Reconstrução tomográfica de imagens SPECT a
partir de poucos dados utilizando variação total

João Guilherme Vicente de Araujo

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

João Guilherme Vicente de Araujo

Reconstrução tomográfica de imagens SPECT a partir de poucos dados utilizando variação total

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências – Ciências de Computação e Matemática Computacional. *VERSÃO REVISADA*

Área de Concentração: Ciências de Computação e Matemática Computacional

Orientador: Prof. Dr. Elias Salomão Helou Neto

USP – São Carlos
Junho de 2017

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados fornecidos pelo(a) autor(a)

Ar

Araujo, João Guilherme Vicente de
Reconstrução tomográfica de imagens SPECT a
partir de poucos dados utilizando variação total /
João Guilherme Vicente de Araujo; orientador Elias
Salomão Helou Neto. -- São Carlos, 2017.
84 p.

Dissertação (Mestrado - Programa de Pós-Graduação
em Ciências de Computação e Matemática
Computacional) -- Instituto de Ciências Matemáticas
e de Computação, Universidade de São Paulo, 2017.

1. SPECT. 2. CSP. 3. Reconstrução simultânea. 4.
Variação total. I. Helou Neto, Elias Salomão,
orient. II. Título.

João Guilherme Vicente de Araujo

**Tomographic reconstruction of SPECT images from few data
using total variation**

Master dissertation submitted to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP, in partial fulfillment of the requirements for the degree of the Master Program in Computer Science and Computational Mathematics. *FINAL VERSION*

Concentration Area: Computer Science and Computational Mathematics

Advisor: Prof. Dr. Elias Salomão Helou Neto

**USP – São Carlos
June 2017**

*Este trabalho é dedicado à minha família que,
nas suas limitações, foi o suporte que me trouxe até aqui.*

AGRADECIMENTOS

Primeiramente, agradeço imensamente à Joice A. Fernandes por ter sido minha parceira na busca desse sonho, pois mesmo nos momentos difíceis fez-me acreditar que tudo era possível. Nesse percurso foi na sua presença que encontrei força e inspiração.

Agradeço também ao Prof. Elias Salomão Helou Neto que me instruiu durante esses anos à pesquisa científica e me apresentou a beleza dos métodos de reconstrução tomográfica.

Deixo aqui meus agradecimentos especiais à todos os amigos que fizeram parte da minha trajetória acadêmica: Alfredo R. Jorge, Pâmella Sátiko, Roberta Nunes, Milenna Midori, Camila Milano, Aline Diniz e todos aqueles que proporcionaram dias felizes em São Carlos.

*“I did not think; I investigated.
(Wilhelm Conrad Röntgen)”*

RESUMO

ARAÚJO, J. G. V. **Reconstrução tomográfica de imagens SPECT a partir de poucos dados utilizando variação total**. 2017. 84 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

Para realizar a correção de atenuação em uma tomografia computadorizada por emissão de fóton único (SPECT, em inglês) é necessário medir e reconstruir o mapa dos coeficientes de atenuação utilizando uma leitura de um tomógrafo de transmissão, feita antes ou simultaneamente à leitura de emissão. Essa abordagem encarece a produção da imagem e, em alguns casos, aumenta consideravelmente a duração do exame, sendo a imobilidade do paciente um fator importante para o sucesso da reconstrução. Uma alternativa que dispensa a leitura de transmissão é reconstruir tanto a imagem de atividade quanto o mapa de atenuação somente através dos dados de uma leitura de emissão. Dentro dessa abordagem propusermos um método baseado no algoritmo criado por Censor, cujo objetivo é resolver um problema misto de viabilidade côncavo-convexo para reconstruir simultaneamente as imagens. O método proposto é formulado como um problema de minimização, onde a função objetivo é dada pela variação total das imagens sujeita à viabilidade mista de Censor. Os testes foram feitos em imagens simuladas e os resultados obtidos na ausência de ruídos, mesmo para uma pequena quantidade de dados, foram satisfatórios. Na presença de dados ruidosos com distribuição de Poisson o método foi instável e a escolha das tolerâncias, nesse caso, ainda é um problema aberto.

Palavras-chave: SPECT, CSP, Reconstrução simultânea, Variação total.

ABSTRACT

ARAÚJO, J. G. V. **Tomographic reconstruction of SPECT images from few data using total variation**. 2017. 84 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

In order to perform attenuation correction in single photon emission computed tomography (SPECT), we need to measure and reconstruct the attenuation coefficients map using a transmission tomography scan, performed either sequentially or simultaneously with an emission scan. This approach increases the cost required to produce the image and, in some cases, increases considerably the scanning time, therefore the patient immobility is an important factor to the reconstruction success. An alternative that dispense the transmission scan is reconstruct both the activity image and the attenuation map only from emission scan data. In this approach we proposed a method based on the Censor's algorithm, which objective is to solve a mixed convex-concave feasibility problem to reconstruct simultaneously all images. The method proposed is formulated as a minimization problem, where the objective function is given by the total variation of the images subject to Censor's mixed feasibility. In the simulations, artificial images were used and the obtained results without noised data, even for small amount of data, were satisfactory. The method was unstable in the presence of Poisson distributed noise and the tolerance choice, in this case, is an open problem yet.

Keywords: SPECT, CSP, simultaneous reconstruction, Total variation.

LISTA DE ILUSTRAÇÕES

Figura 1	– Geometria da transformada de Radon. A linha tracejada é a reta usada para o cálculo de $\mathcal{R}[f](\theta, t')$. É possível ver também a representação de $\mathcal{R}_\theta[f](t)$ sobre o eixo t . Imagem retirada de (HELOU, 2009).	23
Figura 2	– Sinograma do <i>phantom de Shepp-Logan</i> calculado nos ângulos $\theta \in [0, 2\pi]$. A linha pontilhada indica a posição da função $\mathcal{R}_\theta[f](t)$ destacada na Figura 1. Imagem retirada de (HELOU, 2009).	24
Figura 3	– Leitura da Tomografia por Transmissão: (a) feixes paralelos; (b) feixes divergentes.	26
Figura 4	– Leitura da Tomografia por Emissão: (a) PET; (b) SPECT.	27
Figura 5	– Discretização do domínio $[-a, a] \times [-a, a]$ com destaque na representação do pixel 1. Somente a borda vermelha pertence ao pixel 1.	30
Figura 6	– Criar imagem que tenha as fontes P_1 e P_2 .	36
Figura 7	– Representação do <i>phantom de Shepp-Logan</i> em resolução 512×512 .	45
Figura 8	– Esquerda: Resultado de 100 iterações do EM. Direita: Resultado de 2000 iterações do EM.	46
Figura 9	– Comparação entre as iterações do EM e do OS-EM com 5 subconjuntos sem sobreposição.	47
Figura 10	– Comparação entre os tempos de execução entre os algoritmos EM e OS-EM para vários níveis.	47
Figura 11	– Erro numérico dos algoritmos EM e OS-EM para vários níveis utilizando o <i>phantom de Shepp-Logan</i> com imagem objetivo.	48
Figura 12	– Gráfico do erro do EM a cada 3 iterações em comparação com o erro de cada iteração do EM+TV com 3 iterações do EM e 8 do TV.	48
Figura 13	– Representação do mapa de atenuação descrito por $\frac{H(1/2-r)+H(1/3-r)}{2}$ em coordenadas polares, onde $H(x)$ é a função de Heaviside.	49
Figura 14	– Resultado de 100 iterações do EM sem atenuação (esquerda) e com a atenuação dada (direita).	49
Figura 15	– Reconstrução sem atenuação e 64 vistas. Da esquerda para a direita: 300 iterações do EM; zoom da iteração 300 do EM; 100 iterações do EM+TV, com $\alpha = 4 \times 10^{-3}$; zoom da iteração 100 do EM+TV.	50
Figura 16	– Reconstrução com atenuação e 64 vistas. Da esquerda para a direita: 300 iterações do EM; zoom da iteração 300 do EM; 100 iterações do EM+TV, com $\alpha = 4 \times 10^{-3}$; zoom da iteração 100 do EM+TV.	50

Figura 17 – A imagem de atividade (esquerda) e o mapa atenuação (direita) do modelo de estudo proposto. Representam, simplificadaamente, uma seção do tórax humano e um tumor no coração.	68
Figura 18 – Gráficos dos valores máximos dos tamanhos de passo τ_i e ν_i do algoritmo CSP de forma que cada subiteração esteja mais próxima possível do conjunto factível côncavo e convexo, respectivamente. A resolução de dados utilizada foi 40×32	70
Figura 19 – A imagem de atividade (esquerda) e o mapa atenuação (direita) da iteração 3000 do algoritmo CSP com resolução de dados 40×32 e valores máximos τ_i e ν_i de tamanhos de passo. As imagens de atividade e atenuação iniciais foram constantes com valores 10cm^{-1} e $0,028\text{cm}^{-1}$, respectivamente. . . .	71
Figura 20 – O resultado das iterações 80, 300, 1000 e 3000 do algoritmo TV-CSP para o modelo do tórax, onde o conjunto de dados tem resolução 40×32 , 40 vistas ao longo de 360° e 32 detectores em cada vista. A tolerância foi $\varepsilon = 1e - 5$ e os tamanhos de passo $\tau = 1,0$ e $\nu = 2,0$	71
Figura 21 – O resultado de 3000 iterações do algoritmo TV-CSP para o modelo do tórax com as resoluções de dados da Tabela 3. A tolerância foi $\varepsilon = 1e - 5$ e os tamanhos de passo $\tau = 1,0$ e $\nu = 2,0$	72
Figura 22 – Gráfico dos resíduos médios (superior esquerdo), das variações totais (superior direito) e dos erros nas imagens de atividade (inferior esquerdo) e atenuação (inferior direito) simulados através 3000 iterações do TV-CSP com $\nu = 2,0$ e $\tau = 1,0$ utilizando as resoluções propostas no texto e 3000 iterações do CSP com resolução 40×32 e parâmetros máximos para ν e τ . .	73
Figura 23 – Primeira linha: sinograma original (esquerda); e sinograma ruidoso (direita) com distribuição de Poisson e erro relativo de 3.9%. Segunda linha: a imagem de atividade (esquerda) e o mapa atenuação (meio) da iteração 1500 do algoritmo TV-CSP com resolução de dados 40×32 , tamanhos de passo para o CSP $\tau = 1,0$ e $\nu = 2,0$, tamanho de passo inicial para o método de subgradientes $\lambda_0 = 2,0$ e fatores de multiplicação $\alpha = 40$ e $\beta = 1,5$. Mapa de atenuação (direita) utilizando $\beta = 2,0$. As imagens de atividade e atenuação iniciais foram constantes com valores $10\alpha \text{ cm}^{-1}$ e $0,028\alpha \text{ cm}^{-1}$, respectivamente.	75
Figura 24 – A imagem de atividade (esquerda) e o mapa de atenuação (direita) da iteração 1500 do algoritmo TV-CSP com resolução de dados 40×32 , tamanhos de passo para o CSP $\tau = 1,0$ e $\nu = 2,0$, tamanho de passo inicial para o método de sugradients $\lambda_0 = 2,0$ e fator de multiplicação $\alpha = 40$. As tolerâncias foram tomadas, raio a raio, como a distância entre o sinograma original e o ruidoso.	76

LISTA DE TABELAS

Tabela 1 – Informações das elipses que compõem a imagem de atividade.	67
Tabela 2 – Informações das elipses que compõem o mapa de atenuação.	68
Tabela 3 – Taxa de descrição baseada no suporte da função do sinograma para diferentes resoluções.	69

SUMÁRIO

1	INTRODUÇÃO	21
1.1	Tomografia Computadorizada	21
1.2	Transformada de Radon	22
1.2.1	<i>Transformada de Radon Atenuada</i>	23
1.3	Tomografia por Transmissão	25
1.4	Tomografia por Emissão	26
1.5	A Discretização do Problema	29
1.5.1	<i>O Problema Inverso</i>	32
1.5.2	<i>Algoritmo do Traçado de Raios</i>	33
2	RECONSTRUÇÃO DA IMAGEM DE ATIVIDADE SPECT	39
2.1	Métodos Iterativos	39
2.1.1	<i>Algoritmo EM</i>	39
2.1.2	<i>Algoritmo OS-EM</i>	41
2.1.3	<i>Regularização EM com Total Variation</i>	42
2.2	Resultados	44
2.2.1	<i>Reconstrução sem atenuação</i>	45
2.2.2	<i>Reconstrução com atenuação</i>	46
2.2.3	<i>Reconstrução através de poucas vistas</i>	48
3	IMAGENS DE ATIVIDADE E ATENUAÇÃO VIA SPECT	51
3.1	Reconstrução simultânea: ML	52
3.2	Reconstrução da imagem de atividade através das Condições de Consistência	54
3.2.1	<i>Método do Bronnikov</i>	55
4	RECONSTRUÇÃO SIMULTÂNEA UTILIZANDO TV-CSP	59
4.1	Problema misto de viabilidade côncavo-convexo utilizando CSP	61
4.2	Algoritmo TV-CSP	64
4.3	Resultados	67
4.3.1	<i>Reconstrução com dados ruidosos</i>	73
4.4	Conclusão	75
5	CONCLUSÃO FINAL	77

REFERÊNCIAS	81
-----------------------	----

INTRODUÇÃO

1.1 Tomografia Computadorizada

Em 1895, o físico alemão Wilhelm Conrad Röntgen descobriu os raios x e notou que eles poderiam atravessar a matéria, sendo captados por um filme sensível situado atrás da mesma e que distingue as intensidades dos raios incididos. Quanto maior o número atômico do meio percorrido, maior é a taxa de absorção da radiação, logo, um raio chega mais intenso ao filme quanto menor for o número atômico do meio e vice-versa. Dessa maneira, diversas intensidades são registradas na imagem gerada no filme possibilitando a visualização de elementos internos da matéria.

A principal aplicação da radiografia tem por finalidade diagnósticos médicos, pois é uma técnica não invasiva e relativamente de baixo custo. Um exemplo clássico são os ossos do esqueleto humano que, por atenuarem muita radiação, são alvos fáceis de estudos. Contudo, pode haver estruturas com taxa de absorção baixa que estão na proximidade de ossos ou tecidos mais densos. Sendo assim, essas estruturas seriam omitidas de uma visualização clara no filme. Uma solução para esse tipo de problema é reconstruir tridimensionalmente o objeto de estudo, obtendo informações completas sobre os elementos internos, pois além da densidade e forma teríamos uma localização espacial desses tecidos.

A palavra *tomografia* é derivada do grego antigo *tomos*, que significa “corte” ou “fatia”, e *graphō*, que significa “escrita”. Essas fatias são reconstruídas matematicamente utilizando a leitura de radiografias em um número suficiente de ângulos, onde tais fatias são transversais ao eixo de rotação do equipamento radiográfico e se “empilhadas”, representam o objeto tridimensionalmente. Esse procedimento matemático – reconstruir as fatias através de informações do interior do objeto de estudo em diversos ângulos utilizando algum mecanismo de leitura – é chamado de *tomografia computadorizada*.

Neste texto, vamos abordar duas classes de tomografia computadorizada, a *tomografia*

por transmissão e a tomografia por emissão, com foco na segunda classe. O aparato matemático para a modelagem dessas classes será apresentado na próxima seção, a *transformada de Radon*. Contudo, neste capítulo não trataremos da reconstrução em si, mas do modelo de obtenção dos dados.

1.2 Transformada de Radon

O conceito da transformada de Radon (TR) foi apresentado em 1917 por Johann Radon (RADON, 1986) que depois serviu para modelar o problema mais simples em tomografia computadorizada, onde se utilizavam raios x para obter a imagem tomográfica. Nesta seção iremos apresentar os conceitos básicos da modelagem desta transformada.

A maioria das técnicas de tomografia computadorizada busca recuperar a imagem de uma fatia de um dado objeto tridimensional, ou seja, recuperar uma função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ que determina alguma propriedade interna do objeto naquela região, por exemplo, no caso do corpo humano, o coeficiente de atenuação dos tecidos ou a atividade metabólica dos órgãos internos.

Essa recuperação é feita através do conhecimento das integrais de linha de f ao longo de retas no plano da fatia, sendo possível estender esse conceito para três ou mais dimensões através de integrais sobre hiperplanos, mas nesse texto trataremos somente sobre a TR em duas dimensões. Assim, desejamos recuperar f a partir da seguinte função de (θ, t) :

$$\mathcal{R}[f](\theta, t) := \int_{\mathbb{R}} f(t\mathbf{v}_{\theta} + s\mathbf{v}_{\theta}^{\perp}) ds. \quad (1.1)$$

onde $\mathbf{v}_{\theta} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$ e $\mathbf{v}_{\theta}^{\perp} = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}$. Denotaremos $\mathcal{R}_{\theta}[f](t) := \mathcal{R}[f](\theta, t)$ e quando não houver ambiguidade quanto à f , tomaremos $p_{\theta}(t) := \mathcal{R}_{\theta}[f](t)$, onde p_{θ} é conhecida como a *projeção* de f com relação ao ângulo θ .

A aplicação $f \mapsto \mathcal{R}[f]$ é conhecida como *transformada de Radon* e uma condição suficiente para a sua existência é que f pertença ao espaço de Schwartz ou *espaço das funções de decréscimo rápido* em $\mathbb{R}^2$ (NATTERER, 1986), i.e., $f \in \mathcal{S}(\mathbb{R}^2)$.

Para compreender melhor o significado geométrico dessa aplicação vejamos um exemplo. A função f será representada como uma imagem no plano, onde cada ponto do plano indicará o valor da função de acordo com uma escala de tons de cinza (ou cores), assim como na Figura 1. Nesse exemplo usamos o *phantom de Shepp-Logan* (SHEPP; LOGAN, 1974), uma imagem artificial composta por uma soma de funções indicadoras de elipses. A escala das cores para com os valores é dada pela barra vertical à direita da figura. Além dos eixos coordenados canônicos x e y podemos observar o eixo t , cuja inclinação com relação à horizontal é dada pelo ângulo θ . Nesse eixo determina-se os valores das integrais de linha sobre as retas que são perpendiculares a t e distam t' da origem levando em consideração o sentido de t , em particular a linha tracejada. Como é possível ver, o gráfico sobre o eixo t representa a projeção da função f sobre ele e

algumas características da imagem em questão podem ser visualizadas, como por exemplo a acentuação dos valores na borda.

A partir de sua definição na [Equação 1.1](#) fica claro que $\mathcal{R}[f]$ é 2π -periódica em θ e, como a integral é sobre toda a reta, podemos chegar à conclusão de que $\mathcal{R}[f](\theta + \pi, -t) = \mathcal{R}[f](\theta, t)$. Logo, restringiremos os valores de θ para o intervalo $[0, \pi)$, pois qualquer outra informação fora desse intervalo seria redundante. Na [Figura 2](#) é possível ver, como uma imagem no plano $\theta \times t$, a representação de $\mathcal{R}[f]$ para o *phantom de Shepp-Logan*. Esta representação é conhecida como *sinograma* e para ilustrar a sua propriedade de reflexão foi plotado $\theta \in [0, 2\pi]$. A linha pontilhada no sinograma representa a função $p_\theta(t)$ destacada na [Figura 1](#).

Para reconstruir a função f , devemos, a partir do sinograma ([Figura 2](#)) encontrar a imagem descrita na [Figura 1](#), ou seja, calcular a inversa $\mathcal{R}^{-1}$. Para o cálculo da inversa $\mathcal{R}^{-1}$ existe uma fórmula analítica conhecida desde 1917 e publicada pelo próprio Johann Radon ([RADON, 1986](#)). Nessa fórmula são utilizados a transformada de Fourier e o operador dual de $\mathcal{R}$.

Figura 1 – Geometria da transformada de Radon. A linha tracejada é a reta usada para o cálculo de $\mathcal{R}[f](\theta, t')$. É possível ver também a representação de $\mathcal{R}_\theta[f](t)$ sobre o eixo t . Imagem retirada de ([HELOU, 2009](#)).

1.2.1 Transformada de Radon Atenuada

A transformada de Radon, como definida acima, modela um raio que tem como trajetória uma reta num determinado plano e é atenuado pela função f no decorrer dessa trajetória. Essa

Figura 2 – Sinograma do *phantom de Shepp-Logan* calculado nos ângulos $\theta \in [0, 2\pi]$. A linha pontilhada indica a posição da função $\mathcal{R}_\theta[f](t)$ destacada na Figura 1. Imagem retirada de (HELOU, 2009).

função f , na prática, determina o coeficiente de atenuação em cada ponto de uma fatia do objeto de estudo, ou seja, o *mapa de atenuação*. A representatividade matemática da atenuação não é explícita na definição da Equação 1.1, pois, de certo modo é incluída nos dados. Contudo, existem classes de tomografia onde, para cada ponto do interior do objeto, são emitidos raios que vão sendo atenuados durante sua trajetória até os detectores, percorrendo assim, uma semirreta. A quantidade de raios emitidos em um ponto é chamada de *atividade*, que nesse caso é a imagem que desejamos reconstruir.

Assim, a *transformada de Radon atenuada* (TRA) é definida da seguinte forma:

$$\mathcal{R}^\mu[f](\theta, t) := \int_{\mathbb{R}} f(t\mathbf{v}_\theta + s\mathbf{v}_\theta^\perp) \exp\left(-\int_{[s, \infty]} \mu(t\mathbf{v}_\theta + r\mathbf{v}_\theta^\perp) dr\right) ds, \quad (1.2)$$

onde o par (θ, t) determina a reta de integração, $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ representa a imagem de atividade, $\mu: \mathbb{R}^2 \rightarrow \mathbb{R}$ o mapa de atenuação e os vetores $\mathbf{v}$ e $\mathbf{v}^\perp$ são descritos da mesma forma que na Equação 1.1. Perceba que se tomarmos $\mu \equiv 0$, a transformada de Radon atenuada se resume à TR convencional.

A TRA é 2π -periódica em θ também, mas geralmente, ao contrário da TR,

$$\mathcal{R}^\mu[f](\theta, t) \neq \mathcal{R}^\mu[f](\theta + \pi, -t),$$

ou seja, os valores de θ podem cobrir o intervalo $[0, 2\pi)$, sem que a informação seja redundante.

Uma solução analítica para a inversa da TR atenuada foi desenvolvida recentemente, (NOVIKOV, 2002), porém nos deteremos aos algoritmos iterativos, pois possuem um gama de

aplicabilidade maior do que os métodos analíticos. E quanto às aplicações das duas transformadas, na próxima seção trataremos de apresentar alguns aspectos básicos de sua utilização em tomografia.

