

Processamento de Imagens

Paulo César Moraes de Menezes

13 de março de 2024

Sumário

1	Algumas informações iniciais sobre o documento	3
2	Créditos	3
3	Introdução	3
3.1	O que é Processamento Digital de Imagens?	3
3.1.1	Processos de baixo nível	4
3.1.2	Processos de médio nível	4
3.1.3	Processos de alto nível	4
3.2	As origens do Processamento Digital de Imagens	4
3.3	Aplicações do Processamento Digital de Imagens	5
3.3.1	Imagens Raio-Gamma	6
3.3.2	Imagem de Raio-X	6
4	Fundamentos de Imagens Digitais	6
4.1	Elementos da percepção visual	6
4.1.1	A visão humana	6
4.1.2	Formação da Imagem no Olho	9
4.1.3	Adaptação ao brilho e discriminação de cores	10
4.2	Luz e o Espectro do campo Eletromagnetico	15
4.3	Detecção e Aquisição de imagem	17
4.3.1	Obtenção de imagem através de um único sensor	17

1 Algumas informações iniciais sobre o documento

Este documento terá todo o meu estudo direcionado ao tema de Processamento Digital de Imagens, a ideia é adicionar informações que foram adquiridas tanto na disciplina Processamento de Imagens (DCE536) quanto em outras fontes de conhecimento. A ideia é que este documento seja uma espécie de guia para o meu estudo, e que possa ser útil para outras pessoas que estejam interessadas no tema. Além disso dedico o meu estudo para aprofundar o meu conhecimento em Inteligência Artificial, que é a área que pretendo seguir na minha carreira profissional.

2 Créditos

Algumas das informações presentes nesse documento foram extraídas do conteúdo fornecido pelo professor Luiz Eduardo da Universidade Federal de Alfenas. Quero agradecer a ele por compartilhar seu conhecimento na área. Além disso, algumas outras informações serão extraídas de fontes externas, que serão devidamente referenciadas. Além disso, os textos são nota de base do livro "Digital Image Processing" de Rafael C. Gonzalez e Richard E. Woods.

3 Introdução

3.1 O que é Processamento Digital de Imagens?

Uma imagem pode ser definida como uma função de duas dimensões $f(x, y)$, onde x e y são coordenadas espaciais em um plano e a amplitude de f em qualquer par de coordenadas (x, y) é chamada de intensidade ou nível de cinza da imagem naquele ponto. Quando os valores de x , y e f são todos finitos e discretos, chama-se a imagem de imagem digital. No que diz respeito ao campo de processamento digital de imagens (PDI), o termo refere-se a processar uma imagem por meio de um computador digital. É válido destacar que uma imagem digital é composta por finitos elementos, cada elemento possui um valor de intensidade e posição associado a ele.

Com essas informações em mente, também é válido destacar um dos grandes responsáveis por dar vida a capacidade de reconhecer imagens: a visão humana. A visão humana é um dos grandes sentidos que o ser humano possui, e é também um dos mais complexos. A visão é capaz de capturar e processar informações visuais de forma extremamente rápida e eficiente. Entretanto, a visão humana não é perfeita, e também possui limitações, como por exemplo não ser capaz de enxergar certos comprimentos de onda de luz, ou ser limitado pelo espectro EM (Eletromagnético) que é capaz de capturar.

Contudo, as máquinas conseguem capturar quase todo o espectro EM, e com isso é possível processar imagens de forma extremamente eficiente, e é aí que entra o Processamento Digital de Imagens, que é a área que estuda técnicas e métodos para processar imagens de forma eficiente, e é uma das áreas que possui grande importância na área de Inteligência Artificial, pois é uma das áreas que estuda a capacidade de reconhecer padrões em imagens, que é uma das capacidades que o ser humano possui.

As máquinas conseguem operar em faixas de frequência que o ser humano não consegue, como por exemplo raios-x, raios gama, entre outros.

É comum existir um debate do que seria processamento de imagens e visão computacional. Esse debate é comum, pois as duas áreas possuem muitos pontos em comum, e muitas vezes são utilizadas de forma intercambiável, entretanto, é válido destacar que as duas áreas possuem diferenças, e que a visão computacional é uma área que estuda como as máquinas podem ser capazes de interpretar e entender o mundo visual, enquanto que o processamento de imagens é uma área que estuda como as imagens podem ser processadas de forma eficiente. O processamento de imagens concentra-se principalmente em técnicas e algoritmos para manipular e melhorar imagens digitais. Isso inclui operações como filtragem, realce de borda, remoção de ruído, entre outros, com o objetivo de melhorar a qualidade ou extrair informações úteis das imagens.

3.1.1 Processos de baixo nível

Esses processos são conhecidos também como processos primitivos ou pré-processamento, e são processos para reduzir ruído, realçar bordas, corrigir desfoque, entre outros.

Um processo de baixo nível é caracterizado pelo fato de que a sua entrada e saída são imagens, e que não existe um conhecimento prévio sobre o conteúdo da imagem, ou seja, o processamento é feito de forma cega, sem saber o que está sendo processado.

3.1.2 Processos de médio nível

O processo de médio nível envolve técnicas de segmentação, que é o processo de dividir a imagem em partes, ou seja, é o processo de dividir a imagem em partes que são de interesse, e partes que não são de interesse. Essa segmentação pode ser feita separando a imagem em regiões ou em objetos. Além disso esse tipo de processo envolve uma descrição sobre os objetos segmentados de modo a reduzi-los de uma forma adequada para o processamento computacional e a classificação de objetos individuais.

