

UTILIZAÇÃO DE SIFT NA CORRESPONDÊNCIA ENTRE IMAGENS AÉREAS E CENAS TERRESTRES DE CONTROLE

ADILSON BERVEGLIERI¹

ANTONIO MARIA GARCIA TOMMASELLI²

Univ Estadual Paulista - Unesp

Faculdade de Ciências e Tecnologia – FCT

¹Programa de Pós-Graduação em Ciências Cartográficas – Bolsista FAPESP DR

²Departamento de Cartografia, Presidente Prudente - SP

aberveg@yahoo.com.br, tomaseli@fct.unesp.br

RESUMO - No campo da Fotogrametria e Visão Computacional, a extração de características e a correspondência de imagens são técnicas frequentemente utilizadas. Entretanto, quando as diferenças geométricas e radiométricas são expressivas, o relacionamento correto entre duas imagens torna-se complicado. Para contornar esse problema, a técnica SIFT (*Scale Invariant Feature Transform*) utiliza características invariantes à escala e parcialmente invariantes à rotação, mudanças de iluminação e ponto de vista. Neste trabalho, realiza-se uma aplicação fotogramétrica que consiste em estabelecer correspondência multiescala entre imagens aéreas e terrestres nadirais, utilizando a técnica SIFT, de maneira que as cenas terrestres possam ser usadas para orientar imagens aéreas. Experimentos foram realizados e demonstram que é possível localizar automaticamente pelo menos um ponto em comum para estabelecer a correspondência das imagens. Por fim, apresentam-se algumas estratégias no sentido de aumentar a robustez da aplicação, para posteriormente possibilitar um refinamento subpixel.

Palavras-chave: Correspondência por características, Cenas de controle, Localização automática.

ABSTRACT – In the Photogrammetry and Computer Vision fields, feature extraction and image matching are techniques widely used. However, when geometric and radiometric variations are noticeable, the correct relationship between two images becomes troublesome. To solve this problem, the SIFT (*Scale Invariant Feature Transform*) technique uses features that are invariant to changes in scale and partially invariant to rotation, changes in illumination and view point. This paper performs a photogrammetric application based on multiscale image matching. By means of the SIFT technique, nadir terrestrial images could be used to compute the orientation of aerial images. Experiments were performed showing that it is possible to automatically locate at least one homologous point match the images. Finally, some strategies are presented in order to increase the robustness of the application and to enable a further step of subpixel refinement.

Keywords: Feature-based matching, Control scenes, Automatic location.

1 INTRODUÇÃO

A extração de características e a correspondência de elementos são tarefas essenciais em Fotogrametria e Visão Computacional, quando o interesse é localizar pontos similares em duas ou mais imagens. Porém, vários problemas, como distorções geométricas, variação de escala, perspectiva, iluminação, orientação etc., dificultam a determinação do relacionamento correto entre conjuntos de dados, principalmente no caso de cenas com grandes diferenças entre si, como na comparação de cenas terrestres com aéreas para a localização automática de pontos de controle.

Com a intenção de contornar tais problemas, pode-se usar o vetor descritor da técnica SIFT (*Scale Invariant*

Feature Transform), o qual armazena informações extraídas sobre gradientes locais (magnitude e orientação). Este possibilita localizar pontos comuns entre imagens com base em características invariantes, ou seja, que não dependem de condições geométricas e radiométricas, como ocorre com os métodos de correspondência por área (correlação, erro, *Least Squares Matching* etc.) e com os operadores de interesse de Förstner (FÖRSTNER, 1986) e Harris (HARRIS e STEPHENS, 1988) (LINGUA et al, 2009).

Além dessas características, segundo Lowe (1999), a técnica SIFT apresenta vantagens por ser invariante à escala e parcialmente invariante à rotação, às mudanças de iluminação e ao ponto de vista da câmara. Em razão disso, suas características tornam-se atraentes para a

Fotogrametria e podem ser usadas na correspondência fotogramétrica.

