

PERCEPÇÃO VISUAL HUMANA

Guido Stolfi – 04/2008 – Televisão Digital

1. Estrutura do Olho Humano

À primeira vista, a estrutura básica do olho corresponde a um instrumento óptico simples: como acontece em uma câmara fotográfica, a imagem real de um objeto observado é projetada por uma lente convergente sobre uma superfície constituída de células sensíveis à luz. Fibras nervosas levariam então a imagem ao cérebro através de impulsos elétricos.

Esta descrição simplista, originada das descobertas de Kepler e Descartes relativas à fisiologia do globo ocular, mostra-se na realidade muito distante da real dimensão do fenômeno: a percepção visual é um processo de *reconstrução* da realidade exterior, realizado *pelo córtex cerebral*, a partir de informações fragmentadas captadas pelos olhos. O estudo deste processo de percepção deve então considerar em primeiro lugar as características e limitações do olho, na qualidade de um sensor óptico peculiar.

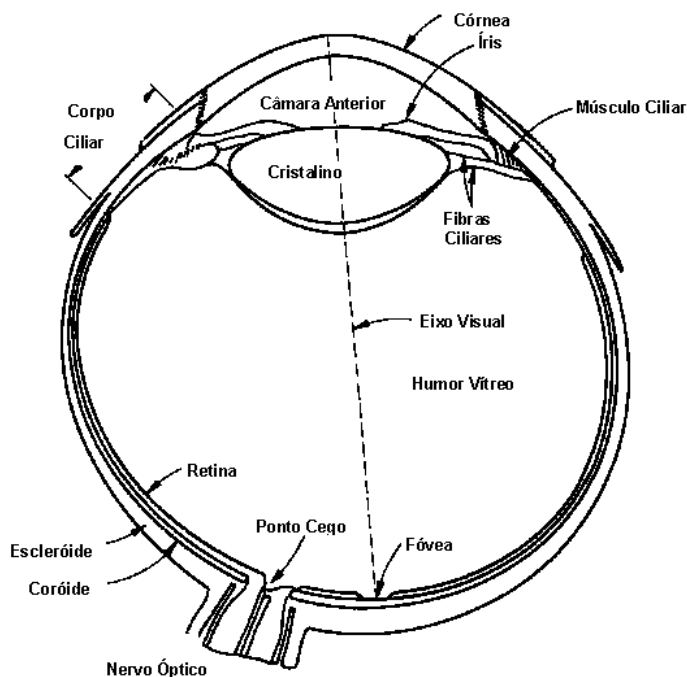


Fig. 1.1 – Estrutura do Olho

A figura ao lado apresenta uma seção horizontal do olho humano. A córnea, em conjunto com o cristalino, forma uma lente que projeta sobre a retina uma imagem real dos objetos visualizados.

Os músculos ciliares ajustam a distância focal equivalente (entre 14 e 17 mm, aproximadamente), através da deformação do cristalino. A íris serve de diafragma, ajustando a nitidez da imagem e a quantidade de luz que entra pela abertura da pupila.

O humor vítreo é uma solução aquosa de proteínas, que, além de prover nutrientes para o cristalino, atua como um filtro, absorvendo luz ultravioleta e infravermelha.

Os olhos são mantidos em constante movimento por meio de um conjunto de músculos (fig. 1.2). Algumas funções destes movimentos são:

- **Convergência binocular:** focalizar ambos os olhos sobre um mesmo objeto, compondo uma imagem estereoscópica (com percepção de profundidade);
- **Exploração do campo visual:** a observação de detalhes de um ambiente ou imagem é feita pela fóvea; a imagem projetada na retina é deslocada pelos movimentos da cabeça e dos globos oculares, de modo que os pontos de interesse da imagem são explorados pela fóvea;
- **Acompanhamento de objetos em movimento:** a percepção de movimento envolve tanto a detecção de deslocamentos da imagem projetada sobre a retina (movimento

retinal) quanto o acompanhamento ocular de um objeto, procurando manter a sua imagem estática em relação à retina (movimento ocular);

- **Transformação de variações espaciais em variações temporais:** Uma imagem permanentemente projetada sobre a retina, na mesma posição, desaparece da nossa percepção em pouco tempo. O nosso mecanismo de visão mantém os olhos em constante movimento, de forma a renovar continuamente os estímulos luminosos.

Através de pequenos movimentos exploratórios (denominados *movimentos sacádicos*), o processo visual transforma diferenças espaciais de luminância em variações temporais de iluminação retinal, que provocam o disparo de informações para o cérebro. A maioria das fibras do nervo óptico transmite impulsos disparados apenas por variações de luminosidade. O cérebro reconstrói a imagem observada a partir dessas informações, cancelando nesse processo todas as irregularidades estáticas do campo visual ocular, como por exemplo, as sombras projetadas pelos nervos e vasos sanguíneos na retina. Assim sendo, a composição da cena, ou *percepção visual* propriamente dita, dá-se efetivamente no cérebro, a partir de informações elementares fornecidas pelo olho.

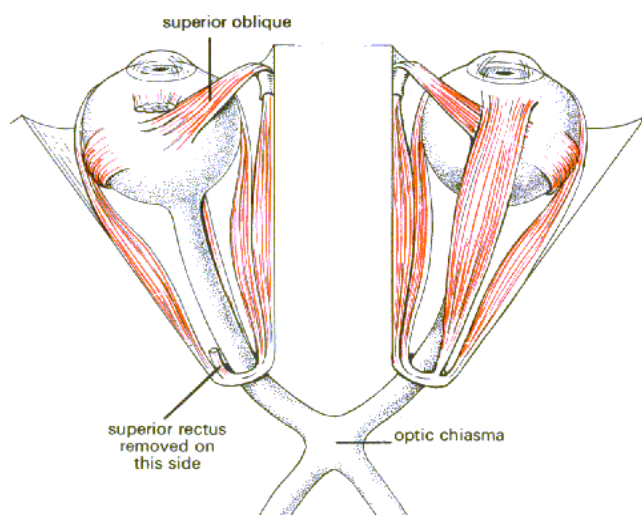


Fig. 1.2 - Musculatura do Globo Ocular

1.1 Retina

A retina é uma película constituída de células nervosas interligadas; pela sua estrutura e constituição, podemos dizer que é uma extensão do córtex cerebral. As células receptoras de luz, situadas na parte externa da retina, possuem pigmentos (*rodopsina* e *iodopsina*) que, ao absorverem a luz, sofrem uma alteração química que dispara impulsos nervosos.

