

Visão Computacional aplicada à Localização e Mapeamento Simultâneos de Robôs Subaquáticos

Paulo Drews Jr , Silvia Botelho

Fundação Universidade Federal do Rio Grande - Departamento de Física - NAUTEC

Av. Italia Km 8 - Rio Grande - RS - Brasil

paulo,silviacb@ee.furg.br

Resumo

O uso de veículos subaquáticos autônomos para inspeção visual é um campo promissor da robótica. Devido à dificuldade de localizar o robô e mapear o ambiente simultaneamente (SLAM), este trabalho propõe o uso de visão computacional e de mapas topológicos. Utilizando uma câmera de vídeo como fonte sensorial, essa abordagem é composta por dois estágios principais: i) aplicação do SIFT para extração de características em seqüências de imagens e ii) uso de mapas auto-organizáveis. O sistema desenvolvido foi validado em situações reais e simuladas, usando robôs reais em testes online. A precisão e a robustez obtidas em condições subaquáticas desfavoráveis, como variação de iluminação e ruído, conduzem a uma original e eficiente técnica de SLAM.

1. Introdução

Sendo o ambiente marinho rico em recursos biológicos e minerais, assim sua exploração é relevante e atrativa. Avanços tecnológicos vêm conduzindo ao desenvolvimento de veículos autônomos subaquáticos (*autonomous underwater vehicles* - AUVs) para realização das mais diferentes tarefas [9]. Devido às restrições associadas ao ambiente subaquático, bem como à longa duração das missões, busca-se diminuir a interferência humana na realização das atividades do veículo. Diversos tipos de sensores vêm sendo utilizados de forma a permitir uma percepção eficiente do meio, todos eles de custo elevado, tais como *Doppler Velocity Log* (DVL), *Acoustic Transponder Networks* (ATN) e *Inertial Navigation Systems* (INS) [3]. Entretanto, o conjunto de incertezas relacionadas ao meio ainda impedem a existência de uma solução de alto desempenho a custo exequível.

Por outro lado, AUVs quando utilizados para inspeção visual são equipados com uma câmera presa ao robô [3]. Essa câmera realiza a captura de imagens do fundo do mar, por meio das quais podem ser determinados marcos naturais (características naturalmente presentes no meio). Os

marcos são representados por descritores que compõem os mapas do ambiente navegado, viabilizando a localização e mapeamento simultâneos do robô (*simultaneous localization and mapping* - SLAM) ao longo da sua navegação. Assim, frente ao desafio, não trivial, associado ao problema de SLAM subaquático, o uso de informação visual surge como uma solução criativa, robusta, de baixo custo e de bom desempenho.

2. Trabalhos Relacionados

A proposta aqui apresentada é inovadora, embora existam diversos trabalhos na área de SLAM visual subaquático. Muitos dos trabalhos existentes são baseados em mosaicos. Um dos pioneiros em mosaicos subaquáticos, [8] propõe a busca de contornos por meio do laplaciano da gaussiana (LoG). A correlação por meio desse operador permite maior tolerância à iluminação não-uniforme. Embora o sistema desenvolvido operasse em tempo real, o movimento do robô era restrito a uma trajetória em colunas previamente conhecidas ("zigzag"). Outro trabalho que permitiu localização e mapeamento utilizando mosaicos, foi desenvolvido por [10]. Esse detectava características por meio do método KLT.

Embora tenham sido usados os mais diversos métodos para detecção de características em ambiente subaquático, não foram encontrados trabalhos usando o SIFT com essa finalidade. Uma abordagem com SIFT para SLAM visual foi realizada por [11]. Estes usaram o SIFT em um sistema visual estéreo para determinar marcos visuais, juntamente com a informação de odometria. Os testes foram feitos em ambiente terrestre estruturado com mapas conhecidos *a priori*. [2] apresentam a abordagem mais semelhante à proposta por este trabalho. Tais autores realizam SLAM visual com mapas topológicos baseados em aparência. Para isso, o método SIFT é usado em imagens omnidirecionais. Esta abordagem foi validada com um robô móvel em ambiente terrestre estruturado.

