

# Processamento de Imagens Heterogêneas

JACQUES FACON<sup>1</sup>

<sup>1</sup>CEFET-PR - Centro Federal de Educação Tecnológica do Paraná  
CPGEI - Curso de Pós-Graduação em Engenharia Elétrica e Informática Industrial  
Av. Sete de setembro, 3165  
80230-901 Curitiba-PR, Brasil  
b.facon@cefet.anpr.br

**Abstract.** How can we analyse and interpretate heterogenous images? The use of classical techniques doesn't resolve the problem. This paper presents solutions based on combinations of morphological operations. We show how to delete the heretogenous informations to return filtered and segmented images. We demonstrate the results of this process applied on deteriorated images.

## 1 INTRODUÇÃO

Uma imagem contém, de forma geral, diversas informações que podem ser, em função da interpretação, consideradas como picos e vales. A detecção desses faz parte das tarefas comuns da área de Processamento de Imagens. No entanto, pela natureza das imagens analisadas, esta operação não constitui uma tarefa trivial. A manipulação de imagens de diversas origens (aéreas, médicas, ecográficas, industriais, etc...) demonstra que a informação relevante está, de forma geral, misturada com várias informações ruidosas (perturbações meteorológicas, resíduos de corante, ruído ecográfico, poeira, etc...). Acontece que essas perturbações podem ser de um tamanho importante. Para simplificar as explicações, chamaremos de *fundo* as informações ruidosas, já que essas são confinadas no sinal portador. A imagem, onde encontram-se misturados o fundo ruidoso e a informação relevante, será chamada de *Imagem Heterogênea*.

A diversificação da própria área de Processamento de Imagens exige ferramentas, facéis de uso e automatizáveis, filtrando e segmentando esse tipo de imagens. A filtragem, como a segmentação, exige a detecção da informação relevante. Como conseguir essa detecção quando os picos e vales estão misturados com diversos ruídos que chegam a ter um tamanho importante? Os filtros clássicos (filtro da média, da mediana, de Deriche, de Nagao-Matsuyama, filtros espaciais, limiarização), de forma geral, são eficientes para um ruído do tipo pimenta e sal (*salt and pepper*), mas fracassam quando o tamanho do ruído cresce. Técnicas de detecção de bordas (gradiente, zero-crossing) colocam em evidência, além da informação relevante, bordas inúteis correspondentes a este ruído.

Geometricamente falando, a operação de extra-

ção de picos (vales) pode ser interpretada da mesma maneira que a noção de máximo (mínimo). A dificuldade dos filtros clássicos em extrair este tipo de informação provém da incapacidade em discernir as noções de máximo e de mínimo. Um pico  $P$  (vale  $V$ ) extraído é uma região conexa de pixels tal como:

- cada pixel  $p$  vizinho da região  $P$  ( $V$ ) satisfaz:

$$f(p) < \text{mínimo} (> \text{máximo})\{f(q) : q \in P\}$$

- cada pixel  $p$  vizinho da região  $P$  ( $V$ ) satisfaz:

$$\text{máximo}\{f(q) : q \in P\} - \text{mínimo}\{f(q) : q \in P\} < v$$

Além destes parâmetros, um terceiro critério permite definir completamente os picos (vales): o tamanho (ou largura). Filtrar e segmentar uma imagem heterogênea deve então consistir em definir ferramentas que associam noções de tamanho às de nível de cinza (altura e profundidade). Uma tal ferramenta existe, é a Morfologia Matemática [Serra-1982]. A Morfologia Matemática consiste em fazer interagir um conjunto desconhecido (imagem) com um conjunto previamente definido (elemento estruturante). O elemento estruturante, além de definir a região de interação, define, por suas dimensões, o critério tamanho. Os critérios máximo e mínimo são definidos através das duas operações morfológicas elementares que são a erosão e a dilatação que agem, a partir do elemento estruturante  $g$ , sobre a imagem  $f$  da seguinte maneira:

$$\begin{aligned} f_{\text{erodido}}(x) &= f(x) \text{ ero } g(x) \\ &= \text{Min}\{f(y) : y \in D[g_x]\} \end{aligned}$$

$$\begin{aligned} f_{\text{dilataado}}(x) &= f(x) \text{ dil } g(x) \\ &= \text{Max}\{f(y) : y \in D[g_x]\} \end{aligned}$$