Ainda que o georreferenciamento direto permita a redução do espaço de busca pelos elementos de controle na imagem a poucos pixels, usar a técnica SIFT conjuntamente pode contribuir para a melhoria da acurácia nos processos de aerotriangulação, pois os descritores são distintos, robustos à oclusão e não exigem segmentação (TOTH et al, 2011), favorecendo as aplicações de correspondência ou reconhecimento de objetos. Nos casos em que o georreferenciamento da imagem aérea for de baixa qualidade, a técnica SIFT deve permitir a busca em um espaço de centenas de pixels.

Outra motivação para o uso da SIFT é que, apesar de ser amplamente utilizada na área de Visão Computacional, no contexto da Fotogrametria, para aplicações mais acuradas, o seu uso ainda é pequeno (LINGUA et al, 2009).

Nesse sentido, o trabalho objetiva estudar a viabilidade de localizar automaticamente cenas de controle (de alta resolução) em cenas não orientadas (de menor resolução), como, por exemplo, uma cena terrestre nadiral (coletada por uma câmara super-grande angular ou *fish-eye*) e uma imagem aérea de alta resolução ou, ainda, usar uma ortoimagem aérea de alta resolução para orientar uma imagem de satélite de baixa resolução. Em outros termos, o propósito é localizar uma cena sem depender de um georreferenciamento preciso da imagem a ser orientada, usando a técnica SIFT. Conseqüentemente, o espaço de busca da cena de controle na imagem não orientada pode ter seu tamanho ampliado sem comprometer o resultado.

Além disso, também se apresenta uma forma simples para coletar cenas de controle, a qual considera, na imagem terrestre, um gabarito orientado, próximo ao ponto de rastreamento com GPS.

Sobre tais cenas podem ser aplicados diversos métodos para a localização automática de pontos, os quais variam conforme o tipo de primitiva utilizada. Geralmente, conforme Heipke (1996), os métodos são classificados como: métodos baseados em áreas, que consideram correlações entre os tons de cinza dos pixels em subimagens, e métodos baseados em características, cujos atributos são extraídos e associados a elementos de interesse (pontos, retas, contornos, polígonos etc.), permitindo fazer relações entre imagens.

2 A TÉCNICA SIFT

A técnica de processamento de imagens SIFT detecta e extrai descritores locais de uma imagem. Estes devem ser encontrados em quantidades para cobrir densamente a cena e possibilitar o reconhecimento dos objetos. A geração dos descritores SIFT baseia-se em quatro etapas principais, segundo Lowe (2004): detecção dos extremos no espaço-escala; localização dos pontos-chave; atribuição da orientação e construção do descritor. Na sequência, essas etapas são apresentadas com um exemplo de cena terrestre.

2.1 Etapa 1: Detecção dos extremos

A primeira etapa, a detecção de extremos, consiste em encontrar máximos e mínimos locais por meio de uma filtragem em cascata com a função DoG (*Difference of Gaussian*), buscando as características que sejam mais estáveis. A DoG, Equação (1), é obtida pela subtração das imagens filtradas com escalas variadas por uma constante k , ou seja, é a diferença entre as imagens borradas por um filtro gaussiano nas escalas σ e $k\sigma$. Isto permite detectar variações de intensidades, como bordas.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \quad (1)$$

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma),$$

em que: $L(x, y, \sigma)$ é uma imagem gerada a partir da convolução (*) de um filtro gaussiano $G(x, y, \sigma)$ com a imagem de entrada $I(x, y)$.

Cada conjunto, composto pelas imagens da filtragem gaussiana e pelas resultantes das diferenças (DoG) entre elas, forma uma oitava, como exemplifica a Figura 1. Após, a imagem filtrada com σ tem seu tamanho reduzido à metade e serve de entrada para a próxima geração de oitavas, formando uma pirâmide. O número de oitavas não é fixo.

