

DESENVOLVIMENTO DE SISTEMA PARA GERAÇÃO DE MOSAICO DE IMAGENS VISANDO LOCALIZAÇÃO E NAVEGAÇÃO DE ROBÔS MÓVEIS

Edson CAVALCANTI Neto (1), Guilherme Costa HOLANDA (1), Antonio Themoteo VARELA (1), André Luiz Carneiro de ARAÚJO (1), Auzuir Ripardo de ALEXANRIA (1);

(1) IFCE - Instituto Federal de Educação, Ciência e Tecnologia do Ceará, Av. Treze de Maio, 2081 - Benfica - Fortaleza/CE, +55 (85) 3307-3666,

{edsoncavalcanti@ifce.edu.br, guilhermecosta@ifce.edu.br, andreluiz@ifce.edu.br, themoteo@ifce.edu.br, auzuir@ifce.edu.br}

RESUMO

A Robótica tem se desenvolvido bastante nos últimos anos, em especial a Robótica móvel. Os robôs móveis são constituídos dos subsistemas a seguir: percepção, localização, navegação e controle de movimento. Mosaicos podem ser definidos como um conjugado de imagens de escala aproximada, de uma determinada região, recortadas e montadas, artisticamente ou tecnicamente de forma que passe a idéia de que tal conjunto é uma única imagem. Em Visão Computacional, a geração de mosaicos a partir de imagens é o processo usado quando a cena a ser representada é maior que a imagem fornecida por uma câmera, de tal forma que duas ou mais imagens são unidas recompondo a cena. A partir do mosaico, é possível realizar a navegação no ambiente em que o robô está situado. Utilizando a *Scale-invariant feature transform* (SIFT), pode-se rastrear pontos em imagens, inclusive em tempo real, e ser efetuada a composição de mosaicos. As simulações foram desenvolvidas utilizando-se a SIFT e implementadas em ambiente MATLAB. Os resultados obtidos são bastante promissores e coerentes com o esperado. A simulação desenvolvida se aproxima do real, pois as imagens obtidas não possuem níveis distantes de profundidade, escala e luminosidade.

Palavras-chave: Geração de Mosaicos, SIFT, Localização de robôs.

1. INTRODUÇÃO

Dentre os sistemas de um robô móvel o sistema de percepção se qualifica dentre os mais importantes. Este sistema é o responsável pela aquisição de dados do ambiente que o robô atua (SIEGWART, 2004). Estes dados, além de serem armazenados como parte da própria missão do robô, são trabalhados de forma a serem utilizados na geração de informações relativas à localização do robô em seu ambiente, através de seu sistema de localização. De posse de sua localização, o robô é capaz de gerar um mapa do ambiente que atua e comparar sua posição local com uma posição mais global no mapa, sendo este gerado ou entregue pelo seu desenvolvedor ou usuário do sistema.

A capacidade de navegar utilizando-se o referido mapa, apoiado, na missão que deve desempenhar é tarefa extremamente necessária e realizada pelo seu sistema de navegação. Este último informa qual trajetória o sistema de controle de movimento deve desempenhar.

O mosaico pode ser definido como um conjugado de fotos de escala aproximada, de uma determinada região, recortadas e montadas, artisticamente ou tecnicamente de forma que dê a impressão de que tal conjunto é uma única imagem (WOLF, 1983). O mosaico pode ser utilizado como mapa de localização e navegação para o robô móvel.

De acordo com Calheiros (2010), mosaicos são imagens geradas a partir de várias outras imagens parciais de um mesmo local. Essas imagens podem ser obtidas a partir da captura de vídeo onde várias imagens em sequência ou não.

A necessidade de criação de mosaicos tornou-se importante com o decorrer do tempo, sobretudo quando os primeiros satélites começaram a aquisição de imagens da Terra. Essas imagens tinham o propósito de criar

um mapa topográfico do planeta, mas somente com uma imagem não era possível a visualização completa e detalhada necessária, então a criação de uma imagem maior gerada a partir de outras menores, o mosaico, foi utilizada para realizar a criação desse mapa (LUCAS, 2009).

