

Correspondência de Pontos em Imagens de Satélite

Gustavo Augusto Mascarenhas Goltz^{1,2}, Elcio Hideiti Shiguemori²

¹Instituto Nacional de Pesquisas Espaciais, INPE – São José dos Campos – SP – Brazil

²Instituto de Estudos Avançados, IEAv - São José dos Campos – SP – Brazil

{gustavo, elcio}@ieav.cta.br

Abstract. *In this work the SIFT algorithm – Scale Invariant Feature Transform – was used to generate local descriptors of satellite images, in order to find corresponding points between pairs of images, with each pair of images representing the same scene.*

Resumo. *Neste trabalho o algoritmo SIFT – Scale Invariant Feature Transform – foi utilizado para gerar descritores locais de imagens de satélite, com o objetivo de encontrar pontos correspondentes entre pares de imagens, sendo que cada par de imagens representa uma mesma cena.*

1. Introdução

A busca automática por pontos correspondentes entre duas imagens que representam a mesma cena tem grande utilidade em visão estéreo, navegação autônoma, restreio de alvos, entre outras aplicações [Belo, 2006]. O conhecimento de pontos correspondentes entre bandas espectrais diferentes que representam a mesma região terrestre auxilia em tarefas de interpretação automática de imagens, na identificação de características de interesse [Gonzalez e Woods, 2000].

A eficiência do uso de descritores gerados pelo algoritmo SIFT [Lowe, 2004] foi constatada em trabalhos anteriores [Goltz e Shiguemori, 2008] aplicando-o na busca de pontos correspondentes em imagens aéreas e imagens de ambientes internos. A idéia deste trabalho foi verificar o funcionamento da técnica anterior em imagens de satélite, em situações pré-estabelecidas, utilizando uma implementação do SIFT em C++, implementada por [Vedaldi, 2005]. Abaixo o algoritmo SIFT é descrito resumidamente, evidenciando seus pontos principais.

2. SIFT

O algoritmo SIFT - *Scale Invariant Feature Transform* - foi desenvolvido por David Lowe em 1999. Este algoritmo detecta e constrói descritores locais de uma dada imagem, sendo este descritor relativamente invariante a transformações de escala, rotação, iluminação e baixa variação de perspectiva.

De forma geral, pode-se dividir este algoritmo em 4 etapas principais [Belo, 2006]:

- Detecção de extremos;

- Localização de pontos-chave;
- Definição de Orientação;
- Descritor dos pontos-chave.

A detecção de extremos se faz ao passar um filtro gaussiano na imagem e subtraindo a imagem filtrada da original. Então filtra-se a imagem original mudando o parâmetro σ do filtro gaussiano e subtraindo o resultado da primeira filtragem. Isso se repete para vários valores de σ , múltiplos do valor de σ inicial. O filtro produz imagens com escalas diferentes, e a subtração identifica os pontos extremos da imagem, ou seja, suas bordas.

A localização de pontos-chave se faz analisando o gradiente do *pixel* em relação a vizinhança-8 da imagem filtrada usando σ' , mais as variações desta mesma região das imagens filtradas com σ e σ'' , sendo $\sigma < \sigma' < \sigma''$ (aqui pode-se imaginar um cubo de aresta igual a 3 *pixels*, com o ponto que está sendo analisado no centro deste cubo). Pontos com menor variação são eleitos como pontos-chave, entre aqueles pontos identificados no processo de detecção de extremos.

Com o objetivo de se construir um descritor independente de rotação, atribui-se uma orientação ao ponto-chave através da utilização de histogramas da região do ponto-chave, e também calcula-se a magnitude deste ponto.

O descritor de cada ponto-chave é construído com base na região vizinha deste ponto, dada pela imagem filtrada a qual este ponto pertence. É construído um histograma para oito direções diferentes usando a magnitude de cada pixel, sendo esta magnitude atenuada por uma função gaussiana. A normalização do descritor deixa-o invariante à iluminação.

Ao se calcular dois conjuntos de descritores, cada um pertencente a uma imagem, sendo que as duas imagens representam a mesma cena, pode-se encontrar pontos correspondentes entre as duas imagens utilizando estes descritores. Este processo é explicado a seguir.

3. Geração de Correspondências

Como os descritores gerados pelo SIFT são vetores, a distância euclidiana é utilizada para encontrar pontos correspondentes, pois dois pontos iguais a princípio devem possuir descritores iguais, ou no mínimo descritores muito semelhantes.

