

Aplicação de *Blending* em Múltiplas Resoluções para Mosaico de Imagens de Sensoriamento Remoto

Vantier Veronezi Bagli, Leila Maria Garcia Fonseca (orientador)
Instituto Nacional de Pesquisas Espaciais
Divisão de Processamento de Imagens
São José dos Campos – SP - Brasil
{vantier, leila}@dpi.inpe.br

Abstract

Image mosaic is a process of merging two or more images in order to obtain a larger one. Therefore, a global view of some area can be obtained. Two phases must be considered in the mosaic process: image registration and image blending. Taking into account that the images are already well registered, this work presents a blending process based on multi-resolution decomposition for seamless mosaicking of remote sensed images. Besides, the graph-cut approach has been used to look for the best splicing curve in the overlapping area.

1. Introdução

O mosaico de imagens tem a finalidade de juntar duas (ou mais) imagens (georreferenciadas e/ou registradas) para gerar uma imagem maior. Desta forma, uma visão completa das cenas coletadas pelos sensores pode ser obtida. Este processo é utilizado quando a área de estudo é maior que a cena disponibilizada pelo satélite. Como este caso é geralmente comum, o processo de mosaico torna-se indispensável em diversas aplicações.

Mesmo que o processo de registro seja bem aplicado, alguns problemas podem ser observados na área de transição entre as imagens. Se os valores de níveis de cinza nesta área forem diferentes, torna-se necessário determinar novos valores, de modo a suavizar a área de transição entre as imagens. O método empregado para solucionar este problema é chamado na literatura de *blending*, que consiste em determinar os valores de níveis de cinza na área de sobreposição entre as imagens para suavizar a transição entre as imagens mosaicadas. Um método muito utilizado é o da média ponderada dos níveis de cinza

dentro da área de sobreposição [1]. Em alguns casos a área de transição pode continuar perceptível mesmo aplicando métodos de interpolação. Este problema pode ocorrer devido a vários fatores: diferenças de textura, resolução, registro não adequado, área de transição ou de sobreposição pequena, linha de corte por feições lineares, entre outros.

Alguns autores têm proposto a análise em multiresolução para amenizar estes problemas na zona de transição [2, 3 e 4]. A idéia básica desta técnica é que ao decompor uma imagem em múltiplas resoluções, objetos de diferentes tamanhos (áreas de pastagens, edificações, etc) podem ser analisados separadamente nos diferentes níveis de resolução. Em cada nível de resolução o processo de interpolação pode ser mais específico, ou seja, o tamanho da zona de transição para a interpolação é escolhido de acordo com o nível de resolução.

Um trabalho que impulsionou a pesquisa para a geração de mosaicos utilizando análise por multiresolução foi proposto por Burt e Adelson [2]. As imagens que formam o mosaico são decompostas em múltiplas resoluções através da pirâmide Laplaciana. Assim, mosaicos com áreas de transição pouco perceptíveis e com baixo custo computacional podem ser obtidas. Hsu e Wu [3] propõem decompor as imagens através da Transformada Wavelet Discreta (TWD) e combiná-las de maneira similar ao proposto por Burt e Adelson. Usando também a TWD, Su et al [4] propõem um método em que o usuário interfere no processamento através da indicação de um parâmetro (λ) que serve para balancear dois termos: fidelidade dos dados e suavização na área de transição.

Os resultados apresentados nestes trabalhos mostram que a aplicação do processo de *blending* em múltiplas resoluções pode gerar mosaicos com melhor qualidade, comparados com a simples interpolação por média ponderada sobre as imagens originais.

