

UM MÉTODO PARA EXTRAÇÃO DE PONTOS HOMÓLOGOS EM PARES DE IMAGENS ESTEREOSCÓPICAS DE LARGA ESCALA

A METHOD FOR EXTRACTION OF HOMOLOGOUS POINTS IN PAIRS OF LARGE SCALE STEREOSCOPIC IMAGES

Cleiston Rodrigues da Silva^a, Marcus Vinícius de Souza Ferraz^a e Alessandra Martins Coelho^b

^a*Programa de Pós-Graduação em Modelagem Computacional, Universidade Federal de Juiz de Fora, Campus Universitário, Rua José Lourenço Kelmer, s/n - São Pedro, Juiz de Fora, MG, Brasil*
<http://www.ufjf.br>

^b*Departamento de Ciência da Computação, Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas Gerais, Campus Rio Pomba, Av. Dr. José Sebastião da Paixão s/n - Lindo Vale, Rio Pomba, MG, Brasil* <http://www.riopomba.ifsudestemg.edu.br>

Palavras-chave: Extração de características, SIFT, Visão computacional, estereoscopia.

Resumo. Este trabalho apresenta um método para extração de características utilizando o algoritmo SIFT em pares de imagens estereoscópicas de alta resolução. Devido a limitações do SIFT, este não pode prover uma solução direta, portanto um método para dividir as imagens em blocos menores é proposto. Os pontos extraídos de cada bloco são processados e combinados com o objetivo de gerar uma solução global. O *K-Nearest Neighbor* é usado para selecionar correspondências. Para tornar o processo de busca mais rápido, utiliza-se o *randomized kd-tree*. Um filtro para eliminação de falsas correspondências é desenvolvido. Os resultados obtidos pelo método proposto são comparados com os resultados obtidos pelo método L^2SIFT , que também realiza o processo de divisão em blocos, utilizando o *dataset* de Toronto. O método proposto obtém desempenho superior, encontrando um número maior de correspondências.

Keywords: Feature Extraction, SIFT, Computer Vision, Stereoscopy.

Abstract. This work presents a method for extraction of characteristics using the SIFT algorithm in pairs of high resolution stereoscopic images. Due to limitations of SIFT, it can not provide a direct solution, therefore a method to divide the images into smaller blocks is proposed. The points extracted from each block are processed and combined in order to generate a global solution. The K-Nearest Neighbor is used to select matches. To make the search process faster, the randomized kd-tree is used. A filter for eliminating false matches is developed. The results obtained by the proposed method are compared with the results obtained by the L^2SIFT method, which also performs the blocking process, using the Toronto dataset. The proposed method obtains superior performance, finding a greater number of correspondences.

1 INTRODUÇÃO

A detecção e correspondência de características é um componente essencial de muitas aplicações de Visão Computacional, tais como o alinhamento de imagens e o estabelecimento de um conjunto denso de correspondências para a geração de um modelo tridimensional (Szeliski, 2010). A extração de características é também um dos principais tópicos em fotogrametria. Este processo consiste na extração de características de interesse de duas ou mais imagens do mesmo objeto e da correspondência dessas características em imagens adjacentes (Lingua et al., 2009).

A fotogrametria aérea e a estereoscopia vêm sendo aplicadas a estudos geotécnicos permitindo a obtenção de medições de escala, altura, área, comprimento e perímetro de objetos a partir de pares de imagens 2D (Coelho, 2012). Para a reprodução precisa de um modelo 3D, a partir de pares de imagens estereoscópicas, é necessário a extração de um grande número de correspondências confiáveis. Diversos algoritmos para a extração de características invariantes à escala e rotação foram propostos na literatura, dentre eles o *Scale Invariant Feature Transform* (SIFT) (Lowe, 2004), o *Speed Up Robust Features* (SURF) (Bay et al., 2006) e o *Oriented FAST and Rotated BRIEF* (ORB) (Rublee et al., 2011). Falsos positivos podem surgir no processo de correspondência entre as características de cada imagem. O uso de restrições geométricas podem ser utilizadas para eliminação de falsas correspondências. Jiang e Jiang (2019), por exemplo, utiliza restrições geométricas e fotométricas através da triangulação de Delaunay.