Na retina temos duas classes de células sensíveis à luz:

- **Cones:** Células sensíveis à intensidade e à cor (compreendem três sub-tipos: *L*, *M* e *S*, sensíveis respectivamente ao amarelo, verde e azul); são responsáveis pela Visão Fotópica, ou visão de luz intensa. Dos cerca de 7.000.000 de cones existentes na retina, metade situam-se na Fóvea. Muitos cones são ligados individualmente a terminações nervosas; seu diâmetro varia de 1 a 5 μm .
- **Bastonetes:** Células sensíveis apenas à intensidade, responsáveis pela Visão Escotópica, ou visão em condições de pouca luz. Há cerca de 75 a 150 milhões de bastonetes na retina, mas são praticamente inexistentes na Fóvea. Por serem ligados em grupos às terminações nervosas, formam conjuntos mais sensíveis à luz do que os Cones, porém com menor capacidade de resolução de detalhes. A visão escotópica não proporciona sensação de cor.

A figura 1.3 mostra detalhes das terminações nervosas na retina e suas conexões com os cones e bastonetes. Os gânglios são interligados entre si lateralmente, através de dendritos. Estas ligações permitem que a retina efetue um pré-processamento da informação visual, afunilando os estímulos provenientes das cerca de 150.000.000 de células sensoras para aproximadamente 1.000.000 de fibras contidas no nervo óptico.

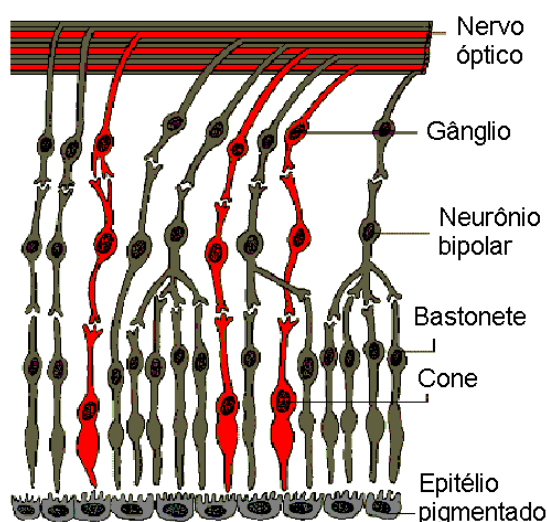


Fig. 1.3 - Estrutura da Retina

A quantidade de impulsos nervosos disparados pelas células sensoras é aproximadamente proporcional ao logaritmo do iluminamento retinal. Em média, são necessários no mínimo 5 a 10 fótons para provocar um impulso nervoso, em condições de baixa iluminação.

O processamento efetuado pelos gânglios remove essa proporcionalidade: a maioria das fibras do nervo óptico transporta informações de *variação relativa* de luminosidade. Algumas fibras disparam impulsos ao detectarem aumento de luz; outras sinalizam diminuição de luz; outras ainda sinalizam as duas situações. Poucas fibras transportam impulsos proporcionais à luminosidade estática.

1.2 Fóvea

A **Fóvea** é uma região especializada da retina, com área aproximada de 1,5 mm², subtendendo um ângulo de visão transverso de cerca de 2 graus. Dentro desta região temos a Visão Central, predominantemente Fotópica, com resolução de detalhes de luminosidade de cerca de 1 a 2 minutos de grau (correspondendo a detalhes de 1 a 2 mm a uma distância de 3m).

A resolução de detalhes de cor (ou crominância), no entanto, é de apenas de 5 a 10 minutos de grau. Fora da fóvea, temos a região de Visão Periférica, que se caracteriza por menor resolução espacial, pouca percepção de cor e maior sensibilidade à luz e movimento.

Na Figura 1.4 temos a distribuição aproximada das células (Cones e Bastonetes) em função da distância angular, medida em relação à fóvea. O Ponto Cego corresponde à região de entrada do nervo óptico no olho, onde não há células receptoras.



Visualização do Ponto Cego

O ponto cego ocupa uma região de até 4 graus de extensão no nosso campo visual. O fato de que não percebemos sua existência, em situações normais, mostra que o cérebro constrói a percepção do campo visual através do processamento de informações *ativas* (impulsos nervosos) enviadas pela retina. A *ausência de informação* não significa *ausência de luz* na imagem.