1.3 Tomografia por Transmissão

Como o próprio nome sugere, esse tipo de tomografia computadorizada utiliza-se da transmissão de um sinal através de um objeto, sendo essa transmissão feita paralelamente a uma fatia dele. Esse sinal é lançado com uma determinada intensidade, mas, conforme atravessa o objeto, atenua-se e é gravado no filme sensível com uma intensidade menor ou igual à inicial. Apesar do registro da intensidade final, estamos interessados na razão entre essas duas intensidades, pois elas fornecem as exponenciais negativas dos valores das integrais do coeficiente de atenuação linear ao longo das trajetórias dos sinais. Com esses dados em mãos podemos reconstruir matematicamente a função que determina o coeficiente de atenuação em cada ponto dessa fatia do objeto, conforme veremos mais adiante.

A trajetória do sinal depende da modalidade de tomografia que estamos utilizando, porém muitas modalidades tomográficas utilizam-se de trajetórias descritas por retas, como por exemplo a *tomografia por raios x*. Neste caso as integrais de linha seriam sobre retas, o que torna o modelo tratável através da transformada de Radon. Sem perda de generalidade, vamos usar o caso da tomografia por raios x para representar o conjunto de modalidades de tomografia por transmissão.

Denotaremos respectivamente por I_e e I_d as intensidades de emissão e detecção de um determinado feixe de raios x. Assim, para calcular a integral de linha sobre a reta que descreve o raio vamos usar a lei de Beer-Lambert simplificada ([PARNIS; OLDHAM, 2013](#)), ou seja,

$$-\log \frac{I_d}{I_e} = \mathcal{R}[f](\theta, t), \quad (1.3)$$

onde $\mathcal{R}[f]$ é a transformada de Radon e o par (θ, t) determina a reta em questão. Através dessa equação podemos obter amostras de $\mathcal{R}[f]$ e métodos para encontrar f a partir dessas amostras serão o foco deste texto em seções posteriores.

Para reconstruir uma imagem tomográfica é necessário termos informações sobre as integrais de linha de todas as retas do plano e isso não é possível fisicamente, portanto vamos trabalhar somente com amostras de $\mathcal{R}[f]$. Essas amostras podem ser obtidas de muitas maneiras, aqui vamos descrever apenas duas.

Para o esquema da esquerda na [Figura 3](#) são necessários diversos emissores dispostos em uma reta, ou somente um que se mova ao longo da mesma. O segundo aparelho é mais barato, porém prolonga o tempo de exame. Além disso, para se obter os dados necessários para a reconstrução, a leitura deve ser feita em diversos ângulos distintos, ou seja, todo o aparelho

de emissores e detectores deve ser girado por algum mecanismo. Esse tipo de geometria de aquisição é conhecido como leitura por *feixes paralelos*.

O esquema à direita nos dá uma opção muito interessante: um único emissor capaz de iluminar uma região em formato de leque e um conjunto de detectores posicionados em um círculo. Dessa forma, os aparelhos que utilizam desta geometria são mais baratos e não teriam acréscimo no tempo de exame se comparados com a leitura por feixes paralelos. Essa geometria é conhecida como leitura por *feixes divergentes*, e, atualmente, é o esquema utilizado nos tomógrafos por raios x.

Figura 3 – Leitura da Tomografia por Transmissão: (a) feixes paralelos; (b) feixes divergentes.

1.4 Tomografia por Emissão

Em tomografia por transmissão queremos reconstruir aspectos físicos do interior de um objeto. No caso do corpo humano, a anatomia dos órgãos, pois estes possuem, entre si, coeficientes de atenuação distintos possibilitando assim sua visualização por esse tipo de técnica. Contudo, há casos de exames tomográficos onde a anatomia não é tão importante. Em muitas situações é necessário obter informações sobre o metabolismo dos tecidos ou acompanhar a dinâmica de um medicamento *in vivo*.

A *tomografia por emissão* consiste em detectar eventos ocorridos no interior de um objeto de estudo. Esses eventos de alguma forma representam alguma atividade interna que ao acontecer libera um sinal que pode ser detectável no exterior do objeto, ou seja, a detecção acontece de dentro para fora. Isso pode ser facilmente visto em aplicações médicas, onde é introduzido no paciente um radiofármaco, uma substância marcada com um isótopo radioativo

Figura 4 – Leitura da Tomografia por Emissão: (a) PET; (b) SPECT.

e que faz parte do processo químico corporal que será estudado. Após determinado tempo, ao participar do processo químico, o elemento radioativo decai e emite radiação detectável por equipamentos apropriados. A partir dos dados detectados, por exemplo, podemos identificar uma região de alto nível de metabolismo, um tumor, e obter dados como posição e formato para uma possível extração ou análise de crescimento. Para essa modalidade tomográfica existem duas técnicas principais: a *tomografia por emissão de pósitrons*, conhecida como PET, e a *tomografia por emissão de fóton único*, ou SPECT.

Na PET, quando o radioisótopo sofre decaimento, é emitido um pósitron, uma antipartícula do elétron com carga oposta. O pósitron emitido viaja alguns milímetros até perder sua energia cinética e interagir com o elétron mais próximo. Nessa interação ambas partículas se aniquilam produzindo, assim, um par de fótons de radiação gama com energia bem definida movendo-se em direções opostas (aproximadamente) para fora do corpo. Os detectores são dispostos de maneira a envolver o paciente, conforme ilustrado na [Figura 4a](#). A detecção é quase simultânea para um par de fótons e isso define a região do espaço onde a emissão foi ocorrida. Nessa mesma figura, essa situação é exemplificada pela coincidência detectada no par (13,31) de detectores. Uma vez que a detecção não é exatamente simultânea, há a necessidade de inserir não somente informações sobre as intensidades detectadas, mas o tempo em que os detectores se excitaram. Esse tipo de informação facilita no caso de ocorrer duas ou mais detecções de pares de fótons (quase) simultâneas.

Um problema encontrado nessa técnica, o mais evidente, é causado pela atenuação completa de fótons no interior do paciente. Na [Figura 4a](#), um dos pares tem um fóton totalmente atenuado no corpo e o outro isoladamente detectado no elemento de número 7. Esse efeito

pode adicionar muito ruído ao problema quando esse fóton isolado é detectado (quase) simultaneamente com um par de fótons, pois não é possível saber qual dos três deve ser eliminado dos dados. Uma maneira de tentar eliminar essa situação é ter informações adicionais sobre a imagem, como por exemplo, a leitura de uma tomografia por transmissão. O espalhamento é outro efeito que pode acrescentar ruídos aos dados, pois causa detecções erradas, tais como a dos detectores (21, 40). Porém, é fácil de detectar essa ocorrência, uma vez que num espalhamento o fóton perde muita energia.

Como estamos tratando de uma fatia do objeto, algumas emissões em uma leitura PET não serão registradas porque algum dos fótons não está em rota de colisão com o arranjo de detectores, perdendo assim, uma grande quantidade de dados. Para amenizar esse problema, podemos utilizar diversos anéis de detectores arranjados na forma de um cilindro e considerar detecções entre anéis, resultando em uma leitura PET 3D dos dados. O método de reconstrução mais conhecido para esse tipo de dados tridimensionais é o algoritmo de retroprojeção filtrada 3D (3DRP) (KINAHAN; ROGERS, 1990). Contudo, o tempo computacional do 3DRP é muito elevado e estratégias de reconstrução 3D aproximada, como os algoritmos *rebinning*, transformam os dados 3D obtidos em conjuntos de dados 2D contendo cada um o respectivo sinograma aproximado de uma fatia do objeto e assim, a imagem 3D é recuperada aplicando métodos de reconstrução 2D para cada fatia (DEFRISE *et al.*, 1997). Entretanto, neste texto iremos abordar somente reconstruções bidimensionais.

No caso da reconstrução da imagem de tomografia por transmissão, a fonte de emissão é conhecida, então é possível usar a Equação 1.3 para a amostragem dos dados. Como na PET isso não é possível, pois não temos a localização da emissão, vamos desconsiderar a largura da tira que une um par de detectores qualquer e assumir que o número de detecções contadas nesse par é proporcional, exceto por um fator relacionado à atenuação, à integral da concentração de radioisótopo ao longo da reta que os liga. Em termos matemáticos, essa suposição torna-se:

$$n_{i,k} e^{\mathcal{R}[\mu](\theta_{i,k}, t_{i,k})} = \mathcal{R}[f](\theta_{i,k}, t_{i,k}), \quad (1.4)$$

onde o par de índices subscritos (i, k) indica os detectores envolvidos na contagem, $n_{i,k}$ o número de coincidências detectadas e o par $(\theta_{i,k}, t_{i,k})$ o ângulo e a distância relativa à origem da reta formada pelo par de detectores (i, k) . Os coeficientes de atenuação linear do objeto são representados pela função μ e a função que desejamos reconstruir é representada pela função de atividade f . Os valores de $\mathcal{R}[\mu]$ podem ser facilmente obtidos utilizando a leitura de um tomógrafo por transmissão, deixando o problema idêntico ao apresentado na Equação 1.3. A diferença é que o controle sobre a geração dos dados em PET é muito menor, e isso pode ser visto pela aquisição imprecisa dos dados.

Além da PET, temos a SPECT, que tem o mesmo princípio básico de funcionamento: a utilização de radioisótopos. Contudo, a diferença é que, ao invés de emitir um pósitron, eles geram um único raio gama ao decair (cf. Figura 4b), e a detecção é feita através de câmeras que possuem colimadores, estruturas em forma de orifícios cilíndricos, que indicam a direção de

percurso do fóton. Porém, esse esquema pode desperdiçar muitos fótons, pois as colisões podem ocorrer na lateral do colimador. Assim, uma leitura tridimensional dos dados é impraticável, pois teríamos muitas dessas ocorrências entre os anéis, tendo que recorrer a uma sequência de reconstruções das fatias bidimensionais do objeto.

Os problemas relacionados à trajetória/detecção encontrados em PET também aparecem na SPECT (cf. [Figura 4](#)). O efeito de atenuação não acrescenta ruídos aos dados, pois o fóton não chega a ser detectado, e o espalhamento pode ser amenizado devido ao fato do fóton perder muita energia, contudo, uma desvantagem da SPECT é que grande parte dos fótons com trajetórias coplanares à camera não são contabilizados, algo que não acontece na técnica PET.

Novamente, como em PET, a aquisição dos dados será feita pela contagem de detecções feitas em cada orifício de uma determinada câmera. A separação entre atenuação e transformada de Radon realizada na [Equação 1.4](#) faz sentido, pois a união dos caminhos percorridos pelos fótons é representada pela reta que liga ambos detectores, ou seja, toda a atenuação naquela reta interfere na contagem dos fótons no par de detectores. No caso SPECT, a atenuação na direção de detecção do fóton só interfere durante sua trajetória, assim temos que usar a transformada de Radon atenuada, $\mathcal{R}^\mu[f]$, definida na [Equação 1.2](#):

$$n_{\theta,t} = \mathcal{R}^\mu[f](\theta,t),$$

onde $n_{\theta,t}$ é o número de fótons contados na posição t com a câmera no ângulo θ , μ é o mapa de atenuação e f é a função a ser reconstruída. Com o conhecimento de μ podemos reconstruir f analiticamente a partir de $\mathcal{R}^\mu[f]$, sendo este um problema mais complexo do que o de inverter $\mathcal{R}[f]$. Essa solução analítica foi publicada em 2002 ([NOVIKOV, 2002](#)), mas antes disso já haviam métodos iterativos que reconstruíam f dado μ ([ERDOĞAN; GUALTIER; FESSLER, 1999](#); [BROWNE; PIERRO, 1996](#); [SHEPP; VARDI, 1982](#); [HUDSON; LARKIN, 1994](#)). Além disso, como veremos no [Capítulo 3](#) e [4](#), é possível obter a reconstrução sem que se conheça o mapa de atenuação, mas tais métodos são numericamente instáveis e dependem de um bom ajuste de seus parâmetros.

Como conclusão das modelagens apresentadas, uma tomografia por emissão de pósitrons possui uma maior precisão apesar dos ruídos, pois detecta todos os fótons em qualquer direção. Enquanto isso, a SPECT tem como vantagem o menor custo, pois utilizam-se radioisótopo de meia-vida mais longa, tornando possível a instalação de tomógrafos mais distantes de centros produtores de radiofármacos.

1.5 A Discretização do Problema

Nesta seção, vamos apresentar a discretização da transformada de Radon. Primeiramente, note que a TR é linear, ou seja, $\mathcal{R}[\alpha f_1 + \beta f_2] = \alpha \mathcal{R}[f_1] + \beta \mathcal{R}[f_2]$, $\forall \alpha, \beta \in \mathbb{R}$, e assim, em dimensão finita, pode ser representada por uma matriz, que denotaremos por $\mathbf{R}$. A discretização

da função f a ser reconstruída será feita no seu domínio, que, usualmente, é o retângulo $(-a, a) \times (-a, a)$, pois como os objeto de estudo são de tamanho limitado, então para qualquer $\mathbf{x} \in \mathbb{R}^2$ onde $\|\mathbf{x}\|_\infty \geq a$, com a suficientemente grande, temos $f(\mathbf{x}) = 0$, ou seja, nos limitaremos à reconstrução dos valores nos pontos internos ao retângulo. Esses pontos serão representados por elementos na malha quadrada $\{\bar{x}_i\} \times \{\bar{y}_j\}$, onde $\bar{x}_i = -a + (i-1)\Delta\xi$ e $\bar{y}_j = -a + (j-1)\Delta\xi$, $\Delta\xi = \frac{2a}{n}$, n o número de partições do intervalo $[-a, a]$ e $i, j = 1, 2, \dots, n+1$. Os centros dos *pixels* são dados por (x_r, y_s) , com $x_r = (\bar{x}_r + \bar{x}_{r+1})/2$, $y_s = (\bar{y}_s + \bar{y}_{s+1})/2$ e $r, s = 1, 2, \dots, n$. Deste modo, podemos representar cada pixel por um número inteiro k (cf. Figura 5), onde $k = (s-1) \cdot n + r$ para o ponto $\mathbf{x}_k = (x_r, y_s)$ e $k = 1, 2, \dots, n^2$. As amostras de f serão dadas nesses pontos centrais dos pixels, gerando um vetor $\mathbf{f} \in \mathbb{R}^P$, onde $\mathbf{f}_k = f(\mathbf{x}_k)$ e $P = n^2$ é o número de pixels. Definindo o k -ésimo pixel como o conjunto

$$\mathcal{P}_k = \left\{ \mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x} - \mathbf{x}_k\|_\infty \leq \frac{\Delta\xi}{2} \right\} \setminus \mathcal{U}_k, \quad (1.5)$$

onde

$$\mathcal{U}_k = \left\{ \mathbf{x} \in \mathbb{R}^2 : (\mathbf{x} - \mathbf{x}_k)^T \mathbf{e}_1 = \frac{\Delta\xi}{2} \text{ ou } (\mathbf{x} - \mathbf{x}_k)^T \mathbf{e}_2 = \frac{\Delta\xi}{2} \right\}$$

é o conjunto que contém os pontos da aresta superior e lateral direita do pixel e $\mathbf{e}_1, \mathbf{e}_2$ são os elementos da base canônica de $\mathbb{R}^2$.

Figura 5 – Discretização do domínio $[-a, a] \times [-a, a]$ com destaque na representação do pixel 1. Somente a borda vermelha pertence ao pixel 1.

Utilizando as funções indicadoras

$$\phi_k = \begin{cases} 1, & \text{se } \mathbf{x} \in \mathcal{P}_k; \\ 0, & \text{caso contrário,} \end{cases}$$

podemos representar f através de uma combinação linear da base $\{\phi_1, \phi_2, \dots, \phi_P\}$, isto é

$$f(\mathbf{x}) \approx \sum_{j=1}^P \phi_j(\mathbf{x}) \mathbf{f}_j, \quad \|\mathbf{x}\|_\infty < a. \quad (1.6)$$

Deste modo, queremos reconstruir o vetor $\mathbf{f} \in \mathbb{R}^P$, que representa a discretização de f .

Para ter conhecimento da transformada de Radon é necessário ter os valores de suas integrais de linha sobre as retas determinadas pelo par (θ, t) , conforme a [Figura 1](#). Para valores de θ fora do intervalo $[0, \pi)$ temos informações redundantes (ver §5, [Seção 1.2](#)) e para valores de t tal que $|t| \geq \sqrt{2}a$, as retas geradas por (θ, t) não intersectam o interior do retângulo e, conseqüentemente, os valores de f em seus pontos serão zero. A partir disso tomaremos a discretização de $\mathcal{R}[f]$ no domínio $\{\theta_m\} \times \{t_q\}$, com $m = 1, 2, \dots, v$ e $q = 1, 2, \dots, r$, onde r é o número de raios em cada vista e v é o número de vistas. Assim como na discretização da f os dados da TR serão vetorizados por $\mathbf{g} = (g_i)$, onde $g_i = g(\theta_m, t_q)$, $i = (q-1) \cdot v + m$, $i = 1, 2, \dots, D$ e $D = vr$.

A integral de linha sobre a reta determinada por (θ_m, t_q) na função indicadora ϕ_j é zero se não intersecta $\mathcal{P}_j$, caso contrário é igual ao comprimento da intersecção, pois ϕ_j é constante em $\mathcal{P}_j$. Portanto, tomando σ_{ij} como o comprimento da intersecção do i -ésimo raio passando pelo pixel j temos, pela linearidade da TR, que

$$g_i = g(\theta_m, t_q) = \sum_{j=1}^P \mathcal{R}[\phi_j](\theta_m, t_q) \mathbf{f}_j = \sum_{k=1}^P \sigma_{ik} \mathbf{f}_k, \quad i = 1, 2, \dots, D. \quad (1.7)$$

A partir da equação acima podemos escrever a aplicação da transformada de Radon da seguinte forma:

$$\mathbf{R}\mathbf{f} = \mathbf{g}, \quad \mathbf{R} \in \mathbb{R}^{D \times P} \quad (1.8)$$

onde $\mathbf{R}$ é a matriz que descreve a TR, $\mathbf{f}$ é a imagem desejada e $\mathbf{g}$ é o sinograma que contém os dados da transformação. Os elementos da matriz $\mathbf{R}$ da TR são dados por $r_{ij} = \sigma_{ij}$.

Para a transformada de Radon atenuada vamos discretizar a função do coeficiente de atenuação, μ , assim como fizemos para a função de atividade na [Equação 1.6](#):

$$\mu(\mathbf{x}) \approx \sum_{j=1}^P \phi_j(\mathbf{x}) \boldsymbol{\mu}_j, \quad \|\mathbf{x}\|_\infty < a,$$

onde $\boldsymbol{\mu} \in \mathbb{R}^P$ e $\boldsymbol{\mu}_j = \mu(\mathbf{x}_j)$. Logo, utilizando a discretização da f temos o seguinte:

$$\begin{aligned} g_i = g(\theta_m, t_q) &= \int_{\mathbb{R}} f(t\mathbf{v}_\theta + s\mathbf{v}_\theta^\perp) \exp\left(-\int_{[s, \infty]} \mu(t\mathbf{v}_\theta + r\mathbf{v}_\theta^\perp) dr\right) ds \\ &\approx \int_{\mathbb{R}} \sum_{j=1}^P \phi_j(t\mathbf{v}_\theta + s\mathbf{v}_\theta^\perp) f_j \exp\left(-\int_{[s, \infty]} \sum_{k=1}^P \phi_k(t\mathbf{v}_\theta + r\mathbf{v}_\theta^\perp) \boldsymbol{\mu}_k dr\right) ds. \end{aligned}$$

Colocando os somatórios para fora das integrais obtemos a seguinte aproximação:

$$\begin{aligned} g_i = g(\theta_m, t_q) &\approx \sum_{j=1}^P \mathbf{f}_j \int_{\mathbb{R}} \phi_j \left(t\mathbf{v}_\theta + s\mathbf{v}_\theta^\perp \right) \exp \left(- \sum_{k=1}^P \boldsymbol{\mu}_k \int_{[s,\infty]} \phi_k \left(t\mathbf{v}_\theta + r\mathbf{v}_\theta^\perp \right) dr \right) ds \\ &= \sum_{j=1}^P \mathbf{f}_j \sigma_{ij} \exp \left(- \sum_{k \in K_{ij}} \boldsymbol{\mu}_k \sigma_{ik} \right), \end{aligned} \quad (1.9)$$

onde $K_{ij} \in \{1, 2, \dots, P\}$ contém os índices dos pixels intersectados pelo raio i no segmento que parte do pixel j em direção ao detector. Assim, no caso da TR atenuada, os elementos da matriz $\mathbf{R}_\mu$, que representa a transformação, são descritos da seguinte forma

$$r_{ij} = \sigma_{ij} \exp \left(- \sum_{k \in K_{ij}} \sigma_{ik} \boldsymbol{\mu}_k \right). \quad (1.10)$$

Tomando $a_{ij} = \exp \left(- \sum_{k \in K_{ij}} \boldsymbol{\mu}_k \sigma_{ik} \right)$ como um elemento de uma *matriz de atenuação* $\mathbf{A}_\mu$, que é obviamente uma aplicação não-linear com relação a $\boldsymbol{\mu}$, e usando o produto de Hadamard ($\odot$) podemos escrever a aplicação da TR atenuada da seguinte forma:

$$\mathbf{R}_\mu \mathbf{f} = (\mathbf{R} \odot \mathbf{A}_\mu) \mathbf{f} = \mathbf{g}, \quad \mathbf{R}_\mu, \mathbf{R} \in \mathbb{R}^{D \times P} \quad (1.11)$$

onde, quando conhecemos $\boldsymbol{\mu}$, é a aplicação linear que descreve a TRA. O produto de Hadamard acima foi utilizado para representar a TRA na sua primeira implementação, mas devido ao alto custo computacional desse método abortamos essa abordagem. Para facilitar a notação, ao longo do texto vamos utilizar a matriz $\mathbf{R}$ para denotar tanto a TR quanto a TRA e quando for necessário, faremos distinção.

É fácil notar que as dimensões desses problemas, [Equações 1.8 e 1.11](#), são grandes, basta tomar imagens de resolução baixa, por exemplo, 1024×1024 e teremos, caso a atenuação seja conhecida, 2^{20} variáveis. Além disso, a matriz do sistema é naturalmente esparsa, pois para um determinado pixel $\mathcal{P}_k$ e um ângulo θ_m , poucos serão os raios que intersectarão o pixel. Contudo, há uma preocupação maior, a aquisição dos dados, pois ela pode afetar de maneira importante a solução do problema. Ainda nessa seção vamos abordar superficialmente este assunto.

