

6

Implementação e Análise dos Resultados

Este capítulo mostra a implementação do modelo apresentado para a visualização estereoscópica em curta distância, e apresenta a análise dos resultados obtidos.

6.1

Processamento

A cena utilizada no processamento, por ser uma cena natural, inclui texturas, sombras, oclusão, problemas de iluminação, inerentes ao ambiente natural e às limitações de webcam, em tempo-real. Estas variáveis não estão sendo limitadas no programa, já que o uso de webcams estereoscópicas pretende dar liberdade aos usuários quanto a seu ambiente, iluminação e cores na cena, assim como liberdade de movimento.

O sistema captura as imagens de duas webcams, sincroniza-as e exibe suas imagens em tempo real. A técnica de anaglifo é usada para a visualização do par estéreo, tanto em cinza como colorido, com filtros vermelho e ciano. Anaglifo foi a técnica escolhida por ser a de mais baixo custo computacional, e pela simplicidade dos óculos. Apesar da técnica ser conhecida por não reproduzir bem as cores, o resultado obtido não mostra uma perda significativa de cores.

A resolução usada para visualizar as imagens é de 640 x 480 pixels para gerar uma qualidade de imagem melhor. É importante observar que esta é uma resolução bastante boa para a simplicidade do hardware utilizado. Em vídeo, esta resolução é a máxima admissível para as webcams de primeira linha, e a taxa máxima é de 30 fps, mas nunca atingem isso, a menos de um computador PC com uma configuração mais sofisticada.

6.1.1

Software

Os softwares utilizados foram:

- linguagem C, para o desenvolvimento do programa
- OpenGL para o processamento gráfico

- Glut e Glui para interface
- DirectX para captura e sincronismo das câmeras

O programa foi implementado no ambiente Microsoft Visual C++, na plataforma Windows.

6.2

Hardware utilizado

Os equipamentos usados foram (Figura 6.1):

- PC Intel Pentium 4, 524 MB de RAM, 1,60 GHz de CPU
placa de vídeo NVIDIA GForce2MX440, 512Mb RAM
placa USB2, para obter banda suficiente para sincronizar as duas webcams
- um par de webcams Creative Webcam Live!Pro, com as seguintes especificações:
 - cabo USB
 - sensor de imagem VGA CMOS
 - resolução de vídeo - até 640 x 480 pixels
 - resolução de imagem - modo parado: 1280 x 960 pixels
 - formatos de vídeo: RGB24, I420
 - controle de exposição: automático
 - campo de visão: 57 graus +/- 5%
- óculos de papel com filtro ciano-vermelho

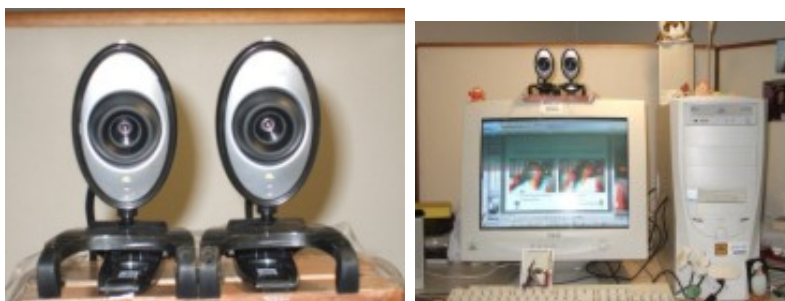


Figura 6.1: Par de câmeras web e PC Pentium4

6.3

Interface do visualizador

A Figura 6.2 mostra a tela do programa. O grupo *Modules* contém os botões para os diferentes módulos de processamento, permitindo a escolha da visualização no Módulo1, que tem a geometria proposta aplicada apenas ao espaço das câmeras, e com projeção no plano; e no Módulo2, com a geometria proposta aplicada ao espaço das câmeras e ao espaço do observador, visualizando-se as imagens estéreo no *DVNP*. O botão para o Módulo Convencional, irá funcionar como o modelo convencional se as webcam tiverem sua calibração modificada, paralelizadas por exemplo. É possível escolher a visualização em anaglifo cinza ou colorido, após a escolha do Módulo.

O grupo de botões *Parameters* permite alterar a resolução da imagem a ser visualizada (largura e altura), e a distância ao observador, que ao ser alterada, o *DVNP* será imediatamente redesenhado, em tempo real, para se adaptar à nova distância do observador.

