

Teoria espaço-escala: abordagem morfológica

Leyza Baldo Dorini ¹
Neucimar Jerônimo Leite ²

Resumo: Abordagens multi-escala vêm sendo amplamente utilizadas em diversas aplicações de análise e processamento de sinais, sendo fundamentais em casos onde não existem informações preliminares sobre a escala de observação apropriada. A idéia básica consiste em criar uma família de sinais derivados, permitindo assim a análise de diferentes níveis de representação para escolha daqueles que exibem as características de interesse. A teoria espaço-escala é uma destas abordagens. A partir dela, podem ser estabelecidas as condições necessárias para a definição de transformações que possibilitem a manipulação de características presentes em diferentes níveis de maneira consistente. Este trabalho traz uma breve revisão das principais abordagens espaço-escala utilizadas na área de processamento de imagens, bem como de suas propriedades fundamentais.

1 Introdução

Ao se extrair características de interesse de uma imagem, é de extrema importância considerar que determinadas informações apenas fazem sentido sob certas condições de observação. Um exemplo típico é a diferença entre a análise do galho de uma árvore e da floresta a qual ela pertence. Enquanto o primeiro deve ser observado de uma distância variando de alguns centímetros a alguns metros, a floresta deverá ser observada a alguns quilômetros de distância, por exemplo. Este fato está relacionado ao conceito de escala.

A uma imagem digital pode-se associar, em princípio, as escalas externa e interna, correspondendo ao seu tamanho e à sua densidade de amostragem, respectivamente. Entretanto, no contexto deste trabalho, o termo escala refere-se aos diferentes tipos de estruturas presentes em uma imagem. Em outras palavras, dependendo da relação entre o tamanho de tais estruturas e o tamanho (e forma de aplicação) dos operadores utilizados nas transformações, diferentes informações se tornam explícitas.

Contudo, a escolha da escala de observação mais adequada não é uma tarefa trivial. Embora ela seja conhecida em certas situações específicas, como no caso de uma intervenção por parte de um especialista, por exemplo, isso não é possível na maioria dos demais casos.

Este problema motivou o desenvolvimento das abordagens multi-escala, amplamente utilizadas em diversas aplicações de análise e processamento de sinais (incluindo, portanto,

¹Departamento de Informática, UTFPR

²Instituto de Computação, Unicamp

{leyza@dainf.ct.utfpr.edu.br, neucimar@ic.unicamp.br}

imagens). A idéia básica consiste em criar uma família de sinais (imagens, no caso deste trabalho) derivados cujas estruturas são sucessivamente eliminadas, definindo assim um conjunto de sinais simplificados. Desta forma, é possível analisar os diferentes níveis de representação e utilizar aqueles que exibem apenas as características de interesse (Figura 1).

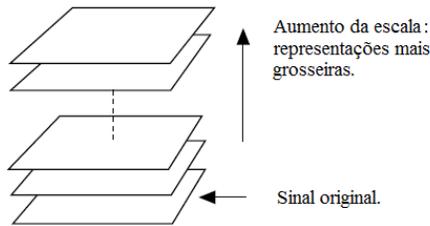


Figura 1. Representação multi-escala de um sinal.

Os principais problemas referentes ao uso de métodos multi-escala se devem a dificuldade de relacionar informações significativas do sinal através dos diferentes níveis. Em [27], Witkin propôs uma nova abordagem, denominada espaço-escala, na qual uma característica de interesse descreve um caminho contínuo na representação, possibilitando que estruturas da imagem presentes em diferentes escalas possam ser manipuladas de maneira consistente. Assim, é possível relacionar informações obtidas em diferentes níveis de observação, bem como determinar sua localização precisa no sinal original.

Um questionamento que pode surgir é por que representar em múltiplos níveis uma informação que já está presente no sinal original. Uma das principais razões é que a simplificação sucessiva remove detalhes indesejados, tais como ruído ou estruturas não-significativas, facilitando o processamento de tarefas futuras. Além disso, por lidar explicitamente com o parâmetro de escala, esta abordagem possibilita manipular apenas das estruturas significativas da imagem que se tornam explícitas em cada nível. Aliado ao fato que o aumento da escala está relacionado a redução da quantidade de informação, isto também implica em um aumento da eficiência computacional.

Desde a introdução da teoria espaço-escala, diversas formulações têm sido propostas. Este artigo faz uma breve revisão das principais abordagens e propriedades [1, 6, 9]. A Seção 2 apresenta as idéias básicas da teoria e a Seção 3 discute a formulação linear inicial proposta por Witkin, assim como os axiomas que determinam as condições necessárias para constituição de uma abordagem espaço-escala. Na Seção 4 é discutida a extensão para o caso bidimensional e na Seção 5 são introduzidas as abordagens não-lineares. Dentre estas, destacam-se os espaços-escala baseados em morfologia matemática, discutidos na Seção 6 e foco principal deste trabalho.

2 Visão geral da teoria espaço-escala

A teoria espaço-escala estabelece as condições necessárias para a definição de transformações que possibilitem a manipulação de características presentes em diferentes níveis de representação de maneira consistente. Uma propriedade fundamental é que a transformação para um nível mais grosseiro não introduza novas estruturas, ou seja, características pertencentes em uma determinada escala também estão presentes em todas as escalas mais finas (ou seja, nas escalas mais baixas).

Tal propriedade conduz a uma simplificação do sinal (da imagem) original, considerando que a quantidade de características diminui e que as remanescentes em um dado nível correspondem a simplificações daquelas existentes inicialmente. Contudo, não é necessário que suas localizações sejam exatamente as mesmas através das escalas. É preciso apenas que elas definam caminhos contínuos, ou seja, basta que seja possível rastrear a localização de uma característica em cada escala. Desse modo, características em diferentes níveis estão relacionadas se elas pertencem ao mesmo caminho no espaço-escala resultante.

Outra propriedade exigida para abordagens espaço-escala é a fidelidade [1]. Sejam f a imagem original, \mathbf{x} a localização do ponto sendo observado e $[T(\mathbf{x}, \sigma)](f)$ a representação da observação na escala σ . A fidelidade requer que a transformação espaço-escala convirja para o sinal original conforme σ tende a zero, isto é,

$$\lim_{\sigma \rightarrow 0} [T(\mathbf{x}, \sigma)](f) = f(\mathbf{x}). \quad (1)$$

Invariância Euclidiana também é desejada, ou seja, uma translação e/ou rotação da imagem original deve implicar em um espaço-escala igualmente transladado/rotacionado.

Observe que os requisitos para definição de um espaço-escala conduzem naturalmente a uma simplificação do sinal em consideração. As propriedades de continuidade e “não introdução” de novas características permitem a sua manipulação consistente, enquanto a fidelidade indica que a filtragem inicia no sinal original e fica mais intensa com o aumento da escala. Além disso, de acordo com a condição de invariância Euclidiana, o procedimento de filtragem não deve depender de nenhum conhecimento preliminar sobre a orientação ou o posicionamento do sinal utilizado.

Ao contrário de outras abordagens multi-escala, tais como pirâmides, a representação espaço-escala preserva a mesma resolução espacial (mesma amostragem) em todos os níveis. Isso permite o acesso imediato aos dados de interesse em qualquer escala sem a necessidade de processamento adicional. Aliado à propriedade de continuidade, esse fato permite rastrear de forma direta as características de interesse através da representação, bem como determinar sua localização exata no sinal original. Embora possuam propriedades em comum, as abordagens espaço-escala diferem na forma com que geram as representações. A seguir, serão discutidas algumas delas.

