

Fundamentos da Imagem Digital

por Antonio Escaño Scuri

(scuri@tecgraf.puc-rio.br)

Tecgraf/PUC-Rio

Janeiro de 1999

Descrição

Apresentamos aqui os conceitos básicos envolvidos em Processamento de Imagens Digitais e procuramos dar uma visão geral da área. Mas, não serão apresentadas técnicas de processamento em profundidade, nem pelo lado matemático, nem pelo lado computacional.

Por isso mesmo, serve tanto a pessoas da área técnica que estão iniciando seu aprendizado em Processamento de Imagens, ou pessoas de Editoração Eletrônica que queiram aprender os conceitos que estão por traz de programas como o Adobe PhotoShop e o Corel PhotoPaint.

Assim, qualquer um com conhecimentos básicos em computadores, pode assistir ao curso. Mas, o conceito matemático de função é largamente utilizado e é fundamental para o entendimento de algumas noções.

Os conceitos apresentados servem também como uma preparação para um curso mais avançado de Processamento de Imagens. Inclusive vários conceitos são tratados superficialmente por esse motivo e para não tornar o texto ilegível para pessoas que não sejam da área técnica. Mas, uma das maiores dificuldades aqui é definir o que deve realmente ser superficial e o que não deve. Isso só a evolução desse material, através do retorno de pessoas que tenham acesso a ele, poderá dizer.

No futuro pretendemos criar mais dois textos, um chamado “Técnicas de Processamento de Imagens Digitais” com enfoque matemático e computacional, e outro “Aplicações de Imagens Digitais” com enfoque prático envolvendo aplicações científicas, incluindo vários estudos de casos.

Ementa

Dividimos a apresentação em 3 tópicos básicos, mais 3 tópicos complementares e por fim, 3 tópicos de técnicas e aplicações. Sugerimos que sejam vistos inicialmente os três primeiros tópicos para melhor compreensão dos demais.

Conceitos Básicos

Apresentamos alguns dos conceitos básicos necessários para o entendimento dos outros tópicos. Situamos “Processamento de Imagens” e “Computação Gráfica”, apresentamos o “Paradigma dos Quatro Universos”, e mostramos algumas noções de contínuo e discreto.

Resolução Espacial

Quando realizamos a passagem para o mundo digital, a primeira grande etapa para imagens é a amostragem. Aqui vamos explorar um pouco mais essa questão, assim como seu problema inverso, o de reconstrução da imagem original. Como veremos reconstrução é sinônimo de interpolação. Veremos também uma definição do importante conceito de *aliasing*.

Cor

O que é cor? Como representar cor no computador? Quais os padrões estabelecidos? Essas e outras perguntas serão abordadas aqui de maneira introdutória. Falaremos também sobre Luminância e Crominância, assim como outras propriedades de cor, e por fim veremos algumas técnicas de quantização.

Captura

Apresenta alguns dos dispositivos que usamos para passarmos a imagem para o mundo digital. No caso *Scanners*, *Frame/Vídeo Grabbers* e Câmeras Digitais.

Visualização

Uma vez a imagem tendo sido digitalizada, é imediato o desejo de reconstruí-la em alguma superfície de visualização. As mais comuns são monitores e impressoras. Sobre estes mostraremos algumas peculiaridades.

Armazenamento

Embora existam vários formatos de arquivos de imagem (TIFF, PCX, GIF, JPEG) todos armazenam a mesma imagem digital. Para falar sobre isso, mostraremos antes como uma imagem é armazenada na memória do computador.

Processamento

As funções de processamento de imagens digitais podem ser classificadas em duas classes, por escopo ou por resultado. Por escopo, está relacionado com técnicas de processamento. Por resultado, está relacionado com os tipos de resultado de cada técnica.

Software

Aqui procuramos dar uma visão geral de alguns programas de processamento de imagens, procurando situá-los dentro do contexto dos conceitos apresentados. Tais como Adobe PhotoShop, Corel PhotoPaint, Paint Shop Pro, Lview Pro, ACDSee, XV, GIMP.

Sumário

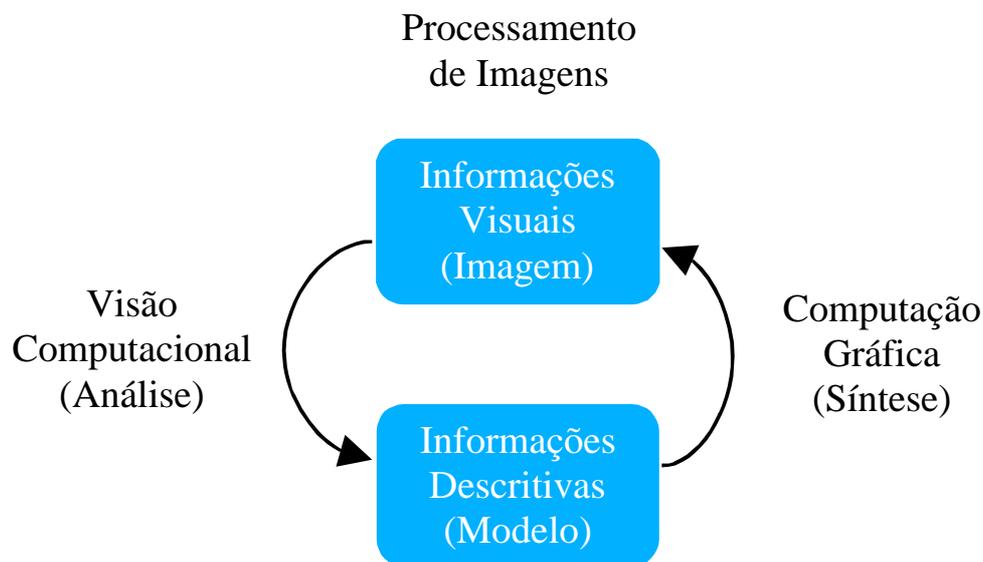
1. Conceitos Básicos	7
1.1. Computação Gráfica e Processamento de Imagens.....	7
1.2. Tópicos de Processamento de Imagens	9
1.3. Contínuo x Discreto.....	9
1.4. O Paradigma dos Quatro Universos	11
2. Resolução Espacial	15
2.1. Topologia de um Pixel.....	16
2.2. Reconstrução	18
2.3. Aliasing.....	20
3. Cor	22
3.1. Processos de Formação de Cor.....	23
3.2. Amostragem e Reconstrução de Cor	24
3.3. Cor no Paradigma dos 4 Universos	26
3.4. Padronização para Cor.....	27
3.5. Sistemas de Cor.....	32
3.6. Sistemas de Gerência de Cor.....	37
3.7. Quantização com Menos de 24 bits.....	39
4. Captura	42
4.1. Scanners	42
4.2. Digitalizadores de Sinal de Vídeo	43
4.3. Câmeras Digitais.....	44
5. Visualização	45
5.1. Monitores	45

5.2. Impressoras	47
6. Armazenamento.....	48
6.1. Compressão	50
6.2. Formatos de Arquivo	51
6.3. Imagens Grandes e Multiresolução	52
7. Processamento.....	54
7.1. Restauração/Realce.....	55
7.2. Segmentação	55
7.3. Extração de Atributos	55
7.4. Classificação	56
7.5. Operações Pontuais.....	56
7.6. Operações Locais.....	58
7.7. Operações Globais.....	59
8. Software	61
9. Aplicações.....	64
10. Internet Links	66
11. Bibliografia	67

1. Conceitos Básicos

1.1. Computação Gráfica e Processamento de Imagens

Antes de mais nada, precisamos nos situar dentro do contexto de Computação Gráfica. Estamos trabalhando basicamente com dois tipos de informação: Visual e Descritiva. Informação visual é por exemplo a imagem vista na tela do computador. Informação descritiva é referente ao modelo matemático que representa os objetos visualizados. A área de Processamento de Imagens abrange operações que são realizadas sobre imagens e que resultam em imagens. A área de Computação Gráfica abrange as operações de síntese de imagem, ou seja, a geração de uma visualização do modelo. A área de Visão Computacional abrange as operações de análise dos objetos contidos na imagem e a geração de modelos matemáticos desses objetos.



Aqui iremos ver várias funções de Processamento de Imagens e um pouco da interface desta área com as áreas de Visão Computacional e Computação Gráfica.

Na área de Computação Gráfica se encaixam programas de arquitetura, design e simulação gráfica, tais como as utilizadas em filmes com animação por computador. Programas comerciais que representam muito bem esta área são o AutoCad e o Studio 3D, ambos da AutoDesk, e o Corel Draw da Corel.

Desta maneira, o modo de encararmos os programas das diversas áreas se torna diferente de acordo com o nosso objetivo. Por exemplo, a planta baixa de uma casa pode ser desenhada a partir de suas medidas, sem a necessidade de uma fotografia aérea da casa ou, até mesmo, da existência da casa. Por outro lado, para analisarmos uma amostra de bactérias no microscópio precisamos da imagem da amostra, de onde podemos obter alguma informação, sem a necessidade de conhecermos a posição exata de cada bactéria.

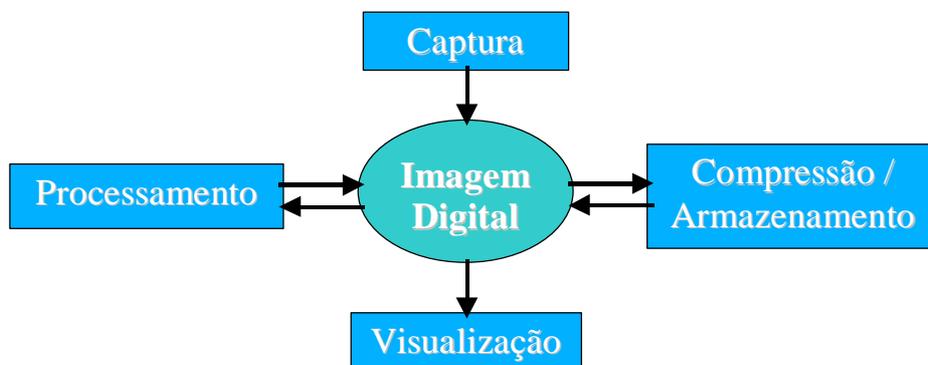
Esta diferença de objetivos vai levar à modelos de trabalho diferentes. Em processamento de imagens usa-se o modelo de matricial e a computação gráfica geralmente se baseia no modelo de objetos vetoriais. Neste modelo os objetos são armazenados apenas a partir da descrição das coordenadas de seus vértices, sejam elas espaciais ou planares (três ou duas dimensões, respectivamente). Dessa maneira utiliza-se um sistema de coordenadas Cartesiano, onde os objetos podem ser escalados, rotacionados e transladados com maior liberdade para cada objeto. O modelo matricial utiliza uma matriz de dados para armazenar a informação de cor em cada ponto da imagem, onde o sistema de coordenadas é obviamente uma grade de números inteiros que descrevem a posição na matriz. Portanto, no modelo matricial não há distinção dos objetos contidos na imagem. Além disso, armazenar a matriz que contém a imagem geralmente exige muito mais memória que armazenar a descrição vetorial.

A diferença de modelos fica explícita quando apresentamos o paradigma dos 4 universos, que nos ajudará a situarmos todos esses conceitos.

1.2. Tópicos de Processamento de Imagens

Uma vez tendo nos situado no contexto de Processamento de Imagens podemos agora desdobrar essa área nos tópicos que serão abordados. O ponto central de discussão será, é claro, a Imagem Digital. Falaremos sobre dois aspectos importantes que definem a imagem: Resolução e Cor. Feito isso, podemos atacar os pontos de Captura e Visualização, Armazenamento e por fim Processamento. Esses outros pontos tem por objetivo esclarecer algumas confusões na área e preparar o aluno para cursos mais avançados.

A figura a seguir ilustra a relação entre os diversos tópicos.



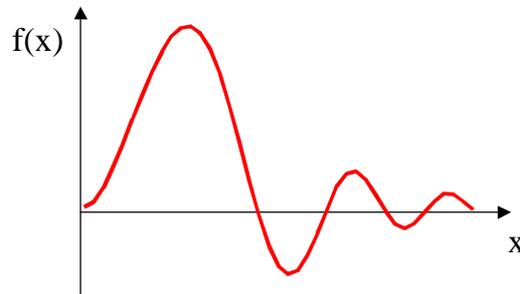
1.3. Contínuo x Discreto

Para podermos introduzir o conceito de imagem digital é necessário antes falarmos sobre funções contínuas e funções discretas. Podemos pensar em uma função contínua como um corda sem emendas que se estende indefinidamente (ou seja infinita) e por mais que você corte essa corda, sempre poderá cortá-la um pouco menor indefinidamente (ou seja infinitesimal). Matematicamente estamos falando de uma função real definida na reta toda (em \mathfrak{R}), onde para cada valor de x fornecido sempre existe um único valor de $f(x)$ obtido.

Formalmente:

$$f : \mathfrak{R} \rightarrow \mathfrak{R}$$
$$x \rightarrow f(x)$$

Colocando isso em um gráfico:



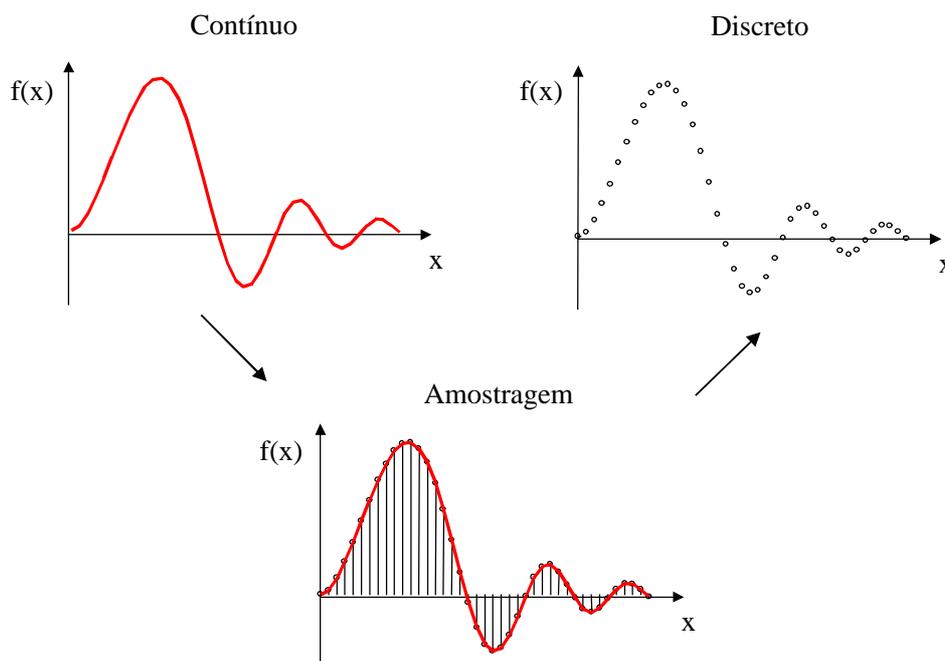
O computador só é capaz de armazenar bits, um valor que pode ser 0 ou 1. Para simplificar as coisas juntou-se 8 bits formando uma palavra chamada de byte. Um byte então pode assumir $2^8 = 256$ valores diferentes, variando de 0 a 255. Quando falamos de sistemas operacionais de 16 bits e 32 bits, estamos falando da maior palavra que aquele sistema pode processar de uma vez, por isso sistemas de 32 bits são mais eficientes que os de 16.

Assim, o computador trabalha sempre com números inteiros ou na melhor das hipóteses com uma aproximação de um número real, chamada de ponto flutuante (isso porque o número de bits dedicado para as casas decimais é flexível de acordo com o número). Por consequência, não é possível representar uma função contínua no computador. Podemos apenas simulá-la.

O processo para trazer uma função contínua para o computador é discretizando-a (ou digitalizando-a), ou melhor, tomando valores pontuais ao longo de x e guardando o valor de $f(x)$ correspondente (é claro que o eixo $f(x)$ também é contínuo, assim também precisaremos discretizá-lo). O processo de discretização do eixo x (o domínio) é chamado de Amostragem, o do eixo $f(x)$ (o contradomínio) é chamado de Quantização.

Em resumo, a discretização de qualquer sinal contínuo primeiro passa por uma amostragem e depois passa por uma quantização. Além disso, não podemos armazenar um sinal que se estenda indefinidamente, portanto o sinal digital também é limitado a um intervalo do domínio.

A amostragem que vemos na figura é a mais comum de todas e mais popular, implementada na grande maioria de dispositivos de captura. É chamada de uniformemente espaçada, pois cada amostra é tomada em intervalos iguais. Embora existam outras técnicas de amostragem que utilizam menos amostras onde a função é monótona e mais amostras onde a função apresenta mais irregularidades.

