

UNIDADE 4 - Segmentação

Na etapa de segmentação, busca-se um particionamento da imagem em regiões de forma a separar elementos de interesse, para o problema a ser resolvido, de elementos que são irrelevantes para o problema. Em alguns casos, em problemas que envolvem contagem ou reconhecimento de múltiplos objetos, a segmentação, além de separar os elementos irrelevantes, separa os objetos de interesse em regiões distintas. Na Figura 22 temos algumas imagens de entrada para a segmentação na parte superior, e o resultado esperado da segmentação na parte inferior. Exemplos de segmentos gerados por um possível segmentador estão delimitados por uma linha vermelha e as regiões irrelevantes em cada caso foram pintadas de branco. Quando não existe região irrelevante e a segmentação deve apenas agrupar áreas similares marcamos cada região de um mesmo grupo com a mesma cor (terceira imagem). O número de segmentos e o que precisa ser segmentado é altamente dependente do problema e para uma mesma imagem podem existir várias alternativas. No caso da imagem do bebê, por exemplo, se o interesse fosse em identificar a expressão facial o segmentador teria que manter apenas a face. Se fosse para identificar a direção do olhar, apenas os olhos e para analisar todos os tipos de objetos na cena, teríamos que manter a banheira e a cadeira em segmentos próprios. Na imagem dos elefantes o interesse poderia ser na vegetação e não nos elefantes, ou então apenas nos filhotes de elefantes.



Figura 22: Exemplos de segmentação

Nas próximas seções trataremos de alguns grupos importantes de técnicas que podem ser aplicadas para segmentar imagens.

4.1 Segmentação por limiarização

Através da análise de histogramas de tons de cinza ou de cores é possível definir um ou mais limiares para agrupar os pixels de uma imagem. Por exemplo, no histograma de tons de cinza, poderíamos definir o tom de número 127 como limiar e fazer com que todos os pixels das imagem com tons entre 0 e 127 pertençam a um grupo ou segmento enquanto os pixels com tons entre 127 e 255 pertençam a um outro grupo ou segmento. Nesse caso, a imagem seria segmentada em dois grupos, não necessariamente contíguos, um contendo as regiões mais claras e o outro as mais escuras da imagem. Usando mais de um limiar podemos dividir os pixels em mais de dois grupos. Os limiares podem ser definidos manualmente, experimentando-se diferentes valores e observando-se o resultado, ou automaticamente através das diversas técnicas disponíveis, incluindo algumas baseadas em aprendizagem automática a partir de amostras do tipo de segmento que se espera obter. Os vídeos referenciados nas Figura 9 e 16 mostram exemplos de segmentação por limiarização. Em alguns problemas envolvendo segmentação em dois grupos, como é o caso, por exemplo, de problemas onde temos um objeto que se destaca em relação ao fundo da imagem e queremos simplesmente separar o fundo e o objeto de interesse, é comum termos como resultado da segmentação uma imagem chamada binária, com apenas dois tons, um representando o fundo (geralmente com valor 0) e o outro representando o objeto ou a frente (geralmente com valor 255). Esse é um caso especial de segmentação chamado de binarização.

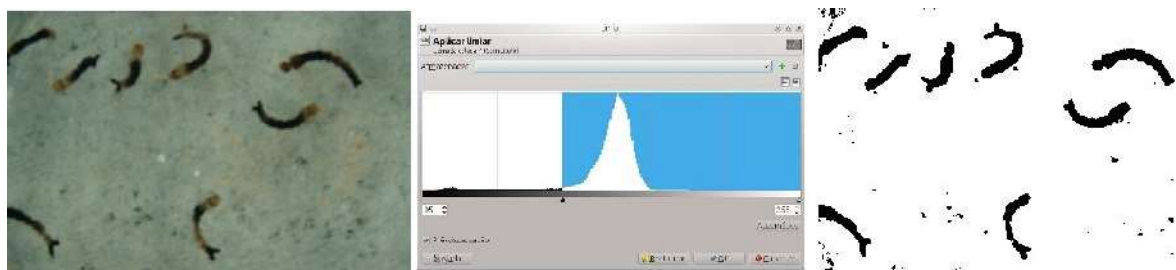


Figura 23: Aplicação de binarização em uma imagem com larvas do *Aedes aegypti*

Na Figura 23 mostramos uma imagem de larvas do *Aedes aegypti* (à esquerda), o seu histograma de tons de cinza com o limiar de 95 escolhido para

