

Capítulo 3:

Difusão Anisotrópica

Resumo e nossas contribuições

O espaço de escala é uma das teorias utilizadas para a análise multi-escala de imagens e sinais. A técnica do espaço de escala linear gera as imagens em resoluções grossas fazendo convolução da imagem original com um núcleo gaussiano ou, equivalentemente, usando a imagem original como a condição inicial de um processo de difusão. Esta abordagem possui um defeito sério: é difícil obter a localização acurada das arestas importantes nas escalas grossas. A difusão anisotrópica foi proposta para superar esta dificuldade. Nela, os coeficientes da difusão são escolhidos de forma a encorajar a suavização intra-região e evitar a suavização inter-região. Com isso, os ruídos são eliminados e a imagem é simplificada ao mesmo tempo em que mantém as arestas nítidas. Temos utilizado a difusão anisotrópica em várias aplicações do Processamento e Análise de Imagens. Este capítulo descreve as teorias do espaço de escala linear e da difusão anisotrópica, e as nossas contribuições científicas nesta área.

Primeiro, descrevemos a teoria do espaço de escala linear e mostramos o efeito de deslocamento das arestas nas escalas grossas. Em segundo lugar, descrevemos a teoria da difusão anisotrópica (a clássica e aquela baseada na estatística robusta) e mostramos duas aplicações: a filtragem de sinais do sensor de aceleração e a detecção de arestas em imagens. Em terceiro lugar, descrevemos o melhoramento do algoritmo de reconstrução tomográfica de máxima entropia através da difusão anisotrópica robusta. Por fim, descrevemos o melhoramento do modelo linear geral (um processo de

detecção das áreas ativadas do cérebro em imagens de ressonância magnética funcional) usando a difusão anisotrópica robusta.

As nossas contribuições na área da difusão anisotrópica e do espaço de escala são:

- 1) *Melhoramento do algoritmo de reconstrução tomográfica máxima entropia (MENT) através da difusão anisotrópica robusta (RAD)*: Esta contribuição científica foi publicada em [Cn13] e encontra-se submetida em [Su03]. Nesta tese, ela está documentada na seção 3.4. O principal responsável por esta contribuição foi meu ex-orientando de doutorado Harold I. A. Bustos. Temos outros trabalhos publicados na área de tomografia [Ci07; Cn11; Cn09] mas que não estão documentados nesta tese.

Resumo: A máxima entropia (MENT) é uma técnica de reconstrução amplamente conhecida baseada na otimização da entropia. Se somente uma pequena quantidade de dados de aquisição estiver disponível, este algoritmo converge para uma imagem ruidosa e borrada. Propomos um melhoramento a este algoritmo que consiste em aplicar alternadamente a reconstrução MENT e a difusão anisotrópica robusta (RAD). Testamos esta idéia para a reconstrução de dados paralelos em ângulo completo, mas a idéia pode ser aplicada para qualquer cenário de aquisição de dados. A nova técnica tem gerado imagens surpreendentemente claras com arestas nítidas, mesmo utilizando uma quantidade de dados de projeção extremamente pequena.

- 2) *Melhoramento do modelo linear geral utilizado na detecção de áreas ativadas do cérebro a partir das imagens de ressonância magnética funcional (fMRI) utilizando RAD*: Esta contribuição foi publicada em [Cn12] e encontra-se submetida em [Su01]. Nesta tese, ela está documentada na seção 3.5. Eu fui o principal responsável por esta contribuição, contando com a colaboração do Prof. H. Z. Cho da University of California, Irvine.

Resumo: O imageamento por ressonância magnética funcional (fMRI) consegue mapear de forma não invasiva as áreas do cérebro com as atividades neuronais aumentadas sem o uso de um agente de contraste exógeno. A baixa

taxa de relação sinal-ruído das imagens fMRI torna necessário o uso de sofisticadas técnicas de processamento de imagens, tais como o mapa estatístico paramétrico (SPM), para detectar as áreas de cérebro ativadas. As nossas pesquisas levaram a uma nova técnica para obter um SPM limpo a partir dos dados fMRI ruidosos. Ela está baseada na RAD. Uma aplicação direta da RAD à fMRI não funciona, em grande parte devido à falta de bordas nítidas entre as regiões ativadas e não-ativadas. Para superar esta dificuldade, propomos calcular o SPM a partir da fMRI ruidosa, obter os coeficientes de difusão no espaço SPM, e então efetuar a difusão nas imagens fMRI utilizando os coeficientes previamente calculados. Estes passos são iterados até a convergência. Resultados experimentais utilizando a nova técnica geraram SPMs surpreendentemente nítidos e sem ruídos, com alto grau de significância estatística.

- 3) *Uso da RAD para filtrar sinais de sensores.* Para ilustrar o processo de RAD, utilizei nesta tese uma adaptação de um artigo, elaborado principalmente pelo meu orientando de mestrado Marco A. A. de Melo, onde se usa a RAD para filtrar os sinais de sensor de aceleração [Su06] (seção 3.3).
- 4) Também temos algumas contribuições em filtro nebulosos no espaço de escala [Cn08] e no uso da RAD para segmentar imagens coloridas [Ci09], mas que não estão documentadas nesta tese.

3.1 Introdução

Percebemos os objetos no mundo como tendo estruturas em escalas grossas e finas. Uma floresta pode parecer simplesmente um amontoado verde quando vista de distância. À medida que nos aproximamos, começamos a distinguir as árvores individuais, os troncos, os galhos, as folhas, as nervuras das folhas, os orvalhos sobre as folhas, etc. Assim, a multi-escala constitui uma noção natural da percepção visual. A representação multi-escala de uma imagem em forma de pirâmide foi desenvolvida já

na década de 70. Nesta estrutura, quanto mais grossa for a escala, menos pixels conterá a imagem.

Em 1983, Witkin [Witkin, 1983] propôs que a escala poderia ser considerada como um parâmetro contínuo, generalizando a noção de pirâmide. A idéia essencial desta abordagem é muito simples: dada uma imagem digital Q , essa imagem na escala σ é a convolução da Q com a máscara gaussiana de desvio-padrão σ . Esta teoria é denominada de espaço de escala gaussiano ou linear. A imagem Q na escala $\sigma=0$ é a própria imagem original. À medida que se vai da escala fina para a escala grossa, a imagem se torna cada vez mais “borrada”.

A convolução com a máscara gaussiana de desvio-padrão σ pode ser vista como a solução da equação de condução de calor, onde o valor da imagem original Q num ponto (x, y) é a temperatura inicial nesse ponto, o tempo decorrido é $t = \sigma^2 / 2$, e a imagem Q na escala σ representa as temperaturas no instante t . Assim, a convolução gaussiana é um processo de difusão isotrópica. Isotrópico significa “aquele que apresenta as mesmas propriedades físicas em todas as direções”, segundo [Aurélio, 1999].

As pesquisas subseqüentes levaram a diferentes formas de simplificar a imagem original, utilizando filtros diferentes da convolução gaussiana. Por exemplo, Jackway e Deriche [Jackway and Deriche, 1996] propuseram o uso de operadores morfológicos, resultando no espaço de escala morfológico.

Uma outra forma de simplificar imagens foi proposta por Perona e Malik [Perona and Malik, 1987; Perona and Malik, 1990], e teve um grande impacto científico. Eles propuseram o uso da difusão anisotrópica, substituindo a difusão isotrópica. No espaço de escala linear (que utiliza a difusão isotrópica para simplificar uma imagem), uma imagem em escala grossa torna-se borrada e as arestas deslocam-se espacialmente de uma escala para outra. Utilizando a difusão anisotrópica, as arestas continuam nítidas mesmo em escalas grossas e permanecem na mesma posição mesmo mudando de escala.

Na formulação da difusão anisotrópica de Perona-Malik, existe uma função chamada parada-na-aresta (edge stopping function) g , que controla a intensidade da difusão de acordo com o gradiente do ponto que deve sofrer difusão. A função parada-na-aresta possui um parâmetro de escala σ que, em conjunto com o gradiente, indica se a difusão deve ser forte ou fraca. A correta escolha da função parada-na-aresta e da escala afetam de forma decisiva o resultado da filtragem da imagem. Perona e Malik sugeriram duas funções parada-na-aresta, sem apresentar uma justificativa fundamentada para a escolha.

A difusão anisotrópica robusta (RAD) [Black et al., 1998] foi proposta como um melhoramento da difusão anisotrópica de Perona-Malik. Esta técnica assume que a entrada é uma imagem constante por regiões corrompida pelo ruído gaussiano aditivo com média zero e pequena variância. O objetivo é estimar a imagem original a partir dos dados ruidosos. Black et al. usaram a estatística robusta para resolver este problema, e propuseram o uso da função “Tukey’s biweight” como a função parada-na-aresta, de acordo com a teoria estatística adotada. Na prática, a RAD converge mais rapidamente e conserva ainda melhor as bordas do que a difusão de Perona-Malik.

A RAD mostra-se útil em diversas aplicações de Processamento e Análise de Imagens. Ela é um excelente detector de arestas. Também é um ótimo filtro de ruídos aditivos, que preserva as bordas ao mesmo tempo em que elimina os ruídos. Quando utilizado como um filtro, a RAD procura estimar a imagem original constante por regiões a partir da sua versão corrompida pelo ruído gaussiano aditivo. Esta característica torna-a extremamente eficiente em diversas aplicações.

Organização deste capítulo

O restante deste capítulo está organizado como segue. A seção 3.2 apresenta os conceitos básicos do espaço de escala linear (ou gaussiano), subdividido em casos unidimensional e bidimensional. A seção 3.3 apresenta os conceitos básicos do espaço de escala não-linear gerado pela difusão anisotrópica, incluindo a difusão anisotrópica robusta (RAD). A fim de ilustrar o processo da difusão anisotrópica, apresentamos

na seção 3.3 duas aplicações: a filtragem de sinais unidimensionais do sensor de aceleração e a detecção de arestas em imagens. A seção 3.4 apresenta o melhoramento do algoritmo MENT-estendido através da RAD, denominada MENT reconstrução-difusão. A seção 3.5 apresenta o melhoramento obtido no processamento da ressonância magnética funcional (fMRI) através do uso de RAD. Finalmente, na seção 3.6 apresentamos as nossas conclusões.

3.2 Espaço de Escala Linear

Caso unidimensional

Nesta subsubseção, vamos explicar o espaço de escala linear seguindo de perto o texto [Velho et al., 2000]. Historicamente, a teoria do espaço de escala foi elaborada primeiro para os sinais unidimensionais e depois estendida para as imagens. Para definir o espaço de escala, vamos definir antes dois conceitos básicos.

Definição (convolução): Dadas duas funções $f, g : \mathbb{R} \rightarrow \mathbb{R}$ a convolução é definida:

$$f(x) * g(x) = (f * g)(x) = \int_{-\infty}^{\infty} f(u)g(x-u)du$$

Definição (gaussiana): A distribuição normal $N(\mu, \sigma)$, onde μ é a média e σ é o desvio-padrão, é definida através da função gaussiana:

$$g(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

A figura 3.1a mostra a função $g(x,0,1)$ e as suas derivadas. Costuma-se adotar $\sigma^2 = 2t$ e $\mu = 0$ para obter a notação:

$$G_t(x) = \frac{1}{\sqrt{4\pi t}} \exp\left[-\frac{x^2}{4t}\right]$$

Note que $\lim_{t \rightarrow 0} G_t(x)$ resulta no impulso de Dirac.

Com isso, já podemos definir o espaço de escala:

Definição (espaço de escala): Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ um sinal unidimensional. O espaço de escala deste sinal é a função $F : \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R}$ (representada por $F(x,t) = F_t(x)$) que é a solução da equação de calor:

$$\begin{cases} \frac{\partial F_t(x)}{\partial t} = \frac{\partial^2 F_t(x)}{\partial x^2} \\ F_0(x) = f(x) \end{cases}$$

Afirmção: O espaço de escala de $f : \mathbb{R} \rightarrow \mathbb{R}$ pode ser obtido através das convoluções com gaussianas:

$$F_t(x) = G_t(x) * f(x)$$

O espaço de escala gaussiano tem as seguintes propriedades básicas [Velho et al., 2000]:

- *Linearidade:* A transformação L_t que leva o sinal original $f(x)$ ao espaço de escala $F_t(x)$ é linear, isto é, $L_t\{f + \lambda g\}(x) = L_t\{f(x)\} + \lambda L_t\{g(x)\}$.
- *Invariância por translação:* Se T é uma translação qualquer, o espaço de escala de $T(f)$ é $T(F)$, isto é, $G_t(x) * T f(x) = T(G_t(x) * f(x))$.
- *Causalidade:* O sinal f é “simplificado” com o aumento da escala, isto é, os cruzamentos de zero não aumentam com o aumento de t .

A causalidade é a propriedade que permite afirmar que um sinal é simplificado pela convolução por gaussiana. Informalmente, uma “aresta” do sinal (isto é, um ponto do sinal onde há uma transição abrupta de nível) corresponde ao cruzamento de zero da segunda derivada. Em sinais unidimensionais, os cruzamentos de zero da segunda derivada desaparecem à medida que o sinal é filtrado por máscaras gaussianas com desvios-padrões cada vez maiores. Por outro lado, nunca pode aparecer um cruzamento de zero da segunda derivada numa escala grossa σ_g , se ela não estava presente em todas as escalas σ mais finas que σ_g . Uma aresta pouco acentuada não pode ser detectada numa escala grossa, enquanto que numa escala bem fina todas as arestas podem ser detectadas. Formalizando este conceito, temos:

Definição (cruzamento de zero): Um cruzamento de zero de uma função contínua $f(x)$ é um intervalo fechado $[a, b]$ (possivelmente com $a=b$) tal que:

$$\begin{cases} f([a, b]) = 0 \\ \lim_{x \rightarrow a^-} \text{ sinal}(f(x)) = - \lim_{x \rightarrow b^+} \text{ sinal}(f(x)) \neq 0 \end{cases}$$

Proposição (causalidade do cruzamento de zero): Dada uma função $f(x)$ contínua, considere o seu espaço de escala gaussiano $F_t(x)$. O número de cruzamentos de zero de $F_t(x)$ não aumenta à medida que t cresce.

Corolário: Se $f(x)$ é diferenciável, então o número de máximos e mínimos de $F_t(x)$ não aumenta à medida que t cresce.

Demonstração: Os máximos e mínimos (com relação a x) de $F_t(x) = G_t(x) * f(x)$ são os cruzamentos de zero de

$$\frac{\partial F_t(x)}{\partial x} = G_t(x) * \frac{\partial f(x)}{\partial x}$$

que é o espaço de escala de $f'(x)$. ■

De forma semelhante, pode-se demonstrar que os cruzamentos da segunda derivada de f (as arestas) não aumentam à medida que t cresce.

Proposição: $F_t(x)$ é uma função suave (infinitamente diferenciável) para qualquer $t > 0$ fixo.