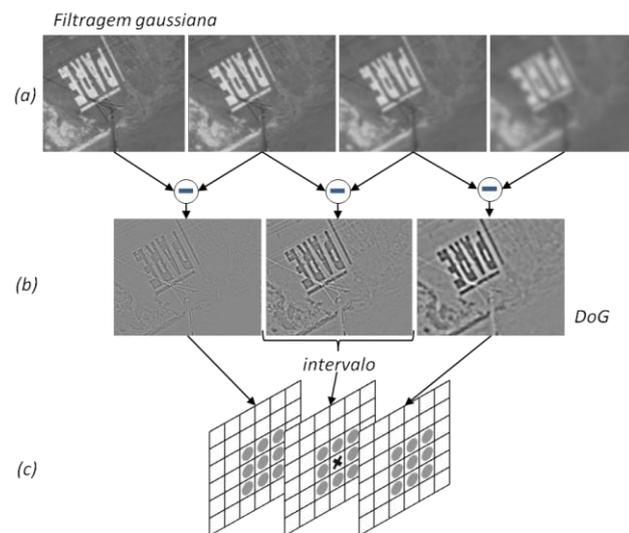


Figura 1: Exemplo de uma oitava – (a) Imagens filtradas; (b) Imagens DoGs, sendo $k = 2^{1/2}$ e $\sigma = 0,5$; (c) Um pixel com seus 26 vizinhos.

Depois da geração da pirâmide, usando-se a imagem do intervalo DoG, cada pixel é comparado com seus vizinhos da própria imagem e das adjacentes, sendo, neste caso, a vizinhança igual a 26. Então, faz-se a seleção dos pontos candidatos a serem pontos-chave.

2.2 Etapa 2: Localização dos pontos-chave

Nesta etapa, realizam-se as localizações precisas dos pontos-chave. Todos os pontos candidatos, detectados na etapa anterior, são ajustados para a localização, escala e razão das curvaturas principais, conforme Lowe (2004). Isto permite rejeitar os pontos de baixo contraste (logo, sensíveis a ruídos) ou aqueles mal localizados ao longo de uma aresta.

Para aumentar a qualidade da correspondência ajusta-se uma superfície de segunda ordem ao ponto de amostragem local, permitindo interpolar a localização do máximo. Esta abordagem usa uma expansão por série de Taylor (de grau 2) da função $D(x, y, \sigma)$, deslocada de modo que a origem seja o ponto de amostragem da região.

A partir do ponto sob análise, o valor da função no ponto e suas derivadas são avaliados. Com base na Equação (2), estima-se a localização do extremo \hat{x} :

$$\hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x} \quad (2)$$

Como resultado, tem-se um sistema linear 3×3 , solucionado por mínimos quadrados. Se $\hat{x} > 0,5$, indica que o extremo está mais próximo a outra região. Neste caso, realoca-se o ponto e uma nova interpolação é realizada. Ao final, o deslocamento de \hat{x} é adicionado à localização do ponto em análise, resultando no valor da função no extremo $D(\hat{x})$.

Com a finalidade de refinar os pontos de baixo contraste e aumentar a estabilidade, Lowe (2004) recomenda rejeitar os resultados inferiores a algum limiar previamente definido ($|D(\hat{x})| < \text{limiar}$). Mediante testes experimentais, Lowe (2004) sugere um limiar de 0,03 para pixel da imagem no intervalo [0, 1]. Como exemplo, a Figura 2(a) mostra o resultado quando não se define nenhum limiar de rejeição (204 pontos-chave), enquanto que a Figura 2(b) exibe o resultado após a aplicação do limiar de 0,03, gerando 41 pontos-chave.

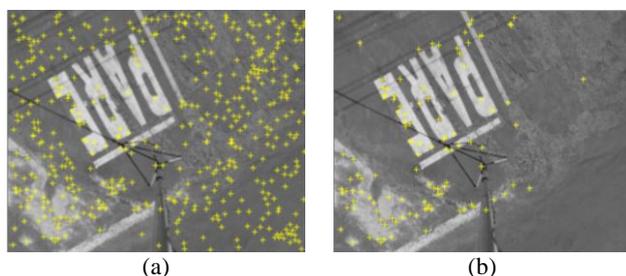


Figura 2: Imagens (215×215), considerando 3 oitavas – (a) Sem limiar de rejeição, encontrados 204 pontos-chave; (b) Com limiar 0,03, encontrados 41 pontos-chave.

Entretanto, fazer o refinamento por um limiar ainda não é suficiente, pois a função DoG possui altos valores de resposta ao longo das arestas, ficando sensível a pequenos ruídos.

