

Processamento Digital de Imagens: experimentos com algoritmos de segmentação

Borges, C. D. B.

Resumo A automatização de tarefas com auxílio de processamento de imagens depende fortemente da capacidade de algoritmos para extrair informações úteis do emaranhado de *pixels* que se apresenta. Um dos sub-processos mais relevantes durante a extração dessas informações é a segmentação, procedimento que visa a separar os *pixels* da imagem em regiões disjuntas, geralmente de acordo com algum critério de similaridade. O presente trabalho faz uma breve análise acerca da performance de quatro técnicas clássicas de segmentação: Crescimento de Regiões, *Watershed*, *K-means* e Limiarização de Otsu.

1 Introdução

A segmentação de imagens é uma técnica amplamente utilizada para o propósito de particionar uma imagem em diversos conjuntos disjuntos de *pixels*. O critério utilizado para esse particionamento varia de acordo com a aplicação. No entanto, pode-se citar a separação de *pixels* pertencentes a diferentes objetos como uma das aplicações de maior necessidade e relevância. O presente trabalho avalia o desempenho de quatro técnicas de segmentação: Crescimento de Regiões, *Watershed*, *K-means* e Limiarização de Otsu. O restante desta seção é dedicado a uma breve apresentação de cada uma das técnicas. As ferramentas de software utilizadas e os métodos de pré-processamento aplicados são comentados na seção 2. Nas seções 3 a 8 são apresentados os resultados e observações feitas acerca das segmentações obtidas. Uma discussão geral é realizada na seção 9 e as conclusões são apresentadas na seção 10.

1.1 Crescimento de Regiões

A segmentação por crescimento de regiões baseia-se na expansão de conjuntos a partir de *pixels* denominados

sementes. Cada segmento da imagem inicia com apenas um *pixel* semente. A cada iteração do algoritmo, os *pixels* vizinhos àqueles já pertencentes ao conjunto são avaliados e é feita uma decisão sobre sua inclusão no conjunto. O quesito usado para inclusão pode ser assumir diferentes formas, sendo a condição de limiar uma das mais simples e amplamente usadas. O algoritmo termina sua execução quando os *pixels* da borda do conjunto já não possuem nenhum vizinho que possa ser incluído de acordo com a condição definida.

Os parâmetros de entrada para execução da segmentação por regiões são o conjunto de posições de *pixels*-semente e o limiar a ser utilizado no critério de parada. A escolha desses valores pode ser feita manualmente ou automaticamente, através de análises estatísticas sobre a imagem em questão. Neste trabalho, no entanto, esses valores foram escolhidos manualmente.

1.2 *Watershed*

A técnica de *Watershed* baseia-se no tratamento da imagem em escala de cinza como uma superfície topográfica tridimensional contendo vales, planícies e montes. A variação mais clássica desse método é o *Watershed* por inundação, técnica que consiste em, metaforicamente, inundar os vales da superfície da imagem até que as “águas” oriundas de diferentes vales entrem em contato, delimitando assim, de maneira simplificada, as bordas das regiões.

Para que o *Watershed* seja executado, é importante definir uma etapa de pré-segmentação, que modela a imagem original como uma superfície topográfica em escala de cinza, cujos vales localizem-se sobre os objetos de interesse. A partir da representação topográfica da imagem, pode-se calcular os centros dos vales como os pontos de mínimos locais da superfície. O processo de inundação será iniciado sobre esses pontos e executado de maneira paralela.

1.3 *K-means*

K-means é uma técnica desenvolvida originalmente para agrupamento ou clusterização de vetores. O método consiste no reposicionamento iterativo de um conjunto de K centroides no espaço dimensional dos vetores a serem agrupados. Na utilização do *K-means* para segmentação de imagens, cada *pixel* é representado como um vetor que armazena suas informações de cor e posição. Esses vetores são então utilizados como amostras para a execução do algoritmo de agrupamento *K-means*, cujo resultado final é um conjunto de centroides posicionados no espaço vetorial dos *pixels* da imagem, de tal forma que a variância entre os membros de cada agrupamento seja minimizada. O pertencimento de um pixel a um *cluster* é dado através do grupo ao qual pertence o seu centroide mais próximo.

O parâmetro de entrada para execução da segmentação via *K-means* é o valor de K , ou seja, o número desejado de agrupamentos/segmentos. É possível também calcular o valor de K automaticamente, como fazem alguns algoritmos de clusterização adaptativa. Neste trabalho utiliza-se uma versão adaptativa do *K-means*.

