

# Segmentação e Classificação de Embarcações em Eclusas

Fagner de Assis Moura Pimentel

Orientadora: Michele Fúlvia Angelo

Coorientador: Diego Gervasio Frías Suárez

{fagnerpimentel,diegofriass}@gmail.com,  
mfangelo@ecompu.uefs.br

Mestrado

Programa de Pós-graduação em Ciência da Computação

Ano/semestre de ingresso no programa: 2012.2

Época esperada de conclusão: 2014.1

Exame de qualificação - 21/10/2013

Previsão de defesa - Outubro de 2014

***Abstract.** This project it's about identification and classification of ships in river environments using and comparing techniques for segmentation like Background Subtraction and classifiers like Artificial Neural Networks and Support Vector Machine. This study is expected to define a set of techniques that best fit the segmentation and classification of ships in river environments in order to optimize the flow of ships at locks located on the Tietê river as future application.*

**Keywords.** Segmentation, Classification, Ship, Fluvial Environment, Computer Vision, surveillance.

## **1. Problema de pesquisa e caracterização da contribuição**

A hidrovía Tietê-Paraná no estado de São Paulo, conta com seis eclusas, localizadas nas barragens de Barra Bonita, Bariri, Ibitinga, Promissão e Nova Avanhandava. Estas eclusas são mantidas e operadas pela concessionária AES/Tietê [AES]. A identificação das embarcações é feita pelo operador da eclusa, e o controle de prioridade para entrar na eclusa se dá de forma manual, sem qualquer tipo de otimização.

Com base neste contexto, a proposta deste projeto é desenvolver um sistema de segmentação e classificação de embarcações em ambientes fluviais a partir de imagens extraídas do sistema de vigilância da concessionária AES/Tietê que utilizam câmeras do tipo Pan-Tilt-Zoom, visando o monitoramento e controle automático das mesmas durante as operações de eclusagem e possibilitando uma menor ocorrência de erros no controle e menos gastos desnecessários, como por exemplo, no tempo de espera das embarcações e vertimento da água utilizada.

Este projeto tem como objetivos específicos o estudo e a análise de técnicas de subtração de background que serão utilizadas na fase de segmentação das embarcações; a construção de um dataset com ground truth do tipo bounding box contendo os vários tipos de embarcações que navegam no rio Tietê para a fase de treinamento e testes dos classificadores; a utilização de técnicas de rastreamento para minimizar os custos da segmentação nos vídeo; a comparação e a seleção de técnicas de visão computacional que melhor se adequem a segmentação e classificação de embarcações em ambientes fluviais, baseado nos resultados dos testes.

## **2. Fundamentação teórica e trabalhos relacionados**

### **2.1. Segmentação**

No contexto de segmentação de embarcações, foram encontrados principalmente trabalhos relacionados a métodos de segmentação baseados em região (region-based) como a subtração de background. A seguir são apresentados alguns destes métodos de segmentação.

Em [Luo et al. 2006] é aplicada a subtração de background juntamente com um duplo thresholding para a segmentação de embarcações. Como acontece normalmente em imagens com background dinâmico, não somente o movimento das embarcações como também da água e das árvores são captados pela subtração de background. O duplo threshold, um alto e um baixo, é utilizado visto que um único threshold nem sempre é capaz de diferenciar um objeto móvel do fundo dinâmico. A ideia é explorar a coerência espacial do objeto na imagem resultante da subtração. Primeiramente, é passado um threshold alto determinando um subconjunto de pixels do objeto, em seguida um threshold baixo contendo um subconjunto maior de pixels. O resultado final é dado pelos pixels de baixo threshold que se interligam recursivamente aos pixels de alto threshold.

