

MAPEAMENTO DE PONTOS CARACTERÍSTICOS POR COERÊNCIA DE TRIÂNGULOS

ROGÉRIO YUGO TAKIMOTO*, EDSON KENJI UEDA*, MARCOS DE SALES GUERRA TSUZUKI*,
EMILIANO GONÇALVES DE CASTRO*, THIAGO DE CASTRO MARTINS*, FÁBIO KAWAOKA TAKASE†

**Laboratório de Geometria Computacional, Departamento de Engenharia Mecatrônica e de Sistemas
Mecânicos, Escola Politécnica da Universidade de São Paulo
Avenida Prof. Melo Moraes, 2231, CEP 05508-030 - São Paulo, SP, Brasil.*

†*Mind Tecnologia e Conhecimento, São Paulo, Brasil*

Emails: takimotoyugo@gmail.com, ueda.edson@gmail.com, mtsuzuki@usp.br,
emiliano@usp.br, tcmartins@gmail.com, fktakase@gmail.com

Abstract— The 3D reconstruction of a scene using two uncalibrated images is a challenging task. The 3D reconstruction of the scene is realized in three steps: feature point determination by the SIFT (*Scale Invariant Feature Transform*), mapping of the feature points, and refinement of the mapping by the RANSAC (*RANdom SAmples Consensus*) with epipolar geometry evaluation. The determination of the correct epipolar geometry is a combinatorial problem, where the best set of 7 pairs of feature points is searched. As the RANSAC is a non deterministic algorithm that can not distinguish between incorrect and correct feature points mappings. In this work, it is assumed that three pairs of feature points preserve its topological orientation independently of how they are seen, what turns the RANSAC more robust. The error analysis is done using the epipolar distance, then the metric reconstruction is presented. Some results comparing the conventional RANSAC and the improved RANSAC are shown. The improve RANSAC showed a better performance.

Keywords— Fundamental matrix, epipolar geometry, 3D reconstruction, RANSAC, SIFT.

Resumo— A reconstrução 3D de uma cena utilizando duas imagens não calibradas é uma atividade complexa. A reconstrução 3D automática é realizada pelo encadeamento de três passos: determinação de pontos característicos pelo SIFT (*Scale Invariant Feature Transform*), mapeamento dos pontos característicos, e refinamento do mapeamento pelo RANSAC (*RANdom SAmples Consensus*) com avaliações da geometria epipolar. A determinação da geometria epipolar correta é um problema combinatório, em que se procura pelo melhor conjunto de 7 pares de pontos característicos. Como o RANSAC é um algoritmo não determinístico que não consegue discernir pares de pontos característicos incorretos dos corretos. Neste trabalho, assume-se que três pares de pontos característicos preservam a sua orientação topológica independentemente da forma como são vistos, tornando o RANSAC mais robusto. É feita a análise dos erros pela distância epipolar, em seguida a reconstrução métrica é apresentada. Alguns resultados comparam o RANSAC convencional e o proposto.

Palavras-chave— Matriz fundamental, geometria epipolar, reconstrução 3D, RANSAC, SIFT.

1 Introdução

A recuperação da informação 3D utilizando imagens 2D pode ser realizada através de duas abordagens: varredura por laser (Takimoto et al., 2013) e através da visão computacional. Na visão computacional, a partir de duas imagens é possível recuperar informações que podem ser usadas para a reconstrução 3D. Existem dois métodos principais representantes desta abordagem. O primeiro depende de uma prévia calibração da câmera para determinar sua posição relativa a um sistema de coordenadas de referência (Ito, 1991). Numa calibração tradicional, utiliza-se um objeto de referência espacial. A principal vantagem desta abordagem é a sua precisão, no entanto, a sua principal desvantagem ocorre em sistemas onde as características ótica e geométrica da câmera mudam dinamicamente. A segunda abordagem é baseada na recuperação da geometria epipolar a partir de duas imagens não calibradas. Uma aplicação utilizando a geometria epipolar foi primeiramente publicada por Longuet-Higgins (1981). A obtenção da informação 3D a partir de imagens pode

ser utilizada em diversas aplicações, por exemplo, na navegação de robôs móveis (van der Zwaan et al., 2002), medicina, arqueologia e arquitetura (Dias et al., 2006).

