

TECSUB – LOC: Localização e mapeamento baseado em visão para robôs subaquáticos

Silvia. S. Botelho, Gabriel. L. Oliveira, Emanuel Estrada

Centro de Ciências Computacionais - FURG, Rio Grande, RS – silviacb@furg.br

Centro de Ciências Computacionais - FURG, Rio Grande, RS – gabrielleivas@gmail.com

RESUMO: O uso de veículos autônomos para inspeção visual é um promissor campo da robótica. As imagens capturadas podem também auxiliar em sua localização/navegação. Nesse contexto, este trabalho propõe o uso de técnicas de visão computacional para localização e mapeamento de robôs autônomos. Supondo o uso de câmeras de inspeção, nossa proposta é composta de dois estágios: o uso de visão computacional através do algoritmo SIFT para extrair pontos característicos em uma seqüência de imagens em um robô subaquático e o desenvolvimento de mapas topológicos para localização e navegação. A integração desses sistemas permitirá ao sistema desenvolvido simultaneamente fazer a localização e mapeamento do ambiente. Um conjunto de testes com robôs reais foi feito, revelando como resultados uma boa precisão e robustez para uma série de condições, iluminação e ruído, mostrando ser uma técnica de SLAM promissora e original.

PALAVRAS-CHAVE: veículos autônomos para inspeção visual

ABSTRACT: The use of Autonomous Underwater Vehicles (AUVs) for visual inspection tasks is a promising robotic field. The images captured by the robots can also aid in their localization/navigation. In this context, this paper proposes an approach to localization and mapping problem of underwater vehicle. Supposing the use of inspection cameras, this proposal is composed of two stages: the use of computer vision through the algorithm SIFT to extract the features in underwater image sequences and the development of topological maps to localization and navigation. The integration of such systems will permit simultaneous localization and mapping of the environment. A set of tests with real robots was accomplished, regarding online and performance issues. The results reveal an accuracy and robust approach to several bottom conditions, illumination and noise, leading to a promissory and original SLAM technique.

KEYWORDS: Autonomous Underwater Vehicles

1. INTRODUÇÃO

A utilização de tecnologias para exploração de ambientes não estruturados desconhecidos de difícil acesso para exploração humana, como o ambiente subaquático, faz com que robôs móveis remotamente controlados sejam cada vez mais usados. O ambiente subaquático, em particular, é

uma promissora área de exploração tecnológica que oferece possibilidades para pesquisas que vão desde inspeção de estruturas submersas ao mapeamento do próprio ambiente em questão [GAR01]. Assim sendo, o desenvolvimento de sistemas de inspeção visuais aplicados a veículos não tripulados apresenta-se como uma interessante área de pesquisa, com aplicação direta à Fundação Universidade Federal do Rio Grande no que diz respeito às atividades relacionadas ao ecossistema costeiro.

O uso de Unmanned Underwater Vehicles (UUVs - veículos não tripulados para navegação subaquática) para a criação de mapas visuais do ambiente constitui-se em uma importante ferramenta para a exploração subaquática [FLE00, GON98]. Estes se dividem em: (i) Veículos Operados Remotamente (Remote Operated Vehicles- ROVs), os quais requerem um piloto humano para conduzir a sua navegação; e (ii) Veículos Autônomos Subaquáticos (Autonomous Underwater Vehicles - AUVs), os quais têm a capacidade de realização de missões de alto nível sem intervenção humana explícita. Porém, tendo em vista o alto custo e a limitada autonomia de navegação e energia de um UUV complexo, em determinadas missões, o uso de um conjunto de robôs torna-se uma estratégia mais atrativa, por precisarem de menos tempo para percorrer uma mesma área comparada com um único robô realizando tal tarefa. [BOT04, BOT05].

Em suma, este projeto apresenta a proposta de desenvolvimento de um sistema distribuído para geração de mosaicos (mapas visuais), auxiliando na localização e navegação autônomas de robôs móveis de forma online, onde as informações podem ser visualizadas e utilizadas no momento em que estiverem disponíveis. Estes mosaicos tratam-se de mapas gerados pelo alinhamento de consecutivas imagens (frames) [BOT04] capturadas por câmeras fixadas na estrutura do robô, a serem utilizados ao final do processo de inspeção como fontes ricas de informação a físicos, oceanógrafos, engenheiros, dentre outros. Sendo construídos de forma contínua por informações coletadas do ambiente explorado, podem ser ainda aproveitados como mapas de referência no processo de navegação/localização do próprio veículo [FLE00], ou seja, o AUV pode fazer uso do mosaico que está sendo construído para guiar a sua própria navegação em tempo real de inspeção.