Além disso uma outra característica desse tipo de processo está no fato de que as entradas são imagens, mas as saídas são atributos extraídos dessas imagens, como por exemplo, a forma, a cor, a textura, entre outros.

3.1.3 Processos de alto nível

Esse tipo de processo envolve dar sentido ao conteúdo da imagem, ou seja, envolve a interpretação do conteúdo da imagem, e é um dos processos mais complexos, e que envolve a utilização de técnicas de inteligência artificial, como por exemplo, redes neurais, entre outros.

3.2 As origens do Processamento Digital de Imagens

Uma das primeiras aparições das imagens digitais foi no jornal industrial no começo dos anos 1920, na qual foi enviado através de submarinos entre Londres e Nova York, e foi a primeira vez que uma imagem foi transmitida de forma eletrônica. Esse feito reduziu para horas um trabalho que levava semanas para ser feito.

Essas capturas pioneiras tinham capacidade de reconhecer até 5 tons de cinza. Esse valor foi aumentando até que em 1929 alcançou a incrível marca de 15 tons de cinza.



Figura 1: Imagem digital produzida por fita codificada.

Um grande fator que limitou o avanço das imagens digitais foi a falta de tecnologia para armazenar essas imagens, é nítida a relação de dependência entre poder computacional e armazenamento de imagens.

Ao longo do século XX, o PDI foi se desenvolvendo, e foi se tornando uma área de grande importância para a sociedade, e com a medida que a tecnologia computacional evoluía, principalmente durante a corrida espacial, o PDI foi se tornando uma área de grande importância para a sociedade.



Figura 2: Primeira imagem digital da Lua, em 1964.

Além disso, durante a década de 70, o PDI começou a ter aplicações em medicina, e com isso o PDI começou a ter um grande impacto na sociedade, principalmente com o desenvolvimento de Tomografias Computadorizadas, que é uma das grandes aplicações do PDI.

3.3 Aplicações do Processamento Digital de Imagens

Atualmente praticamente todas as áreas da sociedade possuem aplicações do PDI, e é uma área que exerce grande influência na sociedade.

3.3.1 Imagens Raio-Gamma

A vasta maioria dos usos nesta área é para aplicações de medicina nuclear e observações astronômicas.

A ideia para a medicina nuclear é aplicar isótopos radioativos em um paciente, e com isso é possível observar o decaimento desses isótopos. Desse modo, é possível coletar imagens do interior do corpo humano, e com isso é possível diagnosticar doenças.

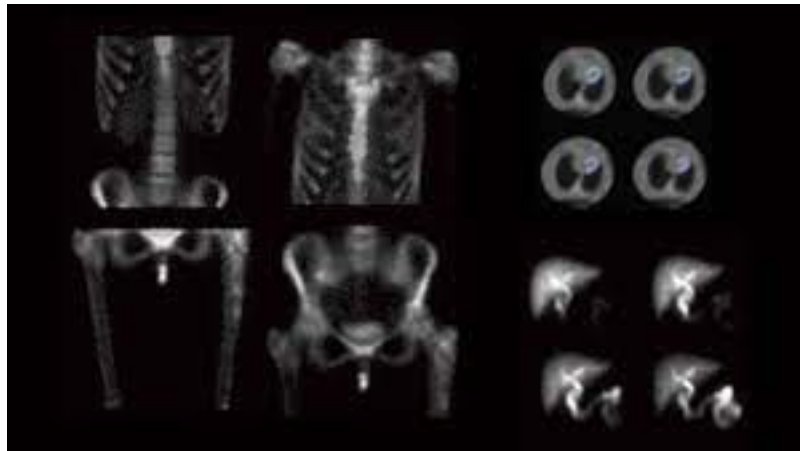


Figura 3: Gamma Ray Image

3.3.2 Imagem de Raio-X

Esta é uma das aplicações mais antigas do PDI, e é uma das aplicações mais conhecidas do PDI. Usada principalmente para diagnóstico médico, mas pode ser utilizada em outras áreas, como por exemplo, para inspeção de bagagens em aeroportos.

4 Fundamentos de Imagens Digitais

4.1 Elementos da percepção visual

É notório que, baseado no que foi visto até então, o campo do processamento digital de imagens é baseado em alguns elementos matemáticos e físicos, entretanto a análise e a intuição humana de como as coisas funcionam é de grande importância para o desenvolvimento de algoritmos e técnicas de processamento de imagens.

Com isso em mente é fundamental compreender as limitações da percepção humana, e como a percepção humana é capaz de interpretar e entender o mundo visual.

4.1.1 A visão humana

A figura 5 exemplifica a seção transversal do olho humano. O olho é quase uma esfera, com diâmetro aproximado de 24mm.

Três membranas principais cobrem o olho:



Figura 4: Raio-X

- A córnea é um tipo de tecido mais rígido, e é a parte que cobre a parte frontal do olho.
- Tem-se a esclera que encobre o restante do globo ocular.
- A coróide fica na parte de baixo da esclera e é responsável por fornecer oxigênio e nutrientes para o olho. A importância da coróide fica mais nítida ainda à medida que quaisquer danos a ela podem causar danos irreversíveis à visão.
- O coróide é dividido em duas partes: O corpo ciliar e o íris. O corpo ciliar é responsável por produzir o humor aquoso, que é um líquido que preenche a câmara anterior do olho. O humor aquoso é responsável por manter a pressão do olho e nutrir a córnea e o cristalino. A íris é a parte colorida do olho, e é responsável por controlar a quantidade de luz que entra no olho. A pupila de entrada pode variar de $2mm\ f/8,3$ em ambientes bem iluminados até $8mm\ f/2,1$ em ambientes escuros.
- O cristalino é composto de cerca de 65% de água e 35% de proteínas. Ele é responsável por focar a luz que entra no olho na retina. Em alguns casos extremos, o cristalino pode ser coberto por uma membrana chamada catarata, que pode causar cegueira. Ela é responsável por absorver cerca de 8% do espectro da luz visível.
- A membrana mais interna, a retina é responsável por converter a luz que entra no olho em sinais elétricos que são enviados para o cérebro. A retina é composta de dois tipos de células fotorreceptoras: os cones e os bastonetes. Os cones são responsáveis por detectar a cor e funcionam melhor em ambientes bem iluminados. Os bastonetes são responsáveis por detectar a luz e funcionam melhor em ambientes escuros. Uma outra estrutura importante é a Fóvea, que é responsável por fornecer a visão central e é composta apenas por cones.

