

Avaliação de métodos de casamento de imagens para mosaico de imagens orbitais

Emiliano Ferreira Castejon ¹
Carlos Henrique Quartucci Forster ²
Leila Maria Garcia Fonseca ¹
Etoe Marcari Junior ¹

¹ Instituto Nacional de Pesquisas Espaciais - INPE
12227-010 - São José dos Campos - SP, Brasil
{castejon,leila,etore}@dpi.inpe.br

² Instituto Tecnológico da Aeronáutica - ITA
12228-900 - São Jose dos Campos, SP – Brasil
forster@ita.br

Abstract. The increasing number of satellite images has reinforced the need for automatic image mosaicking and registration methods. Several techniques for automatic image registration have been developed - Fonseca and Manjunath et al (1996); Zitová and Flusser (2003). One of the phases in the registration and mosaicking processes is the image matching that uses similarity measures to identify control points in the images. Similarity measures such as correlation, means squared difference and morphological have been used to register remote sensing images. However, these measures are more adequate to register images with similar spectral responses. Some authors have proposed the use of a metric based on mutual information to register images acquired in different time and different sensor. In order to evaluate some similarity measures, this paper describes a comparative study and a practical evaluation of three matching metrics based on: Correlation, Mutual Information (MI) and Scale Invariant Feature Transform (SIFT). Real situations problems such as the registration of temporal, spectral and cloud coverage images are analyzed and tested. The results showed that the method based on MI got the best performance in spite of its high computational cost.

Palavras-chave: remote sensing, image processing, register, mosaic, sensoriamento remoto, processamento de imagens, registro, mosaico.

1. Introdução

Registro de imagens é um processo comumente utilizado em sensoriamento remoto que tem como objetivo efetuar correções geométricas nas imagens de forma a tornar possível a sua utilização em conjunto com outras imagens provenientes de sensores diferentes ou adquiridas em diferentes períodos no tempo.

A operação de mosaico de imagens tem a finalidade de juntar duas ou mais imagens para gerar uma imagem maior. Este processo torna-se inevitável quando uma área geográfica de estudo é maior do que a cena disponibilizada pelo sensor. Desta forma, através do processo de mosaico é possível obter uma abrangência total do espaço geográfico de interesse – Bagli (2004).

Os processos de mosaico e registro automático envolvem, basicamente, três fases: identificação de feições candidatas a pontos de controle, casamento dos pontos de controle e transformação geométrica para corrigir a imagem de ajuste em relação a uma imagem de referência. Para realizar o casamento de feições diversas medidas de similaridade podem ser utilizadas - Fonseca e Manjunath (1996). Uma das métricas mais utilizadas é o índice de correlação, que tem mostrado desempenho superior aos métodos baseados em morfologia matemática e diferença de níveis de cinza - Fonseca et al (2006). Recentemente, o método baseado na informação mútua (*mutual-information*) têm sido testado para imagens médicas e de sensoriamento remoto - H. et al. (2003) e Viola e Wells (1997).

Portanto, este trabalho tem como objetivo analisar o desempenho de três critérios de similaridade em termos do número de acertos e erros no casamento das feições: índice de correlação (IC), informação mútua (MI) e SIFT (*Scale Invariant Feature Transform*).

2. Processo de casamento de imagens

No método manual o processo de casamento consiste na seleção de pontos de controle pelo usuário. Os pontos de controle mais comuns são o cruzamento de rios e estradas, pontos salientes de uma curva ou borda, interseção de áreas, etc.

Para garantir a boa qualidade da correção geométrica, é necessário que um grande número de pontos de controle espalhados uniformemente sobre as imagens sejam identificados. É muito importante que cada ponto seja marcado com precisão nas duas imagens o que é um trabalho exaustivo e demorado. O registro manual fica sujeito a erros se o processo é feito sem experiência ou sem a devida atenção.

Os métodos semi-automáticos ou automáticos oferecem um conjunto de ferramentas (algoritmos e recursos computacionais), que facilita e torna mais rápida a operação de identificação de pontos de controle - Fonseca e Manjunath (1996); Fedorov et al (2002). Neste tipo de sistema são usados algoritmos de casamento, que usam medidas de similaridade para identificar os pontos de controle.

