



INSTITUTO SUPERIOR TÉCNICO
Universidade Técnica de Lisboa

COR e LUZ

por

João Manuel Brisson Lopes

Departamento de Engenharia Informática

texto elaborado para a disciplina de

Computação Gráfica

Licenciatura em Engenharia Informática e de Computadores

publicado em Maio de 2003

reeditado em Dezembro de 2008, Abril 2013

Este texto, elaborado no contexto da disciplina de Computação Gráfica da Licenciatura em Engenharia Informática e de Computadores do Instituto Superior Técnico, foi originalmente concebido para fazer parte de um conjunto de textos sobre Computação Gráfica, apresentando-se agora como um texto independente.

Contacto do autor: brisson@ist.utl.pt

© 2000, 2001, 2003, 2008, 2013 J. M. Brisson Lopes & IST

Cor e Luz

1 Introdução

A cor está presente em tudo o que observamos e é assim um elemento essencial em Computação Gráfica onde desempenha funções múltiplas na visualização de imagens e cenas. Em primeiro lugar, a cor confere realismo às imagens e cenas visualizadas. Quanto maior for a fidelidade da reprodução da cor maior será o realismo e a naturalidade dos resultados observados. A cor transmite igualmente informação sobre as condições de iluminação e a forma dos objectos visualizados, acentuando os seus contornos. Isto significa que, para obter uma reprodução fiel da cor, deverão ser empregues modelos de cor correctos e precisos.

A cor desempenha também um papel importante na visualização de propriedades associadas aos objectos visualizados. Os diferentes níveis de elevação dos mapas orográficos são assinalados por cores diferentes, em que as cores de tom verde assinalam áreas de baixa altitude e as cores de tom acastanhado as de maior altitude. A variação da cor nestas cartas transmite ainda a noção de declives mais ou menos acentuados. A cor permite visualizar o não visível adicionando informação aos objectos visualizados, como é o caso dos mapas de temperatura em que as cores avermelhadas assinalam as zonas mais quentes e as cores com tons azuis as áreas mais frias. A cor desempenha também a função de destacar e chamar a atenção para a informação e dar-lhe significado: o vermelho significa perigo, o amarelo cuidado.

Este capítulo começa abordando a cor de forma intuitiva, fazendo apelo aos conceitos comuns na linguagem do dia-a-dia, para concluir que esta não é precisa e é necessário maior rigor no tratamento da cor. Para isso, abordam-se a natureza da luz (e da cor) e a forma como ela é percebida pelo subsistema humano de visão. O conhecimento adquirido permite então modelar a cor através de modelos exactos (CIE) ou adaptados aos dispositivos de saída gráfica (RGB, CMY, etc.). Estes modelos orientarão a especificação da cor e a sua reprodução. Analisam-se em seguida os problemas derivados das limitações e características dos dispositivos de saída gráfica, apresentando soluções para a sua resolução. O emprego e escolha de cores, quer sob o ponto de vista estético, quer sob o ponto de vista de reforço à visualização é tratado de seguida.

2 Caracterização Perceptual da Cor

Na base de qualquer teoria da cor estão os fenómenos associados à cor e à forma como as cores são percebidas e distinguidas. A linguagem corrente emprega termos como cor, cor pura, cor saturada, luminosidade e brilho de uma cor com sentidos bastante latos. O termo *cor*¹ pretende distinguir entre azul, verde, amarelo, etc. A *saturação* da cor, por vezes denominada pureza da cor ou simplesmente saturação, indica o afastamento da cor observada de um cinzento com a mesma intensidade luminosa. Uma cor vermelha ou azul puras são cores altamente saturadas, enquanto um rosa e as cores denominadas de pasteis são cores pouco saturadas.

A *luminosidade* é a intensidade da luz reflectida pela superfície dos objectos, enquanto que, para marcar a diferença, o *brilho* é a quantidade de luz emitida pelas superfícies de objectos luminosos como o Sol ou uma lâmpada. Esta diferenciação não é exacta na língua portuguesa pois, por exemplo, é comum falar do brilho de um espelho que, na verdade, é um objecto reflector perfeito.

A modelação e reprodução da cor obrigam a uma maior correcção do que aquela permitida pela linguagem corrente. Esta maior correcção pode ser obtida de forma empírica como é o caso do modelo de cor de Munsell. Este modelo estabelece padrões de cor. Tais padrões são empregues na determinação dos parâmetros das cores que o modelo define por comparação entre amostras das cores a determinar com os padrões. O modelo de Munsell estabelece como parâmetros a cor (hue), o valor (a luminosidade) e a croma (saturação). O modelo de Munsell é intrinsecamente subjectivo dado que depende do julgamento pessoal de cada observador ajuizar se duas cores são ou não idênticas. Este processo é bastante impreciso pois depende do tamanho relativo das amostras comparadas, da cor da luz circundante, da iluminação geral do ambiente em que a medida é realizada e ainda da forma como a iluminação é efectuada².

A noção de cor em artes visuais encontra-se, por razões históricas, intimamente ligada aos pigmentos das tintas utilizadas. Em artes visuais empregam-se termos como *cor pura*, *pastel*, *sombreado* e *tom*. O *sombreado* é o resultado da adição de um pigmento negro a um pigmento de cor pura em que a pigmento negro retira luminosidade ao pigmento puro. Uma cor pastel resulta da adição de um pigmento branco a uma cor, diminuindo a sua saturação. Um tom resulta da adição simultânea de um pigmento negro e um pigmento branco a uma cor, reduzindo a saturação e a luminosidade proporcionadas pelo pigmento da cor original. Num sistema adaptado a artes visuais obter-se-ia então a definição da cor pela especificação dos pigmentos de cores puras, dos pigmentos branco e negro e das proporções em que deveria ser realizada a sua mistura para que se obtivessem os tons pretendidos. O modelo de Ostwald, apresentado na figura 2.1 é um exemplo de um modelo do espaço de cor com base nestas noções.

Os modelos anteriores são modelos inteiramente baseados na percepção da cor e incluem um grau de subjectividade que não é desprezável e, por isso, impede uma definição precisa da cor e a sua correcta reprodução. Para ultrapassar estas limitações há que compreender o que é a luz e como ela é percebida e interpretada.

1 O termo equivalente em língua inglesa é hue.

2 É bem conhecido que a cor de uma peça de vestuário avaliada dentro de uma loja com luz artificial difere da cor percebida quando a mesma é avaliada à luz do dia.

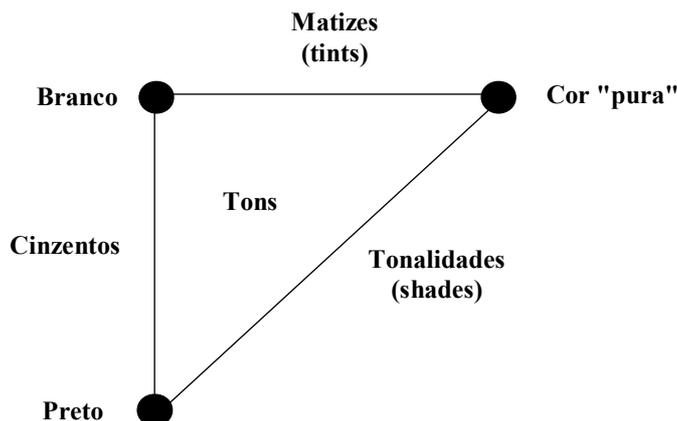


Figura 2.1 – Modelo de Ostwald

3 Espectro Visível

O espectro de radiação electromagnética ocupa uma grande gama de frequências (ou comprimentos de onda) que se estende desde as baixas frequências das ondas de rádio para comunicação submarina às muito altas frequências correspondentes aos raios gama. O espectro electromagnético é habitualmente dividido em bandas segundo os efeitos ou o tipo de utilização de cada banda. A figura 3.1 apresenta algumas destas bandas em função do comprimento de onda.

Os animais possuem órgãos receptores, os olhos, que são sensíveis à radiação electromagnética numa banda estreita de comprimentos de onda denominada espectro visível, pois estes órgãos interpretam como luz a radiação electromagnética dentro desta banda. Os limites do espectro visível variam consoante a espécie. A banda visível pelo olho humano situa-se entre 350 nm e 700 nm, mas estes limites variam de pessoa para pessoa. Os gatos e os insectos detectam luz numa banda mais larga que se estende à zona do ultra violeta ($\lambda < 400$ nm) e à zona dos infravermelhos ($\lambda > 700$ nm). Répteis como as cobras vêem principalmente na zona dos infravermelhos e são praticamente insensíveis ao que chamamos luz.

O ser humano interpreta os efeitos da absorção dos fotões como algo a que chama luz. A cor é essencialmente um conceito subjectivo próprio do ser humano e consiste na interpretação que o sistema sensorial e o cérebro atribuem aos diferentes comprimentos de onda da luz recebida ao interpretarem os estímulos nervosos provocados pela absorção dos fotões da radiação electromagnética com comprimento de onda compreendido entre 350 e 700 nm.

3.1 Cores Visíveis e Cores Espectrais

Cores espectrais são cores a que correspondem comprimentos de onda bem determinados do espectro visível. A observação de um espectro visível mostra que nele não se encontram todas as cores visíveis. Com efeito, existe uma gama de cores que habitualmente colocamos entre o azul e o vermelho e que inclui a cor magenta. Esta gama de cores está ausente do espectro visível pois tal equivaleria a ligar comprimentos

de onda menores (azuis) aos maiores (vermelhos) sem passar pelos comprimentos de onda intermédios (verdes e amarelos).

Também a cor branca não é uma cor do espectro visível porque não corresponde a nenhum comprimento de onda preciso, antes resulta da detecção simultânea de radiação visível em muitos comprimentos de onda, com uma intensidade mais ou menos uniforme.

Os exemplos apresentados mostram que existem mais cores visíveis do que cores espectrais. As cores não espectrais são devidas à mistura de luzes com diferentes comprimentos ou gamas de comprimentos de onda. A cor magenta é uma das misturas de cor mais simples que resulta da adição de duas cores do espectro, o azul e o vermelho. Para a produzir basta combinar a luz proveniente de dois focos em que cada um deles emite luz de cada uma daquelas cores. Neste caso a cor resultante é produzida por *adição*. Mas a cor magenta pode também ser produzida por *subtração*. Se se fizer incidir luz branca sobre uma superfície que absorve os comprimentos de onda visíveis intermédios (correspondentes aos verdes), a luz reflectida pela superfície ficará reduzida às componentes com comprimentos de onda nas zonas do azul e do vermelho. O cérebro humano interpretará então a superfície como sendo de cor magenta.

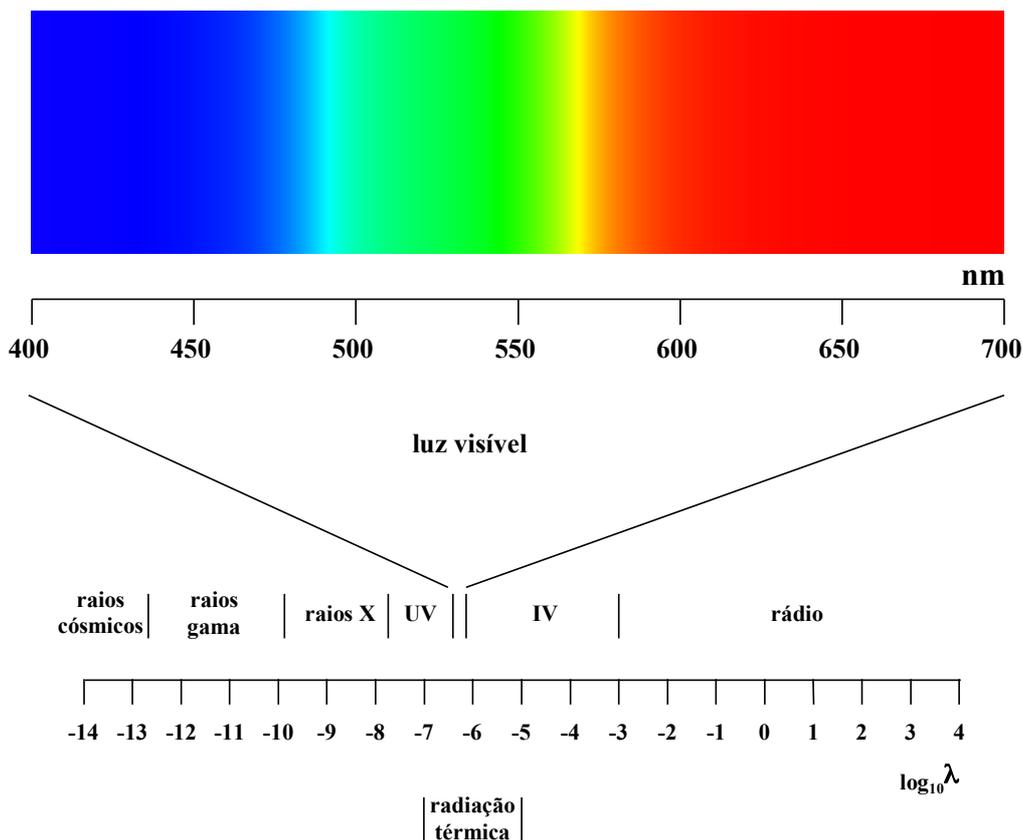


Figura 3.1 – Espectro electromagnético mostrando as bandas de comprimento de onda principais e a banda correspondente à luz visível.

3.2 Fisiologia da Visão

O funcionamento do olho humano é semelhante ao funcionamento dos equipamentos de registo de imagem, pelo menos na recepção e detecção da luz. O olho humano (veja-se a figura 3.2) foca a luz que recebe e fá-la convergir na retina onde se forma uma imagem invertida. Na retina existem células sensíveis à luz, os bastonetes e os cones, que transformam os fotões que absorvem em impulsos nervosos. Estes impulsos são comunicados ao cérebro através do nervo óptico. O cérebro recebe estes impulsos e processa-os, interpretando-os como imagens. Estas persistem durante algum tempo e, se um número suficiente de imagens se formar no cérebro por unidade de tempo, o cérebro interpreta as diferenças entre imagens sucessivas como informação sobre o movimento dos objectos visionados

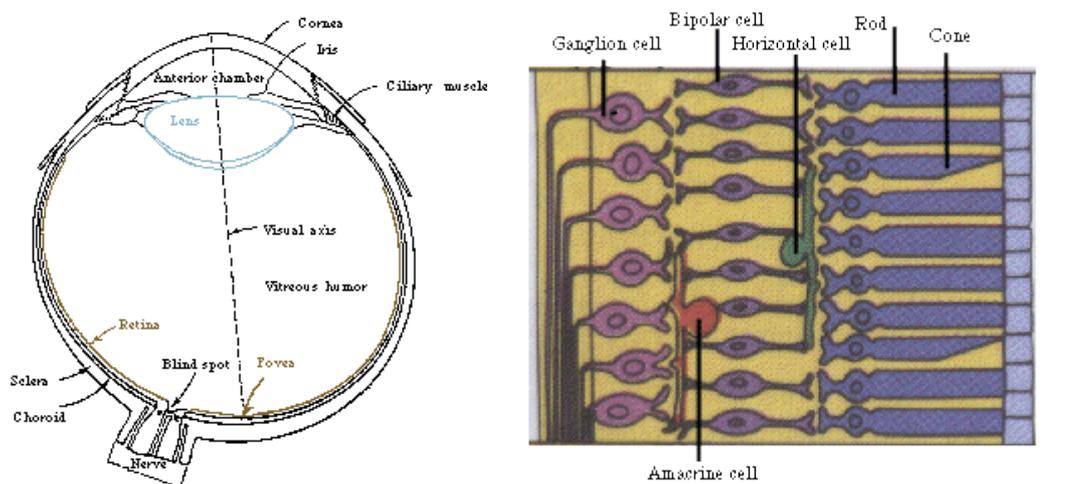


Figura 3.2 – Olho humano (Fischler e Firsheim, 1987) à esquerda e estrutura da retina (Hubel, 1988) à direita, mostrando cones e bastonetes.

O olho humano possui cerca de 115 a 120 milhões de bastonetes, de forma aproximadamente cilíndrica, e cerca de 6,8 milhões de cones. Os bastonetes, com cerca de 60 μm de comprimento e 2 μm de espessura, são sensíveis à intensidade luminosa em toda a gama de comprimentos de onda a que o olho humano é sensível. Os bastonetes proporcionam a detecção das formas dos objectos, a visão nocturna e a informação necessária à orientação. A detecção das formas dos objectos resulta da identificação das suas arestas através da descontinuidade da luminosidade nas arestas devida às diferentes luminosidades provenientes de faces dos objectos com diferentes orientações espaciais. Sendo sensíveis à luz em todos os comprimentos de onda, os bastonetes não conseguem discriminar entre luz recebida num comprimento de onda e luz recebida noutro comprimento de onda diferente, isto é, não detectam a cor.

Os cones, como o seu nome indica, são células da retina com forma aproximadamente cónica, mais curtas do que os bastonetes e um pouco mais largas (6 a 7 μm na sua secção média). Ao contrário dos bastonetes, os cones são sensíveis à luz apenas em certas gamas de comprimento de onda. Assim, existem cones sensíveis à luz na zona do vermelho (cones do tipo ρ), na zona do verde (cones do tipo γ) e na zona do azul (cones do tipo β). As gamas de comprimentos de onda a que cada tipo de cone é sensível estão parcialmente sobrepostas tal como a figura 3.3 apresenta. É por este facto que é errado falar em cones de uma dada cor, como se cada tipo de cone fosse sensível a um

comprimento de onda único e bem determinado³. Os cones necessitam de níveis de luminosidade mais elevados do que os bastonetes e, por esta razão, o olho humano não é capaz de detectar a cor dos objectos em condições de iluminação muito fraca como à noite.

