



Universidade Do Minho

Mestrado Integrado Em Engenharia Biomédica

Bibliotecas Digitais

QUESTIONÁRIO

Braga, Junho 2013



Universidade Do Minho

Mestrado Integrado Em Engenharia Biomédica

Bibliotecas Digitais

QUESTIONÁRIO

Docente: Joaquim Macedo

Discentes: Hugo Gomes (58536)

Marta Paes Moreira (54459)

Telma Veloso (58521)

Conteúdo

1	Compressão de Imagens	4
2	Imagens, Câmaras e Cor	7
3	Análise do Conteúdo da Imagem	14

1 Compressão de Imagens

1. Para efeitos de compressão de imagem, são passíveis de exploração as redundâncias estatística, espacial, estrutural, de conhecimento e psico-visual.

2. Os tipos de compressão disponíveis podem ser sub-divididos em dois domínios genéricos: sem perdas e com perdas. A compressão sem perdas, conforme a designação indica, reporta-se a uma técnica através da qual não se verificam quaisquer perdas de dados no decorrer do processo de compressão, correspondendo o ficheiro comprimido a uma cópia idêntica ao ficheiro na sua forma original, contrariamente à compressão com perdas, que descarta parte dos dados de forma a minimizar a quantidade de informação a ser processada e/ou transmitida, resultando na diminuição progressiva da fidelidade do ficheiro comprimido ao ficheiro original. Assim, enquanto que a compressão sem perdas apresenta a vantagem da reversibilidade, na grande maioria das vezes, a compressão com perdas permite a produção de um ficheiro comprimido significativamente mais pequeno que vai igualmente ao encontro dos requisitos da aplicação em questão. Naturalmente, esta apresenta a desvantagem de ser irreversível, sendo desadequada a quaisquer aplicações que requeiram o processamento e/ou transmissão de dados na íntegra, como sejam as que envolvem imagens médicas e de satélite.

3. A codificação *run-length* consiste numa técnica de compressão em que *runs* de dados – sequências nas quais valores semelhantes de dados ocorrem em múltiplos elementos de dados consecutivos – são armazenados enquanto valores individuais de dados e contados individualmente. Assim, a codificação é feita utilizando um *run* de cada vez, em vez de se processar *pixel a pixel*, e revela-se especialmente eficaz em situações que envolvam dados que contenham um grande número de *runs*.

4. A codificação preditiva, também conhecida como DPCM (*Differential Pulse Code Modulation*), baseia-se na predição de cada amostra a ser codificada, utilizando valores de amostras previamente codificadas, codificando (e transmitindo ou armazenando) apenas o erro de predição. Esta técnica baseia-se no pressuposto de que o sinal a codificar é redundante, isto é, que tem correlação e memória.

5. Assumindo a inexistência de erro de quantificação do sinal, o cálculo do erro da saída previsível para a imagem fornecida baseia-se na aplicação do preditor de 1ª ordem aos elementos da primeira linha e coluna da matriz e do preditor de 3ª ordem aos restantes elementos. A saída prevista (M_{out}) é então calculada, a título de exemplo, da seguinte forma:

$$M_{out1,2} = (25 \times 0.97) = 24.25$$

$$M_{out1,3} = (24 \times 0.97) = 23.28$$

...

$$M_{out2,2} = [(30 \times 0.9) - (25 \times 0.81) + (24 \times 0.9)] = 28.35$$

$$M_{out2,3} = [(48 \times 0.9) - (24 \times 0.81) + (23 \times 0.9)] = 44.46$$

...

Posteriormente, o erro da saída previsível (M_{DPCM}) é calculado subtraindo os valores obtidos para a saída à matriz original da imagem:

$$M_{DPCM1,1} = X$$

$$M_{DPCM1,2} = (24 - 24.25) = -0.25$$

$$M_{DPCM1,3} = (23 - 23.28) = -0.28$$

...

