 1993] são exemplos de sistemas desse tipo (Figura 7.37). Para garantir a imersão completa, é necessário que o usuário seja devidamente rastreado, para que tanto sua translação como seus movimentos de cabeça sejam refletidos no ambiente virtual (i.e., o ambiente deve ser gerado com o ponto de vista alterado em função do movimento do usuário). Os capacetes de RV têm a vantagem de serem mais baratos que os sistemas tipo CAVE, mas estão limitados a um usuário e são invasivos, podendo levar a cansaço físico e desconforto. As CAVEs são menos invasivas que os capacetes e permitem a participação de mais de um usuário (embora apenas um deles tenha sua posição rastreada, i.e., o ambiente virtual é visualizado de acordo com o ponto de vista de apenas um dos usuários). Além do alto custo, soluções do tipo CAVE exigem muito espaço para instalação dos equipamentos,

freqüentes manutenções para calibração dos monitores e sofisticados dispositivos de interação.

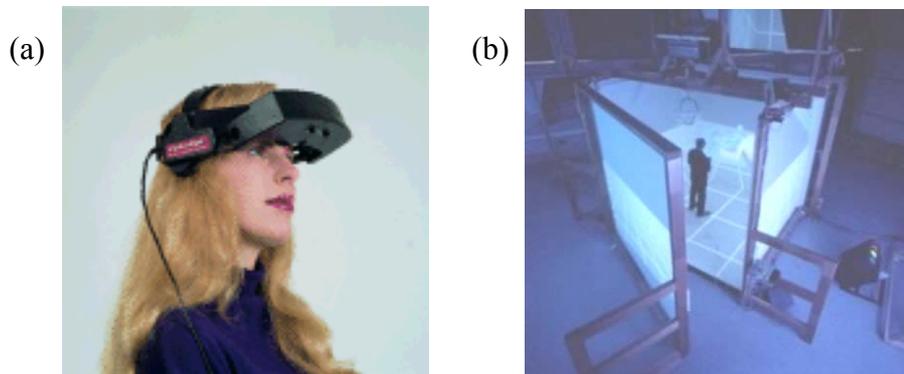


Figura 7.37 – Exemplos de sistemas imersivos: (a) capacete; (b) CAVE.

Com relação à estereoscopia, ela deve ser tratada com cuidado em todos os tipos de sistemas acima mencionados, para que ela não comprometa a imersão. Nos sistemas baseados em monitor e nos parcialmente imersivos, a existência de uma área de projeção limitada pode causar a perda do estereó caso os objetos cruzem a borda da tela. Nos sistemas imersivos e parcialmente imersivos, a violação da estereoscopia também pode ocorrer quando os objetos virtuais aparecem muito próximos do usuário (entre sua mão e a tela e projeção). Nesse caso, a informação da estereoscopia indica que o objeto deveria encobrir a mão do usuário, mas na verdade é a mão que encobre o objeto, projetado na tela. Isso gera uma informação contraditória de profundidade, que não consegue ser tratada pelo cérebro.

7.3.2. Dispositivos de Entrada

Uma das conseqüências do advento da RV foi a necessidade de se redefinir o paradigma de interface homem-computador. O sistema tradicional mouse-teclado-monitor foi substituído por dispositivos que permitissem maior imersão do usuário no ambiente virtual e o manuseio de todas as potencialidades dessa nova tecnologia [Silva et al. 2004].

Os dispositivos de entrada podem ser classificados em três tipos [Sherman e Craig 2003]: controladores físicos, de rastreamento de corpo e de reconhecimento de voz e som. Nesse texto, trataremos apenas dos dois primeiros tipos.

Os controladores físicos variam desde simples *joysticks* até complexos sistemas de simulação para aplicações específicas (por exemplo, um simulador de cabine de avião). Esses dispositivos podem fornecer três tipos de informação para o sistema de RV: analógica, digital e posicional. A informação analógica é gerada continuamente por um dispositivo, como por exemplo, o sinal enviado por um *joystick* quando sua haste é movimentada. A informação digital é binária, como por exemplo, a gerada ao apertar um botão do *joystick*. A informação posicional é obtida através dos dados vindos do dispositivo ou da conversão da informação analógica. O resultado é composto por três coordenadas de translação e pelos três ângulos de Euler, ou por um quatérnio [Foley et al. 1993].

Os dispositivos de rastreamento de corpo complementam a interação do usuário com o ambiente virtual. São eles que fornecem ao sistema de RV a informação sobre o posicionamento do usuário, permitindo a geração das imagens “corretas” (isto é, projetadas para o ponto de vista do usuário). Eles também servem para rastrear a posição da mão do

usuário ou de outros objetos usados por ele para a interação com os objetos do ambiente virtual.

Existe uma variedade de dispositivos de rastreamento, que utilizam várias tecnologias diferentes, entre eles, os eletromagnéticos, mecânicos, acústicos, inerciais e ópticos.

Os dispositivos eletromagnéticos usam campos magnéticos para medir a posição e orientação do usuário. O sistema é composto por transmissor e receptor em forma de bobina. São sistemas precisos e rápidos, mas estão muito sujeitos a interferências do meio físico (objetos metálicos) e têm alcance limitado devido ao campo magnético gerado.

Os dispositivos de rastreamento mecânicos medem ângulos e distâncias entre juntas. Dada uma posição conhecida, todas as outras podem ser determinadas pela relação entre as juntas. Uma das vantagens deste tipo de dispositivo é a facilidade de adicionar a capacidade de *force feedback*, restringindo o movimento do usuário ao aplicar força contrária a ele (simulando colisão com objetos virtuais, por exemplo). São dispositivos de alta precisão e não sofrem interferências do meio externo, mas podem restringir a movimentação do usuário, caso fiquem presos ao chão ou possuam muitas juntas.

Os dispositivos acústicos utilizam ondas sonoras, geralmente ultra-sônicas, para medir a distância. Para isso é usado um par transmissor/receptor (alto-falante/microfone). Os métodos mais usados são o cálculo do tempo de vôo e coerência de fase. Em ambos, o objetivo é converter tempo em distância. Este tipo de dispositivo é mais barato que os anteriores, mas possui uma certa latência devido à baixa velocidade de propagação do som e está sujeito a interferências do meio (ruídos e obstruções no caminho entre transmissor e receptor). Devido ao baixo alcance das ondas sonoras, o espaço de trabalho é limitado.

Os dispositivos inerciais utilizam magnetômetros passivos, acelerômetros e girômetros. Os magnetômetros passivos medem o campo magnético do ambiente (geralmente da Terra) e fornecem medidas angulares. Os girômetros fornecem medida angular mais precisa e os acelerômetros fornecem medidas lineares. Todos são baseados na segunda lei de movimento de Newton ($F = ma$ e $M = I\alpha$), sendo assim, o sistema deve integrar a leitura para obter a velocidade e a posição [Silva et al. 2004]. São dispositivos de boa precisão, sem interferência do meio externo e sem limitação física de espaço de trabalho. A desvantagem é que as integrações para cálculo da posição podem gerar erros.

A Figura 7.38 mostra exemplos de dispositivos usando as tecnologias acima mencionadas.

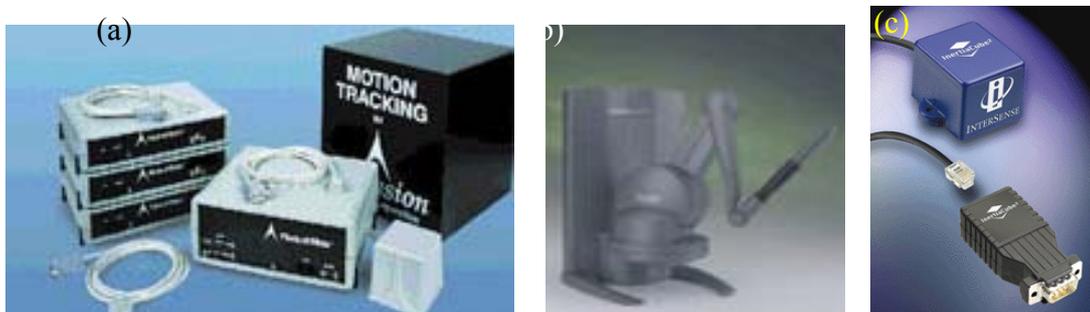


Figura 7.38 – Exemplos de dispositivos de rastreamento: (a) Flock of Birds da Ascension Technology (magnético); (b) Phantom da Sensable Technologies (mecânico); (c) InertiaCube2 da InterSense (inercial).