A distribuição de cones e bastonetes na retina não é uniforme. No centro da retina existe uma pequena depressão com cerca de 1,5 mm de diâmetro, denominada "*fovea centralis*", onde apenas existem cones. Em torno desta depressão localiza-se uma área com a forma de uma coroa circular com cerca de 1 cm de diâmetro exterior, a "*macula lutea*", que possui cones e bastonetes. No resto da retina existem apenas bastonetes cuja densidade vai diminuindo à medida que aumenta a distância ao centro da retina. Esta distribuição de cones e bastonetes divide a visão em visão central, que permite identificar pormenores e cores e realizar a focagem do olho, e em visão periférica que apenas detecta formas e movimentos de objectos sem grande detalhe e sem cor.

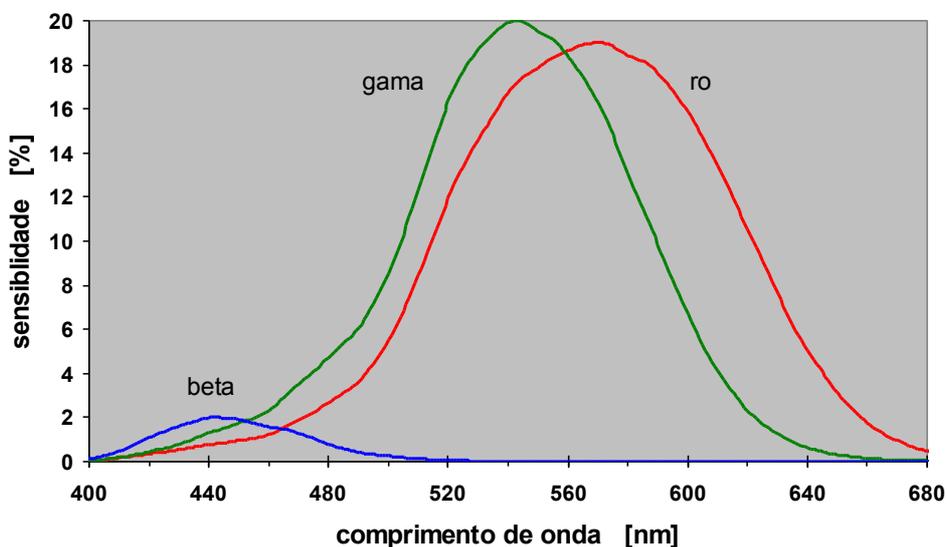


Figura 3.3 – Sensibilidade dos cones da retina em função do comprimento de onda.

A tabela 3.1 apresenta as gamas de comprimento de onda a que os três tipos de cone são sensíveis, o comprimento de onda em que cada um deles apresenta maior sensibilidade (λ_{\max}), a sua distribuição relativa média e a fracção da luz incidente que absorve. Verifica-se que o número de cones do tipo ρ existentes na retina é quase o dobro do número de cones do tipo γ . O número de cones do tipo β é muito inferior ao número de cones de qualquer dos outros dois tipos. Como se depreende da tabela 3.1 e da figura 3.3, a maior sensibilidade do olho humano deverá verificar-se na gama de comprimentos de onda detectada pelos cones γ e ρ , a zona intermédia entre o vermelho e o verde. Medidas experimentais confirmaram esta hipótese e permitiram determinar que a gama de sensibilidade máxima combinada dos três tipos de cones está centrada em torno do comprimento de onda de 550 nm, que corresponde à cor amarela. De modo

³ Alguns autores preferem designar por L, M e S os três tipos de cones para identificar, respectivamente, os cones sensíveis a maiores comprimentos de onda (L de large), intermédios (M de middle) e curtos (S de short).

inverso, é possível também concluir que é na gama de comprimentos de onda mais baixos (zona do azul) que existe menor sensibilidade à cor por parte de todos os tipos de cones.

Tipo de cone	Cor principal	Distribuição relativa [%]	Gama detectada [nm]	λ da maior sensibilidade [nm]	Fracção de luz absorvida a λ_{\max} [%]
β	azul	4	350-550	440	2
γ	verde	32	400-660	540	20
ρ	vermelho	64	400-700	580	19

Tabela 3.1 – Distribuição relativa dos três tipos de cones da retina, gamas detectadas e características dos comprimentos de onda de absorção máxima.

3.3 Teoria dos Três Estímulos

No seu conjunto, os cones da retina permitem detectar luz na gama de comprimentos de onda situada entre 350 e 700 μm . Estes valores poderão variar de pessoa para pessoa. Cada tipo de cone é sensível à luz numa gama limitada de comprimentos de onda. Esta sensibilidade não é uniforme, existindo um comprimento de onda para o qual a sensibilidade de cada tipo de cone atinge um máximo (veja-se a figura 3.3), e é praticamente nula para comprimentos de onda nos limites da gama a que é sensível. Assim, os comprimentos de onda inferiores a 400 nm são detectados apenas pelos cones de tipo β , ou seja, estes comprimentos de onda são interpretados como azuis indistintos. De igual modo, os comprimentos de onda superiores a 680 μm , que são detectados apenas pelos cones de tipo ρ , confundem-se e são percebidos como sendo o mesmo tom de vermelho.

Os comprimentos de onda compreendidos entre 400 e 680 nm são, em geral, detectados pelos três tipos de cones, mas cada tipo de cone detecta um dado comprimento de onda com uma sensibilidade diferente, o que dá origem a impulsos nervosos de diferentes intensidades. À luz com um comprimento de onda de 500 nm, que corresponde, aproximadamente, à cor ciano, correspondem sensibilidades dos cones de tipo β , γ e ρ de cerca de 20%, 30% e 10% das respectivas sensibilidades máximas. Para um comprimento de onda de 550 nm (amarelo, aproximadamente) teremos respostas de 0%, 99% e 80%, respectivamente. É a diferença entre as respostas dos três tipos de cones que permite interpretar diferentes comprimentos de onda como correspondendo a cores diferentes. Esta interpretação é conhecida como *teoria dos três estímulos*. Esta teoria postula que qualquer cor (comprimento de onda) do espectro visível pode ser reproduzida através da adição dos resultados obtidos pelo estímulo dos três tipos de cones de forma diferente. A cor percebida depende unicamente da relação entre os três estímulos.

A teoria dos três estímulos tem também outra consequência que é a verificação de que da combinação de cores diferentes pode resultar uma cor única, como é o caso da cor

ção. Chama-se *cores metâmeras* ou *metâmero* a duas cores espectralmente diferentes mas que produzem a sensação de uma mesma cor.

3.4 Outras Consequências na Percepção da Cor

As características do olho humano atrás descritas têm consequências na percepção da forma dos objectos e na sua cor.

A primeira consequência respeita à capacidade de distinção entre cores puras próximas. Esta capacidade de discriminação está relacionada com a grandeza absoluta e relativa das respostas produzidas por cada tipo de cone a cada comprimento de onda. Estas afectam a diferença que deve existir entre dois comprimentos de onda próximos para que as respectivas cores, embora semelhantes, possam ser percebidas como cores diferentes. Esta diferença parte de um valor de cerca de 10 nm a 400 nm e passa sucessivamente pelos máximos e mínimos apresentados na tabela 3.2 até um máximo de, aproximadamente, 11 nm a cerca de 660 nm.

λ [nm]	Discriminação $\Delta\lambda$ [nm]	Obs.
400	10	máximo
425	2	mínimo
445	4,5	máximo
475	1	mínimo
525	4	máximo
560	1	mínimo
660	11	máximo

Tabela 3.2 - Valores aproximados das diferenças máximas e mínimas de comprimento de onda para que duas cores puras contíguas sejam percebidas como cores distintas em função do comprimento de onda aproximado.

O resultado desta discriminação variável em função do comprimento de onda é o número de cores puras distintas que, no total, o olho humano é capaz de discriminar e que corresponde a 128 cores. Por outro lado, uma análise das respostas combinadas dos três tipos de cones da retina leva-nos a considerar que deverá existir maior facilidade de discriminar entre cores sombreadas na zona do amarelo e uma menor facilidade na zona do azul. Com efeito, experiências realizadas permitiram determinar que o olho humano consegue distinguir um máximo de 23 cores sombreadas na zona do amarelo e 16 cores sombreadas na zona do azul. Como o olho humano consegue igualmente distinguir entre cerca de 130 níveis de saturação, é fácil então concluir que o olho humano é capaz de discriminar cerca de 380 000 ($128 \times 23 \times 130$) cores diferentes.

Vimos já que a visão central detecta movimentos dos objectos e é muito mais sensível aos seus pormenores. A sensibilidade aos pormenores dos objectos é essencial para a detecção das formas destes através do reconhecimento das arestas que, por sua vez, desempenha um papel importante no subsistema inconsciente responsável pela focagem do olho. A visão periférica apenas detecta formas difusas e movimentos de modo

grosseiro. Esta última função está relacionada com a detecção de movimentos susceptíveis de fazer perigar a segurança dos indivíduos.

A informação de cor produzida pelos cones não é suficiente para a detecção das arestas e formas dos objectos, mas pode reforçar positivamente a informação proveniente dos bastonetes quando, por exemplo, existam diferenças de cor em torno das arestas. No entanto, em situações em que a diferença de cor entre objectos não é facilmente perceptível, a informação de cor proveniente dos cones tem o efeito de dificultar a percepção de arestas e formas em vez de a facilitar⁴.

4 Modelação da Cor

Os modelos de cor derivados da experiência do dia-a-dia empregues em artes visuais são subjectivos dado que dependem de factores múltiplos tais como o juízo pessoal e diferenças fisiológicas dos observadores. Estes modelos heurísticos são ainda dependentes da situação concreta em que são realizadas as avaliações porque a iluminação do objecto e a iluminação geral, o contexto em que a cor é avaliada e o tamanho da amostra de cor fazem variar a sua percepção.

A modelação da cor assume extrema importância em Computação Gráfica pois a apresentação de imagens em dispositivos de saída gráfica necessita de informação sobre a cor dos objectos das cenas. Os formatos intermédios têm então que guardar esta informação de uma forma coerente e precisa que introduza o menor número de erros ou falhas de precisão. Tendo em conta a teoria dos três estímulos, é imediato que qualquer modelo de cor deverá possuir três parâmetros. O problema essencial da modelação da cor está na definição de que grandezas devem ser associadas a cada um dos três parâmetros. Esta escolha, que é inclusivamente influenciada pela natureza dos dispositivos de saída gráfica, origina a existência de vários sistemas de cor, como veremos de seguida. Mas antes disso é necessário precisar o significado das grandezas e termos a empregar.

A luz emitida segundo um comprimento de onda preciso produz uma cor pura do espectro visível. Na realidade, exceptuando fontes de luz como os lasers, as fontes de luz emitem-na com uma determinada energia em muitos comprimentos de onda numa gama centrada à volta de um comprimento de onda dominante⁵. Se a energia emitida pela fonte for baixa, a cor aparecerá como um sombreado mais ou menos escuro, se a energia for alta, a cor será percebida como uma cor viva e brilhante. Adicionando à cor emitida uma luz branca, cuja densidade energética espectral é mais ou menos uniforme em todos os comprimentos de onda, a cor tornar-se-á tanto mais pálida e esbranquiçada quanto maior for essa densidade espectral em relação à energia da cor pura. O resultado será uma cor totalmente branca no caso limite em que as energias espectrais forem iguais.

⁴ É por esta razão que a detecção visual de objectos distantes no alto mar, onde não existem diferenças de cor significativas, é confiada a indivíduos daltónicos dado que estes, porque não detectam tons de verde ou de vermelho, recebem menos informação de cor e são consequentemente menos confundidos.

⁵ No corpo negro, o comprimento de onda dominante depende da temperatura do corpo emissor através da lei de Wien.

Na figura 4.1, a luz branca possui uma densidade energética espectral e_1 e a cor pura uma densidade energética espectral de e_2 . Quando $e_1 = 0$, diremos que a *pureza da excitação luminosa* é de 100%, enquanto esta pureza será de 0% se $e_1 = e_2$. A pureza da excitação luminosa será então

$$\frac{e_2 - e_1}{e_2} \tag{4.1}$$

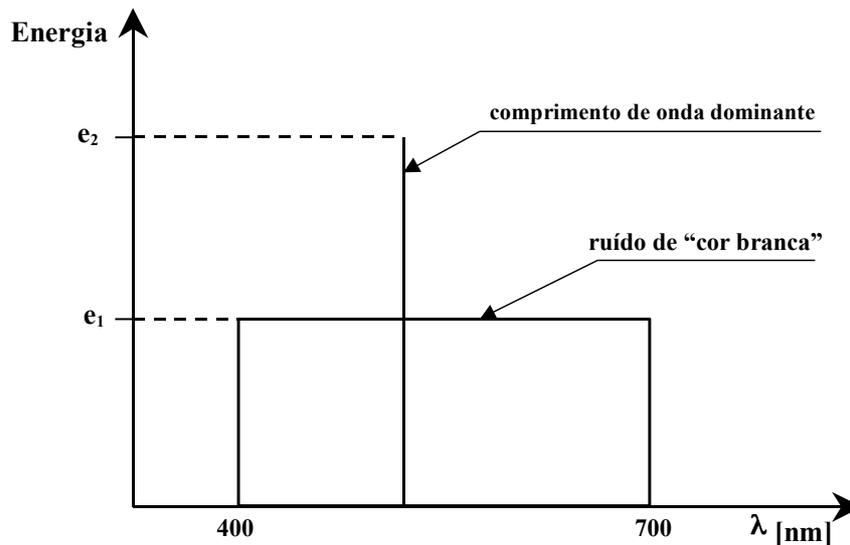


Figura 4.1 – Modelo simples de cor

Podemos identificar imediatamente os três parâmetros quantitativos deste modelo simples: comprimento de onda dominante, luminância (energia da luz emitida) e pureza da excitação luminosa (saturação). A estes parâmetros quantitativos correspondem termos perceptuais. Esta correspondência é a seguinte:

Termo quantitativo	Termo perceptual
Comprimento de onda dominante	Cor
Pureza da excitação luminosa	Saturação
Luminância	Luminosidade (de objectos reflectores) Brilho (de objectos emissores)

4.1 Modelos de cor

A teoria dos três estímulos parece indicar que bastará emitir três cores primárias que sejam detectáveis pelos três tipos de cones da retina para que se possam reproduzir todas as cores visíveis, variando simplesmente a proporção das quantidades de luz emitida por cada uma das fontes primárias. Estas proporções seriam determinadas pelas curvas de resposta característica de cada um dos três tipos de cones.

Na prática, as três cores primárias emitidas por cada um dos tubos de raios catódicos de um monitor (vermelho a 700 nm, verde a 546 nm e azul a 436 nm) não correspondem às

cores detectadas pelos cones. Há então que modificar as funções peso aplicadas a cada uma das componentes primárias emitidas. Estas novas funções peso apresentam valores negativos em algumas gamas de comprimento. Isto significa que, com um monitor, não é possível reproduzir todos os comprimentos de onda de luz visível, isto é, não é possível reproduzir todas as cores do espectro visível pela combinação ponderada de luzes vermelha, verde e azul. Existem portanto cores que não podem ser simplesmente reproduzidas em monitores a cores pela adição ponderada das cores vermelha, verde e azul.

4.2 Modelo CIE

A incapacidade de modelos baseados na mistura de cores vermelhas, verdes e azuis poderem representar todas as cores do espectro visível levou a que a CIE (Commission Internationale de l'Éclairage) criasse em 1931 um modelo com três cores primárias denominadas X, Y e Z, que substituíam as cores primárias vermelha, verde e azul, e que possuíam funções peso (ou de mistura) de X, Y e Z, intrinsecamente positivas e tais que passasse a ser possível representar todas as cores do espectro visível. Estas funções, as funções \bar{x}_λ , \bar{y}_λ e \bar{z}_λ , são chamadas *funções de ajustamento da cor* ou CMF, Colour Matching Functions (veja-se a figura 4.2) foram calculadas a partir das funções de mistura para as cores vermelha, verde e azul e apresentam a particularidade de a função \bar{y}_λ se encontrar ajustada à resposta combinada do olho humano em função do comprimento de onda. A relação entre as cores primárias CIE e as cores vermelha, verde e azul é dada por

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0,49 & 0,31 & 0,20 \\ 0,17697 & 0,81240 & 0,01063 \\ 0 & 0,01 & 0,99 \end{bmatrix} \begin{bmatrix} r \\ g \\ b \end{bmatrix} \quad (4.2)$$

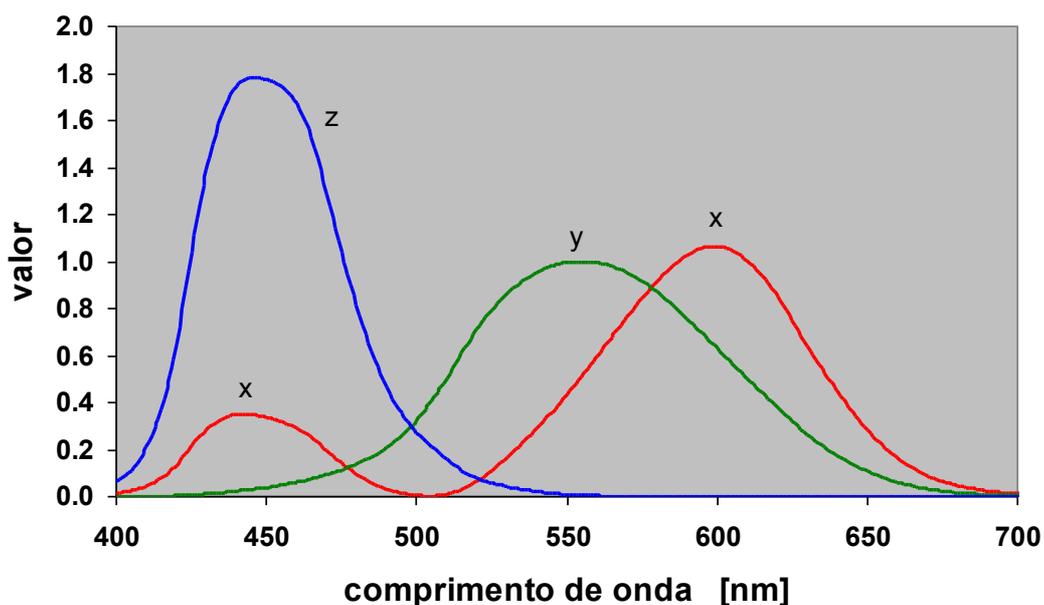


Figura 4.2 – Funções CMF x, y z.