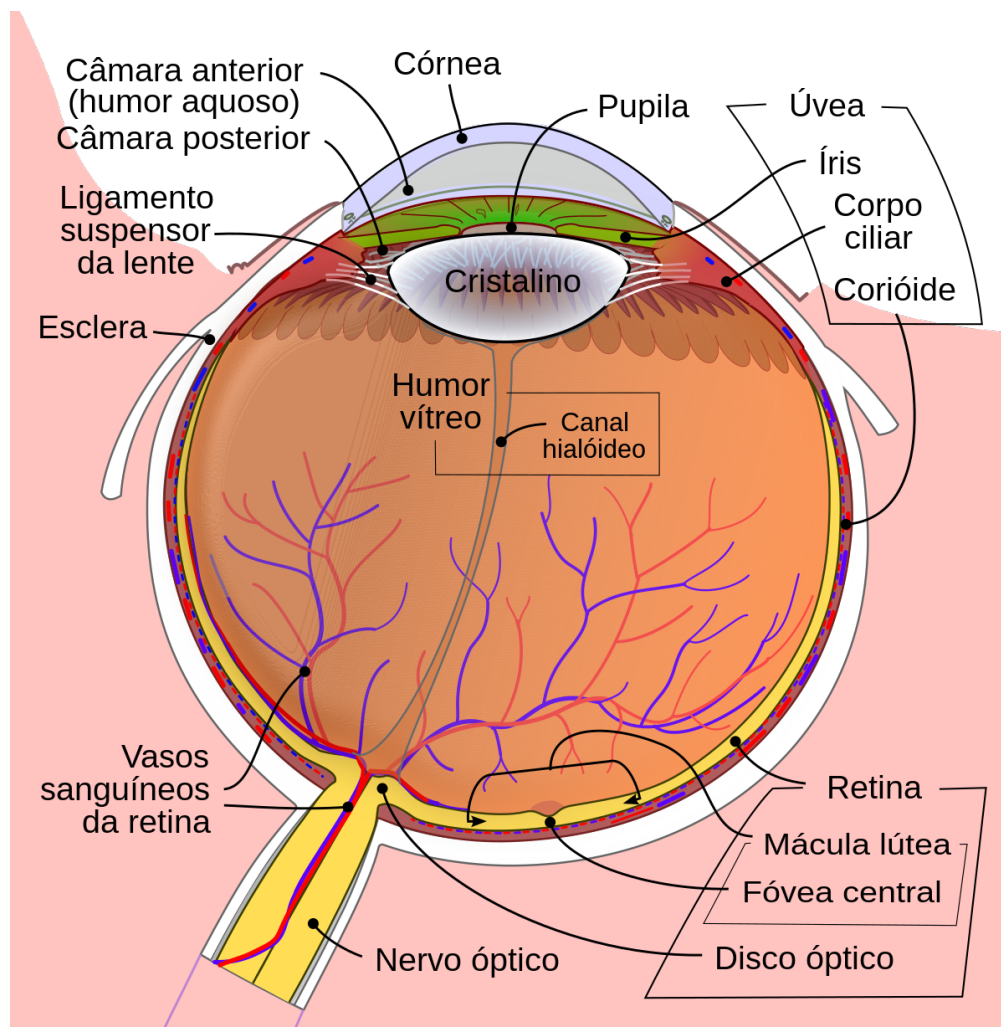


Figura 5: Seção transversal do olho humano.

Uma outra característica importante do olho humano se diz respeito a cones e bastonetes, que são responsáveis por capturar a luz e transformar em sinais elétricos que são enviados para o cérebro.

Cones: Em cada olho formam um total de cerca de 6 milhões de cones, e são responsáveis por capturar a luz e transformar em sinais elétricos que são enviados para o cérebro. Os cones são responsáveis por capturar a cor, e são responsáveis por capturar a luz em ambientes bem iluminados. Esse tipo de visão é chamada de visão fotópica.

Bastonetes: Em cada olho formam um total de cerca de 120 milhões de bastonetes, e são responsáveis por capturar a luz e transformar em sinais elétricos que são enviados para o cérebro. Os bastonetes são responsáveis por capturar a luz em ambientes escuros, e são responsáveis por capturar a luz em ambientes escuros. Esse tipo de visão é chamada de visão escotópica.