A recuperação da geometria epipolar entre câmeras não calibradas usualmente é realizada em três etapas: determinação dos pontos característicos, mapeamento dos pontos característicos, e estimação da geometria epipolar. Neste trabalho o SIFT (Lowe, 1999) é utilizado para determinar pontos característicos. O RANSAC será utilizado para determinar a geometria epipolar pelo mapeamento de pontos característicos (Fischler and Bolles, 1981). Esta é uma tarefa difícil, pois o RANSAC não modela o processo de mapeamento, e é uma caixa preta que cria diversas tentativas de correspondência sem direção.

Assim, neste trabalho, propõe-se um aprimoramento de modo a guiar o RANSAC com agrupamentos de pontos característicos segundo triângulos coerentemente orientados. É assumido que três pontos característicos visíveis em cenas próximas precisam preservar a sua orientação independentemente da forma como eles são vistos. A



Figura 1: Descritores SIFT.

geometria epipolar é determinada utilizando pelo menos 7 pares de pontos característicos. A informação 3D é recuperada utilizando as matrizes de câmera e uma referência adicional é fornecida para encontrar o fator de escala. Este artigo está estruturado da seguinte maneira: alguns conceitos básicos são apresentados na seção 2 (pontos característicos determinados pelo SIFT, cálculo da matriz fundamental por 7 pares de pontos correspondentes e refinamento da correspondência inicial pelo RANSAC), na seção 3 é apresentado o RANSAC com melhorias (incluindo a coerência topológica), a seção 4 apresenta a análise de erros, a seção 5 apresenta a reconstrução métrica, os resultados são apresentados na seção 5 e as conclusões na seção 6.

2 Determinação dos Pontos Característicos e Cálculo da Matriz Fundamental Utilizando Pontos Correspondentes

Os pontos característicos são determinados pelo SIFT (Lowe, 1999) que representa pontos característicos por meio de descritores (os descritores são vetores de características da imagem ou de determinadas regiões da imagem e podem ser usados para comparar regiões em imagens diferentes). Os vetores de características são invariantes à escala, translação, rotação e parcialmente invariantes à iluminação e transformação afim como a projeção 3D. O SIFT foi proposto por Lowe (1999) para extrair as características distintas e invariantes da imagem. A Fig. 1 mostra alguns descritores SIFT detectados na imagem.

Após a determinação dos pontos característicos, a correspondência entre os pontos característicos é feita através da semelhança métrica entre os descritores. Dentre essas métricas, as mais comuns incluem a soma do quadrado da diferença, a soma absoluta das diferenças e a correlação normal.

Supondo que existe uma aproximação inicial

para o mapeamento dos pontos característicos, a geometria epipolar (Hartley and Zisserman, 2000) fornece informações para identificar erros no mapeamento e refinar a aproximação inicial para o mapeamento dos pontos. Segundo esta geometria, qualquer ponto m no espaço 3D, forma com os centros de projeção das imagens um plano, que interceptará os dois planos de projeção (ou as duas imagens) em uma linha que passará necessariamente pelos pontos epipolares. A expressão matemática que relaciona os pontos correspondentes em imagens diferentes é (Luong et al., 1993)

$$m^T \cdot \mathbf{F} \cdot m' = 0 \quad (1)$$

onde \mathbf{F} é uma matriz 3×3 chamada de matriz fundamental. m é um ponto da imagem e m' é o seu ponto correspondente na outra imagem. Em particular, para $m = (x, y, 1)^T$ e $m' = (x', y', 1)^T$, cada par de pontos correspondentes determina uma equação linear em função dos elementos de \mathbf{F} . Considerando-se n pares de pontos correspondentes e denominando \mathbf{f} como um vetor de 9 elementos composto pelos coeficientes de \mathbf{F} , obtém-se

$$\mathbf{A} \cdot \mathbf{f} = 0 \quad (2)$$

onde \mathbf{A} representa um sistema homogêneo de equações lineares, e \mathbf{f} pode ser determinado a menos de um fator de escala. Para que exista uma solução, \mathbf{A} deve ter posto no máximo 8, caso em que a solução é única, determinada pelo gerador do espaço nulo à direita de \mathbf{A} .