binarizar a imagem (ao centro) e imagem binarizada resultante (à direita). Algumas medições, como área e diâmetro por exemplo, podem ser obtidas diretamente das imagens binarizadas através de um módulo que pode ser chamado de analisador de blobs ou de partículas⁷ (que realiza diversos tipos de medições das “manchas pretas” resultantes da binarização). Também é possível contar o total de blobs (manchas), operação bastante útil em diversos tipos de problemas envolvendo imagens microscópicas, por exemplo.

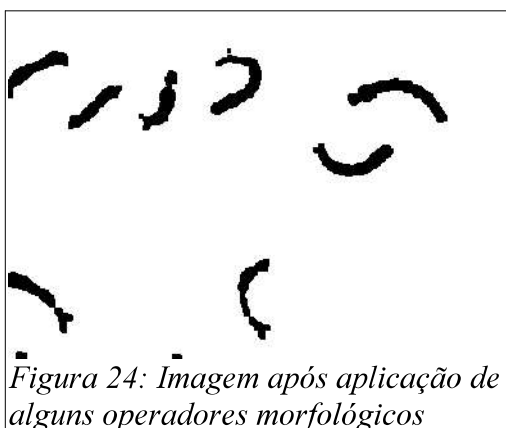


Figura 24: Imagem após aplicação de alguns operadores morfológicos

É possível perceber que o resultado da segmentação por limiarização/binarização na Figura 23 ainda não é perfeito, se considerarmos que o problema é, por exemplo, contar o total de larvas e medir os seus tamanhos. No entanto, utilizando operadores morfológicos sobre a imagem binarizada é possível melhorá-la para reduzir os ruídos representados pelas pequenas manchas pretas

que surgem na imagem devido ao fundo do recipiente contendo as larvas, que é um fundo marmorizado (na verdade, o recipiente é transparente e está sob uma bancada marmorizada). A Figura 24 mostra a imagem binarizada após a aplicação do operador morfológico de fechamento. A Figura 25 mostra uma das telas de uma sequência de três vídeos que explicam como implementar uma técnica de limiarização que serve para segmentar pixels de uma imagem que correspondem à pele humana.



Figura 25: Implementação de limiarização utilizando OpenCV. Vídeo disponível em <http://youtu.be/ctKa5ppxAK0>

⁷ Não confundir com o **filtro** de partículas apresentado na unidade sobre Rastreamento

4.2 Segmentação por agrupamento

Existe uma sub-área da aprendizagem de máquina chamada de aprendizagem não-supervisionada e o problema de segmentação pode ser facilmente formulado como um problema de aprendizagem não-supervisionada⁸. Com isso, é possível utilizar as técnicas de aprendizagem não-supervisionada, chamadas também de técnicas de aglomeração (**clustering**) ou agrupamento para segmentação em imagens. As técnicas de aglomeração tem como objetivo agrupar elementos em uma quantidade geralmente pré-definida de grupos utilizando alguma medida de similaridade. Os elementos aqui considerados podem variar desde a simples tonalidade de cada pixel da imagem ou cor do pixel até propriedades mais complexas que podem ser extraídas utilizando as técnicas de extração de atributos tratadas na Unidade 5 deste texto. No caso das tonalidades, que são simples valores numéricos, a medida de similaridade poderia ser o valor absoluto da diferença entre duas tonalidades. Já no caso das cores, cada pixel é representado por três números, e a simples diferença aritmética não pode ser diretamente aplicada como medida de similaridade.

