



UMA TÉCNICA PARA DETERMINAÇÃO DE LINHAS EPIPOLARES EM IMAGENS PANORÂMICAS TERRESTRES

L. F. Castanheiro¹, A. M. G. Tommaselli², M. B. Campos³, A. Berveglieri⁴

¹Bolsista PIBIC-CNPq, Faculdade de Ciências e Tecnologia – FCT, Pres. Prudente – SP

²Prof. Titular, Departamento de Cartografia, Faculdade de Ciências e Tecnologia – FCT, Pres. Prudente – SP

³Bolsista Capes, Programa de Pós-Graduação em Ciências Cartográficas

⁴Bolsista Pós-Doutorado FAPESP, Faculdade de Ciências e Tecnologia – FCT, Pres. Prudente – SP

Comissão IV – Sensoriamento Remoto, Fotogrametria e Interpretação de Imagens

RESUMO

Com o crescente aumento de aplicações utilizando sensores omnidirecionais surgiu a necessidade de adaptar processos fotogramétricos clássicos, como a correspondência de imagens, às diferentes características geométricas das imagens panorâmicas. As técnicas de correspondências desenvolvidas para imagens perspectivas convencionais geram resultados insatisfatórios ao serem aplicadas em imagens panorâmicas. Nesse sentido, este artigo apresenta testes e comparações com três tipos de operadores de interesse: SIFT (*Scale-Invariant Feature Transform*), MOPS (*Multi-Scale Oriented Patches*) e SURF (*Speeded-Up Robus Features*), aplicados em imagens panorâmicas. Apresenta, ainda uma técnica de redução do espaço de busca e filtragem de falsas correspondências entre imagens *fisheye* baseada na geometria epipolar. Os resultados mostram o potencial da técnica proposta na otimização do uso dos operadores de interesse em imagens panorâmicas.

Palavras-chave: Fotogrametria, Geometria Epipolar, *Fisheye*, Imagens Omnidirecionais.

ABSTRACT

The number of applications using omnidirectional systems have increased in the last years, requiring modifications in classical photogrammetric processes, such as imaging matching, considering the different geometry of panoramic images. Usually, image matching techniques designed for perspective images are applied on panoramic images, resulting in unsuitable solutions. In this regard, this paper presents tests and assessments of three types of interest operators: SIFT (*Scale-Invariant Feature Transform*), MOPS (*Multi-Scale Oriented Patches*) and SURF (*Speeded-Up Robus Features*). A technique to reduce the search space and to filter false matching using epipolar geometry is also presented. The results show the potential of the technique proposed to optimize the use of interest operators in panoramic images.

Keywords: Photogrammetry, Epipolar Geometry, *Fisheye*, Omnidirectional Images.

1- INTRODUÇÃO

Um dos problemas fundamentais nas áreas de Fotogrametria e Visão Computacional é a identificação e correspondência automática de feições homólogas em um ou mais pares de imagens, utilizadas em diversos processos fotogramétricos como, calibração de câmara, fototriangulação e reconstrução 3D de objetos. A correspondência de imagens é um problema mal condicionado, isto é, a solução pode não existir, devido à oclusão da feição em um dos pares de imagens, ou não ser única, devido a eventuais padrões repetitivos (Heipke, 1996), sendo estes os principais desafios enfrentados pelos algoritmos de correspondência automática.

O desafio pode ser ainda maior em imagens panorâmicas devido à grande variação de escala entre as cenas, à resolução espacial desuniforme, à distorção radial acentuada e à geometria interna não perspectiva (Ray, 2002). Em imagens obtidas a partir de câmaras com geometria *fisheye*, por exemplo, a lente é semi-esférica, sendo projetada para o plano imagem utilizando projeções como a estereográfica, equidistante, ortogonal, dentre outras (Hughes et al., 2010). Estas características não são comuns na maioria dos operadores de interesse e nos métodos de correspondência de imagens desenvolvidos para imagens perspectivas, o que pode causar uma deterioração na solução quando aplicados em imagem

panorâmicas. Uma forma de minimizar o problema de mal condicionamento e reduzir o custo computacional é restringir a região de busca das feições homólogas na imagem. Boas aproximações para a solução podem ser obtidas considerando estratégias como a busca hierárquica, busca sobre a linha epipolar conjugada, busca vertical, dentre outras (Wolf e Dewitt, 2000). A restrição do espaço de busca também pode minimizar a complexidade da correspondência em imagens panorâmicas. No entanto, as características dessas imagens devem ser consideradas para que uma boa aproximação da solução seja obtida.