Uma outra classe de dispositivos de rastreamento usa informação visual para rastrear o usuário. Os dispositivos ópticos normalmente usam técnicas de visão computacional para determinar a posição do objeto a partir da imagem obtida com câmera(s) de vídeo. Outro equipamento óptico bastante utilizado é o composto por emissores de luz, como os diodos (LEDs) ou laser. Esse tipo de dispositivo tem como vantagem o baixo custo e a mobilidade que dá ao usuário (não possui fio). A desvantagem é a velocidade de captura dependente do dispositivo utilizado (câmeras NTSC, por exemplo, capturam imagens a 30 quadros por segundo) e a precisão relativamente baixa (muito dependente da técnica de visão utilizada). Uma outra restrição é que o objeto sendo rastreado deve sempre estar dentro do campo de visão da câmera. Para reduzir essa limitação, são usadas câmeras adicionais, de modo a tentar ter o objeto sempre dentro do campo de visão de pelo menos duas delas. A Figura 7.39 ilustra o local preparado com quatro câmeras emissoras de infra-vermelho para capturar a posição de um dispositivo óptico sobre a mesa. O dispositivo é recoberto com material retro-reflexivo, para retornar o sinal infravermelho na direção de cada câmera. Como existem quatro câmeras, o sistema funciona enquanto o dispositivo é visto por duas ou mais câmeras [Silva 2004].

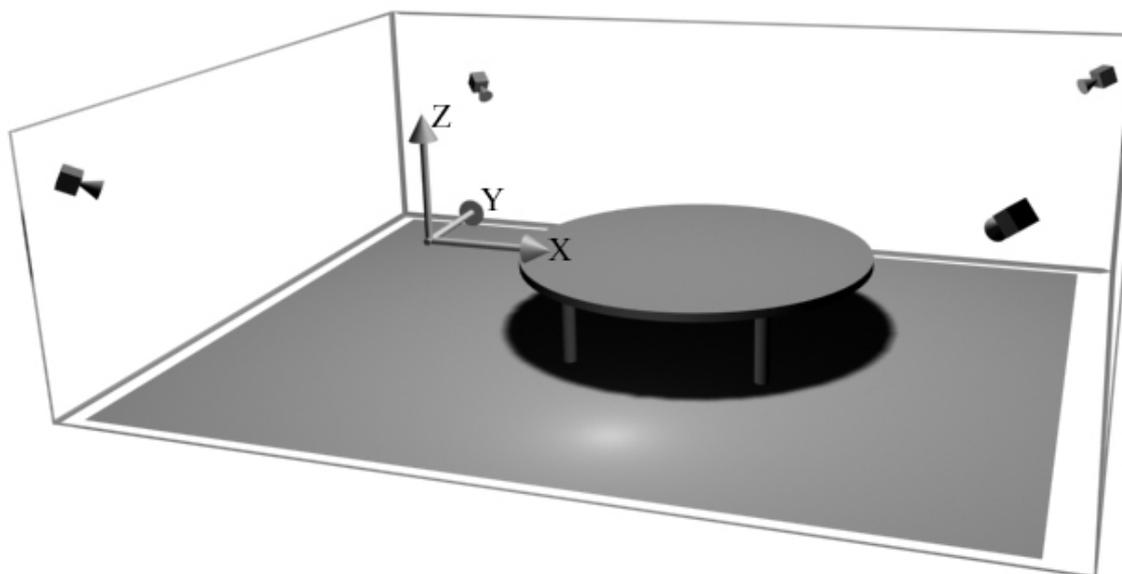


Figura 7.39 – Modelo esquemático de local de uso de dispositivo óptico.

7.3.3. Ferramentas para Desenvolvimento

Existe uma ampla gama de ferramentas para desenvolvimento de aplicações RV, desde bibliotecas gráficas de mais baixo nível, como o OpenGL [OpenGL 2004], até linguagens descritivas de grafos de cena, como a VRML (*Virtual Reality Modeling Language*) [Web3D Consortium 1997].

Grafos de Cena

Os grafos de cena são de suma importância para o desenvolvimento de aplicações de RV. Eles podem ser definidos como ferramentas conceituais para representação de ambientes virtuais tridimensionais nas aplicações de computação gráfica [Walsh 2002]. Um grafo de cena é formado por nós conectados por arestas compondo um grafo acíclico direcionado. Cada nó possui um conjunto de atributos que podem, ou não, influenciar seus nós

conectados. Os nós são organizados de uma maneira hierárquica correspondendo semântica e espacialmente com o mundo modelado.

Os nós podem ser divididos em três categorias: nó raiz, nós intermediários que são chamados de nós internos ou nós de agrupamento e os nós folha que estão localizados no final de um ramo. O nó raiz é o primeiro nó do grafo e todos os outros nós estão ligados a ele direta ou indiretamente. Os nós internos possuem várias propriedades, sendo o uso mais comum o de representar transformações 3D (rotação, translação e escala). Os nós folha contêm, geralmente, a representação geométrica de um objeto (ou dados de áudio, quando o grafo de cena possuir esse recurso) [Silva et al. 2004].

Os grafos de cena implementam um princípio chamado de herança de estado. Os nós internos armazenam o estado do sistema, onde estado significa a posição e a orientação dos objetos no ambiente virtual e seus atributos de aparência. A herança de estado é uma propriedade dos grafos de cena que determina que cada nó deve herdar as propriedades de estado de todos os seus ancestrais no grafo até a raiz.

Considerando o modelo de uma casa e admitindo que seus objetos estejam modelados em relação à origem do sistema de coordenada, podemos acrescentar para cada quarto da casa uma translação que irá posicioná-lo corretamente em relação à casa. A casa pode ser ainda rotacionada, por exemplo, a fim de ficar voltada para uma determinada direção. O grafo para essa cena é ilustrado na Figura 7.40. Devido à herança de estado, todas as geometrias identificadas pelos nós raízes herdarão as propriedades dos seus ancestrais e serão posicionadas corretamente. A herança de estado é, portanto, uma ferramenta bastante útil para organização da cena 3D.

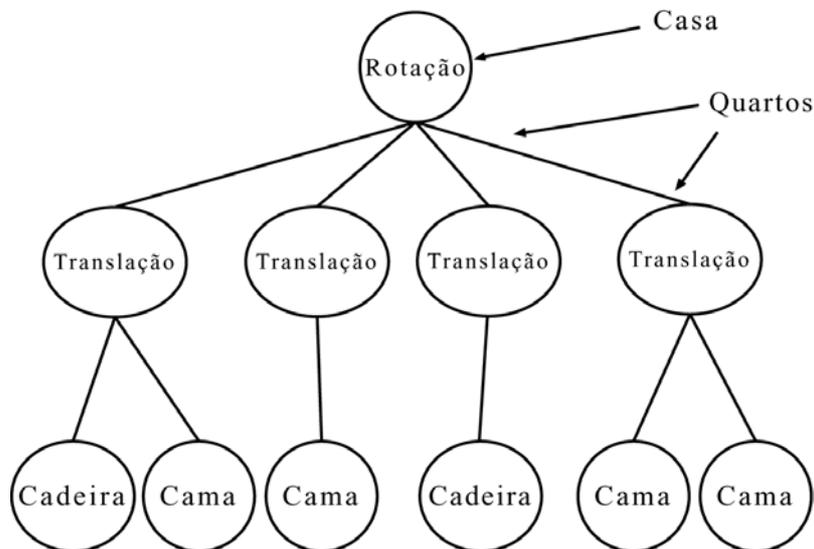


Figura 7.40 – Exemplo de grafo de cena.

Além da organização hierárquica da cena, a grande vantagem do grafo de cena é a implementação de uma série de otimizações, que são essenciais para a visualização em tempo-real. Essas otimizações servem principalmente para reduzir a quantidade de vértices que são enviados para a renderização na placa gráfica. O grafo de cena pode embutir recursos de descarte por volume de visão (objetos fora do *frustum* do observador), por oclusão (objetos ocultos por outros) e por tamanho (objetos muito pequenos em cenas complexas). Também é comum que os grafos de cena tenham recursos de níveis de detalhe (*LOD – Level of Detail*), cuja idéia é a de que objetos muito distantes do observador podem

ser renderizados com menor qualidade visual (menor número de polígonos) do que objetos mais próximos [Luebke et al. 2002].