O algoritmo SIFT não pode prover uma solução direta para a extração de características e correspondências usando imagens de alta resolução obtidas por câmeras métricas em fotogrametria aérea (Sun et al., 2014). Não é possível alocar memória suficiente quando as imagens são muito grandes, pois o algoritmo precisa construir uma série de pirâmides para criar a escala de espaço gaussiano e DoG (*Difference of Gaussian*). Com o propósito de contornar a limitação de memória do SIFT, Sun et al. (2014) propôs o L^2SIFT para o processamento de fotografias aéreas de alta resolução. As fotografias são reduzidas para se obter, através do SIFT, os *seed points*, que serão usados pelo método linear *Singular Value Decomposition* (SVD) (Umeyama, 1991). Este último retorna os parâmetros de transformação que serão usados para dividir as imagens em blocos correspondentes menores, que podem ser computados pelo SIFT. Para determinar a área de sobreposição estereoscópica, a qual será dividida em blocos, são usadas a direção do voo, a orientação da câmera e os parâmetros da transformação. Calcular a área de sobreposição é importante para melhorar a eficiência do algoritmo (Sun et al., 2014).

Coelho (2012) propôs uma metodologia para análise da susceptibilidade de deslizamentos de terra, a partir da extração de pontos homólogos de aerofotogramas, através do uso de estereoscopia e computação gráfica. Os pontos foram extraídos aplicando-se o SIFT em pares de imagens reduzidos. Verificou-se que o algoritmo não detectou pontos em regiões rochosas ou de mata densa, importantes para melhor representação do terreno. Observou-se também que, por se tratar de uma área muito extensa, a seleção de uma área menor poderia trazer resultados mais satisfatórios. Pensando nisso, é proposto neste trabalho o desenvolvimento de um método para automatizar a criação de recortes menores de aerofotogramas, visando maximizar o número de pontos obtidos. Assim, os recortes são processados individualmente pelo SIFT, e as correspondências obtidas são agrupadas com o objetivo de construir uma solução global.

2 EXTRAÇÃO DE CARACTERÍSTICAS COM O SIFT

SIFT é um algoritmo utilizado para a extração de características invariantes à escala, rotação e parcialmente invariantes à mudança de iluminação e posicionamento de câmera (Lowe, 2004). Foi desenvolvido com base em modelos teóricos de visão biológica (Wei et al., 2008). As características, ou descritores, são altamente distintivos, permitindo que pontos homólogos sejam detectados com alta probabilidade de acerto. A obtenção de características segue as etapas descritas nas subseções a seguir.

2.1 Detecção de extremos no espaço escala

O primeiro estágio para detecção de *keypoints* é identificar localizações e escalas que podem ser repetidamente atribuídas a diferentes visões do mesmo objeto. Detectar localizações que são invariantes à mudança de escala da imagem pode ser alcançado pesquisando-se características estáveis em todas as escalas possíveis, usando uma função contínua conhecida como espaço escala (Witkin, 1984). A função *Difference of Gaussian* (DoG) é implementada para encontrar potenciais pontos (*keypoints*) invariantes a escala e rotação.

2.2 Localização de *Keypoints*

A cada localização candidata, um modelo detalhado é criado para determinar localização e escala. Após encontrar um candidato a *keypoint*, deve-se realizar uma análise dos dados locais, escala e razão das principais curvaturas. Essa análise permite rejeitar pontos que possuem baixo contraste (sensíveis a ruído) ou mal localizados ao longo de uma borda. Cada ponto amostrado é comparado aos seus oito vizinhos próximos e os nove vizinhos na escala acima e abaixo. Este será selecionado somente se for o maior ou o menor de todos os seus vizinhos. O processo de detecção de máximos e mínimos locais é ilustrado na Figura 1.

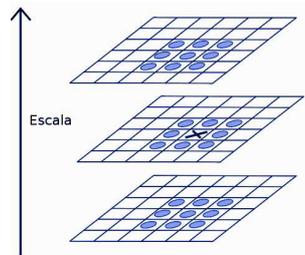


Figura 1: Detecção de máximos e mínimos locais.

2.3 Atribuição de orientação

Ao atribuir uma orientação consistente para cada *keypoint*, baseando-se em propriedades locais da imagem, o descritor pode ter uma representação relativa à sua orientação, atingindo assim invariância à rotação. Um histograma de orientações, de 36 posições, é gerado a partir da orientação dos gradientes. Cada amostra adicionada ao histograma é ponderada pela magnitude do gradiente. Os picos no histograma de orientações correspondem as direções dominantes dos gradientes. O maior pico e aqueles que estão acima de 80% do maior pico são usados para criar novos *keypoints* com a orientação correspondente.