1.4 Limiarização de Otsu

A Limiarização de Otsu é uma técnica de segmentação bastante simples, pois consiste apenas em uma limiarização comum com um limiar definido automaticamente através da análise do histograma da imagem. O método de busca por um limiar ótimo é realizado com um escaneamento de todos os valores possíveis e a posterior escolha do valor que minimiza a variância intra-classe resultante. A depender das características da imagem e de seu histograma, uma segmentação mais adequada só poderá ser realizada com dois ou mais limiares, evento para o qual existe uma solução de busca definida pelo método de Limiarização Múltipla de Otsu. Neste trabalho, o número de limiares utilizados foi escolhido de acordo com as características observadas na imagem.

2 Metodologia

O código relativo a este trabalho foi desenvolvido no software MATLAB R2017a. Os *scripts* podem ser encontrados no arquivo compactado anexado junto a este relatório. Para os métodos de Crescimento de Regiões e *K-means* foram utilizados *scripts* disponíveis na plataforma File Exchange da Mathworks [1, 2]. As outras técnicas, *Watershed* e Limiarização de Otsu foram implementadas com funções disponíveis no MATLAB R2017a.

No caso do Crescimento de Regiões, as sementes e limiares foram definidos manualmente de acordo com a imagem. Antes da aplicação do algoritmo, as imagens foram convertidas para escala de cinza e submetidas a uma filtragem mediana com *kernel* 3×3 .

O pré-processamento para aplicação do *Watershed* envolveu uma filtragem por cor no espaço HSV, centralizada na cor do objeto de interesse. Também foram aplicadas as operações morfológicas de abertura seguida de dilatação com elementos estruturantes no formato de disco. Os tamanhos dos elementos estruturantes foram escolhidos de maneira diferente para cada imagem.

Para o *K-means*, o único pré-processamento foi a aplicação de uma filtragem mediana 3×3 . Já no caso da Limiarização de Otsu, a imagem foi convertida para escala de cinza e submetida a uma filtragem mediana.

As quatro técnicas foram aplicadas sobre seis imagens de teste, exibidas na Figura 1, e seus desempenho foram avaliados visualmente. As avaliações feitas, juntamente com as considerações levantadas durante a implementação, são relatadas nas observações feitas durante a exposição dos resultados e nas discussões finais.



Figura 1 As seis imagens de teste utilizadas no experimento.

