

Segmentação: Watershed

29 de maio de 2002

Nina S. T. Hirata
 Depto. de Ciência da Computação
 IME - USP

[Este texto é baseado em partes da dissertação de mestrado de Roberto Hirata Jr., intitulada “Segmentação de Imagens por Morfologia Matemática”, defendida em 1997 no IME-USP.]

Referências : [1, 2]

1 Alguns conceitos básicos

Distância geodésica: Seja $A \subset \mathbf{E}$ e d uma métrica sobre \mathbf{E} . A **distância geodésica** entre x e y é definida por

$$d_A(x, y) = \min\{l(P) : P \text{ é um caminho entre } x \text{ e } y \text{ totalmente contido no conjunto } A\}$$

onde $l(P)$ corresponde ao comprimento do caminho P .

Sejam $x, y \in A$, $B \subset A$, $x \notin B$. A **distância geodésica do ponto x ao conjunto B** é definida por

$$d_A(x, B) = \min\{d_A(x, y) : y \in B\}$$

Zona de influência geodésica: Seja B um conjunto formado por B_1, B_2, \dots, B_k , k componentes conexas de A . A **zona de influência geodésica** da componente B_i é definida por

$$iz_A(B_i) = \{p \in A, \forall j \in [1, k] / \{i\}, d_A(p, B_i) < d_A(p, B_j)\}$$

ou seja, todos os pontos de A que estão mais próximos de B_i do que qualquer outra componente B_j .

SKIZ: Os pontos que não pertencem a nenhuma zona de influência geodésica pertencem ao esqueleto por zonas de influência (“Skeleton by Influence Zones”) ou **SKIZ** de B dentro de A , que é denotado $SKIZ_A(B)$, isto é,

$$SKIZ_A(B) = A / IZ_A(B),$$

onde $IZ_A(B) = \bigcup_{i \in [1, k]} iz_A(B_i)$.

A figura 1 ilustra o conceito de distância geodésica e zona de influência geodésica.



Figura 1: Distância Geodésica e Zona de influência geodésica.

2 Linhas de Partição de Águas

Uma imagem em níveis de cinza pode ser vista como um relevo onde encontramos partes que se assemelham a vales e outras a montanhas. Quando chove, a medida que a água vai enchendo os vales e formando poças, elas vão crescendo e juntando-se a outras poças, formadas pela acumulação d'água em vales vizinhos.

Bacias de captação: O nome que se dá aos vales onde acumulam-se as poças.

Linhas de Partição de Águas : corresponde ao lugar geométrico formado pelos pontos onde ocorre a junção das poças.

Mínimo Regional: Seja f uma imagem níveis de cinza, dizemos que $R_h \subset \mathbf{E}$ é um **Mínimo Regional** de altitude h se R_h é um conjunto conexo de pontos de \mathbf{E} tal que $f(p) = h, \forall p \in R_h$ e $f(p) > h$ para todos os pontos da borda externa de R_h .

Chamamos de $RM_h(f)$ o conjunto dos mínimos regionais de f de altitude h (ou seja, todas as regiões planas de altitude h , cercadas por "paredes laterais" mais altas).

Notação : Sejam h_{min} e h_{max} , respectivamente, o menor e o maior valor de f , i.e., $h_{min} = \min\{f(x) : x \in \mathbf{E}\}$ e $h_{max} = \max\{f(x) : x \in \mathbf{E}\}$.

A seção inferior de f no nível h é o conjunto definido por,

$$f^h = \{x \in \mathbf{E} : f(x) \leq h\}$$

Bacias de Captação de f : É o conjunto $X_{h_{max}}$ obtido pela relação de recorrência,

$$X_{h_{min}} = f^{h_{min}}(f)$$

$$X_{h+1} = RM_{h+1}(f) \cup IZ_{f^{h+1}}(X_h)$$

para todo $h \in [h_{min}, h_{max} - 1]$.

Linhas de Partição d'Águas As **Linhas de Partição d'Águas** da função f é o conjunto complementar das Bacias de Captação de f .

A figura 2 mostra a dinâmica da relação de recorrência acima, para dois níveis.

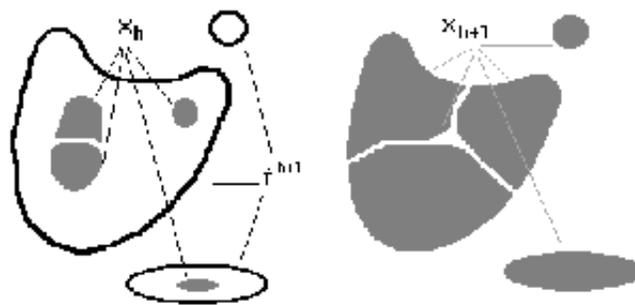


Figura 2: Dinâmica da recorrência

3 Paradigma de segmentação baseada em watershed

Os métodos existentes para segmentar uma imagem utilizam basicamente duas idéias. Uma delas é achar os contornos dos objetos na imagem. A outra, é agrupar pontos que tenham **características** semelhantes até que o objeto de interesse seja reconstruído. O Paradigma de Beucher é um método baseado na extração dos contornos dos objetos a serem segmentados. O problema de identificação de contornos pode ser resolvido utilizando-se o operador LPE (linha de partição de águas), em conjunto com outros operadores morfológicos que servem para “preparar” a imagem. Essa preparação deve ocorrer para o realce e identificação dos objetos que devem ser segmentados.