O processo de criação de mosaicos é realizado por 4 etapas, em que a primeira é a captura das imagens que constituem o mosaico. A segunda etapa consiste no pré-processamento digital das imagens. Na terceira etapa é realizado um tratamento na imagem a partir de um sistema computacional para identificar pontos de interesse e semelhantes nas imagens. Na última etapa é efetuada a criação do mosaico (VARGAS, 2004).

Na primeira etapa, na captura das imagens é necessário ter alguns cuidados, pois dependendo do modo de aquisição, iluminação e posição da câmera o mosaico será criado defeituosamente.

Na segunda etapa, no pré-processamento de imagens é utilizado um algoritmo de reajuste de histograma e um filtro de imagem para facilitar o terceiro passo em que será realizado o processo de detecção dos pontos de interesse (GONZALEZ, 2008).

Na terceira etapa, será realizado um tratamento da imagem a partir do algoritmo *Scale-invariant feature transform* (SIFT) que possui como objetivo principal para esse projeto obter os pontos-chaves e descrevê-los (GOEDEME, 2007).

Após essa etapa, é realizada a quarta e última etapa. Nesta, é realizada a criação do mosaico, onde são utilizadas as imagens que constituem o mosaico e o vetor com as coordenadas dos pontos de interesse de cada imagem. Primeiramente é utilizado o sistema de homografia para realizar a comparação dos pontos das imagens e obter a relação das imagens. A partir desta relação, o programa sobrepõe as imagens formando assim uma imagem maior, que é o resultado da junção das imagens menores (MACHADO, 2004).

O objetivo principal deste trabalho é o desenvolvimento de um sistema para a captura de imagens do ambiente e a geração do mosaico. Este trabalho descreve os passos utilizados para a obtenção de um mosaico, desde a detecção de pontos críticos até a composição, propriamente dita, do mosaico.

Este trabalho é organizado em revisão bibliográfica sobre pré-processamento de imagens, SIFT e distância euclidiana na segunda seção. A metodologia usada para alcançar os resultados é apresentada na terceira seção. Os resultados obtidos são apresentados na quarta seção.

2. COMPOSIÇÃO DE MOSAICO

O mosaico é utilizado para a navegação e localização do robô móvel por meio de imagens adquiridas através de uma câmera instalada no próprio. O módulo de navegação, de posse do mapa do local fornecido pelo usuário, do mosaico obtido, da posição do robô neste mapa e da missão a ser cumprida, define a trajetória que o robô deve realizar.

No decorrer da montagem de forma técnica de um mosaico, diversas dificuldades podem surgir. As principais dificuldades são: não coincidência geométrica das feições imageadas e o desnível angular entre as linhas correspondentes. Estes podem gerar prejuízos tanto na geometria quanto na visualização do conjugado de imagens obtido.

Os principais métodos necessários para a implementação e entendimento do sistema apresentado são o pré-processamento das imagens, a detecção de pontos críticos, a SIFT e a distância Euclidiana.

2.1 Pré-processamento de imagens

A primeira etapa é transformar a imagem que está em escala RGB de cores para a escala em níveis de cinza. Este processo é útil para que se possa haver uma continuidade radiométrica entre os *pixels* (HASEGAWA, 2004).

Após esta etapa, a imagem é submetida a um processamento por meio de máscaras, matrizes, para uma detecção combinada dos pontos. A utilização de uma formulação multimáscara torna possível o desenvolvimento de um método para determinar se um *pixel* é mais provavelmente um ponto isolado ou parte de uma linha ou de uma borda (GONZALEZ, 2008). As máscaras que empiricamente atingiram melhor resultado foram as máscaras ortogonais de base do subespaço de bordas, conforme a Figura 1.

