

Emprego de Análise em Multiresolução para Mosaicagem de Imagens de Sensoriamento Remoto

Vantier Veronezi Bagli
Divisão de Processamento de Imagens
Instituto Nacional de Pesquisas Espaciais
São José dos Campos – SP – Brasil
vantier@dpi.inpe.br

Leila Maria Garcia Fonseca
Divisão de Processamento de Imagens
Instituto Nacional de Pesquisas Espaciais
São José dos Campos – SP – Brasil
leila@dpi.inpe.br

Resumo

Mosaico de imagens é o processo de juntar duas ou mais imagens para gerar uma imagem maior. Duas fases são envolvidas neste processo: registro de imagens e *blending*. O processo de *blending* consiste em combinar as imagens na área de sobreposição de tal forma que a emenda entre as imagens fique visualmente imperceptível. Mesmo que as imagens estejam bem registradas, a área de transição entre as imagens pode ficar visível e as nuvens e sombras podem gerar artefatos e borramento na região de junção. Assim, este trabalho apresenta, de forma sucinta, um método de *blending* e remoção de nuvens/sombras baseado na decomposição em múltiplas resoluções. A geração de uma linha de corte, que considera informação de textura, pode ser empregada para combinar as duas imagens de uma maneira mais otimizada.

Palavras-chave: *wavelet*, *blending*, mosaico, remoção de nuvens, linha de corte.

1. Introdução

O mosaico de imagens tem a finalidade de juntar duas (ou mais) imagens 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.

O processo de mosaico envolve, basicamente, duas fases: registro de imagens e a aplicação de *blending*. Mesmo quando o processo de registro é bem feito, 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 das imagens. Um método muito utilizado é o da média ponderada dos níveis de cinza dentro da área de sobreposição [8]. A Figura 1a ilustra o problema existente na área de transição entre duas imagens mosaicadas e a Figura 1b um mosaico com aplicação de *blending* por média ponderada.



(a)



(b)

Figura 1. Em (a) e (b) tem-se, respectivamente, mosaico gerado sem e com a aplicação de *blending*, e a seta vermelha representa um exemplo da área de transição.

Em alguns casos, a área de transição pode continuar perceptível mesmo quando o método de interpolação pela média ponderada é usado. 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 inadequada, entre outros. A Figura 2 são partes da Figura 1b que ilustram alguns dos problemas que podem existir mesmo com a aplicação de *blending*.



Figura 2. Problema na área de transição: linha de corte é visível.

Alguns autores têm proposto a análise em multiresolução para amenizar os problemas na zona de transição [1, 5, 10]. A ideia básica desta técnica é que ao decompor uma imagem em múltiplas resoluções, diferentes frequências podem ser analisadas 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, 1983. As imagens que formam o mosaico são decompostas em múltiplas resoluções através da pirâmide Laplaciana. Hsu e Wu, 1996 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., 2003 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.

Considerando que o sistema visual humano tem uma maior capacidade em detectar mudanças em áreas homogêneas (baixas frequências), a construção de uma linha de corte sobre regiões de altas frequências pode melhorar o processo de *blending*. Esta linha de corte é a linha que determina o local que as imagens serão combinadas (costuradas). No trabalho de Efros e Freeman, 2001, as amostras são combinadas utilizando uma linha de corte obtida por uma função de mínimo erro calculada na área de sobreposição. Kwatra et al.,

2003, propõem uma técnica, chamada *graphcut*, na qual utilizam informações de textura para a realização de fotomontagem. Esta linha é determinada por uma função de mínimo custo, na qual as informações dos *pixels* adjacentes na zona de transição são usadas.

Nuvens e sombras, além de apresentarem um problema estético no mosaico e encobrir algumas áreas da imagem que possuem informações importantes, podem prejudicar o processo de *blending*. A Figura 3 mostra este problema, em que as nuvens/sombras foram misturadas com a pastagem, gerando assim um objeto indefinido.



Figura 3. Problema ocorrido pela presença de nuvens na área de sobreposição.

