UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS
Departamento de Engenharia Elétrica

Lúcio André de Castro Jorge

Metodologia de fusão de vídeos e sons para
monitoração de comportamento de insetos

São Carlos
2011

Lúcio André de Castro Jorge

Metodologia de fusão de vídeos e sons para monitoração de comportamento de insetos

Tese apresentada à Escola de Engenharia de São Carlos – USP, como parte dos requisitos para obtenção do título de Doutor em Ciências, Programa de Engenharia Elétrica.

Área de concentração: - Processamento de Sinais e Instrumentação

Orientador: Prof. Dr. Valentin Obac Roda

São Carlos

2011

Trata-se da versão corrigida da tese. A versão original se encontra disponível na EESC/USP que aloja o Programa de Pós-Graduação de Engenharia Elétrica.

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica preparada pela Seção de Tratamento
da Informação do Serviço de Biblioteca – EESC/USP

J82m Jorge, Lúcio André de Castro.
 Metodologia de fusão de vídeos e sons para monitoração
de comportamento de insetos. / Lúcio André de Castro
Jorge ; orientador Valentin Obac Roda. São Carlos, 2011.

 Tese (Doutorado - Programa de Pós-Graduação em
Engenharia Elétrica e Área de Concentração em
Processamento de Sinais e Instrumentação) -- Escola de
Engenharia de São Carlos da Universidade de São Paulo,
2011.

 1. Fusão. 2. Vídeo. 3. Som. 4. *Wavelet-multifractal*.
5. Seleção de características. I. Título.

FOLHA DE JULGAMENTO

Candidato: Engenheiro **LUCIO ANDRÉ DE CASTRO JORGE**

Título da tese: **Metodologia de fusão de vídeos e sons para monitoração de comportamento de insetos.**

Data da defesa: 02/09/2011:

Comissão Julgadora:

Resultado:

Prof. Associado **Valentin Obac Roda (Orientador)**
(Escola de Engenharia de São Carlos/EESC)

APROVADO

Prof. Dr. **Maximilian Luppe**
(Escola de Engenharia de São Carlos/EESC)

Aprovado

Prof. Associado **Adilson Gonzaga**
(Escola de Engenharia de São Carlos/EESC)

APROVADO

Dr. **Adolfo Nicolas Posadas Durand**
(International Potato Center)

APROVADO

Profª. Titular **Agma Juci Machado Traina**
(Instituto de Ciências Matemáticas e de Computação/ICMC/USP)

APROVADO

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:
Prof. Titular **Denis Vinicius Coury**

Presidente da Comissão de Pós-Graduação:
Prof. Associado **Paulo Cesar Lima Segantine**

Dedido este trabalho
À Deus,
à Nossa Senhora,
à minha esposa Paula e ao meu filho Eduardo,
e aos meus pais.

Agradecimentos

Em primeiro lugar, agradeço à Deus, pela minha vida, pelos meus dons e pelas vezes que me carregou no colo aliviando as pedras do caminho.

Aos meus amores Paula e Eduardo que torceram, lutaram, choraram, comemoraram, sofreram, apostaram, me deram força, amor, carinho, ânimo, incentivo e acima de tudo estiveram sempre juntos comigo nesta jornada.

Aos meus pais Edward e Ana Dolores, por nunca me deixarem desistir de nada e ensinarem a olhar sempre para frente, pelo amor, carinho e orações.

Ao meu amigo, orientador e incentivador Professor Valentin Obac Roda pela paciência, orientação e oportunidade.

À minha família, de quem sempre tive incentivo.

À Embrapa pelo apoio sempre presente.

E para não correr o risco de esquecer de ninguém especial, agradeço a todos amigos pelo amizade, força, compreensão, incentivo, críticas, sugestões, etc....

Resumo

Este trabalho apresenta uma nova abordagem para fusão de vídeo e som diretamente no espaço de atributos visando otimizar a identificação do comportamento de insetos. Foi utilizado o detector de Harris para rastreamento dos insetos, assim como a técnica inovadora Wavelet-Multifractal para análise de som. No caso da Wavelet-Multifractal, foram testadas várias *Wavelet*-mães, sendo a *Morlet* a melhor escolha para sons de insetos. Foi proposto a *Wavelet* Módulo Máximo para extrair atributos multifractais dos sons para serem utilizados no reconhecimento de padrões de comportamento de insetos. A abordagem *Wrapper* de mineração de dados foi usada para selecionar os atributos relevantes. Foi constatado que a abordagem Wavelet-multifractal identifica melhor os sons, particularmente no caso de distorções provocadas por ruídos. As imagens foram responsáveis pela identificação de acasalamento e os sons pelos outros comportamentos. Foi também proposto um novo método do triângulo como representação simplificada do espectro multifractal visando simplificação do processamento.

Palavras-chaves: fusão de sensores, seleção de características, Wavelet-multifractal

Abstract

This work presents an innovative video and sound fusion approach by feature subset selection under the space of attributes to optimally identify insects behavior. Harris detector was used for insect movement tracking and an innovative technique of Multifractal-Wavelet was used to analyze the insect sounds. In the case of Multifractal-Wavelet, more than one mother-wavelet was tested, being the Morlet wavelet the best choice of mother-wavelet for insect sounds.

The wavelet modulus maxima was proposed to extract multifractal sound attributes to be used in pattern recognition of an insect behavior. The wrapper data mining approach was used to select relevant attributes. It has been found that, in general, wavelet-multifractal based schemes perform better for sound, particularly in terms of minimizing noise distortion influence. The image features only determine the mating and the sound other behaviors. A new triangle representation of multifractal spectrum was proposed as a processing simplification.

Keywords: fusion, feature selection, Wavelet-multifractal, wrapper

Lista de Figuras

Figura 1.1 - Tela do Xbug mostrando um rastreamento do movimento do inseto. (Fonte: manual do usuário do Xbug disponível na Embrapa Biotecnologia).....	3
Figura 1.2 - Exemplos do sistema de monitoramento do comportamento animal, Ethovision, gerando um etograma digital. (a) Interface mostrando demarcações ao longo do tempo para o comportamento de uma vespa: ataque, busca, descanso, caminhar, etc.; (b) comportamento de ratos: rastros pela parte central do corpo ou pelo nariz (Fonte: < http://www.noldus.com/animal-behavior-research/products/the-observer-xt >. Acesso dia 20/10/2010).....	5
Figura 1.3 - (a) Tela do software SACAM, desenvolvido no início deste trabalho de doutoramento, e devidamente registrado no INPI (Fonte: Manual de Utilização do SACAM); (b) Exemplo de rastreamento.	6
Figura 1.4 - Reconhecimento Automático de Fala.....	9
Figura 1.5 - Modelo Oculto de Markov com Seis Estados sendo Quatro Emissores (Fonte: KOTELLY, 2003).	10
Figura 1.6 - (a) Forma de onda do acasalamento do percevejo (<i>Euschistus heros</i>) monitorado neste trabalho e (b) o espectrograma do sinal (Frequência x Tempo).....	13
Figura 2.1 - Linha do tempo de trabalho com Fluxo Ótico	20
Figura 2.2 - Função peso $W(x,y)$. (Fonte: SIMAKOV; FROLOVA, 2004)	22
Figura 2.3 - Elipse $E(u,v) = \text{constante}$	23
Figura 2.4 - Representação de $\lambda_2 \times \lambda_1$ mostrando como podem ser detectados os pontos de cantos numa imagem (Fonte: SIMAKOV; FROLOVA, 2004)	23
Figura 2.5 - Determinação de Cantos em função de R (Fonte: SIMAKOV; FROLOVA, 2004)	24