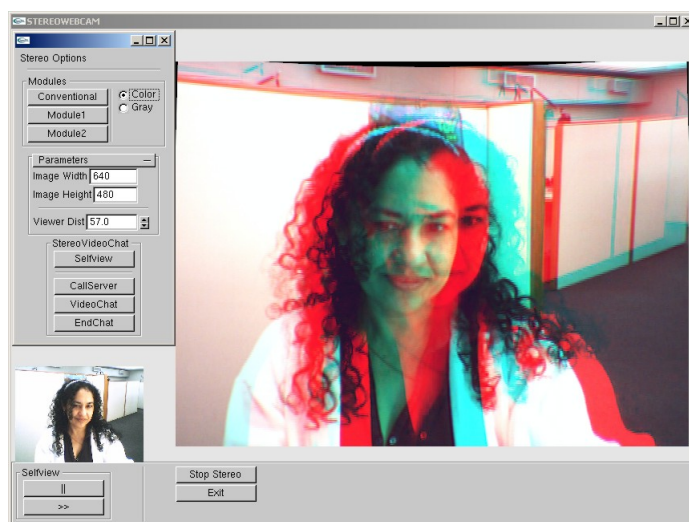


Figura 6.2: Interface do programa

O grupo de botões *StereoVideoChat* são as opções que efetuam a conexão com o servidor, e conseqüentemente o usuário remoto para *EstéreoVideoChat*; sua funcionalidade será descrita na Seção 6.3.1.

A janela pequena no canto inferior esquerdo mostra a imagem capturada de apenas uma webcam, para permitir a visualização da própria pessoa, como é usual em todo programa de webcam. Os botões logo abaixo desta janela permitem “congelar” e liberar a imagem desta janela.

A maior área à direita é destinada à visualização estereoscópica da própria pessoa, quando o programa estiver executando localmente, e a visua-

lização estereoscópica da pessoa remota, quando em conexão, em tempo real. A imagem nesta janela tem dimensão horizontal de 20cm e vertical de 14cm.

Os botões abaixo da janela de estéreo: *Stop Stereo* interrompe a visualização da imagem estéreo e o *Exit* efetua a saída do programa.

A taxa em *frames por segundo* (fps), obtida durante o processamento é mostrada no canto esquerdo superior da tela de estéreo, em tempo real.

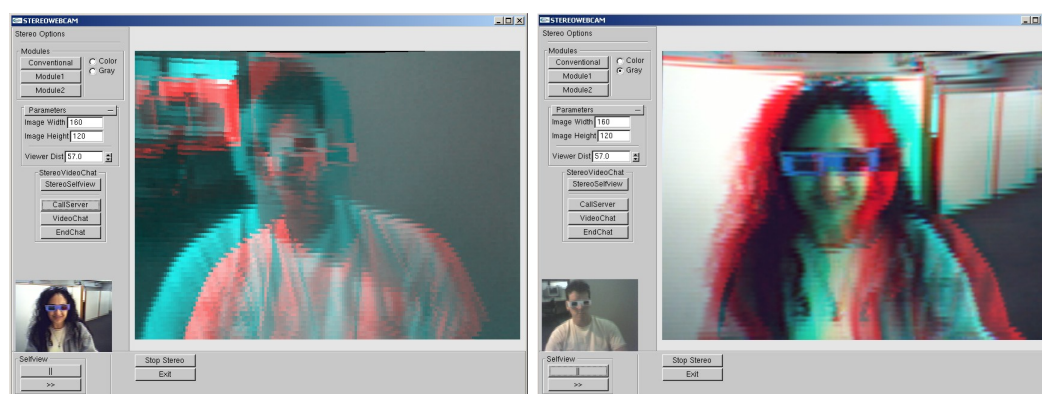
6.3.1

Módulo de Transmissão do vídeo estereoscópico

A transmissão do vídeo estereoscópico foi implementada usando a arquitetura cliente-servidor, e permite a conexão de duas pessoas para a visualização em *StereoVideoChat*, em tempo real.

