

5

O estéreo natural e o modelo de estéreo proposto

“Entender o sistema de visão estéreo natural é útil na construção de sistemas artificiais similares. O efeito das características naturais no estéreo computacional é ainda discutido” (35).

A visão estéreo humana é um campo de pesquisa bastante complexo e um dos mecanismos da visão mais pesquisados (35); o interesse aqui é obter, a partir de algumas características específicas deste processo natural, ferramentas para promover melhorias na abordagem computacional.

Para isso, serão apresentados os fatores da visão humana relacionados com a percepção de profundidade, tanto do ponto de vista fisiológico quanto psicofísico.

A psicofísica estuda a relação entre os estímulos físicos e a forma como estes são captados sob certas condições. A psicologia cognitiva tenta explicar o mecanismo de estéreo em um nível mais abstrato do que a neurofisiologia (35).

Este capítulo apresenta pontos importantes do complexo processo natural da visão estéreo humana, relevantes para a abordagem deste trabalho, e descreve o modelo de processamento estéreo proposto.

5.1

Conceitos Básicos da visão

“A percepção visual tem sido freqüentemente comparada à operação de uma câmera. No entanto, esta analogia é quebrada rapidamente porque as câmeras não capturam o que o sistema visual realmente faz, que é criar uma percepção tridimensional do mundo, que é diferente das duas imagens projetadas sobre a retina” (34).

A visão estéreo humana se dá resumidamente em dois processos: a *visão binocular* e a *estereopsia*.

O princípio da binocularidade se baseia no fato do ser humano ter dois olhos perfeitamente alinhados e capazes de enxergar imagens semelhantes que possam fundir-se em uma imagem. Esta binocularidade, e a disparidade

retiniana gerada por ela, é o que nos garante a sensação espacial e de profundidade dos objetos e pessoas.

Os olhos capturam imagens $2D$ dos objetos em torno de nós, visão binocular, e o cérebro usa estas imagens para recuperar a descrição da estrutura $3D$ do ambiente. Estereopsia é o processo responsável por esta reconstrução da dimensão de profundidade no mundo visual (48).

A parte da fisiologia binocular referente ao tratamento das imagens pelo cortex visual não será abordada neste trabalho devido a sua complexidade. Apenas os aspectos da geometria da visão binocular, que correspondem à aquisição das imagens no espaço do mundo até a retina, serão abordados.

5.1.1 Só um pouco de fisiologia do olho

Alguns conceitos relacionados com a fisiologia do olho se fazem necessários para permitir a compreensão de parâmetros que serão considerados no modelo a ser proposto.

A Figura 5.1 ilustra a fisiologia do olho. A **fóvea** é uma região, de maior acuidade ou precisão visual, localizada no centro da mácula e é responsável pela visão central de forma, que é necessária nos humanos para ler, olhar televisão ou cinema, dirigir, e desempenhar qualquer atividade para a qual os detalhes visuais sejam de primária importância (61).

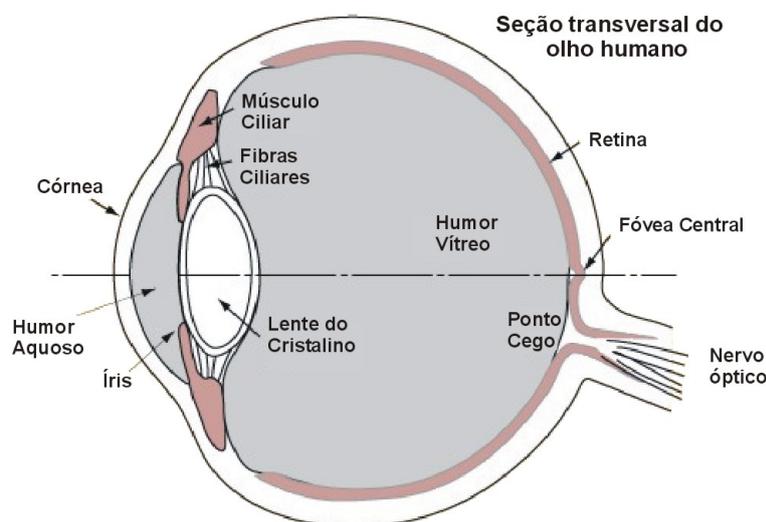


Figura 5.1: Diagrama do olho humano (45)

Quando focamos em um objeto, então, o objeto, o centro da pupila e a fóvea formam uma linha reta, que é o eixo ótico. O centro da fóvea está localizado exatamente no eixo ótico.

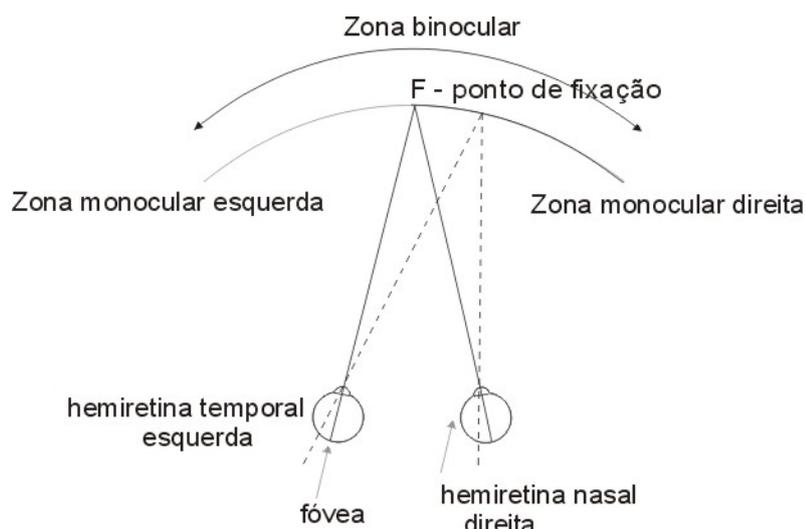


Figura 5.2: Campo visual

O **campo visual** é a visão capturada pelos dois olhos, sem movimento da cabeça. As fôveas de ambos os olhos fixam sobre um mesmo ponto no espaço, chamado *ponto de fixação*. A partir deste ponto, as metades esquerda e direita do campo visual são definidas e chamadas de zona monocular esquerda e zona monocular direita, respectivamente (Figura 5.2). As hemiretinas temporal e nasal são definidas em relação à fóvea, que é a região central da retina com mais alta acuidade visual (34).

A **retina** contém receptores sensíveis à luz (fotoreceptores) que constituem o primeiro estágio da percepção visual. Os fotoreceptores podem efetivamente ser vistos como transdutores convertendo energia de luz para impulsos elétricos (sinais neuronais). Sinais neuronais originados nestes receptores são repassados a outros neurônios que levam aos centros do cortex visual. Fotoreceptores são funcionalmente classificados em cones e bastonetes (*rods*). Os bastonetes são sensíveis à penumbra e ausência de cor (visão noturna), enquanto os cones respondem à claridade, luz cromática (luz do dia). A retina contém 120 milhões de bastonetes e 7 milhões de cones, e estes estão organizados concentricamente (16).

Os pontos no campo visual projetam em posições diferentes na retina. Estas projeções são medidas como distância angular em relação à fóvea, em qualquer direção. Esta distância angular é chamada de **excentricidade retiniana**. Na prática, os olhos movem para uma variedade de posições de fixação, tanto lateralmente quanto em profundidade, quando o observador move a fóvea para inspecionar uma cena visual (57).

5.1.2

Visão espacial

As dimensões das características da retina são usualmente descritas em termos de dimensões da cena projetada, e em termos de graus do **ângulo visual**, definido como:

$$A = 2 \arctan \frac{T}{2D},$$

onde T é o tamanho do objeto da cena e D é a distância do ponto nodal do olho ao objeto. Alguns ângulos visuais conhecidos são apresentados na Tabela 5.3, reproduzida de (16).

Objeto	Distância	Ângulo visual
dedo indicador	comprimento do braço	1.5 – 2°
sol ou lua		0.5° ou 30' de arco
moeda de um quarto de dólar	comprimento do braço	2°
moeda de um quarto de dólar	85m	1' (1 minuto de arco)
moeda de um quarto de dólar	5km	1''(1 segundo de arco)

Figura 5.3: Medidas de ângulos visuais conhecidos (16)

A Figura 5.4 mostra a distribuição retiniana dos receptores cones e bastonetes, relacionados com o ângulo visual em graus. A fóvea contém 147.000 cones/mm² e um número bem menor de bastonetes. Em torno de 10 graus de excentricidade retiniana, o número de cones cai bruscamente para menos de 20.000 cones/mm²; enquanto a 30 graus, o número de bastonetes na periferia cai para 100.000 bastonetes/mm² (16). Esta distribuição ilustra o que foi dito a respeito da visão precisa em torno da fóvea.

De uma forma genérica, o campo visual inteiro corresponde a uma área de 23.400 graus quadrados, definida por um elipsóide, cujo eixo maior horizontal subtende um ângulo visual de 180 graus, e o eixo menor vertical subtende um ângulo de 130 graus (16).

A região de mais alta acuidade visual, a fóvea, subtende um ângulo de 2 graus; a parafóvea, região ainda de alta precisão visual, abrange um ângulo em torno de 4 a 5 graus, e depois disso a precisão visual cai bruscamente. A 5 graus, a precisão visual é de apenas 50%, e o chamado campo visual útil abrange até 30 graus de ângulo visual, Figura 5.5. O resto do campo visual tem poder de resolução muito baixo e é usado, na sua maioria para percepção de movimento no ambiente (16).

Tyler (57) reforça este conceito ao afirmar que a precisão de estéreo é excelente na fóvea e que se torna muito pobre depois de aproximadamente 20

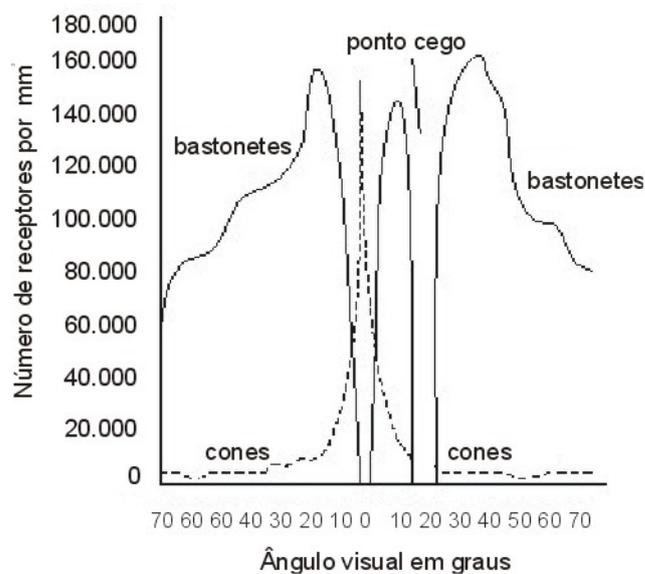


Figura 5.4: Densidade aproximada de cones e bastonetes (16)

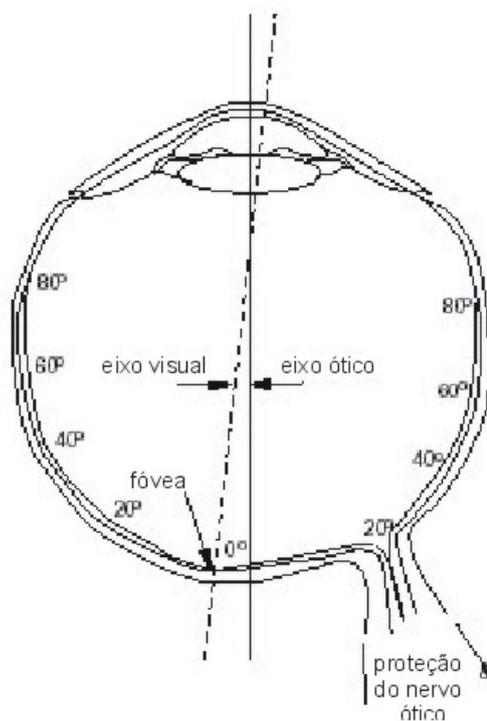


Figura 5.5: Graus do ângulo visual (16)

graus de excentricidade retiniana.