Figura 2.6 - Aplicação da DoG no espaço imagem e escala	25
Figura 2.7 – Máximos e mínimos da Diferença de Gauss. (Fonte: LOWE, 2004).....	26
Figura 2.8 – Gradiente da Imagem. (Fonte: LOWE, 2004)	29
Figura 2.9 – Descritor do <i>keypoint</i> . (Fonte: LOWE, 2004)	29
Figura 3.1 - Sinais exemplos de séries de tempo: (a) um ruído gerado numericamente; (b) eletrocardiograma, (c) ruído gerado numericamente (Fonte: < http://www.physionet.org/tutorials/multifractal/behavior.htm > Acesso dia 20/03/2009).	35
Figura 3.2 - Exemplos de funções <i>Wavelet</i> -mãe	37
Figura 3.3 - Escalamento de uma função <i>Wavelet</i> -mãe	38
Figura 3.4 - Translação aplicada a uma <i>wavelet</i>	39
Figura 3.5 - Cálculo do fator C de relação entre a <i>wavelet</i> -mãe e o sinal.....	39
Figura 3.6 - Translação da função <i>wavelet</i> -mãe.....	40
Figura 3.7 - Escalamento da função <i>wavelet</i> -mãe.....	40
Figura 3.8 - Representação do espectro <i>Wavelet</i> , ou escalograma, resultante com todos os coeficientes calculados. (Fonte: ADDISON, 2002).....	41
Figura 3.9 - (a) Sinal de som de um inseto macho e (b) o escalograma da TWC do sinal (a) em aplicando a <i>Wavelet</i> -mãe da Figura 3.10, obtidos pelo programa desenvolvido neste trabalho.	42
Figura 3.10 - Função <i>Wavelet</i> -mãe de Morlet (MISITI et al., 2002)	43
Figura 3.11 - Processo de filtragem de um sinal	45
Figura 3.12 Exemplo de filtragem com <i>downsampling</i> de um sinal senoidal ruidoso..	45
Figura 3.13 - Representação e formação do MMTW de um sinal qualquer. (Fonte: http://cas.enscm.fr/~chaplais/wavetour_presentation/Regularite/Detection_of_singularities.html , acessado em 20/06/2009).....	47
Figura 3.14 - Comparação entre a geometria fractal e a geometria euclidiana. (Fonte: JORGE et al., 2008).....	48

Figura 3.15 - Modelos de fácil entendimento, (a) um brócolis, (b) é o detalhe das estruturas do brócolis que demonstram a auto-similaridade e (c) é uma folha de samambaia, demonstrando o mesmo. (Fonte: < http://pt.wikipedia.org/wiki/Fractal > acessado em 20/05/2008).	49
Figura 3.16 - Exemplo de uma construção de um fractal determinístico. (Fonte: JORGE et al., 2008).....	49
Figura 3.17 - Auto-similaridade em um sinal (Fonte: GRAPS, 1995).....	50
Figura 3.18 - Utilização de caixas no método “Box-counting” para o cálculo da dimensão fractal, $D_h = 1,39$ (Fonte: POSADAS, 2001).....	51
Figura 3.19 - Uma aplicação do método de análise fractal em uma estrutura 2D de um sistema poroso em solos. Os poros representam-se com áreas negras, e ϕ é o tamanho dos quadrados, ou escala. Os quadrados que estão sendo especificados na figura mostram os extremos contidos em um quadrado. (Fonte: POSADAS et al., 2001).	52
Figura 3.20 - Espectro multifractal.....	57
Figura 4.1 - Níveis de fusão: (a) básico ou por dados; (b) intermediário ou por características e (c) pós-decisão ou por decisão.	63
Figura 4.2 -.Espaço de busca (estados) (BLUM; LANGLEY, 1997).....	64
Figura 4.3 -.Abordagem embutida (FERREIRA ; JORGE, 2007).....	68
Figura 4.4 -.Abordagem Filtro (FERREIRA ; JORGE, 2007)	69
Figura 5.1 - Etapas de desenvolvimento da fusão de som e imagem.....	73
Figura 5.2 - Vista geral do local do experimento: (a) distribuição de microfones e suporte da câmera de vídeo (b) autofalante invertido, sendo usado como microfone.	74
Figura 5.3 - Tomada de uma arena to tipo circular utilizada com 2 insetos. (a) vista lateral sob um auto-falante, e (b) vista superior com microfone de eletreto e superfície de polímero excelente resistência.....	75
Figura 5.4 - Etograma dos comportamentos reprodutivos de machos e fêmeas . (Fonte: Embrapa Recursos Genéticos).....	76

Figura 5.5 - Sons do <i>Euschistus heros</i> notificados por Moraes et al. (2005) cantos (a) e (b) de fêmea, (c) e (d) de macho, (e) de rivalidade dos machos e (f) macho após recusa de cópula pela fêmea (Fonte: Moraes et al.,2005).	77
Figura 5.6 - Exemplos dos sons capturados no experimento e os respectivos espectrogramas calculados.	77
Figura 5.7 - Sequência de vídeo típica capturada nos experimentos.	78
Figura 5.8 - Imagens de diferenças de quadros consecutivos do vídeo da Figura 5.7, bem como a cada 9 a 10 quadros.	79
Figura 5.9 - Imagem mostrando resultado de movimento utilizando o Fluxo ótico Lucas-Kanade com média de 8 frames consecutivos.	80
Figura 5.10 - Utilizando a diferença entre dois quadros, (a) na vertical e (b) na horizontal, por Lucas-kanade, (1981).	80
Figura 5.11 - Detecção de movimento com GMM adaptado.	81
Figura 5.12 - Keypoints nos quadros nos instantes de tempo (a) t e (b) $t+1$ determinados pela transformada SIFT.	82
Figura 5.13 - Pontos característicos obtidos na transformada SIFT e Harris, interligados pelo algoritmo RANSAC.	83
Figura 5.14 - Exemplo típico de um relatório gerado para cada trilha do experimento.	86
Figura 5.15 - Fluxograma de ações no software SACAM para cada experimento.....	87
Figura 5.16 - Algoritmo da Transformada Wavelet contínua.....	88
Figura 5.17 - Wavelet-mães para cantos de fêmea: (a) Morlet (c) Chapéu Mexicano (e) Shanon, com os respectivos espectros TW.	90
Figura 5.18 - Wavelet-mães para cantos de macho: (a) Morlet (c) Chapéu Mexicano (e) quarta derivada Gaussiana, com os respectivos espectros TW.	91
Figura 5.19 - Gráfico da variação da entropia do comportamento dos insetos ao longo do tempo. Linhas: vermelha: macho, azul: fêmea, preta: acasalamento, verde: fêmea	

com som de macho induzido, roxa: macho com som de fêmea induzido, amarela: rivalidade, cinza: pós-tentativa acasalamento.....	94
Figura 5.20 – Exemplo do Triângulo multifractal	100

Lista de Tabelas

TABELA 5.1 – Espectros Wavelet com módulo máximo e espectros fractais para exemplos de canto de macho, canto de fêmea, acasalamento e rivalidade.	92
TABELA 5.2 - Dados das emissões de sons de insetos do programa, onde temos os diferentes tipos de comportamento. Os parâmetros de Δf , $\Delta f_{\text{máx}}$ e Δf_{min} são meramente para ver a simetria da curva, α é o coeficiente de Hölder, o D_0 é a dimensão fractal propriamente dita, D_1 e D_2 são as dimensões de entropia e de correlação, respectivamente. [] números negativos; - dado não existente; q=0, 0.	93
TABELA 5.3: Resultados da análise dos experimentos com e sem o Wrapper	98
TABELA 5.4: Precisão de cada classe analisada	99
TABELA 5.5 - Resultados da análise dois experimentos com e sem o Wrapper	100
TABELA 5.6 - Precisão de cada classe analisada no terceiro experimento.....	100

Lista de Siglas

API	<i>Application Programming Interface</i>
AM	Aprendizado de Máquina
AR	Auto regressivo
CFS	<i>Correlation-based Feature Selection</i>
CWT	Transformada Wavelet Contínua
DFT	Transformada Discreta de <i>Fourier</i>
DLT	Método da transformação linear direta
<i>DoG</i>	Função de Diferença de Gaussianas
DTW	<i>Dynamic time warping</i>
DWT	Transformada <i>Wavelet</i> Discreta
FDP	Função de distribuição de probabilidade
FFT	Transformada rápida de Fourier
FSS	<i>Feature Subset Selection</i>
GMM	<i>Gaussian Mixture Model</i>
HMM	<i>Hidden Markov Models</i>
LPC	<i>Linear Predictive Coding</i>
MFCC	<i>Mel-Frequency Cepstral Coefficients</i>
MMTW	Módulo Máximo da Transformada <i>Wavelet</i>
PLP	<i>Perceptual Linear Prediction</i>
RANSAC	<i>Random Sample Consensus</i>
RNA	Rede Neural Artificial
SACAM	Software para Análise de Comportamento de Animais em Movimento
SIFT	<i>Scale-Invariant Feature Transform</i>
STFT	<i>Short-Time Fourier Transform</i>
TCF	Transformada Contínua de <i>Fourier</i>