Demonstração: Note que

$$\frac{\partial^n F_t}{\partial x^n} = \frac{\partial^n}{\partial x^n} (G_t(x) * f(x)) = \frac{\partial^n G_t(x)}{\partial x^n} * f(x)$$

existe pois a função gaussiana $G_t(x)$ é suave para qualquer $t > 0$. ■

Caso bidimensional

Definição (normal): A distribuição normal bidimensional $N(x_0, y_0, \sigma)$, onde (x_0, y_0) é a média e σ é o desvio-padrão, é definida através da função gaussiana:

$$g(x, y, x_0, y_0, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left[\frac{-(x-x_0)^2 - (y-y_0)^2}{2\sigma^2}\right]$$

A figura 3.1b mostra a função $g(x, y, 0, 0, 1)$, e as figuras 3.1c-3.1f mostram o seu módulo do gradiente, as suas derivadas parciais, e o seu laplaciano. Costuma-se adotar $\sigma^2 = 2t$ e $\mu = 0$ para obter a notação:

$$G_t(x, y) = \frac{1}{4\pi t} \exp\left[-\frac{x^2 + y^2}{4t}\right].$$

Definição (espaço de escala): Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ uma imagem 2-D. O espaço de escala desta imagem é a função $F : \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow \mathbb{R}$ (denotada $F_t(x, y)$) que satisfaz a seguinte equação diferencial parcial ou equação de calor bidimensional:

$$\begin{cases} \frac{\partial F_t(x, y)}{\partial t} = \nabla^2 F_t(x, y) = \frac{\partial^2 F_t(x, y)}{\partial x^2} + \frac{\partial^2 F_t(x, y)}{\partial y^2}, \\ F_0(x, y) = f(x, y) \end{cases}, \quad (3.1)$$

Afirmção: A solução da equação diferencial parcial acima pode ser expressa como uma convolução com gaussianas bidimensionais:

$$F_t(x, y) = G_t(x, y) * f(x, y),$$

Proposição (separabilidade): A convolução acima pode ser calculada através de duas convoluções com gaussianas unidimensionais:

$$F_t(x, y) = G_t(y) *_{(y)} \left(G_t(x) *_{(x)} f(x, y) \right).$$

Esta propriedade permite acelerar a computação do espaço de escala gaussiano para as imagens.

Além da linearidade e da invariância por translações, o espaço de escala gaussiano bidimensional possui a invariância por rotações.

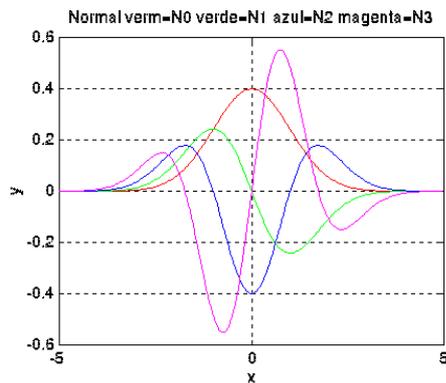
Proposição: Seja f uma imagem qualquer e $g = R_\theta f$ a rotação de f por ângulo θ . Então o espaço de escala G_t de g é a rotação por ângulo θ de F_t , isto é:

$$g = R_\theta f \Rightarrow G_t = R_\theta F_t.$$

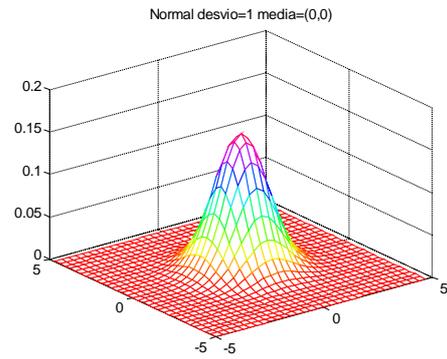
Infelizmente, o princípio de causalidade não vale para as imagens 2-D. Velho et al. [Velho et al., 2000] afirmam: “Tentemos agora entender o que será o princípio da causalidade em 2-D. Note que não faz sentido falar em número de cruzamentos de zero de uma imagem, já que em geral os cruzamentos de zero de uma imagem formam um conjunto de curvas, não um conjunto discreto de pontos. Por outro lado, pode-se falar do número de máximos e mínimos locais de uma imagem genérica (ou de um sinal n -dimensional). No entanto, *não é verdade* que o número de pontos críticos diminua com a escala no espaço de escala de uma imagem qualquer.”

Na prática, qualquer imagem digital está definida em um subconjunto finito de \mathbb{Z}^2 , em vez de \mathbb{R}^2 . Assim, é necessário discretizar de alguma forma o espaço de escala espacialmente. Muitas técnicas de discretização têm sido utilizadas para esta tarefa, por exemplo, a gaussiana amostrada, a gaussiana integrada e a gaussiana verdadeiramente discreta (obtida utilizando a função modificada de Bessel). Por outro lado, não é estritamente necessário discretizar o espaço de escala no tempo, pois é possível calcular “sob encomenda” qualquer pixel em qualquer escala real no espaço de escala discretizado espacialmente $\mathbb{Z}^2 \times \mathbb{R}_+$. Porém, é computacionalmente vantajoso pré-calcular o espaço de escala para algumas escalas fixas, obtendo o espaço de escala discretizada espacial e temporalmente $\mathbb{Z}^2 \times \mathbb{Z}_+$. Veja [Velho et al., 2000; Lindeberg, 1994] para maiores detalhes.

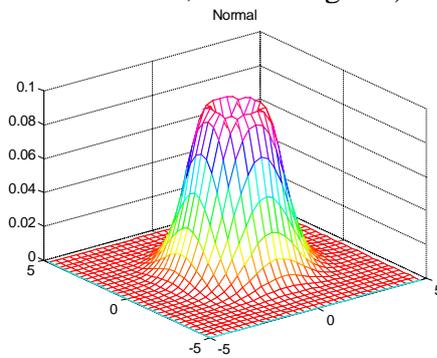
A figura 3.2 mostra a detecção de arestas de uma imagem no espaço de escala linear. A imagem original sofre convoluções com as gaussianas de diferentes desvios-padrões, gerando as imagens em diferentes escalas (primeira coluna). Calculando a convolução da imagem original com o laplaciano da gaussiana com diferentes desvios-padrões, obtém-se a segunda coluna (onde está ilustrado somente o sinal das imagens resultantes: preto indica positivo e branco indica negativo). Aplicamos os operadores morfológicos (dilatação seguida pela operação ou-exclusivo) sobre as imagens da segunda coluna para obter as arestas em diferentes escalas (terceira coluna).



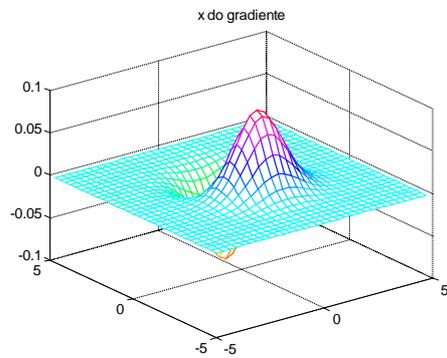
(3.1a) Função gaussiana com $\sigma=1$ (vermelho) e suas 1ª, 2ª e 3ª derivadas (respectivamente em verde, azul e magenta).



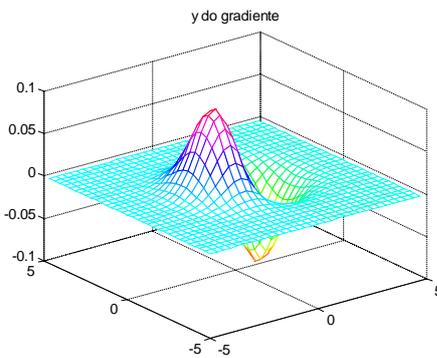
(3.1b) Função gaussiana G bidimensional com $\sigma=1$.



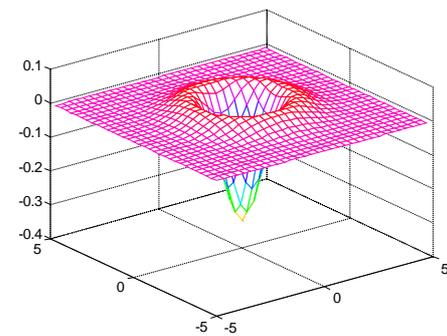
(3.1c) Módulo do gradiente da gaussiana $\|\nabla G(x, y)\|$.



(3.1d) Derivada parcial x da gaussiana $\partial G(x, y) / \partial x$.



(3.1e) Derivada parcial y da gaussiana $\partial G(x, y) / \partial y$.



(3.1f) Laplaciano da gaussiana $\nabla^2 G(x, y)$.

Fig. 3.1: Funções gaussianas unidimensional, bidimensional e suas derivadas.

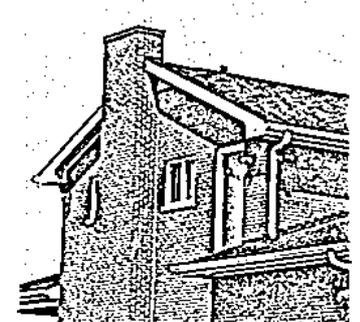
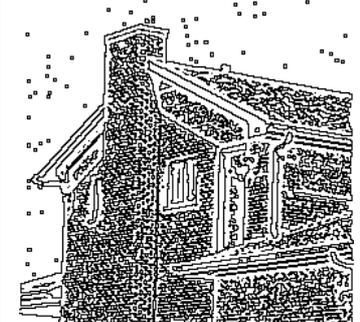
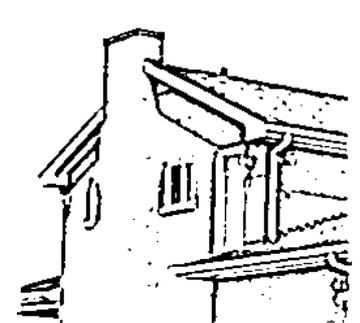
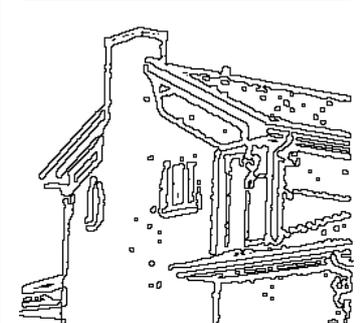
Imagem “casa.tga” no espaço de escala linear.	Sinal do laplaciano da imagem. Preto indica positivo e branco indica negativo.	Cruzamentos de zero do laplaciano (ou arestas).
 <p>(3.2a) Imagem original ($\sigma \rightarrow 0$)</p>		
 <p>(3.2b) $\sigma=1,0$</p>	 <p>(3.2c) $\sigma=1,0$</p>	 <p>(3.2d) $\sigma=1,0$</p>
 <p>(3.2e) $\sigma=1,5$</p>	 <p>(3.2f) $\sigma=1,5$</p>	 <p>(3.2g) $\sigma=1,5$</p>

Fig. 3.2: Continua na próxima página.

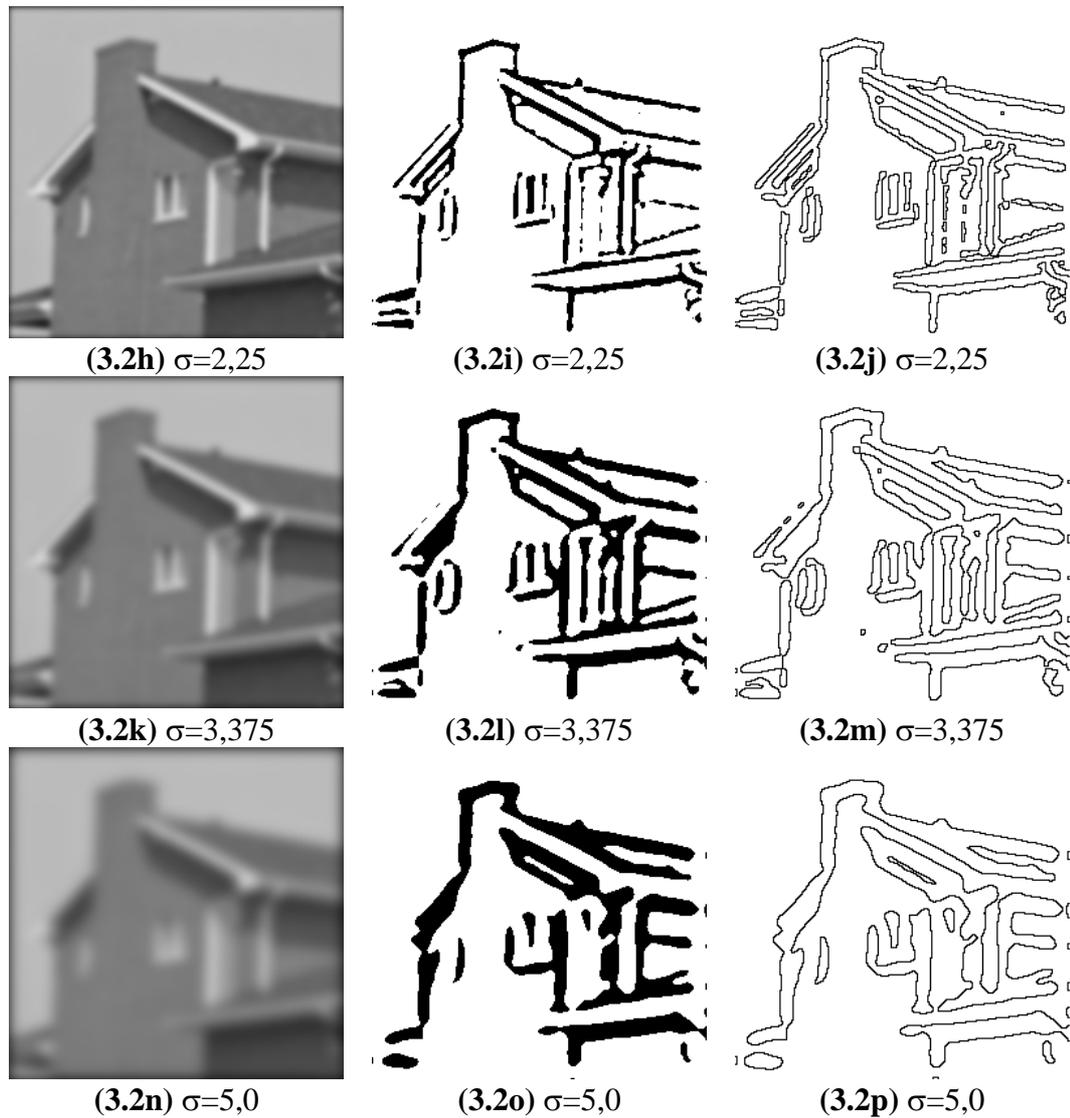


Fig. 3.2: Detecção de arestas no espaço de escala linear. A imagem “casa.tga” no espaço de escala (primeira coluna), o sinal do laplaciano da imagem (segunda coluna) e as arestas ou os cruzamentos de zero do laplaciano da imagem (terceira coluna). Quanto σ cresce, as arestas menos importantes deixam de ser detectadas. Note que as arestas deslocam-se espacialmente à medida que σ cresce.

3.3 Difusão Anisotrópica

O espaço de escala linear possui muitas propriedades matemáticas atraentes. Porém, nas escalas grossas, a imagem torna-se borrada e as arestas deslocam-se espacialmente. Para manter as arestas nítidas, ao mesmo tempo em que se filtram os ruídos e os detalhes pouco importantes, Perona e Malik definiram o espaço de escala não-linear anisotrópica [Perona and Malik, 1990] modificando a equação diferencial parcial (3.1):

$$\begin{cases} \frac{\partial F_t(x, y)}{\partial t} = \nabla \cdot [g(\|\nabla F_t(x, y)\|) \nabla F_t(x, y)] \\ F_0(x, y) = f(x, y) \end{cases}$$

onde $\|\nabla F_t(x, y)\|$ é o magnitude do gradiente da F_t , e g é uma função “parada-na-aresta” (edge stopping function).