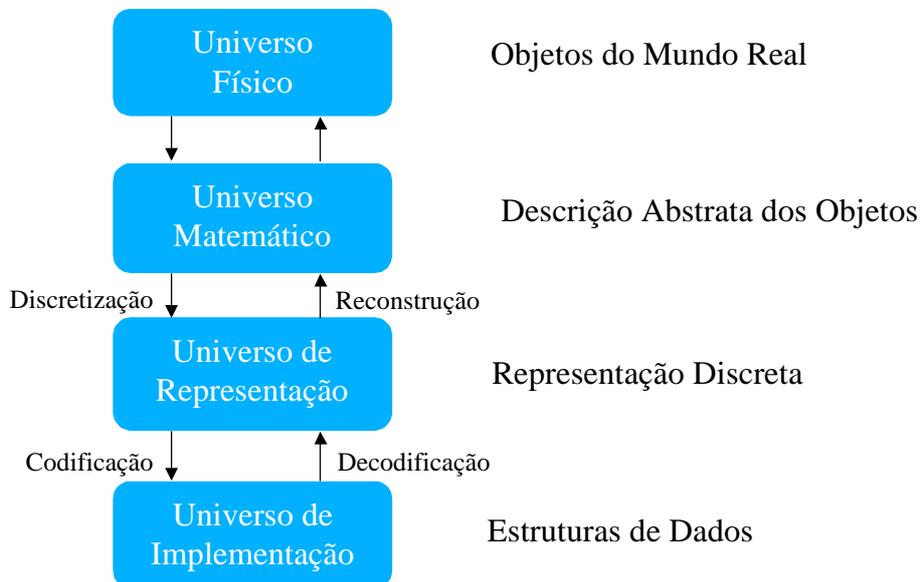


1.4. O Paradigma dos Quatro Universos

Esse paradigma vai nos ajudar a entender esse processo de sair do mundo real, onde os sinais são contínuos, e ir para o mundo do computador onde tudo é discreto. O paradigma especifica 4 universos. O universo Físico, onde estão os objetos do mundo real. O universo Matemático, onde formulamos descrições abstratas desses objetos. O universo de Representação vai permitir trazer essas descrições abstratas para o mundo digital, e é onde se dará a discretização dos sinais contínuos. O universos de Implementação é onde é feita a codificação do sinal discretizado na memória do computador através de uma estrutura de dados.

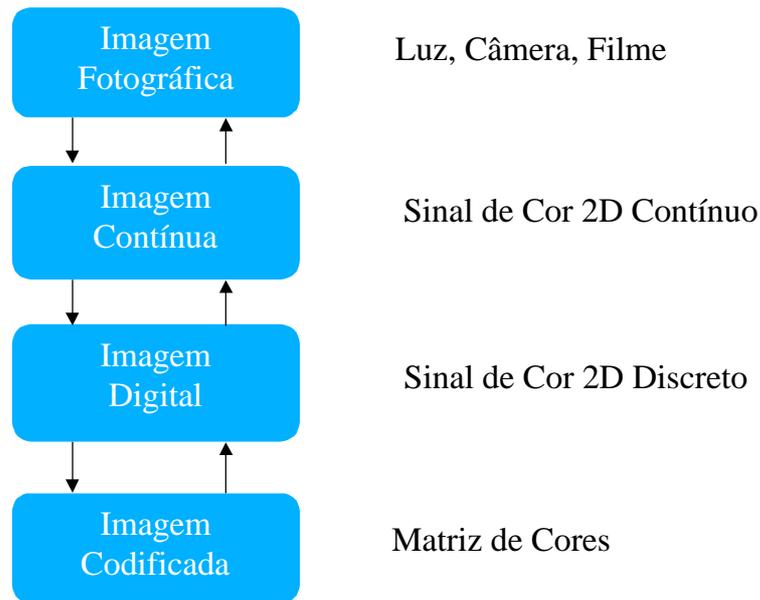
Assim, para estudar um determinado fenômeno, ou objeto da natureza, no computador, nós precisamos associá-lo a um modelo matemático e então encontrar uma representação discreta para esse modelo que pode ser implementada no computador.

A codificação dos dados pode ser decodificada e depois reconstruída para podermos voltar com o fenômeno estudado para o mundo real.



Para imagens também precisamos fazer esses mesmos passos.

A imagem fotográfica é obtida do mundo real através de câmeras ou sensores que captam luz. A imagem capturada em um filme fotográfico representa bem a imagem real que queremos transportar para o computador. Repare que o filme define um plano limitado por um retângulo, onde cada posição nesse plano contém a informação de cor relativa aquela posição, ou seja, a imagem neste caso é um sinal de cor 2D contínuo, onde o domínio é o plano e o contradomínio é o espaço de cor. A discretização disso é muito simples, segue os mesmos princípios da discretização de sinais 2D contínuos quaisquer.



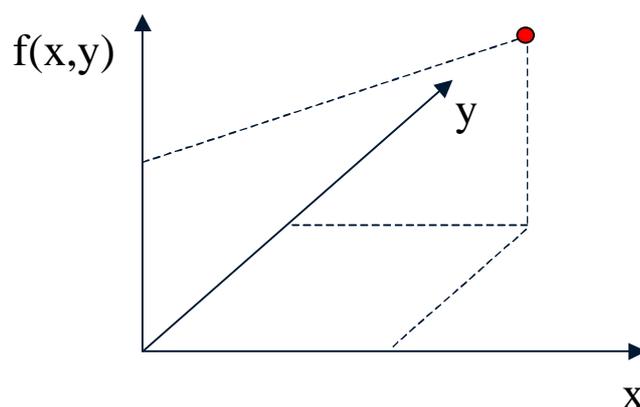
A imagem contínua é modelada matematicamente pela função:

$$f(x, y) = \text{Cor no ponto } (x, y)$$

$$x \in [0, X] \text{ e } y \in [0, Y]$$

, onde x e y são números reais, limitados ao intervalo de 0 a X , e de 0 a Y , respectivamente.

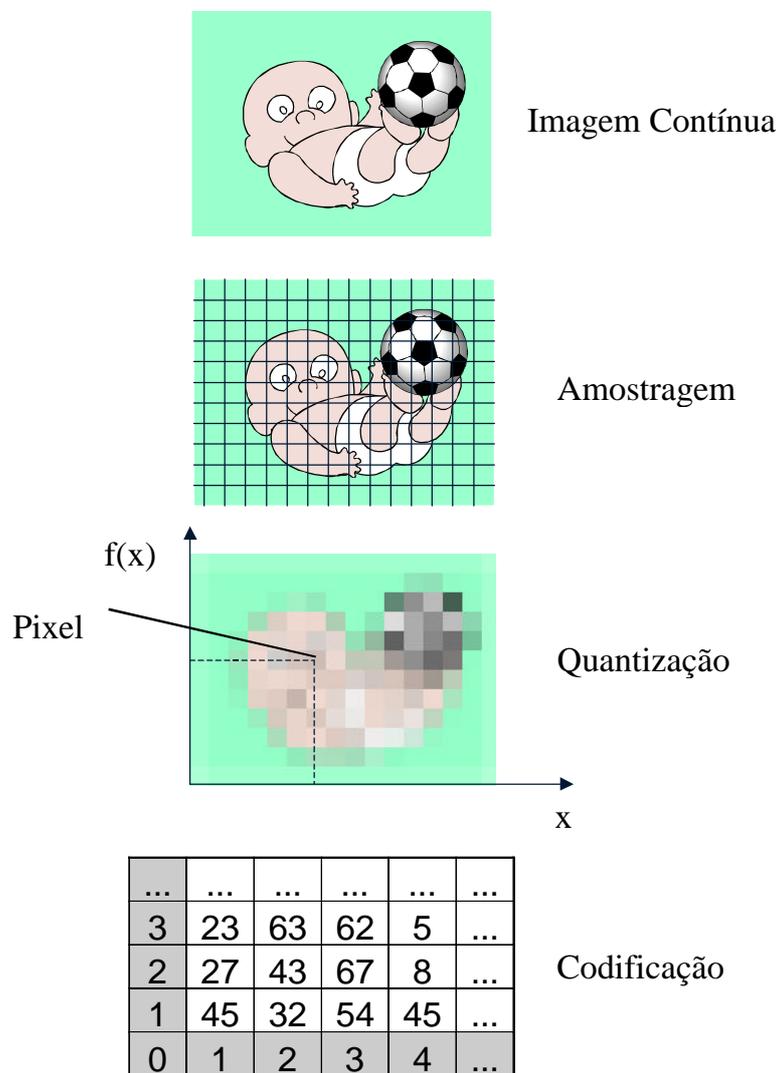
Colocando isso em um gráfico:



A imagem digital é obtida a partir de uma amostragem e de uma quantização dessa função, como já foi visto, e pode ser representada pela mesma função, só que neste caso, x e y são números inteiros. Uma vez amostrado o plano temos uma matriz de valores que representam a cor. Cada elemento dessa matriz é chamado de Pixel (abreviação de Picture Element).

Mas, ainda não falamos nada sobre a quantização da cor. Na realidade cor também é um fenômeno físico para o qual precisamos fazer os mesmos passo que fizemos para o plano da imagem. Precisamos descobrir a definição de cor, suas representações e como trazê-las para o computador. Por enquanto, vamos deixar essas discussões de lado e atacar a questão de amostragem do plano 2D onde as cores estão distribuídas.

A figura a seguir ilustra o processo de discretização da imagem e mostra uma possível codificação da cor para cada elemento da matriz.



2. Resolução Espacial

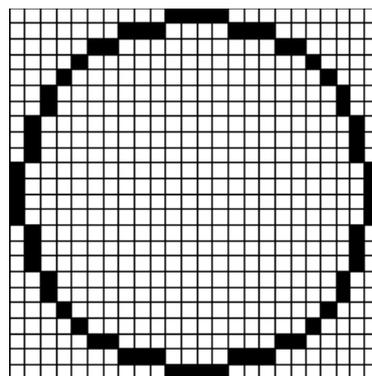
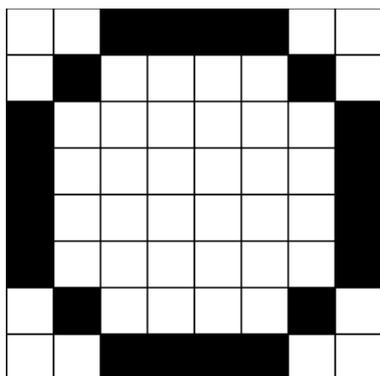
Ao ser digitalizada a imagem assume um tamanho adimensional, em pixels. Mas, pode-se conhecer uma medida de qualidade da amostragem, conhecendo-se a razão entre o número de pixels obtido e o tamanho da imagem real no filme ou equivalente. A isso chama-se de Resolução. Em geral, é medida em pontos por polegada ou DPI (dots per inch), mas pode ser também em pontos por centímetro ou DPC, ou ainda em qualquer outra unidade equivalente. Também é simples estabelecer a relação: número de pixels = resolução x tamanho real.

Para entender melhor isso, pode-se fazer a seguinte questão: Dado o tamanho da imagem (em cm, in, etc...), quantos valores discretos vão ser tomados? Escolhendo-se a resolução a resposta é dada pela equação acima.

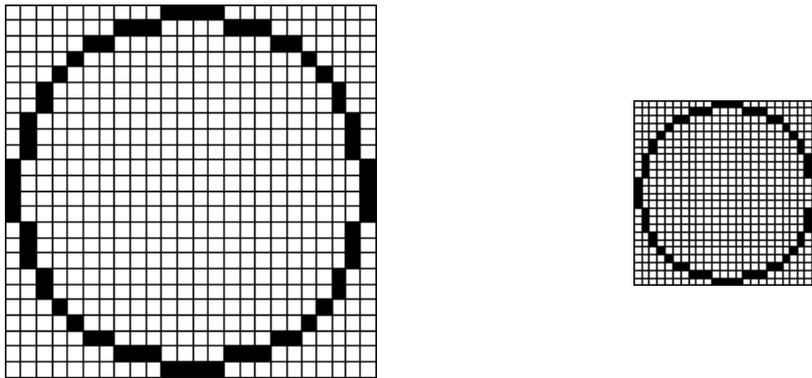
Veja algumas resoluções típicas: Monitor Comum - 72 DPI; Scanner - 1200 DPI; Impressora Jato de Tinta Popular - 600 DPI

Como temos duas dimensões, podemos definir uma resolução horizontal e uma vertical. Quando nada se diz a respeito disso, quer dizer que são iguais (a grande maioria das vezes), caso contrário são necessários dois valores. Uma vez que são diferentes, a razão de aspectos da área entre um pixel e outro não é 1:1 e portanto não é um quadrado. Isso gera algumas confusões, pois no monitor a razão é 1:1 e esta imagem quando visualizada aparecerá deformada.

Quando duas imagens com tamanhos reais iguais são capturadas com resoluções diferentes, naturalmente terão número de pixels diferentes e na tela aparecerão com tamanhos diferentes. Na figura a seguir mostramos um círculo amostrado em duas resoluções, mas exibido no seu tamanho original.



Quando duas imagens de tamanhos reais diferentes são capturadas com resoluções diferentes de tal forma que gerem imagens digitais com o mesmo número de pixels, quando visualizadas no monitor aparecerão com o mesmo tamanho na tela. No exemplo a seguir mostramos as imagens proporcionais ao seu tamanho original.



Quando queremos realizar alguma medida sobre os objetos contidos na imagem surge um outro parâmetro, a escala. Pode-se defini-la matematicamente idêntica a resolução só que as unidades serão as mais variadas possíveis. Caso não seja conhecida, é fácil de se obtê-la. Mede-se algumas distâncias reais na imagem digital e calcula-se a média das razões entre distância e número de pixels, obtendo-se a escala. A partir desse momento outras distâncias podem ser feitas na imagem com qualidade aceitável. É claro que se os objetos não estejam em um mesmo plano, as medidas são muito mais complexas e envolvem transformações projetivas.

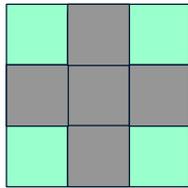
2.1. Topologia de um Pixel

As amostras do plano são pontuais e adimensionais, mas os dispositivos para reconstrução não deixam um espaço entre uma amostra e outra. Assim, muitas vezes iremos representar o pixel, não como um ponto, mas como um quadrado preenchido com a cor do pixel, que está no seu centro. Por sinal, o exemplo anterior já usa essa representação.

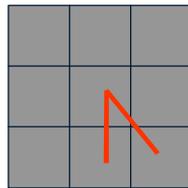
Como vimos, pixels são normalmente quadrados, gerando uma grade (*grid*) regular devido à amostragem uniformemente espaçada. Um problema que aparece nesse momento é: Quais são os vizinhos de um determinado pixel? Essa pergunta é fundamental para algoritmos de preenchimento de áreas e algoritmos de determinação de continuidade de objetos na imagem.

Nesta grade regular podemos definir dois tipos de vizinhança: 4 conectada (onde os vizinhos são equidistantes) ou 8 conectada (onde nem todos são equidistantes, os vizinhos na

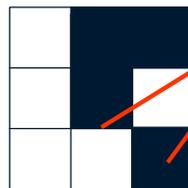
diagonal são mais distantes que os vizinhos nas laterais). Normalmente usa-se uma combinação das duas vizinhanças, definindo 4 conectada para os objetos e 8 conectada para o fundo.



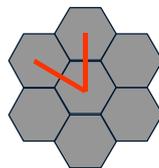
Vizinhança 4 Conectada



Vizinhança 8 Conectada

Vizinhos 8 conectados,
mas não 4 conectados

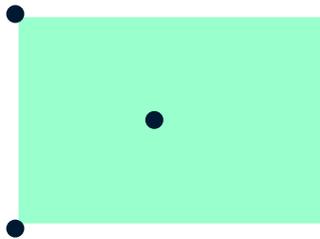
Para criar uma vizinhança onde todos os vizinhos fossem equidistantes e só houvesse uma definição, temos que mudar a grade regular para uma grade hexagonal. São poucos algoritmos implementados com essa vizinhança, pois como a imagem está em geral amostrada em um grade regular, deve-se convertê-la para a grade hexagonal, realizar o processamento e convertê-la de volta para a grade regular.



Vizinhança Hexagonal

Já que estamos falando da grade não poderíamos deixar de falar sobre a origem do sistema de coordenadas. Localizada no canto superior esquerdo na grande maioria dos sistemas de processamento de imagem, pode na realidade ser posicionada em qualquer lugar da imagem, até mesmo fora dela.

Essa localização preferencial para o canto superior esquerdo, com eixo y invertido em relação ao padrão da matemática, é herdada dos próprios dispositivos de visualização, onde a primeira linha a ser redesenhada na tela é a do topo da tela, que segue o padrão de escrita ocidental, da esquerda para a direita e de cima para baixo.