$$\begin{array}{cc}
 \begin{bmatrix} 1 & \sqrt{2} & 1 \\ 0 & 0 & 0 \\ -1 & -\sqrt{2} & -1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & -1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & 0 & -1 \end{bmatrix} \\
 \text{(a)} & \text{(b)} \\
 \\
 \begin{bmatrix} 0 & -1 & \sqrt{2} \\ 1 & 0 & -1 \\ -\sqrt{2} & 1 & 0 \end{bmatrix} & \begin{bmatrix} \sqrt{2} & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -\sqrt{2} \end{bmatrix} \\
 \text{(c)} & \text{(d)}
 \end{array}$$

Figura 1 – Máscaras utilizadas para o pré-processamento das imagens. a) detecção de linhas horizontais, b) linhas verticais, e c) e d) linhas diagonais.

A primeira máscara faz a detecção das linhas horizontais e a segunda faz detecção das linhas verticais. As terceira e quarta fazem detecção das linhas em diagonal, porém, não foram utilizados na implementação, pois com apenas a utilização das duas primeiras máscaras já se obtém um resultado satisfatório, no caso analisado.

O processo de filtragem de imagens consiste na convolução da imagem com cada máscara ilustrada na Figura 1. Após o processo de filtragem obtém-se duas subimagens, uma contendo apenas as linhas horizontais e outra apenas as linhas verticais. Ao multiplicar a imagem de linhas verticais, que é uma matriz de *pixels*, ponto a ponto pela imagem de linhas horizontais encontram-se apenas os pontos comuns às duas imagens, obtendo-se, assim, outra subimagem contendo os pontos críticos da imagem analisada. Esse processo pode ser observado na Figura 2.

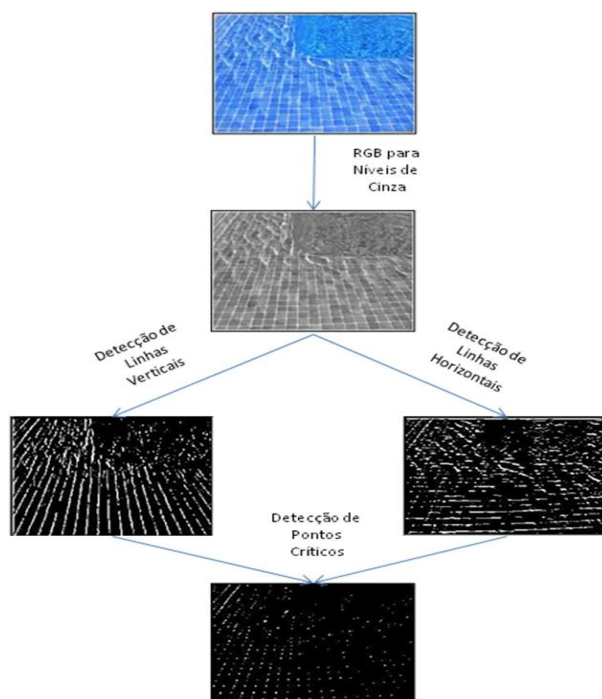


Figura 2 – Sequência de etapas para a localização de pontos críticos na imagem.

2.2 Scale-invariant feature transform (SIFT)

SIFT é um método robusto para extrair e descrever pontos-chaves. As características extraídas por esta técnica são invariantes a translação, rotação, escala e parcialmente invariável a mudanças na iluminação e perspectiva (LOWE, 2004; GOÉDEME et al., 2007; ANGELI, 2008; TEIXEIRA, 2008). Basicamente, o algoritmo é realizado em quatro etapas:

- Detecção de extremos: faz-se uma busca em todas as escalas e localizações de imagens com diferenças de filtros gaussianos (DOG - *difference of gaussian*) para a identificação de pontos de interesse invariáveis a escala e rotação;
- Localização dos pontos-chaves: nos locais onde se encontra um extremo, ajusta-se um modelo detalhado para determinar a localização exata e a escala. Pontos-chaves são selecionados com base em medidas de estabilidade;
- Definição da orientação: define-se uma orientação para cada ponto-chave utilizando os gradientes locais da imagem; e
- Descrição dos pontos-chaves: mede-se o gradiente local de cada ponto-chave, com base na vizinhança.