Para que este problema não ocorra, torna-se necessário um processo de remoção de nuvens. Para realizar este processo são envolvidas duas etapas. A primeira baseia-se na identificação destes objetos e a segunda na substituição deles por dados provenientes da imagem que não possui nuvens. Algumas metodologias para identificação de nuvens e sombras em uma imagem têm sido propostas [2, 9, 12]. O processo de substituição pode ser realizado pelo mesmo esquema de multiresolução utilizado por Hsu e Wu, 1996 em fotomontagens.

Portanto, o objetivo deste trabalho é apresentar uma metodologia de mosaico de imagens que combina a técnica de geração da melhor linha de corte e a análise em múltiplas resoluções para suavizar a área de transição como também remover as nuvens e sombras presentes nesta área. As técnicas propostas por Hsu e Wu, 1996 e Kwatra et al., 2003 são adaptadas e combinadas para gerar mosaico de imagens de satélites. Esta metodologia faz parte da dissertação de mestrado na área de computação aplicada no INPE. Resultados preliminares são apresentados para realizar o mosaico de duas imagens do *LANDSAT-5*, adquiridas em datas diferentes. Um processo de remoção de nuvens e sombras foi testado utilizando o mesmo esquema de multiresolução em imagens *LANDSAT*.

2. Análise em múltiplas resoluções

Na decomposição por wavelet uma imagem pode ser decomposta através de um banco de filtros [7]. São usados filtros passa-altas e passa-baixas, representados como H e L, respectivamente. Fazendo uma análise no domínio da frequência, a imagem LL (*Low-Low*) corresponde às componentes de baixas frequências da imagem original, e as imagens LH (*Low-High*), HL (*High-Low*) e HH (*High-High*) representam, respectivamente, as componentes de altas frequências ao longo da direção vertical (bordas horizontais), horizontal (bordas verticais) e diagonal (cantos) na resolução 2^l . A Figura 4 mostra as disposições destas imagens na representação através da decomposição wavelet em duas dimensões.

LL	HL	HL	HL
LH	HH		
LH		HH	
LH		HH	

Figura 4. Disposição das imagens (HH, HL, LH e LL) obtidas por decomposição wavelet em três níveis.

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 área de transição (emenda) das imagens. A Figura 5 representa um exemplo de *blending* por média ponderada simples, aplicado diretamente sobre a imagem original. Na área de transição, 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* desta área será a soma dos níveis de cinza dos dois *pixels* ponderados pelos pesos das duas imagens.

No algoritmo de média ponderada baseado em multiresolução, a média é realizada em cada nível de resolução dentro de uma área de transição específica, como ilustra a Figura 6. Os diferentes comprimentos das zonas de transição estão relacionados com o nível de resolução representados. 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. O ponto de intersecção entre a reta pontilhada e a contínua indica um ponto da linha de corte.

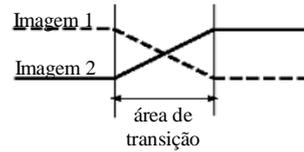


Figura 5. *Blending* por média ponderada sobre as imagens originais.

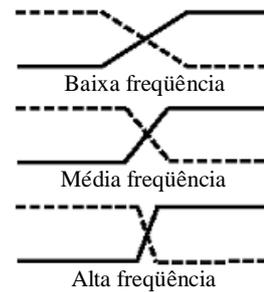


Figura 6. Áreas de transições de *blending* em múltiplas resoluções.

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 (valor 1) e excluídas (valor 0) no processo de composição [1, 5]. Esta máscara 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, 1996 apresentam um método que gera mosaico de imagens baseado na decomposição por transformada wavelet (TW). O procedimento é similar ao proposto em Burt e Adelson com a diferença que, neste caso, em cada nível de resolução tem-se três sub-bandas (HL, LH e HH) mais a sub-banda LL pertencente ao último nível. Com isso, além de se utilizar frequências similares no processo de interpolação em múltiplas resoluções, a utilização da TW também permite que feições com direções similares sejam interpoladas no mesmo esquema.