3. Sistemas Visuais em Ambiente Subaquático

Mesmo que se identifique diversas situações de aplicação do sistema aqui proposto, algumas restrições precisam ser respeitadas para a sua utilização, são elas: *i)* a navegação do veículo deve ser, aproximadamente, restrita a um plano, por exemplo, paralela ao fundo do mar, ou paralela a um casco de navio, com a câmera de vídeo orientada para tal plano; *ii)* a turbidez da água deve permitir uma visibilidade mínima; *iii)* assume-se que o terreno subaquático é razoavelmente plano, ou seja, tenha desníveis desprezíveis com relação à distância da câmera ao solo; *iv)* a condição de iluminação deve ser adequada; *vi)* a imagem precisa ter características distintivas mínimas; *vii)* o ambiente a ser explorado não pode ser excessivamente dinâmico, de forma a impossibilitar a inferência de movimento entre imagens.

4. Um Sistema para SLAM Visual

O sistema proposto neste trabalho é composto por diversos módulos, como mostrado na figura 1. Cada módulo será detalhado nas seções a seguir.



Figura 1. Overview do sistema proposto.

4.1. Pré-Processamento

As distorções causadas pelas lentes da câmera podem ser representadas por uma aproximação radial e tangencial. Como a componente radial causa a maior distorção, a maioria dos trabalhos desenvolvidos somente corrige esta componente [4] [3]. Em ambientes subaquáticos surge mais uma distorção, causada pela difração tanto na água como dentro da própria câmera [12]. Para um melhor desempenho, esta abordagem é implementada em uma LUT (*Lookup Table*).

4.2. SIFT

O SIFT é um método robusto para extrair e descrever características de uma imagem [7]. O algoritmo é composto

por 4 etapas. A **detecção de extremos no espaço de escala** na qual é feita a busca de características em todas as escalas e localizações de imagens com diferença de filtros gaussianos (DoG), estas invariantes a escala. A **localização de pontos-chaves** onde é determinada a localização exata e escala de cada *keypoint*, nessa etapa são definidos os melhores pontos para o mapeamento através do gradiente. Também é feita a **definição da orientação** para cada *keypoint* por meio dos gradientes locais ao ponto, obtendo-se invariância a rotação. Por fim, são determinados os **descritores dos pontos-chaves** onde o gradiente local de cada ponto-chave é medido, utilizando-se a vizinhança do ponto. Estas medidas são transformadas em uma representação que permite tolerância a níveis significativos de distorção e mudança de iluminação.

4.3. Matching e Estimação da Matriz de Homografia

O melhor candidato para correlacionar cada ponto-chave é encontrado pela identificação de seu vizinho mais próximo. Entretanto, muitos pontos de uma imagem não têm correspondentes na imagem anterior. Para eliminar esses falsos *matches*, utiliza-se um método que compara a menor distância com a segunda menor distância [7]. Além disso, o sistema usa a estimação da matriz fundamental para remover *outliers*, ou falsos verdadeiros. Essa matriz é estimada pelos métodos robustos RANSAC e LMedS, separadamente. Então, são usados os resultados obtidos por ambos os métodos, para validar cada correlação.

A correlação entre pontos de duas imagens distintas permite a determinação do movimento relativo entre elas. Os n pares de pontos são usados para determinar a matriz de homografia, H , que permite estimar o movimento entre estas imagens [5].

4.4. Mapas Topológicos

Neste trabalho, é proposto o uso de mapas topológicos, ou seja, mapas que baseiam-se na noção de um espaço não quantitativo. Eles são construídos usando os descritores dos pontos extraídos pelo SIFT e a pose adquirida da homografia. Tais informações são usadas em mapas auto-organizáveis, baseados nas redes de Kohonen [6]. A estrutura do mapa consiste de nodos, ou neurônios, tais conectados por arestas, que representam a idéia de vizinhança. Associado com cada neurônio existe um vetor de pesos, que tem a mesma dimensão do descritor do SIFT, além da posição e orientação.

Durante a navegação, os descritores são apresentados ao mapa. Quando um novo caso de treinamento é apresentado, o mapa determina se um neurônio deve ser criado ou atu-

alizado, eles ainda podem ser excluídos ou confirmados. O mapa topológico, depois de suficientemente treinado, permite a navegação de duas formas: por posições objetivo ou por marcos visuais. A partir da posição atual do robô, um algoritmo de busca em grafos Dijkstra ou algoritmo A^* , é usado para buscar a posição alvo, dependendo do desempenho desejado e do tamanho do mapa.