Considerando que o sistema visual humano é sensível a linhas de corte reta (área de transição), a escolha adequada da linha de corte pode melhorar o processo de *blending*. No trabalho de Efron e Freeman [5], as junções das amostras são efetuadas utilizando uma linha de corte que é obtida por uma função de mínimo erro calculada na área de sobreposição. O erro mínimo é obtido pela diferença dos pixels pertencentes a esta área. Dentro deste contexto, Kwatra et al [6], propuseram uma técnica chamada *graphcut* na qual eles utilizam informações de textura para a realização de fotomontagem. Assim, a idéia básica deste processo é construir uma linha de corte nas regiões que possuem altas frequências (detalhes). Esta linha é determinada por uma função de mínimo custo, na qual as informações dos pixels adjacentes na zona de transição são usadas.

Portanto, o objetivo deste trabalho é apresentar uma metodologia de mosaico de imagens que combina a técnica de *graphcut*, para identificar a melhor linha de corte, e a análise em múltiplas resoluções para suavizar a área de transição preservando os objetos de alta frequência. Este trabalho faz parte da dissertação de mestrado do curso de computação aplicada do INPE que está sendo desenvolvida na Divisão de Processamento de Imagens – DPI/INPE.

2. Análise em múltiplas resoluções

Neste trabalho, resolução espacial de uma imagem é definida como a capacidade de discriminação de informação de um dado objeto de interesse. Por exemplo, em uma imagem de baixa resolução alvos de baixo contraste ou alvos pequenos podem aparecer borrados ou com pouca informação de detalhes se a resolução da imagem não for suficiente para representar a informação do objeto. Por outro lado, em uma imagem de alta resolução os objetos aparecem com boa definição dos contornos e, portanto, com maior informação de detalhes. Portanto, se os objetos são pequenos ou possuem baixo contraste pode ser necessário examiná-los em alta resolução. Por outro lado, se os objetos são grandes ou possuem alto contraste uma visão mais grosseira é suficiente. Se ambos os tipos de objetos aparecem em uma imagem, pode ser vantajoso analisá-los em múltiplas resoluções. Desta forma, a decomposição de uma imagem em múltiplas resoluções permite que a imagem seja representada e analisada em diferentes níveis de resolução.

A decomposição de uma imagem em múltiplas resoluções pode ser realizada por um banco de filtros [7]. Os filtros usados na decomposição são

representados como \tilde{G} e \tilde{H} , e os filtros de síntese como G e H . Os filtros G (\tilde{G}) e H (\tilde{H}) representam os filtros passa-altas e passa-baixas, respectivamente. O esquema de decomposição e reconstrução de uma imagem por transformada wavelet é ilustrado na Figura 1 e 2.

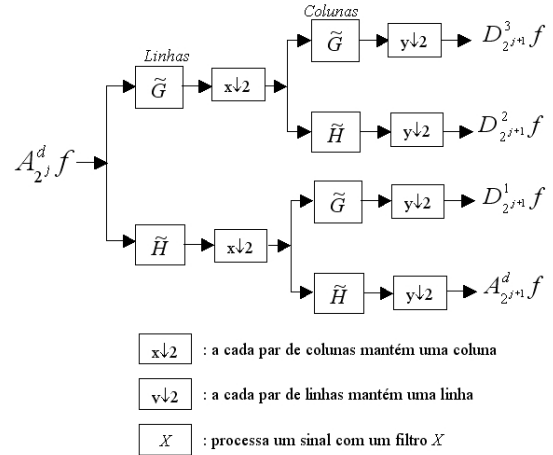


Figura 1. Decomposição de uma imagem $A_{2^j}^d f$ nas sub-imagens $A_{2^{j+1}}^d f$, $D_{2^{j+1}}^1 f$, $D_{2^{j+1}}^2 f$ e $D_{2^{j+1}}^3 f$.

