

Avaliação das técnicas de filtragem de Kalman e suavização Rauch-Tung-Striebel para a melhoria da qualidade de imagens tomográficas de solos agrícolas

Marcos A. M. Laia
Universidade Federal de São Carlos
marcoslaia@gmail.com

Abstract

This work presents the space variant noise filtering of tomographic projections based on the Kalman filter. For development and filter selection it were evaluated different modalities of the Kalman filter included the use of artificial neural network. Results were analyzed by means of Improvement in Signal to Noise Ratio (ISNR) measurements, which were obtained in a region of interest (ROI) on the resultant images, reconstructed with the use of a backprojection algorithm. In this context the results qualified the extended Kalman filter with neural network (joint estimation) having three layers (multilayer perceptron) as the best option for filter for soil science tomography.

1. Introdução

A qualidade de uma imagem é um dos principais requisitos necessários para sua análise, sendo desejável que o objeto reconstruído esteja bem próximo da amostra que foi ensaiada. O uso de algoritmos desenvolvidos para as áreas de conhecimento humano tem crescido consideravelmente e o processamento de imagens tem melhorado a informação visual para análise e interpretação ou percepção automática por máquinas. Na interpretação humana, temos o uso de imagens de raios-X na área médica, na geologia, nas imagens de satélites, restauração de figuras na arqueologia, na melhoria de imagens físicas, na medicina nuclear, na melhoria de imagens transmitidas por ondas de rádio; enquanto na percepção automática de máquinas temos o reconhecimento automático de faces, caracteres, impressões digitais, visão computacional, controles de robôs por supervisão, processamento automático de imagens de satélites para reconhecimento de queimadas, de mudanças de clima e de formação de tempestades e furacões.

No processamento de imagens, os conceitos buscam: a) o modelamento do funcionamento do olho humano; b) estudo de transformadas de imagens, como a Transformada de Fourier e outras transformadas separáveis de imagens; c) a aplicação filtros que visam recuperar uma imagem e utilização de máscaras para que a imagem processada seja mais aplicável do que a imagem original (real); d) compressão de imagens, para que uma imagem digital seja mais próxima da real com uma redução significativa dos dados; e) segmentação de imagens, para detecção de bordas, de pontos, de linhas e descontinuidade, cujo propósito é divisão da imagem em partes para sua análise; f) reconhecimento e interpretação de imagens, como técnicas de análise baseada em redes neurais, com o propósito de análise e reconhecimento de padrões e de características da imagem [1]. Uma imagem de qualidade permite observar dados que podem ser fundamentais para a aplicação da qual se utiliza o processamento. A qualidade das imagens tomográficas tem extrema importância para a classificação de elementos presentes no solo e análise de porosidade no solo para estudos de técnicas para o aumento da produção agrícola tal como uso na preservação do solo.

Este trabalho tem como objetivo descrever técnicas desenvolvidas para melhorar a qualidade de imagens tomográficas aplicando filtros nos sinais tomográficos antes da reconstrução das imagens.

1.1. Tomografia computadorizada

A aplicação da tomografia computadorizada (TC) na ciência do solo para investigação de propriedades físicas do mesmo é de fundamental importância para o estudo do transporte da água e dos solutos nesse meio. A transmissão direta de raios X ou γ possibilita uma grande contribuição para resolver os diversos problemas da área de solos com resultados obtidos numa escala da ordem de milímetros, enquanto muitas

respostas são esperadas em nível de partículas, macroporos e microporos [2].

Largamente utilizada nas áreas médicas, o uso da TC em ciência do solo foi introduzido por Petrovic [3], Hainsworth e Aylmore [4] e por Crestana [5]. Petrovic demonstrou a possibilidade de usar um tomógrafo computadorizado de raios X para medir a densidade de volumes de solos, enquanto Crestana demonstrou que a TC pode solucionar problemas ligados aos estudos da física da água no solo. A partir desses estudos, originou-se um projeto que envolveu o desenvolvimento de um tomógrafo para a ciência do solo [6].

O uso do computador na tomografia é essencial para a reconstrução da imagem a partir de projeções. Uma abordagem sobre o histórico da tomografia pode ser visto em [7].

A tomografia computadorizada de raios X e γ na escala milimétrica utilizada no domínio da ciência do solo foi introduzida pelo CNPDIA [8], permitindo ver secções transversais do interior de amostras sem que seja necessário cortá-las ou fazer qualquer outro tipo de invasão, utilizando o princípio da atenuação da radiação nos meios materiais. Vários trabalhos foram desenvolvidos para o aprimoramento da visualização das imagens adquiridas e do algoritmo de reconstrução, bem como do hardware desenvolvido [9][10] [11][12].

O processo de tomografia envolve a ação de iluminar um objeto em várias direções proporcionais, após armazenagem, extraindo um conjunto de valores que representam amostras de um conjunto de projeções. As projeções podem ser provenientes de diversas fontes como, por exemplo, dos tradicionais raios X, dos raios γ , da ressonância magnética ou do ultra-som. Com base na intensidade emitida pela fonte de raios X e na intensidade captada por um detector na outra extremidade da reta de propagação, pode-se determinar o peso da atenuação devido ao objeto que se faz presente entre fonte e o detector. Este peso da atenuação é um dado crucial para o processo de reconstrução, o qual possibilita deter um mapa de coeficientes de atenuação linear da secção transversal do objeto. Cada dado armazenado é na realidade uma média de alguns parâmetros inerentes do caminho de propagação do raio de iluminação.

O mapa de coeficientes é representado através de pixels cujos valores são dados pelos chamados números CT (nomenclatura decorrente de Computerized Tomography). Esses números são normalizados em função do coeficiente de atenuação da água. Em outras palavras definem-se os números CT por:

$$\text{Número CT} = \frac{\mu - \mu_{H_2O}}{\mu_{H_2O}} \times 1000 \quad (1)$$

onde μ é o coeficiente de atenuação do corpo analisado. Com esse número é possível a obtenção de um mapa de coeficientes de atenuação, o que permite uma análise mais detalhada do corpo em estudo. Em medicina convencionou-se que para a água o número CT é igual a 0 (zero).