Perona e Malik discretizaram (espaço-temporalmente) a sua equação de difusão anisotrópica acima como:

$$I(s, t+1) = I(s, t) + \frac{\lambda}{|\eta_s|} \sum_{p \in \eta_s} g(\|\nabla I_{s,p}(t)\|) \nabla I_{s,p}(t), \quad (3.2)$$

onde:

- $I(s, t)$ é a imagem discretizado espacial e temporalmente;
- s denota a posição de pixel numa grade discreta 2-D;
- t agora denota o passo de tempo discreto (número de iterações, $t \geq 0$);
- a constante $\lambda \in \mathbb{R}^+$ determina a velocidade de difusão (normalmente $\lambda = 1$);
- η_s representa o conjunto de vizinhos espaciais do voxel s . Para imagens 2-D, normalmente quatro pixels vizinhos são considerados: norte, sul, leste e oeste. Para imagens 3-D, seis voxels são normalmente considerados (os quatro voxels já mencionados mais os voxels “em cima” e “embaixo”);
- $\nabla I_{s,p}(t)$ é a magnitude do gradiente da imagem I no ponto s na direção (s, p) na iteração t : $\nabla I_{s,p}(t) = I(p, t) - I(s, t)$, $p \in \eta_s$.

Perona e Malik sugeriram usar uma das duas funções parada-na-aresta abaixo (que vamos denotar por g_1 e g_2):

$$g_1(x) = \frac{1}{1 + \frac{x^2}{2\sigma^2}}$$

$$g_2(x) = \exp\left[\frac{-x^2}{2\sigma^2}\right]$$

A correta escolha da função g e da escala σ afeta substancialmente o quanto as descontinuidades serão preservadas.

Black et al. [Black et al., 1998] propuseram recentemente a difusão anisotrópica robusta (RAD). Esta técnica assume que a entrada é uma imagem constante por regiões corrompida pelo ruído gaussiano aditivo com média zero e pequeno desvio-padrão. O objetivo é estimar a imagem original a partir do dado ruidoso. Black et al. usaram a estatística robusta para resolver este problema. Eles calcularam uma imagem I que satisfaz o seguinte critério de otimização:

$$\min_I \sum_{s \in I} \sum_{p \in \eta_s} \rho_\sigma(I(p) - I(s))$$

onde $I(s)$ é o valor da imagem I no pixel s , η_s é a vizinhança espacial do pixel s , ρ é uma norma de erro robusta e σ é um parâmetro de escala. A equação acima pode ser resolvida pelo sistema (3.2), fazendo $g(x) = \rho'(x)/x$. Black et al. escolheram a função “Tukey’s biweight” como a norma de erro ρ , de acordo com a teoria da estatística robusta. A correspondente função parada-na-aresta, que denotaremos como g_3 , é:

$$g_3(x) = \begin{cases} \left[1 - \frac{x^2}{\sigma^2}\right]^2, & |x| \leq \sigma \\ 0, & \text{caso contrário} \end{cases}$$

Para ter uma noção intuitiva da RAD, considere uma imagem constante por regiões, corrompida pelo ruído. A RAD executa a média da vizinhança intra-região, e evita

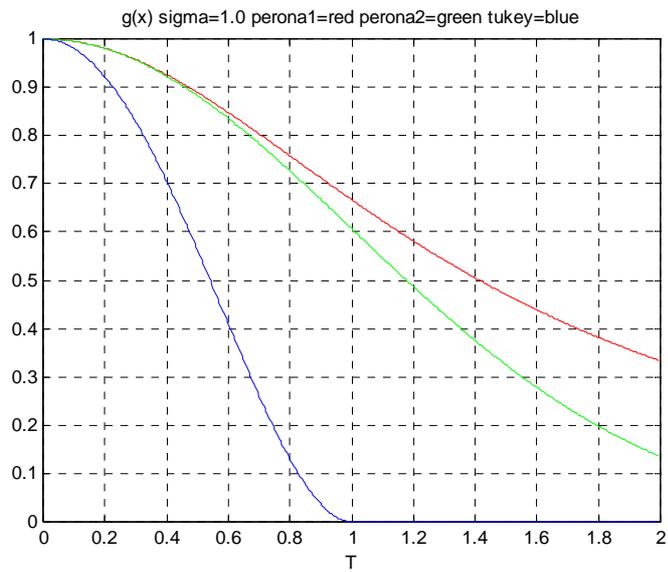
calcular a média inter-região. Assim, este processo atenua os ruídos ao mesmo tempo em que preserva as arestas entre as diferentes regiões nítidas.

A figura 3.3a mostra as três funções parada-na-aresta. Repare que as três estão em escalas diferentes, de forma que é necessário normalizá-las para poder compará-las. Para isso, considere a função $\psi(x) = xg(x) = \rho'(x)$. Esta função é denominada função de influência na estatística robusta e indica o quanto o erro cometido por uma medida particular (e quantificado pela norma de erro ρ) influencia na solução. A figura 3.3b mostra as 3 funções de influência correspondentes às 3 funções parada-na-aresta. Para normalizar as 3 funções parada-na-aresta, os pontos de máximo das 3 funções de influência foram calculados, e as funções ψ_1 e ψ_2 foram ajustadas de forma que os seus pontos de máximo coincidam com o ponto de máximo da ψ_3 ($x = \sqrt{0,2}$). Fazendo isso, obtivemos as funções g_1 e g_2 normalizadas abaixo. A função g_3 não foi alterada.

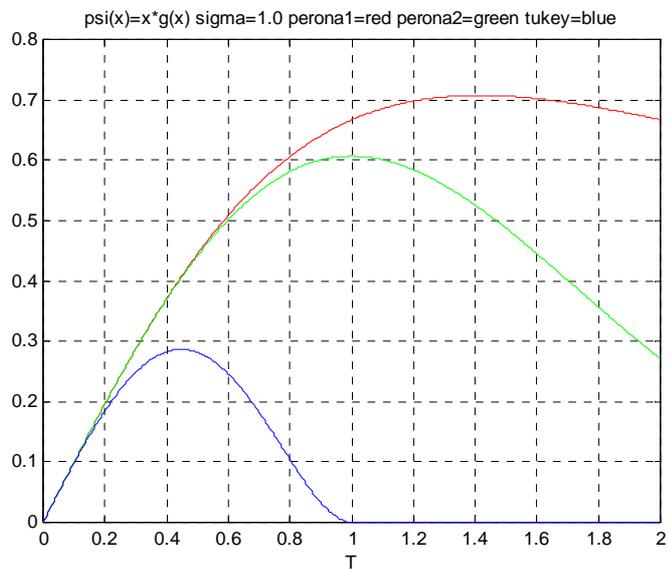
$$g_1(x) = \frac{1}{1 + \frac{5x^2}{\sigma^2}}$$

$$g_2(x) = \exp\left[\frac{-5x^2}{2\sigma^2}\right]$$

As figuras 3.3c e 3.3d mostram as funções parada-na-aresta e de influência normalizadas.

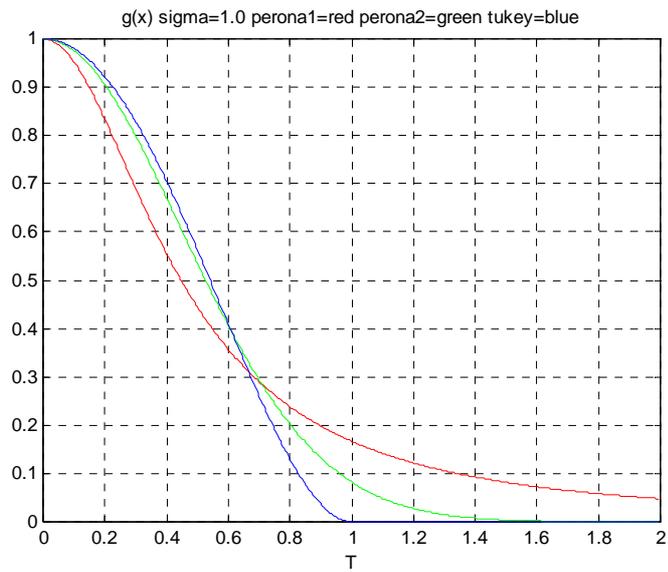


(3.3a) Funções parada-na-aresta não-normalizadas com $\sigma=1$: g_1 (Perona-Malik 1, em vermelho), g_2 (Perona-Malik 2, em verde) e g_3 (Tukey's biweight, em azul).

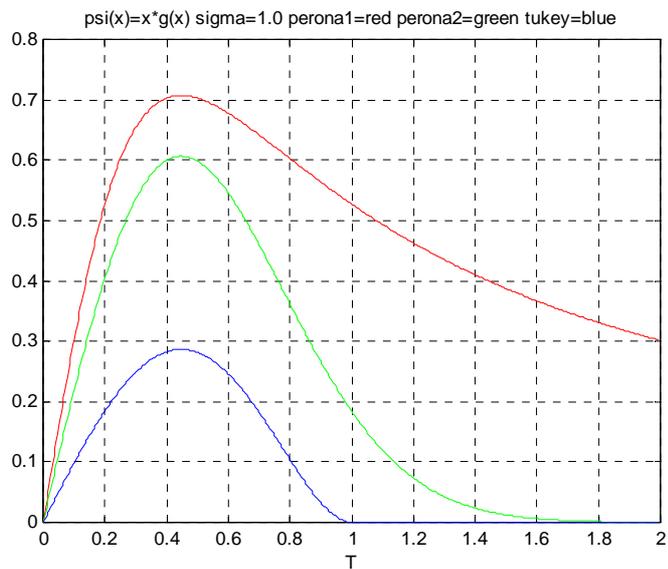


(3.3b) Funções de influência não-normalizadas: ψ_1 (Perona-Malik 1, em vermelho), ψ_2 (Perona-Malik 2, em verde) e ψ_3 (Tukey's biweight, em azul).

Fig. 3.3: Continua na próxima página.



(3.3c) Funções parada-na-aresta normalizadas com $\sigma=1$: g_1 (Perona-Malik 1, em vermelho), g_2 (Perona-Malik 2, em verde) e g_3 (Tukey's biweight, em azul).



(3.3d) Funções de influência normalizadas: ψ_1 (Perona-Malik 1, em vermelho), ψ_2 (Perona-Malik 2, em verde) e ψ_3 (Tukey's biweight, em azul).

Fig. 3.3: Funções parada-na-aresta e de influência, antes e depois da normalização.

Filtragem de sinais unidimensionais pela difusão anisotrópica

Esta subsubseção é uma adaptação resumida do artigo [Su06] elaborado principalmente pelo orientando de mestrado Marco A. A. de Melo.

Vamos ilustrar o uso da difusão anisotrópica, utilizando-a para filtrar sinais unidimensionais. O objetivo é filtrar os sinais do sensor de aceleração ADLX202E da Analog Devices [Analog, 2000], utilizado em sistemas automotivos para determinar a velocidade do veículo [Shih and Weinberg, 2001]. Inicialmente, iremos analisar o efeito das escolhas da função parada-na-aresta e escala na filtragem de um sinal artificial semelhante aos sinais do acelerômetro e contaminado com um ruído gaussiano. O uso de um sinal artificial permite calcular o erro cometido pelo filtro, pois neste caso dispomos do sinal sem ruído com o qual o sinal filtrado pode ser comparado. Usando um sinal real do acelerômetro, não há como calcular o erro, pois não temos acesso ao sinal sem ruído.

Como vimos na seção anterior, as funções parada-na-aresta dependem de um parâmetro de escala σ . Testamos duas maneiras de calcular automaticamente esta escala a partir do sinal ruidoso I . A primeira, denominada escala robusta, está baseada em estatística robusta e foi proposta por [Black et al., 1998]:

$$\begin{aligned}\sigma_1 &= 1,4826 \text{MAD}(\nabla I) \\ &= 1,4826 \text{median}_I \left[\left| \|\nabla I\| - \text{median}_I (\|\nabla I\|) \right| \right]\end{aligned}$$

onde “MAD” denota o desvio absoluto mediano e a constante deriva do fato de que MAD de uma distribuição normal com média zero e variância unitária é $1/1,4826 = 0,6745$.

A segunda forma de calcular a escala provém da análise utilizada em sistemas de transmissão de pulsos contaminados por ruído gaussiano, onde a informação está contida na amplitude e na posição dos pulsos [Carlson, 1986]. Esta análise pode ser aplicada aos sinais do acelerômetro, pois também neste caso a informação está presente na amplitude e no instante dos pulsos. Carlson sugere usar o limiar de detecção

de pulso maior ou igual a duas vezes o desvio-padrão σ_n do ruído gaussiano. Neste trabalho, testamos a seguinte escala:

$$\sigma_2 = 2\sigma_n.$$

O erro entre o sinal filtrado e o sinal sem ruído foi calculado usando a Raiz da Média Quadrática (RMS):

$$\text{RMS}(S, R) = \sqrt{\frac{1}{N} \sum (S_n - R_n)^2}$$

onde:

- S_n é a amostra n do sinal filtrado.
- R_n é a amostra n do sinal original sem ruído.
- N é o número total de amostras dos sinais.

A figura 3.4a mostra o sinal sem ruído gerado artificialmente e a figura 3.4b mostra esse sinal contaminado com ruído gaussiano com desvio-padrão 0,2315. O sinal original e o sinal com ruído foram amostrados em 50 Hz. O erro RMS entre os dois sinais é 0,2314. Portanto, se o erro de algum sinal filtrado for maior que esse valor, o filtro estará introduzindo mais erro, em vez de diminuí-lo. As escalas σ_1 e σ_2 foram calculadas e são respectivamente 0,198 e 0,463.

Os erros obtidos com a filtragem do sinal por difusão anisotrópica usando as funções parada-na-aresta g_1 , g_2 e g_3 estão mostrados nas tabelas 3.1 e 3.2. Na tabela 3.1 foram feitas 50 iterações da equação (3.2) e na tabela 2 foram feitas 100 iterações. Para cada função, são mostrados os 3 erros observados usando as escalas σ_1 , σ_2 e a escala ótima σ^* . As escalas ótimas foram escolhidas manualmente para se obter o menor erro RMS.

Conforme esperado, em ambos os casos o menor erro foi obtido usando a função de Tukey g_3 . A função Perona-Malik g_2 também gerou erros baixos, bastante próximos aos de Tukey. A figura 3.4c mostra o sinal filtrado pela difusão anisotrópica com a melhor qualidade. Pode-se notar que a informação útil foi preservada, ao mesmo tempo em que os ruídos indesejáveis foram eliminados.

Utilizando a função g_1 , os erros aumentaram de 50 para 100 iterações (tabelas 3.1 e 3.2, para as escalas fixas σ_1 e σ_2). Isto é causado pela não-convergência do processo de difusão anisotrópica, quando se utiliza a função g_1 . Neste caso, se o número de iterações for muito grande, o sinal filtrado converge para um sinal com amplitude constante. A figura 3.4d mostra esse fenômeno, onde os pulsos tiveram os picos de-cepados após 100 iterações.