As matrizes completas assim obtidas, correspondentes à saída previsível (M_{out}) e ao erro que afecta essa saída (M_{DPCM}), respectivamente, figuram abaixo.

$$M_{out} = \begin{bmatrix} 25 & 24.25 & 23.28 & 22.31 \\ 24.25 & 28.35 & 44.46 & 32.67 \\ 29.1 & 50.4 & 35.82 & 26.91 \\ 33.95 & 38.25 & 4.86 & 37.35 \end{bmatrix} \quad M_{DPCM} = \begin{bmatrix} X & -0.25 & -0.28 & -4.31 \\ 5.75 & 19.65 & -5.46 & 7.33 \\ 5.9 & -6.4 & -10.82 & 14.09 \\ -3.95 & -18.25 & 18.14 & -15.35 \end{bmatrix}$$

6. Enquanto processo isolado, a codificação baseada na entropia não apresenta um ganho significativo, contrariamente ao que acontece quando é combinada com outros métodos, como seja com a transformada da imagem, tornando-se significativamente mais eficaz.

7. O teorema de Parseval, também conhecido como Teorema de Energia de Rayleigh, descreve uma das propriedades fulcrais da transformada unitária, enunciando que a soma (ou integral) do quadrado de uma dada função é igual à soma (ou integral) do quadrado da sua transformada, ou seja, que a energia total no domínio da frequência é igual à energia total no domínio espacial.

8. Comparativamente, a DCT (*Discrete Cosine Transform*) apresenta uma compactação de energia superior à DFT (*Discrete Fourier Transform*).

9. O cálculo das transformadas DFT e DCT foi efectuado recorrendo aos comandos `fft(x)` e `dct(x)`, respectivamente. Após a eliminação dos quatro coeficientes mais baixos,

os valores obtidos foram:

DFT	DCT
65.0000	24.5677
-4.7726	2.7941
-4.7726	0.9175

Assim, estabelecendo o valor médio dos valores de energia contemplados, verifica-se que, efectivamente, a transformada DCT apresenta uma maior compactação de energia – neste caso específico, cerca de o dobro da compactação conseguida pela transformada DFT.

10. O Fenómeno de *Gibbs* diz respeito à forma peculiar como a série de *Fourier* de uma função composta periódica e continuamente diferenciável se comporta na presença de uma descontinuidade de salto: a soma parcial de ordem n da série de *Fourier* apresenta oscilações de grande magnitude nas proximidades do salto, o que pode, potencialmente, elevar o máximo da soma parcial acima do máximo da própria função. Este fenómeno é considerado uma das causas de artefactos em processamento de sinal, sendo responsável, numa imagem, pela perda de contraste quando os coeficientes de alta frequência apresentam erros de quantificação.

11. A compressão de imagem baseada no vector de quantização consiste numa técnica de processamento de sinal que permite modelar as funções de densidade de probabilidade através da distribuição de vectores protótipo, baseando-se o seu funcionamento na segmentação de um amplo conjunto de vectores – uma imagem, por exemplo – em blocos com, aproximadamente, o mesmo número de vectores na sua proximidade, cada um dos quais representados por um centróide. Neste tipo de codificação, existe um dicionário que gera etiquetas utilizadas para rotular cada um dos blocos, etiquetas essas que são armazenadas na imagem compactada em detrimento do armazenamento dos referidos blocos, requerendo um menor número de *bits* na representação.

12. Por sua vez, a compressão de imagens por fractais baseia-se no conceito de fractal – a imagem de uma textura ou forma expressa enquanto uma ou mais fórmulas matemáticas, denominadas de códigos fractais – para determinar o conjunto de fractais capazes de descrever ou representar uma imagem digital. Esta técnica é sobretudo adequada a imagens naturais e de texturas, uma vez que apresentam alguma similaridade interna.

2 Imagens, Câmaras e Cor

1. Numa passagem do mundo 3D para representação de imagens 2D perde-se informação relativa a ângulos e profundidades - 3D indica que existe a noção das três dimensões - altura, largura e profundidade. Enquanto que 2D, duas dimensões, perde-se informação que só com a representação 3D é possível saber e visualizar.

2. Uma câmara *pinhole* é uma câmara simples sem a habitual lente e em seu lugar apenas uma pequena abertura, basicamente, é uma caixa-escura (à prova de luz) com um pequeno buraco num dos lados. A luz da cena a capturar passa pela abertura e projecta uma imagem invertida da cena no lado oposto da caixa. Curioso referir que o olho humano, quando exposto em ambientes muito luminosos, tem um comportamento muito semelhante a este método.