Uma vantagem direta que o grafo de cena traz com todas as otimizações que ele implementa é a melhoria do desempenho da aplicação. Além disso, o grafo de cena traz outras vantagens para uma aplicação de RV:

- **Produtividade:** o grafo de cena gerencia toda a parte gráfica, reduzindo as várias linhas de código que seriam necessárias para implementar a mesma funcionalidade utilizando uma interface de programação baixo nível, como a OpenGL.
- **Portabilidade:** os grafos de cena encapsulam todas as tarefas de baixo nível necessárias para renderizar a cena e ler e escrever arquivos, reduzindo, ou até mesmo extinguindo, a quantidade de código que é específica de alguma plataforma. Assim, se o grafo de cena for portátil, imediatamente toda a aplicação será portátil, sendo necessário apenas uma nova compilação ao se mudar de plataforma.
- **Escalabilidade:** os grafos de cena são feitos para funcionar em configurações simples baseadas em computadores de mesa e placas gráficas aceleradoras convencionais ou em hardware complexo, como cluster de máquinas gráficas, ou sistema multiprocessado (*multipipe*).

Dos vários grafos de cena existentes, podemos destacar o SGI OpenGL Performer [Silicon Graphics 2004] e o OpenSceneGraph [Osfield e Burns 2003].

O OpenGL Performer, da Silicon Graphics, é uma interface de programação para o desenvolvimento de aplicações gráficas 3D em tempo real. Funciona em plataformas IRIX, Linux e Windows, baseado na biblioteca gráfica OpenGL. O Performer é uma excelente solução de grafo de cena e renderização em tempo real. Porém, é uma solução comercial de alto custo, desenvolvida especialmente para os equipamentos da Silicon Graphics. Em outros equipamentos, boa parte das otimizações que são extremamente importantes, não estão presentes.

O OpenSceneGraph é outra interface de programação construída sobre OpenGL. Ele é responsável pela gerência do grafo de cena e otimizações gráficas. Ele é multi-plataforma, gratuito e de código aberto, e tem uma comunidade bastante ativa em torno. Por ser orientado a objetos, o OpenSceneGraph é bastante extensível, permitindo ao usuário a criação de novas funcionalidades sem a necessidade de modificação do seu código fonte.

Ambientes de Desenvolvimento de Aplicações de RV

Os sistemas de RV precisam de um software para apresentar a aplicação para o usuário final. Para facilitar a criação de aplicações de RV, várias soluções foram sendo criadas com o intuito de prover um ambiente de desenvolvimento para RV. Um ambiente de desenvolvimento permite que o usuário se concentre no desenvolvimento da aplicação de RV ao invés de se preocupar na gerência do ambiente de RV. A Figura 7.41 ilustra como os ambientes de desenvolvimento simplificam o desenvolvimento das aplicações de RV, retirando do código da aplicação o tratamento para os diferentes tipos de dispositivos de entrada e saída.

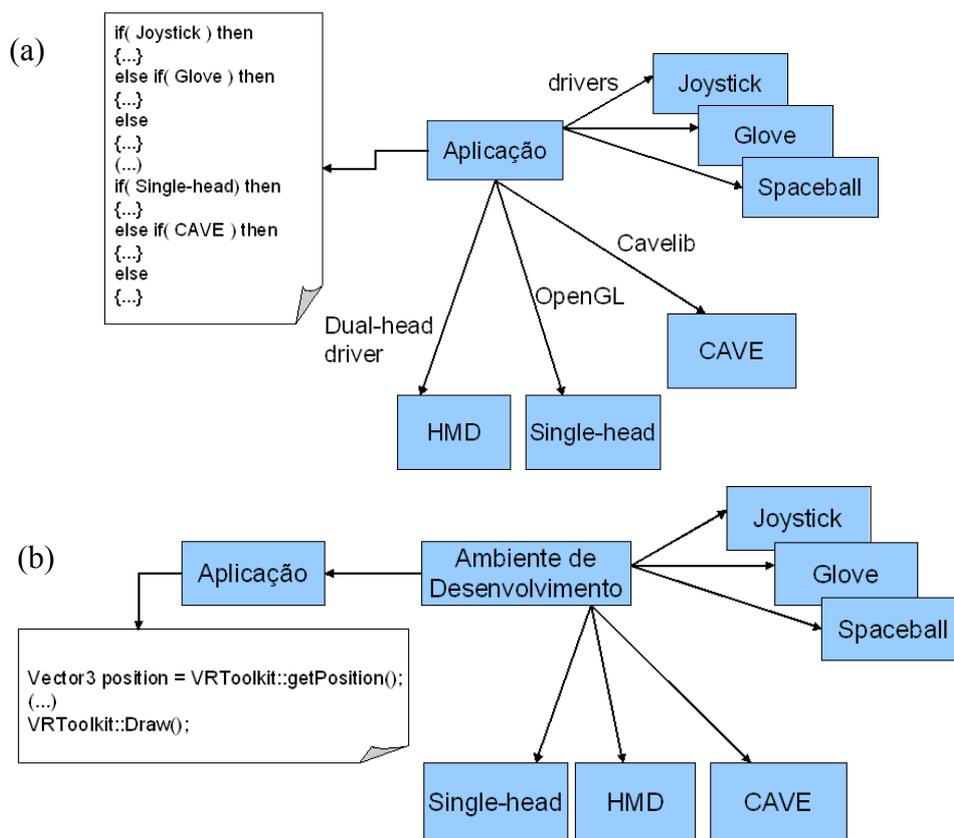


Figura 7.41 – (a) Desenvolvimento sem ambiente de abstração de dispositivos; (b) Desenvolvimento via ambiente de abstração.

O ambiente de desenvolvimento de RV define uma arquitetura que inclui componentes para gerência de dispositivos de entrada, apresentação das saídas visuais, auditivas e outras, além de um mecanismo de configuração. Alguns ambientes ainda estendem a arquitetura a fim de gerenciar também recursos como alocação de memória, multi-tarefas e comunicação.

Dentre os vários ambientes de desenvolvimento existentes, destacaremos o VRJuggler [Bierbaum 2000] e o ViRAL (desenvolvido pelo Tecgraf/PUC-Rio).

O VRJuggler é uma biblioteca de RV, com código aberto, multi-plataforma e independente de dispositivo. Ela fornece uma abstração dos dispositivos de entrada e saída. Junto a ela, qualquer biblioteca de grafo de cena (por exemplo, o Performer ou o OpenSceneGraph) pode ser usada para renderizar a cena. Por ser uma biblioteca de abstração de dispositivos, o desenvolvedor não precisa se preocupar com o dispositivo de entrada que o usuário irá utilizar para interagir com a aplicação, ou como ele irá visualizá-la. A escolha dos dispositivos é feita através de arquivos de configuração. Baseado nesses arquivos, a biblioteca carrega os *drivers* adequados. A grande vantagem dos arquivos de configuração é permitir que diferentes dispositivos sejam combinados de várias maneiras sem necessidade de recompilar a aplicação. Por exemplo, o usuário pode configurar a aplicação para rodar com um *joystick* em uma CAVE ou, apenas mudando alguns parâmetros do arquivo, utilizar um mouse 3-D imerso em um capacete.

O ViRAL (*Virtual Reality Abstraction Layer*) é outra ferramenta utilizada para facilitar o desenvolvimento de aplicações de RV. As aplicações criadas com o ViRAL são independentes de dispositivo porque ele abstrai o contexto onde elas serão executadas. As

aplicações que utilizam o ViRAL não precisam saber, por exemplo, em quantas janelas, com quantos usuários ou com quais dispositivos ela irá executar. Todas essas variáveis são configurações definidas pelo operador.

A existência de uma interface gráfica com o usuário é um dos diferenciais do ViRAL em relação ao VRJuggler (Figura 7.42). Além disso, o ViRAL pode ser usado de duas maneiras: como a aplicação principal ou embutido. Ao ser executado como a aplicação principal, ele carrega as suas janelas e menus e os diversos *plug-ins* criados pelos desenvolvedores, sendo um deles a aplicação propriamente dita. Um *plug-in* é um arquivo utilizado para alterar, melhorar ou estender as operações de uma aplicação principal (no caso, o ViRAL). A segunda maneira de se utilizar o ViRAL é embuti-lo em uma aplicação. O ViRAL funciona como escravo e a aplicação principal é o mestre. O ViRAL possui seis sistemas que podem ser utilizados embutidos em uma aplicação, sendo eles, o sistema de ambiente, usuário, janela, dispositivo, cena e *plug-in*.