Essas medidas são transformadas para uma representação que é tolerante a níveis significativos de distorção e mudança na iluminação. A partir da obtenção de pontos-chaves das duas imagens, estes podem ser comparados através de cálculo de distância euclidiana, encontrando assim os pares correspondentes de pontos. Assim, é possível rastrear os pontos em imagens, inclusive em tempo real, efetuando-se a composição de mosaicos.

2.3 Distância Euclidiana

A distância euclidiana é também conhecida por distância métrica. Geometricamente, esta distância representa a distância entre dois pontos, ou seja, dados os pontos $P = (P_1, P_2, \dots, P_n)$ e por $Q = (Q_1, Q_2, \dots, Q_n)$, em um espaço euclidiano n-dimensional, a distância euclidiana D pode ser calculada

$$D = \sqrt{\sum_{i=1}^n (P_i - Q_i)^2} \quad [\text{Eq. 01}]$$

em que (P_i, Q_i) são as coordenadas no plano dos pontos P e Q dos descritores, respectivamente. A distância euclidiana é obtida a partir do triângulo de pascal, sendo a distância euclidiana a hipotenusa do triângulo (LUPPE, 2005).

A distância euclidiana é necessária, pois é utilizada para achar o ponto correspondente de um ponto crítico de uma imagem em outra. É realizada uma comparação utilizando a distância euclidiana entre os pontos das duas imagens que estão sendo analisadas, os pontos são correspondentes se a distância entre eles seja menor do que a distância entre eles e outros pontos.

3. METODOLOGIA

O algoritmo de captura de imagens, monitoramento e criação do mosaico é implementado na plataforma de simulação MATLAB 2008.

Primeiramente é capturada a primeira imagem do mosaico, posteriormente a partir desta é retirada uma imagem de busca para ser comparada com o quadro atual. Após essa comparação se o ponto médio sair da área central da imagem, a imagem anterior é salva e o quadro atual passa a ser a imagem a ser comparada.

Depois de capturadas essas imagens passam por um pré-processamento digital para melhorar o funcionamento do algoritmo. Posteriormente, é utilizado a *Scale-invariant feature transform* (SIFT), que é um método robusto para extrair e descrever pontos-chaves da imagem. De posse de tais pontos, é empregada a equação matemática da distância euclidiana nos pontos chaves da esquerda com o ponto da direita, tendo a menor distância o ponto correspondente. E, por fim, é mostrada a composição final do mosaico. O fluxograma das etapas do processo de criação do mosaico pode ser observado na Figura 3. As imagens utilizadas nesse trabalho para a geração do mosaico são ilustradas na Figura 5.

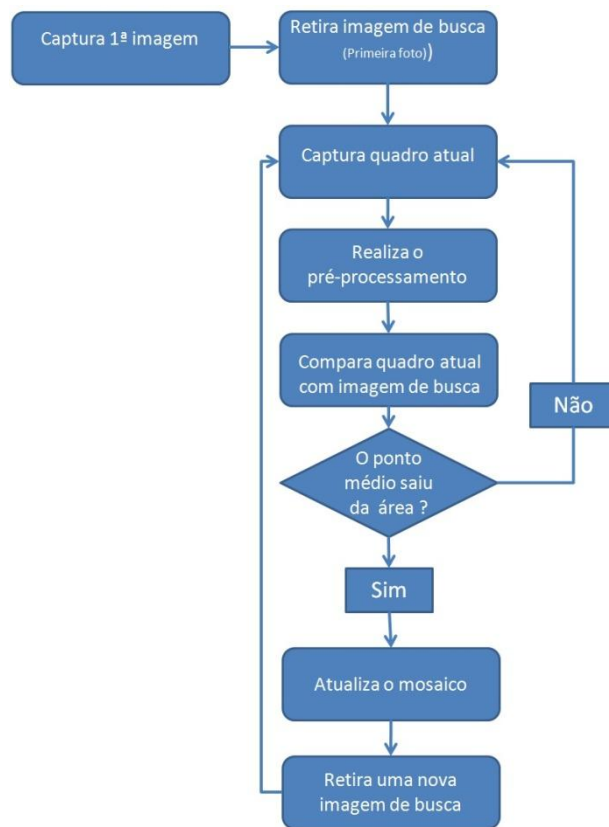


Figura 3 – Fluxograma do programa.