Contornos, em geral, são descontinuidades na imagem e estes podem ser realçados através dos operadores de diferenciação. Normalmente, usa-se o operador gradiente morfológico para isso.

A aplicação do gradiente morfológico em uma imagem em níveis de cinza resulta em uma outra imagem em níveis de cinza com picos e vales onde estão as descontinuidades; tanto maior será a altura dos picos quanto maior forem as descontinuidades. É importante notar que o operador gradiente é muito sensível a ruídos, i.e., a variações de intensidade nos níveis de cinza da imagem, daí a grande quantidade de picos e vales na imagem (partes escuras e claras).

A aplicação do operador LPE sobre a imagem do gradiente resulta, em geral, numa super-segmentação da imagem (maior número de linhas de partição do que de objetos na imagem). A figura 3 mostra uma imagem de células de músculo, o inverso do resultado do operador gradiente morfológico, donde queremos segmentar as células escuras ¹, e a supersegmentação resultante do operador LPE.

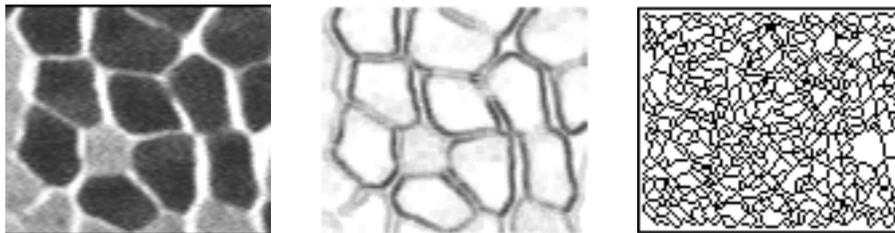


Figura 3: Tecido de músculo, gradiente invertido e supersegmentação.

Filtrar a imagem original nem sempre é uma boa solução para forçar o desaparecimento dos ruídos pois pode acarretar o desaparecimento de contornos importantes. A filtragem posterior à aplicação do operador LPE sobre o resultado do gradiente, com intuito de diminuir a super-segmentação, também pode acarretar o desaparecimento de contornos.

A técnica mais eficiente para resolver o problema acima é a **Mudança de Homotopia do Gradiente**, que consiste em, antes da aplicação do operador LPE, filtrar a imagem resultante da aplicação do gradiente para eliminar as descontinuidades responsáveis pela super-segmentação.

3.1 Mudança de Homotopia do Gradiente

Para fazer a mudança de homotopia do gradiente é preciso introduzir uma nova fase ao processo de segmentação que depende do conhecimento prévio dos objetos a serem segmentados, i.e., temos que introduzir a informação de quais bacias são importantes e quais não, baseado nas características dos objetos que queremos segmentar. Essa informação é introduzida através da escolha (automática, ou não) de regiões conexas de pontos da imagem, que serão chamados de **marcadores**. A finalidade de um marcador será indicar qual objeto deve ser segmentado.

¹É fácil ver que a aplicação um simples threshold nessa imagem pode separar as células escuras do fundo da imagem, mas não as separa entre si. Daí termos de usar outras ferramentas de segmentação.

A mudança de homotopia do gradiente pode ser entendida como um processo de aplainamento de relevos, na direção externa em relação aos vales marcados. Este processo elimina os vales que não são de interesse. A figura 4 mostra uma função h e a função f resultante da aplicação do gradiente morfológico em h , respectivamente. O objetivo será segmentar o objeto representado pelos valores entre 2 e 22 de h . Note que há um vale importante correspondente a essa região em f e vários outros vales devidos a ruídos na imagem. A aplicação do operador LPE sobre f resultaria numa imagem com excesso de contornos, como mencionado acima, pois cada vale corresponde a uma bacia de captação. A função g , mostrada na fig. 5, foi obtida por uma limiarização da função h que segmentou apenas o ponto mais baixo do vale. Ela será a imagem do marcador para o objeto que queremos segmentar. O resultado final da mudança de homotopia é mostrado ao lado do marcador.

