

Mosaico de Imagens baseado na Análise em Múltiplas Resoluções

Vantier Veronezi Bagli, Leila M. G. Fonseca

Instituto Nacional de Pesquisas Espaciais – (INPE)

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. Two phases are 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 a multi-resolution decomposition for seamless mosaicking of satellite images. The generation of a cut line, which considers texture information, can be used to look for the best splicing curve in the overlapping area.*

Resumo. *Mosaico de imagens é o processo de juntar duas ou mais imagens para gerar uma imagem maior. Duas fases são envolvidas no processo de mosaico: 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. Considerando que as imagens estão bem registradas, este trabalho apresenta um método de blending baseado na decomposição em múltiplas resoluções. A geração de uma linha de corte, que considera informação de textura, pode ser empregada para buscar a melhor linha de corte para combinar as duas imagens.*

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 operação 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 [Meunier e Borgmann 2000]. 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.

Alguns autores têm proposto a análise em multiresolução para amenizar os problemas na zona de transição [Burt e Adelson 1983], [Hsu e Wu 1996] e [Su et al 2001]. A idéia 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.

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. 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.

2. Análise em Múltiplas Resoluções

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. Se objetos pequenos e grandes aparecem em uma imagem, pode ser vantajoso analisá-los em múltiplas resoluções. 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 [Mallat 1989]. 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.

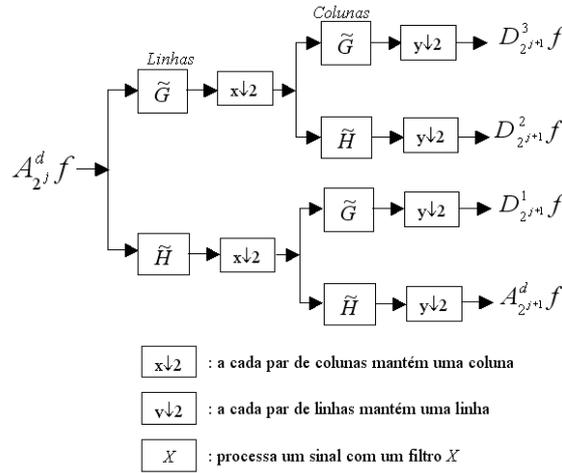


Figura 1. Decomposição de uma imagem $A_{2^j}^d f$ nas sub-imagens $A_{2^{j+1}}^d f$, $D_{2^{j+1}}^1 f$, $D_{2^{j+1}}^2 f$ e $D_{2^{j+1}}^3 f$.

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).

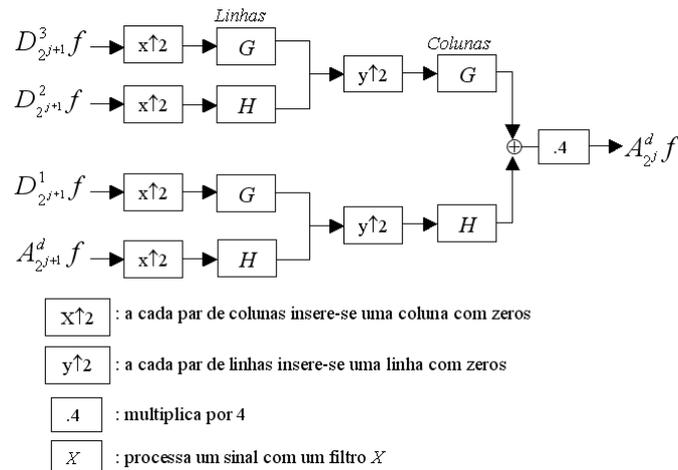


Figura 2. Reconstrução de uma imagem de aproximação discreta $A_{2^j}^d f$ a partir de uma imagem de aproximação $A_{2^{j+1}}^d f$ e das imagens de detalhes $D_{2^{j+1}}^k f$ ($k=1, 2, 3$).

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 3 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.

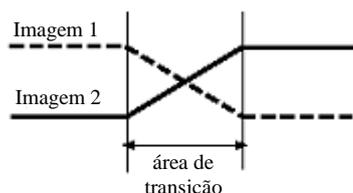


Figura 3. *Blending* por média ponderada.

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, como ilustra a Figura 4. 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.

