FISHCV - Reconhecimento de Espécies de Peixe utilizando Visão Computacional em Dispositivos Móveis

Uéliton de Paula Freitas

1 Antecedentes e Justificativa

Dispositivos móveis estão cada vez mais presentes na vida humana. Muitas vezes, possuem capacidade de processamento superior a de computadores *Desktop*. Novos modelos são elaborados e lançados a cada ano. Juntamente com o crescimento do poder de processamento, as câmeras estão cada vez mais sofisticadas e com maior definição das imagens. Com o melhor desempenho dos dispositivos móveis, aplicações que não eram viáveis (devido a grande demanda de processamento e ausência de periféricos viáveis, como câmeras com boa definição) passaram a ser desenvolvidas. Assim, aplicações que utilizam processamento de imagens, como as desenvolvidas no âmbito da Visão Computacional, passaram a ser exploradas com maior intensidade.

Visão Computacional é o conjunto de métodos e técnicas através dos quais os sistemas computacionais interpretam as imagens. Muitas vezes as técnicas de visão computacional são custosas em termos de processamento. Com o melhoramento dos dispositivos móveis estas técnicas podem ser utilizadas. Bibliotecas como a *OpenCV*, uma biblioteca com funções e métodos da Visão Computacional, fornecem suporte para estes tipos de dispositivos. A biblioteca *OpenCV* fornece métodos para o desenvolvimento de aplicações de reconhecimento de padrões, entre outras funções.

O reconhecimento de padrões em peixes é essencial para especificar qual espécie o mesmo esta classificado. Padrões como cor, nadadeiras, formato do peixe definem as diferentes espécies de peixes. Contudo, é difícil classificar os peixes a partir de imagens devido a vários motivos, como por exemplo: imagens são obtidas de peixes em movimento, o ângulo da imagem pode não ajudar a definir o peixe, a iluminação do ambiente da imagem dificulta a detecção do peixe.

Assim, um banco de dados é necessário para armazenar várias imagens dos peixes em diferentes perspectivas. A base de dados deve conter as imagens das espécies de peixes do Mato Grosso do Sul, devido a este trabalho estar relacionado apenas a espécies encontradas no estado citado. As imagens devem conter exemplos de peixes em vários tipos de iluminação, em diferentes tamanhos e estágio de vida. Com uma boa base de dados é possível desenvolver e testar bons classificadores.

Com as informações necessárias para identificar as espécies dos peixes, este projeto tem por objetivo desenvolver uma aplicação em dispositivos móveis para classifica-las. A aplicação, a priori, tem foco em usuários que possam identificar espécies de peixes em aquários. Qualquer usuário poderá identificar espécies de peixes do Mato Grosso do Sul em aquários residênciais ou comerciais. Posteriormente, a aplicação poderá ser implementada para estudos científicos.

Algumas das características marcantes que diferenciam espécies de peixes, são as formas e cores de cada espécie. Assim, para classificar (identificar) as espécies de peixes, serão utilizadas técnicas de processamento de imagens da Visão Computacional, focadas em formas e cores das espécies dos peixes do Mato Grosso do Sul.

Uma das técnicas utilizadas para classificar objetos baseados em formas e cores é o algoritmo *Bag of Words* em imagens coloridas. O algoritmo *Bag of Words* em imagens coloridas será utilizado para classificar as espécies de peixes na aplicação proposta juntamente com outros algoritmos.

2 Objetivos

2.1 Geral

Este trabalho tem como principal objetivo o desenvolvimento de uma aplicação para dispositivos móveis que classifique diversas espécies de peixes. Geralmente são utilizadas câmeras fixas para realizar tal reconhecimento. Esta aplicação irá realizar o reconhecimento de algumas espécies de peixes do Mato Grosso do Sul utilizando a câmera do dispositivo móvel, e algoritmos de classificação tais como, o algoritmo *Bag of Words* em imagens coloridas.

2.2 Específicos

- Aprofundamento e atualização da revisão de literatura, atavés de artigos relacionados a detecção e classificação de espécies de peixes.
- Construção de um banco de imagens dos peixes do Mato Grosso do Sul.
- Implementação do protótipo da aplicação.
- Validação do protótipo.
- Registro e divulgação de resultados.