1.5.1 O Problema Inverso

A estrutura da matriz $\mathbf{R}$ depende da aquisição dos dados. Por exemplo, se tomarmos o número de raios menor do que o número de pixels, teremos mais incógnitas que equações, ou seja, é muito provável que haja infinitas soluções. De modo geral, teríamos dois casos frequentes, a existência de muitas soluções ou nenhuma. Porém, tão importante quanto o número de soluções, temos o provável mal-condicionamento do problema, que pode ser causado, por exemplo, por uma detecção de ângulo limitado, fazendo o número de condição da matriz do sistema ser muito alto. Deste modo, o sistema acima tem uma grande chance de ser um problema *mal-posto* ou um *problema inverso* segundo a definição de Hadamard ([ENGL; HANKE; NEUBAUER, 2000](#)), ou seja, que uma das seguintes propriedades seja falsa para o sistema ([Equação 1.8](#)):

1. Para qualquer conjunto de dados $\mathbf{g}$ existe uma solução;
2. Para qualquer conjunto de dados $\mathbf{g}$ a solução é única;
3. A solução depende continuamente dos dados $\mathbf{g}$.

No caso linear discreto, as duas primeiras propriedades implicam na terceira, mas computacionalmente de nada vale a continuidade nos dados se o número de condição da matriz $\mathbf{R}$ é muito alto, pois poderíamos encontrar uma solução totalmente diferente da desejada.

Uma possível solução para contornar os problemas impostos pela instabilidade do sistema na [Equação 1.8](#) pode ser expressa como:

$$\begin{aligned} \mathbf{x} \in \arg \min \quad & c(\mathbf{x}) + \gamma r(\mathbf{x}) \\ \text{s.a.:} \quad & \mathbf{x} \in \mathbf{X}, \end{aligned} \quad (1.12)$$

onde $\mathbf{x}$ é a imagem reconstruída e $c(\mathbf{x})$ é uma função que deve forçar a consistência da solução com relação aos dados e geralmente é representada por $c(\mathbf{x}) = F(\mathbf{R}\mathbf{x}, \mathbf{g})$ com F tal que

$$F(\mathbf{x}, \mathbf{z}) = F_{\mathbf{z}}^* \Leftrightarrow \mathbf{x} = \mathbf{z},$$

onde $F_{\mathbf{z}}^*$ é o valor mínimo de F para um dado $\mathbf{z}$. Assim, a função F resolve o problema da inexistência de soluções quando $\mathbf{R}\mathbf{f} = \mathbf{g}$ é impossível e, além disso, essa função pode também resolver a questão da multiplicidade. Porém, independentemente do método utilizado para encontrar o minimizador de F , não há garantias de estabilidade para esta solução. Assim, se faz necessária a introdução do termo $\gamma r(\mathbf{x})$ que, além de estabilidade, irá favorecer soluções mais suaves. O parâmetro $\gamma > 0$ determina quanto desta estabilidade/suavidade será imposta à reconstrução. Essa estratégia é conhecida como *regularização* e r é o *funcional regularizador*, para mais detalhes sobre a teoria de regularização de problemas inversos, veja ([ENGL; HANKE; NEUBAUER, 2000](#)).

Na literatura a função mais fácil de analisar é a de quadrados mínimos, $c(\mathbf{x}) = \|\mathbf{R}\mathbf{x} - \mathbf{g}\|_2^2$, assim, para cada r determinado, temos uma regularização diferente, por exemplo, para a *regularização de Tikhonov*, temos $r(\mathbf{x}) = \|\mathbf{x}\|_2^2$. Entretanto, além da escolha do funcional regularizador, é necessário ter uma estratégia para a escolha do parâmetro γ , que é geralmente dada por $\gamma(\delta, \mathbf{g})$, onde δ é o nível de ruído nas medidas, ou seja, $\|\mathbf{R}\mathbf{x}^* - \mathbf{g}\| = \delta$, $\mathbf{g}$ são os dados do problema e $\mathbf{x}^*$ é a solução real, ou seja, a imagem que gerou os dados. Outro aspecto importante no modelo é o conjunto $\mathbf{X}$, pois restringe as possíveis soluções dando uma informação *a priori* sobre a solução do problema. Por exemplo, as técnicas tratadas neste texto assumem $\mathbf{X} = \mathbb{R}_+^P$, onde P é o número de pixels da imagem a ser reconstruída.

1.5.2 Algoritmo do Traçado de Raios

A fim de diminuir a complexidade algorítmica de executar a operação $\mathbf{R}\mathbf{f}$ ou $\mathbf{R}^T\mathbf{g}$, o algoritmo de traçado de raios ([HAN; LIANG; YOU, 1999](#)) encontra somente os elementos da

matriz cujo valor é não nulo. Para isso, irá determinar de forma recursiva os pixels por onde um determinado raio atravessa e os comprimentos de suas intersecções. Esses pixels são dados em ordem de travessia da fonte até o detector. Com isso, podemos calcular a transformada de Radon (ou a TR atenuada) discreta com um custo computacional mais baixo do que a forma matricial, pois em seu cálculo, para determinado raio, são somente usados os pixels intersectados pelo mesmo.

Como estamos trabalhando com imagens discretas, queremos reconstruir os valores de f e/ou μ em cada um de seus pixels independentemente de seu domínio, por isso não importa onde esses pixels estejam distribuídos no plano. Portanto, de modo geral, distribuiremos os pixels na malha de pontos $\{0, 1, \dots, N_x\} \times \{0, 1, \dots, N_y\}$, onde N_x e N_y são os números de pixels na horizontal e vertical respectivamente e, de maneira semelhante a definição da [Equação 1.5](#), cada pixel (i, j) , com $i = 0, 1, \dots, N_x - 1$ e $j = 0, 1, \dots, N_y - 1$, é descrito pelo conjunto

$$\mathcal{P}_{ij} = \{(x, y) \mid i \leq x < i + 1, j \leq y < j + 1\}.$$

Uma vez que está determinada a discretização do problema, podemos usar o algoritmo do traçado de raios para calcular $\mathbf{R}\mathbf{x}$ e $\mathbf{R}^T\mathbf{g}$.

Em uma tomografia por transmissão, o cálculo da integral de linha sobre a reta L é definido por

$$g = \sum_{\mathcal{P}_{ij} \cap L \neq \emptyset} l(i, j) \mu(i, j), \quad (1.13)$$

onde $\mu(i, j)$ é o coeficiente de atenuação e $l(i, j)$ é o comprimento da intersecção da reta com o pixel (i, j) . Para simulações de PET, o cálculo feito de maneira semelhante,

$$g = \left(\sum_{\mathcal{P}_{ij} \cap L \neq \emptyset} l(i, j) f(i, j) \right) \exp \left[- \sum_{\mathcal{P}_{i'j'} \cap L \neq \emptyset} l(i', j') \mu(i', j') \right], \quad (1.14)$$

onde $f(i, j)$ é a função de densidade do radiofármaco no paciente. No caso de uma simulação por SPECT, dado um pixel (i, j) tal que $\mathcal{P}_{ij} \cap L \neq \emptyset$, então vamos definir $L_{i,j}^+$ como a semirreta contida em L que parte do pixel (i, j) em direção ao detector. Assim, o cálculo dos dados na SPECT é dado por

$$g = \sum_{\mathcal{P}_{ij} \cap L \neq \emptyset} l(i, j) f(i, j) \exp \left[- \sum_{\mathcal{P}_{i'j'} \cap L_{i,j}^+ \neq \emptyset} l(i', j') \mu(i', j') \right]. \quad (1.15)$$

Portanto, se encontrarmos os pixels que fazem parte do trajeto de um raio e o comprimento l em cada um deles, podemos calcular as três modalidades de tomografia citadas anteriormente.

Logo, tomemos um raio que parte de $P_1 = (x_1, y_1)$ para $P_2 = (x_2, y_2)$, conforme a [Figura 6](#). Tomaremos uma parametrização da reta onde o valor do parâmetro é equivalente à distância de P_1 ao ponto $(x(\lambda), y(\lambda))$, ou seja,

$$\begin{aligned} x(\lambda) &= x_1 + (\lambda/L)(x_2 - x_1), \\ y(\lambda) &= y_1 + (\lambda/L)(y_2 - y_1), \end{aligned} \quad (1.16)$$

onde λ é 0 em P_1 e L em P_2 com $L = \|P_2 - P_1\|_2$. Assim, o algoritmo do traçado de raios (HAN; LIANG; YOU, 1999) segue os seguintes passos:

1. Inicialize o traçado de raios pelo cálculo do parâmetro λ inicial e os índices (v_x, v_y) do primeiro pixel do raio.

- a) Calcule os limitantes do parâmetro λ cujo o ponto na reta está contido na malha, isto é, $\lambda_{\min}$ e $\lambda_{\max}$:

$$\begin{aligned}\lambda_{\min} &= \max \left\{ 0, \min \{ \lambda_x(0), \lambda_x(N_x) \}, \min \{ \lambda_y(0), \lambda_y(N_y) \} \right\}, \\ \lambda_{\max} &= \min \left\{ 0, \max \{ \lambda_x(0), \lambda_x(N_x) \}, \max \{ \lambda_y(0), \lambda_y(N_y) \} \right\},\end{aligned}\quad (1.17)$$

onde

$$\begin{aligned}\lambda_x(i) &= \frac{L}{x_2 - x_1}(i - x_1), \\ \lambda_y(j) &= \frac{L}{y_2 - y_1}(j - y_1)\end{aligned}$$

são os valores do parâmetro para as intersecções com as retas $x = i$ e $y = j$, respectivamente.

- b) Calcule os índices (v_x, v_y) do primeiro pixel do raio da seguinte forma

$$v_x = \begin{cases} \lfloor x_{\min} \rfloor, & \text{se } x_{\min} \text{ não é inteiro ou } x_1 \leq x_2, \\ x_{\min} - 1, & \text{se } x_{\min} \text{ é inteiro e } x_1 > x_2, \end{cases}\quad (1.18)$$

onde

$$x_{\min} = x_1 + \lambda_{\min} \frac{x_2 - x_1}{L}$$

e v_y é calculado de modo semelhante.

- c) Tome o valor do parâmetro atual $\bar{\lambda}$ como $\lambda_{\min}$.

2. Dado o pixel (v_x, v_y) , encontre os valores dos parâmetros nas próximas intersecções com as retas perpendiculares aos eixos x e y :

$$\begin{aligned}\lambda_x &= \frac{L}{x_2 - x_1}(u_x - x_1), \\ \lambda_y &= \frac{L}{y_2 - y_1}(u_y - y_1),\end{aligned}\quad (1.19)$$

onde

$$u_x = \begin{cases} v_x + 1, & \text{se } x_1 \leq x_2, \\ v_x - 1, & \text{se } x_1 > x_2, \end{cases} \quad \text{e} \quad u_y = \begin{cases} v_y + 1, & \text{se } y_1 \leq y_2, \\ v_y - 1, & \text{se } y_1 > y_2. \end{cases}$$

Se em algum dos casos o denominador for zero tome o valor do parâmetro maior do que $\lambda_{\max}$.

3. Encontre $\lambda_\xi = \min\{\lambda_x, \lambda_y\}$, onde ξ é algum dos eixos. Se $\lambda_\xi > \lambda_{\max}$, calcule o comprimento do último pixel

$$l = \lambda_{\max} - \bar{\lambda}$$

e termine o traçado de raios.

4. Calcule o comprimento da intersecção no pixel (v_x, v_y) :

$$l = \lambda_\xi - \bar{\lambda}. \quad (1.20)$$

5. Atualize os índices do próximo pixel e o parâmetro atual: $v_\xi = u_\xi$ e $\bar{\lambda} = \lambda_\xi$. Retorne ao passo 2.

Figura 6 – Criar imagem que tenha as fontes P_1 e P_2 .

Note que este algoritmo atravessa menos que $N_x + N_y$ pixels da malha e o número de linhas de comando entre o passo 2 e o passo 5 são contantes, logo, dado um raio, o algoritmo de traçado de raios determina os seus pixels de intersecção em um tempo de ordem linear que depende de N_x e N_y . Além disso, o algoritmo determina de forma recursiva e ordenada os pixels (i, j) e os comprimentos $l(i, j)$ necessários para o cálculo discreto dos dados da tomografia por transmissão e da PET, representados nas Equações 1.13 e 1.14, respectivamente. No cálculo da SPECT, Equação 1.15, para cada pixel (i, j) encontrado pelo algoritmo teríamos que percorrer a semirreta L_{ij}^+ com o mesmo procedimento, o que fará com que o traçado de raios tenha ordem quadrática. Um modo de resolver este problema é obter os pixels em ordem decrescente: basta trocar o ponto da fonte P_1 pelo ponto do detector P_2 e acumular a integral de linha sobre L_{ij}^+ conforme os pixels vão avançando.

No próximo capítulo vamos apresentar métodos iterativos para a reconstrução de imagens de atividade de tomografias por emissão. Para isso, dada as imagens da distribuição dos

coeficientes de atenuação e/ou da função de densidade do radiofármaco, podemos construir os dados (sinogramas) aproximados dos modelos de tomografia citados nas Equações 1.13 - 1.15 e assim simular uma situação real de captação dos dados. A partir disso, vamos utilizar os métodos e comparar os resultados obtidos com a imagem original.

RECONSTRUÇÃO DA IMAGEM DE ATIVIDADE SPECT

Neste capítulo serão apresentados os métodos iterativos que foram objetos de estudo para a reconstrução das imagens de atividade de tomografias por emissão do tipo SPECT onde o mapa de atenuação é dado. Aqui, apresentaremos algumas aplicações dos métodos em dados com poucas vistas e julgaremos qualitativamente a efetividade dos mesmos.

2.1 Métodos Iterativos

A partir de um conjunto de dados $\mathbf{g}$ obtidos através da contagem das detecções de um aparelho de tomografia por emissão, temos que resolver o sistema $\mathbf{R}\mathbf{x} = \mathbf{g}$, onde $\mathbf{R}$ é a transformada de Radon atenuada discreta e $\mathbf{x}$ a imagem desejada. Os elementos da matriz $\mathbf{R}$ serão dados como na [Equação 1.10](#). O mapa de atenuação, que está presente no cálculo da matriz $\mathbf{R}$ foi omitido, pois é suposto seu conhecimento. O sistema a ser resolvido é um problema inverso e sua solução é instável, por isso, para solucionar esse problema usaremos a regularização da [Equação 1.12](#). Contudo, a não ser nos casos abordados por Novikov ([NOVIKOV, 2002](#)), essa solução não pode ser obtida analiticamente e, assim, se faz necessário o uso de métodos iterativos.

2.1.1 Algoritmo EM

Este algoritmo ([SHEPP; VARDI, 1982](#)) traz para o cálculo da solução do problema da [\(1.12\)](#) a natureza estatística da emissão radioativa para obter melhores reconstruções, deste modo é um dos métodos mais apropriados para este tipo de problema. A variável aleatória em questão é a contagem de detecções em cada detector, sendo este número um inteiro não-negativo. Logo, a distribuição de Poisson representa muito bem os dados e assumiremos por simplicidade que cada um dos dados é uma variável aleatória independente com distribuição de Poisson. Dada

a imagem $\mathbf{x}$ a ser reconstruída, o valor esperado para $\mathbf{g}$ é $\mathbf{R}\mathbf{x}$, isto é, $\mathbf{g} \sim \text{Poisson}(\mathbf{R}\mathbf{x})$. Como não temos conhecimento sobre o parâmetro da distribuição, vamos tomar o valor que melhor se adapta aos dados obtidos. Para isso, vamos maximizar a probabilidade desses dados observados $\mathbf{g}$ dado $\mathbf{R}\mathbf{x}$, ou seja, maximizar a função

$$l(\mathbf{x}) = p(\mathbf{g}|\mathbf{R}\mathbf{x}) = \prod_{i=1}^D p(\mathbf{g}_i|\mathbf{R}_i\mathbf{x}) = \prod_{i=1}^D \frac{(\mathbf{R}_i\mathbf{x})^{\mathbf{g}_i} e^{-\mathbf{R}_i\mathbf{x}}}{\mathbf{g}_i!},$$

onde $l(\mathbf{x})$ é a função de verossimilhança de $\mathbf{g}$, $\mathbf{g}_i$ o número de detecções no detector de índice i , $\mathbf{R}_i$ é a i -ésima linha da matriz $\mathbf{R}$ e $\mathbf{R}_i\mathbf{x} = \sum_{j=1}^P r_{ij}\mathbf{x}_j$.

O algoritmo EM (*Expectation-Maximization*) (SHEPP; VARDI, 1982) busca maximizar o logaritmo da função de verossimilhança dos dados, que é expressa, a menos de uma constante, por:

$$\log l(\mathbf{x}) = L(\mathbf{x}) := \sum_{i=1}^D (\mathbf{g}_i \log(\mathbf{R}_i\mathbf{x}) - \mathbf{R}_i\mathbf{x}).$$

Logo, para podermos escrever o modelo de máxima verossimilhança na forma da Equação 1.12, basta tomar $r(\mathbf{x}) \equiv k \in \mathbb{R}$, ou seja, constante, $\mathbf{X} = \mathbb{R}_+^P$ e $c(\mathbf{x}) = -L(\mathbf{x})$. Assim, com essa escolha de $c(\mathbf{x})$ e desconsiderando as restrições, vamos forçar a consistência da solução, pois

$$\nabla c(\mathbf{x}) = -\nabla L(\mathbf{x}) = \mathbf{R}^T \begin{pmatrix} 1 - \frac{\mathbf{g}_1}{\mathbf{R}_1\mathbf{x}} \\ 1 - \frac{\mathbf{g}_2}{\mathbf{R}_2\mathbf{x}} \\ \vdots \\ 1 - \frac{\mathbf{g}_D}{\mathbf{R}_D\mathbf{x}} \end{pmatrix}$$

e $\nabla c(\mathbf{x}) \equiv 0$ se $\mathbf{R}\mathbf{x} = \mathbf{g}$.

Seja $\mathbf{x}^0$ uma imagem inicial pré-determinada, por exemplo, uma imagem em que todos os valores no domínio são dados por uma constante real. Denotaremos por $\mathbf{x}^k$ o estimador para a imagem $\mathbf{x}$, solução do problema, na iteração k . Assim, segue o algoritmo EM:

1. $k \leftarrow 0$, inicializa $\mathbf{x}^k$ com valores positivos;
2. Repita até a convergência de $\mathbf{x}^k$
 - a) $k \leftarrow k + 1$
 - b) *Projeção*. Calcule o valor esperado para o número de detecções em cada detector da seguinte forma:

$$\boldsymbol{\lambda} \leftarrow \mathbf{R}\mathbf{x}^k.$$

- c) *Retroprojeção*. Calcule o próximo iterando como

$$\mathbf{x}_j^{k+1} \leftarrow \frac{\mathbf{x}_j^k}{\sum_{i=1}^D r_{ij}} \sum_{i=1}^D \frac{r_{ij}\mathbf{g}_i}{\boldsymbol{\lambda}_i}, \quad (2.1)$$

para os pixels $j = 1, \dots, P$, onde r_{ij} é definido na equação Equação 1.10.

As iterações expressas em (2.1) convergem monotonicamente (SHEPP; VARDI, 1982) para um ótimo da função $c(\mathbf{x})$, porém essa convergência é razoavelmente lenta e isso tornou o algoritmo inviável para as aplicações da época em que foi criada a técnica. Assim, foi necessário criar algoritmos mais rápidos para o modelo estatístico, por exemplo, o OS-EM.

2.1.2 Algoritmo OS-EM

O algoritmo EM de Shepp e Vardi (SHEPP; VARDI, 1982) requer a cada iteração o cálculo de uma projeção e uma retroprojeção até a convergência para a solução. Denominamos por projeção o produto da forma $\mathbf{R}\mathbf{x}$ e retroprojeção o produto na forma $\mathbf{R}^T\mathbf{g}$. No algoritmo OS-EM (HUDSON; LARKIN, 1994), *Ordered Subsets Expectation-Maximization*, os dados da projeção são agrupados em subconjuntos e o número de subconjuntos é denominado *nível*. Para cada subconjunto é aplicado um processo indêntico ao EM, uma projeção seguida de uma retroprojeção. Além disso, a imagem reconstruída de um determinado subconjunto servirá de imagem inicial para o processo do próximo subconjunto.

Uma iteração do OS-EM é dada pela execução do processo projeção-retroprojeção em todos os subconjuntos. Tomemos $S_1, S_2, \dots, S_t$ como os subconjuntos ordenados de índices dos detectores, escolhidos de forma que $\bigcup_{n=1}^t S_n = \{1, 2, \dots, D\}$ e que $S_n \cap S_m = \emptyset \forall n, m = 1, \dots, t$. Nesse caso, podemos garantir a convergência do algoritmo OS-EM e cada iteração dele terá aproximadamente o mesmo tempo de duração que uma iteração do EM. Um exemplo de subconjuntos ordenados é

$$S_n = \left\{ \left\lfloor \frac{D}{t} \right\rfloor (n-1) + 1, \dots, \left\lfloor \frac{D}{t} \right\rfloor n \right\}, \quad n = 1, 2, \dots, t-1$$

com $S_t = \left\{ \left\lfloor \frac{D}{t} \right\rfloor (t-1) + 1, \dots, D \right\}$

Assumindo os subconjuntos determinados acima e seja $\mathbf{x}^0$ uma imagem inicial pré-determinada. Denotaremos por $\mathbf{x}^k$ o estimador para a imagem $\mathbf{x}$ e por $\hat{\mathbf{x}}^k$ o subiterando na iteração k . Assim, segue o algoritmo OS-EM:

1. $k \leftarrow 0$, inicializa $\mathbf{x}^k$ com valores positivos.
2. Repita até a convergência de $\mathbf{x}^k$
 - a) $\hat{\mathbf{x}}^1 \leftarrow \mathbf{x}^k, k \leftarrow k + 1$
 - b) Para os subconjuntos $n = 1, \dots, t$:
 - i. *Projeção*. Calcule o valor esperado para o número de detecções em cada detector do subconjunto n da seguinte forma:

$$\lambda_i^n \leftarrow \mathbf{R}_i \hat{\mathbf{x}}^n = \sum_{j=1}^P r_{ij} \hat{\mathbf{x}}_j^n$$

para os detectores $i \in S_n$.

ii. *Retroprojeção*. Calcule o próximo iterando como

$$\hat{\mathbf{x}}_j^{n+1} \leftarrow \frac{\hat{\mathbf{x}}_j^n}{\sum_{i \in S_n} r_{ij}} \sum_{i \in S_n} \frac{r_{ij} \mathbf{g}_i}{\lambda_i^n}, \quad (2.2)$$

para os pixels $j = 1, \dots, P$.

c) $\mathbf{x}^k \leftarrow \hat{\mathbf{x}}^{t+1}$.