Na interface do programa, o botão *CallServer*, no grupo *StereoVideoChat*, faz a conexão com o servidor. Se não houver outro usuário conectado, o servidor informa que este usuário deve aguardar uma segunda parte. O botão *VideoChat* pode ser pressionado para tentar iniciar o *StereoVideoChat*, o servidor informará se ainda deve continuar esperando ou iniciará a transmissão caso tenha o segundo conectado. O botão *EndChat* encerra a comunicação remota.

Quando conectados em *StereoVideoChat*, a janela pequena mostra a própria imagem da pessoa local, e a janela de estéreo mostra a imagem estéreo da pessoa remota, como mostra a Figura 6.3. A transmissão está implementada para funcionar com o Módulo2, usando o *DVNP*.



6.3(a): Usuário1

6.3(b): Usuário2

Figura 6.3: *StereoVideoChat*

O desempenho na transmissão é totalmente função do hardware do PC, das webcams, e também do tipo de conexão utilizado como em qualquer VideoChat. Este processamento de transmissão foi feito como uma prova de

conceito para a aplicação que é motivação da tese, por isso não foi usado qualquer algoritmo de compressão para transmissão da imagem.

6.3.2

Visualizando conceitos

ZDL

A imagem 6.4 mostra o ZDL_m , locus de disparidade zero do mundo, na cena capturada pelas câmeras calibradas e projetada na tela de visualização, após sofrer as quatro transformações geométricas.

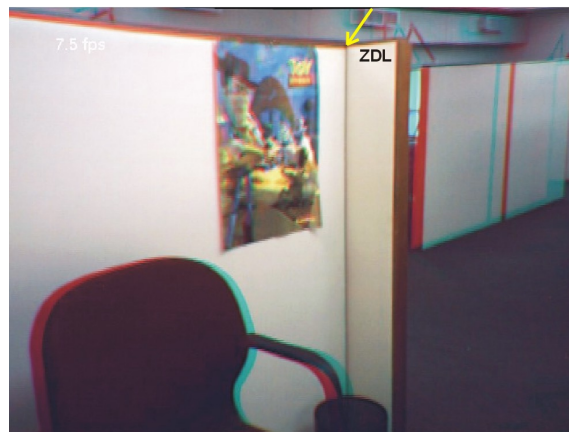


Figura 6.4: ZDL do espaço das câmeras

Observe que as imagens, I_e e I_d estão totalmente sobrepostas no ZDL , horizontal e verticalmente, e que as disparidades próximas ao ZDL são muito pequenas e vão aumentando gradualmente à medida que os planos de profundidade se afastam do ZDL .

Na tela de visualização, o ZDL fica perceptualmente no nível da tela do computador e tudo que está a frente dele tem paralaxe negativo, com a visão estéreo saindo da tela, e toda a cena atrás do ZDL é percebida para dentro da tela.

Display Virtual Não-Planar

Como a câmera virtual que filma o $DVNP$ tem coordenadas diferentes das câmeras reais, isto permite a manipulação do $DVNP$ de forma independente do espaço das câmeras.

As imagens da Figura 6.5 ilustram este conceito, e mostram o $DVNP$ distorcido pela câmera virtual, o que equivaleria ao olhar desviado e torsional do observador. As imagens comprovam o desacoplamento do espaço das câmeras (webcams) e do espaço do observador proposto no modelo.

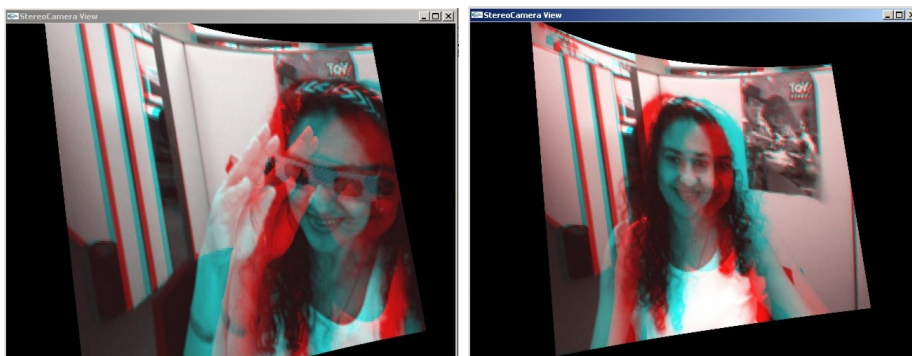


Figura 6.5: *DVNP* em diferentes tomadas da câmera virtual

Ressalta-se que mesmo com a não perpendicularidade da imagem à linha de visão do observador, a fusão das imagens ainda é possível. Mas há uma compressão dos planos de profundidade totalmente diferente dos planos de profundidade obtidos quando a câmera virtual tem a mesma linha de visão das câmeras reais.