3 Abordagem linear

Na formulação inicial, desenvolvida para sinais contínuos unidimensionais, uma família de sinais derivados é gerada através da convolução do sinal original com um núcleo gaussiano de média zero, cujo desvio padrão representa o conceito de escala. Seja $f(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$ o sinal original e $g(\mathbf{x}, \sigma) : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ um núcleo gaussiano (que possui propriedades de suavização). A imagem espaço-escala na escala σ é dada por $T : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$:

$$[T(\mathbf{x}, \sigma)](f) = f(\mathbf{x}) * g(\mathbf{x}, \sigma), \quad (2)$$

em que $*$ denota a operação de convolução e $g(\mathbf{x}, \sigma)$ é o núcleo gaussiano [11]:

$$g(\mathbf{x}, \sigma) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\mathbf{x}^T\mathbf{x}\right). \quad (3)$$

Tal formulação, usualmente denominada espaço-escala gaussiano, consiste na forma canônica para gerar um espaço-escala linear [8, 27]. Koenderink [13] foi o primeiro a observar a sua estrutura diferencial, mostrando que a gaussiana é a função de Green (ou propagador) da equação de difusão. Isso permite que métodos baseados em Equações Diferenciais Parciais (EDPs) possam ser utilizados para gerar esta representação. De fato, as soluções para a equação do calor (ou equação de difusão isotrópica):

$$\frac{\partial T(\mathbf{x}, t)}{\partial t} = \Delta T(\mathbf{x}, t), \quad (4)$$

em que Δ é o Laplaciano, são amostras do espaço-escala gaussiano para um tempo fixo t [11].

A princípio, o projeto de um espaço-escala linear pode parecer arbitrário, levando a crer que qualquer núcleo de suavização conduz a uma representação válida. No entanto, conforme discutido anteriormente, tal procedimento deve satisfazer a um conjunto de propriedades. As condições para o caso linear unidimensional, denominadas axiomas da teoria espaço-escala, são sumarizadas abaixo [17, 19].

1. **Não-criação de novos extremos:** Witkin [27] observou que na representação espaço-escala de sinais unidimensionais contínuos o número de cruzamentos por zero na derivada segunda do sinal diminui com o aumento da escala. Tal propriedade é denominada monotonicidade, considerando que o número de características deve ser um função monotonicamente decrescente em função da escala. Além disso, como a diferenciação comuta com a convolução, o número de extremos em uma derivada espacial de qualquer ordem, calculada a partir da representação espaço-escala, também diminui para escalas crescentes.

As Figuras 2(b)-(e) ilustram o espaço-escala resultante da convolução do sinal unidimensional discreto mostrado na Figura 2(a) com um núcleo gaussiano utilizando desvios padrões crescentes. Pode-se verificar que, com o aumento da escala, os detalhes menos significativos são sucessivamente eliminados levando, conseqüentemente, a uma simplificação progressiva do sinal.

A Figura 2(f) ilustra o gráfico denominado *fingerprint*, o qual mostra a posição das características de interesse (as quais podem variar segundo a abordagem utilizada) em cada nível da representação espaço-escala (os níveis são representados no eixo y). Para facilitar a visualização, considere que tais características equivalem aos máximos locais (pontos que possuem valor maior que todos os seus vizinhos). Observe que os máximos mais significativos persistem até as escalas mais grosseiras e que, uma vez eliminados, eles não tornam a aparecer. Note também que os caminhos definidos através dos diferentes níveis são contínuos e novas estruturas não são criadas, refletindo as propriedades de continuidade e monotonicidade.

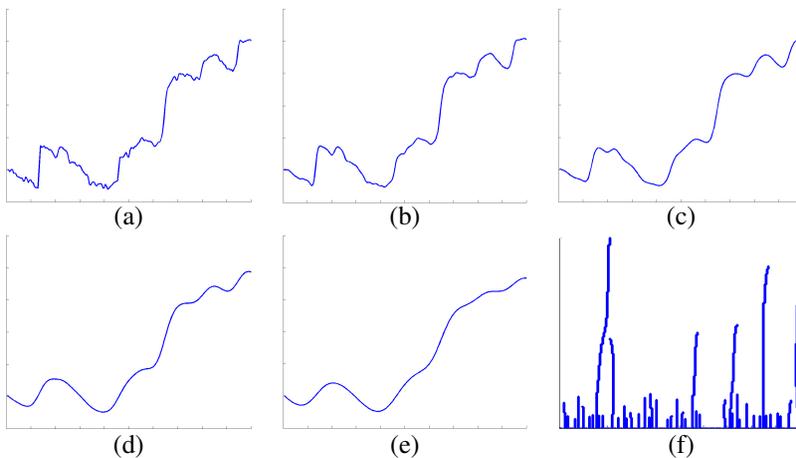


Figura 2. Exemplo de espaço-escala gaussiano: (a) sinal unidimensional aleatório, (b)-(e) representação espaço-escala para escalas crescentes e (f) *fingerprint* considerando máximos locais como características.

- Linearidade e invariância à translação, rotação e escala:** A invariância à rotação e translação garante que todas as posições da imagem são tratadas de forma idêntica, ou seja, a representação espaço-escala não depende de nenhum conhecimento preliminar sobre a posição do sinal original. Isso implica que uma rotação e/ou translação do sinal original deve resultar em um espaço-escala igualmente rotacionado e/ou transladado.

Quando combinadas à linearidade, tais condições permitem o uso de núcleos de convolução como operadores para gerar a representação espaço-escala, trazendo benefícios tais como a implementação eficiente. A propriedade de invariância à escala assegura que núcleos de diferentes escalas são equivalentes em termos qualitativos.

3. **Estrutura de semi-grupo e separabilidade:** a estrutura de semi-grupo garante que o resultado da convolução de dois núcleos manterá a mesma estrutura e que

$$g(\mathbf{x}, \sigma_1) * g(\mathbf{x}, \sigma_2) = g(\mathbf{x}, \sigma_1 + \sigma_2), \quad (5)$$

em que o símbolo $*$ denota a convolução. Tais propriedades permitem que uma representação espaço-escala possa ser construída incrementalmente, ou seja, se uma imagem na escala σ_1 for utilizada como entrada para uma observação em uma escala mais grossa σ_2 , isso é equivalente a realizar uma observação diretamente no nível $\sigma_1 + \sigma_2$. Isso conduz a uma interpretação matemática muito precisa, onde o nível $\sigma_1 + \sigma_2$ pode ser analisado independentemente da escala inicial do sinal [7].

A propriedade de separabilidade, por sua vez, possibilita uma implementação eficiente através da utilização de duas convoluções unidimensionais ao invés de uma única bidimensional.