O raio parte de uma fonte e é detectado por um detector. A função de integração do objeto ao longo do raio é uma integral de linha. Cada integral de linha dos raios paralelos forma uma projeção paralela, a qual pode ser tratada com base na teoria de sinais. Um sinal pode, assim, ser considerado como uma função unidimensional, onde x corresponde às distâncias relativas ao tempo e sua amplitude indica a intensidade no ponto. Então, deve-se assumir um valor no intervalo. Sinais reais são contínuos e podem possuir intensidade também em uma escala contínua. Mas, para haver a possibilidade de análise e tratamento desses sinais, é necessário discretizá-los espacialmente e em intensidade. Este processo chama-se digitalização, tendo como consequência um sinal digital. Estudos sobre conceitos como sinal digital, amostragem e vetor de sinal podem ser encontrados em literatura especializadas [13][14]. Nesse processo, pode haver erros nos dados quantificados que, por sua vez, podem provocar falhas que agem diretamente na qualidade das imagens. Esses erros são chamados de ruídos e serão abordados na próxima seção.

1.2. Ruídos em tomografia

A principal fonte de ruído em imagens TC é a variação quântica (*quantum mottle*), definida como a variação estatística espacial e temporal no número de fótons de raios-X absorvidos no detector. Três processos de interação na radiação com a amostra influenciam na atenuação de feixes de raios-X na TC: o efeito fotoelétrico, o efeito Compton e o efeito de produção de pares. Além das questões relacionadas aos efeitos mencionados que são decorrentes da faixa de energia utilizada na fonte, questões que influenciam a medida tomográfica computadorizada. Dentre elas encontra-se a estatística da contagem de fótons. A probabilidade de detecção de fótons em um intervalo de tempo de exposição t pode ser estimada pela função distribuição de probabilidade de Poisson [15],

$$P_{Poisson} = \frac{(\bar{\phi})^n}{\phi! e^n} \quad (2)$$

onde ϕ é o número de fótons e $\bar{\phi}$ é a medida de fotoelétrons emitidos no intervalo de tempo t , conforme a expressão abaixo:

$$\bar{\phi} = \xi M t \quad (3)$$

onde M é a razão média de fótons (fótons/segundo) e ξ é a eficiência quântica da fotomultiplicadora. A incerteza ou ruído é dada pelo desvio padrão

$$\sigma = \sqrt{\xi M t} = \sqrt{\bar{\phi}} \quad (4)$$

Logo, a relação sinal-ruído, SNR, apresentada pelo sinal incidente, é:

$$SNR = \frac{\bar{\phi}}{\sigma} = \sqrt{\bar{\phi}} \quad (5)$$

Desta relação, avalia-se que, para um número pequeno de fótons, o ruído pode ser considerável, porém, com o aumento, o ruído pode vir a ser desprezível. A emissão de elétrons termiônicos no foto-cátodo da multiplicadora causa um aumento nesse ruído. Considerando que o foto-cátodo emite elétrons aleatoriamente em razão da corrente de cátodo incrementado de

$$\sqrt{(\xi M + M)t} \quad (6)$$

tem-se que a nova relação sinal-ruído é dada pela expressão:

$$SNR = \frac{\sqrt{\xi M}}{\sqrt{\xi M + M}} \quad (7)$$

Na visualização de uma imagem tomográfica, há presença de granulosidade, que é significativa para visualização de objetos de baixo contraste. O termo ruído em imagens tomográficas refere-se à variação dos coeficientes de atenuação sobre o valor médio quando uma imagem é obtida de um objeto uniforme [16]. Outros tipos de ruídos presentes nas imagens TC são os erros de arredondamento no programa de reconstrução (ruído do algoritmo), ruído eletrônico e ruído contribuído pelo sistema de visualização. O ruído eletrônico pode ter origem em dispositivos eletrônicos não ideais, tais como resistores e capacitores não puros, terminais de contatos não ideais, corrente de fuga de transistores, efeito Joule, podendo ser também independente do sinal, como interferência externa (elétrica ou até mesmo mecânica) [17][18][19][20].

Os filtros passa-baixa e mediana são soluções para resolver o problema de sinal ou ruído, mas há perda de informações cruciais. Sistemas com diferentes fontes de ruído não apresentam uma solução com o uso desse filtro. Há, portanto, necessidade do uso de filtros mais complexos que podem ser vistos em [21][22][23], que também trazem sua comparação com uma solução do uso de redes neurais (pré-filtragem) com filtro de Kalman discreto.

2. Filtro de Kalman

O processo a ser discutido resolve o problema geral de estimar um estado $x \in \mathcal{R}^n$ de um processo controlado em um tempo discreto que é gerado por uma equação linear de diferenças estocásticas, ou seja:

$$\begin{bmatrix} x_2 \\ \vdots \\ x_k \end{bmatrix} = A \begin{bmatrix} x_1 \\ \vdots \\ x_{k-1} \end{bmatrix} + B \begin{bmatrix} u_1 \\ \vdots \\ u_k \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ n_{k-1} \end{bmatrix} \quad (8)$$

com uma medição $z \in \mathcal{R}^m$ que é

$$\begin{bmatrix} z_2 \\ \vdots \\ z_k \end{bmatrix} = C \begin{bmatrix} x_1 \\ \vdots \\ x_{k-1} \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_{k-1} \end{bmatrix} \quad (9)$$

As variáveis aleatórias n_k e v_k representam o ruído do processo e da medida (respectivamente). Assume que elas sejam independentes (entre si) e com a distribuição de probabilidade normal:

$$P(n) \sim N(0, Q) \quad (10)$$

$$P(v) \sim N(0, R) \quad (11)$$

Na prática, as matrizes da covariância do ruído do processo Q e da covariância do ruído da medição R podem mudar com cada passo de tempo ou medida, contudo, aqui se assume que elas são constantes.

A matriz $A_{n \times n}$ na equação diferencial (8) relata o estado do tempo anterior do passo $(k-1)$ para o estado atual do passo (k) , na ausência de qualquer função direcionada ou ruído do processo. De acordo com Welch e Bishop [24], na prática $A_{n \times n}$ pode mudar a cada estado, mas é assumido-se como constante. A matriz $B_{n \times 1}$ relata a entrada opcional do controle $u \in \mathcal{R}^1$ para o estado x . A matriz $C_{m \times n}$ é a equação de correção (9) que relata o estado para a medida z_k . Na prática, $C_{m \times n}$ pode mudar, mas também é assumido como uma constante.

O objeto de interesse deste trabalho é entender o uso do filtro de Kalman e o uso do algoritmo para separar um ruído de um sinal. Será utilizada uma visão de alto nível e depois será especificada cada equação do algoritmo.

O filtro de Kalman estima um processo utilizando a forma de realimentação (*feedback*) em um tempo anterior e então obtém realimentação na forma de medidas (ruidosas) no tempo atual. Assim, as equações podem ser divididas em dois momentos: as equações para a atualização do tempo - que são responsáveis pela projeção a frente (no tempo) das estimativas do estado atual e da covariância de erros para obter as estimativas *a priori* para o passo do próximo tempo - e as equações para a atualização das medidas - que são responsáveis pela realimentação, por exemplo, incorporação de uma nova medida na estimativa *a priori* para obter uma estimativa *a posteriori*.

