UNIVERSIDADE FEDERAL DE GOIÁS – REGIONAL CATALÃO
UNIDADE ACADÊMICA ESPECIAL DE MATEMÁTICA E TECNOLOGIA
MESTRADO ACADÊMICO EM MODELAGEM E OTIMIZAÇÃO

André Reis de Geus

CLASSIFICAÇÃO DE INSETOS EM MILHO À GRANEL POR MEIO DE ANÁLISE DE VÍDEOS ENDOSCÓPICOS

DISSERTAÇÃO DE MESTRADO

CATALÃO – GO, 2016

sistema de bibliotecas ufg

UFG

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR AS TESES E DISSERTAÇÕES ELETRÔNICAS (TEDE) NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

1. Identificação do material bibliográfico: **[X] Dissertação** **[ ] Tese**

2. Identificação da Tese ou Dissertação

Autor (a):	André Reis de Geus				
E-mail:	geus.andre@gmail.com				
Seu e-mail pode ser disponibilizado na página? [X]Sim [ ] Não					
Vínculo empregatício do autor	Professor na Universidade Federal de Goiás				
Agência de fomento:	Fundação de Amparo à Pesquisa do Estado de Goiás	Sigla:	FAPEG		
País:	Brasil	UF:	GO	CNPJ:	
Título:	Classificação de Insetos em Milho à Granel por meio da Análise de Vídeos Endoscópicos				
Palavras-chave:	Classificação de insetos; Subtração de plano de fundo; Extração de características; Classificador SVM				
Título em outra língua:	Insects classification in maize by endoscopic vídeo analysis				
Palavras-chave em outra língua:	Insects classification; background subtraction; feature extraction; SVM classifier				
Área de concentração:	Modelagem e Otimização				
Data defesa: (dd/mm/aaaa)	10/03/2016				
Programa de Pós-Graduação:	Programa de Pós-Graduação em Modelagem e Otimização				
Orientador :	Dr. Sérgio Francisco da Silva				
E-mail:	sergio@ufg.br				
Co-orientador:*	-				
E-mail:	-				

*Necessita do CPF quando não constar no SisPG

3. Informações de acesso ao documento:

Concorda com a liberação total do documento [X] SIM [ ] NÃO[1]

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF ou DOC da tese ou dissertação.

O sistema da Biblioteca Digital de Teses e Dissertações garante aos autores, que os arquivos contendo eletronicamente as teses e ou dissertações, antes de sua disponibilização, receberão procedimentos de segurança, criptografia (para não permitir cópia e extração de conteúdo, permitindo apenas impressão fraca) usando o padrão do Acrobat.

André Reis de Geus

Assinatura do autor

Data: 11/03/2016

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

ANDRÉ REIS DE GEUS

CLASSIFICAÇÃO DE INSETOS EM MILHO À GRANEL POR MEIO DE ANÁLISE DE VÍDEOS ENDOSCÓPICOS

Dissertação apresentada como requisito parcial para a obtenção do título de Mestre em Modelagem e Otimização pela Universidade Federal de Goiás – Regional Catalão.

IBiotec, Regional Catalão/UFG (Orientador):
Sérgio Francisco da Silva

CATALÃO – GO

2016

Ficha catalográfica elaborada automaticamente
com os dados fornecidos pelo(a) autor(a), sob orientação do Sibi/UFG.

Reis de Geus, André
Classificação de insetos em milho à granel por meio de análise de vídeos endoscópicos [manuscrito] / André Reis de Geus. - 2016.
LXI, 61 f.: il.

Orientador: Prof. Dr. Sérgio Francisco da Silva.
Dissertação (Mestrado) - Universidade Federal de Goiás, Regional Catalão , Catalão, Programa de Pós-Graduação em Modelagem e Otimização, Catalão, 2016.
Bibliografia.
Inclui siglas, lista de figuras, lista de tabelas.

1. Classificação de insetos. 2. Subtração de plano de fundo. 3. Extração de características. 4. Classificador SVM. I. Francisco da Silva, Sérgio, orient. II. Título.

ANDRÉ REIS DE GEUS

CLASSIFICAÇÃO DE INSETOS EM MILHO À GRANEL POR MEIO DE ANÁLISE DE VÍDEOS ENDOSCÓPICOS

Dissertação apresentada ao Programa de Pós-Graduação em Modelagem e Otimização da Universidade Federal de Goiás – Regional Catalão.

Trabalho **aprovado** em 10 de março de 2016.

Prof. Dr. Sérgio Francisco da Silva
Orientador

Prof. Dr. Marcos Napoleão Rabelo
IMTec/UFG-Regional Catalão

Prof. Dr. Marcos Luiz Henrique
NICIT/UFPE-Centro Acadêmico do Agreste

Catalão – GO
2016

Dedico este trabalho à toda a minha família, especialmente aos meus pais Arthur de Geus e Edna Maria Reis de Geus, que sempre me deram força, incentivo e a orientação necessária para superar todos os obstáculos.

Agradecimentos

Aos professores da graduação e da pós-graduação, especialmente ao meu orientador, prof. Dr. Sérgio Francisco da Silva, pelo direcionamento desta pesquisa. Agradeço também aos professores do departamento, Dr. Marcos Napoleão Rabelo, Dr. Marcos Aurélio Batista e Dr. Tércio Alberto dos Santos Filho não só pelas contribuições nesta pesquisa, mas também pela amizade.

"When you are about to quit, remember why you started." (Autor desconhecido)

RESUMO

GEUS, A. R.. *Classificação de insetos em milho à granel por meio de análise de vídeos endoscópicos*. 2016. 61 f. Dissertação (Mestrado em Modelagem e Otimização) – Unidade Acadêmica Especial de Matemática e Tecnologia, Universidade Federal de Goiás – Regional Catalão, Catalão – GO.

Insetos causam perdas quantitativas e qualitativas significantes em grãos armazenados. Neste cenário, é de vital importância uma identificação rápida de insetos em grãos para que sejam tomadas medidas de controle. Ao invés de coletar amostras de grãos para análise visual/laboratorial, é proposta a realização desta tarefa de identificação de forma automática, usando métodos computacionais para a análise de vídeos endoscópicos. Os vídeos são gravados dentro de armazéns de grãos usando câmera endoscópica. Como o processo de classificação de objetos em movimento em vídeo depende fundamentalmente de uma segmentação de objeto precisa, é proposto um novo método de segmentação por subtração de plano de fundo e comparado seus resultados com os principais métodos da literatura de acordo com um estudo de revisão recente. A principal inovação neste método de subtração de plano de fundo está no processo de binarização que usa dois *thresholds*: um global e um local. Os resultados binarizados são combinados pela adição de detalhes do objeto obtido pelo *threshold* local no resultado do *threshold* global. Resultados experimentais, realizados através de análise visual dos resultados de segmentação e usando um classificador SVM indicam que o método de segmentação proposto produz melhores resultados que métodos do estado da arte atual da literatura de subtração de plano de fundo.

Palavras-chaves: Classificação de insetos, Subtração de plano de fundo, Extração de características, Classificador SVM.

ABSTRACT

GEUS, A. R.. *Classificação de insetos em milho à granel por meio de análise de vídeos endoscópicos.* 2016. 61 f. Master Thesis in Modelling and Optimization – Unidade Acadêmica Especial de Matemática e Tecnologia, Universidade Federal de Goiás – Regional Catalão, Catalão – GO.

Insects cause significant losses of stored grains in both quantity and quality. In the scenary, it is of paramount importance an early identification of insects in grains to take control measures. Instead of sampling and visual/laboratory analysis of grains, we propose to carry out the insects identification task automatically, using computational methods to perform endoscopic video analysis. The videos are recorded inside of grains warehouses by an endoscopic camera. As the classification process of moving objects in video rely heavily on precise segmentation of moving objets, we propose a new method of background subtraction and compared their results with the main methods of the literature according to a recent review. The main innovation of the background subtraction method rely on the binarization process that uses two thresholds: a global and a local threshold. The binarized results are combined by adding details of the object obtained by the local threshold in the result of the global threshold. Experimental results performed through visual analysis of the segmentation results and using a SVM classifier, suggest that the proposed segmentation method produces more accurate results than the state-of-art background subtraction methods.

Keywords: Insects classification, Background subtraction, Feature Extraction, SVM classifier.

LISTA DE FIGURAS

Figura 1 – Silo de metal. 32
Figura 2 – Insetos do milho: (a) Sitophilus zeamais; (b) Tribolium castaneum; (c)Sitotroga cerealella. 32
Figura 3 – Exemplo de classificação entre os dígitos manuscritos 0 e 8. 38
Figura 4 – Imagem de objeto binarizada com código da cadeia sobrescrito na sua borda. 38
Figura 5 – Código diferencial da cadeia 39
Figura 6 – Classificadores lineares em um plano bidimensional. 41
Figura 7 – Classificador linear no espaço de características. 41
Figura 8 – Ambiente simulado para a construção da base de dados. 44
Figura 9 – Imagens da câmera endoscópica: (a) Sem ajuste mecânico; (b) Com ajuste mecânico. 44
Figura 10 – (a) Câmera endoscópica; (b) Ajuste mecânico. 45
Figura 11 – (a) Plano de fundo; (b) Imagem *Frame* atual; (c) máscara; (d) Imagem binarizada do *threshold* global; (e) Imagem binarizada do *threshold* local; (f) Imagem final binarizada. 47
Figura 12 – Reconstrução do contorno usando os coeficientes do descritor elíptico de Fourier 49
Figura 13 – *Design* dos experimentos de classificação usando validação cruzada *k-fold* e o classificador SVM. 49
Figura 14 – Imagens originais e resultados de segmentação obtidos pelos métodos analisados 55
Figura 15 – Continuação da Figura 14 56

LISTA DE TABELAS

Tabela 1 – Taxa de acerto na classificação. 53
Tabela 2 – Matriz de confusão dos experimentos de classificação. 53
Tabela 3 – Métodos de subtração de plano de fundo e seus parâmetros utilizados . . . 54

SUMÁRIO

1 INTRODUÇÃO 31

2 REFERENCIAL TEÓRICO 35
2.1 Métodos de Subtração de Plano de Fundo 35
2.1.1 Métodos Estatísticos 35
2.1.2 Métodos Nebulosos e Neurais 36
2.1.3 Eigenbackground 36
2.2 Métodos de Reconhecimento de Padrões 37
2.2.1 Extração de Características 37
2.2.2 Classificador SVM 40

3 METODOLOGIA 43
3.1 Contextualização do Problema 43
3.2 Método proposto 44
3.2.1 Inicialização de plano de fundo 45
3.2.2 Atualização de plano de fundo 46
3.2.3 Extração e binarização de objeto 46
3.3 Classificação 48

4 RESULTADOS 51

CONCLUSÃO 57

REFERÊNCIAS 59

Capítulo 1

Introdução

A crescente necessidade de comida para suprir a demanda mundial requer que os grãos colhidos sejam mantidos com perdas mínimas até o consumo. Contudo, grãos armazenados são altamente suscetíveis a ataques de insetos. Elias *et al.* (2008) reportam que no Brasil há vários tipos de armazéns de grãos, porém, todos são suscetíveis a ataques de insetos. Como o uso intensivo de inseticidas tem sido desencorajado por fatores como prejuízo à saúde animal e humana, e pela origem de insetos resistentes a inseticidas, a identificação prematura de insetos é fundamental para tomada de decisões com o objetivo de evitar perdas. Ainda de acordo com Elias *et al.* (2008), as perdas anuais de grãos pela infestação de insetos são estimadas em 2 bilhões de dólares, calculada em função da perda de peso, volume e de qualidade.

