

Uma Abordagem Baseada em Visão Computacional para Localização e Mapeamento Simultâneos de Robôs Subaquáticos

Paulo Drews Jr¹, Silvia Botelho¹

¹Departamento de Física – Fundação Universidade Federal do Rio Grande (FURG)
96.203-000 – Rio Grande – RS – Brasil

{paulo, silvia}@ee.furg.br

Abstract. *The use of autonomous underwater vehicles for visual inspection tasks is a promising robotic field. Due to the difficulty of simultaneous robot localization and mapping (SLAM), this paper proposes a solution based on computer vision and topological maps. By means of an inspection camera as sensorial source, this approach is composed by two main stages: i) the application of SIFT to extract the features in underwater image sequences and ii) the use of self-organized maps. The developed system was validated in real and simulated environments, using real robots in online tests. Both accuracy and robustness attained in unfavorable underwater conditions, such as illumination variation and noise, lead to an original and efficient SLAM technique.*

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. Por meio de uma câmera de inspeção como fonte sensorial, essa abordagem é composta por dois estágios principais: i) aplicação do SIFT para extração de características em sequências 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 (possuindo cerca de 97% da biodiversidade) e minerais (grandes reservas de petróleo, ouro, diamante, etc.), a exploração marinha torna-se uma área relevante e atrativa. Avanços tecnológicos vêm conduzindo ao desenvolvimento de veículos autônomos subaquáticos (*autonomous underwater vehicles* - AUVs) para o auxílio e realização das mais diversas tarefas no fundo do mar [McFarlane 2000].

Devido às imposições e restrições associadas ao ambiente subaquático, bem como à longa duração da maioria das missões, busca-se diminuir a necessidade de interferência humana na realização das atividades do veículo, dando a ele maior autonomia. Parte dos desafios associados à autonomia está relacionada à possibilidade do robô saber, com algum grau de certeza, “onde está”, “por onde andou” e “para onde vai”. Tais desafios

decorrem da pouca estruturação desse ambiente, o qual pode ser, muitas vezes, totalmente desconhecido. A imersão em ambiente aquoso também dificulta a percepção e a determinação da pose do robô (posição e orientação), devido, principalmente, a forte turbidez da água, partículas em suspensão e presença de correntes marinhas.

Visando ao desenvolvimento de sistemas de percepção eficazes, diversos tipos de sensores vêm sendo utilizados, todos eles de custo elevado, tais como *Doppler Velocity Log* (DVL), *Acoustic Transponder Networks* (ATN) e *Inertial Navigation Systems* (INS) [Garcia 2001]. Entretanto, o conjunto de incertezas relacionadas ao meio ainda impedem a existência de uma solução de alto desempenho a custo exequível.

Por outro lado, veículos subaquáticos autônomos quando utilizados para inspeção visual em ambientes subaquáticos são, geralmente, equipados com uma câmera *down-looking* presa à estrutura do robô [Garcia 2001]. 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 simultâneo 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 e seu mapeamento surge como uma solução criativa, robusta, de baixo custo e de boa performance.

2. Trabalhos Relacionados

A proposta aqui apresentada é inovadora, embora existam diversos trabalhos na área de SLAM Visual em ambiente subaquático. Muitos dos trabalhos existentes são baseados em mosaicos, usando métodos para extração de características. Ainda assim, há abordagens que utilizam métodos não-baseados em características *Brightness Constancy Model*. Destes, [Negahdaripour and Khamene 2000] destaca-se como um método com baixo custo computacional, entretanto com pouca tolerância a variação de iluminação.

Um dos pioneiros em mosaicos subaquáticos, [Marks et al. 1995] propõe a busca de contornos por meio do operador LoG (do inglês, *laplacian of the gaussian*). 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 [Plakas and Trucco 2000]. Esse detectava características por meio do método KLT, sendo limitado em 30 características por imagem, com uma taxa de 10 a 20 frames por segundos.

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 [Se et al. 2005]. Estes usaram o SIFT em um sistema visual estéreo para determinar marcos visuais, juntamente com a informação de odometria. Estimacão *ego-motion* e o filtro de Kalman também foram usados. Os testes foram feitos em ambiente terrestre estruturado com mapas conhecidos *a priori*.