3. No sistema cartesiano a translação não pode ser expressa por multiplicação de matrizes porém existe outro sistema de coordenadas usado em computação gráfica que permite tais operações, o sistema de coordenadas homogêneas.

As coordenadas homogêneas são úteis porque conseguem definir em matrizes únicas e independentes operações como translação, factor de escala e rotação, isto porque as coordenadas homogêneas possuem mais uma dimensão do que o espaço que representam (exemplo: em $\mathbb{R}^2$ existem 3 coordenadas homogêneas associadas).

Em $\mathbb{R}^2$, se um ponto é representado por (X, Y) , no sistema de coordenadas homogêneas, esse ponto pode ser representado por (x, y, w) onde $x = X/w$ e $y = Y/w$ e que representa um ponto na localização $(x/w, y/w)$. O valor w denomina-se de peso, desta forma um mesmo ponto pode ser representado por um conjunto infinito de combinações.

4.

4.a

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \left(\begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(90) & -\sin(90) & 0 \\ \sin(90) & \cos(90) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x \\ y \\ w \end{bmatrix}$$

4.b

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \left(\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x \\ y \\ w \end{bmatrix}$$

4.c

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \left(\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x \\ y \\ w \end{bmatrix}$$

5. O uso de lentes nas câmaras fotográficas permite focar a luz para o filme fotográfico e desempenha ainda um papel importante para focar propriamente a imagem mediante a distância a que os objectos pretendidos estejam localizados. Isto é conseguido através da reconfiguração da distribuição da luz consoante a forma da lente independentemente do tipo de luz ou até mesmo ondas acústicas.

6. Ponto focal, no campo da Óptica, é uma zona em que os vários raios luminosos que passam por um refractor estão mais juntos e capazes assim de formar uma imagem mais nítida e definida. Este facto pode ser observado quando se utilizam as lentes para concentrar luz solar e aquecer uma determinada área, por exemplo, com o recurso a uma lupa.

7. Profundidade de campo, *depth-of-field*, no que se refere ao cinema de fotografia, é a gama de distâncias em torno do plano focal na qual existe uma nitidez aceitável dos objectos.

De um modo geral, quanto menor for a abertura do diafragma para uma o mesmo valor de distância do objecto fotografado, maior será a distância do plano de foco a que os objectos podem estar enquanto permanecem nítidos. A profundidade de campo depende também do tipo de câmara, abertura, distância e também é influenciada pelo tamanho da impressão e pela distância a que se está a visualizar a imagem.

8. Em Fotografia, à medida que se aumenta a profundidade de campo os objectos mais perto ficam menos pronunciados e os mais distantes apresentam uma maior nitidez. A objectiva da câmara também interfere com a qualidade obtida da imagem à medida que se aumenta a profundidade de campo na medida em que para algumas objectivas o aumento da profundidade de campo implica perda de qualidade na fotografia obtida. Dois exemplos de diferentes profundidades de campo podem ser visualizados nas Figuras 1 e 2, respectivamente baixo e maior valor de profundidade de campo.

9. O campo de visão, *field-of-vision*, é a extensão angular de um determinado ambiente



Figura 1: Exemplo de fotografia com baixo valor de profundidade de campo.



Figura 2: Exemplo de fotografia com maior profundidade de campo.

que é possível ver. A título de exemplo, os animais têm diferentes valores de campo de visão, consoante a sua espécie ou estatuto presa/predador que lhes permitem conseguir características vantajosas às suas necessidades. Os humanos possuem um campo de visão aproximadamente de 180 graus enquanto que alguns pássaros conseguem visualizar 360 graus.

Porém os animais possuem dois receptores visuais, olhos, importante para a percepção da profundidade, esta visão binocular infere perda no campo de visão para cerca de 140 graus nos humanos. O valor de campo de visão também varia verticalmente e não só horizontalmente.

10. Na escolha de uma máquina fotográfica há algumas características que merecem mais atenção, entre elas destacam-se a resolução, exposição, focagem, lente usada e a forma de armazenamento e compressão da imagem.

