

UNIDADE 2 - Formação da imagem, captura e representação

2.1 Visão humana e dispositivos de captura

A visão é talvez o sentido mais complexo do ser humano e que envolve diversos processos que ocorrem não apenas no olho mas também em diversas regiões do cérebro e nos elementos que ligam o olho ao cérebro. De modo bastante simplificado e aproximado grosseiramente, podemos dizer que os dispositivos de captura de imagens, em um sistema de visão computacional, fazem o papel dos olhos, responsáveis por transformar a luz que permeia o ambiente que se observa, e que carrega informações sobre cenas e objetos em três dimensões, em uma imagem de duas dimensões. Essa imagem é depois transportada para o cérebro, que nos computadores, seria representado pelo software ou programa, e que é responsável por extrair desta imagem informações relevantes para a tarefa que precisa ser realizada. Um vídeo explicando o processo de formação da imagem e a relação entre o olho humano e uma câmera digital básica está disponível no link associado à Figura 10.

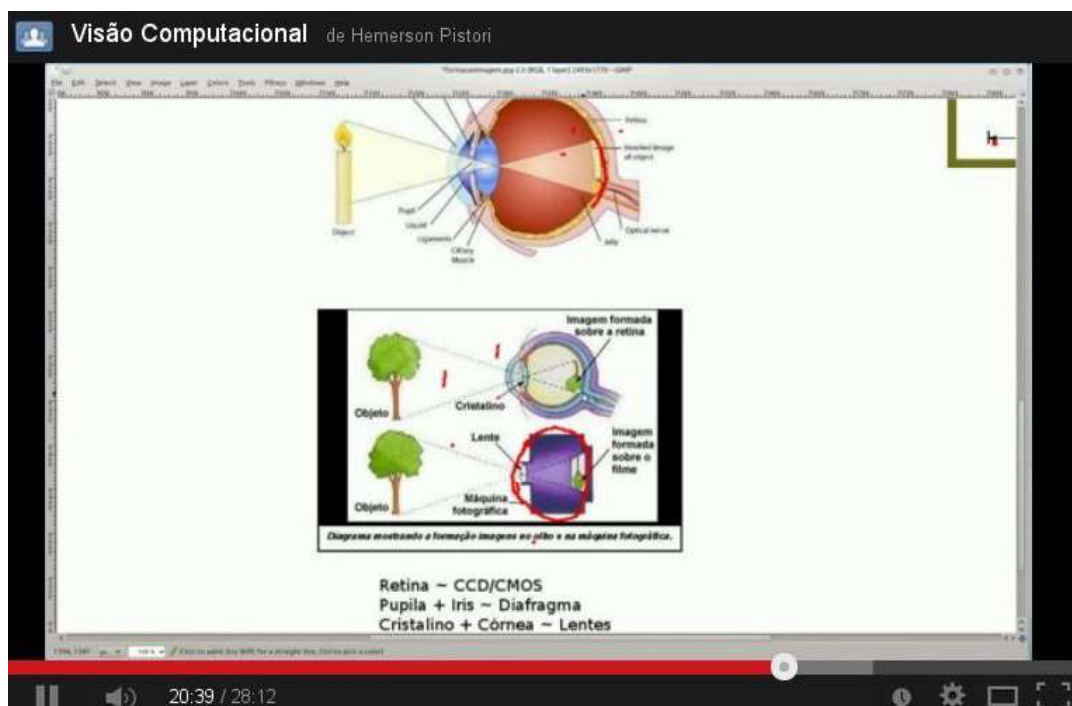
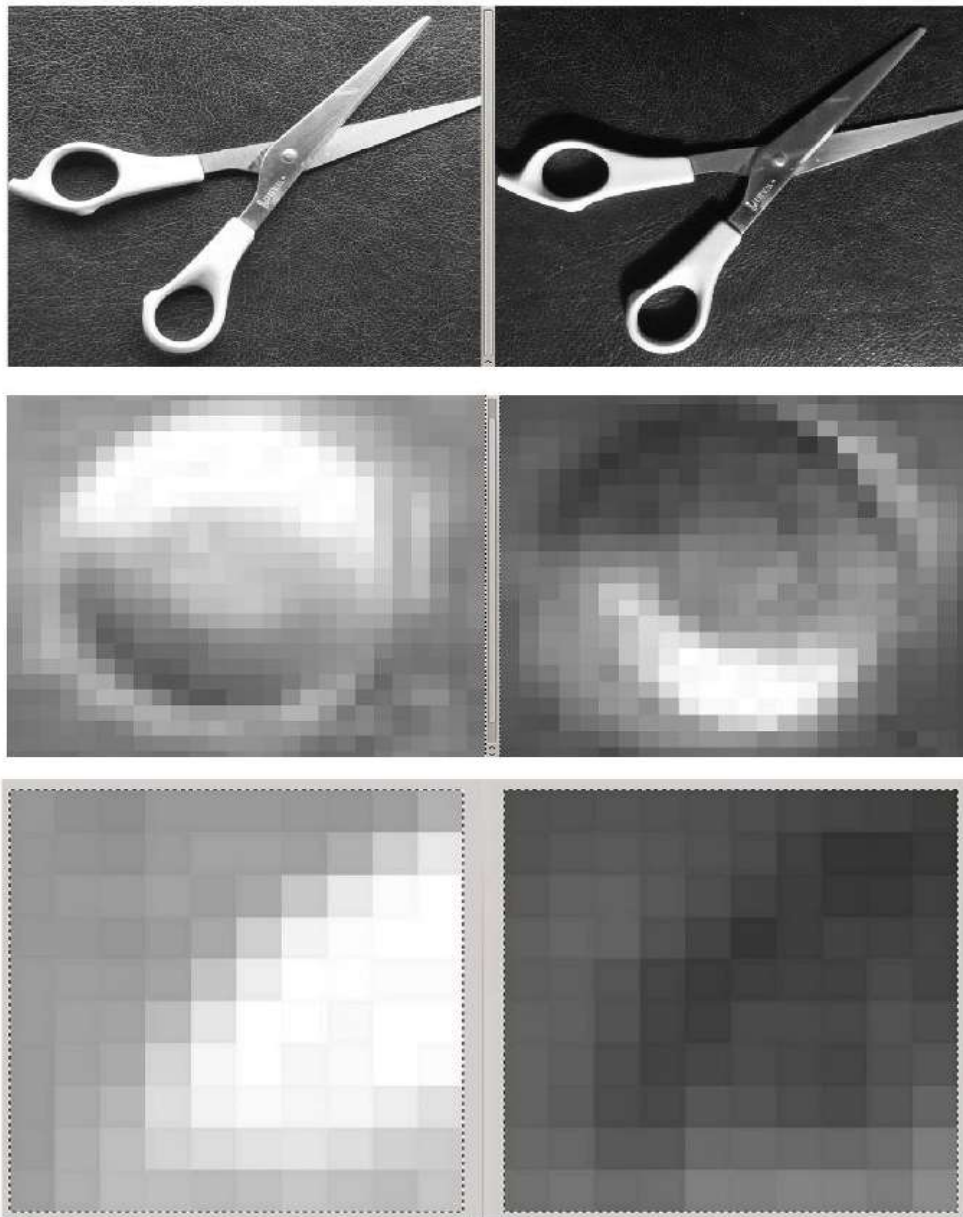