3 Revisão de Literatura

3.1 Visão Computacional

Riesenhuber and Poggio (2000) descreve que para reconhecer uma imagem é necessário vários processos. Há duas tarefas principais: (1) identificação do objeto, (2) categorização do objeto. Em visão computacional existem técnicas relativamente fáceis de identificação de objetos, segundo Riesenhuber and Poggio (2000). Contudo, classificar os objetos é uma tarefa difícil. Na categorização é necessário um conhecimento do objeto e suas possíveis transformações em imagens tais como: ângulos da câmera e iluminação ambiente.

Para a tarefa de classificação dos objetos o algoritmo *Bag of Words* em imagens coloridas será implementado.

3.2 Bag Of Words

Bag of Words é um algoritmo simples utilizado a priori em classificação de contextos de textos. O texto é representado como uma coleção desordenada de palavras, e através da coleção obtida, classifica-se o contexto do texto analizado. A coleção contém o número de ocorrências de cada palavra no texto e é denominada vocabulário. Por exemplo, as palavras de maior frequência em um texto esportivo podem ser: esporte, futebol, vôlei, natação, compeonato, vencedor. Sendo assim, quando deseja-se classificar um novo texto como sendo um texto esportivo ou não, verifica-se se as ocorrências das palavras a ser classificadas, são semelhantes as ocorrências das palavras no vocabulário do *Bag of Words*. Caso o número de ocorrências das palavras são

semelhantes as da coleção do *Bag of Words* o novo texto é classificado como esportivo.

Mesmo com *Bag of Words* tendo origem na classificação de textos, devido a sua flexibilidade, o *Bag of Words* foi também utilizado na classificação visual. Em vez do vocabulário conter palavaras, ele contém características das imagens e/ou dos objetos contidas nelas. Das várias características dos objetos nas imagens, temos as cores dos objetos, bordas e caracteríscas extraidas através de algoritmos complexos de extração de atributos.

Rojas Vigo et al. (2010) mostra o uso de *Bag of Words* em imagens. Este método utiliza cores como um dos principais atributos na classificação de objetos. Primeiramente o estágio de extração de amostra é aplicado, que envolve a região selecionada que contém o objeto. Em seguida todas as regiões selecionadas são representadas por meio de descritores locais, seguida de quantização vetorial em um *codebook* (dicionário) de palavras visuais. Posteriormente cada imagem é representada em um histograma de palavras visuais baseado na variação de cores. Um classificador (tipicamente um SVM não-linear) é então utilizado para reconhecer as categorias com base nos histogramas das imagens. O uso das cores melhoram os resultados tanto na fase de detecção quanto na fase de extração das amostras das imagens.

Wengert et al. (2011) propõe uma assinatura de cores para criar descritores das imagens. O método de assinatura é denominado *Bag-Of-Colors* (BoC). O BoC define uma palheta de cores e um *codeBook* é criado definindo cada assinatura. Os resultados mostram que os descritores criados relevando as cores e utilizando o arcabouço *Bag of Words*, obtiveram bons resultados.

3.3 Técnicas de Identificação de Espécies de Peixes

Para a identificação de peixes, algumas fases devem ser executadas para encontrar a espécie do peixe propriamente dita. Hu et al. (2012), utilizaram dispositivos móveis para a captura de imagens de peixes e então classificar as espécies dos mesmos baseados e sub-imagens. As imagens foram obtidas de celulares e enviadas via MMS (*Multimedia Message System*). Posteriormente as fotos foram divididas em imagens menores, e então através da imagem da pele do peixe obtinha-se a classificação. A classificação foi obtida através da cor, textura e formato das escamas. As características citadas foram as que melhor definem a classificação da espécie visualmente.

As imagens dos peixes trabalhadas neste projeto serão semelhantes as trabalhados por Hu et al. (2012). Porém, a classificação das espécies dos peixes será realizada no dispositivo móvel. Diferente do método utilizado por Hu et al. (2012), em que as imagens são enviadas para um computador e então classificadas.