Origens Mais Comuns

2.2. Reconstrução

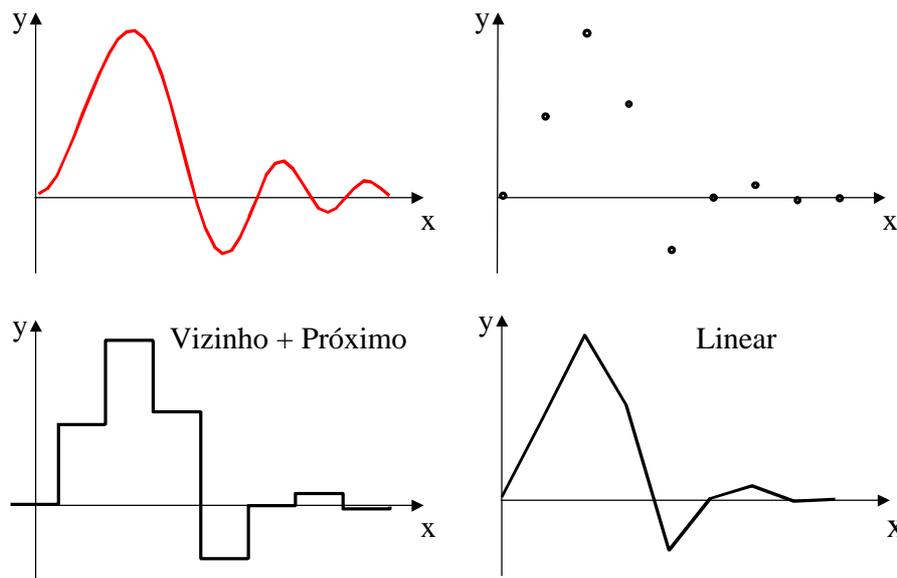
Quando falamos em amostragem imediatamente nós temos o problema inverso: Como reconstruir? Ou melhor podemos fazer a pergunta: Dados os valores discretos, qual a função contínua? Quando estamos no ambiente computacional, também estamos querendo resolver um outro problema que é o de re-amostragem. Isto é, reconstruir e amostrar novamente com outra resolução ou com outro grid regular ou não.

Em matemática, reconstrução é sinônimo de interpolação. O uso desse conceito permite que dado uma seqüência de amostras obtemos o valor em uma posição qualquer do eixo.

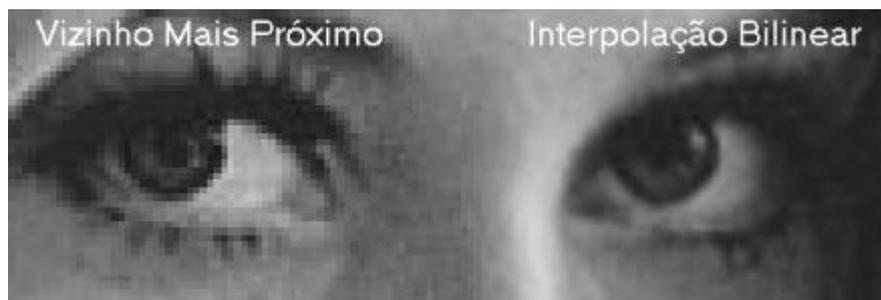
A técnica mais trivial é simplesmente escolher a amostra mais próxima da posição desejada e tomar o seu valor como valor desta posição. Esta técnica é usada muito para a visualização da imagem em maior ou menor proximidade, ou melhor, fazendo Zoom, simulando uma aproximação ou afastamento da imagem em relação ao observador.

Uma outra técnica muito utilizada para fazer reconstrução é a de interpolação linear (bilinear no caso de 2 dimensões). Essa técnica calcula uma média ponderada das amostras mais próximas da posição desejada de acordo com a distância à mesma.

Pela figura a seguir podemos perceber que a interpolação linear irá obter resultados bem melhores que a de vizinho mais próximo. Mas fica claro que podemos obter resultados ainda melhores usando curvas em vez de retas. Uma técnica com esse enfoque que se tornou popular é a interpolação bi-cúbica, que é semelhante a interpolação linear, mas em vez de usar retas, usa funções cúbicas (x^3) para realizar a ponderação.



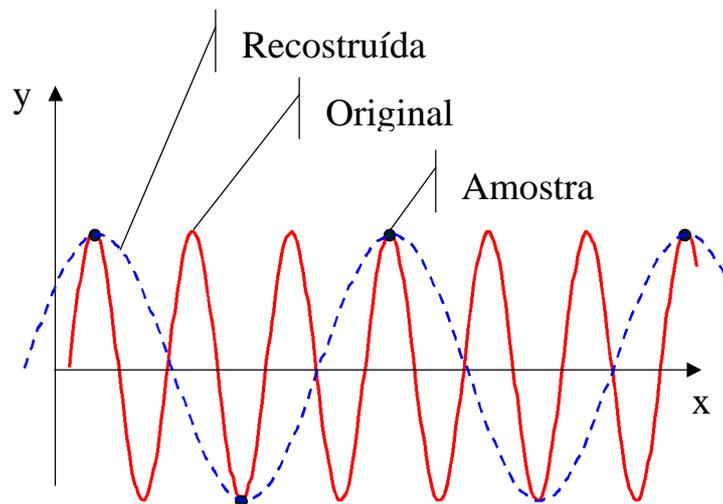
Esta imagem mostra o efeito de uma mudança no tamanho de uma determinada imagem, usando-se a técnica de interpolação linear e a técnica do vizinho mais próximo.



Toda técnica de deformação da imagem, como rotação ou coisas mais genéricas como Warming são implementados usando interpolação (Obs: Morphing = Warming + Blending). Por que no fundo estamos reconstruindo, deformando o modelo contínuo e re-amostrando.

2.3. Aliasing

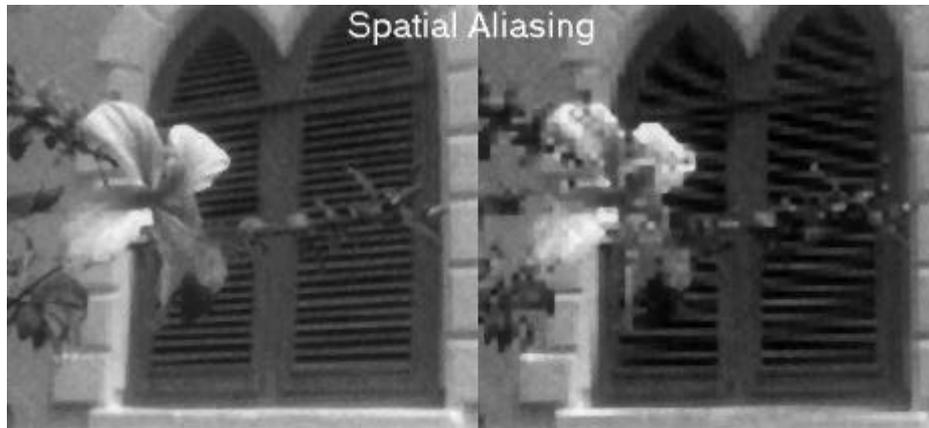
Um fenômeno muito comum em amostragem é o fenômeno de aliasing. Ele acontece quando realizamos uma amostragem muito pobre da função contínua original e quando vamos reconstruí-la obtemos resultados inconsistentes com o esperado. Matematicamente podemos dizer que a frequência de amostragem foi inferior à aproximadamente à maior frequência de variação da função contínua. Podemos ver o fenômeno de aliasing em outras situações como ao filmarmos uma roda de um carro rodando mais rápido do que a câmera é capaz de capturar, neste caso vemos a roda girar para trás, justamente por que a taxa de amostragem está muito baixa, no nosso caso a resolução da imagem.



Para resolver o problema de aliasing temos duas soluções, aumentar a taxa de amostragem, que em geral nem sempre é possível, ou suavizar a função contínua para que não tenha frequências muito altas comparadas com a nossa taxa de amostragem.

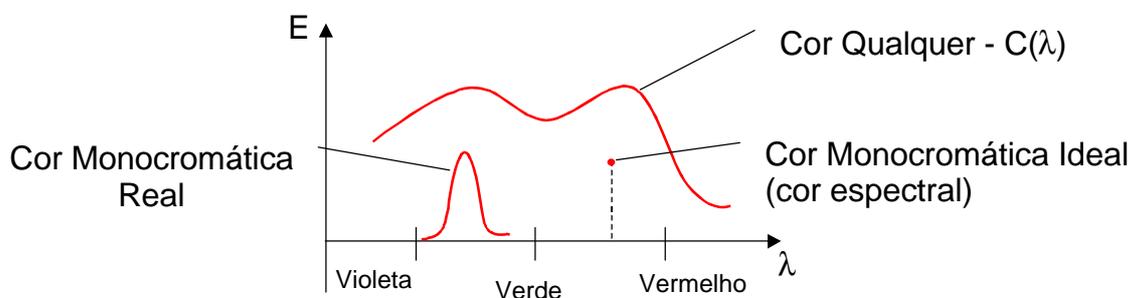
Essa solução, totalmente correta, induziu um erro conceitual que vem se propagando durante anos na literatura. Muitas técnicas de suavização de contornos (para que estes pareçam mais suaves na visualização) são chamadas erradamente de técnicas de anti-aliasing.

Esta imagem mostra o efeito de aliasing espacial em uma imagem que foi amostrada em alta e baixa resolução.



3. Cor

Quando falamos de cor de um ponto de vista físico, na realidade estamos falando de Luz. A luz percorre o espaço, ora se comportando como uma onda, ora como uma partícula, isso caracteriza a natureza dual da luz. Para estudarmos cor iremos observar o comportamento da luz como uma onda. Podemos dizer que neste caso a luz é uma radiação eletromagnética que possui uma certa “energia” para cada comprimento de onda. A função que informa o valor dessa “energia” é chamada de Distribuição Espectral de Potência.



Na figura podemos ver um exemplo de distribuição, e a diferença entre uma cor monocromática ideal (apenas 1 comprimento de onda) e uma cor monocromática real (onde existe um comprimento de onda predominante, mas não é único).

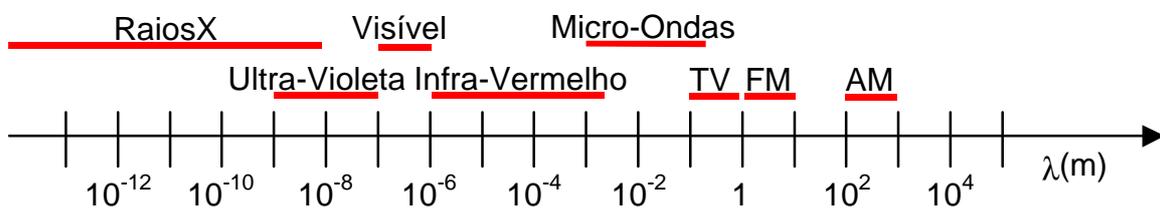
Portanto, luz pode ser modelada matematicamente por uma função unidimensional que depende do comprimento de onda. Mas nós estamos interessados em cor, e neste caso estamos falando da sensação humana de diferentes espectros de luz. Ou seja, a definição de cor neste caso envolve uma característica perceptual da espécie humana. Até porque é sabido que cada espécie possui uma percepção de cor diferente. Consequentemente precisamos entender um pouco mais sobre como o olho humano percebe cor e envia essa informação para o cérebro. (A ciência que estuda cor do ponto de vista físico é chamada de Colorimetria. A ciência que estuda cor do ponto de vista perceptual é chamada de Fotometria.)

No nosso olho existem dois tipos de sensores, cones e bastonetes. Bastonetes são sensíveis a todos os comprimentos de onda ponderadamente, possuem melhor resposta à luz (visão noturna), mas não são capazes de distinguir cor. Já os cones são menos sensíveis à luz (visão direta), mas permitem a distinção de cores por, na realidade, existirem três tipos de cones, sensíveis a

comprimentos de onda próximos a 450nm, 550nm, e 600nm ($\text{nm}=10^{-9}\text{m}$). Isso caracteriza o processo de discriminação de cor do olho, chamado de tricromacidade.

Tanto cones quanto bastonetes permitem com que olho perceba luz entre aproximadamente 380nm e 780nm. Luz de 380nm até 500nm parece azul, de 500nm a 600nm parece verde, e de 600nm a 780nm parece vermelho

É claro que o resto do espectro eletromagnético é ocupado com outros tipos de ondas. A figura a seguir mostra algumas regiões conhecidas e/ou utilizadas por nós.



Portanto existem duas classes de sensores, uma que percebe a intensidade da luz e outra que percebe as diferenças de cor. E é assim que a informação de cor é enviada para o cérebro, em dois canais. Essa decomposição tem o nome de luminância-crominância e é usada em muitas representações de cor.

A cor monocromática ideal é também chamada de cor espectral, já que possui apenas um único comprimento de onda. A cor púrpura não é uma cor espectral, ela contém comprimentos de onda bem próximos a 380nm e outros bem próximos a 780nm. Veremos esse fenômeno da cor púrpura novamente mais adiante. A luz branca possui todos os comprimentos de onda com energia máxima. A luz branca ideal seria um espectro com uma linha horizontal.

A distribuição de cones e bastonetes ao longo da retina é responsável por diversos aspectos da nossa visão. Embora esses aspectos sejam muito interessantes vamos procurar nos concentrar apenas naqueles que nos interessam no momento.

3.1. Processos de Formação de Cor

As cores que nós percebemos são geradas a partir de diversos processos psicoquímicos. Em geral podemos classificá-los em três grandes grupos: aditivos, subtrativos e por pigmentação.

No processo aditivo (o mais simples), vários raios de luz são combinados para formar um novo raio. No processo subtrativo um raio de luz passa por um filtro que elimina alguns

comprimentos de onda resultando numa nova cor. No processo por pigmentação a luz passa por diversos eventos, sendo refletida, transmitida e absorvida, o resultado dessa interação é a nova cor.

Todos os três processos são comuns na natureza. O processo aditivo é usado nos monitores de computadores e televisões. O processo subtrativo aparece quando usamos um projetor para vermos slides em uma tela. E o processo de pigmentação nos permite ver cores em pinturas e trabalhos impressos. Em alguns processos industriais como impressão offset, mais de um processo pode ser utilizado.

Simplificando o que acontece em uma impressora podemos ver a subtração ocorrendo no papel branco, onde toda a luz é refletida, mas onde existe cor apenas a mesma será refletida e todo o resto será absorvido, ou “subtraído” da luz branca incidindo sobre o papel.

De qualquer forma temos um sistema receptivo de cor baseado em cones sensíveis vermelhos, verdes e azuis. Na escola primária aprendemos que as cores “básicas” são “vermelho”, amarelo e “azul”. Na realidade isso está relacionado com a mistura de tintas que poderia se aproximada por um processo subtrativo e esse vermelho na realidade era magenta, e o azul era ciano, cores complementares do vermelho verde e azul.

Nós iremos concentrar nossos estudos sobre o processo aditivo, pois a luz que chega ao nosso olho se combina aditivamente e todos os estudos de colorimetria também usam esse processo.

3.2. Amostragem e Reconstrução de Cor

Observando o modelo de sensores de cor do olho humano, vemos que podemos representar cor usando um número pequeno e finito de sensores, que são responsáveis por captar uma parte do espectro cada um deles. Isso é possível pelo fato de que o olho humano percebe um determinado conjunto de cores como sendo a mesma cor. Esse fenômeno é chamado de metamerismo.

Para podermos representar cor no computador precisamos saber como amostrar e reconstruir cor. No universo físico a amostragem de cor está associada aos receptores e a reconstrução está associada aos emissores.

Usando o mesmo modelo do olho o receptor deve ter um número finito de sensores, que poderiam ser 3, mas vamos apresentar de forma genérica usando N sensores. Matematicamente podemos modelar o receptor usando a curva de resposta espectral de cada sensor i , como $s_i(\lambda)$ e escrevendo:

$$C_i = \int C(\mathbf{I}) s_i(\mathbf{I}) d\mathbf{I}$$

onde C_i é a cor C percebida pelo sensor i .

Para que o emissor possa agora reconstruir a cor C ele deverá ter N sensores que estimulados corretamente produzirão a cor C . Matematicamente podemos modelar o emissor usando a base de primárias de cada sensor ($P_k(\lambda)$) e escrevendo:

$$C(\mathbf{I}) = \sum_{k=1}^N \beta_k \cdot P_k(\mathbf{I})$$

onde β_k são as componentes primárias da cor C .

O que gostaríamos de fazer agora é obter os β_k em função dos C_i .

Substitui-se C na equação do receptor pela fórmula do emissor para cada um dos sensores do receptor. Isso resulta em:

$$C_i = \sum_{k=1}^N \left(\beta_k \cdot \int P_k(\mathbf{I}) s_i(\mathbf{I}) d\mathbf{I} \right)$$

Só que na prática, as curvas de resposta espectral $s_i(\lambda)$ do receptor são muito complicadas de se obter, por isso precisamos de uma outra forma de reconstruir cor.

Primeiro, iremos normalizar as componentes β_k por uma cor de referência $W(\lambda)$ com objetivo de calibrar a escala de cores corretamente para cada sensor. O que resulta em:

$$W(\mathbf{I}) = \sum_{k=1}^N w_k \cdot P_k(\mathbf{I}) \quad T_k = \frac{\beta_k}{w_k}$$

onde T_k são as componentes primárias normalizadas da cor C .