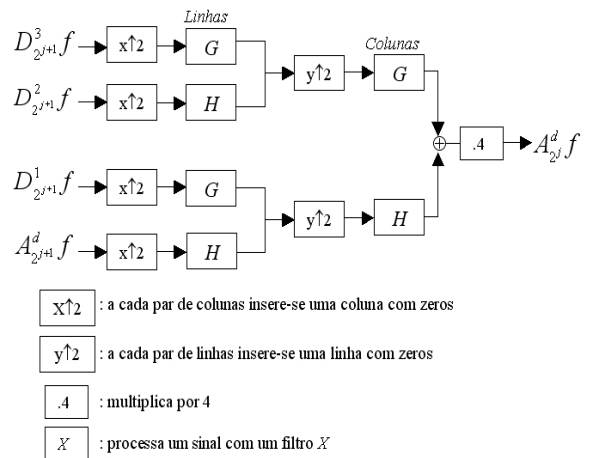


Figura 2. Reconstrução de uma imagem de aproximação discreta $A_{2^j}^d f$ a partir de uma imagem de aproximação na resolução mais baixa $A_{2^{j+1}}^d f$ e das imagens de detalhes $D_{2^{j+1}}^k f$ ($k=1, 2, 3$).

Fazendo uma análise no domínio da frequência, a imagem $A_{2^j}^d f$ corresponde às componentes de baixas frequências da imagem original, e as imagens $D_{2^j}^1 f$,

$D_{2^j}^2 f$ e $D_{2^j}^3 f$ representam, respectivamente, as componentes de altas frequências ao longo das direções vertical (bordas horizontais), horizontal (bordas verticais) e diagonal (cantos) na resolução 2^j . As imagens $A_{2^j}^d f$, $D_{2^j}^1 f$, $D_{2^j}^2 f$ e $D_{2^j}^3 f$ são também chamadas, respectivamente, de bandas LL (Low-Low), LH (Low-High), HL (High-Low) e HH (High-High).

A Figura 3 mostra as disposições destas imagens na representação através da decomposição wavelet em duas dimensões.

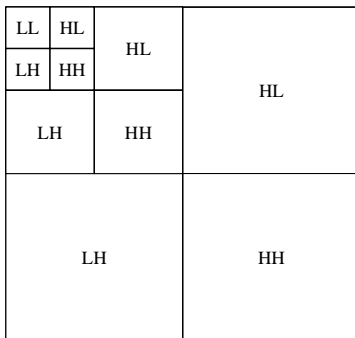


Figura 3. Disposição das imagens (HH, HL, LH e LL) na representação através da decomposição wavelet de uma imagem para três níveis de resolução.

3. Blending em múltiplas resoluções

O método de média ponderada pode ser realizado nos diferentes níveis de resolução e o objetivo principal é ocultar a linha de corte (emenda) na área de transição entre as imagens.

Para um melhor entendimento, a Figura 4 representa o método de média ponderada simples. Nesta área, o peso da imagem 1 decresce de 1.0 para 0.0 e o peso da imagem 2 cresce de 0.0 para 1.0. O novo pixel da área de transição será a soma dos níveis de cinza dos dois pixels ponderados pelos pesos das duas imagens.

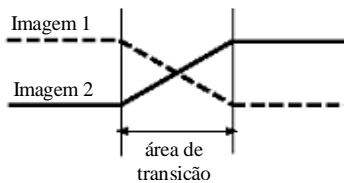


Figura 4. Blending por média ponderada.

No algoritmo de média ponderada, baseado em multiresolução, a média ponderada é realizada em cada nível de resolução dentro de uma área de transição específica ao nível de resolução, como ilustra a Figura 5. Os diferentes comprimentos das zonas de transição

estão relacionados com o nível de resolução na decomposição. Pode-se observar que no nível de resolução mais baixa (poucos detalhes) a área de transição é maior. À medida que a resolução aumenta (mais detalhes) o tamanho da área de transição diminui, evitando assim a geração de artefatos devido à mistura de informação. Portanto, pode-se obter uma interpolação ponderada específica para diferentes níveis de resolução. O ponto de intersecção entre a reta pontilhada e a contínua indica um ponto da linha de corte.