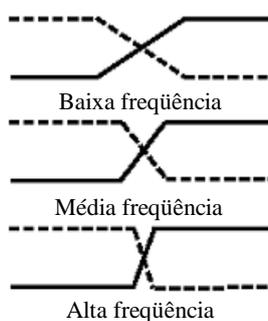


Figura 4. Transição em *blending* por multiresolução.

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 [Hsu e Wu 1996]. 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 (1983) 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 5(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 $u_k \in V$, $k = 1, 2, 3$ e 4 , distintos e vizinhos-4, formando o conjunto $A(v)$ dos vértices adjacentes ao vértice v em G (Figura 5(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, $(u, v), (v, u) \in E$.

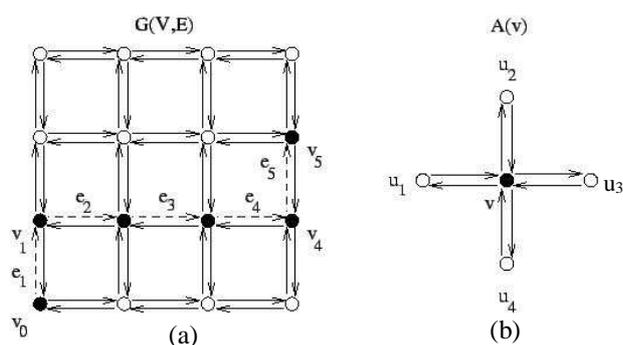


Figura 5. Em (a) o grafo $G(V,E)$ definido para imagem e (b) o vértice com vizinhança 4.

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 5(b).

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

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

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 *REGEMY*, que é um sistema de registro automático desenvolvido pela Divisão de Processamento de Imagens, INPE, em cooperação com a Universidade de Santa Bárbara, Califórnia [Fedorov 2002].

A Figura 6 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.

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.

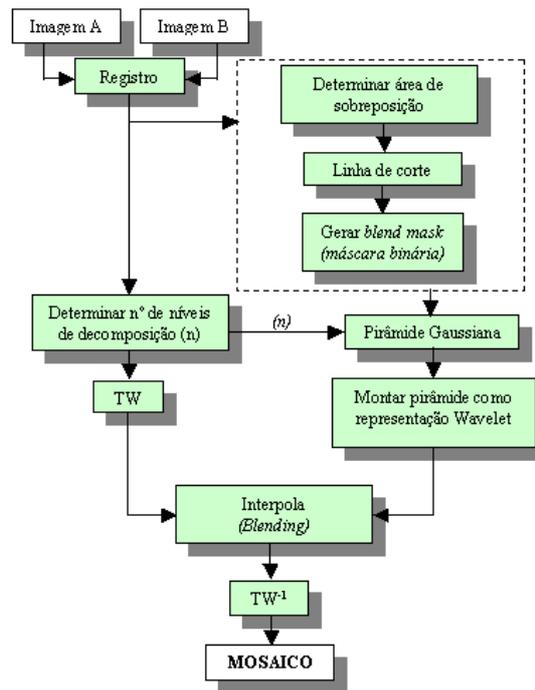


Figura 6. Esquema do mosaico automático.

Para testar o desempenho do método proposto, duas imagens Landsat-5, adquiridas em datas diferentes, são usadas. O resultado preliminar é ilustrado na Figura 7.

A Figura 7 mostra as duas imagens *LANDSAT* (7(a) e 7(b)), que foram mosaicadas utilizando a máscara binária 7(e). 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 7(f) tem-se a imagem resultante obtida pelo algoritmo descrito neste trabalho. Pode-se observar, através dos círculos da Figura 7, 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 7(d)). Neste exemplo, não foi usada nenhuma técnica específica para a geração da linha de corte.

Os métodos implementados serão avaliados por inspeção visual e de forma quantitativa através do índice IQI [Wang e Bovik 2002].