Agora definimos as funções de reconstrução de cor $C_k(\lambda)$ do emissor em relação ao receptor de maneira que as componentes T_k são obtidas com:

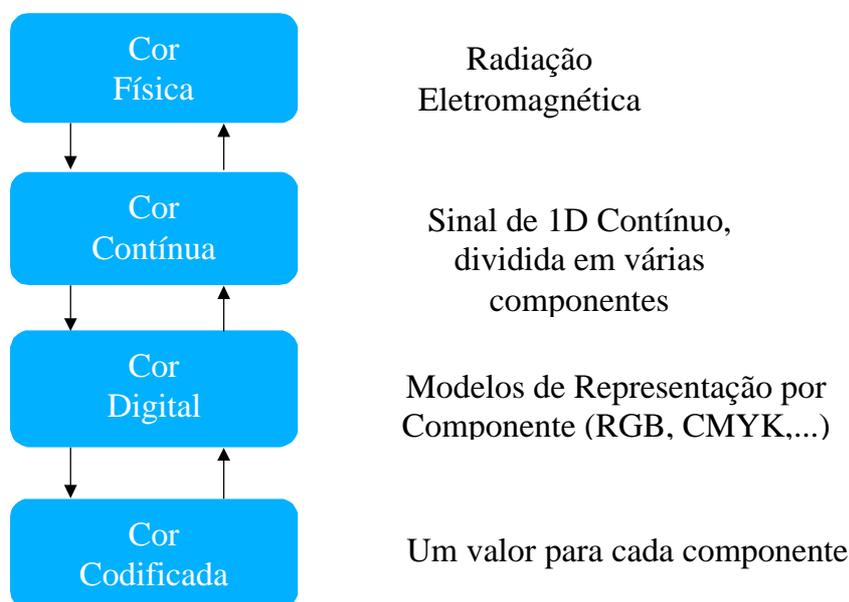
$$T_k = \int C(\mathbf{I}) C_k(\mathbf{I}) d\mathbf{I}$$

Não vamos entrar nos detalhes matemáticos (que já foram muitos até agora) da definição das funções de reconstrução de cor, mas a grande vantagem em usá-las é que em vários casos podem ser facilmente obtidas a partir de experimentos. Podemos interpretar as funções de reconstrução como sendo as coordenadas das cores puras (espectrais) vistas pelo receptor, escritas na base de primárias de cada sensor do emissor.

Quando o receptor é o olho humano essas funções são chamadas de “*Color Matching Functions*”. Mais tarde veremos um dos experimentos realizados para obtenção das funções de reconstrução de cor para o olho.

3.3. Cor no Paradigma dos 4 Universos

Uma vez que sabemos amostrar e reconstruir cor, podemos finalmente situar cor no paradigma dos 4 universos. Cor é uma onda eletromagnética que matematicamente pode ser representada por uma função unidimensional, mas que usamos um número finito, pequeno e discreto de sensores como no olho humano. Cada conjunto de sensores é chamado de modelo de representação, e cada sensor no modelo é chamado de componente.



Quando codificamos cor é muito comum usarmos 1 byte para cada componente. Um byte permite $2^8 = 256$ possibilidades, assumindo valores de 0 a 255. Por isso esses números aparecem tanto em programas de processamento de imagens e não sabíamos de onde vinham. Mas, repare que essa é apenas uma possível representação, por ser a mais comum induz a algumas pessoas pensarem

que é a única, e pior, algumas pessoas pensam que cor no computador se resume aos tais valores de 0 a 255.

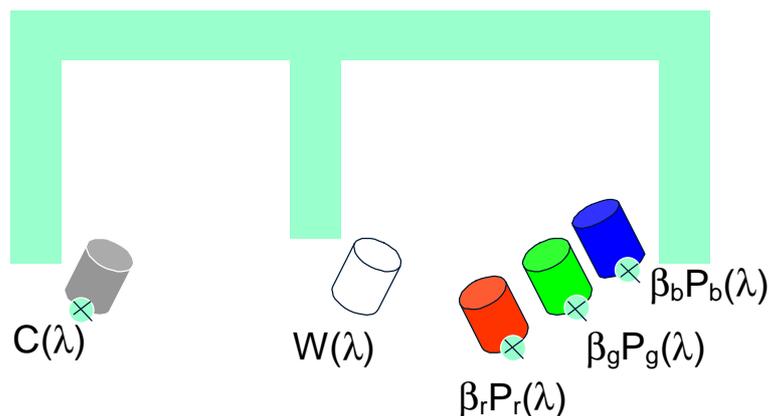
Se assumirmos que são 3 componentes isso totalizaria 2^{24} bits, ou 16 milhões de possibilidades. Alguns experimentos mostram que somos capazes de reconhecer aproximadamente 400 mil cores diferentes, o que na prática resulta em um número menor que 2^{19} bits, ou seja 2^{24} está mais do que suficiente levando em conta os diversos metamerismos.

O modelo de representação de cor que nós vimos é completamente genérico e para que os programas de processamento de imagens possam conversar entre si e interagir melhor com o usuário é preciso estabelecer padrões.

3.4. Padronização para Cor

CIE-RGB

O primeiro padrão baseado nesses conceitos foi estabelecido em 1931 pelo comitê CIE (*Comissão Internacional de l'Éclairage*). Foi chamado de CIE-RGB pois define as primárias $P_r(\lambda)$, $P_g(\lambda)$ e $P_b(\lambda)$ do emissor como sendo 3 cores espectrais. O padrão também define o branco como $W(\lambda) = 1$, e com isso foi realizado um experimento de cor para definir o Observador Padrão, ou seja, obter as funções de reconstrução de cor para esse sistema.



Para tanto, colocou-se 3 luzes monocromáticas baseadas nas primárias definidas, acopladas a um mecanismo que possa variar a intensidade de cada uma delas ($\beta_r P_r(\lambda)$, $\beta_g P_g(\lambda)$ e $\beta_b P_b(\lambda)$). Colocou-se então uma luz branca de referência também já especificada. Ajustando as três luzes de

maneira que juntas produzam a luz branca de referência, guardou-se os valores (w_r , w_g , w_b) para fazer a normalização dos ajustes de cada uma das luzes.

Agora, colocou-se uma outra luz monocromática $C(\lambda)$ com um mecanismo que permita variarmos seu comprimento de onda. Para um determinado número de comprimentos de onda do espectro visível ajustamos a luz a cada comprimento de onda e ajustamos as três luzes de maneira que juntas fiquem perceptualmente idênticas cada uma dessas luzes. Os valores obtidos com $c_r = \beta_r/w_r$, $c_g = \beta_g/w_g$, $c_b = \beta_b/w_b$ para cada um dos comprimentos de onda formam exatamente as funções de reconstrução de cor do sistema CIE-RGB com o receptor sendo o olho humano.

Isso foi feito para um grande número de pessoas escolhidas por sua competência na distinção de cores. Mas durante a experiência descobriram que determinadas cores não conseguiam ser ajustadas pelas três luzes. A solução foi mudar uma das luzes de lugar passando-a para o outro lado, podemos interpretar isso como uma subtração ou melhor um coeficiente β negativo.

Quando visualizamos todas as cores visíveis em 3 dimensões temos o chamado sólido de cor, que neste caso não é um cubo, possui a forma semelhante a de um cone de duas pontas.

CIE-XYZ

O fato de existirem valores negativos no padrão tem uma série de desvantagens, além disso várias outras desvantagens desse sistema fez com que o comitê criasse um outro sistema, baseado nos experimentos do CIE-RGB, mas com algumas características interessantes: todas as componentes das cores visíveis não devem ser negativas, duas primárias devem ter luminância zero (uma das componentes será a própria luminância), um maior número de cores espectrais deve possuir pelo menos uma componente zero. A solução foi buscar uma base de primárias fora do visível, ou seja o sistema também representa cores que não são visíveis, mas tudo bem pois basta não considerá-las.

Fazendo uma transformação das funções de reconstrução de cor do sistema CIE-RGB para o CIE-XYZ temos o gráfico da figura a seguir. Nela vemos também um segundo experimento realizado em 1964 que variou um pouco apenas em relação ao primeiro experimento. Também podemos perceber que a curva de luminância (Y) é muito intensa no verde, mostrando que percebemos mais variações de verde do que de outra cor. O “calombo” na curva da componente X vem das componentes negativas que haviam no padrão CIE-RGB. E por fim reparem que a curva do Z é muito estreita, ou seja podemos usar menos memória para guardar essa componente (isso será usado mais tarde quando falarmos de alguns algoritmos de quantização de cor).

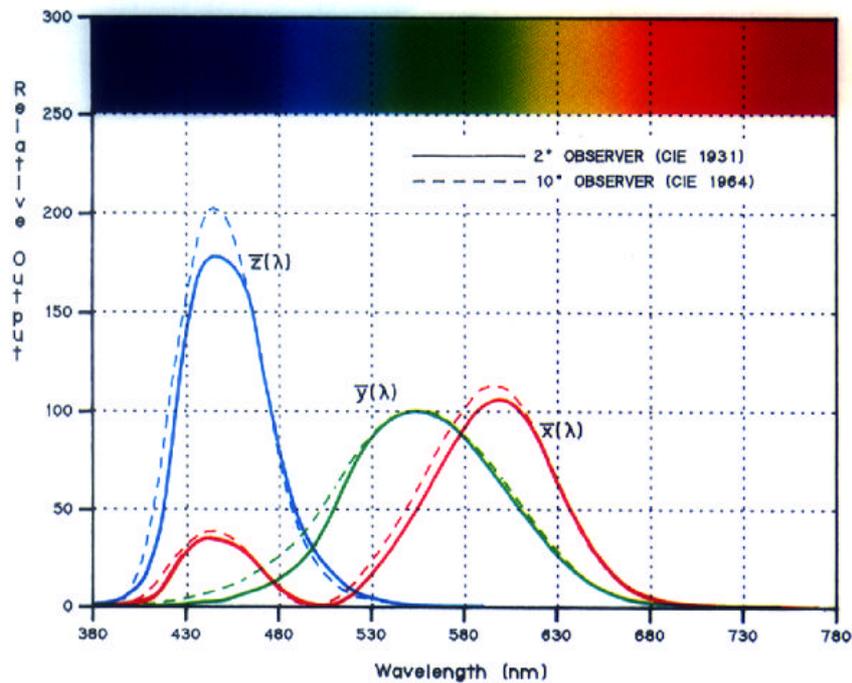
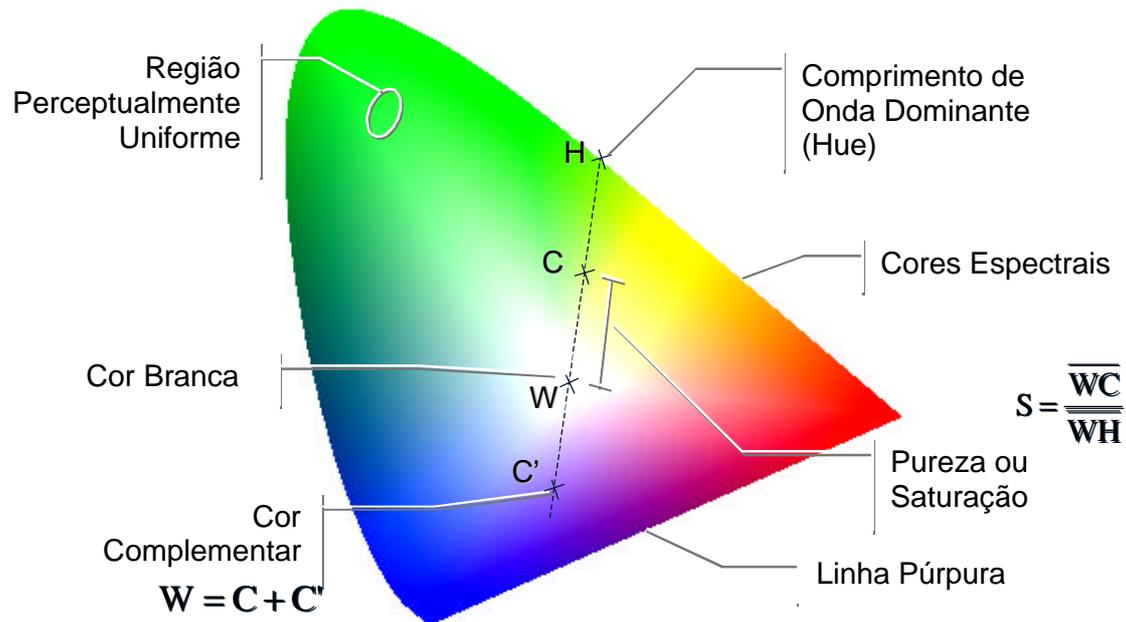


Diagrama de Cromaticidade

Já que luminância só é referente a intensidade da cor, criou-se uma visualização das cores visíveis que envolve somente as componentes X e Z de alguma forma. Na realidade é como se pegássemos todas as cores e projetássemos em um plano de luminância constante Y (Observação: de fato o diagrama é mais fácil de ser entendido por estarmos usando o sistema CIE-XYZ, mas ele foi originalmente definido para o sistema CIE-RGB).

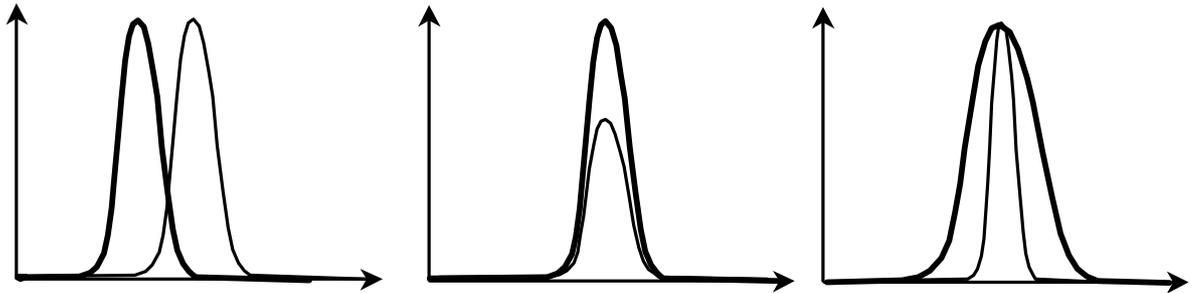
O resultado é a figura a seguir chamada de Diagrama de Cromaticidade. Nele podemos ver uma série de propriedades interessantes de cor.



As cores na borda do diagrama são todas as cores espectrais, cujas componentes são exatamente as funções de reconstrução de cor. No centro temos o branco W a partir do qual temos algumas propriedades. Dada uma cor C, a cor pura, ou o comprimento de onda dominante de C, ou melhor o matiz H de C é dado pela projeção de W, passando por C, na borda do diagrama. A saturação é simplesmente a razão entre as distâncias de W à C e de W a H, ou melhor o quanto uma cor se aproxima da cor pura ou se afasta da ausência de cor (se aproxima do branco). A cor C' que somada a C resulta na cor branca é chamada de cor complementar de C.

No diagrama os dois extremos do espectro visível são conectados por uma linha artificial (que não são cores espectrais) chamada de linha púrpura.

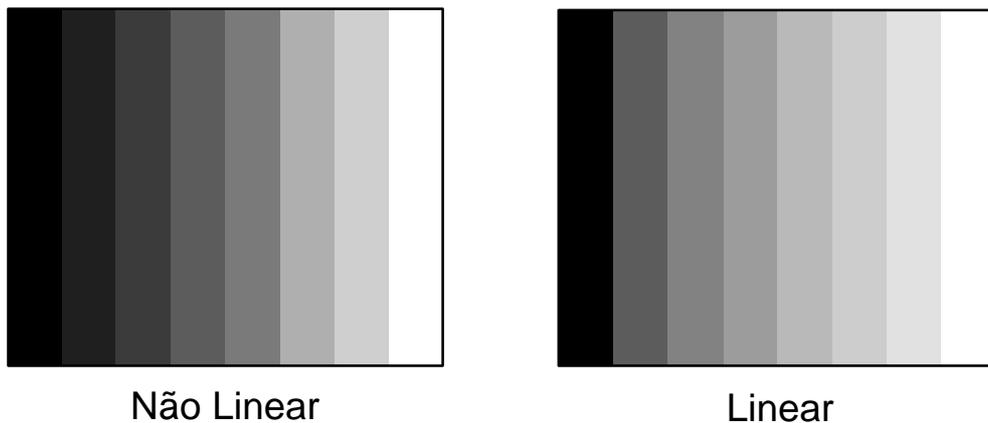
Podemos observar alguns desses fenômenos diretamente no espectro. Já vimos que uma cor monocromática possui uma forma estreita. Se trasladarmos esta forma de lugar estaremos modificando o comprimento de onda dominante ou o matiz. Se aumentarmos sua intensidade estamos aumentando seu brilho. E se alargarmos sua forma estaremos diminuindo sua saturação (se aproximando do branco, ex: tons pasteis).



Um outro aspecto interessante sobre o brilho é que embora este varie de acordo com as condições de iluminação, o olho possui uma capacidade enorme de se adaptar a essas variações. É como se ele mudasse a escala de percepção para poder captar o máximo de detalhes com a luminosidade disponível. A tabela abaixo mostra valores típicos de iluminamento em LUX que variam enormemente e mesmo assim conseguimos “enxergar”.