TDF	Transformada Discreta de <i>Fourier</i>
TDNNs	<i>Time Delay Neural Networks</i>
TF	Transformada de <i>Fourier</i>
TFJ	Transformada de <i>Fourier</i> Janelada ou Transformada de Gabor
TW	Transformada <i>Wavelet</i>
TWC	Transformada <i>Wavelet</i> Contínua
TWD	Transformada <i>Wavelet</i> Discreta

Sumário

AGRADECIMENTOS	VI
RESUMO	VII
ABSTRACT	VIII
LISTA DE FIGURAS	IX
LISTA DE TABELAS	XIV
LISTA DE SIGLAS	XV
SUMÁRIO	XVII
CAPÍTULO 1 INTRODUÇÃO	1
1.1 CONSIDERAÇÕES INICIAIS	1
1.2 ANÁLISE DO COMPORTAMENTO ANIMAL	2
1.3 OBJETIVO	14
1.3.1 Objetivos específicos	14
1.4 CONSIDERAÇÕES FINAIS	14
CAPÍTULO 2 RASTREAMENTO DE OBJETOS EM VÍDEO	16
2.1 CONSIDERAÇÕES INICIAIS	16
2.2 MÉTODOS DE RASTREAMENTO	17
2.2.1 Detectores de Objetos	18
2.2.1.1 Detecção por Subtração	18
2.2.1.2 Modelagem por Mistura de Gaussianas	19
2.2.1.3 Fluxo Ótico	20
2.2.1.4 Detector de Harris	22
2.2.1.5 SIFT (Scale Invariant Feature Transform)	24
2.3 CONSIDERAÇÕES FINAIS	31
CAPÍTULO 3 WAVELET-MULTIFRACTAL	33
3.1 CONSIDERAÇÕES INICIAIS	33
3.2 A TRANSFORMADA WAVELET (TW)	34
3.2.1 A Transformada Wavelet Contínua (TWC)	37
3.2.2 A Transformada Wavelet Discreta	43
3.2.3 Módulo Máximo da Transformada Wavelet (MMTW)	46
3.3 TEORIA FRACTAL	47
3.3.1 Auto-similaridade	48
3.3.2 Determinação da Dimensão Fractal	50
3.3.3 Multifractal	51
3.4 Extração de características em sinais sonoros	57
3.5 Considerações Finais	59

CAPÍTULO 4 FUSÃO	60
4.1 CONSIDERAÇÕES INICIAIS.....	60
4.3 SELEÇÃO DE SUBCONJUNTO DE CARACTERÍSTICAS	63
4.4 SELEÇÃO DE CARACTERÍSTICAS COMO BUSCA HEURÍSTICA	66
4.4.1 Ponto de partida	66
4.4.2 Organização da busca	66
4.4.3 Estratégia de Avaliação.....	67
4.4.4 Critério de parada.....	67
4.5 ABORDAGENS PARA SELEÇÃO DE CARACTERÍSTICAS	67
4.5.1 Embutida	68
4.5.2 Filtro.....	68
4.5.3 Wrapper.....	69
4.6 SELEÇÃO DE CARACTERÍSTICAS COMO FUSÃO DE DADOS.....	70
4.7 CONSIDERAÇÕES FINAIS	71
CAPÍTULO 5 RESULTADOS E DISCUSSÕES	72
5.1 CONSIDERAÇÕES INICIAIS.....	72
5.2 MATERIAIS.....	74
5.3 DETECÇÃO DE MOVIMENTO.....	78
5.4 DETECÇÃO DE SOM	85
5.5 FUSÃO POR SELEÇÃO DE CARACTERÍSTICAS	95
5.5 CONSIDERAÇÕES FINAIS	101
CAPÍTULO 6 CONCLUSÕES E TRABALHOS FUTUROS	102
6.1 CONCLUSÕES	102
6.2 PROPOSTAS DE TRABALHOS FUTUROS	103
REFERÊNCIAS BIBLIOGRÁFICAS.....	105

Capítulo 1

Introdução

1.1 Considerações Iniciais

Existem vários ramos da ciência que estudam o comportamento de animais, principalmente quanto à sua evolução, comparação entre espécies e mesmo para entender o comportamento visando o seu controle. Técnicas tradicionais de monitoramento de animais são subjetivas, manuais e em muitos casos as informações não são simultâneas na descrição do comportamento. Em ambientes agrícolas controlados de produção animal, o comportamento do animal é avaliado com base nas condições ambientais. A detecção de condições prejudiciais ou irregulares, por exemplo, temperatura elevada, pode ser obtida apenas pela observação do comportamento animal, que reage às condições inadequadas de forma diferenciada. Em geral, as unidades produtoras são monitoradas localmente com a supervisão humana (PANDORFI, 2002). Os peritos usam frequentemente a informação visual dos animais para o diagnóstico, mesmo antes de toda a condição de alarme ser detectada por meio de sensores e instrumentos, que auxiliam na análise da resposta das variáveis. O fato é que, às vezes, o especialista que pode diagnosticar corretamente uma condição anormal pode não estar no local para executar uma análise visual. Assim, sistemas automatizados que executam a avaliação e o controle contínuo, em tempo real, principalmente por análise de imagens e outros parâmetros, têm despertado grande interesse, como por exemplo, o conforto térmico dos suínos confinados, baseando-se em testes de padrão de comportamento, como demonstrados por Pandorfi (2002), fornecendo informações pertinentes ao gerente, proprietário ou a um consultor, para a tomada de decisão.

Além da aplicação em unidades de produção, o estudo do comportamento animal tem grande interesse no controle de pragas e doenças, no estudo da própria espécie animal, no entendimento do processo comportamental diante de diferentes situações. Uma forma convencional para o inventário dos comportamentos de animais

utiliza comumente um etograma, que é um inventário dos comportamentos de uma espécie, com os comportamentos detalhadamente descritos e organizados em categorias (instinto, aprendizado, comportamento agressivo, territorial, reprodutivo, alimentar, social e de defesa). Ele coloca o repertório de comportamentos em uma estrutura organizada que permite aos etólogos entenderem como cada comportamento ajuda o animal a sobreviver, acasalar e conseguir alimentos. Os etogramas são também valiosos instrumentos que auxiliam os cientistas na análise da evolução do comportamento, pois facilitam a comparação entre espécies de uma mesma linha evolutiva.

Neste capítulo será feita uma introdução sobre a importância da análise de comportamento de animais mostrando alguns sistemas e metodologias utilizadas.

1.2 Análise do Comportamento Animal

Várias são as formas de observação, em geral, pela análise do movimento, mas tem também detecção de diferentes odores, a emissão de sons, etc. O registro preciso e a análise do movimento realizado em resposta ao estímulo oferecido é difícil, particularmente se taxas de movimento e mudanças na direção são informações essenciais no registro. Desta forma, somente um sistema automático para a análise torna viável o *monitoramento* (JORGE et al., 2005; JORGE et al., 2006).

Um sistema para monitoramento do comportamento deve seguir os movimentos do animal em tempo real e armazenar todas as informações possíveis como: o tempo de residência em cada área da região monitorada, a velocidade, a tortuosidade e a linearidade no deslocamento. Este tipo de monitoramento é praticamente impossível de ser realizado por observação direta e registro simultâneo, sem a utilização de nenhum aparelho específico (JORGE et al., 2006).

