

SISTEMAS NUMÉRICOS E A REPRESENTAÇÃO INTERNA DOS DADOS NO COMPUTADOR

2.0 Índice

2.7	Representação Digital de Áudio, Imagem e Vídeo	2
2.7.1	Sinais Analógicos para representar informações	2
2.7.2	Porque Digitalizar?	3
2.7.3	Digitalização, Amostragem e Quantificação	4
2.7.4	Áudio	6
2.7.5	Vídeos e Imagens Analógicos	7
2.7.6	Representação digital de imagens e vídeos	8
2.7.7	Especificação da Cor	9
2.7.8	Sistema RGB	10

Capítulo 2 – Parte final

2.1 Representação Digital de Áudio, Imagem e Vídeo

Esta seção apresenta como imagens, áudios e vídeos são capturados do mundo real a partir de sinais analógicos e como estes sinais são transformados numa forma digital.

2.1.1 Sinais Analógicos para representar informações

Esta seção apresenta uma introdução às formas de onda e sinais analógicos e seus uso na descrição de informações detectadas pelos sentidos humanos (tato, olfato, gustação, audição, visão).

Informações percebidas e variáveis físicas

Informações que os sentidos humanos podem detectar podem ser descritas como uma ou várias variáveis físicas cujos valores podem ser funções do tempo e do espaço. Note que por informações é entendido aqui como a representação física em termos de estímulos para um sentido humano, e não o significado deste estímulo (conteúdo semântico). Atualmente, em informática, se trabalha principalmente com estímulos auditivos e visuais.

Descrevendo sons com formas de onda

Um som, que atravessa o ar, é uma onda de ar comprimido ou expandido cuja pressão altera no tempo e espaço. Os sons podem ser descritos por valores de pressão que variam apenas com o tempo (valores dependentes do tempo). Um padrão de oscilação sonora é mostrado na Figura 1. Este tipo de representação é chamado de forma de onda (*waveform*). A forma de onda é caracterizada por um **período** e **amplitude**. O **período** (T) é o tempo necessário para a realização de um ciclo completo, isto é, o intervalo de tempo que, num fenômeno periódico, separa a passagem do sistema por dois estados idênticos. A **frequência** (f) é definida como o inverso do período ($f=1/T$) e representa o número de períodos em um segundo. A frequência é normalmente medida em Hz (Hertz) ou ciclos por segundo (cps). A amplitude do som é define um som leve ou pesado.

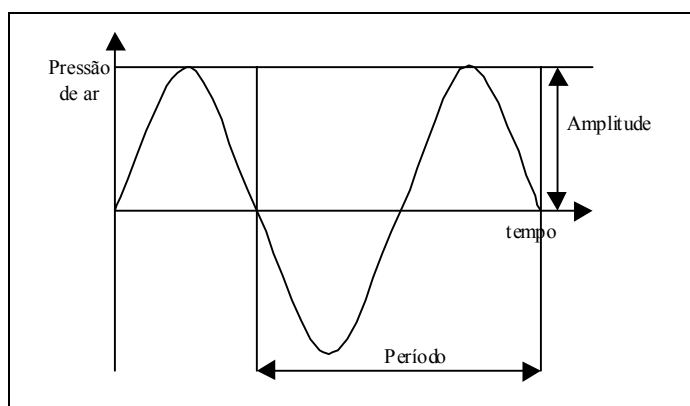


Figura 1. Forma de onda

Como a onda de som ocorre naturalmente, ela nunca é perfeitamente suave ou uniformemente periódica como a forma de onda da Figura 1.

Descrevendo imagens monocromáticas com variáveis físicas

As imagens refletem radiações eletromagnéticas (luz) incidentes que estimulam os olhos do observador. A intensidade de luz é uma função da posição espacial do ponto refletido sob a imagem. Portanto, a imagem pode ser descrita pelo valor da intensidade de luz

que é uma função de duas coordenadas espaciais. Se a cena observada não foi plana, uma terceira coordenada espacial é necessária.

Descrevendo imagens coloridas com formas de onda

Se a imagem não é monocromática, ela reflete diferentes comprimentos de onda. Assim, uma função simples não é suficiente para descrever imagens coloridas. Neste caso, um espectro completo de comprimentos de onda, cada um com sua própria intensidade, torna-se necessário. Assim as imagens teriam que ser descritas pela conjunção de várias funções bidimensionais. Felizmente, o sistema visual humano tem certas propriedades que simplificam a descrição de imagens coloridas.