As equações para atualização do tempo podem ser também pensadas como equações preditoras, enquanto as equações para atualização das medidas podem ser tratadas como equações corretoras. Equações de atualização do tempo (preditor) podem ser dadas por:

$$\begin{bmatrix} \hat{x}_2^- \\ \vdots \\ \hat{x}_k^- \end{bmatrix} = A \begin{bmatrix} x_1 \\ \vdots \\ x_{k-1} \end{bmatrix} + B \begin{bmatrix} u_1 \\ \vdots \\ u_k \end{bmatrix} \quad (10)$$

$$\begin{bmatrix} P_2^- \\ \vdots \\ P_k^- \end{bmatrix} = A \begin{bmatrix} P_1 \\ \vdots \\ P_{k-1} \end{bmatrix} A^T + Q. \quad (11)$$

Nota-se que estas equações projetam as estimativas do tempo e da covariância através do tempo no passo $k-1$ para o passo k .

Equações de atualização das medidas (corretor) podem ser dadas por:

$$\begin{bmatrix} K_2 \\ \vdots \\ K_k \end{bmatrix} = \begin{bmatrix} P_1 \\ \vdots \\ P_{k-1} \end{bmatrix} H^T (H \begin{bmatrix} P_2^- \\ \vdots \\ P_k^- \end{bmatrix} H^T + R)^{-1} \quad (12)$$

$$\begin{bmatrix} \hat{x}_2 \\ \vdots \\ \hat{x}_k \end{bmatrix} = \begin{bmatrix} \hat{x}_2^- \\ \vdots \\ \hat{x}_k^- \end{bmatrix} + \begin{bmatrix} K_1 \\ \vdots \\ K_k \end{bmatrix} \left(\begin{bmatrix} z_2 \\ \vdots \\ z_k \end{bmatrix} - H \begin{bmatrix} \hat{x}_2^- \\ \vdots \\ \hat{x}_k^- \end{bmatrix} \right) \quad (13)$$

$$\begin{bmatrix} \hat{p}_2 \\ \vdots \\ \hat{p}_k \end{bmatrix} = (I - \begin{bmatrix} K_1 \\ \vdots \\ K_k \end{bmatrix} H) \begin{bmatrix} P_2^- \\ \vdots \\ P_k^- \end{bmatrix} \quad (14)$$

A primeira tarefa durante a atualização da medida é computar o ganho de Kalman, K_k . O próximo passo é atualizar para medir o processo para obter z_k , e então gerar um estado estimado *a posteriori* por incorporação à medida na equação do meio. O passo final é obter uma estimada covariância de erro via a equação 14.

2.1. Filtro de Kalman estendido

Uma solução para sistemas não lineares é o filtro de Kalman estendido (FKE) [23]. Analisando a função de predição do filtro de Kalman, equação 10, é possível observar que o filtro se comporta de forma linear, por ser um sistema linear. Aplicando uma função não-linear, pode-se obter uma predição ótima para os próximos estados. Este algoritmo aplica o filtro de Kalman para sistemas não lineares simplesmente linearizando todos os modelos não-lineares, assim as equações do filtro tradicional podem ser aplicadas. O sistema não linear pode ser reescrito na forma:

$$\begin{bmatrix} x_2 \\ \vdots \\ x_k \end{bmatrix} = F \left(\begin{bmatrix} x_1 \\ \vdots \\ x_{k-1} \end{bmatrix}, \begin{bmatrix} u_1 \\ \vdots \\ u_k \end{bmatrix}, \begin{bmatrix} n_1 \\ \vdots \\ n_{k-1} \end{bmatrix} \right) \quad (17)$$

$$\begin{bmatrix} z_2 \\ \vdots \\ z_k \end{bmatrix} = H \left(\begin{bmatrix} x_1 \\ \vdots \\ x_{k-1} \end{bmatrix}, \begin{bmatrix} v_1 \\ \vdots \\ v_{k-1} \end{bmatrix} \right) \quad (18)$$

Para a propagação das variâncias, deve-se conhecer as matrizes jacobianas ou hessianas das funções de transição e observação de estado.

Por não se conhecer a função de transferência ou mesmo aumentar a ordem do filtro, procurou-se utilizar um novo método de estimação onde há a possibilidade de estimar novos estados-parâmetros a partir das cadeias ocultas do modelo markoviano.

2.2. Filtro de Kalman estimação conjunta

O principal problema é quando se deve identificar as funções necessárias para estimação dos estados e dos parâmetros. As equações para predição passam a ser descritas como:

$$x_k = F(x_{k-1}, w_k) \quad (19)$$

$$z_k = H(x_k, v_k) \quad (20)$$

A estimação de parâmetros envolve a determinação de um mapeamento não linear

$$y_k = G(x_k, W) \quad (21)$$

onde x_k é a entrada, W é o peso e y_k é a saída e o mapeamento não linear G é parametrizado pelo vetor W . O mapeamento não linear pode ser feito por uma rede neural artificial, onde W é o peso [26]. O aprendizado corresponde à estimação dos parâmetros de W . O treinamento pode ser feito com pares de amostras, consistindo em uma entrada conhecida e saída desejada (x_k, d_k) . O erro de máquinas é definido pela equação 22. O objetivo do aprendizado é minimizar o erro esperado.

$$e_k = d_k - G(x_k, W) \quad (22)$$

Usando um modelo para treinamento da rede como feito neste trabalho, o FKE pode ser usado para estimar os parâmetros escrevendo uma nova representação estado-espacial:

$$w_k = w_{k-1} + v_k \quad (23)$$

$$y_k = G(x_k, w_k) + e_k \quad (24)$$

onde os parâmetros w_k correspondem a um processo estacionário com uma matriz identidade de transição de estado, governado por um ruído processual v_k (a escolha da variância determina a performance da filtragem). A saída y_k corresponde a uma observação não linear em w_k . O FKE pode ser aplicado diretamente como uma eficiente técnica de segunda ordem para correção dos parâmetros.

Como o problema em questão consiste em trabalhar sobre uma entrada x_k não observada e requer uma união entre a estimação dos estados e parâmetros, deve-se considerar um problema de estimação dual, considerando uma dinâmica discreto-temporal do sistema não linear:

$$x_{k+1} = F(x_k, v_k, W) \quad (25)$$

$$z_k = H(x_k, n_k) \quad (26)$$

onde ambos os estados do sistema x_k e o conjunto dos parâmetros W para o sistema dinâmico devem ser estimados somente do sinal ruidoso y_k .