Neste contexto, o projeto tem proposto uma técnica de SLAM (do inglês Simultaneous Localization and Mapping) baseada no uso do algoritmo SIFT, para extrair as informações necessárias da imagem, em conjunto com o uso de mapas auto-organizáveis construídos a partir de Redes de Kohonen e Estruturas Celulares Crescentes para localização e navegação do AUV.

2. SISTEMA DE VISÃO PARA MAPEAMENTO

Com a finalidade de montar o mosaico, as várias imagens (frames) capturadas de forma consecutiva devem ser alinhadas sucessivamente, resultando em um único mapa visual. Normalmente, o processo pode ser resumido pelos seguintes passos:

- pré-processamento da imagem, objetivando a correção de deformações geométricas e remoção de características inadequadas para processamento;
- detecção de características relevantes de cada imagem;
- correlação de características entre imagens consecutivas;
- estimativa de movimento entre imagens;
- construção do mosaico;

Cada fase da criação do mosaico é abaixo explicada.

2.1 Pré-processamento da imagem

O pré-processamento tem como objetivo corrigir deformações geométricas causadas por imperfeições no sistema óptico e solucionar problemas relativos à iluminação não-homogênea do ambiente. Estas distorções introduzem uma incorreta interpretação do posicionamento real do robô e conseqüentemente, erros de mapeamento e localização.

A distorção causada pelas lentes da câmera na imagem é uma etapa que demanda um certo tempo de processamento, e pode ser modelada por uma aproximação radial e tangencial. Porém, como a componente radial é a que causa maior distorção, a maioria dos trabalhos desenvolvidos até hoje corrige apenas esse problema [GRA02, GAR01], sendo suficiente para a solução do problema. Da mesma forma, optou-se neste trabalho apenas pela correção da distorção radial.

Em ambiente subaquático, há ainda a distorção causada pela difração da água. A Equação 1 mostra um método para resolver este problema [XUN97], onde (m_x, m_y) é o ponto sem distorção radial, (m_0x, m_0y) é o novo ponto sem distorção por difração e u_0 e v_0 são as coordenadas do ponto central.

$$\begin{aligned}m_0x &= m_x + \frac{R_0}{R}(m_x - u_0) \\m_0y &= m_y + \frac{R_0}{R}(m_y - v_0) \\R_0 &= f \tan(\sin^{-1}(1.33 * \sin(\tan^{-1} \frac{R}{f}))) \\R &= \sqrt{m_x^2 + m_y^2}\end{aligned}$$

Equação 1

2.2 Detecção de características relevantes de cada imagem

O algoritmo Scale Invariant Feature Transform – SIFT – é um eficiente filtro para extração e descrição de pontos de características relevantes para a análise de imagens [Lowe04]. Em termos gerais, é um método robusto em relação a ruídos, iluminação imprópria, oclusão e variações mínimas no ponto de vista de captura da imagem, e constitui-se de quatro estágios principais:

- *Detecção de extremos no espaço de escala:* O primeiro estágio utiliza-se do filtro de diferença gaussiana para detectar pontos potencialmente interessantes em um espaço invariante à escala e rotação.
- *Localização de pontos relevantes:* A localização e escala de cada ponto encontrado no estágio anterior são calculadas e pontos relevantes são selecionados baseados em medidas de estabilidade.
- *Definição de orientação dos pontos relevantes:* Considerando características das imagens, como gradientes locais, as orientações dos pontos relevantes são determinadas. Todas as operações futuras serão feitas em relação a esta orientação, além de escala e posição, garantindo invariância em relação a estas transformações.
- *Criação do descritor dos pontos relevantes:* Os gradientes locais da imagem são calculados baseados na vizinhança de cada ponto relevante. Estes cálculos são transformados em uma representação que permite mudanças significantes nos níveis de distorção e iluminação. Sendo assim, o vetor descritor de cada ponto armazena suas características para que um mesmo ponto possa ser encontrado em diferentes regiões de imagens consecutivas.

Um ponto candidato à relevância é um pixel extremo em relação à sua vizinhança, localizado em uma determinada escala de frequência. Esta vizinhança é definida como sendo uma janela $n \times n$, tanto no mesmo espaço de escala quanto nos adjacentes, como mostra a Figura 1. Por meio da operação de Diferença Gaussiana (Equação 2) obtida pela subtração de duas imagens geradas a partir da convolução, chega-se à banda de frequência para cada escala.