4. **Não-realce de extremos locais e gerador infinitesimal:** Quando combinadas, as propriedades de estrutura de semi-grupo e continuidade em relação à escala conduzem a um resultado indicando que a família espaço-escala deve ter um gerador infinitesimal. Isso significa que, sob determinadas condições de regularidade, existe um caso limite do operador, denotado aqui por \mathcal{A} :

$$\mathcal{A} = \lim_{\sigma \rightarrow 0} \frac{[T(\mathbf{x}, \sigma)](f) - f}{\sigma}. \quad (6)$$

Aliada à propriedade de não-realce de extremos locais (assim como o próprio nome indica, esta propriedade requer que o valor de um máximo (mínimo) local não aumente (diminua) com o aumento da escala), tal condição implica que a representação espaço-escala satisfaz a equação de difusão.

Pode-se provar que a única família que satisfaz estes axiomas é a dos núcleos gaussianos, o que torna a abordagem descrita nesta seção a forma canônica de gerar um espaço-escala linear. Além disso, embora os resultados apresentados acima estejam relacionados a sinais contínuos, eles também são válidos para o caso discreto [17]. No entanto, sua extensão para dimensões maiores que um não é possível, conforme discutido a seguir.

4 Espaços-escala em duas dimensões

É possível provar que, para dimensões maiores que um, qualquer núcleo de convolução utilizado para obter a representação espaço-escala introduz novas características com o aumento da escala [16]. Em outras palavras, não existem núcleos triviais em \mathbb{R}^2 ou \mathbb{Z}^2 para os quais a propriedade de monotonicidade seja preservada no caso linear.

Um exemplo que ilustra claramente este fato foi apresentado por Lifshitz and Pizer [16] e é ilustrado na Figura 3. As características de interesse equivalem aos máximos locais. Considere uma imagem composta por dois picos, um deles mais alto que o outro e ambos possuindo uma forma suave bem definida. Eles são conectados de tal forma que a imagem possui um único ponto de máximo, situado no pico de maior altitude (Figura 3(a)). Ao gerar o espaço-escala gaussiano, a região entre eles irá erodir mais rapidamente, fazendo com que o pico mais baixo se torne um máximo local - o que viola a propriedade de monotonicidade (Figura 3(b)).

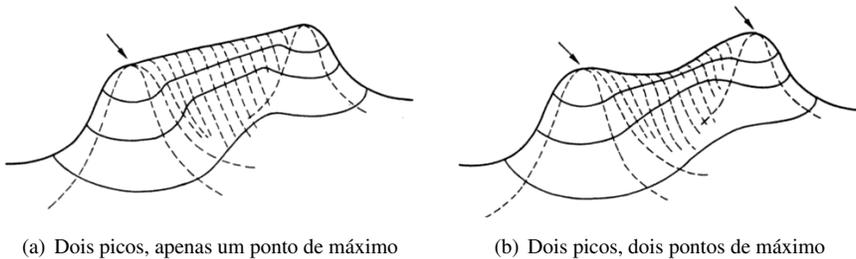


Figura 3. Exemplo ilustrando a criação de um novo máximo local ao gerar um espaço-escala gaussiano em duas dimensões [19].

Sendo assim, para dimensões maiores que um, é preciso considerar um conjunto de condições menos restritivas para constituição de um espaço-escala. Lindenberg [18] observou que, como não é possível evitar que o número de extremos locais aumente, a idéia é restringir o número de criações a poucas ocorrências.

Koenderink [13], por sua vez, definiu uma extensão da propriedade de monotonicidade para duas dimensões, denominada causalidade, segundo a qual novas superfícies (curvas) de nível não devem ser criadas com o aumento da escala. No caso discreto, em que não há uma correspondência direta para curvas de nível, isso implica que os extremos locais devem satisfazer a propriedade de não-realce, segundo a qual o valor de um máximo não deve aumentar (e o de um mínimo não deve diminuir) para escalas crescentes. Observe que a causalidade não impõe condições sobre a criação de novos extremos [4].

É possível provar que as propriedades de homogeneidade, isotropia (as quais basicamente implicam que todos os pontos em diferentes níveis de escala devem ser tratados de forma similar) e causalidade constituem as condições necessárias e suficientes para uma formulação em termos da equação de difusão nos casos uni- e bidimensional [17].

Outro problema relacionado à família de núcleos gaussianos é o borrimento das imagens derivadas, uma limitação para tarefas de mais alto nível que exigem a localização exata dos contornos, tais como reconhecimento de objetos. Este fato pode ser visto como um exemplo do princípio da incerteza, pois o núcleo gaussiano (que constitui um filtro passa-baixa) restringe o espectro de frequência, causando o borrimento no domínio espacial [1]. A Figura 4 ilustra esse fato.

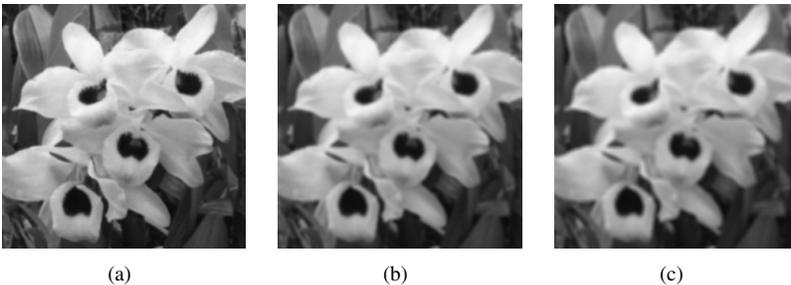


Figura 4. Espaço-escala gaussiano bidimensional utilizando as escalas (desvio padrão do núcleo): (a) 1, (b) 2 e (c) 4.

5 Abordagens não-lineares

De maneira geral, representações espaço-escala não-lineares são geradas com base em transformações de morfologia matemática ou EDPs não-lineares. Uma abordagem amplamente utilizada é a técnica de difusão anisotrópica proposta por Perona e Malik [23], a qual é capaz de reduzir o ruído ao mesmo tempo em que busca preservar contornos e minimiza, assim, o problema de borrimento relacionado ao caso gaussiano.

A idéia principal consiste em diminuir a difusão na direção do gradiente de mais alta magnitude. Desta forma, dependendo do valor de tal magnitude, o operador constitui um filtro passa-altas (em regiões de contorno) ou uma difusão isotrópica (em regiões homogêneas). Além disso, este método preserva a propriedade de monotonicidade para extremos do sinal quando coeficientes de difusão apropriados são escolhidos.

Os autores consideram a seguinte EDP:

$$\begin{cases} \frac{\partial I}{\partial \sigma} = \text{div}[c(|\nabla I|) \cdot \nabla I], \\ I(\sigma = 0) = I_0, \end{cases} \quad (7)$$

em que ∇ é o operador gradiente, div o divergente, $|\cdot|$ denota a magnitude, $c(x)$ é o coeficiente de difusão e I_0 a imagem (ou condição) inicial. Os autores sugerem duas diferentes formas para calcular os coeficientes de difusão:

$$c_1(x) = \frac{1}{1 + (x/k)^2} \quad \text{e} \quad c_2(x) = \exp[-(x/k)^2], \quad (8)$$

em que k é um parâmetro que representa a magnitude do contorno. A primeira equação favorece regiões mais amplas em relação às menores, enquanto a segunda favorece regiões de alto contraste. A Figura 5 mostra as imagens resultantes da aplicação da técnica de difusão anisotrópica para diferentes valores do coeficiente c_1 .