Apesar do crescente número de aplicações em Fotogrametria à curta distância com sensores omnidirecionais, o problema de correspondência em imagens panorâmicas ainda é pouco estudado. Geralmente, as técnicas desenvolvidas para imagens perspectivas convencionais são aplicadas nas imagens panorâmicas (Cruz et al., 2012), resultando em baixos percentuais de detecção, repetibilidade e acurácia. Nesse contexto, este artigo avalia os resultados de alguns operadores de interesse, SIFT (*Scale-Invariant Feature Transform*), MOPS (*Multi-Scale Oriented Patches*) e SURF (*Speeded-Up Robus Features*), aplicados em imagens *fisheye* (Seção 2). Também, são realizados testes com uma técnica de redução do espaço de busca em linhas epipolares conjugadas para imagens panorâmicas, considerando uma superfície de imageamento esférica e o modelo de projeção equidistante (Seções 3 e 4).

2- OPERADORES DE INTERESSE APLICADOS EM IMAGENS PANORÂMICAS

Com o intuito de verificar os principais problemas na correspondência em imagens panorâmicas, foram testados nesse artigo os operadores de interesse SIFT (Lowe, 2004), MOPS (Brown et al., 2005) e SURF (Bay et al., 2008). Em um primeiro momento foram detectadas feições de interesse aplicando os três operadores em um par de imagens panorâmicas. Em seguida, foi realizada a correspondência entre as feições detectadas nas duas imagens. As imagens utilizadas foram do tipo *fisheye* obtidas pela câmara Ricoh Theta S (Theta Developers, 2017). Os processamentos foram feitos nos programas Matlab (Moler et al., 1992) e Fiji (Schindelin et al., 2012). Os critérios de avaliação para cada operador foram a precisão de detecção dos pontos homólogos, velocidade computacional e a insensibilidade a variações de rotação e escala.

2.2- SIFT – *Scale invariant feature transform*

SIFT é uma técnica desenvolvida por Lowe (2004) que detecta e extrai pontos com significativa invariância à escala e rotação. O algoritmo também extrai um grande número de características distintas usando vetores descritores baseados em histogramas de

gradientes e orientações, o que permite uma correspondência mais eficiente (Lowe, 2004). A técnica desse operador é dividida em quatro etapas principais, resumidamente descritas a seguir.

A primeira etapa desse operador é a detecção dos pontos extremos no espaço-escala, que consiste em obter pontos máximos e mínimos locais usando a função DoG (Diferença de Gaussianas) com variação de uma pequena escala para identificar pontos estáveis, que são candidatos a serem pontos-chave. A segunda etapa faz o ajuste da localização dos pontos obtidos na primeira etapa para um ponto de máximo local. São rejeitados pontos com baixo contraste e também pontos mal localizados, como em bordas. Na terceira etapa, o vetor descritor do ponto-chave é representado em relação à orientação com base nos gradientes locais (magnitude e orientação do gradiente local). Por fim, na quarta etapa, os vetores descritores são construídos com base nos histogramas de gradientes ao redor dos pontos-chaves. O resultado final é um conjunto de descritores que representam aquele ponto de interesse invariante com a escala. Então, a correspondência entre pontos extraídos em diferentes imagens pode ser feita pela comparação entre descritores, por exemplo, usando a menor distância euclidiana entre os atributos dos pontos comparados.

