

Esta lista de exercícios está baseada nas referências listadas ao final desse documento.

1) Pesquise a respeito e responda:

- a) Quais são as principais diferenças entre as tecnologias de sensores CCD e CMOS;
- b) Responda justificando. Entre Global Shutter Timing ou Rolling Shutter Timing, qual é o tipo de leitura de sensor mais adequado para uso na aquisição de vídeo de imagens nas seguintes situações:
 - i. Captura de uma cena de pessoas caminhando em um corredor de um aeroporto;
 - ii. Captura de imagens de uma hélice a 900 rpm (rotações por minuto).

- 2)
- a) Qual é a aparência de cada canal RGB quando exibido individualmente?
 - b) É possível se ter uma imagem **visualmente** em tons de cinza por meio de uma estrutura RGB com os 3 canais (três matrizes de brilhos)? Qual seria a característica de cada canal, nesse caso?
 - c) Existem outros sistemas de cor além do RGB? Cite alguns casos e suas características.

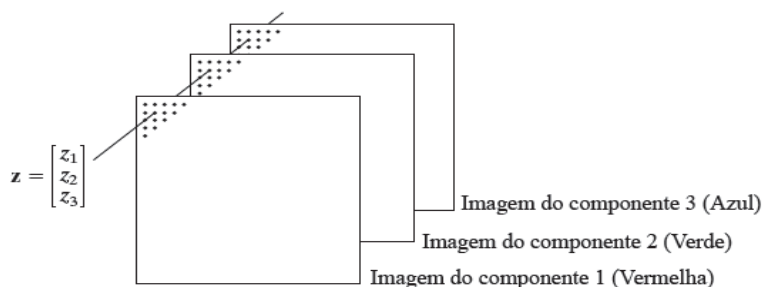


Figura 2.38 Formação de um vetor a partir de valores de pixels correspondentes em três imagens de componentes RGB.

3) Pesquise sobre a captura e representação de imagens em formato RGB (figura abaixo, extraída de Gonzalez e Woods [6]).

- a) É possível converter uma imagem no formato RGB (três matrizes de brilhos) para um formato em tons de cinza (uma matriz de brilhos). Na verdade, você verá que existem pelo menos duas formulações para a conversão de RGB para tons de cinza. Discuta sobre elas e escreva script Python que recebe a imagem de entrada RGB e fornece a equivalente imagem de saída em tons de cinza (<https://www.baeldung.com/cs/convert-rgb-to-grayscale>);
- b) Explique o método da pseudo-cor (falsa cor) usado para a conversão de uma imagem em tons de cinza para uma versão em RGB. Esse método pode ser aplicado na colorização de filmes em “preto e branco” ou para enfatizar certas áreas de uma imagem, por exemplo;
- c) explique o método de cross-fading usado para transições suaves de quadros em vídeo e/ou simular sequências de morfismo entre quadros de um vídeo.

4) Qual o tamanho em pixels de uma imagem gerada na saída de um sensor monocromático (tons de cinza) cuja área de amostragem é de 0,2 cm x 0,3 cm, cada célula desse sensor tem dimensões 0,01mm x 0,02 mm?

5) Suponha que a amostragem realizada por um sensor digital produziu a matriz abaixo, valores em mV (milivolts). Cada valor representa uma intensidade de brilho (tons de cinza) vinda da cena adquirida.

A) Considerando que se trate de um sistema embarcado em um hardware com uma severa restrição de memória. Quantos bits por pixel (bpp) você utilizaria para representar todos as diferentes intensidades de brilho capturadas na matriz amostrada? Discuta o caso e faça a discretização e digitalização dessa imagem, produzindo uma imagem digital de saída.