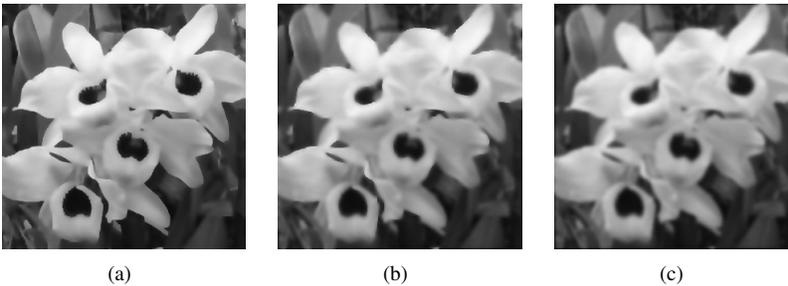


Figura 5. Difusão anisotrópica de uma imagem considerando diferentes coeficientes de difusão: $c_1 = 20$, $c_1 = 50$ e $c_1 = 100$.

Outra abordagem não-linear que obedece ao princípio da causalidade consiste na definição de uma classe especial de espaços-escala que podem ser linearizados através de um isomorfismo no domínio espacial reproduzindo, desta forma, o espaço-escala gaussiano. Tal abordagem, que conduz aos denominados espaços-escala pseudo-lineares, possui aplicações em imagens escalares, vetoriais e multi-espectrais, e pode ser utilizada inclusive em casos em que o espaço-escala gaussiano apresenta inconsistências [4, 5].

Todavia, de forma geral, as técnicas não-lineares mencionadas possuem um custo computacional alto e demandam diversos parâmetros. As abordagens baseadas em morfologia matemática minimizam estes problemas e têm se mostrado eficientes gerando, como abordado a seguir, espaços-escala com interessantes propriedades.

6 Espaços-escala morfológicos

Desde a sua formalização por Matheron e Serra [20, 24, 25], a morfologia matemática vem sendo amplamente utilizada em diversas aplicações de processamento de imagens. No contexto da teoria do espaço-escala, combinações das operações de erosão e dilatação utilizando funções estruturantes não-planares (definidas na próxima seção) têm se mostrado eficientes na definição de diversos espaços-escala não-lineares. A seguir, apresentamos algumas definições básicas relacionadas a morfologia matemática. Na sequência, são discutidas algumas abordagens propostas na literatura.

6.1 Definições básicas de Morfologia Matemática

Para se extrair informações relativas aos objetos de uma imagem de forma precisa, é fundamental determinar adequadamente a vizinhança que será considerada para calcular o valor transformado de cada pixel. Isso depende da escolha apropriada de um elemento de base - um elemento estruturante no caso binário, ou função estruturante para imagens em níveis de cinza - o qual, de maneira simplificada, define o tamanho e a forma da região a ser considerada por uma transformação morfológica.

Tais elementos podem ser classificadas em duas categorias: planares, onde todas as posições possuem o valor zero, ou não-planares, quando os valores associados podem ser distintos, definindo uma altura para a função estruturante. Tais valores podem ser interpretados como “pesos” dados às diferentes direções consideradas [11].

Embora amplamente utilizadas, devido principalmente à facilidade de implementação e resultados satisfatórios em muitos casos, as funções estruturantes planares não conseguem extrair algumas propriedades específicas das imagens. Por exemplo, seu uso transforma as regiões em torno de máximos e mínimos locais em regiões planas, podendo causar o deslocamento da sua posição e constituindo um problema para diversas aplicações [10]. As funções estruturantes não-planares, por sua vez, também podem ser utilizadas na definição de transformações morfológicas dependentes de escala, as quais permitem a análise de diferentes níveis de representação, fator essencial para lidar adequadamente com a natureza multi-escala das imagens.

Para garantir um comportamento de escala condizente tais funções estruturantes, denotadas aqui por B_σ , devem satisfazer a determinadas condições [11], as quais exigem que B_σ seja uma função monotonicamente decrescente ao longo de qualquer direção radial a partir da origem (ou seja, anti-convexa). Para evitar deslocamento de níveis e translação horizontal, as seguintes condições também devem ser observadas [11]:

$$\sup_{\mathbf{t} \in \mathcal{B}_\sigma} \{B_\sigma(\mathbf{t})\} = 0 \text{ e } B_\sigma(\mathbf{0}) = 0. \quad (9)$$

A erosão e a dilatação são as operações fundamentais da morfologia matemática, constituindo a base para a definição de transformações mais complexas [6, 20, 25, 26]. Quando funções estruturantes não-planares são utilizadas, as operações de erosão e dilatação são definidas da seguinte forma:

Definição 6.1 [26] (Dilatação) A dilatação de uma imagem f pela função estruturante $B_\sigma(\mathbf{x})$, $[\delta_{B_\sigma}(f)](\mathbf{x})$, é dada por:

$$[\delta_{B_\sigma}(f)](\mathbf{x}) = \max_{\mathbf{b} \in B_\sigma} \{f(\mathbf{x} + \mathbf{b}) + B_\sigma(\mathbf{b})\}. \quad (10)$$

Definição 6.2 [26] (Erosão) A erosão de uma imagem f pela função estruturante $B_\sigma(\mathbf{x})$, $[\varepsilon_{B_\sigma}(f)](\mathbf{x})$, é dada por:

$$[\varepsilon_{B_\sigma}(f)](\mathbf{x}) = \min_{\mathbf{b} \in B_\sigma} \{f(\mathbf{x} + \mathbf{b}) - B_\sigma(\mathbf{b})\}. \quad (11)$$

A Figura 6 mostra as versões dilatada e erodida de uma imagem. Note que a dilatação enfatiza as regiões claras da imagem (pois toma como base o valor do máximo local na vizinhança do elemento estruturante), ao passo que a erosão enfatiza regiões escuras.

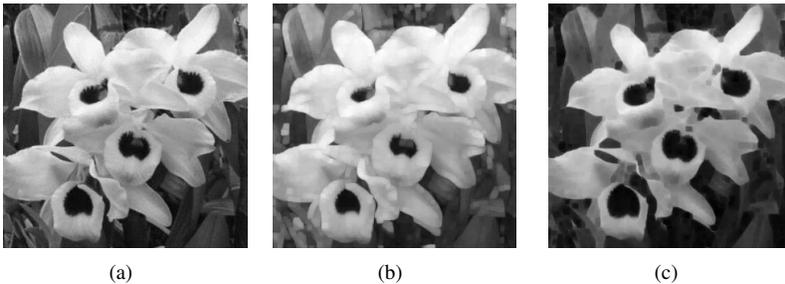


Figura 6. (a) Imagem original, (b) dilatação e (c) erosão de (a) utilizando um elemento estruturante quadrado de tamanho 3×3 .

A dilatação é uma transformação extensiva ($\delta_B(X) \supseteq X$ se o elemento estruturante contém sua origem) e a erosão é anti-extensiva, em que X denota a forma sendo utilizada. Além disso, as duas transformações são invariantes à translação e preservam a homotopia da imagem original. Ambas também são crescentes, implicando assim que, se $X \subseteq Y$, então $\delta_B(X) \subseteq \delta_B(Y)$ e $\varepsilon_B(X) \subseteq \varepsilon_B(Y)$. Pode-se também verificar facilmente que as operações de erosão e dilatação são duais, ou seja, $\delta_B(X) = [\varepsilon_B(X^c)]^c$, em que c denota o

complemento. Entretanto, elas não são auto-duais nem consistem em transformações inversas uma da outra, isto é, $X \neq \delta_B[\varepsilon_B(X)]$.