5.1.3 Percepção de profundidade

As cenas naturais são complexas e contêm uma grande variedade de características que auxiliam na determinação de profundidade dos objetos. No

sistema visual humano, a percepção de profundidade se dá por propriedades monoculares (um olho) e binoculares (dois olhos), ilustradas na Figura 5.6. Uma das maiores tarefas do sistema visual é converter a imagem bidimensional retiniana em três dimensões. Como esta transformação ocorre? Como sabemos o quão distante está um objeto do outro? Como estimamos a profundidade relativa de um objeto tridimensional no campo visual? Estudos psicofísicos indicam que a passagem de duas para três dimensões recai sobre dois tipos de informação: as propriedades monoculares e as propriedades estereoscópicas (34).

Primeiro serão descritas as propriedades monoculares e depois a visão estéreo será tratada com mais detalhes.

Propriedades monoculares

As propriedades monoculares criam a percepção de profundidade à distância. Para distâncias maiores que 30 metros, as imagens da retina vistas por cada olho são quase idênticas, portanto, o sistema comporta-se quase que como se tivesse apenas um olho. Estas propriedades monoculares são:

- *Movimento Paralaxe* - esta percepção de profundidade resulta do movimento do olho. Quando nos movemos, as imagens projetadas pelo objeto no campo visual movem-se através da retina. Objetos mais próximos do que o objeto fixado parecem mover-se mais rapidamente e na direção oposta do que o nosso movimento, enquanto que os objetos mais distantes movem-se mais lentamente e na mesma direção do nosso movimento.
- *Tamanho familiar* - se sabemos por experiência, algo sobre o tamanho de uma pessoa, nós podemos julgar a distância desta pessoa a um ponto.
- *Oclusão* - se uma pessoa estiver ocluindo outra parcialmente, nós assumimos que a pessoa em frente está mais próxima
- *Perspectiva Linear* - quando os objetos de distância conhecida subtendem um ângulo menor, ele é interpretado como mais distante. Linhas paralelas convergem com o aumento da distância tais como estradas, linhas de trem, fios elétricos, etc.
- *Perspectiva de tamanho* - se dois objetos similares parecem diferentes em tamanho, assume-se que o menor está mais distante
- *Sombras, iluminação, textura* - padrões de luz, sombra e textura podem dar a impressão de profundidade. Por exemplo, sombras de cores mais brilhantes tendem ser vistas como mais próximas.

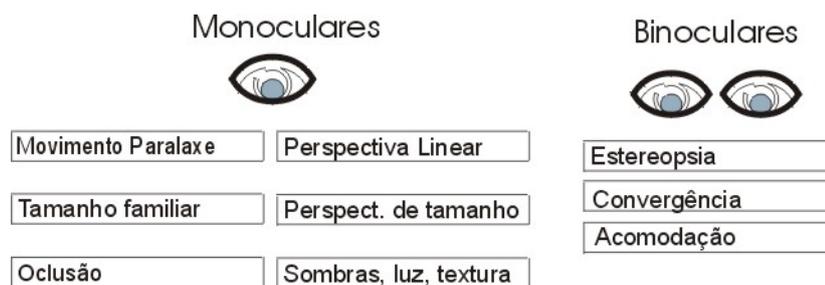


Figura 5.6: Processos para percepção de profundidade

Propriedades binoculares

Outras duas propriedades binoculares importantes são a **acomodação** e **convergência**.

Acomodação é o processo de ajustar o comprimento focal das lentes de forma que o olho se adapte para colocar os objetos próximos ou distantes em foco pelo uso dos músculos, i. e., para focar em uma determinada profundidade ((2),(34)). Para que a acomodação forneça informação de profundidade precisa, o objeto deve estar em foco na retina. Isto quer dizer que de alguma forma o sistema visual deve saber quando um objeto está em foco (46).

A convergência consiste em focar em um objeto pela rotação do globo ocular e ajustar o eixo óptico de cada olho, de forma que eles se encontrem no mesmo ponto.

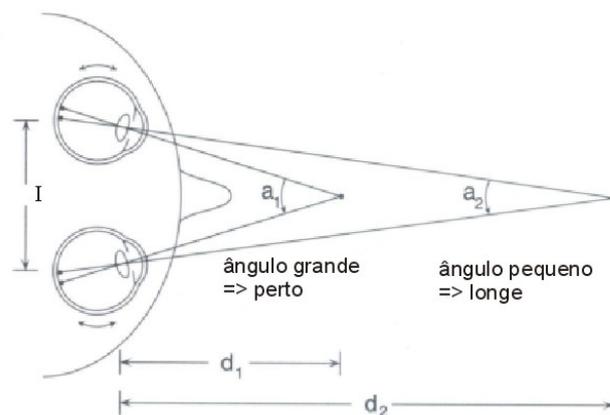


Figura 5.7: Informação de profundidade a partir da convergência (46)

A convergência é uma fonte de informação sobre profundidade, proveniente do quanto os dois olhos giram para dentro ou para fora, para fixar um objeto. Os olhos fixam um dado ponto no espaço externo quando ambos olham diretamente para um ponto, de forma que a luz vinda deste ponto cai no centro de ambas as fôveas, simultaneamente. A Figura 5.7 ilustra a relação de con-

vergência com informação de profundidade, relativas a um ponto de fixação (46).

É genericamente aceito que acomodação e convergência são propriedades de percepção de profundidade menos importantes do que a estereopsia, mas que estas influenciam a percepção de profundidade quando acopladas à visão binocular, especialmente para distâncias do observador inferiores a 2 metros (14).

Versões são os movimentos que envolvem ambos os olhos, mas cada olho se move na mesma direção do outro (61). **Versão pura** é um movimento de versão, em torno de um ponto de fixação simétrica, com ângulo de convergência constante e, após a acomodação, os olhos se movimentam em círculo, ao fazer um movimento de versão pura.

Distância vs. percepção de profundidade

Da literatura de visão humana, é conhecida a alta influência da distância de visao na percepção de profundidade e a dependência da acuidade de estéreo neste aspecto.

À medida que uma pessoa se afasta de um objeto, a imagem retiniana deste objeto gradualmente diminui, e mesmo assim nós não concluímos que o objeto encolheu; de fato, o tamanho do objeto é percebido como constante apesar das mudanças no tamanho da imagem da retina. Este processo é chamado de permanência de escala. A *Lei de Emmert* está diretamente relacionada a esta permanência de tamanho e estabelece que “objetos produzindo imagens retinianas do mesmo tamanho aparentam diferentes em tamanho se eles estiverem localizados em distâncias diferentes”. A lei de Emmert’s está baseada no fato de que o tamanho percebido de um objeto dobra a cada vez que sua distância ao observador cai pela metade e vice-versa (3).

A percepção de diferenças de profundidade diminui com o quadrado da distância do observador para o objeto. Sua capacidade é alta quando o objeto está próximo do observador.

A equação que relaciona distância ao ângulo de convergência é dada por (46)

$$d = \frac{I}{2 \tan(a/2)},$$

onde I é a distância interocular, a é o ângulo de convergência em graus e d é a distância do objeto fixado, baseado na trigonometria de triângulos retos. A equação mostra que se a distância entre os dois olhos é conhecida, o ângulo de convergência do olho pode ser usado pelo sistema visual humano para determinar a distância ao ponto fixado, mas somente para distâncias de poucos

metros. A razão para essa limitação se torna aparente quando o ângulo de convergência é plotado como uma função da distância ao objeto fixado, Figura 5.8.

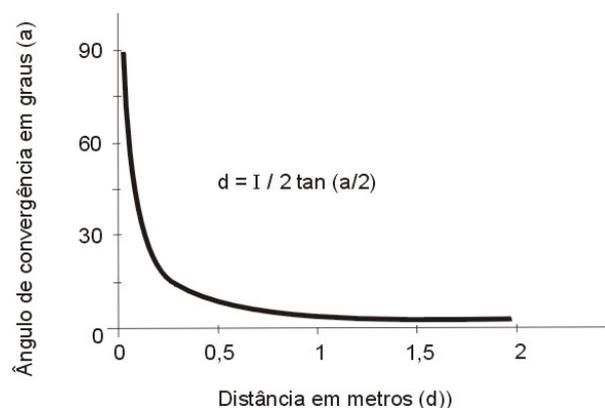


Figura 5.8: Convergência como uma função da distância (46)

Para curtas distâncias, i. e., próximas ao olho, o ângulo de convergência muda rapidamente para pontos que estão apenas levemente diferentes em profundidade, mas para distâncias em torno de 1,8 a 2,4 metros a mudança de convergência é muito pequena, e a partir daí aproxima-se de um limite tendendo a zero (46).

5.2 Geometria da visão binocular

Devido à distância interocular, as projeções de certos pontos nas duas retinas têm posições diferentes. Assim, a visão estéreo está baseada na diferença entre as imagens do olho esquerdo e direito. A diferença destas projeções produz *disparidade binocular*, e em geral depende de muitos fatores, incluindo a estrutura 3D do mundo, ângulo de visão, distância do observador e alinhamento do olho (15).

Disparidades horizontais estão diretamente relacionadas ao posicionamento dos objetos em profundidade, e são conhecidas como a acuidade visual suficiente para a estereopsia. Se um observador fixa o olhar no ponto F , mostrado na Figura 5.9, então as imagens de F cairão em pontos correspondentes nos dois olhos, i. e., nas duas fóveas (15).

O ponto N , mais próximo do que o ponto F , projeta pontos não correspondentes nos dois olhos, e produzirá uma disparidade horizontal representada pelo ângulo a .