5. Implementação, Testes e Resultados

O presente trabalho foi totalmente implementado, sendo validado a partir de um conjunto de testes envolvendo missões simuladas e reais. O sistema foi acoplado ao veículo ROVFURG-II apresentado na figura 2. Esse robô é equipado com uma *Tritech Typhoon*, um sonar *Tritech MiniKing* e um conjunto de sensores (altímetro e acelerômetros). O sistema foi testado em um computador Intel Core 2 Quad Q6600 com 2Gb de RAM DDR2-667. A câmera usada tem padrão NTSC, porém utilizou-se frames de 320x240 com uma taxa de até 29,97 fps.



Figura 2. ROVFURG-II durante testes.

Visando verificar a robustez da proposta, frente a diferentes condições do meio, foram aplicados diferentes filtros às imagens [1], de forma a enaltecer determinadas características subaquáticas, como turbidez, neve marinha, iluminação não-linear e outras. A tabela 1 apresenta os diferentes filtros aplicados.

Característica / Filtro	1	2	3	4	5
Distância da Fonte de Luz (m)	0.2	0.22	0.25	0.25	0.3
Atenuação (%)	0.05	0.05	0.06	0.05	0.05
Ruído Gaussiano (σ)	2	2	2	4	4
Mínimo Nível de Cinza	20	30	20	20	20
Núm. de Flocos de Neve Marinha	30	30	30	30	30

Tabela 1. Características subaquáticas para cada distorção usada nos testes.

5.1. Localização em diferentes condições subaquáticas

O sistema visual foi testado usando as cinco diferentes condições de ambiente subaquático, correspondendo a cada

(filtro i) apresentado na tabela 1. A figura 3, na parte superior, enumera os pontos detectados e correlacionados obtidos em uma missão. Embora o número de pontos e de correlações tenham diminuído com a perda de qualidade devido aos filtros aplicados à imagem, ainda assim foi possível localizar o robô, como mostra a figura 3, na parte inferior. Nesta figura, o movimento de referência é apresentado em azul, executado por um braço robótico composto por um atuador do tipo *harmonic drive* com encoder acoplado, fornecendo uma informação angular a cada 0,000651 segundos, com uma câmera acoplada. Nota-se que a abordagem proposta é robusta à mudanças de condições do ambiente subaquático.

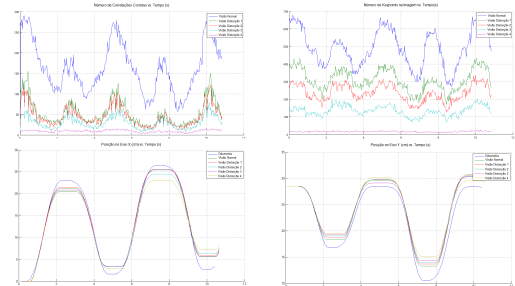


Figura 3. Número de keypoints e correlações verdadeiras, além das posições nos eixos euclidianos, durante movimento de um braço robótico, frente a diferentes filtros.

5.2. Localização Robótica Online

Foram realizados testes para avaliar o desempenho do SIFT, considerando como comparação um algoritmo clássico na localização robótica em ambiente subaquático, o KLT [10]. Em tais testes, utilizou-se a metodologia completa, apenas mudando o algoritmo de localização visual. A figura 4 mostra o resultado de desempenho dos algoritmos. A posição final está representada por um 'X' verde e as medidas em centímetros. SIFT obteve uma taxa média de 4,4 fps (frames por segundos), sem distorção, e 10,5 fps com o filtro 5. O KLT apresentou maiores taxas de frames processados, 13,2 fps e 13,08 fps, respectivamente. Vale notar que o SIFT apresentou um melhor desempenho quando as imagens tinham qualidade inferior, isto se deve ao fato do custo computacional estar relacionado, principalmente, com a determinação dos descritores locais dos pontos, que estão em menor número nesse caso. No entanto, como a dinâmica associada com o movimento é lenta, ambos os métodos podem ser aplicados ao SLAM online.

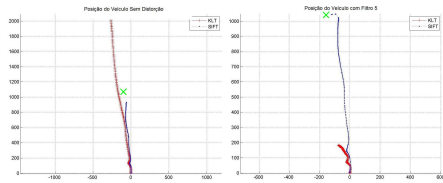


Figura 4. Localização *online* do robô, com e sem distorção.