Figura 10: Vídeo sobre formação da imagem disponível em <http://youtu.be/1b781K1Cghg>

Os dispositivos de captura de imagens disponíveis para aplicação de visão computacional não se restringem às já bastante conhecidas câmeras digitais, que aparecem hoje não apenas como um equipamento separado, mas como parte de celulares, *smartphones*, *tablets* e microcomputadores. Temos dispositivos que vão muito além, em termos de captura de imagens, que a visão humana, como microscópios digitais, aparelhos de ressonância magnética, sonares, tomógrafos, aparelhos de ultrassom, “câmeras” de 180 graus, câmeras multi-espectrais ligadas a satélites e aviões não tribulados, câmeras de infra-vermelho e câmeras térmicas. Todos esses aparelhos geram imagens passíveis de serem analisadas por sistemas de visão computacional visando a automação de alguma tarefa e a escolha do equipamento correto, quando possível, é uma tarefa bastante importante. Além da escolha do dispositivo de captura apropriado para cada tarefa, existem outros aspectos importantes que precisam ser considerados e que podem influenciar no resultado, como o tipo de iluminação utilizado e o posicionamento dos equipamentos, objetos que serão filmados ou fotografados e da fonte de iluminação. A questão da iluminação é especialmente crítica pois como o ser humano possui no seu sistema de visão estratégias poderosas e inconscientes para fazer com que sua capacidade de interpretação das imagens não seja influenciada por mudanças no padrão de iluminação ele tende a pensar que isso não é um problema para um computador. Embora já estejamos avançando bastante na construção de softwares que possam manter seu desempenho em situações diversas de iluminação, posicionamento e tipo de câmera, manter uma iluminação uniforme, sem formação de reflexos e sem interferências (como sombras e trepidações), sempre quando possível, pode facilitar bastante as coisas para a equipe de engenharia ou computação responsável pelo desenvolvimento do sistema de visão computacional. É claro que para algumas aplicações, como por exemplo em um sistema de identificação de animais selvagens por imagens que precisam ser captadas em seu habitat natural, teremos menos controle sobre o ambiente de captura e o software deverá ser mais sofisticado, muitas vezes dependendo de novas pesquisas para colocar o estado da arte em visão computacional em novo patamar.