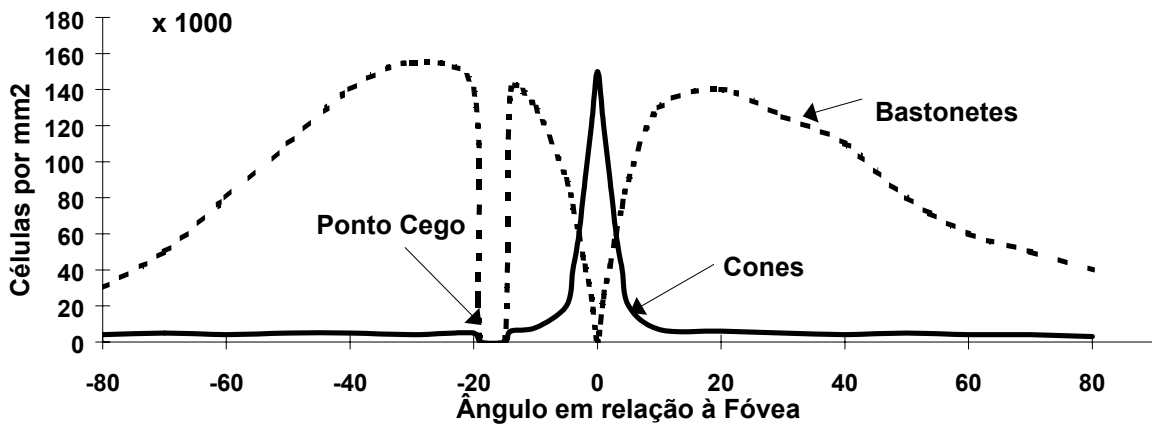


Fig. 1.4 - Distribuição dos Cones e Bastonetes

2. Adaptação à Luminosidade e Faixa Dinâmica

O olho humano pode adaptar-se a uma faixa extensa de níveis de luminosidade, desde sol a pino até a luz das estrelas (Vide figura 2.1). Desde o limite inferior da visão escotópica até o limite de ofuscamento, a nossa percepção tolera variações de luminância da ordem de 10^{10} a 10^{11} . O controle da área de entrada de luz na pupila pela atuação da íris (variando de 3 a 50 mm^2 no máximo) não é suficiente para explicar a extensão desta faixa dinâmica.

Na realidade, quando há iluminação suficiente, a pupila contrai-se de forma a utilizar apenas a área central do sistema óptico (córnea + cristalino), onde as aberrações são menores; abrindo-se totalmente apenas em condições de baixa luminosidade. O diâmetro mínimo de abertura limita a difração da luz num grau consistente com a resolução limite na fóvea.

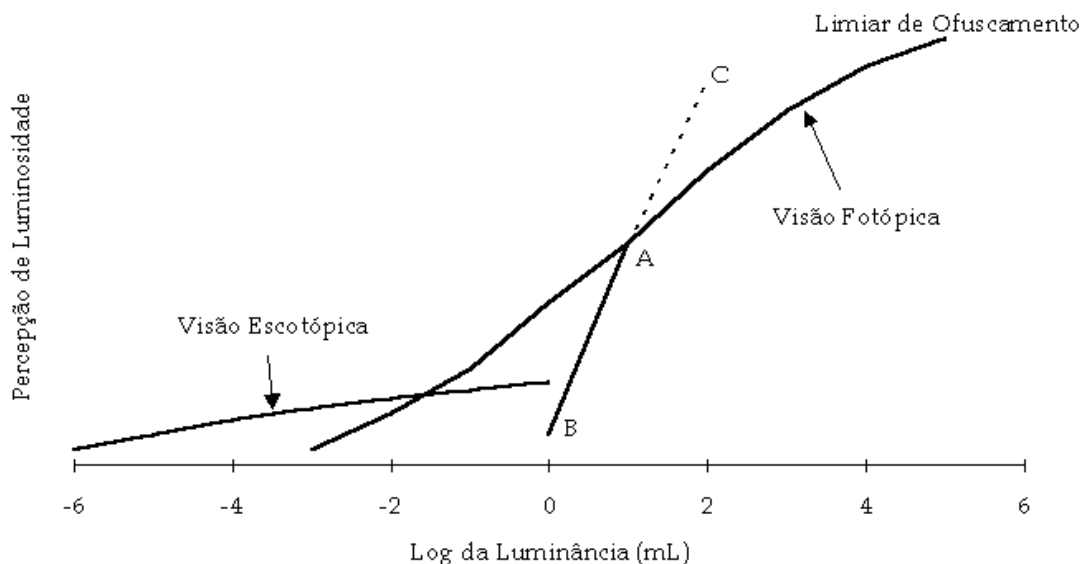


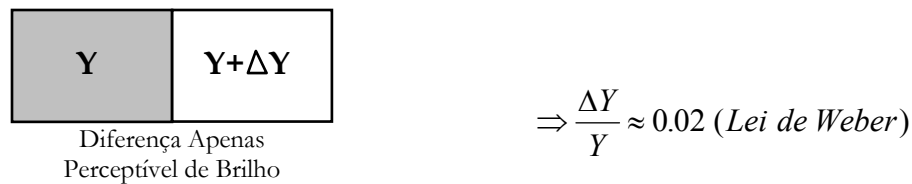
Fig. 2.1 - Faixa Dinâmica da Percepção Subjetiva da Luminosidade

A maior parte da capacidade de adaptação da visão decorre da mudança de sensibilidade das próprias células, devido à dissolução e recomposição do pigmento conforme a luminosidade. Este processo de acomodação não é instantâneo, e podemos dizer que, uma vez que o olho esteja adaptado a um determinado nível de intensidade média (por exemplo o ponto A na fig. 2.1), a faixa dinâmica local reduz-se para algo em torno de 100 a 200 vezes.

Isto significa que um objeto com luminância cerca de 20 vezes menor que **A** é percebido como preto (ponto **B** na fig. 2.1), enquanto que um objeto 5 a 10 vezes mais luminoso que **A** chega a provocar ofuscamento (ponto **C**). Este fenômeno permite que sistemas artificiais de reprodução de imagem (fotografia, TV, cinema etc.) reproduzam satisfatoriamente as sensações subjetivas de claro / escuro, mesmo com relações de contraste de 1 : 100 apenas.

3. Percepção de Luminosidade

A capacidade da visão humana de distinguir diferenças de luminosidade pode ser aproximada pela Lei de Weber, que se baseia em um modelo logarítmico da percepção subjetiva de estímulos. Com relação a dois estímulos de luminância Y e $Y+\Delta Y$, para os quais a diferença de brilho é apenas perceptível, temos:



A figura 3.1 mostra um gráfico da amplitude relativa de luminância correspondente a uma diferença apenas perceptível de brilho (denominada Fração de Weber) em função do nível médio de luminância Y . Dentro de uma faixa apreciável da visão fotópica, esta fração vale aproximadamente 0,02.