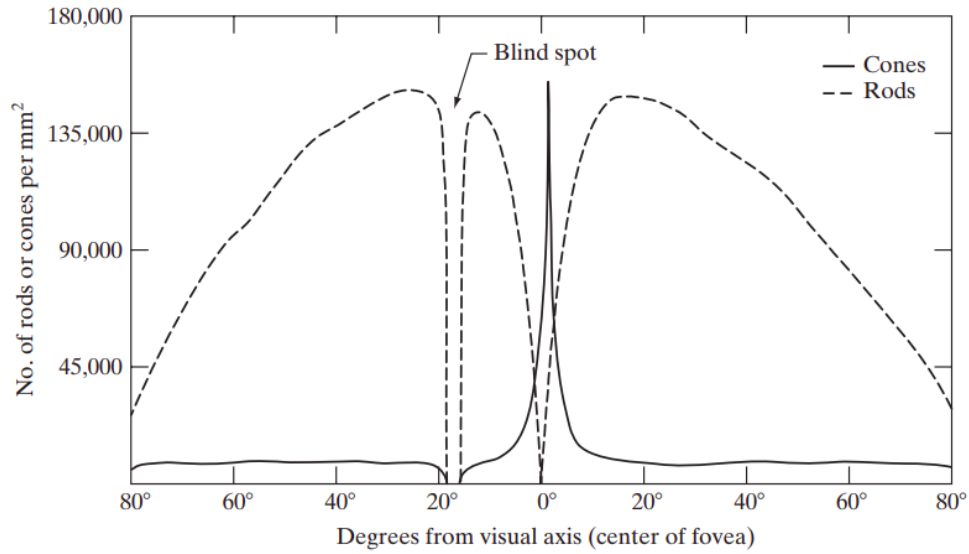


Figura 6: Distribuição de cones e bastonetes na retina.

A figura 8 mostra a distribuição de cones e bastonetes na retina.

4.1.2 Formação da Imagem no Olho

Nas fotografias as lentes possuem um ponto focal já estabelecido e também uma distância focal fixa, enquanto que em caso de filmes ou sensores digitais, a distância focal pode ser ajustada para focar em diferentes distâncias. Já no olho humano, a distância entre a lente e a região da imagem é fixa e o local de obtenção do foco é ajustado por meio de um músculo que altera a forma da lente.

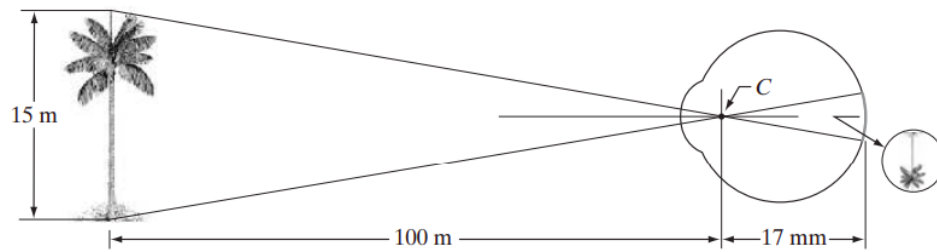


Figura 7: Exemplo da formação da imagem no olho

A distância entre o centro da lente e a retina é de aproximadamente 17mm, e a distância focal da lente é de aproximadamente 15mm. A imagem é formada na retina, e a imagem é invertida, e o cérebro é responsável por inverter a imagem.

A figura 8 mostra como obter dimensões das imagens formadas na retina. Por exemplo, suponha que uma pessoa está olhando uma árvore de 15 metros de altura a uma distância de 100metros. Chamando de h a altura da imagem formada na retina a geometria obtida é a seguinte:

$$\frac{h}{17} = \frac{15}{100} \quad (1)$$

Em outras palavras, a equação geral pode ser escrita como:

$$\frac{h}{d} = \frac{H}{D} \quad (2)$$

Onde h é a altura da imagem formada na retina, d é a distância entre a lente e a retina, H é a altura do objeto real, e D é a distância entre o objeto e a lente.

Como visto na seção anterior a imagem da retina é focada primariamente na fóvea, que é responsável por fornecer a visão central, e é composta apenas por cones.

4.1.3 Adaptação ao brilho e discriminação de cores

Devido as imagens digitais serem exibidas em tons discretos de valores, a habilidade de discriminar cores é uma tarefa a ser considerada para os olhos humanos. A distribuição de níveis de intensidade da luz que um sistema visual humano pode adaptar é vasta, a ordem chega a ser algo em torno de 10^{10} do limiar escotópico ao limite de ofuscamento. Existem algumas evidências de que a subjetividade do brilho é uma função logarítmica da intensidade da luz. A figura 8 mostra a curva

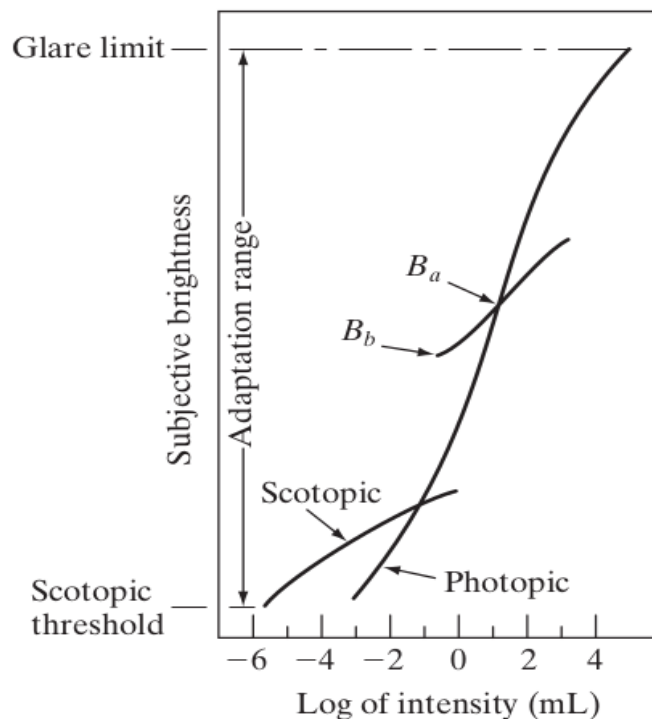


Figura 8: Curva de adaptação ao brilho

de adaptação ao brilho, que é uma curva logarítmica. A sólida e longa curva representa a capacidade de adaptação do sistema. Uma visão fotópica é capaz de discriminar uma faixa de cerca de 10^6 . A transição escotópica-fotópica ocorre em torno de 0.001 a 0.1 Millilamberts.