Ambiente	Valores (lux)
Luz do dia (máximo)	100.000
Luz de dia nublado	10.000
Interior próximo a janela	1.000
Mínimo p/ trabalho	100
Lua cheia	0,2
Luz das estrelas	0,000.3

O olho também possui uma resposta não linear à intensidade. Isso pode fazer com que a escala de tons percebida pelo olho seja não linear em determinados casos, ou seja a gama de tons percebida não corresponde a variações equivalentes de intensidade. Para compensar esse problema realiza-se a chamada Correção Gamma, que nada mais faz do que aplicar a função inversa da resposta do olho para obter uma escala percebida linearmente. Podemos ver as duas situações na figura a seguir.



Um outro experimento de cor verificou que cores próximas perceptualmente se situam em uma elipse no diagrama de cromaticidade. Isso foi usado para definir um outro sistema de cor com características perceptuais.

3.5. Sistemas de Cor

Nós vimos até agora 2 sistemas de cor CIE-RGB e CIE-XYZ, mas existem outros sistemas padronizados ou não padronizados.

Sistemas Perceptualmente Uniformes

A diferença entre duas cores no sistema CIE-RGB pode ser calculada usando-se a definição de diferença euclidiana:

$$d = \sqrt{(r_1 - r_2)^2 + (g_1 - g_2)^2 + (b_1 - b_2)^2}$$

Mas duas cores que são metaméricas podem possuir uma diferença numérica razoável entre si. A idéia é possuir um sistema que usando a mesma equação acima identifique cores

metaméricas e não as distingua numericamente. Poderíamos então, deformar o sistema XYZ para que as elipses sobre o diagrama de cromaticidade, mencionadas anteriormente, se tornem círculos.

Assim, foram criados dois sistemas padrão em 1976, o CIE Luv, que estabelece uniformidade no diagrama de cromaticidade, e o CIE Lab que estabelece uniformidade em um espaço de cor tridimensional. L em ambos os casos está relacionado com a luminância, mas também de uma forma perceptual, e é chamado de *Lightness*.

Sistemas de Vídeo Componente

São sistemas baseados em uma decomposição das componentes RGB em luminância e crominância, onde as componentes de crominância são obtidas subtraindo-se a componente Y das componentes RGB e eliminando-se a componente G pois, como vimos, ela é a que mais colabora na luminância. Assim as componentes ficam Y, R-Y e B-Y.

Os sistemas Y_{pbpr} (Sony Betacam) e $YCbCr$ (Vídeo Digital) são baseados nessa decomposição.

Sistemas de Vídeo Composto (NTSC, PAL, ...)

São baseados na mesma decomposição Y, R-Y e B-Y dos sistemas de cor de vídeo componente, mas para serem transmitidos pelas redes de televisão todas as componentes desses sistemas precisam ser combinadas em um único sinal, por isso são chamados de sistemas de vídeo composto.

Transmissões de vídeo usam os sistemas YUV ou YIQ. Nesses sistemas o sinal de luminância compõe 95% do sinal e a informação de crominância apenas 5%. Isso é possível devido a menor sensibilidade do olho a variações de crominância. Além disso, as televisões preto e branco foram criadas antes das televisões coloridas, o uso dessa codificação permitiu que as televisões preto e branco continuassem a receber o mesmo sinal de televisão mesmo sendo incluída a informação de crominância, que nesse caso é ignorada, pois o sinal não aumentou de tamanho. Esse é um dos grandes milagres da indústria de televisão.

Dentro das siglas de sinais de televisão a letra em destaque (M ou G, etc...) informa o padrão de luminância, a abreviação inicial indica o padrão de cor. Porque o sistema americano e o brasileiro só diferem, essencialmente na cor, é possível reproduzir uma fita de vídeo gravada num sistema em um aparelho de vídeo cassete de outro, e vice-versa, ficando apenas a cor prejudicada.

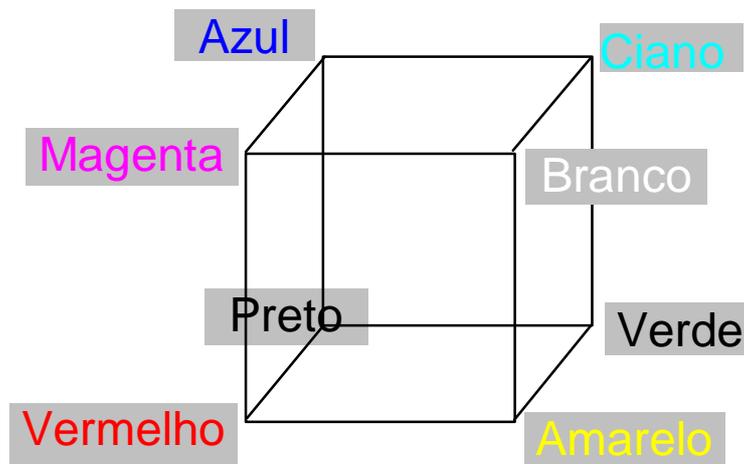
Assim como, fitas européias não conseguem ser reproduzidas aqui no Brasil pois diferem na codificação de luminância.

Sistemas de Dispositivos

São os sistemas usados pelos monitores (mRGB) e scanners (sRGB), cada dispositivo possui o seu próprio sistema. Para permitir a calibração adequada o fabricante deve fornecer os valores das primárias dos fósforos do monitor, por exemplo.

Quando dizemos que uma cor possui componentes $RGB=(255, 0, 36)$, isso não está amarrado a nenhum referencial de cor e perceptualmente pode estimular completamente diferente o olho mudando-se de dispositivo. Para que esses valores tenha significado temos que associa-los a algum sistema de cor padrão, em geral o CIE-XYZ. Por isso a necessidade de calibrarmos o monitor para aplicações que exigem alta qualidade em editoração eletrônica.

Esses sistemas definem exatamente um cubo, visto a seguir:



Por isso, a projeção deles no diagrama de cromaticidade é um triângulo, veja em Sistemas de Gerência de Cor.

Sistemas Complementares

São sistemas usados para os casos que envolvem os processos de formação de cor de forma subtrativa. Por isso são chamados também de Subtrativos, isso gera uma certa confusão pois na realidade imitam o processo de formação de cor subtrativo de uma forma aditiva.

A combinação de RGB em quantidades iguais leva ao branco. Se mudarmos os eixos de referência de lugar rotacionando-os, temos o sistema CMY. Combinando CMY em quantidades iguais leva ao preto. Por isso simula o processo subtrativo.

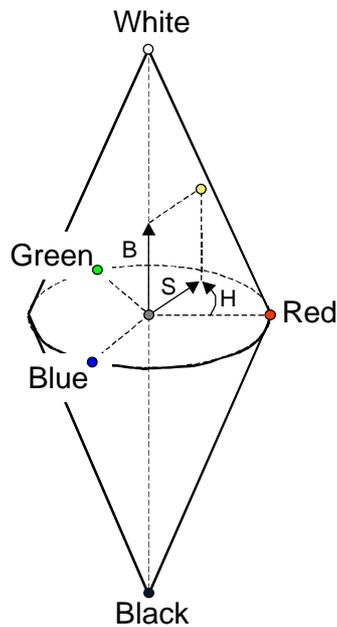
No caso específico de impressoras combinar CMY para obter preto tem uma série de desvantagens, pois precisamos das três cores ao mesmo tempo sobre o papel, gastando muita tinta e mesmo assim não conseguindo um preto adequado, além do que torna o processo de secagem da tinta mais demorado, enrugando o papel. Como tinta preta é muito barata, adiciona-se uma quarta componente que é responsável somente pelo preto K (black). O cálculo dessa componente é feito diretamente a partir das componentes CMY, que posteriormente são alteradas para não conterem informação de preto. Esse cálculo em geral é muito complexo, e é necessário para se obter resultados de alta qualidade gráfica.

Sistemas de Interface (por Coordenada)

São usados para especificar cores em programas. Selecionar valores no sistema RGB é muito pouco intuitivo, embora seja como o olho percebe cor. Utilizando as propriedades de Matiz, Saturação e Brilho (ou alguma grandeza relacionada com a luminância) é muito mais fácil para o usuário especificar um cor e depois obter o valor RGB.

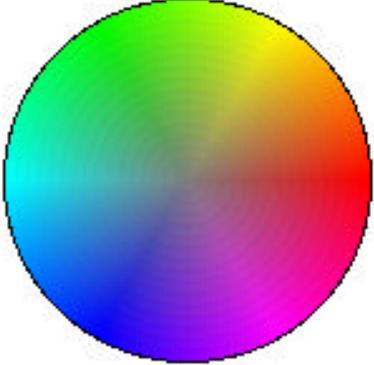
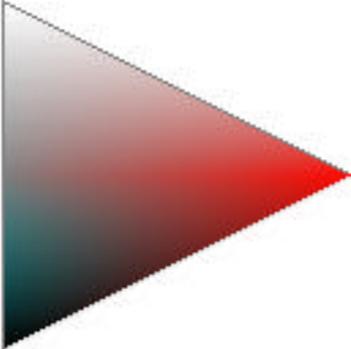
Para isso criou-se uma abstração para o cubo RGB deformando-o em um cone com duas pontas. Essa deformação é extremamente não linear, mas é uma forma simples de representar juntas as três propriedades.

O cone é construído tomando-se o eixo de brilho como o eixo maior do cone, e perpendicular a esse eixo define-se o plano de crominância. Neste plano os matizes possíveis são colocados em um círculo, a linha púrpura não é representada. Quanto mais próximo ao eixo de brilho menor a saturação, mais afastado maior a saturação. No fundo é uma representação simplificada do que está no diagrama de cromaticidade.



Como antigamente precisava-se que a conversão HSB-RGB fosse feita muito rápido, criaram-se aproximações do sistema HSB, onde o cone é substituído por um hexágono (HSL) e pode também possuir apenas uma ponta (HSV).

A tabela abaixo mostra dois cortes no sistema HLS, repare como a intensidade da cor varia para um mesmo valor de L (L é referente à Lightness, mas não tem relação direta com os sistemas CIE-Lab ou CIE-Luv). Pode-se utilizar um desses cortes, e mais um parâmetro escalar para criar a interface com o usuário para selecionar cor. No primeiro caso o usuário escolhe o valor de L e depois escolhe os valores de H e S diretamente no plano de L constante. No segundo caso o usuário escolhe o valor de H e depois os valores de L e S no plano de H constante.

Corte no plano de Lightness = 50%	Corte no plano de Hue = 0°
	

Sistemas de Interface (por Amostra)

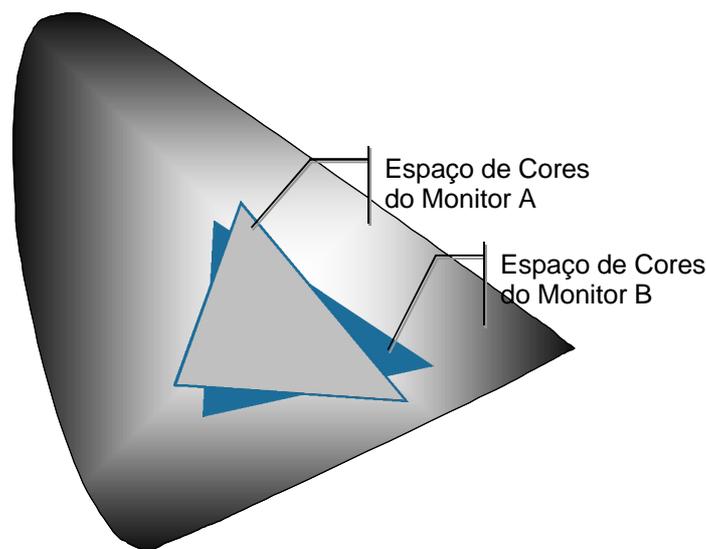
São sistemas usados manualmente por usuário para especificar cor, só que são baseados em conjuntos finitos de amostras de cor, em uma espécie de catálogo.

O primeiro sistema criado foi o de Munsell em 1905, que podemos ver na figura a seguir. Mas existem vários outros: Ostwald, Pantone, Truematch ,Focoltone, ...



3.6. Sistemas de Gerência de Cor

Se nós representássemos todas as cores possíveis em um monitor de computador veríamos um triângulo no diagrama de cromaticidade. Se pegássemos um outro monitor do mesmo fabricante veríamos um outro triângulo numa posição ligeiramente diferente do anterior. Se fosse de um outro fabricante a discrepância pode ser ainda maior. No caso de uma impressora e um monitor a diferença é enorme, já que a região da impressora no diagrama possui até um formato diferente.



Assim vemos que é necessário um mecanismo que nos permita controlar a mudança de um dispositivo para outro, de forma que possamos representar melhor as cores de interesse em cada um desses dispositivos.

Já que existem diversos padrões de cor, ou melhor diversas representações da mesma cor, nada mais natural do que existir um meio de converter de um sistema para outro. Os programas responsáveis por isso são chamados de Sistemas de Gerência de Cor (CMS - *Color Management Systems*) e realizam a conversão entre os diversos sistemas procurando o mínimo de perdas, ou seja o maior número de correspondências possíveis.

Esses programas estão preocupados justamente com a questão de portabilidade de cor, pois quando estamos mudando de um sistema para outro em geral é porque estamos mudando de um dispositivo para outro.

Aqui está um dos grandes pontos da questão de manipulação de cor no computador. Existem diversos desses programas no mercado, mas são caros e precisam de pessoal especializado, além de equipamento adequado para medição de cor. Para que o par Captura-Visualização funcione, todos os dispositivos no seu ambiente de trabalho precisam estar calibrados.

Poucos sistemas operacionais se preocupam diretamente com isso. O pioneiro na área é o MacOS (roda em computadores Macintosh), que desde sua criação já se preocupava com isso. Recentemente, um esforço conjunto de várias empresas levou a criação do padrão ICM (Image Color Matching) para suporte a gerência de cor em aplicações. O Windows 95 já vem com esse recurso e com o lançamento do Windows 98 ficou ainda mais direto seu uso.

O princípio de funcionamento de um sistema de gerência é a criação de perfis de dispositivo que a partir de informações do fabricante e do processo de calibração consigamos obter resultados ótimos, independente do dispositivo usando a informação que está no perfil para realizar a conversão para aquele sistema.

Quando estamos no ambiente de editoração eletrônica esse problema fica óbvio. Se o logotipo de uma determinada marca é um tal vermelho. Esse vermelho tem que parecer o mesmo em qualquer mídia que o logo seja impresso ou visualizado.

3.7. Quantização com Menos de 24 bits

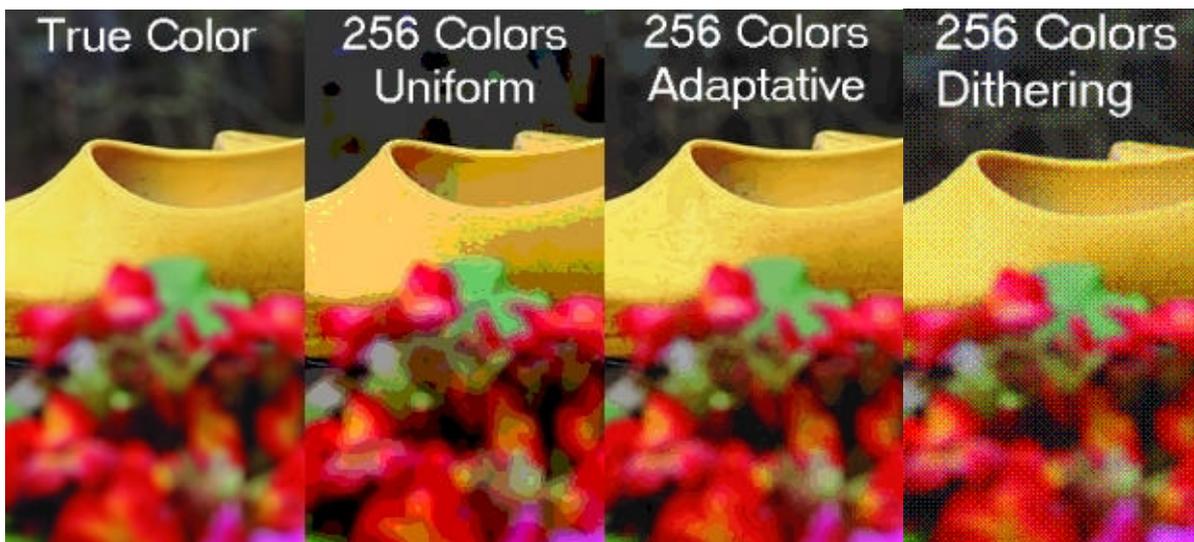
Como vimos bastam 24 bits por pixel (bpp) para visualizarmos uma imagem no computador com ótima qualidade. Mas e se o nosso computador não tiver 24 bpp (16 milhões de cores)? A grande maioria das impressoras imprime somente com 4 (quatro) cores... Uma outra motivação para reduzir o número de bpp é simplesmente para economizar memória, como se fosse uma técnica de compressão de dados com perdas.