Por outro lado, utilizando a função g_3 os erros diminuíram ligeiramente quando o número de iterações aumentou (para as escalas fixas σ_1 , σ_2 e σ^*). O processo de difusão anisotrópica usando g_3 converge para um sinal otimamente filtrado, de forma que a saída não se altera após um número suficiente de iterações. A função g_2 é muito parecida com g_3 , mas mesmo assim observa-se um pequeno aumento de erro ao aumentar o número de iterações (para σ_2).

A partir da tabela 2, podemos extrair as seguintes regras empíricas para calcular a escala ótima σ^* da difusão anisotrópica robusta:

$$\sigma^* = \frac{0,593}{0,198} \sigma_1 = 3,00 \times \sigma_1$$

$$\sigma^* = \frac{0,593}{0,463} \sigma_2 = 1,28 \times \sigma_2$$

A figura 3.4e mostra o sinal filtrado pelo filtro linear Butterworth passa-baixas de quinta ordem, com frequência de corte de 5 Hz. Variamos a frequência de corte até encontrar aquela onde o erro era mínimo. O sinal filtrado mostra uma forte distorção, alterando as amplitudes e os instantes de ocorrência dos picos do sinal. Esta distorção está refletida na alta taxa de erro RMS obtido (0,66135), mais de 8 vezes maior do que o menor erro obtido com a difusão anisotrópica (0,07957) e maior até que o erro do sinal ruidoso não filtrado (0,2314).

Também testamos o desempenho do filtro linear média móvel que consiste em calcular a média aritmética dos valores dos pontos vizinhos:

$$S_n = [Q_{n-1} + Q_n + Q_{n+1}] / 3$$

onde S_n é o sinal filtrado e Q_n é o sinal original com ruído. O erro RMS obtido foi 0,15449.

Aplicamos os resultados obtidos até agora na filtragem de um sinal real do acelerômetro. A qualidade da filtragem deverá ser avaliada apenas visualmente, pois não é possível calcular o erro RMS, uma vez que não dispomos do sinal original sem ruído. O sinal original está mostrado na figura 3.5a e foi fornecido pela Analog Devices. As escalas σ_1 e σ_2 calculadas a partir desse sinal foram respectivamente 0,1927 e 0,506. Usando as duas regras empíricas propostas acima, obtemos os valores candidatos para a escala ótima de 0,578 e 0,648. A figura 3.5b mostra o sinal filtrado pela RAD com escala $\sigma=0,648$. A filtragem limpou fortemente o sinal, permitindo inclusive distinguir duas colisões na região final do sinal, difíceis de serem visualizadas no sinal ruidoso original.

A figura 3.5c mostra o sinal filtrado pelo filtro Butterworth de quinta ordem, com frequência de corte de 5 Hz. Esta filtragem elimina as informações contidas nas amplitudes e nos instantes de ocorrência dos pulsos.

	Perona-Malik g_1	Perona-Malik g_2	Tukey g_3
$\sigma_1 = 0,198$	0,12103	0,19114	0,22150
$\sigma_2 = 0,463$	0,19084	0,11196	0,10764
Escala ótima σ^*	0,10565 ($\sigma^* = 0,142$)	0,09269 ($\sigma^* = 0,374$)	0,08968 ($\sigma^* = 0,594$)

Tab. 3.1: Erros RMS obtidos filtrando o sinal da figura 3.4b pela difusão anisotrópica com 50 iterações, com diferentes funções parada-na-aresta e diferentes escalas.

	Perona-Malik g_1	Perona-Malik g_2	Tukey g_3
$\sigma_1 = 0,198$	0,13658	0,18285	0,22145
$\sigma_2 = 0,463$	0,26024	0,12009	0,10351
Escala ótima σ^*	0,10303 ($\sigma^* = 0,099$)	0,09393 ($\sigma^* = 0,382$)	0,07957 ($\sigma^* = 0,593$)

Tab. 3.2: Erros RMS obtidos com 100 iterações.

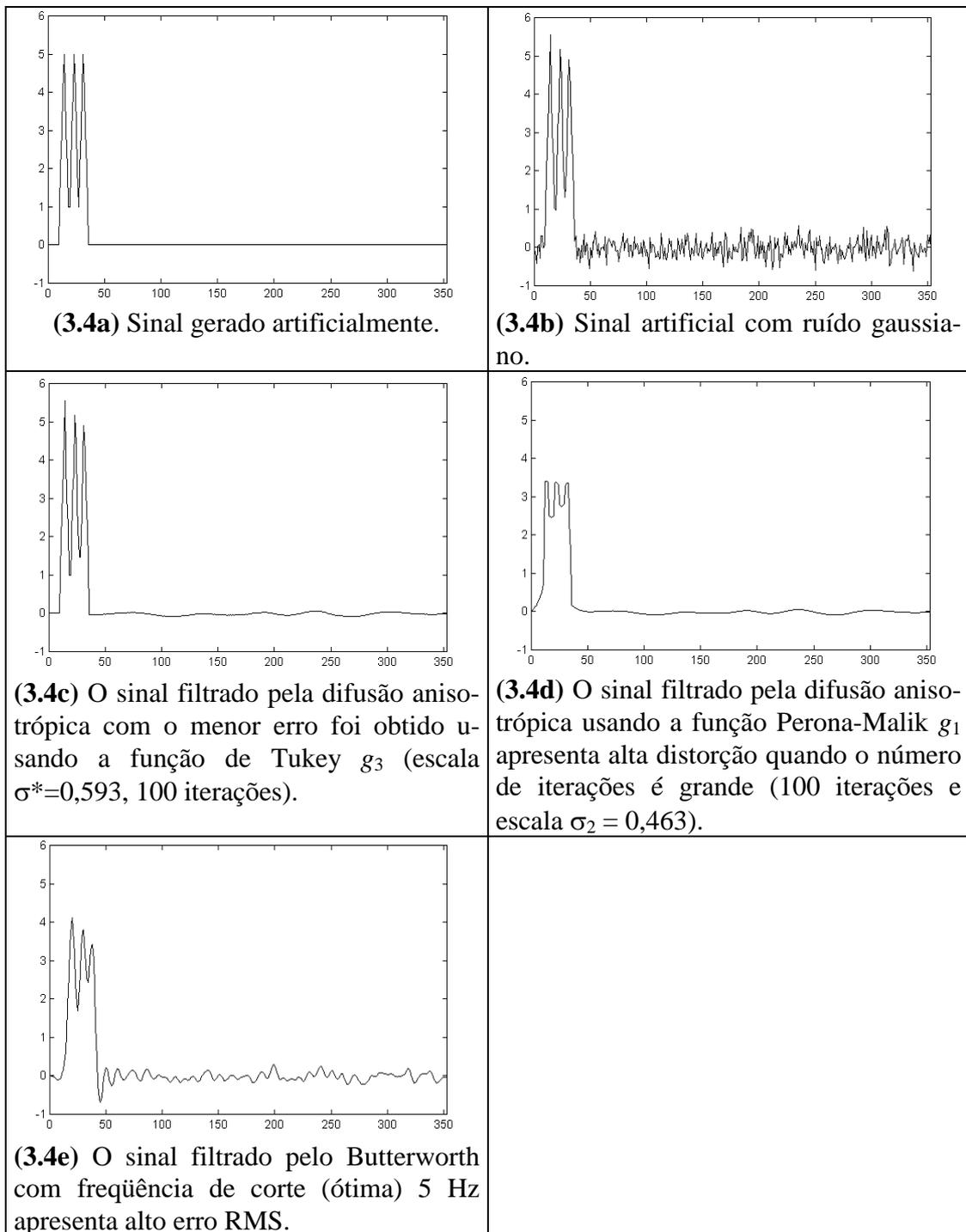


Fig. 3.4: Filtragem de um sinal sintetizado pela difusão anisotrópica.

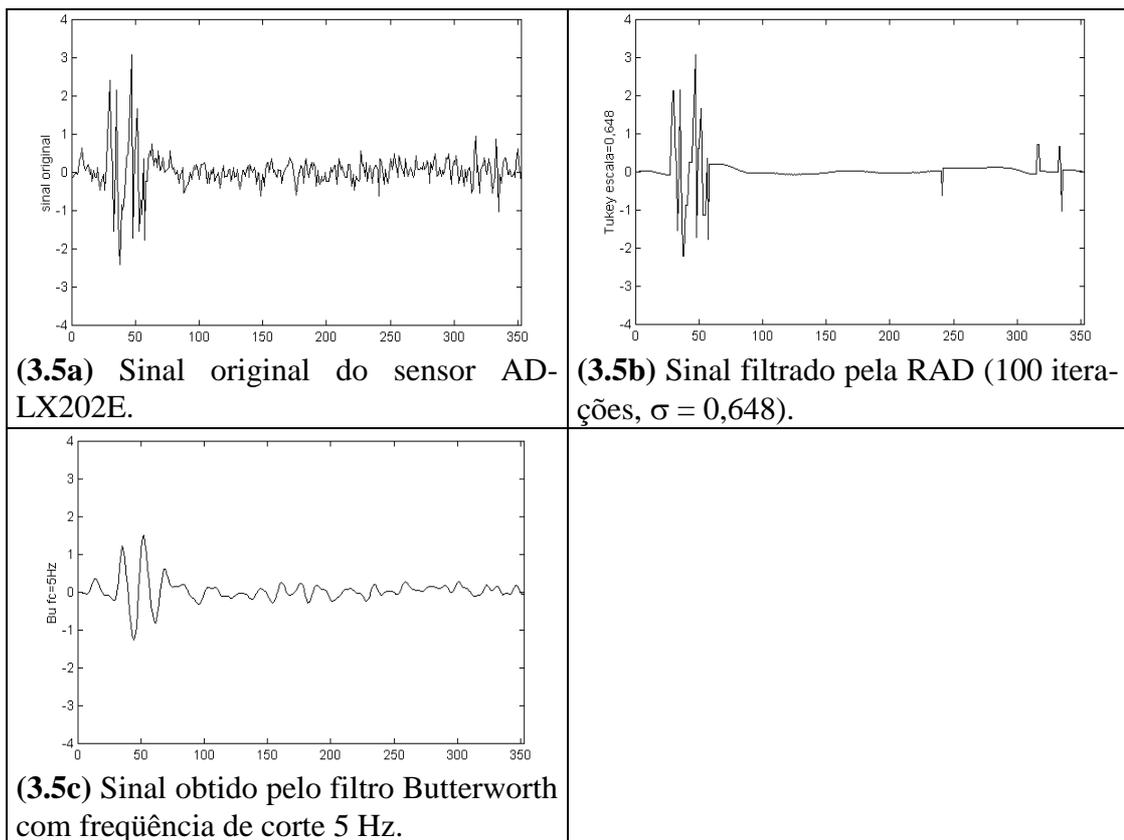


Fig. 3.5: Filtragem do sinal do sensor de aceleração ADLX202E pela difusão anisotrópica e pelo filtro Butterworth.

Detecção de arestas pela difusão anisotrópica

A figura 3.6 mostra a detecção de arestas usando várias funções parada-na-aresta e diferentes escalas σ . O número de iterações foi mantido fixo em $t_{max} = 50$. Uma comparação visual entre as figuras 3.2 e 3.6 permite constatar que a difusão anisotrópica preserva muito melhor a nitidez e a localização das bordas do que a difusão isotrópica.

A figura 3.7 permite constatar a superioridade da função parada-na-aresta de Tukey sobre aquelas de Perona-Malik. Compare as imagens da figura 3.7 (500 iterações) com as imagens da última linha da figura 3.6 (50 iterações). Todas essas imagens foram obtidas usando a escala $\sigma=0,08$. Quando o número de iterações é grande a função g_1 , e em menor grau a função g_2 , borra as arestas. Enquanto isso a RAD (g_3) mantém as arestas perfeitamente nítidas.

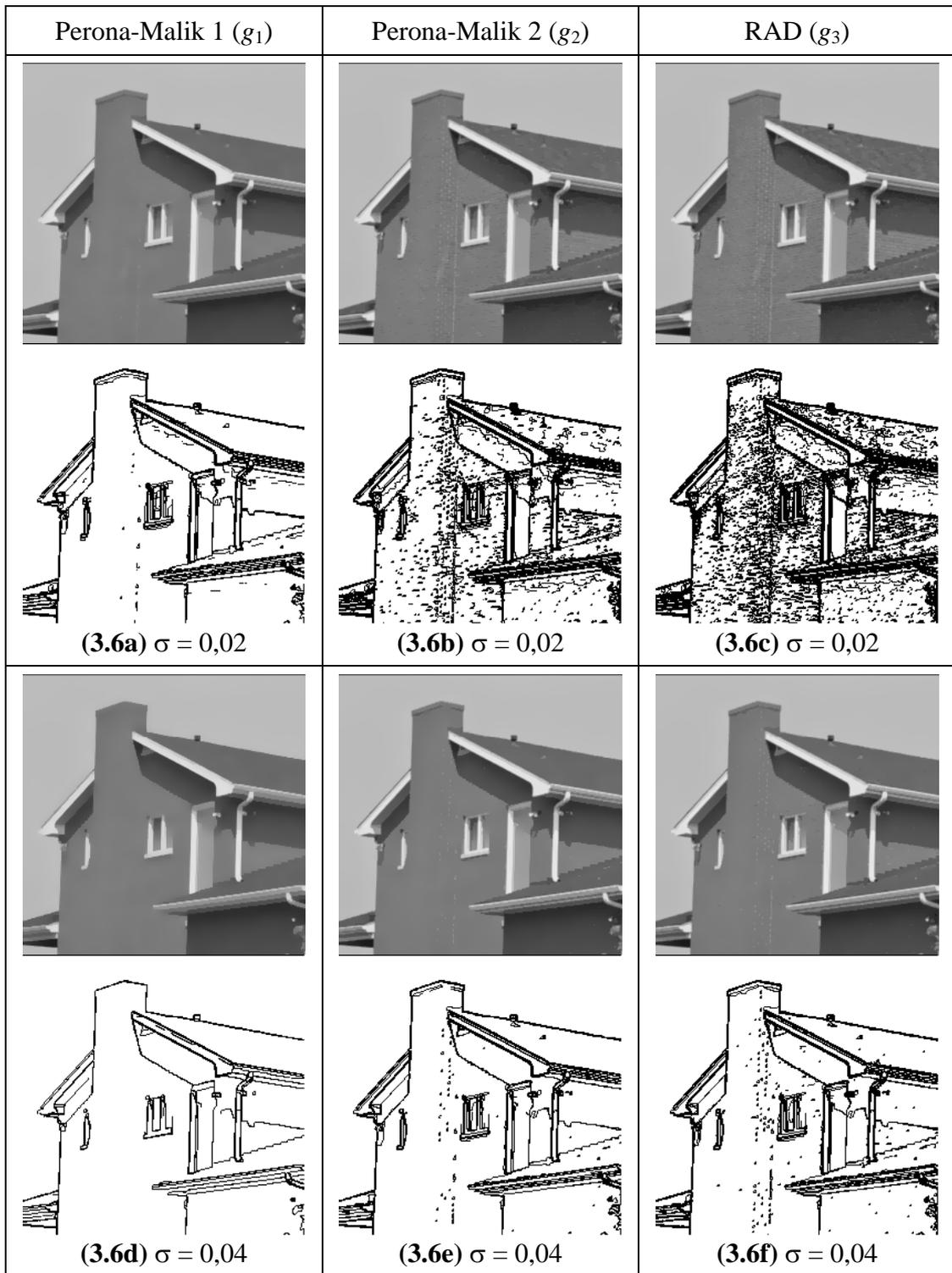


Fig. 3.6: Continua na próxima página.