Para qualquer conjunto de condições de iluminação, o atual nível de sensibilidade a luz do sistema visual é chamado de adaptação ao brilho. Olhando para a figura 8, é possível ver que, por exemplo, o brilho B_a representa uma curva de intersecção curta e representa a faixa de brilho

subjetivo que o olho pode perceber quando está adaptado a este nível. Um fato sobre este intervalo é que ele é extremamente estreito, fechado, ter um nível B_B no qual e abaixo do qual todos os estímulos são percebidos como indistintos negros capazes. A parte superior da curva não é realmente restrita, mas, se for extendido ou levado longe demais, perde seu significado porque intensidades muito mais altas simplesmente elevar o nível de adaptação acima de B_a

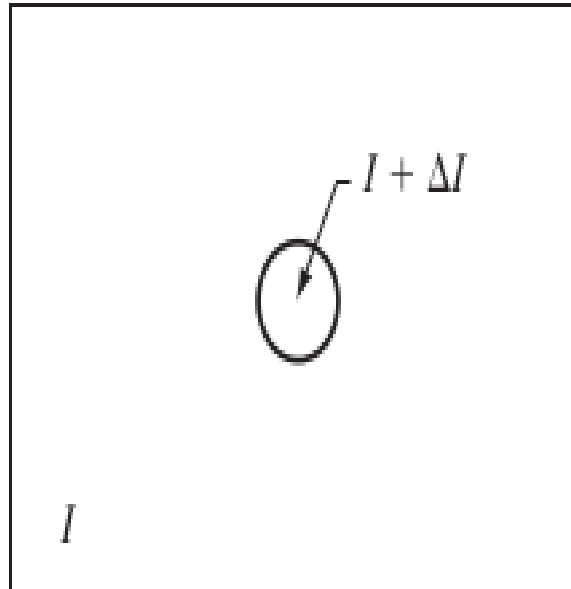


Figura 9: Experimento básico para caracterizar discriminação do brilho

Um outro fato interessante está relacionado com a proporção de Weber. Essa proporção vem da ideia da habilidade do olho humano em conseguir discriminar certas mudanças entre as alterações da intensidade da luz a um certo nível de adaptação considerada interessante. Considere um experimento usado para determinar a capacidade do sistema humano visual para discriminar o brilho. O experimento consiste em ter um olhar subjetivo a um plano uniformemente iluminado em uma área larga o suficiente para ocupar um campo de visão por inteiro. Essa área normalmente é difusa, como um vidro opaco, que é iluminada por uma fonte de luz com intensidade I , e pode ser variada. Nesse campo é adicionado um incremento a iluminação ΔI em uma curta duração, ou seja um flash aparece em um círculo no centro uniforme iluminando o campo, conforme pode ser visualizado na figura 9. Se o ΔI não tiver brilho o suficiente o responsável pelo experimento irá dizer não, indicando uma mudança pouco perceptível. A medida que ΔI fica mais forte, o observador pode dar uma resposta positiva indicando uma mudança perspectiva na luminosidade. Até chegar no ponto de que a resposta é sempre sim. A quantidade de $\Delta I_c/I$ onde ΔI_c é o incremento da iluminação discriminável em 50% de todo o tempo de iluminação no fundo I é o que chamamos de Proporção de Weber. Um pequeno valor de $\Delta I_c/I$ indica uma grande porcentagem de mudança na intensidade requisitada. Isso representa uma discriminação pobre.

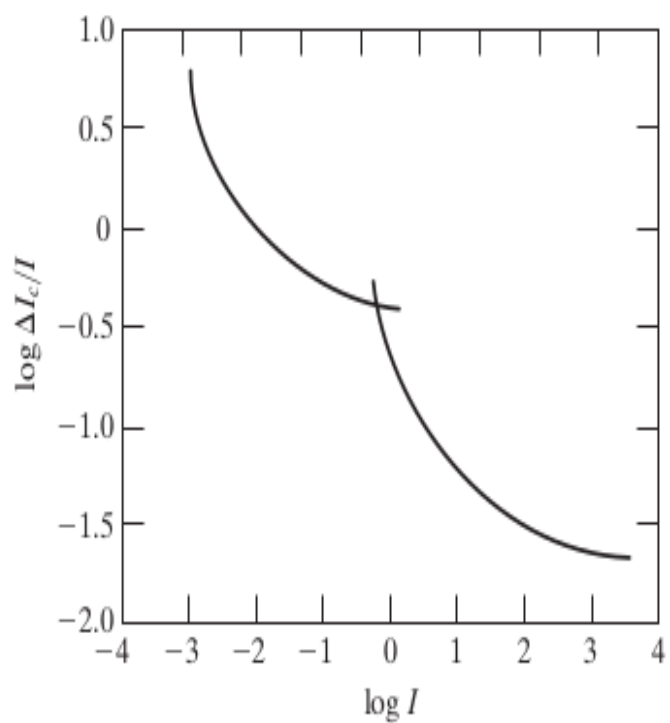


Figura 10: Curva de discriminação do brilho

Uma plotagem do log de $\Delta I_c/I$ como uma função logaritmica de I tem um formato que pode ser visualizado na figura 10