6.8	6.8	6.8	6.8	6.8	6.8
1.8	1.8	1.8	1.8	1.8	1.8
6.8	6.8	6.8	6.8	6.8	6.8
1.8	1.8	1.8	1.8	1.8	1.8
6.8	6.8	6.8	6.8	6.8	6.8
1.8	1.8	1.8	1.8	1.8	1.8

B) Considere que a sua câmera possa realizar digitalizações com as seguintes profundidades de pixels (diferentes resoluções radiométricas): 1bpp, 2 bpp, 4bpp, 8 bpp, 12 bpp. Compare os resultados obtidos para discretização e digitalização dessas imagens em termos da resolução de detalhes das imagens. Discuta um critério de escolha, entre as alternativas de digitalização, se o sistema de processamento de imagens (posterior à captura da cena) tiver restrições de memória.

6.8	6.87	6.89	6.89	6.87	6.8
5.8	5.87	5.89	5.89	5.87	5.8
4.8	4.87	4.89	4.89	4.87	4.8
3.8	3.87	3.89	3.89	3.87	3.8
2.8	2.87	2.89	2.89	2.87	2.8
1.8	1.87	1.89	1.89	1.87	1.8

6) Um mosaico de filtros Bayer é uma matriz de filtros de cores (CFA) para o padrão RGB em uma matriz de fotossensores. Seu arranjo particular de filtros de cores é usado na maioria dos sensores de imagem digital, de chip único, usados em câmeras digitais, filmadoras e scanners para criar uma imagem colorida.

Para a imagem ao lado no padrão Bayer, construa os três canais RGB correspondentes utilizando a técnica (interpolação) baseada em médias de vizinhos explicada em sala.

Descreva como você tratou os pixels nas bordas do padrão durante para a construção dos canais (RGB).

10	130	15	110	15	120
215	40	250	30	250	40
15	255	15	255	15	230
210	30	255	45	250	45
10	115	10	110	10	115
110	30	110	35	115	45

Figura 1: Exemplo de padrão Bayer onde a primeira linha é iniciada em Red-Green e a segunda em Green-Blue

7) Uma das transformações mais básicas consiste em operar uma imagem comparando cada pixel com um valor de limiar (threshold). Normalmente isso é aplicado para separar o fundo da imagem (background) em relação à sua parte mais importante (foreground).

Escreva um script Python que lê o padrão Bayer exibido na Figura 1, converte na imagem colorida RGB (I_{RGB}) e desta para a sua versão em tons de cinza (I_{gray}). Então um operador de limiar (*thresholding*) é aplicado à I_{gray} selecionando os pixels de I_{gray} que se encontrem no intervalo $[L1, L2]$ ($L1$ e $L2$ são tons de cinza fornecidos pelo usuário) formando a imagem de saída $I_{\text{saída}}$.

Exiba as imagens (matrizes) envolvidas.

A seção 3.2.3 do Solomon descreve o operador de Limiar (thresholding).

8) Qual é a relação entre a resolução espacial de uma imagem e o desempenho computacional do processamento de imagens, especialmente para transformações baseadas em pixels?

9) Estime o tamanho do menor ponto circular de diâmetro d que conseguimos discernir, se a página da impressão do ponto estiver a 0,2 m de distância dos olhos.

Suponha que o sistema de visão humano deixa de detectar o ponto quando a sua imagem na fóvea for menor do que o diâmetro de um receptor (cone) nessa área da retina.

Também suponha que a fóvea seja modelada por um sensor de 580x580 cones uniformemente distribuídos em quadrados de tamanhos iguais (ignore alguma linha de separação entre cones).

Considere o sensor com dimensões 1,5 mm x 1,5 mm. A distância focal deve ser considerada da ordem de 17,0 mm. O esquema encontra-se na figura abaixo e a imagem real de uma amostra da região da fóvea encontra-se na Figura 2.