152	144	142	154	154	154	152	152	159	178
151	149	147	157	164	163	161	176	208	233
152	155	156	164	159	169	200	234	252	255
153	153	150	157	188	207	244	253	255	255
157	159	158	165	197	238	265	264	252	252
157	162	164	188	230	255	254	249	255	255
156	162	171	213	247	255	247	250	255	252
158	170	183	222	243	250	246	247	239	225
157	175	191	212	220	227	229	222	210	206
161	176	181	192	188	193	200	194	189	200

84	84	85	83	81	79	73	69	67	62
85	89	91	87	82	76	64	54	52	58
90	97	98	89	79	67	62	56	58	65
96	102	98	82	68	54	62	66	66	69
94	95	85	70	58	63	67	65	75	68
98	96	83	68	63	74	74	72	82	78
99	94	78	68	71	76	73	69	81	87
99	93	77	73	86	84	79	75	87	102
105	103	88	89	106	107	105	103	109	125
113	118	108	111	128	127	132	131	124	140

Figura 11: Na primeira linha temos imagens de uma mesma tesoura obtidas de dia, com luz natural (à esquerda) e a noite com luz artificial (à direita). Na segunda linha ampliamos o ponto que prende as duas lâminas da tesoura nas duas situações de iluminação, na terceira linha ampliamos uma pequena região na parte superior esquerda das imagens da linha 2 e na quarta temos os valores dos pixels

O que chega para ser processado por um sistema de visão computacional a partir do dispositivo de captura de imagens é simplesmente uma ou mais matrizes ou planilhas de números, algo parecido com o que aparece na última linha de imagens da Figura 11. Esse conjunto de números pode ser completamente alterado simplesmente pela troca de uma lâmpada incandescente por uma fluorescente ou por um sistema de iluminação por leds ou então por luz natural. Também influencia muito este conjunto de números a utilização do flash da máquina; o ângulo e a distância da câmera em relação ao objeto e até mesmo os outros objetos que não são de interesse para a solução do problema mas que acabam influenciando no padrão de iluminação que chega até a câmera. A Figura 11 mostra em sua primeira linha a imagem de um mesmo objeto, uma tesoura, obtida com iluminação natural e artificial (lâmpada fria), juntamente com uma pequena parte da matriz de números que serve de entrada para o sistema computacional nos dois casos (quarta linha). Na segunda linha temos uma ampliação do ponto que prende as duas lâminas da tesoura nas duas situações de iluminação e na terceira ampliamos uma pequena região na parte superior esquerda das imagens da linha 2. É possível notar que praticamente nenhum número se repete, mesmo sendo uma imagem do mesmo objeto que pode ser facilmente reconhecido por um ser humano. Cada um desses números é o que convencionamos chamar de pixel. A imagem está em tons de cinza, pois como será detalhado na Unidade 4, imagens coloridas são registradas em um computador através de 3 matrizes e não apenas uma.

Por fim, outro aspecto que pode influenciar a imagem que chega até o sistema de visão computacional é o tipo de arquivo digital utilizado. Para economizar espaço na memória das máquinas, é comum que as mesmas sejam armazenadas em formatos que chamamos de “compactados com perda de informação”, como geralmente ocorre nos arquivos com extensão JPG⁶, por exemplo. O ser humano praticamente não nota que houve perda de informação pois esses formatos foram desenvolvidos justamente considerando o funcionamento do sistema visual humano e buscam minimizar a perda de informação relevante para a interpretação da imagem por um ser humano. No entanto, como o sistema de visão computacional não funciona necessariamente como o do ser humano, informações importantes

⁶ É possível controlar o nível de compactação no formato JPG, quanto maior a compactação, menor o arquivo resultante, mas maior também a perda de informação.

para a máquina podem ser perdidas e por isso, sempre que possível ou até se ter certeza, através de experimentos, que o formato compactado não está prejudicando o sistema, é bom utilizar formatos de arquivos que não descartam informações, como por exemplo, o formato RAW e TIFF.