Por outro lado, para um sistema de reprodução de imagens vale a situação restrita da curva **B A C** da figura 2.1, uma vez que não é objetivo de nenhum processo reproduzir a luminosidade absoluta de uma cena, e sim apenas a faixa dinâmica correspondente em torno do ponto de acomodação. Para essa situação, a resposta subjetiva de brilho \times luminância (denominada "Ligness"), determinada experimentalmente pela CIE (*Comission Internationale de l'Eclairage*), pode ser aproximada pela expressão abaixo (vide gráfico na fig. 3.2):

$$Brilho \text{ Aparente} = \begin{cases} 116(Y / Y_n)^{\frac{1}{3}} - 16 & \text{se } Y / Y_n > 0.008856 \\ 903.3(Y / Y_n) & \text{caso contrario} \end{cases}$$

Onde Y_n = Referência de Branco (Brilho máximo)

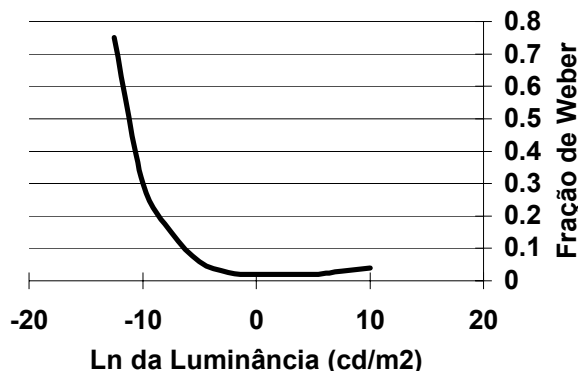


Fig. 3.1 - Sensibilidade para variações de luminância

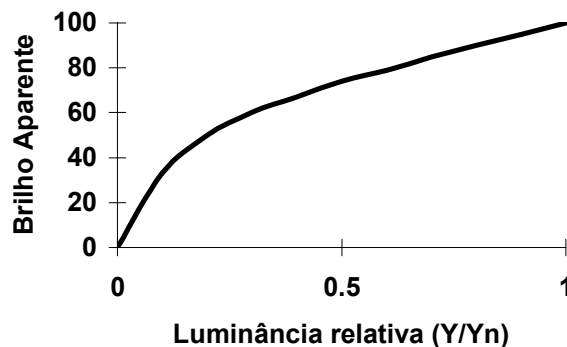


Fig. 3.2 - Percepção de Brilho em torno de um ponto de acomodação

3.1 Correção Gama

Em um sistema de TV é interessante transmitir a informação na forma de um sinal proporcional ao brilho aparente, e não à luminância real, para que o efeito de um eventual ruído aditivo seja percebido com intensidade uniforme para todos os níveis de luminosidade.

Por outro lado, no receptor, temos que considerar ainda a relação não-linear entre a tensão de controle e a luminosidade resultante na tela de um cinescópio (TRC- Tubo de Raios Catódicos) típico. Esta relação é da forma $Y = k \cdot V^\gamma$ onde γ geralmente assume valor igual a 2,2. Considerando a curva da fig. 3.2 aplicada sobre a curva de luminosidade do TRC, obtemos uma relação quase linear entre tensão aplicada e percepção de brilho aparente (ver fig. 3.3).

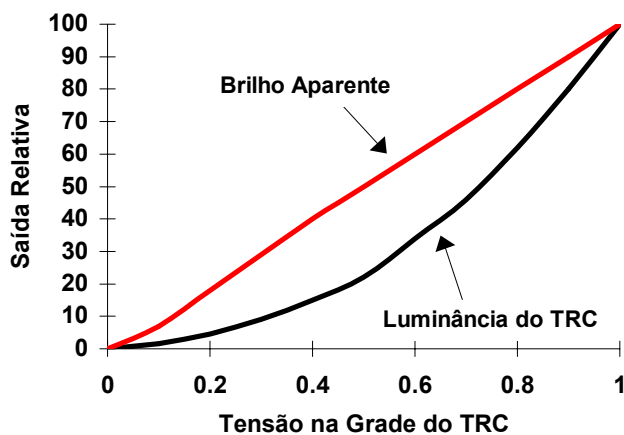


Fig. 3.3 - Brilho aparente e luminância de um Tubo de Raios Catódicos

Considerando essa característica quase complementar, adota-se então uma correção da forma $V = kB_0^{\frac{1}{\gamma}}$ onde V é o sinal a ser transmitido e B_0 é o sinal de entrada, proporcional à luminância da imagem (como fornecido por uma câmara).

Esta conversão (denominada *Correção Gama*) compensa aproximadamente a não-linearidade conjunta do cinescópio e da percepção visual, fazendo com que o sinal transmitido seja praticamente proporcional à *sensação subjetiva de luminosidade*, garantindo no receptor uma sensação de brilho B aproximadamente igual ao brilho original B_0 da imagem.

A característica complementar entre a luminância do TRC e a percepção subjetiva de brilho é uma coincidência, restrita a esta categoria de dispositivos de exibição de imagem. Não se aplica, por exemplo, para visores de cristal líquido, LEDs e outras famílias de dispositivos de visualização. Nestes casos, é necessário simular esta característica artificialmente para manter a proporcionalidade entre o sinal transmitido e a luminosidade aparente.

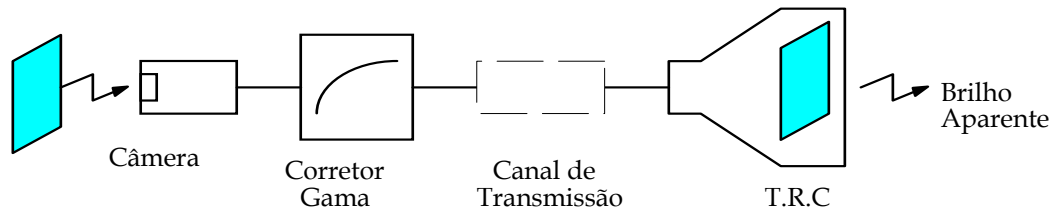


Fig. 3.4 - Correção Gama

4. Resolução Visual

A capacidade de resolução de detalhes espaciais da visão humana pode ser medida por experimentos que avaliam a percepção de contraste (claro/escuro) na visualização de padrões de linhas alternadas. Um destes experimentos é a Carta de Campbell (fig. 4.3). Conforme a largura de cada par de linhas e a distância ao observador, estes padrões representam estímulos com uma determinada *freqüência espacial* (que pode ser expressa em *ciclos por grau*).