Yang (65) define disparidade como a diferença entre o ângulo de convergência associado com o objeto e o ângulo de convergência associado com

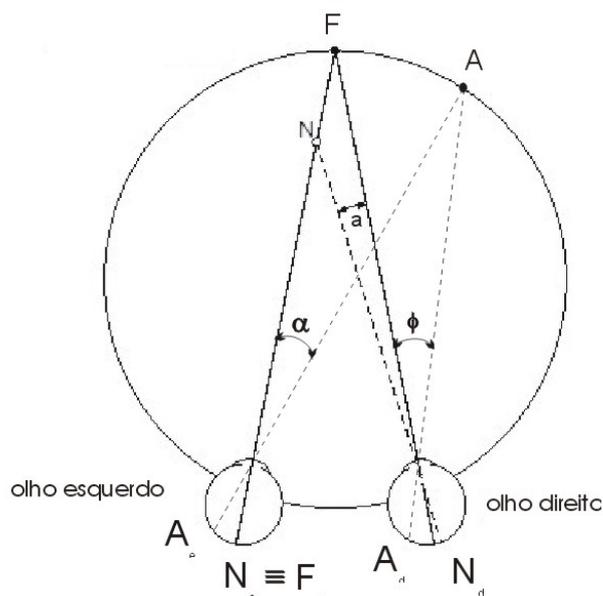


Figura 5.9: Disparidade binocular

o ponto fixado. Na Figura 5.9 este conceito de disparidade fica ilustrado pela diferença dos ângulos $\alpha - \phi$, respectivamente relacionando o ponto A ao ponto de fixação F .

Para convergência simétrica, a profundidade e a disparidade horizontal estão relacionadas pela seguinte expressão simplificada:

$$a \approx I \frac{p}{D^2},$$

onde I é a separação interocular, p é a profundidade de um ponto no espaço relativa ao plano de fixação, e D é a distância de visão (válida se $D \gg p$) (15). É importante observar que a disparidade horizontal é proporcional à profundidade e inversamente proporcional ao quadrado da distância do observador.

Em (12), disparidade é definida como a distância, na direção horizontal, entre os pontos correspondentes das imagens esquerda e direita das imagens retinianas sobrepostas.

Imagens posicionadas na frente do plano de fixação são deslocadas na direção externa de cada retina e têm disparidade cruzada. Imagens atrás do plano de fixação são deslocadas para direção interna à retina e têm disparidade não-cruzada.

Julesz (1960) demonstrou, inventando o estereograma de pontos aleatórios (RDS), nos quais não há qualquer informação sobre superfícies visíveis - exceto a disparidade, que a combinação binocular não precisa acontecer após o reconhecimento de um objeto. A binocularidade aparece

muito antes no sistema visual, antes que qualquer processo complexo de reconhecimento ocorra (48).

No domínio da visão computacional, o problema de combinar as imagens esquerda e direita é extensivamente abordado e muito conhecido como problema de correspondência ou problema de casamento, para possibilitar a reconstrução de informações de profundidade a partir de um par de imagens estéreo. No escopo deste trabalho não será abordada a parte de reconstrução de objetos da cena, mas sim a visualização estereoscópica a partir de um par de imagens estéreo, e portanto o problema da fusão das imagens torna-se principal. Por isso, o interesse estará voltado para analisar como a fusão das imagens ocorre no processo de estéreo natural.

5.2.1

Fusão das imagens

É afirmado pelos pesquisadores que entender quais características da imagem são usadas pelo cérebro para resolver o problema da fusão das imagens é um dos principais tópicos na visão estereoscópica. Ugur Leloglu (35) apresenta alguns fatos da visão estéreo humana que estão proximamente relacionados com a visão estéreo computacional, e diz que apenas as superfícies com um intervalo de disparidade específico, chamado *Área Fusional de Panum*, podem ser fundidas.

Então, torna-se fundamental conhecer as características desta região.

5.2.2

Área Fusional de Panum

A imagem na retina de um olho só pode ser fundida com a imagem na retina do outro olho se as imagens caem na área chamada *Área Fusional de Panum*, que corresponde à região de visão estéreo, mesmo que a disparidade nas imagens retinianas exista; caso contrário, ocorrerá diplopia (visão dupla) (20).

Dentro da área fusional, as duas imagens monoculares causam uma percepção tridimensional única. Objetos fora da área fusional produzem grandes disparidades e geram meias imagens monoculares que normalmente não podem ser fundidas em uma (65).

No espaço do mundo, a área de Panum representa o intervalo total de disparidades que são fusionais, em torno da distância de fixação, sem a ajuda dos movimentos do olho. A presença desta área parece ser a mais simples explicação para a persistência da visão binocular única, apesar das mudanças constantes na disparidade da imagem retiniana (48).

Limites de Fusão

A disparidade máxima sem diplopia (visão dupla) é chamada de Limite Fusional de Panum.

Ugur Leloglu (35) afirma que o intervalo fusional da área de Panum é medido entre 10 – 40 minutos de arco, dependendo dos dados usados.

Já é bem estabelecido pelas ciências que fundamentam a visão de que os limites de fusão variam como uma função de muitos parâmetros, incluindo tamanho da imagem (frequência espacial), frequência temporal, excentricidade, iluminação, envolvimento do mecanismo de vergência, e prática. No entanto, os limites de fusão parecem permanecer relativamente constantes em um intervalo abrangente de luminância e contraste (14).

A Figura 5.10.A ilustra a variação de fusão binocular na área de Panum em função da excentricidade retiniana. A Figura 5.10.B mostra um campo de disparidades geométricas de um plano visto a 20 cm e levemente a frente do ponto de fixação. É importante observar que as disparidades maiores ocorrem nas regiões periféricas, isto representa condições que podem ocorrer durante uma leitura ou escrita (57).

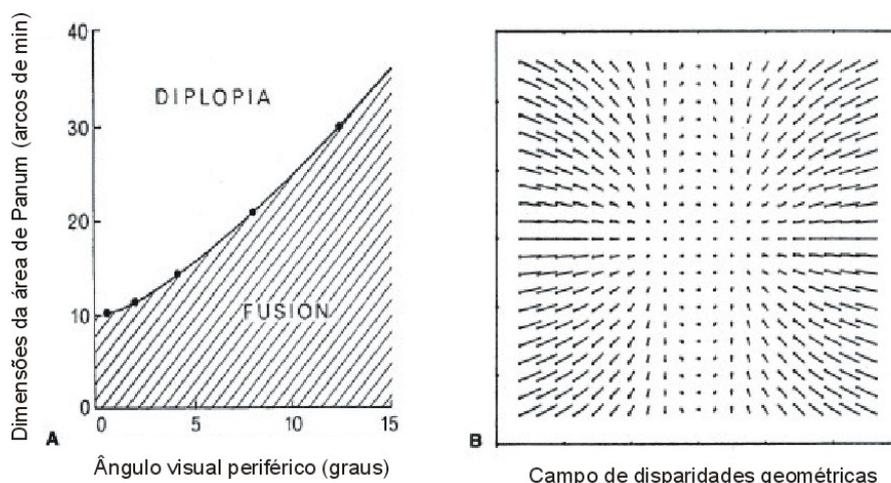


Figura 5.10: Área de Panum e excentricidade da retina (57)

Yei-Yu Yeh, em (14), ao relacionar o limite de fusão com as telas de visualização (displays), fala que o uso de um intervalo de disparidade extremo em uma tela pode resultar em diplopia para porções da imagem que caem fora dos limites de fusão, para fixações de pequenas durações, e pode então exigir grandes e freqüentes mudanças na vergência durante o tempo de varredura na tela. Por outro lado, um intervalo extremo de pequenas disparidades, poderá limitar a utilização de *displays* estereoscópicos mesmo que tais valores possam garantir a fusão de imagens para objetos de vários tamanhos.

5.3

Horopter

Para cada posição de fixação binocular (cima↔baixo, esquerda↔direita, longe↔perto), existe um conjunto de pontos no espaço para os quais a disparidade binocular é zero, significando que estes pontos estão na mesma direção visual no espaço. No espaço, o *locus* de disparidade zero é conhecido como **horopter** (o “horizonte de visão”). Este termo foi introduzido por Aguilonius a aproximadamente quatro séculos passados (57).

O *horopter de pontos*, em geral, é simplesmente uma linha no espaço passando pelo ponto de fixação e conectando todos os pontos no espaço que projetam em pontos correspondentes nas duas retinas, também chamado de *horopter teórico*.

5.3.1

Tipos de Horopter (57)

A definição inicial do horopter foi puramente geométrica e baseada no conceito de correspondência binocular retiniana, com disparidade binocular zero. Este *horopter teórico* deve ser diferenciado do *horopter empírico* medido em um observador, que pode desviar a construção geométrica se existir alguma distorção de correspondência binocular em relação à definição geométrica linear de pontos correspondentes. A diferença entre o horopter teórico e empírico é chamada de *Desvio de Hering-Hillebrand*.

É possível definir um horopter baseado na distância percebida do observador e não só na disparidade binocular propriamente, chamado *horopter de distância*. Este horopter baseia-se no alto nível de percepção envolvendo mais do aparato perceptual, e pode adicionar compensações ou distorções à forma do horopter já definido.

Configurar estímulos em distâncias aparentemente iguais para o observador e com diferentes excentricidades, define o *horopter equidistante*; enquanto que configurar estímulos de forma que os objetos pareçam estar no plano frontoparalelo (abreviado do inglês como AFPP - *apparent fronto-parallel plane*) define o *horopter frontoparalelo*.

O horopter AFPP é usualmente curvo para o observador; a Figura 5.11 ilustra a variação da sua curvatura como função da distância do observador. A distância na qual o horopter AFPP se planifica é chamada de *distância abática* (49).

A curvatura do horopter empírico AFPP é convexa para curtas distâncias do observador. Com o aumento da distância de fixação, a curvatura do horopter

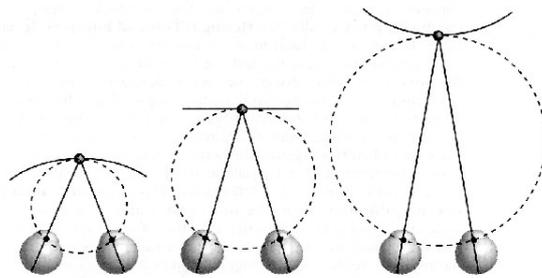


Figura 5.11: O horopter AFPP (10)

muda, variando de convexo para plano, na distância abática. A partir daí, tornando-se côncavo para distâncias maiores.

Um caso especial de interesse envolve o uso de disparidades apenas na direção horizontal. Este caso é equivalente a um horopter medido com linhas verticais e é portanto, conhecido como *horopter longitudinal* por analogia com as linhas verticais de longitude do globo. O horopter longitudinal é um dos mais comumente especificados em textos e é o principal que vem sendo medido empiricamente. Enquanto o horopter de pontos (medido pelo uso de pontos como estímulo) é limitado a uma simples linha no espaço, exceto para fixação no infinito, é importante notar que o horopter longitudinal (medido usando linhas verticais como estímulo) é uma superfície bidimensional curva na forma de um cilindro, que se estende da localização dos dois olhos ao ponto de fixação.