6.4

Imagens estéreo resultantes

As imagens apresentadas têm como objetivo mostrar os resultados obtidos e efetuar a comparação e análise das imagens entre o modelo convencional de câmeras paralelas e o modelo proposto, para validar as hipóteses da tese enunciadas no Capítulo 4.

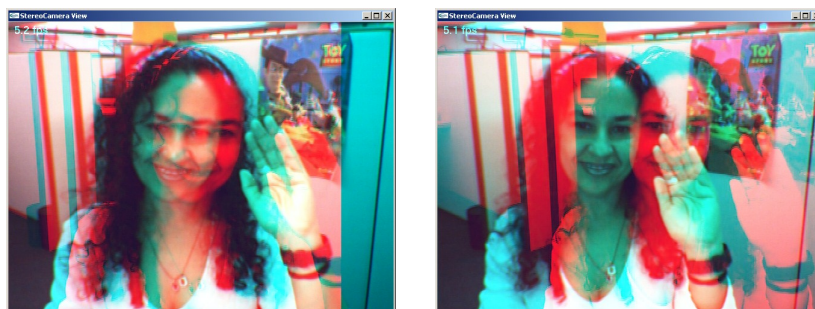
Os primeiros exemplos mostram imagens do modelo convencional com câmeras paralelas, depois apenas imagens com os resultados do Módulo1 e Módulo2 serão apresentadas, porque para calibrar as câmeras para o modelo proposto requer fixá-las cuidadosamente de forma artesanal, por estarem fisicamente separadas e por não haver um protótipo disponível para isso.

Vale lembrar que a distorção de *Crosstalk* ou *ghosting* não foi tratada no processamento das imagens estéreo, e portanto seu efeito será percebido. Isto pode ainda acentuar-se em função da iluminação ambiente que reflete na tela de visualização. Ainda assim, as cores da cena estão bem representadas para a técnica de anaglifo, não caracterizando muita perda de tons de cores. Todas as imagens são anaglifos coloridos ou em tons de cinza, de uma cena complexa.

Reajuste de paralaxe

A Figura 6.6 mostra duas imagens geradas pelo modelo Convencional de câmeras paralelas, uma com o paralaxe adaptado para visualização de curta

distância e a outra com o paralaxe adaptado para longe. É preciso movimentar-se para encontrar o ponto ideal de visualização das imagens estéreo.



6.6(a): paralaxe ajustado para perto

6.6(b): paralaxe ajustado para longe

Figura 6.6: Convencional requer ajustes

A Figura 6.6.a mostra claramente que a imagem com filtro ciano aparece separada na parte direita, sem a imagem de filtro vermelho, ocasionado pela busca de ajuste do paralaxe. Isto ocasiona a perda total do estéreo nesta região da imagem. Nas imagens subseqüentes geradas pelo Módulo1 e Módulo2 do modelo proposto, poderá ser observado que é possível visualizar a imagem variando a posição do observador, sem reajuste de paralaxe.

Fusão das Imagens

As imagens a seguir permitem comparar o processamento Convencional no modelo de câmeras paralelas, com o modelo proposto no Módulo1 e Módulo2, do ponto de vista da fusão. Os exemplos apresentados variam o tipo de iluminação da cena, já que a visualização com webcams deve funcionar para qualquer que seja o tipo de iluminação.

Fusão com luz direta

As imagens foram geradas com luz mais intensa e focada na pessoa, com o objetivo de forçar o excesso de iluminação, por ser um fator que pode dificultar a fusão.