As operações de erosão e dilatação possuem algumas limitações mas, se combinadas de modo apropriado, resultam em transformações morfológicas mais robustas. A erosão, por exemplo, além de eliminar pequenas estruturas, também causa a redução de todas as outras. No entanto, se uma dilatação for aplicada na imagem erodida, parte das estruturas perdidas será recuperada. Tal procedimento, de aplicar uma erosão seguida de uma dilatação, define uma nova transformação denominada abertura [26]:

$$\gamma_B(f) = \delta_{\check{B}}[\varepsilon_B(f)], \quad (12)$$

em que \check{B} denota o transposto de B , $\check{B} = \{-b \mid b \in B\}$. O uso de \check{B} torna o resultado da transformação independente da posição de origem do elemento estruturante, mesmo quando este não é simétrico. O fechamento, que consiste em uma dilatação seguida de uma erosão, tende a recuperar de maneira aproximada a forma inicial das estruturas da imagem que foram dilatadas [26]:

$$\phi_B(f) = \varepsilon_{\check{B}}[\delta_B(f)], \quad (13)$$

em que, novamente, \check{B} denota o transposto de B .

As operações de abertura e fechamento são duais e ambas são crescentes e idempotentes. Além disso, a abertura é anti-extensiva e atua sobre as estruturas claras, enquanto que o fechamento é extensivo e em termos visuais atua sobre as estruturas escuras da imagem. Ambas não preservam a homotopia da imagem original e podem usar tanto funções estruturantes planares quanto não-planares.

Transformações crescentes e idempotentes caracterizam os denominados filtros morfológicos, uma abordagem não-linear que conduz tanto a recuperação de imagens com ruído quanto a remoção de estruturas ou objetos específicos de uma imagem segundo critérios tais como geometria ou contraste [26]. A seguir, são discutidos alguns espaços-escala gerados a partir de transformações de morfologia matemática.

6.2 Erosão/Dilatação Morfológica Multi-escala

Em [10, 11], Jackway introduziu um novo espaço-escala, denominado Erosão/Dilatação Morfológica Multi-escala (*Multiscale Morphological Dilation Erosion* - MMDE), o qual é baseado em um operador do tipo *toggle* e definido por:

$$(f \otimes B_\sigma)(\mathbf{x}) = \begin{cases} [\delta_{B_\sigma}(f)](\mathbf{x}), & \text{se } \sigma > 0, \\ f(x), & \text{se } \sigma = 0, \\ [\varepsilon_{B_\sigma}(f)](\mathbf{x}), & \text{se } \sigma < 0. \end{cases} \quad (14)$$

em que f é a imagem original e $[\delta_{B_\sigma}(f)](\mathbf{x})$ denota a dilatação e $[\varepsilon_{B_\sigma}(f)](\mathbf{x})$ a erosão do pixel na posição \mathbf{x} utilizando uma função estruturante dependente de escala B_σ . Observe que a regra de decisão analisa o sinal da escala para decidir se a imagem será dilatada ($\sigma > 0$) ou erodida ($\sigma < 0$).

A Figura 7 ilustra uma possível representação espaço-escala gerada pelo MMDE. Pode-se provar que as condições necessárias (inclusive monotonicidade) são satisfeitas assumindo-se como características de interesse os máximos locais quando a escala é positiva e os mínimos locais quando ela é negativa [1, 11]. Para tal, é essencial que as funções estruturantes satisfaçam as propriedades especificadas anteriormente na Seção 6.1.

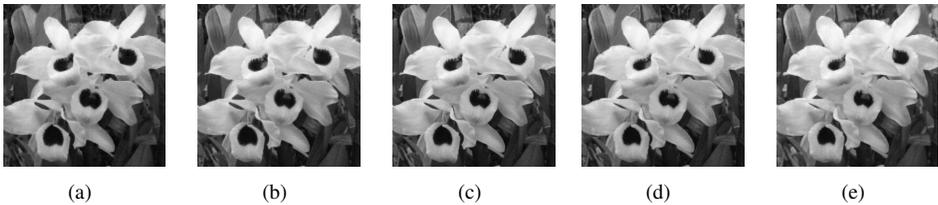


Figura 7. Representação do espaço-escala MMDE. Erosão nas escalas (a) -0.5 e (b) -0.05, (c) imagem original, dilatação nas escalas (d) 0.05 e (e) 0.5.

Contudo, o MMDE apresenta algumas desvantagens. Como as características de interesse são diferentes dependendo do sinal da escala, essa abordagem pode ser vista como dois espaços-escala distintos, onde a noção de escala “negativa” consiste apenas em uma notação conveniente para exibir a representação. Além disso, devido ao viés dos níveis de cinza inerente às transformações extensivas e anti-extensivas (o valor transformado de um pixel é exclusivamente maior ou menor que seu valor original, respectivamente), essa abordagem é sensível a ruído [1].

O operador proposto em [14, 15] explora a idempotência do MMDE, estabelecendo uma relação entre a função estruturante B_σ e os extremos que persistem em uma determinada escala. A fusão dos extremos através das diferentes escalas também é controlada, possibilitando a definição de marcadores que conduzem a bons resultados de segmentação. As principais propriedades do espaço-escala MMDE são preservadas para o caso idempotente [14].

A Figura 8 mostra uma segmentação obtida a partir deste espaço-escala idempotente para duas escalas.

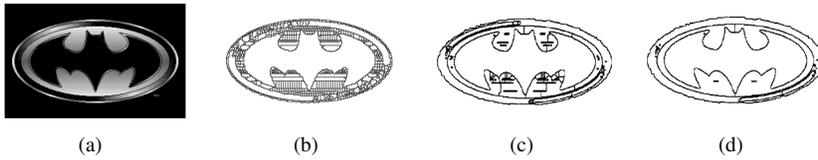


Figura 8. Resultado de segmentação. (a) imagem original, e resultados de segmentação obtidos com base (b) na imagem original e na representação espaço-escala idempotente para (c) $\sigma = 1/18$ e (d) $\sigma = 1/9$. Imagem extraída de [15].

6.3 Espaços-escala baseados em aberturas e fechamentos

A primeira abordagem proposta utilizava imagens binárias e consistia na aplicação da operação de abertura com um elemento estruturante circular de tamanho crescente [2]. As características de interesse eram definidas como os cruzamentos por zero existentes na curvatura do contorno dos objetos da imagem. Posteriormente, verificou-se que as condições necessárias e suficientes para garantir a monotonicidade eram que os elementos estruturantes fossem compactos e anti-convexos [12]. Por dualidade, as mesmas afirmações são válidas para a operação de fechamento [10].

No entanto, embora tais resultados possam ser estendidos para o caso unidimensional através da utilização de funções estruturantes compactas e anti-convexas, a monotonicidade não é válida para a filtragem bidimensional de imagens em níveis de cinza [22]. A Figura 9 ilustra um exemplo de espaço-escala baseado em abertura.