A luz visível é composta de um conjunto de comprimento de ondas eletromagnéticas (do vermelho ao violeta), ou seja, de uma distribuição espectral. Este sinal estimula o sistema visual e cria uma resposta nervosa que é processada pelo cérebro.

Testes realizados mostram que diferentes distribuições espectrais da luz pode dar a mesma resposta visual. Em outras palavras, é possível criar sensações de cores idênticas com diferentes combinações de comprimentos de onda (isto é diferentes combinações de cores).

A teoria da cor foi desenvolvida por Thomas Young no início de 1802 e afirma que qualquer sensação de cor pode ser reproduzida pela mistura em proporções apropriadas de três luzes coloridas monocromáticas primárias. Cores primárias são independentes, isto é, uma cor primária não pode ser obtida misturando outras duas cores primárias. A *Commission Internationale de l'Eclairage* (CIE) recomendou o uso de uma tripla particular de luz monocromática. Cada fonte de luz é definida pelo seu comprimento de onda ($\lambda_1 = 70$ nm, vermelho; $\lambda_2 = 546.1$ nm, verde; $\lambda_3 = 435.8$ nm, azul). Em vez de ser descrita por uma infinidade de funções bidimensionais, qualquer imagem colorida plana pode ser representada por um conjunto de três funções bidimensionais.

2.1.2 Porque Digitalizar?

A utilização de informações na forma digital trás as várias vantagens.

Universalidade de representação

Sistemas computacionais manipulam apenas dados digitais. Quando áudio, imagens, vídeos estão na forma digital, eles podem ser facilmente armazenados e manipulados (processados, transmitidos e apresentados) pelos sistemas computacionais tal qual outros dados. Desta forma, como todas as mídias de apresentação (textos, imagens, som etc.) são codificadas numa única forma (binária), elas podem ser manipuladas de uma mesma forma e pelo mesmo tipo de equipamento. Além disso, informações de áudio e vídeo digitalizadas são facilmente integradas com outros tipos de dados e são de fácil integração com mídia digitais usada nos sistemas computacionais.

Processamento

Informações de áudio e vídeo digitais são processadas, analisadas, modificadas, alteradas, ou complementadas por programas de computador tal qual outros dados. A seguir são apresentados alguns exemplos de processamentos possíveis graças a representação digital de informações de áudio e vídeo [Fluckiger, 95]:

- reconhecimento de conteúdos semânticos (voz, escrita a mão, formas e padrões);
- estruturas de dados, ligações usando apontadores entre elementos de informações podem ser criados para rápida obtenção de informações;
- editores com funções “cut-and-paste” para criar monomídia (p.e. som apenas) ou serem incorporadas a documentos multimídia são possíveis;
- qualidade da informação pode ser melhorada pela remoção de ruídos ou erros como a digitalização, como por exemplo digitalização do som de velhos discos de vinil para criar CDs de áudio, ou de filmes antigos para DVD;
- informações sintetizadas e vídeos podem ser “mixadas”, isto é, misturadas.

Qualidade

Sistemas digitais são mais confiáveis. Sinais digitais são mais tolerantes a ruídos e interferências que os analógicos. Na forma analógica, o valor do sinal é alterado se há ruídos ou interferências. Este erro é acumulativo de um estágio para outro do sistema. Na forma digital, há apenas dois níveis de sinal: alto (1) ou baixo (0). Se o erro causado pela interferência ou ruído é abaixo de um certo limiar, o sinal pode ser reconhecido corretamente. Além disso, em cada estado do processamento digital ou transmissão, o sinal digital é reconstruído, assim erros não são acumulados.

Segurança

Se segurança na comunicação é necessária, a representação digital da informação facilita a criptografia ou outros tipos de codificação.

Armazenamento

A utilização unicamente de mídias digitais permite a existência de um dispositivo único de armazenamento de dados para todas as mídias. As diferenças geralmente estão ligadas a requisitos de tamanho e tipo de arquivo. Imagens e vídeos necessitam de uma maior capacidade de armazenamento que textos ou gráficos. Som necessita de uma capacidade de armazenamento um pouco menor que imagens. Arquivos de imagem e de áudio freqüentemente utilizam recursos de compactação.

Dispositivos digitais apropriados podem ser necessários, tal como CD-ROMs (*Compact Disk-Read Only Memories*) e DVDs (*Digital Video Disk*).