Um sistema de visualização do mosaico é necessário, pois por ele ser uma junção de várias imagens é necessário a diminuição da escala, assim possibilitando a manipulação de imagens na tela do computador. Esse sistema de visualização é ilustrado na Figura 4, onde existem três botões, um para carregar o banco de dados, outro para iniciar o processo de criação do mosaico e outro para poder utilizar a ferramenta de *zoom*.

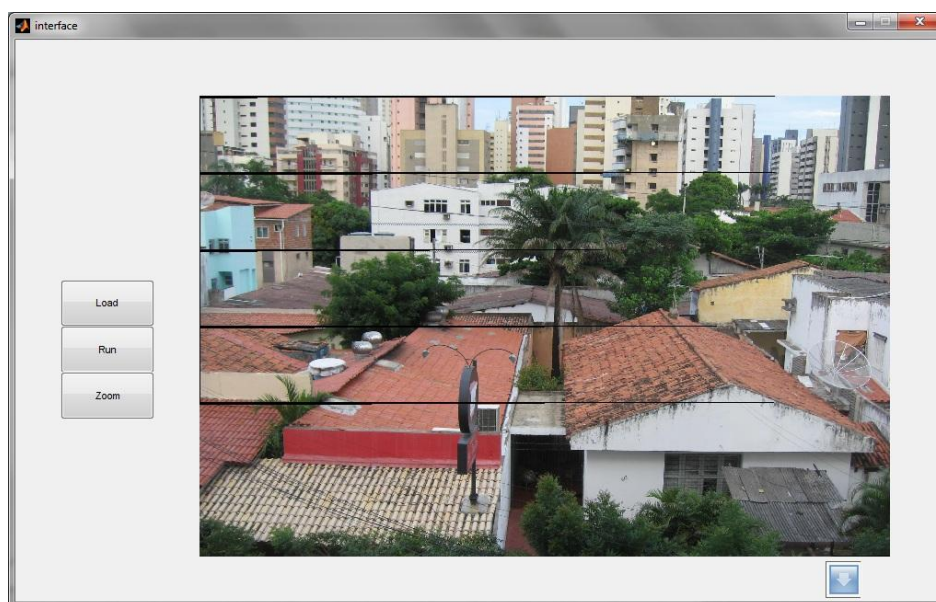


Figura 4 – Interface do programa de visualização de mosaico.

4. TESTES E RESULTADOS OBTIDOS

Para a verificação da rotina no teste realizado foram utilizadas 55 imagens de tamanho 640 x 480 *pixels*. Essas imagens estão ilustradas na Figura 5. Antes da construção do mosaico, as imagens são pré-processadas, utilizando as máscaras da Figura 1 para facilitar a detecção dos pontos críticos. Após o pré-processamento, o algoritmo SIFT é utilizado para achar os pontos críticos e a partir destes pontos e a utilização da distância euclidiana é possível fazer a conexão dos pontos das imagens.



Figura 5 – Conjunto de imagens para a criação do mosaico.

As imagens da Figura 5 são utilizadas para a criação do mosaico da seguinte forma, primeiramente são construídos os mosaicos horizontais, correspondentes ao conjunto de imagens obtido na horizontal. O resultado dos mosaicos horizontais é ilustrado na Figura 6.



Figura 6 – Mosaicos horizontais.



Figura 7– Mosaico final obtido (SNR = 13dB).