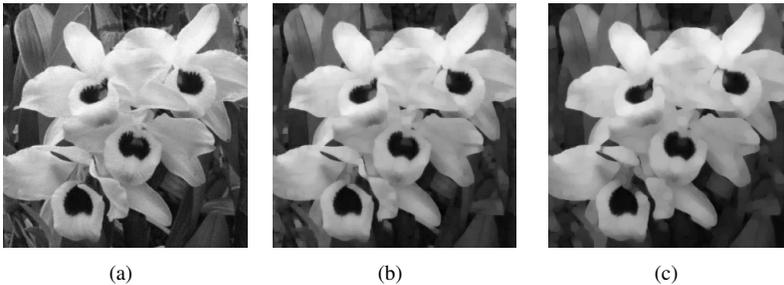


Figura 9. Representação do espaço-escala baseado na operação de abertura: (a) imagem original e abertura utilizando uma função estruturante compacta e anti-convexa de tamanhos (b) 3 e (c) 5.

Um sinal que é estritamente crescente (ou decrescente) em cada subintervalo de tamanho n é denominado um sinal localmente monotônico de grau n , ou lomo- n . O espaço-escala proposto em [1] calcula iterativamente a média entre a abertura e o fechamento de um pixel até chegar a um sinal raiz (as características de interesse são as regiões localmente monotônicas). Entretanto, também não preserva a propriedade de monotonicidade em duas dimensões.

Embora as operações de abertura e fechamento constituam filtros morfológicos (ambas são crescentes e idempotentes), o fato de elas não serem auto-duais pode torná-las sensíveis a ruído. Enquanto a abertura elimina apenas impulsos positivos, por exemplo, o fechamento tem este efeito para os impulsos negativos, o que constitui um viés na transformação.

Para reduzir este problema, podem ser utilizados filtros alternados sequenciais [26], ou seja, pode-se aplicar combinações das operações de abertura e fechamento com funções estruturantes de tamanho crescente de maneira recursiva. A Figura 10 ilustra um exemplo. Assim como no caso anterior, o conceito de escala está relacionado à dimensão (ao tamanho) da função estruturante.

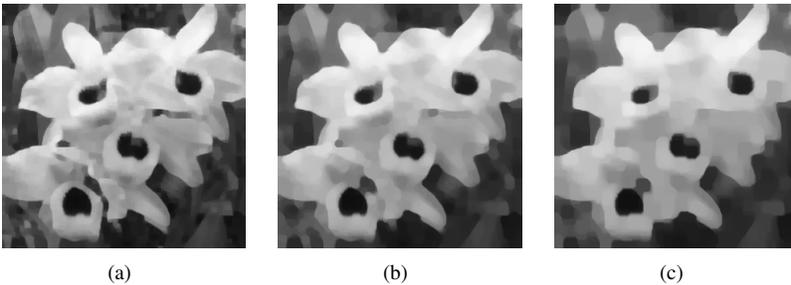


Figura 10. Espaço-escala baseado no filtro alternado sequencial $[\gamma\phi\phi\gamma]_B(f)$ utilizando funções estruturantes de tamanho (a) 3, (b) 5 e (c) 7.

6.4 Espaços-escala baseados em operadores conexos

Uma das grandes desvantagens das operações de abertura e fechamento é que elas causam o deslocamento dos contornos [7, 25]. Para evitar este problema, pode-se utilizar filtros por reconstrução, os quais recuperam apenas a estrutura das componentes que não forem eliminados em uma etapa inicial, preservando assim a forma original dos objetos [26]. Neste caso, os contornos são uma das possíveis características de interesse.

A abertura por reconstrução consiste em uma erosão seguida por uma reconstrução

morfológica por dilatação, ou seja [26]:

$$\gamma_R^{(n)}(f) = R_f[\varepsilon(f)], \quad (15)$$

em que $\varepsilon(f)$ denota a erosão de f com um elemento estruturante de tamanho n e

$$R_f([\varepsilon(f)]) = \delta_f^{(i)}([\varepsilon(f)]) \quad (16)$$

é a reconstrução morfológica por dilatação, em que $\delta_f^{(i)}([\varepsilon(f)])$ representa a i -ésima dilatação geodésica da imagem marcadora $[\varepsilon(f)]$ utilizando f como imagem máscara:

$$\delta_f^{(i)}([\varepsilon(f)]) = \delta^{(i)}([\varepsilon(f)]) \wedge f, \quad (17)$$

ou seja, é o mínimo pontual entre a imagem máscara e a i -ésima dilatação de tamanho n da imagem marcadora. O índice i representa a iteração onde a idempotência é atingida, isto é, $\delta^{(i)}([\varepsilon(f)]) = \delta^{(i+1)}([\varepsilon(f)])$.

Em resumo, pode-se afirmar que a reconstrução por dilatação de $[\varepsilon(f)]$ com relação a f , tal que $[\varepsilon(f)] \leq f$, é definida como a dilatação geodésica de $[\varepsilon(f)]$ com relação à f até a estabilidade.

O fechamento por reconstrução, que consiste em uma dilatação seguida por uma reconstrução morfológica por erosão, é definido por dualidade [26].

A Figura 11 mostra um exemplo de aplicação da operação de abertura por reconstrução utilizando funções estruturantes de tamanho crescente.

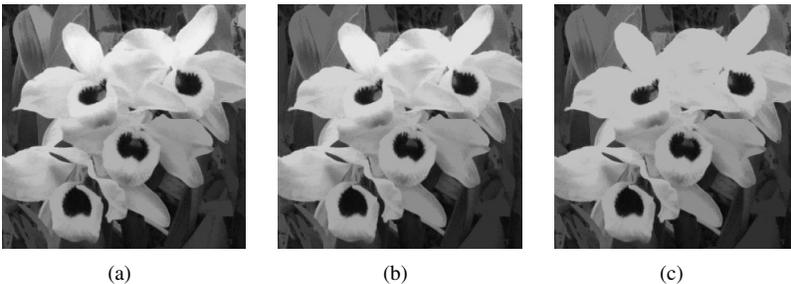


Figura 11. Abertura por reconstrução.

A abertura e o fechamento por reconstrução constituem operadores conexos, os quais transformam uma imagem através da eliminação ou fusão de suas zonas planas (regiões onde o nível de cinza é constante). Isso justifica porque a posição dos contornos é preservada.

Tais operações também podem ser classificadas como *levelings*, uma classe mais genérica de filtros morfológicos introduzida por Meyer [21]. No entanto, filtros por reconstrução não são auto-duais, ou seja, eles tratam o fundo e o primeiro plano da imagem de maneira assimétrica.

6.5 Self-dual Multiscale Morphological Toggle

Em [3], é definida uma nova transformação de imagens com interessantes propriedades de simplificação, a partir da qual detalhes indesejados podem ser eliminados sem que as características de interesse sejam comprometidas de maneira significativa. Esta transformação, denominada *Self-dual Multiscale Morphological Toggle* (SMMT), possui a forma de um operador *toggle* e é definida por:

Definição 6.3 (*Operador SMMT*) Seja $\phi_1^k(\mathbf{x}) = [\delta_{B_\sigma}(f)]^k(\mathbf{x})$, isto é, a dilatação de $f(\mathbf{x})$ com a função estruturante dependente de escala B_σ k vezes. De forma análoga, $\phi_2^k(\mathbf{x}) = [\varepsilon_{B_\sigma}(f)]^k(\mathbf{x})$. Denomina-se operador SMMT a função:

$$[\tau_{B_\sigma}(f)]^k(\mathbf{x}) = \begin{cases} \phi_1^k(\mathbf{x}), & \text{se } \phi_1^k(\mathbf{x}) - f(\mathbf{x}) < f(\mathbf{x}) - \phi_2^k(\mathbf{x}), \\ f(\mathbf{x}), & \text{se } \phi_1^k(\mathbf{x}) - f(\mathbf{x}) = f(\mathbf{x}) - \phi_2^k(\mathbf{x}), \\ \phi_2^k(\mathbf{x}), & \text{caso contrário.} \end{cases} \quad (18)$$

Pode-se mostrar que tal operação possui propriedades espaço-escala ao considerar escalas crescentes e um número de iterações fixo. Como máximos e mínimos da imagem integram ao mesmo tempo, são evitados alguns dos problemas relacionados a outras abordagens espaço-escala baseadas em morfologia matemática, tais como a manipulação de escalas negativas e o viés nos níveis de cinza.