Em [Szpak and Tapamo 2011] é utilizada a subtração de background para o rastreamento de embarcações em ambientes marítimos, onde a maior parte do background é formado pelo oceano. A principal dificuldade neste trabalho está no fato do oceano ser dinâmico e os objetos terem que ser extraídos em um ambiente altamente imprevisível. Neste trabalho foi desenvolvido um algoritmo próprio de subtração de background além de utilizar a título de comparação, algoritmos do tipo Mistura de Gaussiana [Stauffer and Grimson 1999] e Sigma-Delta [Toral et al. 2009].

## **2.2. Extração de Características**

A extração de características resulta em alguma informação quantitativa de interesse ou que seja básica para a discriminação entre as classes de objetos [Gonzalez and Woods 2000]. Segundo [Luo et al. 2006], a forma é a característica mais importante para a detecção de embarcações.

O descritor de forma MPEG-7, baseado em regiões e descrito em [Salembier and Sikora 2002] foi utilizado em [Luo et al. 2006] para extrair as características de forma das embarcações segmentadas. Segundo [Salembier and Sikora 2002], este descritor é ideal para objetos divididos em regiões disjuntas e objetos que contenham buracos em sua forma.

Outro método utilizado para a extração de características de embarcações é o histograma de gradientes orientados (HOG, do inglês Histogram of Oriented Gradients) [Wijnhoven et al. 2010]. A ideia deste método é que a forma dos objetos pode ser caracterizada pela direção e intensidade dos gradientes em uma imagem. Isso é implementado através da divisão da imagem em pequenas regiões (células). As células são divididas em blocos, e para cada bloco é computado o histograma de gradientes orientados sobre os pixels das células [Dalal and Triggs 2004].

## **2.3. Classificadores**

Em aprendizado de máquina, a classificação é o problema de identificação de um conjunto de categorias com base em um conjunto de dados treinados. A seguir são mostrados alguns classificadores.

### **2.3.1. Rede Neural Artificial(RNA)**

Uma RNA é um classificador frequentemente utilizado. Ela é formada basicamente por duas camadas, uma escondida e uma de saída [HAYKIN 2001]. As camadas são compostas por um conjunto de neurônios definidos por funções de ativação. Apesar das RNAs serem lentas para treinar, elas são muito rápidas para serem executadas.

Em [Alves et al. 2004] foi utilizada uma RNA para a classificação de embarcações com 12 neurônios de entrada, 20 neurônios na camada escondida e 5 na camada de saída. Foram utilizadas 240 imagens para o treinamento, variando os ângulos de visão das embarcações. Ao final do treinamento foi obtida uma taxa de acerto de 90.1%. Em seguida foram realizados testes com 41.400 imagens, obtendo uma taxa de acerto de 87.3%.

### **2.3.2. Máquina de Vetor de Suporte (SVM)**

O SVM (do inglês, Support Vector Machine) é um classificador que também pode fazer regressão. Segundo Cortes e Vapnik [Cortes and Vapnik 1995], o SVM implementa a seguinte ideia: um vetor de entrada é mapeado não linearmente para um espaço de várias dimensões de características. Neste espaço, um hiperplano de decisão linear é construído. O algoritmo propõe a criação de hiperplanos que separam as classes de grandes dimensões. Para o treinamento deste classificador é passado um conjunto de exemplos pertencentes a classes distintas. Este treinamento constrói um modelo SVM que representa os exemplos como pontos no espaço, mapeados de forma que sejam claramente separados por um

hiperplano em um espaço multidimensional. Novos exemplos são mapeados no mesmo espaço do modelo e definidos de qual classe pertencem baseados no seu posicionamento em relação ao hiperplano. Este classificador está entre os melhores com dados limitados [Bradski and Kaehler 2008].

Em [Wijnhoven et al. 2010] foi utilizado SVM para a classificação de embarcações a partir de câmeras Pam-Tilt-Zoom e HOG como extrator de características. Para o treinamento do classificador foram utilizadas 280 imagens positivas e exemplos do background extraído como imagens negativas, obtendo uma taxa de acerto de 85%.