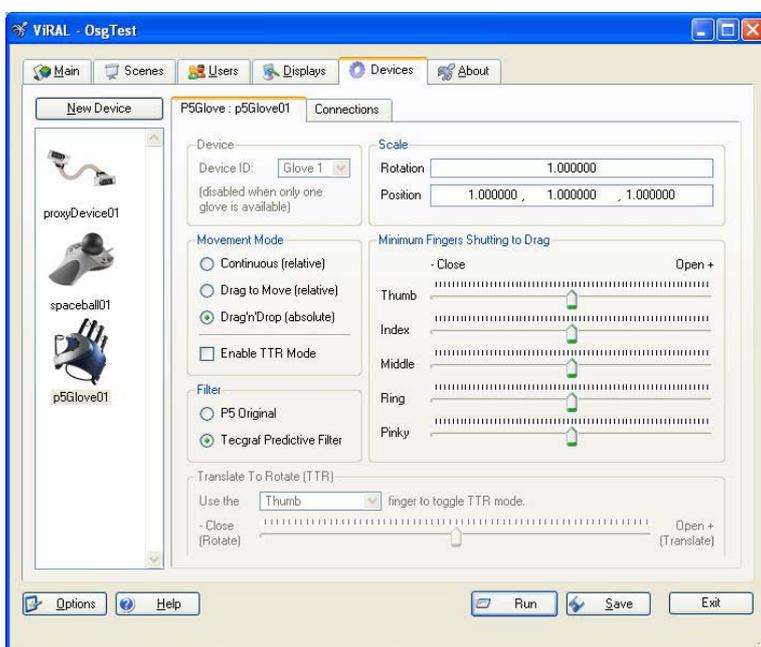


Figura 7.42 – Janela de configuração de dispositivos do ViRAL.

7.3.4. Aplicações de Realidade Virtual

Nesta seção serão apresentadas duas aplicações de RV desenvolvidas no Tecgraf/PUC-Rio: o ENVIRON e o braço de robô virtual.

ENVIRON – ENvironment for VIRtual Objects Navigation

O ENVIRON [Corseuil et al. 2003], [Corseuil et al. 2004] é um visualizador de alta performance, com o objetivo de visualizar em tempo real objetos exportados a partir do 3DSMax e do Microstation [Bentley 2003]. O objetivo é ter estes modelos visualizados em PCs “convencionais”, e não em sofisticadas estações gráficas.

O ENVIRON nasceu da necessidade de visualizar grandes modelos de engenharia (construídos em ferramentas de CAD – *Computer Aided Design*) em RV. Os problemas encontrados neste processo derivam basicamente do fato de que o modelo de engenharia não é construído com a preocupação de ser utilizado para visualização em tempo real. Em alguns casos os modelos são representações visualmente simplificadas, servindo apenas como representação esquemática das características analisadas. Por outro lado, há outros

casos em que os modelos são muito grandes e complexos para a visualização em tempo real. O ENVIRON foi projetado especificamente de modelos de CAD que se encaixam neste último grupo, onde muitos dos componentes, que são importantes no projeto de engenharia, não o são no contexto de RV. Mesmo os componentes que se aplicam à visualização, em grande parte precisam ser otimizados para isto, pois apresentam um detalhamento excessivo e caro para este processo.

Na busca de um caminho ótimo do processo de conversão “CAD para RV” foram detectados alguns obstáculos comuns no que diz respeito à visualização dos modelos. São eles:

- Performance fraca para modelos de grande complexidade: a taxa de quadros por segundo atingida é baixa quando um modelo do porte de uma plataforma de petróleo ou de um avião, por exemplo, é carregado por completo e entra-se em locais de maior densidade de objetos. Os algoritmos de otimização ainda não são suficientes para a quantidade de objetos e detalhes em questão.
- Superfícies curvas complexas não são bem tratadas: algumas superfícies aparecem no modelo CAD na forma de NURBS (*NonUniform Rational B-Splines*), como o casco de um navio, por exemplo. As malhas geradas para este tipo de superfície são muito ineficientes para visualização e demandam um tratamento especial.
- Baixo realismo: o modelo CAD geralmente não possui informações de textura associadas diretamente ao objeto. Aplicar texturas individualmente é um trabalho que não compensa dada a complexidade do modelo.

Na verdade todos esses problemas são originários da diversidade entre um modelo CAD e um modelo RV. Cada um atende a objetivos diferentes. No modelo CAD é realizado um detalhamento da forma para o processo de execução, sem levar em conta se isto vai dificultar a visualização 3D. Não existe ainda um ambiente integrado onde se possa migrar do modelo CAD para o modelo RV e vice-versa, permitindo maiores interações para que ajustes necessários sejam feitos. Os processos de conversão existentes consistem em traduções diretas e imprecisas do CAD para o RV, e normalmente são fortemente dependentes do formato de CAD traduzido.

O ENVIRON foi projetado para atacar parte dos problemas acima mencionados. Ele utiliza o OpenSceneGraph como grafo de cena e, para melhorar a performance, foi implementado um algoritmo de LOD com o objetivo de simplificar e otimizar as malhas para visualização. Para objetos baseados em primitivas simples como cilindros, esferas, linhas, curvas e extrusões e revolução dessas primitivas, o algoritmo de LOD consiste em transmitir para o visualizador as informações paramétricas de geração dessas formas (raio, posição e altura, no caso de um cilindro, por exemplo) e deixar que o visualizador, de acordo com a posição do objeto, gere a malha de triângulos necessária para a visualização. O uso de objetos paramétricos permite uma economia no consumo de memória – objetos longes o suficiente da câmera vão ocupar apenas o espaço necessário para guardar os poucos parâmetros de geração do objeto. Alguns objetos não têm uma representação paramétrica fácil de se trabalhar. Nestes casos o algoritmo implementado consiste em gerar diferentes níveis de detalhamento para a malha de cada objeto, cada um com a quantidade certa de triângulos para uma determinada distância da câmera. A cada quadro, é selecionada a melhor representação, ou seja, a representação com a menor quantidade de vértices que represente fielmente o objeto (Figura 7.43).

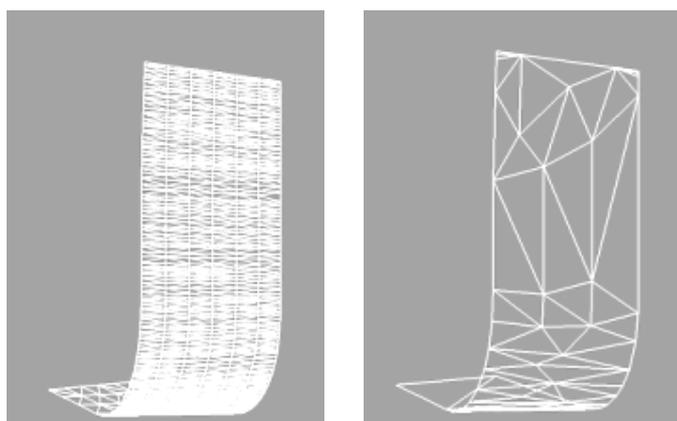


Figura 7.43. Malha otimizada: de 1315 polígonos para 59 polígonos.

Para resolver o problema das superfícies complexas, foi desenvolvido um algoritmo para o tratamento de superfícies complexas (NURBs).

Além da boa representação do modelo, outro componente importante em RV é o encaixe do modelo representado em uma cena dotada de realismo. Neste sentido, foram integradas representações visualmente dinâmicas de céu e de mar [Araújo e Celes 2002]. Com isso evita-se uma simples representação do modelo no espaço vazio. A Figura 7.44 mostra uma janela do ENVIRON mostrando o modelo de uma plataforma de petróleo (a caixa de diálogo mostra a configuração de parâmetros para a estereoscopia).