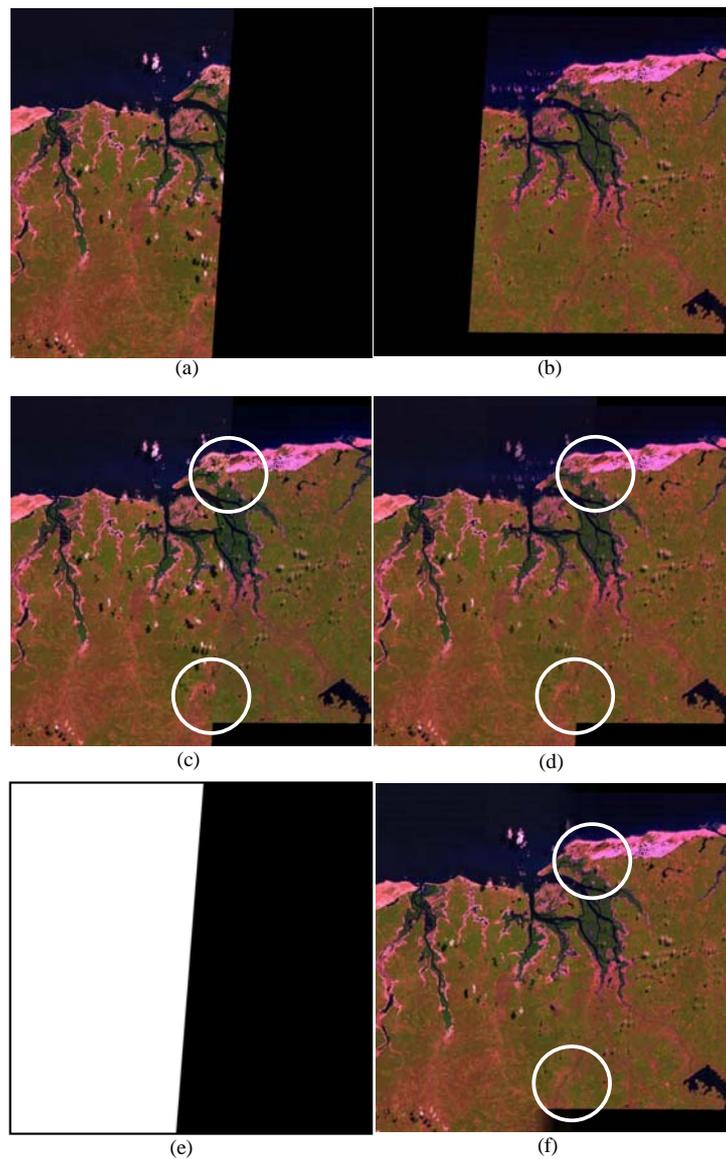


Figura 7. Imagens *LANDSAT-5* registradas em (a) e (b), o mosaico sem tratamento de emenda (c), com aplicação de *blending* por média ponderada simples (d), a máscara de binária (e) e a imagem resultante pelo método proposto (f).

6. Conclusão

Pôde-se observar, através de um resultado preliminar, que o uso da análise de múltiplas resoluções no processo de mosaico de imagens é promissor. 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. O método de *blending* desenvolvido neste trabalho será integrado ao sistema *REGEMY*.

Referências

- Burt, P. J.; Adelson, E. H. (1983), “Multiresolution spline with application to image mosaics”, *ACM Transaction on Graphics*, v. 2, n. 4, p. 217-236.
- Efros, A. A.; Freeman, W. T. (2001), “Image quilting for texture synthesis and transfer”, *In: Proceedings of SIGGRAPH 2001*. p. 341-346.
- Fedorov, D. (2002), “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.
- Fonseca, L. M. G. (1999), “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.
- Hsu, C. T.; Wu, J. L. (1996), “Multiresolution mosaic”, *IEEE Transactions Consumer Electron*, v. 42, p. 981-990.
- Meunier, L.; Borgmann, M. (2000), “High resolution panoramas using image mosaicing”, Stanford University, EE368 (Digital Image Processing) Final Project, Disponível em: <www.moriborg.de/ee/mosaicing/>. Acesso em 31 de janeiro de 2005.
- Kwatra, V.; Schodl, A.; Essa, I.; Turk, G.; Bobick, A. (2003), “Graphcut textures: image and video synthesis using graph cuts”, *In: Proc. ACM Transactions on Graphics, SIGGRAPH 2003*.
- Mallat, S. G. (1989), “A theory for multiresolution signal decomposition: the wavelet representation”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 11, p. 674-693.
- Su, M. S.; Hwang, W. L.; Cheng, K. Y. (2004), “Analysis on multiresolution mosaic images”, *IEEE Transactions on Image Processing*, v. 13, n. 7.
- Wang, Z.; Bovik, A. C. (2002), “A universal image quality index”, *IEEE Signal Processing Letters*, v. 9, n. 3, p. 81-84.