O sistema dinâmico pode ser entendido como uma rede neural, onde W é o conjunto de pesos, e a função F , como uma função de rede neural que utiliza uma entrada x_k . Assim, aplicando essas equações ao FKE, tem-se uma nova função para estimação e para observação. Uma abordagem sobre redes neurais pode ser vista em [25][26][27].

2.3. Algoritmo de suavização Rauch-Tung-Striebel

A solução para o problema de predição linear segue uma maneira direta da teoria básica da propagação da

covariância em relação à média da variável. Nesta seção, será considerado o problema da suavização ótima.

Supondo que exista um conjunto de dados sobre o intervalo do tempo $0 < k \leq N$. A suavização é uma operação offline que envolve a estimação do estado x_k para $0 < k \leq N$, usando todos os dados disponíveis, tanto do passado como do futuro. Assume-se que o tempo final N é fixo para determinar a estimativa do estado ótimo x_k para $0 < k \leq N$, e o dado futuro y_j definida por $k < j \leq N$. A estimação pertinente ao dado passado é conhecida como a teoria da filtragem adiante. Para estar de acordo com a emissão de uma estimação pertinente para o estado futuro, usa-se uma retro-filtragem, que começa no final do tempo N e executa para trás. Deixando como \hat{x}_k^f e \hat{x}_k^b denotam-se as estimativas dos estados obtidos das recursões para frente e para trás, respectivamente. Dadas estas duas estimativas, o próximo passo a ser considerado é como combiná-los em uma estimativa suavizada total \hat{x}_k , que conta com os dados sobre o intervalo de tempo inteiro. O \hat{x}_k é o símbolo utilizado para a estimativa suavizada e não pode ser confundido com o estimado filtrado (*a posteriori*).

Escrevendo a equação de processo como uma recursão para o decrescente k , temos

$$x_k = F_{k+1}^{-1} x_{k+1} - F_{k+1,k}^{-1} w_k \quad (27)$$

onde F_{k+1}^{-1} é o inverso para a matriz de transição $F_{k+1,k}$. O raciocínio para a retro filtragem começa no tempo final N , o que é demonstrado na Fig. 1 (a), e a ser contrastada com a filtragem adiante na Fig. 1 (b). Note-se que a estimativa *a priori* \hat{x}_k^{b-} e a estimativa *a posteriori* \hat{x}_k^b para a retro filtragem ocorrem para a direita e para a esquerda do tempo k , respectivamente, na Fig. 1 (a). Essa situação é o oposto exato do que ocorre no caso da filtragem adiante mostrada na Fig. 1 (b).

Para simplificar a apresentação, nós introduzimos as duas definições:

$$S_k = [P_k^b]^{-1} \quad (28)$$

$$S_k^- = [P_k^{b-}]^{-1} \quad (29)$$

e as duas variáveis intermediárias

$$\hat{z}_k = [P_k^b]^{-1} \hat{x}_k^b = S_k \hat{x}_k^b \quad (30)$$

$$\hat{z}_k^- = [P_k^{b-}]^{-1} \hat{x}_k^{b-} = S_k^- \hat{x}_k^{b-}. \quad (31)$$

Então, construindo o raciocínio da figura, tem-se a derivação das seguintes atualizações para o retro filtro:

Atualização da medida:

$$S_k = S_k^- + H_k R_k^{-1} H_k \quad (32)$$

$$z_k = \hat{z}_k^- + H_k^T R_k^{-1} y_k \quad (33)$$

onde y_k é o observável definido pela equação de medição, H_k é a matriz de medição e R_k^{-1} é a inversa da matriz de covariância do ruído de medição v_k .

Atualização do tempo:

$$G_k^b = S_{k+1} [S_{k+1} + Q_k^{-1}]^{-1} \quad (34)$$

$$S_k^- = F_{k+1,k}^T (I - G_k^b) S_{k+1} F_{k+1,k} \quad (35)$$

$$\hat{z}_k^- = F_{k+1,k}^T (I - G_k^b) \hat{z}_{k+1} \quad (36)$$

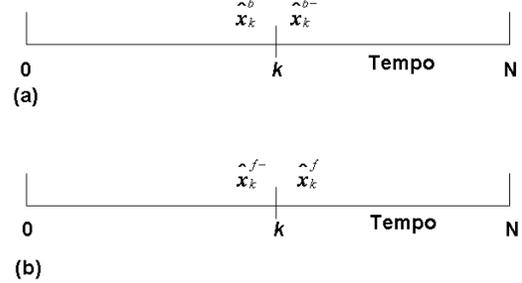


Fig. 1. Ilustrações das atualizações de tempo do suavizador para (a) retro-filtragem e (b) filtragem adiante.

onde G_k^b é o ganho de Kalman para retro-filtragem e Q_k^{-1} é o inverso da matriz de covariância do ruído do processo w_k . O retro-filtro definido pelas equações de medição e atualização do tempo é a formulação de informação do filtro de Kalman. O filtro de informação se distingue do filtro básico de Kalman por propagar a inversa da matriz de covariância de erro melhor do que a matriz de covariância de erro normal.

Sejam os dados observados no intervalo $0 < k \leq N$ para o fixo N e suponha-se que se obtêm as duas seguintes estimativas:

→ A estimativa *a posteriori* adiante \hat{x}_k^f pela operação do filtro de Kalman nos dados y_j para $0 < j \leq k$.

→ A estimativa *a priori* para trás \hat{x}_k^{b-} pela operação do filtro de informação nos dados y_j para $k < j \leq N$.

Com estas duas estimativas e suas respectivas matrizes de covariância de erro conhecidas, o próximo passo de interesse é como determinar a estimativa suavizada de \hat{x}_k e sua matriz de covariância de erro, que incorpora os dados totais sobre o tempo inteiro no intervalo $0 < k \leq N$.