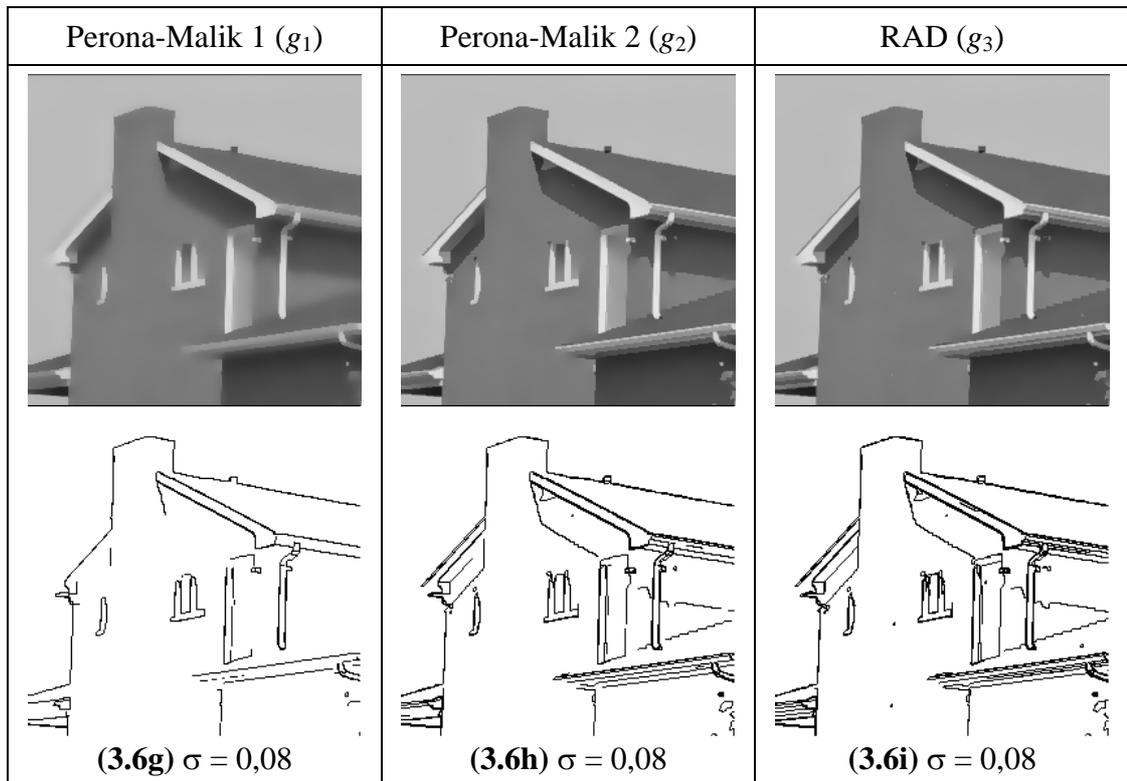


Fig. 3.6: Detecção de arestas usando a difusão anisotrópica com diferentes funções parada-na-aresta e várias escalas σ . O número de iterações foi mantido fixo em $t_{max} = 50$.

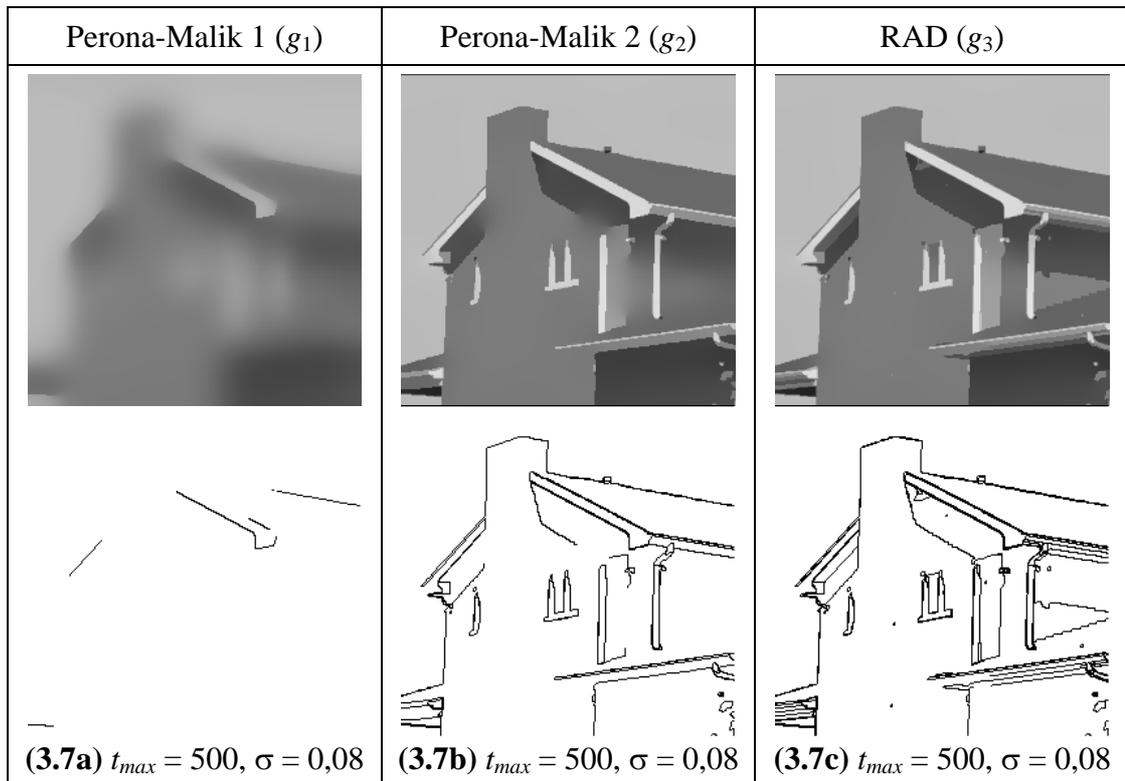


Fig. 3.7: Comportamento da difusão anisotrópica com grande número de iterações ($t_{max} = 500$). A função parada-na-aresta Perona-Malik 1 acaba borrando as arestas. A função de Tukey é a que consegue manter as arestas mais nítidas, pois está baseada na estatística robusta. Compare com a última linha da figura 3.6, onde tínhamos $t_{max} = 50$.

3.4 Melhoramento da Tomografia pela RAD

Introdução

Esta seção descreve uma contribuição científica original nossa. O principal responsável por esta contribuição foi o meu ex-orientando de doutorado Harold I. A. Bustos.

A reconstrução de imagem (ou tomografia) é a técnica usada para obter a distribuição de um meio não observável diretamente através das projeções. Existem muitas técnicas de reconstrução tomográfica, por exemplo, a retro-projeção filtrada, a transformada de Fourier, a reconstrução aritmética [Jain, 1989], a entropia máxima [Reis and Roberty, 1992; Dusassoy and Abdou, 1991], etc. Aqui, estamos interessados nas técnicas tomográficas que geram boas imagens reconstruídas mesmo usando uma pequena quantidade de dados de projeção. Em muitas situações práticas, é vantajoso minimizar a irradiação, para não danificar a amostra sendo irradiada ou para apressar a aquisição dos dados. O algoritmo de máxima entropia (MENT) é uma das melhores técnicas quando somente poucos dados estão disponíveis. Porém, mesmo este método produz imagens borradas com ruídos e artefatos numa situação com dados escassos. Delaney e Bresler [Delaney and Bresler, 1998] propuseram uma técnica tomográfica especialmente projetada para gerar imagens claras a partir de poucos dados. Porém, na prática este algoritmo requer uma quantidade bastante grande de dados para gerar uma imagem nítida, conforme discutiremos mais adiante.

Conforme vimos nas seções anteriores, a difusão anisotrópica é uma técnica bem conhecida usada para filtragem, detecção de arestas e análise multi-escala de imagens. Recentemente, Black et al. descreveram a relação entre a difusão anisotrópica e a estatística robusta, resultando numa técnica teoricamente bem fundamentada denominada difusão anisotrópica robusta (RAD) [Black et al., 1998].

Nesta pesquisa, melhoramos o algoritmo MENT utilizando a RAD. A idéia principal do novo algoritmo é intercalar, em cada passo iterativo do MENT, uma filtragem

RAD. Esta idéia é completamente diferente de simplesmente pós-filtrar com a RAD uma imagem gerada pelo MENT. Como o algoritmo MENT gera imagens muito ruidosas e borradas numa situação de poucos dados, nenhuma pós-filtragem consegue melhorar substancialmente a qualidade da imagem gerada. Porém, incorporando a filtragem RAD no algoritmo MENT, imagens nitidamente reconstruídas podem ser obtidas mesmo usando poucos dados de projeção.

Descrevemos as nossas idéias para um cenário de aquisição de dados em feixe paralelo e ângulo completo (180°) sem nenhum dado em falta. Implementamos e testamos a técnica proposta somente para esta situação. Porém, as idéias desenvolvidas aqui podem ser aplicadas de forma direta para qualquer circunstância de aquisição de dados: feixe paralelo ou em leque, ângulo completo ou limitado, com ou sem dados em falta. Testamos o nosso algoritmo usando uma quantidade extremamente pequena de dados. Sob esta condição, o algoritmo MENT original gera imagens ruidosas e borradas, onde mesmo as arestas importantes não são claramente reconstruídas. O nosso algoritmo gerou imagens nítidas.

Algoritmo de Máxima Entropia Estendida

Minerbo propôs o algoritmo MENT em [Minerbo, 1979] e depois muitos autores propuseram diferentes melhoramentos a este algoritmo. Por exemplo, Dusassoy e Abdou [Dusassoy and Abdou, 1991] introduziram o algoritmo MENT-estendido que pode levar em conta uma informação a priori sobre a imagem a ser reconstruída. Isto é, se uma aproximação f^* da imagem a ser reconstruída f for conhecida, este conhecimento pode ser usado para melhorar a reconstrução.

O funcional de Lagrange abaixo representa o custo funcional do MENT-estendido, sujeito às restrições dos dados de projeção:

$$L(f, \Lambda) = - \iint f(x, y) \log \left[\frac{f(x, y)}{ef^*(x, y)} \right] dx dy \\ - \sum_{j=1}^J \sum_{n=1}^N \Lambda_{j,n} \left[h_{j,n} - \iint f(x, y) \chi_{j,n}(x, y) dx dy \right]$$

onde:

- e é a base neperiana (2,71828...).
- $h_{j,n}$ é a intensidade do n -ésimo raio da projeção j .
- $\Lambda_{j,n}$ é o parâmetro de Lagrange associado à faixa (j, n) . Se os dados fossem completos, este coeficiente seria sempre um.
- $\chi_{j,n}$ é a função indicadora da faixa (j, n) . Esta função é 1 dentro da faixa (j, n) e 0 fora.
- $f^*(x, y)$ é o modelo a priori do objeto $f(x, y)$. Se nenhuma informação a priori estiver disponível, $f^*(x, y)$ pode receber o valor e^{-1} . Neste caso, MENT-estendido torna-se o algoritmo original MENT de Minerbo.

A otimização da equação acima permite-nos achar a solução do problema de reconstrução:

$$f^{(i)}(x, y) = f^*(x, y) \prod_{j=1}^J \sum_{n=1}^N F_{j,n}^{(i)} \chi_{j,n}(x, y), \quad 1 \leq i \leq \zeta \quad (3.3)$$

onde $F_{j,n}$ é a matriz dos parâmetros duais de Lagrange associada à faixa (j, n) . Esses parâmetros são obtidos pelo seguinte sistema iterativo:

$$F_{j,n}^{(i)} = \begin{cases} z h_{j,n}, & i = 0 \\ \frac{h_{j,n}}{\iint f^*(x, y) \prod_{\substack{k=1 \\ k \neq j}}^J \sum_{n=1}^N [F_{k,n}^{(i-1)} \chi_{k,n}(x, y)] \chi_{j,n}(x, y) dx dy}, & 1 \leq i \leq \zeta \end{cases} \quad (3.4)$$

onde z é a largura da faixa (j, n) . Após calcular $F_{j,n}^{(i)}$, eles devem ser inseridos em (3.3) para achar a imagem reconstruída na i -ésima iteração $f^{(i)}(x, y)$.

MENT Reconstrução-Difusão

Vamos supor que a imagem f a ser reconstruída seja constante por regiões. Conforme notamos antes, numa situação com poucos dados, o algoritmo MENT-estendido irá

reconstruir uma versão ruidosa e borrada da imagem f . Esta imagem pode estar tão fortemente corrompida que nenhuma pós-filtragem pode melhorá-la substancialmente. A RAD é um excelente estimador da imagem original f a partir da sua versão corrompida. Porém, utilizando-a como um processo de pós-filtragem, somente um ligeiro melhoramento pode ser obtido. A nossa idéia é usar esta imagem ligeiramente melhorada como o conhecimento a priori f^* do algoritmo MENT-estendido. Isto irá gerar uma imagem reconstruída de melhor qualidade. Esta imagem melhorada pode ser melhorada ainda mais pela RAD e usada como um novo conhecimento a priori pelo MENT-estendido, e assim por diante.

Inicialmente, aplicamos o algoritmo MENT-estendido, iterando a equação (3.4) ζ vezes. Usando a equação (3.3), a imagem reconstruída $f^{(\zeta)}$ é obtida. Esta imagem é filtrada pela RAD, iterando a equação (3.2) uma ou mais vezes. A imagem filtrada é usada como conhecimento a priori f^* pelo MENT-estendido para obter uma nova imagem reconstruída. Esta imagem é novamente filtrada pela RAD, e assim por diante.

Resultados Experimentais

Para mostrar a eficácia da técnica proposta, executamos alguns experimentos. O objeto teste simulado é um cilindro com diâmetro 100 e densidade 5, imerso no meio com densidade 0 (figura 3.8a). Este cilindro contém 5 cilindros menores com diâmetros e densidades variadas.

Experiências consistiram em reconstruir imagens 100×100 a partir de apenas 6 projeções paralelas distribuídas em ângulo completo (180°), cada projeção com 100 raios irradiados. A imagem 3.8b foi obtida pelo algoritmo original MENT-estendido, iterando 10 vezes a equação 3.4. Esta imagem foi filtrada pela RAD ($\sigma=32$, 100 iterações), gerando a imagem 3.8c. Esta imagem foi utilizada como a estimativa inicial do MENT reconstrução-difusão. A reconstrução-difusão (isto é, uma execução de (3.4) seguida por uma execução de (3.2)) foi iterada 9 vezes (com $\sigma=32$), gerando a imagem 3.8d. Sem dúvida, o novo algoritmo gerou uma imagem melhor. As médias das

diferenças absolutas entre a imagem ideal 3.8a e as imagens 3.8b, 3.8c, 3.8d foram respectivamente 12,7%, 12,8% e 7,5%.