Diferentemente do EM, este algoritmo não garante a convergência para uma solução do problema (1.12), pois quando o sistema é inconsistente (não tem solução) o OS-EM não converge, exceto em situações excepcionais que são fora dos padrões dos dados obtidos na prática (HUDSON; LARKIN, 1994). Ainda assim, é possível ver que o OS-EM é relativamente mais rápido que o EM, pois para cada subconjunto a imagem inicial dada é obtida no processo do subconjunto anterior, tendo, de certa forma, informações mais relevantes para o cálculo atual.

Para tomógrafos do tipo SPECT, a coleta dos dados é feita através de projeções em uma câmera numa determinada posição (cf. Figura 4b). Quando a câmera gira em torno do paciente temos uma nova projeção, ou seja, outro conjunto de dados. Assim, a SPECT tem naturalmente uma estrutura que viabiliza o uso do OS-EM para a reconstrução de uma imagem tomográfica. No caso da PET, que coleta todos os dados simultaneamente, o OS-EM pode ser melhor aplicado ao fim de sua execução. Porém, isso não afeta a vantagem computacional de reduzir iterações do OS-EM, pois a SPECT só é mais rápido porque calcula a reconstrução simultaneamente com a aquisição dos dados.

Entretanto, como veremos na Seção 2.2, os algoritmos EM e OS-EM, quando são fornecidos poucos dados, geram uma imagem granulada e imprecisa. Uma solução para esse problema será apresentada na próxima subseção.

2.1.3 Regularização EM com Total Variation

O método de regularização por mínima *Total Variation* (TV) foi proposto por Rudin, Osher e Fatemi (RUDIN; OSHER; FATEMI, 1992; RUDIN; OSHER, 1994) para remover ruídos e borrados em uma imagem preservando todo o seu contorno. O uso da TV é uma técnica amplamente usada em processamento de imagens através da minimização de um funcional, por exemplo, da seguinte forma

$$\mathbf{x} \in \arg \min \int_{\Omega} |\nabla \mathbf{x}| + \alpha \|\mathbf{Ax} - \mathbf{b}\|_2^2, \quad (2.3)$$

onde $\mathbf{A}$ é o operador que gera a imagem borrada, $\mathbf{b}$ é a imagem ruidosa e/ou borrada que descreve os dados obtidos, $\mathbf{x}$ é a imagem reconstruída, Ω é o domínio da função 2D que determina uma imagem no algoritmo, $\alpha > 0$ é o parâmetro da regularização e $\|\mathbf{Ax} - \mathbf{b}\|_2^2$ é o termo que impõe a consistência da solução com relação aos dados do sistema $\mathbf{Ax} = \mathbf{b}$. Nesse caso, a TV é um método

de regularização da forma (1.12), com as funções, parâmetro e restrições que o descrevem dados por: $c(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_2^2$, $r(\mathbf{x}) = \int_{\Omega} |\nabla \mathbf{x}|$, $\gamma = \alpha^{-1}$ e $\mathbf{X} = \mathbb{R}_+^P$.

O propósito de uma regularização combinando a *Total Variation* com o EM é utilizar poucos dados (ou vistas) na reconstrução de uma imagem que tem o ruído descrito pela distribuição de Poisson. A imagem obtida pelo EM utilizando poucos dados é normalmente ruidosa (ver Figura 16), por isso se faz necessário o uso da TV, que, a partir de uma boa escolha do parâmetro α , consegue remover consideravelmente o ruído imposto por este método.

Para que a TV utilize os dados radiográficos, tomemos $\mathbf{A} = \mathbf{R}$ o operador que descreve a transformada de Radon (ou a TR atenuada) e $\mathbf{b} = \mathbf{g}$ os dados obtidos pelas projeções. Além disso, para que haja a combinação dos métodos, a função $c(\mathbf{x})$, que força a consistência dos dados, será dada pelo algoritmo EM e o funcional regularizador $r(\mathbf{x})$ será dado pelo TV, ou seja, para o método de regularização EM+TV,

$$c(\mathbf{x}) = -L(\mathbf{x}) = \sum_{n=1}^D (\mathbf{R}_n \mathbf{x} - \mathbf{g}_n \log(\mathbf{R}_n \mathbf{x})) \quad \text{e} \quad r(\mathbf{x}) = \int_{\Omega} |\nabla \mathbf{x}|.$$

Portanto, a imagem obtida minimiza a variação total e maximiza o logaritmo da função de verossimilhança e método regularizador EM+TV pode ser escrito como

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimizar}} \quad & \int_{\Omega} |\nabla \mathbf{x}| + \alpha \sum_{n=1}^D (\mathbf{R}_n \mathbf{x} - \mathbf{g}_n \log(\mathbf{R}_n \mathbf{x})) \\ \text{s.a.:} \quad & \mathbf{x} \in \mathbb{R}_+^P \end{aligned} \quad (2.4)$$

onde $\alpha > 0$.

Para resolver o problema de minimização EM+TV acima, foi usado um método iterativo semi-implícito de diferenças finitas descrito em (YAN; VESE, 2011). A imagem $\mathbf{x}$ agora será representada por sua forma inicial, uma matriz 2D onde o pixel (i, j) é dado por $\mathbf{x}_{i,j} = (x_i, y_j)$ com $i, j = 1, \dots, n$. O resultado das iterações do EM sem o TV será representado por $\mathbf{x}^{EM}$ e a imagem inicial dada para o EM é $\mathbf{x}^0$. Calculamos assim as iterações $\mathbf{x}^{k+1}$ do TV, com $k \geq 0$, a partir da equação discreta linearizada,

$$\begin{aligned} & - \frac{1}{\alpha} \frac{\mathbf{x}_{i,j}^k}{v_j} \frac{\mathbf{x}_{i+1,j}^k - \mathbf{x}_{i,j}^{k+1}}{\sqrt{\varepsilon + (\mathbf{x}_{i+1,j}^k - \mathbf{x}_{i,j}^k)^2 + (\mathbf{x}_{i,j+1}^k - \mathbf{x}_{i,j}^k)^2}} \\ & + \frac{1}{\alpha} \frac{\mathbf{x}_{i,j}^k}{v_j} \frac{\mathbf{x}_{i,j}^{k+1} - \mathbf{x}_{i-1,j}^k}{\sqrt{\varepsilon + (\mathbf{x}_{i,j}^k - \mathbf{x}_{i-1,j}^k)^2 + (\mathbf{x}_{i-1,j+1}^k - \mathbf{x}_{i-1,j}^k)^2}} \\ & - \frac{1}{\alpha} \frac{\mathbf{x}_{i,j}^k}{v_j} \frac{\mathbf{x}_{i,j+1}^k - \mathbf{x}_{i,j}^{k+1}}{\sqrt{\varepsilon + (\mathbf{x}_{i+1,j}^k - \mathbf{x}_{i,j}^k)^2 + (\mathbf{x}_{i,j+1}^k - \mathbf{x}_{i,j}^k)^2}} \\ & + \frac{1}{\alpha} \frac{\mathbf{x}_{i,j}^k}{v_j} \frac{\mathbf{x}_{i,j}^{k+1} - \mathbf{x}_{i,j-1}^k}{\sqrt{\varepsilon + (\mathbf{x}_{i+1,j-1}^k - \mathbf{x}_{i,j-1}^k)^2 + (\mathbf{x}_{i,j}^k - \mathbf{x}_{i,j-1}^k)^2}} \\ & + \mathbf{x}_{i,j}^{k+1} - \mathbf{x}_{i,j}^{EM} = 0, \end{aligned} \quad (2.5)$$

onde $v_j = \sum_{i=1}^P r_{ij}$. Perceba que na equação acima podemos isolar a variável que representa o pixel (i, j) na iteração $k + 1$.

Por fim, as iterações do algoritmo EM+TV, normalmente, são descritas pela aplicação de 2 ou 3 passos do algoritmo EM, representado pela função $EM(\cdot)$, seguidos de 5 a 8 passos do algoritmo TV (YAN; VESE, 2011) apresentado na Equação 2.5 e que é representado pela função $TV(\cdot)$. A imagem inicial do TV é a obtida pelo EM, deste modo, o algoritmo pode ser descrito da seguinte forma:

1. Inicialize uma imagem positiva $\mathbf{x}^0$ e o número máximo de iterações Q, K e L dos algoritmos EM+TV, EM e TV, respectivamente;
2. $q \leftarrow 0$;
3. Se $q < Q$ faça
 - a) Faça $\hat{\mathbf{x}}^0 \leftarrow \mathbf{x}^q$ e $k \leftarrow 0$;
 - b) Se $k < K$ faça $k = k + 1$,

$$\hat{\mathbf{x}}^k \leftarrow EM(\hat{\mathbf{x}}^{k-1})$$
 e repita o item (b) do passo 3;
 - c) Faça $\hat{\mathbf{x}}^0 \leftarrow \hat{\mathbf{x}}^K$ e $k \leftarrow 0$;
 - d) Se $k < L$ faça $k = k + 1$,

$$\hat{\mathbf{x}}^k \leftarrow TV(\hat{\mathbf{x}}^{k-1})$$
 e repita o item (c) do passo 3;
4. Faça $\mathbf{x}^{q+1} \leftarrow \hat{\mathbf{x}}^L$ e $q \leftarrow q + 1$.

Esse método garante uma solução menos ruidosa para poucos dados fornecidos, porém essa imagem reconstruída depende do valor do parâmetro α atribuído, que é obtido através de teste de tentativa e erro, pois é possível detectar qualitativamente os ruídos gerados por uma tomografia utilizando poucos dados. Esse e os outros resultados obtidos pelos algoritmos apresentados serão expostos na próxima seção.

2.2 Resultados

Para o desenvolvimento dos algoritmos apresentados na seção anterior foi usada a biblioteca RAFT (abreviação de *Reconstruction Algorithms for Tomography*) que tem distribuição gratuita e é escrita em C/C++ (MIQUELES; HELOU, 2007 (acessado em 11 de junho de 2015)). Essa biblioteca contém um conjunto de ferramentas básicas que são necessárias em muitos problemas inversos relacionados à imagiologia médica. Nesta seção vamos apresentar os resultados obtidos nos experimentos numéricos desenvolvidos para testar os algoritmos com

finalidade reconstruir a imagem de atividade do interior de um objeto dado o mapa de atenuação. Assim, apresentaremos os resultados para simulações sem atenuação, que consiste em inverter a transformada de Radon, e os resultados para uma simulação onde o mapa de atenuação é dado. Além disso, devido à instabilidade do sistema $\mathbf{R}\mathbf{x} = \mathbf{g}$ quando temos poucos raios, será fácil de notar que o algoritmo EM+TV gera os melhores resultados.

2.2.1 Reconstrução sem atenuação

Primeiramente, vamos considerar o problema de reconstrução onde não há atenuação, ou seja, estamos interessados em somente reconstruir o mapa de atividade invertendo a transformada de Radon. A imagem teste será o *phantom* de Shepp-Logan (SHEPP; LOGAN, 1974), apresentado na Figura 7, cuja resolução é de 512×512 pixels. A amostragem dos dados será feita sobre 512 ângulos igualmente espaçados entre 0 e π , onde em cada ângulo há 512 raios. Desta forma $\mathbf{R}$ é uma matriz quadrada, pois temos 512×512 raios/equações para 512×512 variáveis.

Figura 7 – Representação do *phantom* de Shepp-Logan em resolução 512×512 .

No algoritmo EM, devemos repetir as iterações até a convergência, porém em termos práticos não conhecemos a imagem a ser reconstruída, então foi tomado um número alto de iterações, no caso 2000, e é possível ver que a reconstrução (cf. Figura 8) se aproxima do *phantom*, porém é necessário mais iterações até a convergência.

O algoritmo OS-EM garante convergência mais rápida para a solução se entre os subconjuntos ordenados não houver sobreposição, pois, caso contrário, algumas projeções seriam incluídas mais de uma vez na soma da retroprojeção (HUDSON; LARKIN, 1994). Nas simulações feitas, os subconjuntos foram criados através do particionamento das vistas, ou seja, dos ângulos, de forma que os subconjuntos tenham a mesma quantidade de raios.

Figura 8 – Esquerda: Resultado de 100 iterações do EM. Direita: Resultado de 2000 iterações do EM.

Todas as simulações deste trabalho foram executadas num computador que possui processador Intel® Core™ i7-2630QM com 2.00 GHz de frequência, 6 GB de memória RAM e sistema operacional Ubuntu 14.04 LTS. A [Figura 10](#) apresenta os tempos de execução para o algoritmo EM e OS-EM (com 2, 5 e 10 subconjuntos). Podemos observar que as iterações de ambos os métodos têm praticamente o mesmo tempo de execução, em particular, para o OS-EM com 5 subconjuntos, em comparação com o EM, temos um acréscimo de 12% no tempo. De forma prática, a [Figura 9](#) apresenta as primeiras iterações do OS-EM com 5 subconjuntos em comparação com as EM de 5 em 5, ou seja, cada iteração do OS-EM com k subconjuntos ordenados sem sobreposição equivale a k iterações do EM e isso pode ser visto na [Figura 11](#).

Para o algoritmo EM+TV foram utilizadas, a cada iteração, 3 iterações do EM seguidas de 8 iterações do TV, o máximo recomendado em ([YAN; VESE, 2011](#)). O parâmetro $\alpha = 0.004$ foi o melhor valor de parâmetro encontrado por tentativa e erro no intervalo $(0; 0,1]$. Através desses testes foi possível ver que a escolha do parâmetro α é muito sensível, mesmo em outros casos.

O algoritmo EM+TV após 3 iterações do EM busca suavizar o resultado obtido utilizando o TV, porém, com o sinograma dado neste teste, as iterações do EM já são suficientemente suaves e o algoritmo torna-se não muito eficiente (cf. [Figura 12](#)).

2.2.2 Reconstrução com atenuação

Agora, vamos considerar o problema de reconstrução tomando o mapa de atenuação previamente dado. O mapa que será utilizado para os testes numéricos é dado pela [Figura 13](#). A imagem a ser reconstruída será a mesma que a dos testes sem atenuação, o *phantom* de Shepp-Logan. Igualmente, a amostragem terá 512×512 raios.

Figura 9 – Comparação entre as iterações do EM e do OS-EM com 5 subconjuntos sem sobreposição.

Figura 10 – Comparação entre os tempos de execução entre os algoritmos EM e OS-EM para vários níveis.

Para o algoritmo EM foram executadas 100 iterações e o resultado obtido é praticamente idêntico ao EM sem atenuação (cf. [Figura 14](#)), a menos de um pequeno borrado na parte inferior, mas isso depende da atenuação tomada. Contudo, a convergência do EM é garantida mesmo com atenuação, assim, executado mais iterações foi obtido um resultado idêntico ao do EM sem atenuação.

Como o OS-EM é um algoritmo que acelera as iterações do EM, os resultados obtidos foram praticamente os mesmos e a convergência foi garantida através da escolha de subconjuntos

Figura 11 – Erro numérico dos algoritmos EM e OS-EM para vários níveis utilizando o *phantom* de Shepp-Logan com imagem objetivo.

Figura 12 – Gráfico do erro do EM a cada 3 iterações em comparação com o erro de cada iteração do EM+TV com 3 iterações do EM e 8 do TV.

ordenados sem sobreposição. Os subconjuntos foram determinados através de partições do conjunto de ângulos do sinograma, tomando todas os detectores em cada uma destas partições. O algoritmo EM+TV obteve resultados semelhantes ao ponto de serem indistinguíveis.

2.2.3 Reconstrução através de poucas vistas

O maior problema de uma reconstrução tomográfica é a amostragem dos dados e isso depende do tomógrafo utilizado. Esse tipo de problema tem dois pontos principais, o número total de raios e os ângulos das projeções/vistas. Quanto ao número total de raios, se este for

Figura 13 – Representação do mapa de atenuação descrito por $\frac{H(1/2-r)+H(1/3-r)}{2}$ em coordenadas polares, onde $H(x)$ é a função de Heaviside.

Figura 14 – Resultado de 100 iterações do EM sem atenuação (esquerda) e com a atenuação dada (direita).

menor que o número de pixels da imagem teremos uma matriz $\mathbf{R}$ retangular, possibilitando a existência de infinitas soluções para o sistema $\mathbf{R}\mathbf{x} = \mathbf{g}$. E quanto aos ângulos, se estes forem limitados em uma certa região, por mais raios que tenhamos nessas projeções, não teremos informações importantes sobre a imagem a ser reconstruída, ou seja, a inversão da transformada de Radon fica inviável, pois os ângulos de amostragem não são distribuídos por todo o domínio.

Para simular uma situação de poucas vistas, vamos considerar o problema de reconstrução sem atenuação utilizando a mesma imagem teste com resolução de 512×512 . Contudo, a amostragem dos dados será feita sobre 64 ângulos igualmente espaçados entre 0 e π , tendo cada ângulo 512 raios. Logo, a matriz $\mathbf{R}$ será retangular.

O algoritmo EM procura uma solução para o sistema $\mathbf{R}\mathbf{x} = \mathbf{g}$. Porém, devido ao mal-condicionamento do sistema, essa solução é instável. Assim, para esse problema, o algoritmo EM+TV se sobressai, pois as iterações instáveis do EM vão sendo suavizadas ao longo das iterações e a solução obtida é mais nítida (cf. Figura 15).

Figura 15 – Reconstrução sem atenuação e 64 vistas. Da esquerda para a direita: 300 iterações do EM; zoom da iteração 300 do EM; 100 iterações do EM+TV, com $\alpha = 4 \times 10^{-3}$; zoom da iteração 100 do EM+TV.

Figura 16 – Reconstrução com atenuação e 64 vistas. Da esquerda para a direita: 300 iterações do EM; zoom da iteração 300 do EM; 100 iterações do EM+TV, com $\alpha = 4 \times 10^{-3}$; zoom da iteração 100 do EM+TV.

IMAGENS DE ATIVIDADE E ATENUAÇÃO VIA SPECT

Nos capítulos anteriores foram apresentados métodos que, a partir do conhecimento do mapa de atenuação, reconstruíam imagens da distribuição de atividade utilizando dados de tomografia por emissão do tipo SPECT. Contudo, para se obter esse mapa de atenuação é necessário reconstruí-lo usando dados de uma tomografia por transmissão, por exemplo, tomografia por raios x (CT). Porém, uma vez que as imagens são obtidas em aparelhos distintos, alinhar o paciente torna-se um processo delicado, adicionando novas dificuldades ao problema. A tendência é usar aparelhos SPECT/CT onde os dois estão acoplados, mas essa estratégia é muito cara de se obter e manter. Mesmo a combinação de ambos aparelhos, CT e SPECT, tem uma série de desvantagens: reduz a quantidade de exames, aumenta a exposição do paciente à radioatividade e dificulta o protocolo de aquisição de dados e o design do leitor ([WELCH *et al.*, 1997](#); [BRONNIKOV, 2000](#)).

Um meio alternativo de obter o mapa de atenuação é reconstruí-lo diretamente através dos dados de emissão, o que simplificaria muito a aparelhagem e a técnica de leitura, ou seja, resolver o seguinte problema para μ e $\mathbf{f}$:

$$\mathbf{R}_\mu \mathbf{f} = \mathbf{g} \quad (3.1)$$

onde $\mathbf{R}_\mu$ é uma matriz resultante de uma aplicação não-linear com relação a μ , μ é o mapa de atenuação, $\mathbf{f}$ a imagem de atividade e $\mathbf{g}$ são os dados. Entretanto, se tivermos mesma resolução para ambas imagens, esse tipo de abordagem duplica o número de variáveis enquanto temos o mesmo número de equações/dados. Além disso, unida ao mal-condicionamento intrínseco do problema, temos a não linearidade com relação à atenuação, tornando o problema ainda mais difícil de se resolver.

Uma gama de autores propuseram técnicas que eliminam completamente a necessidade do uso de leitores de transmissão. Essas técnicas podem ser divididas em duas classes. A primeira

classe aplica as condições de consistência à imagem do operador da transformada de Radon atenuada para obter somente o mapa de atenuação através dos dados de emissão. Em seguida aplica-se um método de reconstrução da imagem de atividade, por exemplo, os que foram citados nos capítulos anteriores. Esse método foi proposto por Natterer (NATTERER, 1981) e foi estudado em diversos artigos (WELCH *et al.*, 1997; BRONNIKOV, 2000; BRONNIKOV, 1999; NATTERER, 1993; NATTERER, 1983). Essencialmente, as condições de consistência são dadas na forma de um funcional que se iguala a zero na imagem do operador do problema, onde tanto o funcional quanto o operador dependem somente da atenuação. A segunda classe de técnicas envolve métodos iterativos para a solução do sistema não-linear (3.1), reconstruindo a atividade e a atenuação simultaneamente. Essa abordagem parece ter sido primeiramente proposta por Censor *et al.* (CENSOR *et al.*, 1979) que usou projeções do subgradiente para resolver um problema misto de viabilidade côncavo-convexo. A partir daí, muitos pesquisadores deram uma grande contribuição para essa classe de algoritmos de reconstrução simultânea (NUYTS *et al.*, 1999; MANGLOS; YOUNG, 1994; KROL *et al.*, 1995; KAPLAN; HAYNOR; VIJA, 1999; DICKEN, 1999).

Nas subseções seguintes vamos apresentar alguns dos métodos de ambas as classes. Para cada um deles daremos uma breve explicação da parte teórica e discutiremos aspectos práticos de sua utilização.

3.1 Reconstrução simultânea: ML

Nos capítulos anteriores, onde se reconstruía somente a imagem de atividade foi apresentado o EM (*Expectation Maximization*) criado por Sheep e Vardi, onde é utilizado o algoritmo de máxima verossimilhança ML (*Maximum Likelihood*) do Dempster (DEMPSTER; LAIRD; RUBIN, 1977). O método que será apresentado a seguir foi proposto por Nuyts em 1983 (NUYTS *et al.*, 1999) e busca reconstruir simultaneamente a atividade e a atenuação utilizando o ML, sendo este chamado de MLAA (*Maximum Likelihood reconstruction for Attenuation and Activity*). Essencialmente, a contagem de fótons nos detectores é dada por um vetor $\mathbf{g}$ e seus valores são descritos como variáveis aleatórias com distribuição de Poisson de parâmetros dados por $\mathbf{R}_\mu \mathbf{f}$, onde μ e $\mathbf{f}$ representam as imagens de atenuação e atividade, respectivamente. Como ambas são desconhecidas, o método ML busca estimá-las maximizando a função de verossimilhança, ou seja, escolher o parâmetro $\mathbf{f}$ que maximiza a probabilidade da escolha dos dados observados. O logaritmo da função de verossimilhança exceto por valores constantes é dado por

$$L(\mu, \mathbf{f}) = \sum_{i=1}^D \left(-(\mathbf{R}_\mu \mathbf{f})_i + \mathbf{g}_i \ln(\mathbf{R}_\mu \mathbf{f})_i \right).$$

A função $L(\mu, \mathbf{f})$ é côncava quando μ ou $\mathbf{f}$ é fixado, mas não é conjuntamente (NUYTS *et al.*, 1999). Portanto, encontrar o máximo global requer estratégias de otimização muito custosas

computacionalmente. Os esforços desse método são limitados a encontrar ótimos locais da função L que, através de algum conhecimento a priori, sejam visualmente úteis.