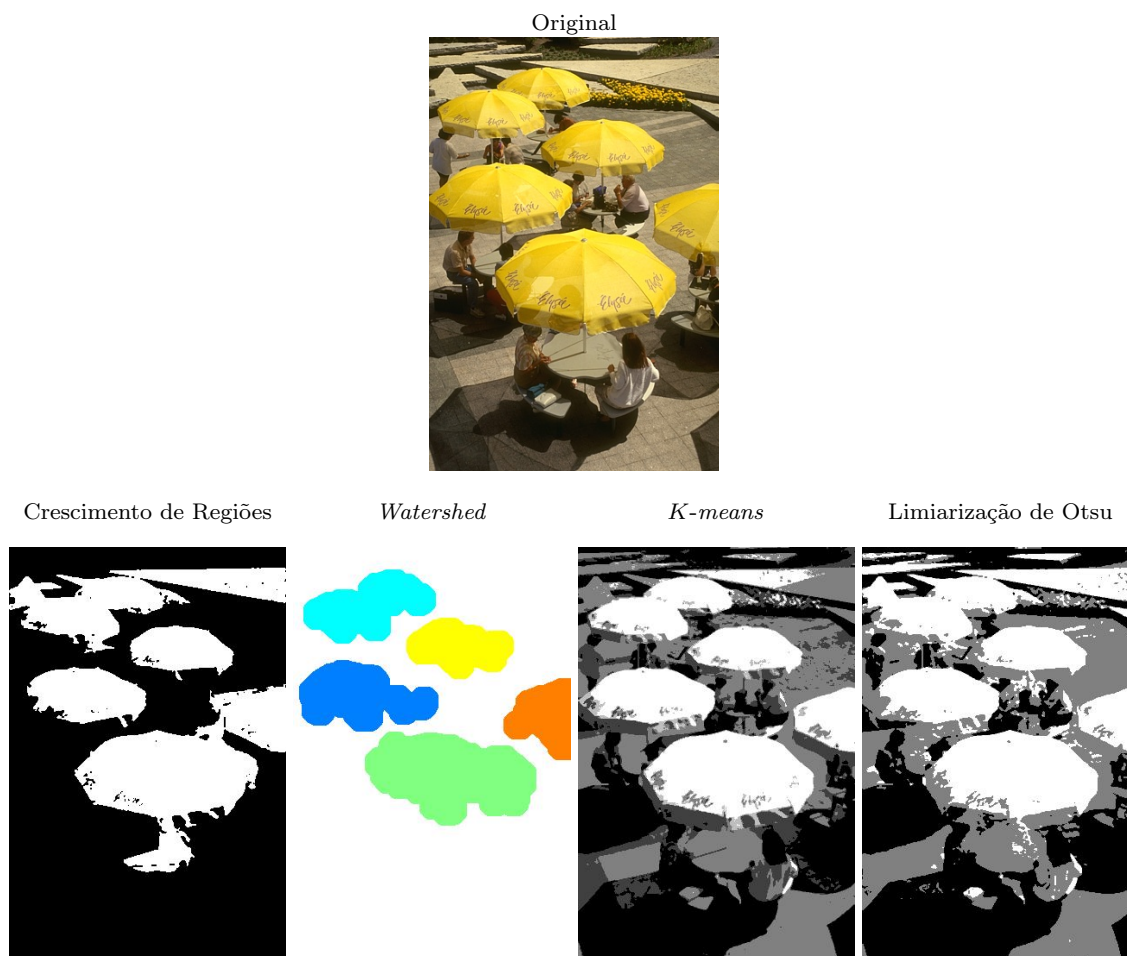


Figura 2 Resultados obtidos para a imagem dos guarda-sóis

3 Resultados da imagem: Os Guarda-Sóis

Os resultados de segmentação obtidos para a imagem dos guarda-sóis são exibidos na Figura 2. Nesta imagem, foi definido o objetivo de segmentar os guarda-sóis do restante dos objetos na imagem.

Na aplicação da segmentação por Crescimento de Regiões, foram selecionadas cinco sementes, uma no centro de cada guarda-sol. O limiar selecionado foi 35. Nota-se que a maior parte dos guarda-sóis foi segmentada corretamente, aparecendo na região clara da segmentação. No entanto, pode ser observado que o crescimento das regiões expandiu-se, em alguns casos, para fora dos objetos desejados. Foi o caso da inclusão errônea de algumas partes da calçada e da mesa. Essa expansão poderia ser controlada pela diminuição do limiar, no entanto, isso também acarretaria na perda de algumas regiões dos guarda-sóis.

A segmentação por *Watershed* mostra resultados interessantes. Usando esse método, houve uma menor inclusão de regiões indesejadas, embora os formatos dos guarda-sóis não tenham sido perfeitamente seguidos.

Além disso, os processos usados no *Watershed* tornam possível a diferenciação dos guarda-sóis sem que haja uma pré-definição de seu número. No caso, foram encontrados cinco guarda-sóis, enquanto na imagem real podem ser vistos seis. Isso ocorreu, pois o algoritmo erradamente uniu os dois guarda-sóis mais distantes.

Na aplicação do *K-means*, o algoritmo adaptativo encontrou cinco *clusters*. Um desses agrupamentos, definido na imagem segmentada pela cor branca, contém os pixels referentes aos guarda-sóis. Pode-se ver que a segmentação dos guarda-sóis ocorre com maior facilidade para suas partes superiores, visto que as regiões de sombra dos guarda-sóis são agrupadas em outra classe pelo algoritmo. Nota-se também que algumas flores, de cor similar à dos guarda-sóis, e uma parte brilhante da calçada são postas no mesmo agrupamento.

A Limiarização de Otsu foi aplicada para busca de dois limiares, portanto, segmentação da imagem em três grupos. Vê-se que os guarda-sóis foram categorizados junto com as regiões mais claras da imagem, que envolvem, além dos próprios guarda-sóis, as partes mais ensolaradas da calçada.



Figura 3 Resultados obtidos para a imagem da feira

4 Resultados da imagem: A Feira

As segmentações realizadas na imagem da feira podem ser visualizadas na Figura 3. Neste caso, o objetivo definido foi realizar a separação dos pimentões, couves-flores e demais frutas e legumes, em relação ao restante da imagem.

As sementes do método de Crescimento de Regiões foram selecionadas sobre cada conjunto de legumes. Vê-se que a expansão ocorreu bem para toda a região de pimentões, couves-flores e verduras exibidos na imagem. Apesar disso, pode ser observado que a expansão ocorreu também para as caixas e para os preços nas placas em alguns casos. Algumas regiões de sombra e brilho nos legumes e verduras foram ignorados.