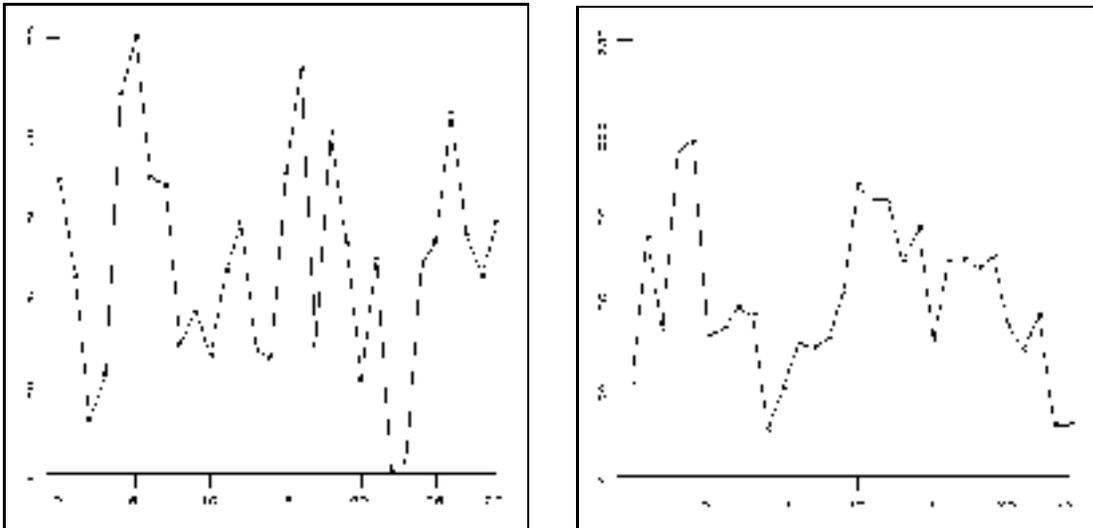


Figura 4: Função unidimensional h e seu gradiente f .

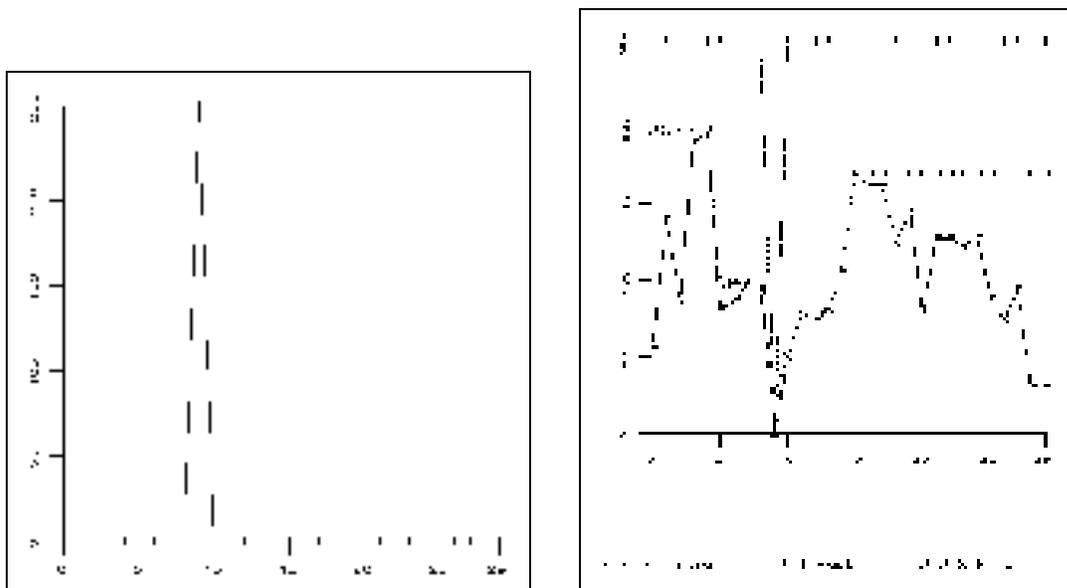


Figura 5: Função Marcadora e resultado final da mudança de homotopia.

No caso do problema das células escuras, a primeira imagem da figura 6 mostra os marcadores para as células escuras, a imagem filtrada pela mudança de homotopia e o resultado após a aplicação do operador LPE, resultando em uma segmentação satisfatória.

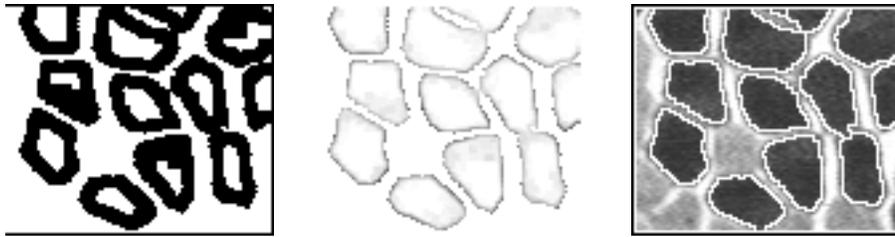


Figura 6: Marcadores, filtragem homotópica e resultado final da segmentação.

O modelo (processo) de segmentação de Beucher pode ser resumido no fluxograma mostrado na fig. 7.