Como alternativa existem alguns softwares para o monitoramento do comportamento dos animais, que permitem a observação em tempo real do animal e medem uma série de parâmetros importantes para a análise das respostas comportamentais. Entre eles o X-bug (COLAZZA et al., 1999) e o Ethovision (NOLDUS et al., 2002a; NOLDUS et al., 2002b), foram especialmente desenvolvidos para pesquisas na área de entomologia. O primeiro é um aplicativo para ambiente Linux, com código fonte não liberado e que apresenta algumas limitações como: somente insetos com comportamento e tamanho próximo ao de vespas parasitas (0,1 a 15 mm) são monitorados com precisão e eficiência, insetos maiores e com movimentos muito

rápidos não são registrados (Figura 1.1); e o segundo é um aplicativo comercial, de alto custo, com ferramentas que permitem o estudo de até 16 animais ao mesmo tempo. Nesse caso, tenta-se fazer um etograma digital semiautomático, fazendo os registros de comportamento após a captura das imagens, apenas pelos vídeos, mostrando uma evolução dos eventos ao longo do tempo (Figura 1.2).

Figura 1.1 - Tela do Xbug mostrando um rastreamento do movimento do inseto. (Fonte: manual do usuário do Xbug disponível na Embrapa Biotecnologia)

Também, numa parceria da Embrapa Instrumentação com a Embrapa Recursos Genéticos foi iniciado um projeto para o estudo do comportamento animal: "Desenvolvimento de Instrumentação para Estudos Comportamentais de Insetos - Novas Estratégias para Manejo de Pragas", dentro do qual se desenvolveu um primeiro protótipo do sistema SACAM (software para análise de comportamento de animais em movimento) (JORGE et al., 2006). Nesse projeto, foram iniciados os estudos dos comportamentos dos insetos em relação às pesquisas com semioquímicos para o controle biológico. A necessidade de avaliar uma grande quantidade de informações da atividade comportamental dos animais leva à tentativa

de encontrar soluções mais automáticas e que são essenciais para a compreensão completa da sua função e dos mecanismos comportamentais.

As técnicas de análise de imagem estão se desenvolvendo rapidamente e cada vez mais se torna viável na análise de comportamento em tempo real. Os resultados preliminares obtidos com o protótipo do SACAM desenvolvido (JORGE et al., 2006) mostraram a viabilidade de monitoramento por imagem para diferentes tamanhos de insetos.

Alguns trabalhos de análise de imagens de vídeo para estudo de animais em movimento têm sido desenvolvidos, sendo um dos primeiros com maior destaque o trabalho desenvolvido por Miller et al. (1982). Uma série de pacotes e ferramentas vem sendo desenvolvidos para diferentes animais. Vigneault et al. (1990) descreveram um sistema para medir bolhas em tempo real que foi adaptado para o monitoramento de insetos. No entanto, este sistema requer um sistema de iluminação relativamente caro e praticamente não é usado pelos pesquisadores. Hoy et al (1996) desenvolveram um software cujo algoritmo tinha como base a obtenção de uma imagem referência sem o inseto, posteriormente usada na comparação com as imagens tomadas com o inseto.

Este é o mesmo processo usado pelo software X-bug, desenvolvido por Colazza et al. (1999) e um dos processos disponíveis no Etheovison (NOLDUS et al., 2002a).

A abordagem inicial para detecção de movimento adotada no SACAM foi o algoritmo de diferenças entre quadros sucessivos de vídeo, calculando-se o valor do tom de cinza médio numa janela de tamanho proporcional ao tamanho do inseto. Um exemplo da interface desenvolvida pode ser visualizado na Figura 1.3a, e os respectivos resultados do rastreamento na Figura 1.3b.

Uma das primeiras aplicações de sistemas de visão computacional para o reconhecimento de movimento foi para o movimento humano. Na literatura, técnicas de detecção de movimento, rastreamento de objetos e análise do movimento tem recebido destaque no campo da visão computacional, sendo as principais aplicações a identificação de marcha, detecção automática de objetos, monitoramento de segurança, recuperação de informações em dados multimídia, reconhecimento de gestos, monitoramento de tráfego, dentre outras aplicações (ZHOU; ZHANG, 2005; YILMAZ et al., 2006). Normalmente os métodos detectam movimento de forma pontual, por região, por detecção de forma ou modelos previamente elaborados.

Figura 1.2 - Exemplos do sistema de monitoramento do comportamento animal, Ethovision, gerando um etograma digital. (a) Interface mostrando demarcações ao longo do tempo para o comportamento de uma vespa: ataque, busca, descanso, caminhar, etc.; (b) comportamento de ratos: rastros pela parte central do corpo ou pelo nariz (Fonte: <<http://www.noldus.com/animal-behavior-research/products/the-observer-xt>>. Acesso dia 20/10/2010).

Figura 1.3 - (a) Tela do software SACAM, desenvolvido no início deste trabalho de doutoramento, e devidamente registrado no INPI (Fonte: Manual de Utilização do SACAM); (b) Exemplo de rastreamento.

Métodos baseados em modelos geralmente possuem um grande custo computacional para a estimativa do movimento, por exemplo, marcha humana baseada no modelo da postura humana. Também estes modelos necessitam de algumas considerações auxiliares, como o conhecimento da cor da pele, a fim de reduzir este custo computacional (GONÇALVES et al., 2006). Por outro lado, há ainda

o método baseado no padrão da aparência para reconhecimento de movimento humano, que não utiliza estimativa da postura. Neste caso, o reconhecimento de movimento pertence a um tipo de reconhecimento de padrões. A técnica GMM (*Gaussian Mixture Model*) é uma das mais usadas (AGGARVAL; CAI, 1999).

A técnica GMM é utilizada para modelar diferentes classes ou gêneros de movimento, como uma função de distribuição de probabilidade (FDP), utilizando uma combinação de FDPs moderadas, ou seja, uma mistura. Os parâmetros do conjunto são ajustados e este processo é iterativo. Após este processo, os GMMs resultantes são usados para determinar as probabilidades de um movimento pertencer a uma das classes (GONÇALVES et al., 2006).

Mais recentemente, utilizada em detecção de movimentos em tempo real, técnicas que detectam pontos característicos, como a transformada SIFT (*Scale-invariant feature transform*) (LOWE, 2004) e detector de Harris (WEIJER et al., 2005), tem sido largamente usadas (STAVENS; THRUN, 2010). Estas técnicas utilizam características obtidas localmente na imagem, sendo invariantes à rotação e à escala. Estas características são robustas com relação à iluminação, aos ruídos, às oclusões e mudanças de ponto de vista na imagem. Além destas propriedades, possuem alta capacidade de distinção, são fáceis de serem extraídos, permitem identificação de objetos de interesse com precisão (LOWE, 2004; STAVENS; THRUN, 2010).

Exemplos de aplicações em monitoramento do comportamento de animais por imagem aparecem em vários trabalhos: insetos (FRY et al., 2000; MÜLLER; ROBERT, 2001; KUMAR et al., 2007; VEERARAGHAVAN et al., 2008; MAURER et al., 2010; KUMAR et al., 2010), serpentes (GONÇALVES et al., 2007), gado de leite (GREEN et al., 2010; BATTIATO, et al., 2010), ratos (GYORY et al., 2010; JHUANG et al., 2010), pássaros (DICKINSON et al., 2010; KO et al., 2010; LÉVESQUE; BERGEVIN, 2010), morcegos (THERIAULT et al., 2010; XIAO; FISHER, 2010), peixes (ZEEUW et al., 2010; CLINE; EDGINGTON, 2010; ROS-SÁNCHEZ et al., 2010), dentre outros.

No entanto, o comportamento animal não está restrito ao movimento realizado pelos animais, mas também existem outros fatores que identificam alterações, quando se deseja entender um dado comportamento. A emissão de sons como forma de comunicação é outro fator importante. Todos os animais, inclusive o homem, se utilizam de sinais sonoros para se comunicar com outros seres da mesma espécie. As comunicações sonoras em animais, em geral, descrevem diferentes comportamentos, como por exemplo, cantos de aproximação, demarcação de área, desconfortos, acasalamento, etc. Animais dependem de seus sentidos vibracionais e acústicos e

habilidades para detectar a presença de ambos predadores e presas e se comunicar com os membros da mesma espécie.

Nos humanos, o processo de comunicação é determinado pela fala, que nada mais é do que a emissão de informações em formato sonoro. Esta informação é complexa do ponto de vista da transmissão de informações e emoções, dependente da língua que o locutor está falando e até do seu sotaque. Na língua portuguesa, por exemplo, existem acima de 10 mil diferentes sons que são emitidos para a comunicação entre os locutores. Outras línguas, como as africanas e a chinesa, a entonação da palavra pode determinar significado completamente diferente à mesma.