2.4 Descritor de *Keypoints*

Um descritor de *keypoint* é criado primeiramente computando-se a magnitude e a orientação na região de seu entorno. Estes são ponderados por uma janela gaussiana, que faz com que os valores dos *pixels* que estão mais afastados tenham menor contribuição para a construção do descritor, enquanto os mais próximos da localização do mesmo tenham maior contribuição. Esses valores são acumulados em um histograma de orientações, contendo oito posições, gerando o descritor.

Para alcançar invariância à rotação, as coordenadas do descritor e a orientação dos gradientes são rotacionados em relação à orientação do *keypoint*, atribuída no passo anterior. O descritor é normalizado para reduzir os efeitos de mudança de iluminação.

2.5 Correspondência de *Keypoints*

A correspondência de *keypoints* se dá pela identificação do vizinho mais próximo de uma base de *keypoints*. O vizinho mais próximo é aquele com a distância euclidiana mínima para o vetor descritor, descrito no passo anterior (Descritor de *keypoints*). Não se conhece algoritmos que possam identificar o vizinho mais próximo em espaços dimensionais grandes que tenham performance aceitável [Muja e Lowe \(2014\)](#) ou que sejam mais eficientes do que uma busca exaustiva [Lowe \(2004\)](#). Para a localização dos vizinhos mais próximos, utilizou-se a biblioteca *Fast Library for Approximate Nearest Neighbors* (FLANN), descrita na próxima subseção.

Após encontrar os vizinhos mais próximos é necessário filtrar as correspondências. Muitos *keypoints* detectados não possuem uma correspondência correta, podendo estes terem sido originados de um ruído ou oclusão. Para cada correspondência, é calculada a distância euclidiana entre os descritores de cada *keypoint* e seus dois vizinhos mais próximos. Se a distância do mais próximo é significativamente menor que a do seu segundo vizinho, então tem-se uma possível correspondência. De acordo com [Lowe \(2004\)](#), a razão (*ratio*) de 0,8 elimina 90% das falsas correspondências enquanto descarta 5% das corretas.

3 FAST LIBRARY FOR APPROXIMATE NEAREST NEIGHBORS (FLANN)

Os algoritmos mais custosos em visão computacional consistem na busca de correspondências em vetores de alta dimensão, também referido como busca por vizinhos mais próximos. Em [Muja e Lowe \(2014\)](#) é proposta uma biblioteca de código aberto para busca aproximada de vizinhos mais próximos. A mesma contém os algoritmos *K-Nearest Neighbor* (K-NN), para determinar os vizinhos mais próximos de cada *keypoint* detectado pelo SIFT, e o *Randomized kd-tree* para acelerar o processo de busca.

3.1 *K-Nearest Neighbor* (K-NN)

Define-se mais formalmente o K-NN, como:

$$KNN(q, P, K) = A, \quad (1)$$

em que:

$$|A| = K, A \subseteq P;$$

$$\forall x \in A, y \in P - A, d(q, x) \leq d(q, y).$$

O K-NN sempre retorna K vizinhos se existirem pelo menos K pontos em P .

3.2 Randomized kd-tree

O *randomized kd-tree* é um algoritmo de busca aproximada indicado para vetores com muitas dimensões (Muja e Lowe, 2009), que possui uma performance aceitável na realização de busca por vizinhos mais próximos. O algoritmo gera múltiplas *randomized kd-trees*, que são analisadas em paralelo. As árvores são construídas de maneira semelhante à *kd-tree* clássica, que divide os dados na dimensão de maior variância, porém no *randomized kd-tree* a divisão dos dados ocorre selecionando-se aleatoriamente uma das N_D dimensões com maior variância.

4 MÉTODO PROPOSTO

Para que os recortes a serem realizados nos pares de imagens correspondam à mesma área é necessário, primeiramente, conhecer o início da área de sobreposição estereoscópica. Identificado o início da área de sobreposição, o próximo passo é definir uma função de correspondência, a qual possa, a partir de um tamanho de corte definido, criar os sub-pares estereoscópicos e, ao final do processo de detecção de pontos homólogos, agrupá-los em uma única solução, como se os mesmos tivessem sido detectados a partir do par de imagens original. O diagrama da figura 2 apresenta simplificada as principais etapas do método proposto.