### **3. Estado atual do trabalho**

Até o presente momento foi realizada a etapa que visa determinar as regiões de interesse de uma dada imagem, ou seja, onde existe uma embarcação. Para isso, foram aplicadas sobre os datasets pre-selecionados, técnicas de subtração de background.

Nesta etapa foi realizado um estudo de performance de 27 algoritmos de subtração de background implementados na biblioteca bgslybrary [Sobral 2013] com o objetivo de escolher os melhores algoritmos que apresentem bons resultados para as imagens que serão utilizadas neste projeto.

Foi utilizado o dataset disponibilizado por [Goyette et al. 2012]. Este dataset possui vídeos separados em diversas categorias. A categoria de background dinâmico, dentre as disponíveis no dataset, é a que mais se assemelha aos vídeos utilizados neste projeto, ela apresenta imagens de barcos, carros, e pessoas anotadas manualmente em um fundo com água e árvores em movimento.

A análise de desempenho destes algoritmos de subtração de background foi realizada com base nas seguintes métricas, bastante estudadas na literatura:

- Recall (Re):  $TP/(TP+FN)$
- Specificity (Sp):  $TN/(TN+FP)$
- False Positive Rate (FPR):  $FP/(FP+PN)$
- False Negative Rate (FNR):  $FN/(TN+FP)$
- Percentage of Wrong Classifications (PWC):  $100(FN+FP)/(TP+TN+FP+FN)$
- Precision (Pr):  $TP/(TP+FP)$
- F-measure:  $2(Pr.Re)/(Pr+Re)$

Tais métricas foram selecionadas de modo a facilitar a criação de um ranking entre os algoritmos utilizados. Para cada métrica, TP é número de verdadeiros positivos, TN é o número de verdadeiros negativos, FN é o número de falsos negativos, e FP é o número de falsos positivos.

## **4. Desenvolvimento necessário para a conclusão**

### **4.1. Coleta e preparo dos dados**

Serão criados datasets com imagens positivas (exemplos de embarcações) e negativas (background) das esclusas de Barra Bonita, Bariri, Ibitinga, Promissão e Nova Avanhandava, localizadas ao longo do rio Tietê e que funcionam 24h por dia. Este datasets serão utilizados para o treinamento dos classificadores e também para a realização dos testes.

Para a criação destes datasets serão selecionados vídeos fornecidos pela concessionária AES/Tietê que possuam exemplos das embarcações que trafegam pelas eclusas em diferentes momentos e com diferentes luminosidades, ângulos e distâncias para a análise de desempenho das técnicas implementadas. As imagens serão obtidas a partir do sistema de vigilância da concessionária AES/Tietê que utilizam câmeras de circuito fechado do tipo Pan-Tilt-Zoom, capazes de realizar movimentos na vertical (Pan), horizontal (Tilt) e de aproximação (Zoom) com resolução de 1440x1080 pixels. As imagens dos datasets deverão possuir anotações feitas manualmente das regiões da imagem (bounding box) correspondentes as embarcações para fins de avaliação de desempenho e posterior comparação com os resultados.

## **4.2. Extração de Características**

Segmentada a embarcação, o próximo passo é extrair as características que a descrevam. As principais características que serão avaliadas para a classificação de embarcação serão forma, como visto em [Luo et al. 2006], cor e textura, no entanto, estas duas últimas terão menor importância. Para a extração destas características serão utilizados os extratores MPEG-7 [Salembier and Sikora 2002] e HOG [Wijnhoven et al. 2010].

## **4.3. Classificadores**

Nesta etapa serão utilizados os classificadores SVM como utilizado em [Dalal and Triggs 2004, Wijnhoven et al. 2010] e RNA como usado em [Luo et al. 2006], previamente treinados com datasets positivos (imagens dos tipos de embarcações) e negativos (imagens onde não aparece nenhuma embarcação). A priori serão classificadas as embarcações voltadas para o turismo e para o transporte de mercadorias por serem as que trafegam com mais frequência nas eclusas. Além dos classificadores SVM e RNA, também poderá ser utilizado o KNN (do inglês k nearest neighbors), como usado em [Luo et al. 2006].