Reconhecendo que o ruído do processo w_k e o ruído de medição v_k são independentes, formula-se a matriz de covariância de erro da estimativa *a posteriori* suavizada \hat{x}_k como se segue:

$$P_k = [[P_k^f]^{-1} + [P_k^{b-}]^{-1}]^{-1} \quad (37)$$

$$= [[P_k^f]^{-1} + S_k^-]^{-1} \quad (38)$$

Mais adiante, invoca-se o teorema da inversão de matriz, que pode ser mostrado a seguir. Deixando A e B serem duas matrizes positiva-definidas relacionadas por

$$A = B^{-1} + CD^{-1}C^T, \quad (39)$$

onde D é outra matriz positiva-definida e C é uma matriz com dimensões compatíveis. O teorema da matriz de inversão demonstra que a inversa da matriz A pode ser expressa como:

$$A^{-1} = B - BC[D + C^T BC]^{-1} C^T B \quad (40)$$

Para o problema em que se trabalha, atribuem-se:

$$A = P_k^{-1} \quad (41)$$

$$B = P_k^f \quad (42)$$

$$C = I \quad (43)$$

$$D = [S_k^-]^{-1} \quad (44)$$

onde I é a matriz identidade. Então, aplicando o teorema de inversão de matriz na equação da matriz de covariância de erro, obtém-se:

$$P_k = P_k^f - P_k^f [P_k^{b-} + P_k^f]^{-1} P_k^f \quad (45)$$

$$= P_k^f - P_k^f S_k^- [I + P_k^f S_k^-]^{-1} P_k^f. \quad (46)$$

Desta equação, encontra-se que a matriz de covariância de erro *a posteriori* suavizada P_k é menor ou igual do que a matriz de covariância de erro *a posteriori* P_k^f produzida pelo filtro de Kalman, o que é naturalmente devido ao fato de que esta suavização utiliza informação adicional contida nos dados futuros. Este ponto é apresentado na Fig. 2, que descreve as variações de P_k , P_k^f e P_k^{b-} com k para uma situação unidimensional.

$$\hat{x}_k = P_k ([P_k^f]^{-1} \hat{x}_k^f + [P_k^{b-}]^{-1} \hat{x}_k^{b-}) \quad (47)$$

Usando as equações anteriores para simplificar, obtém-se:

$$\hat{x}_k = \hat{x}_k^f + (P_k z_k^- - G_k \hat{x}_k^f), \quad (48)$$

onde o ganho suavizador é definido por

$$G_k = P_k^f S_k^- [I + P_k^f S_k^-]^{-1}, \quad (49)$$

que não pode ser confundido com o ganho de Kalman.

O suavizador ótimo já derivado consiste de três componentes:

- Um filtro adiante na forma de um filtro de Kalman;
- Um retro-filtro na forma de um filtro de informação;
- Um suavizador separado, que combina resultados encorpados nos filtros adiante e retrocesso.

O suavizador *Rauch-Tung-Striebel* (RTS), contudo, é mais eficiente do que um suavizador de três partes que incorpora o retro-filtro e um suavizador separado em uma entidade simples. Especificamente, a atualização da medição do suavizador RTS é definida por:

$$P_k = P_k^f - A_k (P_{k+1}^{f-} - P_{k+1}) A_k^T \quad (50)$$

onde A_k é a nova matriz de ganho:

$$A_k = P_k^f F_{k+1,k}^T [P_{k+1}^{f-}]^{-1} - 1. \quad (51)$$

A atualização de tempo correspondente é definida por

$$\hat{x}_k = \hat{x}_k^f + A_k (\hat{x}_{k+1} - \hat{x}_{k+1}^{f-}). \quad (52)$$

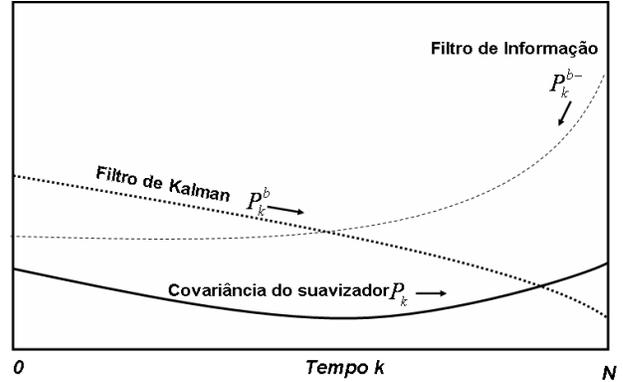


Fig. 2. Ilustração da covariância de erro para a filtragem adiante, retro filtragem e suavizante.

O suavizador RTS procede como a seguir:

1. O filtro de Kalman é aplicado para os dados observáveis em uma maneira adiante, que é $k = 0, 1, 2, \dots$, de acordo com a teoria básica.
2. O suavizador recursivo é aplicado para os dados observados de uma maneira retrocessa, que é $k = N - 1, N - 2, \dots$, de acordo com as equações do RTS.
3. As condições iniciais são definidas por:

$$P_N = P_N^F \quad (53)$$

$$\hat{x}_k = \hat{x}_k^f. \quad (54)$$

3. Modelagem

Com o conhecimento da função não-linear do processo e um filtro de Kalman que suporte funções lineares, se obtém uma melhora significativa no sinal. Como a transformada de Anscombe está presente na equação do processo, não há filtragem do ruído *Poisson*, além de o processo poder ser afetado por outros tipos de ruídos.

Uma solução seria utilizar uma rede neural para promover um melhor mapeamento da função de processo, diminuindo o ruído presente nas projeções. Para uma estimação dos pesos da rede neural em conjunto com as estimações dos estados, pode-se utilizar duas modalidades de filtragem: a estimação dual e a estimação em conjunto. Estas modalidades de filtragem determinam que os pesos iniciais sejam conhecidos, onde o próximo estado é obtido em um

mapeamento linear com o estado anterior. Assim, temos:

$$x_{k+1} = f(x_k, W_k, v_k) \quad (55)$$

Então, utiliza-se um filtro de Kalman para estimar os estados e os pesos. Esta filtragem permite a aplicação em um sistema onde a dinâmica de estado é desconhecida ou caótica (não-determinística).

A Fig. 3 mostra um diagrama de bloco com a implementação do filtro.



Fig. 3. Diagrama de bloco para a filtragem das projeções tomográficas.

Podem-se utilizar as duas modalidades do filtro de Kalman para sistemas não-lineares para compor essa estimação dual. Apesar de ser uma forma mais completa de se estimar, ainda é passível de erros devido ao fato de a estimação do sinal observado ser aproximada.

Uma forma simplificada desse filtro é utilizar a estimação conjunta (uma variação do filtro que combina o filtro de estados aumentados com estimação dual), em que há a possibilidade de estimar, de forma simultânea, os parâmetros necessários para a estimação dos estados do filtro, obtendo em um único filtro a função de estimação e os parâmetros, garantindo uma melhor eficiência no filtro, além de uma implementação simplificada do filtro:

$$\begin{bmatrix} P_{\theta_i}[n] \\ W \end{bmatrix} = \begin{bmatrix} f(P_{\theta_i}[n-1], W, v_k) \\ I * W \end{bmatrix} \quad (56)$$

$$P_{\theta_i}[n] = h(P_{\theta_i}[n], n_k) \quad (57)$$

onde W são os pesos para a função f que pode ser descrita como uma rede neural e I uma matriz identidade.