Pela transformação inversa de (4.2) é possível obter as quantidades de luz vermelha (700 nm), verde (546 nm) e azul (436 nm) necessárias para reproduzir uma qualquer cor do espectro visível a partir das cores primárias CIE. A figura 4.3 apresenta estas quantidades. Existem nesta representação gamas de comprimento de onda em que estas quantidades são negativas o que confirma o que atrás foi afirmado sobre a impossibilidade de os monitores a cores poderem reproduzir fielmente todas as cores do espectro visível.

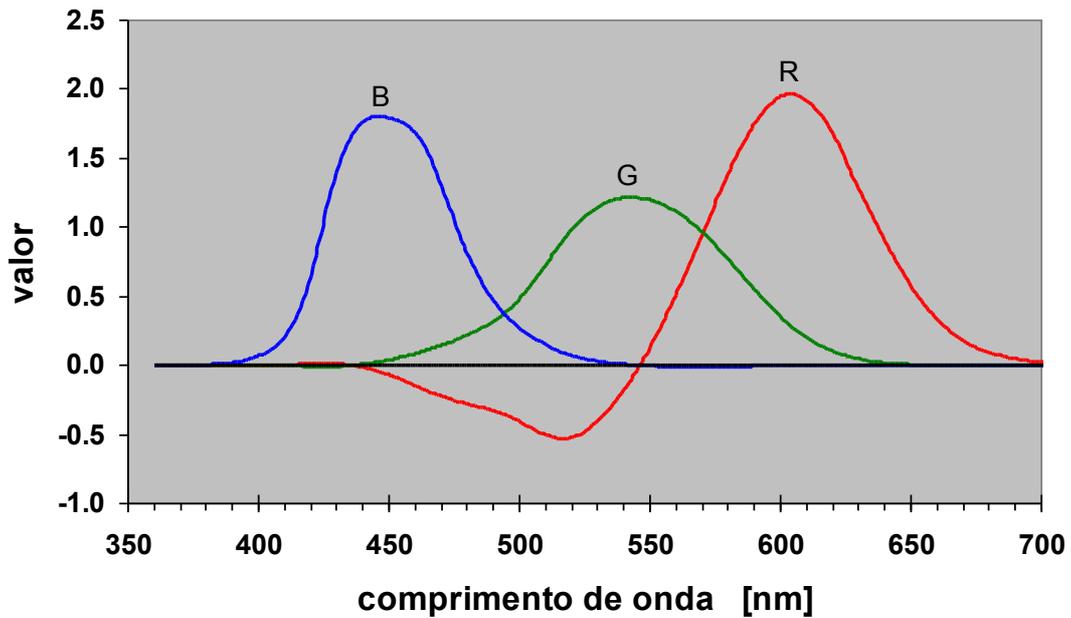


Figura 4.3 – Quantidades RGB necessárias para reproduzir todas as cores do espectro visível. Note-se a existência de quantidades negativas.

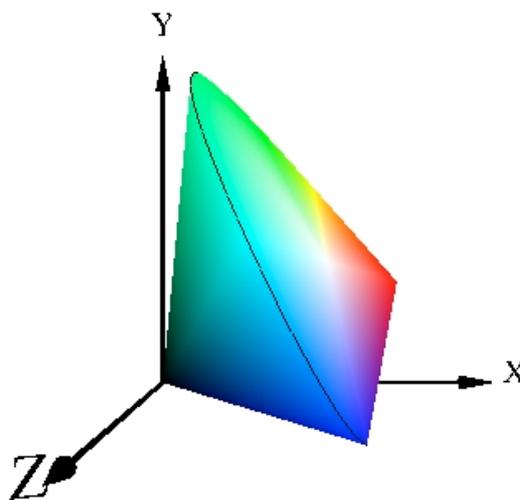


Figura 4.4 – Espaço de cor CIE.

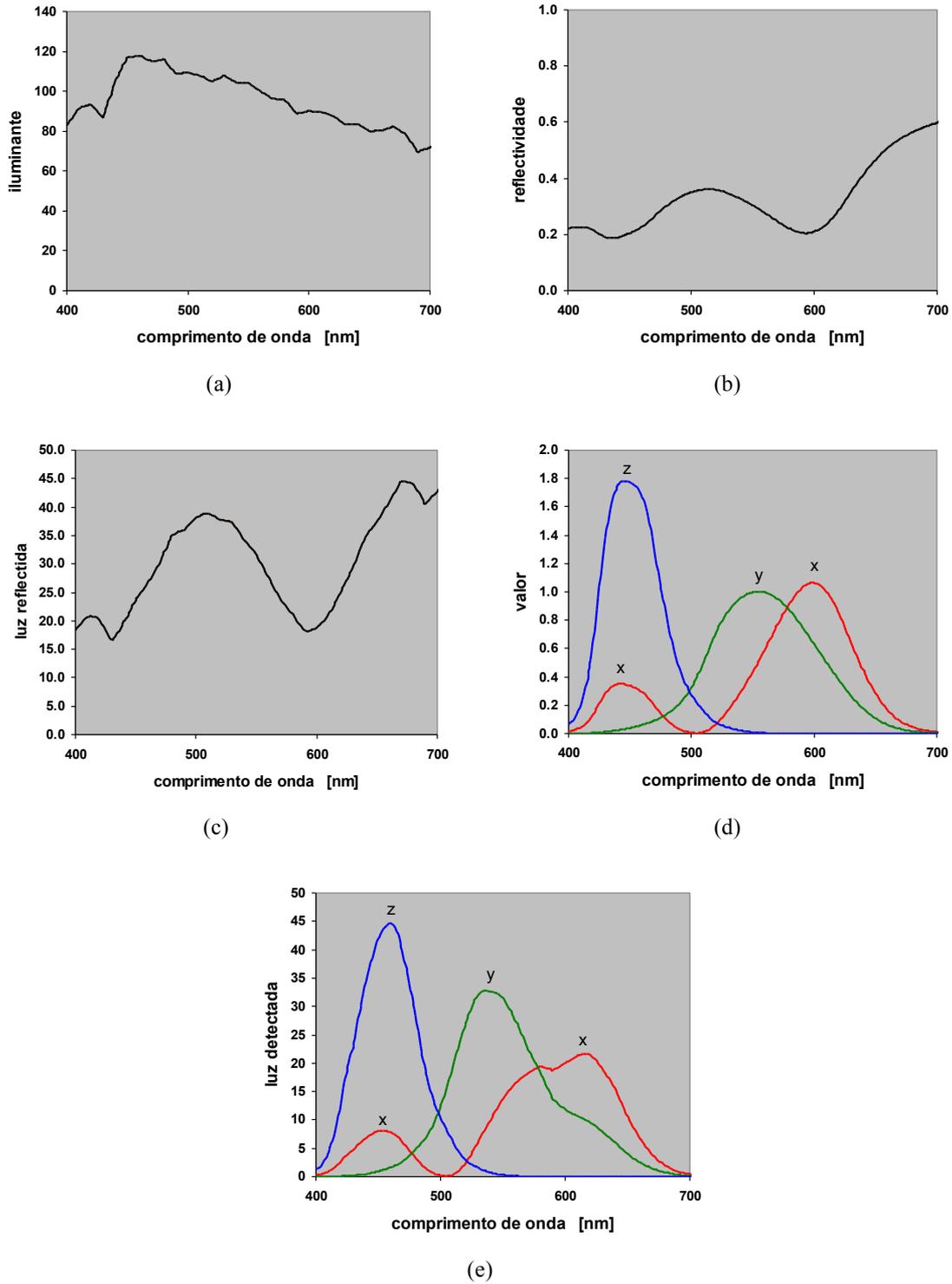


Figura 4.5 – Luz detectada proveniente de uma superfície iluminada pelo iluminante D65 (a). A reflectividade da superfície (b) dá origem à luz reflectada (c) que, ponderada pelas funções CMF (d), resulta nos três estímulos detectados (e).

As funções CMF permitem também calcular a luz detectada pelos cones. A figura 4.5 apresenta o exemplo do que sucede à luz emitida pelo iluminante D65 quando incide numa superfície não emissora e cuja reflectividade depende do comprimento de onda. A

luz reflectida pela superfície apresenta um espectro muito diferente do espectro da luz nela incidente. Ponderando esta luz reflectida pelas funções CMF, obteremos então os valores dos três estímulos correspondentes à luz reflectida pela superfície.

Se A, B e C forem os pesos de cada uma das três cores primárias CIE, teremos que, para a cor D, será

$$D = AX + BY + CZ \quad (4.3)$$

O sub espaço contendo todas as cores visíveis está então localizado no octante em que as coordenadas A, B e C do espaço X,Y,Z são positivas, pois os valores das funções de ajuste são intrinsecamente positivas. Este sub espaço apresenta a forma de um cone com o vértice na origem do espaço X,Y,Z e uma secção em forma de ferradura tal como a figura 4.4 apresenta. O modelo CIE define também uma luz branca de referência, o chamado iluminante C, que corresponde à luz solar quando emitida à temperatura de referência de 6774 K.

Para uma dada cor caracterizada pelos valores A, B e C, podemos definir os chamados *valores de cromaticidade* x, y e z como sendo

$$x = \frac{A}{A+B+C} \quad y = \frac{B}{A+B+C} \quad z = \frac{C}{A+B+C} \quad (4.4)$$

Como $x+y+z=1$, estes valores definem um plano no espaço cromático. Todas as cores possíveis são então projectáveis sobre este plano através de uma projecção cónica em que a recta projectante une a origem do espaço cromático com o ponto representativo da cor a projectar. O *diagrama de CIE de cromaticidade* corresponde então à projecção deste plano no plano X,Y. Para um qualquer ponto deste plano, dados os valores de x e y, o valor de z fica automaticamente determinado uma vez que $z = 1 - x - y$. Os valores originais de A, B e C podem ser recuperados, bastando para tal que se conheça o valor da intensidade luminosa de uma das componentes X, Y ou Z. Logo

$$A = \frac{x}{y} B \quad B = \frac{y}{y} B \quad C = \frac{1-x-y}{y} B \quad (4.5)$$

O diagrama CIE de cromaticidade representado na figura 4.6 apresenta todos os valores de cromaticidade para as cores visíveis dado que a todos os pontos com valores iguais de cromaticidade, mas de diferentes intensidades luminosas, corresponde um único ponto neste diagrama. As cores puras do espectro visível encontram-se localizadas sobre a parte curva do limite do diagrama, com a fonte de luz branca referência localizada no seu interior num ponto cujas coordenadas são, aproximadamente, de 1/3. As cores puras visíveis que não existem no espectro visível, como a cor magenta, encontram-se localizadas sobre o segmento de recta que une os extremos do arco em forma de ferradura.

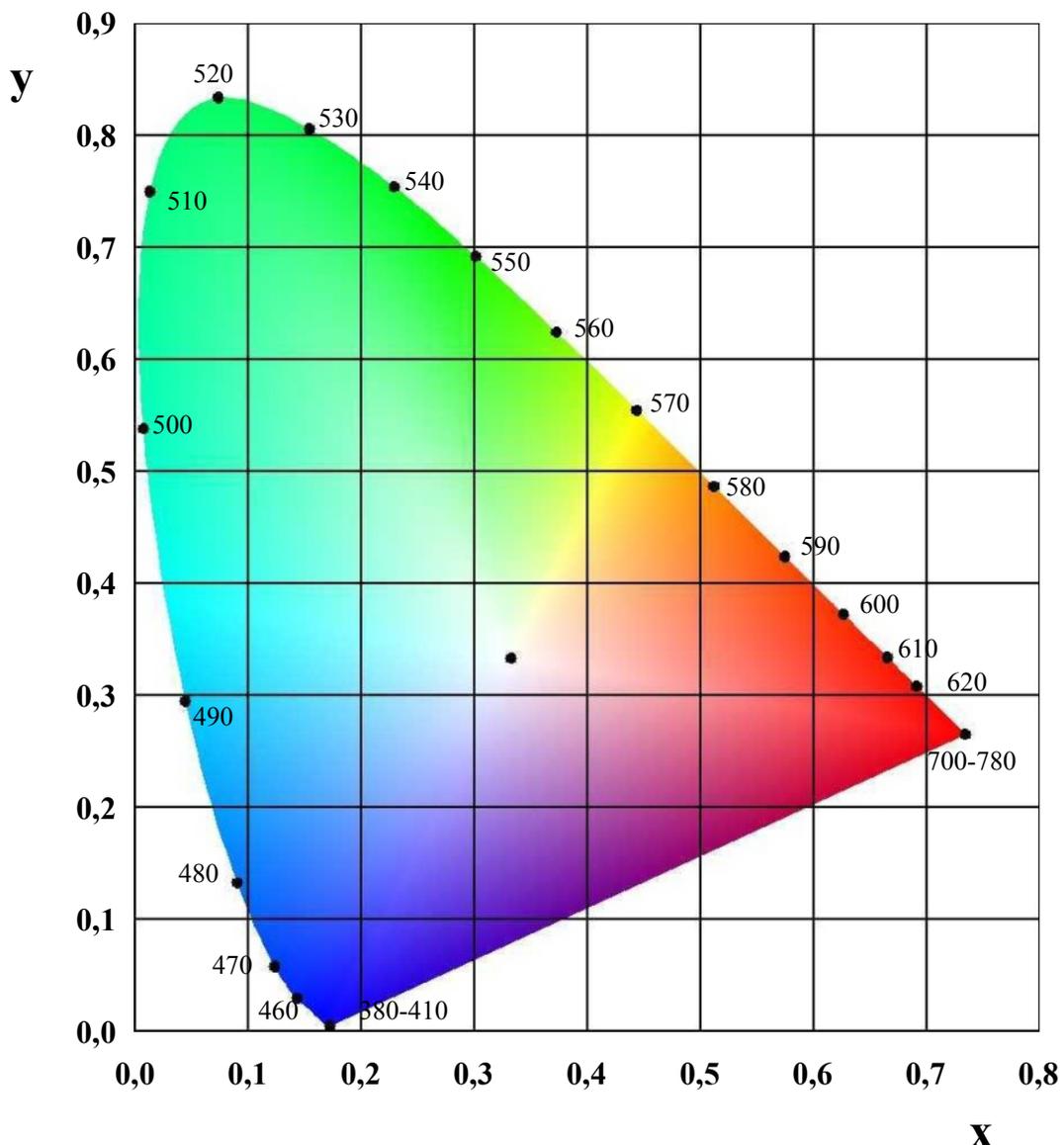


Figura 4.6 – Diagrama CIE de cromaticidade

O diagrama de cromaticidade apresenta algumas propriedades das quais a mais relevante é a sua linearidade. Assim, um segmento de recta que una os pontos representativos de duas cores representa todas as cores possíveis de obter pela mistura dessas duas cores em quaisquer proporções. As proporções de mistura correspondentes à cor de um ponto localizado sobre esse segmento podem ser calculadas a partir da chamada regra da alavanca. Quando essas duas cores forem o iluminante branco de referência e uma cor pura, o quociente entre a distância da cor resultante à cor pura e a distância da cor pura ao iluminante de referência corresponde à saturação da cor. O comprimento de onda dominante desta cor é o comprimento de onda característico da cor pura correspondente.

Outra consequência da linearidade respeita à adição de duas cores de igual saturação localizadas sobre uma linha que contém o ponto representativo da cor branca e que se localizam de lados diferentes da linha relativamente à cor branca. A adição dessas duas cores produz luz branca e, portanto, as duas cores são complementares.

A linearidade do diagrama CIE de cromaticidade permite também a visualização das gamas de cor disponíveis nos dispositivos de saída gráfica. Com efeito, se representarmos no diagrama as três cores primárias típicas de dispositivos como monitores a cores⁶, obtemos um triângulo completamente inscrito dentro do diagrama. Qualquer que seja o triângulo considerado, este nunca poderá compreender todos os pontos interiores do diagrama CIE de cromaticidade. Demonstra-se assim que os dispositivos de saída gráfica do tipo monitor a cores e baseados nas três cores primárias vermelha, verde e azul, nunca poderão reproduzir todas as cores visíveis. Por outro lado, comparando as gamas de monitores a cores com as gamas de impressoras a cores, verifica-se que estas últimas estão normalmente contidas dentro daquelas, o que significa que existem cores que é possível apresentar em monitores a cores, mas que não podem ser reproduzidas por impressoras a cores. Isto implica que as impressoras a cores não podem reproduzir uma imagem tão fielmente como um monitor a cores e, portanto, é necessário ou proceder à redução de cor na transposição para o papel ou utilizar uma gama de cores reduzida nos monitores para que exista uma correspondência o mais fiel possível.

Tal como vimos anteriormente, um espaçamento uniforme de cores no espaço de percepção não corresponde a um espaçamento uniforme em termos de comprimento de onda (veja-se a tabela 3.2). Para resolver este problema, o modelo CIE foi objecto de várias alterações, entre as quais as que resultaram no modelo CIE LUV de 1976.

4.3 Modelo RGB

O modelo RGB é um modelo de cor concebido com base nos dispositivos de saída gráfica com três cores primárias: vermelho, verde e azul. A sigla RGB deriva da junção das primeiras letras dos nomes destas cores primárias em língua inglesa: Red, Green e Blue.