Esse método busca atualizar alternadamente μ e $\mathbf{f}$. Assim, fixando μ em seu valor atual reduzimos o problema em uma tomografia por emissão onde há o conhecimento do mapa de atenuação. Consequentemente, o aumento do valor da função de verossimilhança pode ser garantido aplicando uma única iteração do método ML-EM (SHEPP; VARDI, 1982; LANGE; CARSON, 1984):

$$\mathbf{f}_j^{(k+1)} = \frac{\mathbf{f}_j^{(k)}}{\sum_i r_{ij}} \sum_i \frac{r_{ij} \mathbf{g}_i}{\sum_{j'} r_{ij'} \mathbf{f}_{j'}^{(k)}}, \quad (3.2)$$

onde r_{ij} é representado na Equação 1.10.

No caso da SPECT, não é possível aplicar somente o método ML para reconstruir μ dados $\mathbf{f}$ e $\mathbf{g}$. Assim, Nuyts aplica um algoritmo de gradiente ascendente criado por ele (NUYTS *et al.*, 1999), que é descrito por:

$$\mu_j^{(k+1)} = \mu_j^{(k)} + \frac{\alpha}{N \sum_i q_{ij}} \sum_i q_{ij} \left(1 - \frac{\mathbf{g}_i}{\sum_{j'} r_{ij'} \mathbf{f}_{j'}^{(k)}} \right), \quad (3.3)$$

onde N é o diâmetro da área reconstruída, α é um parâmetro de relaxação e

$$q_{ij} = \sum_{\{k|j \in K_{ik}\}} \sigma_{ij} \sigma_{ik} \exp \left(\sum_{l \in K_{ik}} \sigma_{il} \mu_l \right) \mathbf{f}_k = \sum_{\{k|j \in K_{ik}\}} \sigma_{ij} r_{ik} \mathbf{f}_k.$$

Na equação acima, $K_{ik} \subset \{1, 2, \dots, P\}$ contém os índices dos pixels intersectados pelo raio i no segmento que parte do pixel k em direção ao detector e o valor q_{ij} é interpretado como a quantidade total de fótons emitidos no raio i cuja a trajetória intersecta o pixel j .

Utilizando somente as duas equações, (3.2) e (3.3), não há garantia de uma solução única para o problema. Portanto, um conhecimento a priori, como proposto, guia o algoritmo para soluções mais significativas. No MLAA, foram utilizados os seguintes conhecimentos a priori sobre mapa de atenuação: 1) são esperados valores de atenuação padrões para os músculos, ar, pulmões e outros tecidos; 2) é esperado também que o mapa de atenuação seja localmente suave. Para isso, o conhecimento a priori sobre os valores da atenuação foi implementado como distribuições probabilísticas bi e trimodal, e a suavidade local foi implementada através da distribuição de Gibbs bidimensional. Para detalhes sobre a aplicação do conhecimento a priori ver (NUYTS *et al.*, 1999).

Como no ML-EM, o método MLAA pode ser acelerado utilizando subconjuntos ordenados (OS) (HUDSON; LARKIN, 1994). Para isso basta substituir as somas das linhas de projeção i nas Equações 3.2 e 3.3 por uma soma sobre um conjunto de linhas de projeção. Contudo, no final de cada subiteração há a aplicação do conhecimento a priori, cujo tempo computacional depende do tamanho da imagem e não do subconjunto. Logo, como essa parte não é acelerada, é necessário uma boa escolha para o número de subconjuntos.

A maior desvantagem do MLAA é o número de parâmetros muito alto, pois para cada objeto estudado os valores devem ser escolhidos conforme a necessidade e isso é feito de maneira heurística. Outra desvantagem, que também é encontrada na maioria dos métodos que buscam resolver esse tipo de problema através de reconstrução simultânea, é o efeito *cross-talk*: sobreposição mútua das imagens de atividade e atenuação. Esse efeito ocorre devido à não concavidade da função de verossimilhança. Contudo, em termos práticos, com uma “boa” escolha dos parâmetros, as imagens são visualmente satisfatórias.

3.2 Reconstrução da imagem de atividade através das Condições de Consistência

O grande problema das reconstruções simultâneas é o efeito *cross-talk* entre as imagens de atividade e atenuação reconstruídas, que podem se sobrepor mutuamente. Esse tipo de problema tem como principal motivo o mal-condicionamento intrínseco do sistema (3.1) que implica que possa haver várias soluções ou nenhuma. Assim, Natterer (NATTERER, 1981) propôs uma abordagem que busca determinar o mapa de atenuação μ sem reconstruir a atividade f utilizando as condições de consistência para a imagem do operador da transformada de Radon atenuada. Essa abordagem foi estudada em diversos artigos (WELCH *et al.*, 1997; BRONNIKOV, 2000; BRONNIKOV, 1999; NATTERER, 1993; NATTERER, 1983) e sua maior vantagem é que não há necessidade de se obter informações sobre a imagem de atividade e nenhuma tentativa de reconstruí-la é feita.

Uma vez que se tem os dados gerados por uma tomografia por emissão do tipo SPECT, as condições de consistência são dadas na forma de um funcional que se iguala a zero na imagem do operador da TRA e esse funcional tem como único parâmetro o mapa de atenuação. Assim, ao encontrar uma solução para as condições de consistência obtemos um mapa de atenuação que faz com que o sistema (3.1) torne-se consistente, ou seja, tem pelo menos uma solução para a imagem de atividade que é calculada utilizando os métodos apresentados no Capítulo 2. Contudo, devido a ruídos, erros de discretização e outros fatores físicos, os dados são corrompidos e, juntamente com o mal condicionamento do problema, causam instabilidade para a solução a menos que apliquemos um método de regularização. A ideia dos métodos de regularização é aproximar um problema mal condicionado por um bem condicionado a fim de encontrar uma solução estável que depende dos parâmetros da regularização.

As condições de consistência podem ser aplicadas tanto para o operador contínuo (NATTERER, 1981; WELCH *et al.*, 1997; NATTERER, 1993; NATTERER, 1983) quanto para o discreto (BRONNIKOV, 2000; BRONNIKOV, 1999). Aqui, vamos apenas abordar o caso discreto que, além de ser o foco do texto, tem a vantagem de reduzir os graus de liberdade para a solução do problema. Apresentaremos o conteúdo do trabalho feito por Bronnikov em (BRONNIKOV, 2000) que consiste em encontrar uma raiz para o funcional de consistência

aplicando o método de Newton.

3.2.1 Método do Bronnikov

A discretização de uma tomografia SPECT (Equação 3.1) é representada por um sistema caracterizado pela matriz $\mathbf{R}_\mu \in \mathbb{R}^{D \times P}$, onde D é quantidade de raios utilizados na discretização, P a quantidade pixels e $\boldsymbol{\mu} \in \mathbb{R}^P$ o vetor que representa o mapa de atenuação amostrado. Os elementos da matriz $\mathbf{R}_\mu$ para o caso SPECT são dados na Equação 1.10, onde o vetor $\boldsymbol{\mu}$ é parâmetro da matriz que rege o sistema. Assim, vamos assumir que o número de equações deve ser maior ou igual que o número de variáveis (BRONNIKOV, 2000), ou seja,

$$D \geq P + P,$$

tornando a matriz $\mathbf{R}_\mu$ retangular.

Para o sistema em questão, $\mathbf{R}_\mu \mathbf{f} = \mathbf{g}$, $\mathbf{f} \in \mathbb{R}^P$ e $\mathbf{g} \in \mathbb{R}^D$, temos que qualquer vetor $\mathbf{g}$ pode ser escrito como

$$\mathbf{g} = \bar{\mathbf{g}} + \mathbf{g}_\perp, \quad (3.4)$$

onde $\bar{\mathbf{g}}$ pertence ao subespaço gerado pelas colunas de $\mathbf{R}_\mu$, ou seja, $\text{Im}(\mathbf{R}_\mu)$, e $\mathbf{g}_\perp$ pertence ao seu complemento ortogonal, $\text{Im}(\mathbf{R}_\mu)^\perp$. Dessa forma, o espaço linear $\mathbb{R}^D$ pode ser escrito como uma soma direta desses dois subespaços, ou seja, $\mathbb{R}^D = \text{Im}(\mathbf{R}_\mu) \oplus \text{Im}(\mathbf{R}_\mu)^\perp$ e assim $\mathbf{g}$ tem uma representação única na mesma forma da Equação 3.4. Portanto, o sistema é consistente se $\mathbf{g} \in \text{Im}(\mathbf{R}_\mu)$, o que é equivalente à $\mathbf{g}_\perp = 0$.

Através da pseudoinversa de Moore-Penrose $\mathbf{R}_\mu^+ \in \mathbb{R}^{P \times D}$ da matriz $\mathbf{R}_\mu$ podemos calcular explicitamente os vetores $\bar{\mathbf{g}}$ e $\mathbf{g}_\perp$ da seguinte forma

$$\bar{\mathbf{g}} = \mathbf{P}_\mu \mathbf{g}, \quad \mathbf{g}_\perp = \mathbf{P}_\mu^\perp \mathbf{g},$$

onde

$$\mathbf{P}_\mu = \mathbf{R}_\mu \mathbf{R}_\mu^+, \quad \mathbf{P}_\mu^\perp = \mathbf{I} - \mathbf{R}_\mu \mathbf{R}_\mu^+$$

são matrizes $D \times D$ das projeções ortogonais e $\mathbf{I}$ é a matriz identidade. Usando essa definição de projeção ortogonal, as condições de consistência são escritas da seguinte forma

$$\mathbf{P}_\mu^\perp \mathbf{g} = 0. \quad (3.5)$$

A Equação 3.5 descreve um sistema não linear de equações que pode ser resolvido com relação a $\boldsymbol{\mu}$. Contudo, para implementar as condições de consistência discretas temos que calcular $\mathbf{P}_\mu^\perp$. Um dos métodos mais rápidos para calcular essa projeção é a decomposição QR (GOLUB; LOAN, 2013), onde a matriz do sistema é representada por

$$\mathbf{R}_\mu = \mathbf{Q} \begin{pmatrix} \mathbf{R}_1 \\ 0 \end{pmatrix}, \quad (3.6)$$

onde $\mathbf{Q} \in \mathbb{R}^{D \times D}$ é uma matriz ortogonal e $\mathbf{R}_1 \in \mathbb{R}^{P \times P}$ é uma matriz triangular superior com os elementos da diagonal não-negativos. A matriz $\mathbf{Q}$ pode ser particionada em $\mathbf{Q} = (\mathbf{Q}_1, \mathbf{Q}_2)$, onde $\mathbf{Q}_1 \in \mathbb{R}^{D \times P}$ e $\mathbf{Q}_2 \in \mathbb{R}^{D \times (D-P)}$. Assim, a pseudoinversa é dada por $\mathbf{R}_\mu^+ = \mathbf{R}_1^{-1} \mathbf{Q}_1^T$, e as projeções por

$$\mathbf{P}_\mu = \mathbf{Q}_1 \mathbf{Q}_1^T, \quad \mathbf{P}_\mu^\perp = \mathbf{Q}_2 \mathbf{Q}_2^T.$$

Agora que temos a expressão algébrica para a projeção $\mathbf{P}_\mu^\perp$ podemos verificar, para um determinado mapa de atenuação, se o sistema (3.1) é consistente.

Para resolver o sistema não-linear (3.5) Bronnikov utiliza de um algoritmo iterativo onde cada iteração é dada por

$$\boldsymbol{\mu}^{(k)} = \mathcal{P}_{\mathbb{R}_+^P} \left(\boldsymbol{\mu}^{(k-1)} - \Delta \boldsymbol{\mu} \right),$$

onde $\Delta \boldsymbol{\mu} \in \mathbb{R}^P$ é o passo do algoritmo e $\mathcal{P}_{\mathbb{R}_+^P}$ é um operador de projeção no subconjunto $\mathbb{R}_+^P$, ou seja, impõe a restrição de não-negatividade em todas as coordenadas.

A imagem inicial $\boldsymbol{\mu}^{(0)}$ do algoritmo pode ser o vetor nulo ou uma imagem conhecida a priori. Em cada iteração k do algoritmo, a matriz $\mathbf{R}_\mu$ é recalculada e sua decomposição QR (3.6) é feita. O vetor de passo $\Delta \boldsymbol{\mu}$ é obtido pelo método de Newton e para calculá-lo deve-se resolver a seguinte equação

$$\mathbf{J}_\mu \Delta \boldsymbol{\mu} = \mathbf{g}_\perp, \quad (3.7)$$

onde

$$\mathbf{g}_\perp = \mathbf{Q}_2 \mathbf{Q}_2^T \mathbf{g},$$

e $\mathbf{J}_\mu$ é a matriz Jacobiana do problema, que é escrita como

$$\mathbf{J}_\mu = -\mathbf{Q}_2 \mathbf{Q}_2^T \mathbf{B}_\mu,$$

onde

$$\mathbf{B}_\mu = \left(\frac{\partial \mathbf{R}_\mu}{\partial \mu_1} \tilde{\mathbf{f}} \quad \frac{\partial \mathbf{R}_\mu}{\partial \mu_2} \tilde{\mathbf{f}} \quad \dots \quad \frac{\partial \mathbf{R}_\mu}{\partial \mu_P} \tilde{\mathbf{f}} \right)$$

é uma matriz $D \times P$ e $\tilde{\mathbf{f}} = \mathbf{R}_\mu^+ \mathbf{g}$. Os elementos da matriz $\mathbf{B}_\mu$ podem ser calculados utilizando a Equação 1.10 da seguinte forma

$$b_{ij} = -\sigma_{ij} \sum_{k \in K_{ij}} r_{ik} \tilde{f}_k,$$

onde $K_{ij} \subset \{1, 2, \dots, P\}$ contém os índices dos pixels intersectados pelo raio i no segmento que parte do pixel k em direção ao detector.

Perceba que nenhuma informação sobre a distribuição de atividade é usada para o cálculo das iterações. A solução de quadrados mínimos $\tilde{\mathbf{f}} = \mathbf{R}_\mu^+ \mathbf{g}$, presente na expressão da matriz Jacobiana, é representada através dos dados $\mathbf{g}$ e do mapa de atenuação $\boldsymbol{\mu}$. Assim, diferentemente dos métodos de reconstrução simultânea, somente $\boldsymbol{\mu}$ interfere na função objetivo evitando consequentemente efeitos do tipo *cross-talk* (BRONNIKOV, 2000).

Assim, a partir da solução do sistema linear (3.7) com relação à $\Delta\boldsymbol{\mu}$, temos como definir o passo para cada iteração, porém, como a matriz Jacobiana contém a projeção $\mathbf{Q}_2\mathbf{Q}_2^T$ que é uma matriz $D \times D$ de posto $D - P$, pois $\mathbf{Q}_2$ é uma matriz $D \times (D - P)$ com as colunas ortogonais, então determinar o passo $\Delta\boldsymbol{\mu}$ é um problema mal-condicionado e para encontrar uma solução estável é necessário aplicar um método de regularização.

O método de regularização de Tikhonov foi a regularização utilizada por Bronnikov, a qual consiste em aproximar a solução de mínimos quadrados de modo estável. Essa regularização consiste na minimização do seguinte funcional

$$\|\mathbf{g}_\perp - \mathbf{J}_\mu \Delta\boldsymbol{\mu}\|^2 + \alpha \|\mathbf{L} \Delta\boldsymbol{\mu}\|^2, \quad (3.8)$$

onde $\mathbf{L}$ é uma matriz $P \times P$ que representa a discretização por elementos finitos do operador de Laplace e $\alpha > 0$ é parâmetro de regularização. O segundo termo do funcional da Equação 3.8 é chamado de funcional regularizador e sua função é estabilizar a solução do problema dando, por exemplo, suavidade. A influência desse funcional regularizador é controlada pelo parâmetro de regularização α que é estimado através de tentativa e erro usando simulações de problemas menores. O minimizador do funcional na Equação 3.8 é dado pela solução do seguinte sistema

$$\mathbf{C}_\mu^\alpha \Delta\boldsymbol{\mu} = \mathbf{J}_\mu^T \mathbf{g}_\perp,$$

onde $\mathbf{C}_\mu^\alpha = \alpha \mathbf{L}^T \mathbf{L} + \mathbf{J}_\mu^T \mathbf{J}_\mu$ é uma matriz simétrica $P \times P$. Uma boa escolha do parâmetro de regularização garante que a matriz $\mathbf{C}_\mu^\alpha$ seja definida positiva. Assim, o método mais eficiente de resolver esse tipo de sistema é baseado na decomposição de Cholesky. Aplicando a decomposição de Cholesky é obtido $\mathbf{C}_\mu^\alpha = \mathbf{U}^T \mathbf{U}$, onde $\mathbf{U}$ é uma matriz triangular superior. Portanto, o sistema (3.9) pode ser facilmente resolvido e a solução regularizada $\Delta\boldsymbol{\mu}_\alpha$ é usada para atualizar o vetor do mapa de atenuação.

Esse algoritmo tem como objetivo principal reconstruir de modo estável o mapa de atenuação, ou seja, a imagem de atividade resultante da aplicação da inversa de Moore-Penrose nos dados pode ser instável e muitas vezes inviável em termos práticos. Assim, é necessário a aplicação de um método de reconstrução que busque essa estabilidade, como por exemplo, o algoritmo EM-TV apresentado no Capítulo 2 ou algum método de regularização.

As tentativas de simulações deste algoritmo não foram bem sucedidas, portanto não serão apresentadas. Além disso, não foi possível encontrar na literatura reproduções do algoritmo do Bronnikov sem que tenham sido feitas pelo próprio.

RECONSTRUÇÃO SIMULTÂNEA UTILIZANDO TV-CSP

Neste capítulo vamos apresentar o algoritmo criado para resolver o sistema

$$\mathbf{R}_\mu \mathbf{f} = \mathbf{g}, \quad (4.1)$$

onde $\mathbf{R}_\mu$ é a matriz da discretização do problema, μ é o mapa de atenuação, $\mathbf{f}$ a imagem de atividade e $\mathbf{g}$ são os dados obtidos. Neste caso, nosso esforço será em reconstruir simultaneamente as imagens de atividade e atenuação, $\mathbf{f}$ e μ respectivamente, utilizando somente uma leitura tomográfica do tipo SPECT.

Na literatura, a maioria dos algoritmos de reconstrução simultânea sofrem com o efeito *cross-talk* (e.g., (DICKEN, 1999; NUYTS *et al.*, 1999; KROL *et al.*, 2001; YAN; ZENG, 2009; KROL *et al.*, 2002)), ou seja, a sobreposição mútua das imagens de atividade e atenuação numa espécie de “balanço” entre seus valores. Não estamos interessados em soluções desse tipo, pois não costumam ser representantes da realidade em problemas práticos. O principal fator para tal efeito, que é intrínseco ao problema, é o mal-condicionamento do sistema (4.1), contudo, alguns algoritmos podem agravar esse efeito devido a estrutura de suas funções objetivo, como é o caso do algoritmo MLAA (NUYTS *et al.*, 1999), onde a função objetivo é não-convexa com relação à variável μ .

A estratégia utilizada no algoritmo que será apresentado neste capítulo consiste em dar uma condição para que as imagens reconstruídas tenham estruturas que favoreçam sua utilização na prática. Essa ideia é simples e já foi abordada de maneira semelhante na literatura. Nuyts em (NUYTS *et al.*, 1999) e Yan e Zeng em (YAN; ZENG, 2009) buscaram soluções com algoritmos do tipo ML-EM (*Maximum Likelihood - Expectation Maximization*). Nuyts, estimulou soluções com suavidade local e, através de distribuições de probabilidade bi e trimodal, favoreceu valores de atenuação que eram previamente conhecidos, contudo, esse tipo de estratégia não removeu o efeito *cross-talk* entre as reconstruções. Já Yan e Zeng segmentaram as imagens reconstruídas

pelo ML-EM identificando as regiões de interesse e, aplicando as condições de consistência para somente três valores possíveis de coeficiente de atenuação, buscavam atribuir valores a essas regiões. As imagens reconstruídas por esse último método foram aceitáveis, mas em termos práticos cada paciente possui valores diferentes de atenuação.

A característica comum em imagens de tomografia que será utilizada como condição de aplicabilidade no nosso algoritmo será a de que as imagens que geraram os dados são constantes por partes, pois estamos interessados em regiões que representem órgãos e tumores sem a necessidade de um detalhamento mais complexo. Em contraste com outros métodos desenvolvidos, essa abordagem deixa mais livre os valores que podem ser atribuídos às imagens que queremos reconstruir. A medida de quanto uma imagem é não-constante por partes será dada pela função de *variação total* (TV, abreviatura em inglês), que, apesar de receber o mesmo nome da regularização apresentada [subseção 2.1.3](#), não são equivalentes e será redefinida neste capítulo.

Uma vez que os dados possuem ruídos provenientes de erros de discretização e outros fatores físicos, sabemos que a solução exata do sistema (4.1) pode não ser aquela que esperávamos. Dessa forma, conforme ([CENSOR *et al.*, 1979](#)), estamos interessados em todas imagens de atividade e atenuação que geram dados próximos dos dados obtidos pela leitura SPECT. Para obter pelo menos uma dessas soluções vamos utilizar o método criado por Censor e Lent ([CENSOR; LENT, 1982](#)), o CSP (*Cyclic Subgradient Projection*), para resolver um problema misto de viabilidade côncavo-convexo.

Assim, o algoritmo desenvolvido nessa pesquisa tem como objetivo minimizar a função de variação total TV, que mede a não-constância por partes das imagens, sujeita às condições de viabilidade impostas pelo algoritmo de viabilidade do Yair Censor utilizando o CSP. Essa estrutura criada para o algoritmo foi motivada pelo fato de, num primeiro momento da pesquisa, buscar produzir resultados numéricos do método criado por Censor ([CENSOR *et al.*, 1979](#)) dos quais não há registro na literatura. Após a verificação de resultados pixelados e não-constantes por partes é que se visualizou a importância de, dentre as soluções viáveis, priorizar aquelas que tem características semelhantes às imagens originais.