4. Linha de corte

Kwatra et al., 2003 propõem um método de construir uma linha de corte considerando atributos de textura, conhecida como *graphcut*. 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 7a, 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 $u_k \in V$, $k = 1, 2, 3$ e 4 , distintos e vizinhos-4, formando o conjunto $A(v)$ dos vértices adjacentes ao vértice v em G (Figura 7b). 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, $(u, v), (v, u) \in E$.

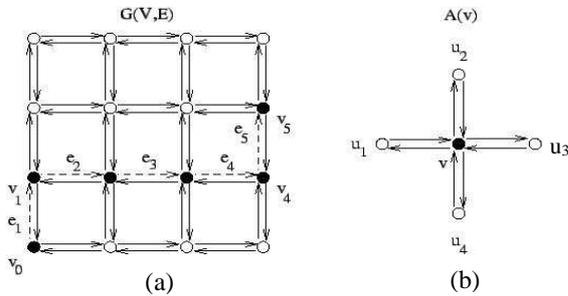


Figura 7. Em (a) o grafo $G(V,E)$ definido para imagem e (b) o vértice com vizinhança 4.

Existem várias formas de calcular o valor de comprimento $l(e)$. Uma forma de encontrá-lo é apresentado em Kwatra et al., 2003, 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 7b.

$$M(s,t,A,B) = \|A(s) - B(s)\| + \|A(t) - B(t)\|. \quad (1)$$

A linha de corte pode ser obtida minimizando a função de custo M (Equação 1). As descontinuidades 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 1 para construir uma nova função:

$$M'(s,t,A,B) = \frac{M(s,t,A,B)}{\|G_A^d(s)\| + \|G_A^d(t)\| + \|G_B^d(s)\| + \|G_B^d(t)\|}, \quad (2)$$

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 e Resultados Preliminares

Baseada na fundamentação teórica apresentada nas seções anteriores, a metodologia proposta consiste em combinar a técnica de *graphcut* e a análise em múltiplas resoluções. A metodologia proposta é dividida em três fases: decomposição das imagens registradas, geração da máscara binária e aplicação do processo de *blending*. O processo de registro é realizado pelo *REGGEEMY*, que é um sistema de registro automático desenvolvido pela Divisão de Processamento de Imagens, INPE, em cooperação com a Universidade de Santa Bárbara, Califórnia [4].

A Figura 8 ilustra o esquema de mosaico proposto neste trabalho. Após o registro das imagens a serem mosaicadas, o número de níveis de decomposição das imagens é definido. Este número pode ser definido utilizando informações das imagens tais como frequência, tamanho da área de sobreposição, resolução, entre outros possíveis parâmetros. Em seguida, as imagens são decompostas através da transformada wavelet baseada na implementação do Mallat, 1989.

Na etapa de geração da máscara binária duas fases são realizadas. A primeira fase consiste em determinar a linha de corte, na área de sobreposição, com menor erro para emendar as imagens. Esta linha de corte pode ser gerada através da técnica de *graphcut*. A segunda fase consiste em construir a pirâmide Gaussiana da máscara com a mesma quantidade de níveis utilizada na decomposição das imagens de entrada.