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad \text{Equação 2}$$

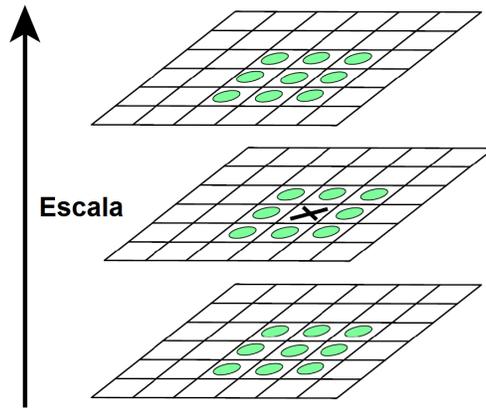


Figura 1

Uma vez que um ponto foi, de fato, considerado relevante, é calculada a sua orientação dominante. Esta é obtida a partir de uma imagem da cascata de filtros, em uma janela de 16×16 pixels, através da Equação 3. Baseado nestes dados é criado, então, um histograma de orientações composto por 36 elementos, em que cada elemento é analisado em função do gradiente da janela analisada. O pico do histograma representa a orientação dominante do ponto em questão.

$$\begin{aligned} \theta(x, y) &= \text{atan2}(G(x, y + 1) - G(x, y - 1), G(x + 1, y) - G(x - 1, y)) \\ m(x, y) &= ((G(x + 1, y) - G(x - 1, y))^2 + (G(x, y + 1) - G(x, y - 1))^2)^{\frac{1}{2}} \end{aligned} \quad \text{Equação 3}$$

A orientação do gradiente previamente calculada na janela do ponto relevante é usada agora para gerar o descritor do ponto. Para que este seja invariante à rotação, a região do gradiente é rotacionada em relação à orientação estabelecida para o ponto. Logo após um histograma de 8 elementos é criado para cada janela de 4×4 pixels. O vetor resultante, de 128 elementos, compõe finalmente o descritor do ponto de relevância.

2.3 Correlação de características entre imagens consecutivas

Normalmente, os candidatos à correlação de um ponto de relevância são pontos próximos ao mesmo, de maneira que o melhor candidato é o ponto que apresenta a menor distância Euclidiana. No entanto, alguns pontos instáveis (outliers) são detectados ao longo do processo, levando a falsas correspondências. Para a eliminação desse problema, um método para comparar a menor distância

com a segunda melhor distancia é usado, selecionando somente correspondentes próximos por treshold [Lowe04].

Além disto, uma Matriz Fundamental estimada pelos métodos RANSAC e LMedS é usada para remoção de outliers, validando cada correspondência encontrada.

2.4 Estimativa de movimento entre imagens

Matriz Homográfica

A matriz homogênea H de transformação projetiva 2D é uma transformação linear com oito graus de liberdade, conforme Equação 4. Esses parâmetros podem representar translação, rotação, escalamento, shear e deformação perspectiva [GAR01].

$$\begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{ou} \quad x' = Hx$$

Equação 4

A matriz de homografia com oito graus de liberdade é a que representa um movimento mais completo. Porém, dependendo do modelo de movimento necessário e do grau de liberdade do mesmo outro tipo de matriz pode ser utilizada. Isto acontece porque, no momento em que se deseja encontrar a matriz H para representá-lo, os pares de pontos correlacionados, por pequenos deslocamentos em suas medidas, vão acabar gerando outros movimentos que não estão sendo realmente executados.

Estimando a Homografia

A correlação de imagens fornece um conjunto de vetores de deslocamento relativo às características de duas imagens obtidas, ou seja, cada vetor representa as coordenadas da mesma característica em ambas imagens. O objetivo de tal etapa de construção do mosaico é determinar o movimento entre tais imagens, que pode ser obtido a partir da homografia.

Como a etapa de correlação fornece um conjunto de pontos correlacionados, estes pontos podem ser usados para se achar a matriz H . Transformando as coordenadas homogêneas em não homogêneas, operando os termos a fim de se obter um sistema linear em termos dos elementos de H e expressando na forma matricial considerando n pares de correspondências, obtém-se a Equação 5:

$$\begin{bmatrix}
 x_1' & y_1' & 1 & 0 & 0 & 0 & -x_1'x_1' & -x_1'y_1' \\
 0 & 0 & 0 & x_1' & y_1' & 1 & -y_1'x_1' & -y_1'y_1' \\
 \square & \square \\
 x_n' & y_n' & 1 & 0 & 0 & 0 & -x_n'x_n' & -x_n'y_n' \\
 0 & 0 & 0 & x_n' & y_n' & 1 & -y_n'x_n' & -y_n'y_n'
 \end{bmatrix}
 \begin{bmatrix}
 h_{11} \\
 h_{12} \\
 h_{13} \\
 h_{21} \\
 h_{22} \\
 h_{23} \\
 h_{31} \\
 h_{32}
 \end{bmatrix}
 =
 \begin{bmatrix}
 x_1 \\
 y_1 \\
 \square \\
 x_n \\
 y_n
 \end{bmatrix}$$