De acordo com Barberato (2001), a quantidade média de perda de grãos no Brasil, estimada pelo Ministério da Agricultura, Pecuária e Abastecimento é de 10% da produção anual. Entretanto, as perdas em qualidade são ainda maiores, uma vez que grãos deteriorados tem valor monetário menor. A maior parte dessas perdas é devido à infestação de insetos. Loeck (2002) e Elias *et al.* (2008) argumentam que a análise periódica de amostras de grãos é uma das medidas mais efetivas para identificar insetos. Contudo, coletar e analisar tais amostras resultam em um alto custo. Considerando um silo metálico real (ilustrado na Figura 1), o processo de amostragem, onde deve-se coletar amostras distribuídas ao longo do interior do armazém, é muito difícil. Este processo pode ser extremamente simplificado pela instalação de câmeras endoscópicas no interior dos silos e, então, usar técnicas de visão computacional para identificar insetos nos grãos, assim como a espécie destes.

Nesta dissertação analisamos a presença de insetos em milho à granel. Plantações de milho cobrem cerca de 20 milhões de hectares no Brazil com uma produção média de 80 milhões de toneladas de grãos por ano. Objetivando evitar perdas durante o período chuvoso, plantadores normalmente colhem o milho em grão uma vez por ano. O milho à granel é armazenado por aproximadamente um ano, ficando altamente suscetível a perca de peso e qualidade devido, principalmente, à infestação de insetos. Estimativas sugerem que, no

Brasil, as perdas pós colheita do milho à granel estão entre 14% do seu peso total. Além das perdas quantitativas em peso e volume, insetos em milho armazenado também estão ligados a contaminação por aflatoxina que podem levar ao envenenamento dos seres vivos que se alimentam dos grãos (DOMENICO *et al.*, 2015). No Brasil, o gorgulho-do-milho (Sitophilus zeamais), o besouro-castanho (Tribolium castaneum) e a traça-dos-cereais (Sitotroga cerealella) são os principais insetos que causam perdas em milho armazenado. Estes insetos são ilustrados na Figura 2. Contudo, a traça-dos-cereais é um inseto de superfície, o qual é facilmente perceptível por observação humana. Assim, nesta pesquisa focamos na classificação do gorgulho-do-milho e do besouro-castanho. Esta decisão também é resultante do fato de uma câmera endoscópica, tal como a que foi usada, não ser um dispositivo adequado para a captura de imagens na superfície dos grãos.

Figura 1 – Silo de metal.

Fonte: EMBRAPA

Figura 2 – Insetos do milho: (a) Sitophilus zeamais; (b) Tribolium castaneum; (c)Sitotroga cerealella.

(a) (b) (c)

Fonte: EMBRAPA

O controle contra insetos do milho tem sido feito principalmente por inseticidas. Entretanto, diversos estudos têm mostrado que insetos do milho resistentes a pesticidas têm

surgido. Desta forma, pesticidas devem ser usados com moderação, quando os insetos aparecerem. Assim, para evitar perdas em grãos armazenados é necessário uma detecção prematura e classificação dos insetos nos grãos para rapidamente aplicar o pesticida correto. Portanto, ferramentas automáticas baseadas em visão computacional são promissoras para detectar objetos em movimento dentro de armazéns de grãos e classificá-los.

Há duas abordagens básicas para a detecção de objetos em movimento em vídeos: fluxo óptico (*optical flow*) e subtração de plano de fundo (*background subtraction*). Resumidamente, técnicas de fluxo óptico (FORTUN; BOUTHEMY; KERVRANN, 2015) calculam vetores velocidade dos objetos em movimento. Estes vetores velocidade, uma vez computados, podem ser usados para uma ampla variedade de tarefas indo de navegação até exploração de ambientes por agentes autônomos. Métodos de subtração de plano de fundo (BOUWMANS, 2014) estimam e mantêm um modelo de plano de fundo (*background*) que é subtraído do *frame* atual para obter o *foreground* (segmentação do(s) objeto(s) em movimento). Esta pesquisa foca em métodos de subtração de plano de fundo pois estes métodos permitem delimitar as fronteiras dos objetos e sabe-se que diferentes espécies de insetos que acometem milho à granel têm diferentes formas.

Vários estudos revelam que na classificação de objetos em vídeos, a segmentação dos objetos em movimento tem um papel fundamental (WANG *et al.*, 2003; CELIK *et al.*, 2007; CHANCHANACHITKUL *et al.*, 2015; NGHIEM, 2014; CHEN; ELLIS; VELASTIN, 2012). Nesta pesquisa é proposto um método de segmentação de objetos em movimento, sendo seus resultados comparados com os métodos mais promissores da literatura de acordo com Sobral e Vacavant (2014). Para a descrição dos objetos foi utilizado o descritor elíptico de Fourier (KUHL, 1982), o qual é robusto ao tamanho do objeto, translação e rotação. Embora vários estudos comparem métodos de subtração de plano de fundo, para o conhecimento dos autores este é o primeiro estudo a comparar tais métodos no contexto de classificação de insetos com base em vídeos endoscópicos. Experimentos com videos reais, obtidos com uma câmera endoscópica revelam que o método de subtração de plano de fundo proposto leva a resultados de classificação mais precisos que métodos alternativos do estado da arte.

O restante desta dissertação está organizado da seguinte maneira. No Capítulo 2 é apresentado o referencial teórico, que descreve os métodos computacionais utilizados. O Capítulo 3 apresenta a metodologia proposta nesse trabalho, a qual consiste do método de segmentação proposto, da descrição dos objetos usando o descritor Elíptico de Fourier e da classificação por *Support Vector Machine* (SVM). O Capítulo 4 apresenta os resultados obtidos dando ênfase na análise das taxas de classificação obtidas e na análise visual dos resultados obtidos pelos métodos de segmentação experimentados, comparados à aqueles do método proposto.

Capítulo 2

Referencial Teórico

2.1 Métodos de Subtração de Plano de Fundo

Vários métodos de subtração de plano de fundo têm sido propostos visando a execução de tarefas de análise de vídeos. Basicamente, todos os métodos de subtração de plano de fundo tentam estimar efetivamente um modelo de plano de fundo a partir de uma sequência temporal de *frames*. O modelo de plano de fundo é primeiramente inicializado e então mantido ao longo da sequência de *frames* do vídeo. Para estimar o *foreground*, isto é, os objetos em movimento, o *frame* corrente é subtraído do modelo de plano de fundo corrente. Na literatura há uma ampla variedade de técnicas para estimar o modelo de plano de fundo. O método mais simples, chamado Diferença Estática de Frames (*Static Frame Difference*), consiste na seleção manual de uma imagem estática no vídeo que representa o plano de fundo, não contendo nenhum objeto em movimento. Entretanto, tal abordagem não é satisfatória em casos onde há mudanças na iluminação, prejudicando a extração dos objetos em movimento (*foreground*) pois este método não conta com um mecanismo de atualização do plano de fundo. Para lidar com a questão de atualização do plano de fundo ao longo da sequência de *frames* do vídeo, vários métodos têm sido propostos envolvendo o uso de abordagens estatísticas, nebulosas (*fuzzy*) e neuro-inspiradadas. A seguir os métodos de maior destaque na literatura serão apresentados.

2.1.1 Métodos Estatísticos

O método de média gaussiana (do inglês, *Gaussian Average*), descrito por Wren *et al.* (1997), modela o plano de fundo baseado na construção ideal de uma distribuição normal dos valores dos últimos n *pixels*. Para cada *pixel*, uma média μ_t e um desvio padrão σ_t são calculados para cada canal de cor. Para cada *frame* t, os *pixels* de I_t podem ser classificados como *foreground* de acordo com a inequação:

$$|I_t - \mu_t| > k\sigma_t \tag{2.1}$$

Um dos mais populares métodos de subtração de plano de fundo é baseado em um modelo probabilístico e paramétrico, proposto por Stauffer e Grimson (1999), posteriormente melhorado por (HAYMAN; EKLUNDH, 2003). Neste algoritmo, chamado Mistura de Gaussianas, as distribuições para a cor de cada *pixel* são representadas pela soma de distribuições gaussianas ponderadas. A cada novo *frame* no tempo t, os parâmetros para todos os *pixels* são atualizados para determinar as variações de cores. De fato, no tempo t, é considerado que o modelo m_t, gerado para cada *pixel* a partir das medidas $\{z_0, z_1, ..., z_{t-1}\}$ de um *pixel*, é correto. A probabilidade que um determinado *pixel* seja pertencente ao plano de fundo, é dada por:

$$P(z_t|m_t) = \sum_{n=1}^{N} \frac{\alpha_n}{(2\pi)^{d/2}|\sum_n|^{1/2}} e^{-1/2(z_t-\mu_t)^T \sum^{-1}(z_t-\mu_t)} \tag{2.2}$$

onde d é a dimensão do espaço de cor das medidas z_t e cada gaussiana n é descrita pela sua média μ_n e matriz de covariância $\sum_n$. As gaussianas são ponderadas pelos fatores α_n onde $\sum_{n=1}^{N} \alpha_n = 1$. $|.|$ denota a matriz determinante. Os canais (R, G, B) de cada pixel são considerados independentes. Para atualizar o modelo de mistura de gaussianas, somente agrega-se tais medidas a uma gaussiana n se:

$$||z_t - \mu_t|| < k\sigma_n \tag{2.3}$$

onde k é 2 ou 3, e σ_n é um vetor representando a variância da distribuição gaussiana de índice n. Posteriormente, as gaussianas são atualizadas a cada novo *frame* conforme um coeficiente de aprendizado.

2.1.2 Métodos Nebulosos e Neurais

Basicamente, uma rede neural é treinada para classificar cada *pixel* da imagem. Para cada *pixel* a rede determina se ele é pertencente ao plano de fundo ou ao *foreground*. Em (CULIBRK *et al.*, 2007), os autores usaram uma rede neural de alimentação direta com 124 neurônios. Esta abordagem se encaixa em redes neurais probabilísticas, já que o plano de fundo é construído através de uma rede neural, enquanto um classificador Bayesiano determina se eles pertencem ao plano de fundo ou não. No trabalho de Maddalena e Petrosino (2009), foi utilizada uma rede não supervisionada do tipo *Self Organizing Map* (SOM) para segmentar os objetos, onde cada *pixel* tem um mapa neural dados por vetores de peso. A principal contribuição deste trabalho é a adição de uma função nebulosa na etapa de aprendizado do plano de fundo.

2.1.3 Eigenbackground

O método Eigenbackground, proposto por Oliver, Rosario e Pentland (2000), é baseado na decomposição de autovalor, aplicada sobre todos os *frames*. Inicialmente, a partir da

amostra de *n frames* são calculados a média μ_b e a matriz de covariância. Os M autovetores mais significativos são armazenados em uma matriz de autovetores ϕ_{Mb} de tamanho $M \times p$, onde p é o número de *pixels* da imagem. Cada novo *frame* I é projetado no autoespaço segundo a Eq. 2.4:

$$I' = \phi_{Mb}(I - \mu_b) \tag{2.4}$$

Para projetar de volta I' no domínio espacial da imagem usa-se a seguinte equação:

$$I'' = \phi_{Mb}{}^{T} I' + \mu_b) \tag{2.5}$$

onde T denota a operação de transposição da matriz.