A grande variação entre os sistemas de imageamento orbitais existentes, cada um com suas características específicas, pode gerar dificuldades para a correlação automática de áreas. O problema é que estes algoritmos podem possuir uma taxa de acertos baixa e prejudicar todo o processo de registro ou mosaico de imagens. Desta forma, é imprescindível que a fase de casamento dos pontos de controle seja robusta o suficiente para gerar um número suficiente de pontos de controle corretos e gerar corretamente a função de transformação geométrica - Machado (1989).

No processo automático de identificação de pontos de controle, a fase de extração de feições na imagem, que são candidatas a pontos de controle é realizada previamente. Técnicas de detecção de bordas, transformada *wavelet* são usadas nesta fase - Fonseca et al (2006).

Ao final do processo de extração de feições têm-se dois conjuntos de feições, um para cada uma das imagens a serem registradas. Estes conjuntos de feições passam pelo processo de casamento que resulta em um conjunto de pontos de controle. De uma forma geral, para cada feição da imagem de referência o algoritmo de casamento procura o seu correspondente, usando medidas de similaridade, na imagem de ajuste.

Neste trabalho, vamos avaliar medidas de similaridade baseadas nos níveis de cinza da imagem. Janelas de níveis de cinza imagem de referência, centradas nas feições extraídas em uma fase anterior, são comparadas com janelas na imagem de ajuste. Assim, os melhores métodos são aqueles que possuem certa independência em relação às variações que possam ocorrer nas imagens a serem processadas, o que aumenta a sua robustez.

2.1 - Métricas de similaridade

A qualidade dos pontos de controle gerados é dependente da métrica e das características das imagens utilizadas. Em situações reais não é trivial encontrar automaticamente bons pontos de controle, pois vários fatores podem influenciar negativamente no procedimento: diferenças espectrais entre as imagens, diferença sazonal, diferença temporal, cobertura de nuvens, problemas de oclusão, etc.

A seguir apresentamos a descrição dos critérios de similaridade avaliados neste trabalho.

Medida de correlação

A medida de correlação consiste na busca do valor máximo da medida do coeficiente de correlação entre duas imagens – Pratt (1974). O coeficiente de correlação que retorna o valor em escala absoluta normalizado entre

Valores no intervalo [-1, 1] é definido por:

$$C(i, j) = \frac{\sum_{l=0}^{k-1} \sum_{m=0}^{L-1} (W_Z(l, m) - \mu_W)(S_{i,j}(l, m) - \mu_S)}{\sqrt{\sum_{l=0}^{k-1} \sum_{m=0}^{L-1} (W_Z(l, m) - \mu_W)^2 \sum_{l=0}^{k-1} \sum_{m=0}^{L-1} (S_{i,j}(l, m) - \mu_S)^2}} \quad (1)$$

Onde:

W_Z - janela da imagem de ajuste

$S_{i,j}$ - janela da imagem de referência

μ_W - média da janela W

μ_S - média da janela S

Esta técnica é mais adequada para registro de imagens com bandas espectrais similares.

Informação mútua

A medida de similaridade baseada na informação mútua é definida em função da entropia de cada imagem e da entropia conjunta entre as duas imagens – H (2003) e Viola e Wells (1997):

$$I(A, B) = H(A) + H(B) - H(A, B) \quad (2)$$

Onde $H(A)$ e $H(B)$ são os valores de entropia para as imagens A e B. $H(A, B)$ é o valor de entropia conjunta entre A e B, calculados da forma:

$$H(A) = \sum -p_A(a) \log[p_A(a)] \quad (3)$$

$$H(B) = \sum -p_B(b) \log[p_B(b)] \quad (4)$$

$$H(A, B) = \sum -p_{A,B}(a, b) \log[p_{A,B}(a, b)] \quad (5)$$

Onde:

p_A : Probabilidade marginal de ocorrências para a imagem A.

p_B : Probabilidade marginal de ocorrências para a imagem B.

$p_{A,B}$: Probabilidade marginal de ocorrências conjuntas para as imagens A e B.