Observando a imagem 6.7.a, a fusão é possível, mas é necessário adaptar-se para encontrar a melhor posição de visualização. Ao varrer os olhos pela cena, através dos planos de profundidade, é natural sentir, pelo olho, a necessidade de readaptação, causando um certo *stress*. O efeito de crosstalk é bastante acentuado. Nesta imagem a iluminação fica reduzida, sobretudo na

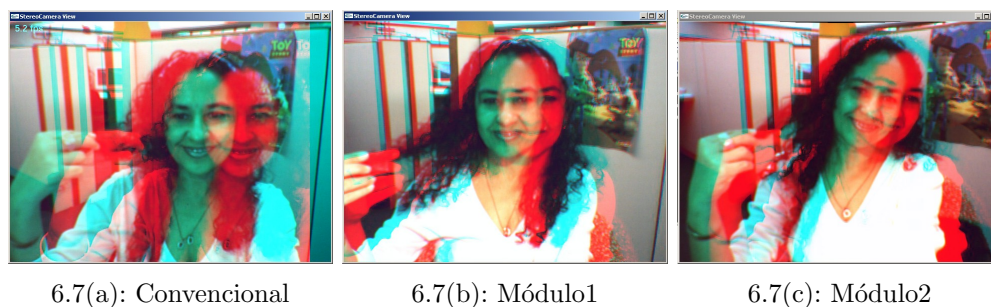


Figura 6.7: Métodos: Convencional, Módulo1 e Módulo2, respectivamente

blusa branca, devido à grande diferença de paralaxe que reduz a sobreposição das cores.

As imagens do Módulo1 e Módulo2, respectivamente, nas Figuras 6.7.b e c, oferecem mais facilidade para fusão imediata, mesmo quando o observador muda o ponto de fixação sobre a cena, e não é necessário procurar a melhor posição de visualização. O crosstalk também está reduzido quando comparado com a imagem convencional.

Fusão com luz indireta

As imagens da Figura 6.8 foram obtidas usando uma iluminação indireta para testar a fusão das imagens nesta condição. A cor de roupa utilizada também seria contra-indicada, pois é azul claro, muito próximo da cor de um dos filtros do anaglifo.



Figura 6.8: Módulo1 e Módulo2

Fusão com meia luz

Estas imagens, Figura 6.9, mostram uma situação de pouca iluminação no ambiente, obtendo-se mesmo assim, boas condições de visualização estereoscópica.



Figura 6.9: Módulo1 e Módulo2

Fusão nas bordas da imagem

Nas imagens da Figura 6.10, pode-se comparar que a facilidade e permanência da fusão é muito semelhante, mas quando o olho aproxima-se das bordas da imagem, principalmente onde está a mão à frente, percebe-se uma perda de fusão na imagem do Módulo1, que não possui as propriedades binoculares incorporadas ao modelo. Na imagem do Módulo2, onde está implementado o *DVNP*, a fusão nas bordas apresenta uma melhoria quando comparada com a imagem do Módulo1.



Figura 6.10: Módulo1 e Módulo2

Cisalhamento

As imagens da Figura 6.11 ilustram a distorção de cisalhamento. Pode-se observar que ao mover a cabeça paralela à tela, movimento no eixo x , a imagem obtida pelo Módulo1 ainda mostra cisalhamento acentuado, ao passo que a imagem obtida pelo Módulo2 apresenta uma redução considerável de cisalhamento.



Figura 6.11: Cisalhamento

Movendo o observador em z

A Figura 6.12 mostra imagens geradas para diferentes posições do observador, para ilustrar a propriedade do *DVNP* (Módulo2) de ajustar a curvatura em função da distância e ângulo de visão do observador.