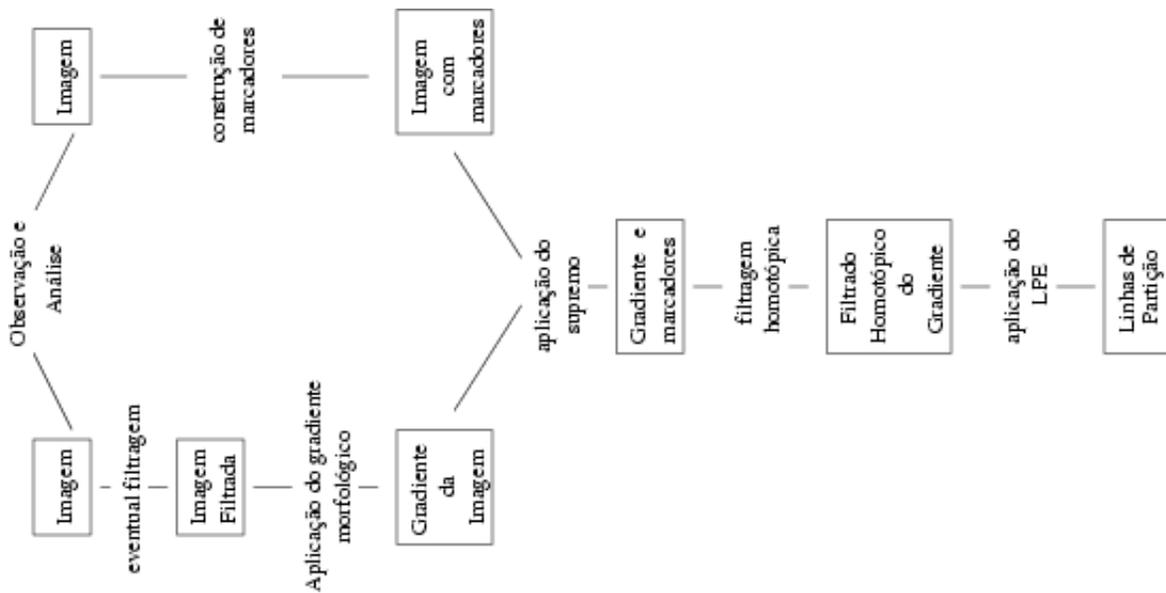


Figura 7: Paradigma de Beucher.

O lado direito do fluxograma corresponde a parte heurística do método, que é o projeto de operadores para achar marcadores na imagem para os objetos de interesse. A parte esquerda corresponde as etapas anteriores à aplicação da mudança de homotopia do gradiente da imagem. Nesta etapa, após a observação e análise visual da imagem, pode-se aplicar filtragens para que o resultado do gradiente seja o melhor possível. Deve-se tomar muito cuidado nisto para evitar o desaparecimento de contornos importantes.

A imagem dos marcadores e do gradiente resultam destas etapas. Em seguida, a imagem dos marcadores deve ser negada e o ínfimo deste resultado com a imagem original deve ser calculada. O resultado do ínfimo será usado para fazer a mudança da homotopia do gradiente. O resultado deste operador será uma imagem com os contornos dos objetos de interesse realçados e sem as discontinuidades devidas aos ruídos. É como se passássemos um trator em cima da superfície, tirando as pequenas colinas e enchendo os pequenos vales, fazendo com que tenhamos apenas uma bacia de captação para cada marcador construído. A aplicação do operador LPE produzirá, então, uma imagem com as linhas de partição bem posicionadas sobre os contornos dos objetos.

A marcação pode ser feita manualmente para algumas imagens, entretanto é comum estarmos interessados em obter uma metodologia robusta para a geração de marcadores, para que ela sirva para outras imagens com características semelhantes.

Através do paradigma de Beucher, **o problema de detecção de contornos que, em geral, é muito complicado, é substituído por um problema mais simples, que é o de achar marcadores para os objetos de interesse na imagem.**

Atualmente, os operadores para produzir os marcadores são projetados de forma “ad-hoc”. É difícil evitar esta parte heurística no processo de segmentação já que a segmentação envolve um conhecimento a priori dos objetos a serem segmentados. Por exemplo, um médico acostumado a examinar o resultado de exames de R-X pode facilmente ver um problema como um cisto na imagem, enquanto um leigo teria dificuldade para interpretar os objetos na imagem. Neste caso, o médico ensinaria o profissional em processamento de imagens para fazer um operador que marcasse e segmentasse cistos na imagem.

4 Exemplos de aplicação

4.1 Separação de Objetos Sobrepostos

Em alguns casos, podemos estar interessados em separar objetos que aparecem sobrepostos na imagem. Por exemplo, podemos ter interesse na contagem de glóbulos brancos no sangue, ou contagem da quantidade ou estimativa do tamanho médio de grãos de soja, ou feijão, etc.

A figura 8 mostra uma série de objetos arredondados sobrepostos (uma imagem binária) e a respectiva função distância, para conseguir um modelo dos objetos que os diferencie pela distância de seus centros às suas bordas.