O modelo RGB descreve as cores como o resultado da adição das três cores primárias, cada uma delas com uma intensidade que pode variar entre 0 e 1. O valor 1 corresponde à intensidade máxima com que a cor pode ser apresentada no dispositivo gráfico e o valor 0 à intensidade mínima⁷. A cor branca corresponde à representação simultânea das três cores primárias, todas à sua intensidade máxima, e a cor preta à cor que é obtida quando todas as cores primárias apresentam intensidade mínima (0).

O modelo RGB está intimamente associado às superfícies emissoras de luz. É por esta razão que este modelo é o modelo quase universalmente empregue pelos equipamentos que manipulam a emissão de luz, tais como os monitores e os televisores a cores. Os filmes fotográficos e cinematográficos, e os registos em vídeo empregam também o modelo RGB no seu funcionamento.

O modelo é omissivo quanto ao que é uma cor primária pura, ou seja, não define qual o comprimento de onda a que corresponde cada uma das três cores primárias. Esta omissão tem consequências na reprodução da cor. Com efeito, verificam-se variações sensíveis de monitor para monitor e, no caso dos televisores, a publicidade menciona muitas vezes a expressão “cores mais naturais”.

⁶ Os pontos correspondentes às cores primárias de cada dispositivo são determináveis por meio do colorímetro.

⁷ Este não corresponde exactamente à não emissão como veremos ao construir escalas de intensidade.

Os três parâmetros do modelo RGB, as intensidades das três cores primárias do modelo, definem um espaço tridimensional com três direcções ortogonais (R, G e B). Constrói-se assim o espaço RGB de cor. As cores deste espaço existem no sub espaço em que $0 \leq (R,G,B) \leq 1$. Este sub espaço corresponde a um cubo de aresta unitária em que o vértice de coordenadas (0,0,0) representativo da cor negra coincide com a origem do espaço e o vértice representativo da cor branca corresponde ao ponto de coordenadas (1,1,1), tal como a figura 4.7 apresenta.

A cada uma das três cores primárias puras corresponde um dos vértices do cubo localizados sobre os eixos do espaço, em que apenas uma das coordenadas não é nula. As cores complementares principais (magenta, amarelo e cião) situam-se nos três vértices restantes e, como veremos, correspondem à adição de duas cores primárias.

Vermelho	=	(1,0,0)		
Verde	=	(0,1,0)		
Azul	=	(0,0,1)		
Preto	=	(0,0,0)		
Branco	=	Vermelho (1,0,0)	+	Verde (0,1,0)
				+ Azul (0,0,1)
				= (1,1,1)
Amarelo	=	Vermelho (1,0,0)	+	Verde (0,1,0)
				= (1,1,0)
Ciã	=	Verde (0,1,0)	+	Azul (0,0,1)
				= (0,1,1)
Magenta	=	Vermelho (1,0,0)	+	Azul (0,0,1)
				= (1,0,1)

Os tons de cinzento correspondem aos pontos situados sobre a diagonal principal em que as três componentes apresentam a mesma intensidade

$$\text{Cinzento} = (x,x,x) \quad \text{com } 0 \leq x \leq 1$$

A designação de cor complementar atribuída ao amarelo, cião e magenta provém da sua localização em vértices do cubo do espaço RGB de cor que são opostos aos vértices das cores primárias (vermelho para o cião, verde para o magenta e azul para o amarelo) e do facto de da adição da cor complementar à respectiva cor primária resultar sempre a cor branca.

Vermelho (1,0,0)	+	Ciã (0,1,1)	=	Branco (1,1,1)
Verde (0,1,0)	+	Magenta (1,0,1)	=	Branco (1,1,1)
Azul (0,0,1)	+	Amarelo(1,1,0)	=	Branco (1,1,1)

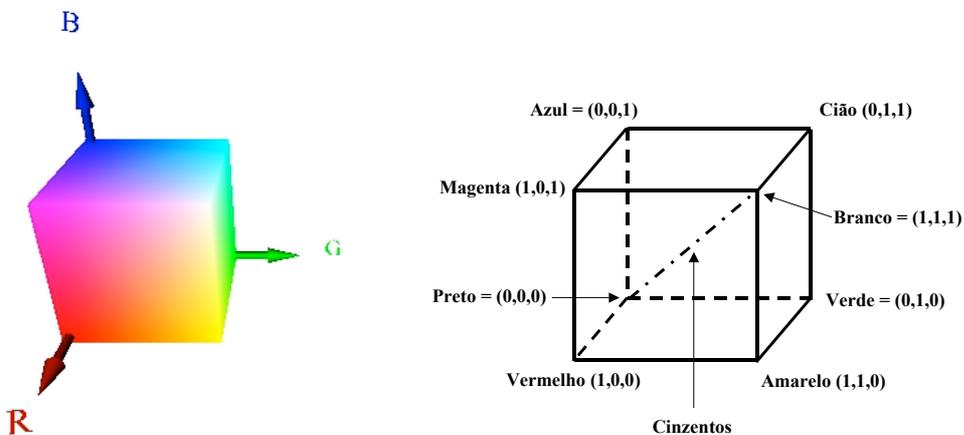


Figura 4.7 – Espaço de cor RGB

Tradicionalmente, as implementações do modelo RGB nos sistemas gráficos empregam valores inteiros entre 0 e 255 para exprimir o valor da intensidade de cada componente em vez de valores reais normalizados entre 0 e 1. Esta idiosincrasia teve origem no facto de o processamento de valores inteiros ter sido muito mais rápido do que o processamento de valores reais nos primeiros sistemas gráficos, além de que a notação com inteiros ser mais simples de escrever e apreender do que a notação com valores reais fraccionários. A discretização em 256 valores de intensidade é mais do que suficiente para o olho humano dado que este consegue distinguir entre um mínimo de 16 intensidades na zona do azul e um máximo de 23 intensidades na zona do amarelo.

4.4 Modelo CMY

O modelo CMY é um modelo de cor baseado nas cores complementares: cião, magenta e amarelo. A sigla CMY provém da junção da primeira letra dos nomes destas cores em língua inglesa: Cyanide, Magenta e Yellow.

O modelo CMY tem por base os fenómenos que se verificam quando a luz incide em superfícies. Estas podem absorver, reflectir ou refractar a luz de forma desigual consoante o comprimento de onda. Quando uma luz branca incide sobre uma superfície, existem gamas de comprimento de onda em que a luz é absorvida pela superfície⁸. A luz correspondente às gamas de comprimento de onda não absorvidas é, em geral, reflectida. A nossa percepção visual da cor da superfície é dada pela cor dessa luz reflectida⁹. Assim, quando a luz natural incide numa superfície que absorve os comprimentos de onda na zona do vermelho, a luz reflectida não terá quaisquer componentes nessa gama e será constituída por apenas verdes e azuis, ou seja, o olho humano detectará a superfície como sendo da cor cião. De igual modo, o olho humano perceberá uma superfície como sendo de cor vermelha se esta absorver comprimentos de onda situados na zona do verde e do azul.

Esta subtracção (por absorção) da luz em determinados comprimentos de onda é a razão pela qual o modelo CMY é também designado por modelo substractivo da cor, em oposição ao modelo RGB que é designado por modelo aditivo da cor. As cores cião, magenta e amarelo são designadas por cores primárias complementares ou cores primárias substractivas dada a forma como resultam da subtracção do, respectivamente, vermelho, verde e azul à cor branca.

O espaço CMY pode ser construído da mesma forma que é construído o espaço RGB. As coordenadas do espaço CMY passam a ser as cores primárias substractivas e, tal como para o espaço RGB, as intensidades de cada componente estão normalizadas num intervalo entre 0 (ausência da componente) e 1 (componente na sua intensidade máxima). As cores ficarão então localizadas dentro de um sub espaço com a forma de um cubo, o cubo CMY.

Quando a intensidade de uma componente CMY é nula, a cor aditiva primária que a complementa não é absorvida e é totalmente reflectida. Se isto se passar com todas as componentes CMY, ou seja, a cor CMY tem como componentes (0,0,0) e isto significa

⁸ A absorção depende de várias características das superfícies como natureza do material, rugosidade, forma da superfície, etc.

⁹ As superfícies também emitem radiação mas, a temperaturas ambientes, esta radiação é emitida na zona dos infravermelhos que não é detectada pelo olho humano e, por esta razão, não é aqui considerada.

que toda a luz branca incidente é reflectida e a superfície é percebida como sendo branca.

De modo semelhante, quando uma componente CMY tem a intensidade máxima (1), a cor aditiva primária complementar é totalmente absorvida. Quando todas as componentes CMY apresentarem intensidade máxima (1,1,1), as cores aditivas primárias complementares correspondentes serão totalmente absorvidas e, conseqüentemente, a superfície será percebida como sendo de cor preta, dado que não reflectirá qualquer luz.

Quando a luz reflectida por uma superfície apresenta a máxima intensidade para o ciano (o vermelho é absorvido) e para a cor magenta (o verde é absorvido) e nula para o amarelo significa que a superfície tem cor azul. Resultados idênticos podem ser obtidos para as cores primárias aditivas verde e vermelho quando se consideram os pares magenta e amarelo e ciano e amarelo, respectivamente. As cores primárias aditivas são então produzíveis pela adição de duas cores subtractivas primárias.

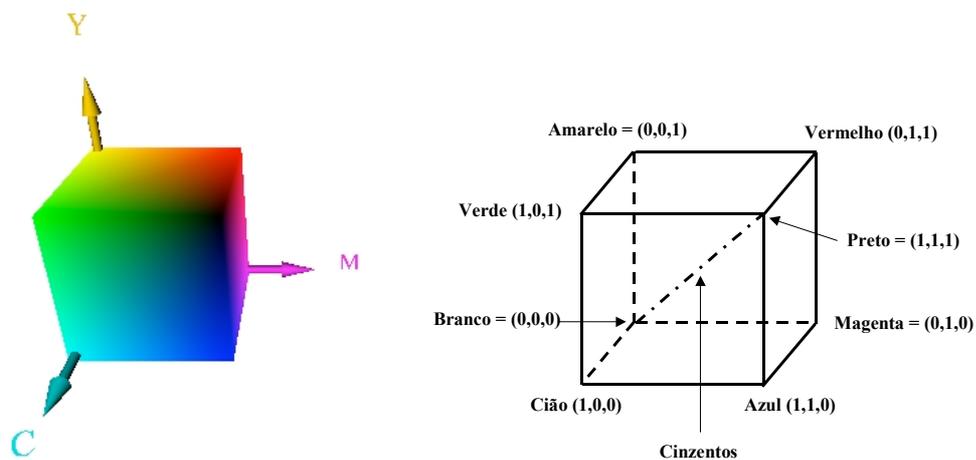


Figura 4.8 – Espaço de cor CMY.

O cubo CMY tem a cor branca (0,0,0) na origem e a cor negra (1,1,1) no vértice oposto, como a figura 4.8 apresenta. As cores primárias subtractivas estão localizadas no cubo CMY nos vértices deste cubo localizados sobre os eixos do espaço CMY, à distância de uma unidade da origem. Os restantes três vértices correspondem à localização das cores aditivas primárias. Assim temos

Ciano	=	(1,0,0)			
Magenta	=	(0,1,0)			
Amarelo	=	(0,0,1)			
Preto	=	(1,1,1)			
Branco	=	(0,0,0)			
Vermelho	=	Magenta(0,1,0)	+	Amarelo(0,0,1)	= (0,1,1)
Verde	=	Ciano(1,0,0)	+	Amarelo(0,0,1)	= (1,0,1)
Azul	=	Ciano(1,0,0)	+	Magenta(1,0,0)	= (1,1,0)

A transformação entre o espaço CMY e o espaço RGB é

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} C \\ M \\ Y \end{bmatrix} \quad \begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4.6)$$

As impressoras a cores empregam o modelo CMY por deposição sobre o papel de tintas correspondentes às cores primárias complementares. Tal como no modelo RGB, o modelo CMY também não define os comprimentos de onda das cores primárias pelo que a reprodução da cor está dependente das tintas empregues e poderá variar de impressora para impressora devido à dificuldade adicional da fabricação de tintas cujas cores correspondam exactamente às cores primárias subtractivas. Por outro lado, a deposição de três tintas correspondentes às cores primárias subtractivas não consegue produzir uma cor negra porque a absorção da gama de comprimentos de onda da luz incidente a ser absorvida por cada tinta não é total. Os fabricantes de impressoras a cores resolveram este problema instalando um quarto tinteiro com uma tinta de cor negra que permite obter uma cor negra muito mais correcta.

4.5 Modelo HSV

O modelo de cor HSV é mais intuitivo que os modelos RGB e CMY. A sua sigla é formada pela primeira letra das palavras de língua inglesa correspondentes às três variáveis do modelo: Hue (cor, matiz, cambiante de cor), Saturation (saturação) e Value (valor). Com estas variáveis, o modelo HSV aproxima-se muito do modelo intuitivo empregue em artes visuais que emprega os conceitos qualitativos de matiz, luz e tonalidade.

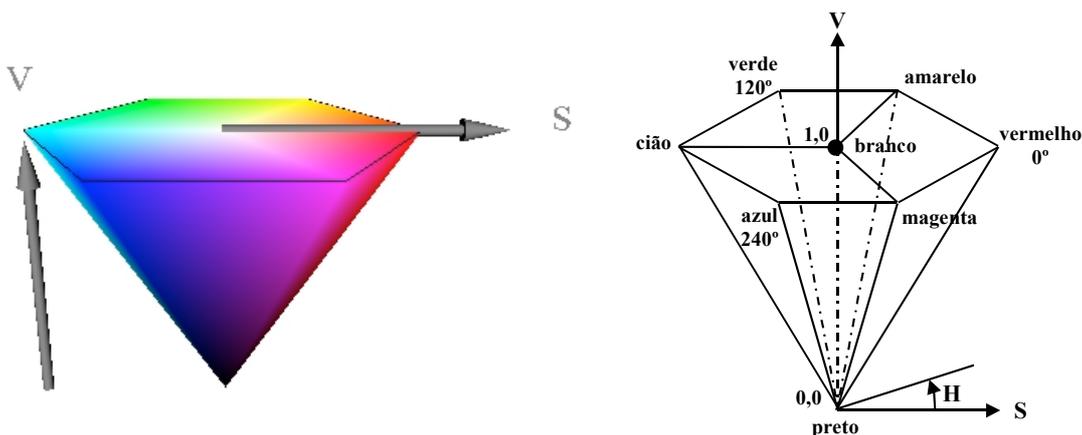


Figura 4.9 – Espaço de cor HSV

A representação do espaço de cor do modelo HSV pode ser facilmente compreendida a partir do cubo do modelo RGB. Colocando o ponto de vista sobre a linha contendo a diagonal principal do cubo RGB e orientando a visão na direcção da origem do espaço RGB, o cubo RGB apresentará uma secção diagonal onde, intercaladas e em planos

diferentes, se encontram as cores primárias aditivas e as cores primárias subtractivas. A cor branca ocupará o centro. Projectando tanto as cores primárias aditivas e subtractivas e a cor branca sobre um mesmo plano à distância de uma unidade da origem e escalando o hexágono de forma a que a distância entre o ponto representativo de qualquer cor primária e o ponto central correspondente à cor branca seja também unitária, obteremos uma pirâmide hexagonal contendo todas as cores do espaço HSV, tal como a figura 4.9 apresenta.

As cores primárias ocupam no espaço HSV os vértices do hexágono e encontram-se dispostas segundo um ângulo correspondente de acordo com a tabela 4.1.

Cor	Ângulo
Vermelho	0
Amarelo	60
Verde	120
Ciã	180
Azul	240
Magenta	300

Tabela 4.1 – Ângulos das cores no espaço HSV.

A saturação de uma cor corresponde à quantidade de cor branca que a cor apresenta. Uma saturação de 1 significa que a cor é pura (está na periferia), enquanto uma saturação de 0 significa que a cor é totalmente branca e, neste caso, o valor do parâmetro H é irrelevante.

O parâmetro V (valor) corresponde à intensidade da cor e varia entre 0 (intensidade nula, ou seja, cor negra em que os valores de H e S são irrelevantes) e 1 (intensidade máxima). Naturalmente, os tons cinzentos encontram-se localizados sobre o eixo da pirâmide em que $0 \leq V \leq 1$ e $S = 0$, sendo o valor de H indiferente.

4.6 Modelo HLS

O modelo HLS é também um modelo muito intuitivo tal como o modelo HVS. A sigla deste modelo deriva das primeiras letras das palavras em língua inglesa Hue (cor, matiz, cambiante de cor), Lightness (luminosidade) e Saturation (saturação), que são os nomes dos três parâmetros deste modelo.

A representação tridimensional do espaço de cor deste modelo é constituída por dois cones unidos pelas bases, tal como a figura 4.10 apresenta. As cores primárias estão situadas no perímetro da base comum e as cores branca e negra no vértice de cada um dos cones. Os tons cinzentos localizam-se sobre o eixo comum dos dois cones.

Tal como no espaço HSV, o parâmetro H (cor) corresponde ao ângulo em que a cor se encontra, tomando a cor vermelha como origem. O parâmetro S (saturação) tem também o mesmo significado que no modelo HSV. As cores puras têm uma saturação de 1, enquanto que as cores com saturação 0 correspondem a cores brancas de intensidade variável (cinzentos) para as quais o parâmetro H não tem qualquer significado. Finalmente o parâmetro L (luminosidade) pode variar entre 0, a que corresponde uma

luminosidade nula, e 1 que equivale à luminosidade máxima que só é possível para a cor branca. Note-se que as cores puras apresentam uma luminosidade de 0,5.