Para ser possível as classificações uma base de dados é necessária para armazenar as informações sobre as espécies de peixes. Uma base de dados específica para análise de espécies de peixes foi criada para analisar as espécies encontradas em Wisconsin (Lyons et al., 2006). A base de dados criada contém algumas informações específicas de cada espécie de peixes como mapas, nome cientifico, familia, espécies similares.

Kresimir Williams and Harms (2012), utilizaram para a detecção dos peixes nas imagens o algoritmo Viola and Jones (VJ)(Viola and Jones, 2001). Posteriormente os peixes foram separados do fundo da imagem. Com a imagem

pré-processada, para a detecção das espécies ele utiliza PCA e SIFT. Segundo os experimentos, o algoritmo SIFT possui melhor resultados na classificação dos peixes para reconhecimento em baixo da água.

Nery et al. (2005), mostram o uso de algoritmos de visão computacional no reconhecimento de espécies em usinas hidrelétricas. Um algoritmo é proposto para escolher as amostras de acordo com a correlação das mesmas. Um classificador bayesiano é utilizado para a classificação das espécies. O classificador baseia-se em cor, textura e forma do peixe para a classificação. As cores médias do peixe mostram-se fundamentais para a classificação. Os experimentos mostram classificações com precisão de até 89% com quatro espécies de peixes.

4 Metodologia

Para a execução deste projeto seguem abaixo as descrições metodológicas dos objetivos específicos apresentadas na Seção 2.2.

4.1 Aprofundamento e atualização da revisão de literatura

Através de consultas aos principais portais de periódicos mundiais, como IEEE Xplore, ACM DL, Science Direct e Scopus, serão identificados artigos com trabalhos correlatos nas áreas de *Bag of Words*, classificação de peixes e cores. Estes artigos serão revisados para complementar o texto apresentado neste plano de trabalho.

4.2 Construção de um banco de imagens de peixes

O banco de imagens conterá várias imagens para cada classe do problema. As classes, neste caso, correspondem à espécies de peixes. As imagens serão capturadas principalmente em aquários, utilizando câmeras de dispositivos móveis. Com apoio de especialistas, cada imagem será anotada, constituindo assim um conjunto de referência para análise de desempenho. Para facilitar o acesso posterior às imagens pelo usuário, as mesmas serão inseridas no software fishcv, através do qual as mesmas poderão ser revisadas e corrigidas. O banco também será disponibilizado através do website do projeto FISHCV (http://trac.gpec.ucdb.br/wiki/site_fishcv) . Exemplos de imagens similares às que serão utilizadas neste plano podem ser vistas na Figura 1 e Figura 2.

4.3 Desenvolvimento do módulo Bag of Words em imagens coloridas

O módulo será desenvolvido em Linguagem C++ tendo como apoio o pacote para Visão Computacional OpenCV (Bradski, 2000) versão 2.4 e o software Weka versão 3.6.9. Serão seguidas as regras definidas pelo grupo de pesquisa e desenvolvimento INOVISAO disponíveis no site do grupo. A metodologia de desenvolvimento de software do INOVISAO tem como base o SCRUM (Sims and Johnson, 2011) com todo o material produzido sob controle de versões utilizando a ferramenta SubVersion. O padrão de documentação de código é baseado no JavaDoc (mesmo para programas em C e C++).

Figura 1: Acará Bandeira

Figura 2: Kinguio

4.4 Validação do módulo Bag of Words em imagens coloridas

O algoritmo *Bag of Words* será comparado com os algoritmos *Bag of Words* em tons de cinza e *Bag of Words* em imagens coloridas utilizando o banco de imagens descrito na seção 4.2. Como técnica de amostragem será adotado a validação cruzada de 10 dobras com 10 repetições disponível no software Weka na versão 3.6.9(Hall et al., 2009). Para cada algoritmo testado, serão calculados os desempenhos médios referentes às metricas recall, precision, f-score e percent correct. Para identificar se os algoritmos testados diferem estatisticamente em relação ao desempenho, considerando-se cada uma das métricas, será utilizado o teste não-paramétrico proposto por (Friedman, 1940) e disponível no software estatístico R , versão NAO DEFINIDA, com cada bloco correspondendo a uma das classes do problema. Caso, utilizando-se um ní-

vel de significância de 95% (p-value < 0.05), seja constatada diferença, um teste post-hoc será realizado e os box-plots resultantes analisados. O teste post-hoc, também disponível no R, tem como base o teste de Wilcoxon com correção para FWER (Family-wise Error Rate) descrito por Hollander e Wolf Hollander (1999).