As operações com o SIFT foram realizadas no Matlab com *scripts* disponibilizados por Lowe (2005). No final do processamento, foram obtidos os vetores descritores e os pontos correspondentes no par de imagens, totalizando 160 pontos homólogos. O resultado dessa correspondência SIFT foi inspecionado visualmente, verificando-se que 44% dos pontos homólogos detectados foram falsas correspondências (Tabela 1). A maior incidência das correspondências incorretas ocorreu na borda da imagem, uma vez que a geometria não perspectiva e a grande distorção radial simétrica das imagens panorâmicas não são consideradas por este operador. A velocidade de processamento do SIFT foi a menor entre os operadores testados (Tabela 1), o que implica em um maior custo computacional utilizando esse operador.

2.3- MOPS – *Multi-scale oriented patches*

O MOPS foi desenvolvido como um algoritmo de supressão não máxima que melhor distribui as feições na imagem (Brown et al., 2005). O primeiro passo desse operador, assim como de outros operadores de correspondência, é a detecção e extração de pontos de interesse. Para encontrar esses pontos, utiliza-se uma função de detecção de canto e, posteriormente, calcula-se a orientação de cada ponto. Para restringir o número máximo de pontos extraídos para cada imagem e, ao mesmo tempo, garantir que os pontos sejam muito bem distribuídos na imagem, foi desenvolvido o ANMS (*Adaptive Non-Maxima Suppression*) (Brown et al., 2005). Por fim, tem-se a etapa da correspondência por características, que

consiste em encontrar um conjunto de candidatos a pontos correspondentes pelo algoritmo do vizinho mais próximo e, então, refina-se as correspondências com um procedimento baseado em análises estatísticas.

Para o MOPS, utilizou-se do programa Fiji que já tem implementado a correspondência de imagens com esta técnica. Esse operador resultou em muitas falsas correspondências, encontrando apenas 5,6% de correspondências corretas, além de detectar um menor número de pontos homólogos, quando comparado aos outros operadores testados (Tabela 1). Para esse operador, não foi possível ter acesso ao tempo de processamento, pois o programa não disponibiliza esse dado.

2.4- SURF – Speeded-up robust features

SURF é um dos operadores capazes de detectar pontos de interesse parcialmente invariantes à escala e rotação, de modo similar ao SIFT, mas com simplificações para acelerar o processamento (Bay et al., 2008). O processamento de busca pelos pontos correspondentes é dividido em três etapas principais. A primeira é a detecção dos pontos de interesse que é feito pela aproximação baseada na matriz Hessiana e também pelo conceito de imagem integral (Viola e Jones, 2001). Após a detecção, é feita a extração dos descritores dos pontos, que descrevem a distribuição de intensidade contida dentro da vizinhança dos pontos de interesse, de modo similar ao SIFT (Bay et al., 2008). Por fim, são determinadas as correspondências entre as duas imagens.

Para o SURF, foi utilizado um *script* implementado no Matlab, gerando, ao final do processo uma imagem com os pontos homólogos detectados. O operador SURF proporcionou um total de 197 correspondências. O percentual de falsas correspondências (45%), foi aproximadamente o mesmo que o SIFT, porém com um tempo de processamento muito menor (Tabela 1).

Nos testes realizados com os três operadores de interesse, o MOPS foi o que apresentou o pior desempenho. Para esse operador, não foi possível obter o código fonte livre, o que limitou os testes com as imagens panorâmicas. O SIFT e o SURF apresentaram uma eficiência similar na correspondência, que pode ser melhorada com a redução do espaço de busca. A Seção 3 apresenta uma técnica para redução do espaço de busca pelas linhas epipolares conjugadas. A Tabela 1 apresenta o resultado geral de cada operador, em que é possível observar as diferenças de desempenho.

Tabela 1 - Resultado dos operadores de interesse.

Operador	Tempo	Total	Corretas	Incorretas
SIFT	89,5 s	160	89(56%)	71(44%)
MOPS	-	71	4(5,6%)	67(94,3%)
SURF	5,5 s	197	109(55%)	88(45%)

3- REDUÇÃO DO ESPAÇO DE BUSCA EM IMAGENS PANORÂMICAS

A geometria epipolar é usada na correspondência entre imagens devido a uma propriedade importante que ela impõe, denominada de restrição epipolar (Zhang, 1995; Schenk, 1999). Essa propriedade é uma restrição geométrica que transforma o problema da correspondência de imagem de bidimensional para unidimensional. A Fig. 1 mostra a geometria epipolar para duas imagens esféricas.