O capítulo será dividido em três seções principais: na primeira, descreveremos o algoritmo de resolução do problema misto de viabilidade utilizando o CSP; em seguida, apresentaremos o algoritmo desenvolvido que busca, dentre todas as soluções viáveis, por imagens constantes por partes; e na última, seção dos resultados, vamos analisar os parâmetros utilizados nos algoritmos e apresentaremos os resultados obtidos para diversas resoluções de dados.

4.1 Problema misto de viabilidade côncavo-convexo utilizando CSP

O seguinte método foi criado por Censor *et. al.* em 1979 (CENSOR *et al.*, 1979) e seu objetivo é encontrar soluções viáveis para o sistema (4.1). O propósito é reduzir esse sistema a um problema misto de viabilidade côncavo-convexo e aplicar o *método de Projeções Cíclicas do Subgradiente* (CSP, abreviação em inglês) que também foi criado por Censor e Lent em 1978, mas somente publicado em 1982 (CENSOR; LENT, 1982).

O sistema (4.1) é gerado a partir da transformada de Radon atenuada discretizada para tomografias SPECT com o mapa de atenuação $\boldsymbol{\mu} \in \mathbb{R}^P$ e cada elemento da matriz não-linear $\mathbf{R}_{\boldsymbol{\mu}} \in \mathbb{R}^{D \times P}$ determinado por

$$r_{ij} = \sigma_{ij} \exp \left(- \sum_{k \in K_{ij}} \sigma_{ik} \mu_k \right), \quad (4.2)$$

onde σ_{ij} é o comprimento da intersecção do raio i no pixel j e $K_{ij} \subset \{1, 2, \dots, P\}$ é o conjunto de índices dos pixels intersectados pela semirreta que parte do pixel j em direção ao detector e que está contida na reta i . Os dados obtidos na leitura são determinados pelo vetor $\mathbf{g} \in \mathbb{R}^D$ e as equações do sistema são expressas da seguinte forma

$$\sum_{j=1}^P r_{ij} \mathbf{f}_j = \sum_{j \in K_i} \sigma_{ij} \mathbf{f}_j \exp \left(- \sum_{k \in K_{ij}} \sigma_{ik} \mu_k \right) = \mathbf{g}_i,$$

onde $K_i \subset \{1, 2, \dots, P\}$ é o conjunto de índices dos pixels intersectados pela reta i .

Para deixar essas equações sem dimensionalidade, as dividimos por algum valor fixado de concentração da atividade, $\bar{f}$, tal que $\hat{\mathbf{g}}_i = \mathbf{g}_i / \bar{f}$ são as medidas modificadas. Assim, fazendo uma mudança de variáveis,

$$\exp(\mathbf{h}_j) = \mathbf{f}_j / \bar{f} \quad (4.3)$$

com $\mathbf{h} \in \mathbb{R}^P$, as equações são reescritas da seguinte forma

$$p_i(\mathbf{h}, \boldsymbol{\mu}) = \hat{\mathbf{g}}_i, \quad i = 1, 2, \dots, D, \quad (4.4)$$

onde

$$p_i(\mathbf{h}, \boldsymbol{\mu}) = \sum_{j \in K_i} \sigma_{ij} \exp \left(\mathbf{h}_j - \sum_{k \in K_{ij}} \sigma_{ik} \mu_k \right). \quad (4.5)$$

Uma propriedade importante sobre as funções $p_i(\mathbf{h}, \boldsymbol{\mu})$ é que para todo i elas são **convexas** em $\mathbf{h}$ e $\boldsymbol{\mu}$ conjuntamente. Além disso, esse fator de adimensionalidade, $\bar{f}$, que será explicado na Seção 4.3, serve para que o algoritmo não gere soluções onde alguns pixels tendam a menos infinito, pois a imagem $\mathbf{h}$ depende logaritmicamente da imagem $\mathbf{f}$ e, se essa segunda tem pixels nulos, teremos interferência na busca por uma solução viável.

O sistema gerado por (4.4) é esparso já que cada função p_i tem derivadas não nulas com respeito a poucas variáveis dentre as $2P$ descritas na modelagem. Ou seja, unindo isso à

imprecisão dos dados obtidos, ao ruído estatístico intrínseco do problema físico e à discretização do modelo, temos que não faz sentido uma busca pela solução exata, uma vez que ela pode não ser desejada em termos práticos. Dessa forma, o método criado por Censor *et. al.*, ao invés de resolver a [Equação 4.1](#), limita seus esforços para encontrar uma solução suficientemente “próxima” da exata encontrando algum ponto viável para o seguinte conjunto de restrições

$$\widehat{\mathbf{g}}_i - \varepsilon_i \leq p_i(\mathbf{h}, \boldsymbol{\mu}) \leq \widehat{\mathbf{g}}_i + \varepsilon_i, \quad i = 1, 2, \dots, D, \quad (4.6)$$

onde ε_i são tolerâncias previamente escolhidas.

Abreviando o par de variáveis, $\mathbf{x} = (\mathbf{h}, \boldsymbol{\mu}) \in \mathbb{R}^{2P}$, e denotando $Q_i(\mathbf{x}) = p_i(\mathbf{x}) - \widehat{\mathbf{g}}_i - \varepsilon_i$ e $R_i(\mathbf{x}) = p_i(\mathbf{x}) - \widehat{\mathbf{g}}_i + \varepsilon_i$, o problema de viabilidade (4.6) é agora interpretado como: encontrar $\mathbf{x} \in \mathbb{R}^{2P}$ tal que

$$Q_i(\mathbf{x}) \leq 0 \text{ e } R_i(\mathbf{x}) \geq 0 \text{ para todo } i = 1, 2, \dots, D, \quad (4.7)$$

onde $Q_i(\mathbf{x})$ e $R_i(\mathbf{x})$ são funções convexas. Esse tipo de problema é chamado de *problema misto de viabilidade côncavo-convexo*.

Para a implementação do algoritmo vamos apresentar o conceito de subdiferencial. Chamamos $\partial f(\mathbf{x})$ de subdiferencial de uma determinada função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ em $\mathbf{x} \in \mathbb{R}^n$ (que é nunca vazio, veja (GILES, 1982)) o conjunto definido como

$$\partial f(\mathbf{x}) := \{\mathbf{f} \in \mathbb{R}^n | f(\mathbf{x}) + \langle \mathbf{f}, \mathbf{z} - \mathbf{x} \rangle \leq f(\mathbf{z}), \forall \mathbf{z} \in \mathbb{R}^n\},$$

onde cada $\mathbf{f} \in \partial f(\mathbf{x})$ é chamado de subgradiente de f em $\mathbf{x}$. Além disso, vamos definir

$$[x]_- = \min\{x, 0\} \text{ e } [x]_+ = \max\{x, 0\}.$$

Assim, para encontrar um ponto viável em (4.7) Censor *et. al.* utilizaram o CSP (CENSOR *et al.*, 1979), e o algoritmo pode ser escrito da seguinte forma

1. Inicialize as imagens com $\mathbf{x}^{(0)} \in \mathbb{R}^{2P}$ arbitrário, o número máximo de iterações $K = cD$, onde $c \in \mathbb{N}$ é o número de ciclos e D é o número de dados, e escolha os parâmetros de relaxação $\tau_k, \nu_k \in (\delta, 2 - \delta]$ com $\delta \in [0, 1]$;
2. $k \leftarrow 0$;
3. $i_k \leftarrow k(\bmod D) + 1$;
4. Projeções do subgradiente:

a) Restrição côncava:

$$\mathbf{x}^{(k+1/2)} \leftarrow \begin{cases} \mathbf{x}^{(k)} - \tau_k \frac{[R_{i_k}(\mathbf{x}^{(k)})]_-}{\|\mathbf{r}^{(k)}\|^2} \mathbf{r}^{(k)}, & \text{se } \mathbf{r}^{(k)} \neq \mathbf{0}, \\ \mathbf{x}^{(k)}, & \text{caso contrário,} \end{cases}$$

onde $\mathbf{r}^{(k)} \in \partial R_{i_k}(\mathbf{x}^{(k)})$ é um subgradiente de R_{i_k} em $\mathbf{x}^{(k)}$ e τ_k é o tamanho de passo;

b) Restrição convexa:

$$\mathbf{x}^{(k+1)} \leftarrow \begin{cases} \mathbf{x}^{(k+1/2)} - v_k \frac{[Q_{i_k}(\mathbf{x}^{(k+1/2)})]_+}{\|\mathbf{q}^{(k)}\|^2} \mathbf{q}^{(k)}, & \text{se } \mathbf{q}^{(k)} \neq \mathbf{0}; \\ \mathbf{x}^{(k+1/2)}, & \text{caso contrário,} \end{cases}$$

onde $\mathbf{q}^{(k)} \in \partial Q_{i_k}(\mathbf{x}^{(k+1/2)})$ é um subgradiente de Q_{i_k} em $\mathbf{x}^{(k+1/2)}$ e v_k é o tamanho de passo;

5. Se $k = K$ então pare, senão $k \leftarrow k + 1$ e volte para o passo 3.

Para o caso de uma restrição puramente convexa, é possível mostrar que os tamanhos de passo v_k escolhidos garantem uma convergência para o conjunto factível convexo (HELOU; PIERRO, 2009). Contudo, não há na literatura uma demonstração para a convergência no caso acima e nem uma condição sobre o tamanho de passo τ_k para que nos aproximemos do conjunto côncavo. O próprio Censor afirma: “temos evidências de que o método se aplica também para o problema misto de viabilidade côncavo-convexo”, quando se referia ao CSP aplicado à viabilidade convexa (CENSOR *et al.*, 1979). Essas evidências ficam claras em termos práticos, onde a reconstrução, na ausência de ruído, é satisfatória.

Os subgradientes $\mathbf{r}^{(k)}$ e $\mathbf{q}^{(k)}$ são facilmente calculados diferenciando a Equação 4.5, pois $\hat{\mathbf{g}}_i$ e ε_i são constantes. Um sugradiente de Q_i num determinado ponto também é subgradiente de R_i neste mesmo ponto. Assim, vamos apresentar somente o cálculo do subgradiente de Q_i . A derivada de p_i com relação à $\mathbf{h}_l$ é dada por

$$\frac{\partial p_i}{\partial \mathbf{h}_l}(\mathbf{h}, \boldsymbol{\mu}) = \sigma_{il} \exp\left(\mathbf{h}_l - \sum_{k \in K_{ij}} \sigma_{ik} \boldsymbol{\mu}_k\right) = \sigma_{il} \exp\left(- \sum_{k \in K_{ij}} \sigma_{ik} \boldsymbol{\mu}_k\right) \exp(\mathbf{h}_l),$$

e assim,

$$\nabla_{\mathbf{h}} Q_i(\mathbf{h}, \boldsymbol{\mu}) = \nabla_{\mathbf{h}} p_i(\mathbf{h}, \boldsymbol{\mu}) = \mathbf{R}_{\mu,i}^T \circ \exp(\mathbf{h}),$$

onde $\mathbf{R}_{\mu,i}$ é a i -ésima linha da matriz $\mathbf{R}_{\mu}$, $(\circ)$ é o produto de Hadamard e $\exp(\mathbf{h})$ é um vetor onde as coordenadas são respectivamente as exponencias das coordenadas de $\mathbf{h}$. Já a derivada de p_i com relação à $\boldsymbol{\mu}_l$ é dada por

$$\frac{\partial p_i}{\partial \boldsymbol{\mu}_l} = -\sigma_{il} \sum_{j|l \in K_{ij}} \sigma_{ij} \exp\left(\mathbf{h}_l - \sum_{k \in K_{ij}} \sigma_{ik} \boldsymbol{\mu}_k\right)$$

e esse subgradiente com relação à $\boldsymbol{\mu}$ não tem uma representação vetorial simples. Portanto, o subgradiente $\mathbf{q} \in \partial Q_i(\mathbf{x})$ do algoritmo acima é representado como

$$\mathbf{q} = \left(\frac{\partial p_i}{\partial \mathbf{h}_1}(\mathbf{x}), \dots, \frac{\partial p_i}{\partial \mathbf{h}_P}(\mathbf{x}), \frac{\partial p_i}{\partial \boldsymbol{\mu}_1}(\mathbf{x}), \dots, \frac{\partial p_i}{\partial \boldsymbol{\mu}_P}(\mathbf{x}) \right)^T.$$

Utilizando algoritmos como o *traçado de raios*, podemos fazer isso em tempo computacional baixo (ver CENSOR, 2006).

Daqui em diante, vamos nos referir ao algoritmo apresentado nesta seção somente como CSP. Portanto, a próxima seção irá abordar o algoritmo TV-CSP, junção do algoritmo de viabilidade CSP com a minimização de um funcional que determina uma característica nas imagens a serem reconstruídas.

4.2 Algoritmo TV-CSP

O algoritmo CSP tem como principal objetivo buscar imagens cujo sinograma esteja próximo aos dados obtidos, pois talvez não haja solução exata ou, se existir, pode não representar a realidade, pois o problema apresentado é não-linear e mal-posto. Ao dar uma folga no sinograma das imagens que queremos reconstruir, aumentamos a chance de que tenhamos uma solução e que, dependendo da tolerância, as imagens originais estejam no conjunto viável do CSP. Mesmo que as imagens originais estejam no conjunto viável não há esforços do algoritmo pela busca dessas imagens. Na seção de resultados é possível ver através da [Figura 19](#) que, mesmo com tamanhos de passo máximos, ou seja, valores que maximizam a aproximação da iteração atual do conjunto factível, o algoritmo atingiu uma solução satisfatória, mas bastante ruidosa.

Assim, se, dentre as imagens do conjunto viável do CSP, procurarmos por aquelas que possuam características semelhantes às imagens originais, podemos diminuir o nosso conjunto de soluções. Com essa ideia, vamos procurar por soluções onde as imagens de atenuação e atividade sejam constantes por partes, pois detectaríamos formato e posição dos órgãos e tecidos. Entretanto, essa estratégia não é indicada em casos onde há a necessidade de visualização das fibras e/ou detalhes físicos dos órgãos do paciente em estudo.

Para medir o quanto uma imagem é constante por partes, vamos utilizar a função de variação total apresentada em ([HELOU; PIERRO, 2009](#)). Assim, vamos assumir que uma determinada imagem $\mathbf{x} \in \mathbb{R}^{mm}$ tem seus valores representados por uma matriz/imagem 2D e cada elemento/pixel (i, j) representado por $x_{i,j}$, então a função de variação total é dada por

$$TV(\mathbf{x}) := \sum_{i=1}^m \sum_{j=1}^n \sqrt{(x_{i,j} - x_{i-1,j})^2 + (x_{i,j} - x_{i,j-1})^2},$$

onde $x_{i,0} = x_{0,j} = 0$. Essa função é convexa e para mostrar tal propriedade, seja $A : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ definida como $A(x, y, z) = (x - y, x - z)^T$, então cada parcela da função TV pode ser escrita como $\|A(x_{i,j}, x_{i-1,j}, x_{i,j-1})\|_2$. Uma vez que A é uma transformação linear e $\|\cdot\|_2$ é convexa, então TV é convexa, pois é uma soma de funções convexas. A vantagem da função TV ser convexa está no fato de que existe uma ampla literatura sobre a minimização de funções convexas que utilizam subgradientes ([GILES, 1982](#); [ALBER; IUSEM; SOLODOV, 1988](#); [SOLODOV; ZAVRIEV, 1998](#); [NEDÍĆ; BERTSEKAS, 2001](#); [BLATT; HERO; GAUCHMAN, 2007](#); [HELOU; PIERRO, 2009](#)).

Assim, o método proposto para a reconstrução simultânea somente com dados SPECT consiste em resolver um problema de minimização da soma das variações totais de $\mathbf{h}$ e $\boldsymbol{\mu}$, onde

$\mathbf{h}$ é definida na [Equação 4.3](#), sujeita as restrições do CSP, ou seja,

$$\begin{aligned} & \text{minimizar} \quad TV(\mathbf{h}) + TV(\boldsymbol{\mu}) \\ & \text{s.a.:} \quad Q_i(\mathbf{h}, \boldsymbol{\mu}) \leq 0, \quad i = 1, 2, \dots, D; \\ & \quad \quad R_i(\mathbf{h}, \boldsymbol{\mu}) \geq 0, \quad i = 1, 2, \dots, D; \\ & \quad \quad \boldsymbol{\mu} \geq 0, \end{aligned} \tag{4.8}$$

onde $\mathbf{h}$ é livre, pois quando $\mathbf{f}$ possui um pixel nulo, o respectivo pixel em $\mathbf{h}$ tem como valor $-\infty$. Contudo, esses pixels dificultam a minimização da função objetivo, pois a imagem de atividade original, constante por partes, possui valores nulos, como pixels fora do corpo, e a variação total da sua respectiva imagem $\mathbf{h}$ é $+\infty$. Para resolver este problema, em termos práticos, utilizaremos os valores nulos do sinograma de dados para diminuir o conjunto de pixels que deverão ser reconstruídos e indentificar, a menos de alguns pixels extras, o suporte das imagens supondo que há atividade em todo objeto de estudo. Esses pixels extras, que não pertencem ao suporte das imagens originais, assumem valores nulos e assim, para identificá-los, escolheremos um valor para o fator de adimensionalidade que faça $\mathbf{h} > 0$ no suporte original. Portanto, durante a execução do algoritmo que será apresentado nesse capítulo, assumiremos que qualquer valor negativo na imagem $\mathbf{h}$ é um pixel que não pertence ao suporte original das imagens, e a partir daí modificaremos o suporte identificado através do sinograma de dados excluindo este pixel.

O TV-CSP, como chamaremos o algoritmo que busca resolver o problema (4.8), será composto de dois passos: no primeiro, será executado um ciclo do CSP para as restrições da minimização, ou seja, daremos um passo em direção ao conjunto viável; no outro, será executada uma iteração do método de subgradientes com tamanho de passo simples para minimizar a função objetivo e, para isso, definiremos o pixel (i, j) da imagem do subgradiente $\mathbf{t}$ da função TV como

$$\begin{aligned} t_{i,j} = & \frac{2x_{i,j} - x_{i-1,j} - x_{i,j-1}}{\sqrt{(x_{i,j} - x_{i-1,j})^2 + (x_{i,j} - x_{i,j-1})^2}} \\ & + \frac{x_{i,j} - x_{i,j+1}}{\sqrt{(x_{i,j+1} - x_{i,j})^2 + (x_{i,j+1} - x_{i-1,j+1})^2}} \\ & + \frac{x_{i,j} - x_{i+1,j}}{\sqrt{(x_{i+1,j} - x_{i,j})^2 + (x_{i+1,j} - x_{i+1,j-1})^2}} \end{aligned}$$

e quando um denominador de uma dessas parcelas for nulo vamos desconsiderá-la. Assim, utilizando a abreviação $\mathbf{x} = (\mathbf{h}, \boldsymbol{\mu}) \in \mathbb{R}^{2P}$, o algoritmo pode ser escrito como:

1. Inicialize as imagens com $\mathbf{h}^{(0)}, \boldsymbol{\mu}^{(0)} \in \mathbb{R}^P$ arbitrários. Sejam K o número máximo de iterações, $c \in \mathbb{N}$ o número de ciclos e D o número de dados, escolha os parâmetros de relaxação $\tau_i, v_i \in (\delta, 2 - \delta]$ com $\delta \in [0, 1)$ e escolha o tamanho de passo inicial λ_0 ;
2. $k \leftarrow 0$;
3. $\mathbf{x}^{(k)} \leftarrow (\mathbf{h}^{(k)}, \boldsymbol{\mu}^{(k)})$;

4. $i \leftarrow 1$;
5. $\mathbf{x}^{(k,i)} \leftarrow \mathbf{x}^{(k)}$;
6. Calcule a próxima subiteração com a projeção do subgradiente (algoritmo CSP do Censor):

a) Restrição côncava:

$$\mathbf{x}^{(k,i+1/2)} \leftarrow \begin{cases} \mathbf{x}^{(k,i)} - \tau_i \frac{[R_i(\mathbf{x}^{(k,i)})]_-}{\|\mathbf{r}^{(k,i)}\|^2} \mathbf{r}^{(k,i)}, & \text{se } \mathbf{r}^{(k,i)} \neq \mathbf{0}, \\ \mathbf{x}^{(k,i)}, & \text{caso contrário,} \end{cases}$$

onde $\mathbf{r}^{(k,i)} \in \partial R_i(\mathbf{x}^{(k,i)})$ é um subgradiente de R_i em $\mathbf{x}^{(k,i)}$ e τ_i é o tamanho de passo;

b) Restrição convexa:

$$\mathbf{x}^{(k,i+1)} \leftarrow \begin{cases} \mathbf{x}^{(k,i+1/2)} - v_i \frac{[Q_i(\mathbf{x}^{(k,i+1/2)})]_+}{\|\mathbf{q}^{(k,i)}\|^2} \mathbf{q}^{(k,i)}, & \text{se } \mathbf{q}^{(k,i)} \neq \mathbf{0}; \\ \mathbf{x}^{(k,i+1/2)}, & \text{caso contrário,} \end{cases}$$

onde $\mathbf{q}^{(k,i)} \in \partial Q_i(\mathbf{x}^{(k,i+1/2)})$ é um subgradiente de Q_i em $\mathbf{x}^{(k,i+1/2)}$ e v_i é o tamanho de passo;

7. Se $i < D$, então $i \leftarrow i + 1$ e volte para o passo 6; senão continue no passo 8.
8. $(\mathbf{h}^{(k)}, \boldsymbol{\mu}^{(k)}) \leftarrow \mathbf{x}^{(k,D+1)}$;
9. Se $\mathbf{h}^{(k)}$ possui pixels com valores negativos, então os tire do suporte das imagens;
10. Calcule a próxima iteração com o método de subgradientes simples:

a) Imagem de atividade:

$$\mathbf{h}^{(k+1)} \leftarrow \mathbf{h}^{(k)} - \lambda_k \mathbf{t}_h^{(k)},$$

onde $\mathbf{t}_h^{(k)} \in \partial TV(\mathbf{h}^{(k)})$;

b) Imagem de atenuação:

$$\boldsymbol{\mu}^{(k+1)} \leftarrow \boldsymbol{\mu}^{(k)} - \lambda_k \mathbf{t}_\mu^{(k)},$$

onde $\mathbf{t}_\mu^{(k)} \in \partial TV(\boldsymbol{\mu}^{(k)})$;

11. Se $k = K$ então pare, senão $k \leftarrow k + 1$ e volte para o passo 3.

Em todos os passos desse algoritmo as variáveis estão sendo manipuladas somente nos pixels que pertencem ao suporte gerado através do sinograma de dados, sendo que esse suporte pode sofrer alterações no passo 9. Na próxima seção, explicaremos em termos práticos o que seria esse suporte gerado com mais clareza.