Quando mais de um número é utilizado como elemento, o mais comum é representar esses elementos como vetores⁹ e utilizar uma das diversas opções de medidas de similaridade entre vetores, como as distâncias Euclidiana, de Manhattan, de

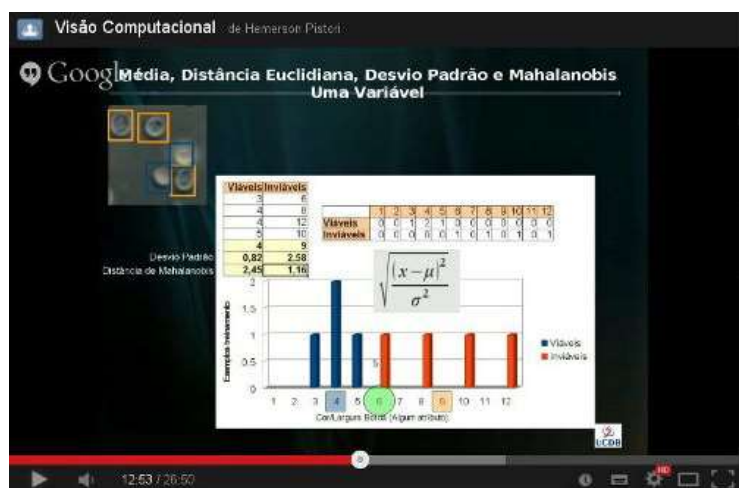


Figura 26: Vídeo sobre distância de Mahalanobis disponível em <http://www.youtube.com/watch?v=0tMw0XEfrw&feature=share>

⁸ Aprendizagem não-supervisionada é também muito estudada na área da computação chamada de mineração de dados (*data mining*)

⁹ Vetor é um conceito da Álgebra e da Física e de maneira muito informal e imprecisa pode ser visto como um sequência de tamanho pré-definido de números.

Mahalanobis (ver Figura 26), de Chebychev, dos Cossenos e da Escavadeira entre muitas outras.