No *Watershed*, a pré-segmentação da cor vermelha impediu que os legumes verdes, amarelos e brancos fossem encontrados. Outra falha aqui observada foi a incapacidade do algoritmo de separar os pimentões vermelhos em suas unidades, visto que apenas um grupo foi encontrado, que engloba todos os elementos.

A aplicação do *K-means* adaptativo resultou na definição de seis clusters. Pode-se observar que uma das classes resultantes são os pimentões vermelhos, outra definiu os pimentões verdes, outra definiu uma coloração específica da madeira das caixas, outra ainda foi capaz de separar as sombras.

A Limiarização de Otsu foi realizada com dois limiares. Vê-se que a categoria de *pixels* escuros representa as sombras da imagem, enquanto a região intermediária engloba os pimentões vermelhos e verdes e algumas colorações de madeira. As regiões mais claras englobaram as placas, as regiões iluminadas das caixas, os pimentões amarelos e os couves-flores.

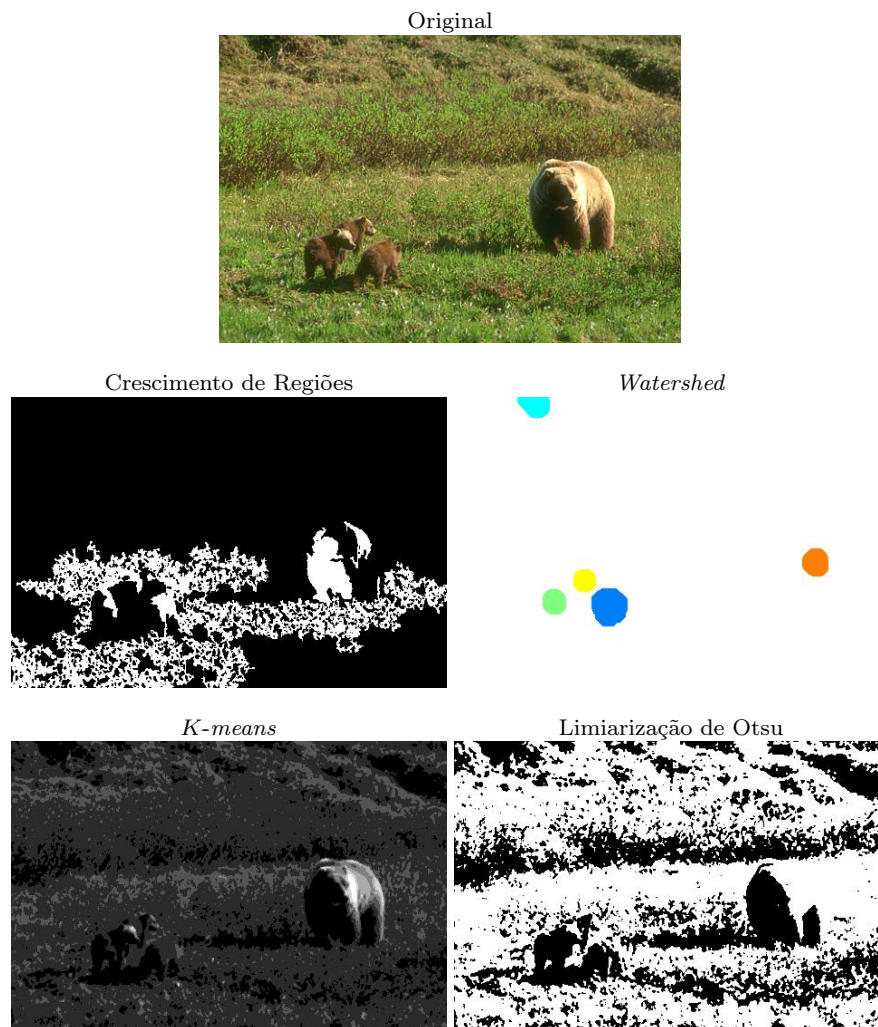


Figura 4 Resultados obtidos para a imagem dos ursos

5 Resultados da imagem: Os ursos

As segmentações feitas com cada uma das técnicas sobre a imagem dos ursos são mostradas na Figura 4. Essa imagem é particularmente difícil, pois além da alta granularidade da região de mata no fundo, há existência de regiões de fundo marrons com forte similaridade de cor em relação à pelagem dos ursos. O objetivo é separar os ursos do plano de fundo vegetal.