Por outro lado, ao variar o número de iterações, uma região maior é analisada no cálculo do valor transformado de cada pixel. Desse modo, extremos mais significativos passam a ter uma maior abrangência, o que conduz a uma homogeneização dos níveis de cinza da imagem.

A seguir, são apresentados alguns exemplos em que a utilização do operador SMMT mostrou-se de fundamental importância para obtenção de resultados satisfatórios na segmentação de imagens em níveis de cinza. Isso se deve principalmente às suas propriedades, as quais conduzem a uma simplificação mais significativa quando comparada àquelas derivadas de métodos tradicionais tais como os filtros por reconstrução.

6.5.1 Segmentação de imagens em níveis de cinza Basicamente, técnicas de segmentação visam particionar uma imagem em um conjunto de regiões homogêneas e não-sobrepostas,

as quais devem representar algo significativo para uma determinada aplicação. Como constituem a base para diversas tarefas de processamento de imagens, tais como reconhecimento e classificação, é essencial que os resultados obtidos sejam precisos.

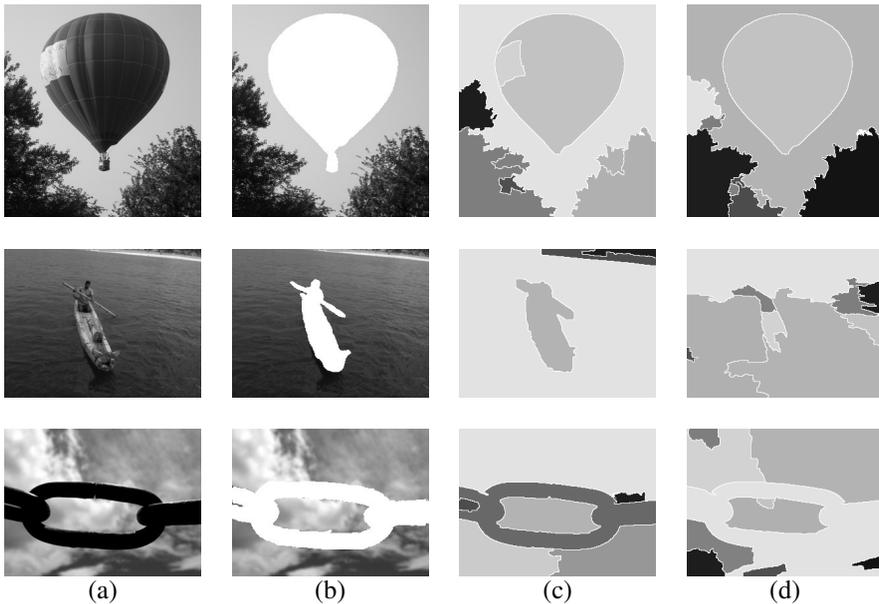


Figura 12. Resultados: (a) imagens originais, (b) segmentação manual e resultados (c) utilizando a imagem processada pelo SMMT e (d) a partir da imagem original.

Embora a fragmentação seja maior quando a imagem transformada pelo SMMT é utilizada, isso não é necessariamente ruim se a segmentação resultante é significativa, assim como ocorre na primeira linha da Figura 12, por exemplo. Além disso, como a distribuição dos níveis de cinza da imagem transformada é mais homogênea, a segmentação final é menos sensível às variações de iluminação, as quais podem influenciar negativamente o resultado, tal como ilustrado na segunda linha da Figura 12.

Além da super-segmentação (segmentação de regiões não significativas), outro problema freqüente é o vazamento, isto é, a situação em que parte do objeto conquista pixels do fundo da imagem ou vice-versa. Isso ocorre principalmente quando os contornos são pouco definidos, mas pode acontecer mesmo se o objeto de interesse tem um contraste completamente diferente do fundo, como no exemplo ilustrado na terceira linha da Figura 12, por exemplo. O resultado obtido através da imagem previamente simplificada não é comprometido devido à regularização realizada pelo SMMT.

6.5.2 Um operador de binarização multi-escala Diferentemente de outras abordagens, na transformação realizada pelo operador SMMT, máximos e mínimos da imagem interagem ao mesmo tempo, conduzindo a uma fusão de regiões que a simplifica de tal forma que importantes estruturas podem ser identificadas mesmo quando da existência de variações nas condições de iluminação. Além disso, por lidar explicitamente com o conceito de escala, tal abordagem possibilita a análise da imagem em diferentes níveis de representação, o que facilita a identificação daqueles contendo o conjunto das características de interesse. Tais propriedades são exploradas para definir uma nova operação de limiarização adaptativa multi-escala:

Definição 6.4 (*Operador de binarização adaptativo multi-escala*) Seja $\phi_1^k(\mathbf{x}) = [\delta_{B_\sigma}(f)]^k(\mathbf{x})$, isto é, a dilatação de $f(\mathbf{x})$ com a função estruturante dependente de escala B_σ k vezes. De forma análoga, $\phi_2^k(\mathbf{x}) = [\varepsilon_{B_\sigma}(f)]^k(\mathbf{x})$. O novo operador de binarização é dado pela função:

$$(f \circ g_\sigma)^k(\mathbf{x}) = \begin{cases} 1, & \text{se } \phi_1^k(\mathbf{x}) - f(\mathbf{x}) \leq f(\mathbf{x}) - \phi_2^k(\mathbf{x}), \\ 0, & \text{caso contrário.} \end{cases} \quad (19)$$

Em resumo, se o valor de um pixel está mais próximo do seu valor erodido atribui-se a ele zero (preto). Caso contrário, atribui-se o valor um (branco), inclusive se a diferença entre o seu valor erodido e dilatado for a mesma.

Observe também que, assim como é o caso para o SMMT, a regra de decisão baseia-se na similaridade do pixel sendo analisado em relação às transformações morfológicas de erosão e dilatação dependentes de escala. Dado que o valor de um máximo (mínimo) local é considerado para determinar o valor dilatado (erodido), pode-se interpretar o resultado da aplicação do operador de binarização como a representação da convergência de um pixel para a região de influência de um extremo local. Em resumo, um pixel arbitrário pode convergir para a região de influência de um mínimo (valor zero), de um máximo (valor um) ou para nenhuma (valor um). Assim como comentado anteriormente, o aumento no número de iterações considera uma vizinhança mais ampla nesta análise.