O horopter longitudinal no plano visual horizontal foi primeiramente definido por Aguilonius como um círculo passando pelos centros óticos dos dois olhos e o ponto de fixação; esta construção usualmente é atribuída a Vieth e subsequente Müller, e portanto denominada Círculo de Vieth-Müller (VMC).

O horopter vertical teórico é também definido como uma linha reta através do ponto de fixação, em convergência simétrica. A linha média do horopter vertical é inclinada para longe do observador na parte superior e a sua parte inferior intercepta o solo, nos pés do observador. Isto ocorre porque os meridianos verticais das retinas são inclinados para fora, no topo em relação ao outro, mesmo quando os meridianos horizontais se sobrepõem. Assim, objetos são percebidos como singulares se eles estão acima do ponto de fixação e um pouco afastados, ou se eles estão abaixo do ponto de fixação e um pouco mais próximo. Se estes objetos estivessem no mesmo plano do ponto de fixação eles gerariam disparidades cruzadas e não cruzadas, respectivamente ((10), (57)).

É curioso que por essa propriedade do horopter vertical inclinamos a parte superior da tela do computador para fora, instintivamente.

A fusão no horopter é também relevante para a prática oftalmológica

porque ela especifica o limite entre fusão e diplopia em qualquer lugar do campo visual, que é um critério importante para o conforto visual.

5.3.2 Horopter e o círculo de Vieth-Müller

O Círculo de Vieth-Müller (VMC) é a geometria do horopter horizontal teórico. Ele é baseado no princípio geométrico de que ângulos iguais inseridos em um círculo (i. e., ângulos inscritos) definem arcos iguais neste círculo quando seus vértices caem na sua circunferência. Então, o arco FP , na Figura 5.12, subtende ângulos iguais, α_L e α_R , nos dois olhos (10).

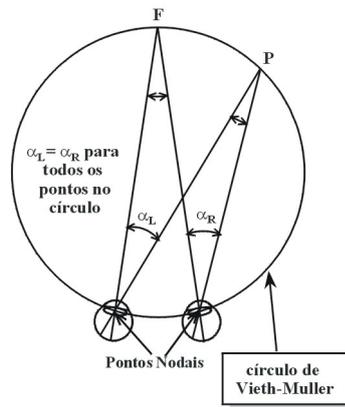


Figura 5.12: Círculo de Vieth-Müller

Genericamente, o círculo de Vieth-Müller descreve o locus de pontos no espaço que produz imagens em pontos geometricamente correspondentes nas duas retinas. Então, por definição, qualquer ponto no círculo de Vieth-Müller terá disparidade horizontal zero. Ao contrário, as imagens dos pontos próximos caem em pontos não correspondentes, com a imagem de um olho lateralmente deslocada da outra (15).

As mudanças na curvatura do horopter empírico espelham mudanças na curvatura do VMC (Figura 5.11).

Propriedades do Círculo de Vieth-Müller (40)

Se o movimento do olho é um movimento de versão pura, o ponto de fixação permanece no mesmo VMC.

Teorema 1 - Se o ponto Q está no círculo de Vieth-Müller, sua disparidade é zero.

Derivado da propriedade geométrica mencionada acima.

Teorema 2 - A disparidade não é zero fora e dentro do círculo de Vieth-Müller (com sinal oposto para fora/dentro).

5.3.3

A curva 3D do Horopter

O caso mais geral do horopter permite a fixação em qualquer ponto, o que inclui convergência assimétrica fora do plano visual. Este caso foi estudado e desenvolvido por Helmholtz, biofísico alemão, em 1866.

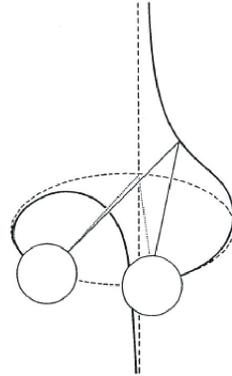


Figura 5.13: Horopter com fixação assimétrica (57)

A Figura 5.13 mostra os olhos fixando um ponto no espaço fora do VMC, i.e., com fixação assimétrica. O VMC e o horopter vertical teórico estão representados pelas linhas tracejadas.

O horopter para fixações assimétricas é uma curva de terceiro grau, que forma uma espiral de uma volta, como mostra a Figura 5.13. Esta curva pode ser construída como o resultado geométrico de se esticar a linha e o círculo de forma a passar pelo ponto de fixação assimétrico.

Do ponto de vista matemático, a curva 3D do horopter está bem descrita em (42) e ilustrada pela Figura 5.14. Sua parametrização cartesiana é dada por:

$$\begin{aligned}x - a &= a \cos t \\y &= b \tan \frac{t}{2}, -\pi < t < \pi, \\z &= a \sin t\end{aligned}$$

e também muito importante, sua projeção no plano xOz , que corresponde ao plano de visão do observador, é dada pelo círculo: $(x - a)^2 + z^2 = a^2$, de raio a , que irá corresponder ao círculo de Vieth-Müller.

A projeção da curva do horopter sobre um cilindro abstrato no espaço permite dissociar os efeitos das rotações oculares das restrições de disparidade inerentes. Desta forma, as cicloversões e ciclovergências torcem o cilindro sem afetar o mapeamento da curva do horopter sobre o cilindro (57).

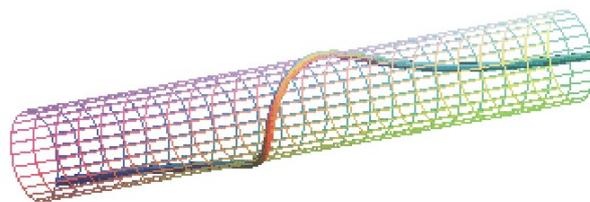


Figura 5.14: Curva 3D do Horopter (42)

Muito recentemente, em Janeiro de 2006, o trabalho de Schreiber, Tweed e Schor (53) estende a definição matemática do horopter para uma superfície cheia no espaço visual, chamada *Horopter estendido*. O horopter estendido quantifica o alinhamento binocular e mostra sua dependência em relação à posição do olho. Reproduz a deformação do horopter teórico em uma forma de espiral (como descrito por Helmholtz) e também descreve um novo efeito de torção ocular, onde o círculo de Vieth-Müller rotaciona fora do plano visual para condições de vergência simétrica em olhares elevados e abaixados, mas o horopter estendido mantém sua localização e forma constantes através das direções dos olhares.

5.4 Experimentos psicofísicos

Esta seção apresenta alguns experimentos psicofísicos existentes na literatura para determinação do horopter, e que são relevantes para ilustrar os conceitos apresentados sobre os diferentes tipos de horopter.

O horopter pode ser medido através de vários métodos. Estes métodos incluem: o *Haplopic*, o *Nonius* e o *Apparent front-parallel plane (AFPP)*. Os métodos Nonius e o AFPP determinam diretamente o horopter longitudinal e o método Haplopic não.

O método Haplopic

O método Haplopic determina os limites interno e externo da visão binocular singular e o horopter é tomado como a linha média destes limites. A Figura 5.15 ilustra o resultado do método Haplopic, para determinar o horopter, considerando a distância do observador a 40cm do objeto alvo (20).

Na Figura 5.15, o eixo vertical representa as medidas, em mm, dos limites superior e inferior da área de Panum. O eixo horizontal mostra os graus de excentricidade retiniana em relação ao ponto de fixação F . O horopter é representado pela curva central tracejada.

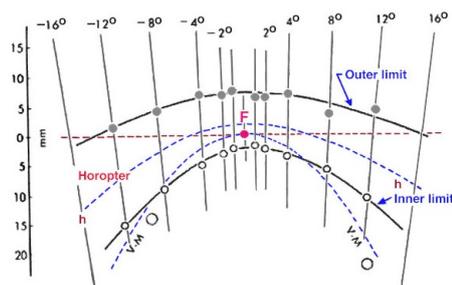


Figura 5.15: O horoptero e a área de Panum (20)

O método Haplopic demonstra a existência da área fusional de Panum, que define a região de visão estéreo.

O método Nonius

O Nonius horoptero é medido pela exposição de apenas a metade de um cilindro de teste para cada olho. Como os dois olhos fixam um cilindro de fixação visto binocularmente, o cilindro de teste é afastado ou aproximado do observador. Quando as duas metades do cilindro de teste são percebidas como alinhadas, elas estão sendo vistas em direções visuais idênticas e, portanto, estimulando pontos correspondentes nos dois olhos (10).



Figura 5.16: Método para medir o horoptero Nonius (10)

O horoptero Nonius é considerado como a técnica mais precisa para medir o horoptero empírico porque mede realmente a direção visual nos dois olhos (49). A Figura 5.16 mostra o teste utilizado para medir o horoptero Nonius.

O método AFPP (10)

O horoptero *Apparent Fronto-Parallel Plane* é produzido da seguinte forma: os cilindros primeiramente são arrumados formando um plano reto perpendicular à linha média de visão do observador, como mostra a Figura

5.17.A. O indivíduo relata então, que os cilindros das extremidades estão afastados para trás, como mostra a Figura 5.17.B. Ao ser solicitado para arranjá-los de forma que, então, eles fiquem em linha reta, o indivíduo desloca os cilindros da extremidade para frente, gerando a distribuição mostrada na Figura 5.17.D, e aí afirma que os cilindros agora estão em linha reta no plano.

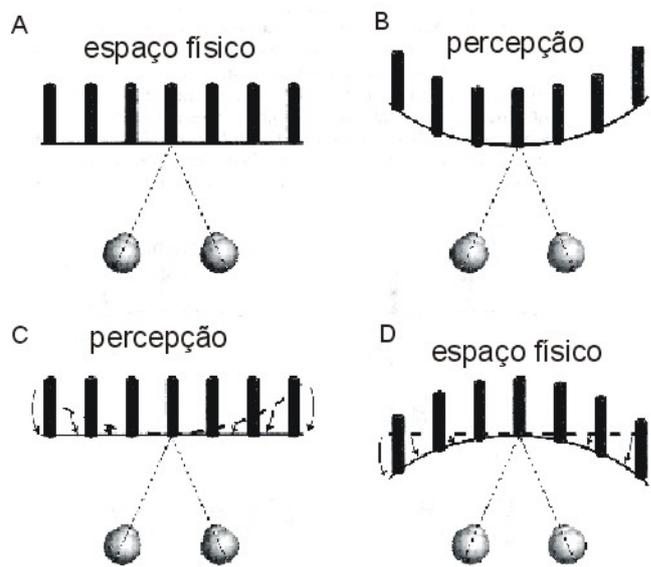


Figura 5.17: O método do horóptero AFPP (10)

Muitos dados apresentados na literatura representando o horóptero são freqüentemente derivados do AFPP, por duas razões: as curvas empíricas mostram os mesmos desvios em relação ao círculo de Vieth-Muller que o horóptero nonius e as medidas são mais fáceis de fazer.