A técnica de Crescimento de Regiões teve um desempenho ruim. Foram selecionadas sementes nas regiões pertencentes aos ursos, no entanto, essas sementes expandiram de maneira desordenada pelo plano de fundo, mesmo com um limiar igual a 15. Para limiares menores, a expansão não ocorria; enquanto limiares maiores espalhavam demasiadamente as regiões.

O pré-processamento do *Watershed* fez uma filtragem por regiões de cor marrom, característica dos ursos. Como pode-se observar, as regiões de urso encontradas

não assumem suas formas corretamente, mas localizam com sucesso suas presenças, inclusive foi capaz de separar os ursos filhotes e reconhecer a existências dos três individualmente. Nesse caso específico, o *Watershed* obteve tanto sucesso que foi capaz de identificar um urso escondido na parte superior da imagem.

O *K-means* adaptativo encontrou sete clusters. A análise da imagem segmentada mostra que a performance foi ruim. Em diversos casos, a mesma categoria contém regiões de urso e de vegetação. Nota-se que a segmentação sofreu bastante com a granularidade do fundo e a presença de regiões de cor similar aos ursos.

A Limiarização de Otsu também não obteve bons resultados. Neste caso, definiu-se apenas um limiar. No entanto, esse algoritmo sofreu devido às similaridades de cor entre regiões da vegetação e os ursos, além das variações de luminosidade e a presença de brilho sobre a pelagem dos ursos.

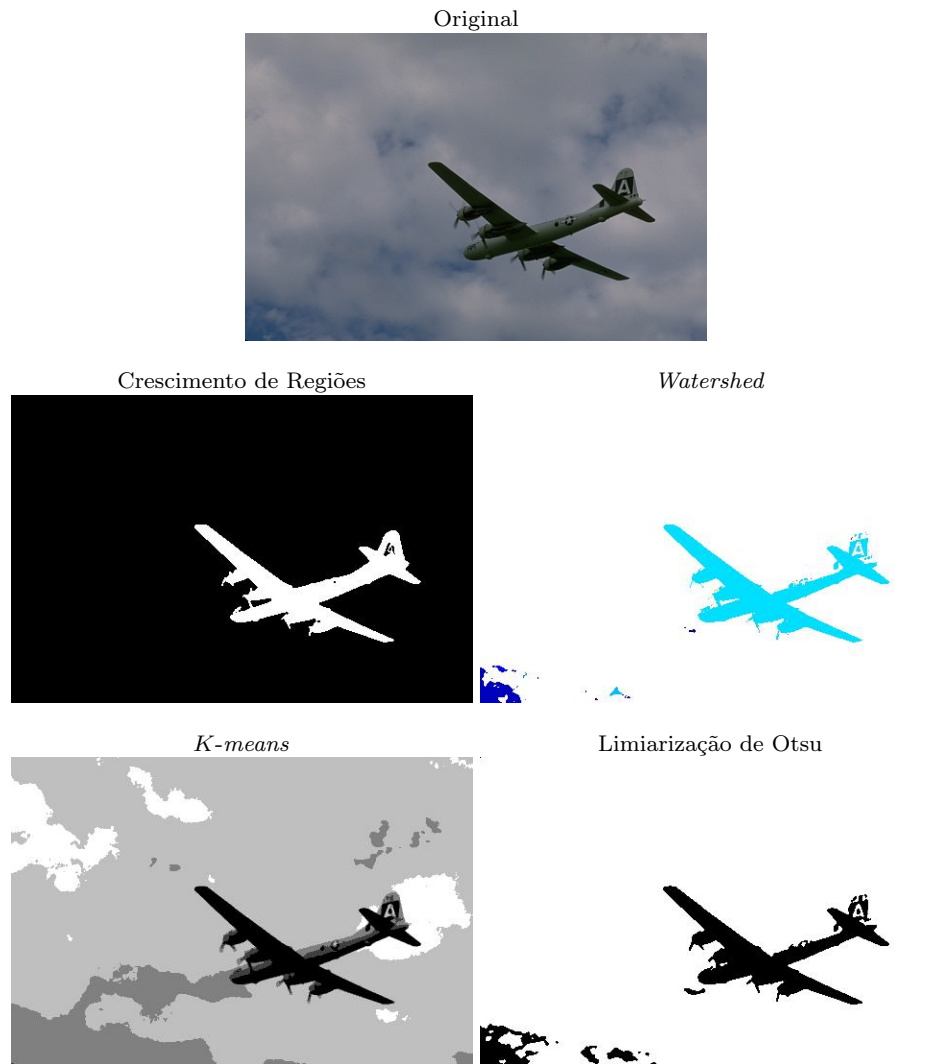


Figura 5 Resultados obtidos para a imagem do avião

6 Resultados da imagem: O avião

As segmentações obtidas para a imagem do avião são mostradas na Figura 5. O objetivo é separar o avião e o céu. Pode-se observar que em todos os casos a segmentação obtém um grau elevado de sucesso, embora ainda haja algumas falhas.