A convergência do algoritmo não foi demonstrada, pois tem como base o algoritmo CSP que não tem demonstração de convergência na literatura e, acrescentando a função objetivo de variação total, fica ainda mais difícil de demonstrar. Contudo, os resultados obtidos no caso da ausência de ruídos tendem, após muitas iterações, à solução original exata mesmo para poucos dados, dando assim evidências de uma possível demonstração de convergência.

4.3 Resultados

Nesta seção apresentaremos os resultados obtidos pelo algoritmo TV-CSP cuja implementação, assim como no [Capítulo 2](#), utiliza a biblioteca RAFT, que tem distribuição gratuita e é escrita em C/C++ (MIQUELES; HELOU, 2007 (acessado em 11 de junho de 2015)). Algumas das funções necessárias para a implementação do TV-CSP não estavam disponíveis na RAFT, mas, através de alterações no código fonte, foi possível obtê-las. Por exemplo, a função que obtém o subgradiente com relação a μ das funções R_i e Q_i foi criada através de alterações no código da função que obtém a transformada de Radon.

Para avaliar a eficiência do algoritmo vamos utilizar um *phantom* numérico, uma imagem artificial, que consiste de um modelo bastante simples de uma seção do tórax humano. A [Figura 17](#) ilustra as imagens originais para o mapa de atenuação e a imagem de atividade. A origem do plano cartesiano se encontra no centro dessas imagens. Cada imagem é composta por uma soma de funções indicadoras de elipses, onde ao interior é atribuído algum valor específico, conforme indicado nas [Tabelas 1 e 2](#). A imagem de atividade, consiste de um anel, representando a atividade no coração, e uma elipse, correspondendo simplificadaamente, aos outros tecidos da seção do tórax. A radiação do anel é cerca de 10 vezes maior do que a da seção do tórax. Nesse modelo, utilizamos 150cm^{-1} como densidade da atividade no anel. O mapa de atenuação é constituído pela elipse da seção do tórax com atenuação equivalente à $0,16\text{cm}^{-1}$, duas elipses menores com coeficiente de atenuação $0,04\text{cm}^{-1}$ representando os pulmões e uma pequena circunferência com a maior atenuação, $0,28\text{cm}^{-1}$, correspondendo à espinha dorsal.

Tabela 1 – Informações das elipses que compõem a imagem de atividade.

Coordenadas do centro*	Eixo maior*	Eixo menor*	Ângulo de rotação	Atividade $\lambda \text{ (cm)}^{-1}$
(0,0; 0,0)	0,855	0,7557	0°	15
(-0,15; 0,4)	0,1414	0,1414	0°	150
(-0,15; 0,4)	0,1	0,1	0°	15

* valores de proporção referentes ao comprimento e à largura da imagem real.

Para se obter uma iteração do algoritmo devemos executar, pelo menos, até duas vezes o par projeção e retroprojeção da transformada de Radon para cada um dos raios/dados do problema, ou seja, o TV-CSP tem um custo computacional muito caro. Assim, precisamos

Figura 17 – A imagem de atividade (esquerda) e o mapa atenuação (direita) do modelo de estudo proposto. Representam, simplificada, uma seção do tórax humano e um tumor no coração.

Tabela 2 – Informações das elipses que compõem o mapa de atenuação.

Coordenadas do centro*	Eixo maior*	Eixo menor*	Ângulo de rotação	Atenuação $\mu \text{ (cm)}^{-1}$
(0,0; 0,0)	0,855	0,7557	0°	0,16
(-0,4; 0,0)	0,5155	0,3251	90°	0,04
(0,4; 0,0)	0,4212	0,3706	90°	0,04
(0,0; -0,5)	0,1	0,1	0°	0,28

* valores de proporção referentes ao comprimento e à largura da imagem real.

diminuir a discretização do modelo do tórax, no caso, escolhemos trabalhar com imagens de resolução 32×32 , ou seja, 1024 pixels.

Daqui por diante, quando falarmos sobre a resolução ou a quantidade de dados num sinograma vamos representá-la na seguinte forma: $V \times D$, onde V é número de vistas igualmente distribuídas ao longo de 360° e D é o número de detectores igualmente distribuídos entre $-\sqrt{2}$ e $\sqrt{2}$ em cada vista, pois estamos assumindo, sem perda de generalidade, que as imagens estão contidas no interior de um quadrado centrado na origem cujo o lado mede dois. As simulações executadas assumiram diferentes resoluções de dados com número de detectores fixo, 32, e com números de vistas iguais à 16, 19, 32 e 40.

Para acelerar os cálculos do algoritmo direcionamos nossos esforços à reconstrução do suporte das funções de atividade e atenuação, ou seja, somente dos pixels não nulos. A identificação do suporte utilizado no algoritmo é feita através dos pixels nulos do sinograma de dados. Cada um desses pixels do sinograma representa uma integral de linha, sobre uma reta, aplicada a uma função de valores não-negativos. Assim, se essa integral é nula, partindo da premissa de que todo pixel do objeto de estudo tem valor de atividade positivo, temos que os pixels utilizados nessa integral serão nulos e assim, se torna possível, a partir de uma imagem constante com valores iguais a 1 nos pixels, a representação do suporte das imagens a serem reconstruídas. Porém, através dessa estratégia, a quantidade de pixels do suporte depende da resolução dos dados do problema, por exemplo, num caso extremo, quando a resolução dos

dados é 4×32 , as vistas serão tomadas sobre os ângulos 0° , 90° , 180° e 270° , e os suportes das imagens serão dados por um retângulo.

Dado um sistema de equações, chamaremos de *taxa de descrição* a razão do número de equações pelo número de variáveis, o que pode ser representado como uma porcentagem e indicar a quantidade de dados obtidos em relação à quantidade necessária para igualar ao número de variáveis. Sistemas com taxa de descrição inferior a 100% são necessariamente mal-postos. Contudo, mesmo que a taxa de descrição seja maior ou igual que 100% não podemos garantir que o sistema seja bem-posto, mas teremos esperança de obter uma boa solução. Por exemplo, um sinograma de resolução 40×32 tem 1280 pixels, e destes, 1028 estão no suporte, assim, como vamos reconstruir imagens de resolução 32×32 , onde cada uma, com o número de pixels reduzido através dos valores nulos do sinograma, terá um suporte com 500 pixels, e assim, criaremos um sistema com taxa de descrição dada por $1028/(2 \times 500) = 102,8\%$. As taxas de descrição das simulações alvo de análise nesse texto estão apresentadas na [Tabela 3](#).

A quantidade de pixels no suporte das funções originais de atividade e atenuação para uma resolução 32×32 é de 492 pixels cada. Mesmo para quantidades altas de dados é provável que o suporte gerado através do sinograma tenha mais pixels que o suporte das funções originais. Esses pixels excedentes tem valor nulo nas imagens originais e, assim, gostaríamos que o algoritmo levasse os valores desses pixels excedentes, na imagem de atividade, para menos infinito, pois a solução buscada pelo CSP trata-se de $\mathbf{h}$, que é o logaritmo da imagem de atividade dividida pelo fator de adimensionalidade, $\mathbf{f}/\bar{\mathbf{f}}$. Contudo, pudemos observar que, para este caso, aproximadamente 100 iterações do TV-CSP são necessárias para que esses pixels fossem os únicos que assumissem valores negativos, porém ainda grandes, normalmente maiores que -3 . Como todos os pixels da imagem de atividade escolhida são maiores que um, o logaritmo desta teria valores maiores que zero, e assim, para todas as simulações, a partir da centésima iteração, atualizamos os valores dos pixels negativos para -30 . No caso de imagens de atividade que, por experiência, sabe-se que teremos valores menores que um, ou seja, com o logaritmo negativo, basta escolher um fator de adimensionalidade adequado para que somente os pixels excedentes assumam valores negativos na imagem $\mathbf{h}$ a ser reconstruída.

Tabela 3 – Taxa de descrição baseada no suporte da função do sinograma para diferentes resoluções.

Resolução dos dados	Pixels no suporte		Taxa de descrição
	Sinograma	Reconstrução (cada)	
40×32	1028	500	102,8%
32×32	828	496	83,47%
19×32	486	496	48,99%
16×32	412	512	40,23%

Agora vamos apresentar os parâmetros utilizados nessas simulações. Primeiramente, o fator de adimensionalidade que divide os dados, $\bar{\mathbf{f}}$, é igual a 1, pois não houve melhoras para outros valores e a imagem de atividade tem todos os pixels maiores que um. As tolerâncias

Figura 18 – Gráficos dos valores máximos dos tamanhos de passo τ_i e v_i do algoritmo CSP de forma que cada subiteração esteja mais próxima possível do conjunto factível côncavo e convexo, respectivamente. A resolução de dados utilizada foi 40×32 .

referentes a solução do problema (4.1), apresentadas em (4.6), foram tomadas constantes em todas as simulações, $\varepsilon_i = 10^{-5}$, $\forall i$.

Os tamanhos de passo $\{\lambda_k\}$ que escolhemos para o método de subgradientes simples da função objetivo são dados por $\lambda_k = \frac{\lambda_0}{k}$, onde, caso não seja informado, tomamos $\lambda_0 = 1$. O algoritmo CSP tem dois tamanhos de passo, τ_i para as restrições côncavas e v_i para as restrições convexas. Sabemos que não temos um limitante para os valores de τ_i . Assim, foram feitas simulações do CSP onde, através do uso da solução original, foram calculados os valores máximos para os parâmetros τ_i de forma que a subiteração atual esteja mais próxima possível do conjunto factível côncavo e, para cada iteração, foi calculado uma média desses valores (cf. Figura 18). Com o propósito de encontrar valores de τ_i que abranjam quantidades de dados diferentes, fizemos o teste para todas as resoluções apresentadas na Tabela 3, mantendo a resolução das imagens reconstruídas como 32×32 . A Figura 19 apresenta o resultado da iteração 3000 do CSP para um sinograma de resolução 40×32 e, τ_i e v_i assumindo valores máximos.

A partir dessas simulações com parâmetros máximos, concluímos que $\tau_i = \tau$ para todo i , onde $\tau \in [0,8; 1,3]$, gera boas soluções. Contudo, essa escolha não garante uma aproximação para o conjunto côncavo, devido à natureza do problema, mas nesse caso funcionou muito bem.

Quanto ao parâmetro v_i , para garantirmos uma aproximação para o conjunto factível convexo, os valores de v_i podem ser, como dito anteriormente, no máximo 2. Nessas simulações foi possível analisar que os valores máximos para os parâmetros v_i , de forma que a subiteração atual esteja mais próxima possível do conjunto factível convexo, podem ser grandes e, como é possível ver na Figura 18, que $v_i = 2$, $\forall i$, é uma boa escolha para acelerar a reconstrução, pois temos uma grande quantidade de valores máximos altos.

Para as imagens de entrada do algoritmo assumimos valores constantes para os pixels do suporte das suas respectivas funções, encontrados através do sinograma. Os pixels que não pertenciam ao suporte foram tomados como nulos. O valor constante utilizado para a atividade e atenuação foram, respectivamente, 10cm^{-1} e $0,028\text{cm}^{-1}$.

Figura 19 – A imagem de atividade (esquerda) e o mapa atenuação (direita) da iteração 3000 do algoritmo CSP com resolução de dados 40×32 e valores máximos τ_i e v_i de tamanhos de passo. As imagens de atividade e atenuação iniciais foram constantes com valores 10cm^{-1} e $0,028\text{cm}^{-1}$, respectivamente.

Figura 20 – O resultado das iterações 80, 300, 1000 e 3000 do algoritmo TV-CSP para o modelo do tórax, onde o conjunto de dados tem resolução 40×32 , 40 vistas ao longo de 360° e 32 detectores em cada vista. A tolerância foi $\varepsilon = 1e-5$ e os tamanhos de passo $\tau = 1,0$ e $v = 2,0$.

Num primeiro momento, estudaremos os resultados do algoritmo para um conjunto de dados desprovido de ruídos, testando a eficiência do algoritmo sob condições ideais. Para isso, apresentaremos os resultados de dados com resolução 40×32 , pois sua taxa de descrição é maior que 1, ou seja, temos mais equações do que variáveis. Mesmo que o sistema seja não-linear, esse é um cenário ideal para reconstrução simultânea. A [Figura 20](#) apresenta algumas iterações do algoritmo TV-CSP. A imagem de atividade é praticamente reconstruída no começo da simulação, com localização, tamanho e formato exatos do tumor. Para o mapa de atenuação, as primeiras iterações são altamente afetadas pelo método de subgradientes da função objetivo, que busca por imagens constantes em todo o suporte da imagem, além disso, há uma grande interferência do efeito *cross-talk*, que desaparece por volta da iteração 150. As estruturas do mapa de atenuação começam a ser definidas depois da iteração 300, onde é possível identificar todos os órgãos

Figura 21 – O resultado de 3000 iterações do algoritmo TV-CSP para o modelo do tórax com as resoluções de dados da Tabela 3. A tolerância foi $\varepsilon = 1e - 5$ e os tamanhos de passo $\tau = 1,0$ e $\nu = 2,0$.

internos e, mais a diante, por volta da iteração 3000, é possível obter resultados semelhantes as imagens originais, validando seu uso clínico.

As simulações dos algoritmos foram executadas num computador que possui processador Intel® Core™ i7-2630QM com 2.00 GHz de frequência, 6 GB de memória RAM e sistema operacional Ubuntu 14.04 LTS. A iteração 3000, apresentada na Figura 20, levou aproximadamente 1 hora e 20 minutos para ser calculada.

Para as outras resoluções de dados, foram executadas 3000 iterações do TV-CSP e os resultados obtidos, exceto pela resolução 16×32 , viabilizaram uma possível utilização clínica em todos os casos, pois as imagens de atividade são idênticas e nos mapas de atenuação todos os órgãos internos, embora os valores nas suas respectivas regiões não sejam constantes, são detectáveis (cf. Figura 21). A resolução 19×32 , para este problema, é a menor resolução de dados com 32 detectores igualmente espaçados que gera uma reconstrução visualmente útil.

Numa execução prática do algoritmo TV-CSP, os erros que comparam as iterações atuais e as soluções, por motivos óbvios, não são possíveis de serem calculados. Assim, nos resta voltar os olhares para os resíduos dos dados, que são calculados pela norma $\|\cdot\|_2$ da diferença entre o sinograma de dados e o sinograma gerado a partir das imagens de uma iteração. Contudo, essa medida não é bem precisa. Suponha que tenhamos reconstruções com resoluções de sinograma distintas e mesmo valor de resíduo em algum momento. Logo, o sinograma da resolução maior estará mais próximo dos dados, pois as parcelas em cada pixel do sinograma gerado serão menores, ou seja, o *resíduo médio* é menor. Além disso, vale a pena lembrar que, um resíduo médio menor não implica numa solução melhor, por isso é importante a minimização da variação total.

Entretanto, como estamos trabalhando com simulações e temos conhecimento sobre as

imagens originais, apresentamos na [Figura 22](#) além do resíduo médio e variação total – que exceto pela variação total original, são informações que obtemos sem o conhecimento das imagens originais – temos os erros para cada uma das imagens que desejamos reconstruir. É possível ver que os erros das imagens de atividade do CSP tem valores altos mesmo tendo reconstruções muito próximas das originais, isso se dá pelo alto valor atribuído aos pixels excedentes no suporte gerado pelo sinograma (cf. [Figura 19](#)), influenciando fortemente nos resíduos médios. Os algoritmos TV-CSP, no passo da otimalidade da função objetivo, atenuam esses valores de forma que assumam valores negativos, e assim, em alguma iteração do algoritmo, sejam forçados a assumir um valor negativo grande, no caso, -30 .

Figura 22 – Gráfico dos resíduos médios (superior esquerdo), das variações totais (superior direito) e dos erros nas imagens de atividade (inferior esquerdo) e atenuação (inferior direito) simulados através 3000 iterações do TV-CSP com $\nu = 2,0$ e $\tau = 1,0$ utilizando as resoluções propostas no texto e 3000 iterações do CSP com resolução 40×32 e parâmetros máximos para ν e τ .

4.3.1 Reconstrução com dados ruidosos

Os resultados apresentados até o momento são reconstruções simultâneas muito próximas das originais e, mesmo que não tenhamos como demonstrar, foi possível ver uma convergência para a solução original. Contudo, essas reconstruções foram calculadas na ausência de ruídos nos dados, o que não exclui o mal-condicionamento num sistema não-linear ou com taxa de descrição baixa. Assim, para avaliar a robustez do TV-CSP, vamos simular os dados ruidosos.

A [Equação 4.2](#) nos permite calcular os dados exatos da seguinte forma:

$$\bar{g}_i = \sum_{j=1}^P r_{ij} \mathbf{f}_j,$$

onde, dado uma atenuação, r_{ij} é um elemento da matriz da transformada de Radon atenuada e f_j é a densidade de atividade no pixel j . Contudo, o cálculo dos dados acima não leva em consideração que o decaimento acontece de acordo com um processo de Poisson com média f_j . Logo, os dados $\bar{g}_i$ da expressão acima são uma combinação de variáveis de Poisson que aqui, por simplificação, tomaremos como independentes. Portanto, os dados de entrada do algoritmo devem ser substituídos por

$$\mathbf{g}_i \sim \mathcal{P}(\alpha \bar{\mathbf{g}}_i), \quad \alpha \geq 1,$$

onde $\mathcal{P}(\lambda)$ é a distribuição de Poisson com média λ e $\alpha \geq 1$ é um fator de multiplicação que interfere no erro relativo aos dados gerados, pois $E[\mathbf{g}_i/\alpha] = \bar{\mathbf{g}}_i$ e $\text{Var}[\mathbf{g}_i/\alpha] = \bar{\mathbf{g}}_i/\alpha$, ou seja, quanto maior α , mais próximo dos dados originais estará os dados gerados.

No CSP buscamos reconstruir imagens em que os valores dos pixels do sinograma difiram no máximo ε_i dos valores originais. Nesse sentido, tomar valores próximos de zero para as tolerâncias ε_i são suficientes. Contudo, no caso onde há ruídos essa escolha não é tão clara. Nos testes utilizando uma tolerância constante para todos os raios, ou seja, $\varepsilon_i = \varepsilon$, não foi possível obter resultados satisfatórios.

O objetivo do CSP com as tolerâncias é dar liberdade para que o sinograma das iterações possa atingir o sinograma original, pois a reconstrução através dos dados ruidosos gera imagens ruidosas ou, na maioria dos casos, leva à uma solução totalmente diferente da esperada. Portanto, como não conhecemos o sinograma $\bar{\mathbf{g}}$, vamos supor que $\alpha \bar{\mathbf{g}}_i$ está no máximo a um múltiplo da raiz quadrada de $\mathbf{g}_i$, ou seja, $|\alpha \bar{\mathbf{g}}_i - \mathbf{g}_i| \leq \beta \sqrt{\mathbf{g}_i}$. Para produzir um ruído razoável aos dados tomamos $\alpha = 40$ e o melhor valor, obtido empiricamente, para β foi 1,5. As imagens reconstruídas pelo algoritmo TV-CSP são apresentadas na [Figura 23](#) (esquerda e meio) e é possível ver que, mesmo que a imagem de atividade tenha ficado satisfatória, o mapa de atenuação não permite a visualização de estruturas que viabilizam uma utilização prática.

A estratégia de escolha da tolerância por um múltiplo da raiz quadrada de $\mathbf{g}_i$ não é suficientemente boa, pois como estamos tratando de distribuições estatísticas, podemos gerar dados que extrapolam a tolerância estabelecida. Nos teste, tomando o parâmetro $\beta = 2,0$, a imagem de atenuação somente identificou a região que representa a coluna do paciente em estudo (cf. [Figura 23](#)). As demais regiões, por influência do método de subgradientes da função objetivo, tenderam a zero, o que implica que a tolerância nesses pixels foi alta demais. Para que o TV-CSP possa ser aplicado em situações práticas é necessário uma escolha mais robusta para as tolerância do sinograma. Assim, utilizamos como tolerância, raio a raio, o valor exato da distância entre o sinograma de dados original e o sinograma de dados ruidosos apresentado na [Figura 23](#) e o resultado obtido ainda não foi satisfatório para a atenuação, mas é possível em ambas imagens identificar o formato e posição de cada órgão/tumor (cf. [Figura 24](#)).

Figura 23 – Primeira linha: sinograma original (esquerda); e sinograma ruidoso (direita) com distribuição de Poisson e erro relativo de 3.9%. Segunda linha: a imagem de atividade (esquerda) e o mapa atenuação (meio) da iteração 1500 do algoritmo TV-CSP com resolução de dados 40×32 , tamanhos de passo para o CSP $\tau = 1,0$ e $\nu = 2,0$, tamanho de passo inicial para o método de subgradientes $\lambda_0 = 2,0$ e fatores de multiplicação $\alpha = 40$ e $\beta = 1,5$. Mapa de atenuação (direita) utilizando $\beta = 2,0$. As imagens de atividade e atenuação iniciais foram constantes com valores $10\alpha \text{ cm}^{-1}$ e $0,028\alpha \text{ cm}^{-1}$, respectivamente.

4.4 Conclusão

O *phantom* utilizado para obter os resultados e avaliar a consistência do algoritmo é simples e não representa as dificuldades de uma situação clínica comum, onde o coração é cercado de vasos sanguíneos, outros órgãos e ossos. É necessário uma reconstrução com dados reais para estudar o comportamento dos algoritmos CSP e TV-CSP em casos mais complexos. Contudo, ao contrário do que se esperava, os resultados apresentados se mostraram bastante satisfatórios.

Na Figura 19 é possível ver o resultado do CSP com tamanhos de passo máximo para um sinograma com resolução 40×32 e taxa descrição igual a 102,8%. Nessa simulação, a criação do suporte das imagens de atividade e atenuação é feita somente através dos pixels nulos do sinograma de dados. Apesar do problema causado por pixels nulos na imagem de atividade original, o CSP apresentou resultados satisfatórios. Contudo, esses resultados ainda estão muito ruidosos para uma aplicação real.

Figura 24 – A imagem de atividade (esquerda) e o mapa de atenuação (direita) da iteração 1500 do algoritmo TV-CSP com resolução de dados 40×32 , tamanhos de passo para o CSP $\tau = 1,0$ e $\nu = 2,0$, tamanho de passo inicial para o método de sugradients $\lambda_0 = 2,0$ e fator de multiplicação $\alpha = 40$. As tolerâncias foram tomadas, raio a raio, como a distância entre o sinograma original e o ruidoso.

O algoritmo TV-CSP foi criado aqui para encontrar, dentre todas as soluções viáveis do CSP, aquela que possui menor variação total entre seus pixels. Os resultados obtidos por esse algoritmo foram excelentes mesmo para uma quantidade baixa de dados que, no caso da resolução 19×32 , temos menos dados do que seria necessário para uma reconstrução SPECT onde o mapa de atenuação é conhecido.