A matriz fundamental será recuperada pelo algoritmo de 8 pontos correspondentes normalizado (Karlstrom, 2007). Este algoritmo envolve a solução por mínimos-quadrados de um sistema de equações. O algoritmo original foi formulado por Longuet-Higgins (1981), e a chave para o seu bom desempenho é a normalização adequada e sistemática dos pontos correspondentes. Neste algoritmo, uma transformação de translação e escalonamento aplicada sobre os pontos das imagens antes da construção do sistema de equações $\mathbf{A} \cdot \mathbf{f} = 0$ garante melhores condicionamento e estabilidade, sem aumento significativo da complexidade do algoritmo.

A normalização consiste na translação e escalonamento de cada imagem de forma que o centróide dos pontos de referência seja deslocado para a origem do sistema de coordenadas da imagem e a distância quadrática média destes pontos em relação à origem seja $\sqrt{2}$. No caso da matriz \mathbf{A} ter posto 7, é possível determinar um conjunto de soluções para a matriz fundamental utilizando-se a restrição de singularidade (já que \mathbf{F} tem 7 graus de liberdade). O caso mais importante é aquele em que são conhecidas apenas 7 pares de pontos correspondentes, gerando uma matriz $\mathbf{A}_{7 \times 9}$, geralmente de posto 7.

Pelo mapeamento de pontos característicos a matriz fundamental pode ser determinada. No en-

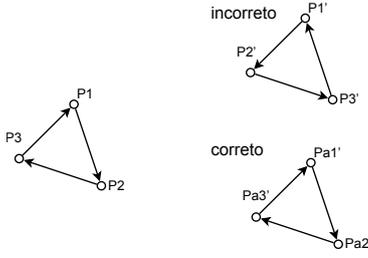


Figura 2: Triângulo à esquerda é referência e à direita existem dois triângulos, sendo um coerentemente orientado em relação à referência e outro incoerentemente orientado.

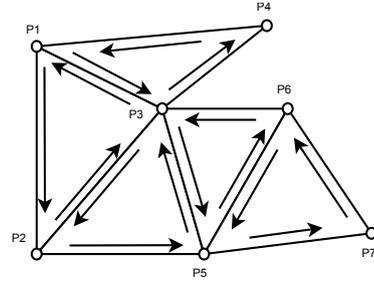


Figura 4: Modelo plano representando 7 pontos característicos selecionados na primeira imagem. Todos os triângulos estão coerentemente orientados.

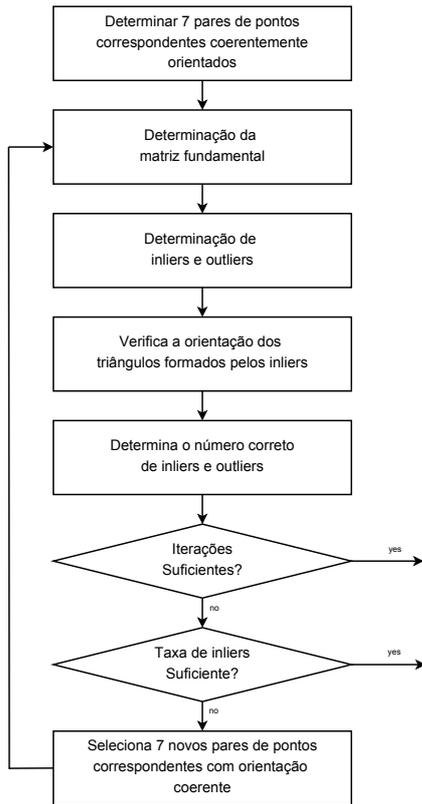


Figura 3: Algoritmo proposto.

tanto, na prática os pontos correspondentes podem estar mapeados de forma incorreta, ou seja, existe a possibilidade de se mapear falsos positivos devido à oclusão e à ambiguidade. O mapeamento incorreto dos pontos característicos são os outliers. Existe a necessidade de se determinar o conjunto de inliers a partir da correspondência para que a matriz \mathbf{F} seja melhor estimada. O RANSAC proposto por Fischler and Bolles (1981) é um estimador comumente utilizado para determinar o mapeamento correto de pontos característicos. Entretanto, o RANSAC não modela o processo de mapeamento, sendo possível que ele convirja para situações não reais.