A complexidade da língua falada humana não fica limitada à quantidade de sons que podem ser emitidos, mas na diferença de pronúncia entre locutores que é muito grande. Estas diferenças são caracterizadas pelos sotaques regionais, condições sociais, grupos sociais, problemas de mobilidade dos órgãos que geram a fala (boca, língua, lábios, dentre outros). O tamanho da laringe é um dos principais fatores que diferenciam a voz masculina da feminina. Para completar, o mesmo locutor não consegue produzir o mesmo som duas vezes da mesma forma.

O fato do som não ser reprodutível torna impossível a utilização de casamento de padrões para o reconhecimento de sons produzidos por humanos. Reconhecimento automático de fala é um dos ramos mais ativos em processamento de sinais nos últimos 30 anos. A complexidade da linguagem humana torna necessária a utilização de técnicas estatísticas no reconhecimento de fala.

Os sons emitidos por animais são similares aos humanos no que tange a sua diversidade mesmo entre os indivíduos da mesma espécie. Em Sven et al. (1996) é demonstrado que é possível distinguir entre indivíduos da mesma espécie de pássaro com técnicas utilizadas na verificação automática de locutor, técnica esta utilizada em biometria em sistemas de segurança.

Técnicas similares de reconhecimento automático de fala, com suas devidas adaptações, podem ser utilizadas no reconhecimento de sons emitidos por animais (POTAMITIS et al., 2006). A escolha adequada de parâmetros de observações é fator determinante para se aumentar o nível de reconhecimento e adaptar a parametrização à espécie estudada.

O processo de reconhecimento automático da fala pode ser resumido pela Figura 1.4, em que um locutor pronuncia uma sequência de palavras, ou seja, codifica um sinal em formato de áudio em uma informação textual. Através da extração de

atributos pode-se representar a fala por um vetor de características e fazer o reconhecimento.

Figura 1.4 - Reconhecimento Automático de Fala.

O locutor pronuncia uma sequência de palavras. Esta sequência é registrada por um microfone e transformada em um sinal elétrico (forma de onda). Esta forma de onda é digitalizada e segmentada em janelas de tipicamente 25 ms para posterior processamento. Usualmente estes segmentos são superpostos já que a informação da fala é redundante (HUANG et al., 2001). A segunda etapa é a parametrização destes segmentos em vetores de observação. Estes vetores de observação reduzem a quantidade de informações presentes na fala ao mesmo tempo em que tendem a reduzir diferenças interpessoais de características vocálicas entre os usuários do sistema. A informação redundante é parcialmente reduzida em sistema de compressão de áudio utilizados em telefonia no qual a informação da fala pode ser comprimida em até 400 bits/s (ERIKSSON; SJOBERG, 1993). MFCC (*Mel-Frequency Cepstral Coefficients*) (DAVIS; MERMELSTEIN, 1980) e PLP (*Perceptual Linear Prediction*) (HERMANSKY, 1990) são as técnicas mais comumente utilizadas na parametrização de sinais de fala. Finalmente o reconhecimento é realizado pela estimativa da sequência de eventos que gerou o sinal de fala. Modelos Ocultos de Markov (*Hidden Markov Models* - HMM) são os mais utilizados nesta etapa (KOTELLY, 2003).

Na Figura 1.5 ilustra-se um HMM utilizado no reconhecimento automático de fala. Os parâmetros a_{ij} representam a probabilidade de transição entre o estado i e o

estado j do modelo. Esta transição é realizada a cada janela de tempo (segmentação) apresentada na etapa anterior. Os parâmetros o_j compõem o vetor de observação do estado j e os parâmetros $b_j(o_j)$ representam a função densidade de probabilidade que usualmente é uma função Gaussiana. Cada estado do HMM é associado a um vetor de médias e a uma matriz de covariâncias, que usualmente pode ser simplificado para uma matriz diagonal caso os vetores de observação sejam descorrelacionados entre si. A sequência de eventos neste caso é dada pelas transições $1 \rightarrow 2$, $2 \rightarrow 2$, $2 \rightarrow 3$, $3 \rightarrow 4$, $4 \rightarrow 5$ e $5 \rightarrow 6$. Somente quatro estados são emissores porque emitem vetores de observação.

O reconhecimento de vocalizações de animais é apresentado em diferentes trabalhos que utilizam desde os Modelos Escondidos de Markov, redes neurais artificiais até o simples cálculo de distâncias (SVEN et al., 1996; KOGAN; MARGOLIASH, 1998; FANDIÑO-MARIÑO; VIELLIARD, 2004; SELOUANI et al., 2005; FAGERLUND; HÄRMÄ, 2005; CHEN; MAHER, 2006; SELIN et al., 2007).

A identificação de espécies de pássaros já existe há algum tempo com a utilização, na maioria dos casos, de sistemas de reconhecimento automático de fala humana, adaptados às vocalizações específicas das aves. O reconhecimento das vocalizações de pássaros é realizado com o objetivo de: monitoramento de padrões de comportamento dos pássaros; monitoramento de padrões de migração para evitar colisões com turbinas (EVANS, 2000) e aeronaves (KWAN et al., 2006); detecção e medição das características vocálicas dos pássaros; detecção e censo de espécies; identificação de indivíduos (FANDIÑO-MARIÑO; VIELLIARD, 2004), dentre outros.

Figura 1.5 - Modelo Oculto de Markov com Seis Estados sendo Quatro Emissores (Fonte: KOTELLY, 2003).

As vocalizações de pássaros podem ser divididas nas seguintes categorias (KREBS; KROODSMA, 1980; CATCHPOLE; SLATER, 1995):

- *Elementos* - são as unidades sônicas elementares da vocalização dos pássaros;
- *Silabas* - incluem um ou mais elementos e tem de alguns milissegundos de duração até centenas de milissegundos;
- *Frases* - são pequenos agrupamentos de silabas;
- *Chamadas* - são sequências compactas de frases;
- *Canções* - são longas e complexas vocalizações.

DTW (*Dynamic time warping*), HMM, redes neurais artificiais e os cálculos de distâncias entre sons de treinamento e de reconhecimento são as técnicas mais utilizadas no reconhecimento de pássaros. DTW foi utilizado para o reconhecimento de espécies de pássaros americanos com sucesso por Sven et al. (1996). A técnica é susceptível a ruído, o que a torna difícil de ser implementada em aplicações reais. HMMs e DTW foram comparados por Kogan e Margoliash (1998). Em condições de pouco ruído, os dois métodos tiveram desempenhos semelhantes, mas com ruído, HMMs tiveram uma performance muito superior devido ao baixo índice de inserções. HMMs tiveram problemas com a classificação de vocalizações de curta duração e unidades sonoras com estrutura mais complexa.

Em Selouani et al. (2005), redes neurais artificiais foram capazes de identificar 16 tipos diferentes de pássaros canadenses com um nível de reconhecimento de 83% utilizando sons de boa qualidade. Neste experimento, 20 coeficientes LPC (*Linear Predictive Coding*) foram utilizados para treinar duas redes neurais artificiais. A primeira era uma rede neural com atraso de tempo TDNNs (*Time Delay Neural Networks*) e a segunda era uma versão de um algoritmo *backpropagation* auto regressivo (AR). Redes neurais artificiais também foram utilizadas para identificar tons do *Poecile atricapillus* por Nickerson et al. (2006).

A classificação dos sons também pode ser realizada utilizando-se os picos espectrais (CHEN; MAHER, 2006), porém o método apresenta problemas quando são classificados sons aperiódicos, quando existe ruído nos sinais, ou quando os componentes espectrais variam muito rapidamente em frequência ou em amplitude. Nestes casos, a trilha do espectro é perdida e o método simplesmente não funciona.

MFCCs foram utilizados em conjunto com as distâncias Euclidiana e de Mahalanobis para a classificação de sons de pássaros que possuem cantos não

harmônicos (FAGERLUND; HÄRMÄ, 2005). Sons harmônicos de pássaros podem ser mais facilmente reconhecidos, principalmente para pássaros comumente encontrados (HELLER, 2005). Wavelets também são utilizadas na identificação de espécies de pássaros em conjunto com redes neurais (SELIN et al., 2007).