Transmissão

Qualquer sistema de comunicação de dados pode ser (potencialmente) utilizado para a transmissão de informações de áudio e vídeo digitais. Porém, algumas dificuldades são causadas pelas características de certas aplicações que exigem um sincronismo temporal preciso (por exemplo, videoconferência).

A vantagem da transmissão digital em relação a transmissão de sinais analógicos é que ela é menos sensível a ruídos, a detecção de erros é mais fácil, processamento corretivo do sinal original e a criptografia são possíveis.

2.1.3 Digitalização, Amostragem e Quantificação

Esta seção apresenta a forma de digitalização dos vários tipos de mídias de apresentação. Digitalização aqui é o processo envolvido na transformação de sinais analógicos em sinais digitais:

- **Sinal analógico** é uma medida física que varia continuamente com o tempo e/ou espaço. Eles são descritos por uma função dependente apenas do tempo ($s=f(t)$), dependente apenas do espaço ($s=f(x,y,z)$), ou dependente do tempo e do espaço ($s=f(x,y,z,t)$). Sinais analógicos físicos (neste caso, som e luz) são convertidos em sinais elétricos equivalentes (uma corrente ou tensão elétrica). A precisão é dependente das características dos sensores que também são chamados de transdutores.
- **Sinais digitais** são seqüências de valores dependentes do tempo ou do espaço codificados no formato binário.

Para a conversão de sinais analógico em digital é necessário a realização de três passos: amostragem, quantificação e codificação. A Figura 2 ilustra o processo de digitalização de um sinal analógico no domínio do tempo.

Amostragem

Nesta etapa um conjunto discreto de valores analógicos é amostrado em intervalos temporais (p.e., para sons) ou espaciais (p.e., para imagens) de periodicidade constante, como apresentado na Figura 2a. A freqüência das amostras é chamada de taxa de amostragem ou freqüência de amostragem. O valor amostrado (amplitude) é mantido

constante até o próximo intervalo, ou seja, até que a próxima amostra seja feita. Isto é realizado por meio de circuitos chamados “sampler and hold” (amostrador e sustentador). Cada uma das amostras é analógica, ou seja, pode ter qualquer valor em um intervalo numérico contínuo com relação ao valor da amplitude (tensão ou corrente elétrica). Porém, é considerada discreta no tempo, isto é, dentro de cada intervalo de tempo, a amostra tem apenas um valor.

O fundamento para o processo de amostragem é o teorema de Nyquist. Este teorema diz que se um sinal analógico contém componentes de frequência até f Hz, a taxa de amostragem deve ser ao menos $2f$ Hz (ou seja, maior ou igual ao dobro da máxima frequência). Na prática, esta frequência é um pouco maior que $2f$ Hz. Por exemplo, a taxa de amostragem do CD-audio é de 44,1 kHz, e dos “tapes” de áudio digital (DAT) é de 48kHz, assim, a faixa audível de frequência de 20 kHz é coberta. Outro exemplo, os componentes principais de frequência da voz humana estão dentro de 3,1 kHz, com isto os sistemas de telefonia analógicos limitam o sinal transmitido a 3.1 kHz; é comum usar uma frequência de amostragem de 8 kHz para converter este sinal em digital.

Quantificação

O processo de converter valores de amostras contínuas em valores discretos é chamado de quantificação. Neste processo nós dividimos o domínio do sinal em um número fixo de intervalos. Cada intervalo tem o mesmo tamanho e recebe um número. Na Figura 2c estes intervalos são numerados de 0 a 7. A cada amostra dentro de um intervalo é atribuído um único valor de amplitude. O tamanho deste intervalo de quantificação é chamado de passo de quantificação. A técnica que utiliza o mesmo passo de quantificação é chamada modulação PCM (*Pulse Coded Modulation*). Algumas vezes, nem todos os valores amostrados são retidos após a quantificação. No caso ilustrado pela Figura 2c, todos os valores amostrados foram retidos.