A pirâmide dupla do espaço HLS pode ser entendida como resultante da pirâmide simples do modelo HSV em que o ponto central da base cone, correspondente à cor branca, foi arrastado para fora, criando deste modo um cone duplo. O cone duplo pode também ser visualizado como uma deformação do cubo RGB em que os vértices das cores primárias aditivas e subtractivas foram projectados num plano perpendicular à diagonal principal do cubo. Este plano localiza-se à distância de 0,5 da origem do espaço. O vértice correspondente à cor branca foi deslocado sobre a diagonal principal de forma a se localizar à distância de uma unidade da origem. Deste modo, poderemos constatar uma maior semelhança entre o cubo do espaço RGB e o duplo cone do espaço HLS do que entre o mesmo cubo e o cone simples do espaço HSV.

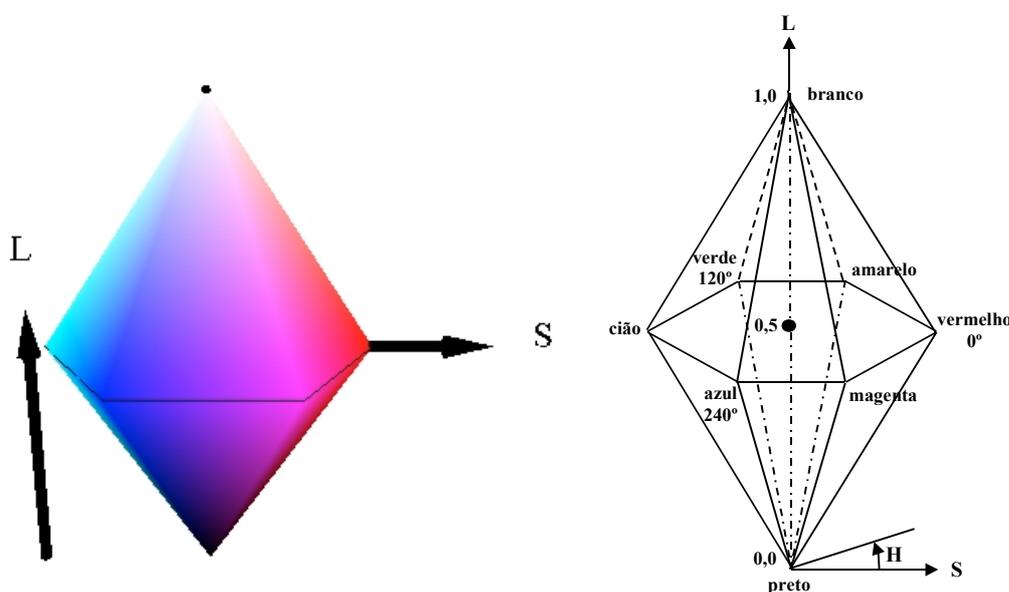


Figura 4.10 – Espaço de cor HLS

4.7 Modelos YIQ e YCbCr (YUV)

Os modelos de cor YIQ e YCbCr foram criados para permitir as emissões dos sistemas de televisão a cores fossem compatíveis com os receptores a preto e branco. O sistema NTSC (National Television Standards Committee), criado em 1953, emprega o modelo YIQ. Por seu lado, os sistemas SECAM (Séquence Electronique Couleur avec Mémoire) e PAL (Phase Alternating Line), de 1961, empregam o modelo YCbCr, definido pela norma CCIR 601-2. Este último modelo é também muitas vezes designado por YUV em Computação Gráfica, no contexto de imagens digitais.

Os dois modelos são muito semelhantes, baseando-se ambos na separação dos sinais de cor RGB em um sinal de luminosidade, ou luminância (Y), e dois sinais de cromaticidade ou diferença de cor. O sinal de televisão correspondente à luminância é transmitido exactamente da mesma forma que o sinal de televisão a preto e branco e, assim, os receptores a preto e branco podiam continuar receber as emissões de televisão

a cores. A definição de luminância consiste na ponderação dos valores das componentes RGB de uma cor por

$$Y = 0,299 R + 0,587 G + 0,114 B \quad (4.7)$$

Note-se a semelhança desta definição com o parâmetro Y do modelo CIE.

No modelo YIQ, os dois parâmetros I e Q contêm a informação correspondente à cor propriamente dita, sendo calculados por diferenças ponderadas entre as componentes vermelha e azul da cor no espaço RGB e a luminância Y tal que

$$\begin{aligned} I &= 0,74 (R - Y) - 0,27 (B - Y) \\ Q &= 0,48 (R - Y) + 0,41 (B - Y) \end{aligned} \quad (4.8)$$

obtendo-se assim a matriz de transformação do espaço de cor RGB para o espaço YIQ

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ 0,596 & -0,275 & -0,321 \\ 0,212 & -0,523 & 0,311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4.9)$$

Desta transformação resulta um espaço YIQ em que Y está compreendido no intervalo $[0;1]$, I em $[-0,596;+0,596]$ e Q em $[-0,523;+0,523]$. Ao parâmetro I corresponde uma gama de variação de cor que vai da cor cião, no seu mínimo, à cor laranja, no seu máximo. Para o parâmetro Q, esta variação verifica-se entre a cor verde e a cor magenta.

O parâmetro Y do modelo YCbCr é idêntico ao parâmetro Y do modelo YIQ. A diferença entre os dois modelos reside na definição dos parâmetros de cromaticidade que, no modelo YCbCr são definidos como as diferenças entre as componentes B e R da cor no espaço RGB e a luminância Y, ou seja

$$\begin{aligned} Cb' &= B - Y \\ Cr' &= R - Y \end{aligned} \quad (4.10)$$

de que resulta uma matriz de transformação. A norma CCIR 601-2 normaliza os intervalos de variação de Cb' e Cr' para um intervalo único, $[-0,5;+0,5]$, pelo que, finalmente, a transformação entre o espaço RGB e o espaço YCbCr é

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ -0,169 & -0,331 & 0,5 \\ 0,5 & -0,419 & -0,0813 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4.11)$$

Modelo		Padrões			
		Vermelho	Verde	Azul	Branco
YIQ	X	0,67	0,21	0,14	0,310
	Y	0,33	0,71	0,08	0,316
YCbCr	X	0,64	0,29	0,15	0,313
	Y	0,33	0,60	0,06	0,329

Tabela 4.2 – Coordenadas CIE dos padrões de cor dos modelos YIQ e YCbCr para calibração de equipamentos de televisão.

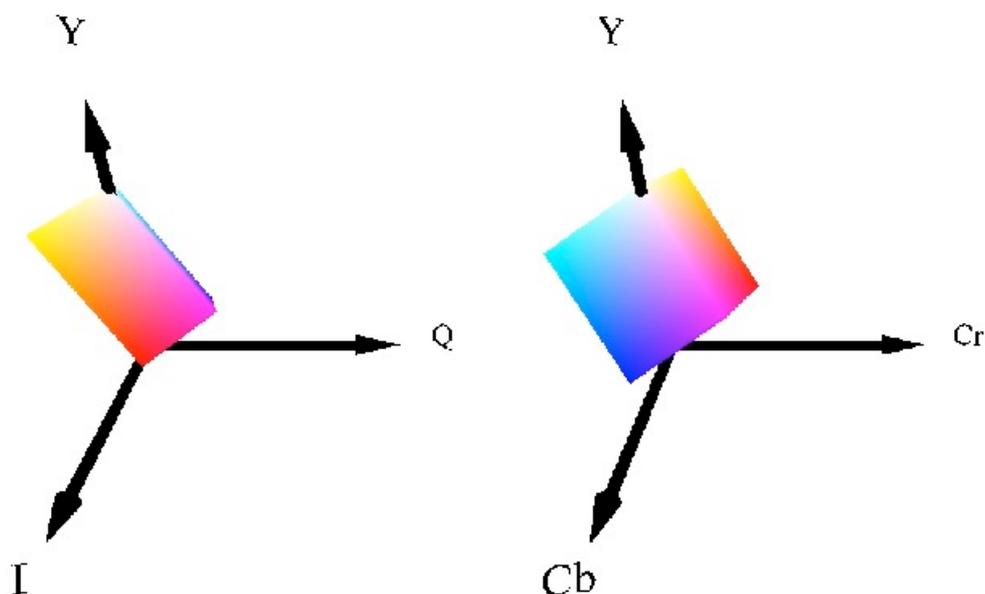


Figura 4.11 – Espaços de cor YIQ e YCbCr.

Os dois modelos de cor diferem ainda nos padrões de cor empregues na calibração dos aparelhos receptores de televisão¹⁰, cuja definição, em coordenadas do modelo CIE, é apresentada pela tabela 4.2.

A figura 4.11 apresenta os espaços de cor dos modelos YIQ e YCbCr que são de forma semelhante (paralelepípedica), em que a cor negra se encontra na origem dos espaços e a cor branca no vértice superior. Os dois modelos diferem, naturalmente, na localização das restantes cores.

O modelo YCbCr, para além da sua aplicação à televisão e ao vídeo, é também empregue por alguns formatos digitais de imagem, dos quais o mais importante é o formato JFIF (também incorrectamente designado por JPEG). A vantagem do emprego deste modelo é, tal como em televisão, a separação da cor em componentes de luminância e de crominância. Como a vista humana é muito mais sensível à luminância do que à crominância, a informação sobre as componentes de crominância de uma dada

¹⁰ Os sistemas de televisão NTSC, SECAM e PAL diferem em ainda mais aspectos do que os do modelo de cor, como número de linhas, frequência das portadoras de sinal, etc.

cor no espaço YCbCr não necessita de uma gama de valores (e precisão) igual à da sua luminância. Se as componentes YCbCr ocuparem originalmente 8 bits cada, é possível então reduzir a representação das componentes de crominância a apenas 2 bits, ou seja, $\frac{1}{4}$ do espaço inicialmente ocupado e a informação sobre uma cor passa a ocupar apenas 12 bits em vez dos 24 bits iniciais. Uma imagem pode assim ser reduzida a metade do seu tamanho original. Esta técnica é designada por *quantização da cor*.

4.8 Escala de Intensidade

Uma característica do olho humano é a sua resposta logarítmica a níveis diferentes de intensidade da luz (luminosidade). Isto permite ao olho humano operar tanto em condições de muito fraca luminosidade como em situações de grande intensidade luminosa. O olho humano é sensível ao quociente entre duas intensidades luminosas e não ao valor absoluto dessas intensidades. Assim, numa escala de intensidade luminosa entre 0 e 1, a diferença entre as intensidades de 0,1 e 0,11 ($0,11/0,1=1,1$) é percebida exactamente como a diferença entre as intensidades de 0,7 e 0,77. Isto significa que as escalas lineares da intensidade luminosa apresentadas pelos modelos de cor têm que ser transformadas em escalas logarítmicas para que o olho humano percepcione correctamente as intensidades pretendidas e as suas diferenças.

Se o quociente entre duas intensidades consecutivas numa escala a definir for

$$I_{j+1} = r I_j \quad (4.12)$$

O parâmetro r é o coeficiente da progressão geométrica da escala de intensidade tal que

$$r = \frac{I_{j+1}}{I_j} \quad (4.13)$$

O valor máximo de uma escala de intensidade normalizada é 1. O valor mínimo deveria ser 0 mas, na prática, tal não é possível devido aos fenómenos de dispersão e reflexão da luz que se verificam nos dispositivos de saída gráfica que emitem luz. A intensidade mínima que se consegue obter em monitores varia entre 0,005 e 0,025. Esta intensidade será representada por I_0 . Poderemos então construir uma escala de intensidade tal que

$$\begin{aligned} I_0 &= I_0 \\ I_1 &= rI_0 \\ I_2 &= rI_1 = r^2I_0 \\ &\dots \\ I_{n-1} &= rI_{n-2} = r^{(n-1)}I_0 \end{aligned} \quad (4.14)$$

Como, numa escala com n valores de intensidade, $I_{n-1} = 1$ porque se trata da intensidade máxima, obteremos

$$\begin{aligned} r &= \left(1/I_0\right)^{\frac{1}{n-1}} \\ I_j &= r^j I_0 = I_0^{\frac{n-1-j}{n-1}} \end{aligned} \quad (4.15)$$

Nº de níveis	I_0	Gama dinâmica, $1/I_0$	r	Níveis
4	0,125	8	2	0,125; 0,25; 0,5; 1
4	0,0625	16	2,5198421	0,0625; 0,15749; 0,39685; 1
256	0,02	50	1,01545948	0,02; 0,0203; 0,0206; ...; 0,94848; 1

Tabela 4.3 – Níveis de intensidade para valores diferentes da intensidade mínima (I_0) e de r

O quociente entre o valor máximo e o valor mínimo de uma escala de intensidade assim construída, $1/I_0$, é característico do dispositivo de saída gráfica e designa-se por *gama dinâmica do dispositivo* (“device dynamic range”). A tabela 4.3 apresenta os valores das intensidades para três escalas de intensidade com os valores de r e I_0 .

Como vimos, o valor de I_0 depende do dispositivo de saída gráfica. Por seu lado, o valor de r depende exclusivamente da capacidade do olho humano distinguir entre duas intensidades contíguas na escala de intensidade. O valor de r limite será então aquele para o qual o olho humano deixar de distinguir duas intensidades contíguas da escala como diferentes. A figura 4.12 apresenta uma fotografia a preto e branco com 231 níveis diferentes de intensidade e a reprodução da mesma fotografia com apenas 64 níveis de cinzento, mostrando que não é possível notar quaisquer diferenças entre as suas imagens. A mesma fotografia encontra-se reproduzida na figura 4.13 com 4, 8, 16 e 32 níveis de cinzento, notando-se diferenças apreciáveis para as imagens com menor número de níveis. No entanto estas diferenças são quase imperceptíveis na imagem com 32 níveis de cinzento o que leva a sugerir que o número mínimo de níveis deverá ser de 64 níveis.



(a)

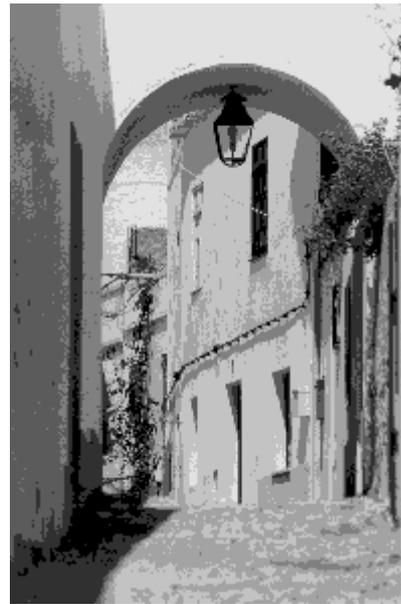


(b)

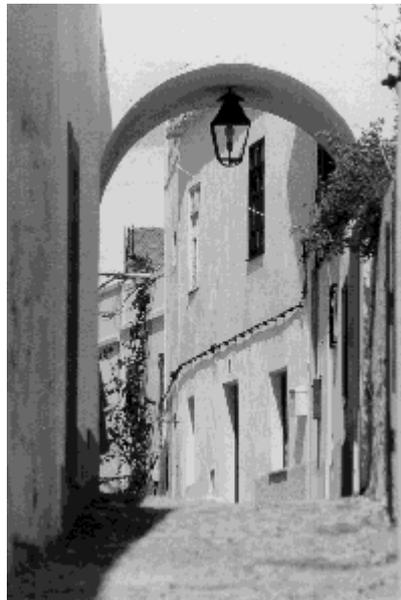
Figura 4.12 – Fotografia original (a) e respectiva reprodução com apenas 64 níveis de intensidade de cinzentos (b).



(a)



(b)



(c)



(d)

Figura 4.13 – Reprodução da fotografia original da figura 4.12 com 4 (a), 8 (b), 16 (c) e 32 (d) níveis de cinzento.

Experiências realizadas mostraram que o valor de r para o qual o olho humano deixa de distinguir diferenças entre intensidades sucessivas de uma escala de intensidades se encontra próximo de 1,01. A tabela 4.4 apresenta valores típicos da gama dinâmica de vários dispositivos de saída gráfica e o correspondente número mínimo de níveis de intensidade para que, com um valor de r igual a 1,01, se obtenham escalas de intensidade que possam produzir a sensação de continuidade nos observadores.

Tipo de media	Gama dinâmica ($1/I_0$)	Nº mínimo de níveis
Monitor	40-200	372-533
Fotografia em papel	100	464
Fotografia em diapositivo	1000	695
Papel revestido (preto e branco)	100	464
Papel revestido (cor)	50	394
Jornal (preto e branco)	10	232

Tabela 4.4 – Gama dinâmica e número mínimo de níveis de intensidade para diferentes tipos de media

Note-se as diferenças entre a fotografia impressa em papel fotográfico, imagens em papel comum e imagens em papel de jornal.

4.9 Correção Gama

Como vimos anteriormente, a construção de uma escala de intensidade é um processo simples uma vez que se conheçam a intensidade mínima e o número de níveis de intensidade pretendidos. No entanto, a resposta não linear característica dos tubos de raios catódicos dos monitores, e também dos registadores de filmes de diapositivos, distorcem a escala construída. Torna-se então necessário corrigir esta distorção. Para isso, iremos analisar o funcionamento de um tubo de raios catódicos de um monitor a cores (as consequências para um deles aplicam-se aos restantes). Em monitores, a intensidade da luz emitida por um tubo de raios catódicos depende do número de electrões que incidem no fósforo do tubo, sendo esta relação

$$I = k N^\gamma \quad (4.16)$$

em que N é o número de electrões que incidem no fósforo e k e γ duas constantes características de cada tubo de raios catódicos. A constante γ é, como veremos, a mais importante. O seu valor varia tipicamente entre 2,2 e 2,5.

O número de electrões do feixe que incide na matriz de fósforo do tubo é linearmente dependente da tensão aplicada ao controlo do feixe de electrões e esta é função linear do valor de controlo, X . Combinando então as constantes desta função linear numa só constante, K , (4.16) transforma-se em

$$I = K X^\gamma \quad (4.17)$$

Para obter a intensidade I_j da escala de intensidade definida é necessário que o valor de controlo seja

$$X_j = (I_j / K)^{1/\gamma} \quad (4.18)$$

Assim, para que o nível de intensidade de ordem j seja representado com a intensidade correcta no ecrã, o tubo de raios catódicos deve receber um sinal de controlo com valor

de X_j correspondente à intensidade I_j pretendida e não pura e simplesmente o valor de I_j .