Dadas duas imagens i_1 e i_2 diferentes, que representam uma mesma cena, obtém-se com o SIFT duas matrizes M_1 e M_2 , que contêm descritores de pontos-chave de i_1 e i_2 , respectivamente. Para cada descritor de M_1 , calcula-se a distância euclidiana com todos os descritores de M_2 . Para cada descritor de M_1 , executam-se as operações abaixo:

d_1 = menor distância euclidiana encontrada.

d_2 = segunda menor distância euclidiana encontrada.

Se $((d_1 \times \text{constante}) < d_2)$

Então a correspondência encontrada é verdadeira.

Senão a correspondência encontrada é falsa .

A utilização do teste acima elimina um grande número de correspondências falsas, pois quando um vetor é relativamente muito semelhante a outro, a distância euclidiana entre ambos é bem menor se comparada a distância com outros vetores (tão menor a ponto de continuar sendo “menor” se multiplicada por um escalar maior que 1).

Após este processamento, o algoritmo acima é repetido trocando-se M_1 por M_2 e vice-versa. Como as correspondências verdadeiras devem ser encontradas em ambos os casos, esta é mais uma forma de eliminar falsas correspondências, tomando somente aquelas que ocorreram em ambos os processamentos.

Na sequência têm-se as regras criadas para avaliar os algoritmos descritos anteriormente (SIFT e algoritmo que busca pontos correspondentes).

4. Metodologia

As imagens utilizadas foram obtidas com o sensor Quickboard em 4 de maio de 2004, referente à cidade de São José dos Campos, São Paulo. Selecionou-se 4 cenas diferentes, sendo que cada uma foi representada em imagens com as bandas 1, 2, 3 e 4, e pancromática. Das 20 amostras de imagens (4 cenas vezes 4 bandas, mais 4 cenas pancromáticas) geraram-se os descritores de cada uma pelo algoritmo SIFT. Para cada cena, a busca por correspondências fez-se aos pares de imagens:

banda 1 – banda 2; banda 1 – pancromat.; banda 2 – pancromat.; banda 4 – pancromat..

banda 1 – banda 3; banda 2 – banda 3; banda 3 – banda 4;

banda 1 – banda 4; banda 2 – banda 4; banda 3 – pancromat.;

Dado o par de imagens, comparou-se através da distância euclidiana seus descritores, encontrando os pontos correspondentes entre as duas imagens.

5. Resultados

Nas quatro cenas analisadas, encontrou-se grande número de pontos correspondentes (variando de 16 a 77 correspondências, com 1 erro) entre as bandas 1 e 2. A figura 1 representa a cena 1 entre as bandas 1 e 2, contendo alguns pontos correspondentes encontrados. O par banda 1 e 3 também obteve um grande índice de acertos. O mesmo ocorreu entre as bandas 2 e 3.



Figura 1. Alguns pontos correspondentes entre as bandas 1 e 2, para a cena 1

Algumas correspondências encontradas entre as bandas 1 e 2 estão representadas na figura 2. O método de busca por pontos correspondentes não obteve êxito entre as bandas 1 e 4, sendo que das 4 cenas, a primeira gerou duas correspondências erradas e nas restantes não foi encontrada correspondência. Resultado semelhante ocorreu entre o par de bandas 2 e 4, e o par banda 1 e pancromática.

