

DEPARTAMENTO DE
ENGENHARIA ELECTROTÉCNICA E DE COMPUTADORES

Síntese por Computador
de
Imagens Estereoscópicas
com
Elevado Realismo

João Miguel Queirós Magno Leitão

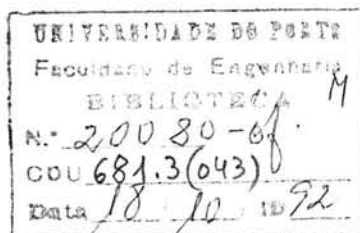
FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Rua dos Bragas, 4099 Porto Codex – PORTUGAL

Faculdade de Engenharia da Universidade do Porto
Departamento de Engenharia Electrotécnica e de Computadores

Síntese por Computador de Imagens Estereoscópicas com Elevado Realismo

João Miguel Queirós Magno Leitão



N.º. 34726

Tese submetida para satisfação parcial
dos requisitos do programa do curso de
Mestrado em Engenharia Electrotécnica e de Computadores,
perfil de Informática Industrial.

Porto, Abril de 1994

043 M

L 548

EX. 2

Tese realizada sob a supervisão do Professor Doutor F. Nunes Ferreira,
Professor Catedrático da Faculdade de Engenharia da
Universidade do Porto

Tese realizada por:

João Miguel Queirós Magno Leitão

Licenciado pela Universidade do Porto em
Engenharia Electrotécnica e de Computadores.

Assistente do Segundo Triénio do
Instituto Superior de Engenharia do Porto.

jml@diana.inescn.pt
tel: +351 2 2094229 / 4000

INESC, CG&CAD
Largo Mompilher, 22
4000 PORTO, Portugal

Prefácio

Desde que terminei a Licenciatura em Engenharia Electrotécnica e de Computadores, em 1989, tenho-me dedicado à síntese de imagem por computador, a partir de modelos tridimensionais de cenas. Desta fase do meu trabalho, integrado no grupo de Computação Gráfica e CAD do INESC, foram especialmente significativos os trabalhos que desenvolvi com o objectivo de optimizar a geração de imagens de elevado realismo, utilizando o algoritmo de Ray-Tracing.

Nos últimos anos, no entanto, têm vindo a desenhar-se fortes perspectivas de trabalho para aplicações em ambientes de realidade virtual. Nestas aplicações, devido ao papel dominante da visão na percepção humana dos ambientes, é de crucial importância a utilização de imagens convincentes e informativas. Por esta razão, e graças às possibilidades concedidas pelos dispositivos de visualização normalmente utilizados nestes ambientes, é comum (e conveniente) utilizar imagens estereoscópicas.

A ideia de realizar uma Tese de Mestrado sobre síntese e visualização de imagens estereoscópicas, surgiu então como forma de me preparar para um dos diversos problemas com que me vou deparar nos ambientes de realidade virtual, mantendo a continuidade com o trabalho anterior sobre síntese de imagem. A implementação de um programa optimizado para obtenção de imagens estereoscópicas realistas por Ray-Tracing, surge também como sequência natural dos anteriores desenvolvimentos deste algoritmo, e face às perspectivas de trabalho futuro.

A relativa novidade da estereoscopia no mundo das Ciências da Computação, a par do conhecimento secular dos aspectos físicos envolvidos, tornaram o tema escolhido para esta tese, no mínimo, aliciante. Além disso, a estereoscopia apresenta-se cada vez mais como uma solução viável e interessante para elevar o nível de realismo e informação das aplicações gráficas. Não deverá ser necessário esperar muitos anos para que os dispositivos de visualização estereoscópica se tornem vulgares em sistemas de baixo custo. A ideia de participar nesta possível revolução dos dispositivos de visualização desde logo me seduziu.

Agradecimentos

Agradeço particularmente ao meu orientador, Professor Doutor Fernando Nunes Ferreira, a confiança depositada e a colaboração prestada durante todo o tempo em que decorreu o curso de Mestrado, bem como todos os conselhos que muito contribuíram para a tese apresentada.

De forma especial, agradeço também a todos os meus colegas do grupo de Computação Gráfica e CAD, do INESC, a ajuda, a disponibilidade e a amizade manifestadas.

Quero ainda agradecer a toda a minha família, em especial à minha mãe e à minha mulher, pela paciência que demonstraram e pelo apoio que sempre me deram.

Por fim, agradeço ao ISEP as facilidades concedidas, ao INESC as condições de trabalho oferecidas, e ao INIC a bolsa de estudo concedida.

Palavras Chave

Ray-Tracing

Síntese de Imagem

Visualização

Realismo

Estereoscopia

Paralaxe

Profundidade

Keywords

Ray-Tracing

Image Synthesis

Visualization

Realism

Stereopsis

Parallax

Depth

Resumo

A visualização de imagens tridimensionais com base no efeito estereoscópico tem sido objecto de interesse para utilizadores e investigadores de diversas áreas. Nos últimos anos, com a vulgarização crescente dos dispositivos de visualização estereoscópica, surgiram diversas aplicações que utilizam o fenómeno da paralaxe para aumentar o realismo das imagens ou para melhorar a transmissão de informação geométrica ao utilizador. As áreas de aplicação incluem medicina, sistemas de controlo e robótica, cartografia, desenho assistido por computador, aplicações militares, e investigação sobre a visão humana.

Nesta tese, vamos analisar a criação de sensação de profundidade a partir de imagens planas, e examinar os diversos métodos publicados que se utilizam para gerar imagens estereoscópicas. Vamos ainda apresentar um novo método, baseado no algoritmo de Ray-Tracing, assim como as diversas técnicas incorporadas com o objectivo de reduzir o tempo de cálculo. Como demonstram os resultados apresentados, obtidos durante as experiências realizadas, o algoritmo desenvolvido permite gerar um par estereoscópico de imagens em apenas 70% do tempo necessário à geração do mesmo par, utilizando o algoritmo tradicional de Ray-Tracing.

Abstract

Users and researchers from different areas have been interested in the visualization of three-dimensional images based on the stereoscopic effect.

In the past few years, after the common use of stereoscopic visualization devices, some applications which use the parallax phenomenon either to increase the realism of images or to improve the transmission of geometrical information to the user, have appeared.

The application areas include medicine, robotics and control systems, cartography, CAD, military and investigation related to human vision.

In this thesis we analyze the creation of depth sensation in flat images and we examine the different published methods that are used to generate stereoscopic images. We also present a new method based on the Ray-Tracing algorithm and several techniques to the propose of reducing the calculation time.

According to performed experiences and their results, this algorithm permits the creation of a stereoscopic pair of images in only 70% of the time needed to the generation of the same pair using the traditional Ray-Tracing algorithm.

Índice

Índice	1
Índice de Tabelas	4
Índice de Figuras	5

Capítulo 1

Introdução.....	8
1.1. Trabalhos Anteriores	9
1.2. Organização da tese	9

Capítulo 2

Informação de Profundidade	11
2.1. Pistas Monoculares.....	11
2.1.1. Luz e Brilho	11
2.1.2. Tamanho relativo	12
2.1.3. Interposição	12
2.1.4. Textura	12
2.1.5. Paralaxe de movimento	13
2.1.6. Perspectiva aérea	13
2.1.7. Perspectiva	13
2.1.8. Depth Cuing	13
2.1.9. Acomodação.....	14
2.2. Estereoscopia	14
2.2.1. Convergência	15
2.2.2. Disparidade Retinal	15
2.2.3. Paralaxe	16
2.2.4. Separação Interaxial.....	18
2.3. Conclusão	18

Capítulo 3

Apresentação de imagens tridimensionais	19
3.1. Técnicas de Visualização de Imagens Tridimensionais.....	20
3.1.1. Free Viewing	20
3.1.1.1. Visualização Paralela	20
3.1.1.2. Visualização Cruzada	22
3.1.2. Métodos de Selecção de Imagem	23
3.1.2.1. Barreiras de Paralaxe	24
3.1.2.2. Folha Lenticular.....	25
3.1.2.2.1. Geração de Imagem para Dispositivos Lenticulares	26
3.1.3. Estereoscópios.....	27

3.1.4. Efeito Pulfrich.....	27
3.1.5. Polarização das Imagens	28
3.1.6. Anáglifos	29
3.1.7. Multiplexagem Temporal das Imagens.....	29
3.1.7.1. CrystalEyes	30
3.1.8. Holografia.....	34
3.1.8.1. Hogramas de Transmissão	35
3.2. Armazenamento de Imagens Estereoscópicas	36
3.2.1. Transparências	36
3.2.2. Vectógrafo.....	37
3.2.3. Gravação Vídeo	39
3.2.4. Estereogramas Aleatórios.....	40
3.3. Conclusão	41

Capítulo 4

Visualizador Montado na Cabeça	43
4.1. Modelo de um sistema óptico monoscópico.....	44
4.2. Modelo de um sistema óptico estereoscópico	48
4.3. HMD's baseados em CRTs	49
4.4. HMD's por Fibra Óptica.....	51
4.5. HMD's com Escrita Directa na Retina	52
4.6. Conclusão	53

Capítulo 5

Síntese de Pares de Imagens.....	54
5.1. Geração de Imagem Monoscópica	55
5.1.1. Projecção Paralela.....	55
5.1.2. Projecção Perspectiva	56
5.2. Geração de imagens estereoscópicas.....	57
5.2.1. Utilização de Rotações.....	57
5.2.1.1. Rotações com Projecção Paralela.....	59
5.2.1.2. Rotações com Projecção Perspectiva	60
5.2.2. Utilização de Shears.....	61
5.2.2.1. Shears com Projecção Paralela.....	62
5.2.2.2. Shears com Projecção Perspectiva	63
5.2.3. Utilização de Direcções de Observação Comuns	65
5.2.3.1. Direcções de Observação Comuns com Projecção Paralela	67
5.2.3.2. Direcções de Observação Comuns com Projecção Perspectiva.....	67
5.3. Conclusão	69

Capítulo 6

Visualização de Imagens Estereoscópicas.....	70
6.1. Rotações com Projecção Paralela.....	75
6.2. Rotações com Projecção Perspectiva	76
6.3. Shears com Projecção Paralela	78
6.4. Shears com Projecção Perspectiva.....	79

6.5. Direcções de Observação Comuns com projecção Perspectiva	80
6.6. Conclusão	84

Capítulo 7

Qualidade das Imagens Estereoscópicas.....	86
7.1. Ortoestereoscopia	86
7.2. Realismo de Imagens Estereoscópicas	87
7.3. Fontes de Erro	90
7.3.1. Crosstalk.....	90
7.3.2. Convergência Incorrecta	91
7.3.3. Acomodação não relacionada com a convergência	91
7.3.4. Ângulo de visão incorrecto.....	92
7.3.5. Não utilização de projecção descentrada	92
7.3.6. Distância interocular ignorada.	92
7.3.7. Posição de observação ignorada.....	94
7.3.8. Distorção óptica ignorada	95
7.3.9. Utilização apenas dos vértices em sistemas não lineares.....	96
7.3.10 Conflito com os limites do ecrã	97
7.4 Animação.....	98
7.4.1 Ecrãs do tipo Raster.....	98
7.4.1.1 Double Buffering	99
7.4.2 Animação de Imagens Estereoscópicas.....	101
7.5 Conclusão	102

Capítulo 8

Ray-Tracing Estereoscópico.....	103
8.1. Ray-Tracing Tradicional.....	103
8.1.1. Utilização da Coerência de Imagem.....	104
8.2. Ray-Tracing Estereoscópico	105
8.2.1. Intersecção Raio/Objecto Filho	108
8.2.2. Cálculo de Iluminação.....	109
8.2.3. Resultados Obtidos	111
8.2.4. Cortes na Árvore de Objectos	113
8.2.5. Cortes na Árvore de Shading.....	115
8.2.6. Raios Reenviados à Origem.....	117
8.2.7. Utilização da Coerência de Imagem.....	121
8.3. Conclusão	124

Capítulo 9

Conclusões e Trabalho Futuro	125
---	------------

Capítulo 10

Bibliografia	127
10.1. Referências Bibliográficas.....	127
10.2. Outra Bibliografia Utilizada	130

Índice de Tabelas

Capítulo 3

Tabela 3.1. Classificação dos métodos de apresentação de imagens tridimensionais.....	20
---	----

Tabela 3.2. Valores utilizados na estação gráfica IRIS Indigo XS24Z.....	33
--	----

Capítulo 4

Tabela 4.1. Resumo das tecnologias disponíveis para varrimento com raio laser.....	53
--	----

Tabela 4.2. Tecnologias para implementação de HMD's.....	53
--	----

Capítulo 7

Tabela 7.1. Classificação das imagens em termos de realismo.....	88
--	----

Tabela 7.2. Aumento do realismo provocado pela adição do efeito estereoscópico para cada tipo de imagem.....	89
--	----

Capítulo 8

Tabela 8.1. Tempos necessários à obtenção de imagens com diversas resoluções, para o caso monoscópico, e estereoscópico com diferentes distâncias inter-axiais.....	112
---	-----

Tabela 8.2. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto no caso monoscópico.....	112
---	-----

Tabela 8.3. Tempos necessários à obtenção de imagens com diversas resoluções, para o caso monoscópico, e estereoscópico com diferentes distâncias inter-axiais.....	114
---	-----

Tabela 8.4. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto para uma imagem monoscópica.....	114
---	-----

Tabela 8.5. Tempos necessários à obtenção de imagens com diversas resoluções, para o caso monoscópico, e estereoscópico com diferentes distâncias inter-axiais.....	116
---	-----

Tabela 8.6. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto para uma imagem monoscópica.....	117
---	-----

Tabela 8.7. Tempos necessários à obtenção de imagens monoscópicas, e estereoscópicas com diferentes distâncias inter-axiais.....	120
--	-----

Tabela 8.8. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto para uma imagem monoscópica.....	121
---	-----

Tabela 8.9. Tempos necessários à obtenção de imagens com diversas resoluções, para o caso monoscópico, e estereoscópico com diferentes distâncias inter-axiais.....	122
---	-----

Tabela 8.10. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto para uma imagem monoscópica.....	123
--	-----

Índice de Figuras

Capítulo 2

Figura 2.1. Luz e Brilho.....	11
Figura 2.2. Tamanho Relativo.....	12
Figura 2.3. Interposição.....	12
Figura 2.4. Textura.....	12
Figura 2.5. Estereoscópio de Wheatstone.....	14
Figura 2.6. Paralaxe nula.....	16
Figura 2.7. Paralaxe positiva.....	17
Figura 2.8. Divergência associada à visualização de imagem com paralaxe superior à distância interocular.....	17
Figura 2.9. Paralaxe negativa.....	18

Capítulo 3

Figura 3.1. Utilização de visualização paralela.....	21
Figura 3.2. Imagem estereoscópica para observar com visualização paralela.....	21
Figura 3.3. Divergência provocada pela utilização de visualização paralela com imagens de grandes dimensões.....	22
Figura 3.4. Visualização cruzada.....	22
Figura 3.5. Imagem estereoscópica observada com visualização cruzada.....	23
Figura 3.6. Imagem estereoscópica para observar com visualização cruzada.....	23
Figura 3.7. Dispositivo de visualização utilizando barreira de paralaxe.....	24
Figura 3.8. Esquema de um panoramograma.....	25
Figura 3.9. Esquema de um dispositivo de visualização utilizando folha lenticular.....	25
Figura 3.10. Obtenção de imagens para dispositivos lenticulares.....	27
Figura 3.11. Projecção de imagens estereoscópicas, utilizando duas câmaras com polarizadores, e óculos polarizados para os assistentes.....	28
Figura 3.12. Organização do frame buffer no modo de visualização estereoscópico.....	31
Figura 3.13. Utilização do sistema CrystalEyes com recurso de um controlador apropriado.....	32
Figura 3.14. Utilização do sistema CrystalEyes com um computador preparado para estéreo.....	32
Figura 3.15. Interferência de ondas de duas fontes distintas.....	34
Figura 3.16. Obtenção de um holograma.....	35
Figura 3.17. Visualização de um holograma.....	35
Figura 3.18. Impressão de um vectógrafo.....	37
Figura 3.19. Polarização da luz, com a tinta de iodo.....	38
Figura 3.20. Vectógrafo estereoscópico.....	38
Figura 3.21. Imagem visível pelo olho esquerdo.....	38
Figura 3.22. Imagem visível pelo olho direito.....	39

Figura 3.23. Estereograma de pontos aleatórios. Para observar esta imagem, deve ser utilizada visualização cruzada.	41
---	----

Capítulo 4

Figura 4.1. Modelo de um sistema de visualização monoscópico.	45
Figura 4.2. Grelha ortogonal utilizada para estudo da distorção introduzida pelos sistemas ópticos.	46
Figura 4.3. Grelha ortogonal distorcida pelo sistema óptico.	46
Figura 4.4. Grelha ortogonal pré-distorcida pela função de correcção D^{-1} .	47
Figura 4.5. Imagem obtida pela pré-distorção de uma grelha ortogal, seguida de uma distorção.	48
Figura 4.6. Modelo óptico de um HMD estereoscópico.	48
Figura 4.7. Diagrama básico de um head-mounted display baseado em CRT's.	50
Figura 4.8. Arquitectura básica do Boom.	51
Figura 4.9. Diagrama básico de um head-mounted display por fibra óptica.	51

Capítulo 5

Figura 5.1. Obtenção da imagem de um ponto, por projecção paralela.	56
Figura 5.2. Obtenção da imagem de um ponto, por projecção perspectiva.	56
Figura 5.3. Utilização de pontos de observação distintos com direcções convergentes.	58
Figura 5.4. Utilização de um único ponto de observação, com objectos rodados.	58
Figura 5.5. Utilização de rotações.	58
Figura 5.6. Utilização de shears.	62
Figura 5.7. Utilização de dois pontos de observação distintos.	65
Figura 5.8. Utilização de um único ponto de observação e objectos deslocados lateralmente.	65
Figura 5.9. Modelo para obtenção de imagens estereoscópicas, utilizando câmaras paralelas.	66
Figura 5.10. Modelo para obtenção de imagens estereoscópicas, utilizando câmaras paralelas.	66
Figura 5.11. Modelo para obtenção de imagens estereoscópicas, utilizando câmaras paralelas.	67
Figura 5.12. Modelo para obtenção de imagens estereoscópicas, utilizando câmaras paralelas.	68

Capítulo 6

Figura 6.1. Visualização de um par estereoscópico de imagens num ecrã plano.	71
Figura 6.2. Visualização de um par estereoscópico de imagens num ecrã plano.	72
Figura 6.3. Visualização de um par estereoscópico de imagens num ecrã plano.	73
Figura 6.4. Sensação de perspectiva inversa devido à ausência de perspectiva. Para observar esta imagem, utilizar visualização cruzada.	76

Capítulo 7

Figura 7.1. Comparação do realismo das diversas imagens.	88
Figura 7.2. Aumento do realismo provocado pela adição do efeito estereoscópico para cada tipo de imagem.	89
Figura 7.3. Variação do ângulo de convergência com a distância interocular.	93
Figura 7.4. Distorção por variação da IPD.	94
Figura 7.5. Distorção causada pela variação da distância de observação.	94
Figura 7.6. Distorção causada pelo movimento lateral da cabeça.	95
Figura 7.7. Utilização de pré-distorção.	96
Figura 7.8. Problemas devidos à pré-distorção apenas dos vértices.	97
Figura 7.9. Objecto cortado pelos limites da imagem.	98

Capítulo 8

Figura 8.1. Geração de objectos adicionais por translação.....	105
Figura 8.2. Objectos pais, e objectos filhos.....	106
Figura 8.3. Armazenamento das descrições dos objectos, partilhando informação sobre material.	106
Figura 8.4. Armazenamento das descrições dos objectos, com partilha de informação geométrica e de informação sobre o material.....	107
Figura 8.5. Obtenção de dois pontos de intersecção para cada raio.....	108
Figura 8.6. Obtenção de dois pontos de intersecção em objectos distintos para cada raio.....	108
Figura 8.7. Árvore de shading completa, obtida para um raio estereoscópico.	110
Figura 8.8. Imagem estereoscópica obtida pelo algoritmo apresentado.	111
Figura 8.9. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto no caso monoscópico.....	112
Figura 8.10. Imagem estereoscópica obtida utilizando cortes na árvore de objectos.	113
Figura 8.11. Erro introduzido pela aplicação de cortes na árvore de objectos.	114
Figura 8.12. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto no caso monoscópico.....	115
Figura 8.13. Aplicação de cortes na árvore de shading.....	116
Figura 8.14. Imagem estereoscópica obtida utilizando cortes na árvore de shading.	116
Figura 8.15. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto no caso monoscópico.....	117
Figura 8.16. Reenvio de raios à origem.....	118
Figura 8.17. Reflexão difusa.....	118
Figura 8.18. Reflexão especular.....	118
Figura 8.19. Reenvio de raios à origem, utilizando um ponto de observação único, e objectos transladados.	119
Figura 8.20. Imagem estereoscópica obtida utilizando reenvio de raios à origem.	120
Figura 8.21. Erro introduzido pela utilização de reenvio de raios à origem.	120
Figura 8.22. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto no caso monoscópico.....	121
Figura 8.23. Imagem estereoscópica obtida utilizando coerência de imagem.....	122
Figura 8.24. Erro introduzido pela utilização de coerência de imagem.....	122
Figura 8.25. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto no caso monoscópico.....	123

Capítulo 1.

Introdução

Nos últimos anos, com a proliferação dos sistemas gráficos computadorizados, tem sido intensa a procura de novos dispositivos e métodos de apresentação de imagens tridimensionais cada vez mais realistas e ricas em informação geométrica. As elevadas resoluções espaciais e cromáticas dos ecrãs actuais permitem visualizar imagens com qualidade quase fotográfica, fornecidas por poderosos sistemas computacionais. Apesar de serem intrinsecamente planas, as imagens mais ou menos realistas comuns possuem determinadas qualidades, que permitem ao seu observador ajuizar a distância a que se encontra dos objectos. No entanto, a sensação de profundidade obtida pela observação de uma destas imagens monoscópicas, ou de uma fotografia tradicional, fica bastante aquém daquela que é obtida no mundo visual. Neste, utilizamos uma outra fonte de informação de profundidade - a estereoscopia. A sensação estereoscópica de profundidade deve-se ao facto de possuímos dois olhos que vêem o mundo de pontos diferentes.

A geração de uma imagem estereoscópica envolve a obtenção de duas vistas, a partir de dois pontos de observação distintos. A geometria básica do algoritmo gerador de imagens estereoscópicas pode ser facilmente explicada utilizando duas câmaras, ou utilizando uma única câmara e aplicando aos objectos transformações distintas para cada uma das vistas. A experiência demonstra que introduzir o efeito estereoscópico a imagens geradas por computador não é uma tarefa muito complicada. A obtenção de uma imagem monoscópica por computador a partir de uma base de dados tridimensional, requer o cálculo de uma projecção perspectiva num ponto único. A obtenção de uma imagem estereoscópica por computador a partir de uma base de dados tridimensional, requer o cálculo de duas projecções perspectivas em pontos distintos.

Apesar de se basear normalmente na simulação de fenómenos físicos, ou pelo menos geométricos, a geração de imagens estereoscópicas é, como toda a criação de imagem, uma arte. Mesmo depois de várias centenas de anos de utilização de estereoscopia, o método ou algoritmo para criação correcta de imagens estereoscópicas não é óbvio, e é por vezes de implementação delicada. No entanto, graças à maior versatilidade que os meios computadorizados apresentam sobre os outros meios de geração de imagem (por exemplo: fotografia), não se devem tolerar imagens estereoscópicas geradas por computador que não sejam simultaneamente agradáveis e de visualização confortável.

Este trabalho pode interessar a todos os que pretendem conceber ou adquirir um sistema de visualização estereoscópica. Aqui encontrarão a descrição das vantagens e

desvantagens na utilização da estereoscopia, os métodos e os dispositivos disponíveis. Podem ainda encontrar nesta tese a descrição de um novo método de geração de imagens estereoscópicas realistas, baseado no algoritmo de Ray-Tracing.

1.1. Trabalhos Anteriores

Existem vários métodos publicados para calcular as imagens direita e esquerda de um par estereoscópico. Em [Grotch83] e em [Wixson83], descrevem-se métodos para criação de pares estereoscópicos de imagens que se baseiam em rotações. Estes métodos, e as distorções que introduzem, foram estudados em [Hodges90b]. Em [Lipscom79] propõe-se a aproximação das rotações por deformações (shears). Vários autores [Akka91], [Southar92], [Penna86], aconselham a utilização de um plano de projecção único para a obtenção das duas imagens, ou a utilização de translações. Estas aproximações derivam do algoritmo das lentes com eixos paralelos, que foi originalmente apresentado em [Baker87]. Mais recentemente foi apresentada uma alternativa para geração de imagens para dispositivos Visualizadores Montados na Cabeça (HMD's) [Robinet91], baseada em projecções descentradas cujos eixos podem não ser paralelos. Em [Hodges88] e em [Hodges90a], encontramos excelentes descrições e análises dos principais métodos publicados.

Existem também várias tecnologias disponíveis para visualizar imagens tridimensionais. Em [Kasper87] e em [Bains93] encontramos descrições das metodologias de concepção de hologramas. Descrições mais ou menos gerais dos sistemas de visualização e apresentação estereoscópica disponíveis, e das tecnologias utilizadas, podem ser encontrados em [Okoshi76], [Lipton82], [Allister91] e em [Love90]. Um resumo dos principais dispositivos do tipo HMD disponíveis no mercado foi recentemente apresentado [King93]. Outras técnicas de apresentação de imagens estereoscópicas menos utilizadas estão descritas em [Jones90] e em [Sher90].

Não são conhecidos trabalhos relevantes anteriores sobre geração de imagens estereoscópicas, utilizando o algoritmo de Ray-Tracing. No entanto, existem numerosos trabalhos sobre optimização do algoritmo de Ray-Tracing [MLeitão93], [Ward91], [Glassne89], cujas ideias, sujeitas a pequenas adaptações, podem servir de base à implementação de um Ray-Tracer estereoscópico eficiente.

1.2. Organização da tese

Neste trabalho propomo-nos descrever e analisar os diversos modelos de geração de imagens estereoscópicas, e tirar conclusões sobre a correcção destes. Vamos estudar o problema para duas situações típicas de visualização: o caso da apresentação de ambas as imagens num mesmo ecrã que possa ser considerado plano; e o caso dum dispositivo tipo HMD.

No capítulo 2, vamos analisar como as pessoas vêem e se apercebem do mundo tridimensional. Vamos ainda, neste capítulo, definir alguns termos como *paralaxe*, *distância interaxial* e *disparidade retinal*, que serão utilizados mais tarde.

No capítulo 3, vamos descrever o funcionamento de vários dispositivos de visualização estereoscópica. Os tipos de imagem a utilizar em cada dispositivo, e os seus métodos de armazenamento são também tratados neste capítulo. Devido à sua especificidade e grande variedade de arquitectura, os dispositivos de montagem na cabeça (HMD) foram considerados separadamente no capítulo 4.

No capítulo 5, analisaremos, de uma forma que nos parece ir mais longe do que os estudos publicados anteriormente, os diversos modelos conhecidos, que podem ser utilizados para gerar imagens estereoscópicas. A análise do processo de visualização sob condições de visualização diversas, e o estudo da adequação de cada um dos modelos de geração será o tema do capítulo 6. Neste capítulo vamos eleger o modelo e os parâmetros de visualização considerados ideais para as condições de visualização consideradas.

No capítulo 7, serão tratados outros aspectos relacionados com a qualidade e os defeitos mais comuns das imagens e animações estereoscópicas, sintetizadas por computador.

No capítulo 8, será descrito o algoritmo base e a implementação realizada de um método de geração de imagens estereoscópicas realistas. O método desenvolvido, baseado no algoritmo de Ray-Tracing, implementa o modelo de geração de imagens estereoscópicas considerado correcto no capítulo 6, e tenta satisfazer os diversos requisitos aconselhados também no capítulo 7.

Infelizmente, não é possível apresentar neste tese imagens verdadeiramente estereoscópicas. Para que o leitor possa ajuizar a qualidade do algoritmo desenvolvido, são apresentados os pares estereoscópicos de imagens obtidos. Estes pares de imagens podem ser visualizados em estereoscopia utilizando visualização paralela (ver 3.1.1.). São ainda apresentadas as imagens de erro, correspondentes à diferença entre as imagens obtidas pelo programa implementado e as obtidas pelo algoritmo tradicional de Ray-Tracing. Todas estas imagens estão disponíveis publicamente por FTP, em [asterix.inescn.pt](http://asterix.inescn.pt/pub/academic/stereo_rt), no directório `/pub/academic/stereo_rt`. O leitor pode assim obter as imagens na sua forma electrónica, e visualizá-las com o auxílio de qualquer dispositivo de visualização estereoscópica baseado em computador.

Capítulo 2.

Informação de Profundidade

Uma imagem bidimensional cria ao seu observador a ilusão de profundidade da cena visualizada, através da apresentação de várias pistas de percepção de profundidade [Kendler74]. As pistas de informação de profundidade monoculares são a base para a percepção de distância em qualquer dispositivo de visualização. Estas pistas são pelo menos tão importantes como o efeito estereoscópico, na criação de imagens que possam ser consideradas na realidade tridimensionais. Todas as pistas monoculares de profundidade podem e devem existir nas imagens estereoscópicas. Uma imagem, com informação de profundidade rica, é mais fácil de perceber quando lhe é acrescentado o efeito estereoscópico.

Neste capítulo vamos descrever as pistas de informação de profundidade e analisar a sua influência na forma como uma cena tridimensional é entendida. Vamos também explicar a criação da sensação estereoscópica de profundidade.

2.1. Pistas Monoculares

2.1.1. Luz e Brilho

A luz e o brilho, ou a ausência destes, fornecem uma informação básica de profundidade. Os artistas sabem bem como criar objectos que pareçam planos ou curvos, simplesmente adicionando efeitos luminosos. A luz e o brilho de uma superfície depende da sua orientação relativamente às fontes luminosas. O conhecimento da orientação das superfícies permite o entendimento da profundidade de um objecto.

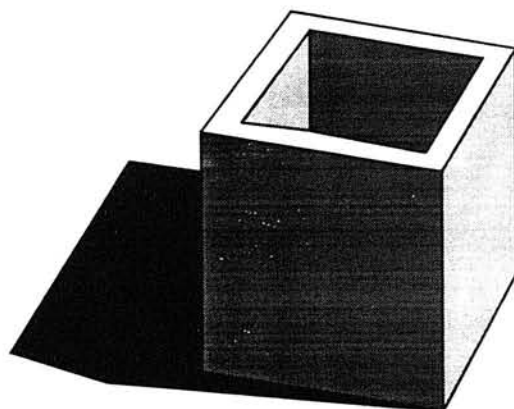


Figura 2.1. Luz e Brilho.

2.1.2. Tamanho relativo

É sabido que os objectos parecem-nos maiores quando vistos de perto, e mais pequenos quando vistos de longe. A memória ajuda-nos a optar pela distância a que se encontram objectos que nos são familiares. Um objecto cujo tamanho seja conhecido, e que se apresente com pequenas dimensões, é considerado como estando longe.

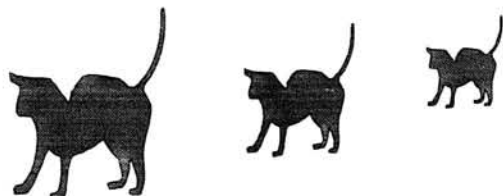


Figura 2.2. Tamanho Relativo.

2.1.3. Interposição

A *interposição* é uma das pistas mais óbvias e é normalmente aquela que é tomada para garantia ou prova das posições relativas dos objectos. Nós entendemos que o livro para onde estamos a olhar está mais perto de nós do que a secretária, porque não a podemos ver através do livro. O livro está interposto entre nós e a secretária, ou entre nós e os objectos que estão mais longe.

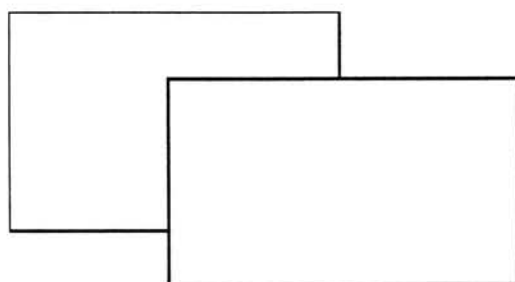


Figura 2.3. Interposição.

2.1.4. Textura

A *textura* é, das pistas monoculares de distância, aquela cuja utilização é mais recente. Todas as outras são conhecidas e utilizadas pelos artistas desde a época da Renascença. Um material com *textura* permite a sensação de profundidade, porque a *textura* é tanto mais visível quanto mais próximo de nós o objecto se encontrar.

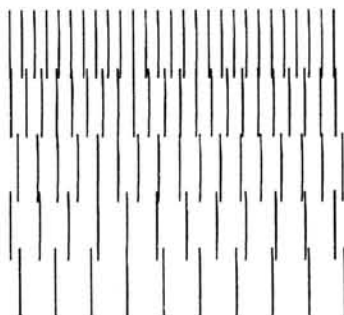


Figura 2.4. Textura.

2.1.5. Paralaxe de movimento

A *paralaxe de movimento* pode ser utilizada no mundo visual ou em ecrãs com imagens animadas, e pode ser obtida, por exemplo, pela rotação de um objecto. Um exemplo familiar da utilização desta pista é o observador que viaja num comboio. Toda a gente já reparou que os postes ou as árvores próximas parecem deslocar-se com velocidade superior à das montanhas distantes.

A *paralaxe de movimento* não pode ser obtida em todas as composições gráficas, como na pintura ou em imagens estáticas. No entanto, o uso desta pista de informação pode produzir um bom efeito de profundidade, independentemente da orientação espacial e do eixo de rotação. Esta independência axial indica que a *paralaxe de movimento* é independente do efeito estereoscópico. Se as duas pistas estivessem relacionadas, só a rotação em torno de um eixo vertical (ou translação na horizontal) forneceria informação de distância. Na realidade vemos que até a rotação em torno de um eixo horizontal permite a sensação de distância desejada. Esta explicação justifica-se porque há quem pense que a estereoscopia e a paralaxe de movimento são manifestações da mesma entidade. A *paralaxe de movimento* (por exemplo, a rotação de um objecto) é uma boa pista para utilizar juntamente com a estereoscopia porque, como acontece com algumas das outras pistas monoculares, faz acentuar o efeito estereoscópico.

2.1.6. Perspectiva aérea

A *perspectiva aérea* é a diminuição da visibilidade com o aumento da distância. Isto pode acontecer devido à névoa entre o observador e o objecto observado. No caso de nevoeiro intenso, não seriam visíveis os objectos que se encontrassem a uma grande distância do observador.

2.1.7. Perspectiva

A *perspectiva* é a relação entre os objectos em primeiro plano e os objectos no fundo. De entre as pistas de informação de distância não estereoscópicas, a *perspectiva* é a mais importante para utilização em sistemas binoculares, porque escala a sensação estereoscópica de profundidade. As imagens geradas por computador, com perspectiva intensa, parecem mais profundas e podem ser mais facilmente entendidas. Se a perspectiva for exagerada, a profundidade da imagem é aumentada.

2.1.8. Depth Cuing

Depth Cuing é o nome de uma técnica disponível em estações gráficas de trabalho. É normalmente implementada simplesmente pela utilização de uma biblioteca gráfica. A utilização desta técnica provoca uma redução da intensidade da iluminação do objecto em função da distância ao ponto de observação. A redução gradual do brilho simula conjuntamente o efeito de perspectiva aérea e a pista da luz e do brilho.

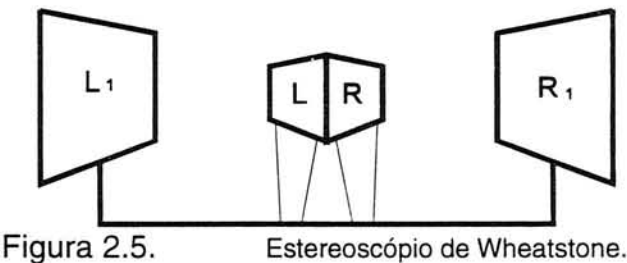
2.1.9. Acomodação

A *acomodação* ou *focagem* é a tensão muscular necessária para ajustar a distância focal de cada um dos olhos a um objecto no espaço. Nas imagens planas, como a distância ao ponto de observação é constante para toda a cena, a acomodação é fixa. A informação de profundidade proveniente desta pista, está portanto em conflito com as restantes pistas de informação de profundidade.

Este conflito entre a acomodação e as outras pistas de informação de profundidade vai continuar a existir, quer em imagens monoscópicas quer em imagens estereoscópicas, enquanto não existirem dispositivos de visualização capazes de apresentar imagens com acomodação variável.

2.2. Estereoscopia

Até há bem pouco tempo, não era utilizada a *estereoscopia* como fonte de informação de profundidade. Esta fonte de informação, difere de todas as outras já referidas, pois trata-se de uma fonte de informação binocular. Ao longo dos anos, homens como Leonardo de Vinci, entenderam que cada um dos olhos via imagens diferentes. Mas foi Wheatstone quem, em 1838, explicou ao mundo, com o seu estereoscópio (figura 2.5) e as suas imagens, que existe uma sensação de profundidade única, a estereoscopia, produzida pela disparidade na retina. Kepler e outros questionaram a razão de não vermos uma imagem duplicada do mundo. Wheatstone explicou que o problema era de facto a solução, demonstrando que a mente humana funde as duas imagens planas obtidas na retina numa só, com estereoscopia.



O estereoscópio de Wheatstone é constituído por dois espelhos rectangulares (L e R) com uma aresta comum, mas formando um pequeno ângulo entre eles. O observador deve encostar o nariz à zona onde os espelhos se unem. Desta maneira, o olho direito vê a imagem colocada na posição R₁ e o olho esquerdo vê a imagem colocada na posição L₁. Wheatstone usava desenhos ou pinturas colocados nas posições L₁ e R₁ para obter o efeito pretendido.

Um dispositivo de visualização estereoscópico, é um sistema óptico cujo elemento final é a mente humana. Este dispositivo funciona pela apresentação, à mente humana, de imagens do mesmo tipo daquelas que os olhos esquerdo e direito vêem no mundo visual.

A estereoscopia fornece, de facto, duas pistas de informação de profundidade: A convergência e a disparidade retinal.

Com base num trabalho com estereogramas de pontos aleatórios, em [Julesz71] classifica-se a disparidade retinal como sendo a fonte de informação mais poderosa na determinação de profundidade relativa. Outros autores, reconhecem igualmente a disparidade retinal como uma importante pista de informação de profundidade, mas fazem notar que a prioridade atribuída a cada uma das possíveis pistas conflituosas, varia de observador para observador. No entanto, é habitualmente aceite que a acomodação e a convergência são pistas de informação de profundidade menos importantes do que a disparidade retinal. A acomodação e a convergência influenciam a percepção de profundidade, quando utilizadas em conjunto com a disparidade retinal, sobretudo para objectos colocados a menos de dois metros do observador.

2.2.1. Convergência

A *convergência* é o movimento coordenado de rotação para dentro dos dois olhos, para focar um objecto no espaço. O ângulo formado pelos dois raios que unem os olhos ao objecto focado, designa-se por *ângulo de convergência*.

Quando se observa uma imagem monoscópica plana, os olhos convergem no plano da imagem. A informação de profundidade resultante é portanto fixa, em concordância com a acomodação, mas em conflito com as outras pistas monoculares de informação de profundidade. Na visualização de uma imagem estereoscópica, a convergência deve estar de acordo com todas as outras fontes de informação de profundidade, com excepção da acomodação que ainda não é possível controlar.

Para que a pista da convergência seja correcta, é necessário utilizar um dispositivo de visualização estereoscópica e colocar correctamente as imagens no ecrã. Uma imagem estereoscópica observada por free-viewing não fornece correctamente a pista da convergência.

2.2.2. Disparidade Retinal

A *disparidade retinal* pode ser facilmente demonstrada com uma experiência simples. Vamos pôr um dedo em frente dos olhos, a uma distância de aproximadamente 30 cm. Quando olhamos para o dedo, os nossos olhos convergem no dedo. Isto quer dizer que os eixos ópticos, de ambos os olhos, se cruzam no dedo. Há vários músculos que são usados para mover os nossos olhos de maneira a centrar a imagem em cada uma das retinas. Se mantivermos os olhos a convergir no dedo, enquanto prestamos atenção a um objecto do fundo, verificamos que o objecto do fundo aparece duplicado. Da mesma maneira, quando estamos a olhar para o objecto do fundo, os olhos convergem nele. O dedo aparecerá agora em duplicado.

Se fosse possível captar as imagens que aparecem em cada uma das retinas e sobrepor-las uma à outra como se de dois diapositivos se tratasse, veríamos duas imagens que quase coincidiam. A *disparidade* é a distância, segundo uma direcção horizontal, entre pontos correspondentes das imagens da retina sobrepostas. Os pontos das imagens da retina, pertencentes ao objecto onde os olhos convergem, têm *disparidade nula*.

A *disparidade retinal* é causada pelo facto de cada um dos nossos olhos ver o mundo de um ponto diferente. Nos adultos, os olhos estão distanciados em média 64 milímetros um do outro. A disparidade é processada pelo cérebro numa única imagem do mundo visual. A capacidade da mente humana para combinar duas imagens diferentes, apesar de similares, é chamada *fusão*. A sensação de profundidade resultante é chamada *estereoscopia*.

2.2.3. Paralaxe

Um ecrã estereoscópico difere de um ecrã plano, unicamente no facto de ser capaz de apresentar os valores de paralaxe dos pontos de imagem. A *paralaxe* produz, como veremos, disparidade nas imagens formadas em cada um dos olhos, permitindo assim a sensação de estereoscopia.

A paralaxe também pode ser medida em termos de um ângulo de visão. Se as projecções para olho esquerdo e direito aparecem no ecrã separadas por uma distância R , e o observador está colocado a uma distância d do ecrã, o ângulo de paralaxe será:

$$\beta = 2 \cdot \arctan \frac{R}{2 \cdot d} \quad (2.1)$$

Dependendo do seu valor, a paralaxe pode ser classificada em quatro tipos diferentes. O caso da *paralaxe nula* (figura 2.6) é aquele em que os pontos homólogos das duas imagens estão exactamente sobrepostos. Neste caso, o ângulo de paralaxe é também nulo.

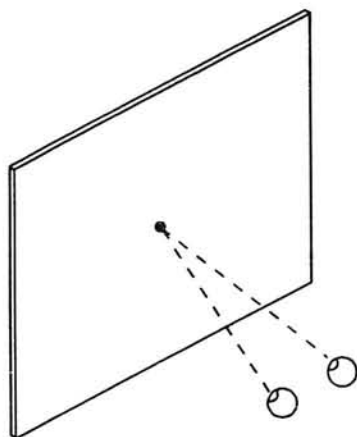


Figura 2.6. Paralaxe nula.

Quando os olhos do observador, distanciados um do outro por uma determinada distância (distância interocular), observam uma imagem com paralaxe nula, os olhos convergem no plano do ecrã. Por outras palavras, os eixos ópticos dos olhos cruzam-se no plano do ecrã. Quando os pontos da imagem têm paralaxe nula, podemos dizer que se trata de um caso de *zero paralaxe setting* (ZPS).

A figura 2.7. ilustra um tipo de paralaxe positiva. Qualquer valor de paralaxe positivo, entre zero e a distância interocular, faz a imagem aparecer do lado de trás do ecrã.