A maior falha ocorreu no *K-means*, pois o algoritmo de escolha adaptativa acabou por selecionar $K = 5$. Por consequência, houve uma supersegmentação da imagem. As outras falhas, menores, ocorreram no *Watershed* e na Limiarização de Otsu, na qual a região inferior esquerda da imagem foi incorretamente marcada: no *Watershed*, essa região foi marcada como um objeto separado do avião e do céu; enquanto em Otsu, a região foi marcada erroneamente como parte do avião.

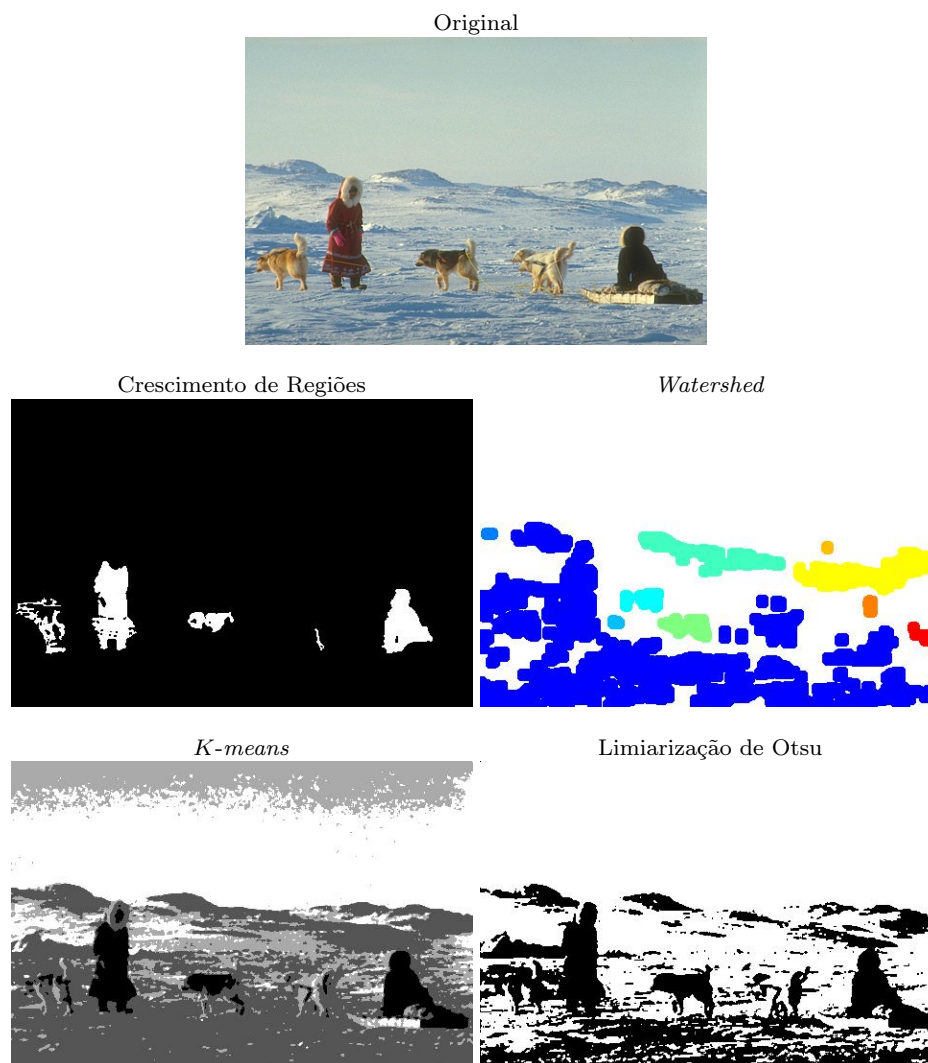


Figura 6 Resultados obtidos para a imagem dos esquimós

7 Resultados da imagem: Os Esquimós

As segmentações realizadas para a imagem dos esquimós é exibida na Figura 6. A princípio, pode-se notar quatro tipos de objetos importantes na cena: o céu, a neve, os esquimós e os cachorros. O objetivo é tentar separar esses objetos.