O desvio de Hering-Hillebrand (10)

O coeficiente de desvio de Hering-Hillebrand, H , é uma medida da curvatura do horóptero empírico no ponto de fixação e seu desvio em relação ao círculo de Vieth-Muller. Quando H é igual a zero, o horóptero coincide com o VMC, quando $H < 0$, o horóptero fica interno ao VMC e quando $H > 0$ indica que o horóptero cai fora do VMC.

Após esta análise da geometria binocular, e pelas hipóteses enunciadas previamente, será proposto a inclusão de parâmetros inerentes à visão binocular e estereopsia no modelo de processamento computacional da visualização estereoscópica de imagens naturais.

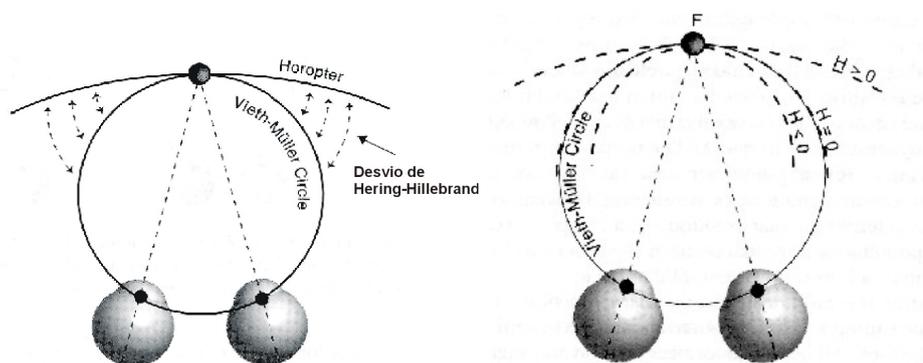


Figura 5.18: O desvio de Hering-Hillebrand (10)

5.5

Um Modelo para visualização estereoscópica em curta distância

Os fatores apresentados como relevantes para a visão estéreo nos estudos psicofísicos, e na fisiologia da visão binocular, não vêm sendo considerados no modelo e cálculo do estéreo computacional, e não estão incorporados diretamente na geometria do modelo utilizado, como exposto no capítulo 4.

Pelo exposto, fica evidente a importância da distância de visão em todo processo da visão estereoscópica, assim como da geometria da visão binocular natural. Como não há na literatura de estereoscopia, uma especificação de processamento computacional diferenciado para visualização estereoscópica de curta ou longa distância, e a aplicação em enfoque é tipicamente uma visualização em curta distância, será proposto um modelo para visualização estereoscópica de curta distância, e em tempo real.

Posicionamento das câmeras

Antes de analisar a geometria de uma webcam em relação a outra, i.e., câmeras paralelas vs. câmeras convergentes (*toed-in*), se faz necessário analisar a posição das webcams em relação ao usuário.

Usuários de webcam sempre estão preocupados com a posição de sua webcam, de forma a obter o que eles consideram o melhor ângulo para mostrar seus rostos. Frequentemente, a webcam é posicionada lateralmente, obtendo uma imagem quase de perfil ao invés da vista frontal de sua face, como mostrado na Figura 5.19.

A Figura 5.20 mostra imagens capturadas de duas câmeras web sincronizadas, e em tempo real, quando as câmeras estão posicionadas lateralmente.

O posicionamento das webcams nesta configuração mostrou que manter a fusão das imagens pelo observador se tornou muito difícil, exigindo que o

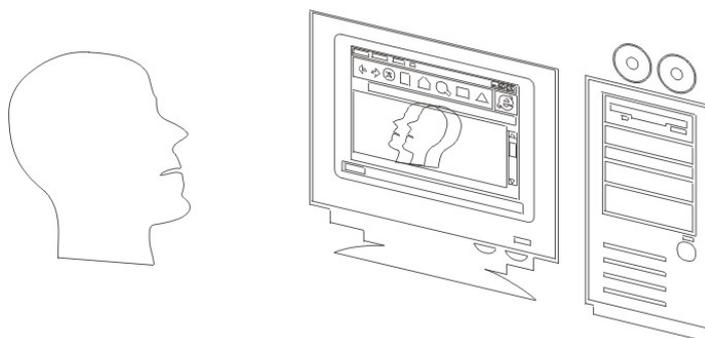


Figura 5.19: Webcams posicionadas lateralmente em relação ao observador



Figura 5.20: Imagens laterais das câmeras 1 e 2

usuário, como observador, mova-se procurando por uma posição melhor para eliminar a diplopia, mesmo usando óculos de filtro.

Esta busca por uma posição melhor que eliminasse a diplopia, levou ao posicionamento frontal das câmeras em relação ao usuário, pois era a indicação natural de paralelizar o usuário em relação as câmeras. A Figura 5.21 ilustra o esquema de configuração do usuário em relação as câmeras, posicionadas em frente a ele(a).

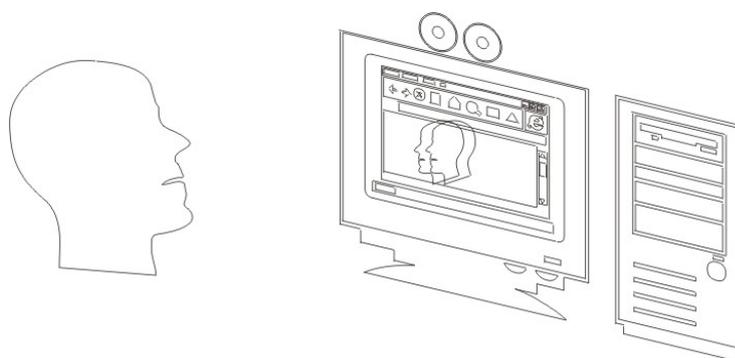


Figura 5.21: Webcams posicionadas em frente ao observador

Este posicionamento frontal das webcams diminuiu a movimentação do observador em busca de uma posição mais favorável à fusão das imagens.

A Figura 5.22 mostra a geometria de ambas as configurações. Para o posicionamento lateral das câmeras, o plano de projeção cai em uma direção diferente da linha de visão do observador.

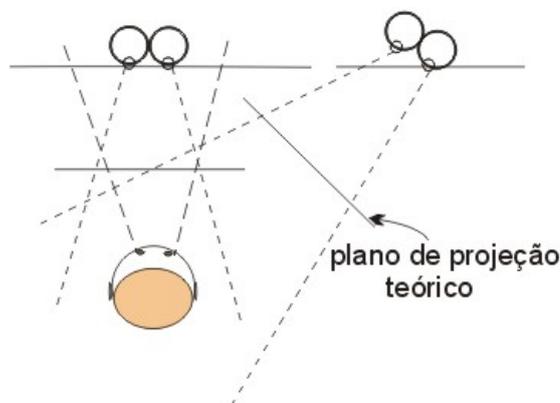


Figura 5.22: Duas geometrias, variando o posicionamento das câmeras

Esta diferença na percepção de fusão das imagens indica que o plano de estéreo precisa estar alinhado com o objeto e o observador; neste caso, alinhado com a pessoa filmada e as câmeras, no espaço do mundo.

Esta análise confirma a dependência entre as posições das câmeras relativas ao observador, de forma a obter a fusão das imagens, e a possibilidade de dependência da geometria do olho para a visualização estereoscópica de curta distância.

Na aplicação com webcams em enfoque não há pós-processamento, a captura e exibição funcionam dinamicamente, em tempo real; por isso, o usuário é o objeto, já que as câmeras funcionam como “olhos” capturando sua imagem; e por outro lado, o usuário é o observador quando ele(a) vê a imagem estéreo gerada, ao mesmo tempo. Neste caso, o objeto e o observador são o mesmo.

Esta interpretação levou a enunciar o *Paradigma do usuário como observador-objeto* (5), ilustrado na Figura 5.23.

Este fato torna a geometria mais complexa porque o processo acontece dinamicamente, dificultando lidar com a interdependência entre as seis variáveis principais, três do sistema de câmera e três do dispositivo de visualização do sistema, descritas na seção 4.3. Note que a distância de visualização do observador para a tela e a distância entre os olhos do observador são mencionadas como variáveis para o dispositivo de exibição.

Assim, se faz necessário que as variáveis para o dispositivo de exibição sejam então desacopladas das variáveis do sistema de câmera, para que mu-

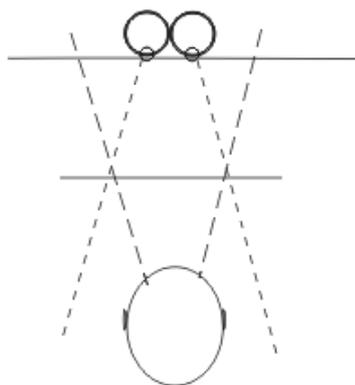


Figura 5.23: O paradigma do usuário como observador-objeto

danças de posicionamento do observador não tenham que refletir diretamente em mudanças nas variáveis das câmeras, como ocorre no modelo convencional.

Portanto, será denominado **espaço do observador** o espaço físico que este ocupa diante da tela de visualização, no mundo, e incluindo as variáveis inerentes ao processo de visualização das imagens na tela; e **espaço das câmeras** ao espaço do mundo, incluindo o campo de visão e as variáveis do sistema de câmeras.

Câmeras paralelas ou convergentes?

Como já foi discutido, muitos trabalhos contra-indicam o uso de câmeras convergentes por introduzir distorções nas imagens, mas busca convergência pela variação de paralaxe das imagens no modelo de câmeras paralelas, o que muitas vezes recai em um processo de tentativas e erros.

Na abordagem proposta, não é possível considerar o observador olhando para o infinito. As imagens estéreo são vistas em curta distância, e como foi exposto, o processo de convergência se intensifica à medida que a distância do observador diminui, exercendo função fundamental neste tipo de visualização.

Na configuração de webcams estereoscópicas, a geometria, com base no horopter teórico, irá influenciar o posicionamento das câmeras para calibração e aquisição das imagens, como também a exibição das imagens (visualização), em função do observador. Portanto, o modelo proposto irá adotar a configuração de câmeras convergentes.

5.5.1

O diagrama do modelo proposto

A estrutura do modelo proposto separa o espaço das câmeras e o espaço do observador com o objetivo de tornar as seis variáveis principais, descritas em 4.3, menos dependentes entre si, ou seja, que as variáveis que são responsáveis

pela visualização possam variar em função do observador, sem precisar que as variáveis do sistema de câmera sejam recalculadas por isso.

Este modelo foi desenvolvido em duas etapas.