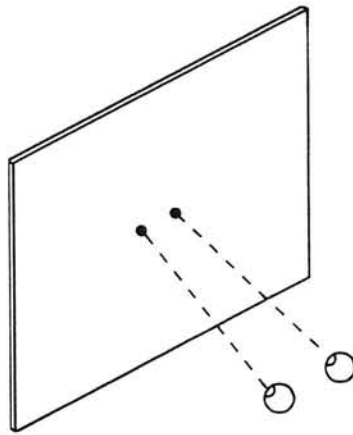


Figura 2.7. Paralaxe positiva.

Um caso particular de paralaxe positiva é aquele em que os eixos de ambos os olhos são paralelos. Isto acontece no mundo visual, quando olhamos para objectos colocados a uma grande distância do observador. Num sistema de visualização estereoscópico, quando a distância entre os olhos é igual ao valor da paralaxe, teremos também os eixos dos olhos paralelos. Veremos mais tarde que em alguns dispositivos de visualização, a utilização de valores de paralaxe muito próximos da distância interocular, produz desconforto para o observador.

Outro tipo de paralaxe positiva é mostrado na figura 2.8. Neste caso, o valor da paralaxe é superior à distância interocular, e portanto os eixos dos olhos são divergentes. A divergência nunca acontece quando olhamos para objectos existentes no mundo visual.

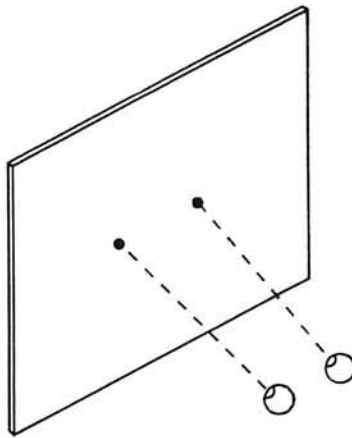


Figura 2.8. Divergência associada à visualização de imagem com paralaxe superior à distância interocular.

Para que os nossos olhos possam fundir imagens com paralaxe superior à distância entre os olhos, é necessário um esforço muscular anormal, o que pode causar desconforto. Não existem razões válidas para a utilização de divergência em imagens estereoscópicas geradas por computador.

Um exemplo de *paralaxe negativa* ou *cruzada* é apresentado na figura 2.9.

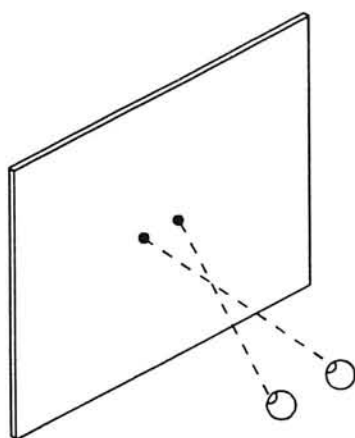


Figura 2.9. Paralaxe negativa.

Neste caso, os eixos dos olhos cruzam-se entre o observador e o ecrã. Os objectos com paralaxe negativa aparecem à frente do plano do ecrã, ou seja, entre este e o observador.

2.2.4. Separação Interaxial

Chama-se *separação interaxial* à distância entre os eixos das lentes utilizadas para obter uma fotografia estereoscópica. O eixo de uma lente é a linha que passa pelo centro óptico da lente e é perpendicular ao plano da imagem. O conceito é o mesmo, quer se trate de lentes reais ou de modelos de geração de imagem por computador.

A paralaxe e, portanto, o efeito estereoscópico, são devidos à existência da separação interaxial. Se as lentes estão muito próximas uma da outra, o efeito estereoscópico de profundidade é reduzido. No caso limite das duas lentes estarem sobrepostas, resulta uma imagem plana. A paralaxe, e portanto o efeito estereoscópico, aumenta quando se afastam as lentes uma da outra.

2.3. Conclusão

As pistas de informação de profundidade monoculares existem tanto nas imagens sintetizadas por computador como no mundo visual. Apesar dos ecrãs tradicionais não poderem fornecer a pista estereoscópica de profundidade, podem apresentar imagens que se assemelham a vistas do mundo tridimensional, utilizando correctamente as pistas monoculares de informação de profundidade. Todas estas pistas, especialmente a perspectiva, podem e devem existir nas imagens estereoscópicas. Uma imagem, com informação de profundidade rica, é mais fácil de perceber quando lhe é acrescentado o efeito estereoscópico.

Capítulo 3.

Apresentação de imagens tridimensionais

Neste capítulo vamos descrever o funcionamento de vários dispositivos de visualização estereoscópica. Os tipos de imagem a utilizar em cada dispositivo, e os seus métodos de armazenamento e difusão serão também tratados. Devido à sua especificidade e grande variedade de arquitectura, os dispositivos de montagem na cabeça (HMD) serão considerados separadamente no capítulo seguinte.

Existem várias tecnologias disponíveis para visualizar imagens tridimensionais. Em [Kasper87] e em [Bains93] encontramos descrições das metodologias de concepção de hologramas. Descrições mais ou menos gerais dos sistemas de visualização e apresentação estereoscópica disponíveis, e das tecnologias utilizadas, podem ser encontrados em [Okoshi76], [Lipton82], [Allister91] e em [Love90]. Um resumo dos principais dispositivos do tipo HMD disponíveis no mercado foi recentemente apresentado [King93]. Outras técnicas de apresentação de imagens estereoscópicas menos utilizadas estão descritas em [Jones90] e em [Sher90].

Vamos começar por resumir as possíveis aproximações para apresentação de imagens tridimensionais, e dividi-las em várias categorias (tabela 3.1). Todos os métodos actualmente disponíveis para imagens digitais, incluindo os métodos holográficos, utilizam pares estereoscópicos ou fatias de imagens.

No nível mais alto, estas técnicas podem ser divididas em “visualização sem ajuda”, ou autoestereoscópicas, e “visualização com ajuda”, na qual é necessário utilizar um dispositivo dedicado para entregar a imagem correcta ao observador. As técnicas autoestereoscópicas podem depois ser divididas nas categorias holográfica e não holográfica. As técnicas de visualização que necessitam de um dispositivo dedicado, são divididas nas categorias dinâmica e estática. As imagens dinâmicas permitem a visualização do movimento.

Representação de imagens tridimensionais

Autoestereoscópicas		Utilizando dispositivo de visualização	
Holográfica	Não-Holográfica	Estática	Dinâmica
Completamente gerada	Free-View	Estereoscópios	Multiplexagem temporal
Multiplexada	Barreira de Paralaxe		Polarização
	Ecrã lenticular		Efeito Pulfrich
			Anáglifos

Tabela 3.1.

Classificação dos métodos de apresentação de imagens tridimensionais.

3.1. Técnicas de Visualização de Imagens Tridimensionais

3.1.1. Free Viewing

O método mais simples consiste em utilizar uma impressora gráfica ou uma plotter para produzir as duas imagens colocadas lado a lado. O par estereoscópico pode então ser visto sem ajuda de nenhum sistema óptico, utilizando uma técnica que se chama *free viewing*. Nesta técnica, cada um dos olhos olha para uma das imagens do par.

Existem dois métodos de se obter a técnica de free viewing: visualização paralela, e visualização transversal ou cruzada.

3.1.1.1. Visualização Paralela

Na visualização paralela, o observador vê a imagem colocada do lado direito com o seu olho direito, e a imagem colocada do lado esquerdo com o seu olho esquerdo (figura 3.1.).

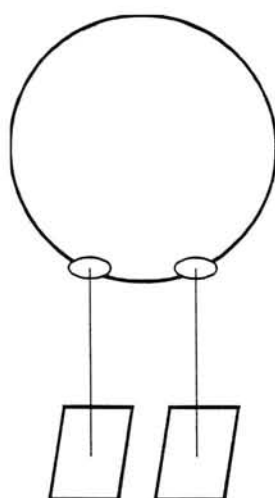


Figura 3.1.

Utilização de visualização paralela.

O cubo da figura 3.2. pode ser visto em estereoscopia utilizando visualização paralela.

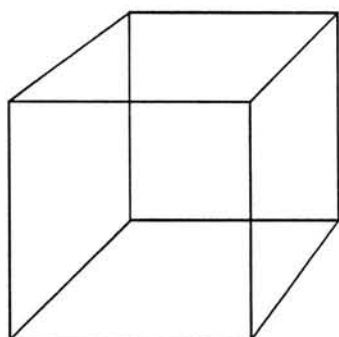


Figura 3.2.

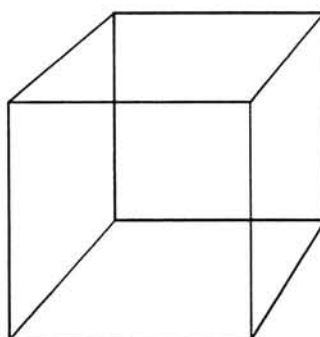


Imagem estereoscópica para observar com visualização paralela.

Para observar esta imagem utilizando visualização paralela, o observador deve estar preferencialmente colocado a cerca de 20 cm de distância. Por vezes é conveniente colocar uma folha de papel entre as perspectivas, de tal maneira que cada olho só possa ver a imagem do lado respectivo.

A utilização de visualização paralela com imagens de largura superior a metade da distância interocular não é aconselhável, pois provoca a divergência das linhas de visão (figura 3.3.). Como esta situação nunca acontece no mundo real, é necessário um esforço muscular anormal para conseguir a divergência. Para além da dificuldade sentida ao observar a imagem tridimensional, a divergência provoca desconforto.

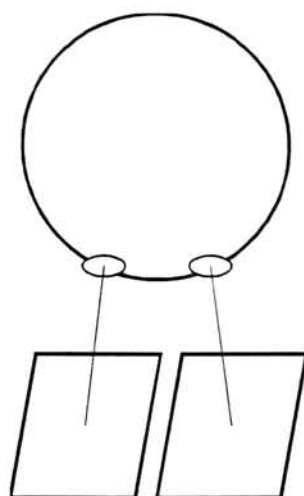


Figura 3.3.

Divergência provocada pela utilização de visualização paralela com imagens de grandes dimensões.

3.1.1.2. Visualização Cruzada

No caso da visualização cruzada, a imagem colocada no lado esquerdo deve ser vista com o olho direito, e a imagem do lado direito deve ser vista com o olho esquerdo (figura 3.4.).

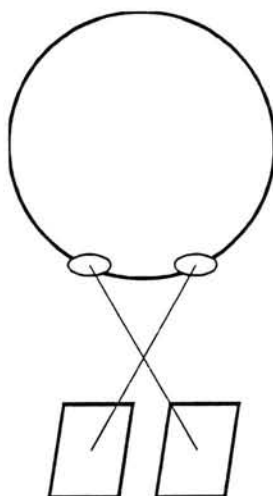


Figura 3.4. Visualização cruzada

Este método é normalmente mais fácil de utilizar e de aprendizagem mais rápida para a maioria das pessoas. Este método tem também a vantagem de permitir visualizar imagens cuja largura ultrapassa metade da distância interocular ($\cong 3.2$ cm). O aumento da largura das imagens apenas altera a distância de cruzamento (convergência) das linhas de visão (figura 3.5.).

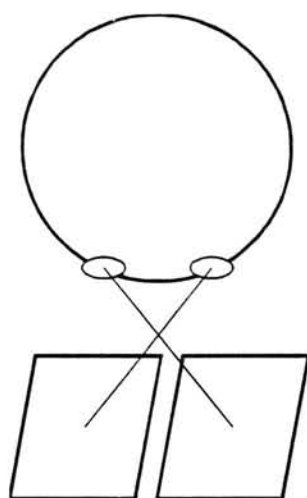


Figura 3.5. Imagem estereoscópica observada com visualização cruzada.

O cubo da figura 3.6. pode ser visto em estereoscopia utilizando visualização cruzada.

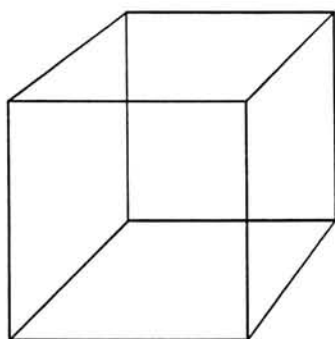


Figura 3.6.

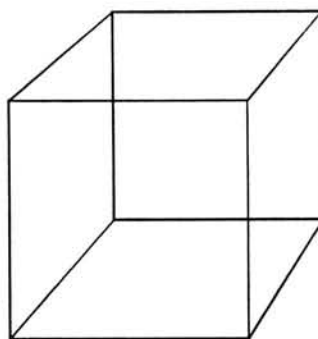


Imagem estereoscópica para observar com visualização cruzada.

Para ajudar a ver a figura utilizando visualização cruzada, pode-se segurar um lápis, de tal maneira que este fique paralelo ao papel e que a ponta fique a meia distância entre o observador e o plano das imagens. Ajusta-se a posição do lápis de tal maneira que, com o olho direito fechado, a ponta do lápis coincida com um ponto no centro da imagem da direita, e que, com o olho esquerdo fechado, a ponta coincida com o mesmo ponto na imagem da esquerda. Depois de posicionado, observa-se a ponta do lápis com ambos os olhos. Agora podemos relaxar os olhos suavemente para prestar atenção à imagem que aparece por trás do lápis. Poderemos então ver uma imagem tridimensional, entre as duas imagens perspectivas planas.

3.1.2. Métodos de Selecção de Imagem

Com excepção para o método de Free-Viewing, descrito anteriormente, todas as técnicas de visualização de imagens estereoscópicas necessitam de hardware dedicado. No entanto, é sempre desejável conseguir visualizar imagens estereoscópicas sem necessidade de utilizar um dispositivo especial para cada observador (por exemplo óculos). Os processos que permitem isso chamam-se autoestereoscópicos. Três métodos autoestereoscópicos são vulgarmente utilizados: a barreira de paralaxe, a folha lenticular e a holografia. Os métodos holográficos serão tratados mais tarde.

3.1.2.1. Barreiras de Paralaxe

Os métodos baseados em barreiras de paralaxe são conhecidos desde o princípio deste século. Os mais conhecidos são o *estereograma de paralaxe* e o *panoramograma de paralaxe*.

Um *estereograma de paralaxe* consiste numa barreira em forma de grelha com sulcos verticais finos, que é colocada em frente de uma imagem previamente preparada para o efeito. Esta imagem deverá conter as vistas, destinadas aos olhos esquerdo e direito, em colunas verticais alternadas. A grelha e a imagem com as perspectivas devem ser colocadas de tal maneira que, com o utilizador colocado na devida posição, cada um dos olhos apenas veja a respectiva perspectiva do par estereoscópico (figura 3.7.).

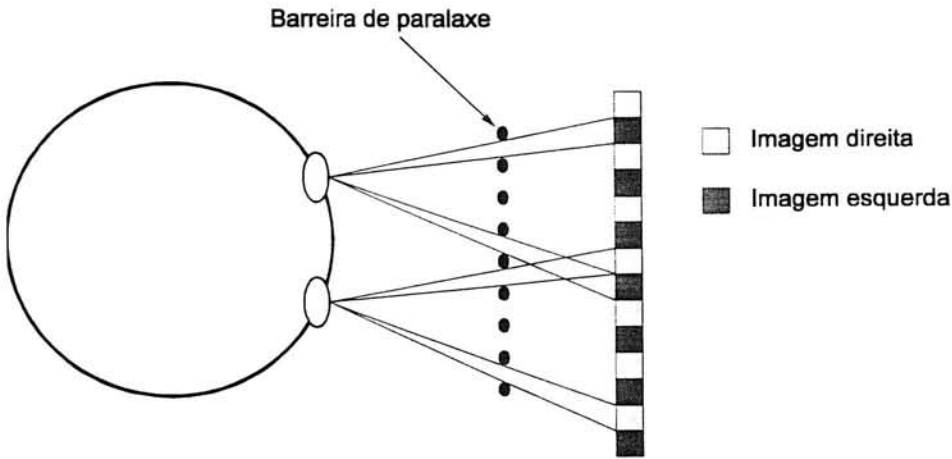


Figura 3.7. Dispositivo de visualização utilizando barreira de paralaxe.

Num *panoramograma*, a distância entre as falhas (sulcos) da grelha é maior, e por isso é possível guardar mais colunas de imagem atrás da grelha. A largura das falhas da grelha não é alterada, para que de cada falha só seja possível observar uma coluna (figura 3.8.). Assim a coluna visível através de cada falha depende da posição do olho e da posição do observador.

Diversos problemas têm impedido as técnicas de barreira de paralaxe de se tornarem mais populares na implementação de sistemas autoestereoscópicos. A maior desvantagem relaciona-se com a falta de brilho resultante, devido à quantidade de luz bloqueada pela barreira. Outro problema pode ser a falta de nitidez, resultante do posicionamento incorrecto da barreira de paralaxe em relação às tiras da imagem. Normalmente as tiras da imagem são muito finas, para não serem notadas pelo observador. Mas tiras excessivamente finas provocam difracção, que também prejudica a nitidez da imagem tridimensional.

Uma característica dos panoramogramas que causa desconforto é o salto da imagem, que se sente quando o observador se desloca, passando de uma perspectiva para a seguinte. As várias perspectivas dos objectos colocados longe do observador aparecem em posições largamente espaçadas no panoramograma, e então o objecto parece saltar de uma posição para outra, quando o ponto de observação é deslocado. Se for possível incluir um número de perspectivas extremamente elevado, o salto torna-se razoavelmente pequeno. No entanto, para não permitir a sobreposição das diferentes

perspectivas, o números destas tem de ser mantido razoavelmente baixo. O efeito desagradável do salto dos objectos pode ser minimizado, impondo uma profundidade de campo pequena.

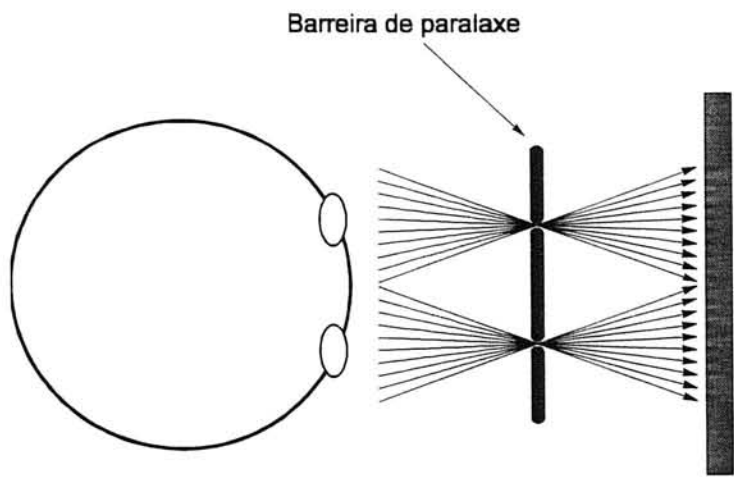


Figura 3.8. Esquema de um panoramograma.

3.1.2.2. Folha Lenticular

Uma folha lenticular é uma série de lentes cilíndricas, normalmente revestida de plástico. Esta folha é projectada de tal maneira, que dois raios luminosos paralelos, que entram pela frente, convergem num ponto da superfície de trás da folha. Para visualizar imagens tridimensionais, coloca-se um ecrã plano com as devidas dimensões encostado à parte de trás das lentes (figura 3.9.). Cada uma das lentes deverá ter atrás uma tira da imagem. Cada um dos olhos será focado pelas lentes numa tira da imagem diferente.

Os dispositivos que utilizam barreiras de paralaxe são normalmente considerados de qualidade superior aos lenticulares, mas a qualidade dos lenticulares depende principalmente da uniformidade e da qualidade das lentes utilizadas. É possível construir folhas lenticulares de grandes dimensões e de elevada qualidade, apesar de ser um trabalho difícil e caro. A qualidade dos ecrãs lenticulares pode ser melhorada, utilizando séries de lentes elípticas em lugar das lentes cilíndricas.

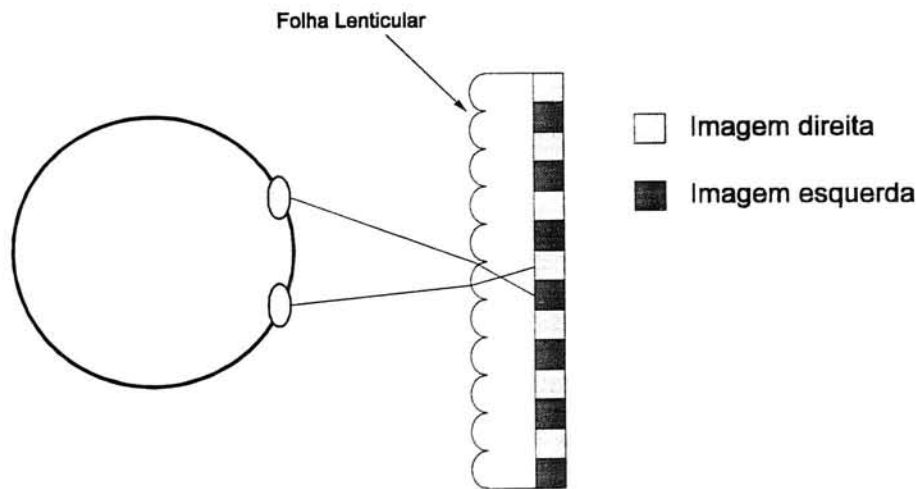


Figura 3.9. Esquema de um dispositivo de visualização utilizando folha lenticular.

Os ecrãs lenticulares possuem pelo menos uma vantagem significativa sobre os seus congéneres de barreira de paralaxe. Uma vez que utilizam refacção em vez da oclusão, o brilho das imagens é superior e mais adequado à utilização em ambientes com luz natural. Os ecrãs lenticulares não obrigam à utilização de iluminações especiais, sendo portanto mais adequados para a generalidade dos processos de armazenamento da imagem. Esta vantagem, combinada com a facilidade de construção de lentes em plástico, justifica o grande número de utilizadores que preferem a utilização dos dispositivos lenticulares aos métodos de barreira de paralaxe. Imagens destinadas a sistemas de visualização lenticulares aparecem vulgarmente em revistas e postais tridimensionais.

Nos métodos de selecção de imagem, a imagem está fisicamente separada do selector, por isso não é necessário que as imagens sejam estáticas. Isto sugere várias possibilidades de aplicação, como os CRTs lenticulares, ou até uma projecção lenticular para um auditório autoestereoscópico. Um auditório deste tipo foi já mostrado publicamente no pavilhão soviético da Expo-90 em Osaka.

3.1.2.2.1. Geração de Imagem para Dispositivos Lenticulares

Um método para gerar as tiras de uma imagem panorâmica, destinada a um dispositivo com selecção da imagem, utiliza uma câmara fotográfica especial. Neste tipo de câmara, o filme é colocado a uma determinada distância atrás de uma série de tiras com a mesma frequência espacial das tiras da barreira de selecção de imagem. Mantendo o obturador aberto, a câmara é movida horizontalmente uma distância não superior ao espaço entre duas tiras consecutivas para não haver sobreposição. Assim, em cada tira é gravada uma imagem contínua, que contém todas as vistas possíveis entre os extremos direito e esquerdo. Uma vantagem desta imagem contínua é a possibilidade de ser vista de qualquer distância e de qualquer ângulo, dentro dos limites determinados.

Na geração computadorizada de imagens panorâmicas, não é possível criar uma imagem contínua, mas podem-se juntar várias vistas para aproximar os elementos contínuos. As diferentes vistas podem ser obtidas em separado. Para fazer a junção, podemos utilizar meios computadorizados, que aproveitem apenas algumas tiras de cada imagem para gerar uma única imagem panorâmica. Alternativamente, podemos usar a própria barreira ou folha lenticular, para criar a imagem panorâmica a partir das várias imagens geradas por meios fotográficos. Para isso, passam-se as imagens a filme, ficando cada uma em sua película. Estas imagens são em seguida divididas em tiras pela projecção, através de uma barreira de selecção ou de uma folha lenticular idêntica à que vai ser utilizada na visualização. As tiras resultantes de cada uma das imagens são entrelaçadas e gravadas consecutivamente na película final. Isto pode ser conseguido, mantendo a posição das películas e deslocando ligeiramente a barreira, ou deslocando o projector de maneira a obter outro ângulo de projecção através da folha lenticular (figura 3.10.).

As vistas discretas gravadas podem ser em qualquer número a partir de duas, no entanto é conveniente ter tantas quantas for possível. A utilização de um grande número de imagens permite não só a possibilidade de visualização estereoscópica a partir de vários pontos, mas também fornece paralaxe de movimento, quando a posição de observação é alterada.

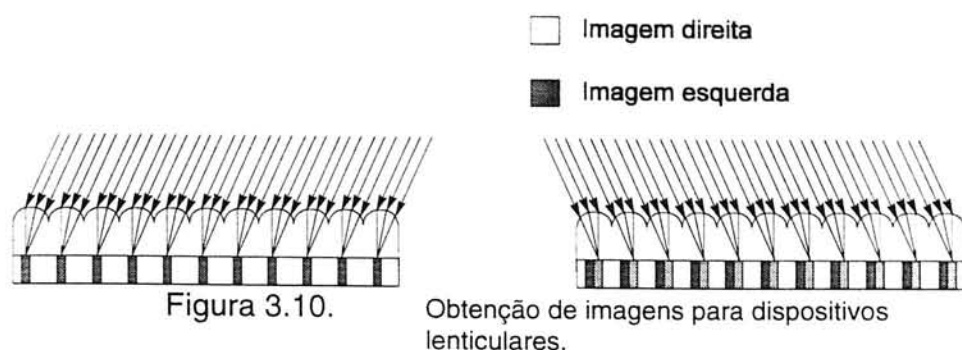


Figura 3.10.

Obtenção de imagens para dispositivos lenticulares.

Um efeito indesejável que aparece associado à paralaxe de movimento é a observação de um salto da posição dos objectos que estão mais distantes do observador. Este efeito pode ser minimizado de diversas formas como, por exemplo, limitando a profundidade de campo na imagem ou incrementando o número de imagens, para que a diferença entre duas vistas consecutivas seja menos sentida.

3.1.3. Estereoscópios

A utilização de free viewing, para observar pares estereoscópicos de imagens, requer prática e é difícil de controlar para a maioria das pessoas. Para facilitar a visualização, desenvolveram-se vários dispositivos, denominados genericamente como *estereoscópios*.

Existem inúmeras arquitecturas de estereoscópios, normalmente baseados em lentes e/ou espelhos. Os mais conhecidos são o estereoscópio de Wheatstone, descrito no capítulo anterior, e o View-Master. Este, permite visualizar pares estereoscópicos de imagens gravadas em pequenos diapositivos, e pode ser adquirido nas lojas de brinquedos com custo pouco elevado. São também conhecidas arquitecturas de estereoscópios especialmente concebidas para visualizar pares estereoscópicos num ecrã de um monitor normal [Gelder91].

3.1.4. Efeito Pulfrich

Um método extremamente simples, que produz um par estereoscópico a partir de qualquer imagem com movimento, consiste na utilização de efeito Pulfrich. Um filtro de densidade neutra colocado em frente de um dos olhos provoca um atraso na entrega da imagem ao cérebro. Com apenas um dos olhos coberto pelo filtro, a mesma imagem será vista por ambos os olhos, mas a partir de posições temporais diferentes. Se a imagem incluir objectos em movimento, cada um dos olhos verá os objectos em posições diferentes das posições vistas pelo outro olho no mesmo instante.

É evidente que a sensação de profundidade resultante da utilização deste método depende fortemente do movimento. Se um objecto está imóvel, aparecerá na mesma posição a ambos os olhos, e portanto não resultará nenhuma sensação de profundidade. Os objectos em movimento aparecerão quer à frente, quer atrás, dos objectos estáticos, dependendo da direcção de translação a que estão sujeitos e de qual o olho escolhido para aplicar o filtro. Este método pode facilmente ser utilizado em animações geradas por computador e gravadas em vídeo. No entanto este método

provoca muitas vezes desconforto, devido à inconsistência das pistas de profundidade e à existência de paralaxe vertical, se a animação não for adequadamente executada.

3.1.5. Polarização das Imagens

A utilização de polarização da luz foi aplicada em primeiro lugar ao cinema estereoscópico, e à visualização, por um auditório, de imagens estereoscópicas em transparências. As duas imagens são enviadas simultaneamente ao observador, mas com polarizações inversas. A separação das imagens é efectuada junto do observador através de uns óculos apropriados.

Quer se trate de cinema ou de transparências estáticas, esta técnica requer dois projectores, polarizadores para cada uma das lentes, um ecrã com superfície metálica (normalmente chamado ecrã de prata), que retenha a polarização da luz depois da reflexão, e óculos polarizados para cada um dos observadores (figura 3.11.).

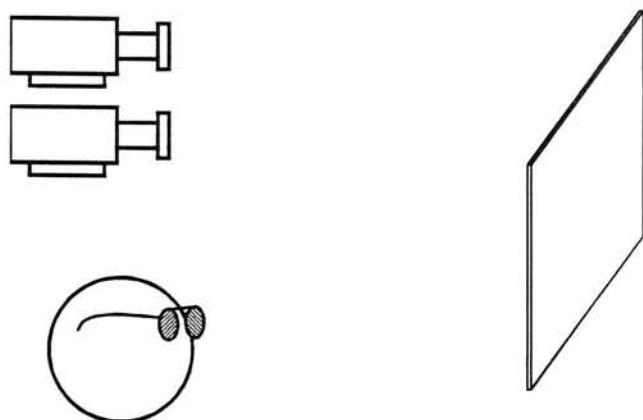


Figura 3.11.

Projectão de imagens estereoscópicas, utilizando duas câmaras com polarizadores, e óculos polarizados para os assistentes.

A utilização de um sistema mecânico-óptico ou electro-óptico para separar duas imagens apresentadas sequencialmente, permite utilizar apenas um projector, ou até aplicar a mesma técnica a um CRT. Em lugar de aplicar um polarizador diferente a cada uma das lentes dos projectores, basta aplicar um polarizador a uma das imagens apresentadas, e um polarizador diferente à imagem apresentada em seguida. Esta estratégia corresponde a aproveitar um canal que utiliza multiplexagem temporal das duas imagens, realizar a desmultiplexagem respectiva, e multiplexar de novo as duas imagens, mas agora por polarização.

Um raio luminoso possui um campo eléctrico que vibra perpendicularmente à sua direcção de propagação. Este campo eléctrico pode ser visto como sendo composto por duas componentes ortogonais. Um polarizador linear absorve a componente perpendicular ao seu eixo de transmissão. Se um segundo filtro polarizador for colocado em frente do primeiro, mas com o eixo de transmissão rodado 90 graus relativamente ao primeiro (orientação cruzada), então os raios luminosos que passam o primeiro filtro serão absorvidos pelo segundo. Se os dois eixos de transmissão forem orientados na mesma direcção (orientação paralela), então a luz que passa o primeiro filtro também atravessa o segundo.

Um par ideal de filtros polarizadores transmitiria 50% da luz visível quando colocados com orientação paralela e 0% em orientação cruzada. Na prática, usando polarizadores de dióxido de crómio, verifica-se cerca de 0,1% de transmissão, em orientação cruzada, e 32% de transmissão, em orientação paralela.

Quando se utilizam polarizadores para visualização estereoscópica, deve-se colocar um polarizador em cada um dos projectores, orientados de maneira que o ângulo entre a polarização das duas vistas, seja de 90 graus. A polarização dos óculos deve corresponder à polarização da projecção, de tal maneira que a polarização de cada um dos olhos seja paralela relativamente à perspectiva respectiva e cruzada relativamente à vista do outro olho.

Um problema com a utilização desta técnica é o aparecimento de imagem fantasma (um olho vê parte da imagem destinada ao outro olho), quando o utilizador move a cabeça de maneira que esta deixa de estar vertical. Isto acontece porque só é possível assegurar a correspondência correcta para uma determinada orientação dos óculos. Por isso, o utilizador deve manter a cabeça na vertical, para conseguir a melhor separação entre as vistas dos olhos esquerdo e direito.

Recentemente surgiram filtros com polarização circular, que permitem eliminar este problema. No entanto, mesmo com polarização circular, o utilizador deve tentar manter a cabeça na vertical, para que a paralaxe não fique diagonal em vez de horizontal.

3.1.6. Anáglifos

Os anáglifos consistem na sobreposição das imagens esquerda e direita, utilizando a cor para as codificar. O observador utiliza óculos com filtros de cor diferente em cada olho, para separar as imagens. Uma vez que as imagens são sobrepostas, não existe nenhuma restrição quanto ao seu tamanho. No entanto, uma vez que são utilizados filtros de cor, não é possível ter imagens tridimensionais coloridas. Nas imagens geradas por computador, os filtros podem ser escolhidos de acordo com a cor da tinta de uma impressora ou de uma plotter.

Esta técnica pode ser utilizada não só em imagens estáticas, mas também em filmes tridimensionais, que são vulgares desde 1950. As maiores desvantagens deste método são a falsa coloração que produz e o desconforto verificado em alguns utilizadores. A qualidade do anáglifo depende fortemente da capacidade de filtragem cromática dos filtros utilizados.

3.1.7. Multiplexagem Temporal das Imagens

Os sistemas de apresentação sequencial das imagens para o olho esquerdo e para o direito, também chamados *sistemas de multiplexagem temporal*, têm conhecido um rápido progresso nos últimos anos. Os primeiros sistemas deste tipo utilizavam obturadores mecânicos para tapar o olho apropriado durante um refrescamento do ecrã. As imagens eram apresentadas alternadamente: Quando o olho esquerdo estava tapado, era apresentada a imagem para o olho direito. Quando o olho direito estava tapado, era apresentada a imagem para o olho esquerdo. Consequentemente, os sistemas deste tipo requerem pelo menos dois frame buffers, um para cada olho, e um

mecanismo que permita comutar os dois frame buffers entre dois refrescamentos consecutivos do ecrã.

Para bons resultados é fundamental dispor de um CRT que possa ser refrescado pelo menos 100 vezes por segundo. Desta forma, cada um dos olhos receberia 50 imagens por segundo. A utilização de taxas de refrescamento inferiores conduz ao aparecimento de flicker, tornando a visualização desconfortável.

Actualmente existem duas técnicas electro-ópticas que substituíram os sistemas mecânicos de apresentação sequencial das imagens. Ambas utilizam LCD para realizar a desmultiplexagem, tapando ou transmitindo uma das imagens. Uma das técnicas realiza a desmultiplexagem junto do observador, através de uns óculos activos apropriados. A outra técnica realiza a desmultiplexagem temporal junto ao ecrã, e multiplexa novamente as imagens mas agora por polarização. A desmultiplexagem é então efectuada por óculos polarizadores passivos usados pelos espectadores.

Existem no mercado diversos sistemas de multiplexagem temporal que utilizam óculos activos para separar as imagens. Os mais conhecidos são o CrystalEyes da StereoGraphics que descrevemos a seguir, e o 3D TV de uma companhia com o mesmo nome. As técnicas utilizadas por estes dois produtos são as mesmas. O 3D TV é um produto de mais baixo custo, destinado essencialmente à utilização em PC, com frequências de varrimento não superiores a 60Hz.

O método dos óculos passivos utiliza um painel LCD colocado em frente do monitor. O observador tem de usar óculos polarizados, parecidos com óculos de sol. O LCD comuta a polarização da luz emitida pelo ecrã, com a mesma frequência e em sincronismo com a apresentação das imagens. Durante a apresentação de uma imagem, a luz é polarizada numa direcção. A lente dos óculos que é polarizada na mesma direcção permite a passagem da luz para o olho respectivo. A lente que é polarizada na direcção inversa não permite a chegada da luz ao olho. Na apresentação da imagem seguinte, o LCD polariza a luz na direcção contrária, permitindo que a imagem chegue ao outro olho. Sistemas de visualização estereoscópica deste tipo, aplicáveis a monitores gráficos de computadores, são comercializados pela Tektronix e pela StereoGraphics.

Esta técnica que utiliza óculos passivos torna-se mais económica para um número elevado de observadores, devido à simplicidade dos óculos passivos utilizados. No entanto, a utilização de óculos activos é mais simples, e economicamente interessante para um só utilizador.

3.1.7.1. CrystalEyes

O primeiro protótipo do sistema CrystalEyes foi criado pela StereoGraphics em 1983. Este sistema utiliza um emissor de infravermelhos localizado no monitor. O emissor difunde um sinal em sincronismo com o varrimento vertical do monitor. O utilizador utiliza uns óculos (Eyewear) que incorporam um receptor de infravermelhos, destinado a captar o sinal difundido pelo emissor. Este sinal de sincronismo é utilizado para controlar os bloqueadores de cristal líquido (LC) em sincronismo com o varrimento vertical. Desta maneira apenas um olho vê em cada instante a imagem presente no monitor. O Eyewear incorpora toda a electrónica necessária ao seu funcionamento. A

alimentação é efectuada através de duas pequenas baterias de Lítio também incorporadas.

O sistema CrystalEyes é baseado na técnica, já conhecida, de apresentação sequencial das imagens destinadas aos olhos direito e esquerdo. No entanto, ao contrário de alguns sistemas mais antigos, este produto mantém o número de imagens por segundo apresentadas a cada olho. Isto é conseguido, sem necessidade de modificação do hardware gráfico, pela duplicação do número de frames apresentados, e pela divisão do frame buffer em duas metades (figura 3.12). A metade de cima contém a imagem destinada ao olho esquerdo, e a metade de baixo contém a imagem destinada ao olho direito.

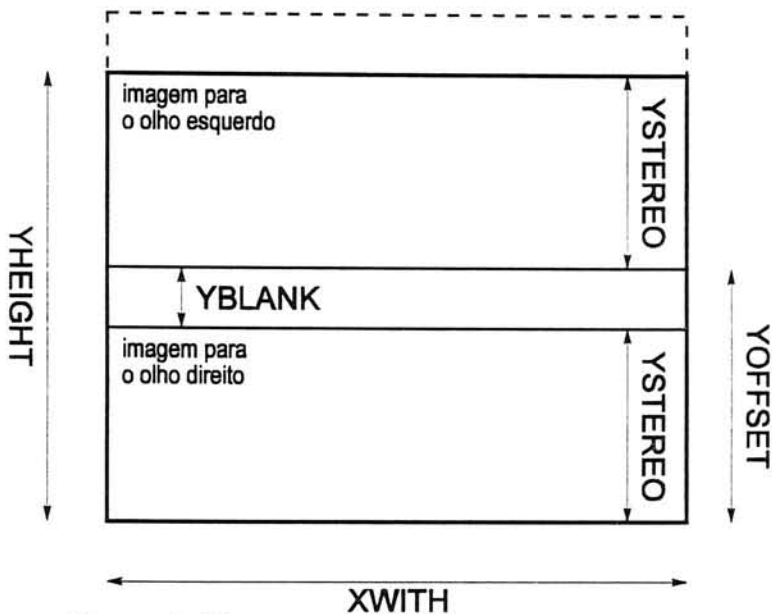


Figura 3.12. Organização do frame buffer no modo de visualização estereoscópico.

Desta maneira, a taxa de refrescamento passa efectivamente de 60 para 120 Hz. Como cada um dos olhos recebe apenas metade das imagens apresentadas, o número de imagens vistas por cada um dos olhos permanece igual a 60 por segundo. Devido à divisão do frame buffer, o número de linhas por imagem passa para metade. A frequência de varrimento horizontal permanece constante.

Se visualizássemos a 60 imagens por segundo, um frame buffer dividido e contendo as imagens de um par estereoscópico, veríamos ambas as imagens anormalmente comprimidas na vertical. A imagem destinada ao olho esquerdo apareceria na parte superior do ecrã, e a imagem destinada ao olho direito apareceria na metade inferior. No entanto, duplicando a frequência de varrimento vertical para 120Hz, as imagens apareceriam sobrepostas, embora escritas sequencialmente uma após a outra, em varrimentos distintos.

Para se poder aplicar o sistema CrystalEyes é necessário dispor de um monitor capaz de, automaticamente, comutar de 60 para 120Hz e ajustar a altura da imagem para

repor o “aspect ratio”¹. Estes monitores detectam os novos impulsos de sincronização vertical, e alteram o varrimento de forma correspondente. Cabe ao adaptador gráfico incluir os novos impulsos de sincronismo vertical. Na maioria dos casos, os adaptadores gráficos não prevêm esta inserção adicional de impulsos de sincronismo, pelo que é necessário intercalar um controlador dedicado entre o adaptador gráfico e o monitor (figura 3.13). Este controlador, que pode ser fornecido opcionalmente pelo fabricante (GDC-3), adiciona o novo impulso de sincronismo a cada varrimento vertical e fornece também a ligação do emissor CrystalEyes.

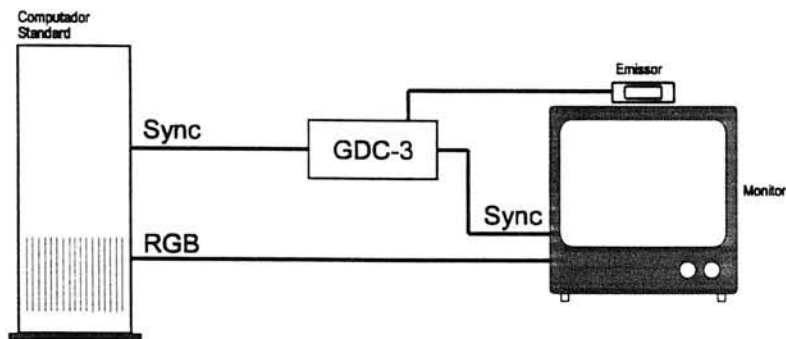


Figura 3.13. Utilização do sistema CrystalEyes com recurso de um controlador apropriado.

Alternativamente, é possível utilizar um computador com um adaptador gráfico preparado para duplicar a frequência de varrimento vertical e permitir a ligação do emissor CrystalEyes (figura 3.14). Um exemplo de um computador nestas condições é o SiliconGraphics IRIS Indigo XS24Z, que foi utilizado nas experiências realizadas com este sistema.

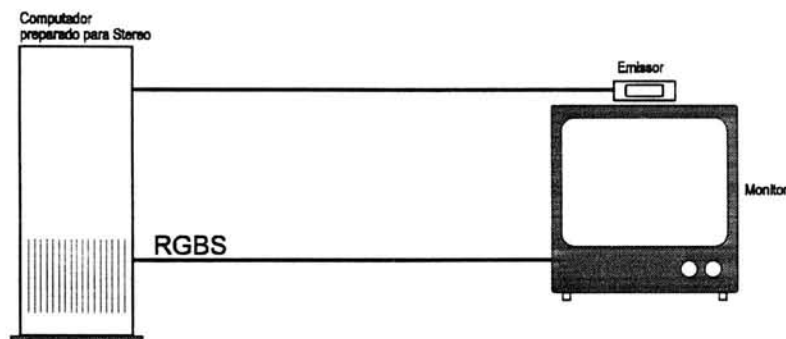


Figura 3.14. Utilização do sistema CrystalEyes com um computador preparado para estereó.

Em cada varrimento vertical são incluídas centenas de linhas horizontais, seguidas de 20 a 40 linhas invisíveis. O monitor normalmente desenha estas linhas durante o retorno do varrimento vertical ao topo do ecrã. Quando se utiliza uma frequência de varrimento vertical dupla da normal, existe necessariamente uma nova zona horizontal que se torna invisível. A altura desta área medida em número de linhas (YBLANK), é igual ao número de linhas normalmente invisíveis. Isto provoca uma duplicação do número de linhas invisíveis, e também o não aproveitamento de uma parte do frame

¹ aspect ratio - relação de aspecto. Relação entre as dimensões horizontais e verticais de um pixel.

buffer. A porção do frame buffer, correspondente aos pixels das linhas que agora ficam invisíveis, não é aproveitada para a imagem visível.