Minimização de distância entre descritores locais SIFT (*Scale Invariant Feature Transform*)

Esta técnica é baseada no algoritmo SIFT descrito em Lowe (1999), que permite a detecção e extração de descritores locais razoavelmente estáveis e invariáveis a mudanças de iluminação – Belo (2006). Um descritor local usa a informação contextual da feição a ser analisada.

O algoritmo SIFT é composto por quatro passos: Detecção de extremos, localização de pontos chave, definição da orientação de cada ponto chave e geração de descritores locais para cada ponto chave. Para o estudo aqui descrito, a etapa de geração de descritores é a de maior relevância. Dada uma janela de imagem de dimensões $K \times K$ esta etapa pode ser realizada pelos seguintes passos:

a) Subdivide-se a janela em pequenas micro-janelas de tamanho $n \times n$ (com n , geralmente, de dimensão 4×4).

b) Para cada sub-janela é criado um histograma direcional que relaciona a orientação e magnitude dos pontos que a compõe. Para um ponto A_{ij} , a magnitude M_{ij} e a orientação R_{ij} são definidas por:

$$M_{ij} = \sqrt{(A_{i,j} - A_{i+1,j})^2 + (A_{i,j} - A_{i,j+1})^2} \quad (6)$$

$$R_{ij} = a \tan 2(A_{i,j} - A_{i+1,j}, A_{i,j+1} - A_{i,j}) \quad (7)$$

c) O descritor local da janela $K \times K$ é então composto pela concatenação dos histogramas direcionais de cada uma das suas micro-janelas.

3. Metodologia

Para analisar o comportamento dos três métodos citados anteriormente, 10 feições pontuais centradas em janelas, extraídas das imagens de referência, são processadas pelos algoritmos de casamento. As feições foram escolhidas de forma a representar os problemas (sazonalidade, temporalidade, etc) mais complicados no processo de casamento. Desta forma, para um par de imagens referência e ajuste, respectivamente, K e S , 10 *kernels*/janelas (figura 1) são manualmente selecionados da imagem K . A escolha destes *kernels* é feita de modo que todos estejam centrados em feições pontuais visualmente expressivas (cruzamentos, edificações, etc.).

O objetivo é determinar a localização de cada *kernel* dentro das 10 regiões de busca na imagem S , mostradas na figura 2. O tamanho das janelas de *kernel* e busca usados nos testes, foi 128×128 e 1000×1000 *pixels*, respectivamente. Cada um dos *kernels* é submetido a uma busca usando as três medidas de similaridade – correlação, informação mútua e SIFT. A busca é realizada por meio de uma busca exaustiva pelo “deslizamento” do *kernel* sobre a área de busca na imagem S . Dentro da área de busca, todos os pontos centrados em uma janela da imagem são testados.

Para eliminar outras possíveis deformações geométricas e/ou de posicionamento geográfico, cada par de imagens selecionado foi inicialmente co-registrado utilizando pontos de controle manualmente selecionados e o modelo determinado por uma transformação afim. Desta forma é possível utilizar o critério do posicionamento geográfico como parâmetro de avaliação que determina se a posição encontrada na imagem de busca S é correta.

Os pares de imagens foram selecionados no catálogo CBERS, disponível na *internet*. As imagens selecionadas cobrem diferentes tipos de alvos e assim são bons exemplos para testar algoritmos de casamento de imagens. Para avaliar os métodos, alguns critérios foram definidos para levar em conta os principais problemas encontrados no processo de registro de imagens baseado em área:

a) Teste 1 – Acurácia – mede o desempenho geral do método quando a área de busca é proporcionalmente muito maior do que o *kernel*. Imagem utilizada: CBERS_2_CCD1XS_20040920_153_126_L2_BAND4.tif (20/09/2004)

b) Teste 2 – Sazonalidade – analisa o comportamento dos métodos em imagens adquiridas em estações do ano diferentes (diferenças radiométricas). Imagens utilizadas:

CBERS_2_CCD1XS_20040130_153_126_L2_BAND4.tif (30/01/2004)

CBERS_2_CCD1XS_20040920_153_126_L2_BAND4.tif (20/09/2004)

c) Teste 3 – Temporalidade – analisa o efeito das diferenças causadas pela diferença de tempo de aquisição entre as duas imagens. Imagens utilizadas:

CBERS_2_CCD1XS_20040704_153_126_L2_BAND4.tif (04/07/2008)

CBERS_2_CCD1XS_20080610_153_126_L2_BAND4.tif (10/07/2008)

d) Teste 4 – Bandas diferentes – analisa a sensibilidade do algoritmo quando as bandas são diferentes. Imagens utilizadas:

CBERS_2_CCD1XS_20040920_153_126_L2_BAND2.tif (20/09/2004)

CBERS_2_CCD1XS_20040920_153_126_L2_BAND4.tif (20/09/2004)

e) Teste 5 – Cobertura de nuvens – analisa o desempenho do algoritmo em imagens com diferentes coberturas de nuvens. Imagens utilizadas:

CBERS_2_CCD1XS_20070126_153_126_L2_BAND4.tif (26/01/2007)

CBERS_2_CCD1XS_20070221_153_126_L2_BAND4.tif (20/09/2004)

A implementação dos três métodos foi realizada usando o suporte para tratamento de dados matriciais geográficos provido pela biblioteca TerraLib (<http://www.terralib.org>), desenvolvida pela Divisão de Processamento de Imagens – DPI – no Instituto de Pesquisas Espaciais – INPE – Brasil.

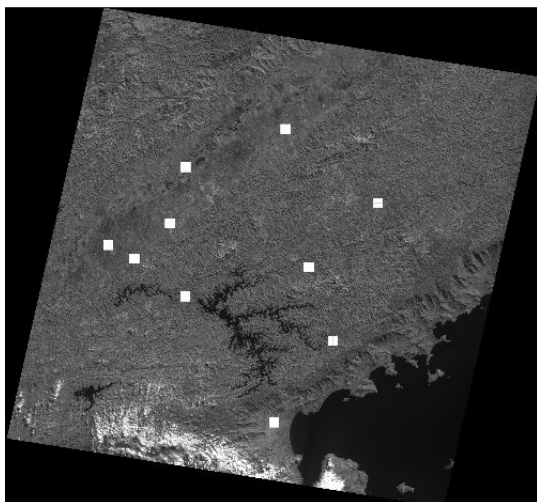


Figura 1 – Exemplos de kernels.

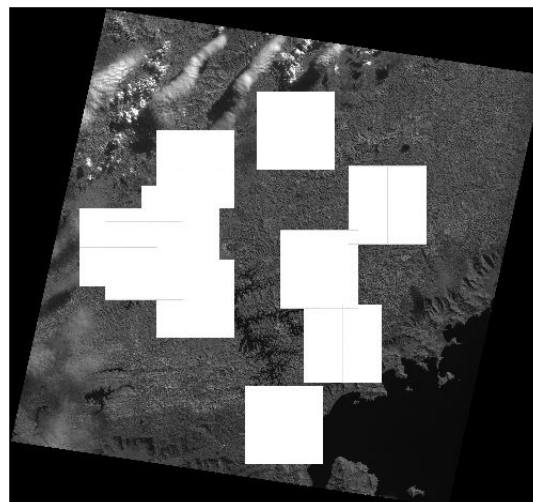


Figura 2 – Exemplos de áreas de busca.

5. Resultados

Como o processo de casamento de imagens por busca exaustiva tem elevado custo computacional usamos *kernels* de dimensões 128x128 pontos e áreas de busca com dimensões 1000x1000 pontos (áreas de busca centradas na posição estimada do *kernel* na imagem de busca). Os acertos e erros são dados pelo número de casamentos certos e errados, respectivamente, dentre os 10 *kernels* selecionados. Para cada *kernel*, busca-se uma janela de mesmo tamanho, na imagem de ajuste, que case com ele. O casamento é realizado pegando-se os pares de janelas cuja medida de similaridade é máxima. No processo de casamento, pode ocorrer dos pares casados serem errados. Consideramos correto o casamento quando a

distância máxima entre o centro do *kernel* e a janela na imagem de ajuste for de 21 metros (aproximadamente 1 *pixel*) de deslocamento nos eixos X e Y.

Os resultados obtidos são mostrados nas figuras 3 a 7 que indicam o número de *kernels* localizados corretamente (ou incorretamente) para cada método. A figura 8 mostra o tempo gasto na execução dos três métodos, avaliado para cada um dos testes. As unidades de tempo utilizadas são dadas em segundos.