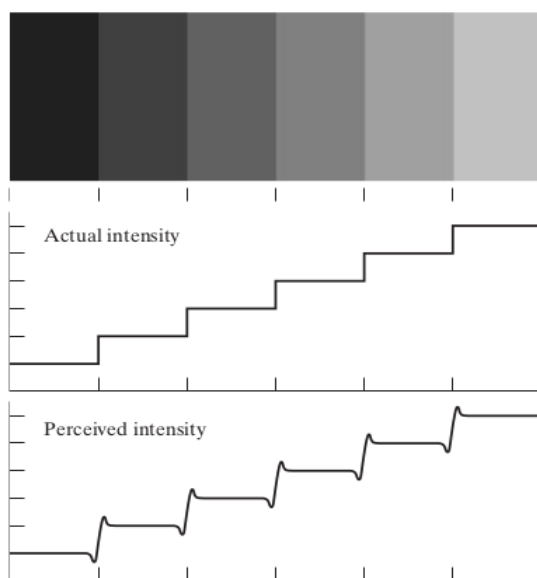


Figura 11: Ilustração do efeito Mach e ilustração da função de intensidade

Efeito Mach: O efeito Mach é uma ilusão visual em que linhas paralelas parecem convergir

em um ponto quando observadas em uma imagem bidimensional. No processamento digital de imagens, isso pode causar distorções perceptuais se não for considerado, exigindo técnicas específicas para corrigir ou mitigar o efeito, como correção de distorção ou uso de algoritmos de interpolação.

Além disso o fenômeno abordado na figura 10 ilustra que o brilho recebido não é unicamente uma função de intensidade. Isso pode ser nitidamente visto na figura 11



Figura 12: Exemplificação do efeito de contraste simultâneo

Um outro fenômeno é o chamado de contraste simultâneo, isto está associado ao fato de que uma região de cor cinza parece mais clara quando colocada em um fundo escuro, e mais escura quando colocada em um fundo claro. Isso é um fenômeno que ocorre devido a interação entre a região e o fundo. Esse fenômeno é ilustrado na figura 12

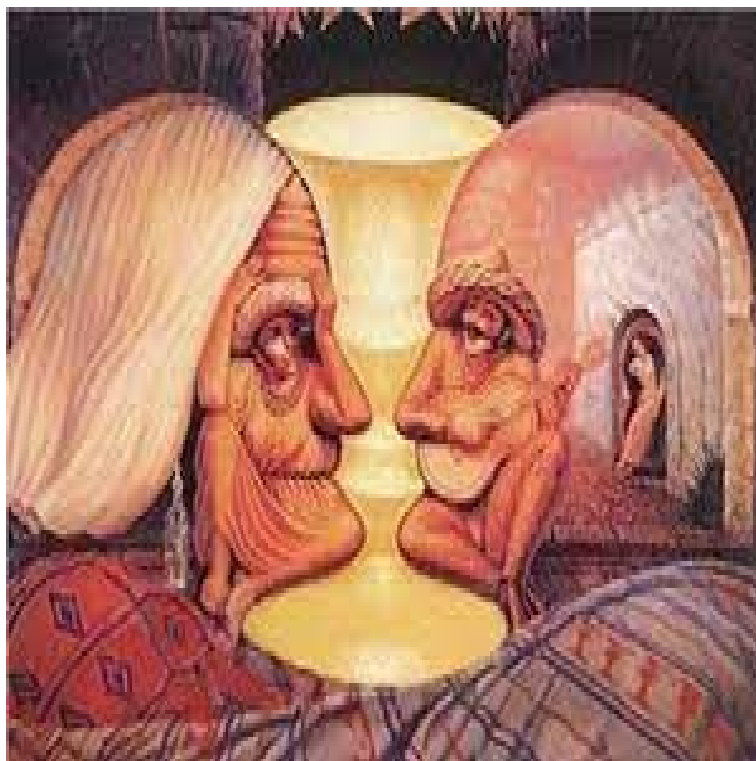


Figura 13: Exemplo de ilusões de óptica

Outro exemplo até um tanto quanto divertido está associado aos fenômenos de ilusões de óptica. Essas ilusões são fenômenos que ocorrem quando o cérebro interpreta a informação visual de uma forma que não corresponde à realidade. Isso ocorre devido a uma série de fatores, como por exemplo, a forma como a luz incide sobre um objeto, a forma como a luz é refletida, entre outros. Alguns exemplos podem ser vistos na 13, onde pode ser visto no rosto do senhor um homem com uma espécie de violão na mão, enquanto que no rosto da mulher pode ser visto uma mulher segurando uma espécie de chapéu. Existem diversos tipos de ilusões de óptica.