Quando a resolução normal do ecrã é *XWITH* colunas por *YHEIGHT* linhas, e o número de linhas normalmente invisíveis é *YBLANK*, a área invisível que separa a vista destinada ao olho esquerdo da vista destinada ao olho direito, é também *YBLANK*. Cada uma das imagens será mapeada numa região do ecrã com *XWITH* pixels de largura e altura igual a:

$$YSTEREO = \frac{YHEIGHT - YBLANK}{2} \tag{3.1}$$

A linha inferior da imagem destinada ao olho direito será mapeada na linha 0. A linha inferior da imagem destinada ao olho esquerdo será mapeada na linha

$$YOFFSET = YSTEREO + YBLANK \tag{3.2}$$

Um ponto com coordenadas (x,y) na imagem destinada ao olho direito, deve então ser mapeado num pixel, cujas coordenadas de ecrã são:

$$\left(XWITH \cdot \frac{x - x_{min}}{x_{max} - x_{min}}, YSTEREO \cdot \frac{y - y_{min}}{y_{max} - y_{min}} \right) \tag{3.3}$$

Por sua vez, um ponto com coordenadas (x,y) na imagem destinada ao olho esquerdo, deve ser mapeado num pixel, cujas coordenadas de ecrã são:

$$\left(XWITH \cdot \frac{x - x_{min}}{x_{max} - x_{min}}, YSTEREO \cdot \frac{y - y_{min}}{y_{max} - y_{min}} + YOFFSET \right) \tag{3.4}$$

(*x_{min}*, *y_{min}*) representa o canto inferior esquerdo do viewport e (*x_{max}*, *y_{max}*) representa o canto superior direito do viewport, expressos em coordenadas do mundo. O canto inferior esquerdo tem como coordenadas (0,0).

Para o computador e adaptador gráfico utilizados (SiliconGraphics IRIS Indigo R4000 XS24Z), os valores das constantes referidas são os apresentados na tabela seguinte:

XWITH	1280
YHEIGHT	1024
YSTEREO	494
YBLANK	36
YOFFSET	530

Tabela 3.2. Valores utilizados na estação gráfica IRIS Indigo XS24Z.

3.1.8. Holografia

A Holografia é uma forma de fotografia sem lentes, que foi descrita pela primeira vez por Gabor, em 1947. Este método baseia-se na gravação em filme, mas aquilo que é gravado não é uma imagem, como acontece na fotografia tradicional. Os negativos produzidos por processos holográficos parecem manchados. Quando devidamente iluminados, os hologramas reconstroem imagens tridimensionais, que podem ser virtualmente indistinguíveis das dos objectos reais que representam. Alterando a sua posição, o observador pode ver os objectos mais próximos de vários lados e o que se encontra por trás deles. Apesar dos hologramas serem gravados em filme preto e branco, a imagem reconstruída a partir deles é colorida, normalmente monocromática. No entanto, também é possível reconstruir imagens policromáticas a partir de hologramas. Existem vários tipos de hologramas que permitem variações nas condições de visualização e nas técnicas de fabrico. O tipo de holograma mais conhecido é o holograma de reflexão. São deste tipo os hologramas que se podem encontrar nos cartões de crédito, nas capas de algumas revistas e de alguns discos, e até como brindes em caixas de produtos alimentares ou detergentes.

A obtenção de imagens tridimensionais é possível porque o holograma armazena mais informação do que a fotografia (o termo holograma significa toda a mensagem). Na fotografia tradicional, uma imagem é gravada num filme sensível à luz. Quando revelado, cada ponto do filme ficará pouco ou muito escuro, dependendo (inversamente) da intensidade ou amplitude das ondas luminosas que o atingiram. Como nos filmes normais esta dependência corresponde a uma proporcionalidade inversa, diz-se que se trata de filmes negativos, e a imagem gravada é uma imagem negativa. Independentemente do facto de ser positivo ou negativo, o filme fotográfico armazena apenas a amplitude das ondas luminosas, perdendo-se a informação da fase dessas mesmas ondas.

Quando a luz proveniente de fontes luminosas diferentes se junta, as ondas interferem uma com a outra, resultando daí uma nova onda que é a soma em valor instantâneo das duas componentes. Normalmente estas componentes possuem frequências diferentes e então a sua soma é variável no espaço e no tempo. No caso especial das frequências serem as mesmas, resulta uma onda estacionária.

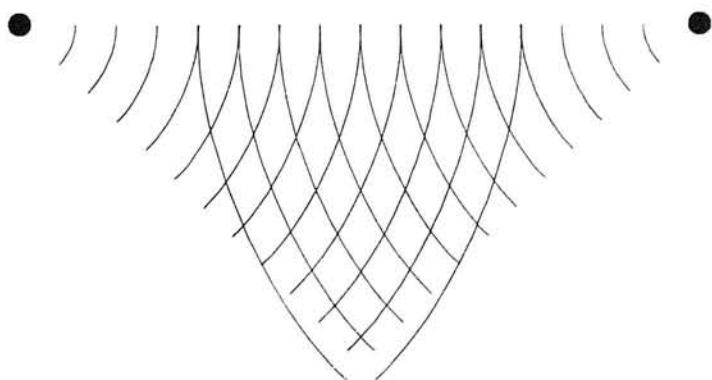


Figura 3.15. Interferência de ondas de duas fontes distintas.

3.1.8.1. Hogramas de Transmissão

Para gravar um holograma de transmissão, divide-se um feixe laser em dois feixes elementares: o feixe do objecto e o feixe de referência. O feixe do objecto é utilizado para iluminar o objecto, e parte da sua luz vai ser reflectida pelo objecto para o filme holográfico. O feixe de referência é aplicado directamente ao filme holográfico. Os dois feixes, juntos, produzem um padrão de interferência, que contém a informação de fase e de amplitude que vai ficar gravada no filme (figura 3.16.).

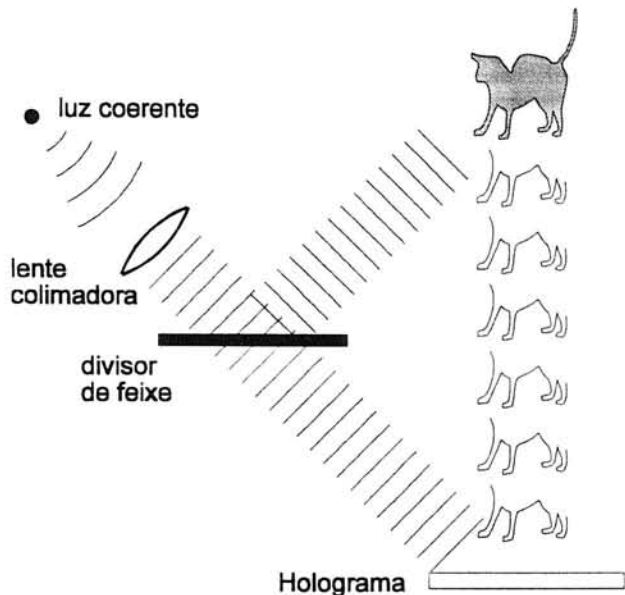


Figura 3.16. Obtenção de um holograma.

Depois de revelado, o filme é chamado holograma de amplitude e as suas áreas escuras servem como modeladores de difracção. Se o holograma for novamente iluminado pelo feixe de referência com o mesmo ângulo de incidência, vai difractar a luz reconstruindo a imagem do objecto original (figura 3.17.).

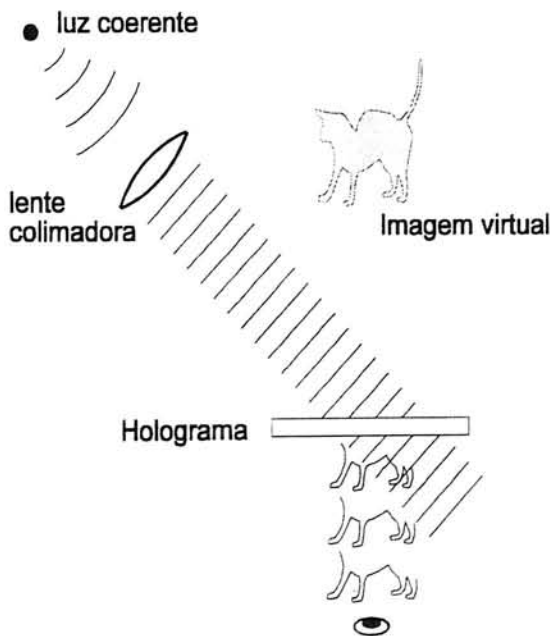


Figura 3.17. Visualização de um holograma.

As ondas luminosas, que eram reflectidas no objecto e que nos permitiam observá-lo, foram gravadas no filme. Quando a imagem é reconstruída, as ondas luminosas visíveis são idênticas àquelas que o observador via, quando olhava para o objecto original.

Para além de hologramas de amplitude, é possível criar hologramas de fase. Em vez de modelarem a amplitude pela oclusão de alguma luz, permitem a passagem de toda a luz, mas com a fase modelada pela variação da espessura do holograma. Os hologramas de fase podem ser obtidos através de uma compressão que corresponda à luminosidade do filme.

Devido às suas capacidades únicas para apresentação de imagens tridimensionais, têm sido feitos esforços desde os anos 60 na aplicação de hologramas a imagens geradas por computador. No entanto, ainda vamos ter que esperar alguns anos até que esta técnica de visualização de imagens tridimensionais seja praticável para aplicações gráficas tradicionais.

3.2. Armazenamento de Imagens Estereoscópicas

Devemos primeiro notar que, para armazenar as imagens geradas digitalmente, podemos necessitar de grandes quantidades de memória. Por exemplo, para armazenar um único par de imagens destinado a um frame buffer com resolução de 1024 por 1024 e 24 bits por pixel, é necessário dispor de 6 megabytes de memória, partindo do princípio que não era utilizada nenhuma compressão. Uma animação de 30 segundos, com 30 pares de imagens por segundo, necessitaria mais de 5 gigabytes. Existem no entanto, algumas técnicas alternativas mais vantajosas, para representar imagens tridimensionais.

3.2.1. Transparências

O filme fotográfico é um método muito popular para guardar e distribuir pares estereoscópicos de imagens, uma vez que permite utilizar simultaneamente elevada resolução espacial e cromática. Além disso, os sistemas ópticos envolvidos são já há muitos anos bem conhecidos. Vimos já alguns métodos que podem ser utilizados para visualizar imagens tridimensionais estáticas em impressão positiva (transparências), que são utilizadas desde há alguns anos em fotografia tridimensional.

Um método comum para apresentar pares estereoscópicos de imagens a um grupo de pessoas (principalmente àquelas que não dominam a arte de free viewing), consiste em captar as imagens esquerda e direita em filme fotográfico (utilizando dois slides 35mm, ou filme cinematográfico no qual as imagens esquerda e direita são guardadas em quadros alternados), e projectar ambas as vistas num único ecrã. A separação da imagens pode ser conseguida pela utilização de óculos com filtros de cor, como no anáglifo, com óculos activos, ou preferencialmente, com luz polarizada e óculos apropriados.

3.2.2. Vectógrafo

Outro método que utiliza polarização é o Vectógrafo da Polaroid. Este método utiliza uma lâmina plástica com duas películas de álcool polivinílico, para gravar um par estereoscópico de imagens. As folhas de álcool polivinílico são o material transparente que é utilizado como base para uma pintura com compostos de iodo, transformando-se em filtros polarizadores.

Um polarizador linear pode ser criado esticando uma película de álcool polivinílico para alinhar as suas moléculas em cadeias finas e longas. Tingindo-a com uma tinta à base de iodo, a película fica com electrões livres, que se movem ao longo das cadeias moleculares. Quando a luz atravessa uma película esticada e tingida, a componente do campo eléctrico, que está alinhado com estas cadeias, provoca a vibração dos electrões ao longo das cadeias. Uma vez que produz trabalho, a energia desta componente do campo eléctrico é absorvida. A componente do campo eléctrico perpendicular às cadeias moleculares é transmitida, uma vez que os electrões não se podem mover nesta direcção e portanto a sua energia não é absorvida.

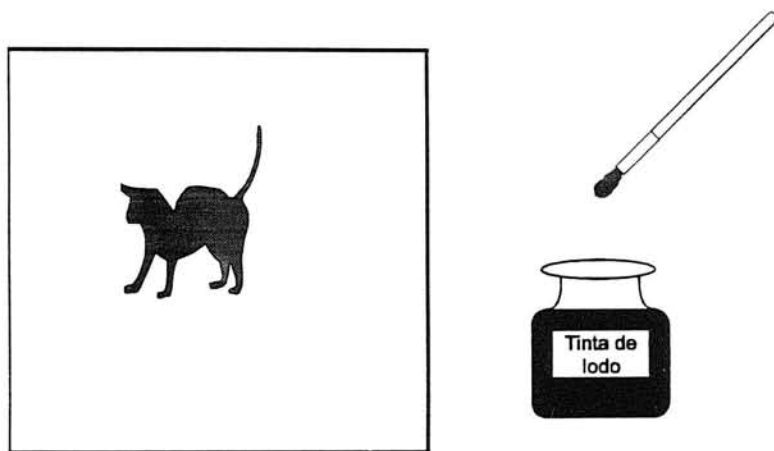


Figura 3.18. Impressão de um vectógrafo

Para obter o vectógrafo estereoscópico, juntam-se duas películas esticadas ainda por pintar, com os eixos de transmissão perpendiculares. Este conjunto das duas películas constitui a lâmina do vectógrafo. Se desenharmos uma imagem, num dos lados da lâmina, com tinta de iodo, a área coberta pela tinta vai polarizar a luz que a atravessa, enquanto todas as áreas que não estão pintadas transmitem a luz sem qualquer polarização. Desenhando na outra face da lâmina, obtemos uma imagem em luz polarizada, mas agora com polarização perpendicular à primeira.

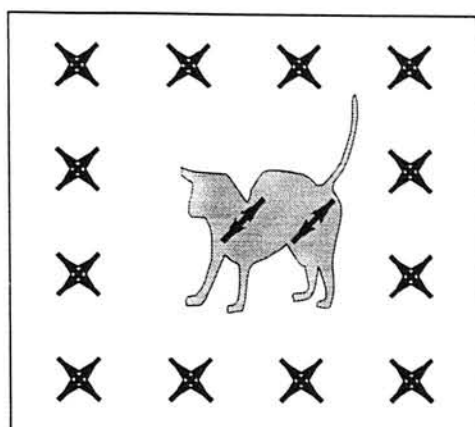


Figura 3.19. Polarização da luz, com a tinta de iodo.

Para imagens estereoscópicas, a perspectiva para o olho esquerdo é gravada na face da frente da lâmina e a perspectiva para o olho direito é gravada na face da trás (Figura 3.20.).

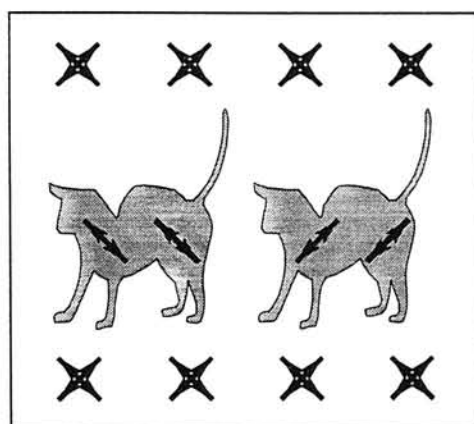


Figura 3.20. Vectógrafo estereoscópico

A imagem tridimensional é vista usando óculos polarizados, para que cada uma destas imagens seja entregue apenas ao olho respectivo. O olho esquerdo verá apenas a imagem gravada na face da frente (Figura 3.21.).

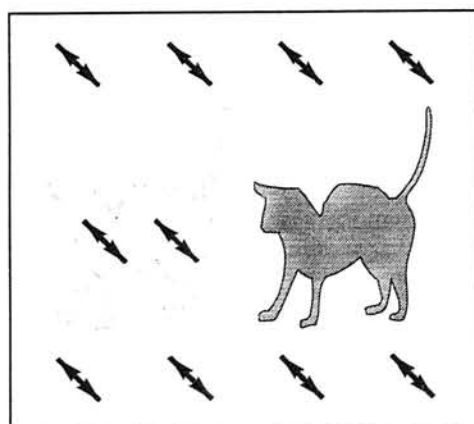


Figura 3.21. Imagem visível pelo olho esquerdo

E o olho direito verá apenas a imagem gravada na face de trás da película (Figura 3.22.).

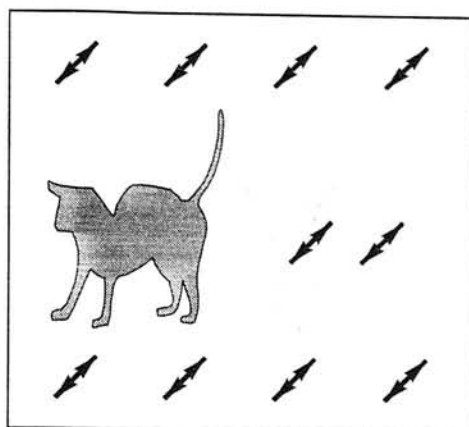


Figura 3.22. Imagem visível pelo olho direito

É também possível mostrar um vectógrafo a um grande número de pessoas equipadas com óculos polarizados, utilizando um projector para projectar o conjunto num ecrã metálico. A utilização de apenas um único projector tem a vantagem de eliminar os problemas de falta de alinhamento das imagens na altura da apresentação.

Para criar um vectógrafo, as imagens podem ser desenhadas por uma plotter. Para resoluções mais elevadas e tons contínuos, pode ser utilizado um processo através de filme fotográfico. Independentemente do processo de gravação do vectógrafo, é necessário garantir o correcto alinhamento das imagens.

Uma vez que é possível fabricar filtros polarizadores coloridos, é também possível obter vectógrafos coloridos. Os métodos de separação de cor são aplicados a cada uma das duas imagens, para se obterem oito imagens de componentes primárias CMYK. Estas imagens elementares são então gravadas com tinta colorida. Eventualmente poderia ser possível utilizar uma impressora de jacto de tinta; no entanto, estas impressoras não são actualmente aplicadas, devido à natureza altamente corrosiva das tintas utilizadas. A impressão a jacto de tinta poderia incluir a vantagem da impressão simultânea nas duas faces, simplificando assim o processo de alinhamento das imagens.

3.2.3. Gravação Vídeo

O gravador de vídeo (video tape) é um dispositivo pouco dispendioso, que permite guardar imagens estereoscópicas de resolução média. Num ecrã com 25 ou 30 varrimentos por segundo interlaçados, podemos considerar a imagem dividida em dois campos: o campo par que é constituído por todas as linhas pares, e o campo ímpar que é constituído pelas linhas ímpares. Cada um dos campos é actualizado trinta vezes por segundo.

Para visualizar as duas imagens de um par estereoscópico, a perspectiva a entregar a um dos olhos é enviada para o campo par, e a perspectiva para o outro olho é enviada no campo ímpar. O ciclo de actualização entrelaçado apresentará não uma, mas duas perspectivas diferentes, que se alternam 50 ou 60 vezes por segundo. Para evitar o aparecimento de imagens fantasma, deve-se utilizar um ecrã com fósforo rápido.

Para se ver a imagem tridimensional, é necessário usar um dispositivo mecânico ou electro-óptico, para tapar, em cada instante, o olho oposto à vista apresentada.

Os sistemas disponíveis comercialmente utilizam óculos activos, com lentes de cristal líquido, para cada um dos observadores. É colocado perto do monitor um emissor de infravermelhos, sincronizado com o ciclo de refrescamento vertical. Os óculos dos observadores sincronizam através da recepção de infravermelhos, para tapar alternadamente o olho esquerdo ou direito. O problema que pode surgir com a utilização deste método é o aparecimento de flicker, uma vez que é a tecnologia que determina a taxa de refrescamento.

Alternativamente, pode ser utilizado um cristal líquido, que é colocado em frente do monitor e é sincronizado com o ciclo de refrescamento vertical. O sistema utiliza polarização circular, de modo a polarizar a perspectiva enviada no campo par num sentido, e a perspectiva enviada no campo ímpar no sentido oposto. O observador utiliza agora óculos passivos, com as lentes polarizadas circularmente e em sentidos contrários, para permitir a cada olho ver apenas a perspectiva respectiva. A utilização deste método pode igualmente levar ao aparecimento de flicker.

3.2.4. Estereogramas Aleatórios

Um estereograma aleatório é uma imagem composta de pontos, caracteres, ou símbolos, aparentemente posicionados de forma aleatória. Esta imagem, quando vista da forma correcta, apresenta uma imagem tridimensional. Os estereogramas aleatórios são muito úteis para o estudo do efeito estereoscópico, porque fornecem uma imagem constituída unicamente pelo efeito estereoscópico de profundidade. Numa imagem deste tipo, não existe coloração das superfícies, nem contornos definidos para os objectos. A forma dos objectos pode apenas ser entendida pela variação da distância a que as superfícies dos objectos se encontram do observador.

Em todos os outros tipos de pares estereoscópicos de imagens, é possível a visualização monoscópica de qualquer uma das imagens componentes do par. No caso dos estereogramas aleatórios, as duas imagens do par só fazem sentido em conjunto, visualizadas uma por cada olho. Quando observadas em separado, ou pelos dois olhos em simultâneo, apenas é visível um padrão pseudo-aleatório de pontos ou símbolos.

Para que a resolução da imagem tridimensional resultante não seja demasiado pobre, os símbolos utilizados devem ser de dimensões mínimas. Por esta razão, os estereogramas aleatórios mais interessantes são os que utilizam pequenos pontos, normalmente da dimensão de um pixel do ecrã. Neste caso, tomam o nome de *estereogramas de pontos aleatórios*, ou *RDS (Random Dot Stereograms)*.

Para além da baixa resolução e da falta de cor dos objectos da imagem obtida, os RDS têm um outro problema: a dificuldade de ser visualizado, principalmente para observadores com pouca prática. Para minimizar esta dificuldade, os pares de pontos podem ser coloridos aleatoriamente. Se for utilizado um ecrã de um computador em vez de papel, pode ser utilizada uma técnica que consiste em alterar continua e aleatoriamente os padrões e/ou as cores dos pontos que constituem o estereograma.

Pode parecer que a estereoscopia requer necessariamente duas imagens separadas, ou pelo menos um método que possa separar uma única imagem em duas, para fornecer a cada um dos olhos uma vista separada. Em 1990, Tyler e Clarke demonstraram que as duas imagens de um estereograma de pontos aleatórios podem ser combinadas numa só. Ao resultado chama-se *SIRDS* - *Single Image Random Dot Stereogram* (Estereograma de pontos aleatórios numa única imagem).

Para criar um SIRDS, basta sobrepor os dois padrões separados de pontos aleatórios, colocando cuidadosamente os pontos de tal maneira que cada um sirva simultaneamente para as duas vistas da imagem. Tudo o que é necessário é impor as restrições adequadas ao padrão de pontos aleatórios. Os pares estereoscópicos de imagens fotográficas ou modelos “wire frame” não podem ser combinados numa única imagem, pois o resultado seria sempre uma imagem dupla. Um exemplo de um estereograma de pontos aleatórios numa única imagem pode ser visto na figura 3.23.

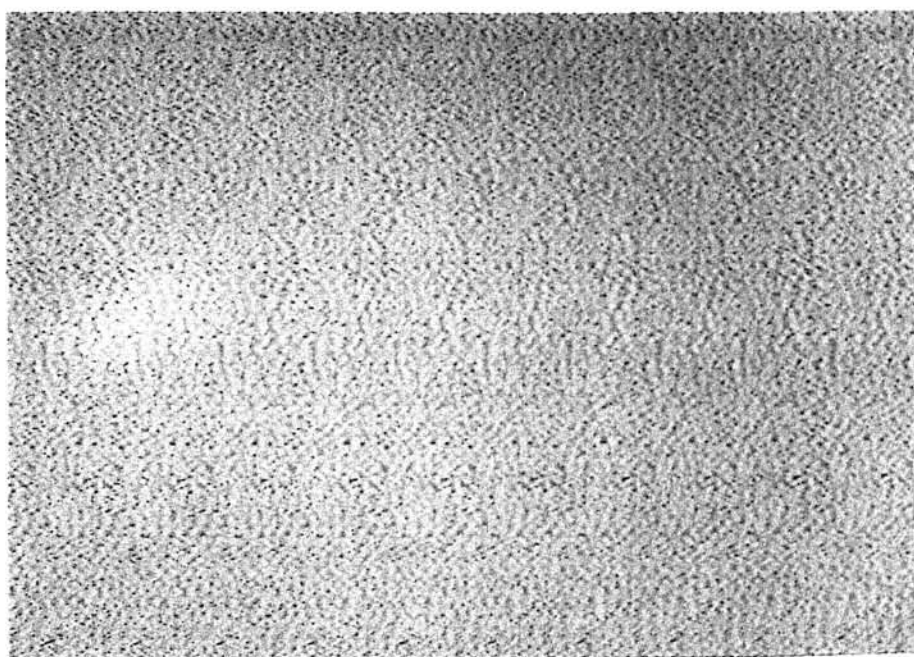


Figura 3.23. Estereograma de pontos aleatórios. Para observar esta imagem, deve ser utilizada visualização cruzada.

A descrição de algoritmos para construção de SIRDS pode se encontrada em [Thimble92].

3.3. Conclusão

Apesar de ainda pouco aplicados a imagens geradas por computador, estão já disponíveis diversos dispositivos que permitem visualizar imagens estereoscópicas. Em virtude de terem sido concebidos especificamente para trabalhar com os sistemas de visualização por computador existentes, os dispositivos que utilizam multiplexagem temporal das imagens são os mais adequados à maioria das aplicações interactivas

tradicionais. O custo e a complexidade destes dispositivos podem ser considerados suportáveis para a maioria das aplicações que não necessitam de apresentar imagens a um grande número de observadores.

Quando o número de observadores é grande, as soluções baseadas na polarização das imagens devem apresentar vantagens, sobretudo em termos de custo. Para impressão e difusão de imagens estereoscópicas, os anáglifos e a utilização de free-viewing são as técnicas mais interessantes. As aplicações com ambientes virtuais, que necessitam de largos ângulos de visão para permitir a sensação de inclusão, requerem um dispositivo de visualização diferente, montado na cabeça no utilizador/participante. São estes dispositivos que iremos tratar no capítulo seguinte.

Capítulo 4.

Visualizador Montado na Cabeça

O primeiro HMD (Head Mounted Display) foi construído por Ivan Sutherland em 1968. Desde então, vários HMD's foram construídos, e alguns deles disponibilizados comercialmente. Associado a cada HMD, foi sempre definido o código de geração de imagem considerado adequado ao dispositivo respectivo.

No primeiro HMD foram utilizados pequenos CRTs monocromáticos [Suther65] [Suther68], acoplados a um sistema óptico. O conjunto permitia obter uma imagem virtual que parecia colocada a 45 cm em frente dos olhos, e um campo de visão de 40 graus. A utilização de um espelho semi-prateado permitia sobrepor as imagens geradas por computador à observação directa do mundo real. Esta facilidade levou mais tarde à aplicação da designação STHMD (See Through Head Mounted Displays), para diferenciar estes dispositivos dos que não permitiam a observação directa do mundo real.

Os HMD's que se seguiram foram de uma maneira geral estereoscópicos. Os HMD's estereoscópicos permitiam o ajuste mecânico da distância interocular, de modo a tornar possível a sua utilização a qualquer pessoa. O software de geração de imagem, desenvolvido especialmente para cada versão de HMD, permitia o ajuste da separação dos centros de projecção. Um HMD essencialmente idêntico foi transferido e utilizado em diversos trabalhos realizados na University of Utah.

Em 1983, Mark Callahan no MIT construiu um STHMD [Callaha83], similar aos já obtidos por Sutherland. Utilizava espelhos semi-prateados montados em armações de óculos, CRTs monocromáticos com 5 cm de diagonal, e um capacete de bicicleta.

Em 1985, Mike McGreevy e Jim Humphries, construíram, na NASA, um HMD baseado em LCD's monocromáticos de televisões de bolso. Neste HMD foram utilizadas as lentes fabricadas pela LEEP para visualização estereoscópica de fotografia [Leep90]. Estes sistemas ópticos oferecem um secção de saída bastante larga, pelo que não é necessário realizar ajustes mecânicos da distância interocular. Para além de ser mais leve que os HMD's baseados em CRT's, este dispositivo diferenciava-se dos seus antecessores pela impossibilidade de o utilizador observar directamente o mundo real. Este HMD foi posteriormente melhorado por Scott Fisher, Warren Robinett e outros.

Durante os anos 70 e 80, Tom Furness dirigiu um programa de desenvolvimento de vários protótipos de STHMD's, destinados à utilização em aviões militares. Nos diversos protótipos foram incorporados sistemas ópticos, desenvolvidos especialmente para o efeito.

O primeiro HMD por fibra óptica foi desenvolvido pela CAE Electronics para ser utilizado em simuladores de voo. O sistema permitia campo de visão binocular horizontal de 135° e vertical de 64° . Dispunha ainda de uma secção central de 25° por 19° com elevada resolução.

Vários protótipos de HMD's têm sido construídos desde 1985 na University of North Carolina em Chapel Hill. Os diversos protótipos incluem modelos com e sem observação directa do mundo real e com sistemas ópticos desenvolvidos especialmente para a aplicação a que se destinam. Actualmente encontra-se em desenvolvimento um HMD com escrita directa na retina [Holmgren92].

Em 1988, a empresa Californiana VPL começou a vender o EyePhone, que foi o primeiro HMD disponível comercialmente. O EyePhone utiliza LCD a cores e os sistemas ópticos LEEP. Não permite a observação directa do mundo real.

Em 1989, a LEEP disponibilizou comercialmente um novo HMD. Utilizava LCD's monocromáticos e os já bem conhecidos sistemas ópticos LEEP. Posteriormente foram introduzidos modelos melhorados que utilizavam LCD coloridos e maior ângulo de visão [Leep90]. Actualmente, a LEEP oferece vários modelos de HMD's, monoscópicos e estereoscópicos, que primam pelo largo ângulo de visão.

Também em 1989, a Reflection Technologies produziu um HMD para um só olho chamado PrivateEye. Este produto utiliza um espelho vibratório e um conjunto linear de LEDs para produzir uma imagem 2D monocromática. O ângulo de visão horizontal é aproximadamente 25° .

Neste capítulo vamos analisar a arquitectura dos HMD's correntes, que utilizam ecrãs LCD. Vamos também referir outras técnicas menos comuns de concepção de dispositivos de visualização montados na cabeça.

4.1. Modelo de um sistema óptico monoscópico

Começamos por considerar um sistema monoscópico de visualização que utiliza um sistema óptico semelhante ao utilizado nos HMD's (figura 4.1).

O nosso objectivo imediato é conseguir relacionar as coordenadas normalizadas, de um ponto de uma imagem, com as coordenadas correspondentes no plano da imagem virtual. Posteriormente faremos uso desta relação para estudarmos o modelo estereoscópico.

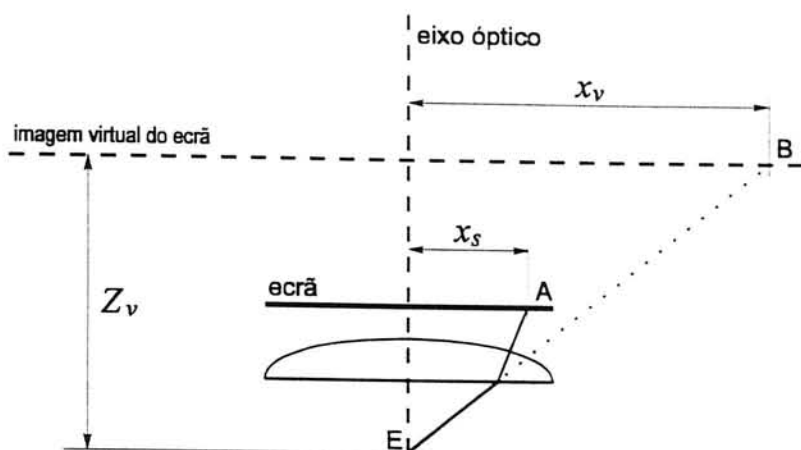


Figura 4.1.

Modelo de um sistema de visualização monoscópico.

As coordenadas do ponto **A**, tomadas com base no eixo da imagem, serão:

$$x_s = x \cdot w_s \quad (4.1)$$

$$y_s = y \cdot h_s \quad (4.2)$$

Nestas equações, w_s e h_s representam respectivamente a largura e a altura da imagem apresentada no ecrã, e x e y as coordenadas horizontal e vertical normalizadas do ponto.

A distância do ponto **A** ao eixo da imagem pode ser dado por:

$$r_s = \sqrt{y_s^2 + x_s^2} \quad (4.3)$$

A representação matemática da distorção óptica depende da natureza do sistema óptico. Para lentes como as LEEP, a aproximação a um polinómio de terceira ordem é adequada.

$$r_{vn} = r_{sn} + k_{vs} \cdot r_{sn}^3 \quad (4.4)$$

Onde k_{vs} representa o valor da distorção geométrica introduzida. A expressão 4.4 pode ser rearranjada algebricamente para:

$$r_{vn} = (1 + k_{vs} \cdot r_{sn}^2) \cdot r_{sn} \quad (4.5)$$

Deve-se notar que nesta aproximação de terceira ordem não foi incluído nenhum termo de segunda ordem. É no entanto sabido [Robinet91] que as aproximações polinomiais das funções de distorção de lentes deste tipo nunca incluem termos de ordem par.

Utilizando as expressões 4.1 e 4.2, podemos expandir a igualdade 4.5 para coordenadas rectangulares, e assim obtermos a posição da imagem virtual do pixel **A**, em função das coordenadas normalizadas de ecrã.

$$(x_{vn}, y_{vn}) = ((1 + (x_{sn}^2 + y_{sn}^2)k_{vs})x_{sn}, (1 + (x_{sn}^2 + y_{sn}^2)k_{vs})y_{sn}) \quad (4.6)$$

Estas coordenadas são medidas no plano da imagem virtual, relativamente ao eixo óptico. As expressões de x_{vn} e de y_{vn} podem ser vistas como funções de duas variáveis. Se chamarmos D à função de distorção, podemos escrever:

$$(x_{vn}, y_{vn}) = D(x_{sn}, y_{sn}) \quad (4.7)$$

Podemos facilmente estudar o efeito que a função D provoca quando aplicada a uma imagem. Para tal, basta considerarmos uma imagem tipo grelha rectangular de linhas horizontais e verticais (figura 4.2), aplicar-lhe a função de distorção D a cada um dos pixels, e observar o resultado.

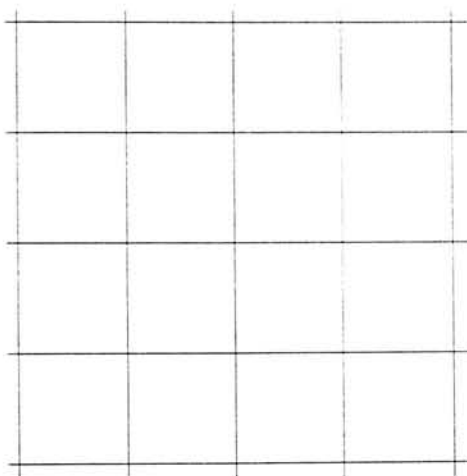


Figura 4.2. Grelha ortogonal utilizada para estudo da distorção introduzida pelos sistemas ópticos.

A imagem distorcida (figura 4.3), deve ser aproximadamente igual à que seria visível num HMD tradicional quando lhe é fornecida a grelha anterior.

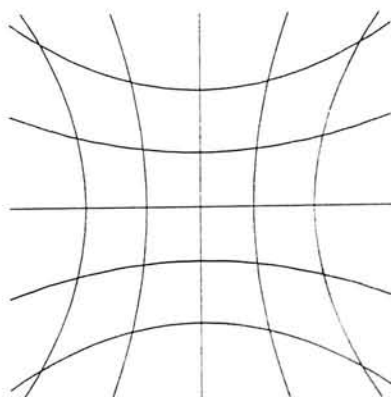


Figura 4.3. Grelha ortogonal distorcida pelo sistema óptico.

Para corrigir a distorção introduzida pelo sistema óptico, devemos pré-distorcer a imagem com a função inversa D^{-1} . Uma imagem pré-distorcida com a função inversa D^{-1} será correctamente apresentada ao utilizador, pois a pré-distorção D^{-1} compensa a distorção D causada pelo sistema óptico.

Há várias formas de representar esta função inversa em sistemas informáticos. Uma forma exacta não é praticável, mas uma aproximação polinomial é possível.

$$r_{sn} = r_{vn} + k_{sv} \cdot r_{vn}^3 \quad (4.8)$$

$$(x_{vn}, y_{vn}) = D^{-1}(x_{sn}, y_{sn}) \quad (4.9)$$

$$(x_{sn}, y_{sn}) = ((1 + (x_{vn}^2 + y_{vn}^2)k_{sv})x_{vn}, (1 + (x_{vn}^2 + y_{vn}^2)k_{sv})y_{vn}) \quad (4.10)$$

Outra possibilidade de representar a função D^{-1} num computador, consiste na utilização de uma tabela de duas entradas para cada um dos valores de saída. No entanto, as limitações de tamanho da tabela provavelmente obrigariam a interpolar valores entre as entradas da tabela para obter os valores desejados.

A imagem da figura 4.4 foi obtida distorcendo através da função D^{-1} a imagem da figura 4.2. Esta imagem deverá aparecer semelhante à da figura 4.2, quando vista através das lentes LEEP.

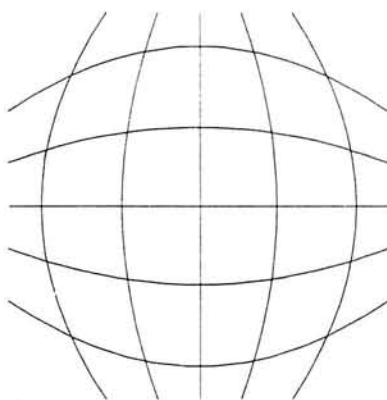


Figura 4.4. Grelha ortogonal pré-distorcida pela função de correcção D^{-1} .

Para verificar que a pré-distorção D^{-1} é compensada pela distorção D , podemos aplicar a distorção D à imagem da figura 4.4, e comparar o resultado (figura 4.5) com a imagem da figura 4.2.

As diferenças entre a imagem da figura 4.5 e a da figura 4.2, devem-se ao facto de a função D^{-1} não ser a inversa exacta da função D , mas sim uma aproximação de 3ª ordem. Uma vez que a função D também é uma aproximação, a comprovação da qualidade da função de pré-distorção proposta só pode ser efectuada usando as próprias lentes do HMD.

Para além da distorção geométrica, os sistemas ópticos introduzem também variações da intensidade luminosa, dependendo da zona do ângulo de visão. Esta variação deve-se ao facto de o aumento introduzido pelo sistema óptico não ser uniforme para todo o ângulo de visão. É possível compensar esta distorção, definindo uma função de correcção $B(x_{sn}, y_{sn})$, que compense a distorção introduzida.

encontra, o utilizador toma como base a sua distância interocular e utiliza os ângulos de convergência das linhas de visão.

Se conhecermos as especificações dos sistemas ópticos, e as posições relativas dos ecrãs e das lentes relativamente aos olhos do utilizador, podemos determinar a posição tridimensional do ponto C em função das coordenadas bidimensionais dos pontos da imagem. Este estudo permitirá definir o modelo adequado à geração de pares estereoscópicos de imagens a utilizar em HMD's.

Os centros dos ecrãs podem estar deslocados dos eixos ópticos uma determinada distância. Os dois eixos ópticos podem divergir um do outro um pequeno ângulo. O desvio entre os centros dos ecrãs e o ângulo de divergência dos eixos ópticos podem ser utilizados para ajustar correctamente a convergência, sem necessidade de afinações subjectivas. Conhecendo-se a posição e a orientação das imagens virtuais dos ecrãs relativamente aos olhos, podemos determinar a transformação a aplicar aos objectos modelados para obtermos o par de imagens correcto. Esta transformação pode ser definida como uma sequência de transformações tridimensionais conhecidas como a translação, a rotação, e a projecção perspectiva descentrada. Podemos desde já prever que a translação será sempre necessária, porque os dois olhos estão em posições diferentes no espaço. A rotação é necessária se os eixos ópticos não forem paralelos. Existirá também uma projecção que será descentrada se os centros dos ecrãs estiverem deslocados dos respectivos eixos ópticos.

4.3. HMD's baseados em CRTs

Os primeiros HMD's construídos utilizavam ecrãs do tipo CRT. Devido ao seu elevado peso e custo, estes ecrãs têm vindo a ser substituídos pelos mais recentes ecrãs de cristais líquidos (LCD). No entanto, nas aplicações em que é fundamental dispor de imagem com resolução elevada, continua a ser utilizada a aproximação baseada em ecrãs do tipo CRT.

Existem, no mercado, pequenos ecrãs CRT's, que podem ser utilizados para resoluções de cerca de 1000 linhas. Na construção da maioria dos HMD's actuais, baseados em CRT's, utilizam-se dois destes pequenos ecrãs, colocados dos lados da cabeça do utilizador (figura 4.7), e orientados para a frente. A imagem é depois reflectida até cada um dos olhos.

Utilizando elementos semi-reflectivos em frente dos olhos do utilizador, é possível manter a visão do mundo exterior, obtendo-se assim um STHMD. Graças ao brilho superior que os CRT's apresentam sobre os LCD's, esta arquitectura é bastante mais adequada à implementação de dispositivos STHMD, sobretudo se for previsível a utilização dos sistemas em ambientes fortemente iluminados ou à luz do dia.

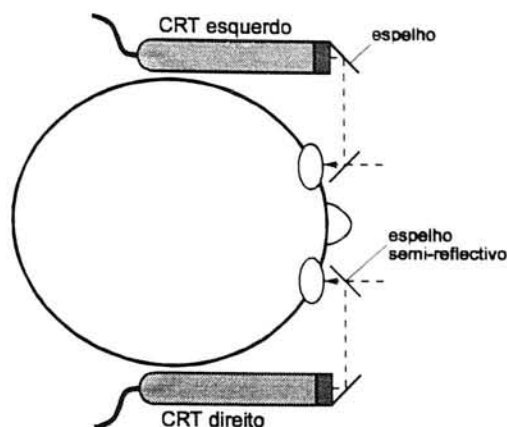


Figura 4.7. Diagrama básico de um head-mounted display baseado em CRT's.

Os ecrãs CRT utilizados são muito semelhantes aos utilizados nas câmaras de vídeo. Uma das maiores limitações da utilização destes ecrãs, é o cuidado necessário no projecto do HMD, devido à presença de tensões muito elevadas (milhares de volt), junto à cabeça do utilizador.

A ausência de cor, é o outro grande problema que os ecrãs CRT's apresentam na construção de HMD's. A concepção de um ecrã CRT policromático em miniatura, está para além das capacidades da tecnologia actualmente disponível. Não é actualmente possível fabricar a máscara de sombras (película metálica colocada directamente atrás do ecrã), com as dimensões mecânicas requeridas.

É possível, no entanto, obter uma imagem policromática utilizando apenas um ecrã CRT monocromático, colocando à sua frente um filtro LCD colorido. Apresentando sequencialmente linhas ou imagens correspondentes a cada uma das componentes de cor (vermelho, verde e azul), e comutando o filtro em sincronismo, o olho entende apenas uma única imagem que corresponde à combinação das várias componentes.