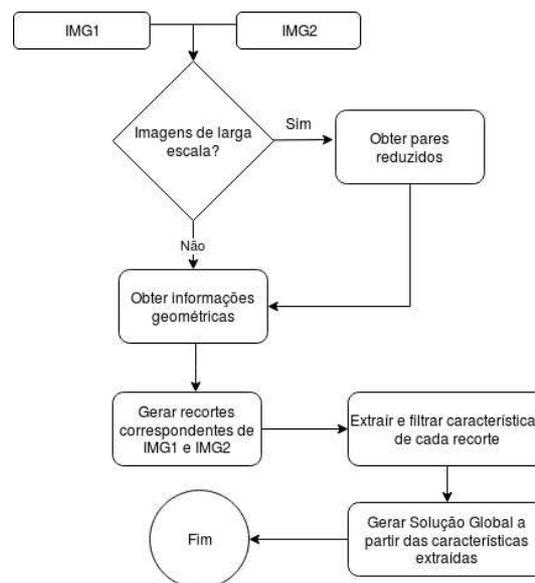


Figura 2: Diagrama simplificado do método proposto.

As imagens devem ser previamente convoluídas por um filtro gaussiano com máscara 5×5 para remoção de ruído. O SIFT é, então, aplicado a cada imagem para detectar seus respectivos *keypoints*, bem como seus descritores associados. Para efetuar a correspondência o algoritmo k-NN é usado. E para reduzir o tempo de busca, o *randomized kd-tree*. O próximo passo é eliminar as falsas correspondências encontradas.

Apesar de se obter correspondências confiáveis usando um valor baixo para o *ratio*, falsos-negativos também são descartados. Pensando nisso, é proposto um filtro que utiliza informações geométricas extraídas dos pares reduzidos para eliminar falsas correspondências.

4.1 Filtro proposto

As informações geométricas são utilizadas para eliminar as falsos positivos. Partindo de um conjunto C contendo pelo menos N_c correspondências confiáveis detectadas pelo SIFT, e de um

valor de *ratio* (r) baixo, busca-se obter informações que serão utilizadas posteriormente para eliminar as falsas correspondências. Começando com $r = 0,2$, incrementa-se iterativamente em $0,1$ até obter N_c correspondências ou alcançar $r = 0,6$.

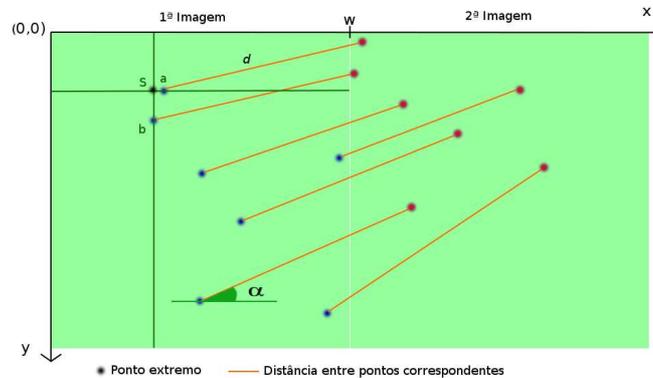


Figura 3: Representação do ponto mínimo (S) e as distâncias, d , entre pontos homólogos.

Para exemplificar, na figura 3 temos a representação de imagens estereoscópicas posicionadas lado a lado em um plano. As informações geométricas que serão extraídas são: $S(x_{min}, y_{min})$, representando o início da área de sobreposição estereoscópica, em que x_{min} e y_{min} são os valores mínimos de cada dimensão dentro do conjunto encontrado; L representa a mediana das distâncias euclidianas entre as coordenadas de *keypoints* homólogos; α é dado pela mediana dos ângulos formados entre as retas que passam pelos pontos homólogos e a reta do eixo das abscissas.

As correspondências que estão dentro das faixas de distâncias $[L * (1 - m), L * (1 + m)]$ e ângulos $[\alpha - \theta, \alpha + \theta]$ são consideradas confiáveis. L e α são parâmetros automáticos e θ e m são as margens de erro. Os valores atribuídos a m e θ , bem como outros parâmetros do método, estão descritos na seção Experimentos.

4.1.1 Recorte dos pares

Os processo de divisão de pares correspondentes se dá através do uso das informações geométricas. Os valores de S e L são escalados para o tamanho original das imagens multiplicando-os pelo fator (λ) de redução das imagens. Portanto, o início da área de sobreposição estereoscópica, bem como a mediana das distâncias entre pontos, são, respectivamente, $S' = \lambda * S$ e $L' = \lambda * L$.