[Mahon and Williams 2004] propuseram um sistema visual para SLAM subaquático usando filtro ótico e filtro de Kalman estendido com o auxílio de um sonar,

com um conhecimento prévio da textura presente no ambiente. [Booij et al. 2007] 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

A proposta deste trabalho é utilizar descritores, obtidos a partir do processamento de imagens subaquáticas, para a localização e mapeamento de robôs autônomos.

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 AUV 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.

3.1. Imagens Subaquáticas

A aplicação de técnicas de visão computacional para imagens subaquáticas trata com alguns problemas adicionais à visão tradicional. Eles surgem, principalmente, devido às propriedades de transmissão do meio. As propriedades ópticas dos corpos dentro da água dependem da interação entre a luz e o ambiente aquático. Esta interação inclui, basicamente, dois processos: absorção e espalhamento [Xu and Negahdaripour 1997].

A absorção é o processo no qual a energia da luz é convertida em outra forma (principalmente calor). O espalhamento é produzido por mudança de direção dos fótons, devido à colisão com partículas em suspensão na água. Esse fenômeno pode ser dividido em retro-dissipação, que surge quando a luz é refletida na direção da câmera, e em pró-espalhamento, que é produzido quando a luz é refletida pelo objeto, alterando sua direção. Esse efeito produz um borramento na imagem do objeto.

Além das questões associadas à iluminação, devido, principalmente, à absorção e iluminação artificial, o processamento de imagens subaquáticas deve considerar outros fenômenos. Um exemplo importante é a *neve marinha*, que ocorre devido a presença de partículas macroscópicas na água, fazendo com que a análise das imagens seja dificultada. Outra questão é a falta de características em imagens de ambientes subaquáticos. As imagens, muitas vezes, não têm características relevantes, ou seja, pontos, linhas ou contornos, comumente explorados pelos sistemas de visão convencional [Garcia 2001].

Entretanto, apesar dos diversos problemas citados, a aplicação de técnicas de visão em ambiente subaquático apresenta uma vantagem importante em relação ao ambiente terrestre. Essa vantagem deve-se à característica estática do fundo do mar [Negahdaripour and Khamene 2000], que facilita a detecção de movimento.

4. Um Sistema para SLAM Visual

O sistema proposto neste trabalho é composto por diversos módulos, como mostrado na figura 1. A imagem é capturada por uma câmera e pré-processada para remoção de distorções. Com a imagem corrigida, o método SIFT é utilizado, visando à determinação de pontos-chaves e seus descritores. Os pontos são ordenados pelo maior gradiente, sendo excluídos pontos muito próximos entre si.

Para determinar a correlação entre os pontos, esses são buscados na imagem anterior. Para isso, usa-se a distância euclidiana entre os descritores. De todos os pontos correlacionados, alguns são falsos e devem ser removidos, de forma a não degradar o desempenho do sistema. Assim, é usada a matriz fundamental com estimação robusta, por meio dos métodos RANSAC e LMedS [Stewart 1999].



Figura 1. Overview do sistema proposto.

Com os pares de pontos correlacionados, o movimento relativo entre os frames é estimado, usando-se a matriz de homografia. Assim, é possível determinar a posição e a orientação, tanto do robô quanto dos melhores pontos-chaves, determinados na etapa do SIFT. Essas informações são adicionadas a um mapa topológico, tornando possível o mapeamento do ambiente e a localização do robô em tempo *online*¹. A seguir detalha-se cada uma das etapas da proposta.

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 [Gracias et al. 2002] [Garcia 2001].

Em ambientes subaquáticos surge mais uma distorção, causada pela difração tanto na água como dentro da própria câmera [Xu and Negahdaripour 1997]. Essa pode ser corrigida por meio da equação 1, onde $R = \sqrt{x_d^2 + y_d^2}$ é a distância do ponto à origem, f é a distância focal, $P_0 = (x_0, y_0)$ é o ponto sem distorção radial, $P_d = (x_d, y_d)$ é o novo ponto, sem distorção adicional, e $P_c = (x_c, y_c)$ é o centro da imagem. Para um melhor desempenho, esta abordagem é implementada em uma LUT (do inglês, *Lookup Table*).