Mesmo quando existe um número reduzido de cores disponíveis para visualização na tela do computador, a representação de cada cor ainda usa 8 bits por componente. Para reduzir o tamanho de cada pixel, substitui-se sua cor por um índice para uma tabela de cores de 8 bits (veja no tópico de Armazenamento).

Para essa redução de tamanho ser efetiva o número de cores da imagem quantizada deve ser menor que 8 bpp, ou no máximo 256 cores diferentes. Portanto, precisamos encontrar dentre as cores existentes na imagem quais são as 256 mais significativas e dada uma cor da imagem qual das 256 vou escolher para substituí-la.

Existem duas técnicas muito conhecidas. A quantização Uniforme é a mais simples, e usa uma tabela de cores fixa que subdivide o cubo RGB, em 256 partes igualmente espaçadas. Dado que uma cor está em uma dessas partes é trivial encontrar a cor equivalente na tabela. Esse algoritmo é excepcionalmente rápido, mas não possui resultados muito bons quando a imagem possui muitas cores. A quantização Adaptativa é mais complexa e usa uma tabela de cor adaptada as cores presentes na imagem. A técnica mais comum é a do corte mediano (Median Cut) que procura dividir o cubo de maneira que, onde existem mais cores existirão mais divisões, ou melhor procura fazer com que cada divisão tenha aproximadamente o mesmo número de cores. Esse algoritmo é um pouco mais lento e produz resultados muito bons.

A imagem a seguir mostra as duas técnicas de quantização, usando uma palheta de cores uniforme e usando uma palheta adaptativa. Além da imagem original e de uma imagem quantizada com palheta uniforme, mas dessa vez usando também a técnica de dithering vista no próximo tópico.



Dithering

No caso de impressoras de 4 cores cada componente CMYK de 8 bits é quantizada para 1 bit (2 cores, onde há tinta ou onde há o fundo branco do papel). Pode se mostrar que nesse caso reduz-se o problema de quatro componentes para imprimir uma componente de cada vez.

A idéia é fazer com que, embora existam apenas 2 cores, o olho perceba meio-tons. Técnicas de meio-tons são usados na indústria desde o fim do século passado.

No computador a técnica mais comum, utilizada para esse propósito, é a de dithering, cujo objetivo é minimizar perceptualmente os efeitos gerados pela quantização de contornos. Embora, essa técnica possa ser usada em várias situações de quantização (como vimos no exemplo anterior), ela é muito importante no processo de quantização para 1 bit. Inclusive alguns métodos de dithering são baseados em técnicas de meio-tons da indústria, e por isso dithering também é conhecido com meio-tom digital.

O funcionamento básico de um algoritmo de dithering é calcular o erro de cor gerado pela quantização de um determinado pixel, e quando existe uma mudança abrupta entre dois níveis de cores quantizadas, substitui-se essa mudança por uma mistura dos dois níveis, com se estivéssemos espalhando o erro de quantização em uma vizinhança do pixel quantizado.

Existem muitas técnicas de dithering implementadas, tais como Bayer, Floyd-Steinberg, difusão de erro, etc.

4. Captura

Existem diversos mecanismos de captura de imagens. Os mais conhecidos são o Scanner, o Digitalizador de Sinal de Vídeo (Frame/Vídeo Grabber) e a Câmera Digital.

4.1. Scanners

Scanners provem as imagens de mais alta resolução, mas em compensação são mais caros. Podem capturar diversos tipos de superfície, as mais importantes são papel e filmes fotográficos. Utilizam o sistema de cor RGB e possuem uma calibração fácil de ser feita. O procedimento de calibração requer um modelo impresso, em geral fornecido pelos programas que aceitam calibração. Esse modelo contém uma série de cores, conhecidas ao programa. Assim, quando a imagem é capturada o programa compara os valores obtidos com os valores esperados e pode então gerar um perfil de correção para aquele scanner.

Os scanners são muito semelhantes à máquinas Xerox (fotocopiadoras). Existem basicamente três tipos: manuais, de mesa ou de cilindro. Em qualquer caso o scanner dispõe de uma fonte de luz em forma de uma linha que varre a imagem impressa e mede a quantidade de luz refletida ou transmitida em cada ponto. A luz captada é convertida em um sinal elétrico através de um conjunto de foto-detetores que também formam uma linha. O sinal elétrico é finalmente digitalizado e enviado ao computador.

No caso de scanners manuais a varredura é feita deslizando o aparelho sobre a imagem a ser capturada. No de mesa, em processo análogo ao da Xerox, a varredura é realizada automaticamente pelo aparelho que desloca a fonte de luz e os foto-detetores. No de cilindro tanto a fonte de luz quanto a imagem se movem para fornecer maior qualidade ao processo. Neste caso a fonte de luz se move dentro do cilindro enquanto a imagem é fixada na sua superfície que gira durante o processo.

A escolha da DPI a ser utilizada é o fator mais crítico no uso do scanner. A primeira coisa que se deve ter em mente é "Qual o objetivo ao digitalizar a imagem?". Uma solução é encarar três tipos de resultados: para visualização somente em monitores, para edição e impressão futura, e para análise científica.

No primeiro caso, a resolução máxima em pixels do monitor vai ditar o tamanho máximo em pixels da imagem. Imagem capturadas para o uso na Internet se encaixam neste caso.

No segundo caso, a maior DPI da impressora a ser utilizada vai ditar a DPI a ser utilizada no scanner durante a captura. Mas, no caso de impressoras a 2 cores, deve-se compensar o tamanho da matriz de dithering. Por exemplo, se for 6x6 em 600 DPI, basta enviar uma imagem de 36 tons de cinza em 100 DPI.

No terceiro caso, onde o objetivo é obter medidas na imagem, o que vai indicar a resolução é a precisão dos valores que se deseja obter no espaço da imagem. Esta precisão indicará a menor distância entre dois pixels, o que por consequência ditará a DPI.

4.2. Digitalizadores de Sinal de Vídeo

Digitalizadores de Sinal de Vídeo dependem do padrão de vídeo utilizado, NTSC, PAL-M, etc. Em geral são de baixa resolução. Alguns permitem capturar um único quadro, chamados de *Frame Grabbers*, outros permite gravar seqüências inteiras em tempo real, são chamados de *Vídeo Grabbers*. Algumas placas gráficas já incorporam, além da saída para o monitor de vídeo, a saída para vídeo analógico compatível com os padrões conhecidos, e possuem também entrada para vídeo analógico e para RF, ou seja, pode-se ligar o vídeo cassete no computador para capturar e gerar novos vídeos, assim como usar o computador como sintonizador de TV (sinal de Rádio Frequência).

Uma imagem de vídeo gerada por uma câmera ou um videocassete pode ser digitalizada diretamente a partir do sinal elétrico que a representa, que em geral está disponível como uma saída do equipamento.

O sinal de vídeo é organizado em linhas de imagem, que contém um sinal analógico cada uma. Portanto, o número de linhas da imagem já é praticamente definido pelo próprio sinal de vídeo; para os sistemas NTSC e PAL-M esse valor é 525 e para o sistema PAL-G é 625. Dentre estas apenas aproximadamente 480 são visíveis (na realidade somente de 240 a 360 linhas do sinal de vídeo contém informação relevante).

O número de colunas é normalmente escolhido de acordo com o número de linhas para manter a razão de aspectos da imagem. A escolha do número de colunas também é limitado pela banda do canal de vídeo, que para o sinal padrão NTSC é da ordem de no máximo 750 pixels por linha. Nas televisões comerciais existem cerca de 440 pontos numa linha horizontal.

A calibração em alguns casos é feita ajustando-se o nível de branco e o nível de preto. Além disso, é preciso não esquecer de regular a própria câmera, realizando o que se chama de "bater o branco", ou seja, para uma determinada luz informar para a câmera o nível de branco para a gravação. As câmeras domésticas tentam realizar essa calibração automaticamente.

Quando se está trabalhando diretamente com o conjunto câmera/digitalizador, após a captura da imagem, pode-se tentar melhorar a qualidade da mesma com um método muito simples. Com a câmera desfocada e/ou com um fundo neutro, como um papel cinza, captura-se uma imagem, esta imagem ao ser subtraída da imagem normalmente capturada elimina sujeiras que porventura estejam na lente ou em um fundo irregular advindo de problemas de alinhamento do equipamento, etc....

4.3. Câmeras Digitais

Câmeras Digitais usam um dispositivo chamado CCD (Charge Coupled Device) que atua como se fosse o filme fotográfico. Alcançam resoluções de 640x480 até milhares de pixels. Os Line Scan CCDs são usados para altas resoluções, com alta linearidade, alta banda dinâmica e readout lento e possuem qualidade excepcional. O sistema de cor é RGB. A calibração pode ser feita exatamente como um scanner, mas a qualidade da imagem agora depende de fatores externos, como iluminação do objeto.

As câmeras de vídeo domésticas que conhecemos utilizam esse dispositivo. A limitação da imagem gerada por estas câmeras está associada ao padrão de vídeo analógico utilizado para armazenar o vídeo na fita.

Mais recentemente com a padronização de vídeo digital, já pode-se armazenar a imagem capturada com qualidade muito superior a do vídeo tradicional. Inclusive algumas câmeras utilizam 3 CCDs para capturar a imagem, um para cada componente de cor R,G e B.

5. Visualização

Há duas formas básicas de visualização, o monitor e a impressora, embora existam outros dispositivos, como por exemplo gravadores de slides.

5.1. Monitores

Atualmente existem diversas tecnologias de monitores, os mais comuns funcionam como as televisões comerciais usando tubo de raios catódicos (C.R.T.). Mas a distância entre os centros dos três fósforos RGB do monitor é de 0,21mm e da TV é de 0,60mm (essa distância é chamada de dot pitch). São chamados de dispositivos matriciais de rastreamento, onde a imagem precisa ser reescrita na tela a cada determinado intervalo de tempo (isto é um dos motivos das imagens digitais serem chamadas de raster images). Por consequência, a imagem gerada para toda a tela do monitor precisa ser guardada em algum dispositivo; este dispositivo é a placa gráfica.

O monitor utiliza o sistema RGB e possui um número determinado de pixels na horizontal e na vertical. Mas a cor é contínua, depende apenas das cores fornecidas pela placa gráfica. Esta irá dizer se estamos trabalhando em 640x480 com 256 cores ou 1280x1024 com 24 bits, por exemplo. Assim, a visualização no monitor é dependente do melhor ajuste entre ele e a placa de vídeo. Geralmente possuem baixa resolução, aproximadamente 72 DPI.

Vimos que a imagem formada na tela, está sendo redesenhada a cada instante. Se este redesenho não é rápido o suficiente aparece o fenômeno chamado de *flickering*. A imagem parece estar piscando. Isso ocorre porque a frequência de rastreamento é muito baixa para a configuração escolhida. É preciso saber se, tanto o monitor, quanto a placa gráfica, suportam uma frequência aceitável para a configuração desejada. Em geral, valores acima de 75Hz são excelentes, e 60Hz é aceitável.

É interessante observar que a razão de aspectos de um pixel na tela é 1x1, mas um ajuste errado do monitor pode provocar distorções. As resoluções de tela são múltiplos de 4x3 (640x480, 800x600, 1024x768) para que justamente resultem em *pixels* quadrados. Para verificar isso, veja se a área de visualização no monitor está na razão 4x3 (cabe lembrar que a razão de uma tela de cinema é 2x1 e das novas televisões de alta definição (HDTV) é de 16x9). Um outra maneira de fazer a calibração de aspectos é simplesmente desenhar um quadrado em *pixels*, usando por

exemplo um programa simples de pintura, e medir esse quadrado com uma régua na tela do monitor, se estiver distorcido então deve ajustar o monitor até que o quadrado esteja correto.

A calibração de cor de monitores requer um dispositivo que é colocado na sua tela e retorna a informação de cor necessária para o programa de calibração. Mas neste caso, a calibração também depende da luz ambiente.

Mesmo não possuindo esse dispositivo pode-se fazer um ajuste mínimo do monitor. Primeiro, coloque na tela uma figura com uma grande área preta e outra branca. Coloque o controle  no mínimo. Reduza o controle  até que o fundo preto da tela fique correto e não com um certo brilho, mas não reduza demais, só o necessário (se aumentando esse controle, logo acima deste nível, a tela ficar com uma cor escura em vez de um cinza escuro, o seu monitor está com um problema de calibração de um dos canhões, deve-se levá-lo à um técnico). Agora, ajuste o controle  de maneira que o branco mostrado fique com uma intensidade adequada para você. Resista à tentação de colocar o brilho do monitor muito alto, isso tem uma série de desvantagens além de piorar a qualidade da imagem.

Os controles  e  são chamados, para monitores modernos, de “Nível de Preto” e “Imagem” (“Black Level” and “Picture”), respectivamente. O controle  ajusta o nível de preto, que faz com que toda a imagem fique mais clara ou mais escura, mas na realidade está maximizando o contraste. O controle  ajusta o nível de branco e por consequência todos os valores intermediários até o preto, sem modificar o nível de preto, isso está na realidade está ajustando o brilho. Por razões históricas, esses símbolos estão associados aos nomes de “Brilho” e “Contraste”, respectivamente, por causa da sua simbologia.

Uma curiosidade sobre monitores e televisão é que as TVs possuem qualidade muito menor que a dos monitores, mas que visualmente as pessoas ficam satisfeitas com o resultado. Isso acontece porque, quando um ponto é impresso na tela da TV, ele na realidade sofre um fenômeno de espalhamento fazendo com que a imagem fique com contornos mais suaves e conseqüentemente agradável para a vista. Além do que, a gradação tonal é muito mais suave por ser contínua, e no caso do monitor está limitada pela placa de vídeo.

5.2. Impressoras

Impressoras usam tecnologias muito variadas. Por isso, o processo de calibração é muito complexo e as vezes não gera bons resultados. Este depende de muitos fatores: papel, tinta, até umidade do ar. E existem uma grande variedade de dispositivos para fazê-lo. Em contra partida, impressoras chegam a altíssimas resoluções.

O sistema de cor na maioria das impressoras é CMYK, mas existem algumas poucas impressoras que usam o modelo RGB. Algumas impressoras operam apenas a 4 cores. Neste caso utilizam resolução espacial para compensar a baixa resolução cromática, com por exemplo técnicas de dithering vistas no tópico de Quantização de Cor.

Por isto existem duas grandes problemáticas envolvendo impressão de imagens. A primeira é relacionada a fidelidade da cor impressa, pois necessariamente precisamos converter de RGB para CMYK. A segunda é que a grande maioria das impressoras imprime somente com apenas 4 cores, que não são combinadas. Assim, calibração de cor e as técnicas de dithering são fundamentais para obtermos resultados de alta qualidade.

As impressoras de jato de tinta possuem resoluções espaciais inferiores as das impressoras laser, mas em contrapartida possuem resolução cromática muito melhor, podendo reproduzir até 16 milhões de cores, enquanto que a grande maioria das impressoras laser possuem apenas 2 cores, o branco do papel e o preto. Neste último caso, para imprimir tons de cinza, a impressora usa resolução espacial para simular tons enganando o olho, usando a proximidade entre os pontos desenhados.

Levando esses fatos em consideração, devemos prestar atenção durante a inserção de imagens em documentos para não utilizarmos tamanhos muito grandes que não vão ser aproveitados pela impressora e vão aumentar o tamanho do documento, assim como aumentar o tempo de impressão, desnecessariamente.

6. Armazenamento

Uma vez capturada a imagem, esta é armazenada na memória do computador como uma seqüência de bytes, ou melhor um mapa de bits (Bitmap). Este armazenamento é feito independente da resolução e número de cores do monitor (Device Independent Bitmap - DIB), embora alguns programas simples, como o PaintBrush do Windows 3.1, que usa um bitmap dependente da configuração de tela.

Como vimos anteriormente, a imagem digital é simplesmente uma matriz de pixels, onde cada pixel é um número (ou 3, etc...) que representa o valor da cor naquela posição.

Quanto maior o número de pixels, maior o tamanho da imagem na memória. Mas qual o tamanho de cada pixel?

Isso vai depender de obviamente do sistema de cor e do número de bits usado para representar cada componente. Se a imagem é monocromática (tons de cinza) terá apenas 1 componente, se for colorida no sistema RGB terá 3 componentes, e assim por diante.

Cada componente será quantizada por um determinado número de bits. Em geral, usam-se apenas 8. A partir da versão 4.0 o sistema Photoshop, por exemplo, passou a tratar imagens com 16 bits por componente. Cuidado para não confundir esse número com o valor de 16 bits de cor de algumas resoluções de display, neste caso usam-se 5 bits para cada componente RGB, e um bit é flexível, normalmente usado para dar ao verde 6 bits. Já em sistemas de processamento de imagens científicas é muito comum tratar imagens com números inteiros com 32 bits que incluem valores negativos, ou até mesmo números de ponto flutuante (valores com várias casas decimais de precisão).