Uma outra forma de realizar o *blending* em múltiplas resoluções é através de formas pré-estabelecidas (máscara binária). A máscara permite estabelecer, previamente, quais partes da imagem serão consideradas e excluídas para compor uma outra imagem [2, 3 e 4].

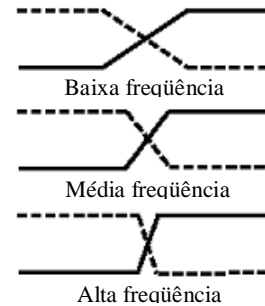


Figura 5. Transição em *Blending* por multiresolução.

Um exemplo de *blending* utilizando máscaras binárias pode ser visto na Figura 6. Duas imagens 6(a) e 6(b), são mosaicadas utilizando a máscara binária 6(c) que define quais partes da imagem 6(b) serão “incluídas” e “excluídas” no processo. Em (d) pode-se observar a imagem resultante obtida em Su et al [4].

Segundo Burt e Adelson [2] e Hsu e Wu [3] a máscara binária S que determina quais pixels pertencerão ao mosaico (valor 1) e quais serão excluídos (valor 0) é convoluída por um filtro passa-baixas h' para todos os níveis da decomposição, obtendo as máscaras $C'_M(S)$, ..., $C'_2(S)$, $C'_1(S)$, onde M representa o número de níveis de decomposição e $C'_i(S)$ ($i=1, \dots, M$) as máscaras que funcionam como funções de ponderação. Assim, uma pirâmide Gaussiana da máscara é gerada, onde a área de transição aumenta em direção ao topo da pirâmide.

Hsu e Wu [3] têm proposto um método para gerar mosaico de imagens baseado na decomposição por transformada wavelet (TW). O procedimento é similar ao proposto em [2] com a diferença que, neste caso, em

cada nível de resolução tem-se quatro sub-bandas (LL, HL, LH e HH) que são processadas.

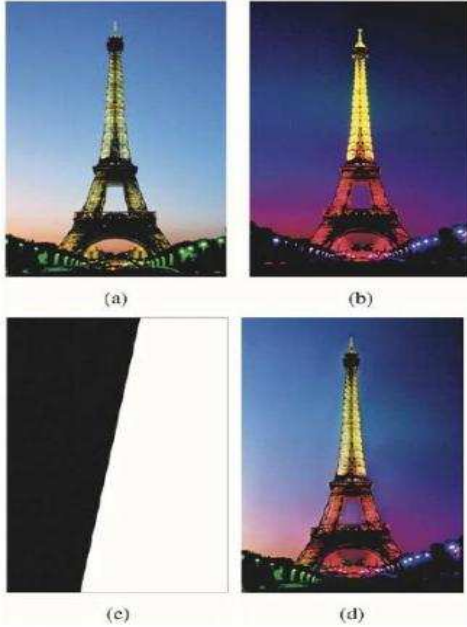


Figura 6. Exemplo de *blending* de duas imagens (a) e (b) utilizando máscara binária (c) para interpolação e definição da linha de corte e em (d) o resultado do processo.

4. Linha de corte

Nas próximas seções (4.1 e 4.2) serão apresentadas algumas técnicas que podem ser utilizadas no processo de *blending*. Algumas destas técnicas podem ser utilizadas para criar uma linha de corte mais adequada para o processo de geração de mosaicos.

4.1 Síntese de imagens

Um método de síntese de imagens, considerando atributos de textura, foi proposto por Efros e Freeman [5] e é conhecido por “*image quilting*”.