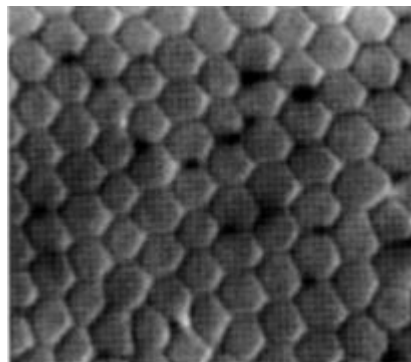
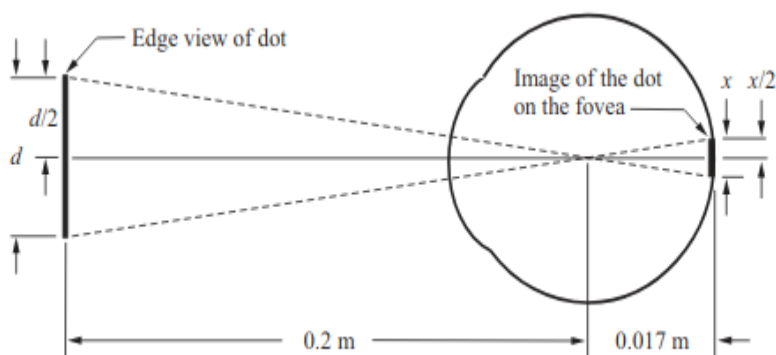


Figura 2: Cones reais na região da fóvea.

10) Considerando a disponibilidade de uma câmera de orifício (pinhole) com distância focal 5 mm e um prédio de 27 metros de altura que pode ter sua imagem capturada por esta câmera, apresentando uma altura de 150 pixels.

É possível utilizar a imagem em questão para estimar a distância do prédio em relação à câmera? Quais parâmetros adicionais são necessários para determinar essa estimativa? Qual seria o modelo geométrico que representa a situação? Estipule tais parâmetros e calcule a distância estimada da câmera ao prédio.

11) Um fabricante de automóveis está automatizando a instalação de componentes dos para-choques dos carros produzidos. Esses componentes são coordenados de acordo com suas cores de forma que os robôs precisam saber a cor de cada carro para selecionar o componente adequado dos para-choques. Os carros são fabricados apenas nas cores: cinza médio, branca, preta, vermelha, azul e verde. Descreva uma estratégia simples, barata e eficiente para a determinação automática da cor do carro baseado na aquisição de imagens.

Não é permitido utilizar filtros coloridos nas lentes das câmeras.

12) Um policial sem conhecimento de programação está analisando um vídeo e precisa estimar a velocidade de um carro no vídeo de 30 *frames per second* (fps), com uma câmera posicionada conforme figura abaixo.

Além da velocidade da câmera, é conhecida a distância entre os pontos x_0 e x_3 exibidos na mesma figura. A partir dos dados disponíveis (vídeo de 30fps, distâncias, etc), descreva como o policial deve proceder para inspecionar o vídeo e determinar a velocidade desejada.

