

Título do Projeto: Valorização do Mel em Ecossistemas Frágeis: Implantação de Denominação de Origem Controlada em Comunidades do Interior do Mato Grosso do Sul.

Título do Plano de trabalho: Aprendizagem Supervisionada com Bag-of-Words para Identificação da Origem de Amostras de Mel.

Orientador: Hemerson Pistori (pistori@ucdb.br)

Orientanda: Carolini Rodrigues. **RA:** 143759

Curso: Engenharia da Computação **Semestre:** 3º

1. Antecedentes e Justificativa

No Brasil existem diversas atividades no setor rural que estão em processo de automatização ou que praticamente já se encontram automatizadas. A apicultura é uma atividade que está se valorizando no mercado e precisa ser mais explorada em grande escala. O mel é um produto que tem proteínas e diversas vitaminas essenciais à nossa saúde (MOREIRA; MARLA, 2001). Este alimento é obtido através do cultivo de colmeias por pequenos produtores. Para melhor utilização desse produto e comercialização é necessário que haja certo controle. Com o devido controle o comércio e a apicultura poderá gerar maiores lucros à economia e gerar maior renda aos produtores. O estado de Mato Grosso do Sul tem grande potencial apícola, pois possui uma fauna e flora abundante o que favorece a criação de abelhas (EMBRAPA, 1986). O projeto no qual este plano de trabalho se insere tem como base o comércio controlado do mel. Controle esse feito através da análise do pólen encontrado em amostras do produto para estabelecimento de sua origem. Estabelecendo-se a origem do produto espera-se que o mesmo possa ser valorizado, como por exemplo, o CONSAD da Serra da Bodoquena no Sudoeste do Mato Grosso do Sul.

Uma abelha inicia o processo de produção de mel quando colhe o néctar de uma planta. Esta produção está diretamente relacionada ao processo de polinização. Quando a abelha suga o néctar da planta alguns grãos de pólen permanecem na abelha, logo esta retorna à colmeia onde produz o mel. Logo quando ao analisar uma amostra do mel produzido nesta colmeia será

detectado o grão de pólen recolhido anteriormente. Baseado nesta informação é possível saber a origem do mel. Uma análise prévia da flora da região da qual se quer verificar o mel produzido é necessária para se realizar as classificações do produto. Atualmente os grãos de pólen presentes nos méis são contados visualmente (DELGADO et al., 2003). Esse processo é bem detalhado e cansativo, podendo acarretar em falhas. Neste projeto está sendo proposta a utilização de modelos, técnicas da visão computacional e inteligência artificial em imagens microscópicas.

Particularmente, neste plano de trabalho, para a identificação do mel a partir de imagens, serão estudadas técnicas baseadas em aprendizagem de máquina supervisionada com *Bag of Words* (LI et al., 2011) que transforma dados não estruturados em um formato estruturado. Primeiramente será necessário o recolhimento de imagens de amostras de mel e a extração de informações de grãos de pólen.

Este trabalho será de grande importância para o projeto, pois um método de rastreabilidade eficaz é um dos principais fatores para a valorização do mel produzido no estado do Mato Grosso do Sul e também para a preservação do ecossistema, pois apicultura é uma atividade ambiental sustentável devido ao fato de só se desenvolver em ambiente preservado.

2. Objetivos

2.1 Geral

Identificar a origem de amostras de mel através da estratégia Bag of Words com a utilização de imagens microscópicas do produto.

2.2 Específicos

Para atingir o objetivo geral definido na seção 2.1, foram estabelecidos os seguintes objetivos específicos:

- ▣ Aprofundamento e atualização da revisão de literatura
- ▣ Construção de um banco de imagens de amostras de mel.
- ▣ Implementação do módulo baseado em Bag of Words.
- ▣ Validação do módulo de baseado em Bag of Words.
- ▣ Registro e divulgação de resultados.

3. Revisão de literatura.

3.1 Categorizações Visuais

Atualmente existe uma grande quantidade de informação visual, principalmente com o crescimento do uso da internet. Formas de busca, classificação e categorização de imagens são de grande ajuda ao usuário. Dada uma classificação eficiente de uma imagem pode-se raciocinar, pesquisar com base na nova imagem.