10.a O ruído está para a fotografia digital como os grãos das películas estão para a

fotografia analógica. O ruído está sempre presente e surge como manchas aleatórias numa superfície originalmente suave, no entanto pode ser tão pequeno que se considere inexistente. As máquinas digitais produzem três tipos de ruído que são os mais comuns - aleatório, de padrão fixo e de bandas - que podem ser observados na Figura 3. Por vezes, o ruído é



Figura 3: Tipos de ruído gerado por máquinas digitais.

produzido com propósitos artísticos, no entanto considera-se que o ruído diminui a qualidade da imagem, uma vez que consiste em interferências no sinal que se pretende obter.

A resolução de uma imagem descreve o nível de detalhe que uma imagem tem. Uma resolução mais elevada significa que a imagem tem maior detalhe. A resolução pode ser definida pela distância mínima entre dois pontos sem que estes fiquem indistinguíveis, isto é, a distância mínima necessária para visualizar ambos os pontos, nitidamente. O termo resolução também é usado como uma contagem de pixels numa imagem digital. Normalmente representa-se por $N \times M$ onde N é número de colunas e M o número de linhas de pixels. Na Figura 4 é possível ver a mesma imagem com diferentes resoluções de pixels. Outra maneira de representar a resolução é pela quantidade de pixels numa imagem. Este valor é calculado através da multiplicação do número de colunas pelo número de linhas e depois dividindo por um milhão, obtendo assim o número de megapixels. Em máquinas digitais compactas, como o sensor de imagem é pequeno, ter uma grande quantidade de megapixels numa superfície pequena torna-se desvantajoso porque os pixels têm tendência a se influenciar mutuamente, dando origem a ruído na imagem obtida.

Algumas máquinas digitais permitem a realização de alguns ajustes e melhoramentos depois de tirada a fotografia. Isto apenas é possível quando a fotografia é tirada em modo JPEG, quando é tirada em modo RAW o utilizador apenas pode fazer os ajustes mais tarde, no computador. A máquina guarda a fotografia alterada como uma cópia, mantendo a

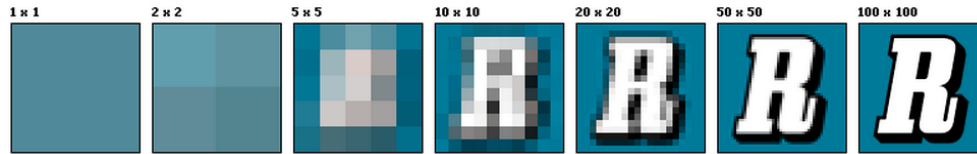


Figura 4: Tipos de ruído gerado por máquinas digitais.

versão original. É possível fazer vários ajustes, entre eles, corrigir os olhos vermelhos, cortar, alinhar, redimensionar e aplicar efeitos de cor.

O objectivo da compressão é reduzir a informação redundante e irrelevante de modo a que a transmissão e o armazenamento da imagem sejam mais eficiente. A compressão permite reduzir o espaço ocupado pelo ficheiro da imagem e pode ter ou não perdas. Quando a compressão tem perdas, não é possível recuperar a imagem original. Este tipo de compressão é usado quando a transferência da imagem é mais importante que a qualidade, como é o caso das máquinas fotográficas digitais, uma vez que gravam informação que o olho humano não é capaz de detectar. A título de exemplo, o formato **png** usa compressão sem perdas, enquanto que o formato **jpg** usa compressão com perdas.

Blooming é um efeito que ocorre quando um pixel num sensor CCD (*charged-coupled device*) é sobrecarregado, isto é, a carga de um pixel ultrapassa o nível de saturação e os pixels vizinhos começam a ser preenchidos por essa carga. A estrutura do sensor CCD permite que os electrões se movimentem mais numa direcção do que noutra, dando origem a faixas verticais, como se pode observar na Figura 5

Por vezes, as máquinas digitais introduzem artefactos indesejados como, por exemplo, ruído colorido em excesso ou *blooming* colorido. Isto acontece quando as definições da máquina apresentam um valor elevado para o ISO. Para evitar a presença de artefactos de cor em fotografias é realizado um ajuste chamado *white balance* que permite tornar neutras as cores neutras.