Para mostrar a superioridade da nossa proposta, vamos definir uma projeção reconstruída $h_{j,n}^{(i)}$, calculada a partir da imagem reconstruída na i -ésima iteração $f^{(i)}$ como:

$$h_{j,n}^{(i)} = \iint f^{(i)}(x, y) \chi_{j,n}(x, y) dx dy .$$

Vamos definir a norma euclidiana $k^{(i)}$ entre as projeções originais $h_{j,n}$ e as projeções calculadas a partir da imagem reconstruída $h_{j,n}^{(i)}$ como:

$$k^{(i)} = \sqrt{\sum_{j=1}^J \sum_{n=1}^N (h_{j,n}^{(i)} - h_{j,n})^2} .$$

A figura 3.9 mostra as normas euclidianas $k^{(i)}$ em diferentes iterações da reconstrução. As primeiras 10 iterações correspondem ao MENT-estendido original e as últimas 9 iterações ao MENT reconstrução-difusão proposto. Sem dúvida, a norma euclidiana converge mais rapidamente utilizando o algoritmo proposto. O parâmetro de escala $\sigma=32$ foi escolhido para maximizar a convergência da norma euclidiana.

Um outro fantom gerado artificialmente está ilustrado na figura 3.10a. Foram irradiadas seis projeções paralelas distribuídas sobre 180° com 100 raios por projeção (600 raios ao todo). O algoritmo MENT-estendido original foi executado sobre estes dados, gerando a imagem 3.10b. Esta imagem foi filtrada pela RAD ($\sigma=50$, 70 iterações), gerando a imagem 3.10c. Olhando esta figura, fica claro que uma pós-filtragem não consegue gerar uma imagem reconstruída nítida. A imagem 3.10c foi usada como a estimativa inicial da imagem a ser reconstruída pela MENT reconstrução-difusão. Este algoritmo foi iterado 70 vezes usando parâmetro de escala $\sigma=50$, gerando a imagem 3.10d. Sem dúvida, o algoritmo proposto gerou a melhor imagem. Todas as imagens têm resolução de 100×100 pixels. As médias das diferenças absolutas entre a imagem ideal 3.10a e as imagens 3.10b, 3.10c e 3.10d foram respectivamente 14,89%, 15,22% e 8,15%. Usando o bem-conhecido algoritmo de retro-

projeção filtrada sobre os mesmos dados, a imagem de baixa qualidade 3.10e foi obtida.

Devido à relevância do trabalho de Delaney e Bresler [Delaney and Bresler, 1998], comparamos rapidamente os nossos resultados com os deles. As experiências descritas em [Delaney and Bresler, 1998] coletam uma projeção a cada grau, em ângulo completo ($-90^\circ, 90^\circ$) ou limitado ($-75^\circ, 75^\circ$), com somente um ou dois ângulos em falta. No nosso caso, uma projeção paralela é coletada a cada 30° , em ângulo completo. Portanto, o nosso problema é muito mais severamente subdeterminado que o problema considerado por Delaney e Bresler. O nosso algoritmo pode reconstruir imagens nítidas a partir dos dados de projeção altamente subdeterminados.

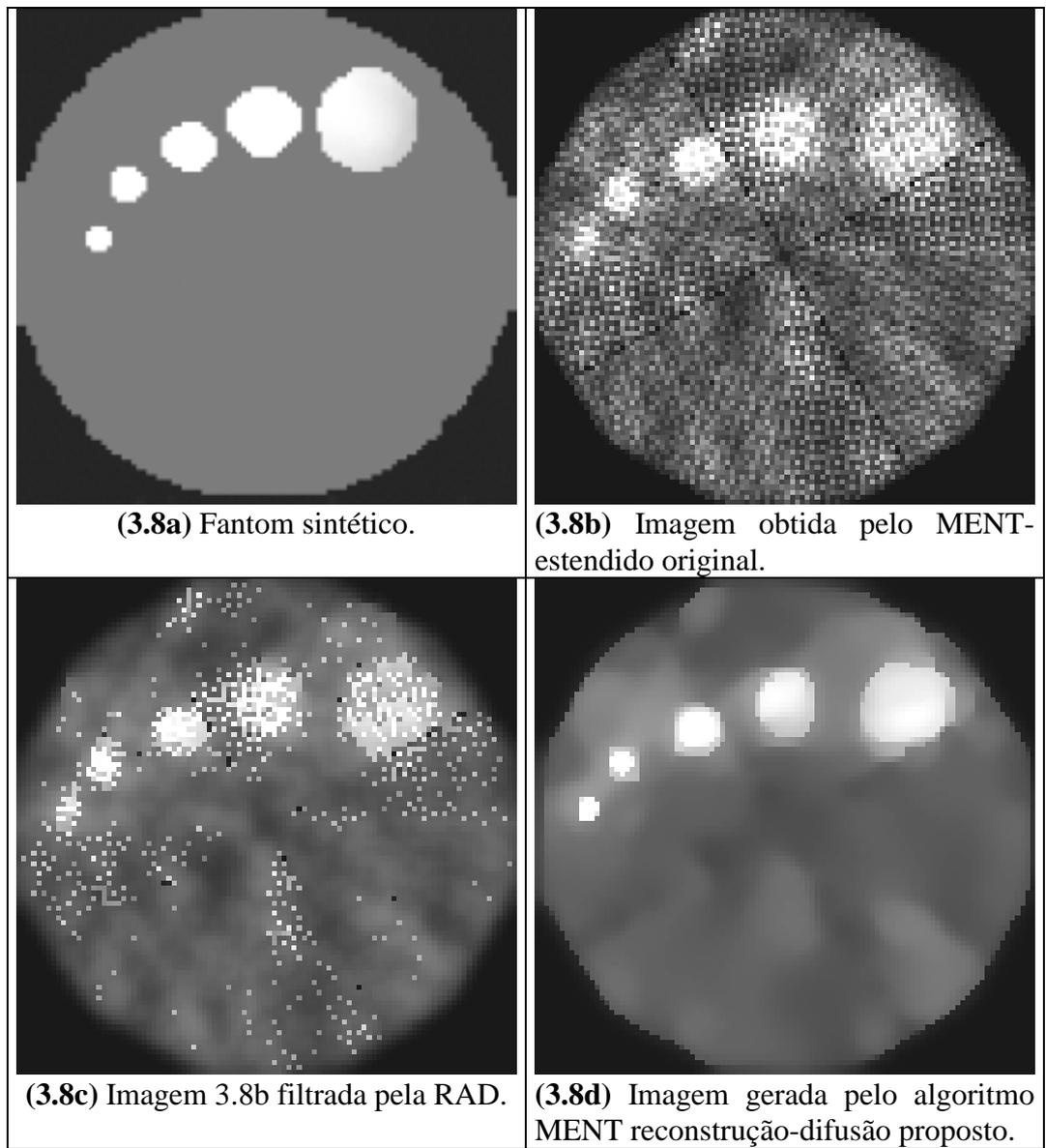


Fig. 3.8: Comparação do algoritmo MENT-estendido com o MENT reconstrução-difusão.

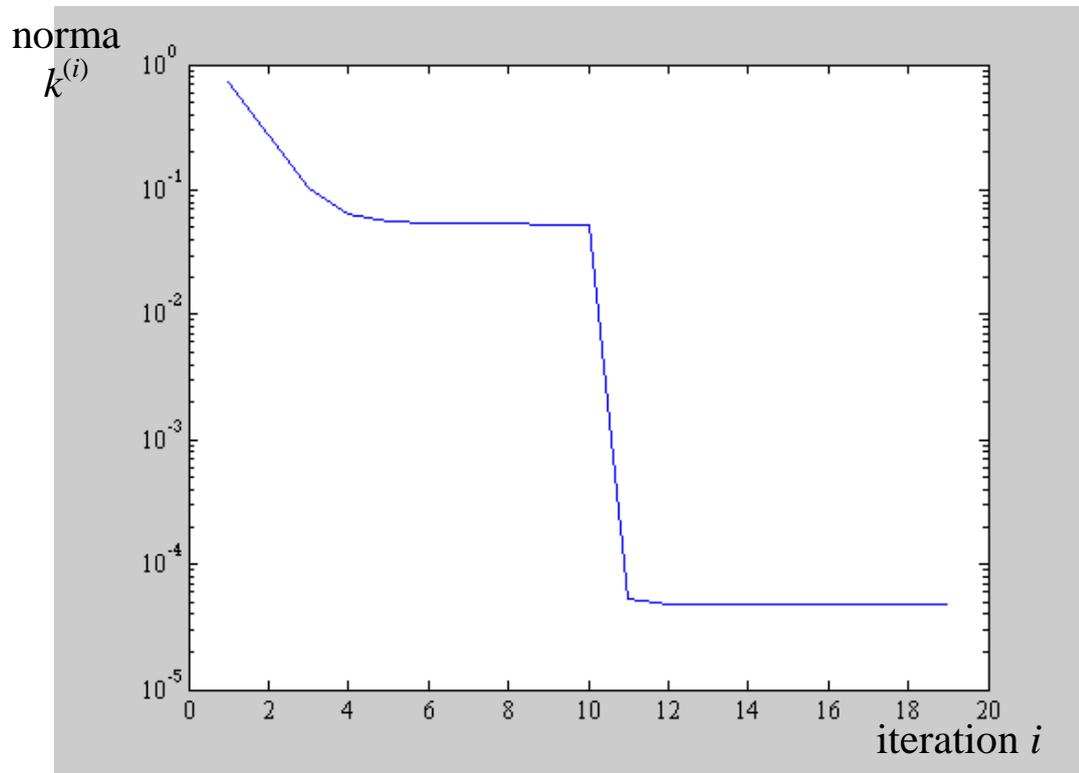


Fig. 3.9: Norma euclidiana $k^{(i)}$ da diferença entre as projeções originais e as projeções calculadas a partir das imagens reconstruídas na i -ésima iteração. As primeiras 10 iterações correspondem ao algoritmo MENT estendido original e as últimas 9 iterações correspondem ao MENT reconstrução-difusão proposto.

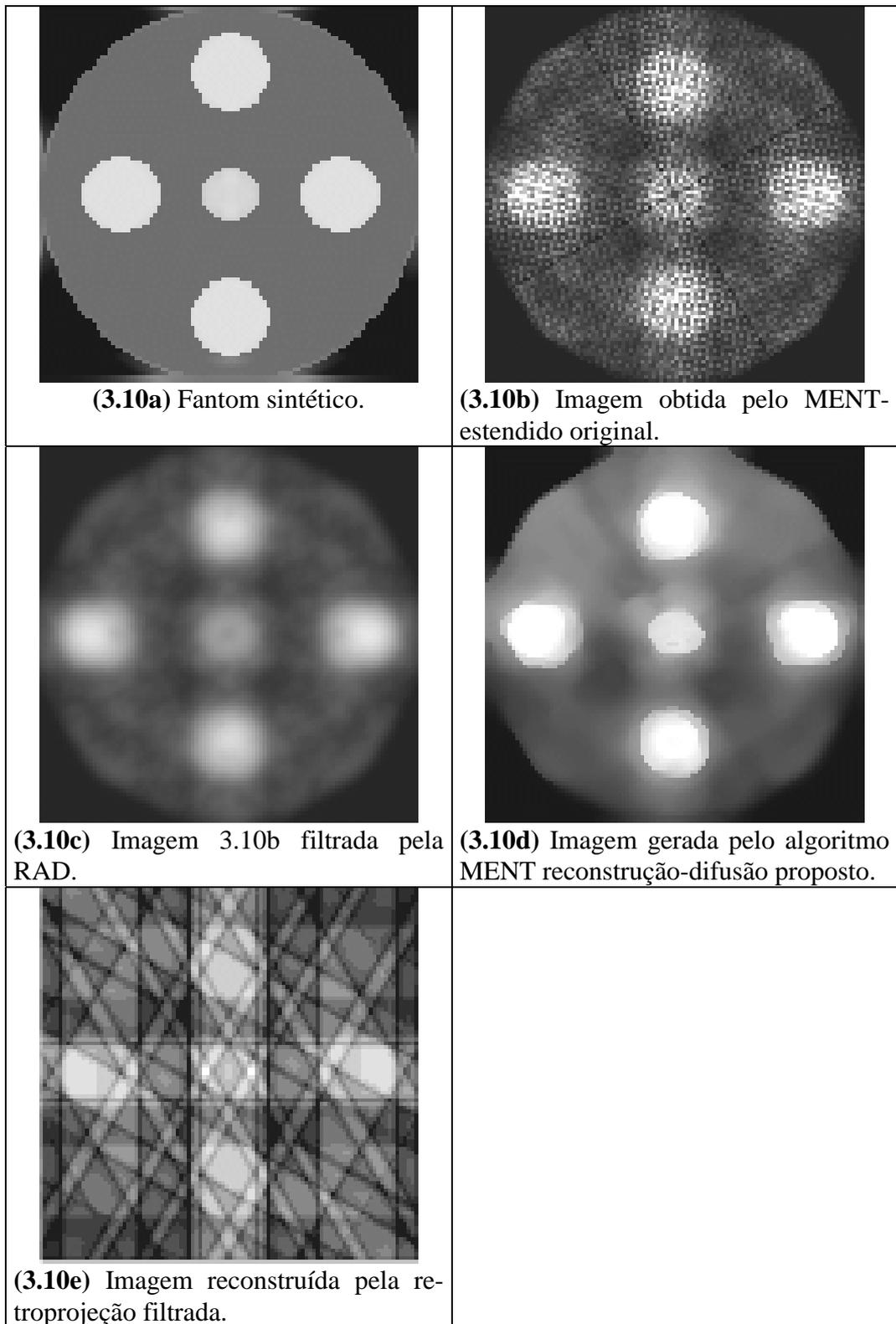


Fig. 3.10: Comparação entre os algoritmos MENT-estendido, MENT reconstrução-difusão e retroprojeção filtrada, utilizando os mesmos dados.

3.5 Melhoramento da fMRI pela RAD

Introdução

Esta seção descreve uma contribuição científica original minha. Eu fui o principal responsável por esta contribuição, e o trabalho foi realizado em colaboração com Prof. H. Z. Cho da University of California, Irvine.

O objetivo do imageamento funcional é mapear as atividades de um cérebro vivo no espaço e no tempo. O “padrão ouro” para medir a atividade celular cerebral envolve a gravação, direta e invasiva, do potencial elétrico da membrana dos neurônios individuais. Porém, tais medidas são limitadas a certas condições experimentais. Para estudos em sujeitos humanos, os métodos não-invasivos tais como PET (positron emission tomography) ou fMRI (functional magnetic resonance imaging) devem ser aplicados.

Desde o início dos anos 80, a PET dominou o campo de neuroanatomia funcional, mas nos últimos 10 anos a fMRI desenvolveu-se numa técnica alternativa e poderosa. Aumentos locais na atividade neuronal provocam a relativa desoxigenação do sangue e o aumento da perfusão, que rapidamente reverte a desoxigenação inicial, levando ao aumento da oxigenação que dura durante vários segundos. A fMRI é sensível a oxigenação do sangue (efeito normalmente abreviado como BOLD - blood oxygen level dependent) e possui a escala espaço-temporal de aproximadamente 1-3 mm e 1 ou mais segundos. Os limites inferiores da resolução efetiva da fMRI são fisiológicos e impostas pela organização espaço-temporal da resposta hemodinâmica evocada (2-5 mm e 5-8 segundos).

Em virtude da meia-vida relativamente longa dos traçadores radioativos utilizados, a PET pode medir somente respostas somadas durante um período de tempo razoavelmente longo. Em contraste, a fMRI pode ser usada de segundo paradigmas periódico (estímulos repetidos) ou relacionado a eventos. No paradigma periódico, o sujeito se

alterna entre os períodos de estímulo e descanso. Neste caso, cada voxel da fMRI consiste numa série temporal que pode ser dividido em “épocas”, sendo cada época o período de tempo que corresponde à atividade seguida pelo descanso. No paradigma relacionado a eventos, o sujeito executa a atividade durante somente um período curto de tempo.