Quando falamos de 8 bits, falamos de 2^8 (256) possibilidades de valores diferentes para cada componente. Assim, se são 3 componentes temos $2^8 \times 2^8 \times 2^8 = 2^{24}$ (também chamada de resolução de cor *True Color*), totalizando aproximadamente 16 milhões de cores possíveis. É claro que imagem alguma usa toda as cores disponíveis ao mesmo tempo.

Esta imagem mostra o efeito de reduzirmos a resolução e/ou reduzirmos o número total de cores de uma determinada imagem.



O tamanho em bytes de uma imagem pode ser calculado com:

$$\text{Tamanho} = \text{Largura} \times \text{Altura} \times \text{Profundidade} \times \text{BPC}$$

onde largura e altura são o número de pixels na horizontal e vertical, respectivamente; profundidade é o número de componentes do sistema de cor (RGB=3, CMYK=4, Tons de Cinza =1, ...); e BPC é o número de bits por componente, geralmente 8, mas pode ser 4 ou até mesmo 1. Um outro termo utilizado é BPP, bits por pixel que corresponde a Profundidade x BPC, totalizando o número de bits efetivo que cada pixel ocupa.

Vejamos alguns exemplos do cálculo do tamanho ocupado em memória por uma imagem:

Exemplo 1:

Imagem de 1024x768 com 24 Bpp = 2,25 Mb

Exemplo 2:

Imagem 15cm x 10cm capturada com 300 DPI em True Color (24 Bpp)

Largura = $15 \times 300 / 2,54 = 1771$

Altura = $10 \times 300 / 2,54 = 1181$

Tamanho = $1771 \times 1181 \times 3 = \sim 6 \text{ Mb}$

Exemplo 3:

A mesma imagem com 600 DPI

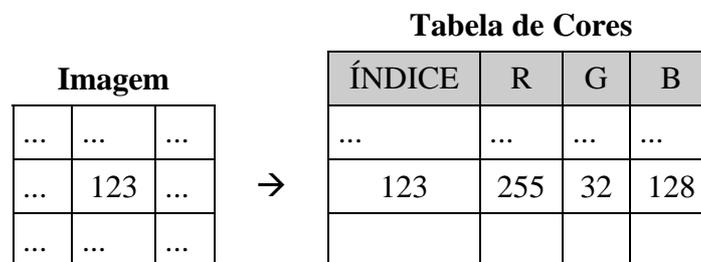
Tamanho = ~24 Mb

Exemplo 4:

Página A4 com 600 DPI = ~100Mb

Historicamente, computadores não possuíam uma interface gráfica tão poderosa quanto temos hoje. Assim, foram definidas simplificações para o modelo RGB, que embora diminuam significativamente o número de cores disponíveis, reduz o tamanho da imagem em memória. Essas imagens são chamadas de indexadas ou coloridas com palheta de cores (*indexed images, palette based images*).

Nessas imagens o valor do *pixel* contém um único número que corresponde a um índice para uma tabela, essa tabela é que conterà o valor da cor em RGB. Imediatamente podemos ver que as 3 componentes são reduzidas para 1, isso já é uma grande economia já que o tamanho da palheta é pequeno perto do tamanho da imagem. As imagens tons de cinza podem ser vistas como um caso particular onde a palheta contém os valores de R, G e B iguais.



6.1. Compressão

Observando os exemplos anteriores vemos que não há computador que resista a tanto pixel. Por isso foram criados, e ainda são muito estudados, métodos de compressão de imagem. Esse métodos podem ser classificados em dois tipos, com perdas e sem perdas. Os métodos com perdas são baseados no fato de que pequenos detalhes podem ser eliminados que perceptualmente não serão notados de imediato.

Dentre os métodos com perdas se destacam o JPEG (Join Photographic Experts Group), criado a partir de uma composição de várias técnicas e adotado amplamente. Um outro método

muito bom é chamado de Fractal, embora pouco use da matemática de fractais. Esse método infelizmente é privado de uma empresa que comercializa seus direitos de uso.

Dentre os métodos sem perdas, podemos enumerar vários: Código de Huffman, RLE (Run Length Encoding), LZW (Lempel-Ziv & Welch), JBIG (Join Bi-level Image Experts Group). Cada um desses métodos utiliza uma técnica diferente que em geral lhe dá o nome.

O mais simples de todos e também o mais fácil de entender, é o RLE. Esse método usa o fato de que imagens podem possuir muitos pixels repetidos em seqüência numa mesma linha. Assim, em vez de armazenar todas essas repetições, armazena-se o valor repetido e o número de repetições. Obviamente esse método só funciona bem para imagens com muitas áreas de uma mesma cor.

6.2. Formatos de Arquivo

Como a memória RAM do computador é volátil, precisamos de algum mecanismo de persistência para a imagem, ou seja, gravá-la em um arquivo. Os detalhes de como gravar imagens em arquivos não são universais e conseqüentemente existem vários padrões.

A tabela a seguir mostra os formatos de arquivo padronizados que já estão bem estabelecidos no mercado:

Formato	Sistema de Cor	Compressão
GIF	RGB 256	LZW
TIFF	RGB*, CMYK, YCbCr, Lab, Luv	RLE, LZW, JPEG, JBIG, Outros
JPEG	RGB, YCbCr, CMYK, Gray	JPEG
PCX	RGB*	RLE
BMP	RGB*	RLE
TGA	RGB*	RLE

Repare que o formato TIFF se destaca entre os demais. Atualmente ele é o formato mais recomendado para armazenar imagens que serão usadas em diversos programas, pois praticamente todos já o reconhecem e sua flexibilidade é enorme.

Dois outros formatos são importantes de serem mencionados, o GIF e o JPEG. Isto porque eles são utilizados como os padrões de imagem na Internet.

O formato GIF armazena apenas imagens RGB com 256 cores ou menos, sendo adequado para imagens de mais baixa qualidade e imagens tons de cinza. Esse formato permite que você defina uma cor que será ignorada pelo Browser, tornando parte da imagem que contém aquela cor, transparente. Uma outra característica interessante é que as linhas no arquivo podem ser gravadas entrelaçadas, fazendo com que o Browser receba algumas linhas de toda a imagem antes de chegarem todas as linhas, permitindo um efeito de que toda a imagem está sendo recebida em níveis de qualidade crescentes. Esse formato possui duas versões 87a e 89a, somente a segunda que permite transparência. O GIF usa compressão LZW para comprimir a imagem. Essa compressão é muito semelhante a usada no formato de arquivo compactado ZIP e possui ótimos resultados. Além disso, esse formato suporta várias imagens num mesmo arquivo, formando opcionalmente uma animação.

O formato JPEG, usa a compressão JPEG que lhe deu o nome, e permite gravar imagens Tons de Cinza, RGB, CMYK e outros sistemas de cor, mas sempre comprimidas com perdas. Embora tenha perdas é o melhor formato para colocar imagens de alta qualidade. Para Internet o efeito da imagem sendo recebida com níveis de qualidade crescentes é feito através de codificação progressiva (*Progressive Encoding*). Não possui transparência.

Um outro formato que está sendo padronizado para a Internet é o PNG (Portable Network Graphics). Cujas compressão é sem perdas e permite imagens com mais alta qualidade que o GIF. Além de não usar a compressão LZW por causa de um problema de direitos autorais com o GIF.

Um erro muito comum é tratar formatos de arquivos de imagem como tipos de imagem. Mas é preciso ficar bem claro que tipos de imagem não estão intrinsecamente associados à formatos arquivos, estes se prestam apenas ao armazenamento das mesmas. Portanto não faz sentido falar em “uma imagem TIFF” e “uma imagem GIF”, como imagens diferentes.

6.3. Imagens Grandes e Multiresolução

Mesmo com os métodos de compressão, para processarmos as imagens, precisamos descomprimi-las para a memória do computador. Quando as imagens são muito grandes comparadas ao tamanho da memória RAM disponível, temos que utilizar de outros recursos para poder processar tal imagem.

O primeiro deles é memória virtual, ou seja, usar disco como se fosse memória RAM. Como não precisamos de toda imagem ao mesmo tempo, deixamos uma parte dela no disco, quando esta parte é requisitada, carregamos a mesma para a memória e descarregamos para o disco partes que não estão sendo utilizadas.

Um outro recurso é utilizar o conceito de multiresolução. O que o próprio nome diz é que teremos a mesma imagem em diferentes níveis de resolução. Dependendo da resolução que o usuário estiver trabalhando, somente um nível é afetado. Quando o usuário muda de nível, as mudanças realizadas naquele nível são propagadas para o novo nível. Somente quando o usuário chega no último nível ou pede para salvar as alterações em disco, é que a propagação é feita para todos os níveis.

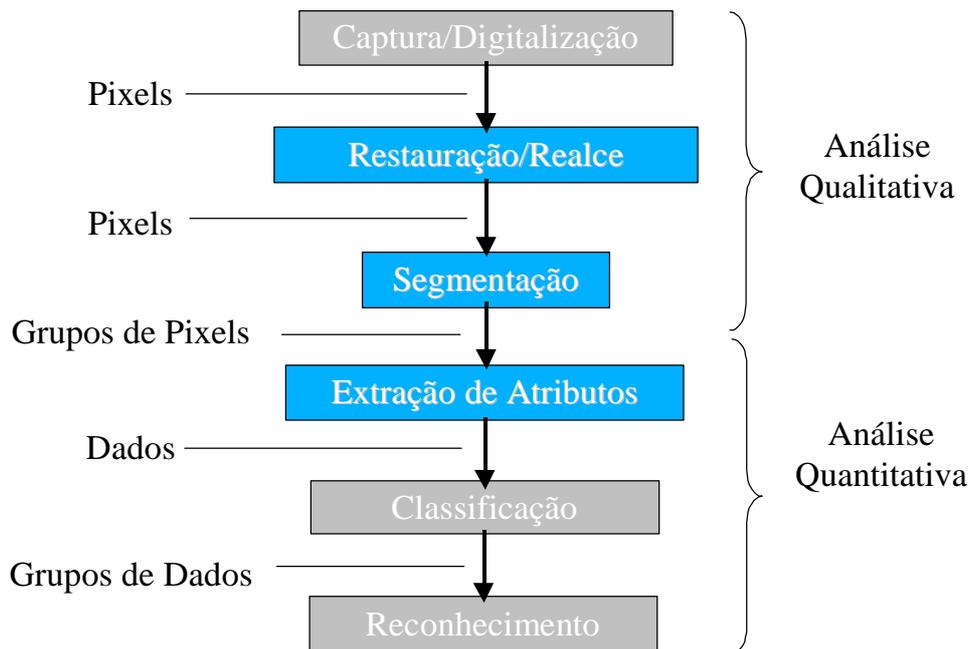
Esta imagem mostra os níveis de uma imagem em multiresolução.



Quando possuímos muitas imagens grandes é comum guardarmos uma versão reduzida para consulta rápida ao conteúdo do arquivo, esta versão reduzida é chamada de *thumbnail*.

7. Processamento

As funções de processamento de imagens digitais podem ser classificadas em duas classes, por escopo ou por resultado. Por escopo, está relacionado com técnicas de processamento. Por resultado, está relacionado com os tipos de resultado de cada técnica.



Através do diagrama mostrando uma seqüência tradicional para processamento da imagem, da captura à classificação, podemos ter uma idéia melhor das classes de processamento por resultado.

Há um fator comum em todas as classes de processamento, este fator é qualidade. Existem duas subdivisões em qualidade de imagem: fidelidade e inteligibilidade. No primeiro caso estamos preocupados em aproximar a imagem processada da imagem original ou de um padrão estipulado que a melhor represente. No segundo caso, nos preocupamos com a informação que conseguimos extrair da imagem, seja pelo olho humano, seja por algum processamento.

Programas para editoração eletrônica precisam ter um cuidado muito grande com a fidelidade na representação da cor da imagem, já programas de processamento de imagens científicas em geral não estão preocupados com a fidelidade da cor e sim a informação contida na imagem.

7.1. Restauração/Realce

Quando estamos no contexto de editoração eletrônica, ficamos normalmente situados apenas no escopo de operações de Realce ou Obscurecimento, e Restauração ou Deformação.

A restauração busca compensar distorções específicas, normalmente gerada no momento de aquisição. Quando se pode identificar experimentalmente a função que representa a deformação ou construir um modelo matemático adequado, é possível buscar a função inversa e aplicá-la sobre a imagem deformada. Por exemplo: Correção de foco, Imagens borradas por movimento. Em todos os casos, a formulação matemática envolvida é extremamente complexa e o custo computacional muito alto. Os resultados podem ter conseqüências somente qualitativas.

7.2. Segmentação

Quando estamos no contexto de processamento de imagens científicas, o mais comum é queremos obter dados relacionados com os objetos presentes na imagem. Então são necessárias as operações de Segmentação que procurarão isolar regiões de pixels, e operações de Extração de Atributos que vão olhar para essas regiões e calcular uma série de parâmetros que as descreverão.

A operação de segmentação mais comum é a limiarização por um tom de corte. Tudo que está acima deste tom vira branco, tudo que está abaixo vira preto, obtendo-se uma imagem binária. A partir desse momento as operações Morfológicas são excepcionalmente úteis, veja em Operações Locais.

7.3. Extração de Atributos

Utilizando imagens binárias é muito simples obter dados relevantes, ou atributos, das regiões segmentadas, tais como:

Número total de objetos.

Propriedades geométricas, do tipo: Área, Perímetro, Centro de gravidade, Largura máxima e mínima

Atributos relacionados à forma, tais como: Circularidade, Concavidade

Propriedades de luminância, tais como: Nível de cinza médio de cada região (1o momento da distribuição), Desvio padrão do nível de cinza (2o momento da distribuição), outros momentos estatísticos da distribuição de cada região (skewness e kurtosis)

Propriedades associadas à textura

As propriedades geométricas são medidas diretamente sobre a imagem binária. As de luminância são obtidas da imagem original, a partir da informação de posição dos diversos objetos, fornecida pela imagem binária.

7.4. Classificação

Uma vez com esses parâmetros coletados, queremos distinguir objetos na imagem agrupando esses parâmetros de acordo com sua semelhança para cada região de pixels encontrada. Feita essa classificação desses parâmetros os objetos estão reconhecidos e podemos agora tomar decisões e relatar fatos relacionados com os objetos do mundo real, ponderando sempre através uma medida de erro da classificação.

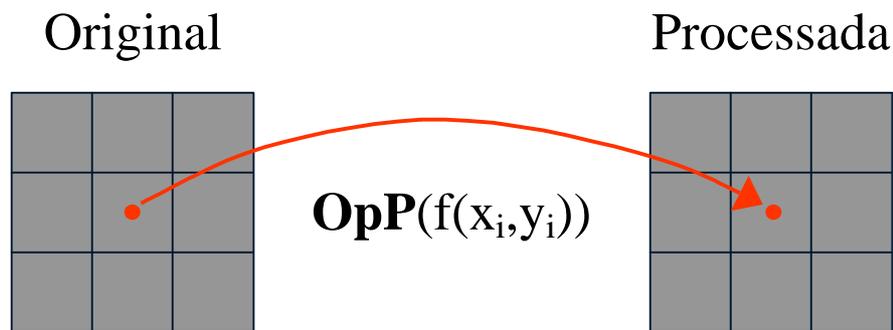
Este processo é muito complexo e existem diversos níveis de automação. Os mais simples implicam em processos de agrupamento estatístico, para os quais a decisão humana é fundamental. Os mais sofisticados permitem ao computador reconhecer diferentes objetos através de técnicas de inteligência artificial com pouca ou nenhuma intervenção humana. Os processos que possuem intervenção são chamados de supervisionados.

Aqui, a palavra classificação não denota nenhum juízo de valor, mas apenas o grupamento em classes dos diversos objetos obtidos na segmentação, cujos atributos já foram medidos.

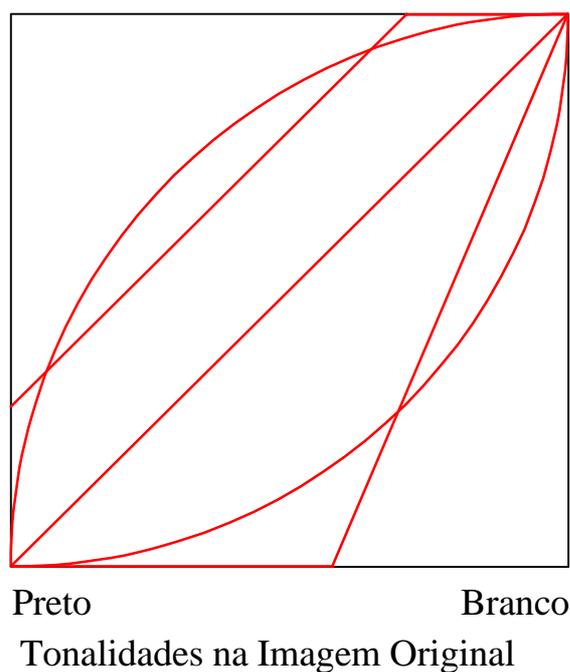
Em geral, vários atributos são necessários para uma correta classificação. Mas, quanto mais atributos, mais complexo se torna o problema. Desta forma, é muito importante realizar uma seleção adequada dos atributos disponíveis, visando otimizar o processo.