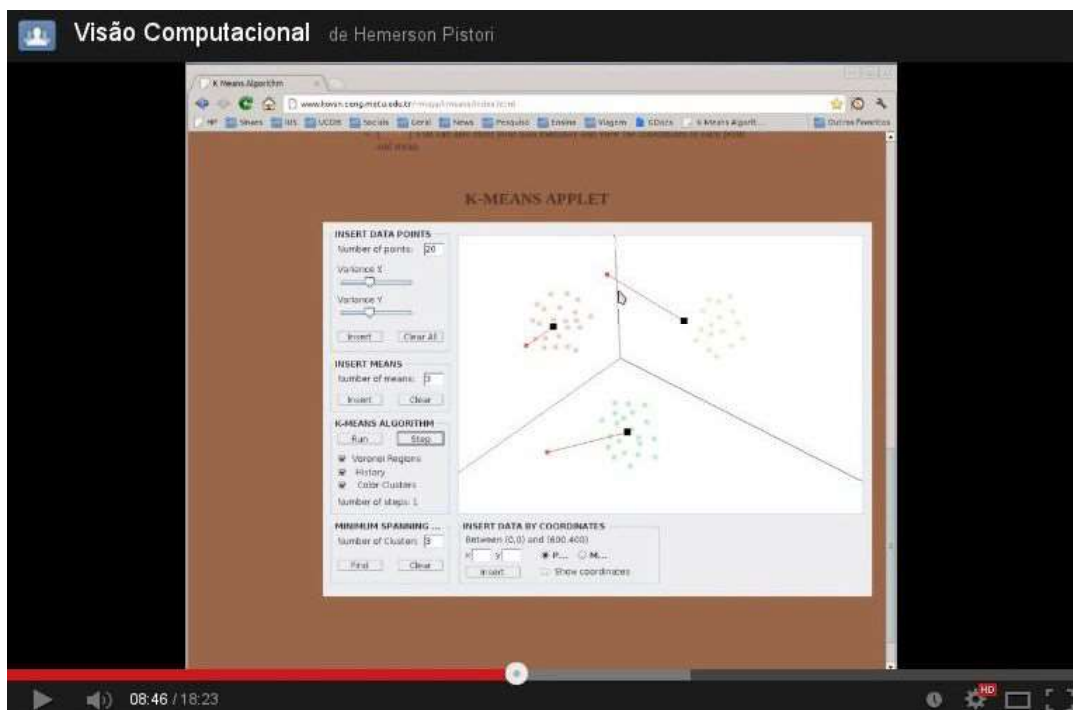


Figura 27: Vídeo sobre o k-médias disponível em <http://youtu.be/cxSLmrT9zk>

Tanto a determinação do número de grupos quanto da medida de similaridade, além dos próprios elementos que serão

considerados nas comparações, são dependentes do problema e na maioria das vezes precisam ser determinados através de experimentos ou de uma combinação entre experimentos e conhecimento prévio sobre o problema e a teoria que fundamenta cada uma das possíveis opções disponíveis. Existem muitas técnicas de aglomeração disponíveis e o vídeo referenciado na

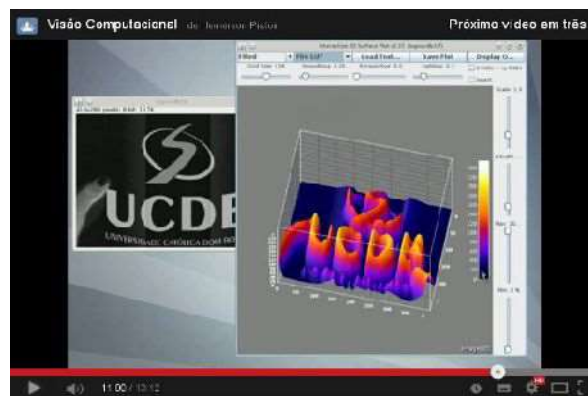


Figura 28: Visualizando uma imagem como sendo uma superfície. Vídeo disponível aqui: http://youtu.be/Xkiv7Lt_eXQ

Figura 27 explica uma das mais famosas chamada de k-médias (do inglês *k-means*)

Uma outra técnica muito utilizada para realizar segmentação por aglomeração é a dos divisores de águas (*watershed*). Para entender essa técnica é preciso primeiro compreender que uma imagem em tons de cinza pode também ser

modelada como um função de duas dimensões, que por sua vez, pode ser visualizada como uma superfície topográfica (como um cordilheira de montanhas, por exemplo). O vídeo indicado na Figura 28 ilustra essa ideia utilizando uma ferramenta do ImageJ.

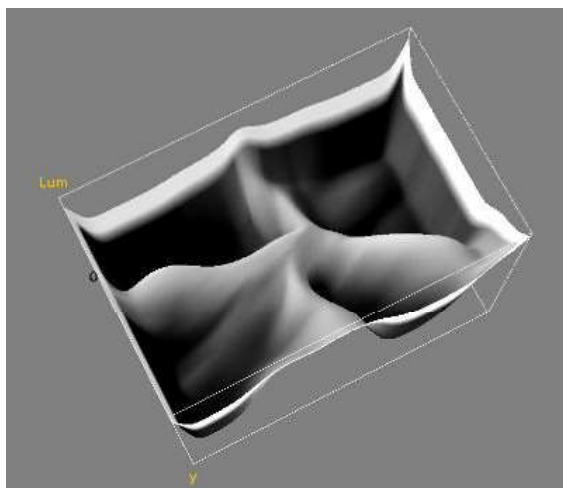


Figura 29: Ilustração de uma imagem visualizada como uma superfície topográfica

A técnica de divisores de água chega até os contornos dos segmentos de uma determinada imagem em um processo que simula uma inundação em uma região com morros e vales dentro de uma caixa, algo como a superfície mostrada na Figura 29.