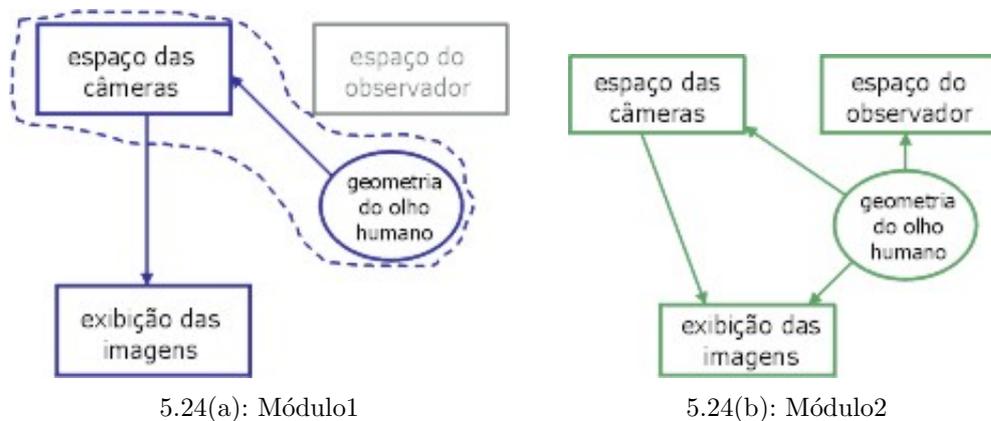


Figura 5.24: Diagrama do Modelo Estéreo Proposto

A primeira, quando a geometria do olho humano é aplicada à geometria do espaço das câmeras, sendo o processo de visualização influenciado apenas indiretamente por esta aquisição. Esta primeira fase de desenvolvimento do modelo tem o objetivo de avaliar o efeito do desacoplamento dos espaços, gradualmente, ainda sem interferir diretamente nas variáveis do espaço do observador; esta etapa é denominada de Módulo1 e seu diagrama é mostrado na Figura 5.24.a.

A segunda etapa (o processamento completo) tem a geometria do olho humano aplicada a ambos os espaços, de câmera e observador (visualização), sendo referenciada como Módulo2 e seu diagrama é apresentado na Figura 5.24.b.

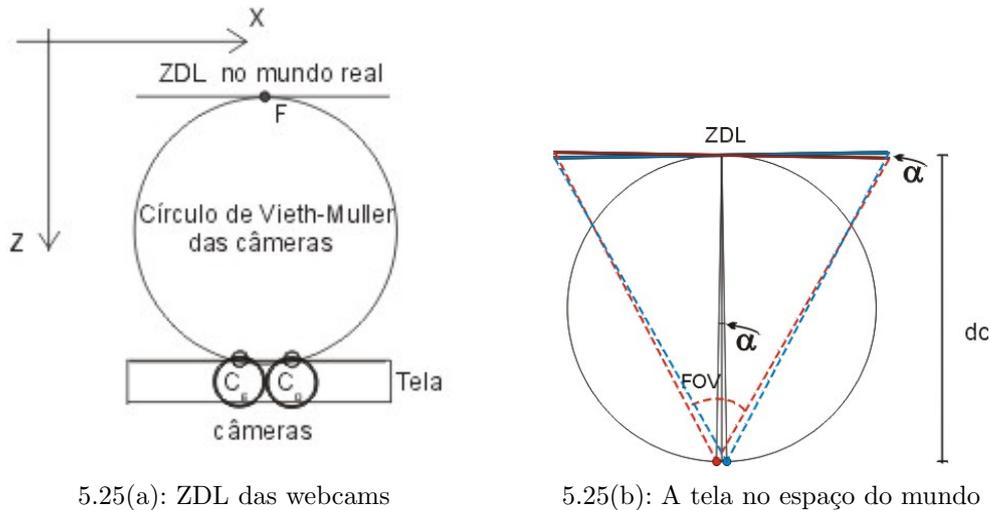
5.6

A geometria do modelo proposto

A geometria da visão binocular, mais especificamente o conceito do horopter, será a base para a concepção do funcionamento dos espaços das câmeras e do observador, que serão descritos a seguir.

Espaço das câmeras

O espaço das câmeras é construído considerando-se a existência do horopter passando pelo eixo óptico das câmeras C_e e C_d , e pelo ponto de fixação F , onde as duas imagens, I_e e I_d , coincidem pela convergência das câmeras, Figura 5.25.a.



5.25(a): ZDL das webcams

5.25(b): A tela no espaço do mundo

Figura 5.25: Espaço das câmeras e o VMC

As câmeras convergem para um Locus de Disparidade Zero (ZDL), no espaço do mundo, onde as imagens das câmeras esquerda e direita se sobrepõem vertical e horizontalmente, introduzindo assim uma região de disparidade zero na imagem estéreo resultante. Esta região de convergência recai sobre uma “tela virtual” no mundo real, e cuja dimensão é determinada em função do campo de visão das câmeras (FOV), Figura 5.25.b.

Para uma webcam, seja W_m a dimensão da “tela” (ZDL_m) no mundo real, então sua dimensão é dada por

$$W_m = 2d_c \tan \frac{FOV}{2}$$

A distância das câmeras ao ZDL, d_c , será grande o suficiente para gerar um ângulo de convergência, α , muito pequeno entre as lentes das câmeras. A determinação desta distância foi assumida em função do espaço de movimentação do observador diante da tela e também influenciada pelo conceito da distância abática, na qual o horopter empírico se planifica, considerando que as imagens adquiridas pelas câmeras são planas. Assim, o espaço de movimentação do observador estará sempre dentro do cone do campo de visão das câmeras, e no qual as disparidades geradas nas imagens estarão limitadas por este intervalo.

No ZDL_m , as imagens das webcams são coincidentes no centro e suas disparidades vão crescendo gradualmente para as bordas das imagens, por causa do ângulo entre elas. Assim, todos os objetos da cena à frente do ZDL_m terão paralaxe negativo na imagem estéreo resultante, e os objetos atrás do ZDL_m terão paralaxe positivo.

A separação interaxial das câmeras será igual à distância média interocular, 6,5cm, e permanecerá constante durante todo o processamento, deixando assim de ser tratada como variável no modelo.

Esta relação entre as imagens, obtida da calibração, permanece fixa durante todo o processamento, não havendo mais modificação do posicionamento das câmeras, e nem posterior manipulação manual do paralaxe no par de imagens estéreo.

Isto caracteriza a modificação na geometria do espaço das câmeras em relação ao modelo convencional.

Espaço do observador

Considerando a distância do observador ao computador como sendo o diâmetro do círculo teórico de Vieth-Müller, que passa pelo ponto nodal dos olhos do observador até o ponto de fixação do objeto observado - no caso a tela do computador - então, as imagens exibidas na tela correspondem ao horopter do observador.

O espaço do observador com base no VMC é mostrado na Figura 5.26.a, onde o diâmetro do círculo é a distância, d_o , do observador à tela. O horopter teórico recai sobre a superfície da tela, correspondendo ao ZDL do observador, ZDL_o , local dos pontos de disparidade zero.

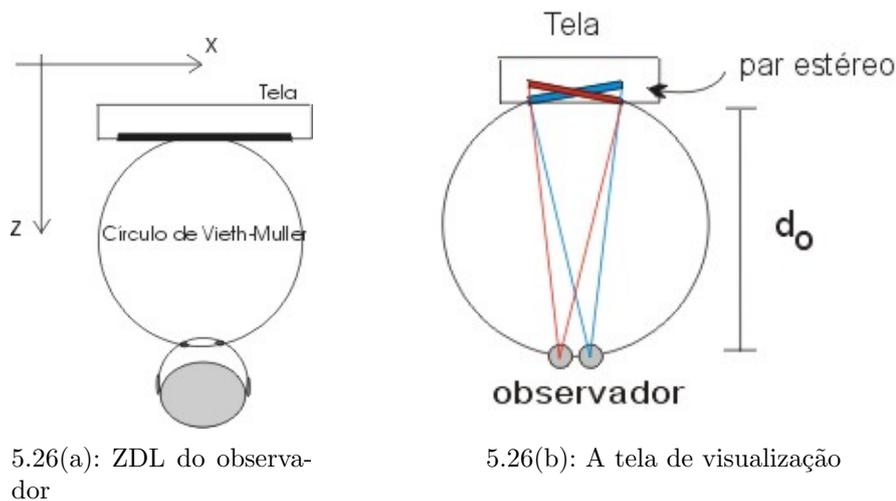


Figura 5.26: Espaço do observador e o VMC

A Figura 5.26.b mostra o par de imagens estéreo na tela, do ponto de vista do observador. As imagens sofrem uma escala para o espaço do observador, mas mantém entre si a mesma relação da aquisição, do espaço das câmeras.

Esta operação de escala do ZDL_m , espaço do mundo, para o ZDL_o , tela de visualização do observador, é dada por

$$S = W_m/W,$$

onde W é a dimensão da tela de visualização da imagem.

5.6.1

Um display virtual não-planar

No modelo proposto, as imagens I_e e I_d , (câmeras esquerda e direita), serão combinadas e tratadas como uma única imagem de textura, com o intuito de manter a relação estabelecida entre as imagens desde o momento em que são capturadas no espaço das câmeras, no mundo real (fase de aquisição).

Com o objetivo de introduzir no modelo um comportamento mais próximo do processamento estéreo natural, entre a visualização das imagens estéreo e o posicionamento do observador, será gerado um objeto gráfico superfície, S , sobre o qual as imagens estéreo serão mapeadas.

Diferentemente do modelo convencional, onde as transformações de translação e projeção ocorrem sobre os objetos gráficos I_e e I_d (imagens), no processo de visualização do novo modelo, todas as transformações de visualização ocorrerão sobre este objeto gráfico superfície, na qual as imagens estão mapeadas.

As imagens combinadas e mapeadas sobre a superfície passam a ter o comportamento da superfície, que será denominada *display virtual* no modelo. Assim, as mudanças que ocorrerem no espaço do observador serão refletidas no comportamento do display virtual, sem com isso causar uma mudança nas variáveis do espaço das câmeras. Fica assim estabelecido o desacoplamento direto dos dois espaços, câmeras e observador.

O Módulo1, como mostrado no diagrama, é uma etapa intermediária de desenvolvimento do modelo e servirá de parâmetro comparativo para análise das imagens estéreo resultantes. Neste módulo, o objeto gráfico superfície é um plano no qual as imagens são projetadas, e portanto a relação dos pixels entre as duas imagens, do momento da aquisição é mantida. Sobre o plano, não haverá qualquer operação em função da movimentação do observador, portanto, uma vez exibido o par estéreo, este ficará invariável qualquer que seja a posição do observador, permitindo que a análise das imagens resultantes recaia totalmente em função da geometria das câmeras (para validação da Hipótese $H1$).

O Módulo2 irá aplicar a geometria do horopter no processo de visualização, para que a visualização das imagens estéreo se torne mais adequada

ao comportamento do observador. A superfície S gerada neste módulo é resultante da síntese dos conceitos sobre o horopter teórico, empírico, e $3D$, e será uma superfície cilíndrica, gerada a partir de um setor circular, com dimensões baseadas no mundo real do observador, segundo a geometria do círculo de Vieth-Müller.

Esta superfície S , será portanto, o **display virtual não-planar, DVNP**, no qual as imagens serão mapeadas. O display virtual não-planar se comportará em função da distância e ângulo de visão do observador e as imagens, como textura, passam a ter o comportamento geométrico da superfície. É importante ressaltar que está sendo incluída também a geometria do horopter vertical neste módulo de visualização.