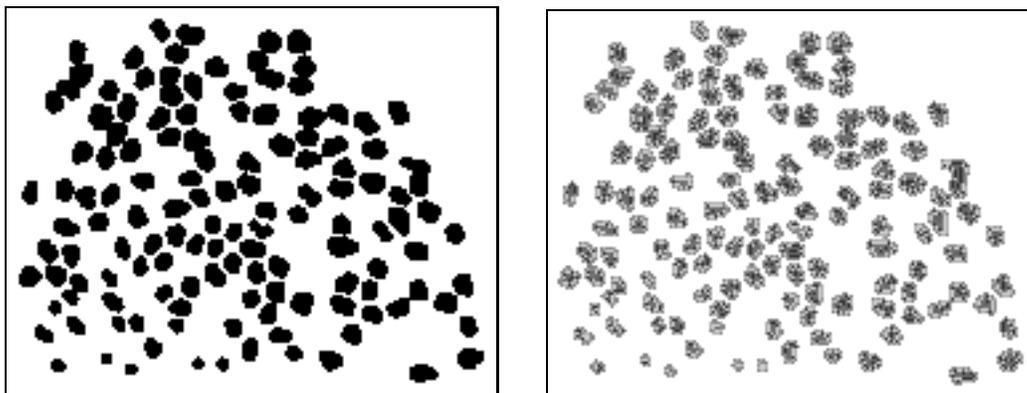


Figura 8: Objetos sobrepostos e Função Distância

Para encontrar marcadores para os objetos, pode-se calcular os máximos locais da função distância. Quando há mais de um máximo local em um objeto, a dilatação pode ser utilizada para uní-los e em seguida uma erosão pode ser aplicada para que os marcadores não fiquem muito grandes. A imagem da fig. 9 mostra os marcadores encontrados pelo operador descrito acima. A imagem foi composta com a imagem original para melhor visualização do posicionamento dos marcadores.

Calcula-se o supremo dos marcadores com o resultado da função distância e aplica-se o operador LPE sobre a imagem negada deste supremo. A imagem ao lado dos marcadores na fig. 9 mostra as linhas de partição d'águas encontradas. Note que as linhas não correspondem às bordas dos objetos, mas elas podem ser usadas para separá-los.

A subtração morfológica da imagem original pela imagem das linhas de partição pode ser usada para separar os objetos que se sobrepõe. Como as bordas podem não ser suaves, pode-se melhorá-las aplicando uma abertura por um disco (no caso foi usado um disco de diâmetro sete em métrica Euclidiana) para alisá-las. A imagem da fig. 10 mostra o resultado da subtração e o resultado final da segmentação.

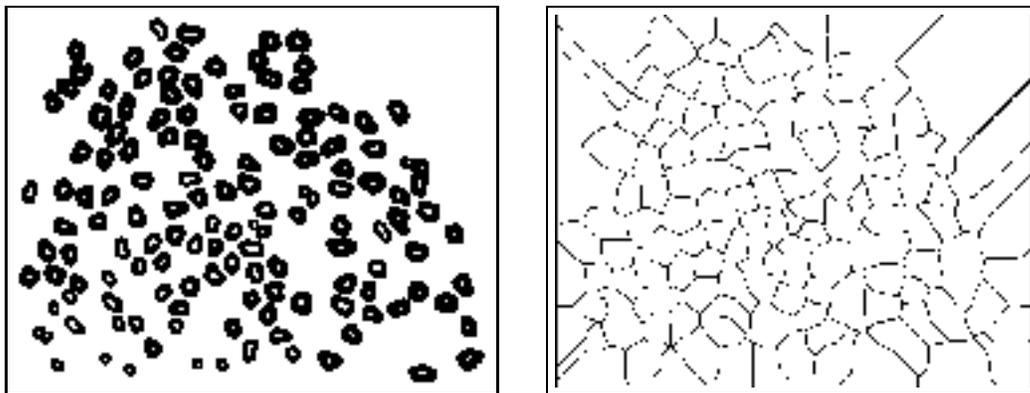


Figura 9: Objetos com seus marcadores e o resultado do operador LPE.

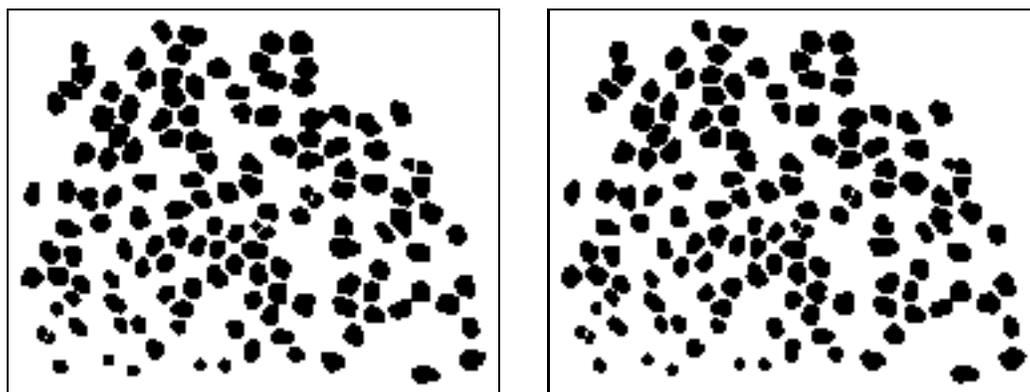


Figura 10: Subtração morfológica e resultado final.

4.2 Calculadora

Segmentação dos caracteres

A primeira imagem da figura 11 mostra a face de uma calculadora. Os pontos que formam as teclas são, em média, mais escuros que os pontos fora das teclas ou os pontos dos caracteres impressos. Estes últimos, por sua vez são, em média, bem claros. Uma limiarização poderia ser tentada para segmentar os dígitos, mas o resultado não é muito bom.