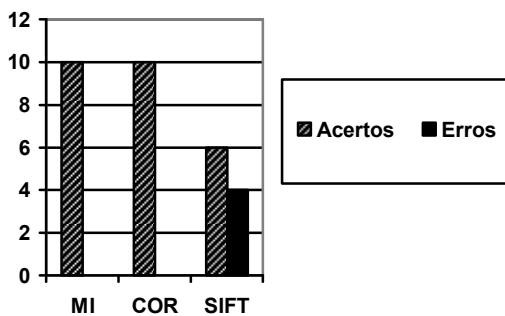


Figura 3 – Teste 1 – Acurácia.

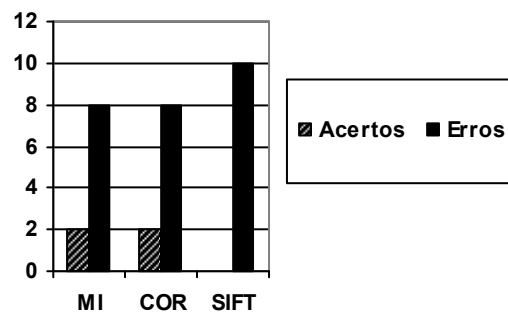


Figura 4 – Teste 2 – Sazonalidade.

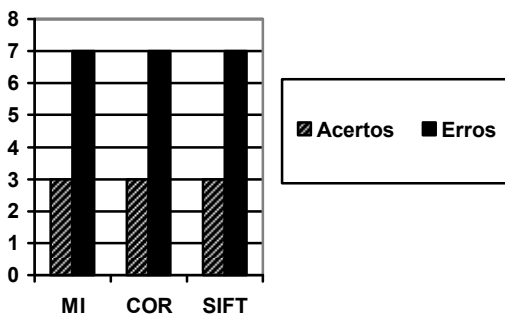


Figura 5 – Teste 3 – Temporalidade.

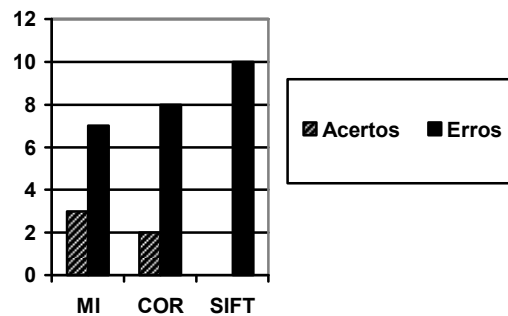


Figura 6 – Teste 4 – Bandas diferentes.

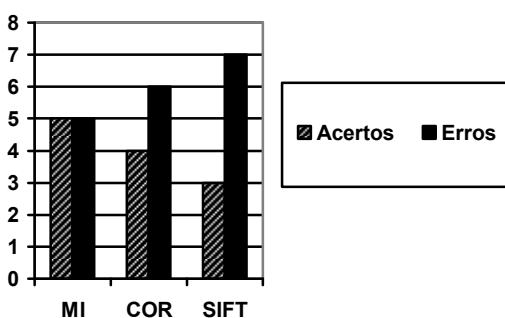


Figura 7 – Teste 5 – Cobertura de nuvens

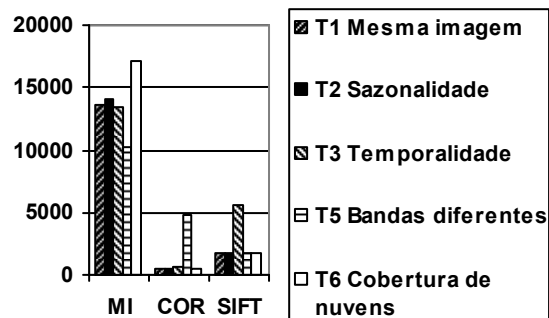


Figura 8 – Tempo de processamento

Pelas figuras 1-7, pode-se observar que, de uma forma geral a taxa de erros foi muito alta, com exceção do Teste 1. É importante observar que em um algoritmo de casamento eficiente, os casamentos errados que são chamados de *outliers* são removidos por meio de técnicas de verificação de consistência – Fonseca (1999). Nos testes realizados neste trabalho esta fase não foi implementada. De uma forma geral, pode-se observar, que o método baseado em MI

teve um desempenho um pouco melhor do que a correlação, nos casos de imagens com nuvens e bandas diferentes.