Blocos de tamanho $n \times n$ são recortados na primeira imagem, a partir da coordenada (s_x, s_y) . O ponto correspondente, (t_x, t_y) , é obtido a partir da equação 2.

$$\begin{aligned} t_x &= s_x + L' * \cos \alpha - w \\ t_y &= s_y + L' * \sin \alpha \end{aligned} \quad (2)$$

Os pontos (s_x, s_y) e (t_x, t_y) correspondem, respectivamente, ao primeiro *pixel* do bloco de origem e destino do par de imagens. As áreas de maior altitude podem não ser encontradas dentro do bloco da segunda imagem, devido ao deslocamento de áreas de maior paralaxe. Optou-se

então por expandir o segundo bloco para encontrar correspondências também nessas áreas. Os blocos foram expandidos em cada aresta pelo valor de e , obtido pela equação 3.

$$\begin{aligned} e_x &= \kappa * n * \cos \alpha \\ e_y &= \kappa * n * \sin \alpha \end{aligned} \quad (3)$$

O ponto (e_x, e_y) indica o quanto cada aresta do segundo bloco deve ser expandida, κ é um valor definido entre 0 e 1, e n é o comprimento dos lados dos recortes. Os blocos da segunda imagem são preenchidos com fundo preto caso estejam fora do domínio da mesma, além disso, índices negativos recebem o valor zero. A figura 4 ilustra o processo de divisão e correspondência de blocos homólogos.

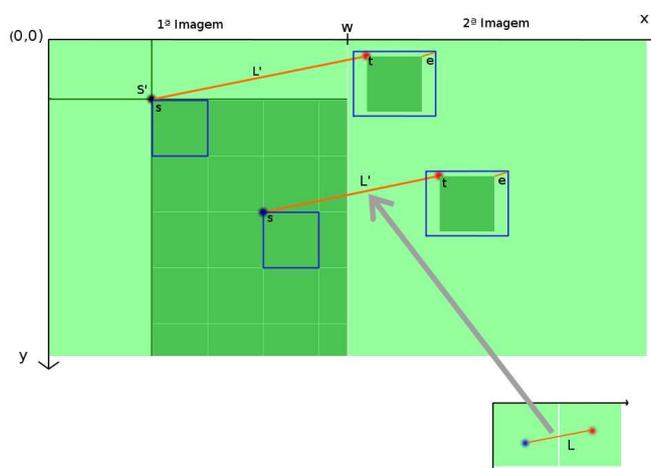


Figura 4: Exemplo de correspondência entre blocos homólogos.

4.1.2 Geração da solução global

Para cada bloco criado, armazena-se a coordenada do primeiro *pixel* em relação à imagem original. Esta coordenada será somada às coordenadas dos *keypoints* extraídos de cada bloco após a eliminação de falsas-correspondências, pelo filtro proposto. Com isso as soluções locais tornam-se uma solução global para o par de imagens original.

5 EXPERIMENTOS

Para validar o método proposto foram realizados experimentos com dois conjuntos de imagens. Foram processados os aerofotogramas de Nova Friburgo, Rio de Janeiro (figura 5), utilizadas previamente nos trabalhos de Fernandes (2008), do Carmo (2010) e Coelho (2012). O método foi aplicado aos pares de imagens de tamanho original (10928×11096) e reduzidos (1093×1110).

Treze imagens do *dataset* de toronto, com resolução de 11500×7500 , também foram processadas. Estas estão publicamente disponíveis pelo *ISPRS Test Project on Urban Classification and 3D Building Reconstruction* (Rottensteiner et al., 2013).

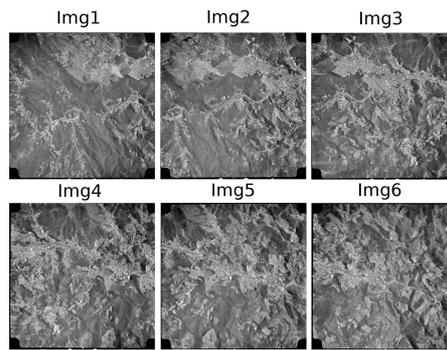


Figura 5: Imagens aéreas de Nova Friburgo (RJ).