Figura 12: Utilização do tabuleiro de xadrez para calibração

2.2 Calibração

Existem diversas aplicações de visão computacional que exigem que o sistema seja capaz de realizar medições, como comprimentos, áreas e distâncias, tanto como etapas intermediárias de algum processamento quanto para relatar ao usuários estas medições. Acontece que no processo de captura e formação da imagem digital acontecem fenômenos que fazem com que o mapeamento das dimensões e distâncias dos objetos presentes na imagem digital, para os mesmos objetos no ambiente real, não seja trivial. Além da própria projeção em perspectiva, que faz, por exemplo, que objetos que estejam mais distantes da câmera apareçam em tamanho menor na imagem, temos distorções relacionadas com as lentes utilizadas ou com os sensores das câmeras, entre outros. Um exemplo extremo são imagens geradas por lentes grande angular ou olho de peixe que causam efeitos similares ao apresentado na Figura 12 à direita. Na etapa de calibração o sistema, geralmente baseando-se na imagem de algum objeto de dimensões conhecidas e facilmente identificável, como uma tabuleiro de xadrez, determina as transformações geométricas necessárias para recuperar as distâncias e dimensões originais dos objetos na cena. Também são utilizadas informações técnicas sobre os sensores da câmera digital, lentes e outros elementos que afetam a geração da imagem digital. Desde que não ocorram depois grandes mudanças no tipo e posicionamento da

câmera utilizada, este procedimento não precisa ser repetido para cada nova imagem a ser processada. É fundamental, no entanto, tomar muito cuidado com a questão da calibração pois se isto não for bem feito os resultados do sistema podem ser invalidados. Existem softwares que te dão a opção de imprimir em uma folha em branco o desenho (E.g.: uma parte de um tabuleiro de xadrez) que deverá ser colocado no ambiente em que o sistema irá operar para depois de detectá-lo realizar a calibração.

2.3 Ruído

Ruídos são elementos ou artefatos presentes em uma imagem digital que não contribuem e muitas vezes atrapalham o sistema de visão computacional na análise daquilo que é relevante para a solução de um problema. Eles podem ser introduzidos tanto no próprio ambiente real que está sendo filmado ou fotografado, como por exemplo algum tipo de sujeira que por falta de cuidado contaminou uma amostra microscópica que seria analisada, ou por falhas e deficiências nos equipamentos e acessórios usados na captura da imagem. Lentes podem ter pequenos riscos, sensores podem falhar e mesmo sem falhas, pela própria maneira com que são fabricados, introduzem na imagem final efeitos que não estão relacionados com a imagem original. O tipo de iluminação utilizado também pode gerar ruídos que para a visão humana são imperceptíveis mas que alteram significativamente os valores dos pixels das imagens originais. Reflexos e sombras, dependendo do problema, também podem ser considerados como ruídos indesejáveis, em outros casos, podem ser informação importante. Sombras, por exemplo, podem ser úteis para estimar a altura de objetos quando temos informações sobre a posição da fonte de iluminação e vários objetos na imagem. A transmissão de dados e a compactação das imagens com perda de informação também podem ser fontes de ruído.

Na Figura 13 temos uma folha de soja de um experimento que precisa medir a área que foi comida por uma lagarta. É possível ver vários pontinhos pretos nas regiões da folha que foram comidas. Esses pontos são fezes da lagarta e para o problema de medir a área podem ser considerados como ruído.



Figura 13: Folha de soja após herbivoria

Agora, se o problema fosse determinar a quantidade de fezes produzida pela lagarta, já não seriam considerados ruídos. Para simular o efeito do ruído provocado por determinados equipamentos, existem softwares que permitem que introduzamos de forma controlada diferentes tipos de ruído. Assim, é possível realizar experimentos controlados com diferentes técnicas de redução de ruído para identificar a melhor forma de lidar com o problema. A Figura 14 mostra a imagem da folha de soja após a inserção de alguns tipos de ruído. Na próxima unidade iremos tratar de técnicas que podem ser utilizadas para reduzir o efeito de ruídos em imagens.



Figura 14: Folha de soja após a aplicação de diferentes técnicas para simular a presença de ruído