Se fizermos alguns furos na parte de baixo dessa caixa e começarmos a bombear água através desses furos ou então colocarmos essa caixa na chuva para que comece a se encher d'água aos poucos, as regiões mais baixas encherão primeiro e as

mais altas por último e alguns lagos serão formados durante o processo. Se pararmos de colocar água no exato momento em que dois lagos estão prestes a transbordar e se unir em um único lago maior, teremos uma faixa ou linha de “terra” separando esses dois lagos. É justamente essa linha que chamamos de divisor de águas. Dependendo da topologia poderemos ter vários lagos com vários divisores de água e é isto que o algoritmo de segmentação simula. Os segmentos resultantes serão os lagos e em muitas aplicações estes lagos irão corresponder a contornos de objetos de interesse. Um dos problemas com este método é determinar quantos furos devem ser feitos, em quais locais e o quanto de água devemos bombear para dentro da caixa. Existem diversas propostas para determinar automaticamente valores para esses parâmetros e uma das implementações mais conhecidas é baseada em morfologia matemática.

4.3 Segmentação por ajuste de modelos

Na segmentação por ajuste de modelos partimos de algum tipo de descrição de entes geométricos, como linhas, círculos e elipses e busca-se encontrar nas imagens contornos de objetos que se enquadrem nessas descrições. Um grupo de técnicas bastante popular para este tipo de segmentação são as chamadas *snakes* que trabalham com modelos deformáveis. São também chamadas de técnicas de contorno ativo. A ideia do modelo deformável, e que recebe o apelido de cobra (*snake*) justamente por isso, é partir de uma curva desenhada dentro ou fora do objeto a ser segmentado, e ir aos poucos “contorcendo” ou deformando essa curva até que ela se ajuste perfeitamente ao contorno do objeto que precisa ser segmentado. A Figura 30 mostra uma sequência de imagens com um modelo inicial (curva em verde) sendo deformado (curva em vermelho) por uma técnica do tipo *snake* até conseguir contornar a lâmina de uma faca, que nesse caso, era o objeto a ser segmentado. Neste exemplo, temos uma segmentação semi-automática, pois a curva inicial foi desenhada pelo usuário, no entanto, em alguns casos, é possível automatizar completamente o processo usando informações aproximadas sobre a posição do objeto alvo ou alguma outra técnica para gerar a curva inicial a ser deformada.

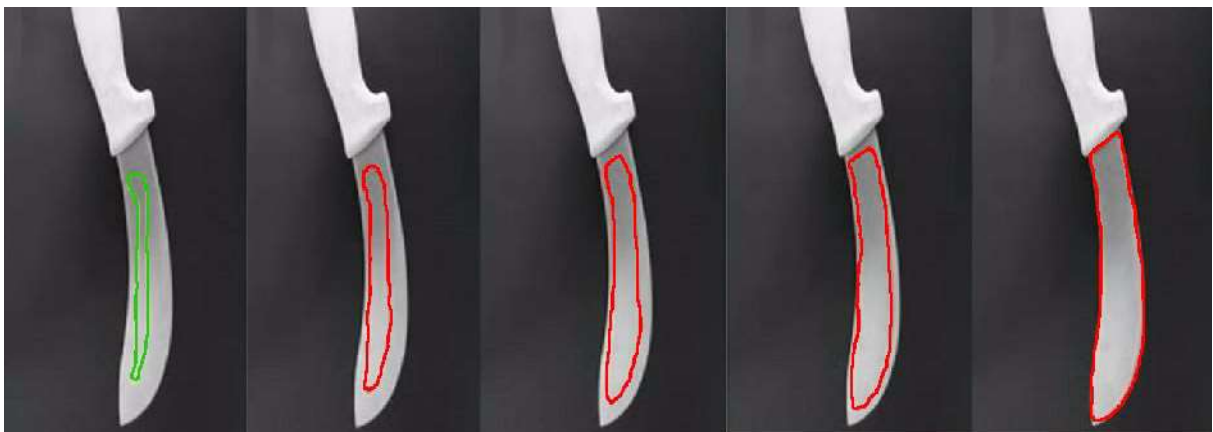


Figura 30: Exemplo de aplicação da técnica de segmentação por contorno ativo

Outra técnica bastante conhecida, principalmente para segmentação de linhas, círculos e elipses, é a transformada de Hough. Essa técnica é aplicada em imagens binárias, quando temos apenas 2 tipos de pixels, os que pertencem ao fundo (preto) e os que pertencem a objetos (branco). A transformada de Hough funciona