Para solucionar esse problema, calcula-se a razão entre curvaturas, partindo do pressuposto que: um pico mal definido na função DoG forma uma grande curvatura principal do lado oposto da borda e uma pequena curvatura na direção perpendicular. Assim, obtêm-se tais curvaturas a partir de uma matriz Hessiana (H) 2×2 em cada localização e escala do ponto-chave.

Como os autovalores de H são proporcionais às curvaturas, basta calcular a razão entre o traço (soma dos autovalores) e o determinante (produto dos autovalores) de H .

Caso o determinante seja negativo, descarta-se o ponto sob análise. Para definir quais pontos serão pontos-chave, aplica-se um limiar na razão de curvaturas.

Até este passo, apenas foram selecionados os extremos. As duas próximas etapas referem-se à orientação desses pontos e à construção do descritor.

2.3 Etapa 3: Atribuição da orientação

A atribuição é baseada em propriedades locais de cada ponto-chave, o que possibilita alcançar invariância à rotação da imagem.

Nesse momento, seleciona-se a imagem filtrada L que tenha escala gaussiana σ mais próxima e de oitava referente ao ponto em análise. Isto garante que os cálculos sejam realizados em escala invariante. Em cada região, os pontos da imagem L têm seus gradientes calculados, obtendo-se a magnitude $m(x, y)$ e a orientação $\theta(x, y)$.

Constrói-se então um histograma de orientações para os pixels da região do ponto-chave. Lowe (2004) indica 36 valores para θ , abrangendo 360° de orientações, divididos em 10 intervalos. Aos pontos vizinhos são dados três pesos, sendo o primeiro, conforme uma função de distância normalizada entre a orientação dos pixels e a orientação do ponto-chave; o segundo, com base na magnitude $m(x, y)$ e o terceiro, usando uma janela gaussiana circular com o valor de σ 1,5 vezes maior do que a escala do ponto-chave. Com esses pesos, segundo Lowe (2004), o histograma é atualizado. Os máximos no histograma de orientação representam as direções dominantes dos gradientes locais. Além do pico do histograma, também são usados, para definir a orientação, outros picos com valor acima de 80% em relação ao maior. No final, ainda se aplica um ajuste parabólico aos três valores mais próximos de cada pico, a fim de interpolar a posição do máximo. Desse modo, pode-se construir o descritor.

2.4 Etapa 4: Construção do descritor local

As operações anteriores atribuíram, a cada oitava, pontos-chaves para localizações, escala e orientação. Calculam-se, então, com normalização em relação à orientação, os descritores que representam as regiões de cada ponto-chave cujos passos, segundo Lowe (2004), consistem em (ver Figura 3):

- Selecionar a imagem filtrada L de σ e a oitava referentes ao ponto-chave;
- Rotacionar as coordenadas do descritor e as orientações gradientes em relação à orientação do ponto-chave;
- Aplicar uma ponderação gaussiana com σ' igual à metade da janela do descritor para atribuir um peso à magnitude de cada ponto vizinho, dando menos ênfase aos gradientes mais afastados;
- Definir $k \times k$ pixels em cada $n \times n$ regiões em torno do ponto-chave;
- Montar, em cada região, um histograma para 8 direções com base nas magnitudes dos pixels;
- Construir o descritor a partir do vetor que armazena o histograma. Para o caso de um

conjunto 4×4 de histogramas com 8 células de acumuladores, gera-se um vetor de tamanho $4 \times 4 \times 8 = 128$, para cada ponto-chave;

- E, finalmente, normalizar o vetor para a unidade de comprimento e aplicar um limiar de ajuste, fazendo com que as direções mais acentuadas não prevaleçam na representação do descritor. Depois, normaliza-se o vetor novamente. Este passo trata da invariância à iluminação.