A percepção relativa de contraste em função da freqüência espacial caracteriza a “resposta em freqüência” da visão; a curva correspondente para um observador padrão pode ser vista na figura 4.1. Podemos constatar que a resolução de crominância (distinção entre tonalidades de cor com mesma luminosidade aparente) é muito menor que a resolução de luminância (contraste claro/escuro). O "pico" na resposta em freqüência da luminância, em torno de 3 a 4 ciclos por grau, evidencia uma maior sensibilidade para *variações locais* de luminosidade.

Dois mecanismos de processamento neuronal, realizados pela retina, explicam certas características da resposta em freqüência da visão:

- *inibição lateral*: uma célula receptora que estiver fortemente iluminada provoca a inibição da resposta das células vizinhas, através das interconexões horizontais entre os gânglios. Este efeito corresponde a um filtro passa-altas, que amplifica a percepção de contornos nas imagens.
- *campos receptivos*: as mesmas interconexões podem atuar, em níveis de luminosidade reduzidos, de forma a agrupar os estímulos provenientes de vários receptores, aumentando a sensibilidade e a relação sinal/ruído, porém reduzindo a resolução (equivalente a um filtro passa-baixas).

O efeito da inibição lateral pode ser relacionado com o fenômeno das "Bandas de Mach" (figura 4.2), pelo qual a visão humana parece exagerar a amplitude de transições de luminosidade entre regiões homogêneas adjacentes (equivalente a um *overshoot* no brilho aparente).