Anteriormente já foram feitos estudos sobre a detecção de imagem, focando em rostos, carros e pedestres (OSUNA; FREUND; GIROSI, 1997). As formas de detecção de imagem requerem um alinhamento preciso e manual e a separação destas imagens em diferentes pontos de vista (LI et al., 2002).

Existe um algoritmo usado para detectar pontos de interesse a partir de uma imagem, o SURF (BAY et al., 2008). Para cada ponto identificado o SURF gera um vetor de 64 posições. O SURF é um algoritmo muito usado a fim de gerar respostas rápidas. Um método utilizado para categorização de imagens é o Bags of Keypoints (CSURKA et al.; 2004). Suas principais vantagens é a eficiência computacional, sua resistência a mudanças de iluminação e sua simplicidade. Contudo existem alguns problemas como a detecção e reconhecimento de imagem.

A categorização visual traz três principais problemas:

Reconhecimento: É a distinção de objetos particulares, ou seja, o reconhecimento de imagens estruturalmente distintas.

Conteúdo de Recuperação Baseada em Imagens: Refere-se à recuperação de imagens de baixo nível. Esses recursos são obtidos a partir da consulta de uma imagem.

Detecção: Baseia-se se partes da categoria visual estão presentes em uma determinada imagem.

Existem três principais passos a serem seguidos na abordagem de categorização de imagens. Em primeiro lugar as características locais são extraídas a partir de imagens de vídeo ou quadros chaves traduzidas para os descritores de recursos, em seguida as palavras visuais são aprendidas por

clustering. Em segundo lugar, os números de ocorrência de palavras visuais são contados e as duas relações contextuais são medidas a partir de estatísticas. Logo as imagens são representadas e classificadas.

3.2 Bag of Words

O Bag of Words (BOW) tem origem na classificação de palavras, que com o tempo foi se tornando eficaz na categorização visual devido sua flexibilidade. As configurações e recursos do Bag of Words tem sido muito utilizadas desde que foi introduzido, pois possui uma categorização de alto desempenho (YANG, 2007).

O BOW utiliza-se do modelo de arco para a representação de imagens. A princípio o modelo de arco tem como base, ou entrada, um documento de texto do qual se retira todas as palavras formando um vetor que dá origem à um dicionário (MATSUBARA, 2003). Em representação de visual os pontos extraídos a partir de um conjunto de imagens formam um dicionário. Com o agrupamento de imagens realizado agora cada característica é buscada através de uma palavra visual, assim representando uma imagem. Nesse processo o vocabulário ajuda a converter as características locais em um vetor, o que é conveniente para a aprendizagem padrão.

Para conseguir isso existem três etapas: a detecção de recurso, descrição característica e geração de codebook. Após a detecção das características de uma imagem esta é captada. Detectores de pontos, ou descritor é usada para converter as informações em um vetor. Os detectores mais usados são o SIFT e o SURF. Agora cada imagem é uma coleção de vetores. O último passo é gerar um codebook, ou também chamado de dicionário visual. É gerada uma “palavra chave” a partir de coisas em comum em uma imagem, como por exemplo, manchas. Esse processo é um método de agrupamento ou clustering. Assim uma imagem é buscada ou mapeada através dessas palavras chaves.

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 Aprendizagem Supervisionada, SURF, Bag Of Words, Categorização Visual e Classificação de imagens microscópicas. Estes artigos serão revisados para complementar o texto apresentado neste plano de trabalho.

4.2. Construção de um banco de imagens de amostras de mel

Para ajustar os parâmetros do módulo que será desenvolvido e também para testar seu desempenho, conforme será descrito na seção 4.4., um banco de imagens será construído. O banco conterá no mínimo 270 imagens. As imagens serão capturadas diretamente do pote mel e transferidas para uma placa. Esse processo é chamado de plaqueamento. Através deste processo as amostras podem ser analisadas microscopicamente, utilizando um microscópio do tipo tradicional marca Bel acoplado com um dispositivo de captura com lentes 400X. 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 em um banco de dados, através do qual as mesmas poderão ser revisadas e corrigidas. O banco também será disponibilizado através do website do projeto Mel¹. Exemplos de imagens similares às que serão utilizadas neste plano podem ser vistas nas Figura 1, 2 e 3.