A baixa razão sinal-ruído nas imagens fMRI obriga o uso de sofisticadas técnicas de Processamento e Análise de Imagens para detectar as áreas ativadas do cérebro. Em primeiro lugar, os dados devem passar através de transformações espaciais para corrigir o movimento da cabeça do sujeito durante a aquisição de fMRI. Se a experiência envolver sujeitos diferentes, os dados devem além disso ser normalizados, isto é, as imagens devem ser arqueadas de forma que todas elas se conformem a algum cérebro padrão.

Depois das transformações espaciais, as análises estatísticas são efetuadas. Muitos procedimentos estatísticos diferentes foram propostos para analisar os dados fMRI dependentes do nível de oxigenação [Lange et al., 1999; Gold et al., 1998]. Um dos procedimentos estatísticos mais populares é o modelo linear geral [Friston et al., 1995]. Neste modelo, o usuário especifica manualmente uma “matriz de projeto” (design matrix) e faz uso da regressão linear múltipla para estimar os parâmetros, isto é, determinar quão bem a série temporal de cada voxel se encaixa dentro da matriz de projeto especificada. Estes parâmetros são então utilizados para computar a significância estatística de um efeito. Estas estatísticas, dispostas espacialmente, formam o mapa estatístico paramétrico (SPM - statistical parametric map). As notas de curso [Friston, 1997] são uma boa referência sobre o processamento de fMRI através do modelo linear geral.

Mesmo com todos esses aparatos de processamento de imagens, uma fMRI ruidosa sempre dá origem a um SPM ruidoso. Os filtros passa-baixas simples não podem ser usados indiscriminadamente pois eles borram as arestas das áreas ativadas. As técnicas tradicionais de filtragem que preservam as arestas também não podem ser usadas pois não existe uma fronteira clara entre as áreas ativadas e não-ativadas.

Na literatura existem muitos trabalhos para atenuar o ruído e aglutinar as regiões ativadas nos dados fMRI [Goutte et al., 1999; Ardekani and Kanno, 1998; Kershaw et al., 1999; Chuang et al., 1999; Friston et al., 1994]. Em particular, Solé et al. [Solé et al., 2001] propuseram recentemente uma técnica denominada “média anisotrópica” (anisotropic averaging). Esta técnica foi inspirada na difusão anisotrópica, introduzida por Perona e Malik [Perona and Malik, 1990]. A média anisotrópica calcula um conjunto inicial de voxels claramente ativados utilizando os coeficientes de correlação. Este conjunto é então utilizado para construir uma complexa “medida de similaridade” para calcular os pesos da média ponderada. Apesar de Solé et al. tentar explicar a definição da sua medida de similaridade com argumentos intuitivos, somos impelidos a perguntar se não existiria uma forma mais natural e simples de definir essa medida. Além disso, a sua técnica pode ser usada somente para processar fMRI periódica, pois a sua medida de similaridade está baseada na transformada de Fourier da série temporal de cada voxel. Num fMRI relacionado a eventos, a transformada de Fourier não faz o mínimo sentido.

Nos trabalhos [Cn12; Su01], propusemos uma outra técnica para obter um SPM nítido a partir de fMRI ruidosa utilizando o modelo linear geral. Em vez de definir uma medida de similaridade altamente complexa baseada no conjunto de voxels claramente ativados, usamos a magnitude do gradiente dos parâmetros estimados como argumentos para calcular os coeficientes de difusão. Substituímos a média anisotrópica pela difusão anisotrópica robusta [Black et al., 1998]. Esta técnica pode ser usada para processar fMRI tanto periódico como relacionado a eventos.

Modelo Linear Geral

O modelo linear geral é simplesmente uma equação que relaciona o que se observa com o que se esperaria observar, expressando as observações como uma combinação linear dos componentes esperados e algum erro residual. O modelo linear geral pode ser escrito como [Friston et al., 1995]:

$$Y = X\beta + \varepsilon .$$

Geralmente, todas as variáveis envolvidas na equação acima são matrizes. Porém, para simplificar a exposição, iremos supor que Y é um vetor coluna das observações, β é um vetor coluna dos parâmetros, e ε é um vetor coluna dos erros. X é a matriz de projeto (design matrix) com uma linha por observação e duas colunas: A primeira coluna é o parâmetro do modelo e a segunda coluna é “fantoche”, inteiramente preenchido com 1, cuja finalidade é corrigir a média das observações. O modelo linear geral assume que os erros ε_j são variáveis aleatórias normais, independentes e identicamente distribuídas.

A seguinte equação executa a estimação de mínimos quadrados dos parâmetros:

$$\hat{\beta} = (X^T X)^{-1} X^T Y.$$

$\hat{\beta}$ é na verdade um vetor coluna com duas linhas. Porém, somente a primeira coluna $\hat{\beta}_1$ é útil, pois $\hat{\beta}_2$ é “fantoche”. Chamaremos a imagem obtida dispendo espacialmente os parâmetros obtidos $\hat{\beta}_1$ como EPM (estimated parameters map). Um EPM B pode ser transformado num SPM (statistical parametric map) através de alguns cálculos. Por exemplo, dividindo o valor de um voxel de B pelo seu respectivo erro padrão, a estatística t de Student é obtida. O mapa paramétrico estatístico das estatísticas t de Student é denotado como SPM{t}.

O seguinte exemplo numérico clarifica essas idéias:

$$\begin{bmatrix} 50 \\ 51 \\ 60 \\ 62 \\ 51 \\ 52 \\ 62 \\ 63 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \end{bmatrix}$$

A primeira matriz Y é uma série temporal dos valores de um voxel. Vamos supor que os valores foram adquiridos a cada 5 segundos. A segunda matriz X é a matriz de projeto. A primeira coluna indica que gostaríamos de detectar uma ativação que se repete a cada 20 segundos e dura 10 segundos. A segunda coluna é “fantoche”, colocada na matriz de projeto para corrigir a média dos valores de Y . Estimando os parâmetros, obtemos $\hat{\beta}_1 = 10,75$ e $\hat{\beta}_2 = 51,00$. O alto valor de $\hat{\beta}_1$ indica que muito provavelmente este voxel está ativado.

Para transformar $\hat{\beta}_1$ numa distribuição t de Student t_1 com r graus de liberdade, aplicamos a equação abaixo:

$$t_1 = \frac{\hat{\beta}_1}{\sqrt{\varepsilon^T \varepsilon / r}} = 10,947$$

onde $r = 7$ é a quantidade de observações menos um. Isto significa que, se o voxel considerado fosse completamente não relacionado com a matriz de projeto, existe uma probabilidade $5,9 \times 10^{-6}$ de resultar um parâmetro maior que 10,75 por acaso.

Média Anisotrópica

A difusão anisotrópica foi aplicada com sucesso para MRI estrutural [Gerig et al., 1992]. Porém, esta técnica não pode ser aplicada diretamente em fMRI, principalmente devido à falta de bordas claras entre as regiões ativadas e não-ativadas. Também não pode ser aplicada diretamente em SPM, pois SPM é normalmente muito ruidoso e filtrá-lo pela difusão anisotrópica somente aumentaria as imperfeições do SPM.

Fazer uma convolução da fMRI com uma máscara para suavização geralmente aumenta a razão sinal a ruído. Porém, este procedimento também borra as arestas entre as regiões ativadas e não-ativadas. Portanto, é desejável efetuar somente a suavização intra-região, evitando a suavização inter-região.

Solé et al. [Solé et al., 2001] propuseram esta idéia e chamaram-na de média anisotrópica (anisotropic averaging). Consiste em calcular uma média seletiva da vizinhança do sinal. Seja I uma imagem fMRI e seja $I(s, n)$ o valor do voxel na posição espacial s e no volume (isto é, na aquisição ou na observação) $n \in [1 \dots N]$. A média anisotrópica irá aos poucos modificar I . Vamos denotar imagem fMRI na iteração $t \geq 0$ como $I(s, n, t)$, sendo $I(s, n, 0) = I(s, n)$. Então, a média aritmética consiste em calcular:

$$I(s, n, t+1) = \frac{1}{\sum_{p \in \eta_s} w(s, p)} \sum_{p \in \eta_s} w(s, p) I(p, n, t),$$

para todos os voxels s , todos os volumes n e as iterações $t \geq 0$. Como antes, η_s representa o conjunto dos vizinhos espaciais do voxel s .

Seja $I(s)$ a série temporal do sinal no voxel s . A medida de similaridade Ψ determina os pesos $w(s, p)$, utilizando a série temporal $I(s)$ que sofrerá o processo de média anisotrópica e a série temporal do seu voxel vizinho $I(p)$:

$$w(s, p) = \Psi(I(s), I(p)).$$

Esta medida de similaridade permite-nos distinguir voxels ativados daqueles não-ativados. Permite-nos calcular uma média seletiva, combinando somente os sinais da mesma classe. Solé et al. propuseram calcular um conjunto inicial Ω de voxels claramente ativados escolhendo os voxels com alto coeficiente de correlação com a matriz de projeto. Depois, os espectros de Fourier dos voxels em Ω são calculados para definir a função de medida de similaridade Ψ . O espectro de Fourier da série temporal de cada voxel s é também calculado para avaliar a similaridade entre s e os voxels em Ω . O procedimento todo é altamente complexo e os leitores são encaminhados a [Solé et al., 2001] para maiores detalhes. Gostaríamos de perguntar: “todos esses cálculos são realmente necessários?” Mais, “como esta técnica pode ser aplicada no protocolo de aquisição de fMRI relacionado a eventos?”

Algoritmos

Propomos uma abordagem diferente, motivada diretamente pela difusão anisotrópica robusta, para filtrar fMRI. O nosso método é mais simples, pode ser aplicado a fMRI relacionado a eventos, e tem gerado SPMs surpreendentemente nítidos. Este método também aumentou a significância estatística do SPM, o que nos permite decidir com mais confiança se um voxel está ativado ou não. A nossa técnica está descrita abaixo.

Seja dada uma fMRI I e uma matriz de projeto X . Seja $I(s, n)$ o valor de I na posição espacial s e no volume $n \in [1..N]$. Vamos denotar a imagem fMRI na iteração $t \geq 0$ como $I(s, n, t)$, com $I(s, n, 0) = I(s, n)$.

Usando a fMRI I e a matriz de projeto X , estime os parâmetros $\hat{\beta}$ para cada voxel s , como descrevemos anteriormente. Estes parâmetros, dispostos espacialmente, formam o EPM B . Vamos denotar como $B(s, t)$ o valor de B no voxel s e iteração t . A magnitude do gradiente de $B(s, 0)$ será usada como o argumento da função “parada-na-aresta” g para calcular os coeficientes de difusão $g\left(|\nabla B_{s,p}(0)|\right)$ no instante $t = 0$, onde:

$$\nabla B_{s,p}(t) = B(p, t) - B(s, t), \quad p \in \eta_s.$$

Estes coeficientes são usados para executar a difusão em fMRI $I(s, n, 0)$, gerando a fMRI difundida $I(s, n, 1)$ no instante $t = 1$. $I(s, n, 1)$ é então usada para estimar o novo EPM $B(s, 1)$ na iteração $t = 1$. Estes passos são repetidos até que a média do valor difundido esteja abaixo de algum limiar predefinido. É também possível especificar o número de iterações desejado, em vez de definir um limite para o valor médio difundido. A seguinte equação descreve este processo:

$$I(s, n, t+1) = I(s, n, t) + \frac{\lambda}{|\eta_s|} \sum_{p \in \eta_s} g\left(|\nabla B_{s,p}(t)|\right) \nabla I_{s,p}(t),$$

para todos os voxels s , todos os volumes n e as iterações $t \geq 0$.

A melhor função “parada na aresta” g é a função biweight de Tukey. Note que a escolha correta do parâmetro de escala σ da função de Tukey é essencial para gerar bons resultados.

Quando o processo de difusão terminar em alguma iteração t_f , o EPM $B(s, t_f)$ pode ser transformado em SPM{t} usando o procedimento descrito anteriormente.

Resultados e Dados Experimentais

A figura 3.11 mostra uma parte de um fantom fMRI simulado com $10 \times 10 \times 3$ voxels por volume e 64 volumes. Todos valores dos voxels eram 500 originalmente. O ruído gaussiano com média zero e desvio-padrão 10 foi somado aos valores originais. Os volumes 3, 4, 7, 8, 11, 12, ... têm um quadrado 6×6 ativado no centro do volume, com dois furos não-ativados de 4 voxels cada um. Os voxels ativados tiveram seus valores aumentados de 20.

O SPM{t} obtido pelo modelo linear geral sem filtragem é apresentado na coluna à esquerda da figura 3.12. Entre os voxels ativados, o menor valor foi 1,48 e o maior 2,78. Usando a distribuição t de Student com 63 graus de liberdade, podemos inferir que existe uma probabilidade 7,2% de um voxel não-ativado assumir valores maiores que 1,48 por acaso. Entre os voxels não-ativados, o menor e o maior valores foram -0,93 e 0,74, respectivamente.

O SPM{t} obtido filtrando fMRI com o método proposto (usando a função de Tukey com $\sigma = 10$) é mostrado na coluna direita da figura 3.12. Note que o SPM filtrado é completamente sem ruído, e as arestas estão perfeitamente preservadas. Todos os voxels ativados apresentaram valores aproximadamente 12,1, significando que a confiança estatística melhorou consideravelmente. Virtualmente, é impossível que um voxel não-ativado assumira valores tão altos por acaso (probabilidade menor que 10^{-16}). Os valores dos voxels não-ativados ficaram no intervalo de -0,22 a 0,45.

A figura 3.13 mostra fMRI real com $79 \times 95 \times 68$ voxels por volume e 12 volumes. Os volumes foram realinhados para corrigir o movimento da cabeça do paciente. Depois,

os volumes foram embaralhados aleatoriamente para remover qualquer sinal de ativação que possa estar presente. Ativamos artificialmente pequenas regiões esféricas nos volumes 3, 4, 7, 8, 11 e 12 aumentando o valor dos voxels em 3%. A coluna esquerda da figura 3.13 mostra três fatias do volume 1, sem ativação. A coluna direita da figura 3.13 mostra três fatias do volume 3, com regiões ativadas artificialmente. Note que as áreas ativadas são completamente invisíveis a olho nu.

A coluna esquerda da figura 3.14 mostra o $SPM\{t\}$ obtido sem filtragem. O voxel ativado com o menor valor tinha valor 1,3. Assumindo a distribuição t de Student com 11 graus de liberdade, um voxel não-ativado pode assumir valor maior que 1,3 com probabilidade 11%. A coluna direita da figura 3.14 é obtida limiarizando $SPM\{t\}$ na altura 2,2. Note que muitas áreas não-ativadas foram falsamente detectadas como ativadas (e vice-versa).

A figura 3.15 mostra o SPM obtido filtrando fMRI com o método proposto (coluna esquerda) e imagens limiarizadas correspondentes (coluna direita). A maioria do ruído foi removida e a imagem limiarizada está perfeita: não há nem voxels não-ativados falsamente detectados como ativados, nem voxels ativados falsamente detectados como não-ativados. O voxel ativado com o menor valor tem valor 2,4, uma melhoria considerável sobre 1,3 anterior.