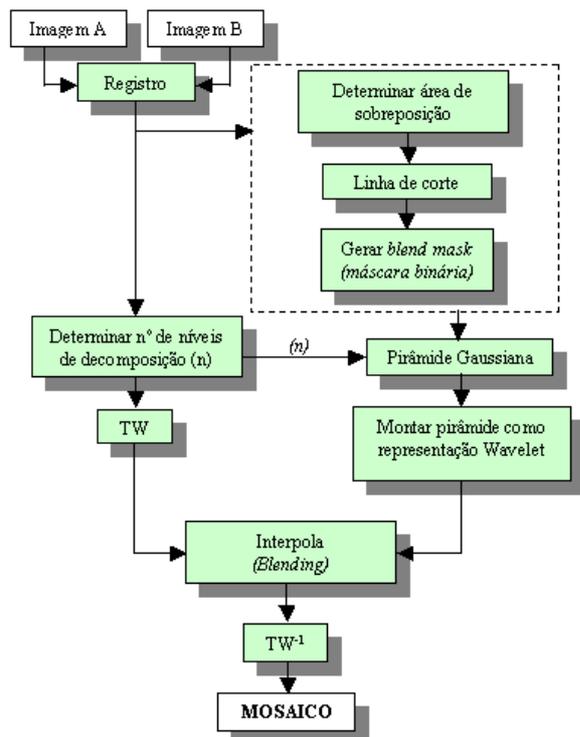


Figura 8. Esquema do mosaico automático

A última etapa consiste em aplicar o processo de interpolação em múltiplas resoluções (*blending*) utilizando as imagens obtidas pela TW e as máscaras obtidas pela pirâmide Gaussiana. Como o processamento será realizado no domínio da transformada wavelet é necessário calcular a transformada inversa para obter o mosaico.

Para testar o desempenho do método proposto, duas imagens *LANDSAT-5*, adquiridas em datas diferentes, são usadas. O resultado preliminar é ilustrado na Figura 9.

A Figura 9 mostra as duas imagens *LANDSAT-5* (9a e 9b), que foram mosaicadas utilizando a máscara binária 9e. A “linha” que separa a parte branca da preta na máscara binária é a linha de corte. Neste caso, a linha foi construída manualmente. A decomposição da máscara através da pirâmide Gaussiana define uma área de transição para cada nível de decomposição. Em 9f pode-se observar a imagem resultante obtida pelo algoritmo descrito neste trabalho. Pode-se observar (círculos vermelhos da Figura 9) que o uso de *blending* em múltiplas resoluções por TW produz resultados melhores do que o algoritmo que usa a média ponderada simples no processo de *blending* (Figura 9d). Neste exemplo, não foi usada nenhuma técnica específica para a geração da linha de corte.

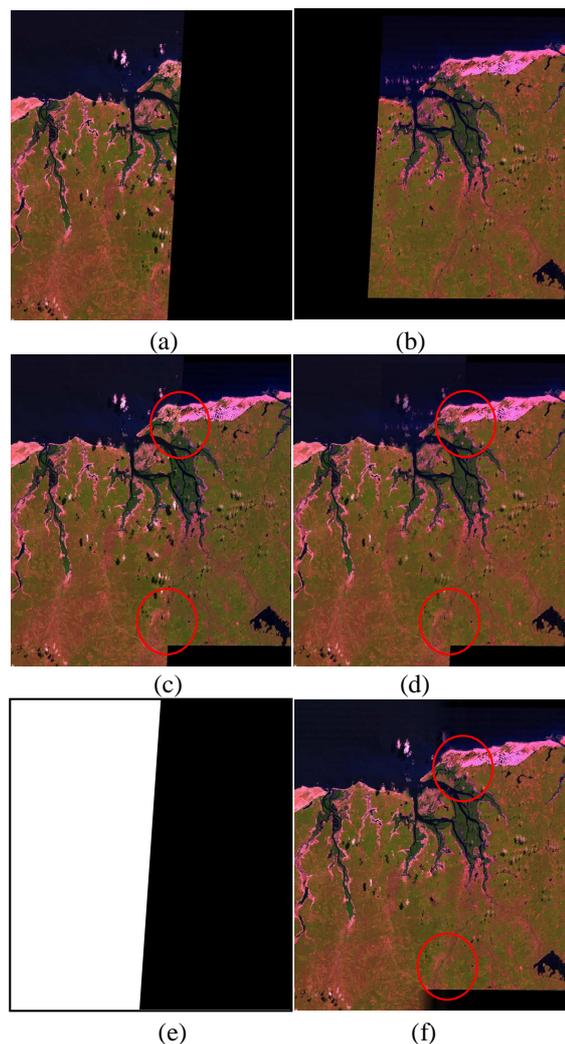


Figura 9. Imagens *LANDSAT-5* registradas em (a) e (b), o mosaico sem tratamento de emenda (c), com aplicação de *blending* por média ponderada simples (d), a máscara de binária (e) e a imagem resultante pelo método proposto (f).