Na aplicação do Crescimento de Regiões, foi colocada uma semente em cada ser vivo da cena: os esquimós e os cachorros. Foi utilizado o limiar 18. Vê-se que apenas um esquimó da direita foi bem segmentado. Ademais, houve a segmentação parcial do esquimó da esquerda e do cachorro mais escuro, no centro. Nos outros cachorros, a expansão das regiões foi defeituosa.

No *Watershed*, a segmentação também não obteve bons resultados. De fato, é difícil extrair algum significado da segmentação nesse caso.

Na segmentação via *K-means*, o algoritmo adaptativo selecionou $K = 4$, um valor razoável. No entanto, a segmentação também ocorreu de maneira defeituosa. Nota-se que boa parte do céu foi segmentada corretamente, exceto a parte superior, cujo gradiente de coloração fez com que se confundisse com algumas regiões de neve. Os esquimós foram identificados e segmentados, com exceção das regiões brancas do capuz, que foram confundidas com a neve.

Enfim, a segmentação por Limiarização de Otsu foi aplicada com a utilização de apenas um limiar. Os esquimós e o cachorro do centro foram quase corretamente segmentados, no entanto, a segmentação do plano de fundo mostrou-se bastante irregular.

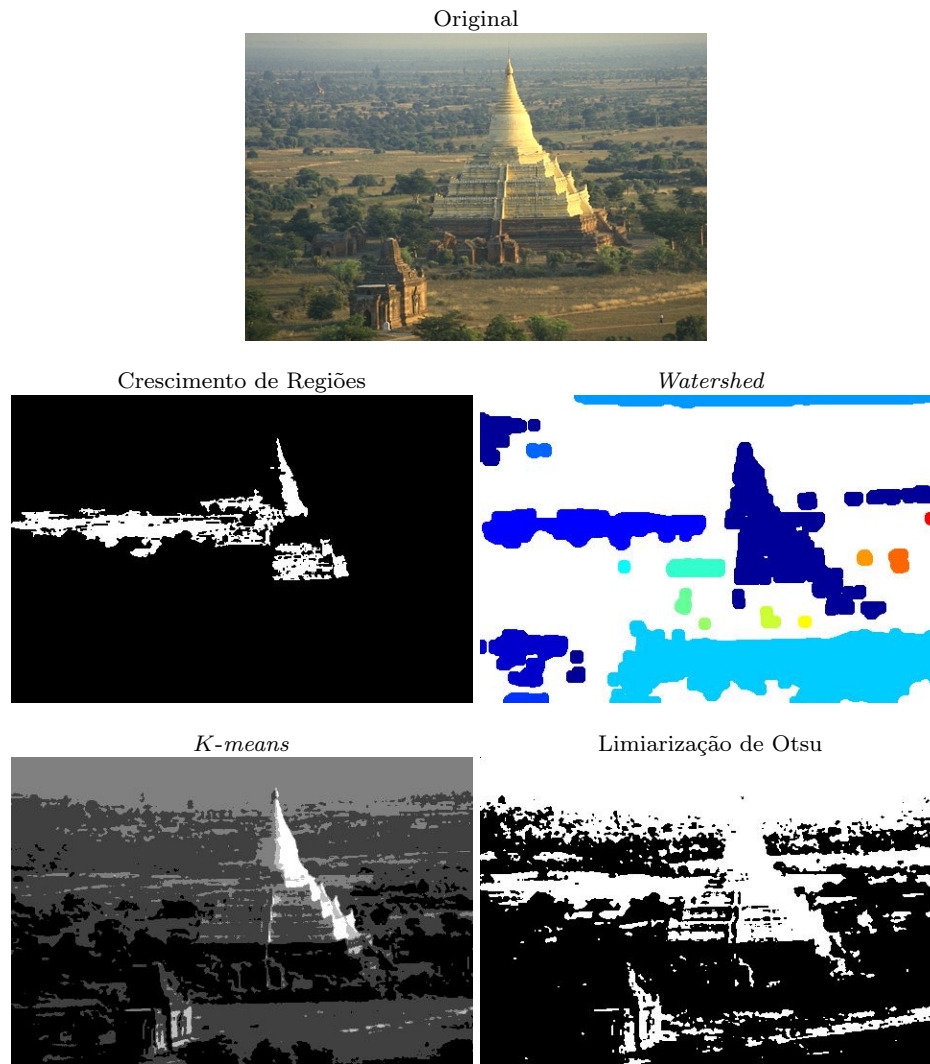


Figura 7 Resultados obtidos para a imagem do templo

8 Resultados da imagem: O Templo

Os resultados de segmentação alcançados para a imagem do templo são mostrados na Figura 7. Neste caso, o objeto principal é separar os templos do plano de fundo de campos e vegetação. A maior dificuldade nessa segmentação é a elevada presença de bordas e diferenças de iluminação. Outra questão é a similaridade de cor entre o templo pequeno e demais partes do terreno.