As propriedades de distributividade e de combinação da erosão e da dilatação permitem definir outras duas operações que são a abertura:

$$f \text{ abe } g = (f \text{ ero } g) \text{ dil } \tilde{g}$$

e o fechamento:

$$f \text{ fec } g = (f \text{ dil } g) \text{ ero } \tilde{g}$$

Explorando as propriedades dessas quatro operações, serão apresentadas ferramentas morfológicas compostas, sucessivamente os tophats por abertura, por fechamento, por reconstrução, e os Hdome e Hbasin.

## 2 TOPHAT CLÁSSICO

As operações de abertura e de fechamento tem a facultade de reduzir o ruído de tamanho inferior ao do elemento estruturante usado. Acontece que, dependendo do tamanho do ruído, esse não pode ser removido sem remover também as informações relevantes. Explorando esta aparente dificuldade, a filtragem e segmentação vão consiste em duas etapas, a eliminação e a recuperação.

### 2.1 TOPHAT POR ABERTURA

No caso de detecção de picos, um modo de pensar consiste em usar uma combinação entre uma imagem original e a imagem correspondente aberta [Serra-1982]. A abertura deve ser efetuada de tal maneira a eliminar a informação relevante para conservar somente o fundo ruidoso. Subtraindo este resultado parcial à imagem original, obtemos a regeneração da informação relevante e eliminação do ruído. A transformação *Tophat1*, que segue este princípio, expressa-se da seguinte maneira:

$$\text{Tophat1}(f) = (f) - (f \text{ abe } g)$$

Como a abertura é um processo anti-extensivo, o seu resultado fica abaixo do sinal original, portanto a transformação *Tophat1* é sempre positiva.

A figura 1-a) e -b) ilustra o processamento de uma imagem heterogênea de retina do olho humano. Como os vasos sanguíneos representam a informação do tipo "pico", foi empregado um Tophat por abertura. A abertura da imagem, a partir de um elemento estruturante 3x3 em cruz, com 40 iterações eliminou a informação relevante e preservou as regiões heterogêneas (incluindo a mancha). Por subtração, estas regiões sumiram devolvendo os vasos com mais nitidez (figura 1-b).

### 2.2 TOPHAT POR FECHAMENTO

De forma análoga, a detecção de vales consiste em usar uma combinação entre uma imagem original e a imagem correspondente fechada [Serra-1982]. O processo de fechamento permite eliminar a informação relevante e conservar as partes ruidosas. Realizado com o elemento estruturante adequado, as partes ruidosas ficam bem nítidas de tal maneira que, durante a subtração, a informação relevante seja regenerada e o ruído eliminado. Esta transformação dual do *Tophat1*, chamada de detector de vales *Tophat2*, é definida da seguinte maneira:

$$\text{Tophat2}(f) = (f \text{ fec } g) - (f)$$

Como o fechamento é um processo extensivo, o seu resultado da transformação *Tophat2* é sempre positivo.

Caso o elemento estruturante  $g$  plano contenha  $g = 0$  no seu domínio, podemos estabelecer o resultado particular seguinte:

$$\begin{aligned} \text{Tophat1}(-f) &= (-f) - ((-f) \text{ abe } g) \\ &= (f \text{ fec } g) - (f) \\ &= \text{Tophat2}(f) \end{aligned}$$

A figura 1-c) ilustra o processamento de um texto parcialmente escondido pela sobreposição de uma assinatura e degradado pela adição de uma mancha. Desejando recuperar o texto que representa uma informação do tipo "vale", foi empregado um Tophat por fechamento. Durante o fechamento da imagem, a partir de um elemento estruturante 3x3 em cruz, com 5 iterações, foram preservados a assinatura e a mancha permaneceram enquanto a informação relevante sumiu. A subtração devolveu um texto mais nítido (figura 1-d) e suprimiu as regiões heterogêneas.