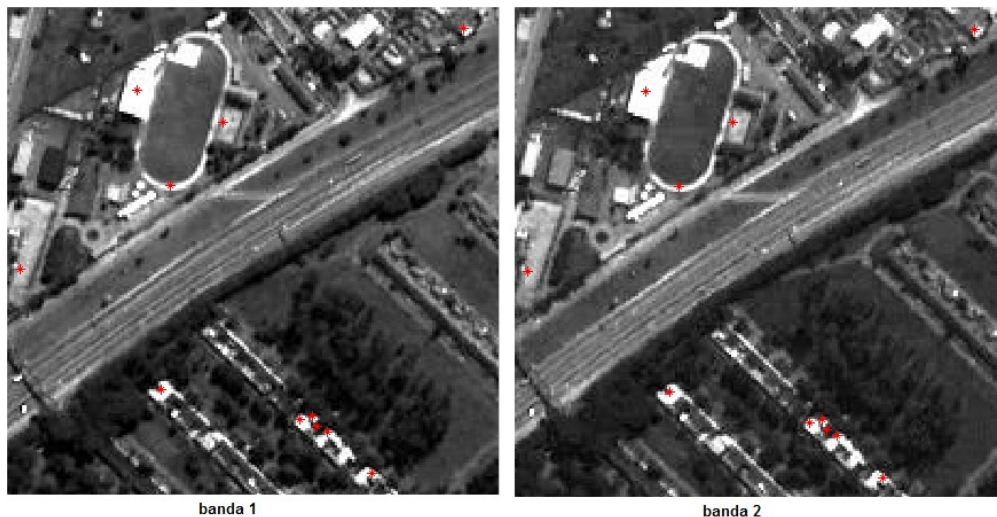


Figura 2. Alguns pontos correspondentes entre as bandas 1 e 2, para a cena 2

Os descritores da imagem pancromática tiveram um número maior de correspondência com a banda 4 (como visto na figura 3), seguido da banda 3.

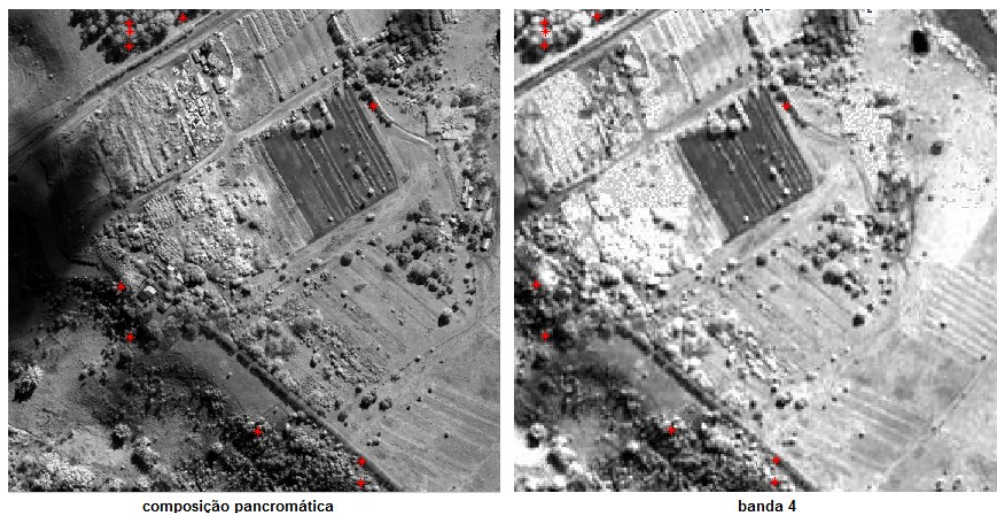


Figura 3. Alguns pontos correspondentes entre a composição pancromática e a banda 4, para a cena 3

Todos os resultados estão representados nas tabelas 1, 2, 3 e 4, contendo os dados das cenas 1, 2, 3 e 4 respectivamente. Algumas correspondências encontradas entre as bandas 3 e 4 na cena 4 estão apresentadas na figura 4, onde nota-se um erro do método aplicado, o qual afirmou que o ponto mais acima da imagem à direita se corresponde ao ponto mais acima da imagem à esquerda.



Figura 4. Alguns pontos correspondentes entre as bandas 3 e 4, para a cena 4

Tabela 1. Correspondências encontradas para a cena 1

Pares de imagens:	Número de pares de pontos correspondentes:	Pares de pontos corretos:	Pares de pontos errados:	% de Acerto:
Banda 1 - Banda 2	16	16	0	100
Banda 1 - Banda 3	8	8	0	100
Banda 1 - Banda 4	2	0	2	0
Banda 1 - Pancrom.	3	1	2	66,67
Banda 2 - Banda 3	25	24	1	96
Banda 2 - Banda 4	1	0	1	0
Banda 2 - Pancrom.	0	0	0	-
Banda 3 - Banda 4	2	2	0	100
Banda 3 - Pancrom.	3	3	0	100
Banda 4 - Pancrom.	13	11	2	84,62