4.2 Luz e o Espectro do campo Eletromagnético

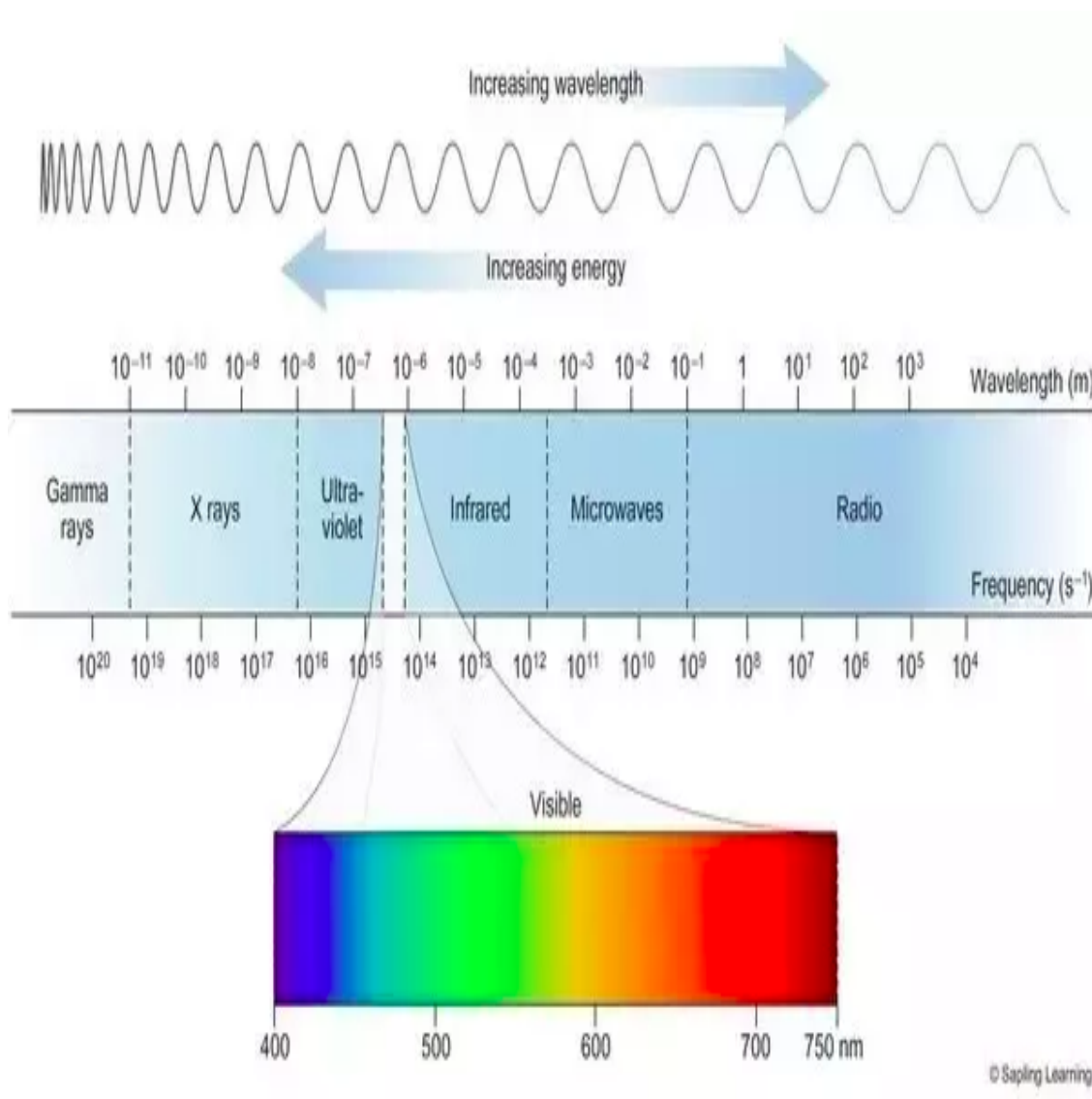


Figura 14: Espectro do campo eletromagnético

Conforme pode ser evidenciado na figura 14, o espectro visível é apenas uma pequena parte do espectro eletromagnético. O campo do espectro eletromagnético pode ser representado por comprimentos de onda, frequência ou energia. Comprimento de onda é representado por λ e a frequência pode ser representada por ν . A equação que relaciona comprimento de onda e frequência é dada por:

$$\lambda = \frac{c}{\nu} \quad (3)$$

Onde c é a velocidade da luz no vácuo, que é aproximadamente $3.0 \times 10^8 \text{ m/s}$. A energia de um fóton é dada por:

$$E = h\nu \quad (4)$$

Onde h é a constante de Planck, que é aproximadamente $6.626 \times 10^{-34} J.s$. A energia de um fóton é diretamente proporcional a frequência, e inversamente proporcional ao comprimento de onda.

A frequência é abordada em Hertz(Hz), e é uma medida de ciclos por segundo. A frequência é uma medida de quão rápido um fóton oscila.

Ondas eletromagnéticas podem ser vistas como propagação sinusoidal de campos elétricos e magnéticos que são perpendiculares entre si e perpendiculares a direção de propagação da onda.