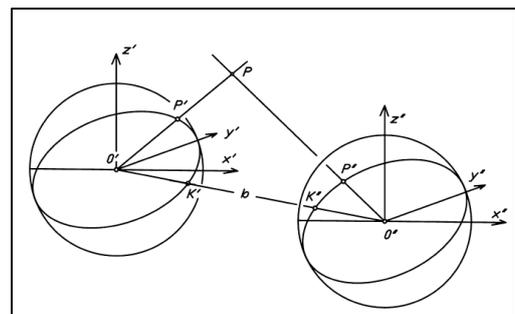


Fig. 1 – Geometria epipolar em imagens esféricas.

Fonte: Fangi e Nardinocchie (2013).

A técnica de redução do espaço de busca pela linha epipolar em imagens *fisheye* parte de um ponto de interesse na imagem da esquerda e define uma linha de busca na imagem da direita. Devido à geometria complexa da imagem *fisheye*, alguns arranjos devem ser feitos, como a projeção do ponto em uma superfície esférica cujo raio é a distância focal da câmara (f) (Fig. 2).

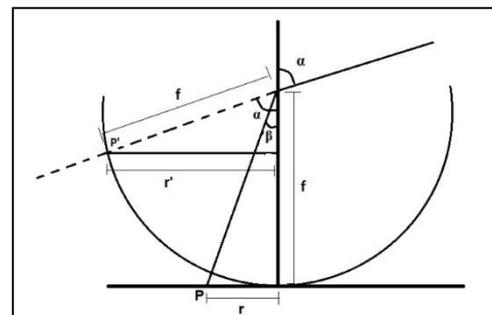


Fig. 2 - Projeção do ponto p na superfície esférica.

Na Fig. 2, p é o ponto na imagem *fisheye* com coordenadas (x,y,-f), r é a distância do ponto ao centro da imagem, α é o ângulo do raio incidente, p' é o ponto sobre a superfície esférica com coordenadas (x_{esf}, y_{esf}, z_{esf}) e r_{esf} é a projeção de r na esfera.

O ângulo de incidência do raio luminoso para um modelo de lente *fisheye* do tipo equidistante é dado pela Equação 1, e as coordenadas do ponto projetado na esfera são dadas pelas Equações 2, 3 e 4.

$$\alpha = r/f \quad (1)$$

$$x_{esf} = r_{esf} \cos(\theta) \quad (2)$$

$$y_{esf} = r_{esf} \sin(\theta) \quad (3)$$

$$Z_{esf} = f \cos(\alpha) \quad (4)$$

em que, $r_{esf} = f \sin(\alpha)$; e $\theta = \arctg(x/y)$.

A partir das coordenadas do ponto projetado na superfície esférica, faz-se a projeção para vários planos Z no espaço objeto, determinando-se, assim, as suas coordenadas no espaço objeto (X,Y) usando as equações de colinearidade na forma inversa e adotando um sistema local, com origem na foto da esquerda. Com as coordenadas no espaço objeto local, o próximo passo é determinar as posições desses pontos na imagem da direita, e para isso, deve-se calcular a matriz relativa entre as imagens e os elementos da base, como mostram as Equações 5 e 6, respectivamente. Parte-se do princípio de que são conhecidos os parâmetros de orientação exterior de ambas as imagens, o que pode ser obtido por um sistema de navegação.

$$R^R = R^e(R^d)^{-1} \quad (5)$$

$$[B_x \ B_y \ B_z]^T = R^e [X_0^d - X_0^e \ Y_0^d - Y_0^e \ Z_0^d - Z_0^e]^T \quad (6)$$

A determinação da linha epipolar conjugada na imagem da direita pode ser feita utilizando o modelo de projeção para lentes *fisheye*, como já mencionado na Seção 1. Nesse caso, utilizou-se o modelo equidistante, compatível com o modelo de projeção da câmara Ricoh Theta S, utilizada na aquisição das imagens. As Equações 7 e 8 mostram as equações do modelo equidistante, desenvolvidas a partir da condição expressa na equação (1).