Um corpo de prova ou phantom é um objeto confeccionado especialmente para caracterização e calibração do tomógrafo ou outros experimentos. Assim, para a caracterização dos resultados da pesquisa, um phantom homogêneo cilíndrico foi construído com um material composto de Nylon (Fig. 4).

O phantom heterogêneo foi construído com uma base de plexiglass em forma cilíndrica, com apenas quatro comprimentos cilíndricos para inserção de materiais.

Os *phantoms* têm 60 mm de diâmetro, sendo que as amostras de cálcio e água possuem 19 mm de diâmetro e as amostras de fósforo e alumínio, 22 mm de diâmetro, como ilustrado na Fig. 5. As tomografias destes corpos de prova foram imageadas utilizando o minitomógrafo de primeira geração.

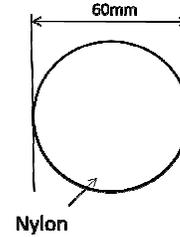


Fig. 4. Diagrama para construção do phantom homogêneo.

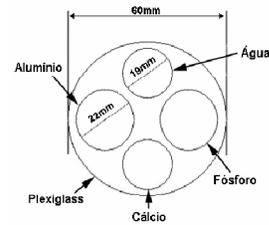


Fig. 5. Diagrama para construção do phantom heterogêneo.

4. Resultados obtidos

Os resultados obtidos com a aplicação do filtro para os *phantoms* podem ser vistos nas Fig. 6 e Fig. 7. Utilizou-se neste trabalho uma rede neural do tipo perceptron multicamadas com duas camadas, sendo a camada de entrada com 1 neurônio, 10 neurônios na camada oculta com a função de transferência ou ativação sigmoideal e um neurônio na camada de saída com a função de transferência linear.

Na Fig. 6 foram utilizados os *phantoms* de baixa resolução (sinal de baixa amostragem). Em (a) e (b), temos os *phantoms* originais, ou seja, sem a filtragem dos sinais tomográficos antes da sua reconstrução, com uma amostra homogênea e outra heterogênea.

Em seguida, utilizou-se o filtro de Kalman discreto com as variâncias Q e R com valores de 0,5 e 1, em (c) e (d). Em (c) é possível perceber uma melhoria devido ao fato de a distribuição dos valores dos pixels ser mais uniforme. Em (d), além da melhoria, é perceptível um pouco de distorção na parte inferior das amostras.

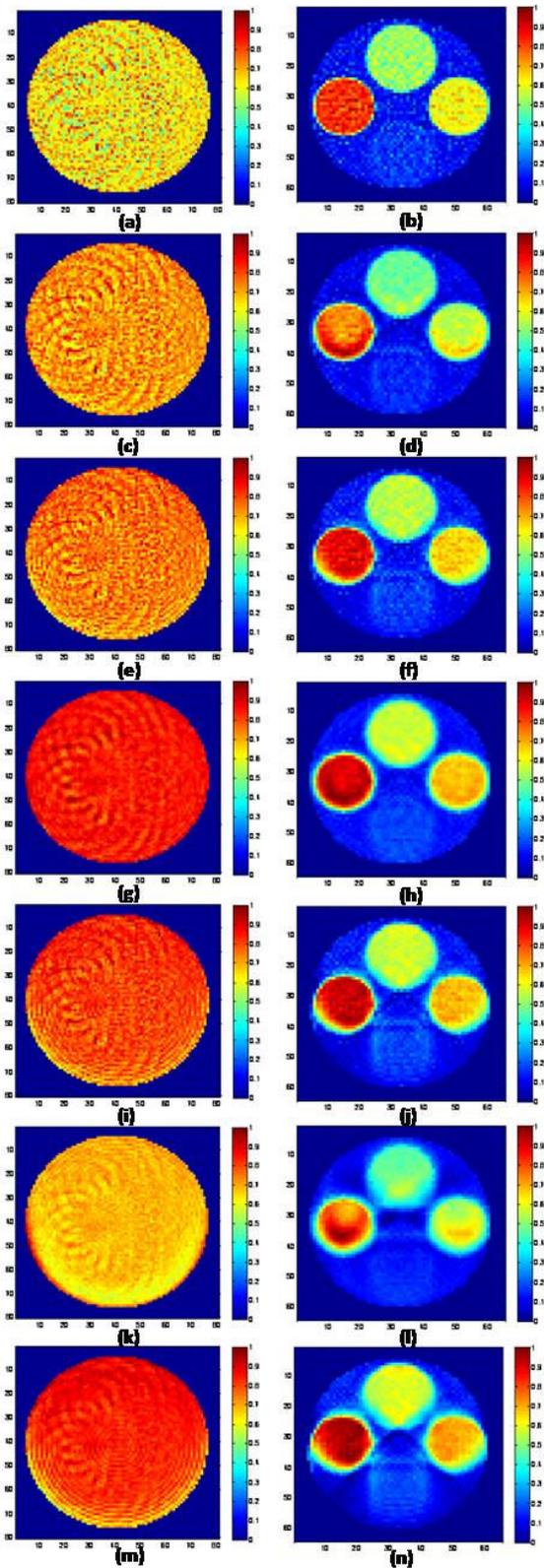


Fig. 6. *Phantoms* utilizados para a calibração do tomógrafo.