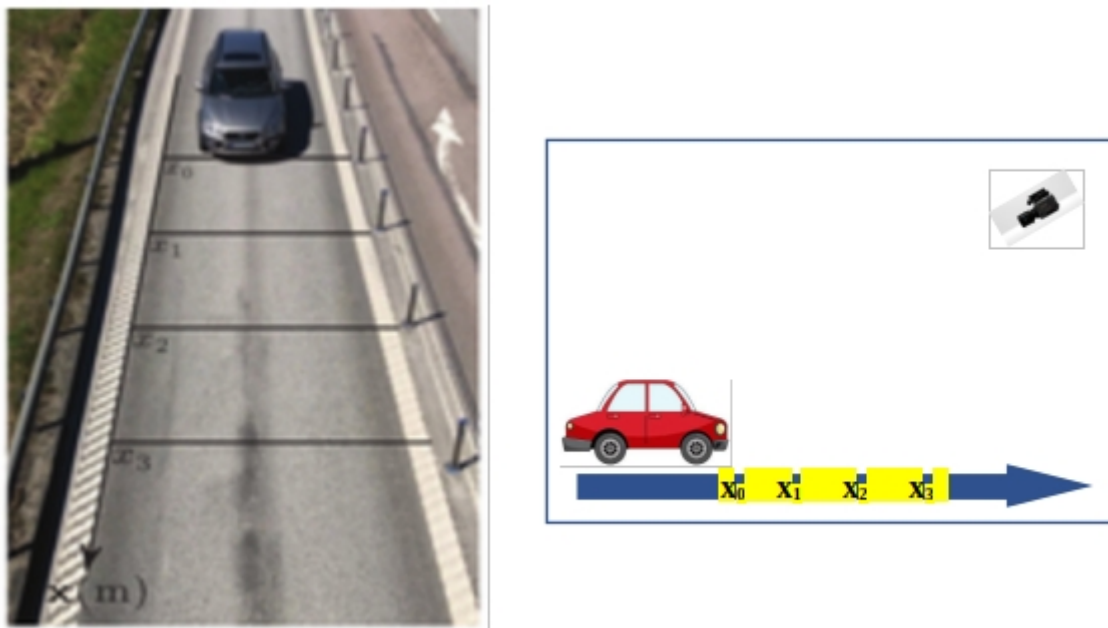


Figura 3: Um quadro (frame) do vídeo que exhibe o deslocamento do carro na direção do eixo X exibido na figura.

13) Dada uma peça retangular, com regiões escuras e claras e escuras, paralela ao plano-de-imagem (Figura 4). O vetor normal ao plano da peça é colinear ao eixo **OZc** (eixo focal). O foco (centro de projeção) encontra-se na origem $[0,0,0]^T$ e o eixo **OZc** intercepta a peça (Figura 4).

O centro de projeção encontra-se na origem $[0,0,0]^T$, a distância focal d entre o centro de projeção f e o plano de imagem é igual a 5 mm .

Cada pixel do sensor é um quadrado de lado $7,5 \times 10^{-6}\text{ m} = 0,0075\text{ mm}$.

Coordenadas do centro do sensor: $0x=0y=1024$ **pixels**

A distância do objeto (entre o centro de projeção e a folha) é de 1.500 mm .

O sensor é binário: um pixel está aceso (preto) ou apagado (branco). Um pixel está aceso se um ou mais raios projetores incidirem sobre o mesmo, caso contrário o pixel é apagado (branco).

Tem-se como **objetivo** a análise da imagem capturada visando a identificação dos **detalhes** do objeto, para a tomada de decisões em um projeto mecânico. Os detalhes em questão são vistos como reentrâncias no objeto (Figura 4). As coordenadas 3D (em mm) dos oito pontos identificados no objeto (Figura) constam na Tabela-1 abaixo:

Tabela 1 – Oito pontos marcados na Figura 4.

1.	a :	650,7	2.000,0	1.500,0
2.	b :	653,5	2.000,0	1.500,0
3.	c :	650,7	1.990,0	1.500,0
4.	d :	653,5	1.990,0	1.500,0
5.	e :	645,3	500,3	1.500,0
6.	f :	645,0	500,3	1.500,0
7.	g :	645,3	500,0	1.500,0
8.	H :	645,0	500,0	1.500,0

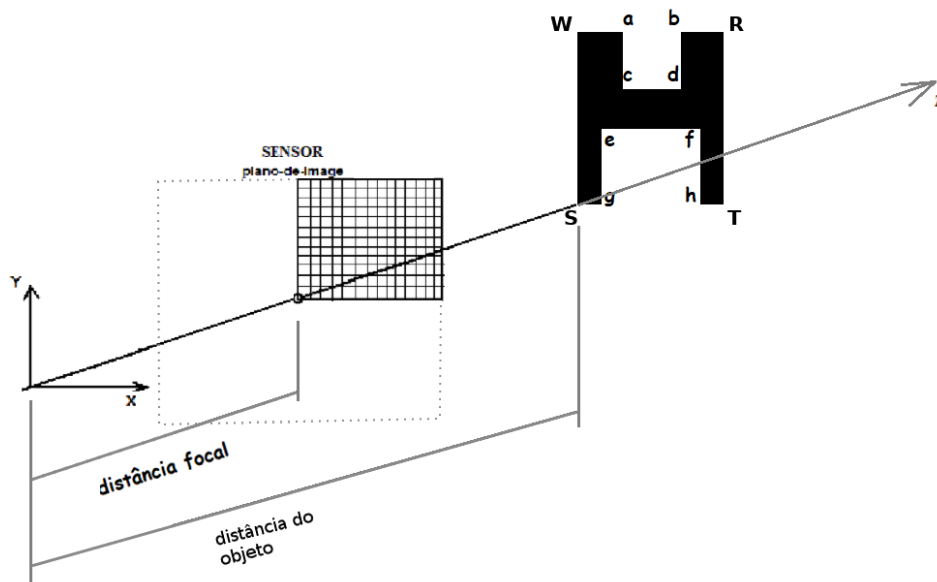


Figura 4: Esquema utilizado na questão.