11. Apesar do grande espectro de radiações electromagnéticas, apenas as radiações cujo comprimento de onda está entre os 380 e 750 nm é que são visíveis ao olho humano. Vide Figura 6.

12. A cor é uma percepção visual provocada pela acção de um feixe de fotões sobre células específicas da retina (cones). A cor de um material é determinada pelo valor médio das frequências (ou comprimento de onda) das radiações que reflecte. Um objecto possui uma determinada cor quando não absorve a radiação cuja frequência (ou comprimento de onda) corresponde a essa cor. Consulte a Figura 7 para ver os intervalos de cada cor.

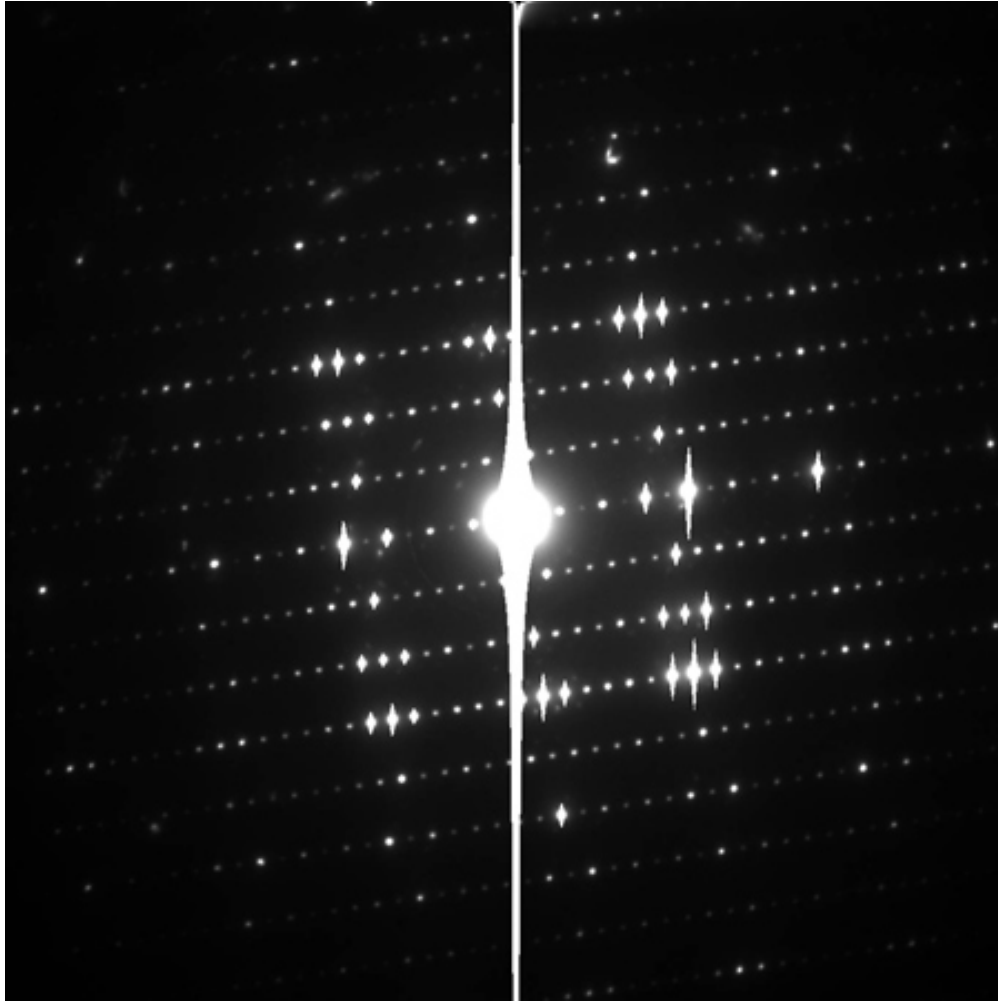


Figura 5: Faixas verticais originadas pelo efeito de *blooming*.



Figura 6: Espectro da luz visível.