Equação 5

ou $Ah = b$. A equação 6 pode ser resolvida com quatro ou mais pares de correspondências e os oito elementos da matriz h podem ser calculados a partir da estratégia pseudo-inverse least squares. Entretanto, essa solução envolve a inversão da matriz, o que pode falhar porque a matriz pode ser singular ou muito próxima à singular [GAR01]. Por esta razão utiliza-se a técnica chamada de Singular Value Decomposition (SVD) [PRE02], que trata de resolver um conjunto de equações ou matrizes que podem ser singulares ou muito próximo a isso. O método SVD pode ser definido pelo seguinte teorema: qualquer matriz $m \times n$ cujo número de linhas m é maior ou igual ao seu número de colunas n , pode ser escrito como o produto de uma matriz $m \times n$ coluna-ortogonal U , uma matriz $n \times n$ diagonal W com elementos positivos ou zero (valores singulares) e a transposta de uma matriz $n \times n$ ortogonal V . Então a matriz A da pode ser escrita na forma da Equação 6:

$$A_{m \times n} = U_{m \times n} W_{n \times n} V_{n \times n}^T$$

Equação 6

Assim, a solução SVD para a Equação 6 seria definida pela equação 7:

$$h = V [diag \ 1 / W_j] U^T b$$

Equação 7

Antes de encontrar a matriz H é necessário gerar uma matriz H_N sendo a homografia dos pares de pontos de correlação normalizados. Finalmente, a matriz H é encontrada a partir da equação 8:

$$H = T^{-1} H_N T'$$

Equação 8

sendo T a matriz de transformação para a normalização dos pontos de interesse da imagem I e T' a matriz de normalização dos pontos de interesse da imagem I' .

2.5. Construção do Mosaico

A construção de mosaicos é baseada nas decisões de como e quando incluir novas imagens e na definição de suas posições globalmente, a fim de que possa ser criado visualmente. O conceito de

construção do mosaico deste trabalho permite não só que apenas um veículo gere seu mosaico, mas também permite a construção de um mosaico a partir de vários veículos operando em paralelo, tudo sob a mesma estrutura de programação, para a geração de um mosaico distribuído.

Tendo definida a matriz de transformação projetiva H entre a imagem anterior I e a imagem atual I' , e sabendo-se que a média das distâncias entre os pontos de interesse m e m' das respectivas imagens é diferente de zero, então pode-se registrar globalmente a imagem atual I' , adicionando-a à estrutura do mosaico juntamente com a matriz H , agora referenciada por ${}^k H_{k+1}$, onde k faz referência à imagem I e $k+1$ à imagem I' . A transformação projetiva global da imagem ${}^{k+1}$ no mosaico pode ser definida pela Equação 9.

$${}^1 H_{k+1} = \prod_{i=1..k} {}^i H_{i+1}$$

Equação 9

Sabendo a transformação projetiva global de cada imagem pertencente ao mapa visual, a equação 10 calcula onde cada pixel (posição $m^{[k+1]}$) da imagem ${}^{k+1}$ do mosaico deve ser escrito no mosaico final (posição $m^{[1]}$). Os pontos $m^{[k+1]}$ variam em $x \in [1..largura]$ e $y \in [1..altura]$, sendo $largura$ e $altura$ as dimensões da imagem capturada.

$$m^{[1]} = {}^1 H_{k+1} m^{[k+1]}$$

Equação 10

3. LOCALIZAÇÃO POR MAPAS TOPOLÓGICOS

Os vetores extraídos pelo SIFT são utilizados para compor o mapa topológico obtido pelo uso de mapas auto-organizáveis baseados em Kohonen Neural Networks [KOH01] e Estruturas Celulares Crescentes [FRI93].

Mapa de Kohonen é um tipo de rede neural artificial treinada através de aprendizado não supervisionado. Um mapa de pequena dimensão, normalmente de duas, discretiza o espaço de entrada das amostras de treinamento, tentando preservar as propriedades topológicas do mesmo. A estrutura de uma rede de Kohonen consiste de componentes chamados nodos ou neurônios. Associado a cada nodo há um vetor de pesos da mesma dimensão dos vetores dos dados de entrada e uma posição no espaço. Os nodos são conectados por linhas, resultando em um grid 2D.