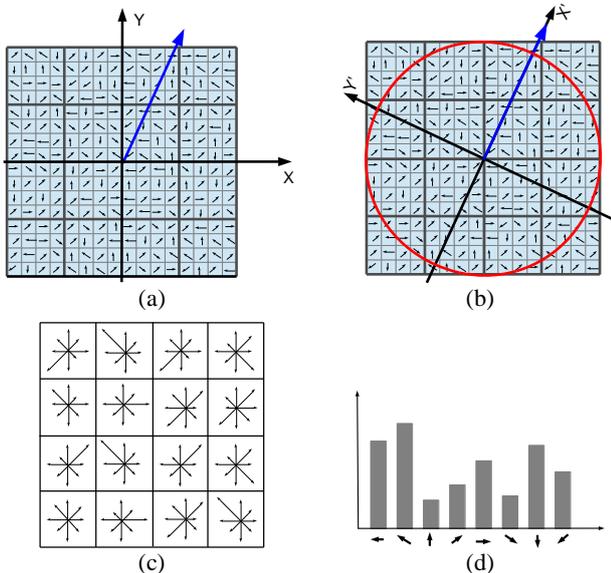


Figura 3: Construção de um descritor: (a) Gradiente do ponto-chave; (b) Rotação em relação ao ponto-chave e janela circular de ponderação gaussiana; (c) As magnitudes são acumuladas no histograma. Região 4×4 ($n = 4$ e $k = 4$); (d) Histograma de direções.

A Figura 4, que é uma cena terrestre nadiral, mostra 111 pontos-chave como vetores indicando escala, orientação e localização, construídos após os refinamentos e usando o vetor descritor de tamanho igual a 128.



Figura 4: Orientação, escala e localização de 111 pontos-chave.

Assim sendo, constrói-se o descritor para cada ponto-chave de uma imagem. Logo, como resultado, tem-se um conjunto de descritores, que podem ser usados na comparação com outras imagens para fins de correspondência.

3 MATERIAIS E MÉTODOS

A abordagem tem como finalidade estabelecer correspondência entre imagens multiescalas de diferentes

sensores ópticos. Neste caso, procura determinar elementos de correspondência entre cenas terrestres nadirais e imagens aéreas.

Para isso, foram utilizadas imagens aéreas adquiridas por uma câmara digital Hasselblad – 60 megapixels, apresentando um GSD (*Ground Sample Distance*) de aproximadamente 10 cm e um conjunto de cenas terrestres tomadas por uma câmara digital Fuji Finepix S3Pro com lente *fish-eye*, cujos GSDs variam entre 2 e 3 mm.

A forma de aquisição dessas cenas terrestres está retratada na Figura 5. A partir da imagem aérea, são identificados os locais adequados para a determinação do apoio de campo. O local deve conter algum elemento de cena que seja fotointerpretabil. As coordenadas de um ponto distinguível são determinadas por rastreamento GPS. Em seguida, a câmara com lente *fish-eye*, em posição nadiral, é levantada acerca de 4 m de altura por meio de um bastião apoiado em um tripé. São tomadas duas fotos.



Figura 5: Forma de aquisição das cenas terrestres georreferenciadas.

A primeira tem inserido na cena um gabarito em forma de 90° , nivelado em relação ao solo e com alvos pré-definidos. Sua função é servir como referência para o cálculo do GSD e também como eixos coordenados XYZ para a orientação da imagem. Por meio de uma bússola, uma das hastes é direcionada ao norte magnético e, assim, tem-se a informação de orientação aproximada.

Posteriormente, os instrumentos de medição são retirados da cena, e a imagem que se destina a fazer correspondência é adquirida, sem movimentar a câmara. Terminada as aquisições, pode-se calcular a orientação da imagem por meio do processo de ressecção espacial e retificar as imagens, corrigindo-se tanto a inclinação da câmara como o efeito de distorção provocado pela lente *fish-eye*. É importante ressaltar que os parâmetros de orientação interior da câmara com a lente *fish-eye* foram previamente determinados em um processo de calibração usando o modelo de Conrady-Brown, em um campo de calibração terrestre.

A partir das imagens retificadas, cada par de imagens (um *template* terrestre e um recorte da área de busca na imagem aérea) é montado, aplicando-se a técnica SIFT. Como resultado, adquire-se um conjunto de pontos-chave, os quais devem ser verificados, pois falsas correspondências podem ocorrer. Assim, objetivando minimizar esses efeitos, neste trabalho, implementou-se

em Matlab uma estratégia de busca por triângulos semelhantes entre os pares de imagens, com base na combinação de todos os pontos-chave, gerando uma procura aproximada pelos ângulos, os quais se mantêm independentemente da escala ou rotação utilizada.