### 3 TOPHAT POR RECONSTRUÇÃO

Vimos anteriormente que, através da abertura e do fechamento, podíamos construir uma nova imagem contendo somente o ruído ou as informações relevantes. As operações de erosão e de dilatação eram rigorosamente em números iguais. O fato de usar o mesmo número chega a criar padrões inexistentes na imagem original e gerar distorções. Uma outra ideia consiste em controlar, de forma diferente, o número de erosões e de dilatações. Neste caso, a imagem original é reconstruída. Detetar picos e/ou vales pode ser realizado por comparação entre a imagem original e a imagem correspondente reconstruída. Por Tophat por reconstrução, entendemos o controle diferenciado do número de iterações de cada operação.



### 3.1 TOPHAT POR RECONSTRUÇÃO POR ABERTURA

Uma técnica *controlável* consiste em reconstruir a erosão controlada do sinal  $f$  [Dougherty-1992]. A detecção dos picos do sinal  $f$  é então a operação  $A_r(f)$  definida, através de abertura por reconstrução cinza do sinal  $f$  *ero g*, da seguinte forma:

$$A_r(f) = f - \text{abe}_f^{n_{\max}g}(f)$$

Esta equação significa que o Tophat por reconstrução por abertura consiste em subtrair à imagem inicial a imagem reconstruída por abertura. A imagem reconstruída provém da convergência da dilatação geodésica por  $g$  do sinal erodido  $f$  *ero g*. A dilatação geodésica segue o critério de estabilidade  $n_{\max}$  que representa a última dilatação ( $n_{\max}$ ) geodésica para a qual o sinal processado permaneceu abaixo do sinal original  $f$ .

A figura 2-a) ilustra o processamento de uma imagem de placa de circuito impresso, de origem ruidosa, degradada pela adição de uma mancha. Desejando recuperar as trilhas e as ilhas da placa que representam uma informação do tipo "pic", foi usado um Tophat por reconstrução por abertura. Após erosão, com 2 iterações com um elemento estruturante 3x3 em cruz, o processo convergiu com 1 iteração de dilatação, a partir de um elemento estruturante 3x3 em cruz. Durante este processo, as regiões heterogêneas permaneceram enquanto a informação relevante sumiu. A etapa de subtração, ao mesmo tempo que suprimiu as regiões heterogêneas, devolveu trilhas e as ilhas mais nítidas (figura 2-b).

### 3.2 TOPHAT POR RECONSTRUÇÃO POR FECHAMENTO

Da mesma maneira, podemos detectar os vales do sinal  $f$  em termos de fechamento por reconstrução da dilatação do sinal  $f$ . Esta operação  $F_r(f)$  define-se da seguinte forma [Dougherty-1992]:

$$F_r(f) = \text{fec}_f^{n_{\max}g}(f) - f$$

O Tophat por reconstrução por fechamento consiste em subtrair à imagem reconstruída por fechamento a imagem inicial. A imagem reconstruída provém da convergência da erosão geodésica por  $g$  do sinal dilatado  $f$  *dil g*. A erosão geodésica segue o critério de estabilidade  $n_{\max}$  que representa a última erosão ( $n_{\max}$ ) geodésica para a qual o sinal processado permaneceu abaixo do sinal original  $f$ .

A figura 2-c) ilustra o processamento de uma imagem microscópica de células degradada pela adição de uma mancha. Desejando recuperar as células

que representam uma informação do tipo "vale", foi empregado um Tophat por reconstrução por fechamento. Após a dilatação iterada 3 vezes, com um elemento estruturante 3x3 em cruz, o processo de erosão, com um elemento estruturante 3x3 em cruz, convergiu com 1 iteração. As regiões heterogêneas permaneceram enquanto a informação relevante sumiu. Por subtração, as regiões heterogêneas desapareceram e as células permaneceram mais nítidas (figura 2-d).