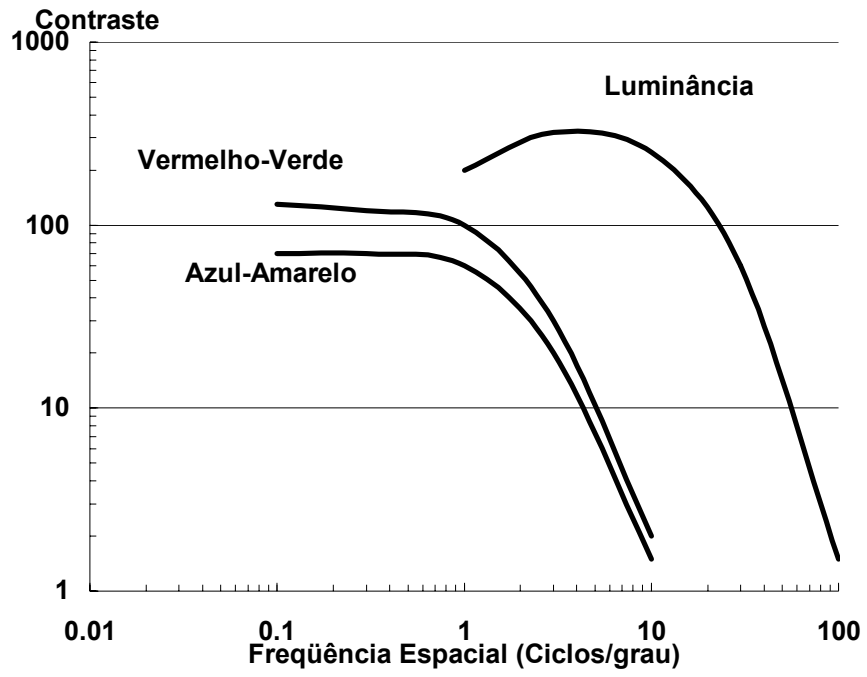


Fig. 4.1 - Resolução espacial da visão humana

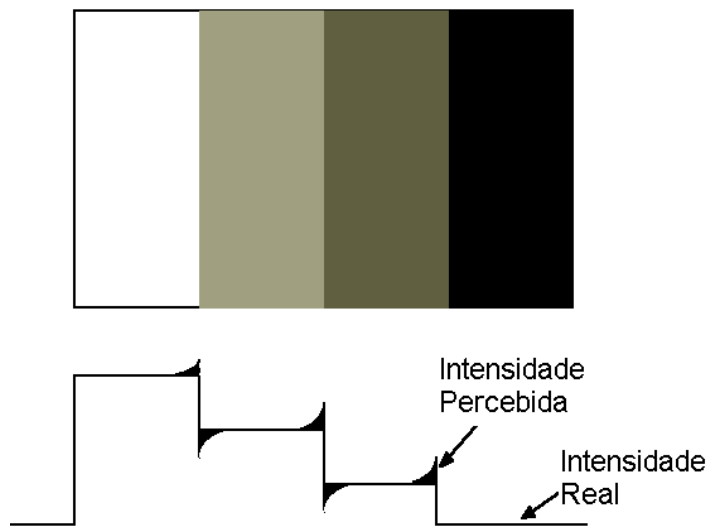


Fig. 4.2 - Bandas de Mach

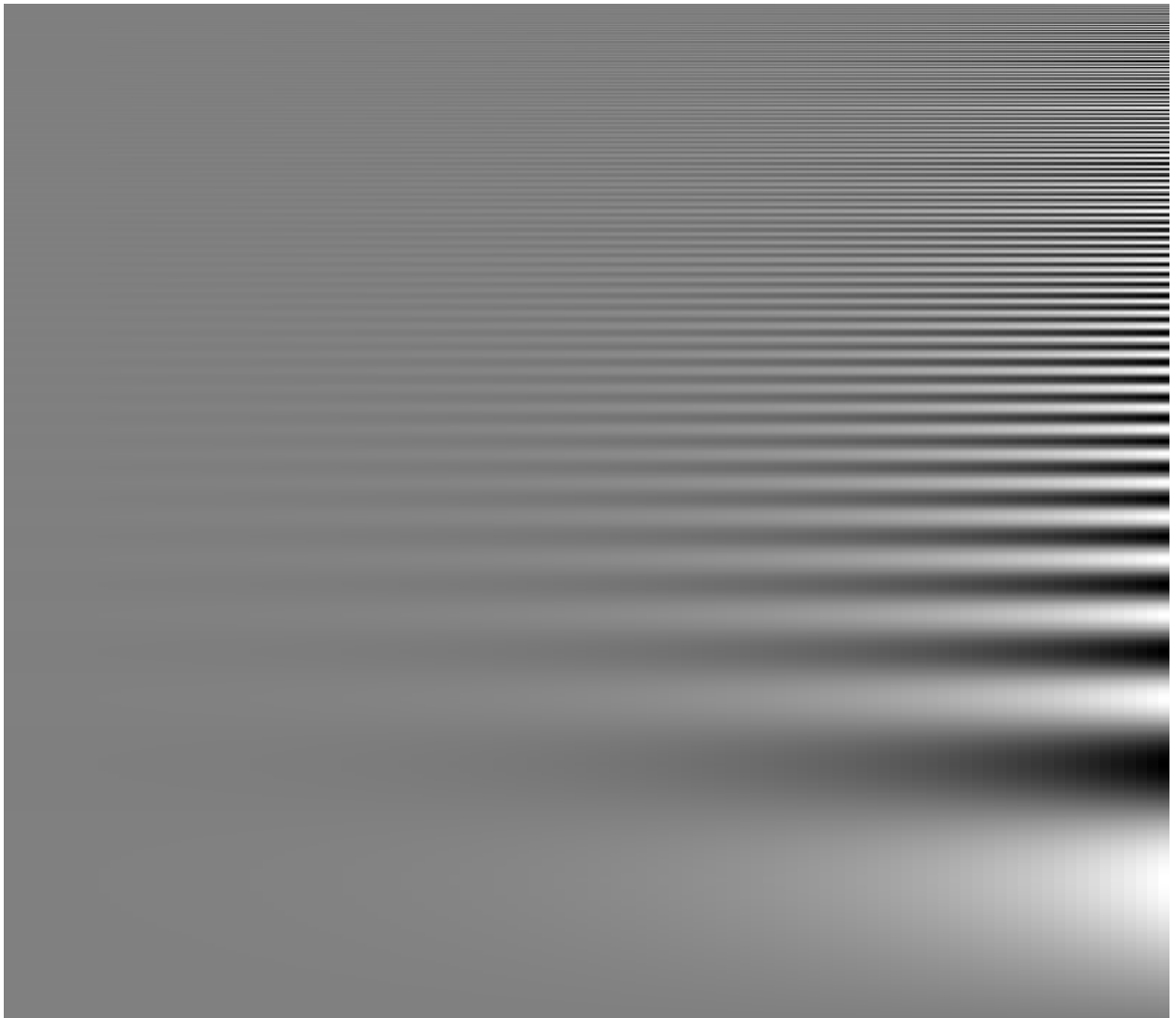


Fig. 4.3 – Carta de Campbell

5. Resolução Temporal e Cintilação

Outro aspecto importante da visão, relevante em Televisão e Cinema, refere-se à *Frequência Crítica de Cintilação*, definida como aquela acima da qual não é mais perceptível a presença de modulação na luminância de uma imagem. Dependendo das condições de iluminação, do campo da visão (periférica ou central), do ângulo aparente do objeto sujeito à cintilação, e do próprio indivíduo, esta frequência varia de 15 a 100 Hz aproximadamente. Na figura 5.1 temos a frequência crítica apresentada em função da luminância média, para objetos pequenos (1 grau) e grandes (20 graus), para visão central e periférica. Na figura 5.2, vemos a sensibilidade à cintilação de texturas em função da dimensão dos detalhes (frequência espacial).