13. Segundo o modelo HSB - *hue, saturation, brightness* - a saturação é o grau de pureza, ou seja, o grau de mistura entre a matiz de uma determinada cor e a cor branca. Enquanto que, segundo o modelo HSL - *hue, saturation, luminosity* -, a saturação é a proporção entre a quantidade de uma determinada cor em relação ao cinzento médio. Quanto menos cinzento,

Cores do espectro visível		
Cor	Comprimento de onda	Frequência
vermelho	~ 625-740 nm	~ 480-405 THz
laranja	~ 590-625 nm	~ 510-480 THz
amarelo	~ 565-590 nm	~ 530-510 THz
verde	~ 500-565 nm	~ 600-530 THz
ciano	~ 485-500 nm	~ 620-600 THz
azul	~ 440-485 nm	~ 680-620 THz
violeta	~ 380-440 nm	~ 790-680 THz

Figura 7: Cores do espectro visível.

mais saturada é a cor. Vide Figura 8.

Gradação de saturação no modelo HSL				
matiz 100% puro	75% de saturação	saturação média	25% de saturação	0 de saturação

Gradação de saturação no modelo HSV				
matiz 100% puro	75% de saturação	saturação média	25% de saturação	0 de saturação

Figura 8: Saturação segundo dois modelos diferentes.

14. Brilho consiste na intensidade da radiação emitida por um objecto, no entanto também pode ser a quantidade de luz, medida em watts/cm^2 .

15. RGB é um sistema de cores aditivas, formado por vermelho (**R**ed), verde (**G**reen) e azul (**B**lue), onde luzes destas cores são adicionadas segundo diferentes combinações para dar origem a outras cores. Este sistema é usado em monitores de televisão e computadores, enquanto que nas impressoras é usado o modelo CMYK (de cores subtractivas). É possível representar uma cor de forma numérica com este modelo, indicando a quantidade de vermelho, verde e azul que contém. Quando as três componentes estão no máximo, o resultado é branco e quando estão no mínimo, o resultado é preto.

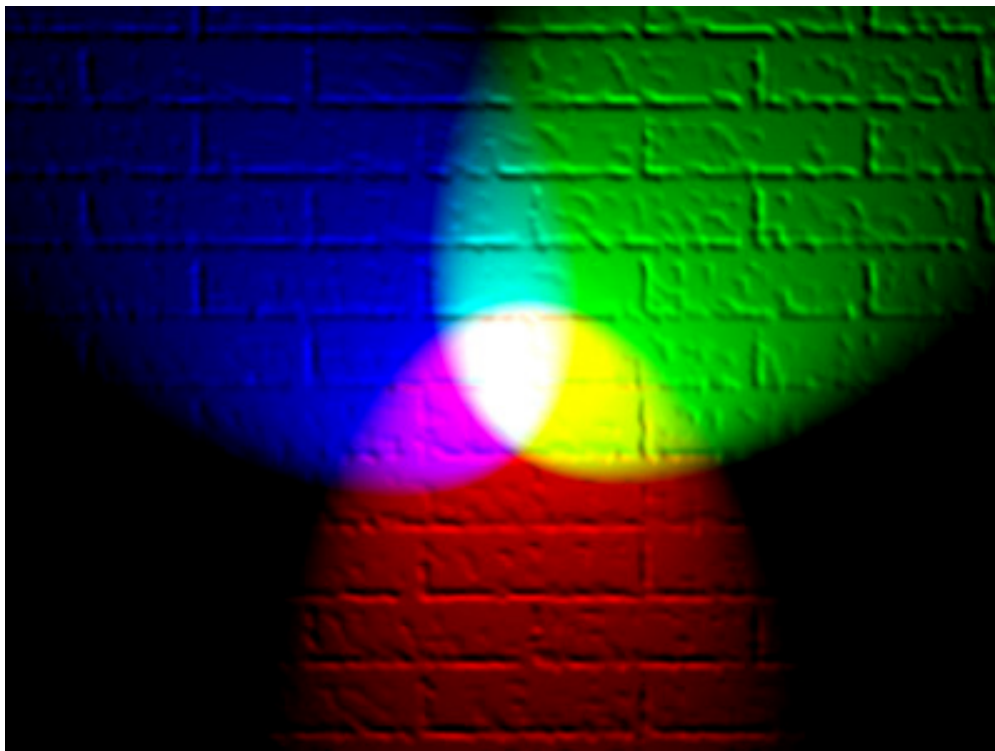


Figura 9: Representação do modelo aditivo RGB.