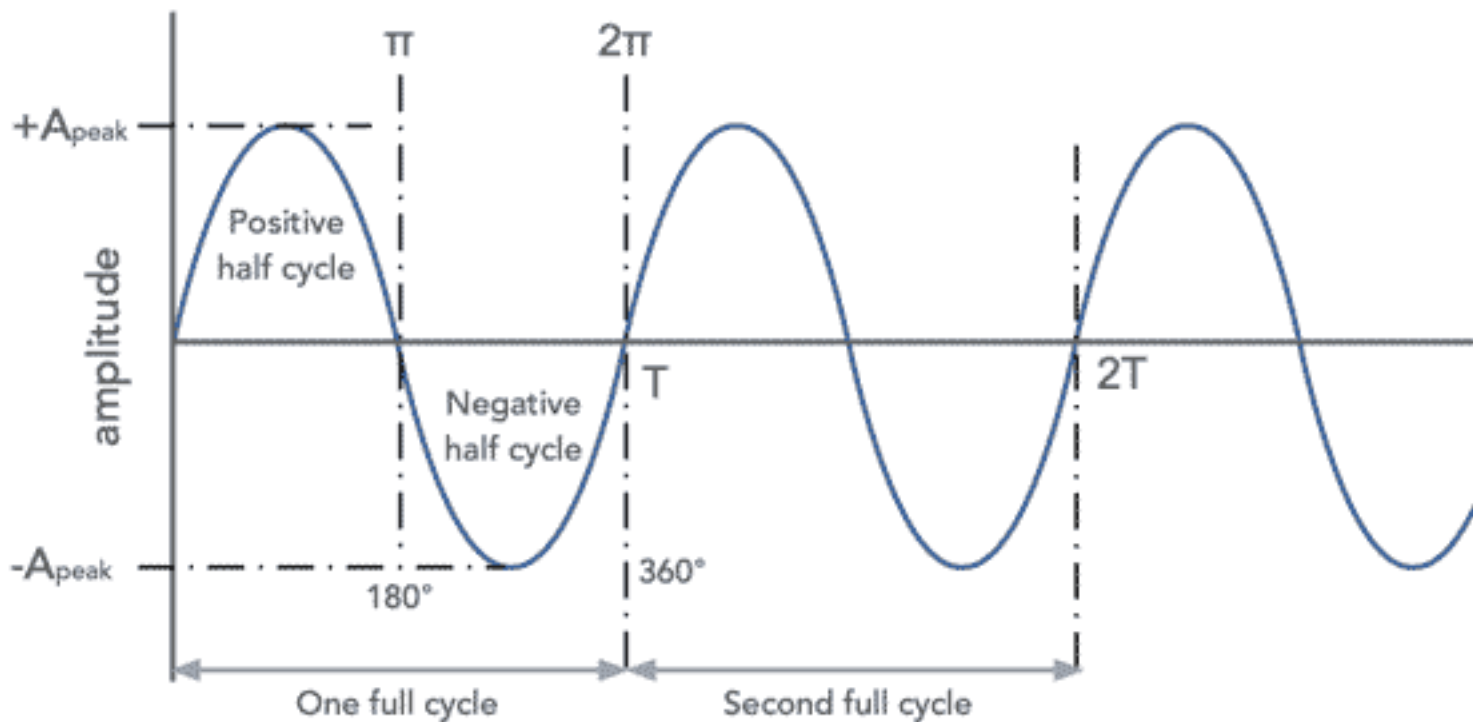


Figura 15: Exemplo de onda sinusoidal

Cada pacote de energia é chamado de fóton. É visto na equação 4 que a energia de um fóton é diretamente proporcional a frequência, e inversamente proporcional ao comprimento de onda.

Com essas informações em mente tem-se um fato: Por que ondas de raio gamma são mais perigosas que ondas de rádio? A resposta está na equação 4, que mostra que a energia de um fóton é diretamente proporcional a frequência, e inversamente proporcional ao comprimento de onda. Com isso, ondas de raio gamma possuem mais energia que ondas de rádio, e com isso são mais perigosas.

A faixa de cores pode ser vista na imagem 14, e é uma faixa que vai de aproximadamente $400nm$ até $700nm$. A luz branca é composta por todas as cores, e a cor branca é a cor que é percebida quando todas as cores são misturadas. As cores que são enxergadas são as cores que são refletidas pelos objetos.

4.3 Detecção e Aquisição de imagem

A grande maioria das imagens analisadas são geradas pela combinação de luz e sensores. A luz atua como um sinal que é capturado por um sensor, e é convertido em um sinal elétrico que é processado por um computador. Por exemplo um exame de raio-x é um exemplo de imagem que é gerada pela combinação de luz e sensores.

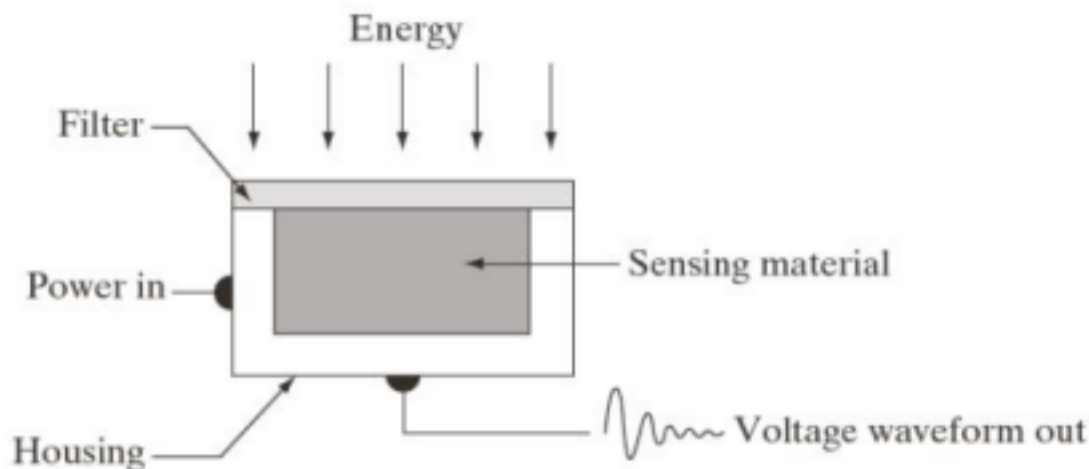


Figura 16: Exemplo de um sensor de imagem simples

Observando a imagem 16, vê-se que a energia luminosa é passada por uma espécie de filtro que é responsável por separar a luz em suas componentes de cor, e cada componente de cor é passada por um sensor que é responsável por converter a luz em um sinal elétrico.

Existem alguns outros sensores, como por exemplo, Sensores Lineares e Sensores de Matriz.

4.3.1 Obtenção de imagem através de um único sensor

O mais comum componente desse tipo de sensor é o chamado fotodiodo, que é um componente que é responsável por converter a luz em um sinal elétrico. O uso de um filtro nesse componente consegue aumentar a seletividade espectral do sensor, por exemplo, um filtro que melhor absorve a luz azul consegue capturar melhor a luz azul.