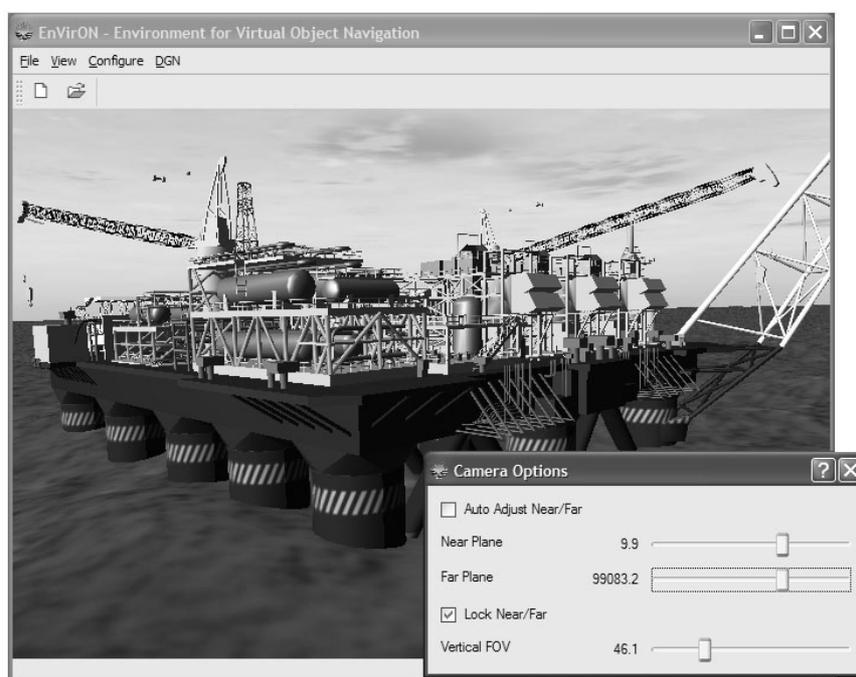


Figura 7.44. Janela do ENVIRON.

Braço de robô virtual

Esta aplicação foi desenvolvida com o ViRAL e tem como objetivo simular uma operação com um braço de robô virtual. A operação em si foi bastante simplificada, com o objetivo de permitir que usuários leigos utilizem o sistema. O braço de robô, por outro lado, segue o modelo real, inclusive com os movimentos das juntas.

A simulação foi criada como uma cena do ViRAL, o que permitiu seu uso em diversas configurações de RV, desde monitor mono, até salas com multi-projeção estéreo. Além disso, o robô pode ser manipulado com diferentes dispositivos de entrada: foram testados luva, mouse 3D e wand (uma espécie de joystick, com capacidade de rastreamento da posição e orientação da mão do usuário). Os dados da extremidade do braço, passados pelo dispositivo de entrada, servem de entrada para um programa de cinemática inversa, que fará os cálculos das posições das outras juntas, gerando o movimento do braço mecânico. A Figura 7.45 ilustra o esquema de funcionamento da aplicação do braço de robô.

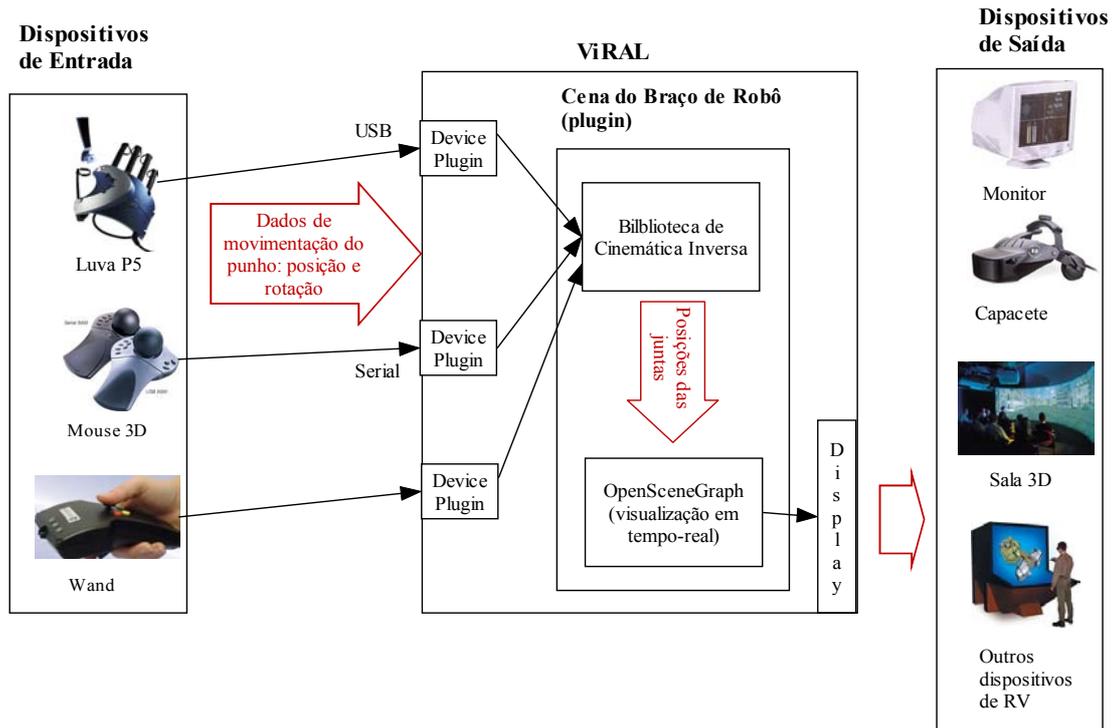


Figura 7.45. Esquema de funcionamento do braço de robô virtual.

A cena virtual, com o braço do robô e os objetos que são manipulados, foi gerada no 3DSMax, que pode ser lido pela biblioteca de visualização em tempo-real. O braço de robô, em particular, foi convertido e ajustado para o 3DSMax a partir do modelo real de projeto, construído em ferramenta de CAD pelo CENPES/Petrobras. A Figura 7.46 mostra quadros da simulação gerada.

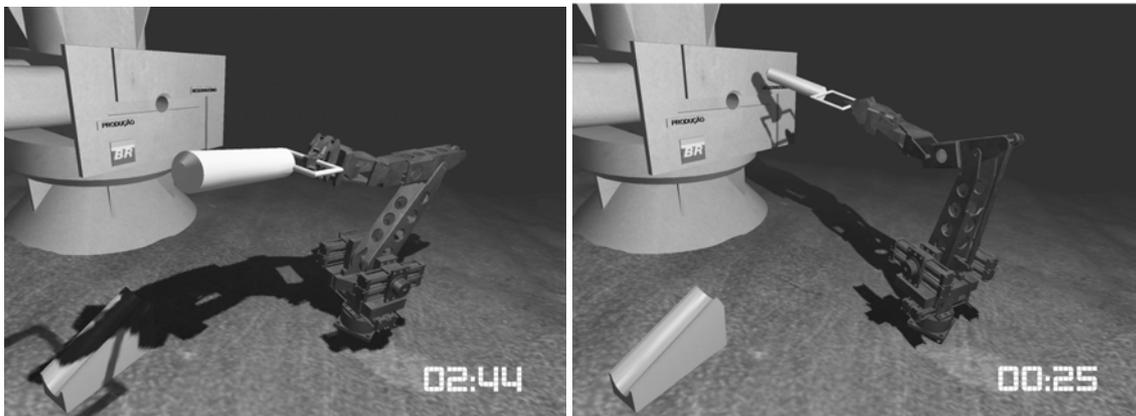


Figura 7.46. Braço de robô virtual.

7.4 Realidade Aumentada

A Realidade Aumentada (RA) é vista como uma variação da RV. Em RV, o usuário é imerso em um ambiente sintético e não participa do mundo real a sua volta. A RA permite que o usuário veja o mundo real com objetos virtuais superpostos ou combinados com ele. Portanto, a RA suplementa a realidade, ao invés de substituí-la completamente. Para o usuário, os objetos reais e os virtuais coexistem no mesmo espaço. A Figura 7.47 mostra um exemplo característico de RA onde objetos reais e virtuais compõem um ambiente, com as informações virtuais se sobrepondo ao mundo real.



Figura 7.47 – Informação virtual sobre o mundo real como auxílio a navegação.
Fonte: How Stuff Works (<http://www.howstuffworks.com>).

A possibilidade de combinar representações virtuais com o mundo real permite dar ao usuário informações adicionais sobre o mundo real que não podem ser obtidas pelos sentidos humanos. As aplicações possíveis para a RA envolvem, por exemplo o concerto de componentes internos de um sistema mecânico, cujas informações de manutenção são virtualmente mostradas sobre as peças a serem reparadas. Na medicina, a RA também pode ser bastante útil, ao indicar, por exemplo, onde o cirurgião deve fazer a incisão em um paciente [Sherman e Craig 2003]. Os militares também têm empregado a RA no intuito de prover à tropa informações vitais sobre seus arredores. Obviamente, a RA também tem grande potencial na educação e entretenimento.

Os componentes mais utilizados em ambiente de RA são os capacetes de RA e os equipamentos de rastreamento. Os ambientes de RA também podem ser executados em sistemas baseados em monitor e de projeção, através da combinação da imagem virtual e da imagem real capturada por uma câmera de vídeo.