A linha de corte construída pelo método *image quilting*, consiste em minimizar as diferenças de textura que existe entre as sobreposições das amostras de imagens que serão unidas. A linha de corte por mínimo erro é calculada de maneira simples. Se dois blocos B_1 e B_2 possuem áreas de sobreposição vertical B_1^{ov} e B_2^{ov} , respectivamente, o erro da área é definido como $e = (B_1^{ov} - B_2^{ov})^2$. Assim, a linha de corte é construída utilizando a função de menor erro acumulativo E :

$$E_{i,j} = e_{i,j} + \min(E_{i-1,j-1}, E_{i-1,j}, E_{i-1,j+1}) \quad (1)$$

A construção da linha de corte na área de sobreposição de dois blocos de texturas similares pode ser realizada utilizando programação dinâmica a partir das medições de mínimo erro.

A Figura 7 ilustra um exemplo de uma linha de corte que utiliza a função de menor erro acumulativo. Como pode ser observado, a área de junção (costura) das duas amostras ficou praticamente imperceptível. Desta forma, a utilização desta técnica no processo de *blending* para mosaicos de imagens de sensoriamento remoto, pode gerar um resultado melhor.

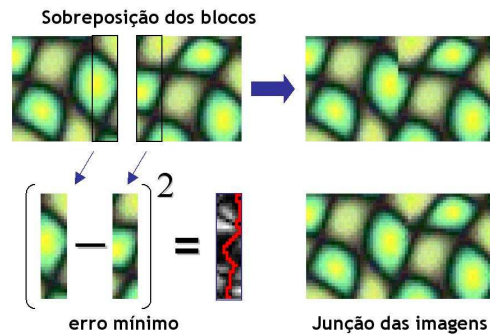


Figura 7. Junção de amostras de textura utilizando linha de corte gerada pelo menor erro.

4.2 Graphcut

Outra forma de construir uma linha de corte considerando atributos de textura é através da técnica *graphcut* [6]. A idéia básica deste processo, é construir uma linha de corte nas regiões que possuem altas frequências (detalhes). Desta forma, confundindo o sistema visual humano nas regiões de altas frequências, a linha de corte pode ficar mais imperceptível.

Considere uma imagem vista como um grafo orientado $G = (V, E)$ apresentado na Figura 8(a), onde os vértices dos *pixels* são os vértices do grafo e cada aresta de *pixel* define dois arcos com orientações opostas. $G = (V, E)$ consiste de um conjunto finito V de $|V|$ vértices e um conjunto E de $|E|$ arcos. Cada vértice $v \in V$ pode ter até quatro vértices $u_k \in V$, $k = 1, 2, 3, 4$, distintos e vizinhos-4, formando o conjunto $A(v)$ dos vértices adjacentes ao vértice v em G (Figura 8(b)). Um arco $e = (v, u)$ em E é um par ordenado de vértices adjacentes em G que inicia no vértice v , termina no vértice $u \in A(v)$ e tem associado um valor de comprimento $l(u,v)$, ou $l(e)$, no intervalo $[0, C_{max}]$ onde C_{max} é o maior custo associado a qualquer arco do grafo $G(V,E)$. Entre dois vértices adjacentes, $v, u \in V$, existem dois possíveis arcos, (u, v) , $(v, u) \in E$.

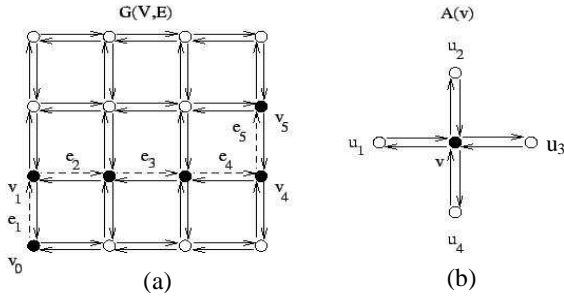


Figura 8. Em (a) o grafo $G(V,E)$ definido para imagem e (b) o vértice com vizinhança 4.