Comparado aos resultados encontrados na literatura para as mesmas imagens de teste, visualmente, o TV-CSP apresenta resultados muito mais próximos das imagens originais mesmo com uma quantidade baixa de dados, tornando-o, na ausência de ruído, mais eficiente. O comportamento do algoritmo para dados ruidosos não foi tão preciso, pois as imagens de atenuação reconstruídas não identificaram as regiões do objeto de estudo. Esse problema possivelmente é causado pela escolha inadequada das tolerâncias do sinograma.

O maior problema encontrado no algoritmo proposto é que não temos garantia de convergência, pois temos uma projeção do subgradiente em um conjunto côncavo e assim não podemos encontrar um limitante para os tamanhos de passo τ . Seria necessário uma investigação mais profunda para uma escolha de tamanhos de passo que abrangessem uma grande quantidade de problemas. A escolha de tamanhos de passo decrescente se mostrou bastante interessante nesse sentido.

CONCLUSÃO FINAL

Neste trabalho, primeiramente, estudamos as tomografias por transmissão onde os raios possuem uma trajetória retilínea. Essa técnica de tomografia é modelada pela transformada de Radon (TR), que leva a função f definida no plano na função $\mathcal{R}[f]$ definida no espaço das retas no plano, onde seu valor em uma determinada reta é igual a integral de linha sobre a reta em questão. A inversa dessa transformada possui uma fórmula analítica publicada pelo próprio Radon (RADON, 1986) e a implementação tanto da transformada de Radon quanto da sua inversa está disponível em várias linguagens de programação científica.

A fim de introduzir os conceitos e ferramentas da área de pesquisa, focamos na implementação de métodos iterativos de reconstrução do mapa de atenuação de tomografias por transmissão utilizando métodos como o *Expectation Maximization* (EM) (SHEPP; VARDI, 1982), o *Ordered Subset Expectation Maximization* (OS-EM) (HUDSON; LARKIN, 1994) e o *Expectation Maximization and Total Variation* (EM-TV) (YAN; VESE, 2011). As implementações foram feitas em MATLAB e não apresentaram dificuldades, pois a linearidade da transformada de Radon possibilita sua interpretação matricial quando o problema é discretizado.

Estudando outra classe de tomografia, as tomografias por emissão, abordamos duas técnicas: as tomografias por emissão de pósitron (PET, em inglês) e as tomografias por emissão de fóton único (SPECT, em inglês). Nessas técnicas, uma vez que se é conhecido o mapa de atenuação, busca-se reconstruir a imagem de atividade, uma imagem que determina a ocorrência média de um evento biológico no corpo de um paciente. Decidimos estudar mais profundamente a segunda técnica, a SPECT, pois é modelada pela transformada de Radon atenuada, uma versão mais abrangente da transformada de Radon. A transformada de Radon atenuada é linear com relação à imagem que queremos reconstruir, nomeadamente a imagem de atividade, e não-linear com relação ao mapa de atenuação. Atualmente os tomógrafos SPECT precisam do mapa de atenuação como dado de entrada para reconstruir a imagem atividade.

Dando continuidade a pesquisa buscamos implementar os algoritmos EM, OS-EM e

EM-TV para o caso SPECT. Nesse ponto da pesquisa optamos por migrar para a linguagem C/C++ utilizando a biblioteca RAFT (MIQUELES; HELOU, 2007 (acessado em 11 de junho de 2015)) que dispõe do algoritmo do traçado de raios para acelerar de forma excepcional os cálculos utilizados nos algoritmos citados. A RAFT possui funções que calculam a transformada de Radon simples e sua retroprojeção. Assim, a partir de modificações dessas funções foi possível criar as versões para o caso atenuado da transformada.

As reconstruções das imagens de atividade nas simulações dos algoritmos estudados foram feitas para dois tipos de dados: um com quantidade suficiente para que a matriz da transformada de Radon fosse quadrada; e outro com poucos dados, tendo assim, mais variáveis que equações. No primeiro caso, o algoritmo EM conseguiu em 2000 iterações obter uma reconstrução muito próxima da imagem original onde os todos os elementos da imagem de teste, o phantom de Shepp-Logan (SHEPP; VARDI, 1982), são claramente identificados. Contudo, para essa simulação foram necessárias mais de duas horas de execução do algoritmo. A estratégia utilizada para acelerar as iterações do EM foi utilizar o algoritmo OS-EM, que divide os dados obtidos em subconjuntos ordenados de forma que o EM seja aplicado em cada um, onde a imagem de entrada para uma subiteração num subconjunto é a imagem de saída do EM aplicado ao subconjunto anterior. Cada iteração do OS-EM tem aproximadamente a mesma complexidade algorítmica do EM (HUDSON; LARKIN, 1994) e, sendo t o número de subconjuntos, a iteração n do OS-EM equivale, aproximadamente, à iteração $t \times n$ do algoritmo EM. Portanto, fica claro que numa reconstrução SPECT onde o mapa de atenuação é conhecido, o OS-EM é mais indicado por sua rapidez.

O segundo caso que estudamos foi quando temos menos dados que variáveis, ou seja, muitas variáveis são livres. Nessa situação foi possível identificar nas reconstruções dos algoritmos EM e OS-EM as posições e os formatos dos elementos da imagem de teste, contudo, essas soluções foram muito ruidosas devido a pequena quantidade de dados utilizada para resolver o sistema. O algoritmo EM-TV busca, através de uma regularização, encontrar uma solução estável para esse problema, ou seja, onde as regiões não sejam tão ruidosas como no EM. Contudo, a qualidade da reconstrução regularizada do EM-TV depende de uma escolha subjetiva do parâmetro α que dependerá da imagem a ser tratada. Para encontrar um valor eficiente de α nos casos simulados foram testado 1000 valores igualmente espaçados no intervalo $(0, 1)$.

Como o EM-TV tem o algoritmo EM como um de seus passos, a lenta convergência do mesmo foi herdada. Uma alternativa que não foi explorada na pesquisa é a de acelerar as iterações do EM-TV utilizando o algoritmo OS-EM no lugar do EM. Contudo, dentre os algoritmos abordados neste texto, o mais indicado para uma situação com pouca incidência de radiação ou dados é o EM-TV.

Embora a incidência de radiação, ou quantidade de dados, possa ser reduzida nas tomografias SPECT, ainda existem problemas como o custo da aparelhagem, o tempo de utilização e o mal alinhamento das imagens para a reconstrução. Assim, buscamos estudar algoritmos de

reconstrução simultânea somente através de dados SPECT.

Motivados pelos benefícios da técnica de reconstrução simultânea realizamos uma pesquisa bibliográfica sobre o tema a fim de encontrar algoritmos eficientes e buscar melhorá-los. No [Capítulo 3](#) apresentamos alguns desses algoritmos estudados. Houve a tentativa de implementação de somente um deles, o Método de Bronnikov ([BRONNIKOV, 2000](#)), que consiste em encontrar uma imagem de atenuação que produza um sistema consistente a partir dos dados obtidos. Na teoria, o método parece perfeito, mas ao tentarmos implementar utilizando os mesmo dados e imagens teste, não obtivemos tais resultados. Contudo, não temos provas sobre a ineficiência do método, pois ainda é possível que a implementação tenha falhado.

Na busca por algoritmos de reconstrução simultânea, encontramos um algoritmo criado por Yair Censor ([CENSOR et al., 1979](#)) que neste texto é chamado de CSP (*Cyclic Subgradient Projections*) onde não havia ainda sido revisado. Um algoritmo antigo que não garante convergência. Num primeiro momento, tínhamos como objetivo reproduzir os resultados do algoritmo que, até então, não existiam. O maior problema do algoritmo é que um de seus dois parâmetros não garante a convergência do método para o conjunto côncavo e talvez isso tenha afastado qualquer tentativa de reprodução do mesmo. A surpresa foi que os resultados foram melhores do que o esperado tornando-o assim, o algoritmo alvo de estudos desse trabalho.

A mudança de variáveis realizada no CSP inviabiliza reconstruções de imagens de atividade que possuam valores nulos. Utilizamos assim, o conceito de suporte de uma imagem, ou seja, o conjunto de pixels não nulos da mesma. Buscando um suporte idêntico ao da imagem que queremos reconstruir utilizamos os pixels nulos dos sinograma de dados. Através dessa estratégia, somente alguns dos pixels que não estão no suporte da imagem original deixaram de ser detectados como nulos. Isso fez com que as reconstruções melhorassem bastante e ficassem próximas do ideal, mas as regiões que representavam os órgãos da imagem teste ainda eram muito ruidosas e os pixels que não pertenciam ao suporte da imagem original assumiram valores de atividade excessivamente altos.

O algoritmo CSP possui dois parâmetros: τ e ν , que são os tamanhos de passo das restrições côncava e convexa, respectivamente. Um tamanho de passo ν pertencente ao intervalo $(0, 2)$ garante que a próxima subiteração esteja mais próxima conjunto convexo. Contudo, essa garantia não existe para a escolha de τ . Para determinar uma valor aceitável para τ , foi simulado vários testes com imagens diferentes e uma boa escolha está num valor pertencente ao intervalo $(0, 1)$.

Como o CSP é um método de resolução do problema misto de viabilidade côncavo-convexo, estávamos interessados, dentre todas as soluções viáveis, nas soluções livres de ruídos. Assim, criamos um problema de minimização restrita, onde a restrição é exatamente a restrição abordada pelo CSP e a função de minimização é a TV (*Total Variation*). Essa escolha de função de minimização foi dada pelo fato de que a função TV possibilita a reconstrução de imagens que possuam divisões menos suaves entre as regiões do interior do objeto de estudo. Contudo, é

possível que a função de minimização possa ser qualquer uma que busque por reconstruções menos ruidosas ou até mesmo métodos de processamento de imagens.

Com a busca por imagens constantes por partes, os pixels que não pertenciam ao suporte das imagens originais ficaram com os valores negativos, que no caso específico da mudança de variáveis escolhida, indica que seus valores devem ser nulos. Portanto, ao longo da execução do algoritmo criado neste trabalho, o TV-CSP, os pixels extras são removidos e isso faz com que as próximas iterações sejam melhoradas.

Uma das limitações dos algoritmos CSP e TV-CSP é que devemos somente trabalhar com imagens de resolução baixa devido sua complexidade algorítmica. Para encontrarmos uma reconstrução aceitável de uma imagem pequena, 32×32 , temos que esperar cerca de 1 hora de processamento.

A grande descoberta foi a capacidade de reconstrução das imagens a partir de poucos dados do TV-CSP. Numa das simulações foi utilizado menos dados do que seria necessário para uma reconstrução onde é conhecido o mapa de atenuação e mesmo assim o TV-CSP conseguiu reconstruir imagens excelentes. Isso foi um indício de que realmente, os dados SPECT trazem informações sobre o mapa de atenuação. A dificuldade da resolução desse tipo de problema é a não-linearidade com relação a imagem de atenuação que gera muitas complicações, mas ao buscarmos por soluções que tenham determinadas características presentes nas imagens originais diminuimos a quantidade de possíveis soluções.

Para deixar o problema mais próximo da realidade adicionamos ruídos com distribuição de Poisson no sinograma de dados original. As tolerâncias ε_i do TV-CSP não podendo mais assumir valores próximos de zero, foram escolhidas de forma a abordar o desvio padrão da distribuição de Poisson, mas isso deu muita folga para o conjunto de restrições possibilitando soluções com grandes regiões de valores constantes baixos de forma a minimizar a variação total. Então, podemos concluir que o TV-CSP, mesmo que tenha obtido resultados interessantes para o caso sem adição de ruído, ainda não é um algoritmo que possa ser utilizado na prática. Talvez, modificando a estrutura da função de minimização de forma que mais características da imagem original possam ser alvo de busca, o problema com ruído possa ser resolvido.

Vale relembrar que os resultados do CSP apresentados neste texto são as primeiras reproduções do algoritmo criado por Censor e, a partir desses resultados, podemos criar expectativas sobre a criação de métodos de reconstrução simultânea, pois mesmo sem garantias de convergência foi possível reconstruir imagens idênticas as originais eliminando assim a necessidade de uma reconstrução do mapa de atenuação.

REFERÊNCIAS

- ALBER, Y. I.; IUSEM, A. N.; SOLODOV, M. V. On the projected subgradient method for nonsmooth convex optimization in a Hilbert space. **Mathematical Programming**, v. 81, p. 23–25, 1988. Citado na página 64.
- BLATT, D.; HERO, A. O.; GAUCHMAN, H. A convergent incremental gradient method with a constant step size. **SIAM Journal on Optimization**, v. 18, n. 1, p. 29–51, 2007. Citado na página 64.
- BRONNIKOV, A. V. Numerical solution of the identification problem for the attenuated Radon transform. **Inverse Problems**, v. 15, n. 5, p. 1315, 1999. Citado nas páginas 52 e 54.
- _____. Reconstruction of attenuation map using discrete consistency conditions. **IEEE Trans Med Imaging**, v. 19, n. 5, p. 451–462, 2000. Citado nas páginas 51, 52, 54, 55, 56 e 79.
- BROWNE, J.; PIERRO, Á. R. D. A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography. **IEEE Transactions on Medical Imaging**, v. 15, n. 5, p. 687–699, 1996. Citado na página 29.
- CENSOR, Y. Computational acceleration of projection algorithms for the linear best approximation problem. **Linear Algebra and its Applications**, v. 416, n. 1, p. 111–123, 2006. Citado na página 63.
- CENSOR, Y.; GUSTAFSON, D.; LENT, A.; TUY, H. A new approach to the emission computerized tomography problem: simultaneous calculation of attenuation and activity coefficients. **IEEE Transactions on Nuclear Science**, v. 26, p. 2275–2279, 1979. Citado nas páginas 52, 60, 61, 62, 63 e 79.
- CENSOR, Y.; LENT, A. Cyclic subgradient projection. **Mathematical Programming**, v. 24, p. 233–235, 1982. Citado nas páginas 60 e 61.
- DEFRISE, M.; KINAHAN, P. E.; TOWNSEND, D. W.; MICHEL, C.; SIBOMANA, M.; NEWPORT, D. F. Exact and approximate rebinning algorithms for 3-D PET data. **IEEE Transactions on Medical Imaging**, v. 16, n. 2, p. 145–158, 1997. Citado na página 28.
- DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm. **Journal of the Royal Statistical Society**, v. 39, n. 1, p. 1–38, 1977. Citado na página 52.
- DICKEN, V. A new approach toward simultaneous activity and attenuation reconstruction in emission tomography. **Inverse Problems**, v. 15, p. 931–960, 1999. Citado nas páginas 52 e 59.
- ENGL, H. W.; HANKE, M.; NEUBAUER, A. **Regularization of inverse problems**. [S.l.]: Kluwer Academic Publishers, 2000. Citado nas páginas 32 e 33.
- ERDOĞAN, H.; GUALTIER, G.; FESSLER, J. A. An ordered subsets algorithm for transmission tomography. **Physics in Medicine and Biology**, v. 44, n. 11, p. 2835–2851, 1999. Citado na página 29.

- GILES, J. R. **Convex analysis with application in the differentiation of convex functions**. [S.l.]: Boston : Pitman Pub, 1982. Citado nas páginas 62 e 64.
- GOLUB, G. H.; LOAN, C. F. V. **Matrix Computations**. [S.l.]: The Johns Hopkins University Press, 2013. Citado na página 55.
- HAN, G.; LIANG, Z.; YOU, J. A fast ray-tracing technique for TCT and ECT studies. **IEEE Nuclear Science Symposium**, v. 3, p. 1515–1518, 1999. Citado nas páginas 33 e 35.
- HELOU, E. S. **Algoritmos Incrementais com Aplicações em Tomografia Computadorizada**. Tese (Doutorado) — Universidade Estadual de Campinas, São Carlos, SP, 5 2009. elias@icmc.usp.br. Citado nas páginas 15, 23 e 24.
- HELOU, E. S.; PIERRO, Á. R. D. Incremental subgradients for constrained convex optimization: a unified framework and new methods. **SIAM Journal on Optimization**, v. 20, n. 3, p. 1547–1572, 2009. Citado nas páginas 63 e 64.
- HUDSON, H. M.; LARKIN, R. S. Accelerated image reconstruction using ordered subsets of projection data. **IEEE Transactions on Medical Imaging**, v. 13, n. 4, p. 601–609, 1994. Citado nas páginas 29, 41, 42, 45, 53, 77 e 78.
- KAPLAN, M. S.; HAYNOR, D. R.; VIJA, H. A differential attenuation method for simultaneous estimation of SPECT activity and attenuation distributions. **IEEE Transactions on Nuclear Science**, v. 46, p. 535–541, 1999. Citado na página 52.
- KINAHAN, P. E.; ROGERS, J. G. Analytic 3D image reconstruction using all detected events. **IEEE Transactions on Nuclear Science**, v. 36, n. 1, p. 964–968, 1990. Citado na página 28.
- KROL, A.; BOWSHER, J. E.; MANGLOS, S. H.; FEIGLIN, D. H.; TORNAL, M. P.; THOMAS, F. D. An EM algorithm for estimating SPECT emission and transmission parameters from emission data only. **IEEE Transactions on Medical Imaging**, v. 20, n. 3, p. 218–232, 2001. Citado na página 59.
- KROL, A.; ECHERUO, I.; SOLGADO, R. B.; HARDIKAR, A. S.; BOWSHER, J. E.; FEIGLIN, D. H.; THOMAS, F. D.; LIPSON, E.; COMAN, I. L. EM-IntraSPECT algorithm with ordered subsets (OSEMIS) for nonuniform attenuation correction in cardiac imaging. **SPiE Proceedings**, v. 4684, p. 1022–1027, 2002. Citado na página 59.
- KROL, A.; MANGLOS, S. H.; BOWSHER, J. F.; YOUNG, T.; BASSANO, D. A.; THOMAS, F. D. Attenuation compensation in SPECT cardiac imaging using EM-IntraSPECT method. **Journal of Nuclear Medicine**, v. 36, p. 50P, 1995. Citado na página 52.
- LANGE, K.; CARSON, R. EM reconstruction algorithm for emission and transmission tomography. **Journal of Computer Assisted Tomography**, v. 8, p. 306–316, 1984. Citado na página 53.
- MANGLOS, S. H.; YOUNG, T. M. Constrained IntraSPECT reconstruction from SPECT projections. **Proceedings of the 1993 IEEE-MIC**, p. 1605–1609, 1994. Citado na página 52.
- MIQUELES, E. X.; HELOU, E. S. **Reconstruction Algorithms for Tomography - RAFT**. 2007 (acessado em 11 de junho de 2015). <<http://www.milab.ime.unicamp.br/raft/files/intro-txt.html>>. Versão 0.1.0. Citado nas páginas 44, 67 e 78.

NATTERER, F. The identification problem in emission computed tomography. In: **Mathematical Aspects of Computerized Tomography**. [S.l.]: Springer Berlin Heidelberg, 1981, (Lecture Notes in Medical Informatics, v. 8). p. 45–56. Citado nas páginas 52 e 54.

_____. Computerized tomography with unknown sources. **SIAM Journal on Applied Mathematics**, v. 43, p. 1201–1212, 1983. Citado nas páginas 52 e 54.

_____. **The mathematics of computerized tomography**. [S.l.]: Wiley, 1986. Citado na página 22.

_____. Determination of tissue attenuation in emission tomography of optically dense media. **Inverse Problems**, v. 9, p. 731–736, 1993. Citado nas páginas 52 e 54.

NEDÍČ, A.; BERTSEKAS, D. P. Incremental subgradient methods for nondifferentiable optimization. **SIAM Journal on Optimization**, v. 12, n. 1, p. 109–138, 2001. Citado na página 64.

NOVIKOV, R. G. An inversion formula for the attenuated x-ray transformation. **Arkiv för Matematik**, v. 40, n. 1, p. 145–167, 2002. Citado nas páginas 24, 29 e 39.

NUYTS, J.; DUPONT, P.; STROOBANTS, S.; BENNINCK, R.; MORTELMANS, L.; SUE-TENS, P. Simultaneous maximum a posteriori reconstruction of attenuation and activity distributions from emission sinograms. **IEEE Transactions on Medical Imaging**, v. 18, n. 5, p. 393–403, 1999. Citado nas páginas 52, 53 e 59.

PARNIS, J. M.; OLDHAM, K. B. Beyond the Beer–Lambert law: The dependence of absorbance on time in photochemistry. **Journal of Photochemistry and Photobiology A: Chemistry**, v. 267, p. 6–10, 2013. Citado na página 25.

RADON, J. On the determination of functions from their integral values along certain manifolds. **IEEE Transactions on Medical Imaging**, v. 5, n. 4, p. 170–176, 1986. Citado nas páginas 22, 23 e 77.

RUDIN, L.; OSHER, S. Total variation based image restoration with free local constraints. **Proceedings of IEEE International Conference on Image Processing (ICIP)**, I, p. 31–35, 1994. Citado na página 42.

RUDIN, L.; OSHER, S.; FATEMI, E. Nonlinear total variation based noise removal algorithms. **Physica D**, v. 60, p. 259–268, 1992. Citado na página 42.

SHEPP, L.; VARDI, Y. Maximum likelihood reconstruction for emission tomography. **IEEE Transactions on Medical Imaging**, MI-2, p. 113–122, 1982. Citado nas páginas 29, 39, 40, 41, 53, 77 e 78.

SHEPP, L. A.; LOGAN, B. F. The Fourier reconstruction of a head section. **IEEE Transactions on Nuclear Science**, v. 21, n. 3, p. 21–43, 1974. Citado nas páginas 22 e 45.

SOLODOV, M. V.; ZAVRIEV, S. K. Error stability properties of generalized gradient-type algorithms. **Journal of Optimization Theory and Applications**, v. 98, n. 3, p. 663–680, 1998. Citado na página 64.

WELCH, A.; CLACK, R.; NATTERER, F.; GULLBERG, G. T. Toward accurate attenuation correction in SPECT without transmission measurements. **IEEE Transactions in Medical Imaging**, v. 16, p. 532–540, 1997. Citado nas páginas 51, 52 e 54.

YAN, M.; VESE, L. A. Expectation maximization and total variation based model for computed tomography reconstruction from undersampled data. **SPIE Proceedings**, v. 7961, p. 79612X–79612X–8, 2011. Citado nas páginas [43](#), [44](#), [46](#) e [77](#).

YAN, Y.; ZENG, G. L. Attenuation map estimation with SPECT emission data only. **International Journal of Imaging Systems and Technology**, v. 19, n. 3, p. 271–276, 2009. Citado na página [59](#).