Responda justificando:

A) Qual é a área do quadrilátero delimitado pelos pontos W, R, T e S em mm^2 e pixels^2 ?

W(0,0;2000,0;1500,0),
 R(1304,2;2000,0;1500,0),
 T(1304,2;0,0;1500,0),
 S(0,0;2000,0;1500,0)

B) A imagem capturada é útil ao propósito (objetivo) citado anteriormente? Escreva o script Python que implementa os cálculos aplicados na solução da questão e faça upload no Moodle.

C) Caso seja oferecida uma outra câmera com o mesmo tamanho de sensor (mesmas dimensões de altura e largura), qual dos sensores abaixo seriam da sua escolha? Justifique sua resposta:

- S1: cada *pixel* é um quadrado de lado 0,00042 mm.
- S2: cada *pixel* é um quadrado de lado 0,085 mm.

14) Cite exemplos de operações lógicas e aritméticas entre imagens. Cite pelo menos um exemplo de uso prático de alguma dessas operações.

15) Ainda sobre limiarização, a determinação do valor de limiar (ou limiares) pode ser feita de diferentes formas. Existe a maneira baseada em inspeção “visual” feita pelo especialista usuário do sistema, mas também há estratégias automáticas de determinação do limiar, tais como o método isodata e o método de Otsu.

Recomendação bibliográfica: Gonzales-Barron e Butler [7]

A partir do resultado fornecido para a questão anterior (canais RGB obtidos), escreva/implemente o script Python que converte a imagem para tons de cinza. Sobre essa matriz em tons de cinza, determine e aplique um limiar (*thresholding*) via algoritmo Isodata (conforme discutido em sala). Exiba a imagem de entrada e a de saída.

Bibliografia geral:

1. Material no site do Moodle, anotações de aula;
2. Representação (Cap 1) In: Solomon, C. e Breckon, T. Fundamentos de Processamento Digital de Imagens. LTC. 2013. Livro eletrônico (e-book) disponibilizado pela biblioteca do CCT via Uniform Resource Locator (URL): <https://ebookcentral.proquest.com/lib/bibliotecaudesc/reader.ac>
3. Formação (Cap 2, **seção 2.3 em diante**) In: Solomon, C. e Breckon, T. Fundamentos de Processamento Digital de Imagens. LTC. 2013. Livro eletrônico (e-book) disponibilizado pela biblioteca do CCT via Uniform Resource Locator (URL): <https://ebookcentral.proquest.com/lib/bibliotecaudesc/reader.ac>
4. Pixels (Cap 3 **até a seção 3.2**) In: Solomon, C. e Breckon, T. Fundamentos de Processamento Digital de Imagens. LTC. 2013. Livro eletrônico (e-book) disponibilizado pela biblioteca do CCT via Uniform Resource Locator (URL): <https://ebookcentral.proquest.com/lib/bibliotecaudesc/reader.ac>
5. Helio Pedrini, H e Schwartz W. R.. Análise de Imagens Digitais - Princípios, Algoritmos e Aplicações. Ed. THOMSON. 2008;
6. Gonzalez, R. e Woods, R. Processamento Digital de Imagens. Ed. Edgar Blucher. 2000.
7. Ursula Gonzales-Barron, Francis Butler, A comparison of seven thresholding techniques with the k-means clustering algorithm for measurement of bread-crumbs features by digital image analysis, Journal of Food Engineering, Volume 74, Issue 2, 2006, Pages 268-278, (<https://www.sciencedirect.com/science/article/pii/S0260877405001652>)

Há vários links exibidos no pdf da aula de fundamentos, os quais são úteis para essa lista.