Existem várias formas de calcular o valor de comprimento $l(e)$. Uma forma de encontra-lo é apresentado em Kwatra et al [6], onde são utilizados, em cada vértice, valores dos níveis de cinza (NC) de um par de *pixels* em uma mesma posição. Assim, tendo $A(s)$ e $B(s)$ como valores de NC dos *pixels* na posição s das imagens A e B, respectivamente, defini-se uma função de custo M . Esta função fornece o valor de $l(e)$ entre os *pixels* adjacentes s e t , por exemplo, os nós v e u_3 da Figura 8(b).

$$M(s,t,A,B) = \|A(s) - B(s)\| + \|A(t) - B(t)\| \quad (2)$$

A linha de corte pode ser obtida minimizando a função de custo M (Equação 2). As discontinuidades e bordas são mais perceptíveis em regiões de baixas frequências do que em regiões de altas frequências. Com isto pode-se utilizar os gradientes em cada direção (horizontal e vertical) e a equação 2 para construir uma nova função:

$$M'(s,t,A,B) = \frac{M(s,t,A,B)}{\|G_A^d(s)\| + \|G_A^d(t)\| + \|G_B^d(s)\| + \|G_B^d(t)\|}, \quad (3)$$

onde d é o indicador da direção do gradiente que possui a mesma direção que o arco entre s e t . G_A^d e G_B^d são os gradientes das partes A e B ao longo da direção d . Desta forma M' dificulta a passagem da linha de corte nas regiões mais homogêneas.

5. Metodologia

Baseada na fundamentação teórica apresentada nas seções anteriores, a metodologia proposta no desenvolvimento da dissertação é dividida em três fases: decomposição das imagens registradas,

determinação da linha de corte e aplicação do processo de *blending*.

A Figura 9 ilustra o esquema de mosaico automático que está sendo desenvolvido neste trabalho.

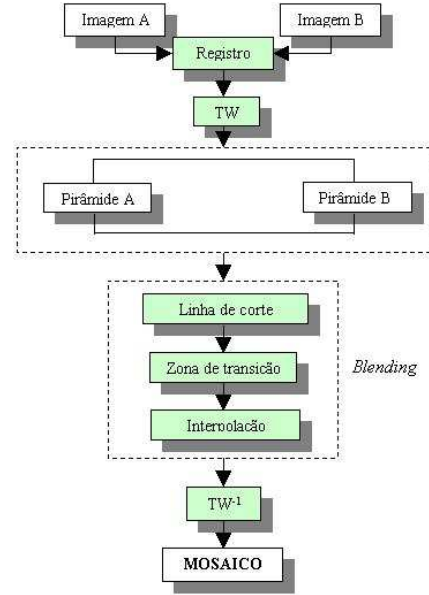


Figura 9. Esquema do mosaico automático.

Inicialmente as imagens a serem mosaicadas são decompostas em diferentes níveis de resolução através da transformada wavelet. Nesta fase dois problemas devem ser resolvidos: a seleção da wavelet e o número de níveis de decomposição. Em relação a decomposição em multiresolução pretende-se usar a transformada wavelet baseada na implementação do Mallat [7].

Na etapa de *blending* duas fases são realizadas. A primeira fase consiste em determinar a linha de corte, na área de sobreposição, com menor erro para fazer a emenda das imagens. Normalmente, este caminho coincide com as bordas dos objetos ou com as regiões de maior textura para confundir o sistema visual humano. Esta linha de corte pode ser gerada através das técnicas descritas nas seções 4.1 e 4.2. A segunda etapa consiste em realizar a interpolação radiométrica dentro da região de transição.

A Figura 10 ilustra o processo de *blending* usando o esquema de multiresolução.

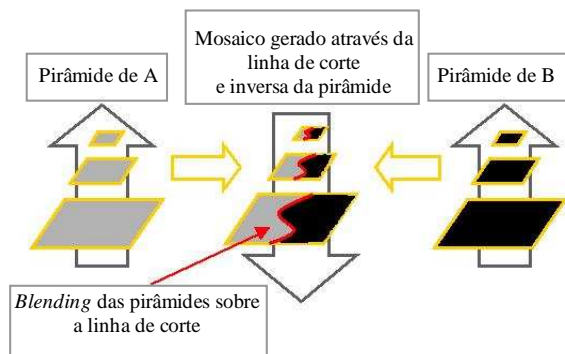


Figura 10. Esquema do mosaico automático.