Os *pixels* pertencentes ao *foreground* são determinados por $|I - I''| > \theta$, onde θ é o limiar selecionado.

2.2 Métodos de Reconhecimento de Padrões

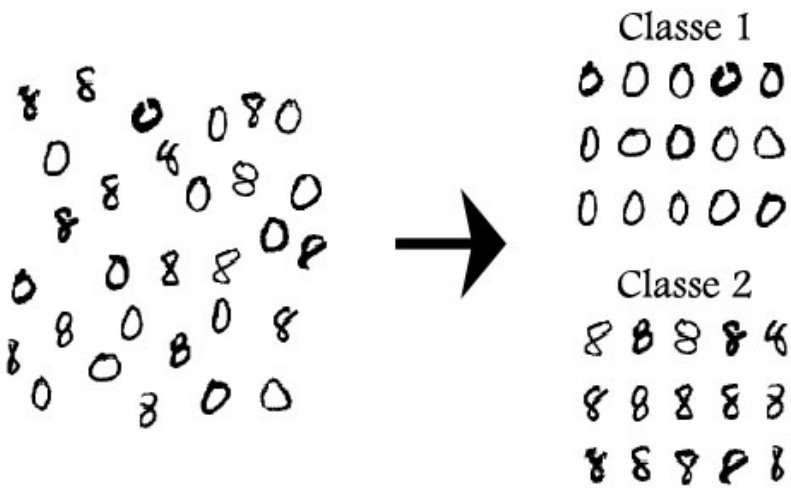
A capacidade humana de reconhecimento de objetos (ou padrões) é um processo aparentemente automático, tendo em vista que conseguimos facilmente distinguir o aroma de uma flor, do aroma de uma cebola; um numeral manuscrito "0", de um numeral manuscrito "8". Diariamente fazemos isso ao reconhecer faces ao nosso redor, porém, é difícil determinar uma sequência de passos para realizar esta tarefa. Cada face de uma pessoa é composta de um padrão composto de estruturas (olhos, nariz, boca, etc), com determinado tamanho e formato, localizadas em certas posições da face, também chamadas de características.

O processo de reconhecimento de padrões é dividido basicamente em duas etapas: a etapa de extração de características deve estar apta para sintetizar padrões específicos que melhor descrevem as categorias (ou classes) de objetos; e a etapa de classificação, que separa os objetos em diferentes classes com base em suas características. Uma classe é uma coleção de objetos semelhantes, mas não necessariamente idênticos, que são distinguíveis dos objetos de outras classes. A Figura 3 ilustra um exemplo de classificação entre duas classes. As seções a seguir descrevem o processo de extração de características e o classificador usado na metodologia desta dissertação.

2.2.1 Extração de Características

Aqui é descrito resumidamente o descritor elíptico de Fourier, introduzido em (KUHL, 1982). O descritor elíptico de Fourier recebe como entrada um contorno fechado, que é primeiramente representado pelo código diferencial da cadeia. Esta representação inicial resulta em uma sequência de vetores de comprimento unitário e codificação direcional como

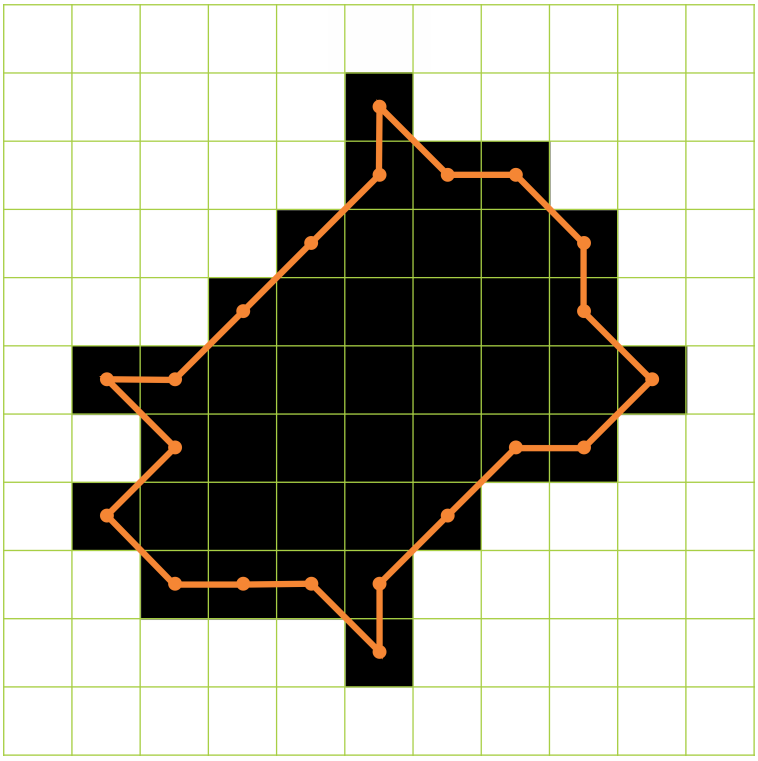
Figura 3 – Exemplo de classificação entre os dígitos manuscritos 0 e 8.



Fonte: Elaborado pelo autor.

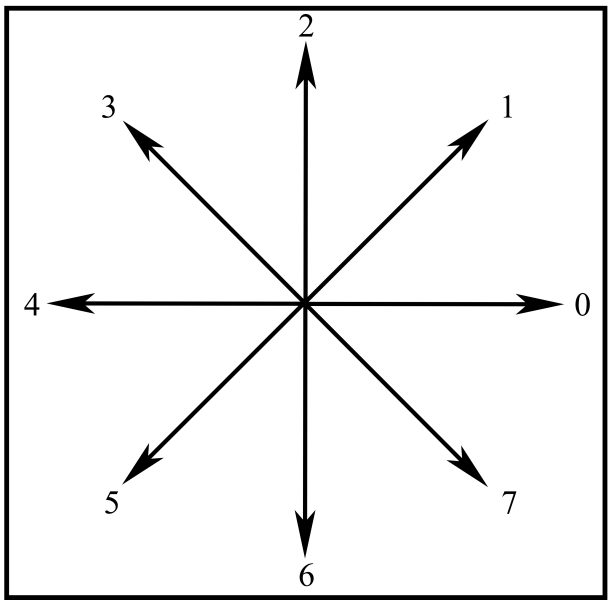
mostrado no exemplo da Figura 5, obtida através do objeto binarizado da Figura 4. *Pixels* de borda individuais são usados para descrever o código da cadeia, começando no *pixel* superior à esquerda do contorno e traçando a borda em sentido horário. A enumeração do código da cadeia é completada quando volta-se ao ponto de partida. Usando o código da cadeia é obtida uma série de Fourier truncada do contorno fechado através das equações:

Figura 4 – Imagem de objeto binarizada com código da cadeia sobrescrito na sua borda.



Fonte: Elaborado pelo autor.

Figura 5 – Código diferencial da cadeia começando no canto superior do objeto da Figura 4.



Código da cadeia:
70767545563443130111 2

Fonte: Elaborado pelo autor.

$$x_N(t) = A_0 + \sum_{n=1}^{N} a_n \cos\left(\frac{2n\pi t}{T}\right) + b_n \sin\left(\frac{2n\pi t}{T}\right) \tag{2.6}$$

$$y_N(t) = C_0 + \sum_{n=1}^{N} c_n \cos\left(\frac{2n\pi t}{T}\right) + d_n \sin\left(\frac{2n\pi t}{T}\right) \tag{2.7}$$

onde t é o passo requerido para cruzar um *pixel* ao longo do contorno fechado, tal que $t_{p-1} < t < t_p$ para valores de código de p dentro do intervalo $1 \leq p \leq k$, onde k é o número total de códigos descrevendo o contorno da borda do objeto. n é o número de harmônicas de Fourier requeridas para gerar a aproximação da borda (cada harmônica tem quatro coeficientes). T é o número de passos necessário para percorrer todo o contorno, $T = t_k$. A_0 e C_0 são os coeficientes de *bias*. Estes coeficientes são relacionados a translação de imagem. N é o número total de harmônicas elípticas de Fourier necessárias para gerar uma aproximação precisa da borda. Em nossos experimentos, atribuímos zero para A_0 and C_0 seguindo a sugestão de (NETO *et al.*, 2006), para fazer com que os coeficientes se tornassem invariantes a translação.

Para cada harmônica, os quatro coeficientes harmônicos a_n, b_n, c_n e d_n são definidos como:

$$a_n = \frac{T}{2n^2\pi^2} \sum_{p=1}^{K} \frac{\Delta x_p}{\Delta t_p} \left[\cos\left(\frac{2n\pi t_p}{T}\right) - \cos\left(\frac{2n\pi t_{p-1}}{T}\right)\right],$$

$$b_n = \frac{T}{2n^2\pi^2}\sum_{p=1}^{K}\frac{\Delta x_p}{\Delta t_p}\left[\sin\left(\frac{2n\pi t_p}{T}\right) - \sin\left(\frac{2n\pi t_{p-1}}{T}\right)\right],$$

$$c_n = \frac{T}{2n^2\pi^2}\sum_{p=1}^{K}\frac{\Delta y_p}{\Delta t_p}\left[\cos\left(\frac{2n\pi t_p}{T}\right) - \cos\left(\frac{2n\pi t_{p-1}}{T}\right)\right],$$

$$d_n = \frac{T}{2n^2\pi^2}\sum_{p=1}^{K}\frac{\Delta y_p}{\Delta t_p}\left[\sin\left(\frac{2n\pi t_p}{T}\right) - \sin\left(\frac{2n\pi t_{p-1}}{T}\right)\right],$$

onde k é o número de passos requeridos para percorrer os primeiros p components do código da cadeia. Δx_p e Δy_p são as mudanças nas projeções x e y do código da cadeia, respectivamente, no componente p. Estes valores podem ser 1, 0 ou -1, dependendo da orientação do componente p. Δt_p é a mudança de passo requerida para atravessar o componente p do código da cadeia. A principal característica deste descritor é que ele é robusto à rotação, translação e resolução de imagem.