¹O tempo *online* é obtido pelo descarte dos frames adquiridos pela câmera enquanto o sistema está ocupado, processando o frame corrente.

$$R_0 = f \tan(\sin^{-1}(1.33 * \sin(\tan^{-1} \frac{R}{f}))); P_0 = P_c + \frac{R_0}{R}(P_d - P_c) \quad (1)$$

4.2. SIFT

O SIFT é um método robusto para extrair e descrever pontos-chaves de uma imagem [Lowe 2004]. O algoritmo é composto por 4 etapas:

1. **Detecção de extremos no espaço de escala:** Nessa etapa, é feita a busca em todas as escalas e localizações de imagens com diferença de filtros gaussianos (DOG, do inglês *difference of gaussian*), visando identificar pontos de interesse invariáveis à escala e rotação.
2. **Localização de pontos-chaves:** Para cada localização em que foi detectado um extremo, um modelo detalhado é ajustado para determinar a localização exata e a escala. Pontos-chaves são selecionados baseando-se em medidas de estabilidade. Nessa etapa são definidos os melhores pontos para o sistema de mapeamento, por meio de medidas de gradiente.
3. **Definição de orientação:** A orientação é definida para cada ponto chave por meio dos gradientes locais da imagem. Toda operação, a partir de então, será feita com relação a dados da imagem transformados em relação à orientação e escala de cada ponto-chave. Desta maneira, obtém-se invariância a estas transformações.
4. **Descritor dos pontos-chaves:** O gradiente local de cada ponto-chave é medido, utilizando-se a vizinhança do ponto. Estas medidas são transformadas para uma representação que permite tolerância a níveis significativos de distorção e mudança de iluminação.

4.3. Matching

O melhor candidato para correlacionar cada ponto-chave é encontrado pela identificação de seu vizinho mais próximo. Esse é definido como o *keypoint* com a menor distância euclidiana entre os descritores dos pontos. 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 [Lowe 2004].

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.

4.4. Estimação da Matriz de Homografia

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 imagem.

A posição global do robô pode ser estimada pela equação 2, onde ${}^1H_{k+1}$ é a matriz de homografia entre a imagem I_1 e a imagem I_{k+1} . A matriz 1H_1 é definida pela matriz identidade, visto que, inicialmente, considera-se o robô na origem e sem rotação.

$${}^1H_{k+1} = \prod_{i=1}^k {}^iH_{i+1} \quad (2)$$

4.5. 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, desenvolvidos a partir de redes neurais de Kohonen [Kohonen 2001] e de estruturas celulares crescentes (GCS) [Fritzke 1993]. A estrutura do mapa consiste de nodos, ou neurônios. Os nodos são conectados por arestas, que representam a idéia de vizinhança. Associado com cada neurônio existe um vetor de pesos, que tem a mesma dimensão do vetor de dados de entrada, além da posição e orientação.

Durante a navegação, os descritores são apresentados ao mapa. O algoritmo de aprendizado é baseado no conceito de vizinho mais próximo. Quando um novo caso de treinamento é apresentado, o mapa determina o nodo que apresenta menor distância ao vetor apresentado. Durante o processo, neurônios são criados, excluídos e 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 Colour Underwater Video Camera*, 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 [Arredondo and Lebart 2005], de forma a enaltecer determinadas características subaquáticas, como turbidez, neve marinha, iluminação não-linear e outras. A tabela 1 apresenta os diferentes filtros aplicados.

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

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

Característica	Filtro 1	Filtro 2	Filtro 3	Filtro 4	Filtro 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

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.

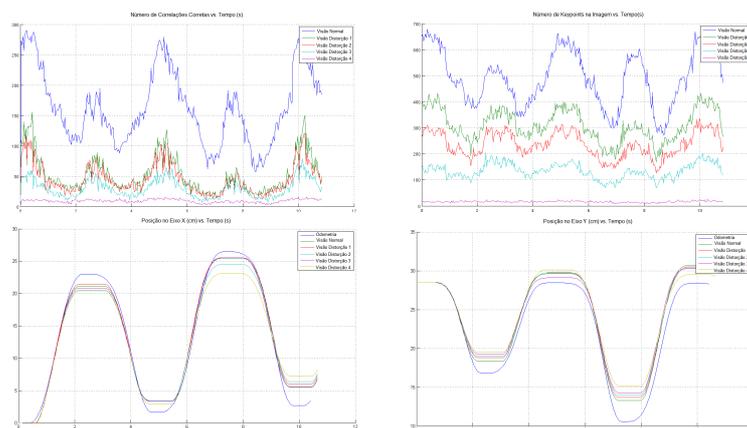


Figura 3. Número de *keypoints* e correlações verdadeiras , além das posições nos eixos euclidianos, durante movimento de um braço robótico, frente a diferentes filtros.