## **4.4. Comparação dos resultados obtidos com as técnicas implementadas**

Após a implementação das técnicas e treinamento dos classificadores, serão realizados os testes com as imagens e vídeos separados para esta finalidade. Em seguida serão feitas comparações destes resultados visando selecionar as técnicas que melhor se adequam a segmentação e classificação de embarcações. Nesta etapa serão utilizadas as curvas ROC (do inglês Receiver Operating Characteristic), comumente utilizada para medir desempenho de classificadores e Precision-Recall, que permite visualizar as diferenças entre a precisão e a taxa de acerto.

## **5. Avaliação dos resultados**

A partir dos dados extraídos das métricas de avaliação dos algoritmos, foram feitas comparações a fim de determinar qual algoritmo melhor se aplica ao projeto. Para cada um dos vídeos presentes no dataset de background dinâmico pre-selecionado, foram feitas 27 avaliações referentes a cada algoritmo da biblioteca bgslibrary utilizando as métricas apresentadas anteriormente. Para cada métrica foi feita a média aritmética dos resultados de todos os vídeos do dataset.

O algoritmo Pixel Based Adaptive Segmenter (PBAS) de [Hofmann et al. 2012] foi o que apresentou os melhores resultados, sendo então escolhido para a realização da segmentação de embarcações em ambientes fluviais neste projeto.

## Referências

- AES. AES/tietê. [www.aesbrasil.com.br](http://www.aesbrasil.com.br). Acesso em: 13 de julho de 2013.
- Alves, J., Herman, J., and Rowe, N. C. (2004). Robust Recognition of Ship Types from an Infrared Silhouette.
- Bradski, G. and Kaehler, A. (2008). *Learning OpenCV: Computer Vision with the OpenCV Library* (Google eBook). "O'Reilly Media, Inc."
- Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.
- Dalal, N. and Triggs, W. (2004). Histograms of Oriented Gradients for Human Detection. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR05*, 1(3):886–893.
- Gonzalez, R. C. and Woods, R. E. (2000). *Processamento de imagens digitais*. Edgard Blucher.
- Goyette, N., Jodoin, P.-M., Porikli, F., Konrad, J., and Ishwar, P. (2012). Changedetection.net: A new change detection benchmark dataset. In *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 1–8. IEEE.
- HAYKIN, S. S. (2001). *Redes neurais*. Bookman.
- Hofmann, M., Tiefenbacher, P., and Rigoll, G. (2012). Background segmentation with feedback: The Pixel-Based Adaptive Segmenter. In *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 38–43. IEEE.
- Luo, Q., Khoshgoftaar, T., and Folleco, A. (2006). Classification of Ships in Surveillance Video. In *2006 IEEE International Conference on Information Reuse & Integration*, pages 432–437. IEEE.
- Salembier, P. and Sikora, Thomas/Manjunath, B. (2002). Introduction to MPEG-7: Multimedia Content Description Interface.
- Sobral, A. (2013). {BGSLibrary}: An OpenCV C++ Background Subtraction Library. In *IX Workshop de Visão Computacional (WVC'2013)*, Rio de Janeiro, Brazil.
- Stauffer, C. and Grimson, W. (1999). Adaptive background mixture models for real-time tracking. In *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, volume 2, pages 246–252. IEEE Comput. Soc.
- Szpak, Z. L. and Tapamo, J. R. (2011). Maritime surveillance: Tracking ships inside a dynamic background using a fast level-set. *Expert Systems with Applications*, 38(6):6669–6680.
- Toral, S., Vargas, M., Barrero, F., and Ortega, M. (2009). Improved sigma-delta background estimation for vehicle detection. *Electronics Letters*, 45(1):32.
- Wijnhoven, R., van Rens, K., Jaspers, E., and de With, P. (2010). Online Learning for Ship Detection in Maritime Surveillance. pages 73 – 80.