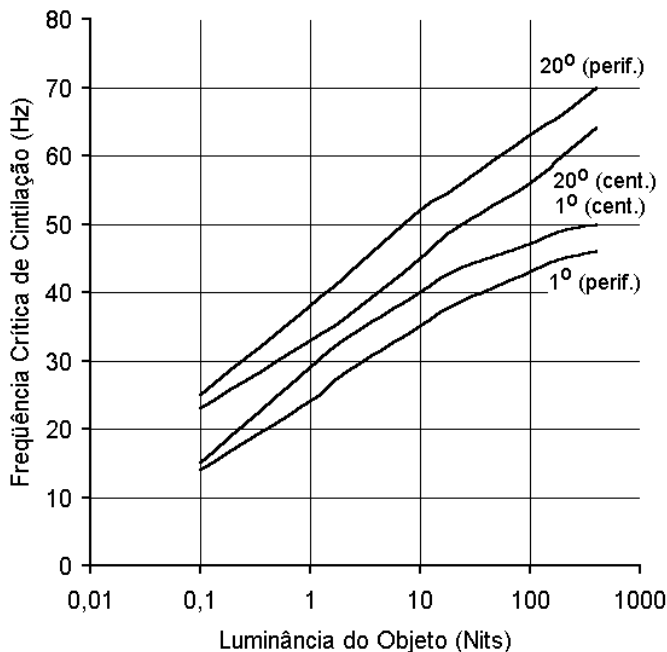


Fig. 5.1 - Percepção de Cintilação

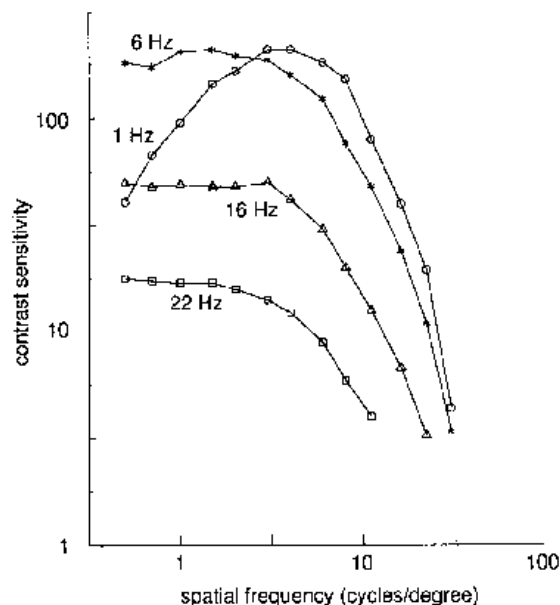


Fig. 5.2 - Sensibilidade à Cintilação de Detalhes

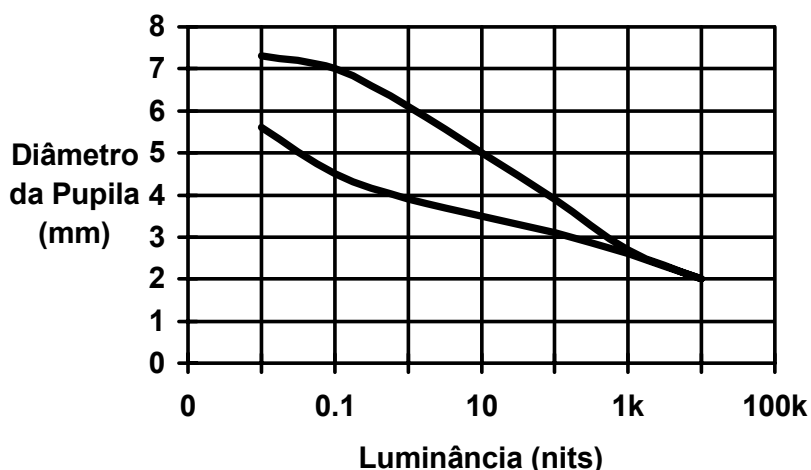


Fig. 5.3 - Diâmetro da Pupila em função da luminância ambiente

A tabela 5.1 apresenta a relação entre a luminância da tela de monitores de vídeo e as correspondentes frequências críticas de cintilação, válidas para 95% da população (ISO/TC159/1987), para campos de visão de 10 e 70 graus. Para os níveis de luminância típicos de telas de TV, a frequência crítica chega a ser da ordem de 70 Hz. As variações nos limites devem-se ao fato de que o diâmetro da pupila varia de um indivíduo para outro, especialmente para níveis de luminância abaixo de 1000 nits (vide figura 5.3).

A frequência crítica de cintilação está relacionada com o tempo necessário para que o cérebro avalie a quantidade média de impulsos nervosos disparados pelas células da retina, de forma a medir a luminosidade da imagem observada.

Luminância da Tela (Nits)	Área da Pupila (mm ²)	Iluminamento Retinal (Trolands)	Frequência Crítica p/ 10° (Hz)	Frequência Crítica p/ 70° (Hz)
10	10.75 ~ 19.63	215 ~ 392	48.4 ~ 54.4	60.6 ~ 66.9
30	9.08 ~ 15.21	544 ~ 912	57.6 ~ 62.8	70.2 ~ 75.5
100	8.04 ~ 11.34	1608 ~ 2268	68.9 ~ 71.9	81.8 ~ 84.9
300	7.07 ~ 9.62	4242 ~ 5772	78.2 ~ 81.3	91.4 ~ 94.6
1000	6.16	12320	88.9	102.4
3000	4.52	27120	96.7	110.5
10000	3.14	62800	105.1	119.2

Tabela 5.1 - Relação entre Luminância da Tela de Monitores de Vídeo e Frequências Críticas de Cintilação, para 95% da População (ISO/TC159/1987) - Campos de Visão de 10 e 70 graus

6. Percepção de Movimento

Dois mecanismos independentes são responsáveis pela percepção de movimento:

- **Movimento retinal:** um objeto em movimento em relação ao ponto de visão projeta uma imagem que se desloca sobre a retina. Os receptores sensíveis a variações de luminosidade enviam sinais ao cérebro, que processa estas informações, extraindo a direção, tamanho e contornos do objeto a partir da correlação desses estímulos; porém proporcionando baixa resolução de detalhes. Este processo é particularmente importante na visão periférica.
- **Movimento ocular :** através do movimento conjunto da cabeça e do globo ocular, sob comandos do cérebro, a visão percebe características de movimento de um objeto sendo acompanhado pela fóvea. Ou seja, sua imagem, projetada na retina, permanece estacionária, excetuando-se os movimentos sacádicos superpostos ao movimento médio do objeto. Nestas condições, as características de resolução espacial da visão são mantidas, apesar do movimento.

7. Percepção de Distância

Vários mecanismos, cada um ativo dentro de uma certa faixa de alcance, contribuem para que a visão humana determine distâncias na visualização de objetos:

- Focalização do cristalino: fornece informação intensa para curtas distâncias (menos de 1 metro);
- Convergência binocular: informação fortíssima para curtas distâncias, com resolução da ordem de milímetros dentro do alcance físico; decresce até distâncias médias (~100 m);
- Paralaxe de movimento: depende da velocidade e extensão do movimento; em condições de visualização normal, com o observador parado, é intensa para distâncias curtas (alguns metros)
- Fator de escala de objetos conhecidos: fornece estimativa forte para qualquer distância (centímetros a quilômetros), porém depende de aprendizado;
- Texturas e contraste: variações em texturas, saturação de cores e contraste da imagem fornecem estimativas razoáveis para grandes distâncias (quilômetros).

A *consistência* das informações proporcionadas por todos esses mecanismos, simultaneamente, determina a intensidade da ilusão de realidade que pode ser produzida por um sistema de visualização. Por exemplo, objetos distantes representados em uma fotografia, examinada a

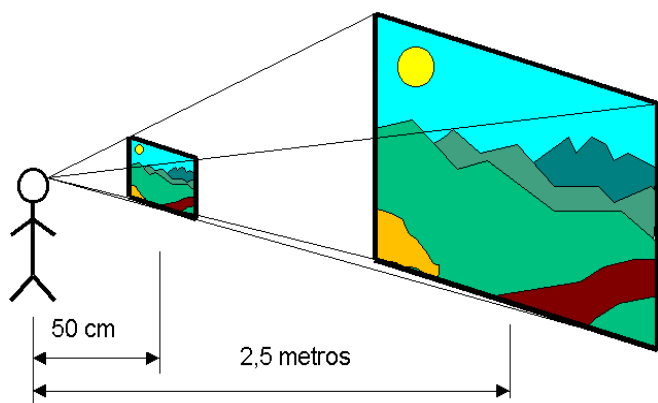


Fig. 7.1 - Diferenças de Percepção de Distâncias

curta distância, não são aceitos como “reais”, uma vez que a convergência, focalização e paralaxe identificam claramente os objetos como sendo próximos, enquanto que o fator de escala e as texturas representadas correspondem a objetos distantes. Essas informações conflitantes levam o cérebro a rejeitar a percepção da existência de objetos “distantes” na fotografia, destruindo a ilusão de realidade. A mesma foto, porém, ampliada e observada a uma distância maior, pode dar melhor impressão de realidade, mesmo que os objetos subentendam o mesmo ângulo visual (figura 7.1).