Esta técnica permite obter um ecrã policromático de elevada resolução sem necessidade de muito hardware adicional, mas requer uma frequência de varrimento de ecrã três vezes maior que a normal. Isto implica a utilização de componentes electrónicos mais dispendiosos, específicos para utilização em aplicações de alta velocidade. No entanto, mesmo com esta limitação, estão a ser desenvolvidos vários projectos baseados nesta técnica de filtragem. Um pequeno ecrã policromático com resolução VGA (640×480), foi desenvolvido pela Tektronix, e espera-se que possa ser produzido com preço inferior a USD \$1000. Este dispositivo deve ficar disponível comercialmente nos próximos anos, e permitirá a construção de HMD's com custos idênticos e padrões de qualidade superiores aos disponíveis actualmente.

Uma arquitectura alternativa, também baseada em CRT's, é um dispositivo fabricado pela Fake Space Labs, denominado *Boom*. A figura 4.8 apresenta o esquema básico de um destes dispositivos especiais.

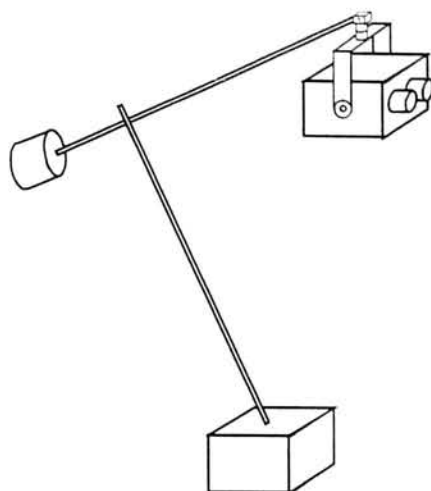


Figura 4. 8. Arquitectura básica do Boom

Graças à utilização de um contra-peso, os problemas devidos ao tamanho e ao peso dos CRT's são minimizados. Assim, é possível utilizar CRT's coloridos, e de muito maior resolução.

4.4. HMD's por Fibra Óptica

Utilizando cabos com agrupamentos de fibras ópticas, é possível criar, com relativa simplicidade, HMD's de elevada resolução (figura 4.9). Dispositivos deste tipo foram já implementados por companhias como Polhemus Labs e CAE-Link.

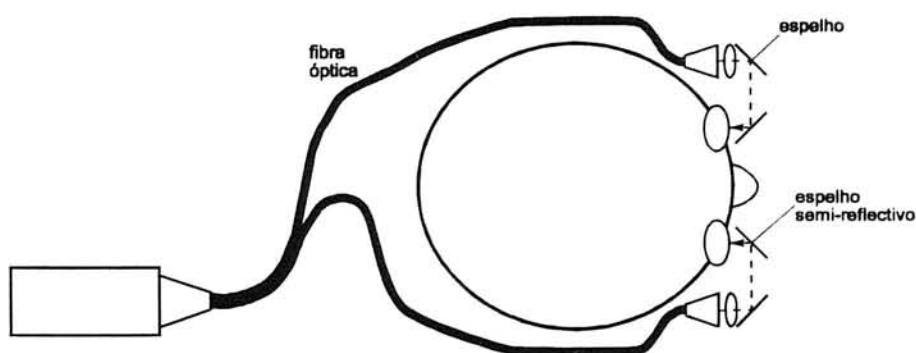


Figura 4.9. Diagrama básico de um head-mounted display por fibra óptica.

A imagem colorida de elevada resolução é projectada num dos extremos do cabo de fibras ópticas. Este cabo, com cerca de 3 centímetros de diâmetro, contém normalmente mais de um milhão de fibras ópticas. A imagem que surge no outro extremo do cabo é reflectida até aos olhos do observador.

Apesar de ser conceptualmente simples, a construção de um par destes cabos, opticamente correctos, tem um custo muito elevado, e fortemente dependente do número de fibras utilizadas. Por esta razão, o uso desta arquitectura tem sido restringido apenas aos HMD's mais avançados, destinados às aplicações onde o factor mais significativo não é o preço, mas sim a resolução e o baixo peso.

4.5. HMD's com Escrita Directa na Retina

A resolução pobre e o peso elevado são, actualmente, os dois maiores problemas técnicos dos dispositivos HMD's. Esta constatação fornece um forte incentivo na pesquisa de soluções tecnológicas alternativas que possam ser superiores aos CRTs e LCD's para utilização em HMD's.

Com o PrivateEye, ficou demonstrado que um pequeno e leve display monocromático de média resolução pode ser construído pela modulação dos elementos de um conjunto unidimensional, e utilizando um espelho móvel para gerar a imagem bidimensional virtual. O conjunto de LED's do PrivateEye é constituído por 280 elementos, cuja extorsão dá origem a uma imagem com 720×280 pontos de resolução. A resolução deste dispositivo pode ser melhorada pela utilização de mais elementos no conjunto de LED's, e pela modulação e varrimento mais rápidos. No entanto, a obtenção de cor é problemática porque seria necessário utilizar conjuntos de LED's de cor azul, e os LED's azuis, mesmo discretos, estão disponíveis apenas desde os últimos anos.

Os CRTs e os LCD's formam a imagem num conjunto bidimensional de elementos. O PrivateEye faz um varrimento unidimensional de um conjunto de elementos também unidimensional para formar a imagem bidimensional. Uma terceira possibilidade é utilizar um único elemento com possibilidade de modulação muito rápida, e produzir um varrimento segundo uma grelha bidimensional, para formar a imagem 2D. A ideia é semelhante àquela que deu origem à concepção interna dos CRTs. Para a construção de um dispositivo deste tipo, a utilização de laser parece adequada, uma vez que os emissores deste tipo de raios podem ser modelados com a velocidade requerida. Dispositivos de visualização deste tipo, têm sido alvo de estudo na University of North Carolina, Chapel Hill, NC [Holmgren92].

Uma grande vantagem de sistemas deste tipo é que permitem criar a imagem directamente na retina, não necessitando portanto de um ecrã ou plano de projecção para apresentar a imagem ao olho. A este processo chama-se *escrita directa na retina* ou DRW (Direct Retina Write). Pensa-se que a utilização de DRW permitirá eliminar muitas camadas de transmissão existentes entre a cena e a sensação visual, pois permite um percurso luminoso mais directo entre a informação e a visão. Por outro lado, esta simplificação de processos talvez permita a construção de HMD's de elevada resolução mas com pouco peso, de modo a permitir uma utilização confortável.

As tecnologias actuais que poderiam ser utilizadas para provocar um varrimento com um raio laser são: espelhos poligonais rotativos, espelhos orientados por galvanómetros, espelhos com deflecção piezoelétrica, e deflectores acusto-ópticos. Na tabela seguinte temos um resumo das características destas tecnologias de varrimento. O preço para cada dispositivo apresentado em US Dólares, é meramente indicativo e refere-se ao ano de 1992.

Tecnologia	Resolução (deflecção max)	Frequência	Vendedor	Custo	Comentários
Acusto-óptica	$>10^3$ (alguns graus)	$>10^4$ Hz	IntraAction	\$5000	Mais promissor para HMD's. Cor implica vários deflectores.
Polígono rotativo	$>10^3$ (dezenas de graus)	$>10^4$ Hz	Lincoln Laser	\$2000	Ideal para varrimentos rápidos. Dimensões impróprias para HMD's.
Galvanómetro	$>10^3$ ($>60^\circ$)	>100 Hz	General Scanning	\$1000	Adequado para deflecções lentas.
Piezoeléctrica	$>10^3$ ($<1^\circ$)	$>10^3$ Hz	Physik Instrumente	\$6000	Pequenas deflecções

Tabela 4.1. Resumo das tecnologias disponíveis para varrimento com raio laser.

4.6. Conclusão

A qualidade de imagem oferecida pelos HMD's correntes, baseados em ecrãs LCD, é muito inferior ao que seria desejável para aplicações de realidade virtual convincentes. É portanto de esperar que apareçam em breve, no mercado, HMD's com arquitecturas diversas. As arquitecturas alternativas ao uso de LCD's, que surgem como mais promissoras, são as baseadas em CRT's, as que utilizam fibra óptica, e as que efectuem escrita directa na retina (DRW). A tabela 4.2 apresenta um resumo das características fundamentais destas tecnologias.

Tecnologia	Resolução	Peso/Mobilidade	Custo	Comentários
LCD	baixa	leve / boa	reduzido	Solução vulgarmente disponível.
CRT	razoável	elevado / boa	médio	Solução disponível. Cor ainda é problema.
Fibra Óptica	boa	leve / má	elevado	Já implementado. Preço muito elevado.
DRW	boa	leve / boa	elevado	Muito promissor. Ainda em desenvolvimento.

Tabela 4.2. Tecnologias para implementação de HMD's.

Apesar dos desenvolvimentos de algumas destas técnicas indicarem ser possível obter um HMD leve e de elevada resolução, ainda não é possível a sua construção com custos adequados à comercialização em grande escala. Não é ainda possível indicar, com algum grau de certeza, qual será a técnica dominante na construção dos futuros HMD's. É até possível que se mantenha o domínio das arquitecturas baseadas em LCD's, por longos anos, verificando-se apenas uma evolução na resolução e na qualidade destes dispositivos.

Capítulo 5.

Síntese de Pares de Imagens

Existem vários métodos publicados para calcular as imagens direita e esquerda de um par estereoscópico. Em [Grotch83] e em [Wixson83], descrevem-se métodos para criação de pares estereoscópicos de imagens que se baseiam em rotações. Estes métodos, e as distorções que introduzem, foram estudados em [Hodges90b]. Em [Lipscom79] propõe-se a aproximação das rotações por deformações (shears). Vários autores [Akka91], [Southar92], [Penna86] aconselham a utilização de um plano de projecção único para a obtenção das duas imagens, ou a utilização de translações. Estas aproximações derivam do algoritmo das lentes com eixos paralelos, que foi originalmente apresentado em [Baker87]. Mais recentemente foi apresentada uma alternativa para geração de imagens para dispositivos Visualizadores Montados na Cabeça (HMD's) [Robinet91], baseada em projecções descentradas, cujos eixos podem não ser paralelos. Em [Hodges88] e em [Hodges90a], encontramos excelentes descrições e análises dos principais métodos publicados. De [Hodges88] podemos também retirar algumas citações de publicações anteriores que permitem rever a evolução histórica dos algoritmos utilizados:

- 1979 "As imagens estereoscópicas complementares são geradas por rotações sobre um eixo vertical que produzem, quando visualizadas, uma efectiva sensação de profundidade de campo."
- 1979 "As transformações estereoscópicas ... são shears e não rotações como se poderia esperar."
- 1983 "As duas vistas são ligeiramente diferentes..."
- 1986 "As duas vistas de um par estereoscópico podem ser produzidas utilizando duas posições diferentes para os olhos e um único plano de projecção."
- 1986 "...pode ser apropriado rodar o objecto para preparar a vista para o segundo olho, em vez de o transladar."
- 1987 "Os pares estereoscópicos gerados por computador são por vezes produzidos pela rotação de alguns graus de um objecto ou da cena. Nós desaconselhamos fortemente este procedimento."
- 1989 "Nós recomendamos a produção das duas vistas como perspectivas deslocadas lateralmente ao longo do eixo x, seguidas de um deslocamento lateral das imagens resultantes para obter a convergência correcta."

Neste capítulo propomo-nos descrever e analisar os diversos métodos, e tirar conclusões sobre a correcção destes. Vamos estudar o problema para duas situações

típicas de visualização: o caso da apresentação de ambas as imagens num mesmo ecrã que possa ser considerado plano; e o caso dum dispositivo tipo Head Mounted Display. O estudo aqui apresentado vai mais longe do que os estudos comparativos anteriores, porque engloba as duas situações de visualização mais comuns, e a aplicação de todos os métodos de geração de imagem com projecções paralela e perspectiva.

Em qualquer dos casos, a geração da imagem será conseguida através de uma descrição tridimensional da cena a visualizar. Ao espaço tridimensional que contém a cena chamaremos espaço de modelação, e às respectivas coordenadas cartesianas chamaremos coordenadas de modelação.

Consideremos um sistema com o observador no ponto $E=(0,0,0)$ e um plano de imagem paralelo ao plano x-y e colocado a uma distância z_{pp} segundo o eixo z. As conclusões que serão tiradas a partir do estudo efectuado para este sistema de coordenadas serão válidas para qualquer outro. Vamos também considerar coordenadas de ecrã normalizadas, para que os resultados sejam aplicáveis a qualquer dispositivo de visualização. Portanto, um ponto do ecrã terá coordenadas (x,y) em que $-1 \leq x \leq 1$ e $-1 \leq y \leq 1$.

Começamos por rever o caso de geração de imagens planas e depois estudaremos o caso estereoscópico. Em qualquer dos casos vamos considerar a utilização de projecção paralela e de projecção perspectiva.

5.1. Geração de Imagem Monoscópica

Comecemos por considerar um ponto $P_0 = (x_0, y_0, z_0)$, e uma câmara com ângulos de abertura fov_x e fov_y . Num sistema monocular, a projecção de P_0 no plano da imagem seria o ponto Q com coordenadas normalizadas de imagem (x,y):

5.1.1. Projecção Paralela

Considerando que a direcção de observação é paralela ao eixo dos z, para obter a projecção paralela do ponto $P_0 = (x_0, y_0, z_0)$ no plano $z = z_{pp}$, basta eliminar a componente z das coordenadas de P_0 :

Em coordenadas normalizadas de imagem o ponto $Q = (x, y)$ será:

$$x = \frac{x_0}{\tan \frac{fov_x}{2}} \quad (5.1)$$

$$y = \frac{y_0}{\tan \frac{fov_y}{2}} \quad (5.2)$$

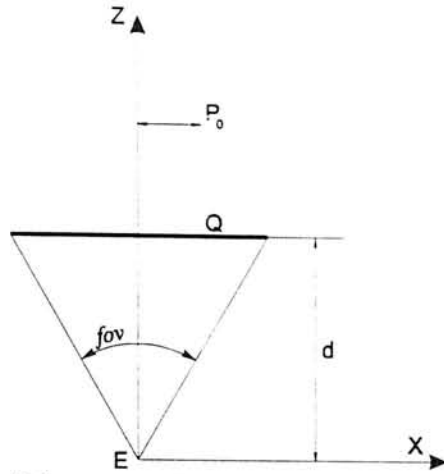


Figura 5.1. Obtenção da imagem de um ponto, por projecção paralela.

Ou, em notação matricial, e utilizando coordenadas homogêneas:

$$\begin{bmatrix} x & y \end{bmatrix} = \begin{bmatrix} x_0 & y_0 & z_0 \end{bmatrix} \cdot \begin{bmatrix} \cot \frac{fov}{2} & 0 \\ 1 & \cot \frac{fov}{2} \\ 0 & 0 \end{bmatrix} \quad (5.3)$$

5.1.2. Projecção Perspectiva

Estando o observador colocado na origem das coordenadas, podemos aplicar uma transformação perspectiva centrada no ponto (0,0,0).

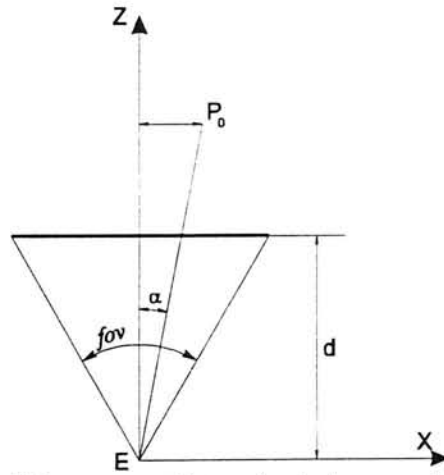


Figura 5.2. Obtenção da imagem de um ponto, por projecção perspectiva.

O ponto $Q = (x, y)$ será agora, em coordenadas normalizadas de imagem:

$$x = \frac{x_0}{z_0} \cdot \frac{1}{\tan \frac{fov}{2}} \quad (5.4)$$

$$y = \frac{y_0}{z_0} \cdot \frac{1}{\tan \frac{fov_y}{2}} \quad (5.5)$$

Ou, em notação matricial, e utilizando coordenadas homogéneas:

$$P = \begin{bmatrix} \cot \frac{fov_y}{2} & 0 & 0 & 0 \\ 0 & \cot \frac{fov_y}{2} & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \cot \frac{fov_y}{2} \end{bmatrix} \quad (5.6)$$

5.2. Geração de imagens estereoscópicas

Consideremos agora um sistema estereoscópico, com os centros de projecção E_l e E_r , tais que:

$$E_r = -E_l = (x_E, 0, 0) \quad (5.7)$$

Para cada método de geração de imagem vamos obter os pontos imagem de um ponto genérico P_0 , para cada uma das vistas do par estereoscópico. À imagem do ponto P_0 na vista destinada ao olho esquerdo chamaremos P_{il} , e à imagem do ponto P_0 na vista destinada ao olho direito chamaremos P_{ir} . As coordenadas normalizadas destes pontos imagem serão:

$$P_{il} = (x_{il}, y_{il}) \quad (5.8)$$

$$P_{ir} = (x_{ir}, y_{ir}) \quad (5.9)$$

Consideraremos sempre que as coordenadas normalizadas de imagem variam entre -1 e 1, e que a sua origem se encontra no centro da imagem respectiva. A paralaxe de imagem para o ponto P_0 pode ser definida como a distância entre as imagens direita e esquerda deste mesmo ponto. A paralaxe é então um vector bidimensional R , tal que:

$$\mathbf{R} = (R_x, R_y) = P_{il} - P_{ir} \quad (5.10)$$

5.2.1. Utilização de Rotações

A ideia da utilização de rotações para obter pares estereoscópicos de imagens surgiu pela constatação de que, quando um observador olha para um objecto, os seus olhos estão focados num ponto particular desse objecto. Na figura 5.3 podemos ver o que acontece quando o observador olha para um cubo (a aresta com risco grosso representa a face da frente do cubo).

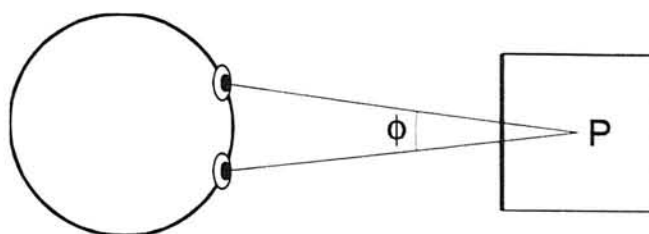


Figura 5.3.

Utilização de pontos de observação distintos com direcções convergentes.

Os olhos do observador estão rodados para dentro, de modo a que as duas linhas de visão se cruzem no objecto.

Surgiu então a ideia de rodar a cena para a direita um ângulo $\phi / 2$ em torno de um ponto P, para obter a vista do olho direito, e rodar a cena para a esquerda um ângulo $\phi / 2$ em torno do ponto P, para obter a vista do olho esquerdo (figura 5.4).

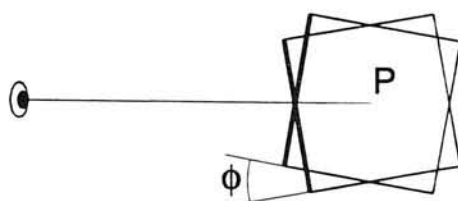


Figura 5.4.

Utilização de um único ponto de observação, com objectos rodados.

Vamos considerar o centro de projecção em $(0,0,0)$, e o plano de projecção paralelo ao plano x-y e colocado a uma distância d do centro de projecção segundo o eixo z.

Vamos aplicar rotações em torno do eixo vertical que passa no ponto $(0,0,r)$. A rotação será de $-\phi / 2$ para obtermos os pontos que darão origem à imagem para o olho direito e de $+\phi / 2$ para obtermos a imagem para o olho esquerdo (figura 5.5).

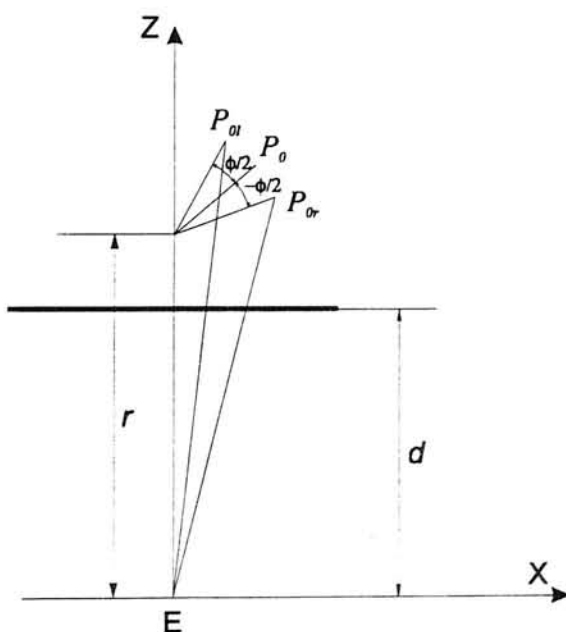


Figura 5.5.

Utilização de rotações.

O ponto $P_0 = (x_0, y_0, z_0)$ dará assim origem aos pontos $P_{0l} = (x_{0l}, y_{0l}, z_{0l})$ e $P_{0r} = (x_{0r}, y_{0r}, z_{0r})$, tais que:

$$x_{0l} = x_0 \cdot \cos(\phi/2) - (z_0 - r) \cdot \sin(\phi/2) \quad (5.11)$$

$$x_{0r} = x_0 \cdot \cos(\phi/2) + (z_0 - r) \cdot \sin(\phi/2) \quad (5.12)$$

$$y_{0l} = y_{0r} = y_0 \quad (5.13)$$

$$z_{0l} = (z_0 - r) \cdot \cos(\phi/2) + x_0 \cdot \sin(\phi/2) + r \quad (5.14)$$

$$z_{0r} = (z_0 - r) \cdot \cos(\phi/2) - x_0 \cdot \sin(\phi/2) + r \quad (5.15)$$

Para obtermos a imagem do P_0 , vamos agora projectar os pontos P_{0l} e P_{0r} no plano de projecção. Conforme foi dito, esta projecção pode ser paralela ou perspectiva. Vamos estudar os dois casos em separado.

5.2.1.1. Rotações com Projecção Paralela

Aplicando uma projecção paralela, conforme especificado em 5.1.1, e segundo a direcção do eixo z ao ponto P_{0l} , obtemos as coordenadas normalizadas de imagem:

$$x_{il} = \frac{x_0 \cdot \cos(\phi/2) - (z_0 - r) \cdot \sin(\phi/2)}{\tan \frac{fov_x}{2}} \quad (5.16)$$

$$y_{il} = \frac{y_0}{\tan \frac{fov_y}{2}} \quad (5.17)$$

E, aplicando a mesma transformação ao ponto P_{0r} , obtemos:

$$x_{ir} = \frac{x_0 \cdot \cos(\phi/2) + (z_0 - r) \cdot \sin(\phi/2)}{\tan \frac{fov_x}{2}} \quad (5.18)$$

$$y_{ir} = \frac{y_0}{\tan \frac{fov_y}{2}} \quad (5.19)$$

Como $y_{ir} = y_{il}$, não existe paralaxe vertical. A paralaxe horizontal de imagem é:

$$R_{ix} = x_{ir} - x_{il} = \frac{2}{\tan \frac{fov_x}{2}} \cdot (z_0 - r) \cdot \sin(\phi/2) \quad (5.20)$$

Como era de esperar, para valores de ϕ pequenos e positivos, a paralaxe (R_{ix}) cresce com a profundidade z_0 .

5.2.1.2. Rotações com Projecção Perspectiva

Aplicando a projecção perspectiva, especificada em 5.1.2, aos pontos P_{ol} e P_{or} definidos pelas equações 5.11 a 5.14, obtemos as coordenadas normalizadas dos pontos imagem, que são para a vista direita:

$$x_{ir} = \frac{1}{\tan \frac{fov_x}{2}} \cdot \frac{x_0 \cdot \cos(\frac{\phi}{2}) + (z_0 - r) \cdot \sin(\frac{\phi}{2})}{(z_0 - r) \cdot \cos(\frac{\phi}{2}) - x_0 \cdot \sin(\frac{\phi}{2}) + r} \quad (5.21)$$

$$y_{ir} = \frac{1}{\tan \frac{fov_y}{2}} \cdot \frac{y_0}{(z_0 - r) \cdot \cos(\frac{\phi}{2}) - x_0 \cdot \sin(\frac{\phi}{2}) + r} \quad (5.22)$$

e para a vista esquerda:

$$x_{il} = \frac{1}{\tan \frac{fov_x}{2}} \cdot \frac{x_0 \cdot \cos(\frac{\phi}{2}) - (z_0 - r) \cdot \sin(\frac{\phi}{2})}{(z_0 - r) \cdot \cos(\frac{\phi}{2}) + x_0 \cdot \sin(\frac{\phi}{2}) + r} \quad (5.23)$$

$$y_{il} = \frac{1}{\tan \frac{fov_y}{2}} \cdot \frac{y_0}{(z_0 - r) \cdot \cos(\frac{\phi}{2}) + x_0 \cdot \sin(\frac{\phi}{2}) + r} \quad (5.24)$$

Com os valores das equações 5.21 e 5.23, e sabendo que

$$2 \cdot \sin \frac{\phi}{2} \cdot \cos \frac{\phi}{2} = \sin \phi \quad (5.25)$$

podemos determinar a paralaxe de imagem para o ponto P_0 .

$$R_{ix} = \frac{1}{\tan \frac{fov_x}{2}} \cdot \frac{(x_0 + (z_0 - r)^2) \cdot \sin \phi + 2 \cdot r \cdot (z_0 - r) \cdot \sin(\frac{\phi}{2})}{((z_0 - r) \cdot \cos(\frac{\phi}{2}))^2 - (x_0 \cdot \sin(\frac{\phi}{2}) + r)^2} \quad (5.26)$$

A paralaxe vertical de imagem R_{iy} é igual a $y_{ir} - y_{il}$, e pode ser calculada utilizando os valores das equações 5.22 e 5.24.

$$R_{iy} = \frac{y_0}{\tan \frac{fov_y}{2}} \cdot \left(\frac{1}{(z_0 - r) \cos(\frac{\phi}{2}) - x_0 \sin(\frac{\phi}{2}) + r} - \frac{1}{(z_0 - r) \cos(\frac{\phi}{2}) + x_0 \sin(\frac{\phi}{2}) + r} \right) \quad (5.27)$$

$$R_{iy} = \frac{1}{\tan \frac{fov_y}{2}} \cdot \frac{2 \cdot x_0 \cdot y_0 \cdot \sin(\frac{\phi}{2})}{((z_0 - r) \cdot \cos(\frac{\phi}{2}) + r)^2 - x_0^2 \cdot \sin^2(\frac{\phi}{2})} \quad (5.28)$$

Analisando o numerador desta expressão, podemos constatar que a paralaxe vertical é nula sempre que pelo menos uma das coordenadas, x_0 ou y_0 , seja zero. Além disso, podemos também provar que a paralaxe vertical é uma função crescente com o produto $x_0 \cdot y_0$, e decrescente com z_0 . Uma vez que a paralaxe vertical aumenta quando o denominador da fracção tende para zero, podemos prever o aparecimento de valores de paralaxe muito elevados quando:

$$\left((z_0 - r) \cdot \cos(\phi/2) + r\right)^2 \equiv x_0^2 \cdot \sin^2(\phi/2) \quad (5.29)$$

ou seja, quando x_0 se aproxima de:

$$\frac{(z_0 + r) \cdot \cos(\phi/2) + r}{\sin(\phi/2)} \quad (5.30)$$

Como o ângulo de rotação ϕ é normalmente pequeno, então $\cos(\phi/2) \equiv 1$. Podemos então dizer que a paralaxe vertical é muito elevada quando:

$$x_0 \equiv \frac{z_0 + r \cdot (1 - \cos(\phi/2))}{\sin(\phi/2)} \quad (5.31)$$

Como ϕ é sempre pequeno, $\sin(\phi/2)$ é também pequeno. Isto implica que, para que a expressão 5.31 seja verificada, x_0 tem de ser muito maior que z_0 . Isto significa que a paralaxe vertical apresentará valores mais elevados para os objectos próximos do ponto de observação que apareçam nos bordos da imagem.

Valores de ϕ mais pequenos originam menor paralaxe vertical, mas também diminuem a paralaxe horizontal que produz o efeito estereoscópico.

5.2.2. Utilização de Shears

Como o ângulo utilizado na rotação para obter as imagens esquerda e direita é muito pequeno (normalmente $\phi \leq 7^\circ$), tem sido sugerida a utilização nas rotações das aproximações:

$$\sin(\phi/2) = \phi/2 \quad (5.32)$$

e

$$\cos(\phi/2) = 1 \quad (5.33)$$

O ponto $P_0 = (x_0, y_0, z_0)$ daria agora origem ao ponto $P_{0l} = (x_{0l}, y_{0l}, z_{0l})$ que seria usado para obter a imagem esquerda:

$$x_{0l} = x_0 - (z_0 - r) \cdot \phi/2 \quad (5.34)$$

$$y_{0l} = y_0 \quad (5.35)$$

$$z_{0l} = z_0 + x_0 \cdot \frac{\phi}{2} \quad (5.36)$$

e ao ponto $P_{0r} = (x_{0r}, y_{0r}, z_{0r})$ que seria usado para obter a imagem direita:

$$x_{0r} = x_0 + (z_0 - r) \cdot \frac{\phi}{2} \quad (5.37)$$

$$y_{0r} = y_0 \quad (5.38)$$

$$z_{0r} = z_0 - x_0 \cdot \frac{\phi}{2} \quad (5.39)$$

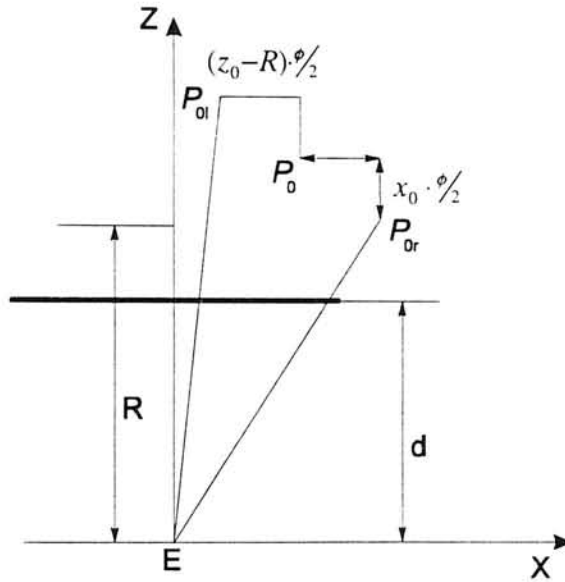


Figura 5.6. Utilização de shears.

Esta aproximação pode ser vista como uma deformação (shear) ao longo do eixo z, seguida de uma outra deformação em x.

5.2.2.1. Shears com Projecção Paralela

Aplicando uma projecção paralela, conforme especificado em 5.1.1, e segundo a direcção do eixo z ao ponto P_{0l} , obtemos as coordenadas de imagem:

$$x_{il} = \frac{x_0 - (z_0 - r) \cdot \frac{\phi}{2}}{\tan \frac{fov_x}{2}} \quad (5.40)$$

$$y_{il} = \frac{y_0}{\tan \frac{fov_y}{2}} \quad (5.41)$$

E, aplicando a mesma transformação ao ponto P_{0r} , obtemos:

$$x_{ir} = \frac{x_0 + (z_0 - r) \cdot \phi/2}{\tan \frac{fov_y}{2}} \quad (5.42)$$

$$y_{ir} = \frac{y_0}{\tan \frac{fov_y}{2}} \quad (5.43)$$

Uma vez que no processo de visualização se vai utilizar deslocamento horizontal das imagens, podemos eliminar o factor $r \cdot \phi/2$, pois é constante para toda a cena. As coordenadas horizontais das imagens do ponto P_0 , seriam agora:

$$x_{il} = \frac{x_0 - z_0 \cdot \phi/2}{\tan \frac{fov_y}{2}} \quad (5.44)$$

$$x_{ir} = \frac{x_0 + z_0 \cdot \phi/2}{\tan \frac{fov_y}{2}} \quad (5.45)$$

Como $y_{ir} = y_{il}$, não existe paralaxe vertical. A paralaxe horizontal de imagem é:

$$R_{ix} = x_{ir} - x_{il} = \frac{z_0 \cdot \phi}{\tan \frac{fov_y}{2}} \quad (5.46)$$

Como era de esperar, para valores de ϕ pequenos e positivos, a paralaxe (R_{ix}) cresce com a profundidade z_0 .

5.2.2.2. Shears com Projecção Perspectiva

Aplicando a projecção perspectiva, especificada em 5.1.2, aos pontos P_{0l} e P_{0r} definidos pelas equações 5.34 a 5.38, obtemos as coordenadas normalizadas dos pontos imagem, que são para a vista direita:

$$x_{ir} = \frac{1}{\tan \frac{fov_y}{2}} \cdot \frac{x_0 + (z_0 - r) \cdot \phi/2}{z_0 - x_0 \cdot \phi/2} \quad (5.47)$$

$$y_{ir} = \frac{1}{\tan \frac{fov_y}{2}} \cdot \frac{y_0}{z_0 - x_0 \cdot \phi/2} \quad (5.48)$$

e para a vista esquerda:

$$x_{il} = \frac{1}{\tan \frac{fov_y}{2}} \cdot \frac{x_0 - (z_0 - r) \cdot \phi/2}{z_0 + x_0 \cdot \phi/2} \quad (5.49)$$

$$y_{il} = \frac{1}{\tan \frac{fov_y}{2}} \cdot \frac{y_0}{z_0 + x_0 \cdot \frac{\phi}{2}} \quad (5.50)$$

Com os valores das equações 5.47 e 5.49, podemos determinar a paralaxe de imagem para o ponto P_0 .

$$R_{ix} = \frac{\phi}{\tan \frac{fov_y}{2}} \cdot \frac{x_0 + (z_0 - r)^2 + r \cdot (z_0 - r)}{(z_0 - r)^2 - (x_0 \cdot \frac{\phi}{2} + r)^2} \quad (5.51)$$

A paralaxe vertical de imagem R_{iy} é igual a $y_{ir} - y_{il}$, e pode ser calculada utilizando os valores das equações 5.48 e 5.50:

$$R_{iy} = \left(\frac{y_0}{z_0 - x_0 \cdot \frac{\phi}{2}} - \frac{y_0}{z_0 + x_0 \cdot \frac{\phi}{2}} \right) \cdot \frac{1}{\tan \frac{fov_y}{2}} \quad (5.52)$$

$$R_{iy} = \frac{2 \cdot x_0 \cdot y_0 \cdot \frac{\phi}{2}}{z_0^2 - x_0^2 \cdot (\frac{\phi}{2})^2} \cdot \frac{1}{\tan \frac{fov_y}{2}} \quad (5.53)$$

Analisando o numerador desta expressão, podemos constatar que a paralaxe vertical é nula sempre que pelo menos uma das coordenadas, x_0 ou y_0 , seja zero. Além disso, podemos também provar que a paralaxe vertical é uma função crescente com o produto $x_0 \cdot y_0$, e decrescente com z_0 . Uma vez que a paralaxe vertical aumenta quando o denominador da fracção tende para zero, podemos prever o aparecimento de valores de paralaxe muito elevados quando:

$$z_0^2 \cong x_0^2 \cdot (\frac{\phi}{2})^2 \quad (5.54)$$

Podemos então dizer que a paralaxe vertical é muito elevada quando x_0 se aproxima de:

$$\frac{z_0}{\frac{\phi}{2}} \quad (5.55)$$

Como ϕ é normalmente muito pequeno, para que a expressão 5.55 seja verificada, x_0 tem de ser muito maior que z_0 . Isto significa que a paralaxe vertical apresentará valores mais elevados para os objectos próximos do ponto de observação que apareçam nos bordos da imagem.

Tal como acontecia no caso da utilização de rotações com projecção prespectiva, valores de ϕ mais pequenos originam menor paralaxe vertical, mas também diminuem a paralaxe horizontal que produz o efeito estereoscópico.

5.2.3. Utilização de Direcções de Observação Comuns

Outra aproximação possível para produzir um par estereoscópico de imagens consiste na utilização de dois pontos de observação distintos, e um plano de projecção único. As imagens obtidas por este método correspondem àquelas que são obtidas por duas câmaras paralelas, ligeiramente deslocadas horizontalmente uma da outra (figura 5.7).

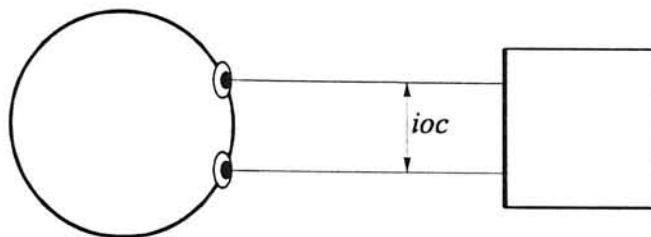


Figura 5.7.

Utilização de dois pontos de observação distintos.

Os mesmos resultados podem ser obtidos utilizando uma única câmara e deslocando os objectos horizontalmente para obter as duas vistas. A imagem esquerda é obtida a partir do deslocamento para a direita de todos os objectos. A imagem direita é obtida a partir do deslocamento para a esquerda de todos os objectos (figura 5.8).

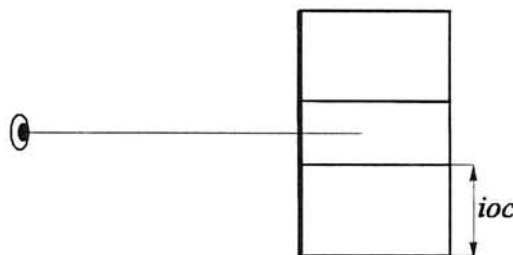


Figura 5.8.

Utilização de um único ponto de observação e objectos deslocados lateralmente.

Vamos continuar a considerar dois sistemas de coordenadas de imagem coplanares e independentes, com as origens colocadas nos eixos das respectivas imagens (figura 5.9). O ponto P_0 será projectado em cada um dos ecrãs, dando origem ao ponto P_{ir} na imagem do lado direito, e ao ponto P_{il} na imagem do lado esquerdo.

$$P_{ir} = (x_{ir}, y_{ir})$$

$$P_{il} = (x_{il}, y_{il})$$

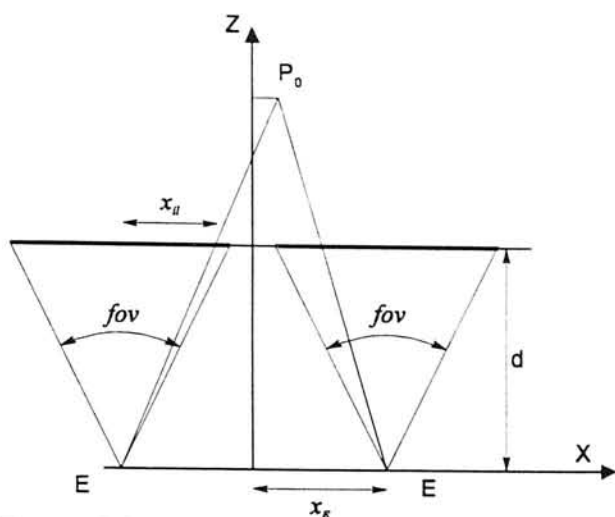


Figura 5.9.

Modelo para obtenção de imagens estereoscópicas, utilizando câmaras paralelas.

Veremos que é possível obter os mesmos resultados pela projecção, num mesmo plano, dos pontos P_{or} e P_{ol} (figura 5.10). O ponto P_{or} é obtido por deslocamento para a esquerda do ponto P_0 , e a sua projecção faz parte da imagem direita. O ponto P_{ol} é obtido por deslocamento para a direita do ponto P_0 , e a sua projecção faz parte da imagem esquerda.

$$P_{or} = (x_0 - x_E, y_0, z_0) \quad (5.56)$$

$$P_{ol} = (x_0 + x_E, y_0, z_0) \quad (5.57)$$

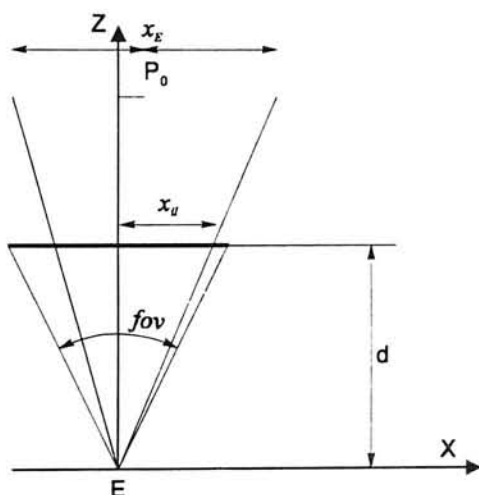


Figura 5.10.

Modelo para obtenção de imagens estereoscópicas, utilizando câmaras paralelas.

5.2.3.1. Direcções de Observação Comuns com Projecção Paralela

Aplicando uma projecção paralela, conforme especificado em 5.1.1, e segundo a direcção do eixo z, aos pontos P_{0l} e P_{0r} definidos pelas equações 5.56 e 5.57, obtemos as coordenadas normalizadas de imagem:

$$x_{il} = \frac{x_0 + x_E}{\tan \frac{fov_l}{2}} \quad (5.58)$$

$$x_{ir} = \frac{x_0 - x_E}{\tan \frac{fov_r}{2}} \quad (5.59)$$

$$y_{il} = y_{ir} = \frac{y_0}{\tan \frac{fov_y}{2}} \quad (5.60)$$

Para obter o vector de paralaxe de imagem, basta aplicar na equação 5.10 as expressões 5.58, 5.59 e 5.60:

$$\mathbf{R}_i = (R_{ix}, R_{iy}) = \left(\frac{2 \cdot x_E}{\tan \frac{fov_x}{2}}, 0 \right) \quad (5.61)$$

Apesar de não ser nula, a paralaxe horizontal é constante para todo o espaço tridimensional. Isto significa que a sensação estereoscópica de profundidade, resultante da visualização de imagens obtidas por este método, é fixa. A imagem aparece toda no mesmo plano. Este plano poderá estar à frente ou atrás do plano do ecrã, dependendo das condições de visualização. A utilização de translações com projecção paralela não é portanto útil para criar pares estereoscópicos de imagem.

5.2.3.2. Direcções de Observação Comuns com Projecção Perspectiva

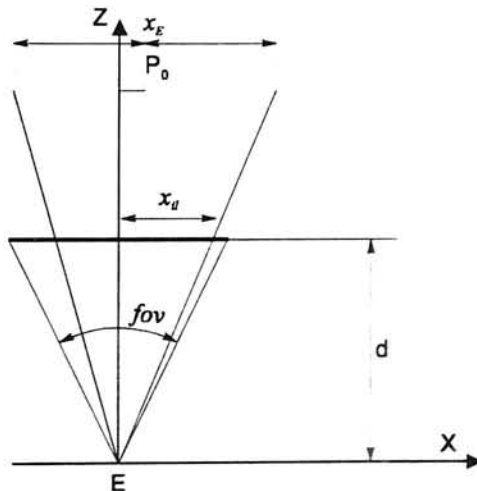


Figura 5.11. Modelo para obtenção de imagens estereoscópicas, utilizando câmaras paralelas.

Aplicando a projecção perspectiva, especificada em 5.1.2, aos pontos P_{ol} e P_{or} definidos pelas equações 5.56 e 5.57, obtemos as coordenadas normalizadas dos pontos imagem, que são:

$$x_{il} = \frac{x_0 + x_E}{z_0} \cdot \frac{1}{\tan \frac{fov_x}{2}} \quad (5.62)$$

$$x_{ir} = \frac{x_0 - x_E}{z_0} \cdot \frac{1}{\tan \frac{fov_x}{2}} \quad (5.63)$$

$$y_{il} = y_{ir} = \frac{y_0}{z_0} \cdot \frac{1}{\tan \frac{fov_y}{2}} \quad (5.64)$$

Estes resultados podem também ser obtidos pela aplicação de duas projecções perspectivas ao ponto P_0 . Uma projecção centrada no ponto E_r permite obter o ponto P_{ir} que pertence à imagem do lado direito, e uma projecção centrada no ponto E_l permite obter o ponto P_{il} da imagem do lado esquerdo (figura 5.12).