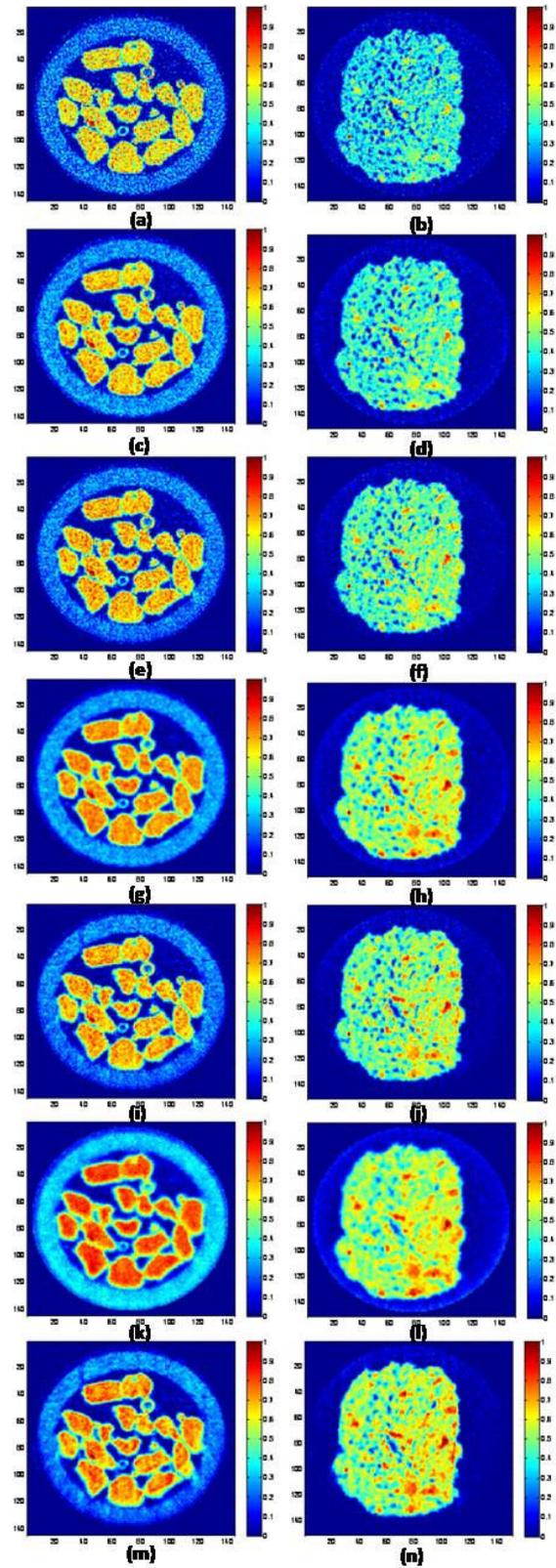


Fig. 7. *Phantoms* com amostras reais de solo.

Essa distorção é devida à natureza do filtro de ter uma filtragem on-line, conforme a filtragem do sinal que tende a um valor ótimo do ganho de Kalman. O uso da rede neural no filtro de Kalman permitiu que apenas o ruído do sistema fosse filtrado, ou seja, o ruído branco.

Esta modalidade do filtro permite uma melhor convergência, ou seja, os valores tendem a chegar numa solução ótima de forma mais rápida do que na modalidade padrão. A pequena distorção encontrada no fundo da imagem (e) e em cada amostra (f) é função do uso de vários neurônios na camada oculta, assim, há uma perda de qualidade no início da filtragem de cada sinal.

O algoritmo de suavização RTS após aplicação do filtro na modalidade discreta permite que, no final da filtragem, os valores ótimos sejam reaplicados ao longo do sinal a partir do final até o início. Na imagem (g) e, principalmente, na imagem (h), é possível uma melhor distribuição das imagens sem distorção, pois os sinais foram filtrados de forma igual. No caso da aplicação do algoritmo com o filtro de Kalman com redes neurais, houve uma uniformização, mas com pouca melhoria em relação ao algoritmo em sua forma pura, como se pode observar nas imagens (i) e (j).

Como segundo experimento, resolveu-se aumentar os valores das variáveis Q e R para 0,1 e 2, respectivamente, e aplicar os filtros em conjunto com o suavizador RTS. O objetivo é a diminuir a incerteza do ruído do processo e aumentar a incerteza do ruído de observação, com o objetivo de produzir uma melhor filtragem. Como o algoritmo de Kalman discreto acaba filtrando o ruído Poisson, é perceptível uma pequena distorção do sinal nas amostras heterogêneas (k) e uma menor variância na imagem homogênea (l). Já as imagens resultantes com o filtro de Kalman com redes neurais ainda apresentaram pequenas distorções (m) e (n), mas sem um aumento ou presença de artefatos. Para o uso em aplicações reais, utilizaram-se *phantoms* obtidos a partir de tomografias de solo como areia e um torrão de terra. Essas imagens servem para fazer um comparativo com o sistema real. Também foram aplicadas as mesmas variâncias Q e R (0,5 e 1) nestes *phantoms*.

Nas imagens da Fig. 7, é possível perceber que as amostras possuem uma qualidade maior do que as anteriores, com pequenos detalhes visíveis, o que ocorre por terem uma amostragem maior. O objetivo é manter esses detalhes e amenizar o ruído nessas imagens.

Nas imagens (a) e (b), se encontram as duas amostras originais, de areia e o torrão, respectivamente. A imagem (a) consiste em vários grãos de areia em um tubo enquanto a imagem (b), de um torrão para análise de porosidade.

Em (c) e (d) foi aplicado o filtro de Kalman discreto e percebe-se uma melhoria na qualidade da imagem de ambas as amostras em relação a imagem original, mas também uma perda de detalhamento na imagem do borrão, pois o filtro não consegue definir o que é ruído ou parte da amostra, já que o sinal tende a ser caótico e, apenas trabalhando com o valor anterior do sinal, não permite uma abordagem como um todo.

Em (e) e (f) foi aplicado o filtro de Kalman com redes neurais. Os resultados foram pouco melhores do que o filtro anterior, havendo uma melhor filtragem e mantendo detalhes.

Em (g) e (h), foram aplicados o filtro de Kalman discreto e o suavizador RTS. Os grãos de areia estão com uma variância menor no interior, mas na imagem do torrão alguns detalhes foram perdidos.

Nas imagens (i) e (j) foram aplicados o filtro de Kalman com redes neurais e o suavizador RTS. É possível perceber que o apenas o ruído gaussiano da equação de observação (do detector) foi novamente filtrado, o que permitiu ainda manter detalhes dos valores dos poros (por serem altos) na imagem do torrão e presença de ruído nos grãos de areia.

Como nos *phantoms* de calibração, foram utilizados novos valores nas variâncias de Q (0,1) e R (2) e o uso do suavizador RTS para gerar novas amostras filtradas. Em (k), com o uso do filtro de Kalman discreto, todos os tipos de sinais são filtrados, tratados como uma distribuição gaussiana, pois o valor da incerteza da observação e a confiança na estimação do sinal foram ampliados e apresentaram a melhor imagem do experimento, mantendo os pequenos detalhes. Em (l) houve uma diminuição dos números de poros junto com os ruídos. Em (m) é possível perceber a presença do ruído Poisson e em (n) a presença da grande maioria dos poros do torrão. Estes últimos filtros utilizaram o filtro de Kalman com redes neurais.

5. Conclusões

A pré-filtragem dos sinais tomográficos é uma técnica que permite a filtragem dos sinais tomográficos antes de sua reconstrução.

O uso do filtro de Kalman discreto permitiu que os sinais fossem pré-filtrados, apresentando uma redução considerável de ruídos. Em sinais tomográficos cujos sinais não se apresentam de maneira não uniforme, o resultado não foi tão preciso, pois apenas a informação do momento anterior não era capaz de processar uma estimativa mais próxima ao valor ideal. A filtragem também apresenta um aspecto importante quanto à filtragem de ruídos Poisson. O filtro tende a tratar o ruído do emissor de fótons do tomógrafo como se fosse do receptor, assim tratando como um ruído branco.