$$x_{sd} = -f \frac{X_c}{\sqrt{(X_c)^2 + (Y_c)^2}} \arctg\left(\frac{\sqrt{(X_c)^2 + (Y_c)^2}}{Z_c}\right) \quad (7)$$

$$y_{sd} = -f \frac{Y_c}{\sqrt{(X_c)^2 + (Y_c)^2}} \arctg\left(\frac{\sqrt{(X_c)^2 + (Y_c)^2}}{Z_c}\right) \quad (8)$$

em que, X_c , Y_c , Z_c são as coordenadas do ponto no sistema de referência fotogramétrico (Equação 9).

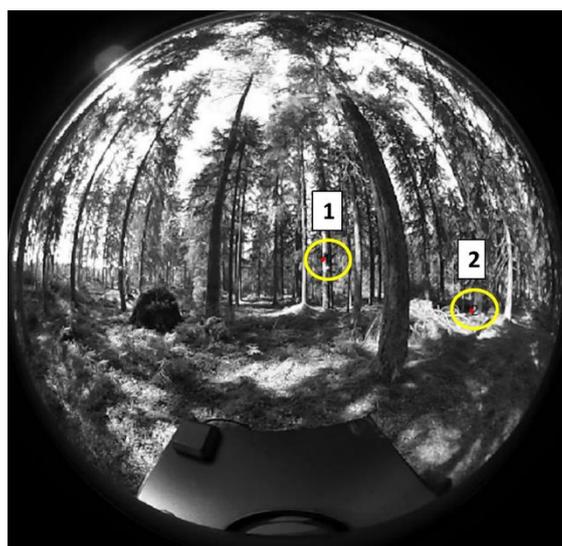
$$[X_c \ Y_c \ Z_c]^T = R^R [X - B_x \ Y - B_y \ Z - B_z]^T \quad (9)$$

4 – EXPERIMENTOS E RESULTADOS PRELIMINARES

No intuito de verificar as correspondências obtidas pela técnica SIFT, elaborou-se um algoritmo para o cálculo da linha epipolar conjugada na imagem da direita a partir de um ponto identificado como ponto de interesse pelo SIFT na imagem da esquerda, utilizando a técnica apresentada na Seção 3. O código foi implementado no *software* Matlab 2013, e as imagens utilizadas nesse experimento também são do tipo *fisheye*, obtidas pela câmara Ricoh Theta S.

Foram inspecionados dois pontos obtidos pelo operador SIFT com diferentes distribuições na imagem (centro e borda). A Fig. 3 apresenta em (a) os pontos de interesse identificados pelo operador SIFT na imagem da esquerda. Os resultados obtidos com o operador e com a linha epipolar conjugada para os

pontos 1 e 2, são apresentados em (b) e (c) respectivamente.



(a)



(b)



(c)

Fig. 3 – Filtragem pela linha epipolar conjugada: (a) Localização dos pontos na imagem da esquerda obtidos pelo SIFT. (b) Resultado da linha epipolar para o ponto 1 e (c) para o ponto 2.

O ponto 1 é uma falsas correspondências geradas pelo operador, o que pode ser observado na Fig. 3.b. Fazendo a filtragem pela linha epipolar conjugada para o ponto 1, observa-se que a linha não coincide com o ponto correspondente identificado pelo SIFT. Portanto, usando a restrição epipolar o ponto 1 seria descartado, diminuindo as falsas correspondências geradas pelo operador SIFT usando imagens panorâmicas, e reduzindo os erros nos processos fotogramétricos subsequentes. A Fig. 3.c apresenta os resultados obtidos para o ponto 2, sendo esta uma correspondência correta. Uma vez que os parâmetros de orientação exterior usados para o cálculo da linha epipolar conjugada são valores aproximados, o ponto homólogo, não deve ser buscado apenas sobre a linha epipolar, mas sobre um polígono em volta dela, calculado pela propagação das covariâncias dos parâmetros de orientação exterior. Portanto, se um polígono for determinado no entorno da linha epipolar conjugada, obtida pelo algoritmo implementado, o ponto 2 estará incluso no respectivo polígono e portanto seria classificado como uma correspondência correta.