¹ O endereço do website do projeto Mel é trac.gpec.ucdb.br/wiki/site_melvic

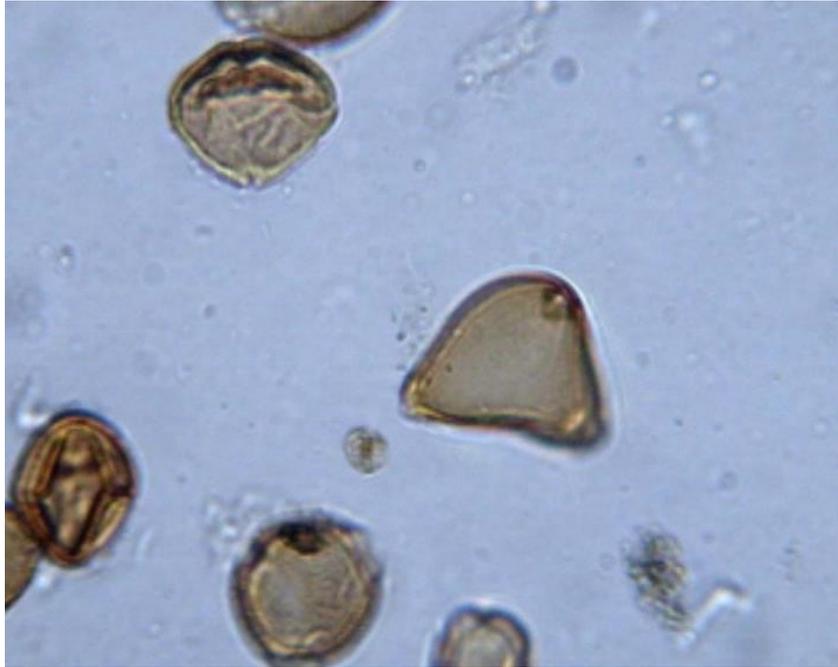


Figura1. Imagem obtida de amostras de mel contendo grãos de pólen.



Figura2. Imagem obtida de amostras de mel contendo grãos de pólen de diferentes espécies.



Figura3. Imagem de uma amostra de mel com grãos de Pólen tipo Serjania.

4.3. Desenvolvimento do módulo Bag of Words

O módulo será desenvolvido em Linguagem C++ tendo como apoio o pacote para Visão Computacional OpenCV (BRADSKI, 2000) versão 2.4.1 e o software Weka versão 3.6. Serão seguidas as regras definidas pelo grupo de pesquisa e desenvolvimento INOVISAO disponíveis no site do grupo². A metodologia de desenvolvimento de software do INOVISAO tem como base o SCRUM (SIMS; JOHNSON, 2011) com todo o material produzido sob controle de versões utilizando a ferramenta SubVersion³. O padrão de documentação de código é baseado no JavaDoc (mesmo para programas em C e C++).

² O site do INOVISAO está em www.gpec.ucdb.br/inovisao e as instruções para desenvolvedores pode ser acessada através do link “trac”, neste mesmo site, ou diretamente em trac.gpec.ucdb.br.

³ O software de controle de versões subversion é apresentado em <http://subversion.apache.org/>.

4.4. Validação do módulo Bag of Words

Os algoritmos implementados serão comparados entre si utilizando o banco de imagens descrito na seção 4.2. Como técnica de amostragem será adota a validação cruzada de 10 dobras com 10 repetições disponível no software Weka⁴ na versão 3.6 (HALL et al., 2009). Para cada algoritmo testado, serão calculados os desempenhos médios referentes às métricas *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 R⁵, versão 2.15.1, com cada bloco correspondendo a uma das classes do problema. Caso, utilizando-se um nível de significância de 95% ($p\text{-value} < 0.05$), seja constatado 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 (1999).

4.5. Registro e divulgação de resultados

Será produzido 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 Latex⁶ 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.

⁴ O Weka é um software livre e gratuito disponível em <http://www.cs.waikato.ac.nz/ml/weka/>.