Embora a técnica SIFT, em princípio, seja invariante à escala e rotação, as imagens foram reamostradas com rotações e escalas aproximadas para efeitos de testes, sendo as aéreas, superamostradas para o GSD de 5 cm, e as terrestres, subamostradas para o GSD de 2 cm, como se apresenta na próxima seção.

4 EXPERIMENTOS E RESULTADOS

Para determinar a correspondência entre duas imagens, devem-se encontrar pontos que sejam comuns entre ambas. A técnica SIFT, como mostrado na Seção 2, detecta pontos de interesse por meio de descritores. Então, tendo-se os descritores das imagens, a tarefa de determinar as correspondências consiste em encontrar os descritores em uma imagem e seus melhores equivalentes na outra. A correspondência candidata para cada ponto-chave é determinada mediante a identificação de seu vizinho mais próximo em um banco de dados de pontos-chave. Tal vizinho mais próximo é definido como o ponto-chave com distância Euclidiana mínima em relação ao vetor descritor invariante.

Na sequência, são mostrados quatro testes de correspondências entre as cenas terrestres e aéreas com base na técnica desenvolvida e implementada em Matlab por Lowe (2005), cujos resultados demonstram situações diferentes.

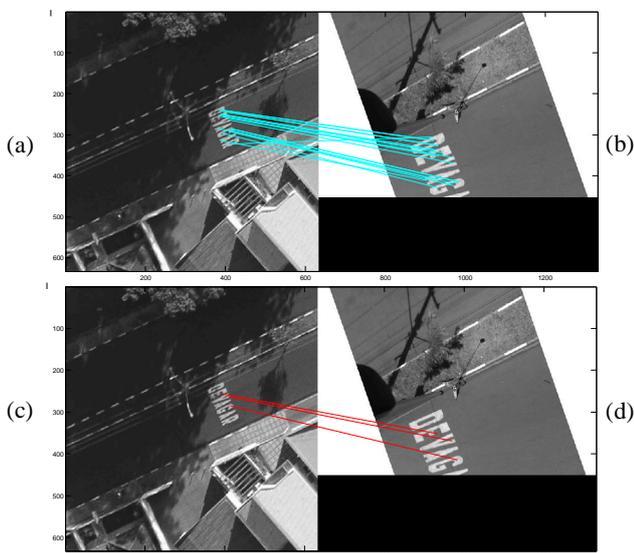


Figura 6: Caso favorável – (a) correspondência entre uma imagem aérea (sem restrição de tamanho) e (b) uma imagem terrestre. Entre (c) e (d) foram identificadas 19 relações, sendo destacados somente 3 casos corretos.

No primeiro teste, como mostra a Figura 6(a), tem-se um recorte de uma imagem aérea com cerca de 1000 m² e uma imagem terrestre com aproximadamente 125 m². Este é um caso desejável, pois apresenta um espaço de busca grande, sem necessidade de restringir o tamanho

e com várias ocorrências de relacionamentos corretos entre as imagens, dos quais se destacam somente três, após a busca por triângulos semelhantes. É importante ressaltar que, para o propósito deste estudo, basta apenas uma correspondência correta.

Já no segundo teste, Figura 7, pode-se perceber que há vários relacionamentos corretos, bem como outros incorretos, mas o uso da filtragem por triângulos semelhantes conseguiu encontrar 3 pares de correspondência corretos.