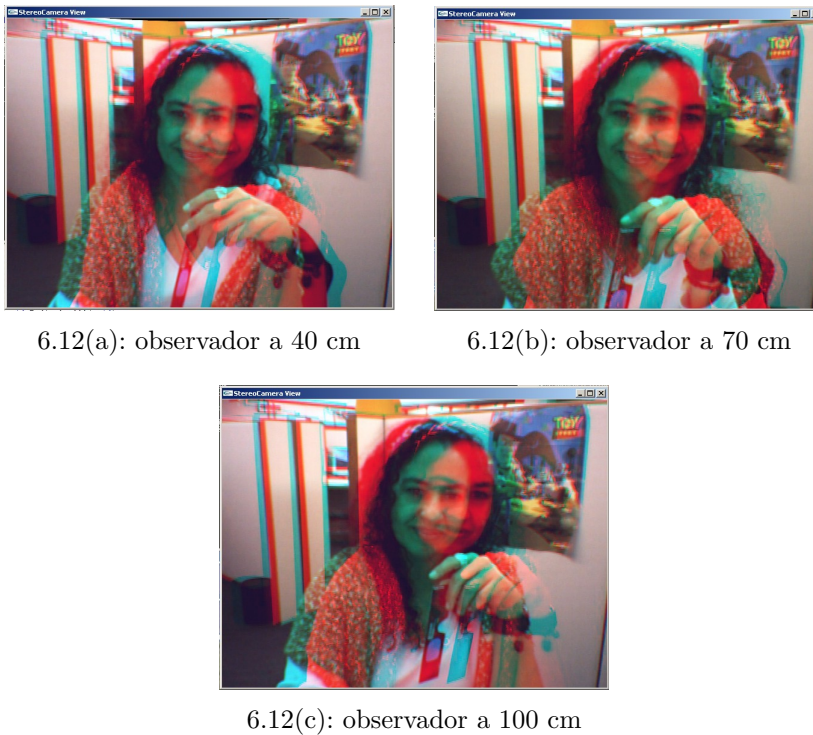


Figura 6.12: Diferentes planos de profundidade

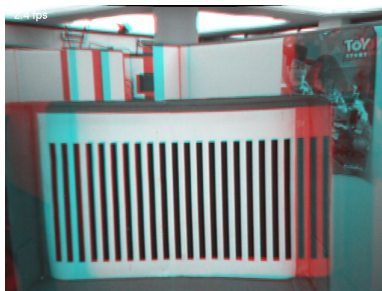
Essa mudança na curvatura do *DVNP* gera uma variação no paralaxe das imagens. Para avaliar o resultado, o observador deve-se posicionar na distância indicada na legenda da figura e observar que a distância relativa

entre os objetos da cena variam, por exemplo entre mão e rosto, causando um efeito mais estéreo na imagem gerada com a distância correspondente.

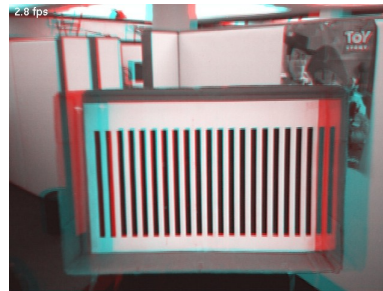
Perspectiva

A Figura 6.13.a mostra a distorção de perspectiva que ocorre na imagem gerada pelo método convencional, com as câmeras paralelas. A imagem visualizada em curta distância, mostra um objeto muito próximo às câmeras. As câmeras paralelas distorcem a perspectiva do objeto próximo e ainda dificultam também a visualização estereoscópica nas bordas da caixa.

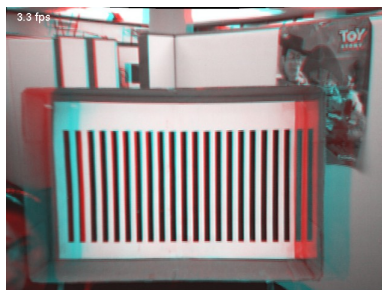
Esta distorção de perspectiva melhora parcialmente pela geometria das câmeras convergentes no Módulo1, Figura 6.13.b. A distorção é compensada pela geometria usada para projeção das imagens no Módulo2, Figura 6.13.c.



6.13(a): Convencional



6.13(b): Módulo1



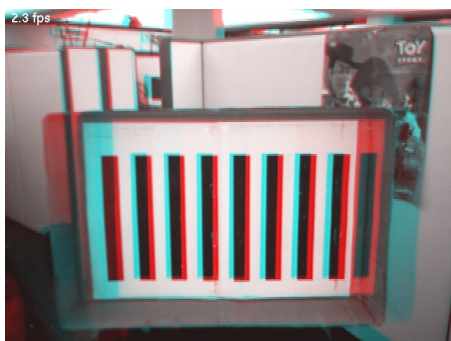
6.13(c): Módulo2

Figura 6.13: Perspectiva e planos de profundidade

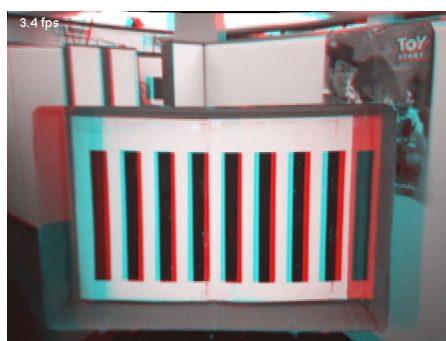
Outro ponto interessante refere-se aos planos de profundidade. Na Figura 6.13.a os planos de profundidade mostram-se mais comprimidos. Na Figura 6.13.c, Módulo2 os planos de profundidade têm uma distribuição mais coerente, muito perceptível principalmente da divisória para a coluna (região de paralaxe positivo), onde a compressão mostra-se bem menor do que na imagem obtida das câmeras paralelas.