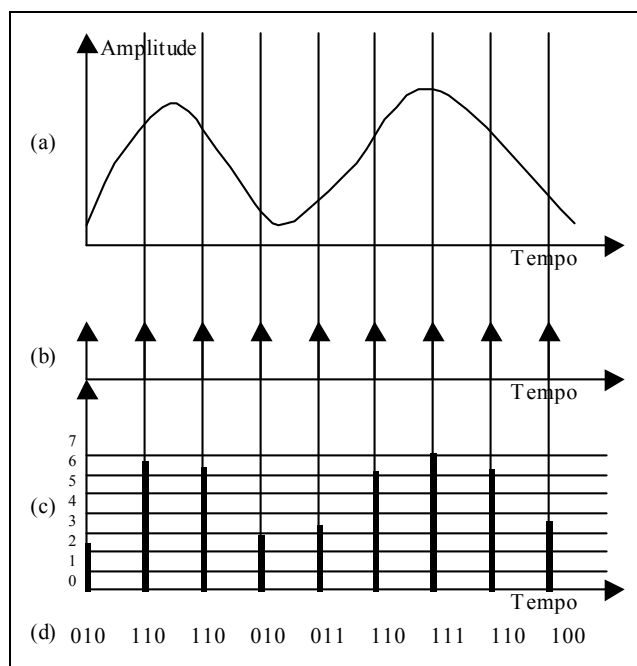


Figura 2. Conversão A/D [Lu, 96]: (a) sinal analógico; (b) pulsos de amostragem; (c) valores amostrados e intervalos de quantificação; (d) seqüência digital

Codificação

A codificação consiste em associar um conjunto de dígitos binários, chamado de *code-word*, a cada valor quantificado. No caso da Figura 2d, oito níveis de quantificação são usados. Estes níveis podem ser codificados usando 3 bits, assim cada amostra é representada por 3 bits.

Em algumas aplicações de telefonia, a digitalização da voz humana utiliza 16 bits por amostra, que então leva a 2^{16} ou 65.536 passos de quantificação. Em outras aplicações

de compressão de voz, algumas vezes, apenas 8 bits são necessários, produzindo apenas 256 passos de quantificação.

Taxa de bits

Taxa de bits é definida como o produto entre taxa de amostragem e o número de bits usados no processo de quantificação. É dada em bits por segundo – bps – ou kilobits por segundo – kbps. Por exemplo, supondo uma frequência de 8k Hz e 8 bits por amostra, a taxa de bits necessária à telefonia é igual a $8000 \times 8 = 64$ kbps.

Conversão analógica/digital e digital/analógica

Em sistemas computacionais as informações são representadas internamente no formato digital. Mas, humanos reagem a estímulos sensoriais físicos, ou seja, analógicos. Assim a conversão digital-para-analógico (ou conversão D/A) é necessária na apresentação de certas informações (Figura 3).

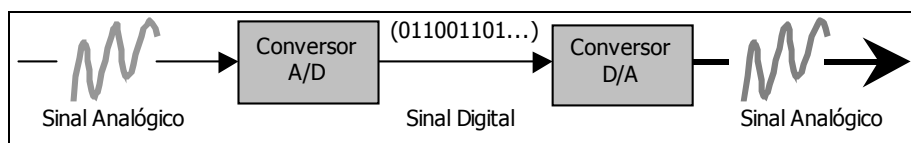
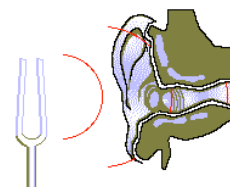


Figura 3. Conversão analógico/digital e digital/analógica

2.1.4 Áudio

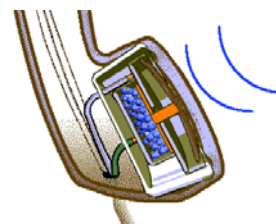
Áudio é causado pelo distúrbio da pressão de ar que alcança o tímpano. Quando a frequência do distúrbio de ar está na faixa de 20 Hz a 20.000 Hz ele é audível. A maioria dos sistemas trabalham com esta faixa de frequência. Outro parâmetro usado para a medição do som é a amplitude (medido em decibéis - dB), variação que causa o som leve ou pesado. Por exemplo, o limiar da dor é de 100 a 120 dB.



A onda sonora é uma onda contínua no tempo e amplitude. A onda apresentada na Figura 1 pode ser um exemplo de onda sonora.

Representação digital de áudio

A forma de onda de áudio é convertida em um sinal elétrico contínuo (analógico) por um microfone. Este sinal elétrico é medido normalmente em volts. Para que sistemas computacionais processem e comuniquem sinais de áudio, o sinal elétrico deve ser convertido em um sinal digital. O mecanismo que converte o sinal de áudio em analógico em digital é chamado de Conversor Analógico para Digital (CAD).