5.2. Localização Robótica em Tempo Online

Alguns testes foram realizados para avaliar a *performance* do SIFT, considerando como comparação outro algoritmo bastante usado em localização robótica em ambiente subaquático, o KLT [Plakas and Trucco 2000], na versão de Shi e Tomasi [Shi and Tomasi 1994]. Os testes usando KLT foram realizados com a implementação de [Birchfield 2007] com algumas modificações, como a busca de novos pontos a cada cinco imagens processadas, pois facilmente se perdem as características em condições subaquáticas.

A figura 4 mostra o resultado de desempenho dos algoritmos SIFT e KLT. SIFT obteve uma taxa média de 4,4 fps (frames por segundos), sem distorção, e uma taxa de 10,5 fps com o uso do **filtro 5**. O KLT apresentou maiores taxas de frames processados, 13,2 fps e 13,08 fps, respectivamente. Vale notar que o SIFT apresentou uma *performance* melhor 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. O KLT manteve sua performance quase constante. No entanto, como a dinâmica associada com o movimento do veículo subaquático é lenta, ambos os métodos podem ser aplicados ao SLAM em tempo *online*.

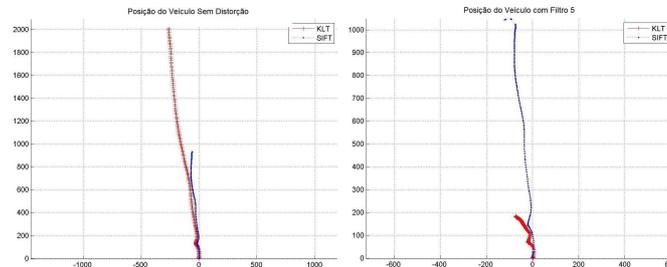


Figura 4. Localização do robô real *online*, com e sem distorção.

5.3. Robustez a Variação de Escala

Foram realizados testes para determinar a robustez do sistema proposto a variação brusca de escala. Assim, foi realizado um movimento de translação na câmera com a utilização contínua e variante do *Zoom*, simulando um movimento do robô em condições críticas de variação de escala.

A figura 5 mostra o resultado do SIFT e KLT, com a adição de filtros críticos relacionados à qualidade da imagem (tabela 1). O resultado do SIFT foi superior ao KLT, embora apresentando um pequeno erro em Y, em condição extrema. Nos testes, o SIFT apresentou uma taxa média de 6,22 fps, quando usando a imagem capturada pela câmera; de 7,31 fps, usando o **filtro 1**; e 10,24 fps usando o **filtro 5**. Já o KLT apresentou 12,5, 10,2 e 11,84 fps, respectivamente.



Figura 5. Localização com movimento de translação e escala, com e sem distorção.

5.4. Mapas Topológicos Auto-organizáveis

Testes para validar o sistema de mapeamento proposto foram realizados. A figura 6 mostra o mapa final, usando imagens adquiridas durante a navegação do veículo. Este mapa pode ser usado para localizar o veículo, e auxiliar a navegação por meio de alvos visuais, validando o SLAM e questões online associadas à tarefa de inspeção do AUV.

6. Conclusão

O trabalho apresentou uma nova abordagem para localização e mapeamento simultâneos de robôs subaquáticos, utilizando apenas a informação visual em tempo *online*. O uso

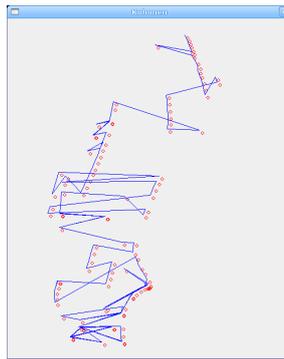


Figura 6. Mapa topológico gerado pelo movimento do ROVFURGII.

de visão computacional como fonte sensorial mostrou-se uma solução de baixo custo, com bom desempenho e robusta. Tal sistema pode ser utilizado em tarefas de inspeção autônomas, ou no auxílio ao controle em malha fechada do robô, no caso de existir um operador remoto.