5- CONCLUSÃO

Devido à complexidade das imagens *fish-eye*, a correspondência usando operadores invariantes à escala é mais apropriada. Apesar dos três operadores testados nesse artigo serem significativamente invariantes à escala, alguns problemas, como o alto percentual de falsas correspondências e a distribuição não ótima dos pontos detectados nas imagens ainda permanecem. A partir dos testes realizados pelo cálculo da linha epipolar conjugada para a redução do espaço de busca, conclui-se que esse procedimento ajuda a reduzir as correspondências incorretas e também reduz o custo computacional no processo de busca por correspondências. Neste artigo apenas foi plotada a linha epipolar conjugada sendo recomendado passo adicional de propagação das covariâncias, delimitando assim um polígono no entorno da linha determinada para a redução do espaço de busca. Recomenda-se um estudo experimental para determinar os erros nos parâmetros de orientação exterior e seu impacto na determinação da janela de busca para câmaras panorâmicas.

AGRADECIMENTO

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq - processo 162235/2017-8) pela concessão da bolsa no programa PIBIC. E a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP – Processo 2013/50426-4).

REFERÊNCIAS BIBLIOGRÁFICAS

- Bay, H. et al. Speeded-up Robust Features (SURF). Computer vision and image understanding, v.110, n. 3, p 346-359, 2008.
- Brown, M.; Szeliski, R.; Winder, S. Multi-Image Matching Using Multi-Scale Oriented Patches. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Proceedings CVPR, IEEE, 2005. P. 510-517.
- Cruz, J. M.; Bogdanova, I.; Paquier, B.; Bierlaire, M.; Thiran, J. P. Scale invariant feature transform on the sphere: Theory and applications. International journal of computer vision, v. 98, n. 2, p. 217-241, 2012.
- Fangi, G.; Nardinocchie, C. Photogrammetric Processing of Spherical Panoramas. The Photogrammetric Record 28(143): 293-311. 2013
- Heipke, C. Overview of image matching techniques. In: Workshop on the Application of digital photogrammetric workstations, 1996, Lausanne. Proceedings... Lausanne: OEEPE, 1996.
- Hughes, C.; Denny, P.; Jones, E.; Glavin, M. Accuracy of fish-eye lens models. Applied Optics, v. 49, n. 17, p. 3338-3347, 2010
- Lowe, D. G. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, vol. 60, no. 2, pg. 91-110. 2004
- Lowe, D. G. Demo software: SIFT keypoint detector. Matlab, 2005. Disponíveis em: < <http://www.cs.ubc.ca/~lowe/keypoints/>>. Acesso em: 22 nov 2016.
- Moler, C.; Gilbert, J. R.; Schreiber, R. Sparse Matrices in MATLAB: Design and Implementation. SIAM Journal on Matrix Analysis and Applications, p. 333-356. 1992
- Ray, S. F. Applied photographic optics. 3ª Edição. Focal Press, Oxford – UK, 2002.
- Schenk, T. Digital Photogrammetry: background, fundamentals, automatic orientation procedures. Laurelville: TerraScience, 428 p. 1999
- Schindelin, J; et al. Fiji: an open-source platform for biological-image analysis. Nature methods, v. 9, n. 7, p. 676-682. 2012
- Theta Developers. Proprietary Technical Information: Lens Parameter Information. Disponível em: < <http://theta360.guide/community-document/community.html>> Acessado em 19 de abril de 2017.
- Viola, P. A.; Jones, M. J. Rapid object detection using a boosted cascaded of simples features. Accepted Conference on Computer Vision and Pattern Recognition. Cambridge, 2001
- Wolf, P.R.; Dewitt, B.A. Elements of Photogrammetry: with Applications in GIS, ed. 3. McGRAW-HILL, 2000
- Zhang, Z. et al. A Robust Technique for Matching two Uncalibrated images Through the Recovery of the Unknown Epipolar Geometry. Artificial Intelligence, Essex, UK, v.78, n1-2, p.87-119, 1995.