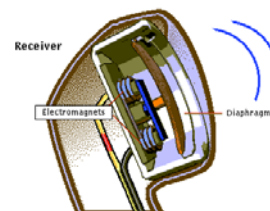


Áudio digital necessita ser amostrado continuamente em uma taxa fixa. Cada amostra é representada por um número fixo de bits. A tabela a seguir mostra a taxa de amostragem e o número de bits usados para cada amostra para várias aplicações de áudio. Relembrando, quanto maior a taxa de amostragem e maior o número de bits por amostragem, maior é a qualidade do áudio restituído, mas com isso maior é a taxa de bits. Note na tabela que para áudio estéreo, tal como CD-áudio, dois canais são necessários.

Aplicações	Nº de canais	Taxa de amostragem	Bits por amostragem	Taxa de bits
CD-Audio	2	44.1 kHz	16	1,41 Mbps
DAT	2	48 kHz	16	1,53 Mbps
Telefone Digital	1	8 kHz	8	64 Kbps
Rádio digital, long play DAT	2	32 KHz	16	1,02 Mbps

Para a apresentação do áudio digitalizado é necessário realizar a transformação de uma representação artificial do som em uma forma de onda física audível pelo ouvido humano. Para isto são utilizados Conversores Digital para Analógico (CDA).

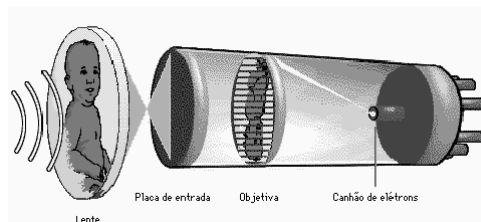
Normalmente os conversores CAD e CDA são implementados em uma única placa. Um exemplo de placa de áudio é Creative Sound Blaster AWE64, possibilitando até 16 bits por amostras, produzindo áudio qualidade CD.



2.1.5 Vídeos e Imagens Analógicas

Captura e reprodução de imagens e vídeos analógicos

As imagens são capturadas usando câmeras da seguinte maneira: as lentes da câmera focam uma imagem de uma cena em uma superfície foto-sensível de sensores CCD (*Charge-Coupled Device*); o brilho de cada ponto é convertido em uma carga elétrica por uma camada de material foto-sensível, estas cargas são proporcionais ao brilho nos pontos; a superfície foto-sensível é rastreada por um feixe de elétrons para capturar as cargas elétricas, devendo ser feito rapidamente antes que a cena mude. Desta maneira a imagem ou cena é convertida em um sinal elétrico contínuo.



Nesta seção, por simplificação considera-se apenas a captura e reprodução de vídeos monocromáticos, onde apenas um sinal de luminância é produzido (apenas a luminosidade é capturada, temos a imagem em tons de cinza). Neste caso são usadas **câmeras de Luminância**, que captam a imagem em tons de cinza, e gera um sinal só com a luminância da imagem. A imagem é gerada por um CCD monocromático que capta o tom de cinza que incide em cada célula do circuito. Este tipo de câmera é utilizada em geral para aplicações em visão computacional e nos casos onde a informação sobre a luminosidade da imagem é suficiente.

O dispositivo de apresentação de imagens mais utilizado é o tubo de raios catódicos (CRT). Eles são usados nos aparelhos de TV e monitores de computadores. Há uma camada de fósforo fluorescente no interior do CRT. Esta camada é rastreada por um feixe de elétrons na mesma forma do processo de captura na câmera. Quando o feixe toca o fósforo ele emite luz durante um curto instante. O brilho da luz depende da força do feixe. Quando quadros repetem-se suficientemente rápidos a persistência da visão resulta na reprodução de um vídeo. Em implementação prática, o sinal elétrico enviado da câmera para o dispositivo de apresentação deve conter informações adicionais para assegurar que o rastreamento esteja sincronizado com o rastreamento do sensor na câmera. Esta informação é chamada *sync information*.

Vídeos e Imagens Coloridos

Todos os sistemas de TV a cores são baseados na teoria *Tristimulus* de reprodução da cor. Esta teoria afirma que qualquer cor pode ser reproduzida com a mistura das três cores primárias: vermelho, verde e azul (sistema aditivo). Para capturar imagens coloridas, uma câmera divide a luz nos seus componentes vermelho, verde e azul. Estes três componentes de cor são focalizados em sensores de vermelho, verde e azul, que convertem estes três componentes em sinais elétricos separados.