3 Melhorias para o RANSAC

Este trabalho propõe uma melhoria no RANSAC ao incluir uma consistência topológica entre os triângulos mapeados. Três pontos característicos que estejam presentes em ambas as imagens devem ser coerentemente orientados. A Fig. 2 exibe três triângulos, sendo que o triângulo $[P_1, P_2, P_3]$ está na primeira imagem, e que os triângulos $[P'_1, P'_2, P'_3]$ e $[Pa'_1, Pa'_2, Pa'_3]$ estão na segunda imagem. Considere duas possíveis correspondências entre pontos característicos entre as duas imagens: $C_1 = [P_1 \leftrightarrow P'_1, P_2 \leftrightarrow P'_2, P_3 \leftrightarrow P'_3]$ e $C_2 = [P_1 \leftrightarrow Pa'_1, P_2 \leftrightarrow Pa'_2, P_3 \leftrightarrow Pa'_3]$. A correspondência definida por C_1 é incorreta pois ambos os triângulos possuem orientações incoerentes. Por outro lado, C_2 está correta. É relevante mencionar que pontos característicos colineares não podem ser utilizados na determinação da matriz fundamental.

Assim, os mapeamentos entre pontos característicos são definidos pelo algoritmo do vizinho mais próximo (Muja and Lowe, 2009) e por triângulos coerentemente orientados.

A Fig. 3 fornece um fluxograma do algoritmo proposto. O algoritmo inicia selecionando 3 pares de pontos característicos coerentemente orientados, em seguida um novo ponto é acrescentado, sempre formando triângulos coerentemente orientados. Fig. 4 exibe um conjunto de 7 pontos coerentemente orientados.

No segundo passo, a matriz fundamental é estimada e os *inliers* e *outliers* são identificados. Os *inliers* são pares de pontos característicos que estão em acordo com a geometria epipolar, enquanto que *outliers* são pares que estão em desacordo. Para este fim, um limiar é definido. Os inliers são agrupados em triângulos coerentemente orientados na imagem de referência (como exibido na Fig. 4) e verifica-se se eles estão mapeados a triângulos coerentemente orientados na outra imagem. Os inliers que estão coerentemente orientados nas duas imagens são identificados, e os outros são classificados como outliers.

A taxa de *inliers* é definida por $p_{in} = \frac{n_{in}}{N}$

onde n_{in} é o número de *inliers* e N é o número total de pares de pontos característicos. Se p_{in} é maior que um limiar então o algoritmo convergiu.

4 Análise de Erros

Como o ruído afeta de forma significativa o cálculo de \mathbf{F} , e pelo fato de existir a possibilidade de falsos positivos estarem presentes nos dados coletados, é muito importante avaliar o erro no cálculo de \mathbf{F} para que a informação geométrica possa ser interpretada corretamente. Além disso, a covariância da matriz fundamental pode ser usada para calcular a incerteza da reconstrução projetiva. É possível quantificar o erro a partir da distância epipolar (em pixels), que é a distância entre um ponto na imagem m e a sua linha epipolar correspondente l dado pela matriz \mathbf{F} . Dado \mathbf{F} , um par de pontos característicos $m \leftrightarrow m'$ e as linhas $l = \mathbf{F}^T m'$ e $l' = \mathbf{F} m$, a distância (em pixels) entre os pontos característicos e a linha pode ser calculada por

$$d = \frac{|m^T \mathbf{F}^T m'|}{\sqrt{(f_1^T m)^2 + (f_2^T m)^2}} \quad (3)$$

$$d' = \frac{|m'^T \mathbf{F} m|}{\sqrt{(f_1^T m')^2 + (f_2^T m')^2}} \quad (4)$$

onde f_j é a coluna j da matriz \mathbf{F} e f^j é a linha j da matriz \mathbf{F} . A soma da distância para N pontos correspondentes é

$$\sum (d(m'_i, \mathbf{F} m_i)^2 + d(m_i, \mathbf{F}^T m'_i)^2) \quad (5)$$

Nota-se que o erro é calculado considerando todas as correspondências obtidas na imagem, não somente os 7 usados no cálculo de \mathbf{F} . O RANSAC não minimiza o erro diretamente, mas utiliza o erro para determinar *inliers* e *outliers*.