Então, o Módulo2 irá incorporar parâmetros da visão binocular no processamento computacional da visualização estereoscópica, para validar a Hipótese $H2$; e, como mantém as mesmas características do espaço das câmeras implementadas no Módulo1, este Módulo implementa as Hipóteses $H1$ e $H2$, para validar a Hipótese $H3$.

5.6.2

Transformações de coordenadas

O modelo de processamento estéreo convencional ocorre em três transformações de coordenadas, descritas na Seção 4.3. No modelo aqui proposto, o processamento ocorre em quatro transformações de coordenadas, que serão descritas a seguir.

Como o Módulo1 mapeia as coordenadas dos sensores em um plano, sem modificar as coordenadas das imagens, equivale em transformações ao modelo convencional. As transformações e o comportamento do modelo serão descritos em relação ao **display virtual não-planar**.

Os parâmetros intrínsecos das câmeras não foram considerados no processo de calibração devido à característica da aplicação, pois há uma imensa variedade de webcams, e para estas câmeras não há muita especificação técnica disponível. É interessante deixar a solução o mais independente possível das características da webcam.

A calibração das câmeras envolve os parâmetros extrínsecos de posição (distância das câmeras ao ZDL_m), orientação (ângulo de convergência das câmeras, α), e distância interaxial das câmeras, I_c . Os parâmetros intrínsecos de especificação de lente e distância focal das câmeras serão relaxados, e apenas o campo de visão (FOV) será utilizado para cálculo da dimensão do ZDL_m , pois esta informação de FOV é mais simples de obter para as webcams.

A Figura 5.27 mostra a seqüência de transformações geométricas do modelo, descritas a seguir:



Figura 5.27: 4 transformações geométricas

1. das coordenadas X_m, Y_m, Z_m do objeto no mundo real para as posições X_{s_e}, Y_{s_e} e X_{s_d}, Y_{s_d} nos dois sensores de imagens das câmeras.
2. das coordenadas dos sensores para o **display virtual não-planar**, X_c, Y_c, Z_c , em coordenadas do mundo; nesta transformação ocorre o mapeamento cilíndrico das imagens na superfície.

Os parâmetros t e s mapeiam as imagens do espaço de textura para a superfície S . As coordenadas paramétricas do cilindro são dadas por

$$x_c = r \cos t, y_c = s, z_c = r \sin t,$$

onde o eixo y representa a altura do cilindro e z é o eixo perpendicular à superfície.

3. das coordenadas 3D da superfície para as coordenadas X_t e Y_t da tela;
4. da tela para coordenadas X_o, Y_o, Z_o percebidas pelo observador.

Então, com uma transformação geométrica a mais, a de número 2, introduzida pelo modelo, gera-se o display virtual, que permitirá efetuar alterações na visualização das imagens, em função do espaço do observador, sem alterar a calibração inicial do espaço das câmeras. Este processo será descrito a seguir.

Comportamento do Display virtual não-planar

O processamento estereoscópico sob a abordagem do novo modelo pode ser considerado como um processo no qual um objeto 3D é primeiro filmado como duas imagens 2D, I_1 e I_2 , por duas câmeras reais. As imagens 2D filmadas, são combinadas e mapeadas em uma superfície 3D cilíndrica, que é o DVNP. Este display virtual é então filmado por uma outra câmera virtual,

e a imagem resultante desta câmera virtual é exibida na tela, em coordenadas 2D.

A Figura 5.28 ilustra o diagrama deste processo.

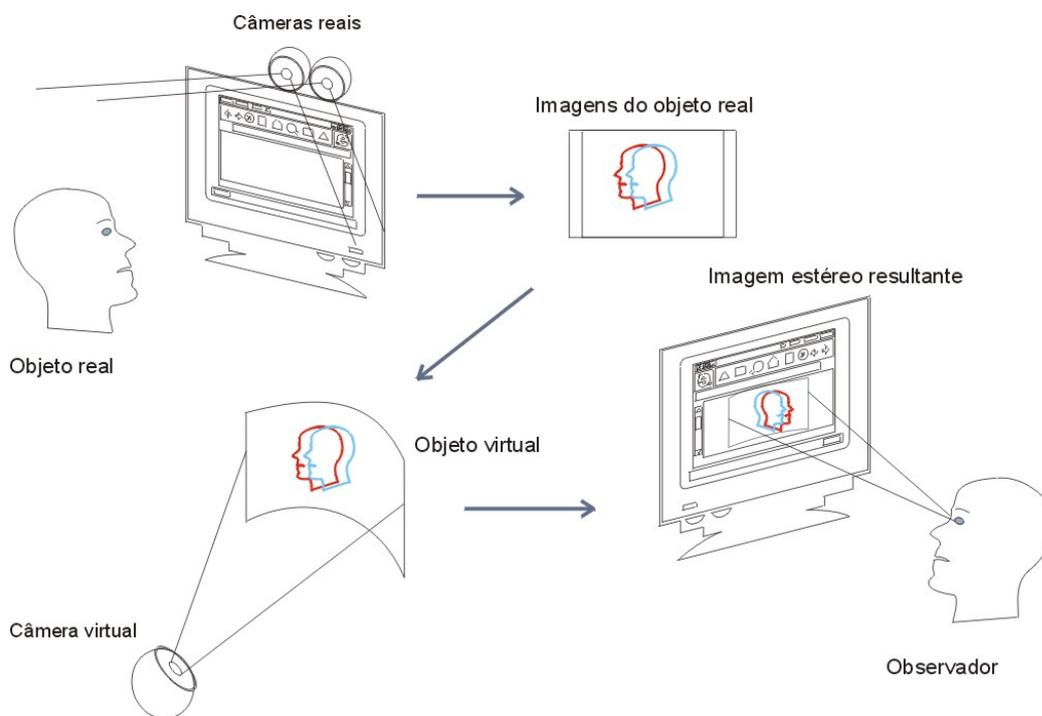


Figura 5.28: O display virtual não-planar - *DVNP*

Com este processo, os dados da câmera virtual diferem dos dados da câmera real, permitindo a manipulação do display virtual em função do observador. A distância da câmera virtual ao display virtual não-planar não é a mesma distância das câmeras reais ao ZDL_m e sim, será a mesma do observador à tela do computador; e o ângulo da câmera virtual, FOV , será o mesmo ângulo de visão do observador.

A Figura 5.29 mostra a relação da câmera virtual com o observador.

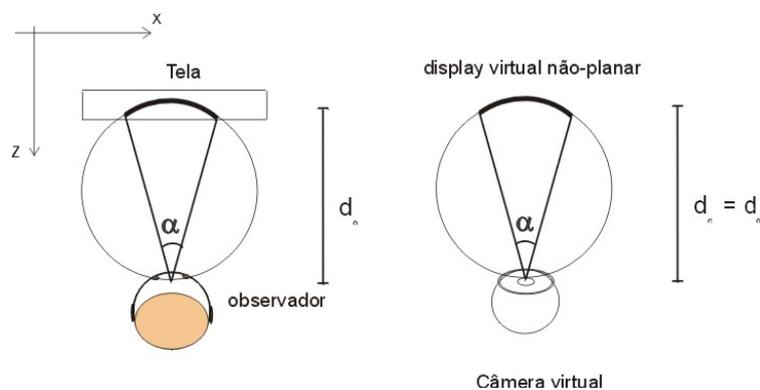


Figura 5.29: Câmera virtual e observador

Desta forma, o modelo pode trabalhar com a visualização das imagens estéreo função do observador, sem recalculas as variáveis do espaço das câmeras. Com isso, espera-se que a imagem exibida fique mais adaptada ao sistema visual natural do observador, por ter o formato do horopter adequado à sua distância de visualização.

Relacionamento convergência/acomodação no modelo

Como citado na literatura de estereoscopia, uma das principais causas de desconforto na visualização de imagens estéreo se dá pela ausência do casamento de convergência e acomodação no modelo.

A cada movimentação do observador, o display virtual não-planar simula o comportamento do horopter do observador, variando sua curvatura em função do ângulo de convergência e da distância do observador à tela de visualização. A câmera virtual representa o olho ciclópico do observador. Assim, o modelo incorpora uma simulação do relacionamento convergência/acomodação, não existente no modelo convencional, mostrado na Figura 5.30.

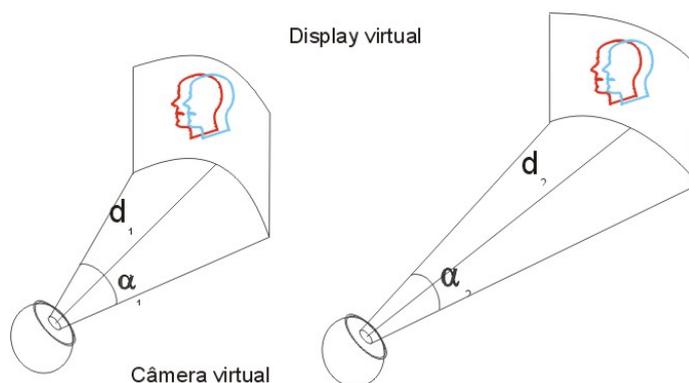


Figura 5.30: O olho ciclópico do observador

Mantendo fixo o comprimento do display virtual não-planar, $DVNP$, para

$$\alpha_1 > \alpha_2 \Rightarrow d_1 < d_2,$$

i.e., a curvatura do display tende a planificar quanto menor o ângulo de visão e maior a distância do observador à tela.

Conseqüentemente, o paralaxe nas imagens varia pelo comportamento natural da superfície: quando a curvatura aumenta, ocorre uma aproximação dos pontos vizinhos (definição de corda) e assim diminuindo a distância euclidiana entre eles, que representa o paralaxe destes pontos; quando a curvatura diminui, tendendo a planificar a superfície, ocorre um afastamento desses mesmos pontos, aumentando o paralaxe entre eles.

A redução no paralaxe das imagens, no modelo convencional, pode ser obtida diretamente pela aproximação da distância interaxial das câmeras. Da mesma forma, o aumento de paralaxe pode ser obtido pelo afastamento da distância interaxial das câmeras. A mesma interpretação pode ser feita em relação à manipulação direta das imagens pelas teclas de navegação, para gerar um aumento ou redução direta no paralaxe das imagens, como sugerido para o modelo convencional, com o objetivo de fazer com que o observador consiga visualizar melhor o estéreo.

Assim, o modelo proposto, está simulando o que seria uma mudança indireta nas variáveis do espaço das câmeras, através da variação da curvatura da superfície e fazendo uma variação no paralaxe das imagens em função do observador, de forma mais natural.

5.7

Compensando as distorções

Esta seção compara os modelos convencional e proposto em relação às distorções e às restrições impostas pelo modelo convencional. É importante também que cada distorção seja analisada pela causa e conseqüência.