A atual proposta do projeto opera no espaço de características invariantes à escala ao invés do espaço da imagem. Durante o processo de navegação, vetores de características são apresentadas aos mapas auto-organizáveis. O algoritmo de aprendizagem é baseado no conceito de aprendizado de vizinho mais próximo (nearest-neighbor). Quando uma nova entrada é obtida, o mapa topológico determina o vetor de características do nodo de referencia que melhor corresponde ao vetor de entrada. Como o sistema usa diversos vetores de características associados a cada imagem, o algoritmo de nearest-neighbor é aplicado a cada vetor separadamente. Os resultados deste algoritmo são então combinados com um esquema simples de votação unânime. Nodos podem ser criados ou deletados durante o processo de aprendizado.

O mapa topológico final permite duas possibilidades de navegação: através de posições alvo ou

através de metas visuais. Da posição corrente, algoritmos de busca em grafo, como o Dijkstra [DIJ59] ou o A* [DEC85] podem ser usados para achar o caminho ideal até o ponto desejado.

4. TESTES E RESULTADOS

Diferentes características subaquáticas foram aplicadas às imagens de teste, como turbidez, neve marinha, iluminação não uniforme, entre outras, simulando diferentes condições, como mostra a Figura 2.



Figura 2

De maneira geral, o sistema visual foi testado em 5 diferentes ambientes subaquáticos, correspondendo cada condição de distorção a uma coluna da Tabela 1, que apresenta as características de cada tipo de situação introduzida nos testes.

Distorção	1	2	3	4	5
Distância à fonte de iluminação (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
Nível mínimo de cinza	20	30	20	20	20
Numero de flocos de neve marinha	30	30	30	30	30

Tabela 1

A Figura 3 mostra os pontos de relevância detectados e correlacionados obtidos em uma seqüência de navegação visual.

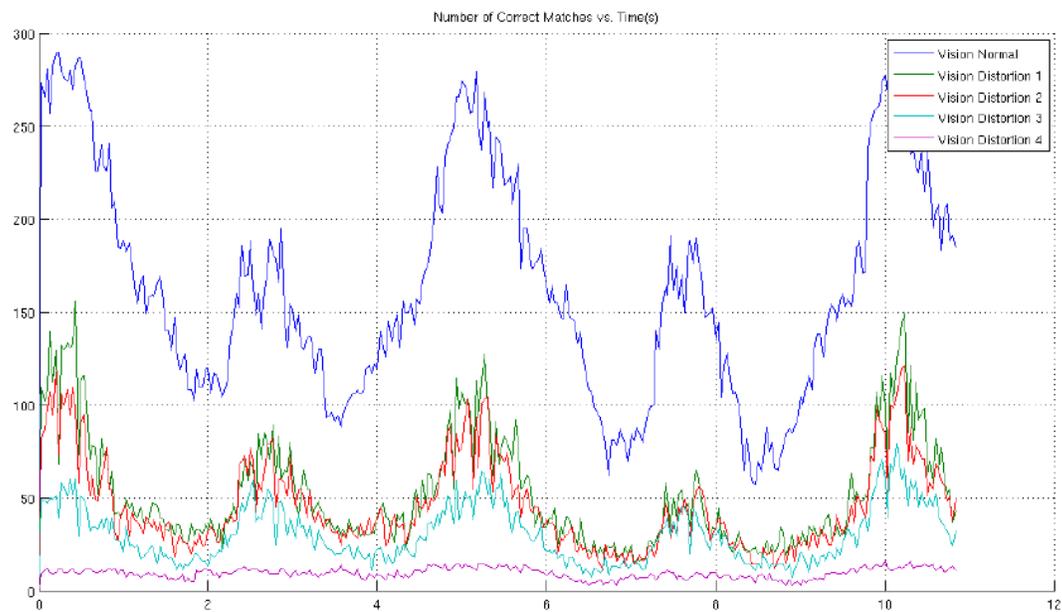
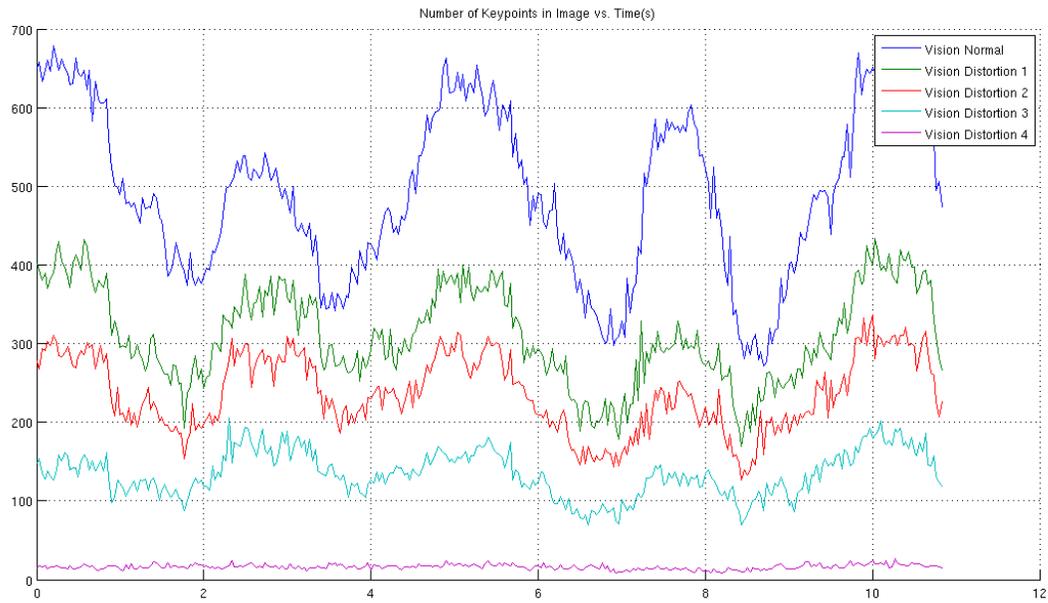