4.5 Registro e divulgação de resultados

Serão produzidos um resumo, um relatório final e no mínimo um artigo científico com resultados finais deste plano. O artigo será submetido para um evento ou revista da área de Visão Computacional. Será utilizada a ferramenta Latex para produção dos textos visando facilitar a adaptação dos mesmos para as regras utilizadas em periódicos e eventos da área da computação e que geralmente disponibilizam modelos em Latex.

5 Cronograma

Tabela 1: Cronograma de 2013

Atividade	Bimestre - 2013					Bimestre - 2014
	2	3	4	5	6	1
Atividade 1.1:Identificação de trabalhos correlatos.	x					
Atividade 1.2:Revisão de trabalhos correlatos.	x					
Atividade 2.1:Preparação do ambiente de captura de imagens.		•				
Atividade 2.1:Anotação das Imagens.		•				
Atividade 2.3: Preparação do banco de dados para armazenamento e recuperação das imagens.		•				
Atividade 3.1: Implementação de técnicas baseadas em *Bag of Words*.			•	•	•	
Atividade 3.2: Documentação dos códigos gerados			•	•	•	
Atividade 4.1: Realização dos experimentos.					•	•
Atividade 5.1: Preparação do artigo, relatório final e resumo.						•

Referências

Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

Friedman, M. (1940). A Comparison of Alternative Tests of Significance for the Problem of m Rankings. *The Annals of Mathematical Statistics*, 11(1):86–92.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., e Witten, I. H. (2009). The weka data mining software: an update. *SIGKDD Explor. Newsl.*, 11(1):10–18.

Hollander, M. e Wolfe, D. (1999). *Nonparametric statistical methods*. Wiley Series in Probability and Statistics - Applied Probability and Statistics Section. Wiley.

Hu, J., Li, D., Duan, Q., Han, Y., Chen, G., e Si, X. (2012). Fish species classification by color, texture and multi-class support vector machine using computer vision. *Comput. Electron. Agric.*, 88:133–140.

Kresimir Williams, C. R. e Harms, J. (2012). Automated techniques for detection and recognition of fishes using computer vision algorithms. *Report of the National Marine Fisheries Service Automated Image Processing Workshop. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-F/SPO-121, 48 p.*

Lyons, J., Hanson, P., e White, E. (2006). A photo-based computer system for identifying Wisconsin fishes. *Fisheries*, 31(6):269–275+.

Nery, M. S., Machado, A. M., Campos, M. F. M., Padua, F. L. C., Carceroni, R., e Queiroz-Neto, J. P. (2005). Determining the appropriate feature set for fish classification tasks. In *Proceedings of the XVIII Brazilian Symposium on Computer Graphics and Image Processing*, SIBGRAPI '05, páginas 173–, Washington, DC, USA. IEEE Computer Society.

Riesenhuber, M. e Poggio, T. (2000). Models of object recognition. *Nature America*.

Rojas Vigo, D. A., Shahbaz Khan, F., van de Weijer, J., e Gevers, T. (2010). The impact of color on bag-of-words based object recognition. In *International Conference on Pattern Recognition*, páginas 1549–1553.

Sims, C. e Johnson, H. (2011). *The Elements of Scrum*. Dymaxicon.

Viola, P. e Jones, M. (2001). Robust real-time object detection. In *International Journal of Computer Vision*.

Wengert, C., Douze, M., e Jégou, H. (2011). Bag-of-colors for improved image search. In *ACM Multimedia'11*, páginas 1437–1440.