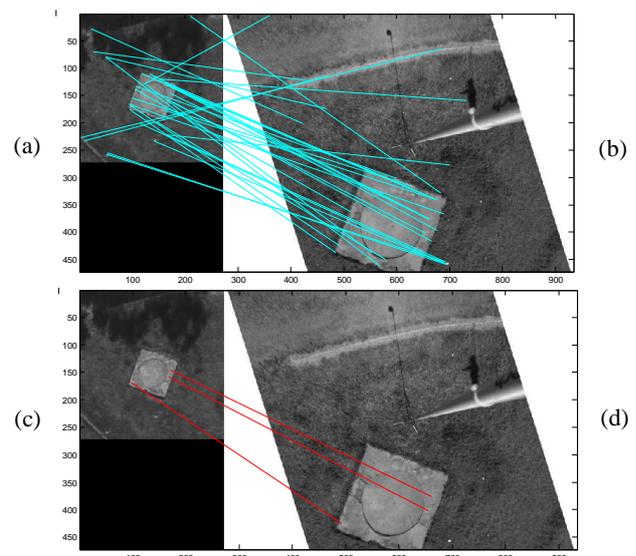


Figura 7: Total de 34 correspondências entre (a) e (b), localizando-se 3 casos corretos, em (c) e (d), pela triangulação.

No teste cujos resultados são mostrados na Figura 8, foram encontradas 160 relações entre ambas as imagens. Nota-se que a busca por triângulos semelhantes identificou 3 pares, porém um desses pares não é correto e deveria ter sido eliminado.

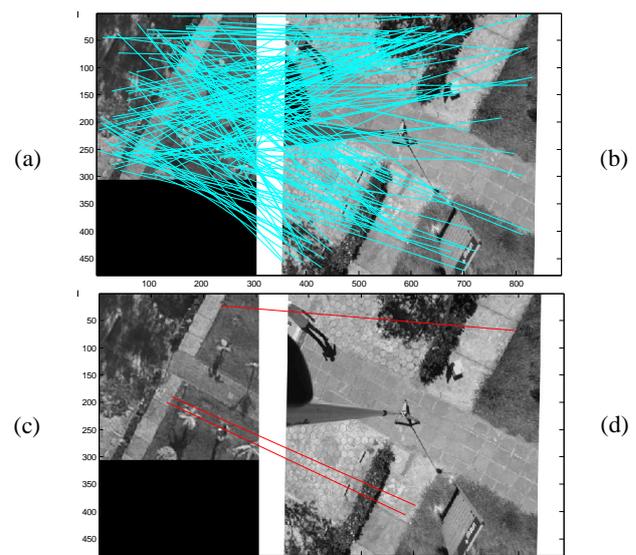


Figura 8: Identificação de 160 relações em (a) e (b), sendo encontrados 3 pares de correspondência em (c) e (d), dos quais um é falso.

Mais uma situação é exposta na Figura 9. Mesmo tendo encontrado 239 relacionamentos, a busca por meio de triângulos resultou em 3 pares, dos quais somente 1 é correto, o que também requer algum procedimento adicional para identificá-lo.

A procura por pares de correspondência foi baseada em triangulação, mas há outras possibilidades que também podem contribuir na minimização dos erros de relacionamentos, tais como aplicação de transformações, RANSAC, relaxação etc.

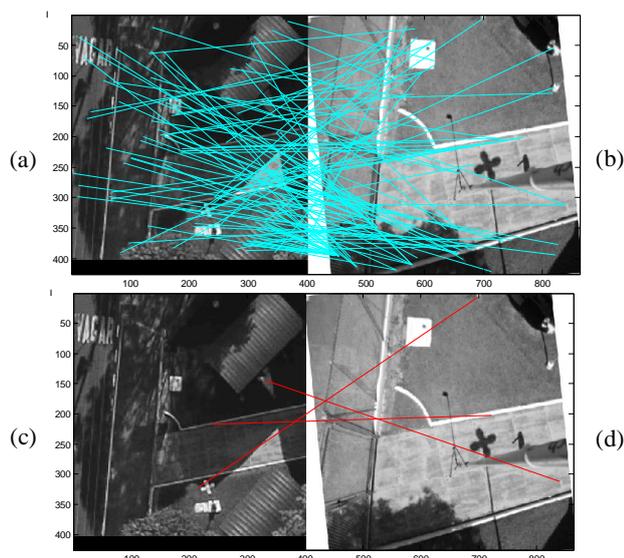


Figura 9: Caso em que a busca por meio de triângulos, entre (a) e (b), encontra somente uma correspondência correta, entre (c) e (d).