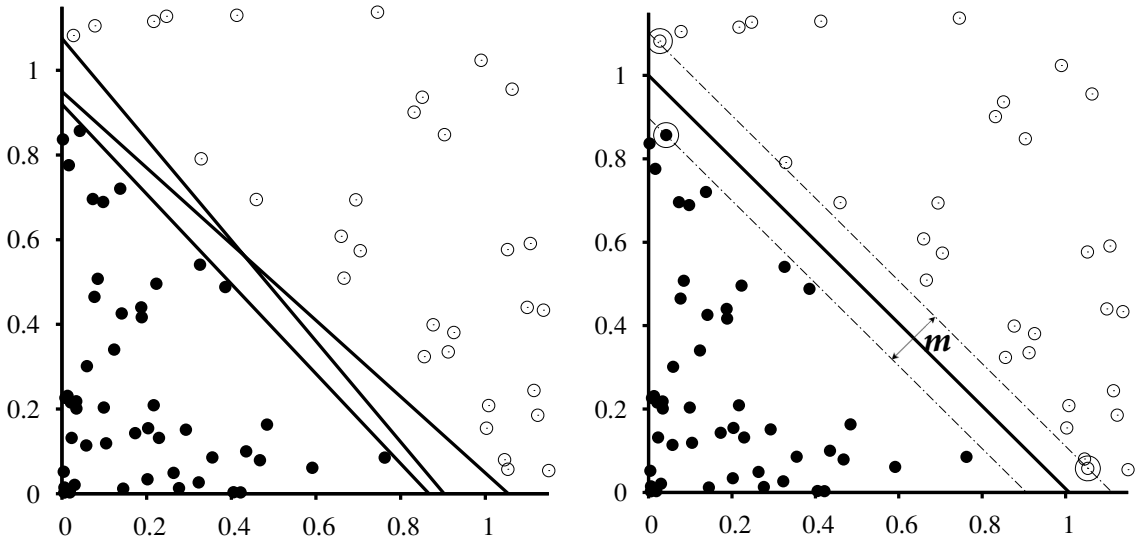
2.2.2 Classificador SVM

O classificador *Support Vector Machine* (SVM), proposto por (CORTES; VAPNIK, 1995), tem sido frequentemente mencionado na literatura por obter bons resultados de classificação em diversos domínios de aplicação (VALERA; VELASTIN, 2011) (MOUNTRAKIS; IM; OGOLE, 2011). A ideia básica é encontrar uma função que separe os dados rotulados, em um hiperespaço d-dimensional. Em problemas de duas classes linearmente separáveis, SVMs determinam um classificador linear f que melhor separa os dados de treinamento. A Figura 6.(a) mostra exemplos de classificadores lineares.

Vários limites podem ser usados em um classificador linear, porém, a escolha do limite pode ser melhor determinada através da maximização da margem m ilustrada na Figura 6.(b), que é a distância entre os valores mais próximos de cada classe, também chamados de vetores de suporte (*support vectors*).

Em casos onde as classes não são linearmente separáveis, é necessário o cálculo de uma função *kernel*, que recebe dois pontos x_i e x_j do espaço de entradas e computa o produto escalar $\Phi(x_i)\Phi(x_j)$ no espaço de características. Para tanto, é necessário escolher uma função de *kernel* adequada, bem como os parâmetros dessa função. A Figura 7 mostra um separador circular em duas dimensões que se torna um separador linear em 3 dimensões.

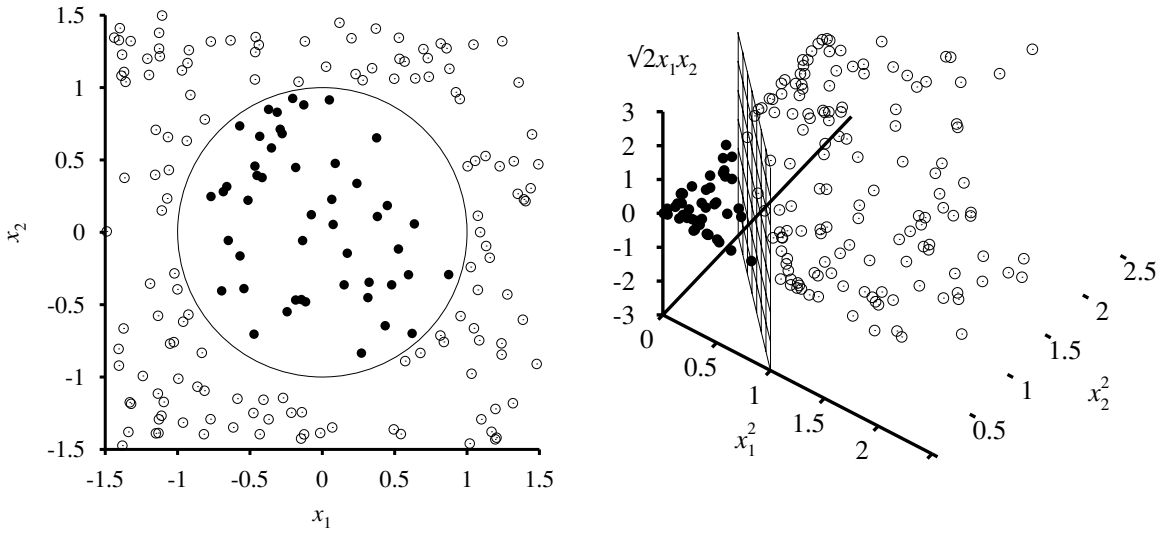
Figura 6 – Classificadores lineares em um plano bidimensional: (a) Duas classes de pontos (pontos pretos e brancos) e três separadores lineares. (b) Um separador linear com maior margem m distante dos *support vectors* (pontos com círculos maiores).



(a) (b)

Fonte: Figuras adaptadas de (RUSSELL; NORVIG, 2009)

Figura 7 – Classificador linear no espaço de características. (a) Separador circular ${x_1}^2 + {x_2}^2 = 1$. (b) Mapeamento de (a) em um espaço tridimensional $({x_1}^2, {x_2}^2, \sqrt{2}x_1x_2)$ com um hiperplano separador.



(a) (b)

Fonte: (RUSSELL; NORVIG, 2009)

Dentre os *kernels* mais utilizados estão a função de base radial, em inglês, *radial basis function* (RBF) e polinômios. O kernel RBF possui o parâmetro de margem C que determina um valor aceitável entre a maximização da margem de separação e a minimização do erro

de classificação, e o parâmetro γ que controla o raio do *kernel*. Já o *kernel* polinomial possui o parâmetro D que determina o grau do polinômio.

Capítulo 3

Metodologia

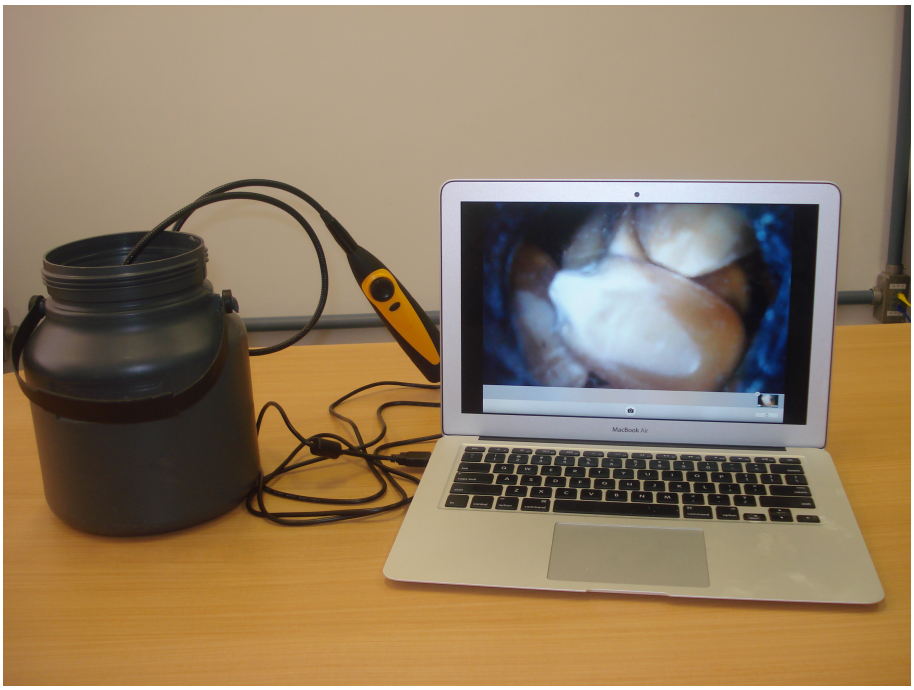
3.1 Contextualização do Problema

A tarefa de identificação e classificação de insetos em milho à granel necessita de constante monitoramento dos grãos no interior dos armazéns que normalmente consistem de grandes silos metálicos (veja Figura 1), obrigando o uso de uma grande quantidade de câmeras endoscópicas para uma cobertura satisfatória da área, que inviabilizaria economicamente o estudo desta dissertação. Contudo, foi construído um ambiente simulado contendo milho à granel e usado uma câmera endoscópica de baixo custo para a construção da base de dados. O ambiente simulado (reservatório de milho contendo as duas espécies de insetos), juntamente com a câmera endoscópica e o computador usado na gravações dos vídeos são ilustrados na Figura 8.

Os primeiros testes utilizando a câmera endoscópica produziram imagens destorcidas conforme pode ser visto na Figura 9(a). Em busca de melhorias na imagem, foram utilizados métodos de *deblurring* (NAYAR; BEN-EZRA, 2004; BAR; KIRYATI; SOCHEN, 2006), que não produziram resultados satisfatórios, já que a causa do fenômeno de borramento era o ajuste da distância focal da câmera, dado pela distância da lente da câmera até a cena (grãos de milho). Como a câmera é introduzida por uma espécie de sonda, os grãos de milho se aproximavam demasiadamente da lente, resultando em uma distância focal menor que a requerida pelo equipamento e, consequentemente no barramento das imagens. Para solucionar o problema, foi realizado um ajuste mecânico para distanciar a lente da câmera dos grãos, conforme pode ser visto na Figura 10. Tal ajuste consiste basicamente na introdução de um pequeno cano na extremidade do endoscópio. Esta solução obteve bons resultados, porém, provocou perdas nas bordas do plano da imagem, o que pode ser visualizado pela tonalidade escura nas extremidades das imagens, conforme ilustrado na Figura 9(b).

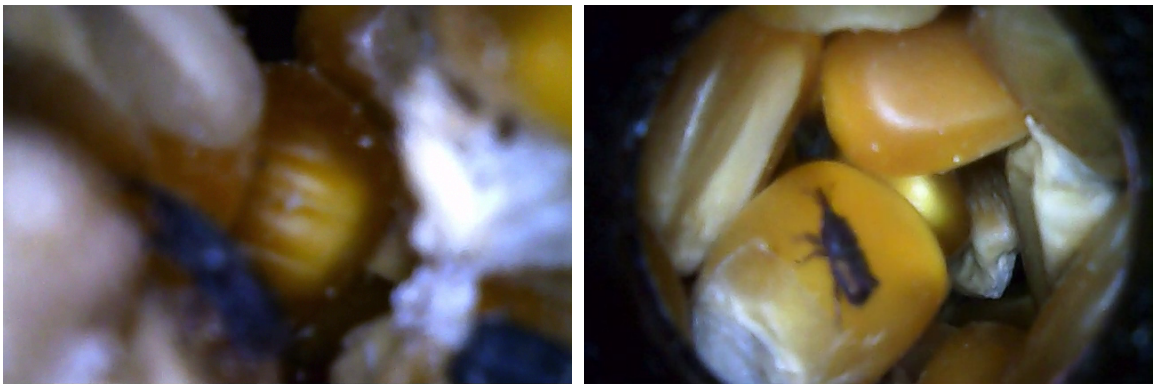
A partir desta configuração, foi construída uma base de dados composta de vídeos de curta duração contendo algum inseto se movimentando. Cada vídeo foi submetido a segmentação de objetos em movimento por subtração de plano de fundo, e manualmente,

Figura 8 – Ambiente simulado para a construção da base de dados.



Fonte: Elaborado pelo autor.

Figura 9 – Imagens da câmera endoscópica: (a) Sem ajuste mecânico; (b) Com ajuste mecânico.