5 Reconstrução Métrica

Após encontrar a matriz fundamental entre duas imagens, ela será utilizada em conjunto com informações adicionais para realizar a reconstrução real da cena. A reconstrução da cena é definida pelas suas matrizes de câmera, com as quais pode-se determinar a localização 3D dos pontos característicos das imagens.

Nesse processo de reconstrução, as matrizes de câmera \mathbf{P} e \mathbf{P}' , assim como os pontos 3D M_i são calculados de modo que

$$m_i = \mathbf{P} \cdot M_i \text{ e } m'_i = \mathbf{P}' \cdot M_i \quad (6)$$

para um conjunto de pontos correspondentes entre duas imagens $\{m_i \leftrightarrow m'_i\}$, onde tanto as matrizes de câmera (\mathbf{P}, \mathbf{P}'), como a posição e a orientação dos pontos 3D M_i são desconhecidos.



Figura 5: Imagens utilizadas na recuperação da geometria epipolar.



Figura 6: Linhas conectando os pontos característicos correspondentes obtida através do SIFT.

Caso haja pontos correspondentes suficientes para permitir que a matriz fundamental seja calculada unicamente, a cena pode ser reconstruída com uma ambiguidade projectiva. A ambiguidade na reconstrução pode ser reduzida fornecendo-se informações adicionais sobre a câmera ou cena.

Para cada imagem, tem-se $m = \mathbf{P} \cdot M$, $m' = \mathbf{P}' \cdot M$, e estas equações podem ser combinadas para formar um sistema $\mathbf{A} \cdot M = 0$, que é linear em M . Inicialmente, elimina-se o fator de escala homogêneo através do produto vetorial, resultando em três equações para cada ponto da imagem, duas das quais são linearmente independentes. Para a primeira imagem, $m' \cdot (\mathbf{P} \cdot M) = 0$, tem-se

$$\begin{aligned} x(p_3^T \cdot M) - (p_1^T \cdot M) &= 0 \\ y(p_3^T \cdot M) - (p_2^T \cdot M) &= 0 \\ x(p_2^T \cdot M) - y(p_1^T \cdot M) &= 0 \end{aligned} \quad (7)$$

onde os vetores p_i^T são linhas de \mathbf{P} . Uma equação da forma $\mathbf{A} \cdot X = 0$ pode ser reescrita

$$\mathbf{A} \cdot M = \begin{bmatrix} x(p_3^T \cdot M) - (p_1^T \cdot M) \\ y(p_3^T \cdot M) - (p_2^T \cdot M) \\ x'(p_3'^T \cdot M) - (p_1'^T \cdot M) \\ y'(p_3'^T \cdot M) - (p_2'^T \cdot M) \end{bmatrix} = 0 \quad (8)$$

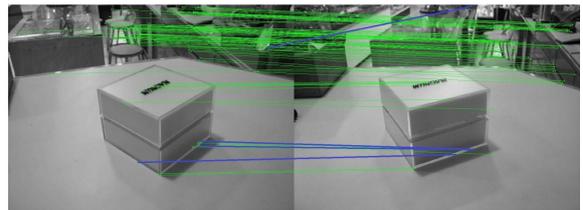


Figura 7: Linhas conectando os pontos característicos correspondentes obtido pelo RANSAC (falso positivos são mostrados em azul).

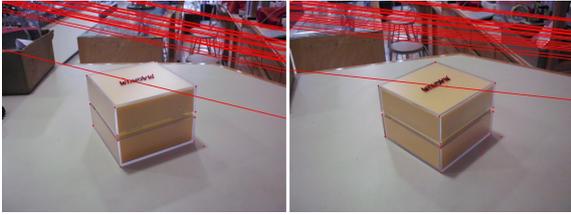


Figura 8: Linhas epipolares obtidas a partir da matriz fundamental.

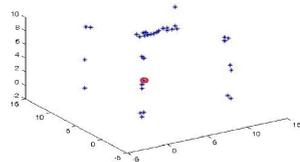
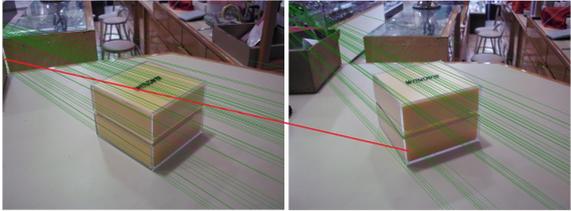


Figura 9: Linhas epipolares obtidas utilizando o método proposto. As linhas vermelhas indicam os falsos positivos e os círculos vermelhos indicam os falsos positivos no espaço 3D.