As imagens foram previamente suavizadas por um filtro gaussiano de máscara 5×5 . A aplicação deste tornou a execução do método mais rápida, além de resultar em um número maior de correspondências encontradas.

5.1 Parâmetros utilizados

Os parâmetros do método proposto foram determinados levando em consideração o número de pontos obtidos, a maneira como estavam dispersos e o tempo total de execução do método. Assim, os parâmetros foram escolhidos experimentalmente, sendo os apresentados na tabela 1, os que mais se adequaram às imagens testadas.

Parâmetro	θ	m	κ	N_c	n	λ
Valor	2	0,2	0,2	10	500	10

Tabela 1: Valor dos parâmetros utilizados para os experimentos.

5.2 Expansão dos blocos

Na divisão dos blocos utiliza-se a mediana das distâncias entre pontos homólogos quando as imagens são dispostas lado a lado no plano. Com isso, regiões de maior variação de paralaxe, ou seja, regiões de maior ou menor altitude, podem estar situadas do lado de fora dos limites do segundo bloco. Assim, a expansão dos blocos é importante para garantir que todas as regiões da imagem sejam analisadas pelo SIFT. Na figura 6 são apresentados os recortes de uma região da imagem original, realizados manualmente, após a geração da solução global. Os itens (a), (b) e (c) representam, respectivamente, os valores de κ iguais a 0 (sem expansão), 0, 1 e 0, 2.

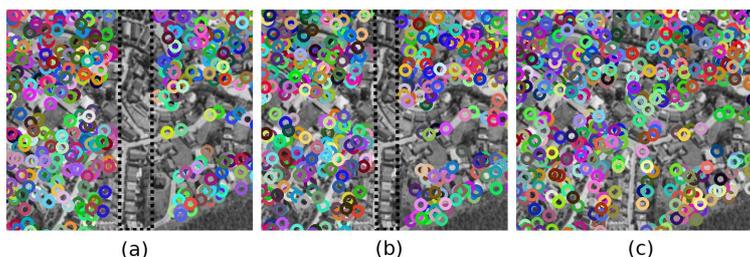


Figura 6: Pontos encontrados em blocos sem expansão.

Em Sun et al. (2014) são realizados recortes de 500×500 , com expansão da segunda imagem em 50 pixels em cada direção. O método aqui proposto utiliza o ângulo (α), descrito na seção

4.1, para realizar a expansão do bloco. Assim, a expansão do bloco é feita na direção de maior variação, evitando que o SIFT detecte *keypoints* que não possuem correspondência, ou seja, evita-se que *keypoits* sejam originados de regiões não relacionadas, os quais podem contribuir para a geração de falsas correspondências. Na figura 7 são apresentados os recortes de uma sequência presente nas imagens de Nova Friburgo. Da esquerda para direita, encontra-se o recorte da primeira imagem (500×500) e seu correspondente expandido na segunda imagem (698×500).

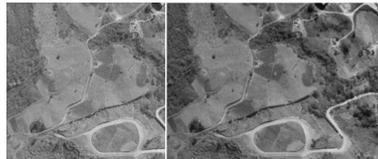


Figura 7: Exemplo de blocos da primeira imagem e o bloco expandido da segunda imagem.

6 RESULTADOS

Os resultados encontrados foram comparados com os resultados apresentados em Coelho (2012) e Sun et al. (2014). Na tabela 2 estão apresentados os resultados obtidos pelo método proposto e os resultados obtidos por Coelho (2012), com o uso das imagens de Nova Friburgo. Conforme pode ser observado, o método proposto não foi muito eficiente ao usar as imagens em formato reduzido, tendo encontrado uma quantidade inferior de pontos, em torno de 60% a menor, do que os pontos encontrados em Coelho (2012).

Para as imagens em tamanho original (sem redução de *pixel*) os resultados obtidos pelo método proposto foram muito significativos, sendo superior a 1000% nos cinco testes.

Pares de imagens	Coelho (2012)	Método proposto	
	(imagens reduzidas)	imagens reduzidas	imagens originais
<i>Img1 e Img2</i>	3785	1059	39728
<i>Img2 e Img3</i>	3729	1422	53154
<i>Img3 e Img4</i>	3414	1377	49062
<i>Img4 e Img5</i>	4192	1594	60834
<i>Img5 e Img6</i>	3916	1278	58863

Tabela 2: Comparação dos resultados encontrados em Coelho (2012) e pelo método proposto.