## 4 OPERADORES Hdome E Hbasin

Nos processos anteriores de Tophat por reconstrução, o primeiro passo consistia em controlar, através da escolha do elemento estruturante, a reconstrução da imagem. Uma outra técnica de detecção de picos e/ou vales consiste em modificar a intensidade da imagem original e reconstruí-la controlando o número de iterações.

### 4.1 OPERADOR Hdome

Seja um sinal  $f$  onde queremos detectar e localizar os picos. Cada ponto deste sinal  $f$  pode ter o seu nível de cinza diminuído de um certo valor  $v$ . Diminuindo o sinal inteiro  $f$  pelo valor  $v$ , definimos uma operação notada  $(f - v)$ .

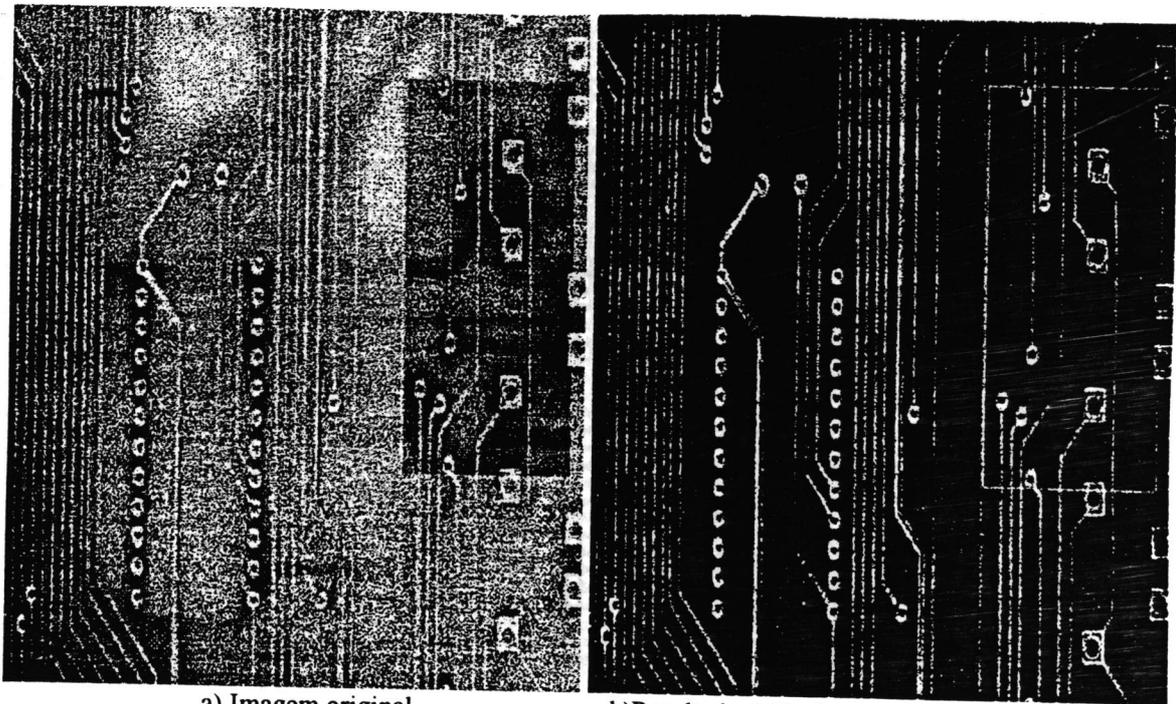
A detecção dos picos do sinal  $f$  pode ser efetuada pela operação  $D_v(f)$  definida, em termos de dilatação geodésica cinza do sinal diminuído  $(f - v)$ , da seguinte forma [Vincent-1993]:

$$D_v(f) = f - \text{dil}_f^{n_{\max}g}(f - v)$$

Neste caso, a dilatação geodésica segue o critério de estabilidade  $n_{\max}$  que representa, como no caso da abertura por reconstrução cinza, a última dilatação ( $n_{\max}$ ) geodésica para a qual o sinal processado permaneceu abaixo do sinal original  $f$ .