Este sistema de equações é sobre-determinado, uma vez que a solução é dada a menos de um fator de escala. A solução corresponde ao vetor singular relativo ao menor valor singular de \mathbf{A} e pode ser resolvido através da SVD (*Single Value Decomposition*) de \mathbf{A} . Este método não é projetivo-invariante e necessita de uma referência adicional para encontrar o fator de escala e a solução correta. Isto pode ser obtido através do conhecimento da localização de alguns pontos no sistema de coordenadas global.

6 Resultados

A recuperação da geometria epipolar de maneira automática envolveu a implementação de três algoritmos: o SIFT para a identificação dos pontos característicos, o cálculo da matriz fundamental e o RANSAC para o refinamento dos pontos correspondentes. A partir da geometria epipolar, recuperada com esses algoritmos, foi possível determinar corretamente as coordenadas global de um objeto na cena. As imagens utilizadas foram capturadas movendo-se a mesma câmera (vide Fig. 5).

Inicialmente o SIFT foi aplicado em ambas as imagens para se obter um conjunto de pontos correspondentes. Em seguida um mapeamento inicial foi determinado, o resultado deste mapeamento pode ser visualizado na Fig. 6. Neste ma-

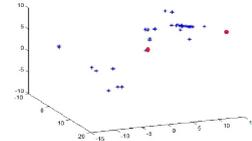
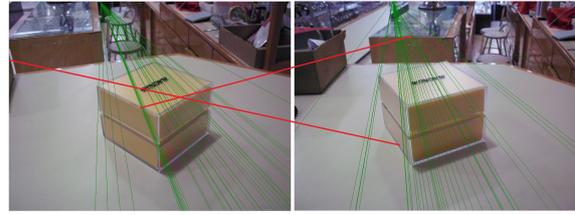


Figura 10: Linhas epipolares obtidas utilizando o método de RANSAC. As linhas vermelhas indicam os falsos positivos e os círculos vermelhos indicam os falsos positivos no espaço 3D.

peamento inicial existem alguns falsos positivos sendo necessário refinar a solução. A solução foi refinada pelo RANSAC com avaliação da geometria epipolar a partir do conjunto de pontos correspondentes (vide Fig. 10). O número de iterações para convergência do RANSAC foi reduzido com a nova proposta em que a orientação topológica entre triângulos mapeados é preservada. Ao final do processo, a geometria epipolar das duas imagens foi recuperada. A Fig. 8 mostra as linhas epipolares obtidas. É possível comparar ambos os algoritmos através da análise da Tabela 1, que mostra o número de falsos positivos juntamente com o erro para um dado threshold. É possível notar que o RANSAC utilizando o método proposto fornece menores números de falsos positivos e erros, apresentando melhoras na qualidade dos dados obtidos. Além disso, foi possível verificar que o erro no mapeamento dos pontos afeta a recuperação da matriz fundamental. À medida que aumenta os falsos positivos detectados como positivos, o erro da matriz fundamental aumenta levando a uma reconstrução com erros. A partir da matriz fundamental e após remover a ambiguidade projetiva, os pontos no espaço 3D foram reconstruídos. Os pontos 3D recuperados nesse processo geram uma nuvem de pontos que podem ser usados para modelar a superfície 3D do objeto. As Figs. 9 e 10 exibem os pontos 3D do mesmo objeto considerando diferentes vistas. É possível visualizar que o número de falsos positivos nas linhas epipolares é menor utilizando o método proposto quando comparado com o método de RANSAC clássico.