Para realizar o Crescimento de Regiões, foram posicionadas sementes em diversos *pixels* do templo maior, incluindo pontos de iluminação e de sombra. Além disso, foi selecionado um limiar igual a 12. A expansão das sementes não foi efetiva. Uma possível explicação é a presença de bordas de alto gradiente, impedindo a propagação das sementes. Simultaneamente, apesar da expansão não ter ocorrido em alguns pontos, em outros a

propagação deu-se de maneira errônea, atravessando as barreiras do templo e alcançando o plano de fundo.

O *Watershed*, por sua vez, foi capaz de segmentar corretamente metade do templo maior. O terreno em frente ao templo também foi segmentado com relativa eficácia. No entanto, as outras regiões de vegetação e o templo menor não foram identificados.

O *K-means* convergiu com $K = 5$. No entanto, a segmentação também não obteve boa qualidade. A parte iluminada do templo foi identificada como uma classe, mas sua região de sombra foi confundida com a vegetação ao redor.

A Limiarização de Otsu também teve resultados ruins. Apesar de sua segmentação ter abrangido quase a totalidade do templo maior, este ainda foi confundido com as regiões de fundo.

9 Discussão dos resultados

Os experimentos, juntamente à análise feita, permitiram identificar alguns desafios relacionados à segmentação. Primeiramente, deve-se considerar que os algoritmos aqui trabalhados fundamentam suas decisões de segmentação em informações de cor. Devido a isso, diferenças de iluminação sobre objetos representam uma fonte de dificuldade, pois alteram a cor aparente do objeto. A segunda maior dificuldade encontrada foi o alto nível de granularidade de algumas regiões, como nas imagens dos ursos e do templo.

O Crescimento de Regiões obteve bons resultados na imagens dos guarda-sóis, da feira e do avião. Nesses casos, foram responsáveis pelo sucesso: a uniformidade das cores dos objetos e a escolha de uma semente bem posicionada, com uma cor representativa. Nos outros casos, a expansão das regiões esbarrou nos obstáculos da alta granularidade e das diferenças de luminosidade.

O *Watershed* obteve resultados bons na imagem dos guarda-sóis, pois permitiu sua identificação individual e com poucos erros. Outro caso interessante foi na imagem dos ursos, pois mostrou-se capaz de identificar cada um deles individualmente, apesar de não ter conseguido se ajustar às suas formas.

As segmentações com *K-means* obtiveram resultados razoáveis, sendo os melhores vistos na imagem dos guarda-sóis e da feira, no qual foi capaz de separar os guarda-sóis e os pimentões sem nenhuma informação acerca de objetos de interesse. Nas outras imagens, seu desempenho foi medíocre.

A Limiarização de Otsu obteve bons desempenhos na imagem dos guarda-sóis e do avião. Sua principal dificuldade é a alta sensibilidade às condições de iluminação. Caso haja uma grande variação de iluminação em um mesmo objeto, como no caso do urso e do templo, a técnica tende a falhar.

10 Conclusões

O presente trabalho fez uma breve análise acerca de quatro técnicas de segmentação: Crescimento de Regiões, *Watershed*, *K-means* e Limiarização de Otsu. Seus desempenhos foram averiguados sobre seis imagens contendo cenas variadas.

Durante a análise, foram identificadas algumas das principais dificuldades da segmentação. Entre elas, foi possível perceber a influência das variações bruscas de iluminação e alta granularidade de algumas regiões. Viu-se também que, entre as técnicas estudadas, nenhuma se adéqua com eficácia a todas as imagens. Portanto, conclui-se que existe uma forte necessidade de estudo

das propriedades das imagens de interesse e adaptação de uma cadeia de pré-processamento específica para cada uma delas.

Referências

1. D.J. Kroon. Region growing (2008). URL <https://www.mathworks.com/matlabcentral/fileexchange/19084-region-growing>
2. A. Dixit. K-means segmentation (2014). URL <https://www.mathworks.com/matlabcentral/fileexchange/52697-k-means-segmentation>