Esta transformação da intensidade pretendida no valor correcto de controlo do tubo de raios catódicos designa-se por *correção gama*. O nome tem origem no nome da letra grega do expoente da curva característica de resposta dos tubos de raios catódicos.

Se o monitor ou a placa gráfica realizarem a correção gama automaticamente, bastará fornecer o valor da intensidade pretendida.

Que consequências tem não realizar a correção gama? Como as intensidades são especificadas por valores normalizados compreendidos entre 0 e 1, os valores que resultam da sua elevação a um expoente positivo maior do que 1 serão inferiores. Um monitor apresentará então cada componente RGB mais escura do que deveria ser. Tomemos como exemplo um píxel cujas componentes RGB deverão ter as intensidades (0,8; 0,2; 0,2). Com $\gamma = 2,5$ e se não se aplicar a correção gama, as intensidades realmente apresentadas no ecrã corresponderão ao trio (0,5724; 0,0179; 0,0179). Não só cada uma das componentes RGB resulta mais escura do que o pretendido, como a relação entre os valores das componentes, que era inicialmente (4:1:1), passou a ser (32:1:1). Além do escurecimento da imagem houve igualmente uma mudança da tonalidade da cor original. No caso presente, a cor original que possuía uma tonalidade avermelhada, passou a ter uma tonalidade castanha, tal como a figura 4.14 apresenta. Na mesma figura, à direita, pode-se observar o resultado da aplicação de uma correção com $\gamma=1,7$, um valor muito comum na maioria das placas gráficas existentes.

Como atrás ficou dito, a correção gama depende das características específicas de cada monitor. Dada a grande escolha de monitores existente para computadores IBM-PC não existe qualquer norma para realizar a correção gama neste tipo de equipamentos, embora existam muitas placas gráficas que a realizam automaticamente, quer pelo conhecimento do modelo do monitor, quer ainda de forma cega aplicando uma correção com um valor do expoente gama arbitrado.

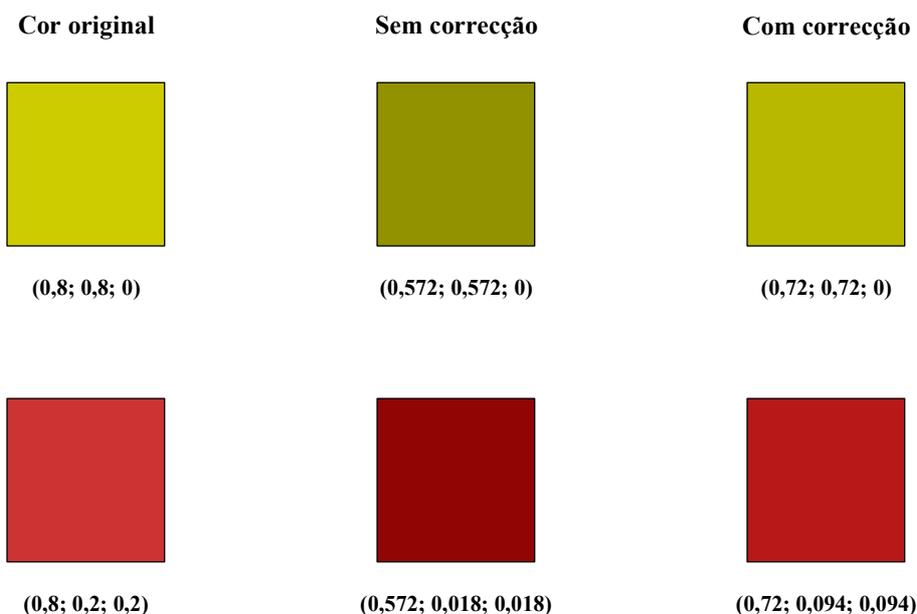


Figura 4.14 – As cores sem correção gama são mais escuras que as cores originais. Uma correção com $\gamma=1,7$ recupera parte da intensidade e tonalidade original das cores.

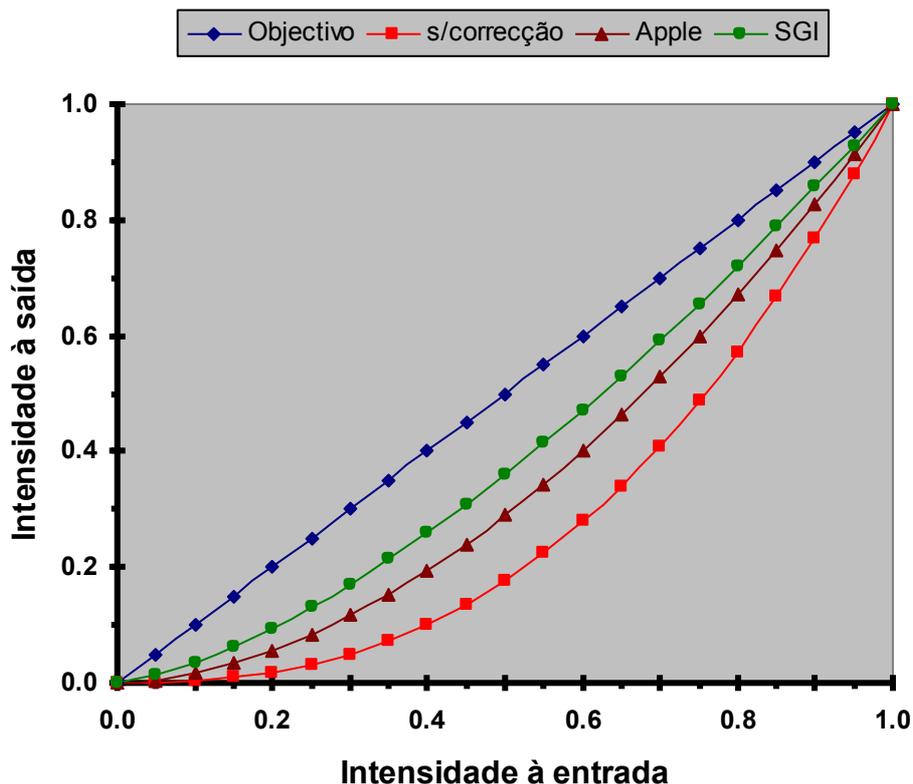


Figura 4.15 – Correção gama automática dos equipamentos Apple (1,4) e SGI (1,7). As intensidades corrigidas encontram-se entre as intensidades não corrigidas ($\gamma=2,5$) e o objectivo (valor totalmente corrigido).

Para os equipamentos SUN, não existe também nenhuma norma mas, dependendo da placa gráfica, é possível ajustar a correção gama através de programas utilitários do sistema operativo. O equipamento SGI apresenta uma correção gama por omissão com um expoente igual a 1,7 que pode também ser alterado por utilitários do sistema. Por seu lado, os equipamentos Apple aplicam automaticamente uma correção gama com um expoente gama cujo valor é de 1,4. A figura 4.15 apresenta os efeitos da correção gama aplicada automaticamente por equipamentos Apple e SGI no caso de um monitor com $\gamma=2,5$. O objectivo (valor totalmente corrigido) encontra-se representado pela linha de resposta linear ideal em que a intensidade reproduzida é a mesma do que a especificada à entrada. Quando não é realizada qualquer correção, as intensidades obtidas no monitor serão mais escuras do que o pretendido. As correções gama aplicadas pelos equipamentos Apple (1,4) e SGI (1,7) encontram-se entre as duas curvas de resposta anteriores e permitem obter valores de intensidade mais próximos dos valores pretendidos.

Da aplicação ou não da correção gama com valores variáveis do expoente resulta que uma imagem pode ser apresentada com cores correctas numa plataforma e com cores erradas noutra plataforma diferente. Este problema tornou-se mais notado com o advento da World Wide Web e realçou a necessidade de se encontrar um compromisso que permitisse a apresentação da mesma imagem com idêntica qualidade em plataformas diferentes. A solução de compromisso mais simples que consiste na aplicação de uma correção intermédia às imagens antes da sua publicação na WWW não produziu os resultados esperados devido à dispersão dos expoentes empregues e

contribuiu para o agravamento do problema. Uma outra solução consistiu na proposta de um novo formato de imagem, o formato *PNG* (Portable Network Graphics), contendo informação sobre a correcção gama já aplicada às imagens. No entanto, este formato continua a ser pouco utilizado devido ao enorme número de imagens já publicadas nos formatos mais populares suportados pelos navegadores (GIF e JPEG).

5 Meios-Tons e Tramas

Existem actualmente dispositivos de saída gráfica que permitem apenas dois níveis de intensidade. Uns, como os monitores monocromáticos, encontram-se quase em desuso, enquanto novos equipamentos nestas condições, como os PDA (Personal Data Assistants), telemóveis e telecomandos com visores gráficos, começam a invadir o mercado. Incluem-se também nesta categoria equipamentos como as impressoras a preto e branco e outros que, embora permitam mais do que dois níveis de intensidade, estão limitados a profundidades de 2 ou 3 bits por píxel e, portanto, não permitem a reprodução das intensidades intermédias entre as intensidades disponíveis por forma a obter um número total de intensidades próximo do número de intensidades mínimo, que, como vimos, é de cerca de 64. As impressoras as cores, quer por jacto de tinta quer a laser, apresentam também este problema que, na ausência do emprego de técnicas apropriadas, as impede de empregar o número mínimo de níveis de intensidade (e cor) necessário para obter imagens com fidelidade de cor aceitável.

Para resolver este problema desenvolveram-se as técnicas de “*meios-tons*” ou de *tramas* ou ainda de *pontos agrupados ordenados*. Estas técnicas baseiam-se no fenómeno de integração espacial do olho humano que não consegue distinguir pontos individuais aglomerados em pequenas áreas quando essas áreas são visionadas a distância suficientemente grande. Uma pequena área constituída por pontos negros e brancos será então percebida com uma intensidade cinzenta uniforme. A mesma área será percebida como sendo de cor amarela se os seus pontos forem alternadamente verdes e vermelhos.

Estas técnicas são exaustivamente aplicadas em jornais por meio de padrões (ou tramas) distintas com forma variável e cujo número pode atingir 60 a 80 formas. Aplicam-se também na impressão de revistas e livros, se bem que neste caso o problema seja menos grave do que no caso de jornais.

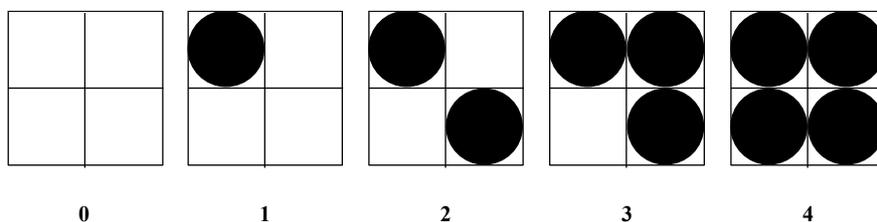


Figura 5.1 – Tramas para áreas com 2x2 píxeis.

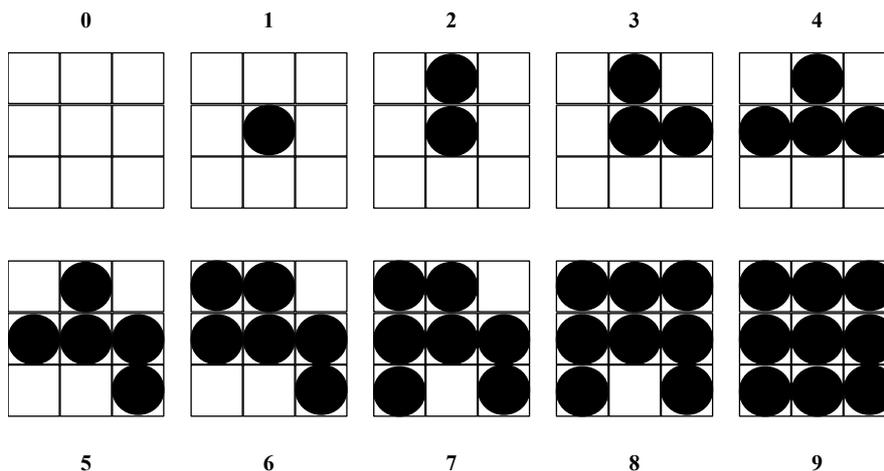


Figura 5.2 – Conjunto de tramas para áreas com 3x3 píxeis. A figura apresenta uma solução possível. Outras soluções são igualmente possíveis, desde que sigam as regras para a construção de grupos de tramas.

Fundamentalmente, as técnicas de meios-tons consistem no agrupamento de vários píxeis contíguos de uma imagem numa única área e na distribuição de pontos negros e brancos nessa área. O número de píxeis agrupados determina o número de níveis de intensidade atribuível a uma dessas áreas. Uma área de 2x2 píxeis permite 5 níveis de intensidade, enquanto uma área de 3x3 píxeis permitirá que esse número se eleve para 10. Em geral, uma área de nxn píxeis permite dispor de n^2+1 níveis de intensidade, estando incluída neste número a intensidade mínima (todos os pontos são negros) e a intensidade máxima (todos os pontos são brancos)¹¹. A figura 5.1 apresenta as tramas correspondentes às cinco intensidades possíveis com uma área de 2x2 píxeis. A figura 5.3 mostra o resultado da aplicação das tramas da figura 5.1 à fotografia original apresentada na figura 4.12. Se se comparar a imagem da figura 5.3, que apresenta 5 níveis de intensidade, com a imagem (a) da figura 4.13, que apresenta 4 níveis, verifica-se que a maior diferença entre as duas imagens reside na perda de resolução da primeira imagem em relação à segunda.

O preço desta técnica é a diminuição da resolução das imagens a reproduzir. Com efeito, a aglomeração em áreas de nxn píxeis em píxeis únicos tem a consequência de diminuir a resolução das imagens n vezes, tanto na horizontal como na vertical. Além disso, a substituição da intensidade de cada píxel original por uma intensidade média de todos os píxeis de uma dada área produz uma integração espacial da intensidade que poderá resultar numa sensação de zonamento. Convém portanto poder partir de imagens com resolução n vezes superior à resolução com que se pretende apresentar essas imagens, mas nem sempre isto é possível.

¹¹ Note-se que se os píxeis se encontrarem igualmente espaçados horizontal e verticalmente, uma área de nxn píxeis tanto pode ser considerada como tendo os lados orientados segundo a horizontal e a vertical, ou fazendo ângulos de 45° com estes.



Figura 5.3 – Reprodução da fotografia da figura 4.12(a) obtida à custa de tramas para áreas de 2x2 píxeis. Note-se a degradação da resolução da imagem obtida.

5.1 Regras para a Construção de Tramas

Os padrões (ou tramas) não devem introduzir nas imagens efeitos visuais, tais como linhas horizontais. Por esta razão a trama correspondente a uma intensidade de 50% apresentada na figura 5.1 tem os pontos de igual intensidade segundo a diagonal e não na horizontal. Se esta última tivesse sido a solução escolhida, qualquer zona com a intensidade de 50% apareceria com o aspecto de riscada por linhas horizontais. Ao colocarmos os pontos de igual intensidade sobre uma diagonal, evitamos este efeito de continuidade dado que, segundo a diagonal, os pontos se encontram mais afastados.

Existe um conjunto de regras a seguir na construção de padrões e que são

- Os padrões não devem introduzir efeitos (artefactos) visuais em áreas contíguas de igual intensidade (linhas horizontais, por exemplo).
- Se um ponto de um padrão estiver activo (“aceso”) para um dado nível de intensidade, esse ponto deverá manter-se activo em todos os níveis com intensidade superior. Esta regra destina-se a evitar efeitos de zonamento através da minimização das diferenças entre padrões de intensidades sucessivas.
- Os pontos activos devem crescer do centro do padrão para a periferia, para evitar a sensação do crescimento dos pontos.
- Os pontos activos devem ser contíguos. Esta regra destina-se a evitar os problemas verificados em impressoras laser e registadores de filme que têm dificuldade em imprimir (ou sensibilizar) pontos isolados. Os dispositivos de saída gráfica baseados em tubos de raios catódicos não necessitam seguir esta regra dado que estes dispositivos não têm quaisquer problemas em representar pontos (píxeis) isolados.

Estas regras definem a ordem pela qual os pontos dos padrões correspondentes a intensidades consecutivas devem ser activados e a respectiva aglomeração. É esta a

razão pela qual estes padrões (ou tramas) são chamadas de “pontos agrupados ordenados”. No caso de dispositivos gráficos de tubos de raios catódicos são possíveis soluções com padrões apresentando pontos activos isolados. Neste caso, os padrões são do tipo de “pontos isolados ordenados”.

A figura 5.2 apresenta uma sequência de tramas com a dimensão de 3×3 píxeis construída de acordo com as regras enunciadas. Na figura 5.4 apresentam-se duas tramas que violam a 1ª regra (introdução de efeitos visuais espúrios) e a última regra (não contiguidade dos pontos), respectivamente.