Ambos os algoritmos (RANSAC convencional e aprimorado) convergem para resultados distintos. O RANSAC convencional encontrou um conjunto de mapeamento de pontos característicos que definiu uma taxa de inliers suficiente para determinar a convergência do algoritmo para uma situação não realista. Enquanto que para todas as

Tabela 1: Comparação do RANSAC com o método proposto (PMet) para a “caixa na superfície”. Th = Threshold, FI = número de falsos positivos, GE = Erro geométrico (em pixels), T = Tempo de processamento (segundos)

Th	RANSAC			P Met		
	FI	GE	T	FI	GE	T
0,0001	4	0.66	3.0	1	0.50	12.5
0,0002	15	1.21	3.4	3	0.93	15.3
0,0003	7	2.46	3.8	4	1.45	17.0
0,0004	19	1.76	4.8	2	1.19	18.0
0,0005	10	1.72	5.4	1	1.36	21.5
0,0006	21	3.86	5.7	1	1.66	21.0
0,0007	10	3.04	5.7	4	1.79	21.0
0,0008	20	3.02	5.7	3	1.80	20.8
0,0009	6	3.28	5.6	2	1.04	20.4
0,0010	11	3.99	5.8	2	1.80	20.3
0,0020	18	7.82	5.7	5	3.01	20.4
0,0030	24	10.43	5.7	4	5.61	20.3
0,0040	24	26.34	5.8	2	6.03	20.3
0,0050	22	23.66	5.6	7	5.88	20.1
0,0060	30	29.58	5.6	8	6.53	20.2
0,0070	27	26.97	5.6	7	9.88	20.2
0,0080	30	46.63	5.3	19	1.99	20.2

execuções do RANSAC aprimorado, a convergência ocorreu para resultados reais.

7 Conclusões e Trabalhos Futuros

A partir dos resultados foi possível verificar que os algoritmos SIFT e RANSAC podem ser utilizados para recuperar os pontos 3D de uma cena. Além disso, a análise do sistema de coordenadas global mostrou que a reconstrução métrica do objeto foi obtida corretamente. O método proposto baseado em triângulos coerentemente orientados se mostrou eficiente na eliminação de mapeamentos incorretos, e a geometria epipolar foi determinada com menos iterações quando comparado ao algoritmo RANSAC convencional. Para trabalhos futuros, é necessário melhorar a qualidade da reconstrução 3D aplicando uma reconstrução da superfície.

Agradecimentos

Rogério Yugo Takimoto é suporta pela FAPESP (processo 2011/22402-8), Marcos de Sales Guerra Tsuzuki e Thiago de Castro Martins possuem suporte parcial do CNPq (respectivamente, processos 309.570/2010-7 e 306.415/2012-7), Edson Kenji Ueda é suportado pela CAPES e Emiliano Gonçalves de Castro é suportado pela FUN-DESPA.

Referências

- Dias, J. M. S., Bastos, R., Correia, J. and Vicente, R. (2006). Semi-automatic 3d reconstruction of urban areas using epipolar geometry and template matching, *Computer-Aided Civil and Infrastructure Engineering* **21**: 466–485.
- Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Commun. ACM* **24**: 381–395.
- Hartley, R. and Zisserman, A. (2000). *Multiple View Geometry in Computer Vision*, Cambridge University Press.
- Ito, M. (1991). Robot vision modelling-camera modelling and camera calibration, *Advanced Robotics* **5**: 321–335.
- Karlstroem, A. (2007). *Estimação de posição e quantificação de erro utilizando geometria epipolar entre imagens*, Mestrado, EPUSP, São Paulo.
- Longuet-Higgins, H. C. (1981). A computer algorithm for reconstructing a scene from two projections, *Nature* **293**: 133–135.
- Lowe, D. G. (1999). Object recognition from local scale-invariant features, *Proc Int Conf Comp Vision*, Corfu, Greece, pp. 1150–1157.
- Luong, Q. T., Deriche, R., Faugeras, O. and Papadopoulos, T. (1993). On determining the fundamental matrix: Analysis of different methods and experimental results, *Technical Report 1894*, INRIA.
- Muja, M. and Lowe, D. G. (2009). Fast approximate nearest neighbors with automatic algorithm configuration, *Int Conf Comp Vision Theory Appl*, pp. 331–340.
- Takimoto, R. Y., Vogelaar, R., Ueda, E. K., Sato, A. K., Martins, T. C., Gotoh, T., Kagei, S. and Tsuzuki, M. S. G. (2013). 3d reconstruction using low precision scanner, *11th IFAC Workshop on Intelligent Manufacturing Systems*, pp. 239–244.
- van der Zwaan, S., Bernardino, A. and Santos-Victor, J. (2002). Visual station keeping for floating robots in unstructured environments, *Robotics and Autonomous Systems* **39**: 145–155.