(a) (b)

Fonte: Elaborado pelo autor.

foram selecionados *frames*, onde os insetos aparecem completamente (por inteiro), para a extração de características e classificação. A seguir são apresentados o método de subtração de plano de fundo proposto e o delineamento dos experimentos de classificação.

3.2 Método proposto

Para a classificação dos insetos, optou-se por diferenciá-los com base em seu formato, já que as espécies presentes na base de dados apresentam diferenças principalmente nas

Figura 10 – (a) Câmera endoscópica; (b) Ajuste mecânico.

(a) (b)

Fonte: Elaborado pelo autor.

patas e cabeça. Em busca de melhores resultados na etapa de extração de características do inseto, foi desenvolvido um método de subtração de plano de fundo que detalha mais precisamente o contorno do inseto, principalmente os detalhes de partes pequenas como as patas e antenas. A seguir são descritas as três etapas que descrevem o método proposto.

3.2.1 Inicialização de plano de fundo

Uma etapa importante em todo método de subtração de plano de fundo é a inicialização do modelo de plano de fundo, também conhecida como *bootstrapping*. Esta etapa deixa de ser trivial quando não se tem uma sequência inicial de *frames* sem movimento de objetos para definir o plano de fundo. O *bootstrapping* deve consumir poucos *frames*, de forma a não perder a detecção de objetos no vídeo. Além disso, tal modelo deve ser inicializado da forma mais precisa possível pois um *bootstrapping* inadequado reflete nos processos de detecção de objetos e de atualização do plano de fundo, consecutivos. Em resumo, o modelo de plano de fundo deve ser criado consumindo um número mínimo de *frames*, da forma mais precisa possível. O *bootstrapping* utilizado neste artigo é o mesmo de (CALDERARA *et al.*, 2006), que particiona a imagem em blocos de 16×16 *pixels* e somente adiciona um dado bloco para o modelo de plano de fundo se não houver evidencia de objetos em movimento naquele bloco, isto é, uma alta fração dos pixels não mudarem de valores entre frames consecutivos. Se mais de 90% dos *pixels* no bloco, têm no *frame* corrente, valores muito próximos daqueles do *frame* anterior, considera-se que não há movimento de objeto(s) no bloco. Se acontecer de um bloco ser considerado estático (sem movimento de objetos) por mais de cinco vezes consecutivas, o bloco é setado como plano de fundo e não é mais analisado. O processo continua até que todos os blocos sejam setados como plano de fundo, ou até o ponto em que o final do vídeo é atingido.

3.2.2 Atualização de plano de fundo

Em nosso cenário de aplicação, os insetos estão em constante movimento e, frequentemente, eles deslocam peles de milho, as quais devem ser, posteriormente, agregadas ao plano de fundo. Assim, a atualização do plano de fundo se torna um passo essencial. Esta é a motivação para o uso de um parâmetro de taxa de aprendizagem para atualização do plano de fundo. Depois do passo de inicialização (*bootstrapping*), o plano de fundo é atualizado a cada *frame* no instante t no vídeo. A cada novo *frame* $I^{(t)}$ atualiza-se o plano de fundo $BG^{(t)}$ conforme a Eq. 3.1.

$$BG^{(t)} = (\alpha * I^{(t)}) + ((1-\alpha) * BG^{(t-1)}) \tag{3.1}$$

onde α é a taxa de aprendizagem que determina o quão rápido o plano de fundo absorve as mudanças na cena. Nos experimentos deste trabalho, como não há mudanças significativas de iluminação e os insetos movem rapidamente, foi considerado um valor de α relativamente baixo, a saber 0.01. Contudo, é esperado que uma atualização do parâmetro α em função da velocidade dos objetos produza melhores resultados. Este método de atualização de plano de fundo foi publicado em (GEUS *et al.*, 2015a).

3.2.3 Extração e binarização de objeto

O algoritmo proposto calcula inicialmente uma máscara pela diferença entre o plano de fundo e o *frame* corrente. Este processo é feito *frame* a *frame*. A diferença entre o *frame* corrente e o plano de fundo é dada por:

$$M^{(t)}(i,j) = \frac{(I^{(t)}(i,j) - BG^{(t)}(i,j)) * i^T}{3} \tag{3.2}$$

onde i^T é o vetor identidade, $I^{(t)}(i,j)$ e $BG^{(t)}(i,j)$ são vetores de componentes de cores (R,G,B), relativos ao *pixel* de coordenadas (i,j). Assim, $M^{(t)}$ é uma matriz bidimensional no instante de tempo t, que contem valores próximos de zero para regiões da imagem onde não há evidência de movimento. A Figura 11.(c) mostra um exemplo de tal máscara, gerada a partir da subtração do plano de fundo na Figura 11(a) do *frame* atual na Figura 11(b).

Então, a máscara é binarizada utilizando dois diferentes limiares (*thresholds*): um limiar global definido empiricamente, para identificar *pixels* que ocorrem em alta variação de intensidade; e um limiar local para identificar *pixels* pertencentes ao objeto com uma pequena variação de intensidade comparada ao plano de fundo local. O limiar global é representado por T_{global} e o limiar local empregado é o introduzido em (SAUVOLA; PIETIKAINEN, 2000), portanto chamado de $T_{sauvola}(i,j)$ devido ao nome do primeiro autor.

O cálculo da máscara binária usando os limiares global e local são dadas pelas Eqs. 3.3 e 3.4, respectivamente. Para calcular o limiar local, foi utilizada uma janela de tamanho 55 ×

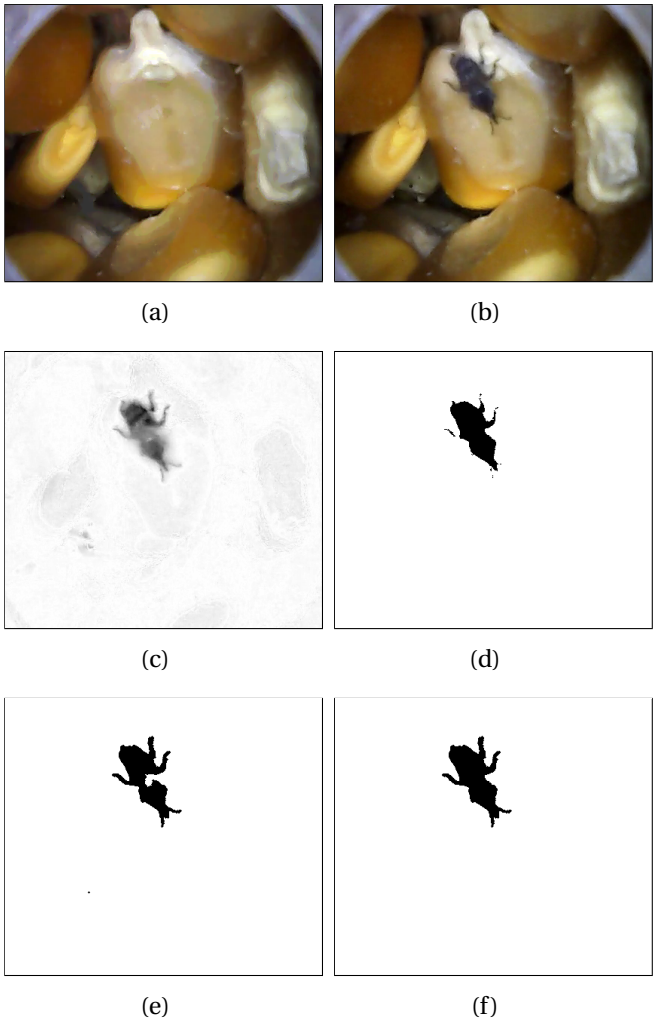


Figura 11 – (a) Plano de fundo; (b) Imagem *Frame* atual; (c) máscara; (d) Imagem binarizada do *threshold* global; (e) Imagem binarizada do *threshold* local; (f) Imagem final binarizada.

55 *pixels*. Resultados preliminares utilizando apenas o *threshold* global foram publicados em (GEUS *et al.*, 2015a) e (GEUS *et al.*, 2015b).

$$B^{(t)}_{global}(i,j) = \begin{cases} 1, & \text{se } M^{(t)}(i,j) > T_{global} \\ 0, & \text{caso contrário} \end{cases} \tag{3.3}$$

$$B^{(t)}_{local}(i,j) = \begin{cases} 1, & \text{se } M^{(t)}(i,j) > T^{(t)}_{sauvola}(i,j) \\ 0, & \text{caso contrário} \end{cases} \tag{3.4}$$

A máscara binarizada final é calculada pela composição das duas máscaras binarizadas B_{global} and B_{local}. A operação de composição é dada pela Eq. 3.5. A idéia da Eq. 3.5 é pegar *pixels* com alta variação de intensidade em relação ao plano de fundo, que tem alta probabilidade de pertencer a um objeto em movimento, e agregá-los com regiões espacialmente conectadas ao objeto, que possuem uma variação local de intensidade significativa

em relação ao pano de fundo.

$$B^{(t)}_{final} = B^{(t)}_{global} \cup \{R \in B^{(t)}_{local}\} \tag{3.5}$$

tal que R são regiões 8-conectadas à $B^{(t)}_{global}$.

A Figura 11 ilustra o processo de binarização. As Figuras 11(d) e (e), respectivamente, mostram os resultados dos *thresholds* local e global considerando a máscara da Figura 11(c). A Figura 11(f) mostra o resultado composto, onde é claramente perceptível que os detalhes dos objetos obtidos no *threshold* local são adicionados à segmentação obtida pelo *threshold* global.

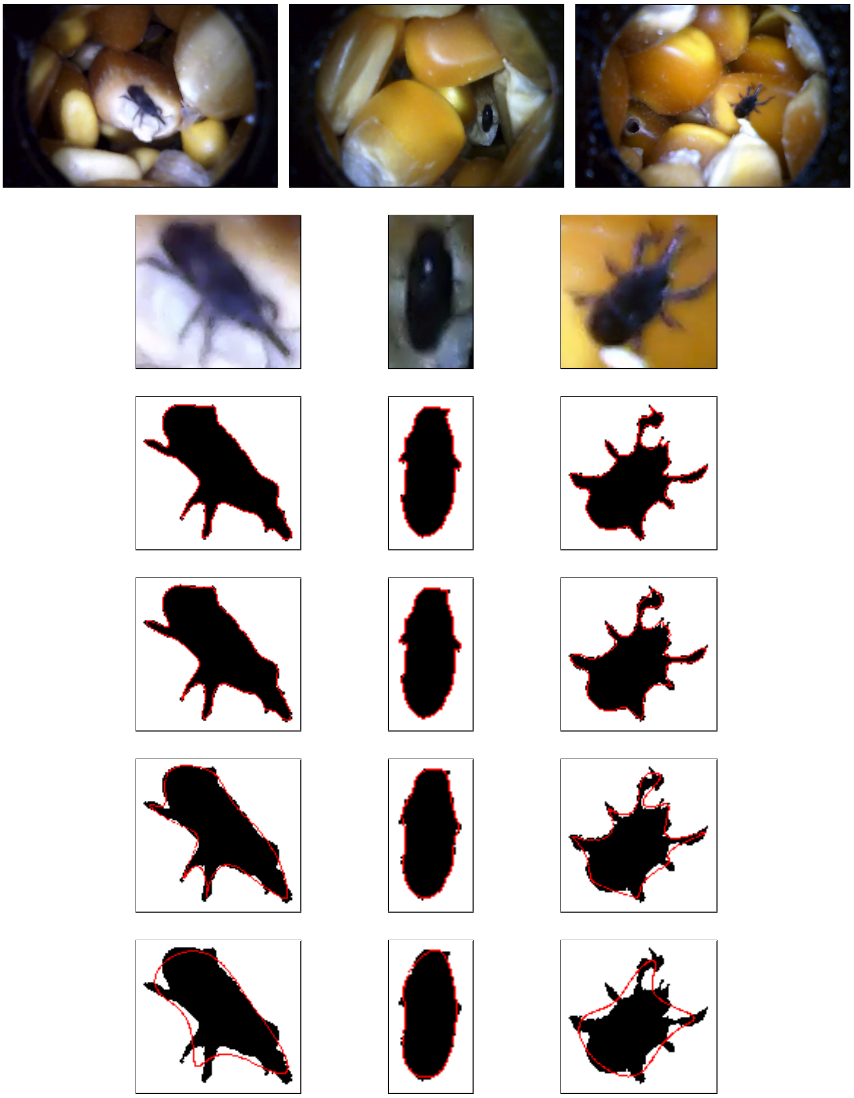
3.3 Classificação

Para estimar a qualidade do método de segmentação proposto nessa dissertação, forem realizados experimentos de classificação comparando-o com os melhores métodos de segmentação baseados em subtração de plano de fundo, listados por Sobral e Vacavant (2014). Não foram realizados experimentos específicos de análise da acurácia de segmentação, pois foi utilizada uma câmera de baixa resolução para a construção da base de dados, o que dificulta a determinação do *ground-thuth* que delimita os insetos.