5.2 Matrizes de Tramas

Designam-se por matrizes de tramas as matrizes cujos elementos indicam a ordem pela qual os pontos das tramas devem ser activados. Das figuras 5.1 e 5.2 resulta que as respectivas matrizes, genericamente identificadas por $D^{(n)}$, são, segundo Baer,

$$D^{(2)} = \begin{bmatrix} 0 & 2 \\ 3 & 1 \end{bmatrix} \qquad D^{(3)} = \begin{bmatrix} 6 & 8 & 4 \\ 1 & 0 & 3 \\ 5 & 2 & 7 \end{bmatrix} \qquad (5.1)$$

Júdice e outros desenvolveram a seguinte relação de recorrência para a geração da matriz $D^{(n)}$ a partir da matriz $D^{(n/2)}$ e da matriz de ordem n, U^n , cujos elementos são todos unitários,

$$D^{(n)} = \begin{bmatrix} 4D^{(n/2)} + D_{00}^{(2)}U^{(n/2)} & 4D^{(n/2)} + D_{01}^{(2)}U^{(n/2)} \\ 4D^{(n/2)} + D_{10}^{(2)}U^{(n/2)} & 4D^{(n/2)} + D_{11}^{(2)}U^{(n/2)} \end{bmatrix} \qquad (5.2)$$

e em que $D_{ij}^{(2)}$ é o elemento da linha i, coluna j da matriz de padrões 2×2.

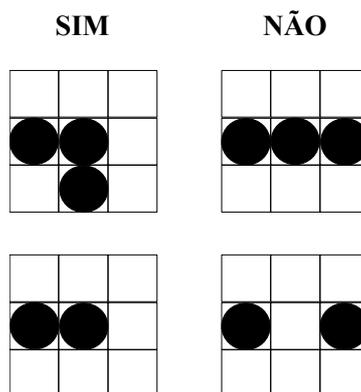


Figura 5.4 – Violação das regras para a construção de tramas: introdução de efeitos visuais indesejados (em cima) e não adjacência dos pontos (em baixo).

5.3 Tramas em Dispositivos com Mais do que Duas Intensidades

As considerações anteriores dizem respeito a tramas para utilização em dispositivos de saída gráfica com apenas dois níveis de intensidade possíveis, os dispositivos monocromáticos. Existem outros dispositivos em que o emprego de tramas se torna necessário para aumentar o número de intensidades reduzido que permitem. São os casos de ecrãs de tubos de raios catódicos com pequena profundidade (2 a 3 bits por pixel) ou das impressoras a cores de tecnologia laser ou de jacto de tinta¹². Nestes casos existem m níveis de intensidade (ou cor), em que $m > 2$, a aplicar em padrões $n \times n$.

Para profundidades de 2 bits e uma trama de 2×2 píxeis, existem 4 píxeis na trama que podem assumir 4 níveis de intensidade e, portanto, teremos $4 \times 3 + 1 = 13$ intensidades disponíveis. Em impressoras a cores a profundidade em bits é substituída pelo número de gotas de tinta que podem ser depositadas em cada píxel. Com duas gotas e três cores obtêm-se oito cores diferentes. Isto demonstra que as impressoras a cores necessitam empregar tramas quando o número de gotas que é possível depositar em cada píxel é reduzido.

5.4 Emprego de Tramas sem Perda de Resolução

Em geral, o emprego de tramas implica perda de resolução da imagem reproduzida em relação à imagem original. É no entanto possível produzir imagens de meios tons sem perda de resolução fazendo apelo a técnicas que utilizam as tramas como máscaras para decidir se cada píxel da imagem final estará activo ou não.

Para simplificar a exposição, consideremos uma imagem original a tons de cinzento com k níveis de intensidade função das coordenadas x, y da imagem, $O(x,y)$, e uma trama $n \times n$ que permite $m = n^2 + 1$ níveis de intensidade. A técnica consiste em três passos

1º passo – A intensidade original é transformada para a escala de intensidades permitida pela trama

$$T(x, y) = f(O(x, y), k, m)$$

2º passo – Determina-se o elemento i, j da trama em função das coordenadas x, y de cada ponto da imagem, tal que

$$i = \text{resto de } x/n$$

$$j = \text{resto de } y/n$$

3º passo – Se $T(x, y) > D_{ij}^n$, o ponto da imagem resultante é activado, senão não o é.

A figura 5.5 apresenta três imagens em que a resolução da imagem original foi mantida pelo emprego desta técnica usando tramas 2×2 , 4×4 e 8×8 . É útil comparar as diferenças entre estas imagens e as imagens (a) e (c) da figura 4.13 e a imagem (b) da figura 4.12. A resolução é a mesma, mas os resultados em termos da mancha obtida são de pior qualidade. No entanto, se compararmos as imagens de figura 5.5 e da figura 5.3,

¹² A expressão “impressora a gotas de tinta” designa mais realisticamente o funcionamento das impressoras a jacto de tinta.

verificaremos que a imagem da primeira apresenta maior qualidade, exactamente como era pretendido.

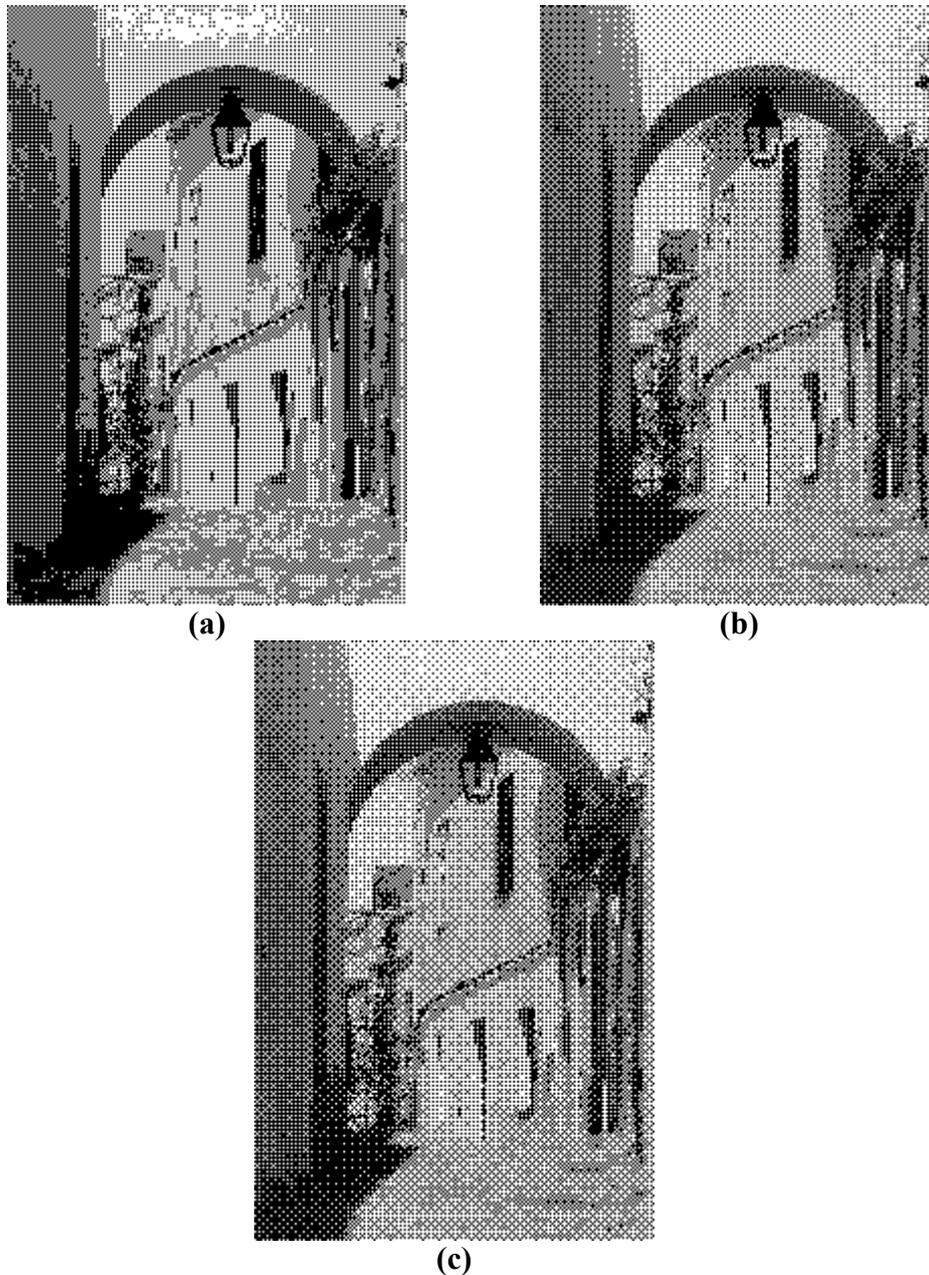


Figura 5.5 – Reprodução da fotografia original da figura 4.12 mantendo a resolução com tramas 2x2 (a), 4x4 (b) e 8x8 (c).

6 Reprodução da Cor

A fidelidade da reprodução da cor é uma exigência natural. Tal como a perda intermitente de partes de palavras numa conversa telefónica impede a percepção correcta da conversa, assim a reprodução da cor em diferentes dispositivos poderá

resultar em cores diferentes das cores originais ou cores que o cérebro humano interpreta como cores não naturais quando se esperaria o contrário¹³. Existem ainda situações em que a reprodução exacta da cor é fundamental como é o caso da construção de arquivos digitais de pinturas que devem permitir recuperar todas as cores por ocasião de restauros.

Os modelos RGB e CMY são os modelos mais utilizados em dispositivos de saída gráfica como monitores e impressoras a cores. Tal como vimos anteriormente, estes modelos não definem que comprimentos de onda correspondem às cores primárias de cada modelo. Isto abre o caminho para a existência de diferenças entre reproduções de imagens produzidas por dispositivos diferentes. Por outro lado, as imagens apresentadas em monitores sofrem dos efeitos de correcções gama incompletas na generalidade dos casos. Existem pois diferenças entre a cor observada nos monitores e a cor apresentada nas imagens impressas em papel, por exemplo. A variação da cor da luz ambiente tem também influência nestes casos.

Para além dos problemas devidos à correcção gama, a reprodução em diapositivos e filmes apresenta problemas próprios. As cores finais obtidas nestes dispositivos dependem ainda da cor da luz emitida pelas unidades de registo, dos filtros de correcção que são empregues, do tipo de filme (propriedades químicas, grão, etc.) e da qualidade e tipo dos reagentes empregues na revelação e fixação, além da temperatura e da duração destas operações. No caso da fotografia sobre papel, os parâmetros de qualidade do papel como textura, brilho e absorção, influenciam igualmente a qualidade dos resultados¹⁴.

A impressão de imagens a cores sobre papel apresenta grandes dificuldades na obtenção de cores fiéis às cores originais, principalmente no caso de imagens fotográficas em que a qualidade dos resultados depende do controlo do processo de impressão (regulação do tempo e do caudal dos jactos de tinta ou “toner”), da qualidade das tintas (ou “toners” coloridos) empregues e as propriedades dos papéis sobre os quais é feita a impressão (textura, absorção, brilho, etc.).

A impressão a cores segundo o modelo CMY processa-se pela deposição no papel de três tintas correspondentes às cores primárias substractivas. Para maior fidelidade, existe uma quarta tinta de cor negra¹⁵. Além da maior fidelidade da cor preta assim obtida, a secagem da tinta depositada será mais rápida porque a quantidade de tinta depositada é menor. Se as impressoras de jacto de tinta pudessem depositar uma quantidade variável de tinta em cada ponto das imagens a reproduzir, estas impressoras poderiam imprimir um elevado número de cores semelhante ao número de cores disponíveis num monitor a cores.

Na realidade, a quantidade de tinta que uma impressora a cores consegue depositar varia discretamente, isto é, a impressora só consegue depositar no papel um número exacto de gotas de tinta, estando limitado o número total de gotas depositáveis em cada ponto. Assim, estas impressoras não podem reproduzir todas as cores que as imagens a imprimir possam possuir e, conseqüentemente, as imagens impressas apresentarão cores

¹³ Numa paisagem, por exemplo.

¹⁴ É comum verificarem-se diferenças nítidas de cor entre duas provas em papel obtidas em momentos diferentes a partir do mesmo negativo.

¹⁵ Porque a absorção da luz não pode na prática ser total, a deposição das três cores primárias substractivas num ponto produzirá sempre um tom cinzento-escuro, mas nunca um verdadeiro tom preto.

que não correspondem às cores originais, podendo inclusivamente apresentar cores pouco realistas que degradam a qualidade das imagens no papel.

Existem dois tipos de estratégias para ultrapassar estes problemas:

- a técnica de tramas policromáticas, usada principalmente na impressão a cores e que obriga à diminuição da resolução
- a técnica de manipulação das tabelas de cor que mantém a resolução e privilegia as cores dominantes das imagens ou zonas das mesmas.

6.1 Algoritmos para Redução do Número de Cores

O problema a resolver consiste em reproduzir uma imagem com uma profundidade original de n bits (2^n cores) num dispositivo cuja profundidade inferior é de m bits (2^m cores) quando $m < n$. O algoritmo desta solução terá que resolver dois problemas:

- selecção das 2^m cores que constituirão o mapa de cor da imagem a apresentar
- como mapear as cores da imagem original nas 2^m cores do mapa de cor da imagem

O algoritmo mais simples consiste em arbitrar um mapa de cor para a imagem segundo um mapeamento que seja apenas função da profundidade disponível. Este algoritmo deve ter em conta a menor discriminação do olho humano para os tons azuis. Tomemos por exemplo o caso de uma imagem com uma profundidade de 18 bits a ser representada num dispositivo de saída que permite apenas 8 bits de profundidade. A imagem original reserva um número igual de bits (6) para cada uma das componentes RGB, com 64 intensidades para o vermelho, 64 para o verde e 64 para o azul. Não é possível dividir igualmente 8 bits por três cores, mas podemos arbitrar 3 bits para o vermelho e 3 bits para o verde e apenas 2 bits para o azul na imagem convertida, tendo em conta a menor discriminação na zona do azul. O resultado é que, tanto para o vermelho como para o verde, as 64 intensidades originais serão reduzidas a 8 intensidades, mas as 64 intensidades de azul passarão a ser representadas por apenas 4 intensidades.

O mapeamento de cores corresponderá à substituição das cores originais pelas cores mais próximas no novo mapa de cores que foi arbitrado. Cores distintas mas próximas da imagem original poderão aparecer na imagem processada mapeadas na mesma cor, sem a garantia de que a cor mais próxima seja semelhante às cores originais. Além disto, dependendo da frequência das cores na imagem original, poderá acontecer que muitas das cores do mapa de cor da imagem transformada não sejam utilizados pela nova imagem. O problema deste algoritmo não se encontra no mapeamento de cores mas sim na definição do mapa de cor que, na maioria das vezes, não é o mais adequado para a imagem em causa. O mapa de cor é gerado pelo algoritmo independentemente das cores existentes na imagem a transformar e da respectiva frequência. É necessário então substituir este algoritmo de mapa de cor único por um algoritmo que apresente um mapa de cor adaptado a cada imagem.

Existem duas famílias de algoritmos adaptativos para a geração de mapas de cor que são:

- Algoritmos de popularidade da cor
- Algoritmos de subdivisão pela mediana

Os algoritmos adaptativos de popularidade baseiam-se na manutenção na imagem transformada das cores mais frequentes da imagem original e na substituição das cores menos frequentes pelas cores mais próximas existentes no mapa de cor da imagem transformada. O algoritmo começa por construir um histograma das cores empregues em cada imagem a transformar e, seguidamente, constrói um mapa de cor com as 2^m cores mais frequentes na imagem original. Os píxeis da imagem original que possuem estas cores serão apresentados na imagem de cor reduzida exactamente com a mesma cor que tinham originalmente, enquanto as cores restantes serão convertidas para a cor do conjunto de 2^m cores que lhe fiquem mais próximas¹⁶. Este algoritmo apresenta as vantagens da sua simplicidade e rapidez.

Os algoritmos adaptativos por subdivisão pela mediana baseiam-se no efeito de integração espacial do olho humano que, ao observar zonas muito próximas, tem tendência a interpretar estas zonas como uniformes. A parte fundamental do algoritmo de subdivisão consiste em determinar essas zonas e atribuir-lhes uma cor predominante. Para isto o algoritmo subdivide sucessivamente uma região da imagem segundo a sua maior dimensão pelo respectivo valor mediano até que existam 2^m regiões e arbitra como cores da imagem de cor reduzida as cores dos centróides de cada uma das regiões. Os algoritmos de subdivisão são mais lentos que os algoritmos de popularidade, mas produzem resultados de maior qualidade. Existem outras estratégias para a subdivisão das regiões, como a que é apresentada por Wan, Wong e Prusinkiewicz¹⁷.

7 Emprego da Cor em Computação Gráfica

O emprego da cor em Computação Gráfica destina-se a múltiplos fins ditados pelos objectivos de cada utilização. A cor desempenha uma função *estética*, isto é, as cores empregues devem transmitir ao observador sensações estéticas que agradem ao observador, não o afastem e ainda lhe facilitem a tarefa de observação. A cor deve criar no observador um estado de espírito tal que predisponha o observador (ou utilizador) a realizar as tarefas a que se propôs. É claro que a noção de estética é subjectiva e dependente da cultura e, por isso, o emprego da cor deve ser planeado em função das preferências estéticas dos observadores sem descuidar os fins daquilo que se pretende transmitir através de imagens ou ecrãs contendo cor.

A cor destina-se igualmente a atrair a atenção do observador para a informação e a realçar conteúdos importantes e natureza dessa informação. A informação em causa poderá ser um texto importante, um aviso sob forma gráfica que emprega um símbolo que a cor faz destacar ou uma zona de uma imagem com cores falsas para realçar a conformidade ou inconformidade com determinados objectivos dos objectos que estão a ser visualizados¹⁸.

A cor deve também proporcionar a transmissão de estados emocionais e estados de espírito. A escolha de cores na pintura exemplifica esta capacidade de transmissão em que as cores suaves transmitem estados de espírito relacionadas com o bem-estar

¹⁶ A distância entre duas cores pode ser calculada como o módulo do vector que une os pontos representativos das duas cores no espaço RGB, por exemplo.

¹⁷ Wan, S., Wong, K. e Prusinkiewicz, P., An Algorithm for Multidimensional Data Clustering, ACM Trans. Mathematical Software, Vol. 14, no. 2, June 1988, pp 153-162.