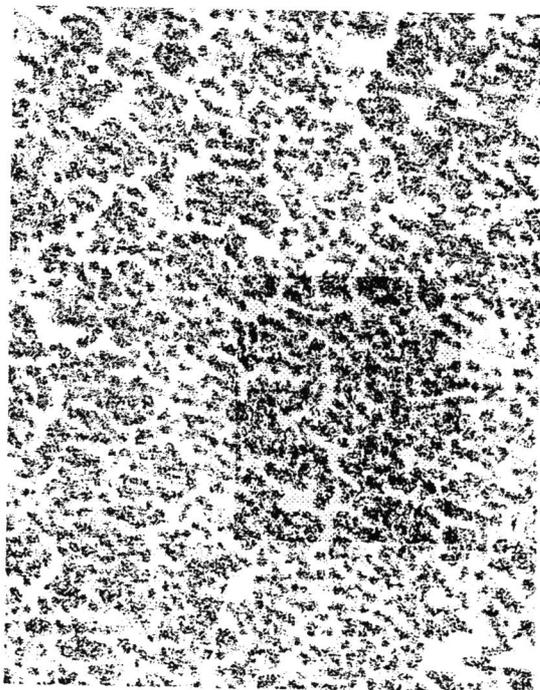
Esta operação  $D_v(f)$  é possível pelo fato que o sinal  $f - v$  está sempre abaixo do sinal  $f$ , condição indispensável para a reconstrução.

A figura 3-a) ilustra o processamento de uma imagem de motivos regulares onde foi colocada uma mancha. Desejando recuperar os motivos regulares que representam uma informação do tipo "pico", foi empregado um Tophat do tipo Hdome. Após ter diminuído o nível de cinza de cada pixel da imagem do valor 10, o processo convergiu com 2 iterações de dilatação. Na subtração entre os sinais original e aberto (figura 3-b)), eliminamos a mancha e recuperamos os motivos (picos).