Outros experimentos, usando a imagem aérea em seu nível de resolução original ou sem qualquer rotação entre as imagens, também foram feitos, mas os resultados demonstraram-se inferiores aos testes apresentados, mostrando que é relevante aplicar uma reamostragem preliminar, com base em valores aproximados.

Logo, pelo que foi mostrado, é possível determinar a correspondência entre imagens multiescalas em diferentes condições de geometria e radiometria, visto que pelo menos um relacionamento correto foi encontrado, nos casos estudados. A extrapolação para os demais casos requer um estudo experimental mais detalhado.

5 CONSIDERAÇÕES FINAIS

A metodologia apresentada tem como propósito estabelecer a correspondência entre imagens multiescalas de diferentes sensores, de maneira que cenas terrestres georreferenciadas possam ser usadas como fonte de controle para imagens aéreas, com alto grau de automação.

Nesse objetivo, basta uma única correspondência correta para se fazer conexão da cena terrestre com a imagem aérea, ou seja, um ponto de ligação já é suficiente para se transferir um ponto georreferenciado da cena terrestre para a imagem aérea. Após a transferência de coordenadas, devem-se aplicar refinamentos com o intuito de atingir a acurácia subpixel.

Pelos testes realizados, verifica-se que pontos de correspondência foram localizados, porém, em algumas situações, ainda há a necessidade de uma estratégia que assegure se o relacionamento é válido, diminuindo a possibilidade de erros.

Embora tenham sido encontrados relacionamentos entre as imagens, outros estudos futuros precisam ser feitos no sentido de analisar a dimensionalidade das cenas de controle e das áreas de busca, pois, uma vez que as cenas terrestres não consideram padrões definidos, suas características dependem da região do ponto, devendo ser determinadas adaptativamente. Já no caso das imagens aéreas, uma alternativa é avaliar a qualidade dos parâmetros aproximados de orientação exterior da câmara, podendo a determinação do espaço de busca ser feita por propagação de erros.

Portanto, são estratégias que devem ser incorporadas ao método de localização automática em desenvolvimento, a fim de que a aplicação fotogramétrica seja mais robusta e possa alcançar a acurácia e a precisão subpixel após refinamentos.

AGRADECIMENTOS

Os autores agradecem à FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) pelo apoio financeiro, na forma de bolsa de estudo – Processo 2010/16954-5.

REFERÊNCIAS

- FÖRSTNER, W. A feature based correspondence algorithm for image matching. In: **Proceedings of Symposium from Analytical to Digital**, Rovaniemi, Finland, 1986, p. 150-166.
- HARRIS, C.; STEPHENS, M. A combined corner and edge detector. In: **Proceedings of 4th Alvey Vision Conference**, Alvey, UK, 1988, p. 147-151.
- HEIPKE, C. Overview of Image Matching Techniques. In: **Proceedings of the OEEPE - Workshop on the application of digital photogrammetric workstations**. Lausanne, Mar. 1996.
- LINGUA, A.; MARENCHINO, D.; NEX, F. Performance analysis of the SIFT operator for automatic feature extraction and matching in photogrammetric applications. In: **Sensors**. v. 9, p. 3745-3766, 2009.
- LOWE, D. G. Object recognition from local scale-invariant features. In: **International Conference on Computer Vision**, Corfu, Greece, September 1999, p. 1150-1157
- LOWE, D. G. Distinctive image features from scale-invariant keypoints. In: **International Journal of Computer Vision**, V. 60, 2, 2004, p. 91-110
- LOWE, D. G.. **Demo software: SIFT keypoint detector**. Versão 4, (2005). Disponível em <<http://www.cs.ubc.ca/~lowe/keypoints/>>. Acesso: 11 maio 2011.
- TOTH, C.; JU, H.; GREJNER-BRZEZINSKA, D. Matching between different image domains. In: STILLA, U.; ROTTENSTEINER, F.; MAYER, H.; JUTZI, B.; BUTENUTH, M. (Ed.) **Photogrammetric Image Analysis: ISPRS Conference**, Pia, 2011. **Proceedings...** Springer-Verlag: Munich, p. 37-47, 2011.