Após segmentação dos insetos em movimento através dos método de subtração de plano de fundo, foi utilizado o descritor elíptico de Fourier (descrito na seção 2.2.1) visando a diferenciação dos insetos com base em seu formato. Em busca da determinação do número mínimo de harmônicas suficientes para representar o objeto segmentado, foi feita a reconstrução do contorno dos objetos através de suas características extraídas. A figura 12 mostra a aproximação do formato dos insetos usando o conjunto de coeficientes a_n, b_n, c_n e d_n, gerados a partir de 30, 20, 10 e 5 harmônicas. Pode ser notado que 20 harmônicas são suficientes para a representar uma aproximação no contorno do objeto. No experimento foram calculados os coeficientes a_n, b_n, c_n e d_n a partir de 30 harmônicas, produzindo 120 coeficientes que representam cada contorno.

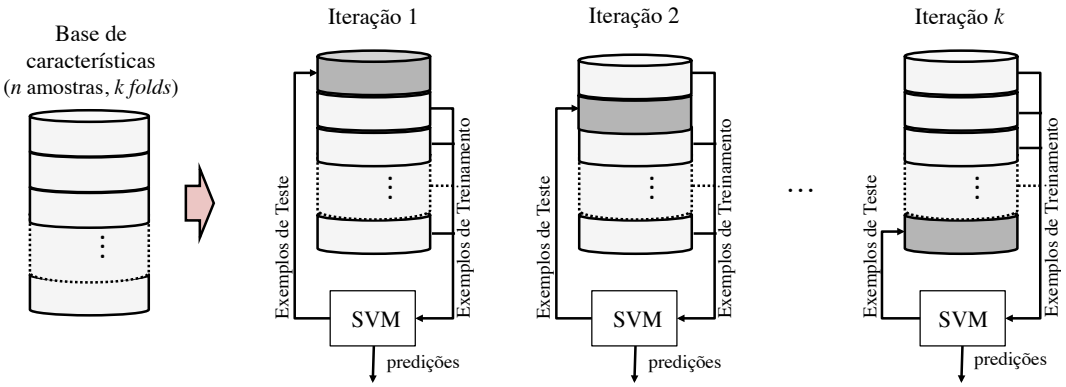
A classificação foi realizada através do método SVM (descrito na seção 2.2.2), utilizando validação cruzada *k-fold*. Na validação cruzada *k-fold*, os dados de características extraídos através do descritor de Fourier são particionados igualmente em k subconjuntos de tamanho aproximadamente igual, também chamados *folds*. Dos k *folds*, apenas um deles é usado para validação, e os $k-1$ *folds* restantes são usados para treinamento. O processo de validação cruzada é repetido por k vezes, onde cada um dos k *folds* são usados apenas uma vez como dado de validação, conforme ilustrado na Figura 13. A vantagem deste método sobre uma simples divisão entre treinamento e teste é que todos os subconjuntos são usados tanto para treinamento quanto para validação.

Figura 12 – Reconstrução do contorno usando os coeficientes do descritor elíptico de Fourier calculados por diferentes números de harmônicas. Primeira linha – imagens originais; Segunda linha – recorte da região contendo inseto; Terceira até sexta linha – imagem binarizada do inseto com aproximação do contorno em vermelho usando 30, 20, 10 and 5 harmônicas, respectivamente.



Fonte: Elaborado pelo autor.

Figura 13 – *Design* dos experimentos de classificação usando validação cruzada *k-fold* e o classificador SVM.



Fonte: Elaborado pelo autor.

Capítulo 4

Resultados

A base de dados utilizada consiste em vídeos de curta duração, gravados no interior de um recipiente de milho à granel, usando uma câmera endoscópica de resolução de 640×426 *pixels*. O uso de uma câmera de baixa resolução é motivado pelo seu baixo custo, visando uma aplicação real em grandes silos de metais, onde seria necessário o uso de várias câmeras endoscópicas para um monitoramento eficaz de insetos.

Para comparar o método proposto, foram tomados os métodos de subtração de plano de fundo da literatura que produzem os melhores resultados, conforme a revisão de Sobral e Vacavant (2014). Neste estudo usamos a biblioteca computacional *BGS Library*, disponibilizada pelos autores citados anteriormente, em https://github.com/andrewssobral/bgslibrary. A Tabela 3 lista os métodos da *BGS Library* usados para comparação, suas referências, e os parâmetros usados para cada um deles. Os parâmetros usados em cada método são os mesmos usados em (SOBRAL; VACAVANT, 2014). Para detalhes específicos de cada método, o leitor pode consultar (SOBRAL; VACAVANT, 2014) e as referências de cada método listadas na Tabela 3. As Figuras 14 e 15 mostram os resultados de segmentação obtidos pelo método de subtração de plano de fundo proposto e pelos métodos selecionados do estado da arte da literatura, em três *frames*. Na Figura 14 são ilustrados: os *frames* originais – na primeira linha; os resultados obtidos pelo método proposto – na segunda linha; os resultados obtidos pelo método de mistura de gaussianas (*Gaussian Mixture Model*) – terceira linha; os resultados obtidos por redes neurais do tipo Mapas Auto-organizáveis Adaptativos (*Adaptive SOM*) – quarta linha; e por fim, na última linha, os resultados do método estatístico *Multi-Layer BGS*. A Figura 15 é uma continuação da Figura 14, que mostra: os *frames* originais – primeira linha; resultados do método baseado em autovalores e autovetores (*Eigenbackground/ SL-PCA)* – na segunda linha; resultados do método básico *Adaptive Background Learning* – na terceira linha; resultados do método baseado em conjuntos nebulosos (*Fuzzy Choquet Integral*) – na quarta linha; e finalmente, resultados do método baseado em *Type-2 Fuzzy* (*Type-2 Fuzzy GMM-UM*) – na quinta linha. Pode-se observar que o método proposto demarca mais precisamente a fronteira dos insetos, sem captar regiões desconexas (ruído) em

torno do inseto. Vale destacar que todos os demais métodos experimentados captam regiões desconexas (ruídos). O fato do método proposto não captar regiões desconexas é explicável pela estratégia de acoplar somente regiões conectadas da limiarização local ao resultado da limiarização global, a qual foi apresentada e discutida no Capítulo anterior. Analisando visualmente os resultados da Figura 14 e 15, constatamos que os métodos de melhores resultados são: o método proposto, *Gaussian Mixture Model* (KAEWTRAKULPONG; BOWDEN, 2002), *Adaptive SOM* (MADDALENA; PETROSINO, 2008) e *Multi-Layer BGS* (YAO; ODOBEZ, 2007), respectivamente.

Para realização da classificação através do SVM, foram selecionados manualmente 602 exemplos de *frames* contendo carunchos e 606 contendo besouros. Cada *frame* selecionado contém somente um besouro ou somente um caruncho. Afim de eliminar pequenas regiões desconexas resultantes na segmentação de alguns métodos, o descritor de Fourier foi aplicado apenas ao maior objeto da máscara binarizada resultante da subtração de plano de fundo. Em termos da precisão de classificação, comparamos o método proposto somente com os métodos *Gaussian Mixture Model* (KAEWTRAKULPONG; BOWDEN, 2002), *Adaptive SOM* (MADDALENA; PETROSINO, 2008) e *Multi-Layer BGS* (YAO; ODOBEZ, 2007), visto que os demais métodos de subtração de plano de fundo não produzem resultados de segmentação satisfatórios, conforme análise visual. Tal fato pode ser notado pelos resultados de segmentação da Figura 14. Os *kernels* selecionados foram Polinomial e RBF com parâmetros determinados por *GridSearch*. *GridSeach* consiste em uma busca exaustiva pelos melhores parâmetros dentro de intervalos especificados para a busca. As taxas de classificação corretas utilizando o método de validação cruzada *k-fold*, com $k = 10$, estão contidas na Tabela 1, e mostram que o método de segmentação proposto apresenta maior taxa de acerto de classificação perante os métodos do estado da arte. Observa-se que a taxa de classificação foi bastante satisfatória para a base construída, a qual continha poucas amostras e estas foram selecionadas com cautela, de modo que os insetos presentes nos frames aparecem por completo e estão próximos ao centro do plano da imagem. Fazendo um comparativo dos resultados de classificação com a análise visual dos resutados de segmentação (Figuras 14 e 15), pode-se notar que o método de segmentação proposto, no caso da base elaborada, confirma sua superioridade em relação aos demais. O método de Mistura de Gaussianas apresenta precisão de classificação condizente com o resultado da análise visual, além de estar em conformidade com estudos da literatura, que o reportam como um dos principais métodos de subtração de plano de fundo. Já o resultado de *Adaptive SOM* comparado ao de *Multi-Layer BGS* é mais difícil de ser explicado; visualmente os resultados de segmentação são ambos imprecisos. Contudo, pode ser notado é que *Adaptive SOM* produz segmentação com bordas bastante irregulares, o que poderia prejudicar a caracterização do objeto pelo descritor Elíptico de Fourier, e consequentemente, produzir resultados de classificação inferiores.

Na Tabela 2 são mostradas as matrizes de confusão para os experimentos de classifi-

cação. Conforme análise das matrizes de confusão pode-se notar que basicamente todos os métodos cometem erros, classificando carunchos como besouros. Isto é justificável pelo fato de que, quando as patas do caruncho estão oclusas, seu formato se parece com o formato do besouro, e pelo fato descritor usado ser robusto à escala dos objetos. Normalmente, o caruncho é maior que o besouro. Contudo, na prática a consideração do tamanho do objeto segmentado como caraterística é inviável pois o tamanho do objeto captado pela segmentação depende da proximidade do inseto à lente da câmera.