Em um sistema de RA, a combinação da imagem real com a virtual pode ser feita por duas tecnologias: óptica e de vídeo [Rolland et al. 1994]. Os dispositivos mais empregados são os visores montados sobre a cabeça do usuário, sejam capacetes, monóculos ou outros. Chamaremos, de forma genérica, esses dispositivos de capacete de RA. O sistema de RA deve alinhar os objetos virtuais com os reais. Esse processo chama-se registro e é um dos problemas mais difíceis de se resolver em RA [Azuma 1997].

Os capacetes de RA com tecnologia óptica funcionam através da colocação de combinadores ópticos na frente dos olhos do usuário. Esses combinadores são translúcidos, de tal maneira que o usuário consiga enxergar o mundo real através dele. Eles também são

parcialmente reflexivos para que o usuário possa ver imagens virtuais geradas por saídas de vídeo acopladas no capacete e refletidas nos combinadores. A Figura 7.48(a) apresenta um diagrama conceitual da tecnologia óptica.

Os capacetes de RA com tecnologia de vídeo funcionam através da combinação de um capacete fechado (como um capacete de RV) e uma ou duas câmeras de vídeo montadas na cabeça do usuário. As câmeras de vídeo irão prover ao usuário a visão do mundo real. A imagem das câmeras é combinada com o gráfico virtual e o resultado é enviado para o visor. A Figura 7.48(b) ilustra esquematicamente a tecnologia de vídeo.

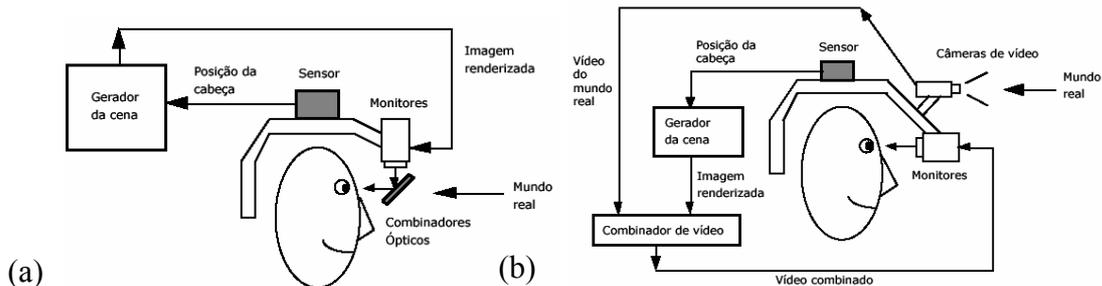


Figura 7.48 –Esquema de funcionamento de capacetes de RA: (a) com tecnologia óptica; (b) com tecnologia de vídeo [Silva et al. 2004].

Os ambientes de RA também podem ser baseados em monitor ou utilizar sistemas de projeção, ao invés do capacete de RA. Uma ou mais câmeras, estáticas ou móveis, filmam o ambiente real. No caso móvel, a câmera pode ser presa a um robô cuja posição está sendo rastreada. A imagem de vídeo do mundo real e a imagem virtual são combinadas da mesma forma dos capacetes de RA com tecnologia de vídeo, porém a imagem resultante é apresentada no monitor ou nas telas de projeção.

Os sistemas ópticos possuem apenas um fluxo de dados, o virtual, enquanto os sistemas de vídeo devem tratar da captura da imagem de vídeo e da geração da imagem virtual. Por conta disso, os sistemas de vídeo adicionam um atraso na visualização da imagem real da ordem de dezenas de milissegundos. A imagem de vídeo deve estar bem sincronizada com a imagem virtual para que não ocorra distorção temporal. O sistema de vídeo apresenta ainda uma distorção proveniente da câmera que deve ser corrigida. Os sistemas ópticos também podem apresentar distorção dependendo das lentes utilizadas, porém são bem menos perceptíveis. Os sistemas de vídeo ainda apresentam desvantagens na resolução da imagem (limitada pela resolução da câmera) e na segurança do usuário (se o sistema falhar, por falta de energia, por exemplo, o usuário perde completamente a visão).

Para o alinhamento dos objetos virtuais com os objetos reais (registro) nos sistemas ópticos a única informação de posição é passada por um dispositivo rastreador que acompanha o movimento da cabeça do usuário. No sistema de vídeo, técnicas de visão computacional permitem recuperar a posição e orientação da câmera.

7.4.1. Aplicação de Realidade Aumentada

Nesta seção será mostrada uma aplicação, desenvolvida no Tecgraf/PUC-Rio, para manipulação de um objeto virtual sobre um marcador real. A idéia é utilizar um padrão desenhado em papel e utilizar técnicas de processamento de imagens para fazer o reconhecimento desse padrão em imagens capturadas por uma câmera, desenhando o objeto virtual corretamente sobre esse padrão. A aplicação usa tecnologia de vídeo, podendo ser visualizada tanto em monitor quanto em capacete (desde que a câmera seja posicionada

sobre o mesmo). A Figura 7.49 mostra as câmeras no mundo real, e o resultado final em RA.



Figura 7.49 –Aplicação de RA: à esquerda estão as câmeras filmando o marcador real, à direita é a imagem final, com a personagem virtual sobreposta à imagem filmada.

7.5 Realidade Virtual Colaborativa

A interligação dos ambientes de visualização é um dos desafios atuais para a evolução destas tecnologias. Aplicações gráficas interligadas, operando em diferentes locais e diferentes tipos de plataformas computacionais, permitirão em breve que especialistas distribuídos geograficamente trabalhem de forma colaborativa como se estivessem no mesmo local. A integração visual de dados elimina as barreiras de comunicação normalmente existentes em projetos, trazendo ganhos de produtividade e qualidade nas tomadas de decisão e permitindo que se reduza substancialmente os prazos e riscos na execução de projetos. Daí o grande interesse atualmente existente nessas tecnologias.

Como será visto a seguir, existem duas grandes classes de desafios quando se fala em RV colaborativa. O mais comum diz respeito à integração de usuários geograficamente distribuídos nos chamados ambientes virtuais colaborativos, ou CVEs (do inglês, *Collaborative Virtual Environments*). O outro tipo de desafio diz respeito ao uso de dispositivos de saída para vários usuários no mesmo local.

7.5.1. Usuários Remotamente Localizados

Os CVEs são simulações de mundos reais ou imaginários onde vários usuários podem interagir em tempo real, compartilhar informações e manipular objetos no ambiente [Singhal and Zyda 1999], [Benford et al. 2001]. Eles representam a convergência de interesses das áreas de RV e CSCW (trabalho colaborativo apoiado por computador, do inglês *Computer Supported Cooperative Work*).

Sistemas experimentais do tipo CVE já são usados há décadas, mas apenas recentemente começaram a sair das esferas acadêmicas e militares. Esse aumento de popularidade se deve principalmente ao rápido aumento da capacidade de processamento das máquinas e sua redução de custos.

Apesar da crescente popularidade, CVEs ainda apresentam uma série de desafios em seu desenvolvimento. Entre esses desafios, podem ser mencionados os problemas relacionados ao gerenciamento de recursos de rede (controle de concorrência, perda de

dados, escalabilidade, etc.), às aplicações gráficas em tempo real (por exemplo, alocação de CPU para geração das imagens) e às aplicações multi-usuários (por exemplo, manutenção de consistência entre os usuários). Também há as dificuldades específicas da área de aplicação do CVE, tais como a integração com grandes bases de dados (por exemplo, para as informações geográficas de simulações militares) e a autenticação de usuários (para aplicações de comércio eletrônico, por exemplo). Quando o campo de aplicação é especificamente a realização de trabalhos colaborativos, somam-se os desafios relacionados à dificuldade de se trabalhar com os objetos do mundo virtual e à necessidade de se criar avatares realistas para ampliar a capacidade de comunicação entre os participantes e seu sentimento de presença [Joslin et al. 2001].

Atualmente, já existe um certo número de aplicações de RV que fornecem recursos para a colaboração, ampliando suas capacidades na direção de se tornarem CVEs. Um exemplo (Figura 7.50) é o Walkinside, um software comercial para visualização de modelos CAD [VRContext 2004]. Ele fornece recursos para que o ambiente virtual seja previamente replicado em várias máquinas conectadas pela Internet e os vários usuários naveguem pelo mesmo modelo. Os avatares podem interagir entre si e com o modelo.