Curvatura do plano de profundidade

Como foi apresentado no Capítulo 4, pelo trabalho de Woods et al. (62), a deformação de curvatura do plano de profundidade é gerada pela primeira transformação geométrica, por conseqüência da convergência das câmeras; isto se reflete no espaço do observador, fazendo com que o observador perceba as bordas das imagens mais distantes que os objetos no centro da imagem.

Neste caso, se as curvas de profundidade são cônicas, para as câmeras convergentes, a projeção da imagem sobre a superfície convexa do cilindro irá compensar esse efeito de distorção das bordas. Esta percepção se assemelha ao teste psicofísico do horopter *AFPP*, mostrado na Figura 5.17.

Keystone

De forma semelhante à curvatura do plano de profundidade, esta deformação é gerada na primeira transformação geométrica, pela convergência das câmeras, e reflete-se na terceira transformação, ou seja, na projeção da tela de visualização (segunda transformação no modelo convencional).

A Figura 5.31 mostra o círculo de Vieth-Müller e as imagens no espaço das câmeras e no espaço do observador. Pela operação de escala entre estes

espaços, o ângulo de convergência das câmeras, que é o ângulo entre as imagens, α , corresponde ao mesmo ângulo entre as imagens na tela, a menos da escala.

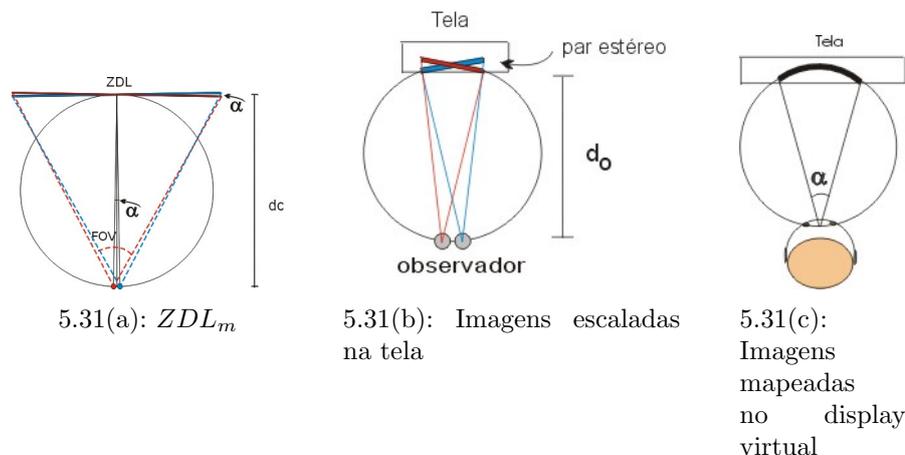


Figura 5.31: Escala dos ângulos de rotação

Com a analogia da câmera virtual como olho ciclópico do observador, tem-se as mesmas dimensões para o espaço da câmera virtual e para o espaço real do observador, e assim como foi mostrado, o ângulo de convergência da câmera virtual e o ângulo de visão do observador são iguais, e é o ângulo aplicado ao setor circular que define o cilindro, Figura 5.31.c.

Então, no novo modelo, quando as imagens são mapeadas na superfície, em coordenadas cilíndricas, estas são rotacionadas pelo ângulo de curvatura da superfície. Pela própria geometria do cilindro, já que a imagem como textura passa a ter o comportamento da superfície, esta distorção é compensada pelo mapeamento cilíndrico.

5.7.1 Análise esquemática das restrições geométricas

Com base nas restrições descritas na Seção 4.3.3, esta análise irá comparar esquematicamente o comportamento do modelo convencional e do modelo proposto, variando o posicionamento do observador em relação à tela de visualização, sob a ótica do horopter.

Observador se move no eixo-x

O movimento lateral do observador gera a distorção de cisalhamento, que é introduzida pelo ângulo de visão do observador, como mostra a Figura 5.32.a; mesmo que as imagens não sofram qualquer mudança, a distância d_o se torna d'_o . O alvo escuro na tela, mostrado na Figura 5.32, representa a imagem visualizada.

Esta distorção acontece no espaço do observador e pela quarta transformação geométrica - da tela para o espaço de percepção do observador. No entanto, ela não é explicitamente introduzida em outras transformações, mas como consequência desta distorção vem a exigência de que os dois centros de projeção (olhos) precisam permanecer centrados em relação a um ponto perpendicular ao centro da tela, caso contrário, gera cisalhamento.

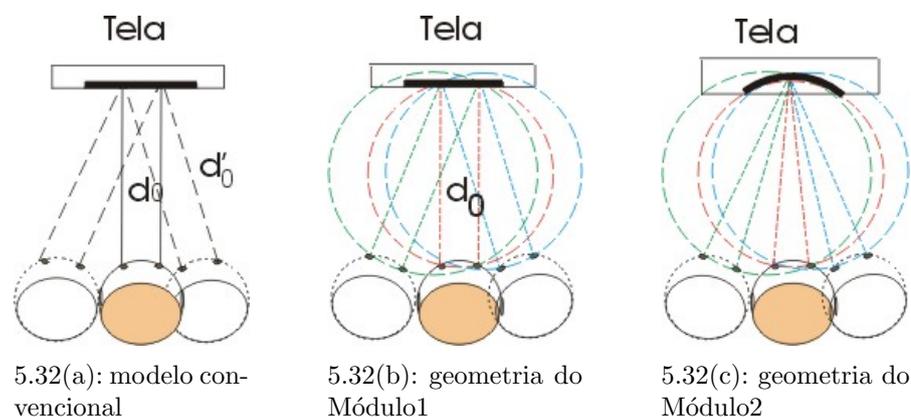


Figura 5.32: Observador movendo-se em x

Neste caso, fica a colocação de que esta distorção de cisalhamento pode estar originada na projeção plana, sobretudo para a visualização em curta distância. A Figura 5.32.b mostra como se comportaria a geometria do observador baseada no VMC, em relação às imagens exibidas em um plano (Módulo1). As linhas que representam a direção visual interceptando a imagem, mostram que apesar da maior parte das imagens caírem no horopter teórico, mesmo quando o usuário se move, ainda existem algumas áreas da imagem que ficam fora da região do horopter, por causa da geometria do plano da imagem.

A Figura 5.32.c representa o display virtual não-planar. Pela própria propriedade geométrica do círculo, o olhar desviado lateralmente ainda recai sobre o horopter e abrange a imagem, não gerando necessidade de reprojeção da mesma. ou de manter o usuário perpendicular à tela.

Observador se move no eixo-z

A restrição do modelo convencional diz que a coordenada z do ZDP precisa ser mudada, após o deslocamento em Z do observador. Ao observar a Figura 5.33.a, vê-se que este deslocamento não muda a perspectiva do observador em relação à tela, e portanto não deveria exigir recálculo dos parâmetros. Mas quando isso ocorre, sobretudo no modelo de câmeras paralelas, existe uma

perda de fusão das imagens e um estresse no olho pelo esforço de buscar a fusão das imagens.

Pela fórmula deduzida das projeções, que diz que ao observador afastar-se da tela, com o aumento de d_o , deve-se aumentar o paralaxe das imagens. Este tipo de reajuste recai então no ponto perceptual do observador, que ao deslocar-se para trás, o paralaxe das imagens não mais corresponde ao esperado pelo sistema visual. Representação que não está incorporada à geometria do modelo convencional.

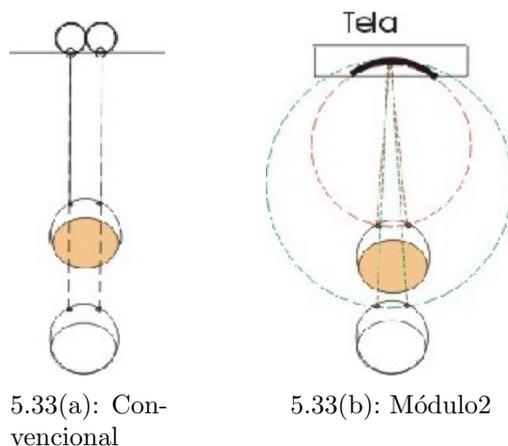


Figura 5.33: Observador movendo em z

No Módulo2 do novo modelo, quando o observador se move no eixo- z , apenas o diâmetro do horopter muda, Figura 5.33.b; se não houver ajuste na imagem, poderá ocorrer uma perda na percepção de profundidade, como acontece na visão normal, mas não ocorre uma perda de fusão das imagens. Ao ajustar a curvatura da superfície para a nova distância do observador, o paralaxe será, por consequência reajustado, aumentando, se d_o aumentar. Mas estas variações são pequenas, e mantêm o paralaxe nas imagens sempre dentro do intervalo gerado pelo cone de visão das câmeras.

O resultado das análises pode ser observado nas imagens resultantes da implementação do modelo, no próximo capítulo. Todas as distorções geradas ou percebidas pela transformação da tela para o olho do observador, requerem uma análise psicofísica precisa, por experimentos com usuários, para permitir a real validação da interpretação dos resultados do modelo. Isto será indicado como diretrizes para próximos trabalhos.

5.8 Diagrama de transmissão do vídeo estéreo

Como o enfoque da aplicação está voltado para as interações remotas, faz-se necessário elaborar a transmissão do vídeo estereoscópico como uma prova de conceito para a proposição da pesquisa.

A construção do modelo proposto com a criação do display virtual não-planar também favorece o processo de transmissão em relação ao modelo convencional. Como foi dito, as imagens são combinadas para serem mapeadas na superfície, resultando em uma imagem de textura.

O display virtual não-planar será gerado no cliente, mas antes de projetá-lo na tela, a superfície será mapeada com a imagem estéreo da pessoa remota. Isto simplifica o processo de transmissão, evitando o processamento de sincronismo entre as imagens esquerda e direita, no recebimento pelo cliente remoto. O sincronismo do servidor fica dedicado ao controle do envio e recebimento da imagem estéreo entre os clientes.

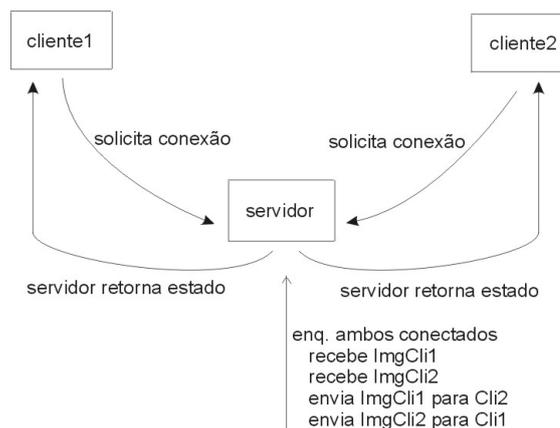


Figura 5.34: Diagrama de transmissão do vídeo estéreo

A transmissão será baseada na arquitetura cliente-servidor, e usará socket. A Figura 5.34 mostra o diagrama de conexão.

O primeiro usuário conecta-se ao servidor, e o servidor informa que este deverá aguardar a conexão do segundo usuário. Quando o segundo usuário conecta-se ao servidor, a transmissão é iniciada, permitindo a visualização estereoscópica da pessoa remota.