Método	SVM Polinomial	SVM RBF
Método Proposto	97.10%	91.63%
Gaussian Mixture Model (MixtureOfGaussianV1BGS)	91.72%	88.32%
Adaptive SOM (LBAdaptiveSOM)	75.16%	75.16%
Multi-Layer BGS (MultiLayerBGS)	90.96%	87.66%

Tabela 1 – Taxa de acerto na classificação.

Proposed	A	B	Acc.
A	594	12	98.02%
B	23	579	96.18%

(a)

Proposed	A	B	Acc.
A	593	13	97.85%
B	88	514	85.38%

(b)

MOG	A	B	Acc.
A	593	13	97.85%
B	87	515	85.55%

(c)

MOG	A	B	Acc.
A	591	15	97.52%
B	126	476	79.07%

(d)

LBA	A	B	Acc.
A	591	15	97.52%
B	285	317	52.66%

(e)

LBA	A	B	Acc.
A	591	15	97.52%
B	285	317	52.66%

(f)

M.Layer	A	B	Acc.
A	593	13	97.85%
B	107	495	82.23%

(g)

M.Layer	A	B	Acc.
A	591	15	97.52%
B	134	468	77.74%

(h)

Tabela 2 – Matriz de confusão dos experimentos de classificação. O rótulo A corresponde à classe Tribolium castaneum (besouro), o rótulo B corresponde à classe Sitophilus zeamais e Acc. sumariza a acurácia por classe. As tabelas da esquerda correspondem ao uso de SVM com kernel polinomial e nas tabelas da direita foi utilizado SVM com kernel RBF.

Tabela 3 – Métodos de subtração de plano de fundo e seus parâmetros utilizados. Os parâmetros utilizados são os mesmos de Sobral e Vacavant (2014).

ID do método	Nome do método	Referências	Parâmetros
Método básico: média e variância ao longo do tempo			
AdaptiveBackgroundLearning	*Adaptive Background Learning*	(SOBRAL; VACAVANT, 2014)	$T = 15$, $\alpha = 0.5$
Método baseado em Fuzzy			
FuzzyChoquetIntegral	Fuzzy Choquet Integral	(BAF; BOUWMANS; VACHON, 2008a)	$T = 0.67$, $LF = 10$, $\alpha_{learn} = 0.5$, $\alpha_{update} = 0.05$, $RGB + LBP$
Método estatístico usando uma Gaussiana			
DPWrenGABGS	*Gaussian Average*	(WREN *et al.*, 1997)	$T = 12.15$, $LF = 30$, $\alpha = 0.05$
Método estatístico usando múltiplas Gaussianas			
MixtureOfGaussianV1BGS	*Gaussian Mixture Model*	(KAEWTRAKULPONG; BOWDEN, 2002)	$T = 10$, $\alpha = 0.01$
Métodos baseados em Type-2 Fuzzy			
T2FGMM_UM	*Type-2 Fuzzy GMM-UM*	(BAF; BOUWMANS; VACHON, 2008b) (BAF; BOUWMANS; VACHON, 2009) (BOUWMANS T. ANF EL BAF, 2010)	$T = 1$, $k_m = 2.5$, $n = 3$, $\alpha = 0.01$
Métodos estatísticos usando características de cor e textura			
MultiLayerBGS	Multi-Layer BGS	(YAO; ODOBEZ, 2007)	Parâmetros originais padrão de (YAO; ODOBEZ, 2007)
Métodos baseados em autovalores e autovetores			
DPEigenbackgroundBGS	Eigenbackground/ SL-PCA	(OLIVER; ROSARIO; PENTLAND, 2000)	$T = 255$, $HS = 10$, $ED = 10$
Métodos Neurais			
LBAdaptiveSOM	Adaptive SOM	(MADDALENA; PETROSINO, 2008)	$LR = 180$, $LR_{training} = 255$, $\sigma = 100$, $\sigma_{training} = 240$, $TS = 40$

Figura 14 – Imagens originais na primeira linha e resultados de segmentação obtidos pelos métodos analisados: segunda linha – método proposto; terceira linha – MixtureOfGaussianV1BGS; quarta linha – LBAdaptiveSOM; quinta linha – MultiLayerBGS; sexta linha – DPWrenGABGS. Os parâmetros usados em cada método são dados na Tabela 3.

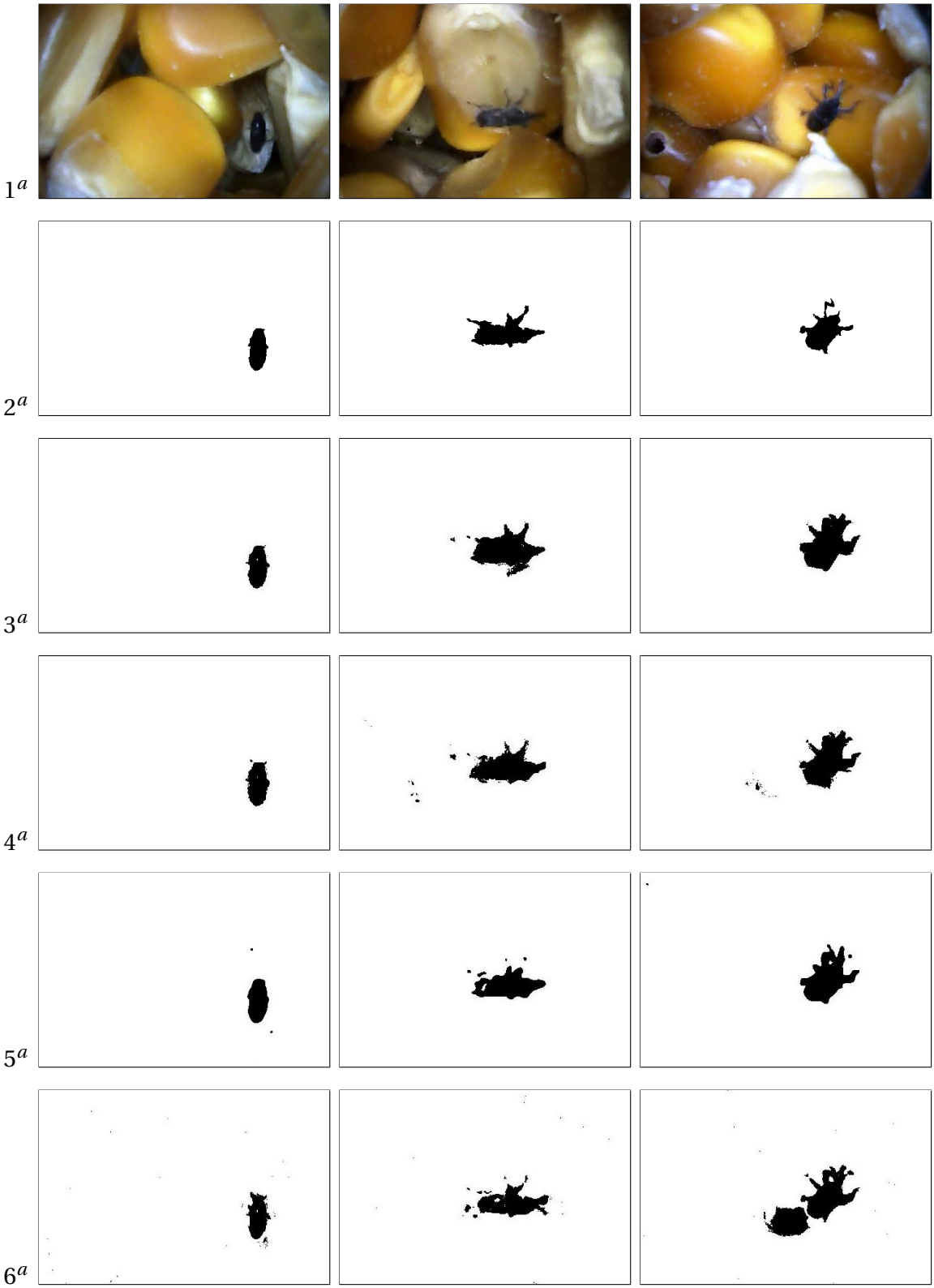
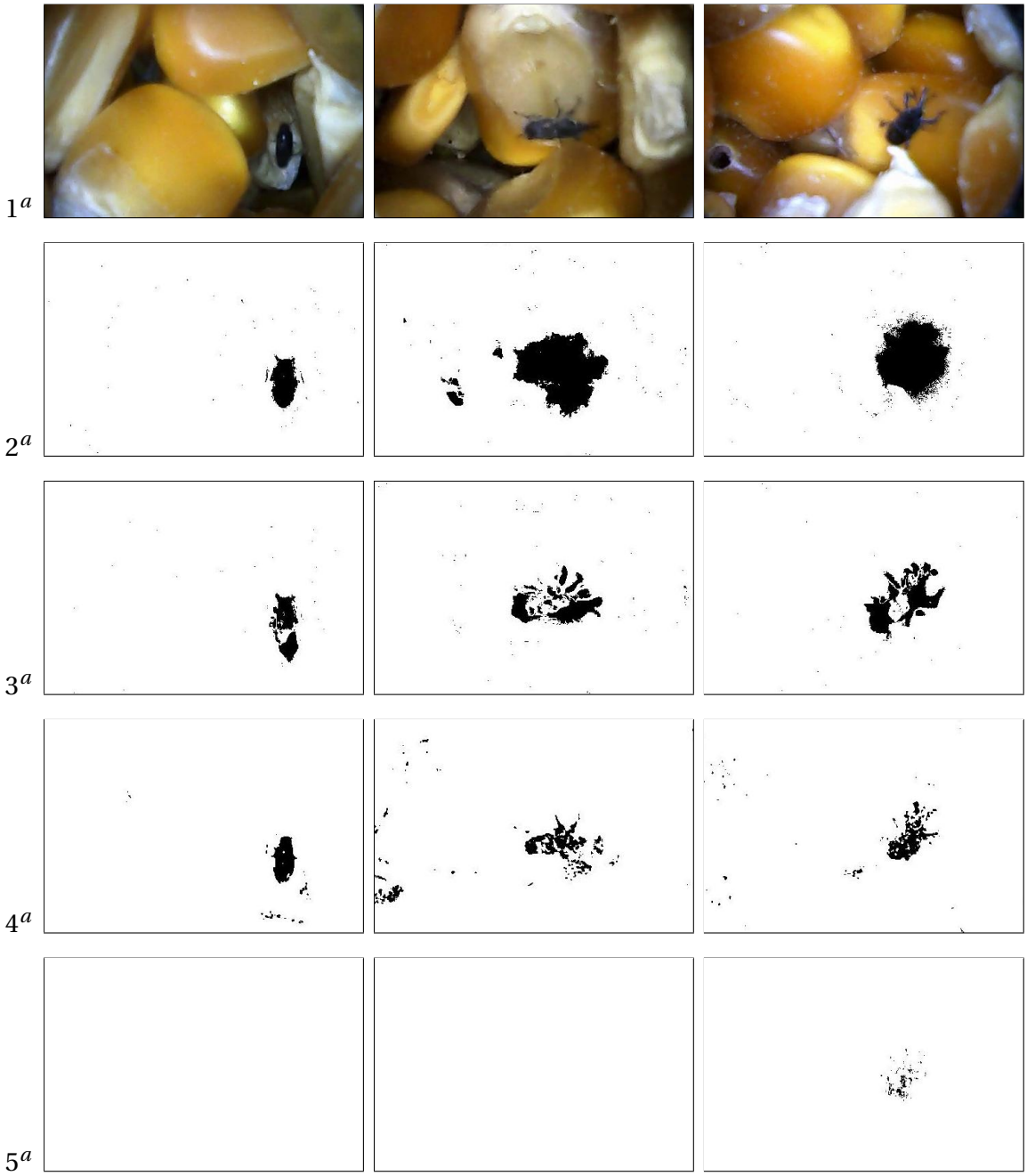


Figura 15 – Continuação da Figura 14 - Imagens originais na primeira linha e resultados de segmentação obtidos pelos métodos analisados: segunda linha – DPEigenbackgroundBGS; terceira linha – AdaptiveBackgroundLearning; quarta linha – FuzzyChoquetIntegral; quinta linha – T2FGMM_UM. Os parâmetros usados em cada método são dados na Tabela 3.