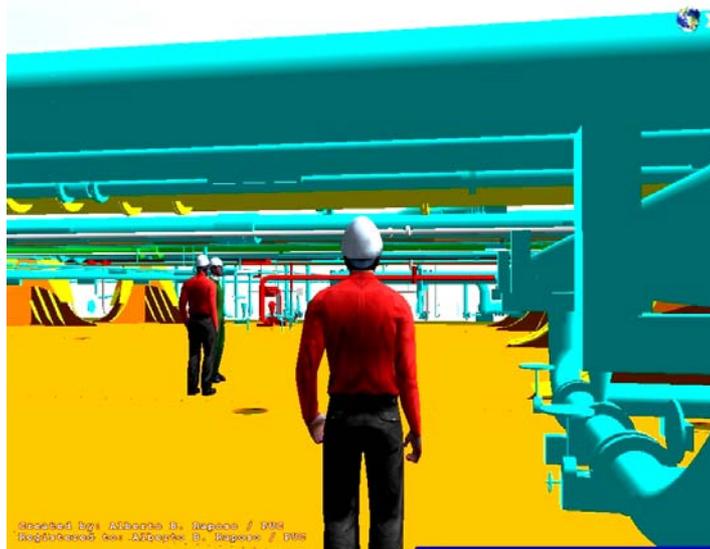


Figura 7.50. Avatares no Walkinside, em modo colaborativo.

7.5.2. Usuários no Mesmo Local

Quando queremos realizar colaboração em ambientes de RV em que os usuários estão no mesmo local, nos deparamos com outro tipo de problema. Nesse caso, a problemática diz respeito à geração das imagens para o ponto de vista correto. Como discutido anteriormente, em ambientes imersivos ou parcialmente imersivos, existe um usuário cuja posição e orientação da cabeça estão sendo rastreadas, de forma a produzir a imagem estéreo correta para o seu ponto de vista. A utilização, porém, fica restrita a apenas um usuário, pois os outros que assistem a mesma projeção a vêem pelo ponto de vista do usuário rastreado, o que pode gerar distorções. A Figura 7.51 apresenta a problemática de dois usuários visualizando o mesmo objeto de pontos de vistas diferentes, sendo que somente um deles está sendo rastreado.

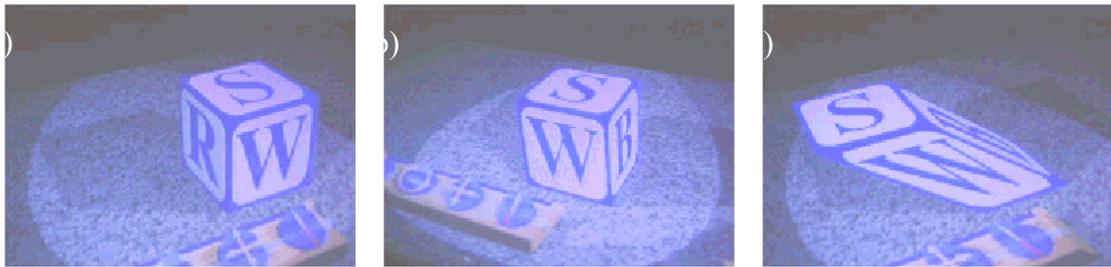


Figura 7.51 (a) A imagem vista por um usuário do lado esquerdo da mesa; (b) a mesma imagem vista do lado direito; (c) A imagem vista por um usuário à esquerda da mesa se a mesma fosse gerada do ponto de vista da direita [Agrawala et al. 1997].

Em [Agrawala et al. 1997] é apresentada uma solução para que dois usuários possam utilizar uma *workbench*. A solução empregada consiste na geração seqüencial de 4 imagens: L1 L2 R1 R2, onde L1 e R1 são as visões do olho esquerdo e direito do primeiro usuário e L2 e R2 para o segundo usuário. Para que o primeiro usuário não veja a imagem do segundo e vice-versa, os óculos 3-D devem ser capazes de fechar as duas lentes simultaneamente. Uma solução alternativa poderia combinar a projeção ativa com a passiva, de modo que a polarização vertical ou horizontal (passiva) filtrasse a projeção ativa do outro usuário. Em ambos os casos, porém, as soluções se limitam a dois usuários.

Uma outra solução para até três usuários trabalhando em uma *workbench* é o IllusionHole [Kitamura 2001]. Essa solução divide a área de projeção total em três áreas distintas e uma máscara colocada sobre a mesa permite que cada um dos usuários, posicionados em lados opostos da mesa, veja apenas a sua área de projeção, com estereoscopia (Figura 7.52).

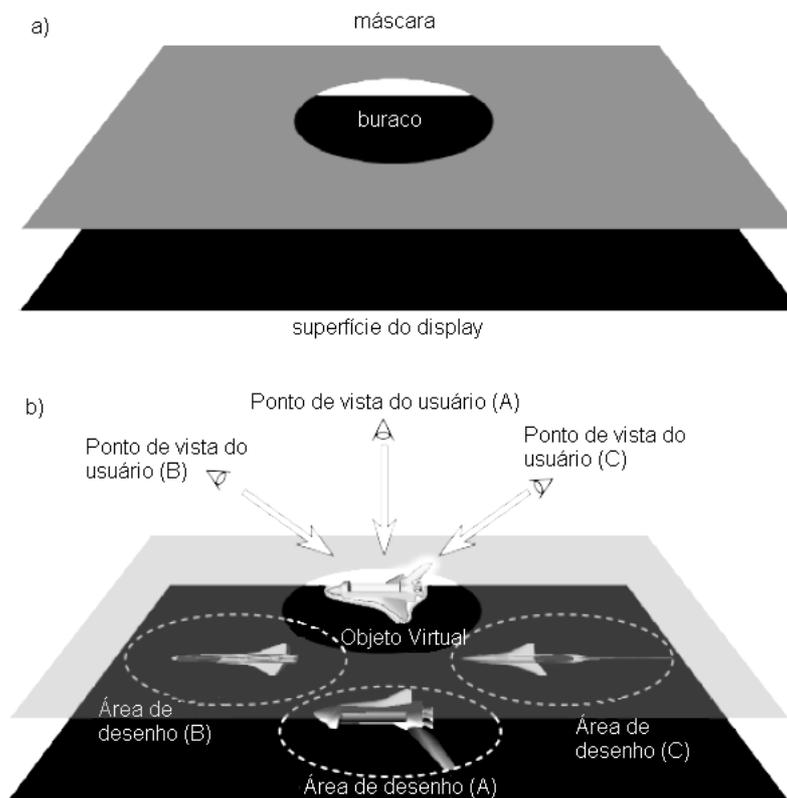


Figura 7.52 IllusionHole: (a) máscara e mesa de projeção; (b) áreas de desenho dos três usuários.

7.5.3. Tele-presença e Tele-imersão

Tele-presença ou tele-imersão é outro conceito frequentemente associado à RV colaborativa. Tele-presença envolve a habilidade de interagir (normalmente via computador) com um ambiente real remoto pelo ponto de vista do dispositivo remoto [Sherman e Craig 2003]. Esta tecnologia é ideal para o controle de robôs em ambientes hostis ou de difícil acesso para o homem (por exemplo, outros planetas, fundo do mar e interior de vulcões). O usuário deve ser capaz de manipular o robô como se estivesse no ambiente remoto. Outra aplicação da tele-imersão é a realização de cirurgias sem a presença física do cirurgião junto ao doente.

A tele-imersão está associada ao espectro de RV porque ela tem como objetivo fazer com que o usuário se sinta fisicamente presente em um ambiente remoto. A presença é simulada ao usuário através de sistemas de *feedback* sensorial: transmissão de imagens, som e sensação tátil, de maneira similar a RV e RA. Além disso, para interagir com o ambiente remoto, a tele-imersão necessita de sistemas para rastrear os movimentos do usuário, assim como em RV. Portanto, a tele-imersão pode ser definida como uma combinação de técnicas de RV imersiva (rastreamento de movimentos e *feedback* sensorial) e de transmissão de imagens em tempo-real.

Um conceito muito próximo da tele-imersão é o de tele-operação. Em ambos os casos, o usuário quer interagir com um ambiente remoto por meio de um dispositivo. A diferença, é que na tele-operação, o usuário não tem a visão de primeira pessoa, isto é, a visão do dispositivo remoto (robô, por exemplo). Tele-operação ocorre, por exemplo, ao pilotarmos um aeromodelo via equipamento de rádio. Para termos tele-imersão nesse caso, seria necessária a transmissão de uma câmera na cabine do avião para que o piloto (no chão) receba o que seria a visão do piloto.