No caso dos insetos, estes apresentam um complexo sistema de comunicação baseado em sinais químicos, visuais e vibracionais. Segundo Ganchev et al. (2007), basicamente os sons emitidos são de diferentes origens:

- *Por estridulação*: gerado pela fricção de partes do corpo. Usualmente produzido por grilos, cigarras, formigas, etc.;
- *Por percussão*: gerado por batida de parte do corpo, como as patas, abdômen, ou cabeça contra um substrato, ou planta, usualmente parecido com o som de um tambor ou de um tapa;
- *Por vibração*: oscilação do corpo como asas no ar, comumente reconhecido como um zumbido, típico de mosquitos, pernilongos, abelhas, etc.;
- *Por contração*: contração rápida e liberação de músculos que operam como membranas, ouvido como uma sequência de toques tipo tambores, produzidos por grilos, gafanhotos, etc.
- *Por expulsão do ar*: eliminação de ar ou fluido através do corpo, usualmente como um assobio ou buzina, produzidos por baratas e gafanhotos.

Uma das pragas da lavoura brasileira, o percevejo ou *Euschistus heros*, tem recebido destaque nos estudos comportamentais (LAUMAN et al., 2005). A sua comunicação é vibracional e ocorre pela propagação, através de plantas, de vibrações geradas pelos movimentos do corpo do inseto. Os insetos, machos e fêmeas, emitem vibrações em diferentes frequências e essas variam de acordo com a função da comunicação, isto é, há diferentes vibrações para os comportamentos como chamamento e acasalamento. Na Figura 1.6, é possível ter uma boa noção dos sinais emitidos pelo inseto.

A Figura 1.6a ilustra a forma de onda gerada durante o acasalamento do percevejo e por meio do respectivo espectrograma (Figura 1.6b) os sinais emitidos pelo inseto no acasalamento ficam limitados na faixa dos 0 a 2 kHz, com alguns sinais de batidas indo até a faixa dos 5 kHz.

(a)

(b)

Figura 1.6 - (a) Forma de onda do acasalamento do percevejo (*Euschistus heros*) monitorado neste trabalho e (b) o espectrograma do sinal (Frequência x Tempo).

Considerando a complexidade da análise de comportamento, as especificidades para cada espécie, seja a humana até diferentes animais, se verifica que o comportamento é a resposta a vários estímulos, visuais, sonoros, olfativos, etc, podendo ser monitorado também por diferentes formas, como por exemplo, o movimento e o som emitido, isoladamente ou se complementando pela fusão de diferentes informações. Por outro lado, a análise de comportamento também implica no monitoramento ao longo do tempo, e consequentemente, ao acúmulo de uma grande quantidade de dados. Se somados diferentes respostas, multiplica-se significativamente a quantidade de dados para serem analisados. Existe a necessidade de avaliar uma grande quantidade de informações da atividade comportamental dos animais, pois também existem muitos comportamentos, o que leva à necessidade de encontrar soluções mais automáticas e que são essenciais para a compreensão completa da sua função e dos mecanismos comportamentais.

Técnicas tradicionais de monitoramento de animais são subjetivas, manuais e em muitos casos as informações não são simultâneas na descrição do comportamento. As técnicas de análise de imagens estão se desenvolvendo rapidamente e começam a se despontar como potencial para a análise de comportamento em tempo real. Além disso, técnicas de reconhecimento de áudio e fala tem tido grande avanço. Assim, a fusão de padrões de movimentação em vídeos e de padrões emitidos de sons aliados às técnicas de inteligência artificial se mostra como uma solução para monitoramentos de informações ao longo do tempo.

1.3 Objetivo

O objetivo deste trabalho é desenvolver uma metodologia de análise e monitoramento do comportamento de insetos comuns das lavouras brasileiras através de imagens e sons, em condições de laboratório.

1.3.1 Objetivos específicos

Neste trabalho são propostos:

- Aplicar a técnica Wavelet-Multifractal para caracterização de sons dos insetos e estabelecer descritor multifractal;
- Obter descritores de movimento a partir de imagens dos insetos;
- Fusão de Som e Imagem para estudo do comportamento de insetos.

1.4 Considerações Finais

O presente trabalho está organizado em 6 capítulos, incluindo as referências bibliográficas, distribuídos da seguinte forma:

Capítulo 1: Apresenta a introdução ao tema do trabalho mostrando a motivação e objetivos.

Capítulo 2: Apresenta uma revisão de técnicas de rastreamento de objetos em vídeos, as tecnologias utilizadas, base teórica e revisão de literatura. Descreve as técnicas de detecção de objetos por detector de Harris, transformada SIFT, subtração de quadros, fluxo ótico e mistura gaussiana.

Capítulo 3: Faz uma revisão teórica e de literatura da técnica utilizada no processamento de sons, a técnica Wavelet-Multifractal que foi a base deste trabalho.

Capítulo 4: Descreve a seleção de características utilizada para a fusão de dados de imagem e som. Faz uma rápida comparação com outras técnicas de fusão encontradas na literatura.

Capítulo 5: Detalha a proposta de fusão de som e imagem apresentando os resultados do comportamento de insetos em acasalamento.

Capítulo 6: Apresenta as conclusões e propostas de trabalhos futuros.

Por fim, são apresentadas as referências bibliográficas.

Capítulo 2

Rastreamento de objetos em vídeo

2.1 Considerações Iniciais

O rastreamento de objetos é uma tarefa importante em visão computacional. Com o aumento do poder computacional e disponibilidade de câmeras de vídeo de boa qualidade, a área de análise de vídeo em tempo real tem se desenvolvido rapidamente, principalmente soluções na detecção de movimento, no rastreamento e na análise do movimento.

O rastreamento de objetos tem sido utilizado no reconhecimento baseado no movimento, principalmente em estudos da marcha humana; na área de segurança, para detecção automática de movimentos suspeitos; na indexação de vídeos, com identificação e detecção de vídeos em bases de dados multimídia; e em sistemas para navegação de veículos, extraíndo estatísticas, evitando obstáculos e planejando novos caminhos, dentre outros (YILMAZ et al., 2006).

De uma forma simples, o rastreamento de objetos visa estimar a sua trajetória ao longo da cena, ou ao longo de vários *frames* de vídeo. No entanto, esta não é uma tarefa tão simples assim, pois em geral, deve mapear o mundo real 3D na tela 2D, as imagens possuem ruídos ou variações na iluminação, os objetos podem sofrer oclusões por obstáculos, podem possuir formas complexas e às vezes serem articulados, além do tempo de processamento, que deve ser próximo do real. No caso dos insetos, as patas podem se movimentar sem que eles se desloquem, tornando os algoritmos mais complexos quando se deseja apenas o movimento do inseto ao longo da área, podem também, se sobrepor, apresentando problemas de oclusão, bem como dificuldade na identificação do número de objetos.

Para simplificar os processos de rastreamento, em geral adotam-se algumas restrições ou condições de contorno, com o conhecimento prévio do tipo de objeto, do número e da forma dos mesmos, bem como da aplicação.

Várias são as abordagens de rastreamento de objetos presentes na literatura, e neste capítulo são descritas as principais técnicas usadas e testadas neste trabalho.

2.2 Métodos de rastreamento

Para qualquer técnica ou método de rastreamento é necessário definir como será representado ou discriminado o objeto a ser rastreado, ou seja, que característica será rastreada. Cada objeto pode ser representado pelo seu histograma, pelo seu contorno, pela sua cor, por seu modelo ou por qualquer outra definição. Em geral os algoritmos variam de acordo com as características que se buscam.

Segundo YILMAZ et al. (2006) os objetos a serem rastreados em uma cena de vídeo são definidos segundo suas características principais, podendo ser representados de diferentes formas, destacando-se principalmente:

- *Pontos*: objetos representados por pontos ou centroides de uma região;
- *Formas geométricas básicas*: retângulos e elipses em geral, representando as características de uma região da imagem;
- *Contornos ou silhuetas de objetos*: definem o contorno do objeto sendo rastreado;
- *Modelos de formas articuladas*: representando um corpo rígido e partes articuladas;
- *Esqueletos*: modelo extraído do eixo médio central de um objeto;
- *Formato padrão (Template)*: formas básicas descrevendo a aparência de um objeto;
- *Distribuição de Densidade de Probabilidade*: aparência do objeto descrita de forma paramétrica como uma gaussiana ou mistura de gaussianas, ou não paramétrica, como os histogramas.