A Figura 13 ilustra o efeito da mudança de parâmetros para uma imagem com iluminação não-uniforme. Como o fundo é mais escuro que os grãos (Figura 13(a)), os pixels pertencentes a ele tendem a se aproximar mais do seu valor erodido quando escalas mais altas são consideradas (pois, neste caso, pequenas variações nos níveis de cinza são suficientes para modificá-los). A aplicação iterativa evita que extremos não significativos - tais como aqueles presentes devido às inconsistências na iluminação - comprometam o resultado (Figura 13(b)).

A Figura 13(c) mostra o resultado ao considerar uma escala menor. Neste caso, regiões com níveis de cinza homogêneos tendem a não ser alteradas, fazendo com que o fundo

não seja transformado. A parte interna, por ser mais clara, aproxima-se mais da transformação definida pela dilatação e também é alterada para o valor um. Por outro lado, os pixels pertencentes aos contornos dos grãos, regiões em que há uma maior variação, estão mais próximos do valor erodido e são associados ao valor zero.

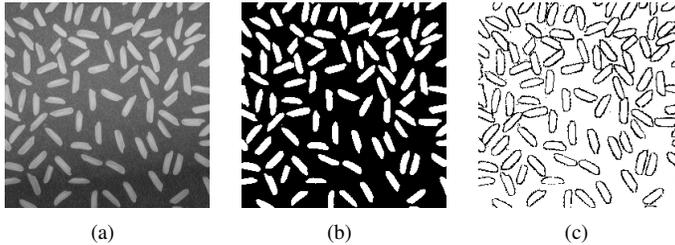


Figura 13. Efeito da alteração dos parâmetros no operador de binarização: (a) imagem original, (b) $\sigma^{-1} = 1$ e $k = 25$ e (c) $\sigma^{-1} = 40$ e $k = 5$.

Portanto, dependendo do objetivo específico, diferentes combinações de parâmetros devem ser utilizadas. Enquanto escalas maiores tendem a criar regiões, escalas menores as delimitam, identificando regiões de contorno.

7 Conclusões

Este trabalho apresentou uma breve revisão da teoria espaço-escala, abrangendo desde as propriedades necessárias até algumas das abordagens existentes na literatura. Como pode-se observar no decorrer do texto, abordagens não-lineares possuem vantagens significativas quando comparadas às lineares. No entanto, nestes casos existe uma grande dificuldade para satisfazer todas as condições necessárias para constituição de um espaço-escala, em especial a monotonicidade.

Além disso, várias abordagens propostas apresentam limitações, tais como o viés nos níveis de cinza, o que pode comprometer de forma significativa o resultado final. Outras utilizam conceitos que influenciam apenas a forma de apresentação dos resultados, mas não a transformação propriamente dita. O uso de uma escala “negativa” no MMDE, por exemplo, indica apenas que operação e quais características de interesse devem ser consideradas (a função estruturante definida para as escalas 15 e -15 é a mesma).

Tais problemas podem ser evitados (ou pelo menos minimizados) com a utilização do operador espaço-escala SMMT. Diferentemente de outras abordagens, máximos e mínimos da imagem interagem ao mesmo tempo, conduzindo a uma fusão de regiões que a simplificam

de tal forma que importantes estruturas podem ser identificadas mesmo quando da existência de variações nas condições de iluminação. Além disso, por lidar explicitamente com o conceito de escala, tal abordagem possibilita a análise da imagem em diferentes níveis de representação, facilitando a identificação daqueles contendo as características de interesse.

8 *

Referências

- [1] J.H. Bosworth and S.T. Acton. Morphological scale-space in image processing. *Digital Signal Processing*, 13:338–367, 2003.
- [2] M.H. Chen and P.F. Yan. A multiscale approach based on morphological filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(12):3377–3386, 1989.
- [3] L. B. Dorini and N. J. Leite. *Transformações de imagens baseadas em Morfologia Matemática*. PhD thesis, Universidade Estadual de Campinas, 2009.
- [4] L. Florack. Non-linear scale-spaces isomorphic to the linear case with applications to scalar, vector and multispectral images. *IJCV*, 42(1/2):39–53, 2001.
- [5] Maas R. Florack, L. and W. Niessen. Pseudo-linear scale-space theory. *International Journal of Computer Vision*, 32(2/3):247–259, 1999.
- [6] H. Heijmans. *Morphological Image Operators*. Academic Press, 1994.
- [7] H.J.A.M. Heijmans and R. van den Boomgaard. Algebraic framework for linear and morphological scale-spaces. *Journal of Visual Communication and Image Representation*, 13:269–301, 2002.
- [8] T. Iijima. Basic theory of pattern normalization for the case of a typical one-dimensional pattern. In *Bull. Electrotech. Lab*, pages 368–388, 1962.
- [9] Weickert J., Ishikawa S., and Imiya A. Linear scale-space has first been proposed in japan. *J. Math. Imaging Vis.*, 10(3):237–252, 1999.
- [10] P. Jackway. *Morphological Scale-Space with Application to Three-Dimensional Object Recognition*. PhD thesis, Queensland University of Technology, 1994.
- [11] P. T. Jackway and M. Deriche. Scale-space properties of the multiscale morphological dilation-erosion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:38–51, 1996.

- [12] B.K. Jang and R.T. Chin. Shape analysis using morphological scale space. In *25th Ann. Conf. on Information Sciences and Systems*, pages 1–4, 1991.
- [13] J.J. Koenderink. The structure of images. *Biological Cybernetics*, 50(5):363–370, 1984.
- [14] N. J. Leite and M. D. Teixeira. Morphological scale-space theory for segmentation problems. In *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing*, pages 364–368, 1999.
- [15] N. J. Leite and M. D. Teixeira. An idempotent scale-space approach for morphological segmentation. In *Mathematical Morphology and its Applications to Image and Signal Processing*, pages 291–300. Kluwer Academic Publishers, 2000.
- [16] L. M. Lifshitz and S. M. Pizer. A multiresolution hierarchical approach to image segmentation based on intensity extrema. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(4):529–540, 1990.
- [17] T. Lindeberg. *Discrete Scale-Space Theory and the Scale-Space Primal Sketch*. PhD thesis, Computational Vision and Active Perception Laboratory (CVAP), Royal Institute of Technology, 1991.
- [18] T. Lindeberg. Scale-space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics*, 21(2):224–270, 1994.
- [19] T. Lindeberg. *Scale-space theory in computer vision*. Kluwer Acad. Pub., 1994.
- [20] G. Matheron. *Random Sets and Integral Geometry*. John Wiley and Sons, 1975.
- [21] Fernand Meyer. Levelings, image simplification filters for segmentation. *Journal of Mathematical Imaging and Vision*, 20(1-2):59–72, 2004.
- [22] K. Park and C Lee. Scale-space using mathematical morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(11):1121–1126, 1996.
- [23] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on PAMI*, 42(12):629–639, 1990.
- [24] J. Serra. *Image Analysis and Mathematical Morphology*. Academic Press, 1982.
- [25] J. Serra. *Image Analysis and Mathematical Morphology, volume 2: Theoretical Advances*. Academic Press, 1988.
- [26] P. Soille. *Morphological Image Analysis: Principles and Applications*. Springer-Verlag, 2003.
- [27] A. P. Witkin. Scale-space filtering: a new approach to multi-scale description. In *Image Understanding*, pages 79–95. Ablex, 1984.