Figura 3: Número de pontos relevantes e correlações

Ainda que o número de pontos tenha diminuído com a perda de qualidade de imagem devido às distorções impostas, continua possível localizar o robô, como visto na Figura 4.

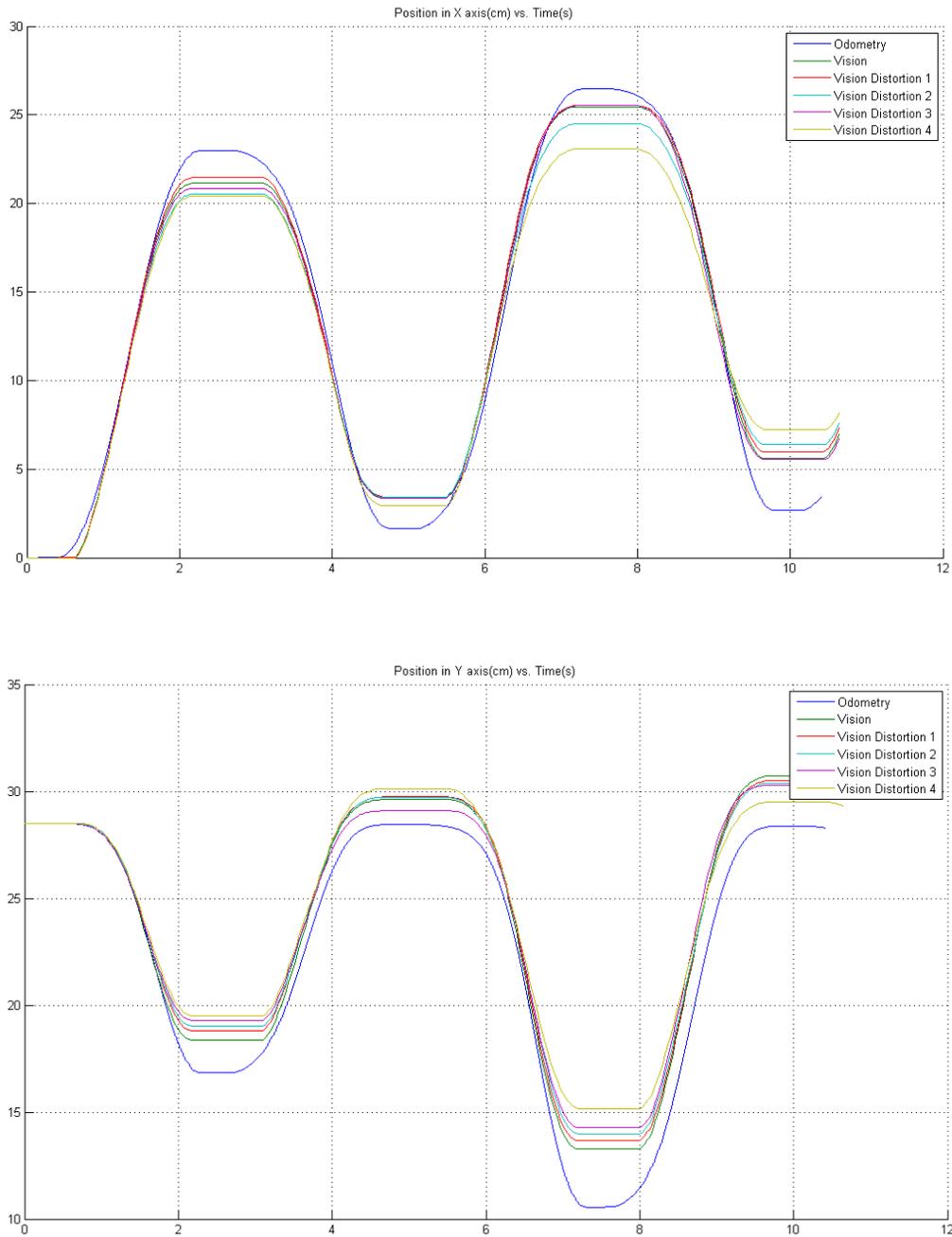


Figura 4: Posições em x e y determinadas por odometria e por sistema visual

Na Figura 4, em azul é mostrado o referencial de movimento, executado por um braço robótico composto por um atuador do tipo Harmonic Drive com um encoder acoplado, fornecendo posições angulares a cada 0.000651 segundos. Anexada a esta estrutura, há uma câmera do tipo Samsung SC-D364 para a captura das imagens de teste. Como pode ser analisado através dos gráficos, a abordagem utilizada mostra-se robusta a variações do ambiente subaquático, conforme o esperado.

Para finalizar, validando o sistema de mapeamento proposto foram realizados testes sobre os mapas topológicos. O último dos resultados do sistema pode ser visto na Figura 5.