Existem outras formas de representação que podem considerar diferentes pontos de visão sobre os objetos, utilizando classificadores estatísticos e modelos ativos, mostrando que, em geral, o método de rastreamento a ser utilizado está diretamente ligado à forma de representação do objeto rastreado.

As diferentes formas de representação dos objetos aliadas às propriedades como a cor, contornos, textura, movimentações de pixels ou fluxos óticos, são escolhidos dependendo da aplicação (YILMAZ et al., 2006). O rastreamento de objetos em movimento inicia com a detecção, que pode ser realizada segundo detectores de pontos, de regiões, modelagem do fundo ou por classificadores supervisionados, dentre outros.

2.2.1 Detectores de Objetos

Um dos algoritmos mais simples e intuitivos para detecção de objetos em movimento é o método da subtração de quadros consecutivos em um vídeo (JORGE et al., 2006). São muito sensíveis a ruídos e mudanças de iluminação. Quando o número de quadros na sequência é grande e existe pequena alteração entre quadros consecutivos, então a modelagem do fundo é mais indicada. Esta técnica é muito usada no contexto de aplicações de segurança, quando a câmera é fixa, e é classificada como métodos preditivos ou não preditivos. Algumas variações podem ser consideradas, partindo da subtração do quadro atual de um quadro padrão, obtido antes do início da detecção, até a subtração dos últimos quadros sucessivos. A mistura de gaussianas (STAUFFER; GRIMSIN, 2000) é o método de detecção baseado na modelagem do fundo da cena que mais se destaca. Variações destes usam funções de Misturas Gaussianas e aparecem em muitas aplicações (BUGEAU; PÉREZ, 2009).

Depois da categoria de detecção por subtração, a detecção pelo fluxo ótico, é um dos mais citados na literatura, destacando-se os métodos propostos por Lucas e Kanade (1981) e Horn e Schunck (1981). Detectores de objetos por pontos salientes têm sido os mais citados recentemente, destacam-se os detectores de Harris e a transformada SIFT. A abordagem desses detectores é extrair quadro a quadro os cantos ou pontos salientes presentes na imagem, seguido de um algoritmo de correspondência entre os pontos detectados em quadros sucessivos para geração dos rastros. Para detectores de pontos destacam-se o detector de cantos de Harris (HARRIS; STEPHENS, 1988), e mais recentemente a transformada SIFT descrita por Lowe (2004).

E por fim, máquina de vetores de suporte, redes neurais e outros algoritmos evolutivos são os que se destacam na detecção por classificadores supervisionados (YILMAZ et al., 2006). A seguir são apresentadas as formulações de cada um destes algoritmos utilizados neste trabalho.

2.2.1.1 Detecção por Subtração

A subtração de imagens ou quadros de vídeos é o método mais utilizado nos primeiros trabalhos de monitoramento de insetos e animais (JORGE et al., 2005). Basicamente, a detecção de movimento é adotada pela simples diferença de mudanças entre dois quadros consecutivos de vídeo $f(x,y,t_i)$ e $f(x,y,t_j)$, obtidos nos instantes de tempo t_i e t_j , respectivamente, conforme equação (2.1):

$$d_{ij}(x, y) = \begin{cases} 1, & \text{para } |f(x, y, t_i) - f(x, y, t_j)| < \theta \\ 0, & \text{caso contrário} \end{cases} \quad (2.1)$$

Na qual o θ é o limiar utilizado para definir se as mudanças foram ou não significativas numa dada janela de análise.

Os primeiros resultados com este algoritmo podem ser observados em Jorge et al. (2006).

2.2.1.2 Modelagem por Mistura de Gaussianas

A Modelagem por Mistura de Gaussianas (GMM) é a modelagem da cor do pixel como mistura de funções gaussianas, cujo objetivo maior é o de detectar variações de iluminação, movimentos na cena e múltiplos objetos. Pixels que não pertencem à distribuição de fundo, identificados por uma função distância, são considerados parte do objeto em movimento.

Seja a probabilidade P do pixel observado X no tempo t , definida por (2.2):

$$P(X_t) = \sum_{i=1}^K w_{i,t} * \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (2.2)$$

Sendo K o número de distribuições, $w_{i,t}$ é a estimativa do peso (porção do dado que é aplicado pela Gaussiana) na i -ésima mistura Gaussiana no tempo t , $\mu_{i,t}$ é o valor médio da i -ésima mistura Gaussiana no tempo t , $\Sigma_{i,t}$ é matriz de covariância na i -ésima Gaussiana no tempo t , sendo η a função Gaussiana de densidade de probabilidade descrita em (2.3):

$$\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}(X_t - \mu)^T \Sigma^{-1} (X_t - \mu)} \quad (2.3)$$

Sendo K determinado pela memória e poder computacional.

Os parâmetros da distribuição que recebem os novos valores observados são atualizados pelas equações (2.4 a 2.6):

$$\mu_t = (1 - \rho)\mu_{t-1} + \rho X_t \quad (2.4)$$

$$\sigma_t^2 = (1 - \rho)\sigma_{t-1}^2 + \rho(X_t - \mu_t)^T (X_t - \mu_t) \quad (2.5)$$

$$\rho = \alpha \eta(X_t | \mu_k, \sigma_k) \quad (2.6)$$

Sendo a correção de pesos das k distribuições no tempo t , dada pela equação (2.7), sendo α a taxa de aprendizado na atualização e $M_{k,t}$, 1 quando for o modelo desejado e 0 quando não :

$$w_{k,t} = (1-\alpha)w_{k,t-1} + \alpha (M_{k,t}) \quad (2.7)$$

Mais detalhes podem ser vistos em Stauffer e Grimsin (2000).

2.2.1.3 Fluxo Ótico

Fluxo ótico é a técnica que mede a velocidade dos pixels de uma imagem comparando os mesmos com o quadro anterior. O deslocamento destes pixels ao longo do tempo por ser usado para estimar o movimento da câmera em aplicações de rastreamento.

Existe uma grande variedade de algoritmos que programam o fluxo ótico, sendo que alguns deles marcaram alguns momentos importantes conforme indica a linha do tempo apresentada na Figura 2.1. Os mais famosos são ainda as técnicas de Lucas e Kanade (1981), de Horn e Shunk (1981), precursores do método do fluxo ótico. Depois disso, Barron et al. (1994) e Galvin et al. (1998) apresentaram comparações completas com os algoritmos que mais se destacam.

Figura 2.1 - Linha do tempo de trabalho com Fluxo Ótico

Dado quadros consecutivos de vídeo, pode-se estimar o movimento através do algoritmo do Fluxo Ótico proposto inicialmente por Lucas e Kanade (1981).

Assim, considerando o brilho constante ao longo de cada quadro no tempo, pequenas variações entre quadros, ou seja, o pixel não se movendo para posições muito distantes e para movimentos coerentes com os vizinhos, podemos definir que a intensidade de um pixel (x,y) no instante $t-1$ da imagem $I(x,y, t-1)$, se desloca na velocidade (u,v) para a posição $(x+u, y+v)$ na imagem $I(x,y,t)$ no tempo t .