Percebe-se que é necessária uma tela de grandes dimensões para proporcionar ilusão convincente de profundidade para o observador.

8. Ilusões Ópticas

Algumas imagens são interpretadas de forma paradoxal pela nossa visão; são as chamadas ilusões ópticas. Mais do que simples curiosidades, elas nos dão pistas sobre o funcionamento dos processos de percepção visual que ocorrem no cérebro.

A ilusão de Muller-Lyers (fig. 8.1) pode ser explicada pelo fato de que a nossa visão interpreta a figura da esquerda como a aresta interior formada por duas paredes, chão e teto; enquanto a figura da direita seria a aresta frontal de um paralelepípedo. A nossa visão procura corrigir a dimensão percebida, considerando que a proximidade afeta a dimensão aparente de um objeto ($ab = cd$ na figura 8.3). O mesmo efeito ocorre na ilusão das Mesas de Shepard (fig. 8.2).

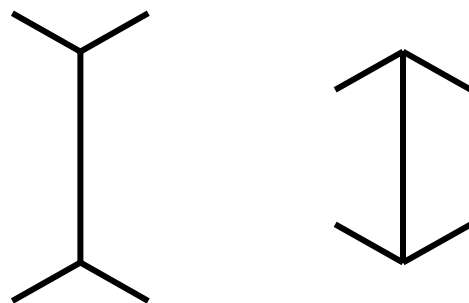


Fig. 8.1 – Ilusão de Muller-Lyers
(as duas linhas verticais são iguais)

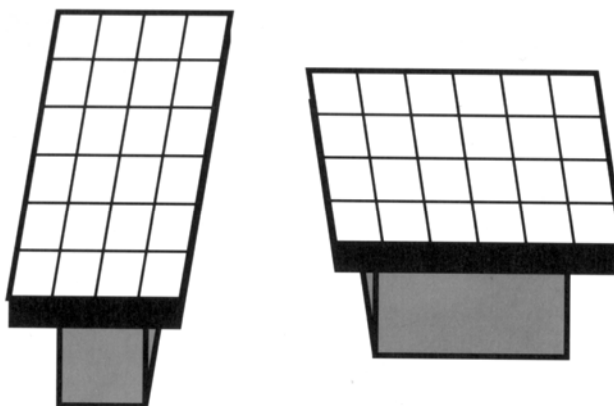


Fig. 8.2 – Mesas de Shepard (os dois tampos são iguais)

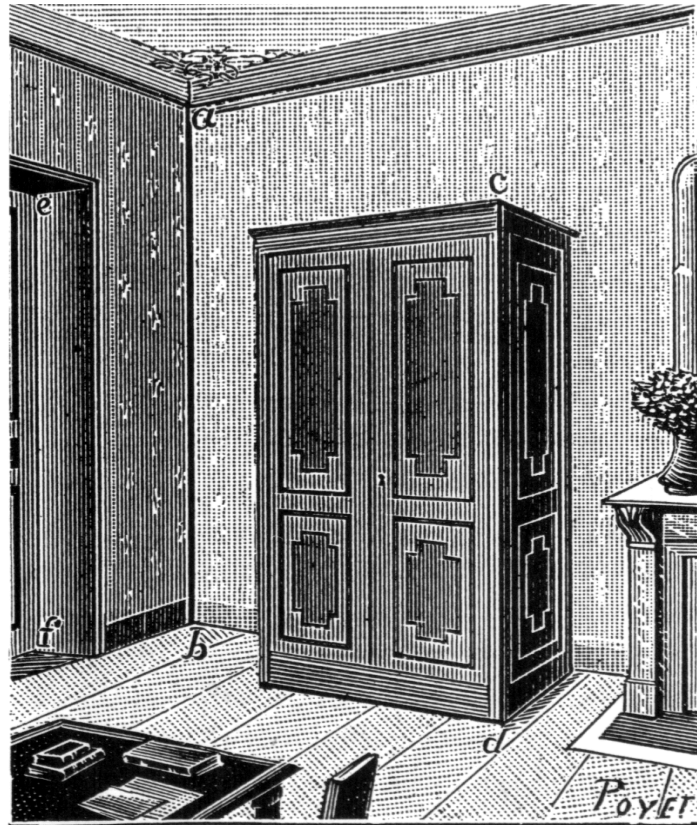


Fig. 8.3 – Justificativa da Ilusão de Muller-Lyers

9. Bibliografia e Referências

Colour Television - S. V. Novakovsy - MIR, 1975

Digital Image Processing - Rafael C. Gonzales, Richard E. Woods - Addison Wesley, 1992

Eye and Brain: The Psychology of Seeing - Richard L. Gregory - Princeton University Press, 1997 (Fig. 1.2)

MPEG Video Compression Standard - John Mitchell et al. - Chapman & Hall, 1996 (Fig. 5.1, 5.2)

Television Engineering Handbook - K. Blair Benson, ed. - Mc Graw-Hill, 1985

Understanding Digital TV - Brian Evans - IEEE Press, 1995

Video Engineering - Andrew F. Inglis, Arch C. Luther - Mc Graw-Hill, 1996

Visual Intelligence - Donald D. Hoffman - W. W. Norton, 1998

Standard Handbook for Electrical Engineers - Donald G. Fink, H. Wayne Beaty, ed. - Mc Graw-Hill, 1993

The Science of Illusions - Jacques Ninio - Cornell University Press, 2001 (Fig. 8.2, 8.3)