7.5. Operações Pontuais

Operações pontuais são operações em que um pixel da imagem resultante depende apenas do mesmo pixel na imagem original. Neste caso a única informação que nós temos é a cor do pixel, por isso muitas das operações pontuais são operações que alteram características de cor e luminância, tais como: brilho, contraste, nível de branco e nível de preto, saturação, correção gamma, limiar (*threshold*), posterização, solarização, negativo, etc.



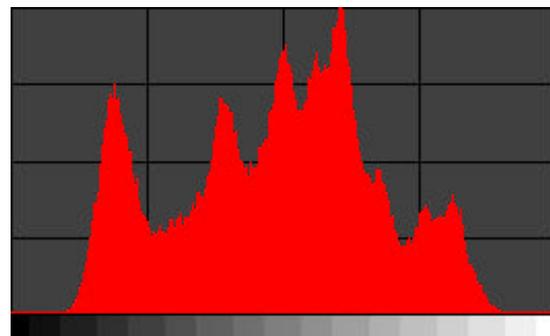
Qualquer operação pontual pode ser visualizada como um mapeamento de pixels da imagem original para a imagem processada. E isso é facilmente representado como um gráfico que relaciona as tonalidades disponíveis na imagem original com as tonalidades disponíveis na imagem processada. Inclusive algumas operações bastante conhecidas são visualizadas nesse gráfico como funções muito simples. A operação pontual de identidade é simplesmente uma reta que mapea todos os tons da imagem original nos mesmos tons na imagem processada. Transladando essa reta estamos simplesmente alterando o brilho na imagem. Alterando sua angulação estamos alterando o contraste, mas existem outras formas de se alterar o contraste como por exemplo usando uma função logarítmica que faria com que tons escuros fosse mapeados em um número maior de tons, realçando esses tons escuros, ou o inverso, uma função exponencial onde tons claros seriam realçados.



Até agora estávamos falando de operações pontuais que envolvem apenas uma imagem original. Estas são chamadas de unárias. Mas podemos definir operações que usam mais de uma imagem original, são chamadas de binárias, ternárias, etc. Por exemplo, uma operação de média de várias imagens capturadas na mesma posição, é uma técnica comum de eliminação de ruído. Mas as operações binárias mais comuns são as operações aritméticas de soma, subtração, divisão, etc de duas ou mais imagens.

Embora as funções pontuais estejam restritas somente ao pixel sendo processado, o processamento em si pode levar em consideração dados globais da imagem, como por exemplo o histograma.

O histograma é uma função estatística da imagem que para cada nível de tonalidade, calcula quantos pixels existem naquela tonalidade. Muitas operações pontuais usam o histograma como parâmetro de decisão para fornecer resultados diferentes para o pixel da imagem processada.

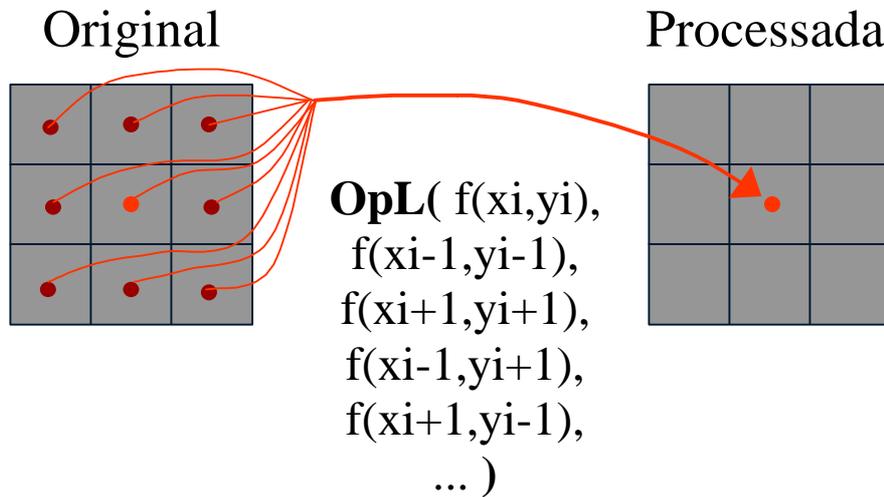


Esta imagem mostra o histograma de uma imagem monocromática.

7.6. Operações Locais

Operações pontuais são operações em que um pixel da imagem resultante depende de uma vizinhança do mesmo pixel na imagem original. As operações locais mais comuns são as operações de filtragem que usam uma convolução com um kernel de dimensão $n \times n$, onde n é tipicamente 3, 5, 7 e 9, mas pode assumir qualquer valor. Por exemplo, no caso de $n=3$ é muito fácil de entender. Cada pixel da vizinhança 3×3 do pixel na imagem original é multiplicado por um valor e todos esses valores são somados resultando no pixel novo. Os valores que ponderamos a vizinhança são

armazenados em uma matriz chamada, neste caso, de kernel de convolução. Operações de Blur, Sharpen, Detecção de Bordas e muitas outras são implementadas assim.



As operações que realizam deformações no grid de pixels utilizam, como vimos, a técnica de interpolação para encontrar o novo ponto. A técnica de interpolação não deixa de ser uma operação local que depende apenas de uma vizinhança, embora algumas técnicas de interpolação muito sofisticadas e em geral muito lentas usam uma vizinhança bem extensa. Dentre as técnicas que usam interpolação podemos citar: Warping, Morphing, Resize, Stretch, Rotação, etc.

Um outro conjunto de operações que também podem ser consideradas como locais são as operações Morfológicas. São muito utilizadas por sistemas de processamento de imagens científicas para a etapa de segmentação, como já mencionado. Embora muito semelhantes a convolução com um kernel, a forma com que os vizinhos são computados é ligeiramente diferente, pois usam operações lógicas, tais como AND e OR.

7.7. Operações Globais

Operações globais são operações em que um pixel da imagem resultante depende de um processamento realizado em todos os pixels da imagem original. Neste grupo de operações estão as transformadas de domínio, tais como a Transformada de Fourier, a Transformada de Wavelets (que na realidade são muitas, dependem de que wavelet está sendo usada) e a Transformada de Hough.

Todas essas transformadas nos dão informações muito interessantes sobre a imagem original. A Transformada de Fourier, por exemplo, é base fundamental para toda teoria de

processamento de sinais e com ela pode-se realizar uma série de operações muito importantes com imagens. A Transformada de Wavelets irá decompor a imagem em uma estrutura de multiresolução, que é muito utilizada para compressão de imagens. A Transformada de Hough procura identificar formas geométricas na imagem, tais como retas e círculos.

Esse grupo de operações é visto somente em programas de processamento de imagens científico.

8. Software

Existem diversos sistemas de processamento de imagens disponíveis no mercado. Eles se distinguem pela forma de distribuição (Comercial, *Shareware*, *Freeware*), pela área de aplicação (editoração eletrônica, análise científica, visualização, edição informal, pesquisa, ...), e pelas plataformas em que estão disponíveis (UNIX, Windows, MAC, ...).

O programa mais bem sucedido, não importando a categoria, é o Adobe Photoshop. Atualmente na versão 5.0, é um dos mais robustos e confiáveis sistemas de processamento de imagens. É voltado para editoração eletrônica e não possui funções de análise de imagens. Possui versões para MAC, Windows e UNIX.

Seguindo essa mesma linha temos o Corel PhotoPaint e o Fractal Design Painter, mas estes só rodam no ambiente Windows. O Fractal Design Painter possui mecanismos muito peculiares de pintura, além de permitir a criação de animações usando seqüências de imagens.

Outros programas na mesma linha, mas que ganharam popularidade por serem *Shareware/Freeware* e não possuírem a complexidade dos sistemas acima, são o Paint Shop Pro e o LviewPro. Um programa semelhante a esses, mas que realiza apenas a visualização é o ACDSee, se tornou popular pela facilidade de fazer *Slide Shows*. Todos esses são somente para Windows.

No UNIX temos o GIMP e o XV, este último muito divulgado na *Internet*. Ambos são *Freeware*.

Quando se trata de análise de imagem e processamento científico, muitos dos programas mencionados são úteis, mas outros assumem um nicho não atacados por estes.

Na linha *Freeware* temos o NIH Image (MAC) e o Image Tool (Windows). No UNIX temos ainda dois grandes sistemas o Khorus e o XITE. Sendo que o Khorus abrange muito mais do que processamento de imagens.

Na linha comercial temos o KS-400 (Windows), e o Digital Micrograph (MAC e Windows), ambos voltados para análise de imagens em microscopia ótica e eletrônica. Temos também outros mais gerais, como o Image Pro, o Optimas, e o WiT. Este último tem uma característica muito interessante de permitir a criação de macros usando uma linguagem visual.

Adobe PhotoShop

Como este é muito difundido, convém comentarmos algumas funções do mesmo associando-as aos conceitos aqui apresentados. Vamos estar comentando funções da versão “4.01”.

No menu **Image** temos as funções **Image Size** e **Canvas Size**, ambas relativas a resolução espacial da imagem.

Na função **Image Size** quando **Resample Image** está marcado pode-se mudar o tamanho em *pixels* da imagem, quando está desmarcado pode-se mudar apenas o tamanho real da imagem sem mudar o número de *pixels*. Junto a este controle você também pode escolher o método de interpolação, quando o número de *pixels* for mudar. Este diálogo é bastante versátil permitindo manter a razão de aspectos da imagem original, e permite também mudar o tamanho em *pixels* baseado no tamanho real e na resolução.

A função **Canvas Size** irá mudar o número de *pixels* da imagem, mas apenas acrescentará *pixels* acima, abaixo ou dos lados de acordo com a posição da imagem original em **Anchor**.

Tome cuidado ao mudar o tamanho da imagem, pois ao reduzi-la estará sempre perdendo informação, e ao aumentá-la corre o risco de ocupar muita memória.

Alguns procedimentos relacionados com calibração dos sistemas de cor utilizados são feitos no submenu **Color Settings** do menu **File**.

Embora suporte 16 bits por componente, além dos tradicionais 8, praticamente nenhuma função fica habilitada quando muda-se para 16 bits. Ainda não entendi o objetivo do suporte a 16 bits já que quase nada pode ser feito.

No submenu **Mode** do menu **Image** você normalmente irá, e deve, usar o sistema de cor RGB, junto com suas variações: **Gray Scale** e **Indexed Color**. Em geral é mais conveniente processar a imagem nesses modos e depois converte-la para algum outro, se necessário. O modo **Bitmap** serve para converter a imagem para apenas 1 bit por pixel.

Quando se está trabalhando no modo RGB pode-se visualizar rapidamente como seria sua imagem em CMYK selecionado a opção **CMYK Preview** do menu **View**.

O sistema CMYK é usado para criar uma separação de cores para a impressão. Também pode-se usá-lo para se trabalhar diretamente sobre imagens CMYK provindas de determinados dispositivos.

O sistema Lab é usado quando se quer trabalhar com imagens de PhotoCD ou quando se quer editar a componente de luminância independentemente dos valores de cromaticidade. O sistema Lab também é usado internamente para converter de um sistema para outro e é o recomendado para o transporte de uma aplicação para outra.

Também tenha cuidado ao utilizar o submenu de **Mode**, pois durante as conversões de um sistema para outro podem ocorrer perdas na imagem. Exceto quando convertendo para **Multichannel** que apenas separa cada uma das componentes para serem trabalhadas independentemente.

Quando a imagem está em tons de cinza, pode-se trabalhar com o modo de **Duotone**, que inclui *monotone*, *duotones*, *tritones* e *quadtones*. Nesses modos a escala de cinzas é substituída por uma escala em um tom, de uma, duas, três ou quatro cores, respectivamente. Este modo é útil para a impressão da imagem quando a tinta preta é substituída por uma outra cor (*monotone*) ou quando usa-se uma segunda cor (*duotone*), por exemplo um cinza, para aumentar a escala tonal da imagem impressa.

Todas as funções do submenu **Adjust** do menu **Image** são operações pontuais e dependem somente da cor do próprio pixel na imagem original. A função **Calculations** do menu **Image** são operações pontuais binárias ou ternárias por precisam de duas ou mais imagens. Quando você junta dois ou mais **Layers** em um único **Layer**, está realizando também uma operação pontual. Já no caso do menu **Filter** a grande maioria das funções são operações locais.

9. Aplicações

Qualquer área onde se possa capturar uma imagem e dela obter resultados é uma área onde pode ser usado um sistema de processamento de imagens. Devido a esta vasta amplitude de áreas de aplicação, sistemas de processamento de imagens são, à priori, inter-disciplinares. Seguem alguns exemplos:

- Sensoriamento Remoto

Centenas de fotografias, enviadas por naves espaciais e satélites em torno da terra, precisam ser avaliadas e catalogadas com fidelidade e eficiência. Fotos de países podem ser usadas para a obtenção rápida de mapas cartográficos e estudos geográficos, tais como análise do solo e da safra agrícola.

- Microscopia.

Análise de imagens provindas de microscópios óticos ou eletrônicos, em áreas que variam desde a medicina até a metalurgia, identificando e classificando partículas.

- Medicina.

Imagens radiográficas de baixa qualidade podem ser realçadas para a obtenção de um diagnóstico mais preciso. Com a introdução da tomografia computadorizada e da ressonância magnética, imagens volumétricas são utilizadas para diversas aplicações, tais como planejamento cirúrgico e diagnóstico.

- Manutenção de obras de arte.

Evitar a degradação com o tempo de obras de arte não é um processo simples, mas possuindo uma imagem da obra original que pode ser processada para ajudar no processo de restauração para se obter um retrato mais fiel o possível. Além disso, existem estudos para a verificação da autenticidade de pinturas através da análise das pinceladas de cada pintor.

- Identificação de impressões digitais.

Através de adequada classificação e armazenamento pode-se identificar uma impressão digital usando um banco de dados associado a um banco de imagens.

- Armazenamento de documentos.

É comum que grandes quantidades de informação fiquem "encalhadas" devido à dificuldade de acesso a quilos de papel que ficam engavetados e deteriorando. Estes textos podem ser transferidos para um sistema de hipertexto que permita uma consulta rápida e dinâmica. Para tanto, é necessário capturar as imagens dos textos e depois realizar um reconhecimento de caracteres, segue então um processo de integração ao hipertexto. Muitas vezes estes textos são documentos oficiais que precisam manter suas características, portanto as próprias imagens são armazenadas.

- Medidas de velocidade de escoamento de fluidos.

Através de partículas suspensas no líquido, pode-se iluminar um corte no escoamento através de um plano de luz, e utilizando uma exposição lenta obtém-se uma fotografia com traços correspondentes às trajetórias das diversas partículas. A partir dos comprimentos dos traços pode-se obter uma medida aproximada da velocidade do fluido.

- Controle de qualidade.

A aplicação de processamento de imagens na área de controle de qualidade é muito vasta. Praticamente qualquer processo industrial que necessite de alguma monitoração ótica ou visual pode ser automatizado com vantagens. O que pode impedir é o grau de dificuldade de interpretar cada imagem, assim como em qualquer outra aplicação de análise de imagens.

Uma aplicação comum é a verificação de placas de circuito impresso, que possuem uma geometria simples e regras fixas.

- Outros: Astronomia, Fotografia, Vídeo, Efeitos Especiais, etc.

10. Internet Links

Logical Vision Ltd - <http://www.logicalvision.com>

Synoptics Ltd - <http://www.synoptics.co.uk>

AccuSoft - <http://www.accusoft.com>

Media Cybernetics Image Analysis Home Page (Image Pro) - <http://www.mediacy.com>

Xite - <http://www.ifi.uio.no/~blab/Software/Xite>

GIMP - <http://www.gimp.org>

XV - <http://www.trilon.com/xv>

NIH Image - <http://rsb.info.nih.gov/nih-image>

ImageTool - <http://ddsdx.uthscsa.edu/dig/itdesc.html>

Poynton's Colour FAQ -

http://www.inforamp.net/~poynton/notes/colour_and_gamma/ColorFAQ.html

A Standard Default Color Space for the Internet (sRGB) -

<http://www.w3.org/Graphics/Color/sRGB.html>

11. Bibliografia

- [BAXES 94] G. A. Baxes; "Digital Image Processing (Principles and Applications)"; John Wiley & Sons, Inc.; 1994.
- [CASTLEMAN 95] K. R. Castleman; "Digital Image Processing"; Prentice-Hall, Inc.; 1995.
- [GOMES 97] J. Gomes e L. Velho; "Image Processing for Computer Graphics"; Springer; 1997.
- [CHAVES 91] E. O. C. Chaves; "Multimídia - Conceituação, Aplicações e Tecnologia"; People Computação Ltda.; 1991.
- [FOLEY 96] J. Foley, A. van Dam, S. Feiner, J. Hughes; "Computer Graphics: Principles and Practice"; Addison-Wesley; 1996.
- [GONZALEZ 92] R. C. Gonzalez, R. Woods; "Digital Image Processing"; Addison-Wesley; 1992.