3 Análise do Conteúdo da Imagem

1. A digitalização é feita em duas partes, primeiro ocorre a amostragem da imagem analógica e depois dá-se a quantização. Na primeira fase, a imagem é percorrida pixel a pixel e são recolhidos valores consoante uma frequência de amostragem constante. Este passo também pode ter o nome de discretização, uma vez que o sinal passa de contínuo a discreto. De seguida, durante a quantização, a cada uma das amostras é atribuído um valor que pertence a um conjunto fixo de valores. As imagens são posteriormente armazenadas como uma matriz tridimensional, com os valores de vermelho (R), verde (G) e azul (B) para cada pixel amostrado anteriormente.

2. A aplicação de filtros em imagens permite obter vários efeitos. Existem vários filtros possíveis de aplicar, tais como a convolução, desfocagem, detecção de contornos, *sharpen*, média e mediana. Por exemplo, a convolução permite filtrar o sinal de entrada de modo a que o sinal de saída tenha frequência limitada. Por sua vez, a detecção de contornos permite encontrar zonas da imagem onde ocorrem mudanças repentinas ou descontinuidades. Por fim, alguns filtros substituem cada pixel pela média (pesada ou não) ou pela mediana da sua vizinhança, dando origem a uma imagem onde não há grande variação entre o valor dos

pixels, isto é, suaviza os valores extremos.

3. Histograma é um gráfico que mede a luminosidade de uma imagem e permite visualizar a quantidade de pixels que existem em cada nível (desde 0 até 255). No lado esquerdo do gráfico encontram-se os pixels escuros (próximos de 0) e no lado direito estão os pixels claros (próximos de 255). Uma imagem deve ter um histograma equilibrado, no entanto, há imagens cujo histograma é desequilibrado propositadamente. Na Figura 10 é possível ver o exemplo de um histograma.

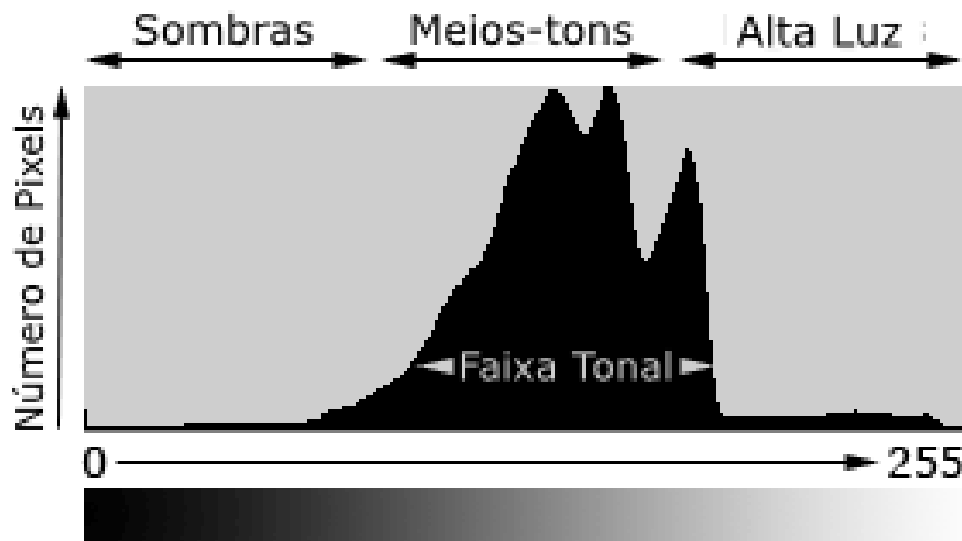


Figura 10: Exemplo de um histograma de uma imagem.

4. A forma de uma imagem é determinada pelos limites externos de um objecto. É possível representar uma forma em torno de uma localização característica através de um histograma da distância dos contornos numa imagem relativamente à localização dada.

5. A textura de uma imagem é um conjunto de medidas calculadas através de processamento de imagem destinado à quantificação de textura de uma imagem. A textura permite obter informação sobre o arranjo espacial da cor ou intensidade numa imagem. A textura pode ser criada artificialmente ou não e é bastante útil na segmentação ou classificação de imagens.