Diversos testes foram realizados com diferentes características subaquáticas. A efetividade da proposta foi validada em um conjunto de cenários reais, com diferentes níveis de turbidez, neve marinha, iluminação não-uniforme, ruído, entre outros. Os resultados mostraram a vantagem de utilizar SIFT, em relação a outros métodos como o KLT, devido à sua invariância à escala e rotação, além desse ser tolerante a variações de iluminação e perspectiva.

Considerando o desempenho temporal do sistema, a proposta pode ser usada para SLAM visual online em AUV, sob condições extremas de ambiente. Vale lembrar que esta técnica pode ser estendida a outros ambientes com as devidas adaptações, como por exemplo o ambiente aéreo, sendo utilizada em veículos aéreos não-tripulados.

A correlação de pontos de interesse foi satisfatória, mesmo com a presença de falsos verdadeiros. O uso da matriz fundamental com estimação robusta, com uso dos algoritmos RANSAC e LMedS mostrou bons resultados. O mapeamento topológico baseado nas redes de Kohonen e nas GCS, apesar de pouco difundido, mostrou potencial para aplicações de SLAM subaquático utilizando informação visual, sendo robusto a impressões sensoriais além de apresentar baixo custo computacional.

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 TECSUB 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

- Arredondo, M. and Lebart, K. (2005). A methodology for the systematic assessment of underwater video processing algorithms. In *IEEE/OES Oceans*, volume 1, pages 362–367.
- Birchfield, S. (2007). KLT 1.3.4 : A C implementation of KLT tracker. <http://www.ces.clemson.edu/stb/kl/>. [Acessado em 12 de Março de 2008.].
- Booij, O., Terwijn, B., Zivkovic, Z., and Krose, B. (2007). Navigation using an appearance based topological map. In *IEEE ICRA*, pages 3927–3932.
- Fritzke, B. (1993). Growing cell structures - a self-organizing network for unsupervised and supervised learning. Technical report, University of California.
- Garcia, R. (2001). *A Proposal to Estimate the Motion of an Underwater Vehicle Through Visual Mosaicking*. PhD thesis, Universitat de Girona.
- Gracias, N., Van der Zwaan, S., Bernardino, A., and Santos-Vitor, J. (2002). Results on underwater mosaic-based navigation. In *MTS/IEEE Oceans*, volume 3, pages 1588–1594.
- Kohonen, T. (2001). *Self-Organizing Maps*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
- Lowe, D. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110.
- Mahon, I. and Williams, S. (2004). SLAM using natural features in an underwater environment. In *IEEE ICARCV*, volume 3, pages 2076–2081.
- Marks, R., Rock, S., and Lee, M. (1995). Real-time video mosaicking of the ocean floor. *IEEE Journal of Oceanic Engineering*, 20(3):229–241.
- McFarlane, J. R. (2000). Underwater technology 2000 rovs and auvs: tools for exploring, exploiting and defending the ocean frontier. *Underwater Technology*, pages 465–471.
- Negahdaripour, S. and Khamene, A. (2000). Motion-based compression of underwater video imagery for the operations of unmanned submersible vehicles. *Computer Vision and Image Understanding*, 79(1):162–183.
- Plakas, K. and Trucco, E. (2000). Developing a real-time, robust, video tracker. In *MTS/IEEE Oceans*, volume 2, pages 1345–1352.
- Se, S., Lowe, D., and Little, J. (2005). Vision-based global localization and mapping for mobile robots. *IEEE Transactions on Robotics*, 21(3):364–375.
- Shi, J. and Tomasi, C. (1994). Good features to track. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 593–600.
- Stewart, C. V. (1999). Robust parameter estimation in computer vision. *Society for Industrial and Applied Mathematics Review*, 41(3):513–537.
- Xu, X. and Negahdaripour, S. (1997). Vision-based motion sensing for underwater navigation and mosaicing of ocean floor images. In *MTS/IEEE Oceans*, volume 2, pages 1412–1417.