7.6 Conclusão

Os sistemas de RV e RA vieram para revolucionar a forma como interagimos com sistemas complexos em computador. As aplicações são muitas e é difícil prever onde os ganhos e os benefícios das tecnologias do espectro de RV serão mais significativos. O certo é que não haverá um único padrão nas interfaces de RV. A tecnologia e as limitações de custos farão com que o tipo de aplicação defina o nível de sofisticação da tecnologia a ser aplicada. Neste cenário, muitas aplicações, na busca de soluções para problemas específicos, acabarão por gerar novos usos e soluções para problemas de outras áreas.

Este curso, ao apresentar alguns aspectos das tecnologias relacionadas ao espectro de RV, mostrou que essas “novas” categorias de aplicações gráficas estão se tornando uma realidade cada vez mais acessível. Começamos mostrando a tecnologia de visualização básica que está por trás da RV, que é a estereoscopia. Depois, foram apresentadas visões gerais da área de RV e RA, as duas tecnologias do espectro de RV que têm despertado mais interesse em pesquisa e desenvolvimento. Em ambos os casos, foram mostrados exemplos de aplicações desenvolvidas pelo Tecgraf/PUC-Rio, utilizando equipamentos gráficos convencionais. Finalmente, foi discutida uma tendência de todas as áreas da computação, e em particular na área de aplicações gráficas, que é a colaboração. Esta tendência se deve parcialmente ao aumento de complexidade das tarefas, que passam a requerer habilidades multidisciplinares, e aos novos paradigmas de trabalho, que envolvem diversos setores de empresas geograficamente dispersas ou até mesmo outras empresas, trabalhando conjuntamente nas diversas fases de elaboração de um produto ou desenvolvimento de um projeto.

Nos Estados Unidos e Europa, vários dos grupos de grande tradição de pesquisa em computação gráfica têm voltado seus esforços para os tópicos relacionados à utilização de ambientes de visualização imersivos, realidade virtual e tele-colaboração em aplicações científicas e tecnológicas.

Com este curso, esperamos motivar os alunos a se aprofundarem nessas áreas, que apresentam vasto campo para pesquisa e tem despertado cada vez mais o interesse da indústria. Para os interessados em conhecer melhor o trabalho dos autores e do Tecgraf, visitem o site <http://www.tecgraf.puc-rio.br>.

Agradecimentos

A pesquisa em Realidade Virtual do Tecgraf/PUC-Rio é apoiada primordialmente pelo CENPES/PETROBRAS, pela FINEP e RNP. Alberto Raposo é bolsista de produtividade em pesquisa do CNPq, processo no 305015/02-8.

Referências

- About.com (2004) The DTI 3D Monitor – real 3D with no glasses. <http://peripherals.about.com/library/weekly/aa031801a.htm>
- Agrawala, M. Beers, A. et al. (1997) “The Two-User Responsive Workbench: Support for Collaboration Through Individual Views of a Shared Space”, ACM SIGGRAPH, p. 327-332.
- Akenine-Möller, T. and Haines, E., Real-Time Rendering, A K Peters, 2002.
- Araujo, M. P. e Celes, W. (2002) “Uma Estratégia para a Visualização do Céu e seus Elementos em Jogos”, I Brazilian Workshop in Games and Digital Entertainment.
- Azuma, R. T. (1997) “A Survey of Augmented Reality”, Presence: Teleoperators and Virtual Environments, p. 355-385.
- Benford, S., Greenhalgh, C. et al. (2001) “Collaborative Virtual Environments”, Communications of the ACM, 44(7), p. 79-85.
- Bentley Systems Incorporated (2003) MicroStation <http://www.bentley.com>
- Bierbaum, A. (2000) VR Juggler: A Virtual Platform for Virtual Reality Application Development. Master Thesis, Iowa State University.
- Burdea, G. and Coiffet, P., Virtual Reality Technology, John Wiley & Sons, 1994.
- Chromatek (2003) ChromaDepth Technologies. <http://www.chromatek.com/>
- Cook, P. R. (2002) “Sound Production and Modeling”, IEEE Computer Graphics and Applications, 22(4) p. 23-27.
- Corseuil, E. T. L., Raposo, A. B. et al. (2003) “Buscando o Uso Operacional de Realidade Virtual em Grandes Modelos de Engenharia”, VI Symposium on Virtual Reality - SVR 2003, p.187-198.
- Corseuil, E. T. L., Raposo, A. B. et al. (2003) “ENVIRON - Visualization of CAD Models In a Virtual Reality Environment”, Eurographics Symposium on Virtual Environments – EG-VE 2004.
- Cruz-Neira, C., Sandin, D. J. and Defanti, T. A. (1993) “Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE”, ACM Computer Graphics, 27(2), p. 135-142.

- Dimension Technologies, Inc. (2004) DTI – The 3-Dimensional Technology Specialists.
<http://www.dti3d.com/>
- Foley, J. D. et al. Introduction to Computer Graphics, Addison-Wesley, 1993.
- Joslin, C., Molet, T. et al. (2001) “Sharing Attractions on the Net with VPark”, IEEE Computer Graphics and Applications; 21(1), p. 61-71.
- Kitamura, Y., Konishi, T., Yamamoto, S. and Kishino, F. (2001) Interactive stereoscopic display for three or more users, Proc. of SIGGRAPH 2001, p.231-239.
- Krueger, M. W., Artificial Reality II, Addison-Wesley, 1991.
- Lipton, L. Foundations of the Stereoscopic Cinema – A Study in Depth. Van Nostrand Reinhold Company, 1982.
- Lipton, L. (1997) “Stereo-Vision Formats for Video and Computer Graphics White Paper”.
http://www.stereographics.com/support/body_stereo_formats.html
- Luebke, D. et al., Level of Detail for 3D Graphics, Morgan Kaufmann; 2002.
- McAllister, D. F. (2002) “Stereo and 3-D Display Technologies”, In Encyclopedia of Imaging Science and Technology, John Wiley & Sons, p. 1327-1344.
- Milgram, P. and Kishino. F. (1994) “A Taxonomy of Mixed Reality Visual Displays”, IEICE Trans. on Information and Systems (Special Issue on Networked Reality), E77-D(12), p.1321-1329.
- OpenGL (2004) OpenGL[®] - The Industry Standard for High Performance Graphics
<http://www.opengl.org>
- Osfield R. and Burns, D. (2003) Open Scene Graph. <http://www.openscenegraph.org>
- Rolland, J., Holloway, R. and Fuchs, H. (1994) “A Comparison of Optical and Video See-Through Head-Mounted Displays”, SPIE Proceedings volume 2351: Telemanipulator and Telepresence Technologies, p 293-307.
- Sherman, W. R. and Craig, A. B., Understanding Virtual Reality: Interface, Application, and Design, Morgan Kaufmann, 2003.
- Silicon Graphics, Inc. (2004) OpenGL Performer™ Getting Started Guide.
<http://www.sgi.com/software/performer/manuals.html>
- Silva R. J. M. (2004) Integração de um Dispositivo Óptico de Rastreamento a uma Ferramenta de Realidade Virtual, Tese de Mestrado, Departamento de Informática, PUC-Rio.
- Silva, R. J. M., Raposo, A. B. e Gattass, M. (2004) “Grafo de Cena e Realidade Virtual”. In Monografias em Ciência da Computação, n.11/04, Editado por C. J. P. de Lucena, Departamento de Informática, PUC-Rio.
- Singhal, S. and Zyda, M. Networked Virtual Environments. Addison Wesley, 1999.
- Stanney, K. M., Handbook of Virtual Environments, Lawrence Erlbaum Associates, 2002.
- StereoGraphics Corporation (1997) StereoGraphics Developers’ Handbook.
http://www.stereographics.com/support/downloads_support/handbook.pdf
- Stone, R. (2000) “Haptic feedback: A potted history, from telepresence to virtual reality”, The First International Workshop on Haptic Human-Computer Interaction, p. 1-7, Springer-Verlag Lecture Notes in Computer Science.

VRContext (2004) WalkInside <http://www.walkinside.com>

Walsh, A. E. (2002) "Understanding Scene Graphs", Dr Dobb's Journal, 27(7), p. 17-26.

Web3D Consortium (1997) VRML97 Functional specification, ISO/IEC 14772-1:1997.
http://www.web3d.org/x3d/specifications/vrml/ISO_IEC_14772-All/ .