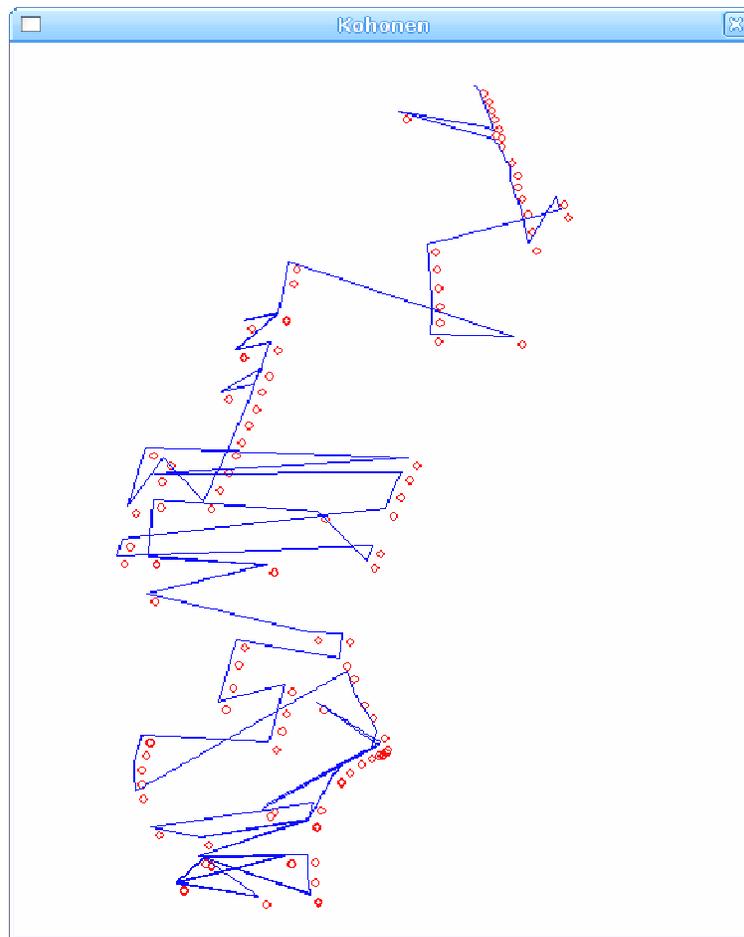


Figura 5: Mapa topológico

REFERÊNCIAS BIBLIOGRÁFICAS

- [BOT04] BOTELHO, S; NEVES, R.; TADDEI, L. Uma Arquitetura para Multi-AUVs Realizando Tarefas de Inspeção Visual. Revista Vetor, vol. 14, nº 1, pp. 30-42, 2004.
- [BOT05] BOTELHO, S; NEVES, R.; TADDEI, L. Localization of a Fleet of AUVs using Visual Maps. Proceedings of the OCEANS Conference, 2005.
- [DEA99] DEAVES, R. Covariance Bounds for Augmented State Kalman Filter Application. IEEE Electronics Letters, 1999.
- [DEC85] Dechter, R. and Pearl, J. (1985). Generalized best-first search strategies and the optimality of a^* , Journal of the Association for Computing Machinery 32(3): 505-536.
- [DIJ59] Dijkstra, E. W. (1959). A note on two problems in connexion with graphs, Numerische Mathematik 1: 269-271.
- [FLE00] FLEISCHER, S. D. Bounded-Error Vision-Based Navigation of Autonomous Underwater Vehicle. Phd Dissertation, Stanford University, 2000.