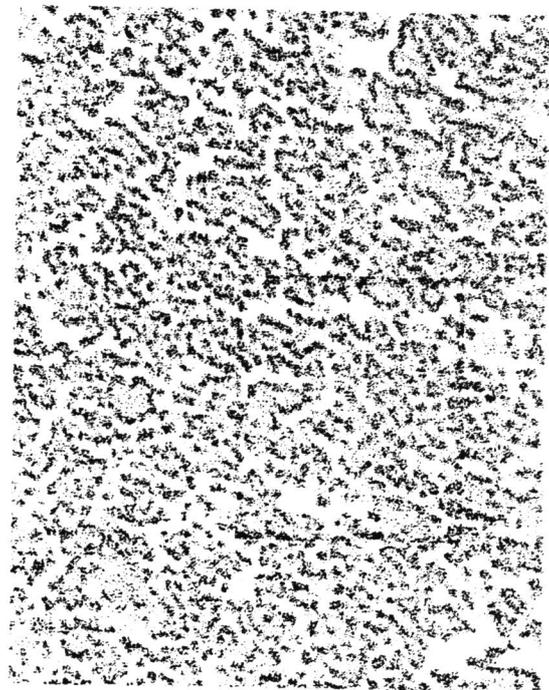


a) Imagem original

b) Resultado do Tophat por reconstrução por Abertura



c) Imagem original



d) Resultado do Tophat por reconstrução por Fechamento

Figure 2: Tophats por reconstrução por abertura e fechamento

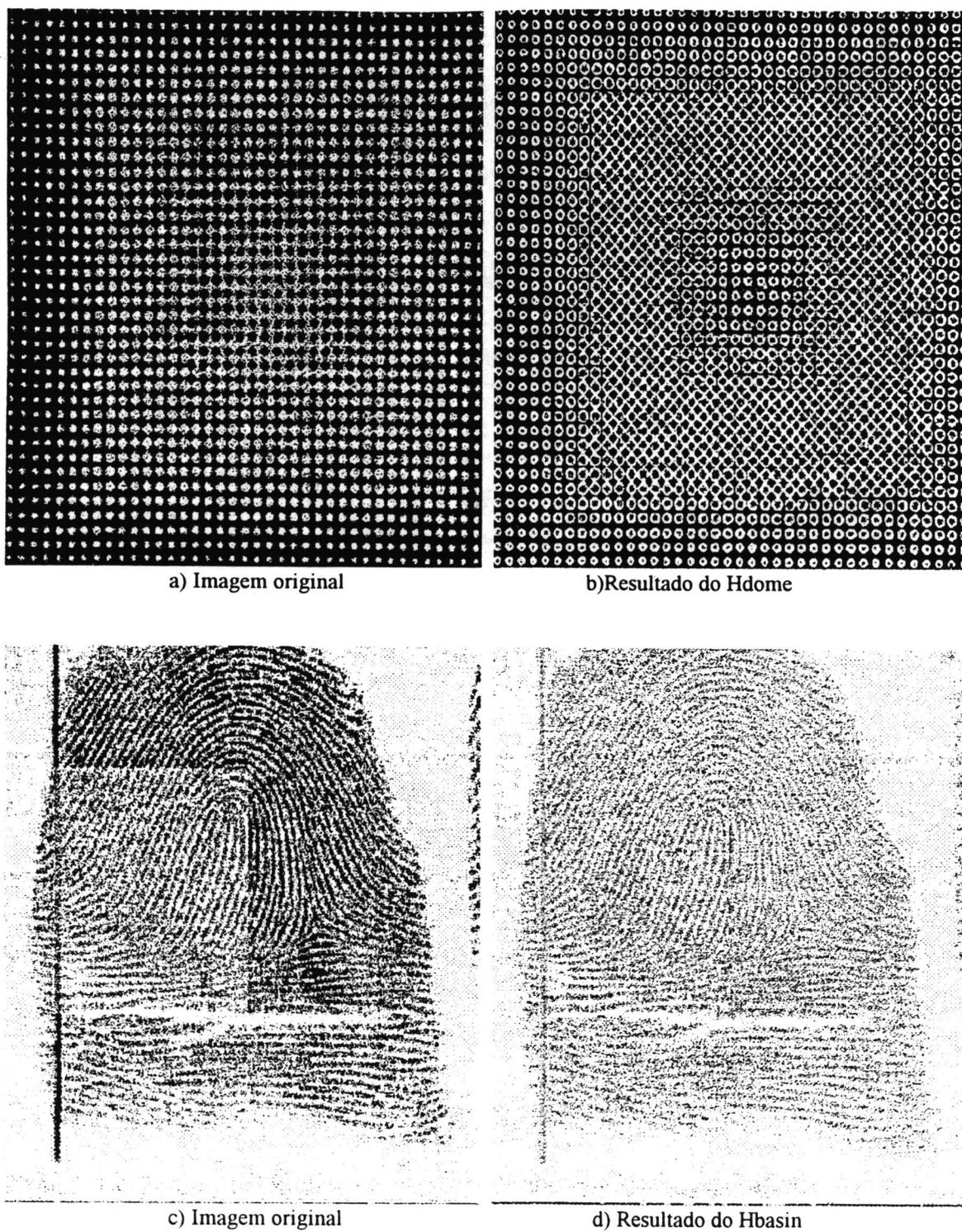


Figure 3: Tophats tipo Hdome e Hbasin

#### 4.2 OPERADOR Hbasin

Da mesma maneira, queremos detectar e localizar os vales do sinal  $f$ . O nível de cinza de cada ponto deste sinal  $f$  pode ser aumentado de um certo valor  $v$ . Aumentando o sinal inteiro  $f$  do valor  $v$ , definimos uma operação notada  $(f + v)$ .