5.3. Mapas Topológicos Auto-organizáveis

Testes para validar o sistema de mapeamento proposto foram realizados. A figura 5 mostra o mapa final, usando imagens adquiridas durante a navegação do veículo. Tal mapa pode ser usado para localizar o veículo, e auxiliar a navegação por meio de alvos visuais além de poder ser utilizado para corrigir a pose do robô, utilizando informações previamente apresentadas ao mapa, validando o SLAM e questões online associadas à tarefa de inspeção do AUV.

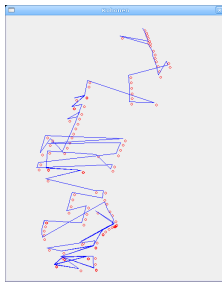


Figura 5. Mapa topológico gerado pelo movimento do ROVFURGII.

6. Conclusão

O trabalho apresentou uma nova abordagem para SLAM subaquáticos, utilizando apenas a informação visual *online*, facilmente extensível a outros ambientes, como aéreo e subterrâneo. Tal sistema pode ser utilizado em tarefas de inspeção autônomas, ou no auxílio ao controle do robô. Diversos testes foram realizados com diferentes condições subaquáticas. A efetividade da proposta foi validada em um conjunto de cenários reais. Os resultados mostraram a vantagem de utilizar SIFT, em relação a outros métodos tradicionais como o KLT, devido à sua invariância à escala e rotação, além desse ser tolerante a variações de iluminação e perspectiva. O mapeamento topológico, apesar de pouco difundido, também mostrou potencial para a aplicação, sendo

robusto a imprecisões sensoriais além de apresentar baixo custo computacional e possibilitar a correção da pose do robô.

Como trabalhos futuros, é proposta uma análise detalhada do sistema de mapeamento topológico, executando um conjunto de testes com diferentes cenários e parâmetros. Também é proposto o uso da informação de escala fornecida pelo SIFT em conjunto com a informação do altímetro do veículo, permitindo a estimação do movimento de profundidade do veículo. A utilização de visão estereoscópica também é uma possibilidade, podendo conduzir a melhores resultados e maior aplicabilidade do método.

7. Agradecimentos

O trabalho apresentado está relacionado ao Projeto TEC-SUB financiado pelo CNPq, e foi inteiramente desenvolvido pelo aluno ao longo da duração de sua bolsa. Agradece-se o auxílio experimental dos pesquisadores dos laboratórios NAUTEC e NuMA da FURG.

Referências

- [1] M. Arredondo and K. Lebart. A methodology for the systematic assessment of underwater video processing algorithms. In *IEEE/OES Oceans*, pages 362–367, 2005.
- [2] O. Booij, B. Terwijn, Z. Zivkovic, and B. Krose. Navigation using an appearance based topological map. In *IEEE ICRA*, pages 3927–3932, April 2007.
- [3] R. Garcia. *A Proposal to Estimate the Motion of an Underwater Vehicle Through Visual Mosaicking*. PhD thesis, Universitat de Girona, 2001.
- [4] N. Gracias, S. Van der Zwaan, A. Bernardino, and J. Santos-Vitor. Results on underwater mosaic-based navigation. In *MTS/IEEE Oceans*, volume 3, pages 1588–1594, 10 2002.
- [5] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004.
- [6] T. Kohonen. *Self-Organizing Maps*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2001.
- [7] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
- [8] R. Marks, S. Rock, and M. Lee. Real-time video mosaicking of the ocean floor. *IEEE Journal of Oceanic Engineering*, 20(3):229–241, July 1995.
- [9] J. R. McFarlane. Underwater technology 2000 rovs and auvs: tools for exploring, exploiting and defending the ocean frontier. *Underwater Technology*, pages 465–471, 2000.
- [10] K. Plakas and E. Trucco. Developing a real-time, robust, video tracker. In *MTS/IEEE Oceans*, pages 1345–1352, 2000.
- [11] S. Se, D. Lowe, and J. Little. Vision-based global localization and mapping for mobile robots. *IEEE Transactions on Robotics*, 21(3):364–375, June 2005.
- [12] X. Xu and S. Negahdaripour. Vision-based motion sensing for underwater navigation and mosaicking of ocean floor images. In *MTS/IEEE Oceans*, pages 1412–1417, 1997.