Tabela 2. Correspondências encontradas para a cena 2

Pares de imagens:	Número de pares de pontos correspondentes:	Pares de pontos corretos:	Pares de pontos errados:	% de Acerto:
Banda 1 - Banda 2	77	76	1	98,70
Banda 1 - Banda 3	61	60	1	98,36
Banda 1 - Banda 4	1	0	1	0
Banda 1 - Pancrom.	0	0	0	-
Banda 2 - Banda 3	81	79	2	97,53
Banda 2 - Banda 4	0	0	0	-
Banda 2 - Pancrom.	1	0	1	0
Banda 3 - Banda 4	3	3	0	100
Banda 3 - Pancrom.	4	4	0	100
Banda 4 - Pancrom.	44	43	1	97,73

Tabela 3. Correspondências encontradas para a cena 3

Pares de imagens:	Número de pares de pontos correspondentes:	Pares de pontos corretos:	Pares de pontos errados:	% de Acerto:
Banda 1 - Banda 2	32	32	0	100
Banda 1 - Banda 3	25	25	0	100
Banda 1 - Banda 4	0	0	0	-
Banda 1 - Pancrom.	0	0	0	-
Banda 2 - Banda 3	23	22	1	95,65
Banda 2 - Banda 4	0	0	0	-

Banda 2 - Pancrom.	0	0	0	-
Banda 3 - Banda 4	0	0	0	-
Banda 3 - Pancrom.	0	0	0	-
Banda 4 - Pancrom.	69	69	0	100

Tabela 4. Correspondências encontradas para a cena 4

Pares de imagens:	Número de pares de pontos correspondentes:	Pares de pontos corretos:	Pares de pontos errados:	% de Acerto:
Banda 1 - Banda 2	25	25	0	100
Banda 1 - Banda 3	7	7	0	100
Banda 1 - Banda 4	0	0	0	-
Banda 1 - Pancrom.	0	0	0	-
Banda 2 - Banda 3	32	32	0	100
Banda 2 - Banda 4	0	0	0	-
Banda 2 - Pancrom.	3	2	1	66,67
Banda 3 - Banda 4	9	8	1	88,89
Banda 3 - Pancrom.	21	21	0	100
Banda 4 - Pancrom.	11	11	0	100

6. Conclusões

Nota-se que um grande número de pontos correspondentes foi encontrado entre os pares de bandas 1 e 2, bandas 1 e 3, bandas 2 e 3, e banda 4 com a imagem pancromática. Pode-se inferir que estes pares de imagens possuem uma correlação maior comparado aos outros, visto que o algoritmo SIFT tem uma dependência grande com o domínio espacial, onde o valor do pixel está intimamente ligado à resposta espectral das cenas para cada banda analisada. Seguindo este raciocínio, os pares de imagens com pouca ou nenhuma correspondência encontrada possuem baixa correlação.

Os resultados também são dependentes do sensor utilizado, pois a correlação entre bandas pode variar de um sensor a outro. Testes com sensores diferentes contribuirão para uma inferência mais exata. Futuramente também pode-se buscar a correspondência entre imagens de sensores diferentes, como também pode-se fazer um pré-processamento nas imagens antes de se gerar seus descritores, a fim de tratar ruído, restaurar, degradar, entre outras operações possíveis, para uma dada finalidade.

7. Referências

- Belo, F. A. W. (2006). "Desenvolvimento de Algoritmos de Exploração e Mapeamento Visual para Robôs Móveis de Baixo Custo, Cap. 3: Transformação SIFT", <http://www.ic.unicamp.br/~siome/teaching/2007/mo446-0207>, Outubro de 2007.
- Gonzalez, R. C. e Woods, R. E. (2000). "Processamento de Imagens Digitais, Cap. 9: Reconhecimento e Interpretação, p. 422-423", Editora Edgard Blücher, São Paulo, 1ª Edição.
- Lowe, David G. (2004). "Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60, 2 (2004), pp. 91-110", <http://www.cs.ubc.ca/spider/lowe/pubs.html>, Outubro de 2007.
- Goltz, G. A. M. e Shiguemori, E. H. (2008). "Aplicação do Algoritmo SIFT em Imagens de Navegação Autônoma", Atividades de Pesquisa e Desenvolvimento – Instituto de Estudos Avançados – IEAv, vol.1 (mar/2007-fev2008), São José dos Campos, 2008.
- Vedaldi, Andrea. (2005). "SIFT++: A Lightweight C++ implementation of SIFT", <http://vision.ucla.edu/~vedaldi/code/siftpp/siftpp.html>, Outubro de 2007.