Após essa etapa, o próximo passo é a construção do mosaico final, ou seja, a construção de um mosaico a partir dos mosaicos obtidos das imagens na horizontal. O mosaico final pode ser observado na Figura 7. O resultado obtido é bastante satisfatório, apesar de algumas falhas no momento da construção do mesmo. A relação sinal ruído (SNR) obtida em relação ao mosaico obtido e a imagem original é de aproximadamente 13 dB. O tempo total para construção do mosaico, que foi aproximadamente de cinco minutos, ainda está elevado tomando em consideração que será embarcado e ser tomado como mapa de navegação de robô autônomo.

5. CONCLUSÃO

Para a utilização do sistema de visão do robô para localização e navegação, a composição de mosaico se torna bastante útil. Este pode utilizar o mosaico para realizar estimativa de localização e tomar decisões relacionadas à navegação. Para isso, se faz uso da SIFT e comparação da imagem obtida atual com mosaico construído ou em construção. Como trabalhos futuros são sugeridos a mudança do sistema da plataforma de simulação MATLAB para o ambiente C++ Builder, a resolução do problema da construção de mosaicos existindo uma diferença de intensidade luminosa muito grande entre as imagens a serem inseridas no mosaico, a otimização do algoritmo para diminuir o tempo do processo de criação do mosaico, a diminuição do erro na hora de montagem do mosaico final e um estudo sobre outros métodos de obtenção de pontos críticos, dentre eles o SURF (*Speeded Up Robust Features*).

6. AGRADECIMENTOS

Agradecimentos aos laboratórios LEM e ITTI.

REFERÊNCIAS

- ANGELI, A.; DONCIEUX, S.; MEYER, J.-A.; FILLIAT, D. **Incremental vision-based topological slam**. IEEE/RSJ International Conference, p. 1031_1036, 2008.
- CALHEIROS, R. N.; PEREIRA J. M.; SANTOS, R. P. dos; FILHO, N. L. D. **Um sistema para reconhecimento de padrões em imagens digitais subaquáticas**. Congresso Regional de Iniciação Científica e Tecnológica, 2010.
- GOEDEMÉ, T.; NUTTIN, M.; TUYTELAARS, T.; GOOL, L. V. **Omnidirectional vision based topological navigation**. International Journal of Computer Vision, p. 219_236, 2007.
- GONZALEZ, R. C.; WOODS, R. E.; EDDINS, S. I. **Digital Image Processing**. 3rd. ed. Knoxville: Gatesmark Publishing, 2009.
- HASEGAWA, J. K. E JÚNIOR, E.R.A., 2004. **Mosaico com imagens digitais**.

LOWE, D. G. **Distinctive image features from scale-invariant keypoints**. International Journal of Computer Vision, 2004.

LUCAS, A. E. S. **Mapeamento e localização baseados em mosaicos visuais**. Dissertação (Mestrado) - Universidade Técnica de Lisboa, 2009.

MACHADO, É. de L.; FREITAS, R. A. C.; SCHNEEBELI, H. A. **Construção de mosaicos de imagens em plataformas robóticas com auxílio de informações sensoriais**. Congresso Brasileiro de Automática (CBA), 2004.

LUPPE, A. C. C. M. e Roda, V.O., 2005. **Arquitetura para transformada de distância e sua aplicação**.

SIEGWART, R.; NOURBAKHSI, I. R. **Introduction to Autonomous Mobile Robots**. The MIT Press, 2004.

TEIXEIRA, E.; CORTEZ, P. **Desenvolvimento de sistema de visão omnidirecional para localização de robôs móveis autônomos**. Dissertação (Mestrado) - Universidade Federal do Ceará, 2008.

VARGAS, A. P.; MADSEN, C. A. B. C. W.; BOTELHO, S. S. C. **Navision – sistema de visão subaquática para navegação e montagem de mosaicos em auvs**. Seminário e Workshop em Engenharia Oceânica SEMENGO, 2004.

WOLF, P.R., 1983. **Elements of Photogrammetry**. McGraw-Hill Book Company