¹⁸ Como uma zona de uma peça que vai quebrar, por exemplo.

(inclusivamente contemplativos) e as cores escuras e sóbrias significam ambientes formais¹⁹. O emprego de cores vivas transmite emoções fortes e, quanto maior for a variedade de cores fortes empregue, tanto maior será a sensação de alegria e festa transmitida.

A cor destina-se também a proporcionar realismo, um aspecto importante na geração de imagens em Computação Gráfica, principalmente na geração de ambientes virtuais como, por exemplo, o estudo de edifícios ainda em projecto em que se pretende verificar a estética visual do edifício e as suas condições de iluminação (natural e artificial, etc.), e a recriação de ambientes tanto actuais como passados e em que o observador se desloca nos espaços entre edifícios, quer para explorar, quer para se ambientar. Outros exemplos são os estudos ergonómicos de painéis de comando de viaturas automóveis e simuladores de voo, onde os cenários também são importantes. Em todos estes casos, as cores apresentadas devem reproduzir o mais fielmente possível as cores naturais sob pena de, se o não fizerem, transmitirem informação cromática que os observadores interpretem erradamente, como pode acontecer em estudos de iluminação, ou como falsa, o que cria situações de rejeição dos ambientes virtuais pelos observadores. Como é óbvio, a necessidade de realismo existe também na reprodução de imagens que devam garantir cores fiéis às cores originais nas reproduções.

A cor destina-se igualmente a *associar áreas* da imagem e a *destacar a organização* dos elementos contidos na imagem e o seu significado. Assim, vários ícones dispostos sobre um fundo comum (da mesma cor ou padrão) são interpretados como um grupo de ícones associados com funções semanticamente afins (criar, abrir e guardar documentos). Em organigramas organizacionais, um conjunto de nomes de serviços com letra da mesma cor em que os nomes estejam inscritos em caixas de cor igual é imediatamente interpretado como fazendo parte do mesmo departamento ou direcção de serviços. Neste caso, os departamentos e o seu número e funções são muito rapidamente apreendidos se as cores dos serviços de cada departamento apresentarem cores próprias e distintas umas das outras.

A cor permite a transmissão de mensagens de informação, aviso, perigo, etc. O emprego da cor nesta codificação deve ser feito com cuidado dada a dependência que o significado da cor tem da cultura. Uma cor pode ter significados diferentes em culturas diferentes, como é o caso da cor negra cujo significado é negativo na cultura ocidental, mas pode assumir um significado totalmente oposto em outras culturas. A tabela 7.1 apresenta os significados mais comuns das cores para a cultura ocidental. Os valores percentuais apresentados pela tabela indicam com que frequência esta cultura atribui a cada uma das cores o significado correspondente. Note-se que existem cores como o vermelho, por exemplo, cujo significado depende do contexto.

¹⁹ Veja-se o caso de “A Ronda da Noite” (Nachtwache) ou a “Lição de Anatomia” de Rembrandt.

Conceito	Cor	%
Parar	Vermelho	100
Avançar	Verde	99
Frio	Azul	96
Quente	Vermelho	95
Perigo	Vermelho	90
Precaução	Amarelo	81
Segurança	Verde	61
Ligado	Vermelho	50
Desligado	Azul	32

Tabela 7.1 – Significados mais correntemente atribuídos a algumas cores pela cultura ocidental. Os valores percentuais indicam a frequência com que cada significado é atribuído à cor correspondente.

A cor *facilita a percepção* pois tem normalmente o efeito de reforçar a identificação da forma dos objectos através da descontinuidade da luminosidade nas arestas dos objectos.

7.1 Princípios Gerais para o Emprego da Cor

O emprego da cor deve ser *moderado e adequado* aos fins em vista. O emprego simultâneo de muitas cores sobrecarrega o sistema de percepção visual com demasiada informação. Isto tem como consequência o desviar da atenção dos elementos de informação essenciais, a não ser que se pretenda uma imagem muito alegre mas, em casos semelhantes, não deve ser apresentado mais nenhum elemento de informação adicional pois será naturalmente perdido. Por outro lado, o emprego de cores muito escuras e semelhantes transmite sentimentos negativos que desmotivam o receptor a tentar perceber a mensagem que lhe é destinada.

Ao tentar construir uma imagem com um conteúdo a transmitir, convém *começar por uma imagem a preto e branco* e, só depois se deve adicionar a cor. Isto permite avaliar se as cores empregues acrescentam ou não valor à eficácia da transmissão do conteúdo. Este procedimento tem benefícios adicionais pois permite ter em atenção casos de daltonismo, ecrãs monocromáticos ou ecrãs e impressoras com muito poucas cores.

A escolha das cores deve *evitar que a cor seja mal interpretada*. Uma mensagem de perigo deve conter a cor vermelha para que a mensagem seja clara e rapidamente entendida, o que não aconteceria se a mensagem fosse verde ou azul. Inversamente, uma mensagem de informação de cor vermelha viva será erradamente interpretada como uma mensagem de perigo. O emprego de cores brilhantes ou muito saturadas pode também confundir a codificação pois este tipo de cores realça as respectivas áreas exagerando a sua importância visual em relação a outras áreas que possam ser de igual importância.

A escolha de cores é um processo iterativo em que se deve testar os resultados com destinatários representativos dos utilizadores de forma a *identificar preferências, detectar problemas e corrigir* os problemas detectados.

O emprego da cor deve proporcionar *satisfação estética*, permitindo a harmonia entre as cores empregues. O conceito de harmonia é subjectivo, mas existem alguns princípios

que podem ser aplicados de forma quase universal. As cores seleccionadas não deverão ser muitas e devem ser tais que se encontrem a igual distância perceptual entre si. A figura 7.1 apresenta um processo simples de escolha das cores derivado da pintura. Nesta figura, as cores puras encontram-se dispostas à roda de uma circunferência²⁰. O processo de selecção consiste em inscrever polígonos com um número de vértices igual ao número de cores pretendidas dentro da circunferência. As soluções mais harmoniosas correspondem à inscrição de polígonos regulares como triângulos equiláteros ou quadrados para a escolha de três ou quatro cores, respectivamente. Uma vez escolhida uma das cores, as restantes cores ficam automaticamente determinadas. Quando se pretendem três cores em que uma delas é o cião (localizada entre verde e azul na circunferência), a escolha mais harmoniosa das outras duas cores recairá numa cor intermédia entre amarelo e laranja e noutra entre vermelho e magenta. No caso de quatro cores, as três cores a seleccionar para além do cião, serão amarelo, uma cor intermédia entre laranja e vermelho e magenta.

Este processo aplica-se também quando se tenham previamente escolhido duas cores. Neste caso a terceira cor será a cor correspondente ao vértice de um triângulo inscrito no círculo e cuja base é constituída pelo segmento de recta que une as duas cores seleccionadas. Quando as cores seleccionadas forem verde e azul, a terceira cor será uma cor intermédia entre vermelho e laranja. Para o caso de quatro cores ter-se-á que inscrever um rectângulo em que um dos lados se encontra definido pelas duas cores seleccionadas. Quando estas cores forem verde e azul, as duas restantes cores serão laranja e vermelho (veja-se a figura 7.1).

Quando a escolha de cores estiver condicionada à partida, obrigando à utilização de cores que se chocam entre si, a utilização de *separadores* ou *molduras de cor negra* para separar as zonas em conflito permite a utilização dessas cores diminuindo a sua agressividade mútua.

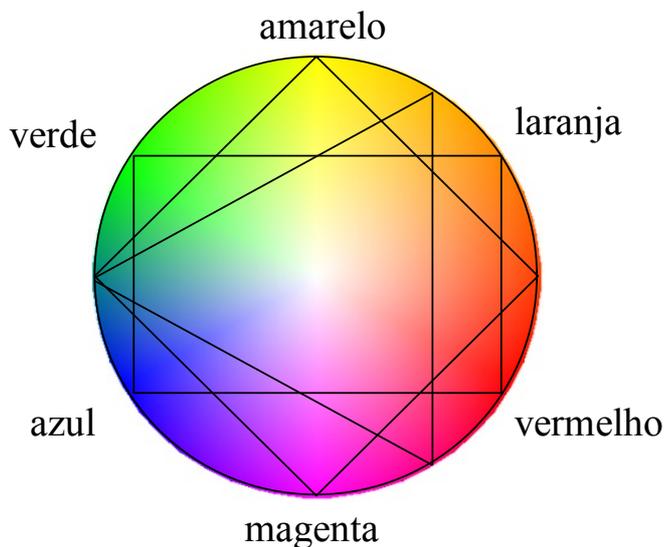


Figura 7.1 – Círculo para a escolha de cores.

²⁰ Não confundir esta disposição com a disposição das mesmas cores nas bases dos cones dos modelos de HSV e HLS, que é diferente.

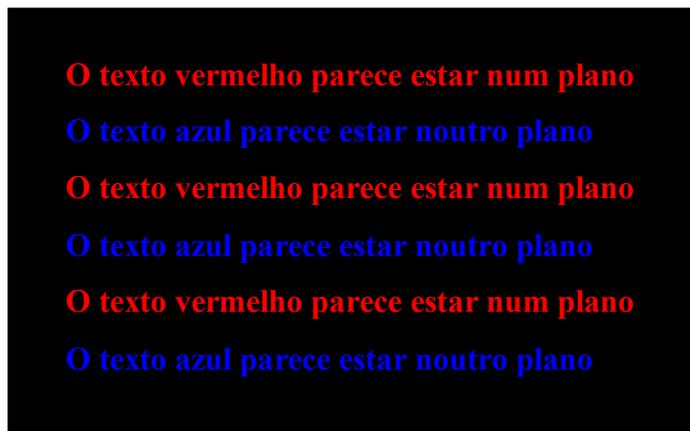


Figura 7.2 – Percepção errada de profundidades diferentes devida ao emprego de cores diferentes muito próximas na imagem.

A dimensão das áreas coloridas deve ser objecto de cuidados especiais. O olho humano apresenta limitações à discriminação de áreas de pequena dimensão e não consegue ver áreas que se apresentem segundo ângulos aparentes inferiores a 20 a 40 minutos de arco. As áreas de grandes dimensões apresentam outros problemas. Quando duas áreas idênticas, uma de cor vermelha e outra verde, são colocadas próximo uma da outra, a área a vermelho parecerá maior do que a área a verde. Outro problema perceptual, que afecta as áreas de maiores dimensões com cores saturadas, deriva da persistência das imagens no sistema visual. Quando a visão é desviada de uma dessas áreas para áreas adjacentes, estas últimas poderão ser percebidas como sendo da cor da área de maior dimensão de que se acabou de desviar o olhar.

Existe ainda um outro fenómeno que ocorre na interpretação de imagens que apresentam cores muito variadas e diferentes posicionadas próximo umas das outras, tal como a figura 7.2 apresenta. O sistema de visão apercebe-se das diferenças de luminosidade entre as várias cores apresentadas, mas tende a interpretar essas diferenças como diferenças de profundidade (distância ao observador). Deste modo, uma área vermelha aparentará encontrar-se mais próxima do observador do que uma área azul que lhe seja contígua.

Em gráficos com poucas cores, a cor de fundo deverá ser uma cor complementar das cores das linhas ou símbolos do gráfico. Quando o número de cores for maior, esta solução deixa de ser possível e a solução mais harmoniosa para a cor do fundo passará pelo emprego de um tom cinzento. Este deverá ser claro se as cores dos gráficos forem cores saturadas e escuro se essas cores forem pouco saturadas.

Os cuidados com as cores de textos englobam a cor do texto e a cor do fundo sobre o qual o texto se encontra. A cor do texto deve criar um contraste nítido com a cor do fundo em termos de tonalidade e intensidade. Assim (veja-se a figura 7.3), não devem ser empregues dois tons de azul, nem azul e preto, dado que se confundem nos limites dos caracteres do texto e tornam difícil a sua leitura. Do lado contrário da escala de intensidade verifica-se o mesmo problema como, por exemplo, se empregam pares de cores como o amarelo e o branco.

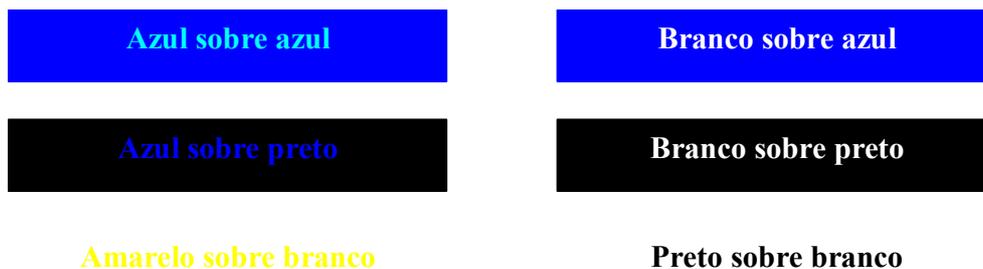


Figura 7.3 – Contraste entre a cor de texto e a cor de fundo. O pouco contraste (à esquerda) dificulta a leitura, enquanto o contraste adequado (à direita) a facilita.

Existe igualmente o problema de, uma vez seleccionadas duas cores suficientemente contrastantes, escolher a cor para o fundo e a cor para o texto. A experiência mostra que texto de cor branca sobre fundo azul é a melhor solução para uma boa leitura, prolongada ou não, e que a solução alternativa mais próxima consiste em texto de cor branca sobre fundo negro. Em geral deveremos escolher a cor menos luminosa para cor de fundo para que o texto, com uma cor mais luminosa, sobressaia do fundo, tal como a figura 7.3 mostra. No entanto, quando se pretender marcar uma parte do texto com uma cor diferente, como o amarelo, o texto marcado não sobressairá suficientemente do texto não marcado porque as duas cores do texto apresentarão luminosidades muito próximas. Uma solução aceitável para este caso (veja-se a figura 7.4) consiste em apresentar a negro sobre fundo branco o texto não marcado, passando o fundo a ser localmente amarelo na zona de texto marcado. Esta solução é superior à que consiste em trocar as cores para marcar texto, ou seja, usar texto branco sobre fundo negro. Deve ainda ser evitado o uso de cores como cores vermelhas e verdes pouco saturadas ou com pouca luminosidade dado que estas cores confundem os indivíduos com daltonismo e lhes dificultam a leitura dos textos.

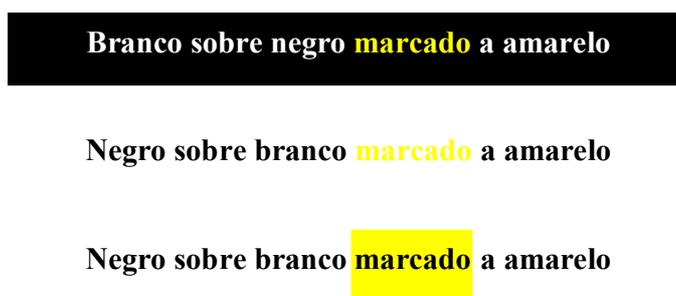


Figura 7.4 – Marcação de texto. O texto marcado a amarelo na linha de baixo sobressai facilmente.

Exercícios

- 1 Apresente a teoria dos três estímulos e relacione o facto de todos os modelos de cor serem função de três parâmetros com esta teoria.
- 2 O modelo de cor da CIE é um modelo baseado em três entidades. Que entidades são estas e qual a sua relação com a teoria dos três estímulos?
- 3 Durante o teste de um monitor a cores empregando a tecnologia de tubos de raios catódicos com tubos para o vermelho, verde e azul, verificou-se que o monitor não permitia reproduzir todas as cores do espectro visível. Que cores não eram reproduzíveis e por que razão (razões) isso não era possível? (sugestão: use o diagrama CIE x,y na sua explicação)
- 4 Apresente a diferença principal entre o modelo de cor da CIE e todos os outros modelos de cor (RGB, CMY, etc.).
- 5 A transformação do modelo de cor RGB para o modelo de cor HSV é ou não uma transformação biunívoca (a um valor RGB corresponde um e um só valor HSV e vice-versa)? Explique porquê.
- 6 Os modelos de cor YCbCr e YIQ apresentam uma diferença fundamental para todos os outros modelos de cor estudados. Qual é essa diferença e que vantagens e desvantagens apresentam em relação aos outros modelos?
- 7 Construa duas escalas de intensidade com 8 intensidades, com I_0 igual a 0,125 e 0,0625 e compare as intensidades que determinar com as intensidades das escalas de 4 intensidades da tabela 4.3.
- 8 Determine para que intensidades de uma dada escala de intensidades é mais importante aplicar a correcção gama (se necessitar, empregue os valores de uma das escalas que determinou no problema anterior e um valor de gama de 2).
- 9 O monitor de um equipamento apresenta um valor de gama de 2,35 e o subsistema de controlo do monitor realiza uma correcção gama empregando um valor pré-definido de 1,9. Que consequências têm estes valores na apresentação de imagens que não tenham sido sujeitas a quaisquer correcções e qual deverá ser o valor da correcção gama a aplicar para que as cores apresentadas no monitor sejam correctas?
- 10 Em que é que a técnica de meios-tons (halftoning) difere da técnica de dithering?
- 11 A matriz de tramas 3×3 de (5.1) não corresponde à sequência de tramas apresentadas na figura 5.2. Verifique se a sequência apresentada pela figura está correcta.
- 12 Empregando a relação (5.2), deduza a matriz de tramas $D^{(4)}$ e verifique se a matriz que calculou obedece às regras para a construção de matrizes de tramas.
- 13 A análise da distribuição de cor numa imagem revelou a existência de 256 níveis de vermelho, 16 níveis de verde e 256 níveis de azul. Pretende-se apresentar esta imagem num monitor que possui apenas 10 bits de profundidade para cada píxel. Apresente duas soluções para a distribuição dos níveis de cor no monitor, sendo uma delas simples e a outra mais complexa.