A redução de *pixels* faz com que informações da imagem sejam perdidas. O mesmo pode acontecer aplicando-se um filtro passa baixa para a redução de ruídos. A aplicação do filtro gaussiano pode justificar a perda de desempenho do método para as imagens reduzidas.

Os resultados obtidos pelo método proposto e os resultados obtidos pelo L^2SIFT , em Sun et al. (2014), são apresentados na tabela 3. Verificou-se que o número de correspondências encontradas foi aproximadamente 306% superior, comprovando novamente a eficácia do método proposto.

7 CONCLUSÃO

Neste artigo apresentou-se um método para extração de características em imagens estereoscópicas de alta resolução, tendo como objetivo maximizar o número de correspondências. Experimentos foram conduzidos para maximizar o desempenho do método proposto.

Parâmetro avaliado	L^2SIFT	Método proposto
Número de correspondências encontradas	113685	462058
Tempo de execução do método (minutos)	18,2	22,2

Tabela 3: Comparativo entre os resultados do L^2SIFT e o método proposto.

Constatou-se que a aplicação do filtro gaussiano diminui o tempo de execução do algoritmo e também maximiza o número de correspondências encontradas. Verificou-se também que a expansão dos blocos na direção de maior variação pode maximizar o número de pontos, uma vez que evita-se a geração de *keypoints* em regiões não relacionadas.

Em trabalhos futuros, os pontos obtidos pelo método proposto serão utilizados para a criação de superfícies tridimensionais. Serão investigados métodos de interpolação de pontos, a fim de realizar a reconstrução de terrenos com maior nível de detalhamento.

REFERÊNCIAS

- Bay H., Tuytelaars T., e Van Gool L. Surf: Speeded up robust features. *Computer vision–ECCV 2006*, páginas 404–417, 2006.
- Coelho A.M. Aplicação da geocomputação a estudos de susceptibilidade de deslizamentos de terra. 2012.
- do Carmo F.P. *Utilização da transformada de características invariante a escala (SIFT) na automatização da obtenção de pontos do Sistema de Imagens Tridimensional Híbrido (SITH)*. Tesis de Mestrado, Universidade do Estado do Rio de Janeiro, Nova Friburgo - RJ, 2010.
- Fernandes S.R. *Ferramenta de visão computacional para processos fotogramétricos*. Tesis de Mestrado, Universidade do Estado do Rio de Janeiro, Nova Friburgo - RJ, 2008.
- Jiang S. e Jiang W. Reliable image matching via photometric and geometric constraints structured by delaunay triangulation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 153:1–20, 2019.
- Lingua A., Marenchino D., e Nex F. Performance analysis of the sift operator for automatic feature extraction and matching in photogrammetric applications. *Sensors*, 9(5):3745–3766, 2009.
- Lowe D.G. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.
- Muja M. e Lowe D.G. Fast approximate nearest neighbors with automatic algorithm configuration. *VISAPP (1)*, 2(331-340):2, 2009.
- Muja M. e Lowe D.G. Scalable nearest neighbor algorithms for high dimensional data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(11):2227–2240, 2014.
- Rottensteiner F., Sohn G., Gerke M., e Wegner J.D. Isprs test project on urban classification and 3d building reconstruction. *Commission III-Photogrammetric Computer Vision and Image Analysis, Working Group III/4-3D Scene Analysis*, páginas 1–17, 2013.
- Rublee E., Rabaud V., Konolige K., e Bradski G. Orb: An efficient alternative to sift or surf. Em *Computer Vision (ICCV), 2011 IEEE International Conference on*, páginas 2564–2571. IEEE, 2011.
- Sun Y., Zhao L., Huang S., Yan L., e Dissanayake G. L2-sift: Sift feature extraction and matching for large images in large-scale aerial photogrammetry. *ISPRS Journal of Photogrammetry and Remote Sensing*, 91:1–16, 2014.
- Szeliski R. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.

- Umeyama S. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 13(4):376–380, 1991.
- Wei W., Jun H., e Yiping T. Image matching for geomorphic measurement based on sift and ransac methods. Em *Computer Science and Software Engineering, 2008 International Conference on*, volume 2, páginas 317–320. IEEE, 2008.
- Witkin A. Scale-space filtering: A new approach to multi-scale description. Em *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'84.*, volume 9, páginas 150–153. IEEE, 1984.