Esse tipo de filtragem nem sempre pode ser a melhor escolha por essa restrição do filtro de não poder tratar o ruído com uma distribuição de Poisson.

O uso do filtro de Kalman estendido com redes neurais permite uma memorização do comportamento anterior do sinal e avalia através desse procedimento qual a melhor estimativa dos estados. É um algoritmo que apresenta uma complexidade maior do que o filtro de Kalman discreto por envolver cálculo de jacobianas para linearização da rede neural e estimativa de parâmetros, fazendo com que o uso do algoritmo seja limitado pela demanda de memória e processamento.

O uso do algoritmo de suavização apresentou melhorias significativas nos filtros pela otimização das antigas amostras após a filtragem. A combinação das duas técnicas permitiu obter imagens de qualidade após a reconstrução dos sinais.

O aumento da confiabilidade no processo e o aumento da incerteza na observação permitiram que houvesse uma melhor filtragem nas projeções do experimento nas projeções, mas esse aumento e uma distância maior entre esses valores podem fazer com que haja distorções (erros na estimativa do próximo valor) e suavização dos sinais, levando a perda de detalhes.

Referências

- [1] Laboratório de Robótica da Universidade de São Paulo, <http://robot.lac.usp.br>, acessado em 27/04/2004.
- [2] S. Crestana, "Técnicas recentes de determinação de características do solo." In: Reunião Brasileira de Manejo e Conservação do Solo e da Água, 10, 1994, Resumos... Florianópolis: Sociedade Brasileira de Ciência do Solo, p86-97, 1994.
- [3] A. M. Petrovic, J.E. Siebert, P.E. Rieke, "Soil bulk analysis in three-dimensions by computed tomographic scanning", *Soil Sci. Soc. Am. J.*, n.46, p.445-450, 1982.
- [4] J. M. Hainsworth, L. A. G. Aylmore, "The use of the computed-assisted tomography to determine spatial distribution of soil water content", *Aust. Journal Soil Res.* N. 21, p.1435-1443, 1983.
- [5] S. Crestana, "A Tomografia Computadorizada com um novo método para estudos da física da água no solo", São Carlos, USP, Tese de Doutorado, 140 páginas, 1985.
- [6] P. E. Cruvinel, "Minitomógrafo de raios X e raios-gama computadorizado para aplicações multidisciplinares". Campinas, UNICAMP. Tese de Doutorado, 329p, 1987.
- [7] M. A. M. Laia, "Filtragem de projeções tomográficas do solo utilizando Kalman e Redes Neurais", São Carlos, UFSCAR, Tese de mestrado, 2007.
- [8] P. E. Cruvinel et al, "X and γ -ray computerized minitomograph scanner for soil science", *IEEE Transactions on Instrumentation and Measurement*, V.39, N.5, p.745-750, October, 1990.
- [9] Y. R. Venturini, "Análise quantitativa da qualidade de imagens digitais com o uso de espectro de Wiener". São Carlos, UFSCar, Tese de Mestrado, 89 páginas, 1995.
- [10] E. R. Minatel, "Desenvolvimento de Algoritmo para Reconstrução e Visualização Tridimensional de Imagens Tomográficas com uso de Técnicas Freqüenciais e Wavelets", São Carlos, UFSCar, Tese de Mestrado, 148 páginas, 1997.
- [11] L. F. Granato, "Algoritmo adaptativo para a melhoria em imagens tomográficas obtidas em múltiplas energias", São Carlos, UFSCar, Tese de mestrado, 135 páginas, 1998.
- [12] N. D. A. Mascarenhas, C. A. N. Santos, P. E. Cruvinel, "Transmission Tomography under poisson noise using the Anscombe Transformation and Wiener Filtering of the projections". *Nuclear Instruments And Methods In Physics Research Section A*, Estados Unidos, v. 423, p. 265-271, 1999.
- [13] E. O. Brigham, "The Fast Fourier Transform and Its Applications", Prentice Hall, New Jersey, 1998.
- [14] E. Deremack, D.G. Crowe, "Optical Radiation Detectors", -John Wiley & Sons, Inc, 300p, 1984.
- [15] W. R. Hender, "The physical principals of computed tomography", EUA, 1983.
- [16] A. D. Ziel, "Noise in measurements"- John Wiley & sons, inc.; 288p, 1976.
- [17] A. J. Duerinckx, A. Macocski, "Polychromatic Streak Artifacts in Computed Tomography Images", *J. Comput. Assist. Tomogr.*, 2,481, 1978.
- [18] P. M. Joseph, R.D. Spital, "A method for correction bone-induced artifacts of CT scanners", 1978.
- [19] G. S. Ibbott, "Radiation therapy treatment planning and the distortion of CT images", *Med. Phys.* 7:261, 1980.
- [20] L. F. Granato, "Algoritmo adaptativo para a melhoria em imagens tomográficas obtidas em múltiplas energias", São Carlos, UFSCar, Tese de mestrado, 135 páginas, 1998.
- [21] M. A. M. Laia, P. E. Cruvinel, "Filtragem de projeções tomográficas utilizando Kalman Discreto e Rede Neurais", *IEEE América Latina*, vol. 6, ed. 1, março, 2008.
- [22] M. A. M. Laia, P. E. Cruvinel, A. L. M. Levada, "Filtragem de projeções tomográficas da ciência do solo utilizando transformada de Anscombe e Kalman", *DINCON'07*, São José do Rio Preto, 2007.
- [23] M. A. M. Laia, P. E. Cruvinel. "Filtragem de projeções tomográficas do solo utilizando Kalman e Redes Neurais numa estimativa conjunta", *DINCON'08*, Presidente Prudente, 2008.
- [24] G. Welch, G. Bishop, "An introduction to the Kalman Filter". Department of Computer Science, University of North Carolina at Chapel Hill. 2004
- [25] P. Trealeaven, M. Pacheco, M. Vellasco, "VLSI Architectures for Neural Networks", *IEEE MICRO*, pp. 08-27, December, 1988.
- [26] E. B. Pizzolato, "Implementação de Redes Neurais no CPER", São Carlos, UFSCar, Dissertação de mestrado, 95 páginas, 1992.
- [27] W. S. McCulloch, W. Pitts, "How we know universals: the perception of auditory and visual forms", *Bulletin of mathematical Biophysics*, vol.9, pp.127-147, 1947.