Os principais problemas a serem investigados são: identificação da linha de corte de menor erro, a escolha do tamanho da área de transição e a aplicação do método de interpolação, que geralmente é através da média ponderada. Como o processamento será realizado no domínio da transformada wavelet é necessário calcular a transformada inversa para obter o mosaico.

Os métodos serão testados em imagens controladas tais como fotos digitais e em seqüências de imagens aéreas e imagens de satélites multi-temporais, que são os casos mais difíceis para realizar o mosaico.

A avaliação será realizada por inspeção visual e de forma quantitativa através do índice IQI [10].

7. Conclusão

Registro e mosaico de imagens são procedimentos muito utilizados nas diversas aplicações na área de sensoriamento remoto, incluindo a área biológica (medicina, odontologia, etc). Pesquisadores que utilizam o sensoriamento remoto dependem da qualidade da imagem para o desenvolvimento eficiente de seus trabalhos. Portanto, os métodos de *blending* que estão sendo desenvolvidos vão contribuir para a melhoria de geração de mosaicos de melhor qualidade beneficiando assim, a comunidade de usuários em sensoriamento remoto. O método de *blending* desenvolvido, neste trabalho, será integrado ao sistema de registro e mosaico automático de imagens, desenvolvido no INPE, em cooperação com a Universidade da Califórnia, Santa Bárbara (UCSB) [9].

8. Referências

- [1] Meunier, L.; Borgmann, M., "High resolution panoramas using image mosaicing", Stanford University, EE368 (Digital Image Processing) Final Project., maio 2000 Disponível em: <www.moriborg.de/ee/mosaicing/>. Acesso em 31 de janeiro de 2005.
- [2] Burt, P. J.; Adelson, E. H., "Multiresolution spline with application to image mosaics", *ACM Transaction on Graphics*, v. 2, n. 4, p. 217-236, out. 1983.
- [3] Hsu, C. T.; Wu, J. L., "Multiresolution mosaic", *IEEE Transactions Consumer Electron*, v. 42, p. 981-990, ago. 1996.
- [4] Su, M. S.; Hwang, W. L.; Cheng, K. Y., "Analysis on multiresolution mosaic images", *IEEE Transactions on Image Processing*, v. 13, n. 7, jul. 2004.
- [5] Efros, A. A.; Freeman, W. T., "Image quilting for texture synthesis and transfer", In: *Proceedings of SIGGRAPH 2001*. Anais: ago. 2001, p. 341-346.
- [6] Kwatra, V.; Schodl, A.; Essa, I.; Turk, G.; Bobick, A., "Graphcut textures: image and video synthesis using graph cuts", In: *Proc. ACM Transactions on Graphics, SIGGRAPH 2003*, Anais, 2003.
- [7] Mallat, S. G., "A theory for multiresolution signal decomposition: the wavelet representation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 11, p. 674-693, jul. 1989.
- [8] Fonseca, L. M. G., "Registro automático de imagens de sensoriamento remoto", São José dos Campos - SP. 219p. (INPE - 7502-TDI/717). Tese (Tese de doutorado em Computação Aplicada) - Instituto Nacional de Pesquisas Espaciais, 1999.
- [9] Fedorov, D., "Sistema semi-automático de registro e mosaico de imagens", São José dos Campos. 153p. (INPE - 9582-TDI/838). Dissertação (Mestrado em Computação Aplicada) - Instituto Nacional de Pesquisas Espaciais, 2002.
- [10] Wang, Z.; Bovik, A. C., "A universal image quality index", *IEEE Signal Processing Letters*, v. 9, n. 3, p. 81-84, mar. 2002.