Em um monitor colorido, há 3 tipos de fósforos fluorescentes – vermelho, verde e azul. Cada tipo de fósforo emite luz um respectivo comprimento de onda quando estimulado por um dos 3 feixes de elétrons do tubo de raios catódicos. Estes fósforos são arranjados de tal forma que cada posição do vídeo (ponto ou pixel) tem 3 tipos de fósforo. A mistura da luz emitida destes 3 fósforos produz um ponto de cor conforme a intensidade de cada cor. Assim, se obtém o sinal RGB.

O sinal RGB (red, green, blue) é separado pelas cores básicas. Com isso, é possível ter uma imagem mais nítida. Ele é utilizado em câmeras e gravadores profissionais, imagens geradas por computador etc.

No sinal de vídeo composto colorido os sinais das cores (RGB) são codificados em um único sinal seguindo um determinado padrão (NTSC, PAL-M, SECAM etc) ;

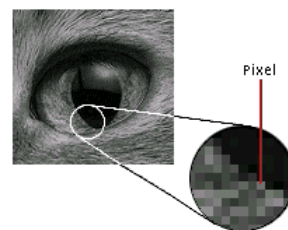
O Y/C (S-video) é composto por duas partes, a luminância (Y) e a croma (C) como isso a imagem tem uma melhor qualidade do que no vídeo composto. É muito usado por vídeos SVHS, "laser disc", DVD e outros aparelhos que geram imagens de boa qualidade (acima de 400 linhas);

2.1.6 Representação digital de imagens e vídeos

A seção anterior discutiu muitos conceitos e nomenclaturas de vídeos analógicos. Esta seção discute a representação digital de imagens e vídeos.

Imagens Digitais

Imagens não são revisáveis porque seu formato não contém informações estruturais. Elas podem resultar de capturas do mundo real (via escaneamento de uma página impressa ou foto, câmeras digitais) ou elas podem ser sintetizadas pelo computador (via programas de *paint*, captura da tela, conversão de gráficos em imagens "bitmap"). Após digitalizadas, as imagens podem ser manipuladas com editores de imagens (por exemplo, Photoshop), que não produzem documentos que retêm a estrutura semântica.



Formatos de Imagens

Imagens no computador são representadas por *bitmaps*. Um bitmap é uma matriz bidimensional espacial de elementos de imagem chamados de pixels. Um pixel é o menor elemento de resolução da imagem. O número de bits disponíveis para codificar um pixel é chamado de profundidade de amplitude (ou de pixel). Exemplos típicos de profundidade de pixel é 1 (para imagens preto&branco), 2, 4, 8, 12, 16 ou 24 bits. O valor numérico pode representar um ponto preto e branco, um nível de cinza, ou atributos de cor (3 valores) do elemento de imagem em imagens coloridas.

O número de linhas da matriz de pixels (m) é chamado de resolução vertical da imagem, e o número de colunas (n) é chamado de resolução horizontal. Denominamos resolução espacial, ou resolução geométrica, ao produto $m \times n$ da resolução vertical pela resolução horizontal. A resolução espacial estabelece a frequência de amostragem final da imagem. Quanto maior a resolução maior deve ser a frequência de amostragem. A resolução espacial dada em termos absolutos não fornece muita informação sobre a resolução real da imagem quando realizada em dispositivo físico. Isso ocorre porque ficamos na dependência do tamanho físico do pixel do dispositivo. Uma medida mais confiável de resolução é dada pela densidade de resolução da imagem que fornece o número de pixels por unidade linear de medida. Em geral se utiliza o número de pixels por polegada, ppi ("pixels per inch") também chamada de dpi ("dots per inch").

Formatos *bitmap* necessitam mais capacidade de armazenamento do que gráficos e textos. Como *bitmaps* ignoram a semântica (significado da imagem), duas imagens de mesma dimensão (altura e largura) ocupam o mesmo espaço. Por exemplo, um quadrado ou uma foto digitalizada, com dimensões idênticas, ocupam o mesmo espaço. Os gráficos, como eles consideram a semântica, ocupam menos espaço.

Imagens e Gráficos Animados

As imagens e os gráficos podem ser apresentados na tela do computador como uma sucessão de imagens/gráficos que podem criar a sensação de movimento.

Uma imagem ou gráfico individual de uma animação é chamado de quadro (ou *frame*). Para ser compreensível, os quadros que compõem a animação devem ser apresentados geralmente em uma taxa aproximadamente fixa. O número de quadros apresentados por

segundo é definido como frequência de quadros e é medido em termos de quadros por segundo (fps – *frames per seconds*).