A melhor maneira de segmentar os picos mais altos de uma imagem, como neste caso, é usar o operador “top hat”, que é o resíduo da reconstrução da imagem original a partir de uma erosão da imagem original. A segunda imagem da fig. 11 mostra o resultado após cinco erosões da imagem original por um quadrado elementar.

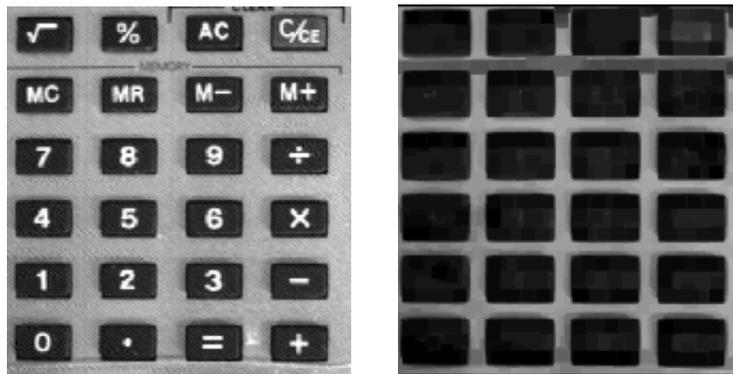


Figura 11: Imagem original e resultado após 5 erosões.

A imagem da fig. 12 mostra o resultado da reconstrução da imagem original pela imagem erodida. O resíduo ou subtração da imagem original pela imagem da fig. 12 é a imagem em níveis de cinza dos dígitos (veja segunda imagem da fig. 12). Uma simples limiarização resolve a segmentação dos dígitos, como pode ser visto na terceira imagem da fig. 12.

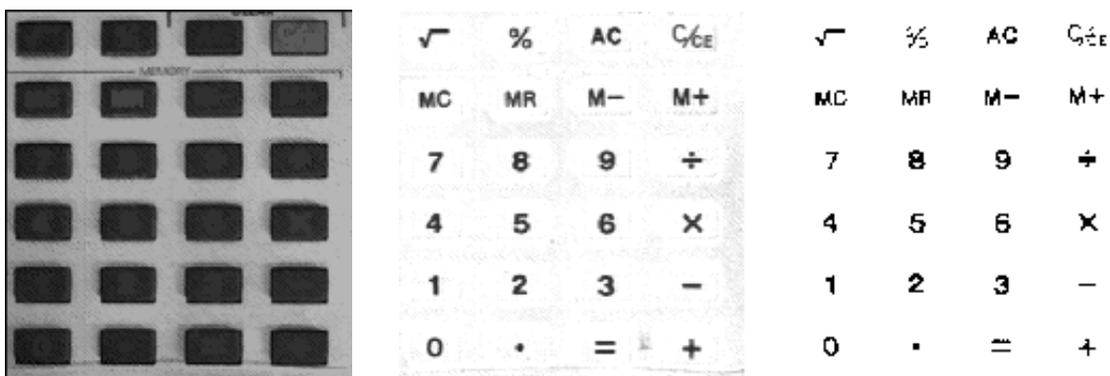


Figura 12: Reconstrução, imagem do resíduo e dígitos segmentados.

Segmentação das bordas das teclas

Como cada tecla tem um caractere impresso, estes serão usados como marcadores internos para as teclas. Para isso, primeiramente a imagem dos dígitos é dilatada 3 vezes por um quadrado elementar para tornar os marcadores conexos.

Para construir os marcadores externos, o tamanho da imagem é aumentada, uma moldura colocada, e a zona de influência geodésica (SKIZ) dos objetos da imagem calculada. O resultado da dilatação pode ser visto na primeira imagem da fig. 13. A segunda imagem é o SKIZ, e a terceira é o supremo das imagens dos marcadores internos e externos para as teclas.

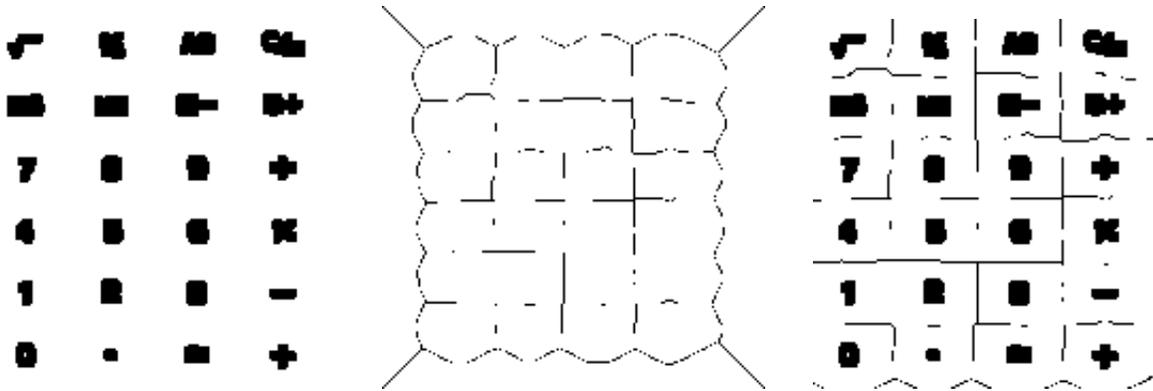


Figura 13: Dilatação dos dígitos, SKIZ da dilatação e marcadores.