através de um processo chamado de votação. Para cada pixel branco da imagem binarizada registra-se um voto para todas as retas, círculos, elipses ou qualquer figura geométrica sendo buscada, que poderiam ter este pixel branco em seu contorno. As figuras geométricas que receberem mais votos serão utilizadas para segmentar a imagem (o que neste caso significa definir os locais onde se encontram retas, círculos, etc). Uma grande vantagem desta técnica é que ela pode encontrar figuras geométricas mesmo em imagens com muito ruído ou com as figuras parcialmente cobertas.

Na Figura 31 temos uma imagem colorida de uma face em que se pretende detectar a região da íris. Ao lado direito, em cima, temos a imagem binarizada, necessária para aplicação da transformada de Hough. Na parte de baixo, à esquerda, temos uma representação do resultado da votação, com os pontos mais claros representando as regiões mais prováveis de se encontrar um círculo centrado nesses pontos (quanto mais claro, mais votos). É possível verificar nessa figura que existem alguns pontos bem claros (bem votados) no centro da íris, mas também temos alguns outros pontos em outras regiões que na imagem binarizada também apresentam formas aproximadamente circulares. A última imagem, à direita, mostra em amarelo os 3 círculos mais votados, e que poderiam servir para alimentar um outro módulo do sistema. Esse outro módulo, usando informações envolvendo cores por exemplo, poderia descartar os dois círculos que não correspondem à íris. Vale notar que os dois círculos falsos foram gerados justamente por um reflexo forte no canto da lente do óculos, que poderia ser evitado com um melhor

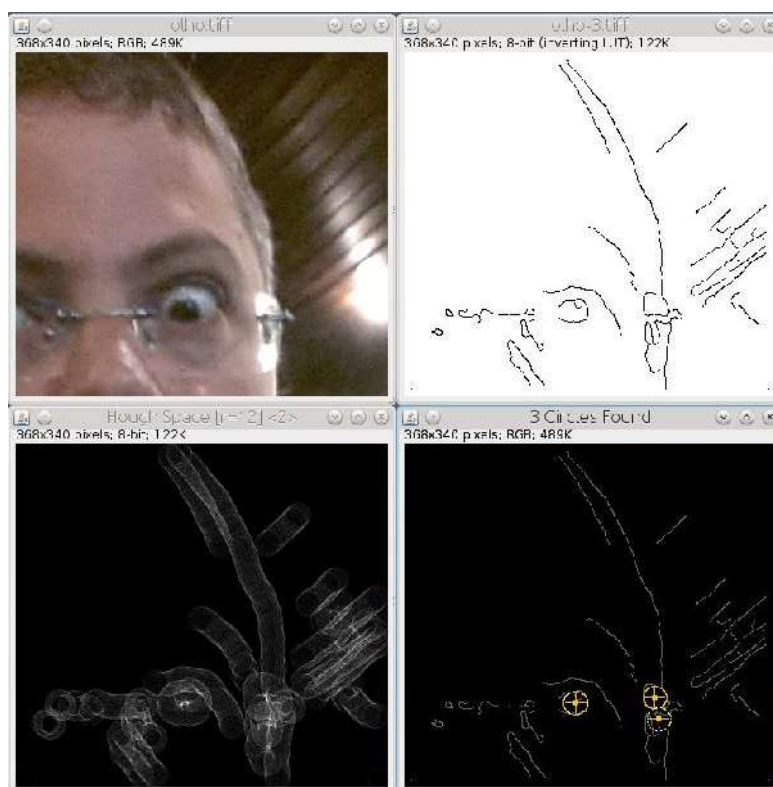


Figura 31: Exemplo de aplicação da transformada de Hough

posicionamento da iluminação ou da câmera, ou ainda com técnicas específicas de redução de ruído.