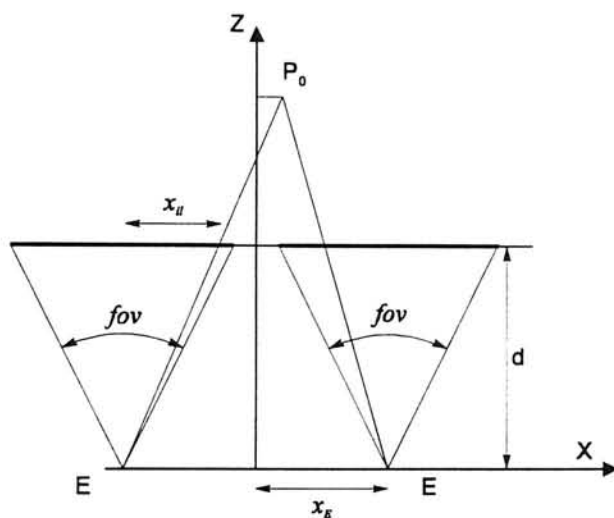


Figura 5.12. Modelo para obtenção de imagens estereoscópicas, utilizando câmaras paralelas.

As coordenadas de imagem dos pontos P_{ir} e P_{il} obtidas por estas projecções são igualmente as já apresentadas em 5.62, 5.63 e 5.64.

Para obter o vector de paralaxe de imagem, basta aplicar na equação 5.10 as expressões 5.62, 5.63 e 5.64:

$$\mathbf{R}_i = (R_{ix}, R_{iy}) = \left(\frac{-2}{\tan \frac{fov_x}{2}} \cdot \frac{x_E}{z_0}, 0 \right) \quad (5.65)$$

5.3. Conclusão

A utilização de rotações ou de shears, com projecção perspectiva, provoca o aparecimento de paralaxe vertical entre as imagens. No capítulo seguinte, demonstraremos ainda que a aplicação das imagens, obtidas por estes métodos, a um ecrã plano, provoca distorção de profundidade. Podemos portanto, desde já, desaconselhar a utilização destes métodos para gerar imagens estereoscópicas para ecrãs planos.

Quando utilizadas com projecção paralela, tanto as rotações como as deformações (shears) apresentam valores nulos para a paralaxe vertical. As deformações apresentam sobre as rotações a vantagem da maior facilidade de cálculo.

A utilização de direcções de observação comuns ou translações, com projecção perspectiva, para obter as duas vistas de uma imagem estereoscópica, não produz paralaxe vertical, ao contrário do que acontece com as rotações e as deformações. Este deverá ser portanto o método mais aconselhado para gerar imagens estereoscópicas com perspectiva, para ecrãs planos.

A utilização de direcções de observação comuns com projecção paralela não é útil para gerar imagens estereoscópicas, uma vez que não permite obter imagens com paralaxe.

Outras críticas e conclusões mais definitivas, sobre a qualidade dos diversos métodos, deverão ser possíveis após o capítulo seguinte. Aí analisaremos o resultado da visualização num ecrã plano, das imagens criadas por cada um dos métodos considerados.

Capítulo 6.

Visualização de Imagens Estereoscópicas

No capítulo anterior analisámos vários modelos de geração de imagens estereoscópicas. Para cada modelo deduzimos a relação entre os pontos tridimensionais modelados e os pontos imagem correspondentes. Vamos agora aplicar as imagens geradas num ecrã plano e estudar a relação entre as imagens e a cena virtual entendida pelo observador. Utilizando as expressões, deduzidas no capítulo anterior, que nos dão as coordenadas e a paralaxe de imagem, vamos ainda determinar a relação entre a cena virtual e a cena modelada, para cada modelo de geração de imagem e dependendo das condições de visualização. Esta relação vai permitir o julgamento da qualidade geométrica do modelo de geração de imagem. O modelo ideal deverá permitir ao observador entender uma cena virtual geometricamente semelhante à cena modelada.

Um sistema de visualização estereoscópico utiliza duas imagens que devem ser entregues ao observador/participante. Cada uma das imagens será entregue ao olho respectivo. O espectador utiliza a informação contida nas imagens para gerar internamente uma representação virtual da cena visualizada.

Cada ponto da imagem esquerda deverá possuir um par respectivo na imagem direita, e vice-versa. As excepções correspondem às imagens dos pontos tridimensionais visíveis apenas por um dos olhos. Cada par de pontos origina um ponto virtual, cuja posição no espaço imaginário depende apenas da posição dos pontos dentro da imagem, e desta relativamente aos olhos do observador. A este espaço imaginário chamaremos espaço virtual, e a um ponto pertencente a este espaço chamaremos ponto virtual.

Para determinarmos a posição de um ponto virtual genérico $P'_i=(x'_i, y'_i, z'_i)$, consideremos a origem do sistema de coordenadas do mundo virtual, colocada sobre o segmento de recta que une os dois olhos do espectador, e exactamente à mesma distância de ambos. O eixo y coincide com a direcção vertical, e o eixo x aponta para a direita.

Consideremos ainda que o observador olha segundo a direcção do eixo z , e que as imagens são colocadas no mesmo plano x - y (mesmo ecrã). Cada imagem vai ser mapeada uniformemente no ecrã, ficando a ocupar uma zona rectangular de área $L \times H$. Para controlar a paralaxe, vamos deslocar horizontalmente as imagens. À distância, medida no plano do ecrã, entre os eixos das duas imagens chamaremos *HIT*

(Horizontal Image Translation). O centro da imagem esquerda vai ser deslocado $HIT/2$ segundo o eixo x . A imagem direita vai ser deslocada $HIT/2$ segundo a mesma direcção, mas em sentido oposto. Será mantido o alinhamento vertical de ambas as imagens com os olhos do observador.

O valor HIT é importante, não só porque controla a paralaxe, mas também porque influencia o ângulo de visão estereoscópico disponível, e o volume visível.

Deve-se notar que o eixo de uma imagem não se encontra obrigatoriamente no seu centro. Isto acontece se a imagem for obtida por projecção descentrada, ou se for uma parte descentrada de uma imagem obtida por projecção centrada. Esta independência entre o eixo de uma imagem e o seu centro geométrico, indica que é possível controlar o valor de HIT sem afectar as posições dos limites das imagens no dispositivo de visualização. Para isso, o valor de HIT terá de ser definido e utilizado durante o processo de geração das imagens.

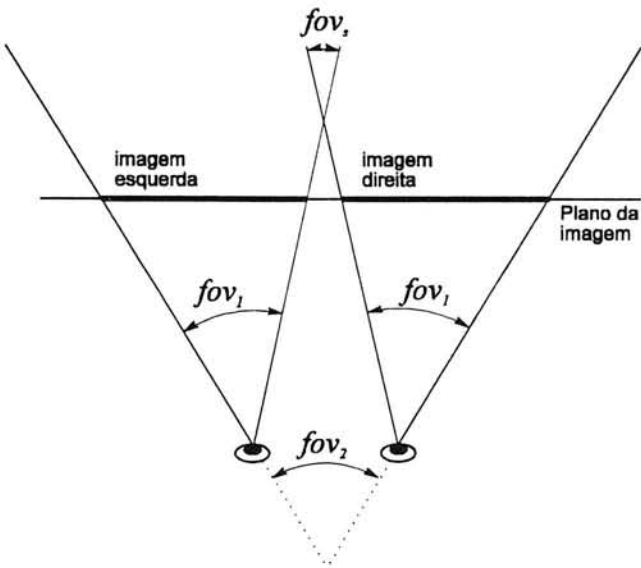


Figura 6.1. Visualização de um par estereoscópico de imagens num ecrã plano.

Como podemos constatar pelas diferenças entre os ângulos das figuras 6.1 e 6.2, as posições relativas dos limites das imagens no dispositivo de visualização afectam os limites do volume visível, tanto para um dos olhos como para ambos.

O *ângulo de visão monocular horizontal* (fov_1) é o espaço visível por cada um dos olhos. O *ângulo de visão binocular horizontal* (fov_2) é o espaço visível por qualquer um dos dois olhos. O *ângulo de visão estereoscópico horizontal* (fov_s) é o espaço visível simultaneamente pelos dois olhos.

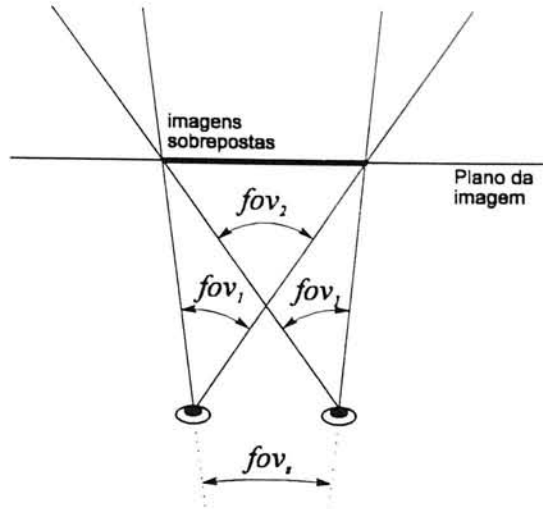


Figura 6.2. Visualização de um par estereoscópico de imagens num ecrã plano.

Cada ponto da imagem esquerda $P_{il} = (x_{il}, y_{il})$ será portanto mapeado num pixel do ecrã $P_{sl} = (x_{sl}, y_{sl})$, de tal maneira que:

$$x_{sl} = x_{il} \cdot \frac{L}{2} - \frac{HIT}{2} \quad (6.1)$$

$$y_{sl} = y_{il} \cdot \frac{H}{2} \quad (6.2)$$

Do mesmo modo, cada ponto da imagem esquerda $P_{ir} = (x_{ir}, y_{ir})$ será mapeado num pixel do ecrã $P_{sr} = (x_{sr}, y_{sr})$, de tal maneira que:

$$x_{sr} = x_{ir} \cdot \frac{L}{2} + \frac{HIT}{2} \quad (6.3)$$

$$y_{sr} = y_{ir} \cdot \frac{H}{2} \quad (6.4)$$

A paralaxe horizontal de ecrã, que vai ser utilizada para determinar a coordenada z do ponto virtual P'_0 , será:

$$R_{sx} = x_{sr} - x_{sl} = HIT + (x_{ir} - x_{il}) \cdot \frac{L}{2} = HIT + R_{ix} \cdot \frac{L}{2} \quad (6.5)$$

Da figura 6.2 podemos obter a coordenada z do ponto virtual P_0 , em função da paralaxe de ecrã:

$$z'_0 = \frac{d \cdot ioc}{ioc - R_{sx}} \quad (6.6)$$

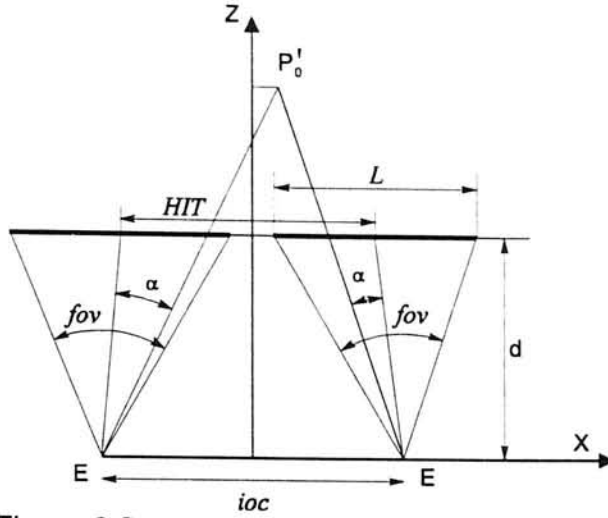


Figura 6.3. Visualização de um par estereoscópico de imagens num ecrã plano.

No mundo visual, a paralaxe dos objectos visíveis é sempre inferior à distância interocular (ioc). Isto acontece porque valores de R_{st} superiores a ioc fazem com que a profundidade z'_0 dada pela equação 6.6 se torne negativa, correspondendo a um ponto visível nas costas do observador. Na realidade, um observador confrontado com imagens com paralaxe superior à distância interocular, entende o ponto virtual como estando muito distante, como acontece para valores de paralaxe iguais a ioc .

Para manter a paralaxe inferior a um valor máximo especificado (R_{smax}) basta controlar o valor de deslocamento horizontal das imagens (HIT) de maneira que:

$$HIT \leq R_{smax} - R_{imin} \cdot \frac{L}{2} \quad (6.7)$$

O valor R_{smax} pode ser especificado como sendo igual à distância interocular ou inferior a esta. Limitar a paralaxe a um valor inferior à distância interocular justifica-se em certos sistemas de visualização estereoscópica, para os quais os valores elevados de paralaxe produzem desconforto. Nesta situação encontram-se por exemplo os sistemas baseados na multiplexagem das imagens num ecrã. Diversos autores aconselham a utilização de valores de paralaxe inferiores a 1.5° , como forma de facilitar a fusão das duas imagens.

É também justificável, por vezes, impor um limite inferior para a paralaxe negativa. Apesar de este limite não existir no mundo visual, pelo menos em termos absolutos, é geralmente aceite que observar objectos demasiado próximos (colocados perto da ponta do nariz), é desagradável. Para manter a paralaxe de ecrã superior a um determinado valor especificado, basta controlar HIT para que:

$$HIT \geq R_{smin} - R_{imin} \cdot \frac{L}{2} \quad (6.8)$$

Os valores típicos para R_{smin} situam-se normalmente entre $-R_{smax}$ e zero.

Da figura 6.3 podemos também obter x'_0 em função de z'_0 e de qualquer das coordenadas de ecrã respectivas:

$$x'_0 = x_{sr} + \frac{z'_0 - d}{d} \cdot \left(x_{sr} - \frac{ioc}{2} \right) \quad (6.9)$$

$$x'_0 = x_{sl} + \frac{z'_0 - d}{d} \cdot \left(x_{sl} + \frac{ioc}{2} \right) \quad (6.10)$$

Do mesmo modo, y'_0 pode ser obtido em função de z'_0 e da respectiva coordenada de ecrã :

$$y'_0 = y_{sr} \cdot \frac{z'_0}{d} = y_{sl} \cdot \frac{z'_0}{d} \quad (6.11)$$

Substituindo z'_0 pelo segundo membro da expressão 6.6:

$$x'_0 = x_{sr} + \frac{R_{sx}}{ioc - R_{sx}} \cdot \left(x_{sr} - \frac{ioc}{2} \right) \quad (6.12)$$

$$x'_0 = \frac{ioc}{ioc - R_{sx}} \cdot \left(x_{sr} - \frac{R_{sx}}{2} \right) \quad (6.13)$$

$$x'_0 = \frac{ioc}{ioc - R_{sx}} \cdot \frac{x_{sr} + x_{sl}}{2} \quad (6.14)$$

$$y'_0 = y_{sr} \cdot \frac{ioc}{ioc - R_{sx}} = y_{sl} \cdot \frac{ioc}{ioc - R_{sx}} \quad (6.15)$$

O que dá, por expansão de R_{sx} , para o seu valor dado pela equação 6.5:

$$x'_0 = \frac{ioc}{\frac{ioc - HIT}{L} - \frac{x_{ir} - x_{il}}{2}} \cdot \frac{x_{ir} + x_{il}}{4} \quad (6.16)$$

$$y'_0 = y_{ir} \cdot \frac{H}{2 \cdot L} \cdot \frac{ioc}{\frac{ioc - HIT}{L} - \frac{x_{ir} - x_{il}}{2}} \quad (6.17)$$

$$z'_0 = \frac{d}{L} \cdot \frac{ioc}{\frac{ioc - HIT}{L} - \frac{x_{ir} - x_{il}}{2}} \quad (6.18)$$

Temos então determinadas as expressões que nos permitem conhecer a posição do ponto virtual em função das coordenadas de imagem do par de pontos respectivo, da

distância interocular do observador e da distância a que o observador se encontra do plano da imagem. Estas expressões vão a seguir ser utilizadas para estudar a variação da posição entendida para o ponto virtual quando é alterada a distância interocular ou a distância de observação para os vários métodos de geração de imagem considerados.

6.1. Rotações com Projecção Paralela

A expressão 5.20 apresenta o valor da paralaxe para um ponto de uma imagem estereoscópica obtida por rotação e projecção paralela. Aplicando esta expressão na equação 6.5 obtemos:

$$R_{xx} = x_{xr} - x_{xl} = HIT + \frac{(z_0 - r) \cdot \sin(\frac{\theta}{2})}{\tan \frac{fov_x}{2}} \cdot L \quad (6.19)$$

Devido ao facto de a paralaxe poder variar ilimitadamente, pela variação da distância ao ponto de observação (z_0), a profundidade dos objectos na cena deve ser limitada, de maneira a que nem a paralaxe positiva nem a negativa excedam a distância interocular. Todos os objectos a visualizar deverão estar contidos entre os planos $z = z_{min}$ e $z = z_{max}$. Simultaneamente, para que a paralaxe se mantenha entre os limites especificados, é necessário que o deslocamento horizontal das imagens (HIT) seja tal que:

$$HIT \leq R_{smax} - \frac{(z_{max} - r) \cdot \sin(\frac{\theta}{2})}{\tan \frac{fov_x}{2}} \cdot L \quad (6.20)$$

$$HIT \geq R_{smin} - \frac{(z_{min} - r) \cdot \sin(\frac{\theta}{2})}{\tan \frac{fov_x}{2}} \cdot L \quad (6.21)$$

As expressões 5.16 a 5.19 apresentam os valores das coordenadas de imagem obtidas para o ponto P_0 , utilizando uma rotação em torno de y , seguida de uma projecção paralela. Aplicando estes valores às equações 6.16, 6.17 e 6.18, obtemos as coordenadas do ponto virtual P'_0 em função da posição do ponto original P_0 e das condições de visualização e de geração de imagem.

$$x'_0 = \frac{ioc \cdot \cos \frac{\theta}{2}}{2 \cdot \tan \frac{fov_x}{2} \cdot \frac{ioc - HIT}{L} - 2 \cdot (z_0 - r) \cdot \sin \frac{\theta}{2}} \cdot x_0 \quad (6.22)$$

$$y'_0 = \frac{H}{2 \cdot L} \cdot \frac{y_0}{\tan \frac{fov_y}{2}} \cdot \frac{ioc}{\frac{ioc - HIT}{L} - \frac{(z_0 - r) \cdot \sin \frac{\theta}{2}}{\tan \frac{fov_x}{2}}} \quad (6.23)$$

$$z'_0 = \frac{d}{L} \cdot \frac{ioc}{\frac{ioc - HIT}{L} - \frac{(z_0 - r) \cdot \sin \phi/2}{\tan \frac{fov_v}{2}}} \quad (6.24)$$

Destas equações é fácil observar que, para valores de ϕ diferentes de zero, x'_0 e y'_0 serão sempre dependentes de z_0 . Isto significa que, utilizando rotações com projecção paralela, as dimensões dos objectos virtuais parecem variar com a sua posição. Este modelo de geração de imagem estereoscópica não pode portanto ser aconselhado para aplicações de simulação visual.

A esta variação das dimensões com a distância chama-se efeito de *perspectiva inversa*. Este efeito é bem conhecido pois aparece sempre que se visualizam imagens (monoscópicas ou estereoscópicas) obtidas por projecção paralela. Como se pode reparar na imagem estereoscópica da figura 6.4, os objectos parecem alargar com a distância.

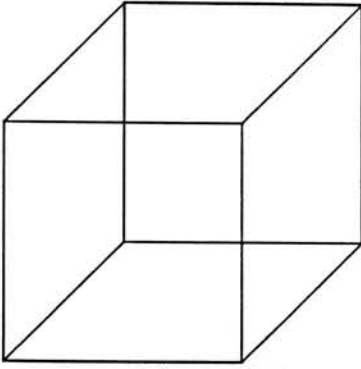
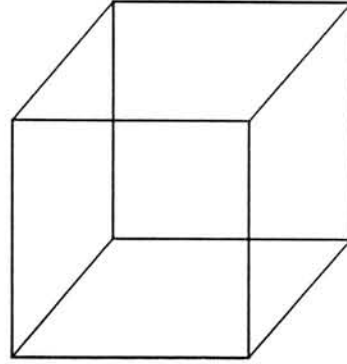


Figura 6.4.



Sensação de perspectiva inversa devido à ausência de perspectiva. Para observar esta imagem, utilizar visualização cruzada.

6.2. Rotações com Projecção Perspectiva

Como vimos anteriormente, as imagens estereoscópicas obtidas com a utilização de rotações e projecção perspectiva possuem paralaxe vertical. Esta paralaxe anormal é própria das imagens obtidas por este método, e não pode ser eliminada pelas condições de visualização. Para além deste problema, as imagens estereoscópicas, obtidas por rotação e projecção perspectiva, apresentam distorção da profundidade, como vamos demonstrar a seguir.

A expressão 5.26 apresenta o valor da paralaxe para um ponto de uma imagem estereoscópica obtida por rotação e projecção perspectiva. Aplicando esta expressão na equação 6.5 obtemos:

$$R_{sx} = HIT + \frac{L/2}{\tan \frac{fov_v}{2}} \cdot \frac{(x_0^2 + (z_0 - r)^2) \cdot \sin \phi + 2 \cdot r \cdot (z_0 - r) \cdot \sin(\phi/2)}{((z_0 - r) \cdot \cos(\phi/2))^2 - (x_0 \cdot \sin(\phi/2) + r)^2} \quad (6.25)$$

Para demonstrarmos que este método de geração de imagens estereoscópicas introduz distorção da profundidade, vamos determinar o lugar geométrico no espaço de modelação, dos pontos situados no plano virtual definido por:

$$z' = \frac{d \cdot ioc}{ioc - HIT} \quad (6.26)$$

A paralaxe de ecrã para os pontos deste plano virtual, pode ser determinada pela resolução da equação 6.6 em ordem a R_{st} .

$$R_{st} = HIT \quad (6.27)$$

Aplicando esta paralaxe à equação 6.25, obtemos a equação que define o lugar geométrico no espaço de modelação, dos pontos pertencentes ao nosso plano virtual.

$$0 = \left(x_0^2 + (z_0 - r)^2 \right) \cdot \sin \phi + 2 \cdot r \cdot (z_0 - r) \cdot \sin(\phi/2) \quad (6.28)$$

Rearranjando esta expressão, podemos obter:

$$0 = x_0^2 + (z_0 - r)^2 + 2 \cdot r \cdot (z_0 - r) \cdot \frac{\sin(\phi/2)}{\sin \phi} \quad (6.29)$$

E, completando de modo a obtermos um quadrado de uma soma:

$$x_0^2 + \left((z_0 - r) + r \cdot \frac{\sin(\phi/2)}{\sin \phi} \right)^2 = \left(r \cdot \frac{\sin(\phi/2)}{\sin \phi} \right)^2 \quad (6.30)$$

Esta expressão corresponde à equação de uma circunferência no plano x-z, com centro no ponto

$$(x_c, z_c) = \left(0, r \cdot \left(1 - \frac{\sin(\phi/2)}{\sin \phi} \right) \right) \quad (6.31)$$

e raio

$$raio = r \cdot \frac{\sin(\phi/2)}{\sin \phi} \quad (6.32)$$

Uma vez que o ângulo de rotação utilizado (ϕ) é normalmente pequeno ($\phi \leq 10^\circ$), podemos aplicar a aproximação

$$2 \cdot \sin(\phi/2) = \sin \phi \quad (6.33)$$

A circunferência pode então passar a ser definida por:

$$(x_c, z_c) = (0, r/2) \quad (6.34)$$

$$raio = r/2 \quad (6.35)$$

Em termos de coordenadas tridimensionais, estes resultados definem uma superfície semi-cilíndrica, com eixo vertical, e cuja intersecção com o plano x-y é a circunferência especificada pelas equações 6.31 e 6.32. Todos os objectos posicionados sobre esta superfície aparecem, ao observador da imagem estereoscópica, num mesmo plano paralelo ao ecrã. Isto provoca uma distorção da profundidade, uma vez que pontos modelados com profundidades diversas, aparecem com profundidades semelhantes.

6.3. Shears com Projecção Paralela

A expressão 5.46 apresenta o valor da paralaxe para um ponto de uma imagem estereoscópica obtida por aplicação de shears e projecção paralela. Aplicando esta expressão na equação 6.5 obtemos:

$$R_{xx} = x_{sr} - x_{sl} = HIT + \frac{z_0 \cdot \phi}{\tan \frac{fov_x}{2}} \cdot \frac{L}{2} \quad (6.36)$$

Para que a paralaxe se mantenha entre os limites especificados, é necessário que o deslocamento horizontal das imagens (*HIT*) seja tal que:

$$HIT \leq R_{smax} - \frac{z_{max} \cdot \phi}{\tan \frac{fov_x}{2}} \cdot \frac{L}{2} \quad (6.37)$$

$$HIT \geq R_{smin} - \frac{z_{min} \cdot \phi}{\tan \frac{fov_x}{2}} \cdot \frac{L}{2} \quad (6.38)$$

As expressões 5.40 a 5.43 apresentam os valores das coordenadas de imagem obtidas para o ponto P_0 , utilizando uma deformação em x, seguida de uma projecção paralela. Aplicando estes valores às equações 6.16, 6.17 e 6.18, obtemos as coordenadas do ponto virtual P'_0 em função da posição do ponto original P_0 e das condições de visualização e de geração de imagem.

$$x'_0 = \frac{ioc \cdot x_0}{2 \cdot \tan \frac{fov_x}{2} \cdot \frac{ioc - HIT}{L} - z_0 \cdot \phi} \quad (6.39)$$

$$y'_0 = \frac{H}{2 \cdot L} \cdot \frac{y_0}{\tan \frac{fov_v}{2}} \cdot \frac{ioc}{\frac{ioc - HIT}{L} - \frac{z_0 \cdot \phi}{2 \cdot \tan \frac{fov_v}{2}}} \quad (6.40)$$

$$z'_0 = \frac{d}{L} \cdot \frac{ioc}{\frac{ioc - HIT}{L} - \frac{z_0 \cdot \phi}{2 \cdot \tan \frac{fov_v}{2}}} \quad (6.41)$$

As considerações sobre perspectiva inversa e paralaxe não limitada feitas à geração de imagens estereoscópicas por rotações e projecção paralela, aplicam-se igualmente neste caso. No entanto, a utilização de shears oferece algumas vantagens sobre a utilização de rotações. Uma vantagem óbvia é a diminuição do número de operações necessárias, de que resulta um tempo de processamento inferior ao necessário para aplicar as rotações.

6.4. Shears com Projecção Perspectiva

Como vimos anteriormente, as imagens estereoscópicas obtidas com a utilização de shears e projecção perspectiva possuem paralaxe vertical. Esta paralaxe anormal é própria das imagens obtidas por este método, e não pode ser eliminada pelas condições de visualização. Para além deste problema, as imagens estereoscópicas, obtidas por shears e projecção perspectiva, apresentam distorção da profundidade, como vamos demonstrar a seguir.

A expressão 5.26 apresenta o valor da paralaxe para um ponto de uma imagem estereoscópica obtida por rotação e projecção perspectiva. Aplicando esta expressão na equação 6.5 obtemos:

$$R_{sx} = HIT + \frac{L/2}{\tan \frac{fov_v}{2}} \cdot \frac{(x_0^2 + (z_0 - r)^2 + r \cdot (z_0 - r)) \cdot \phi}{(z_0 - r)^2 - (x_0 \cdot \phi/2 + r)^2}$$

$$R_{sx} = x_{sr} - x_{sl} = HIT + \frac{(z_0 - r) \cdot \sin(\phi/2)}{\tan \frac{fov_v}{2}} \cdot L \quad (6.42)$$

Para que a paralaxe se mantenha entre os limites especificados, é necessário que o deslocamento horizontal das imagens (HIT) seja tal que:

$$HIT \leq R_{smax} - 2 \cdot \frac{(z_{max} - r) \cdot \sin(\phi/2)}{\tan \frac{fov_v}{2}} \quad (6.43)$$

$$HIT \geq R_{\min} - 2 \cdot \frac{(z_{\min} - r) \cdot \sin(\frac{\phi}{2})}{\tan \frac{fov_x}{2}} \quad (6.44)$$

As expressões 5.16 a 5.19 apresentam os valores das coordenadas de imagem obtidas para o ponto P_0 , utilizando uma rotação em torno de y , seguida de uma projecção paralela. Aplicando estes valores às equações 6.16, 6.17 e 6.18, obtemos as coordenadas do ponto virtual P'_0 em função da posição do ponto original P_0 e das condições de visualização e de geração de imagem.

$$x'_0 = \frac{2 \cdot ioc \cdot x_0 \cdot \cos \frac{\phi}{2}}{\tan \frac{fov_x}{2} \cdot \frac{ioc - HIT}{L} + (z_0 - r) \cdot \sin \frac{\phi}{2}} \quad (6.45)$$

$$y'_0 = \frac{H}{2 \cdot L} \cdot \frac{ioc \cdot y_0}{\tan \frac{fov_y}{2} \cdot \frac{ioc - HIT}{L} + \frac{\tan \frac{fov_y}{2}}{\tan \frac{fov_x}{2}} \cdot (z_0 - r) \cdot \sin \frac{\phi}{2}} \quad (6.46)$$

$$z'_0 = \frac{d}{L} \cdot \frac{ioc}{\frac{ioc - HIT}{L} + \frac{(z_0 - r) \cdot \sin \frac{\phi}{2}}{z_0 \cdot \tan \frac{fov_x}{2}}} \quad (6.47)$$

Para demonstrarmos que este método de geração de imagens estereoscópicas introduz distorção da profundidade, podemos aplicar o mesmo método que utilizamos para o caso de rotações com projecção perspectiva, e assim determinar o lugar geométrico no espaço de modelação, dos pontos situados no plano virtual definido por:

$$z' = \frac{d \cdot ioc}{ioc - HIT} \quad (6.48)$$

Tal como acontece com a utilização de rotações e projecção perspectiva, a geração de imagens estereoscópicas utilizando shears e projecções do mesmo tipo leva ao aparecimento de distorção da profundidade. Todos os objectos que aparecem, ao observador da imagem estereoscópica, num mesmo plano paralelo ao ecrã, pertencem a uma superfície semi-cilíndrica no espaço de modelação.

6.5. Direcções de Observação Comuns com Projecção Perspectiva

A expressão 5.65 apresenta o valor da paralaxe para um ponto de uma imagem estereoscópica obtida por translação e projecção perspectiva. Aplicando esta expressão na equação 6.5 obtemos:

$$R_{st} = x_{st} - x_{st'} = HIT - \frac{L}{\tan \frac{fov_v}{2}} \cdot \frac{x_E}{z_0} \quad (6.49)$$

Para que a paralaxe se mantenha entre os limites especificados, é necessário que o deslocamento horizontal das imagens (*HIT*) seja tal que:

$$HIT \leq R_{s,max} - \frac{L}{\tan \frac{fov_v}{2}} \cdot \frac{x_E}{z_{max}} \quad (6.50)$$

$$HIT \geq R_{s,min} - \frac{L}{\tan \frac{fov_v}{2}} \cdot \frac{x_E}{z_{min}} \quad (6.51)$$

As expressões 5.62, 5.63 e 5.64 apresentam os valores das coordenadas de imagem obtidas para o ponto P_0 , utilizando deslocamento lateral, seguido de uma projecção perspectiva única. Aplicando estes valores às equações 6.16, 6.17 e 6.18, obtemos as coordenadas do ponto virtual P'_0 em função da posição do ponto original P_0 e das condições de visualização e de geração de imagem.

$$x'_0 = \frac{ioc \cdot x_0}{2 \cdot z_0 \cdot \tan \left(\frac{fov_v}{2} \right) \cdot \frac{ioc - HIT}{L} + 2 \cdot x_E} \quad (6.52)$$

$$y'_0 = \frac{y_0}{z_0} \cdot \frac{1}{\tan \frac{fov_v}{2}} \cdot \frac{H}{2 \cdot L} \cdot \frac{ioc}{\frac{ioc - HIT}{L} + \frac{x_E}{z_0 \cdot \tan \frac{fov_v}{2}}} \quad (6.53)$$

$$z'_0 = \frac{d}{L} \cdot \frac{ioc}{\frac{ioc - HIT}{L} + \frac{x_E}{z_0 \cdot \tan \frac{fov_v}{2}}} \quad (6.54)$$

Para analisar a adaptação dos sistemas de geração de imagem e de visualização ao mundo real, vamos utilizar as derivadas parciais da posição virtual, em ordem à posição de modelação original.

$$\frac{\partial x'_0}{\partial x_0} = \frac{ioc}{2 \cdot z_0 \cdot \tan \frac{fov_v}{2} \cdot \frac{ioc - HIT}{L} + 2 \cdot x_E} \quad (6.55)$$

$$\frac{\partial y'_0}{\partial y_0} = \frac{1}{z_0 \cdot \tan \frac{fov_v}{2}} \cdot \frac{H}{2 \cdot L} \cdot \frac{ioc}{\frac{ioc - HIT}{L} + \frac{x_E}{z_0 \cdot \tan \frac{fov_v}{2}}} \quad (6.56)$$

$$\frac{\partial z'_0}{\partial z_0} = \frac{d \cdot ioc}{L} \cdot \frac{\frac{x_E}{z_0^2 \cdot \tan \frac{fov_v}{2}}}{\left(\frac{ioc - HIT}{L} + \frac{x_E}{z_0 \cdot \tan \frac{fov_v}{2}} \right)^2} \quad (6.57)$$

$$\frac{\partial z'_0}{\partial z_0} = \frac{d \cdot ioc}{L \cdot z_0^2} \cdot \frac{x_E}{\tan \frac{fov_v}{2} \cdot \left(\frac{ioc - HIT}{L} \right)^2 + \frac{x_E^2}{z_0^2 \cdot \tan \frac{fov_v}{2}} + 2 \cdot \frac{ioc - HIT}{L} \cdot \frac{x_E}{z_0}} \quad (6.58)$$

Para manter a relação de aspecto entre a altura e a largura dos objectos da cena, é necessário que:

$$\frac{\partial x'_0}{\partial x_0} = \frac{\partial y'_0}{\partial y_0} \quad (6.59)$$

Ou seja, por substituição dos valores das derivadas:

$$\left(1 - \frac{L}{H} \cdot \frac{\tan \frac{fov_v}{2}}{\tan \frac{fov_v}{2}} \right) \cdot \left(\frac{ioc - HIT}{L} + \frac{x_E}{z_0 \cdot \tan \frac{fov_v}{2}} \right) = 0 \quad (6.60)$$

Para verificar a equação 6.60, basta fazer:

$$\frac{L}{H} = \frac{\tan \frac{fov_v}{2}}{\tan \frac{fov_v}{2}} \quad (6.61)$$

ou então:

$$\frac{ioc - HIT}{L} = \frac{x_E}{z_0 \cdot \tan \frac{fov_v}{2}} \quad (6.62)$$

Como a equação 6.62 é de satisfação delicada, uma vez que depende de z_0 , podemos optar por satisfazer apenas a equação anterior. Esta imposição é natural e é habitualmente cumprida, mesmo em imagens monoscópicas. Para tal, basta apenas manter a relação entre a altura e a largura das imagens.

O ideal seria que o factor de escalamento que é aplicado identicamente à largura e à altura da cena, fosse constante para todo o espaço tridimensional, e aplicado também à sua profundidade. No entanto, a derivada da coordenada virtual x em ordem à

coordenada de modelação respectiva, não é constante para todo o espaço tridimensional de modelação. Isto é, a derivada depende do valor de z_0 , o que significa que um objecto com dimensões fixas no espaço de modelação teria dimensões diferentes no mundo virtual, dependendo da sua posição.

Para que isto não aconteça, e uma vez que fov é sempre um valor diferente de zero, teremos de fazer:

$$HIT = ioc \quad (6.63)$$

Desta maneira, as equações 6.52, 6.53 e 6.54 podem ser simplificadas, ficando agora:

$$x'_0 = \frac{ioc}{2 \cdot x_e} \cdot x_0 \quad (6.64)$$

$$y'_0 = \frac{ioc}{2 \cdot x_e} \cdot \frac{H}{L} \cdot \frac{\tan(\frac{fov}{2})}{\tan(\frac{fov}{2})} \cdot y_0 = \frac{ioc}{2 \cdot x_e} \cdot y_0 \quad (6.65)$$

$$z'_0 = \frac{d \cdot \tan \frac{fov}{2}}{L} \cdot \frac{ioc}{x_e} \cdot z_0 \quad (6.66)$$

e as respectivas derivadas:

$$\frac{\partial x'_0}{\partial x_0} = \frac{ioc}{2 \cdot x_e} \quad (6.67)$$

$$\frac{\partial y'_0}{\partial y_0} = \frac{ioc}{2 \cdot x_e} \quad (6.68)$$

$$\frac{\partial z'_0}{\partial z_0} = \frac{d \cdot \tan \frac{fov}{2}}{L} \cdot \frac{ioc}{x_e} \quad (6.69)$$

Numa primeira análise destas derivadas, podemos já constatar que a derivada de cada coordenada virtual, em ordem à coordenada de modelação respectiva, é constante para todo o espaço tridimensional de modelação. Isto é, as derivadas não dependem dos valores de x_0 , y_0 , e z_0 . Isto significa que qualquer objecto vai ser visualizado com todas as suas dimensões constantes, independentemente da sua posição relativa ao observador. No entanto, como a derivada da coordenada z é diferente das restantes, podemos obter um ligeiro escalamento da profundidade do mundo virtual. Desde que assumamos dimensões incompressíveis, este escalamento da profundidade não é muito perturbador. Para eliminar este escalamento, e tornar o sistema perfeitamente homogéneo, é necessário impor condições de visualização muito rígidas para garantir que:

$$\frac{\partial x'_0}{\partial x_0} = \frac{\partial y'_0}{\partial y_0} = \frac{\partial z'_0}{\partial z_0} \quad (6.70)$$

ou seja:

$$\tan \frac{fov_v}{2} = \frac{L}{d} \quad (6.71)$$

Isto significa que para obter imagens estereoscópicas geometricamente correctas, é necessário que o ângulo de abertura da lente da câmara utilizada na geração da imagem seja igual ao ângulo de visualização. Esta imposição não aparece na visualização de imagens monoscópicas e não é normalmente cumprida na fotografia tradicional. O tamanho do papel de uma impressão fotográfica e a distância a que é visualizada são completamente independentes da distância focal da lente utilizada para obter a fotografia. A maioria dos sistemas de visualização estereoscópica, implementados com base num ecrã tradicional, também não cumprem a igualdade de ângulos de abertura imposta pela equação 6.71. Por muito que o programador se esforce, só é possível garantir a igualdade dos ângulos de abertura para uma determinada distância de observação. No entanto o não cumprimento desta igualdade não é grave, pois resulta unicamente no aparecimento de um escalamento segundo o eixo z. Os objectos e a cena vão aparecer com mais ou menos profundidade, mas esta mantém-se independente das outras coordenadas e da posição de observação.

6.6. Conclusão

Devido à existência de paralaxe vertical, a visualização de imagens estereoscópicas obtidas utilizando rotações ou shears, com projecção perspectiva, obriga os olhos a um esforço anormal para conseguir a fusão. Experiências realizadas com estereogramas de pontos aleatórios (RDS) [Julesz77] mostram que, mesmo com ângulos de paralaxe vertical não muito elevados (cerca de $0^\circ 20''$), a fusão das duas imagens é perdida. Além do problema da existência de paralaxe vertical, foi demonstrado que os objectos visualizados com a mesma paralaxe, pertencem a uma superfície semi-cilíndrica no espaço de modelação. As profundidades relativas dos objectos podem portanto aparecer alteradas. A utilização destes métodos para gerar imagens estereoscópicas para ecrãs planos deve então ser evitada.

Quando utilizadas com projecção paralela, tanto as rotações como as deformações (shears) apresentam valores nulos para a paralaxe vertical. Ambos os métodos apresentam paralaxe horizontal ilimitada e podem provocar o efeito de perspectiva inversa. As deformações apresentam sobre as rotações a vantagem da maior facilidade de cálculo.

A utilização de direcções de observação comuns ou translações, com projecção perspectiva, para obter as duas vistas de uma imagem estereoscópica, não sofre de nenhuma das perturbações que ocorrem com as rotações e as deformações. Não existe paralaxe vertical, não existe distorção de profundidade, e o valor máximo possível para

a paralaxe horizontal é limitado. Este é portanto o método correcto para gerar imagens estereoscópicas para ecrãs planos.

Pelo estudo analítico apresentado neste capítulo podemos também concluir que são de fundamental importância as condições de visualização das imagens estereoscópicas, incluindo a posição de observação. O modelo de geração de imagens deve ser afinado e calibrado para o dispositivo e para as condições de visualização utilizados. Foi demonstrado que os eixos ópticos das duas vistas de um par estereoscópico, apresentado num ecrã plano, devem estar colocados a uma distância igual à distância inter-ocular. Foi também demonstrado que a variação da distância de observação implica um escalamento da profundidade da cena virtual.

Capítulo 7.

Qualidade das Imagens Estereoscópicas

Neste capítulo vamos tratar da qualidade de uma imagem estereoscópica, no que diz respeito à sua correcção geométrica, ao nível de realismo e outros aspectos mais ou menos subjectivos. Vamos começar por estudar as condições necessárias para que uma imagem estereoscópica seja considerada geometricamente correcta, e coerente com o mundo real.

No sub-capítulo 7.2, trataremos do nível de realismo das imagens estereoscópicas. Aqui veremos como a introdução do efeito estereoscópico influencia o realismo de uma imagem, e para que níveis de realismo é mais aconselhada a utilização da estereoscopia.

Em 7.3, vamos referir os problemas que ocorrem mais frequentemente na síntese de imagens estereoscópicas por computador. Para cada fonte de imperfeição, será proposto um meio de correcção.

Finalmente, em 7.4, vamos referir os cuidados específicos a ter quando se produzem seqüências animadas de imagens estereoscópicas.

7.1. Ortoestereoscopia

Existem três condições que devem ser satisfeitas para uma imagem poder ser considerada ortoestereoscópica. Uma imagem é considerada ortoestereoscópica se for isomórfica com o mundo real em termos de profundidade estereoscópica e perspectiva.

Em primeiro lugar, a paralaxe deve ser tal que, para ver imagens de objectos muito distantes, os eixos dos olhos do observador têm de estar paralelos. Em segundo lugar, a distância entre os centros de projecção, utilizados para obter as duas vistas, deve ser igual à distância interocular do observador. Por último, a perspectiva deve ser coerente com o mundo real. Isto significa que a distância entre o observador e a imagem deve ser igual ao produto da ampliação da imagem com a distância focal da lente da câmara (ou modelo de câmara) utilizada para obter a imagem estereoscópica. Esta última condição é bastante bem conhecida pelos fabricantes de lentes e é válida para todas as

imagens obtidas fotograficamente ou por processos computacionais que utilizem as regras da perspectiva linear.

Para cumprir a primeira condição, os pontos muito distantes terão valores de paralaxe iguais à distância interocular. Este valor da paralaxe é demasiado grande para poder ser visualizado confortavelmente. É verdade que, no mundo visual, os objectos muito distantes necessitam dos eixos dos olhos paralelos. No entanto, a disparidade entre a acomodação e a convergência, que resulta neste caso num ecrã de pequenas dimensões, introduz desconforto na visualização.