Para aplicar o Paradigma de Beucher deve-se calcular o gradiente da imagem original (mostrado invertido para melhor visualização na fig. 14). O resultado da mudança de homotopia do gradiente da imagem original a partir da imagem dos marcadores é mostrada ao seu lado.

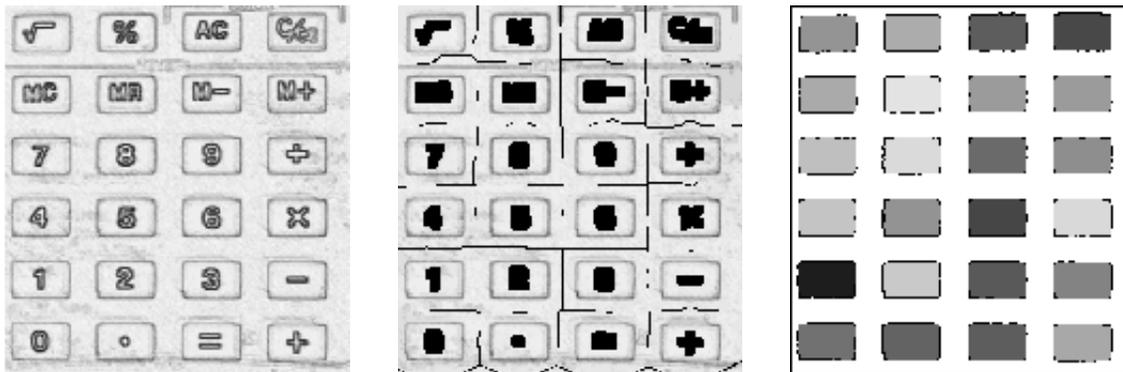


Figura 14: Gradiente-invertido e resultado da mudança da homotopia.

A imagem da fig. 15 mostra o resultado do operador LPE aplicado sobre a imagem resultante da mudança de homotopia. As linhas escuras correspondem às linhas de partição de águas e cada tom de cinza indica uma bacia de captação, que corresponde a uma tecla. A lado dessa imagem é mostrada apenas as linhas de partição, obtida através de uma simples limiarização.

4.3 Segmentação de Arestas para Aplicações Robóticas

Usualmente, um robô que manipula objetos sólidos 3D tem como entrada imagens 2D destes objetos. Para posicionar as garras do robô para que este pegue um objeto é importante saber a posição das arestas do objeto. Isso pode ser feito por um pré-processador que extrai as arestas dos objetos de interesse das imagens de entrada.

As imagens da fig. 16 mostram um sólido visto por uma câmera e o respectivo resultado do operador gradiente (invertida para melhor visualização dos detalhes).

A primeira imagem da fig. 17 mostra o resultado do operador LPE aplicado sobre a imagem do gradiente.

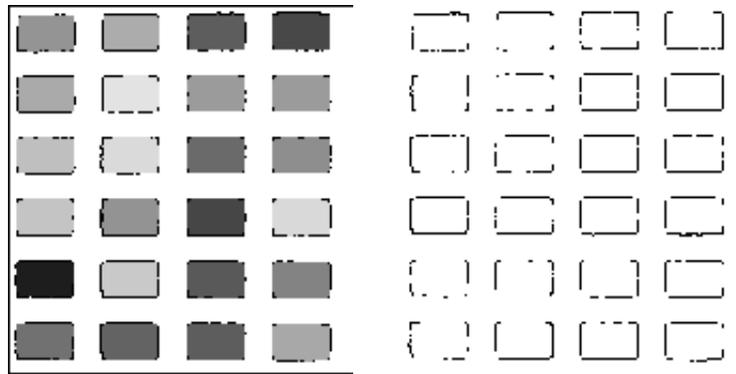


Figura 15: Watershed mais regiões e LPE

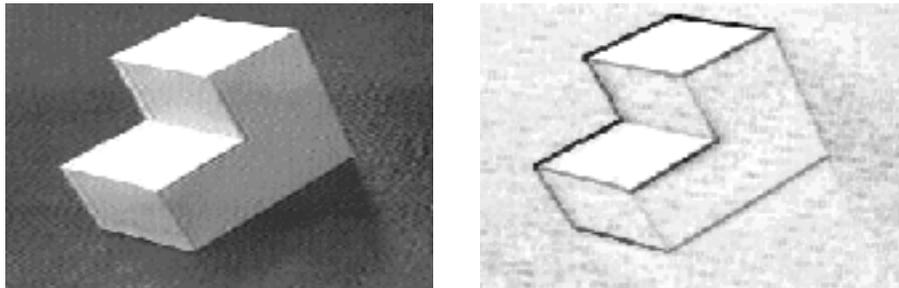


Figura 16: Imagem original e gradiente invertido

Este é mais um exemplo de que o operador LPE não deve ser aplicado sobre a imagem do gradiente sem a filtragem homotópica.