A detecção dos vales do sinal  $f$  pode ser efetuada pela operação  $B_v(f)$  definida, através da erosão geodésica cinza do sinal diminuído  $(f - v)$ , da seguinte forma [Vincent-1993]:

$$B_v(f) = \text{ero}_f^{n_{\max}}(f + v) - f$$

Neste caso, a erosão geodésica segue o critério de estabilidade  $n_{\max}$  que representa, como no caso do fechamento por reconstrução cinza, a última erosão ( $n_{\max}$ ) geodésica para a qual o sinal processado permaneceu acima do sinal original  $f$ .

Esta operação  $B_v(f)$  é possível pelo fato que o sinal  $f + v$  está sempre acima do sinal  $f$ , condição indispensável para a reconstrução.

A figura 3-c) ilustra o processamento de uma imagem de uma impressão digital degradada pela adição de uma mancha. Desejando recuperar as linhas que representam uma informação do tipo "vale", foi empregado um Tophat do tipo Hbasin. Após ter aumentado o nível de cinza de cada pixel da imagem do valor 30, o processo convergiu com 1 iteração de erosão. Portanto, na subtração entre os sinais fechado e original, as linhas (vales) foram ressaltadas (figura 3-d)).

#### 4.3 CONCLUSÕES

Foram apresentadas várias ferramentas de detecção de picos e vales. Através de exemplos, mostramos a eficiência dos processamentos. Quais são as características que podem influir na escolha destes operadores? Um primeiro critério consiste em verificar a "automatizabilidade" dos processos.

O Tophat clássico consiste, em uma primeira etapa, em escolher o elemento estruturante certo que elimina a informação relevante para, na segunda etapa, regenerá-la por subtração. Esta noção de certeza é representada, em realidade, pelo tamanho do elemento estruturante.

No Tophat por reconstrução, o tamanho (ou o número de iterações) do elemento estruturante é escolhido no primeiro processo (erosão ou dilatação clássica). Uma boa detecção provém da noção de convergência da etapa de reconstrução. Neste caso, o que caracteriza a noção de certeza é também o tamanho do elemento estruturante.

No caso do Hdome e Hbasin, a etapa de reconstrução não é mais completamente condicionada pelo

tamanho do elemento estruturante, mas principalmente pelo aumento ou diminuição da intensidade da imagem do valor  $v$ . Isto faz com que a noção de ter-se usado a operação certa dependa da alteração da intensidade.

Destas observações, podemos concluir que enquanto os Tophats clássicos e por reconstrução extraem estruturas pelo tamanho, os operadores Hdome e Hbasin não envolvem nenhuma noção de tamanho e de forma. A detecção efetua-se através do critério "altura" destas estruturas que será no máximo o valor  $v$ .

Uma outra característica que influi na escolha destas ferramentas reside no controle ou na liberdade dos processos morfológicos empregados. O Tophat clássico tem, através do tamanho do elemento estruturante (número de iterações), o total controle das operações, que ou seja na primeira etapa (erosão ou dilatação clássica) ou na segunda etapa (respectivamente dilatação ou erosão).

O Tophat por reconstrução e os processos Hdome e Hbasin têm mais "liberdade" no sentido que o controle das operações é realizado somente na primeira etapa (erosão, dilatação, alteração da intensidade). Na segunda etapa (respectivamente dilatação ou erosão por reconstrução), o tamanho do elemento estruturante (número de iterações) será função da convergência.

A vantagem do Tophat por reconstrução em relação a Hdome e Hbasin reside no controle do aumento (diminuição) do sinal  $f$  a ser reconstruído. Enquanto as operações Hdome e Hbasin exigem a alteração arbitrária da intensidade do sinal  $f$  com um valor geralmente heurístico, os operadores do Tophat por reconstrução realizam um controle mais automatizável desta alteração por erosão ou dilatação.

#### 5 REFERÊNCIAS

- [Dougherty-1992], E. R. Dougherty, "An Introduction to Morphological Image Processing", Volume TT9, SPIE Press, 1992
- [Serra-1982], J. Serra, "Image analysis and mathematical morphology", Academic Press, 1982
- [Vincent-1993], L. Vincent, "Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms", IEEE Transactions on Image Processing, vol 2, No 2, pp 176-201, Abril 1993