A taxa de quadros de uma animação é determinada por 3 fatores:

- A taxa deve ser alta suficiente para produzir a sensação de movimento. Para isto, taxas maiores ou iguais a 25 fps devem ser utilizadas. A tabela abaixo resume as principais frequências de quadro utilizadas atualmente.

Fps	Comentários
<10	Apresentação sucessiva de imagens
10 à 16	Impressão de movimento mas com sensação de arrancos
>16	Efeito do movimento começa
24	Cinema
30/25	Padrão de TV americana/européia
60	Padrão HDTV

- Maior a frequência de quadros utilizada implica em uma mais alta é a largura de banda (faixa de frequência ou largura de canal) necessária à transmissão. Isto pois maior a taxa, maior é o número de quadros que devem ser enviados. Portanto, a rede utilizada pode ditar a frequência de quadros a ser utilizada.
- Problema de frequência de restauração (*refreshing*) de tela: a tela deve ser restaurada 50/60 vezes por segundo para evitar tremulações. Mas se a frequência de quadros for 50 fps, a largura de banda necessária aumentará substancialmente.

Imagens Bitmap Animadas (Vídeo)

Na animação de imagens, cenas são registradas como uma sucessão de quadros representados por imagens *bitmap* possivelmente compactadas. Estas imagens podem ser capturadas da vida real com câmeras ou criadas computacionalmente. A primeira técnica produz o que é chamado de **vídeo**.

Animação de imagens tem as mesmas características que as imagens: falta de uma descrição semântica e necessidade de uma grande capacidade de armazenamento.

Gráficos Animados

O termo gráfico animado ou animação gráfica é utilizado para referenciar apresentação sucessiva de objetos visuais gerados pelo computador em uma taxa suficiente para dar a sensação de movimento. Neste caso, cada atualização é comutada de uma descrição abstrata (equação matemática, por exemplo) em tempo de apresentação.

A principal vantagem das animações gráficas é que elas são mais compactas: elas são descritas por um conjunto de objetos com diretivas temporais (em outras palavras um programa a ser executado em tempo de apresentação). Outra vantagem é que animações gráficas são revisáveis, ou seja, podem ser modificadas. Uma desvantagem desta técnica é que é necessário uma velocidade de processamento suficiente grande para o êxito da apresentação.

2.1.7 Especificação da Cor

Propriedades da Cor

A luz visível é uma forma de radiação eletromecânica que tem um espectro de comprimento de onda variando aproximadamente de 400 nm a 780 nm. Uma luz de diferente comprimento de onda produz uma sensação de cor diferente (p.e. violeta de 380 a 450 nm, azul de 450 a 490 nm, verde de 490 a 560nm). As três propriedades físicas básicas da cor são: luminância (brilho), nuance (cor) e saturação (pureza).

Sistema de especificação de cores

Para comunicar imagens e vídeos coloridos, a cor deve ser especificada usando certo método. A CIE, Comissão Internacional de Iluminação, é o órgão responsável pela

padronização na área de fotometria e colorimetria. Existem vários padrões de cor estabelecidos pela CIE, entre eles o CIE-RGB e o CIE-XYZ.

2.1.8 Sistema RGB

No sistema RGB de representação de cor, uma cor é representada pela intensidade de três cores primárias (teoria Tristimulus): vermelho (Red), verde (Green) e azul (Blue), com cada valor variando de 0 a 255. Exemplos de cores familiares são apresentadas abaixo:

- Branco = 255,255,255
- Vermelho = 255,0,0
- Verde = 0,255,0
- Azul = 0,0,255
- Amarelo = 255,255,0
- Preto = 0,0,0

A representação de imagens coloridas pode ser feita por cores por componente (*true color*), cores indexadas, ou cores fixas. Essa representação vai depender do propósito e dos dispositivos que vão ser usados para trabalhar com essas imagens.

True Color

No sistema *True Color*, cada pixel da imagem é representado por um vetor de 3 componentes de cores (RGB) com um certo número de bits para representar cada componente de cor (resolução de cor). Com isso, quanto maior for a resolução de cor mais qualidade teremos para representar as cores de cada pixel. Geralmente o número de bits para cada componente RGB é igual, ou seja quando temos um pixel sendo representado por 9 bits, usamos 3 bits para cada componente (3-3-3). Mas, pode-se representar com diferentes número de bits as componentes. Por exemplo, uma representação 8 bits/pixel, pode ser usado 3 para componentes R, 3 para G e 2 para B (3-3-2), tal representação em um byte é comumente usado e tira proveito do fato que a percepção humana da componente azul é menos sensível que as outras componentes.