Para encontrar os marcadores para as faces do sólido usamos primeiramente três limiarizações diferentes (uma para as duas faces mais claras, uma para a face em forma de L e para a face mais escura inferior e uma para a face mais escura superior). Em seguida, para cada limiarização escolhemos uma erosão que torna o marcador único para cada face. Para o marcador externo, fizemos uma limiarização que segmenta todo o bloco e depois aplicamos um filtro composto por uma abertura por um disco de diâmetro 8, uma dilatação por um disco de diâmetro 10 e uma erosão por um quadrado elementar. A segunda imagem da fig. 17 mostra o supremo dos marcadores internos e externos obtidos pelas limiarizações e filtrações morfológicas da imagem original. Note que todas as faces contêm um marcador, para que todas as arestas sejam encontradas.

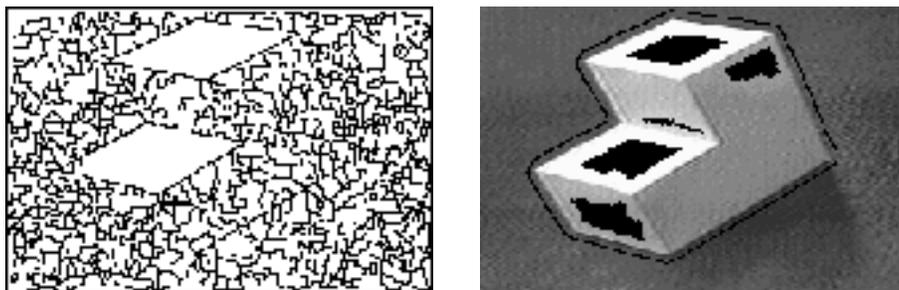


Figura 17: Watershed frustado e marcadores

A primeira imagem da fig. 18 mostra o resultado da mudança de homotopia da imagem do gradiente morfológico usando como marcadores a imagem da fig. 17. A imagem foi invertida para melhor visualização dos detalhes. Note que a imagem está mais suave do que a imagem do gradiente. A segunda imagem da

fig. 18 mostra o resultado do operador LPE. A terceira imagem mostra o resultado de uma limiarização que segmenta apenas as linhas de partição de águas da imagem.

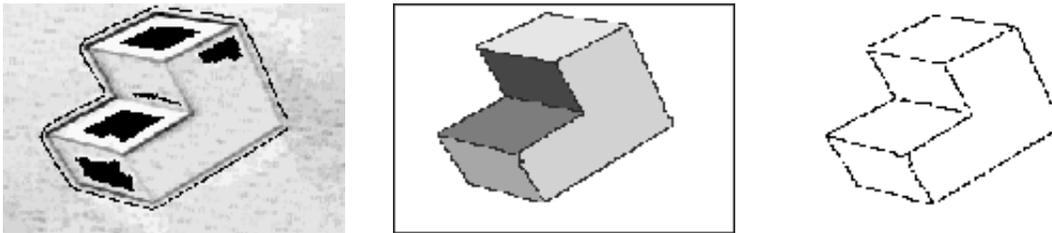


Figura 18: Mudança de homotopia, resultado do Watershed e LPE.

A imagem da fig. 19 mostra o resultado da composição da imagem das linhas de partição com a imagem original.

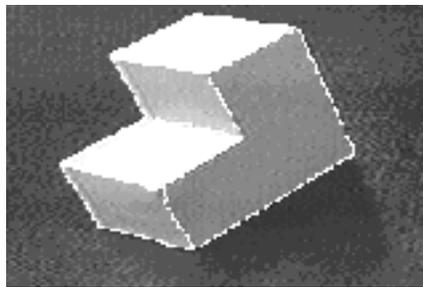


Figura 19: Composição da LPE com a imagem original.

A composição destes operadores resulta em um operador para segmentar arestas de imagens 2D bastante robusto, como pode ser visto nos exemplos que seguem. A fig. 20 mostra os resultados para objetos similares.

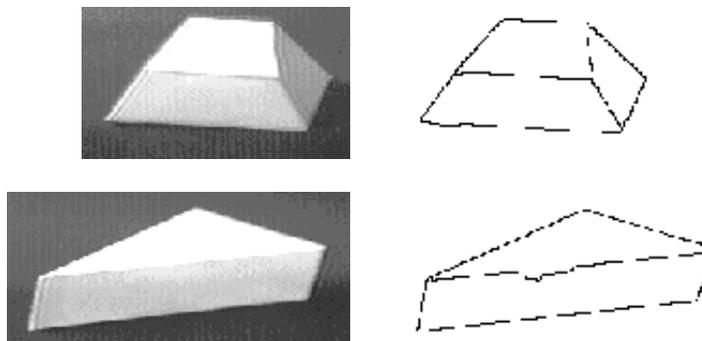


Figura 20: Outro blocos e respectivos LPEs.

Referências

- [1] L. Vincent and P. Soille. Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):583–598, June 1991.
- [2] S. Beucher and F. Meyer. *Mathematical Morphology in Image Processing*, chapter 12. The Morphological Approach to Segmentation: The Watershed Transformation, pages 433–481. Marcel Dekker, 1992.