Conclusão

Como conhecido de muitas pesquisas da literatura, segmentação de objetos tem um papel crucial em sistemas de visão computacional voltados para a classificação de objetos. Neste estudo é apresentado um método de segmentação de objeto baseado na abordagem de subtração de plano de fundo. Resultados experimentais usando uma base de dados de vídeos contendo insetos mostram que o método proposto produz resultados de segmentação e de classificação de insetos, para o problema proposto, superiores aos de outros métodos de segmentação por subtração de plano de fundo do estado da arte da literatura. Foi confirmada também a superioridade do método de subtração de plano de fundo em termos de classificação, onde nossos resultados foram 5.3% mais precisos. Como segmentação tem um papel fundamental em várias aplicações de visão computacional, nossos resultados indicam que o método de segmentação proposto para ser usado com sucesso em outros domínios de aplicações, tais como em sistemas de vigilância. Adicionalmente, conforme é de conhecimentos dos autores, este é o primeiro estudo a comparar métodos de subtração de plano de fundo para a tarefa de classificação de insetos em grãos à granel. Em uma investigação futura planeja-se experimentar o método de segmentação proposto para a tarefa de identificação de pessoas em vídeos de sistemas de vigilância. Além desta investigação pretende-se:

- elaborar novas bases de dados de insetos, maiores e mais complexas, contendo por exemplo, oclusões parciais dos insetos nos *frames* selecionados para compor a base;

- Experimentar técnicas de fluxo óptico (*optical flow*) para ajustar dinamicamente o parâmetro α da Equação 3.1. A idéia inicial é estimar a velocidade que os insetos se movem, por meio de *optical flow*, para em seguida elaborar uma fórmula para o ajuste dinâmico do parâmetro α;

- Experimentar novos extratores de características, uma vez que neste trabalho foi usado apenas o descritor Elíptico de Fourier;

- Pensando em uma composição das características geradas por vários extratores, torna-se interessante o uso de técnicas de redução de dimensionalidade dos dados;

- E por fim, apesar de SVM produzir ótimos resultados de classificação para nossa base, seria interessante fazer um estudo comparativo do desempenho de vários classificadores da literatura atual, incluindo *deep learning*, redes complexas, dentre outras técnicas de classificação.

REFERÊNCIAS

BAF, F. E.; BOUWMANS, T.; VACHON, B. Fuzzy integral for moving object detection. In: *IEEE International Conference on Fuzzy Systems*. Hong Kong: [s.n.], 2008. p. 1729–1736. Citado na página 54.

____. Type-2 fuzzy mixture of gaussians model: Application to background modeling. In: *International Symposium on Advances in Visual Computing*. Las Vegas, NV: [s.n.], 2008. p. 772–781. Citado na página 54.

____. Fuzzy statistical modeling of dynamic backgrounds for moving object detection in infrared videos. In: *IEEE Conference on Computer Vision and Pattern Recognition*. Miami, FL: [s.n.], 2009. p. 60–65. Citado na página 54.

BAR, L.; KIRYATI, N.; SOCHEN, N. Image deblurring in the presence of impulsive noise. *International Journal of Computer Vision*, v. 70, n. 3, p. 279–298, 2006. Citado na página 43.

BARBERATO, C. Sem insetos grãos têm maior valor. *Folha do Paraná*, p. online: http://www.folhaweb.com.br/?id_folha=2–1–114&dt=20010224, 2001. Citado na página 31.

BOUWMANS, T. Traditional and recent approaches in background modeling for foreground detection: An overview. *Computer Science Review*, v. 11, n. 12, p. 31–66, 2014. Citado na página 33.

BOUWMANS T. ANF EL BAF, F. Modeling of dynamic backgrounds by type-2 fuzzy gaussians mixture models. *MASAUM Journal of Basic and Applied Sciences*, v. 1, n. 2, p. 265–276, 2010. Citado na página 54.

CALDERARA, S. *et al.* Reliable background suppression for complex scenes. In: *Proceedings of the 4th ACM International Workshop on Video Surveillance and Sensor Networks*. New York, NY, USA: ACM, 2006. (VSSN '06), p. 211–214. ISBN 1-59593-496-0. Disponível em: <http://doi.acm.org/10.1145/1178782.1178814>. Citado na página 45.

CELIK, T. *et al.* Fire detection using statistical color model in video sequences. *Journal of Visual Communication and Image Representation*, v. 18, n. 2, p. 176 – 185, 2007. Citado na página 33.

CHANCHANACHITKUL, W. *et al.* Local foreground extraction technique for rat walking behavior classification. In: *7th International Conference on Knowledge and Smart Technology (KST)*. Chonburi, Thailand: [s.n.], 2015. p. 54–59. Citado na página 33.

CHEN, Z.; ELLIS, T.; VELASTIN, S. A. Vehicle detection, tracking and classification in urban traffic. In: *15th International IEEE Conference on Intelligent Transportation Systems*. Anchorage, Alaska, USA: [s.n.], 2012. p. 16–19. Citado na página 33.

CORTES, C.; VAPNIK, V. Support-vector networks. *Machine Learning*, v. 20, n. 3, p. 273–297, 1995. Citado na página 40.

CULIBRK, D. *et al.* Neural network approach to background modeling for video object segmentation. *Neural Networks, IEEE Transactions on*, v. 18, n. 6, p. 1614–1627, Nov 2007. ISSN 1045-9227. Citado na página 36.

DOMENICO, A. S. D. *et al.* Análise de trilha da contaminação por aflatoxinas em grãos de milho armazenados. *Pesquisa Agropecuária Brasileira*, v. 50, p. 441 – 449, 06 2015. Citado na página 32.

ELIAS, M. *et al.* Pragas e microorganismos no armazenamento de grãos e derivados. In: ELIAS, M. C. (Ed.). *Manejo tecnológico da secagem e do armazenamento de grãos*. [S.l.]: Ed. Santa Cruz. Pelotas, 2008. Citado na página 31.

FORTUN, D.; BOUTHEMY, P.; KERVRANN, C. Optical flow modeling and computation: A survey. *Computer Vision and Image Understanding*, v. 134, p. 1 – 21, 2015. Citado na página 33.

GEUS, A. R. *et al.* Insects detection in maize by endoscopic video analysis. In: *17th International Conference on Artificial Intelligence*. Las Vegas, NV, USA: [s.n.], 2015. p. 510–515. Citado 2 vezes nas páginas 46 e 47.

____. Segmentacao e classificacao de insetos em milho a granel por meio da analise de video. In: *Simposio Brasileiro de Automacao Inteligente*. Natal, RN, BRA: [s.n.], 2015. p. 1–6. Citado na página 47.

HAYMAN, E.; EKLUNDH, J.-O. Statistical background subtraction for a mobile observer. In: *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. [S.l.: s.n.], 2003. p. 67–74 vol.1. Citado na página 36.

KAEWTRAKULPONG, P.; BOWDEN, R. An improved adaptive background mixture model for real-time tracking with shadow detection. In: *European Workshop on Advanced Video Based Surveillance Systems*. London, UK: [s.n.], 2002. p. 135–144. Citado 2 vezes nas páginas 52 e 54.

KUHL, F. P. Elliptic fourier features of a closed contour. *Computer Graphics and Image Processing*, v. 18, p. 1982, 1982. Citado 2 vezes nas páginas 33 e 37.

LOECK, A. E. *Pragas de produtos armazenados*. [S.l.]: Pelotas: EGUFPEL, 2002. 113p. p. Citado na página 31.

MADDALENA, L.; PETROSINO, A. A self-organizing approach to background subtraction for visual surveillance applications. *IEEE Transactions on Image Processing*, v. 17, n. 07, p. 1168–1177, 2008. Citado 2 vezes nas páginas 52 e 54.

____. A fuzzy spatial coherence-based approach to background/foreground separation for moving object detection. *Neural Computing and Applications*, v. 19, n. 2, p. 179–186, 2009. ISSN 1433-3058. Disponível em: <http://dx.doi.org/10.1007/s00521-009-0285-8>. Citado na página 36.

MOUNTRAKIS, G.; IM, J.; OGOLE, C. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 66, n. 3, p. 247–259, 2011. Citado na página 40.

NAYAR, S.; BEN-EZRA, M. Motion-based motion deblurring. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 26, n. 6, p. 689–698, June 2004. ISSN 0162-8828. Citado na página 43.

NETO, J. C. *et al.* Plant species identification using elliptic fourier leaf shape analysis. *Computers and Electronics in Agriculture*, v. 50, n. 2, p. 121 – 134, 2006. ISSN 0168-1699. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0168169905001560>. Citado na página 39.

NGHIEM, F. B. A.-T. Background subtraction in people detection framework for rgb-d cameras. In: *11th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. Seoul, Korea: [s.n.], 2014. p. 241–246. Citado na página 33.

OLIVER, N.; ROSARIO, B.; PENTLAND, A. A bayesian computer vision system for modeling human interactions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 22, n. 8, p. 831–843, 2000. Citado 2 vezes nas páginas 36 e 54.

RUSSELL, S.; NORVIG, P. *Artificial Intelligence: A Modern Approach*. 3rd. ed. [S.l.]: Prentice Hall, 2009. Citado na página 41.

SAUVOLA, J.; PIETIKAINEN, M. Adaptative document image binarization. *Pattern Recognition*, v. 33, p. 225–236, 2000. Citado na página 46.

SOBRAL, A.; VACAVANT, A. A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos. *Computer Vision and Image Understanding*, v. 122, p. 4–21, 2014. Citado 4 vezes nas páginas 33, 48, 51 e 54.

STAUFFER, C.; GRIMSON, W. Adaptive background mixture models for real-time tracking. In: *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on*. [S.l.: s.n.], 1999. v. 2, p. 252 Vol. 2. ISSN 1063-6919. Citado na página 36.

VALERA, M.; VELASTIN, S. A. Intelligent distributed surveillance systems: a review. *Vision, Image, and Signal Processing*, v. 152, n. 2, p. 192–204, 2011. Citado na página 40.

WANG, L. *et al.* Automatic gait recognition based on statistical shape analysis. *Image Processing, IEEE Transactions on*, v. 12, n. 9, 2003. Citado na página 33.

WREN, C. *et al.* Pfinder: real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 19, n. 7, p. 780–785, 1997. Citado 2 vezes nas páginas 35 e 54.

YAO, J.; ODOBEZ, J. Multi-layer background subtraction based on color and texture. In: *IEEE Conference on Computer Vision and Pattern Recognition*. Minneapolis, MN: [s.n.], 2007. p. 1–8. Citado 2 vezes nas páginas 52 e 54.