⁵ O software R está disponível em <http://www.r-project.org/>. Para o teste de Friedman com análise post-hoc é necessário instalar o

⁶ O editor de textos Latex é livre e gratuito e pode ser obtido em <http://www.latex-project.org/>

5. Cronograma

	2012					2013						
	08	09	10	11	12	01	02	03	04	05	06	07
Atividade 1.1.: Identificação dos trabalhos correlatos	X											
Atividade 1.2.: Revisão dos trabalhos correlatos	X	X	X									
Atividade 2.1.: Preparação do ambiente e captura de imagens		X	X									
Atividade 2.2.: Anotação das imagens			X	X								
Atividade 2.3.: Preparação do banco de dados para armazenamento e recuperação das imagens			X	X	X							
Atividade 3.1.: Implementação de técnicas baseadas em Bag of Words.				X	X	X	X	X				
Atividade 3.2.: Documentação dos códigos gerados				X	X	X	X	X	X			
Atividade 4.1.: Realização dos experimentos							X	X	X	X		
Atividade 5.1.: Preparação do artigo, relatório final e resumo.										X	X	X

6. Referências Bibliográficas

BAY, H.; ESS, A.; TUYTELAARS, T.; GOOL, L. V. "Speeded-up robust Features (surf)," *Comput. Vis. Image Underst.*, vol. 110, pp. 346–359, June 2008.

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

CSURKA, Gabriella; DANCE, Christopher R.; FAN, Lixin. WILLAMOWSKI, Jutta. BRAY, Cédric. *Visual Categorization with Bags of Keypoints*. In Workshop on Statistical Learning in Computer Vision. ECCV.2004.

DELGADO, Fernandez M.; OTERO, Pilar Sá; Carrión, P; CERNADAS, E.; GÁLVEZ, J.F. *Improved Classification Of Pollen Texture Images Using SVM and MLP*. VIIP2003: 3rd IASTED.

EMBRAPA, *Inventário da fauna e flora apícola do pantanal em Mato Grosso do sul*. N° 03 setembro/86 1/18.

FRIEDMAN, M. A comparison of alternative tests of significance for the problem of m rankings. *The Annals of Mathematical Statistics*, v. 11, n. 1, p. 86–92, 1940.

HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. The WEKA Data Mining Software: An Update. *SIGKDD Explorations*, v. 11, n. 1., 2009.

HOLLANDER, M.; WOLF, D. A. *Nonparametric Statistical Methods. 2nd Edition*. New York: John Wiley & Sons, 1999.

KOGLER, Marian. LUX, Mathias. *Bag of Visual Words revisited - An exploratory study on robust image retrieval exploiting fuzzy codebooks*. 25 Jul 2010.

LI, S. Z.; ZHU, L.; ZHANG, Z. Q.; BLAKE, A. ZHANG, H. J.; SHUM, H. *Statistical learning of multi-view face detection, ECCV (European Conference on Computer Vision)*. 2002.

LI, Teng; MEI Tao; KWEON, In-So; Member; *IEEE*; and HUA, Xian-Sheng, Member, *IEEE*. *Contextual Bag-of-Words for Visual Categorization*. *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY*. VOL. 21, NO. 4, APRIL 2011.

LOWE, David G. *Distinctive Image Features from Scale-Invariant Keypoints*. January 5, 2004.

MATSUBARA, Edson Takashi; MARTINS, Claudia Aparecida.;MONARD, Maria Carolina. *PreText: uma ferramenta para pré-processamento de textos utilizando a abordagem bag-of-words*. São Carlos, Agosto 2003.

MOREIRA, Ricardo Felipe Alves; MARLA, Carlos Alberto Bastos de. *Glicídios do mel*. Quim. Nova, Vol. 24, No. 4, 516-525, 2001.

OSUNA, E.; FREUNDA, R.; GIROSI, F. *Training support vector machines: An application to face detection, CVPR (Computer Vision and Pattern Recognition)*.1997.

SIMS, C.; JOHNSON, H. L. *The Elements of Scrum*. Dymaxicon, 2011.

YANG, J.; JIANG, Y. G.; HAUPTMANN, A. G.; Ngo, C. W. "Evaluating bagof-visual-words representations in scene classification," in Proc. Assoc. Comput. Machinery Int. Conf. Multimedia Informat. Retrieval (ACM MIR), Augsburg, Germany, 2007, pp. 197–206.