O processo de identificação e remoção de nuvens e sombras poderá ser realizado no mesmo esquema ilustrado na Figura 8, especificamente na área pontilhada, onde no lugar de geração de linha de corte tem-se o processo de identificação das nuvens [2, 9, 12]. Em seguida será gerada a máscara binária que identifica os locais onde há presença de nuvens e sombras. Na seqüência, a máscara é dilatada para garantir que as bordas das nuvens e sombras não sejam incluídas no processo. Um processo de Hsu e Wu utilizado em fotomontagens foi adaptado para preenchimento das nuvens e sombras. Um resultado preliminar deste processo pode ser visto na Figura 10.

Como no exemplo anterior, a máscara binária que identifica os locais onde há presença de nuvens/sombras foi construída manualmente. As setas vermelhas da Figura 10 (c e d) mostram que a aplicação deste método é promissor, pois houve um ganho de qualidade se compararmos com os métodos de preenchimento direto.

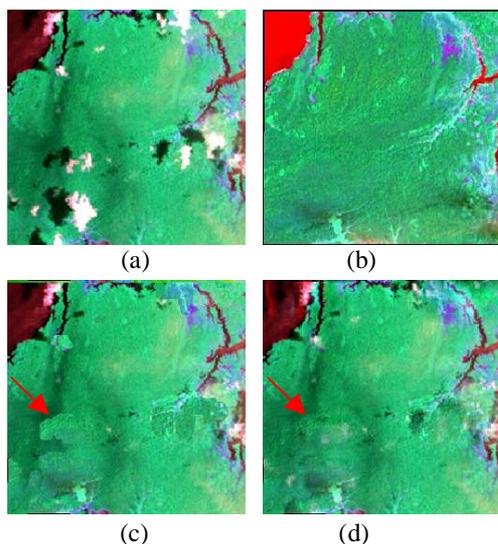


Figura 10. Imagem de interesse com nuvens (a), imagem sem nuvens (b), remoção das nuvens por preenchimento direto (c) e por multiresolução em (d).

As avaliações dos resultados serão realizadas por inspeção visual e de forma quantitativa através do índice IQI [11].

6. Conclusão

Com os resultados obtidos até agora pode-se verificar que os métodos de *blending* que estão sendo desenvolvidos produzem mosaicos de melhor qualidade que os obtidos pelo método da média ponderada simples. Neste trabalho não foi usada nenhuma técnica de busca da melhor linha de corte, apenas foi usada a interpolação ponderada para combinar as duas imagens. Espera-se que com o uso da melhor linha de corte, o resultado possa melhorar ainda mais. Com isto, as técnicas propostas poderão 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) [4].

8. Referências

- [1] 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.
- [2] Carvalho, L. M. T.; "Mapping and monitoring forest remnants: a multi-scale analysis of spatio-temporal data." Wur, Holanda. 138p. (ISBN 90-5808-545-7). Tese (Tese de doutorado em Geociências) - Wageningen University and Research Centre, dez. 2001.
- [3] Efros, A. A.; Freeman, W. T., "Image quilting for texture synthesis and transfer", *In: Proceedings of SIGGRAPH 2001*. Anais: ago. 2001, p. 341-346.
- [4] 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.
- [5] Hsu, C. T.; Wu, J. L., "Multiresolution mosaic", *IEEE Transactions Consumer Electron*, v. 42, p. 981-990, ago. 1996.
- [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] 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.
- [9] Polidório, A. M.; Flores, F. C.; Franco, C.; Imai, N. N.; Tommaselli, A. M. G. "Detecção automática de sombras, nuvens e corpos d'água em imagens cbers e landsat 7 ETM." *In: XII Simpósio Brasileiro de Sensoriamento Remoto*. Anais. Goiânia, Brasil: INPE, 2005
- [10] 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.
- [11] Wang, Z.; Bovik, A. C., "A universal image quality index", *IEEE Signal Processing Letters*, v. 9, n. 3, p. 81-84, mar. 2002.
- [12] Wang, C.; Huang, L.; Rosenfeld, A. "Detecting clouds and cloud shadows on aerial photographs." *Pattern Recognition Letters*, v. 12, n. 1, jan. 1991.