6. Conclusões

Para os testes realizados o método que demonstrou o melhor resultado foi o baseado na informação mútua. Este método, em seus piores casos se equipara ao método de correlação. O método baseado em SIFT mostrou o pior resultado, em todos os testes. É importante observar que para o teste de temporalidade o resultado dos três métodos é semelhante.

Testes adicionais, com *kernels* com dimensões de 40x40 pontos também foram feitos. Este tamanho de *kernel* é o mais comumente utilizado em uma boa parte de algoritmos que se baseiam em informações locais para descrever vetores de características de feições pontuais. Entretanto o desempenho global encontrado foi pior do que para *kernels* de 128x128 pixels, pois existe uma maior confusão devido ao aumento na relação entre a dimensão da área de busca e a área do *kernel*. Uma informação importante é a de que a maior queda no desempenho foi observado para o método SIFT, possivelmente pela própria construção do método. Neste método, quanto menor o tamanho do *kernel*, menor é o número de histogramas direcionais que representa a feição, aumentando a possibilidade de ocorrer casamentos incorretos.

Embora, o método MI tenha obtido os melhores resultados é importante observar que ele exige maior recurso computacional, como mostra a figura 8.

7. Referências

Fonseca, L. M. G.; Fedorov, D.; Manjunath, B. S.; Kenney, C.; Castejon, E.; Medeiros, J. S. Automatic Registration and Mosaicking System for Remotely Sensed Imagery. Revista Brasileira de Cartografia, n. 58/01, p. 49-61, 2006.

Fonseca L. M. G, Manjunath B. S. Registration techniques for multisensor remotely sensed imagery. PE&RS, Vol. 62, No. 9, pp. 1049-1056, Set. 1996.

Fonseca, L. M. G.; Registro automático de imagens de sensoriamento remoto. São José dos Campos - SP. 219p. (INPE - 7502-TDI/717). Tese de doutorado em Computação Aplicada - Instituto Nacional de Pesquisas Espaciais, 1999.

Machado e Silva, A. J. F. Modelos de Correção Geométrica para imagens HRV-SPOT. 1989. p. 239 (INPE-5001-TDL/395). Dissertação de Mestrado em Computação Aplicada - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 1989.

Manjunath B. S., Shekhar C., and Chellappa R., A new approach to image feature detection with applications, Pattern Recognition, vol. 29 (4), Abril/1996, pp. 627-640.

Fedorov, D., Fonseca L. M.G., Kenney C., Manjunath B.S. System for Automatic Registration of Remote Sensing Images. International Geoscience and Remote Sensing Symposium, Toronto, Canada, Junho/2002.

Pratt, W. K. Correlation techniques of image registration. IEEE Transactions on Aerospace and Electronic Systems. v.10 n. 3, p. 353-358, Maio/1974.

H. -m. Chen; M. K. Arora; P. K. Varshney. Mutual information-based image registration for remote sensing data. International Journal of Remote Sensing, Vol. 24, Iss. 18, p. 3701 – 3706, Setembro/2003.

Viola, P.; Wells III, W. Alignment by Maximization of Mutual Information. *International Journal of Computer Vision*. Vol. 24, Num. 2, Setembro/1997.

Lowe, D. G., Object Recognition from Local Scale-Invariant Features. *International Conference of Compute Vision*, p.1150, Vol. 2, 1999.

Belo, F. A. W., Desenvolvimento de Algoritmos de Exploração e Mapeamento Visual para Robôs Móveis de Baixo Custo. Dissertação de Mestrado. Departamento de Engenharia Elétrica da PUC-Rio. Abril de 2006.

Zitová, B.; Flusser, J. Image registration methods: a survey. *Vision and Image Computing*, Vol. 21, Iss. 11, p. 977-1000, 2003.

Bagli, V. V. Mosaico de imagens de sensoriamento remoto. Dissertação (Mestrado em Computação Aplicada) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2004.