Keystone

A distorção de keystone, introduzida nas imagens pela convergência das câmeras, Módulo1, e depois compensada pela geometria do *DVNP*, no Módulo2, pode ser observada na Figura 6.14.



6.14(a): Módulo1



6.14(b): Módulo2

Figura 6.14: Compensação do keystone

6.5

Análise dos Resultados

Os resultados obtidos são bastante significativos em relação às imagens processadas sem a geometria do modelo proposto, tanto em relação à facilidade de fusão das imagens quanto à percepção de profundidade da cena.

Comparando os pontos importantes dos resultados, entre os modelos convencional e o proposto:

– Estéreo Convencional:

ajusta paralaxe x separação das câmeras para obter fusão

a cada profundidade fixada, isto é, diferente ponto de fixação na cena, necessita-se reajustar a imagem (adaptar para o olho)

quando o observador move perpendicular a tela, em z , necessita o reajuste de paralaxe das imagens ou da distância interaxial das câmeras

quando o observador se move paralelo às câmeras, em x - introduz a deformação de cisalhamento na visualização da imagem

– Estéreo com o Modelo proposto:

com a aplicação da geometria proposta, apenas no espaço das câmeras, Módulo1, foi eliminada a necessidade de reajuste de paralaxe

das imagens e da distância entre as câmeras, para cada mudança de posicionamento do usuário

é possível variar o ponto de fixação na cena sem precisar readaptar o olho, nem modificar o *ZDP*

com a geometria aplicada no espaço do observador em função da distância de visualização, foi eliminada a necessidade de reajustar diretamente, pela manipulação das imagens, os parâmetros da imagem (paralaxe), quando o observador move-se perpendicular à tela (movimento em z)

com a geometria aplicada no espaço do observador, o cisalhamento foi praticamente eliminado quando o observador se move paralelo à tela (movimento em x)

a geometria do Módulo2 reduz o keystone introduzido pela convergência das câmeras e melhora a distorção da cena como mostram as Figuras 6.14 e 6.13.

Desta forma, pelos resultados obtidos, confirmam-se as hipóteses da tese enunciadas no Capítulo 4:

- *H1* - Uma geometria adequada para as câmeras web permite a visualização estereoscópica de imagens naturais, em tempo real, sem ajuste contínuo de distância interaxial das câmeras.
- *H2* - A inclusão de parâmetros inerentes à visão binocular e estereopsia no processamento computacional da visualização estereoscópica de imagens naturais melhora a qualidade da imagem estéreo.

Esta geometria também permitiu o desacoplamento entre espaço das câmeras e espaço do observador, e assim, o paralaxe das imagens é adaptado naturalmente em função da distância do observador, pelo comportamento do *DVNP*, não havendo mais uma dependência direta das variáveis do espaço das câmeras.

O Display Virtual Não-Planar, *DVNP*, inova o processamento do estéreo computacional, modificando, por software, a limitação da tela plana e caracterizando um modelo para visualização estereoscópica de curta distância, em tempo real.

H3 - As hipóteses *H1* e *H2* implementadas juntas podem vir a influenciar na redução do desconforto na visualização estereoscópica de imagens naturais.

Há relatos das pessoas que visualizaram as imagens geradas pelo Modelo e exibidas no *DVNP*, de que sentem menos *stress* no olho do que quando observam as imagens geradas pelo modelo convencional. Mas isto não caracteriza

um resultado formal. Para uma validação formal do modelo referente ao desconforto e a confirmação dos resultados da geometria baseada no horopter, sob o ponto de vista do observador, se faz necessário um experimento psicofísico controlado, e isto será indicado como trabalho futuro desta tese.