Especificando uma distância entre os dois centros de projecção igual à distância interocular do observador a que se destina, podemos satisfazer a segunda condição. No entanto, para muitas aplicações como modelação molecular ou simulação de terrenos, isto dá origem a imagens sem interesse. As moléculas são demasiado pequenas e por isso requerem que a separação interaxial das câmaras seja da ordem de grandeza dos átomos. As vistas aéreas de terrenos, por outro lado, requerem uma distância de alguns metros entre os centros de projecção, para que o efeito estereoscópico seja perceptível.

Para além da variedade das dimensões da cena que se pretende visualizar, é preciso ter em conta a variação da distância interocular entre os seres humanos. Conforme é explicado em 7.3.6, quando um par de imagens é obtido para um conjunto de olhos, isso não significa que produza um resultado ortoestereoscópico com a maioria dos outros olhos.

7.2. Realismo de Imagens Estereoscópicas

Como é sabido, a geração de imagens estereoscópicas apresenta desvantagens em termos de tempo de cálculo sobre as congéneres monoscópicas. Os benefícios que se espera obter deste maior investimento em cálculo, são um maior nível de realismo e informação geométrica mais rica. Com o objectivo de classificar o melhoramento introduzido com a estereoscopia, e de estudar a adequação de diversos tipos de imagens à visualização estereoscópica, foram realizadas, nos últimos anos, várias experiências. A metodologia seguida e os resultados obtidos em algumas destas experiências encontra-se publicado em revistas da especialidade [Hodges91].

Nestas experiências, diversas pessoas com alguma experiência de desenho foram utilizadas para classificar relativamente, imagens de diversos tipos. As comparações foram efectuadas interactivamente com a apresentação simultânea de duas imagens de tipos diferentes. Numa primeira experiência, foram apresentados, a um grupo de 10 indivíduos, os pares de imagens por ordem aleatória, e a cada indivíduo foi pedido que seleccionasse, de entre as duas, a imagem que mais facilitava a compreensão geométrica do objecto. Na segunda experiência, pediu-se a cada indivíduo que seleccionasse, de entre as duas imagens de cada par, aquela que considerava mais realista.

Os tipos de imagens apresentadas a cada indivíduo foram wire-frame, hidden line removed (HLR), flat-shaded, vistas projectadas ortogonalmente, wire-frame estereoscópio, hidden line removed estereoscópico e flat shaded estereoscópico. Todas

as imagens estereoscópicas foram geradas utilizando técnicas de projecção com eixos paralelos. A paralaxe horizontal máxima foi mantida inferior a 1.6° para toda a cena.

Os resultados obtidos para esta segunda experiência encontram-se na tabela 7.1. O valor médio corresponde à percentagem de comparações em que a imagem foi classificada como possuindo realismo superior à outra imagem em comparação.

Tipo	Média	sdev
Flat Shaded Stereo	78.13	28.75
Hidden Line Removed Stereo	72.00	18.13
Flat Shaded	62.50	21.88
Wireframe Stereo	45.00	26.25
Hidden Line Removed	42.50	28.75
Wireframe	41.88	30.00
Orthogonal Views	10.00	17.50

Tabela 7.1. Classificação das imagens em termos de realismo.

Estes resultados estão apresentados graficamente no gráfico da figura 7.1. Neste gráfico foram marcados os valores médios obtidos para cada tipo de imagem. A barra, centrada no valor médio, estende-se um desvio padrão para cada um dos lados. A barra corresponde portanto à zona de localização provável de 68% das classificações.

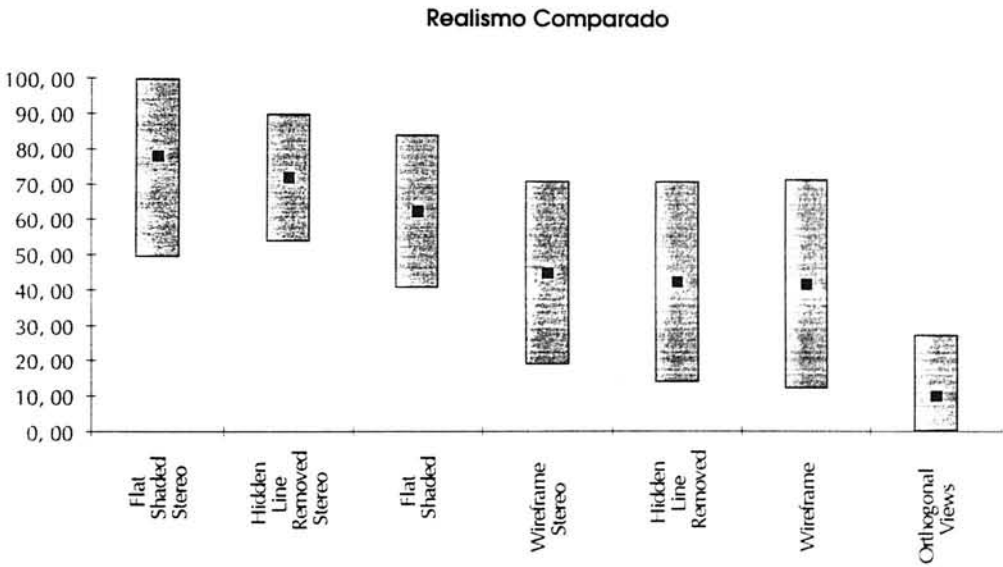


Figura 7.1. Comparação do realismo das diversas imagens.

As imagens do tipo flat shaded estereoscópicas, as HLR estereoscópicas, e as flat shaded, foram consideradas aproximadamente equivalentes e simultaneamente as mais realistas do conjunto estudado. Os tipos wire-frame estereoscópico, HLR e wire-frame monoscópico, foram também considerados aproximadamente equivalentes mas menos realistas que os tipos anteriores. Finalmente, as vistas ortogonais foram unanimemente consideradas as imagens menos realistas.

A classificação das imagens flat shaded como superiores às imagens obtidas por hidden line removal era esperada. As técnicas de hidden-line removal possuem, no entanto, o seu interesse devido à maior velocidade de geração de imagens, quando comparadas com imagens shaded. Pela análise do gráfico da figura 7.1., podemos, no entanto, constatar que as imagens hidden line removed estéreo foram consideradas como sendo pelo menos tão realistas como as imagens flat shaded monoscópicas. A adição de pistas estereoscópicas a uma imagem HLR é computacionalmente menos exigente do que a obtenção de uma imagem flat shaded. Por isso pode ser sugerido, com base nos resultados obtidos, que a informação oferecida pelo shading pode ser substituída pela informação estereoscópica, sem perda perceptível no realismo da imagem. O tempo poupado durante a geração de imagem é significativo. Uma imagem HLR estereoscópica pode ser gerada em tempo real na maioria das máquinas que não são capazes de gerar imagens shaded em tempo real.

Esta conclusão não se aplica, no entanto, às imagens wire-frame sem HLR. Na tabela 7.2. apresentamos para cada um dos tipos flat shaded, HLR e wire-frame, o incremento de realismo esperado quando é adicionado o efeito estereoscópico.

Tipo	Diferença
Flat Shaded	15.63
Hidden Line Removed	29.50
Wireframe	3.13

Tabela 7.2. Aumento do realismo provocado pela adição do efeito estereoscópico para cada tipo de imagem.

Cada um dos valores foi obtido pela diferença entre a classificação do realismo de um tipo de imagem e a classificação do tipo estereoscópico correspondente. Os resultados mostram que a adição do efeito estereoscópico a uma imagem do tipo wire-frame não incrementa significativamente o seu realismo.

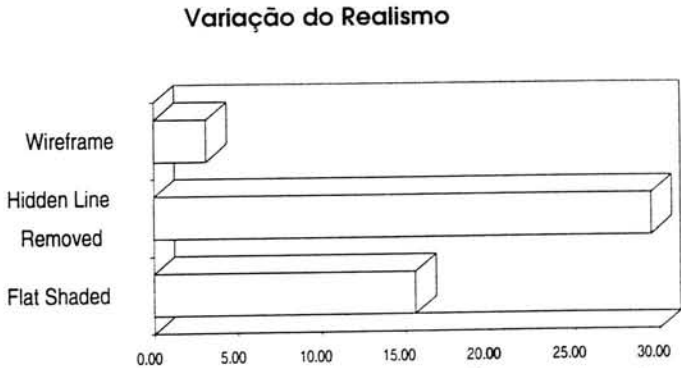


Figura 7.2. Aumento do realismo provocado pela adição do efeito estereoscópico para cada tipo de imagem.

7.3. Fontes de Erro

Existe uma forte correlação entre o movimento da cabeça de uma pessoa e as imagens de um objecto que vão sendo formadas na retina. Esta correlação pode ser explicada com uma simplificação geométrica: as imagens do objecto são projectadas nas retinas. Portanto, as imagens que aparecem nas retinas dependem apenas da forma do objecto e da posição dos dois olhos relativamente ao objecto. Um sistema de visualização estereoscópica tenta reproduzir esta geometria, através da apresentação de imagens distintas a cada um dos olhos, para que os olhos e o cérebro tenham a percepção de visualizar objectos tridimensionais. Se as imagens apresentadas não forem correctas, o observador/participante não entenderá correctamente o objecto simulado. O objecto poderá aparecer distorcido, ou poderá não existir sensação de tridimensionalidade.

Para apresentar ao observador imagens que possibilitem a percepção de objectos não distorcidos, é necessário ter em conta todos os componentes que podem afectar a geometria da imagem final apresentada aos olhos. Estes componentes incluem o software gerador de imagem, os ecrãs, os possíveis sistemas ópticos entre o observador e as imagens, e a posição das imagens relativamente aos olhos do observador. Vamos discutir alguns dos erros comuns em sistemas computadorizados de visualização estereoscópica.

7.3.1. Crosstalk

Chama-se *crosstalk* ao efeito que surge quando, na visualização de uma imagem estereoscópica, um olho vê a imagem que é destinada ao outro olho. Numa imagem estereoscópica perfeita, cada um dos olhos vê apenas a imagem que lhe é destinada. Um caso de estereoscopia sem *crosstalk* é por exemplo a utilização do estereoscópio descrito no capítulo 2. Como o estereoscópio possui as duas imagens fisicamente separadas, não é possível a um olho ver a imagem destinada ao outro olho.

A existência ou não de *crosstalk* depende essencialmente do desempenho do dispositivo de visualização. O efeito de *crosstalk* aparece tipicamente nos sistemas de visualização estereoscópica que utilizam multiplexagem temporal, cromática ou por polarização.

No entanto, é possível minimizar o *crosstalk* sentido pelos utilizadores de um sistema de visualização estereoscópico, manipulando as imagens do par ou alterando alguns parâmetros do algoritmo que as gera. O efeito de *crosstalk* pode ser mais ou menos perceptível, dependendo do brilho da imagem, da cor e, fundamentalmente, da paralaxe e do contraste da imagem. Imagens com elevados valores de paralaxe possuem maior probabilidade de apresentar *crosstalk*. Também imagens com contrastes profundos, como linhas brancas num fundo preto, apresentam normalmente mais *crosstalk*.

Em sistemas de visualização estereoscópica que apresentam as imagens sequencialmente no ecrã do tipo CRT, verifica-se mais *crosstalk* em imagens com elevados valores da componente de cor verde. Isto deve-se ao facto de, nos monitores tradicionais, o fósforo verde possuir maior tempo de remanência. Um processo de tentar reduzir o efeito de *crosstalk* nestes sistemas é a redução da componente de cor verde da imagem.

7.3.2. Convergência Incorrecta

Ambos os olhos são necessários para a visão estereoscópica. Quando os olhos estão focados num objecto distante, as linhas de visão são quase paralelas. Quando os olhos estão focados num objecto próximo, as linhas de visão convergem. Quanto mais próximo se encontrar o objecto, maior a convergência.

De uma maneira geral, um display estereoscópico possui um ecrã para cada olho. Se para gerar a imagem utilizássemos duas câmaras paralelas, as projecções de um ponto situado no infinito ficariam em posições equivalentes em ambas as imagens. Este seria mais ou menos o caso de fotografarmos uma estrela distante de duas posições diferentes mas mantendo a direcção da câmara. Quando estas imagens fossem apresentadas ao observador, este deveria perceber um pequeno ponto luminoso também a uma grande distância. No entanto, em muitos dispositivos de visualização estereoscópica, a utilização directa destas imagens não produziria o efeito desejado. Dependendo do dispositivo, o ponto brilhante poderia aparecer não tão distante como o previsto, ou poderia obrigar à divergência, dificultando assim a fusão das duas imagens.

Para além do problema do alinhamento horizontal relacionado com a convergência e com a divergência, pode ainda existir um desalinhamento vertical entre as imagens dos dois olhos. Esta divergência vertical é fortemente desagradável.

Para criar imagens estereoscópicas com a correcta convergência, quando os eixos ópticos não são paralelos ou quando os centros dos ecrãs estão deslocados dos eixos ópticos, é necessário aplicar transformações correctivas no modelo computacional de geração de imagem.

Nenhuma destas características pode, à partida, ser considerada um erro no projecto ou na construção do dispositivo de visualização, uma vez que é possível obter os efeitos correctos através das transformações computacionais adequadas. Estas características apenas tornam o modelo computacional de geração e/ou transformação de imagem um pouco mais complicado. De facto, a maioria dos dispositivos de visualização possui ecrãs descentrados e muitos HMD's têm eixos ópticos não paralelos.

7.3.3. Acomodação não relacionada com a convergência

Os olhos convergem num objecto próximo, rodando para dentro. Simultaneamente, as lentes dos olhos focam o objecto, criando acomodação à distância a que o objecto se encontra. Portanto, a acomodação está normalmente ligada à convergência. No entanto, nos dispositivos de visualização estereoscópica, cada olho vê a imagem virtual num ecrã. No que diz respeito à focagem, toda a imagem virtual aparece a uma distância constante do olho. Portanto, o utilizador de um dispositivo estereoscópico de visualização deverá aprender a não relacionar a acomodação e a convergência.

Este problema não pode ser ultrapassado com nenhum dos actuais sistemas de visualização estereoscópicos. Até que seja possível construir um ecrã com focagem variável, isto deve ser entendido como limitação da visualização estereoscópica.

Devemos, no entanto notar que, independentemente de serem considerados monoscópicos ou estereoscópicos, todos os dispositivos de visualização existentes (incluindo cinema, televisão e fotografia), fornecem acomodação fixa e portanto não relacionada de todas as outras pistas de informação de distância presentes. O que acontece agora, nos sistemas estereoscópicos, é que, como foi introduzida mais uma pista de informação de distância (a estereoscopia), existe mais uma fonte de informação não relacionada com a acomodação.

7.3.4. Ângulo de visão incorrecto

O ângulo de visão, para um olho, é o ângulo preenchido pela imagem virtual do ecrã. Existe um ângulo de visão horizontal e um vertical, que correspondem ao espaço limitado pelos extremos respectivos da imagem virtual.

7.3.5. Não utilização de projecção descentrada

Se o ecrã é perpendicular ao eixo óptico, mas descentrado relativamente a este, então o olho do observador está descentrado relativamente ao ecrã, e o centro do modelo computacional de projecção deve também ser descentrado. Esta situação requer uma projecção perspectiva descentrada, na qual os limites inferior, superior, direito e esquerdo são especificados independentemente, ao contrário do caso habitual em que cada um destes limites é simétrico do oposto.

Uma projecção perspectiva descentrada pode ser conseguida através da habitual matriz homogénea de transformação 4×4 . Este tipo de projecção perspectiva é pouco comum, mesmo para profissionais de computação gráfica, pelo facto de raramente ser necessária.

Outro erro comum é a utilização de uma simples projecção da cena tridimensional a visualizar, para obter as imagens dos olhos direito e esquerdo. O modelo computacional de geração de imagem deve incluir as transformações que tenham em consideração a posição e a orientação dos ecrãs relativamente aos olhos do observador. Estas transformações incluem sempre uma translação e podem incluir também uma rotação e uma projecção perspectiva descentrada. Em geral a simples utilização da rotação não conduz por si só à obtenção das imagens correctas.

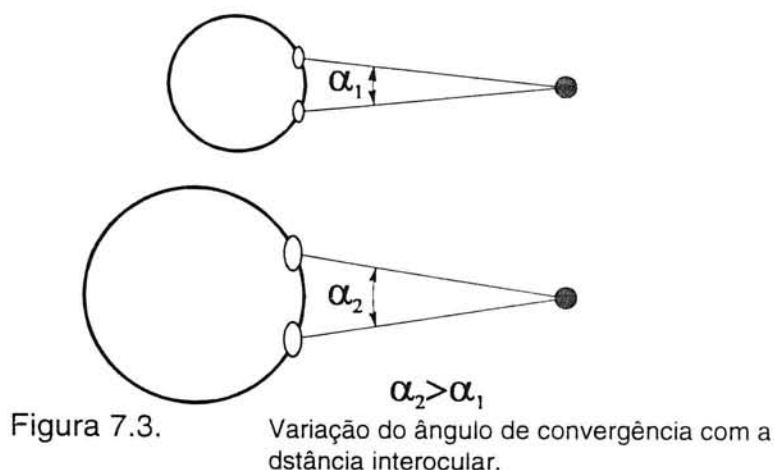
7.3.6. Distância interocular ignorada.

Nos humanos adultos, a distância entre as pupilas (IPD), também chamada distância interocular (IOC), varia normalmente entre 53 e 73 mm, e a IPD média é de cerca de 63 mm. As crianças possuem distâncias interoculares inferiores.

A grande variedade de IPD impõe determinados cuidados na construção de HMD, para que o dispositivo possa ser utilizado por todas as pessoas. A secção de saída (do lado dos olhos) dos sistemas ópticos deve ser suficientemente larga para poder ser usada tanto pelas pessoas de maior IPD, como pelas de menor IPD. Alternativamente, o sistema deve permitir ajuste mecânico da IPD, como acontece nos binóculos. Ambas as soluções têm sido aplicadas nos HMD's. Por exemplo, as lentes LEEP, utilizadas na maioria dos HMD comercialmente disponíveis, possuem uma secção de saída bastante

larga. Por outro lado, os HMD's por fibra óptica, como o CAE, requerem um ajuste mecânico para cada utilizador.

A distância entre os olhos é a linha base, a partir da qual as pessoas medem a distância a que se encontram os objectos no mundo visual. A convergência das linhas de visão desde os olhos até a um ponto fixo, pode ser medida como um ângulo. Como se pode ver na figura 7.3, uma pessoa com os olhos muito distanciados obtém ângulos de convergência diferentes dos obtidos por uma pessoa com distância entre os olhos curta, quando ambas observam um objecto colocado à mesma distância, por exemplo a meio metro.



Apesar de estas duas pessoas terem ângulos de convergência diferentes, ambas entendem o objecto como estando colocado a meio metro de distância. Cada pessoa realiza internamente um processo de calibração da sua IPD.

Quando se utiliza ajuste mecânico da distância entre pupilas, as imagens geradas computacionalmente são enviadas directamente aos olhos do observador. A sua afinação mecânica da IPD não afecta as imagens que são entregues ao observador. Quando não se utiliza ajuste mecânico, mas a secção de saída do sistema óptico é larga, duas pessoas com IPDs diferentes receberão exactamente as mesmas imagens, desde que as imagens virtuais dos ecrãs estejam colocadas no infinito óptico. Neste caso, um deslocamento lateral dos olhos não afecta os ângulos de convergência a que os objectos aparecem. No caso de as imagens virtuais dos ecrãs não estarem colocadas no infinito óptico, a situação é mais complicada, mas duas pessoas com IPDs diferentes continuariam a não obter as imagens correspondentes à sua própria distância interocular.

A solução para este problema consiste em medir a distância entre os olhos para cada utilizador, e utilizá-la no modelo computacional de geração de imagem. Se este processo for efectuado correctamente, todos os utilizadores poderão ver um qualquer objecto simulado à distância correcta.

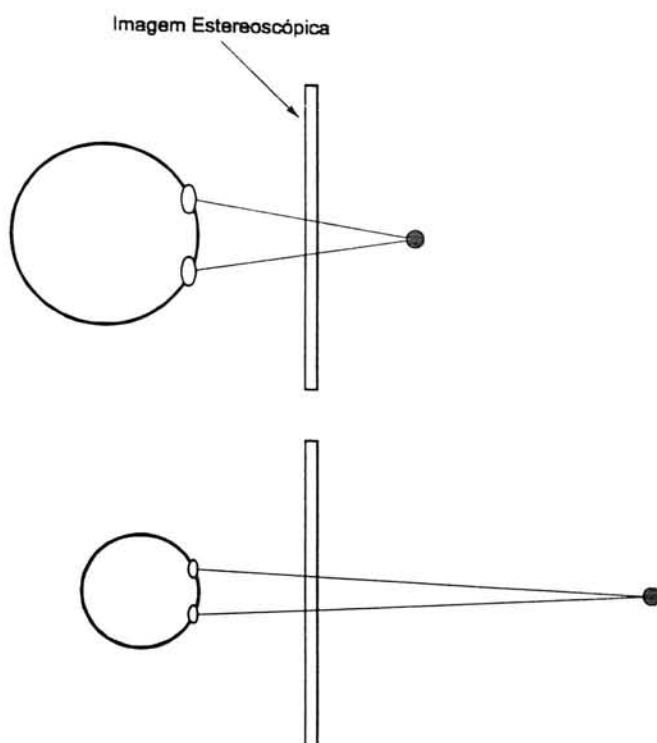


Figura 7.4. Distorção por variação da IPD.

7.3.7. Posição de observação ignorada

A posição em que o observador se coloca para observar uma imagem estereoscópica, ou a posição em que a imagem é apresentada, relativamente à cabeça do observador, influencia a transformação dos pontos da imagem estereoscópica em pontos do mundo virtual (figuras 7.5 e 7.6). Quando se utilizam dispositivos acoplados à cabeça do utilizador, a posição das imagens relativamente aos olhos do observador permanece praticamente constante. Apenas são normais pequenas variações entre utilizadores diferentes. O mesmo não se passa quando se utiliza um écran de um monitor ou uma projecção para uma audiência.

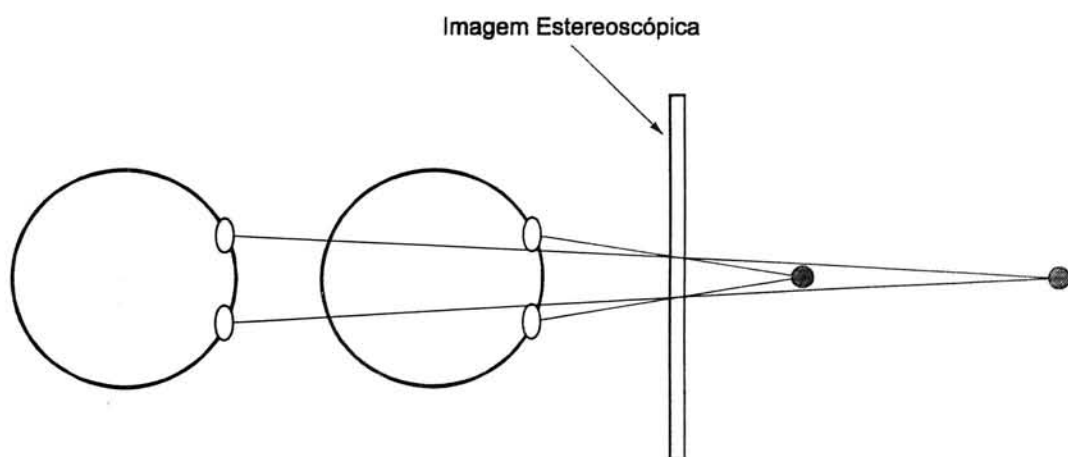


Figura 7.5. Distorção causada pela variação da distância de observação.

O software gerador de imagem deve ter em conta a posição normal de observação da imagem estereoscópica. Quando esta se destina a ser apresentada a uma assistência ou

a um observador móvel, deve ser utilizada a posição média de observação. Apesar de não ser possível obter a imagem tridimensional correcta para toda a audiência, deve ser possível apresentar globalmente imagens agradáveis.

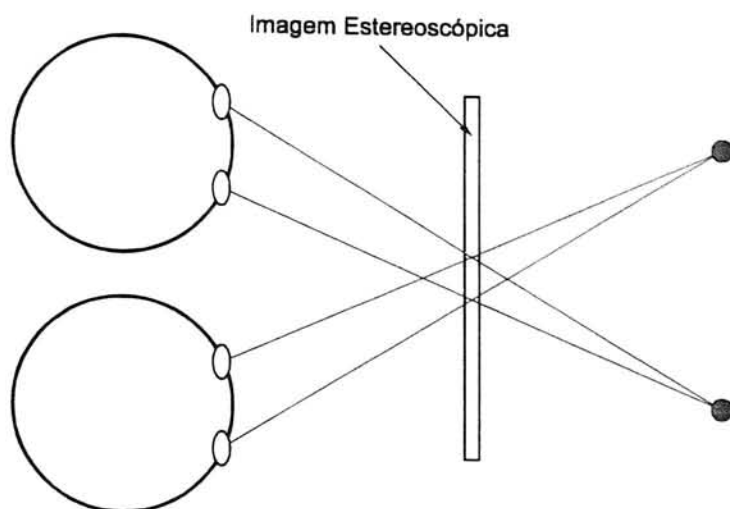


Figura 7.6.

Distorção causada pelo movimento lateral da cabeça.

7.3.8. Distorção óptica ignorada

A distorção óptica aparece normalmente apenas nos dispositivos do tipo HMD. Como, nestes dispositivos, os ecrãs estão demasiado perto dos olhos do utilizador, dificultando a focagem, os fabricantes incluem um sistema óptico entre o olho e o ecrã. O principal objectivo do sistema óptico é o de apresentar a imagem ao utilizador suficientemente distante para que a acomodação seja fácil e, ao mesmo tempo, permitir uma grande ampliação sem alterar muito a imagem. O olho, olhando através do sistema óptico, vê uma imagem virtual do que aparece no ecrã. Na maioria dos HMD disponíveis, a imagem virtual está suficientemente distante para que a focagem seja fácil, e é suficientemente grande para cobrir a parte fundamental do ângulo de visão do utilizador. No entanto, estes sistemas ópticos distorcem não linearmente a imagem, provocando que linhas que eram rectas apareçam curvas na imagem virtual.

Conforme foi dito, a distorção não linear provoca que linhas rectas apareçam curvas. Isto pode ser corrigido pelo sistema gráfico, pela simples pré-distorção das imagens que vão ser enviadas para o dispositivo de visualização. Uma linha recta na imagem virtual pode ser criada traçando no ecrã uma linha curva, de tal maneira que a sua curvatura compense exactamente a distorção introduzida pelo sistema óptico. Para tal, é necessário conhecer a inversa da função de distorção ecrã-para-imagem-virtual, e aplicá-la no sistema gráfico. Isto pode ser conseguido durante a geração de imagem ou após esta, pelo mapeamento de cada pixel para uma nova posição no ecrã, o que é computacionalmente muito dispendioso.

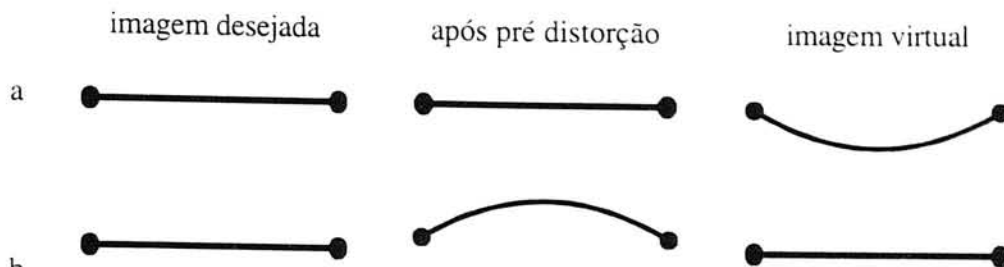


Figura 7.7. Utilização de pré-distorção.

A maioria dos sistemas existentes, que utilizam tipo HMD, limitam-se a ignorar a distorção óptica, porque podem não ter acesso à função de distorção dos sistemas ópticos, e porque a possível correcção do problema traria inconvenientes em termos de desempenho (velocidade/custo). Esta decisão de ignorar a distorção óptica não deixa de ser razoável no actual estado de desenvolvimento de sistemas HMD, pois estes podem ser utilizados mesmo com distorção óptica.

No entanto, um efeito secundário que aparece devido à distorção óptica é o de não ser possível definir um valor correcto para o ângulo de visão (FOV). A distorção não linear implica que a ampliação da imagem varie ao longo do ângulo de visão. Se o ângulo de visão utilizado na geração computacional da imagem for igual ao ângulo de visão físico, os objectos no centro vão aparecer com dimensões erradas. Mas se o ângulo de visão utilizado na geração computacional da imagem for definido de maneira a que pequenos objectos no centro apareçam com as dimensões correctas, os objectos na periferia aparecem mal posicionados. A única maneira de evitar esta escolha ingrata é pré-distorcer a imagem para corrigir a distorção óptica.

É possível minimizar tipos de aberração específicos, projectando sistemas ópticos especialmente para dispositivos HMD. A considerável distorção óptica nos HMD deve-se em parte à utilização de lentes que foram projectadas para outros fins. As lentes LEEP, que são utilizadas na maioria dos HMD disponíveis, foram projectadas para visualização estereoscópica de fotografia, pelo que possuem propositadamente alguma distorção de campo e aberrações cromáticas.

7.3.9. Utilização apenas dos vértices em sistemas não lineares

É vulgar em computação gráfica a representação de primitivas gráficas como linhas ou polígonos, através de um conjunto de pontos (vértices). Nos sistemas mais eficientes, são calculadas todas as transformações necessárias apenas para estes pontos, e só depois são desenhadas as linhas ou os polígonos entre os pontos transformados.

Para além de problemas ao nível do cálculo de iluminação que este método obrigatoriamente acarreta, só será geometricamente correcto se todas as transformações aplicadas aos vértices forem lineares. Parece tentador aplicar a função de pré-distorção apenas aos vértices, e de seguida deixar que o Hardware eficiente mas linear, encha os polígonos. Mas, dessa maneira, apenas os vértices estariam na posição correcta. As arestas dos polígonos continuariam curvas, devido à distorção óptica. A figura 7.8a mostra, num caso simples, o resultado de uma pré-distorção aplicada

apenas aos vértices. A falta de linearidade das arestas seria mais facilmente notada nos polígonos que atravessam uma grande parte do ecrã.

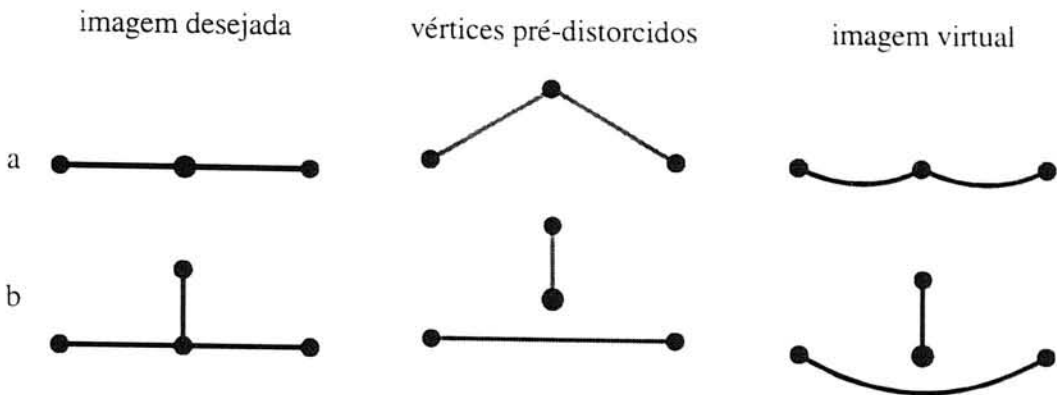


Figura 7.8.

Problemas devidos à pré-distorção apenas dos vértices.

Outro problema que surge com esta aproximação é o da perda de continuidade. Um vértice que toque uma aresta antes de efectuada a pré-distorção, não estará garantidamente colado à aresta na imagem virtual final. Falhas e buracos podem então aparecer por todo o lado. A figura 7.8b ilustra este problema.

7.3.10 Conflito com os limites do ecrã

Os dispositivos de visualização estereoscópica, mesmo os específicos para grandes ângulos de visão, fornecem ângulos de visão inferiores aos que os nossos olhos conseguem captar. Por esta razão, quando observamos uma imagem estereoscópica, há sempre mais alguma coisa exterior à imagem que pode ser vista, ou pelo menos existem direcções para onde podemos olhar que não intersectam a imagem estereoscópica.

Os limites de um ecrã estereoscópico plano são normalmente visíveis. Os limites da zona visível de uma imagem num dispositivo do tipo HMD, são normalmente zonas escuras, onde nada é visível. Quer num caso quer no outro, os limites são sempre entendidos como uma parede mais ou menos plana, no centro da qual existe uma janela, que corresponde à imagem. No caso das imagens estereoscópicas, os objectos podem ser entendidos com estando antes ou depois da janela.

O conflito entre a imagem e os seus limites acontece (figura 7.9) quando um objecto apresentado com paralaxe negativa (antes do plano da imagem) é cortado pelos limites da imagem.

Gera-se um conflito entre a pista estereoscópica, e a pista da oclusão. A primeira indica que o objecto está mais próximo do que a parede que limita a janela. Mas, se parte do objecto se encontra tapado pela parede, então a pista da oclusão indica que o objecto deve estar mais distante do que aquilo que o tapa.



Figura 7.9. Objecto cortado pelos limites da imagem.

A experiência mostra que, nestes casos, a informação estereoscópica de profundidade não é aceite, sobrepondo-se a informação mais evidente da oclusão. A sensação resultante é, no entanto, sempre desagradável, devido à falta de coerência entre as fontes de informação de profundidade.

Este problema contribui para tornar vantajosa a utilização de ecrãs com grandes ângulos de visão para imagens estereoscópicas.

7.4 Animação

7.4.1 Ecrãs do tipo Raster

O circuito electrónico responsável pelo transporte da imagem do frame buffer para o CRT (ou para outro tipo de ecrã utilizado), é chamado *controlador vídeo* (video controller), que pode ser tão simples como alguns componentes electrónicos interligados, ou tão complexo como um sofisticado microprocessador. A *taxa de refrescamento* (refresh rate) pode ser definida como o inverso do tempo requerido para o controlador de vídeo refrescar o ecrã com a imagem do frame buffer, e situa-se tipicamente no intervalo 25-120Hz, para limitar o flicker visível e simultaneamente permitir a visualização suave de movimentos.

Cada imagem completa, que é transferida, pelo controlador vídeo, do frame buffer para o display, é chamada *um quadro* (frame). O termo *quadro* pode também ser utilizado para referir o intervalo de tempo entre dois refrescamentos consecutivos, ou seja, o inverso da taxa de refrescamento.

O controlador vídeo envia os quadros do frame buffer para o display. O frame buffer, por sua vez, recebe a informação do *processador gráfico* (display processor). Nos sistemas simples, o processador gráfico não existe, e é o próprio CPU (Unidade Central de Processamento) que actualiza o frame buffer. Podemos então chamar processador gráfico ao CPU, sempre que este execute as funções próprias do processador gráfico.

Nas implementações actuais de sistemas gráficos, os dados contidos em cada pixel do frame buffer permanecem inalteráveis de um quadro para outro, a menos que

entretanto o processador gráfico recoloca novos dados. Existem diversas aplicações na área da Computação Gráfica, para as quais este processo do tipo memorização do frame buffer é muito útil: o plano de fundo de uma cena pode ser pintado apenas uma vez, e apenas os objectos que podem ser alterados, entre dois quadros consecutivos, necessitam de ser redesenhados (juntamente com os reajustes respectivos do plano do fundo).

Esta técnica adapta-se bem à maioria dos ambientes tradicionais de Computação Gráfica, em que o dispositivo de visualização está fisicamente imóvel e desempenha as funções de uma janela fixa sobre o mundo virtual. No entanto, esta técnica não é facilmente utilizável em aplicações de Realidade Virtual nem na maioria das aplicações que envolvem animação, nas quais o display está acoplado a um objecto (câmara) ou a um participante da cena. Neste caso, toda a informação, incluindo o plano de fundo, pode ser alterada de quadro para quadro, sempre que o observador ou a câmara se movem.

Há diversos problemas que surgem com esta necessidade de actualizar constantemente todo o frame buffer. Um dos mais preocupantes surge quando o processador gráfico actualiza o frame buffer ao mesmo tempo que o controlador de vídeo está a ler o mesmo frame buffer para actualizar o ecrã. Isto resulta normalmente numa excessiva quantidade de flicker visível, pois objectos parcialmente desenhados (e portanto parcialmente apagados), podem ser apanhados durante o refrescamento do display. O erro fica visível até que surja um novo refrescamento.

Para além deste problema, aparece um outro sempre que a parte da imagem que requer actualização seja uma porção significativa da imagem. Neste caso torna-se computacionalmente mais interessante apagar e redesenhar toda a imagem, do que apagar individualmente todos os objectos que necessitam de ser redesenhados. Por isso, a menos que a cena possa ser redesenhada em tempo inferior ao tempo de um quadro (o que não acontece na maioria das situações habituais), o observador veria, não uma sucessão de imagens completas, mas uma sucessão de construções de imagens. Isto é vulgarmente aceitável em aplicações como o CAD¹, onde este processo de construção pode até ser considerado mais informativo. No entanto, isto é inaceitável para aplicações de animação ou imersão convincentes, como no caso da Realidade Virtual.

7.4.1.1 Double Buffering

A solução habitual para estes problemas consiste na utilização de double buffering (armazenamento duplo), ou seja a existência de dois buffers (espaços) de memória, cada um dos quais com capacidade para armazenar uma imagem completa. O controlador vídeo envia para o display a informação que lê de um dos buffers, enquanto o processador gráfico actualiza o outro buffer. Quando o processador gráfico termina a geração da imagem completa, o controlador vídeo é instruído para trocar para o segundo frame buffer (que contém a nova imagem). Simultaneamente com esta troca, é feita também a troca do buffer acessível pelo processador gráfico, passando este a ter acesso ao primeiro frame buffer (contendo a imagem

¹ CAD - Computer Assisted Design (Desenho Assistido por Computador).

desactualizada que estava a ser visualizada). Com esta técnica, o observador vê uma imagem constante durante um certo número de frames, até que uma nova imagem esteja pronta. Nesta altura, a imagem é instantaneamente alterada para a nova imagem, que permanecerá activa até que uma outra nova imagem esteja disponível.

A taxa a que estas novas imagens são completadas pode ser chamada *update rate* (taxa de actualização). É importante notar a diferença entre a taxa de refrescamento e a taxa de actualização, e a relação entre as duas. A taxa de refrescamento é a taxa a que o controlador vídeo lê imagens do frame buffer para o dispositivo de visualização, e é normalmente constante para uma determinada configuração de Hardware de visualização. A taxa de actualização, por outro lado, é a taxa a que as novas imagens completas são produzidas. É normalmente inferior à taxa de refrescamento, e depende fortemente da complexidade da cena a visualizar.

É normalmente preferível efectuar a troca de buffers apenas entre dois refrescamentos consecutivos do dispositivo de visualização, e não durante um refrescamento. Isto é verdade principalmente se o tempo que o processador gráfico gasta a gerar uma imagem completa for comparável ao tempo de um quadro. A não observância desta regra pode levar ao aparecimento de descontinuidades na imagem. Infelizmente, esta restrição implica que o processador gráfico tenha que esperar até ao fim de um refrescamento para começar a gerar uma nova imagem (a não ser que seja utilizado um terceiro buffer), forçando a relação entre a taxa de refrescamento e a taxa de actualização a subir para o valor inteiro imediatamente acima. Isto é mais prejudicial quando a taxa de actualização é próxima da taxa de refrescamento. Se, por exemplo, uma actualização demora apenas 1.2 períodos de refrescamento a ser executada, a sincronização com o frame buffer implica que os restantes 80% do período de refrescamento não sejam utilizáveis para geração de imagem. A taxa de utilização do processador gráfico ficaria então reduzida a:

$$\rho_{pr} = \frac{1.2}{1.2 + 0.8} = 60\% \quad (7.1)$$

É sabido que num sistema de Realidade Virtual, a existência de uma latência baixa é de fundamental importância para que a experiência seja convincente e até não provocadora de enjoos. Nestes sistemas são numerosas e variadas as potenciais fontes de latência. Vamos aqui referir apenas aquelas que são introduzidas directamente pelo sistema gráfico.

Em primeiro lugar, se o processador gráfico gera uma imagem com base na informação posicional dos objectos da cena válida no início do cálculo, esta informação posicional está já desactualizada quando terminar o processo de geração da imagem. Em seguida, a imagem completa tem de aguardar até ao início do próximo refrescamento do ecrã. Quando a imagem chegar ao ecrã, as posições dos objectos que lhe deram origem estarão já desactualizadas um tempo igual ao período de actualização do frame buffer.

Independentemente da troca de buffers ser sincronizada ou não com o varrimento, se a taxa da actualização for inferior à taxa de refrescamento (o caso normal), então aparece a mesma imagem estática no ecrã durante um número inteiro de frames, até

que uma nova imagem esteja pronta. A latência introduzida por esta longa permanência no ecrã é:

$$\tau_{lag} \approx \frac{1}{\tau_{update}} \cdot \int_0^{\tau_{update}} t \cdot dt = \frac{1}{2} \cdot \tau_{update} \quad (7.2)$$

A latência introduzida pelo sistema gráfico é portanto fortemente influenciada pela taxa de actualização.

7.4.2 Animação de Imagens Estereoscópicas

Vamos agora tratar dos cuidados a ter para realizar sequências de imagens estereoscópicas, utilizando um dispositivo de visualização com multiplexagem temporal das duas imagens. Vamos também descrever um programa implementado para fins unicamente experimentais, que permite visualizar uma imagem em tempo real, com possibilidade de alteração interactiva da posição e orientação da câmara estereoscópica.

Para além dos problemas que surgem em qualquer animação, e que foram descritos anteriormente, a apresentação de sequências de imagens estereoscópicas implica mais cuidados ao programador. Tal como no caso monoscópico, é desejável a utilização de buferização dupla (double buffering). Como tratamos de duas imagens, cada um dos buffers será agora duplo, ou simplesmente dividido em duas metades.

No caso de se utilizarem dois buffers para cada uma das imagens do par estereoscópico, é fundamental ter o cuidado de só efectuar a mudança dos buffers em conjunto, e depois de completa a actualização de ambas as imagens. Caso contrário, pode acontecer que, em determinados instantes, os objectos móveis da imagem apareçam em posições diferentes a cada um dos olhos do observador. As inconsistências geométricas introduzidas por este dessincronismo das duas imagens é fortemente perturbador e desagradável.

No caso de se utilizarem buffers comuns para ambas as imagens, o problema da comutação separada não se põe, pois a comutação dos buffers, quando efectuada, afecta obrigatoriamente o conjunto das duas imagens. No entanto, continua a ser fundamental assegurar que a mudança de buffers só ocorra depois de ambas as imagens terem sido actualizadas.

Para além disso, durante a selecção dos parâmetros para a geração das imagens, deve ser tido em conta que estes parâmetros deverão permanecer fixos durante toda a animação, ou pelo menos, não poderão ser alterados em todas as frames, ao sabor dos movimentos da cena. Quando tratámos da geração de imagens monoscópicas, foi dito que se pode manipular a distância interaxial x_E , e o deslocamento horizontal das imagens (*HIT*), em função da gama de profundidades dos objectos da cena, de maneira a manter a paralaxe dentro dos valores considerados aceitáveis. Em aplicações de simulação visual interactiva como a Realidade Virtual, não é normalmente possível determinar o intervalo de profundidade onde os objectos vão aparecer. Assim, não é também possível limitar o valor da paralaxe sem introduzir distorção na profundidade virtual. Por esta razão, aconselhamos a utilização de dispositivos que facilitem a

acomodação desrelacionada da convergência. Nesta situação encontra-se a generalidade dos dispositivos HMD.