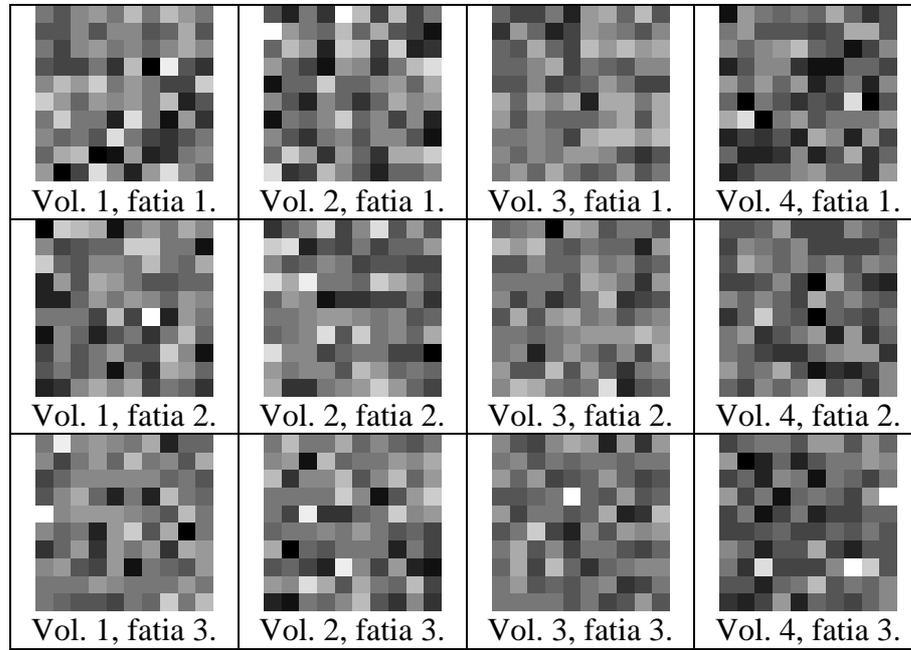


Fig. 3.11: Um fantom fMRI simulado com $10 \times 10 \times 3$ voxels por volume e 64 volumes. Somente os 4 primeiros volumes estão mostrados. Os valores de todos os voxels são 500, somados a um ruído gaussiano com média zero e desvio-padrão 10. Volumes 3, 4, 7, 8, 11, 12, ... tiveram alguns voxels ativados onde 20 foi somado aos valores originais.

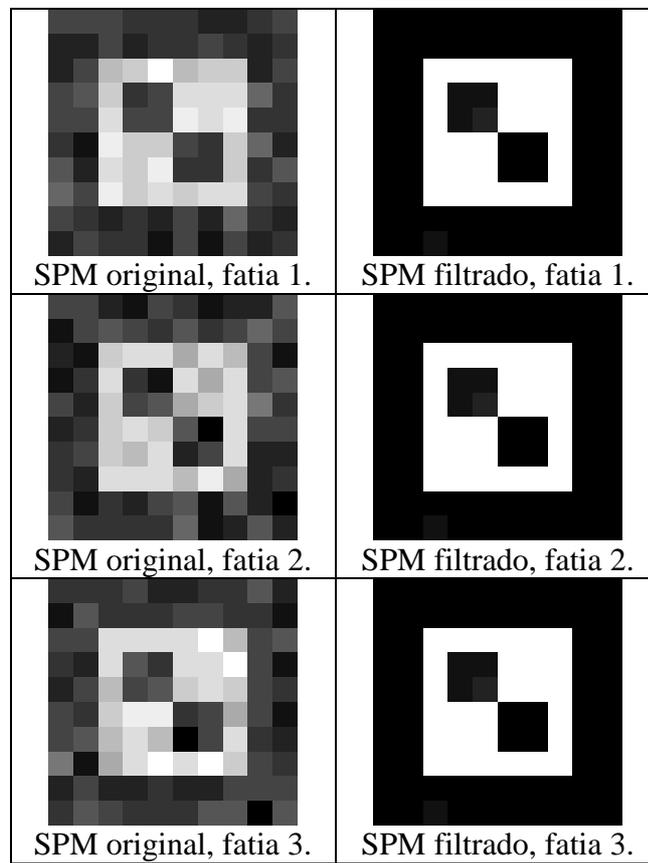


Fig. 3.12: SPM{t} obtido da fMRI da figura 3.11. Coluna esquerda: SPM{t} original. O menor valor de um voxel ativado foi 1,48. Um voxel não-ativado pode assumir um valor maior que 1,48 ao acaso com a probabilidade 7,2%. Coluna direita: SPM{t} obtido pela técnica proposta. Todos os voxels ativados apresentaram valores aproximadamente 12,1. Virtualmente é impossível que um voxel não-ativado assumira valores tão altos por acaso (probabilidade menor que 10^{-16}).

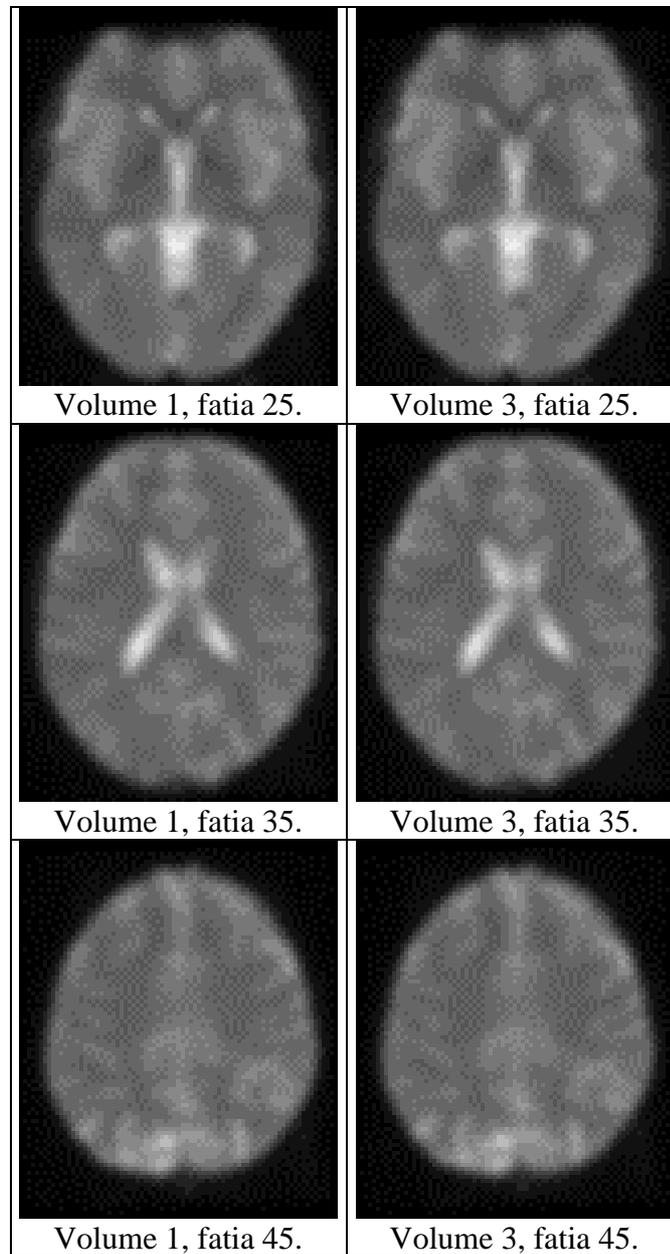


Fig. 3.13: Imagens fMRI reais com áreas artificialmente ativadas. Ativamos artificialmente pequenas regiões esféricas nos volumes 3, 4, 7, 8, 11 e 12, aumentando o valor do voxel em 3%.

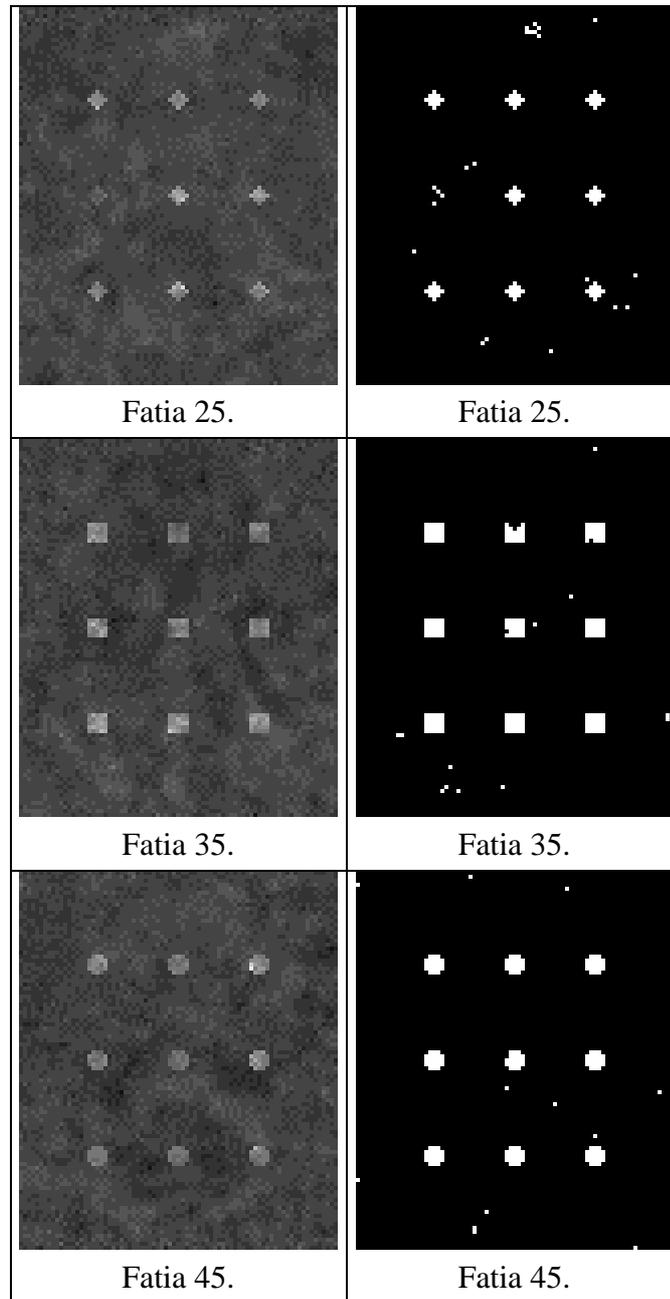


Fig. 3.14: Coluna esquerda: $SPM\{t\}$ obtido da fMRI da figura 3.13 sem filtragem. O menor valor ativado foi 1,3. Coluna direita: Áreas ativadas detectadas limiarizando SPM na altura 2,2. Um voxel não-ativado pode assumir valor acima de 2,2 com probabilidade 2,5%.

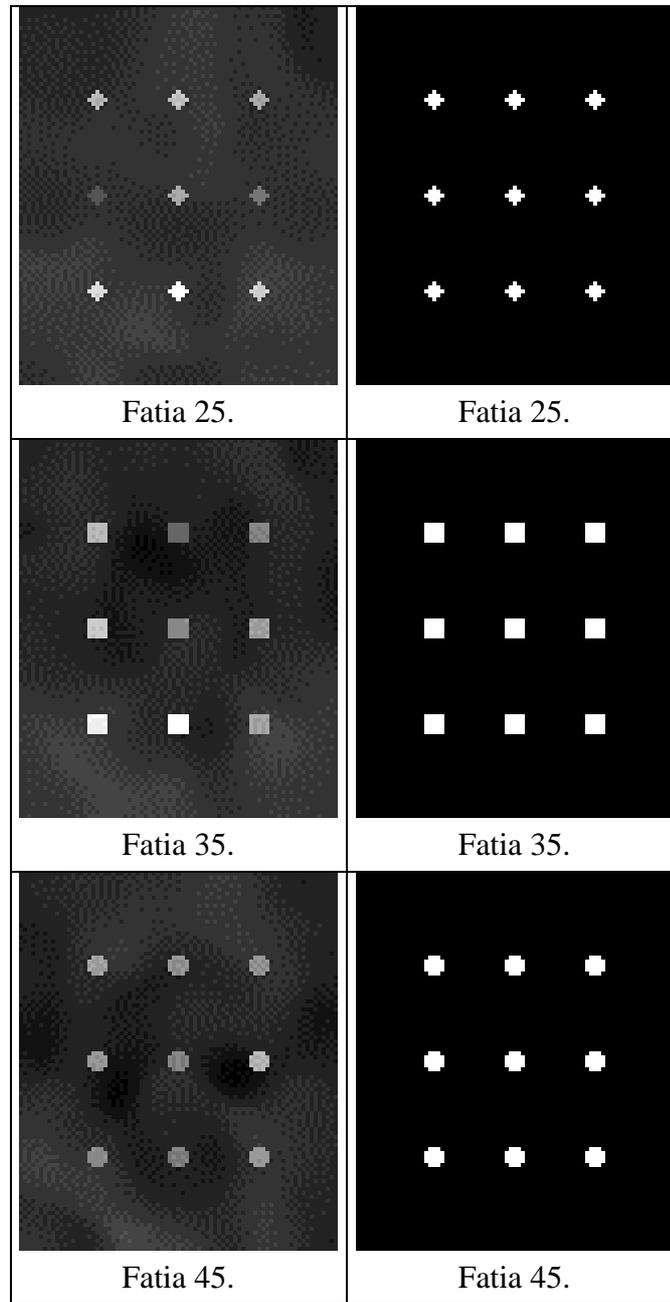


Fig. 3.15: Coluna esquerda: $SPM\{t\}$ obtido da fMRI da figura 3.13 usando a técnica proposta. O menor valor ativado foi 2,4, uma melhoria significativa sobre 1,3 anterior. Coluna direita: Áreas ativadas detectadas limiarizando o SPM em 2,2.

3.6 Conclusões

O objetivo deste capítulo foi apresentar as nossas contribuições científicas na área das aplicações da difusão anisotrópica.

Para isso, expusemos a teoria do espaço de escala linear obtida através da difusão isotrópica e o espaço de escala não-linear obtida pela difusão anisotrópica. Para ilustrar o processo de difusão anisotrópica, mostramos o seu uso na filtragem dos sinais de sensor de aceleração. Demonstramos que a difusão anisotrópica robusta (RAD) é superior às difusões propostas por Perona e Malik na restauração desse tipo de sinal. Depois, ilustramos o uso da difusão anisotrópica na detecção multi-escala das arestas de uma imagem. Também para esta aplicação, concluímos que a RAD é melhor que as técnicas de Perona e Malik.

Em seguida, mostramos a aplicação da RAD no melhoramento da reconstrução tomográfica. Especificamente, o algoritmo MENT-estendido foi melhorado, resultando num algoritmo que denominamos de MENT reconstrução-difusão. Mostramos através dos resultados experimentais que o novo algoritmo gera as imagens mais nítidas que o algoritmo MENT-estendido original. Comparamos também o novo algoritmo com o algoritmo amplamente conhecido retro-projeção filtrada, constatando novamente a superioridade da nossa proposta.

Por fim, aplicamos a RAD para melhorar a detecção das áreas ativadas do cérebro através da análise de imagens fMRI. Mostramos que o uso da RAD pode aumentar a confiabilidade na detecção das áreas ativadas.