- [FRI93] FRITZKE, B. Growing cell structures - a self-organizing network for unsupervised and supervised learning, Technical report, University of California - Berkeley, International Computer Science Institute, 1993.
- [GAR01] GARCIA, R. A Proposal to Estimate the Motion of an Underwater Vehicle through Visual Mosaicking. Phd Dissertation, Universitat de Girona, 2001.
- [GAR02] GARCIA, R.; PUIG, J.; RIDAO, P.; CUFI, X. Augmented State Kalman Filtering for AUV Navigation. IEEE ICRA, 2002.
- [GAR05] GARCIA, R.; CHAROT, F. VLSI Architecture for an Underwater Robot Vision System. Proceedings of the OCEANS Conference, 2005.
- [GRA98] GRACIAS, N.; SANTOS-VICTOR, J. Automatic Mosaic Creation of the Ocean Floor. Proceedings of the OCEANS Conference, 1998.
- [GRA02] GRACIAS, N., VAN DER ZWAAN, S., BERANRDINO, A. and SANTOS-VICTOR, J. Results on underwater mosaic-based navigation, IEEE Oceans Conference, 2002, Vol. 3, pages. 1588-1594.
- [KAL60] KALMAN, R.E. A New Approach to Linear Filtering and Prediction Problems. Transactions of the ASME – Journal of Basic Engineering, pp. 35-45, 1960.
- [KOH01] KOHONEN, T. (2001). Self-Organizing Maps, Springer-Verlag New York, Inc., Secaucus, NJ, USA.
- [LAW80] LAWS, K. I. Textured Image Segmentation. Phd Thesis, Processing Institute, University of Southern California, 1980.
- [Lowe04] Lowe, D. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision. V.60,n. 2, pages 91-110, 2004.
- [MAR94] MARKS, R.; ROCK, S.; LEE, M. Real-time Video Mosaicking of the Ocean Floor. Proceedings of IEEE Symposium on Autonomous Underwater Vehicle Technology, 1994.
- [MAR95] MARKS, R.; ROCK, S.; LEE, M. Real-time Video Mosaicking of the Ocean Floor. IEEE Journal of Oceanic Engineering, vol. 20, n° 3, pp. 229-241, 1995.
- [NEG93] NEGAHDARIPOUR, S.; YU, C. H. A Generalized Brightness Change Model for Computing Optical Flow. Proceedings of the International Conference on Computer Vision, 1993.
- [NEG98] NEGAHDARIPOUR, S.; XU, X.; KHAMENE A. A Vision-System for Real-time Positioning, Navigation and Video Mosaicing of Sea Floor Imagery in the Application of ROVs / AUVs. Proceedings of the 4° IEEE Workshop on Applications of Computer Vision, pp 248-249, 1998.
- [NEG99] NEGAHDARIPOUR, S.; XU, X.; JIN, L. Direct Estimation of Motion from Sea Floor Images for Automatic Station-Keeping of Submersible Platforms. IEEE Journal of Oceanic Engineering, vol. 24, n° 3, pp. 370-382, 1999.
- [PRE02] PRESS, W. H.; TEUKOLSKY, S. A.; VETTERLING, W. T.; FLANNERY, B. P. Numerical Recipes in C: the Art of Scientific Computing. Second Edition, Cambridge University Press, 2002.

- [SHI94] SHI, J.; TOMASI, C. Good Features to Track. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 593-600, 1994.
- [SIN98] SINGH, H.; HOWLAND, J.; YOERGER, D.; WHITCOMB, L. L. Quantitative Photomosaicing of Underwater Imaging. Proceedings of the OCEANS Conference, vol. 1, pp. 263-266, 1998.
- [TAV03] TAVARES, A. Um Estudo sobre a Modelagem e o Controle de Veículos Subaquáticos não Tripulados. Tese de Mestrado, Engenharia Oceânica, FURG, 2003.
- [XUN97] XU, X. and NEGAHDARIPOUR, S. Vision-based motion sensing for underwater navigation and mosaicing of ocean oor images, MTS/IEEE OCEANS Conference and Exhibition, 1997, Vol. 2, pages. 1412-1417.
- [WEL03] WELCH, G.; BISHOP, G. An Introduction to the Kalman Filter. University of North Carolina, 2003.