O tempo de geração de uma imagem estereoscópica, e portanto o período de actualização, é normalmente o dobro do caso monoscópico. A latência introduzida pelo sistema gráfico numa aplicação interactiva, como depende essencialmente da taxa de actualização, aumenta também no mesmo factor.

Esta constatação é por vezes utilizada para desaconselhar a utilização de imagens estereoscópicas em ambientes de tempo real como a Realidade Virtual. É que as desvantagens introduzidas pelo aumento da latência devido à duplicação do período de actualização, podem ser mais importantes do que as vantagens que resultam da utilização de estereoscopia.

Por esta razão, é de fundamental importância tentar diminuir o tempo de geração de imagens estereoscópicas. Se um par estereoscópico de imagens puder ser gerado em tempo inferior ao tradicional dobro do tempo de geração da imagem monoscópica correspondente, então estamos a diminuir o preço a pagar pela utilização da estereoscopia, e a permitir uma utilização mais vasta.

7.5 Conclusão

Apesar de se basear normalmente na simulação de fenómenos físicos, ou pelo menos geométricos, a geração de imagens estereoscópicas é, como toda a criação de imagem, uma arte. Mesmo conhecendo o modelo geometricamente correcto para obter imagens para um determinado dispositivo de visualização, existem vários cuidados que é necessário ter, para garantir uma visualização confortável. Embora muitas vezes subjectivos, alguns destes cuidados são extremamente importantes, podendo até ser considerada vantajosa a introdução de ligeiras perturbações geométricas. No entanto, graças à maior versatilidade que os meios computadorizados apresentam sobre os outros meios de geração de imagem (por exemplo: fotografia), não se devem tolerar imagens estereoscópicas geradas por computador que não sejam simultaneamente agradáveis e de visualização confortável.

Capítulo 8.

Ray-Tracing Estereoscópico

Neste capítulo vamos descrever o método desenvolvido para obtenção de pares estereoscópicos de imagens com elevado realismo, utilizando o algoritmo de Ray-Tracing. A descrição pormenorizada do algoritmo de Ray-Tracing pode ser encontrada em várias publicações [Whitted80], [Glassne89], pelo que nos limitaremos à descrição dos blocos destinados à utilização estereoscópica. Serão também tratados neste capítulo vários métodos desenvolvidos para aceleração da obtenção do par estereoscópico, aproveitando a coerência entre a imagem do olho esquerdo e a do direito.

Referiremos principalmente o caso da obtenção de imagens destinadas à visualização em ecrãs planos. O algoritmo, na forma como é apresentado ou com pequenas alterações, é também aplicável a qualquer outro dispositivo de visualização estereoscópica.

8.1. Ray-Tracing Tradicional

As imagens produzidas pelo algoritmo de Ray-Tracing são reconhecidas pela sua alta qualidade e realismo. O obstáculo que impede a utilização destes algoritmos em sistemas de baixo custo consiste na grande exigência em termos de tempo de processamento.

Têm sido publicados diversos trabalhos focando métodos de aceleração na aplicação de algoritmos de Ray-Tracing [JFoley90]. Alguns métodos, consistindo em simplificações efectuadas sobre o método tradicional de Ray-Tracing, reduzem o tempo de cálculo produzindo uma imagem muito semelhante à imagem que seria obtida pelo método tradicional [MLeitão93], [Ward91]. A economia de tempo conseguida é por vezes significativa mas, mesmo assim, não permite a obtenção de imagens em tempo útil para ambientes interactivos.

Vamos aqui referir apenas alguns métodos de optimização de Ray-Tracing, com o objectivo de, posteriormente, estudarmos a sua aplicação a um Ray-Tracer estereoscópico.

8.1.1. Utilização da Coerência de Imagem

Numa imagem sintetizada por um algoritmo de Ray-Tracing, para um mesmo modelo da cena e da iluminação, o realismo de uma imagem considera-se função do número de pontos calculados (resolução espacial) e do grau de profundidade de cálculo de iluminação (número de reflexões e de refrações). É portanto possível controlar a evolução do realismo influenciando a taxa de amostragem e o máximo nível de profundidade no cálculo da iluminação.

Num algoritmo tradicional de Ray-Tracing [Whitted80], a imagem é amostrada sequencialmente em vários pontos. Para cada ponto é lançado um raio e, recursivamente, são lançados os seus descendentes. Cria-se uma única imagem e podemos considerar que o realismo é constante e pré-estabelecido. Num algoritmo de Ray-Tracing adaptativo, são seleccionados os pontos de amostragem em função da cena. O realismo da imagem pode ser controlado, e é possível posteriormente elevar o nível de realismo da imagem.

A largura média da árvore de “shading”, em qualquer nível, depende do número de pontos amostrados e do número de descendentes de cada raio. Por sua vez, o número de níveis da árvore de “shading” depende da precisão pretendida no cálculo da iluminação, pois a importância de cada raio vai diminuindo com o aumento da profundidade.

Como o tempo de cálculo está também directamente relacionado com o número de raios calculados, é possível reduzir o tempo de processamento pela limitação das árvores de “shading”. Um algoritmo de Ray-Tracing com realismo crescente pode usar estes artifícios para criar imagens parciais, com menor realismo, que possam ser melhoradas com mais tempo de processamento.

Inserido neste contexto, é possível gerar uma sequência de imagens de realismo crescente, para o que, nas primeiras imagens, as profundidades e a largura média das árvores de “shading” são limitadas [MLeitão91]. Esta limitação afecta a qualidade das primeiras imagens, pois a redução da largura média das árvores de “shading” implica uma diminuição do número de pontos calculados (redução na resolução espacial) enquanto que a redução da profundidade implica uma diminuição no número de reflexões e de refrações calculadas. O realismo da imagem final não é afectado por este processo, dado que possibilita a passagem a um nível de realismo superior utilizando cálculos já efectuados.

O utilizador pode controlar a evolução do realismo conforme desejar, por qualquer um dos dois métodos de limitação nas árvores de “shading” ou pela combinação de ambos. Os parâmetros disponíveis para o controlo do realismo de uma imagem parcial são:

- dimensão mínima em “pixels” de uma célula de imagem;
- profundidade máxima da árvore de “shading”;
- limites de uma zona de interesse a melhorar.

Uma vez criada uma imagem com um certo conjunto de parâmetros, é possível redefinir alguns deles e obter uma nova imagem de realismo superior. A definição e a alteração dos parâmetros pode ser feita de modo interactivo pelo utilizador. Para além disso, esta estratégia adaptativa de lançamento de raios, permite obter uma imagem idêntica à imagem final em tempo bastante inferior ao necessário quando se utiliza o processo tradicional de lançamento de raios.

8.2. Ray-Tracing Estereoscópico

Como foi dito e devidamente justificado nos capítulos anteriores, para gerar pares estereoscópicos correctos, para um ecrã plano, os eixos das lentes das duas câmaras devem ser paralelos. Foi também demonstrado, no capítulo 5, que as imagens obtidas por duas câmaras paralelas, observando o mesmo mundo de pontos diferentes, são equivalentes às obtidas por duas câmaras coincidentes, que observam dois mundos distintos, criados pela translação adequada de todos os objectos do mundo de modelação original.

O método proposto baseia-se na duplicação do número de objectos da cena, mantendo assim o número de raios constante. A cada objecto modelado original, aplica-se uma translação para a esquerda para criar um novo objecto (figura 8.1). Este objecto será chamado objecto do lado esquerdo, e vai ser utilizado para gerar a imagem que será entregue ao olho direito. Uma translação de sentido contrário aplicada ao mesmo objecto original dá origem ao objecto do lado direito, que vai ser utilizado para gerar a imagem que será entregue ao olho esquerdo.

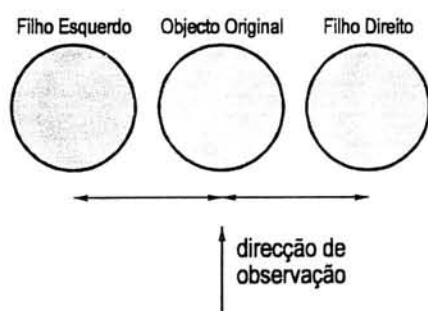


Figura 8.1. Geração de objectos adicionais por translação.

A direcção da translação deve ser simultaneamente perpendicular à direcção de observação e à direcção vertical. O seu versor pode ser obtido pelo produto vectorial normalizado da direcção de observação com o vector que define a vertical da imagem. O módulo da translação corresponde a metade da distância interaxial utilizada na geração da imagem.

Como a translação é aplicada a todos os objectos da cena, teremos um conjunto de objectos filhos com o dobro dos objectos originais. Na figura 8.2, podemos ver um exemplo de uma cena com dois objectos originais (Obj1 e Obj4).

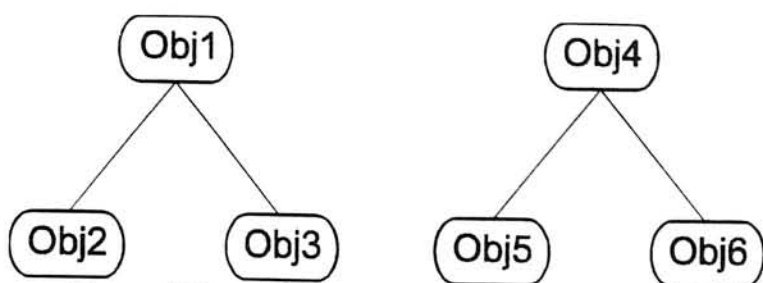


Figura 8.2. Objectos pais, e objectos filhos.

O objecto Obj1 deu origem ao objecto Obj2, através duma translação, e pela translação inversa ao objecto Obj3. O mesmo se passa com Obj4 que dá origem por translação aos objectos Obj5 e Obj6. Referiremos os objectos gerados pela translação dos originais como *objectos filhos*. Obj2 e Obj5 são os filhos do lado esquerdo e Obj3 e Obj6 são os filhos do lado direito. Obj2 é o irmão gémeo de Obj3 e Obj5 é o irmão gémeo de Obj6. Vamos utilizar o nome *primo* para designar a relação entre um objecto filho e todos os outros filhos do mesmo lado. Isto é: Obj2 é primo de Obj5 e Obj3 é primo de Obj6. No entanto Obj2 não é primo de Obj6 nem Obj3 é considerado primo de Obj5.

A geração simultânea das duas imagens necessita que os dois conjuntos de objectos residam simultaneamente em memória, para poderem ser trabalhados em conjunto. Esta duplicação da base de dados de objectos é uma desvantagem a ter em conta quando se opta por esta solução face à utilização de duas câmaras. O aumento de dimensão da base de dados de objectos pode no entanto ser minimizado, duplicando unicamente a informação geométrica de posição de cada objecto.

Na figura 8.3. podemos observar um exemplo de uma estrutura de dados para armazenamento das descrições dos objectos.

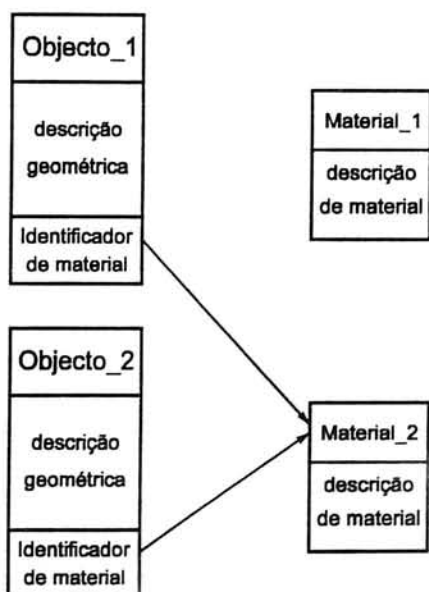


Figura 8.3. Armazenamento das descrições dos objectos, partilhando informação sobre material.

A informação correspondente a cada objecto é armazenada numa estrutura. A informação correspondente ao tipo de material, e que portanto pode ser partilhada, é

colocada numa outra estrutura que é simplesmente referida (apontada) pelas estruturas de todos os objectos que a utilizem. Desta maneira, a duplicação do número de objectos não implica a duplicação do espaço de armazenamento da descrição da cena a visualizar.

Melhores resultados podem ainda ser obtidos partilhando também toda a informação geométrica comum aos dois objectos. Nesta situação podemos considerar por exemplo: o raio de uma esfera ou da base de um cone, o vector de orientação do eixo de um cone, as dimensões de uma caixa, e outra informação geométrica que não depende da posição absoluta do objecto no espaço. Na figura 8.4. podemos ver um exemplo de uma estrutura de dados na qual os objectos gémeos partilham também informação geométrica.

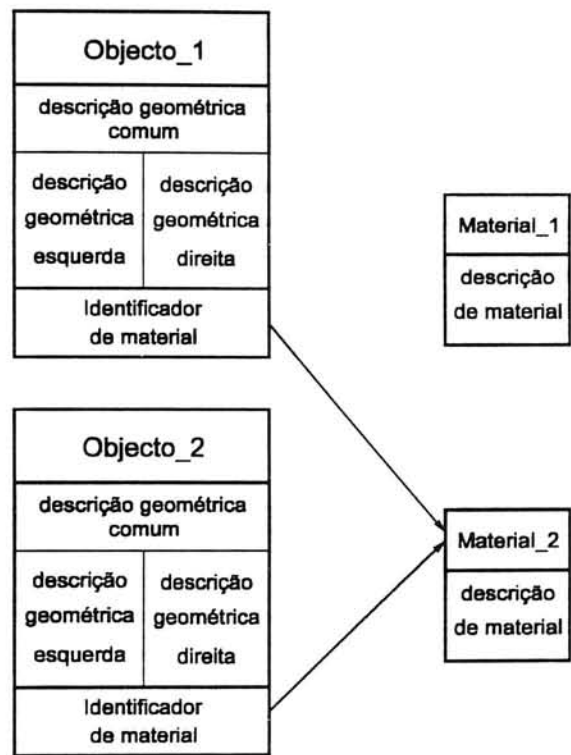


Figura 8.4. Armazenamento das descrições dos objectos, com partilha de informação geométrica e de informação sobre o material.

Toda a informação que é comum aos dois objectos gerados a partir do mesmo original, como a dimensão, a orientação, e a descrição do tipo de material, é agora partilhada. O aumento da dimensão da base de dados é suportável, principalmente quando o objectivo final é a diminuição do tempo de geração do par de imagens.

A utilização de objectos geminados, implica também uma outra desvantagem à partida: o maior tempo de pré-processamento para efectuar as translações e as alocações de memória necessária à criação dos novos objectos. No entanto, como veremos mais tarde, este tempo será normalmente recuperado pela aceleração que pode assim ser introduzida. Aliás, de uma maneira geral, o aumento do pré-processamento como método de aceleração do algoritmo principal é tanto mais interessante quanto maior for o conjunto dos dados a processar. Como no caso de pares de imagens temos, em

princípio, o dobro do número de raios a processar relativamente ao caso monoscópico, deverá ser vantajoso apostar num pré-processamento mais refinado.

Como acontece num Ray-Tracing monoscópico tradicional, os objectos são ordenados hierarquicamente, com base na informação geométrica dos volumes envolventes respectivos. São bem conhecidas as vantagens da ordenação hierárquica dos volumes dos objectos [Glassne89], principalmente na aceleração da determinação da intersecção raio-objecto mais próxima do ponto de partida. A ordenação permite baixar a complexidade do algoritmo. O tempo de cálculo deixa de ser proporcional ao número de objectos para passar a depender apenas do logaritmo deste. Desta maneira, a duplicação do número de objectos não implica uma duplicação do tempo de cálculo, desde que o número de raios lançados permaneça constante.

8.2.1. Intersecção Raio/Objecto Filho

Para cada raio lançado do ponto de observação, devem agora ser obtidos dois pontos de intersecção (I_1 e I_2 da figura 8.5). Um destes pontos deverá pertencer à superfície de um objecto filho do lado esquerdo e outro à superfície de um objecto filho do lado direito.

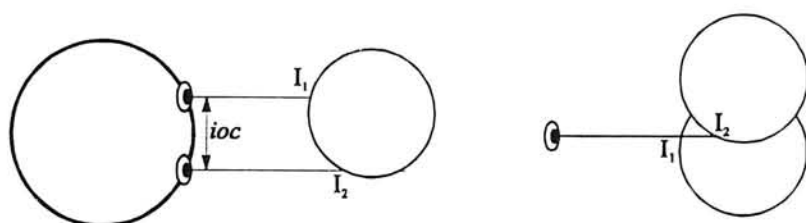


Figura 8.5.

Obtenção de dois pontos de intersecção para cada raio.

Como se observa na figura 8.6, não é obrigatório que estes dois objectos filhos sejam gémeos um do outro, isto é, podem não ter sido gerados a partir de um pai comum.

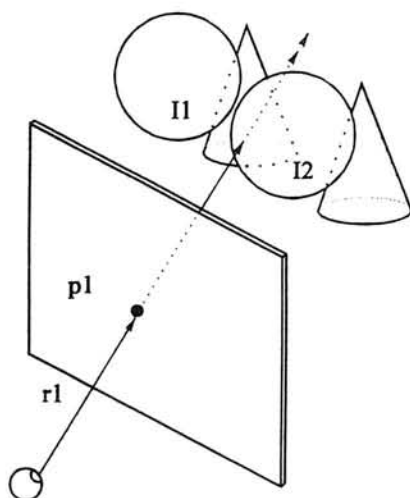


Figura 8.6.

Obtenção de dois pontos de intersecção em objectos distintos para cada raio.

O algoritmo desenvolvido e que é utilizado para determinar o par de intersecções de um raio com todos os objectos filhos dos objectos originais é o seguinte:

Determina-se a intersecção do raio com a caixa envolvente do objecto raiz. No caso da intersecção existir aplica-se o mesmo processo recursivamente a cada um dos seus descendentes. Os objectos cuja caixa envolvente é intersectada e que não possuem descendentes na ordenação hierárquica, são inseridos na lista de objectos intersectados respectiva. Existem duas listas de objectos intersectados, uma para cada lado, ambas ordenadas por ordem crescente de distância.

Quando o processo de determinação das intersecções com caixas envolventes está concluído, determina-se a distância de intersecção com os objectos das listas. Este processo é efectuado separadamente para cada um dos lados e por ordem crescente de distância. Para cada um dos lados, o processo termina quando for confirmada uma intersecção com uma distância inferior a todas as restantes na lista, ou quando toda a lista for percorrida.

```
two_objects intersect(ray)
  cluster_intersect(ray,ROOT_OBJECT)
  for eye:=LEFT to RIGHT
    distance_min[eye]:=real_max
    while queue_not_empty(eye) and bound_distance<distance_min
      queue_read(eye,object,bound_distance)
      if bound_distance<distance_min
        distance:=object_intersect(ray,object)
        if distance<distance_min
          distance_min[eye]:=object
          intersect[eye]:=object
        endif
      endif
    endwhile
  endfor
end
```

```
cluster_intersect(ray,object)
  distance:=bound_intersect(object)
  if distance>0
    if object.type==CLUSTER_TYPE
      for obj:=1 to object->descendentes
        cluster_intersect(ray,object->descendente[i])
      endfor
    else
      queue_insert(object->eye,object,distance)
    endif
  endif
end
```

8.2.2. Cálculo de Iluminação

Para obter os valores da iluminação no par de pontos de intersecção determina-se, em cada um dos pontos, o efeito da iluminação ambiente, de cada uma das fontes luminosas, da reflexão e da transmissão. Para obter os valores de iluminação recebidos no ponto de intersecção segundo as direcções de reflexão e de transmissão, lançam-se

novos raios nestas direcções. Poderá portanto haver quatro raios descendentes do primeiro raio. Cada um destes quatro raios diz apenas respeito a uma das vistas (um dos lados), logo, deverão ser tratados como no algoritmo tradicional.

Será testada a intersecção de cada um destes novos raios apenas com os objectos do lado respectivo. Após verificada a intersecção, cada um destes raios dará origem apenas a um raio reflectido e um raio transmitido, como no algoritmo tradicional de Ray-Tracing. Para não prejudicar este processo, na implementação efectuada manteve-se a estrutura de apontadores que define a ordenação hierárquica para os objectos de um só lado, em paralelo com a ordenação hierárquica global de todos os objectos filhos.

Na figura 8.7. podemos ver um exemplo de uma árvore de shading que pode ser obtida para um raio estereoscópico.

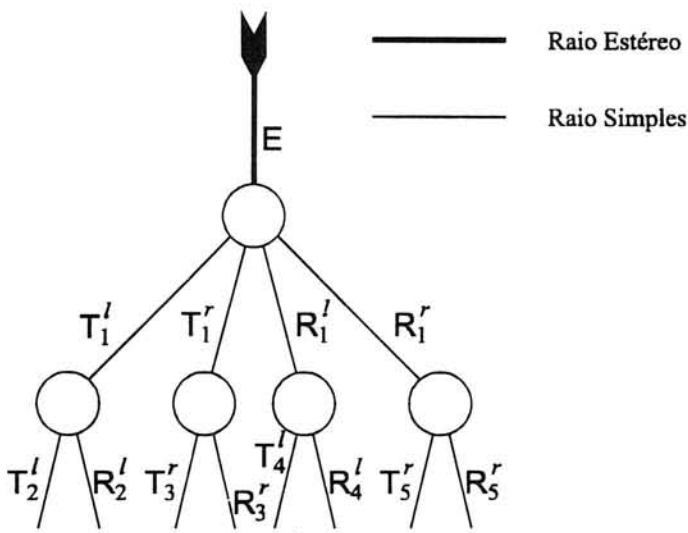


Figura 8.7.

Árvore de shading completa, obtida para um raio estereoscópico.

Para cada ponto de amostragem da imagem, é enviado um único raio (E), que parte de um ponto de observação também único, e que corresponde à posição média dos dois centros de projecção. É testada a intersecção deste raio contra todos os objectos. Em princípio, serão encontrados dois pontos de intersecção distintos; um ponto corresponde à intersecção mais próxima com o conjunto de objectos primos do lado esquerdo, e outro ponto corresponde à intersecção mais próxima com o conjunto de objectos primos do lado direito.

Estes dois pontos correspondem à translação dos pontos de intersecção dos dois raios que partem cada um de seu olho, com os objectos originais. No caso de não ser encontrada nenhuma intersecção neste ponto de amostragem para qualquer das vistas, é atribuída a cor do plano de fundo à imagem respectiva.

Depois de calculados os pontos de intersecção e a influência da luz directa e ambiente incidentes nos pontos de intersecção, é lançado um raio segundo a direcção de transmissão e outro segundo a direcção de reflexão a partir de cada um dos pontos de intersecção. Do ponto de intersecção do raio original com o objecto do lado esquerdo são enviados os raios R_1^l e T_1^l , respectivamente reflectido e transmitido. Do ponto de

intersecção do raio original com o objecto do lado direito são enviados os raios R_1^r e T_1^r , respectivamente reflectido e transmitido.

Poderá haver portanto quatro raios descendentes do primeiro raio. Cada um destes quatro raios diz apenas respeito a uma das vistas (um dos lados), logo, deverão ser tratados como no algoritmo tradicional. Será testada a intersecção de cada um destes novos raios apenas com os objectos do lado respectivo. Após verificada a intersecção, cada um destes raios dará origem apenas a um raio reflectido e um raio transmitido, como no algoritmo tradicional de Ray-Tracing. Para não prejudicar a eficiência deste processo, na implementação efectuada manteve-se a estrutura de apontadores que define a ordenação hierárquica para os objectos de um só lado, em paralelo com a ordenação hierárquica global de todos os objectos filhos.

Nos casos em que os dois objectos intersectados são irmãos gémeos, os pontos de intersecção são coincidentes e as normais às superfícies no ponto intersecção são idênticas, podemos efectuar o cálculo da iluminação em conjunto e atribuir o resultado às imagens de ambos os lados. Isto acontece sobretudo em objectos cuja superfície é paralela à linha interocular, como é normalmente o chão, o tecto, o tampo de uma mesa ou uma parede perpendicular à direcção de observação. Nestes casos, é possível enviar apenas um raio para cada fonte luminosa mais um reflectido e/ou um transmitido para as duas vistas. Estes raios podem ser tratados exactamente da mesma forma que os raios originais enviados do ponto de observação.

8.2.3. Resultados Obtidos

A aplicação das técnicas descritas, a um Ray-Tracing tradicional, permite, normalmente, obter as duas imagens (figura 8.8) de um par estereoscópico num tempo inferior ao necessário para calcular as duas separadamente, através do algoritmo tradicional. As imagens não são alteradas pela aplicação de qualquer das técnicas já descritas.

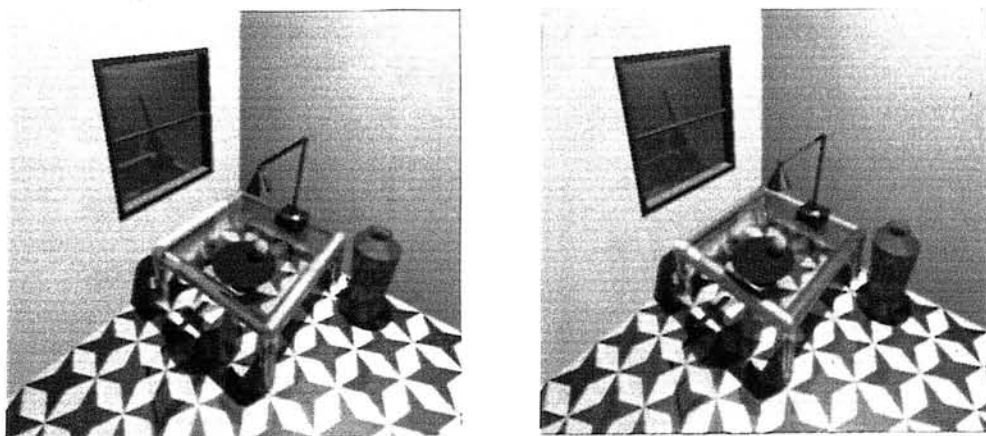


Figura 8.8. Imagem estereoscópica obtida pelo algoritmo apresentado.

Os resultados obtidos para a cena de teste utilizada podem ser observados na tabela 8.1.

Resolução	Mono	distância interaxial				
		0	2	4	6	8
128x128	9.77	16.48	15.28	15.17	16.18	15.00
256x256	38.88	74.17	61.40	61.40	60.98	60.97
512x512	146.40	279.10	243.00	242.73	241.68	240.87
1024x1024	647.95	965.42	987.23	982.05	958.83	957.63

Tabela 8.1. Tempos necessários à obtenção de imagens com diversas resoluções, para o caso monoscópico, e estereoscópico com diferentes distâncias inter-axiais.

A tabela seguinte apresenta a mesma informação, mas agora como percentagem do tempo necessário à geração de duas imagens monoscópicas. Os valores apresentados na coluna Mono, correspondem ao tempo, em segundos, necessário para calcular o par estereoscópico de imagens pelo algoritmo tradicional de Ray-Tracing, e foram obtidos pela duplicação dos valores respectivos da tabela anterior.

Resolução	Mono	distância interaxial				
		0	2	4	6	8
128x128	19.53	84.39%	78.24%	77.65%	82.85%	76.79%
256x256	77.77	95.37%	78.95%	78.95%	78.42%	78.40%
512x512	292.80	95.32%	82.99%	82.90%	82.54%	82.26%
1024x1024	1295.90	74.50%	76.18%	75.78%	73.99%	73.90%

Tabela 8.2. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto no caso monoscópico.

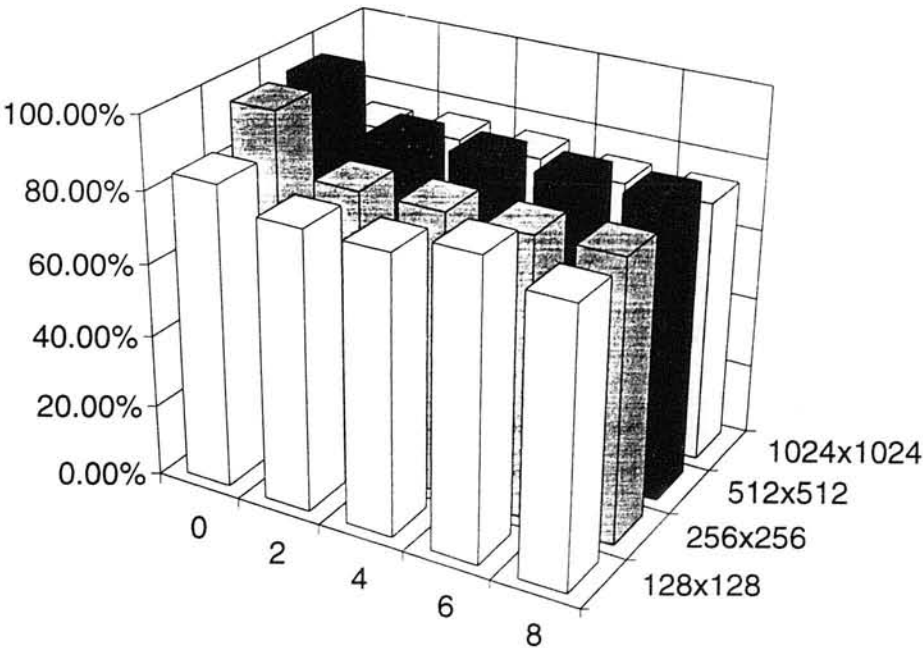


Figura 8.9. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto no caso monoscópico

É possível, no entanto, obter pares estereoscópicos de imagens ligeiramente diferentes dos obtidos pelo algoritmo tradicional, em tempo inferior ao conseguido pela simples aplicação das técnicas descritas.

8.2.4. Cortes na Árvore de Objectos

Uma vez que o efeito estereoscópico só é sentido pelo observador para objectos não muito distantes, podemos prescindir da estereoscopia para os objectos afastados mais do que uma distância previamente definida. Para tal, basta identificar os objectos nessas circunstâncias durante o pré-processamento, e não permitir a geração de filhos a esses objectos originais. Em lugar dos dois filhos gémeos, será utilizado o pai para ambos os lados. Isto permite reduzir o número total de objectos e portanto o número total de testes de intersecção raio/objecto.

As imagens produzidas por este método podem ser consideradas igualmente agradáveis de visualizar. Um exemplo de imagens obtidas com este método pode ser visto na figura 8.10.

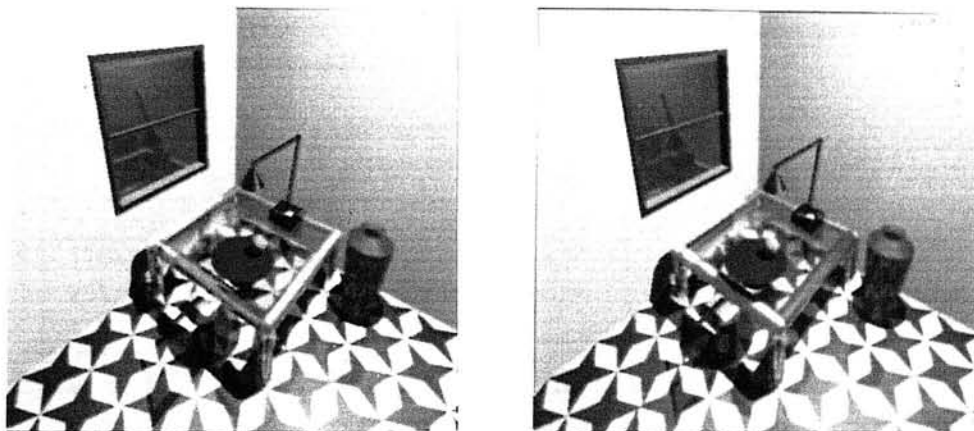


Figura 8.10.

Imagem estereoscópica obtida utilizando cortes na árvore de objectos.

As imagens da figura 8.11. foram obtidas pela diferença entre as imagens da figura 8.10. e a imagem original. Estas imagens correspondem ao desvio introduzido pela aproximação e não devem portanto ser visualizadas em estereoscopia. Podemos notar que as diferenças só aparecem nos contornos dos objectos do fundo.

O utilizador pode negociar uma maior velocidade à custa da diminuição da qualidade da imagem final, simplesmente pela manipulação da distância a partir da qual o efeito estereoscópico é anulado. Valores muito elevados para esta distância permitem obter a imagem geometricamente correcta sem diminuição de tempo. Valores muito baixos implicam a obtenção de um par de imagens idênticas, portanto sem efeito estereoscópico.

Para além da diminuição do efeito estereoscópico, a utilização desta técnica pode introduzir aberrações geométricas, levando ao aparecimento de falhas, buracos, ou sobreposições nos objectos. Isto acontece quando a posição dum objecto relativamente a outro, que deveria estar encostado ou muito próximo, é alterada.

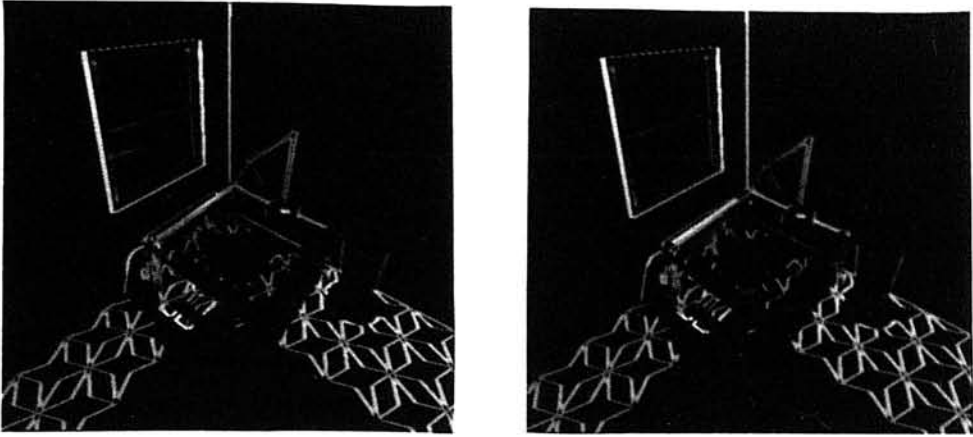


Figura 8.11. Erro introduzido pela aplicação de cortes na árvore de objectos.

Para evitar estes problemas deve-se ter cuidado com a definição da distância de comparação e com o agrupamento dos objectos nestas condições. A técnica pode ser aplicada apenas aos objectos que pertencem a um grupo onde todos os objectos satisfazem as condições de distância ao ponto de observação.

Os resultados obtidos com a utilização deste método podem ser vistos na tabela 8.3.

Resolução	Mono	distância interaxial				
		0	2	4	6	8
128x128	9.77	14.23	14.07	14.05	14.02	14.22
256x256	38.88	56.40	56.08	55.95	56.07	56.10
512x512	146.40	222.93	223.47	222.97	223.03	223.02
1024x1024	647.95	801.28	798.90	798.50	799.52	801.55

Tabela 8.3. Tempos necessários à obtenção de imagens com diversas resoluções, para o caso monoscópico, e estereoscópico com diferentes distâncias inter-axiais.

A tabela seguinte apresenta a mesma informação, mas agora como percentagem do tempo necessário à geração de duas imagens monoscópicas. Os valores apresentados na coluna Mono, correspondem ao tempo, em segundos, necessário para calcular o par estereoscópico de imagens pelo algoritmo tradicional de Ray-Tracing, e foram obtidos pela duplicação dos valores respectivos da tabela anterior.

Resolução	Mono	distância interaxial				
		0	2	4	6	8
128x128	19.53	72.85%	72.01%	71.93%	71.76%	72.78%
256x256	77.77	72.52%	72.12%	71.95%	72.10%	72.14%
512x512	292.80	76.14%	76.32%	76.15%	76.17%	76.17%
1024x1024	1295.90	61.83%	61.65%	61.62%	61.70%	61.85%

Tabela 8.4. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto para uma imagem monoscópica.

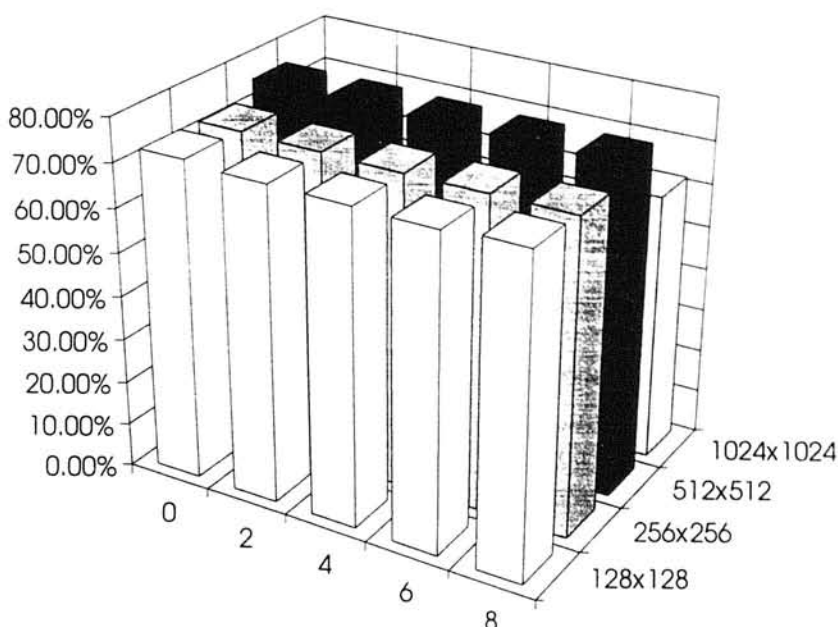


Figura 8.12. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto no caso monoscópico

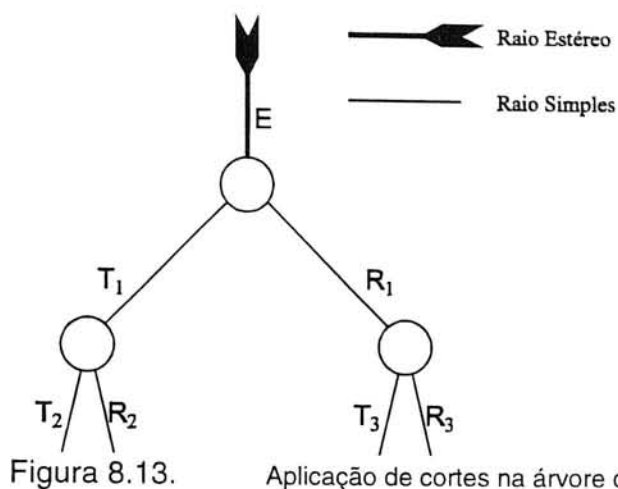
8.2.5. Cortes na Árvore de Shading

É possível conseguir tempos de geração de pares de imagens ainda mais interessantes, embora prejudicando a qualidade da imagem final. No mundo real utilizamos a estereoscopia para determinarmos a distância a que os objectos se encontram. Não é muito importante a existência de efeito estereoscópico nas imagens parcialmente reflectidas em objectos da cena.

Podemos portanto podar a árvore de shading de maneira a não duplicar o cálculo das reflexões e das transmissões. Sempre que dois raios paralelos, partindo cada um do seu olho, atingem a mesma superfície plana do mesmo objecto, podem dar origem a dois raios descendentes por transmissão e por reflexão. Os raios descendentes devido à transmissão do objecto intersectado possuem origens diferentes, mas têm direcções e propriedades luminosas comuns. O mesmo acontece com os raios descendentes por reflexão. A aproximação proposta consiste em substituir cada um destes pares de raios por um outro raio com a mesma direcção e valor, mas com origem no ponto intermédio entre os pontos de intersecção da superfície.

A figura 8.13 mostra a árvore de shading que resulta quando a árvore de shading para um raio é podada logo na primeira intersecção.

Sempre que esta aproximação for aplicável, o tempo de cálculo é diminuído, uma vez que é efectuado o tratamento de um único raio descendente para ambas as imagens, em lugar dos dois raios habituais, um para cada vista.



As imagens produzidas por este método podem ser consideradas igualmente agradáveis de visualizar. Um exemplo de imagens obtidas com este método pode ser visto na figura 8.14.

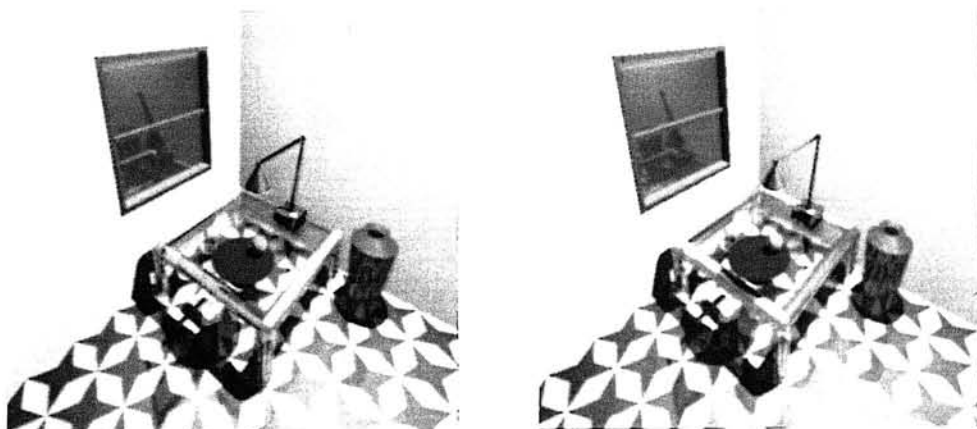


Figura 8.14. Imagem estereoscópica obtida utilizando cortes na árvore de shading.

As diferenças relativamente à imagem original, só aparecem nas reflexões e transmissões. Nota-se bem que os objectos reflectidos no vidro da janela parecem estar no plano desta. Nas imagens anteriores os mesmos objectos reflectidos pareciam estar do lado de fora da janela.

Os resultados obtidos com a utilização deste método podem ser vistos na tabela 8.5.

Resolução	Mono	distância interaxial				
		0	2	4	6	8
128x128	9.77	13.72	14.17	14.42	14.27	14.60
256x256	38.88	66.95	68.13	58.12	58.42	58.97
512x512	146.40	236.60	240.47	231.03	231.85	233.85
1024x1024	647.95	897.48	925.28	916.02	923.28	929.08

Tabela 8.5. Tempos necessários à obtenção de imagens com diversas resoluções, para o caso monoscópico, e estereoscópico com diferentes distâncias inter-axiais.

A tabela seguinte apresenta a mesma informação, mas agora como percentagem do tempo necessário à geração de duas imagens monoscópicas. Os valores apresentados na coluna Mono, correspondem ao tempo, em segundos, necessário para calcular o par estereoscópico de imagens pelo algoritmo tradicional de Ray-Tracing, e foram obtidos pela duplicação dos valores respectivos da tabela anterior.

Resolução	Mono	distância interaxial				
		0	2	4	6	8
128x128	19.53	70.24%	72.54%	73.82%	73.05%	74.74%
256x256	77.77	86.09%	87.61%	74.74%	75.12%	75.83%
512x512	292.80	80.81%	82.13%	78.90%	79.18%	79.87%
1024x1024	1295.90	69.26%	71.40%	70.69%	71.25%	71.69%

Tabela 8.6. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto para uma imagem monoscópica.