O número de bits para representar cada componente fornece a quantidade de cores que podem ser representados nesta nuance. Ou seja, se n é a resolução de cor então a quantidade de níveis possíveis é de 2^n níveis. Por exemplo, uma imagem colorida representada por 12 bits/pixel, com 4 bits para cada componente RGB. Temos então: $2^4=16$ níveis para cada componente de cor RGB, o que nos possibilita representar até 4.096 cores diferentes ($16 \times 16 \times 16 = 4.096$), o que é equivalente a $2^{12} = 4.096$.

Temos alguns padrões de cores nesse formato que são:

Bits/pixel	Padrão	Componente de cor RGB	Máximo de cores
15 bits/pixel	High Color (15 bits)	5 bits/pixel, 32 níveis por componente	32.768 cores
16 bits/pixel	High Color (16 bits)	5/6 bits/pixel, 32/64 níveis por componente	65.536 cores
24 bits/pixel	True Color, (24 bits)	8 bits/pixel, 256 níveis por componente	16.777.216 cores

O padrão com 24 bits/pixel é o mais usado para representar com fidelidade as cores, pois o número de cores que podem ser representadas com essa resolução de cores é maior do que a visão humana pode reconhecer.

Cores Indexadas

Nas cores indexadas, cada pixel é representado por um índice que aponta para uma tabela de cores (paleta) que contem as informações sobre as cores (Figura 4). Temos então um número de cores que podem ser representadas, que é o número de entradas na paleta. A paleta por sua vez, tem em geral 24 bits para representar cada cor no

formato RGB. Dessa forma podemos representar n cores de um conjunto com mais de 16 milhões de cores. Nesse caso, para representar esse tipo de imagem, as informações das cores da paleta devem constar da estrutura além das dimensões e seqüência de índices.

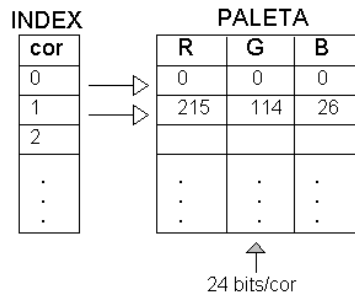


Figura 4. Índice e paleta de cores

O número de cores e a resolução de cor da paleta podem variar. Os dois padrões mais usados são apresentados na tabela abaixo.

Bits/pixel	Padrão	Resolução de cor da paleta (RGB)
4 bits/pixel	16 cores indexadas	24 bits/cor
8 bits/pixel	256 cores indexadas	24 bits/cor

Cores fixas

Nas cores fixas, cada pixel é representado por um índice que aponta para uma tabela de cores fixa. Esse sistema geralmente é usado quando o dispositivo não permite a representação de muitas cores, como no caso de placas de vídeos antigas ou padrões de cores (padrão de cores do MS Windows 3.x, 16 cores). O número de bits para representar um pixel depende do número de cores fixas, ou seja para representar por exemplo 16 cores, são necessários 4 bits/pixel.

Imagens em Tons de Cinza

A representação de imagens em tons-de-cinza é feita discretizando a informação de luminância de cada ponto da imagem. Ou seja, cada pixel contém a intensidade de luminosidade representada em um certo número de bits. Assim, uma imagem com resolução de cor de 8 bits pode representar até 256 níveis de cinza, variando do preto ao branco.

Os padrões mais usados são de 16 e 256 tons-de-cinza, 4 e 8 bits/pixel respectivamente. Representações com mais que 256 tons-de-cinza não são percebidas pela vista humana, ou seja representar uma imagem com 256 níveis é suficiente para a maioria das aplicações.

Imagens Binárias

As imagens binárias são imagens com dois níveis, como preto e branco. São muito usadas por dispositivos de impressão e para representar imagens de documentos monocromáticos. Para representar um pixel de uma imagem binária, como o próprio nome diz, é necessário apenas 1 (um) bit. Essa informação é suficiente para representar cada pixel, ou seja temos uma representação de 1 bit/pixel. Em alguns casos, temos uma informação extra sobre a cor de cada informação, a cor para o bit com valor 0 (zero) e a cor para o bit de valor 1 (um). Essa informação de cor é geralmente é representada em 24 bits/cor no padrão RGB, podendo porém ser representada de outras formas.