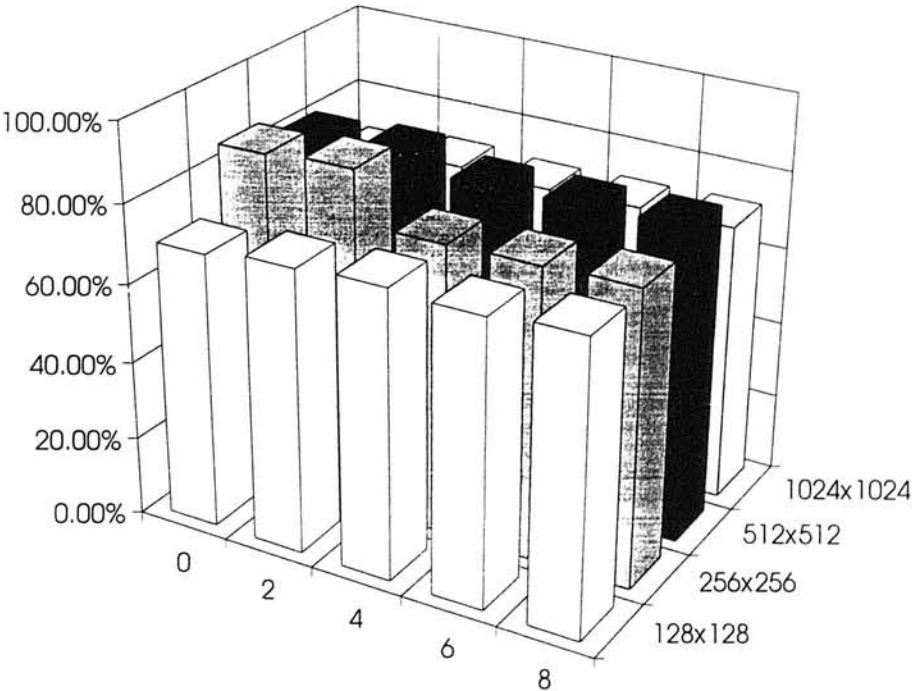


Figura 8.15. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto no caso monoscópico

8.2.6. Raios Reenviados à Origem

Outra técnica que permite fortes acelerações no tempo de geração do par de imagens, consiste em reenviar os raios lançados do ponto de observação, de novo à origem (figura 8.16). Esta é, de entre as técnicas aqui descritas, aquela que maior aceleração permite, embora introduzindo erros na imagem estereoscópica final.

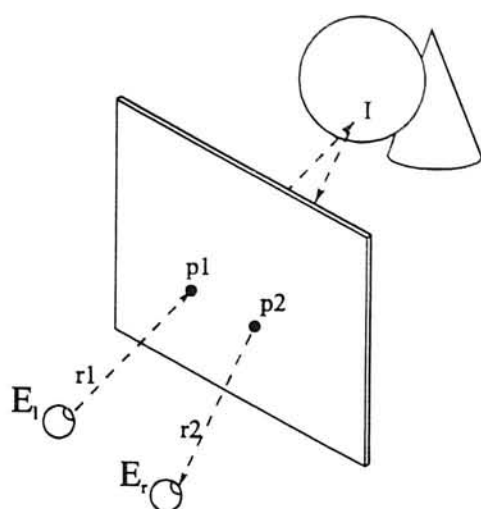


Figura 8.16. Reenvio de raios à origem.

O raio r_1 é enviado a partir do olho esquerdo, através do pixel p_1 . Determina-se o ponto I , que corresponde à intersecção mais próxima com um dos objectos da cena. A iluminação para a atribuir ao ponto p_1 é calculada da forma habitual. Em seguida é enviado um raio desde o ponto I até ao olho direito E_r . A intersecção deste raio com o plano de projecção permite determinar o pixel p_2 , pertencente à imagem do lado direito. A este pixel será atribuído o valor de iluminação já determinado para o pixel p_1 da imagem do lado esquerdo.

Para superfícies perfeitamente difusas, a aplicação desta técnica leva à obtenção das imagens correctas em tempo inferior ao conseguido pela utilização dos métodos anteriormente descritos. No entanto, para objectos que possuam alguma transmissão ou reflexão especular, é introduzido um erro no calculo de iluminação. O erro introduzido é tanto maior quanto mais importante for o efeito especular na superfície considerada.

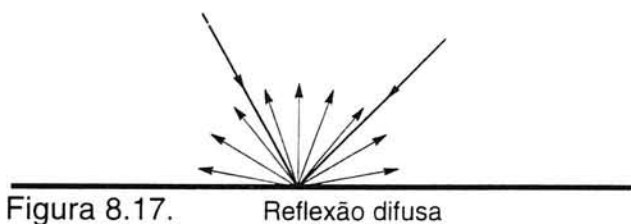


Figura 8.17. Reflexão difusa

Ao contrário do que acontece com a iluminação difusa (figura 8.17), a componente especular da iluminação depende da direcção de observação (figura 8.18). Logo, não é correcto aplicar ao pixel p_2 da imagem do lado direito, o valor de iluminação especular obtido para o pixel p_1 da imagem do lado esquerdo. A aplicação indiferenciada desta técnica introduz portanto erros no cálculo da iluminação especular e transmitida.



Figura 8.18. Reflexão especular

Para que a imagem final não sofra destes defeitos, a técnica de reenvio dos raios ao olho só deve ser aplicada quando todas as superfícies intersectadas nos vários níveis da árvores de shading forem puramente difusas. Em termos práticos, é possível permitir o reenvio dos raios ao olho, desde que os valores dos coeficientes de transmissão e de reflexão especular não sejam significativos.

Tal como foi descrito para a técnica dos cortes na árvore de shading, também aqui é possível fazer depender a aplicação da técnica à comparação dos coeficientes de iluminação do objecto intersectado com um parâmetro de decisão.

Na versão implementada, o reenvio dos raios funciona de maneira um pouco diferente, devido à estrutura de objectos gémeos utilizada (figura 8.19). Um raio enviado do ponto de observação único, através do pixel p1 da imagem do lado esquerdo, intersecta um objecto no ponto I1. O ponto I1 corresponde à intersecção mais próxima do raio com um objecto filho do lado respectivo. O cálculo da iluminação neste ponto é efectuado pelo processo habitual. Com base no ponto I1, é fácil obter o ponto I2, que corresponde ao ponto com a mesma posição relativa, pertencente ao objecto gémeo do objecto intersectado. Do ponto I2 é reenviado um raio para o ponto de observação. A intersecção deste raio com o plano de projecção define o pixel p2, que será usado para atribuir a iluminação calculada.

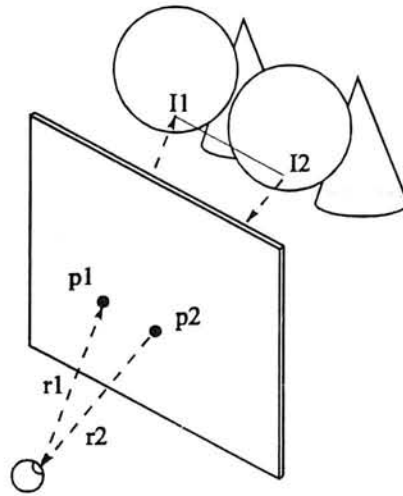


Figura 8.19.

Reenvio de raios à origem, utilizando um ponto de observação único, e objectos transladados.

Uma vez que o segmento que une os pontos de intersecção I1 e I2, tem dimensões e orientação fixas, é possível conhecer o ponto p2 sem necessidade de obter a intersecção do raio reenviado com o plano de projecção. Como a linha interocular é sempre horizontal, as coordenadas verticais dos pixels p1 e p2 são sempre iguais. As coordenadas horizontais diferem da distância entre os dois pixels. Esta distância pode ser obtida por escalamento da distância que une os pontos de intersecção I1 e I2. O factor de escalamento é dado pela relação entre as distâncias do olho ao ponto de intersecção I1 e do olho ao pixel p1.

$$\overline{p1, p2} = \overline{I1, I2} \cdot \frac{\overline{E, p1}}{\overline{E, I1}} \quad (8.1)$$

A distância $\overline{I1, I2}$ corresponde à separação interaxial utilizada no pré-processamento.
Um exemplo de imagens obtidas com este método pode ser visto na figura 8.20.

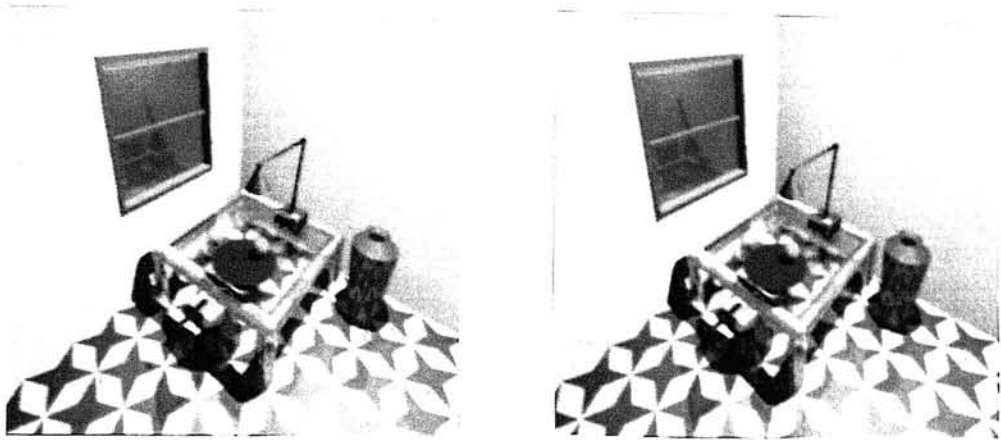


Figura 8.20. Imagem estereoscópica obtida utilizando reenvio de raios à origem.

As imagens da figura 8.21. foram obtidas pela diferença entre as imagens da figura 8.20. e a imagem original. Estas imagens correspondem ao desvio introduzido pela aproximação e não devem portanto ser visualizadas em estereoscopia.

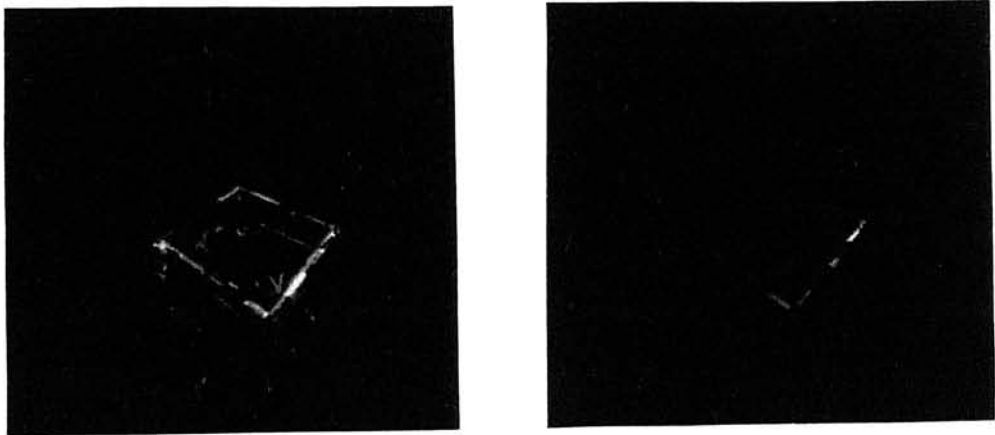


Figura 8.21. Erro introduzido pela utilização de reenvio de raios à origem.

Os resultados obtidos com a utilização deste método podem ser vistos na tabela 8.7.

Resolução	Mono	distância interaxial				
		0	2	4	6	8
128x128	9.77	11.70	11.10	11.20	11.20	12.20
256x256	38.88	48.20	45.30	45.70	51.20	53.40
512x512	146.40	190.80	178.80	180.90	181.90	183.70
1024x1024	647.95	826.00	718.10	722.40	732.00	764.20

Tabela 8.7. Tempos necessários à obtenção de imagens monoscópicas, e estereoscópicas com diferentes distâncias inter-axiais.

A tabela seguinte apresenta a mesma informação, mas agora como percentagem do tempo necessário à geração de duas imagens monoscópicas. Os valores apresentados na coluna Mono, correspondem ao tempo, em segundos, necessário para calcular o par estereoscópico de imagens pelo algoritmo tradicional de Ray-Tracing, e foram obtidos pela duplicação dos valores respectivos da tabela anterior.

Resolução	Mono	distância interaxial				
		()	2	4	6	8
128x128	19.53	59.90%	56.83%	57.34%	57.34%	62.46%
256x256	77.77	61.98%	58.25%	58.77%	65.84%	68.67%
512x512	292.80	65.16%	61.07%	61.78%	62.12%	62.74%
1024x1024	1295.90	63.74%	55.41%	55.75%	56.49%	58.97%

Tabela 8.8. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto para uma imagem monoscópica.

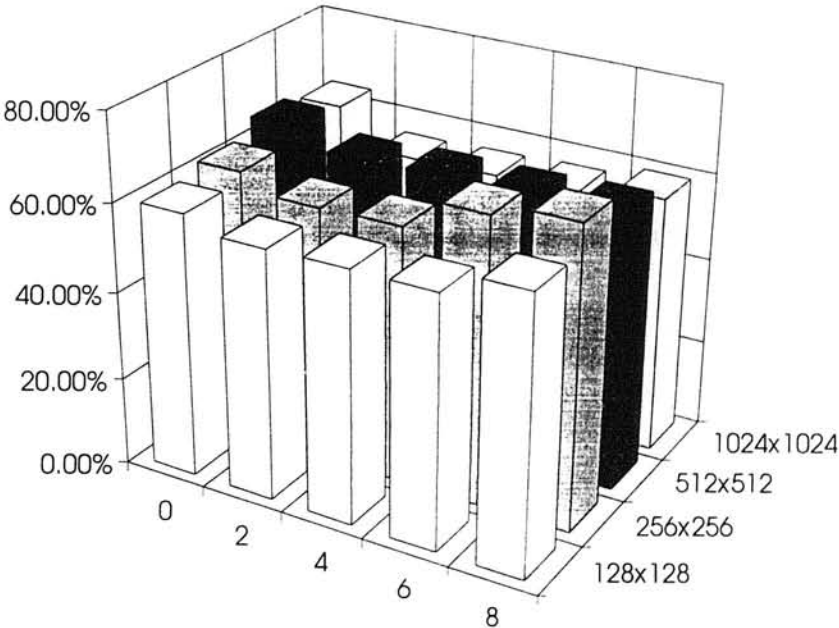


Figura 8.22. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto no caso monoscópico

8.2.7. Utilização da Coerência de Imagem

A aplicação das técnicas que aproveitam a coerência de uma imagem ao algoritmo de Ray-Tracing estereoscópico apresentado, não é difícil. É também pouco complicado introduzir capacidades de obtenção de imagens de realismo crescente. O protótipo desenvolvido incorpora estas duas capacidades, graças aos conhecimentos nesta área adquiridos para o caso monoscópico, em trabalhos anteriores [MLeitão91], [MLeitão92], [MLeitão93], [Sousa90].

A utilização da coerência num algoritmo de Ray-Tracing estereoscópico, permite, tal como no caso monoscópico, obter imagens muito semelhantes às obtidas pelos

algoritmos tradicionais, mas em tempo bastante inferior. A aceleração introduzida com a utilização de selecção de raios adaptativa, está apresentada na tabela 8.9, e as imagens obtidas estão apresentadas na figura 8.23.

Resolução	Mono	distância interaxial				
		0	2	4	6	8
128x128	3.53	6.32	6.63	6.92	6.98	7.23
256x256	8.62	15.02	16.32	16.80	16.88	17.15
512x512	20.95	36.27	39.65	40.12	40.57	40.95
1024x1024	53.48	95.43	103.30	103.90	104.52	105.22

Tabela 8.9.

Tempos necessários à obtenção de imagens com diversas resoluções, para o caso monoscópico, e estereoscópico com diferentes distâncias inter-axiais.

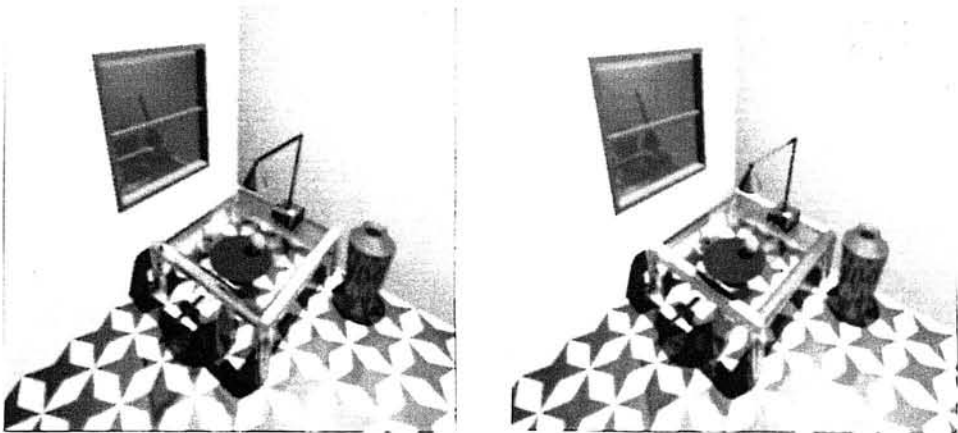


Figura 8.23.

Imagem estereoscópica obtida utilizando coerência de imagem.

As imagens da figura 8.24. foram obtidas pela diferença entre as imagens da figura 8.23. e a imagem original. Estas imagens correspondem ao desvio introduzido pela utilização de coerência de imagem e não devem portanto ser visualizadas em estereoscopia.

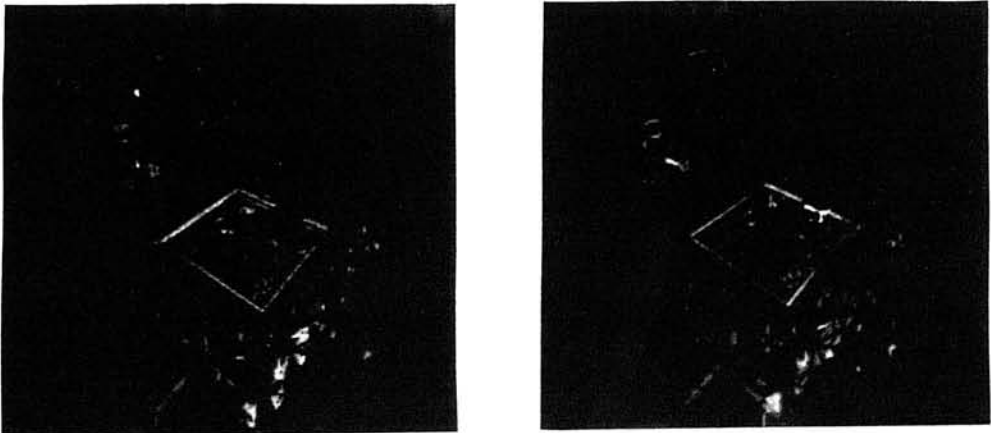


Figura 8.24.

Erro introduzido pela utilização de coerência de imagem.

A tabela seguinte apresenta a mesma informação da tabela 8.9, mas agora como percentagem do tempo necessário à geração de duas imagens monoscópicas. Os valores apresentados na coluna Mono, correspondem ao tempo, em segundos, necessário para calcular o par estereoscópico de imagens pelo algoritmo tradicional de Ray-Tracing, e foram obtidos pela duplicação dos valores respectivos da tabela anterior.

Resolução	Mono	distância interaxial				
		0	2	4	6	8
128x128	7.07	89.39%	93.87%	97.88%	98.82%	102.36%
256x256	17.23	87.14%	94.68%	97.49%	97.97%	99.52%
512x512	41.90	86.56%	94.63%	95.74%	96.82%	97.73%
1024x1024	106.96	89.22%	96.58%	97.14%	97.72%	98.37%

Tabela 8.10. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto para uma imagem monoscópica.

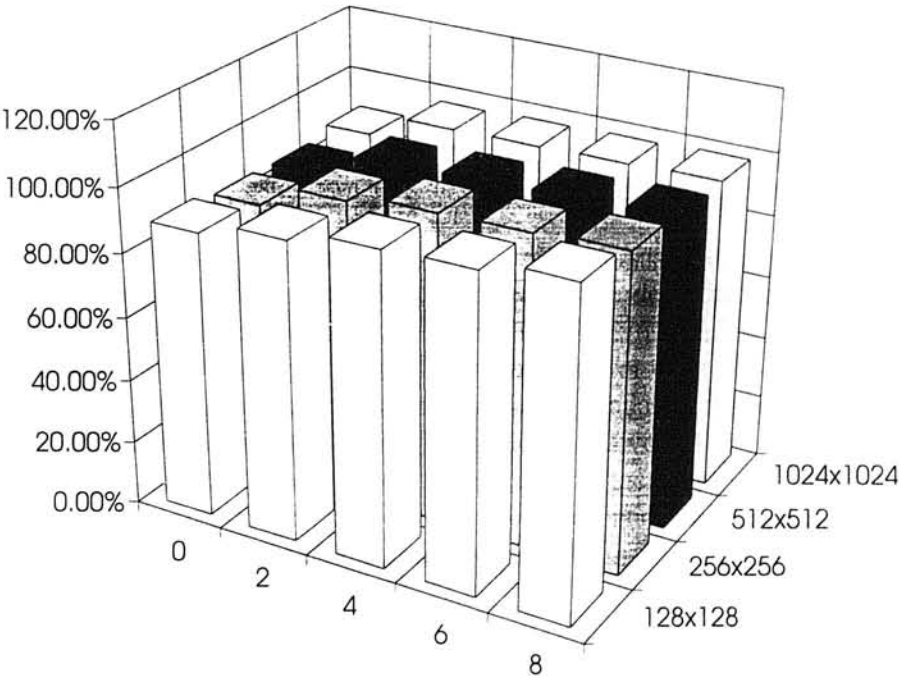


Figura 8.25. Relação entre o tempo necessário à obtenção das imagens estereoscópicas e o dobro do tempo gasto no caso monoscópico

Se compararmos a aceleração introduzida por esta técnica nos algoritmos monoscópico e estereoscópico, podemos verificar que a técnica é mais eficiente no caso monoscópico. Isto deve-se ao facto de o algoritmo estereoscópico já utilizar coerência entre as duas imagens para acelerar o cálculo. Existe portanto coerência que é redundante. Em termos simplificados podemos justificar a situação com o facto de a cor de um pixel poder ser determinada pela coerência que o ponto possui com os vizinhos da mesma imagem, ou pela coerência que possui com o ponto homólogo da outra imagem. O facto de a mesma informação poder ser encontrada por duas vias distintas, não acelera a sua obtenção.

8.3. Conclusão

Os resultados apresentados demonstram que o algoritmo proposto permite gerar um par estereoscópico de imagens em tempo bastante inferior ao algoritmo tradicional de Ray-Tracing. A utilização das técnicas de aceleração apresentadas como facultativas permite ainda uma maior rapidez na geração da imagem estereoscópica, embora por vezes com algum erro introduzido na imagem.

De entre todas as técnicas facultativas apresentadas, a escolha do subconjunto de técnicas facultativas a utilizar em determinada aplicação e a sua parametrização fica a cargo do programador. Nesta escolha, o programador deverá ter em atenção a qualidade requerida para as imagens, o tempo de geração permitido, e ainda o tipo de cenas ou objectos a visualizar. O programador pode ainda passar a decisão sobre as técnicas incorporadas no algoritmo, ao utilizador da aplicação, através da implementação da camada de comunicação respectiva.

De uma maneira geral, podemos afirmar que todas as técnicas facultativas apresentadas são vantajosas para a generalidade das combinações do tipo de cena e da qualidade de imagem requerida. As excepções podem aparecer com a técnica de cortes na árvore de objectos. O erro introduzido por esta técnica pode, em determinadas cenas, ser considerado perturbador, mesmo considerando as vantagens em termos de tempo de cálculo. Esta técnica é mais adequada a cenas com muita profundidade, em que os objectos mais distantes estão isolados dos objectos de primeiro plano.

O algoritmo apresentado, assim como as diversas técnicas facultativas, são directamente aplicáveis à maioria dos dispositivos de visualização estereoscópica disponíveis, incluindo dispositivos do tipo HMD. No entanto, em dispositivos de visualização em que o ângulo entre os eixos ópticos for diferente de zero, à translação aplicada inicialmente aos objectos deve ser adicionada uma rotação com centro na intercepção dos eixos ópticos e com o ângulo respectivo. Nestes casos, a aplicação de cortes na árvore de shading e de reenvio de raios à origem, deve também ser repensada pois aumentará o erro introduzido. Nesta situação encontram-se apenas uma minoria dos dispositivos HMD comercializados [Robinet91].

Capítulo 9.

Conclusões e Trabalho Futuro

O capítulo sobre informação de profundidade, e a descrição dos dispositivos e das tecnologias envolvidas na visualização de imagens estereoscópicas, incluídos nesta tese, devem permitir ao leitor adquirir o conhecimento base necessário à movimentação e trabalho nesta área. A descrição dos dispositivos de visualização estereoscópica, pode ainda ser utilizada como elemento de consulta para, em conjunto com a respectiva pesquisa do mercado, seleccionar o equipamento mais adequado a uma determinada aplicação.

O estudo apresentado sobre os métodos conhecidos de geração de imagens estereoscópicas vai mais longe do que os estudos publicados anteriormente, pois inclui, exaustivamente, os diversos métodos utilizados, com aplicação para cada um, de projecções paralelas e perspectivas. Além disso, contrariamente aos outros estudos comparativos de métodos de geração de imagem, o estudo aqui apresentado inclui também todos os parâmetros que influenciam a visualização das imagens geradas. Desta maneira, todas as conclusões são demonstradas analiticamente, através da relação entre a posição de um ponto no espaço de modelação e a sua posição no mundo virtual.

Foi aqui também tratado o caso de geração de imagens estereoscópicas para dispositivos com largo ângulo de visão, do tipo Head Mounted Display. Para cada caso, é sugerido um método para obter as imagens estereoscópicas correctas. São ainda apresentados os problemas mais comuns na geração de pares de imagens, e alguns conselhos de ordem mais ou menos subjectiva.

O algoritmo apresentado para geração de pares estereoscópicos de imagens, baseado no algoritmo tradicional de Ray-Tracing, permite obter a imagem estereoscópica em tempo bastante inferior ao necessário para gerar o mesmo par pelo algoritmo tradicional. A solução proposta é adaptável à maioria das aplicações que necessitam de imagens com níveis de realismo elevado. Para que isso fosse possível, foram também apresentadas várias técnicas que podem ou não ser incorporadas na implementação final. Algumas destas técnicas permitem ao programador ou ao utilizador negociar uma solução interessante entre um bom tempo de cálculo e uma imagem de excelente qualidade.

O algoritmo proposto é de implementação pouco complexa, pois pode ser implementado a partir de qualquer versão monoscópica do algoritmo de Ray-Tracing. O algoritmo proposto é também compatível com a generalidade das optimizações e

dos melhoramentos publicados para incorporação em Ray-Tracers monoscópicos. Graças à forte ligação do algoritmo apresentado com o Ray-Tracing tradicional, espera-se que esta compatibilidade seja mantida para com todas as optimizações e todos os melhoramentos que deverão aparecer no futuro.

Este trabalho de optimização do algoritmo de Ray-Tracing para a obtenção de imagens estereoscópicas, não pode ser considerado concluído. Apesar dos resultados obtidos serem considerados muito bons, é com certeza possível, e portanto também desejável, obter resultados ainda mais interessantes, em termos de tempo de cálculo e realismo das imagens obtidas. Para estas pesquisas futuras, espera-se que esta tese possa servir como um fiável indicador de alguns caminhos a seguir.

Capítulo 10.

Bibliografia

10.1. Referências Bibliográficas

- [Akka91] Robert Akka, *Creating Stereoscopic Software*, The CrystalEyes Handbook, StereoGraphics.
- [Allister91] David F. McAllister, *3D Displays*, Byte, October 1991.
- [Bains93] Sunny Bains, *Holographic Printing of 3-D Images nears Reality*, Laser Focus World, Vol. 29 No. 4, April 1993.
- [Baker87] J. Baker, *Generating Images for a Time-Multiplexed Stereoscopic Computer Graphics System*, True 3D Imaging Techniques and Display Technologies, Proc SPIE, Vol 761, pp 44-52, 1987.
- [Callaha83] Mark A. Callahan, *A 3D Display Headset for Personalized Computing*. M.S. thesis, Dept. of Architecture, Massachusetts Institute of Technology, 110 pp. 1983.
- [Gelder91] Allen Van Gelder and Jane Wilhens, *Doing it with Mirrors: Low Budget Stereo Graphics*, Visual Data Exploration and Analysis, SPIE Vol 1278, Feb 1991.
- [Glassne89] A. S. Glassner, *An Introduction to Ray Tracing*, Academic Press, ISBN 0.12.286160.4, 1989.
- [Grotch83] S. L. Grotch, *Three Dimensional and Stereoscopic Graphics for Scientific Data Display and Analysis*, IEEE Computer Graphics and Applications 3,8 31-43, Nov 1983.
- [Hodges88] Larry F. Hodges, *Stereoscopic Transformations*, Siggraph 88, Course Notes 21 - Stereoscopic and Multiplanar Computer Graphics.
- [Hodges90a] Larry F. Hodges, *Computing Stereographic Views*, Siggraph 90, Course Notes 6 - Stereographics.

- [Hodges90b] Larry F. Hodges; David F. McAllister, *Rotation algorithm artifacts in stereoscopic images*, Optical Engineering / August 1990 / 29 (8).
- [Hodges91] Larry F. Hodges and Shane Wm. McWhorter, *Stereoscopic Display for Design Visualization*, Signal Processing: Image Communication 4 (1991) 3-13.
- [Holmgren92] Douglas E. Holmgren and Warren Robinett, *Scanned Laser Displays for Head Mounted Displays*, University of North Carolina, Chapel Hill, NC.
- [JDFoley90] J. D. Foley; Andries Van Dam; Steven K. Freiner; *Computer Graphics, Principles and Practice (2^a edition)*, John F. Hughes, Addison-Wesley, 1990.
- [Jones90] Edwin R. Jones, Jr. and A. Porter McLaurin, *VISIDEP: Three-Dimensional Imaging Through Alternating Pairs*, Siggraph 90, Course Notes 6 - Stereographics.
- [Julesz71] B. Julesz, *Foundations of Cyclopean Perception*, Univ. of Chicago Press, Chicago, 1971.
- [Julesz77] B. Julesz, *Recent Results with Dynamic Random-Dot Stereograms*, Three-Dimensional Imaging, SPIE Vol 120, pp 30-35, 1977.
- [Kasper87] Joseph E. Kasper and Steven A. Feller, *The Complete Book of Holograms*, John Wiley & Sons, Inc, ISBN 0-471-62941-3, 1987
- [Kendler74] H. Kendler, *Introdução à Psicologia*, capítulo 4, I volume, 5^a edição, Fundação Calouste Gulbenkian, Lisboa, 1980.
- [King93] Douglas King, *Heads Up*, Computer Graphics World, November 1993.
- [Leep90] *Cyberface II Application Note*, LEEP Systems / Pop-Optix Labs, Waltham, Massachusetts, Aug 1, 1990.
- [Lipscomb79] J. Lipscomb, *Three-Dimensional Cues For a Molecular Computer Graphics System*, PhD. Dissertation, University of North Carolina, Chapel Hill, 1979.
- [Lipton82] Lenny Lipton, *Foundations of the Steroscopic Cinema - A Study in Depth*, Van Nostrand, New York, 1982.
- [Love90] Shaun Love, *Three Dimensional Hardcopy*, Siggraph 90, Course Notes 6 - Stereographics.
- [MLeitão91] J. Miguel Leitão; A. Augusto Sousa; A. Cardoso Costa; F. Nunes Ferreira, *Estratégias e Implementação de Realismo Crescente em Ray-Tracing*, 4^o Encontro Português de Computação Gráfica, Lisboa, 1991.
- [MLeitão92] J. Miguel Leitão; A. Augusto Sousa; A. Cardoso Costa; F. Nunes Ferreira, *Strategies and Implementation of Ray-Tracing with Increasing*

Realism, Second Graphics Workshop on Modelling and Visualisation, Darmstadt - Germany, April 1992.

- [MLeitão93] J. Miguel Leitão; A. Augusto Sousa; A. Cardoso Costa; F. Nunes Ferreira, *Strategies and Implementation of Ray-Tracing with Increasing Realism*, Graphics Modeling and Visualization in Science and Technology, Springer-Verlag, ISBN 3-540-55965-5.
- [Okoshi76] T. Okoshi, *Three-Dimensional Imaging Techniques*, Academic Press, New York, 1976.
- [Penna86] M. A. Penna and R. Petterson, *Projective Geometry and its Applications to Computer Graphics*, Prentice-Hall, 1986.
- [Robinet91] Warren Robinett and Jannick P. Rolland, *A Computational Model for the Stereoscopic Optics of a Head-Mounted Display*, UNC-TR91-009.
- [Sher90] Lawrence D. Sher, *The Oscillating-Mirror Technique for Realizing True 3D*, Siggraph 90, Course Notes 6 - Stereographics.
- [Shirley91] Peter Shirley, Changyaw Wang, *Direct Lighting Calculation by Monte Carlo Integration*, "Second Workshop on Rendering", Barcelona, 1991.
- [Sousa90] A. A. Sousa, A. C. Costa, J. M. Leitão, F. N. Ferreira, *Ray-Tracing Controlável Interactivamente para Produção de Imagens de Realismo Crescente*, 3º Encontro Português de Computação Gráfica, Coimbra, 1990.
- [Southar92] David A. Southard, *Transformations for Stereoscopic Visual Simulation*, Computers & Graphics Volume 16, Number 4, Winter 1992.
- [Sutherl65] Ivan E. Sutherland, *The Ultimate Display*, 1965 Proceedings of the IFIPS Congress, 2, 506-508, 1965.
- [Sutherl68] Ivan E. Sutherland, *A Head-Mounted Three-Dimensional Display*, 1968 Fall Joint Computer Conference, AFIPS Conference Proceedings, 33, 757-764, 1968.
- [Thimble92] Harold W. Thimbleby, Stuart Inglis, and Ian H. Witten *Displaying 3D Images: Algorithms for Single Image Random Dot Stereograms*.
- [Ward91] Gregory J. Ward, *Adaptative Shadow Testing for Ray-Tracing*, "Second Workshop on Rendering", Barcelona, 1991.
- [Whitted80] Turner Whitted, *An Improved Illumination Model for Shaded Display*, Communication of the ACM, 23 (6), June 1980.
- [Wixson83] S. E. Wixson, *Four-Dimensional Processing Tools for Cardiovascular Data*, IEEE Computer Graphics and Applications 3,5 33-59, Aug 1983.

10.2. Outra Bibliografia Utilizada

- [Alliste88] David F. McAllister, *Technologies and Issues for Stereoscopic Display*, Siggraph 88, Course Notes 21 - Stereoscopic and Multiplanar Computer Graphics.
- [Argiro92] Vincent Argiro and William van Zandt *Voxels: Data in 3D*, Byte, May 1992.
- [Balague92] Francis Balaguer, Angelo Mangili, *Virtual Environments*, Computer Graphics Laboratory, Swiss Federal Institute of Technology, Lausanne, ftp.uwashington.edu, 1992.
- [Beaton90] Robert J. Beaton, *Human Factors Engineering of 3-D Stereoscopic Display Systems*, Siggraph 90, Course Notes 6 - Stereographics.
- [Beldie] I. P. Beldie and B. Kost, *Luminance Asymmetry in Stereo TV Images*, Proc SPIE, Stereoscopic Display and Applications, Vol 1457.
- [Bishop92] Gary Bishop and others, *Research Directions in Virtual Environments, report of an NSF Invitational Workshop*, UNC-Chapel Hill, March 23-24, 1992.
- [Bricken90] William Bricken, *Virtual Reality: Directions of Growth. Notes from the Siggraph'90 panel.*, internet paper - sci.virtual-worlds.
- [Brooks92] Frederick P. Brooks, Jr and others, Walkthrough Project - Final Technical Report, TR92-026, University of North Carolina, N.C.
- [Butts90] David R. W. Butts and David F. McAllister, *Implementation of True 3D Cursors in Computer Graphics*, Siggraph 90, Course Notes 6 - Stereographics.
- [Clark92] Jim Clark, *Roots and Branches of 3D*, Byte, May 1992.
- [Cohen85] Michael F. Cohen; Donald P. Greenberg, *The Hemi-Cube, A Radiosity solution for complex enviroments*, Communication of the ACM, 19 (3), June 1985.
- [Cormick87] B. H. McCormick, Visualization in Scientific Computic, Computer Graphics 21, 6, November 1987.
- [Costell92] John P. Costella, *Galilean Antialiasing for Virtual Reality Displays*, School of Physics, University of Melburne, Australia, 25 October 1992.
- [Crow81] F. C. Crow, *Three Dimensional Computer Graphics*, Part 1: Byte, March 1981, pp 54-80; Part2: Byte, Apr. 1981, 11290-302.
- [Davidso92] Clive Davidson, *A physical tour of the imagination*, New Scientist, 29 February, 1992.

- [Eichenl90] Jesse B. Eichenlaub, *Dimension Technologies' Autostereoscopic Method*, Siggraph 90, Course Notes 6 - Stereographics.
- [Fisher90] Scoot S. Fisher and Jane Morril Tazelaar, *Living in a Virtual World*, Byte, July 1990.
- [Fuchs82] H. Fuchs, S. Pizer, Li Ching Tsai, S. Bloomberg and E. Heinz, *Adding a True 3-D Display to a Raster Graphics System*, IEEE Computer Graphics and Applications, Vol 2, Nr 7, pp 73-78, Sep 1982.
- [Harring87] Steven Harrington, *Computer Graphics - A Programing Approach*, McGraw-Hill International Editions, 1987.
- [Hodges85] Larry F. Hodges and David F. McAllister, *Stereo and Alternating Pair Techniques for Display of Computer-Generated Images*, IEEE Computer Graphics and Applications, Vol 5, Nr 9, pp 38-45, Sep 1985.
- [Hollowa87] Richard L. Holloway, *Head-Mounted Display Technical Report*, Technical Report #TR87-015, Dep. of Computer Science, University of North Carolina.
- [Johnson90] Phil Johnson and Richard DeHoff, *Field Sequential Stereoscopic Graphics Systems and Applications*, Siggraph 90, Course Notes 6 - Stereographics.
- [JWest] *Aviary - A Generic Virtual Reality Interface for Real Applications*, Advanced Interface Group, Department of Computer Science, University of Manchester, UK.
- [Kaufman74] Lloyd Kaufman, *Sight and Mind: An Introdution to Visual Perception*. Oxford University Press, New York, 1974.
- [Kirby81] G. H. Kirby and A. Rixon, *Computer Graphics and Stereoscopy for Three-Dimensional Data*, Computer Pysics Comm., Vol 21, pp287-291, 1981.
- [Kreyszi88] Erwin Kreyszig, *Advanced Engineering Mathematics*, Sixth edition, John Wiley & Sons, 1988.
- [Lane82] B. Lane, *Stereoscopic Displays*, Proc. SPIE, Vol 367, pp 20-32, 1982.
- [Leep90] *Cyberface II Application Note*, LEEP Systems / Pop-Optix Labs, Waltham, Massachusetts, Aug 1, 1990.
- [Lipton82] Lenny Lipton, *Foundations of the Steroscopic Cinema - A Study in Depth*, Van Nostrand, New York, 1982.
- [Lipton90] Lenny Lipton, *Some Aspects of Stereoscopic Composition*, Siggraph 90, Course Notes 6 - Stereographics.

- [Lipton91] Lenny Lipton, *Compatibility and selection devices for stereoscopic television*, StereoGraphics Corporation, Image Communication Volume 4 No.1, November 1991.
- [Love88] Shaun Love, *3D Hardcopy*, Siggraph 88, Course Notes 21 - Stereoscopic and Multiplanar Computer Graphics.
- [Nagata90] S. Nagata, *Perceptual Area of Stereoscopic Vision (PASV) and 3D Image Displays*, 3D Forum, Vol 4, Nr 4, pp 42-49, 1990.
- [Norskog] Lance Norskog, *VR Needs*, internet paper, ftp.uwashington.edu. - sci.virtual-worlds.
- [Papatho87] T. V. Papathomas, J. A. Schiavone and B. Julesz, *Stereo Animation for very large data bases: case study - Meteorology*, IEEE Computer Graphics and Applications, 7, pp 18-27, 1987.
- [Pastoor91] Siegmud Pastoor, *3D-television: A survey of recent research results on subjective requirements*, Signal Processing: Image Communication 4(1) 21-32, Elsevier, 1991.
- [Perlmutter83] R. G. Perlmutter and S. S. Friedland, *Computer-generated Holograms in biology and Medicine*, IEEE Computer Graphics and Applications, Vol 3, Nr 5, pp 47-52, 1983.
- [Pimente93] Ken Pimentel, Kevin Teixeira, *Virtual Reality - Through the new looking glass*, Intel/Windcrest/McGraw-Hill, Inc, 1993.
- [Postel85] J. Postel and J. Reynolds, *File Transfer Protocol*. Network Working Group, RFC959, October, 1985.
- [Pountai91] Dick Pountain, *ProVision: The Packaging of Virtual Reality*, Byte, October 1991.
- [RWillia90] R. Don Williams, *Stereographics - Volume 3D Display Technology*, Siggraph 90, Course Notes 6 - Stereographics.
- [Saunders68] B. G. Saunders, *Stereoscopic Drawing by Computer - Is It Orthoscopic?*, Applied Optics, Vol 7, Nr 8, pp 1499-1503, Aug 1968.
- [Sher88] Lawrence D. Sher, *Varifocal Mirror Display*, Siggraph 88, Course Notes 21 - Stereoscopic and Multiplanar Computer Graphics.
- [Skerjan91] R. Skerjanc and J. Liu, *A three camera approach for calculating disparity and synthesizing intermediate pictures*, Image Communication Volume 4 No.1, November 1991.
- [Sousa90b] A. A. Sousa, A. C. Costa, F. N. Ferreira, *Interactive Ray-Tracing for Image Production with Increasing Realism*, Eurographics'90, 1990.

- [Steenbl90] Rochard A. Steenblik, *Chromostereoscopy*, Siggraph 90, Course Notes 6 - Stereographics.
- [Stover82] H. Stover, *True Three Dimensional Display of Computer Data*, Proc SPIE, Vol 367, pp 141-144, 1982.
- [Swillia90] S. P. Williams and R. V. Parrish. *New Computational Control Techniques and Increased Understanding of the Depth-Viewing Volume of Stereo 3-D Displays*, Stereo Displays and Applications, Proc SPIE, Vol 1256, pp 2-11, 1990.
- [Tart91] Charles T. Tart, *Howard Rheingold's VIRTUAL REALITY*, A Review of the book, internet paper, Institute of Noetic Sciences, Sausalito, California 94965, 1991.
- [Tessman89] Thant Tessman, *Perspectives on Stereo*, IRIS Universe, Summer, 1989.
- [Tilton90] Homer B. Tilton, *The Parallactiscope*, Siggraph 90, Course Notes 6 - Stereographics.
- [Valyus66] N. A. Valyus, *Stereoscopy*, The Focal Press, London, 1966.
- [Wallace92] John Wallace and John Fuji *Radiosity*, Byte, May 1992.
- [Wickens90] C. D. Wickens, *Three-Dimensional Stereoscopic Display Implementation: Guidelines Derived from Human Visual Capabilities*, Stereo Displays and Applications, Proc SPIE, Vol 1256, pp 2-11, 1990.
- [Wolf88] R.H.Wolf, C- N.Liu, *Interactive Visualization of 3D Seismic Data: a Volumetric Method*, IEEE Computer Graphics and Applications 8,4 24-30, Jul 1988.
- [Yares92] Evans Yares, *Photo-Realism*, Byte, May 1992.
- [YYeh90] Y. Y. Yeh and L. D. Silverstein, *Limits of Fusion and Depth Judgement in Stereoscopic Color Displays*, Human Factors, Vol 32, Nr 1, pp 45-60, 1990.