Segundo a equação (2.8),

$$I(x, y, t-1) = I(x + u(x, y), y + v(x, y), t) \quad (2.8)$$

Sendo I_x a derivada parcial de I em x e I_y a derivada parcial de I em y , aplicando Taylor pode-se reescrever a equação (2.8), obtendo a equação (2.9):

$$I(x, y, t-1) \approx I(x, y, t) + I_x \cdot u(x, y) + I_y \cdot v(x, y) \quad (2.9)$$

A equação do fluxo ótico pode ser reescrita como sendo (2.10):

$$I_x \cdot u + I_y \cdot v + I_t \approx 0 \quad (2.10)$$

Ou ainda na forma de gradiente (2.11):

$$\nabla I \cdot (u, v) + I_t = 0 \quad (2.11)$$

Considerando que os pixels vizinhos mantêm o movimento coerente, pode-se procurar a solução para o problema da abertura, utilizando, por exemplo, algumas janelas, ou seja, 5x5 pixels, obtém-se 25 equações (2.12) para o pixel p_i :

$$0 = I_t(p_i) + \nabla I(p_i) \cdot [u \ v] \quad (2.12)$$

Da equação (2.12) vem então (2.13):

$$\begin{bmatrix} I_x(p_1) & I_y(p_1) \\ I_x(p_2) & I_y(p_2) \\ \vdots & \vdots \\ I_x(p_{25}) & I_y(p_{25}) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(p_1) \\ I_t(p_2) \\ \vdots \\ I_t(p_{25}) \end{bmatrix} \quad (2.13)$$

Reescrevendo na forma matricial, tem-se:

$$\begin{matrix} A & d & = & b \\ 25 \times 2 & 2 \times 1 & & 25 \times 1 \end{matrix} \quad (2.14)$$

Esta solução leva a mais equações que variáveis desconhecidas. Então, deve-se minimizar $\|Ad-b\|$. Utilizando mínimos quadrados, pode-se dizer que:

$$(A^T A)d = A^T b \quad (2.15)$$

Logo:

$$\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix} \quad (2.16)$$

Em (2.16) a soma representa a soma de todos os pixels na janela $k \times k$.

A matriz de somas é a matriz de $A^T A$ que deve ter inversa, e não pode ser muito pequena, ou seja, os autovalores não podem ser muito pequenos. A razão entre os autovalores também não devem ser muito grandes.

2.2.1.4 Detector de Harris

O detector de cantos de Harris (HARRIS; STEPHENS, 1988) está baseado na matriz de covariância e nos seus autovalores. Quando a matriz possui dois autovalores altos, corresponde a duas direções principais diferentes no gradiente da imagem subjacente. Este detector é calculado rapidamente e eficientemente com uma simples equação (2.17), sendo a resposta guardada para cada pixel da imagem.

Formalmente, seja a função de estimativa de movimento $E(u,v)$, para um deslocamento $[u,v]$, de um ponto (x,y) da imagem $I(x,y)$, definida em (2.17):

$$E(u,v) = \sum_{x,y} w(x,y) [I(x+u,y+v) - I(x,y)]^2 \quad (2.17)$$

Sendo $w(x,y)$ a função peso centralizada em (x,y) (Figura 2.2) e $I(x+u,y+v)$ a imagem deslocada de $[u,v]$.

Figura 2.2 - Função peso $W(x,y)$. (Fonte: SIMAKOV; FROLOVA, 2004)

No caso de $W(x,y)$ gaussiana, tem-se que:

$$G(x,y,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.18)$$

Para pequenos deslocamentos $[u,v]$, pode-se fazer a seguinte aproximação bilinear por expansão de Taylor:

$$E(u,v) \cong [u,v] M \begin{bmatrix} u \\ v \end{bmatrix} \quad (2.19)$$

Sendo M a matriz de covariação 2×2 da derivada da imagem, definida em (2.20).

$$M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (2.20)$$

Pela análise de autovalores de M , determina-se os maiores autovalores λ_1 e λ_2 , maior e menor mudanças de $E(u,v)$ (Figura 2.3).

Figura 2.3 - Elipse $E(u,v) = \text{constante}$

Dependendo dos valores de λ_1 e λ_2 pode-se determinar se o ponto é canto ou não, conforme ilustrado na Figura 2.4.

Figura 2.4 - Representação de $\lambda_2 \times \lambda_1$ mostrando como podem ser detectados os pontos de cantos numa imagem (Fonte: SIMAKOV; FROLOVA, 2004)

Aplicando-se (2.21),

$$R = \det M - k(\text{trace} M)^2 \quad (2.21)$$

Sendo:

$$\det M = \lambda_1 \lambda_2 \quad (2.22)$$

$$\text{trace} M = \lambda_1 + \lambda_2 \quad (2.23)$$

No qual, $k = 0.04-0.06$ é uma constante determinada empiricamente.

O algoritmo de Harris então se resume como:

- 1) *Encontre todos os cantos com $R > \text{limiar}$;*
- 2) *Determine os pontos com máximo local de R .*

Na Figura 2.5, pode-se verificar a resposta do algoritmo em função de R .

Figura 2.5 - Determinação de Cantos em função de R
(Fonte: SIMAKOV; FROLOVA, 2004)

2.2.1.5 SIFT (*Scale Invariant Feature Transform*)

A técnica SIFT foi desenvolvida em 1999 por David G. Lowe, da Universidade Britânica de Columbia – Vancouver (Canadá) (LOWE, 2004).

Esta transformada detecta pontos característicos locais da imagem. As características extraídas da SIFT são invariantes à rotação, à escala e parcialmente invariantes à mudança de iluminação e às pequenas mudanças de perspectivas. São necessárias pelo menos três características corretamente pareadas de cada objeto para uma identificação com confiança.

A primeira etapa da técnica identifica o espaço e as escalas que se mantêm fixos mesmo com diferentes pontos de vista do mesmo objeto. Detectar locais que são invariantes à mudança de escala de uma imagem pode ser efetuado pesquisando por características estáveis através de todas as possíveis escalas, usando uma função contínua de escala, denominada espaço escala (LOWE, 2004).

O espaço escala de uma imagem é definido como uma função $L(x,y,\sigma)$ que é produzida a partir da convolução de uma gaussiana $G(x,y,\sigma)$, com a imagem $I(x,y)$, equação (2.24) :

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \quad (2.24)$$

onde $*$ é a operação de convolução em x e y , e:

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (2.25)$$

Sendo σ o desvio padrão, representando as diferentes escalas $k\sigma$.

Para a detecção dos pontos característicos determina-se primeiramente a função $D(x,y,\sigma)$ a partir da convolução da função da Diferença de Gaussianas (DoG) à imagem $I(x,y)$. A gaussiana DoG pode ser calculada através da diferença de Gaussianas de duas escalas próximas, separadas por uma constante de fator multiplicativo k :

$$D(x,y,\sigma) = (G(x,y,K\sigma) - G(x,y,\sigma)) * I(x,y) = L(x,y,k\sigma) - L(x,y,\sigma) \quad (2.26)$$

A aplicação da DoG é realizada conforme a Figura 2.6.

Figura 2.6 - Aplicação da DoG no espaço imagem e escala

A imagem inicial sofre convoluções incrementais com gaussianas para produzir imagens separadas por um fator constante k no espaço escala. Este processo é chamado de uma oitava e é repetido para um número repetido de oitavas. Variando-se σ , cria-se oitava de imagens com diversos níveis cada um correspondendo a uma escala, ou seja, a um desvio padrão do filtro gaussiano diferente. A imagem com 2σ é reduzida em dimensão, gerando uma nova imagem com a metade da resolução de seu tamanho original, sendo a primeira imagem da próxima oitava. Este processo se

repete várias vezes gerando novas oitavas. A cada oitava, grupo imagens com seguidas convoluções no espaço escala é dividido por um número inteiro s de intervalos, então $k = 2^{1/s}$, (k = uma constante que aparece em todas as escalas) produzindo assim $s+3$ imagens em cada oitava. Imagens em escalas adjacentes são subtraídas para produzir as imagens da *DoG*. Uma vez processado a oitava, é reduzida a amostragem da imagem com duas vezes o valor inicial de σ . A taxa de amostragem σ não é diferente à dada inicialmente, mas o esforço computacional é gradativamente reduzido.

Os *keypoints* são localizados nos máximos e mínimos locais de cada imagem $D(x, y, \sigma)$, que podem ser obtidos comparando-se um pixel (marcado com X) com seus 8 vizinhos na região 3×3 corrente e 9 vizinhos nas escalas adjacentes (acima e abaixo, variando σ), como visto na figura 2.7.

Figura 2.7 – Máximos e mínimos da Diferença de Gauss. (Fonte: LOWE, 2004).

Uma vez que um *keypoint* candidato tenha sido encontrado por comparação de um pixel com seus vizinhos, o próximo passo é detalhar os dados para localização, escala e a taxa das principais curvas. Isto permite que pontos sejam rejeitados por possuir baixo contraste (e sensíveis a ruídos) ou são pouco localizados ao longo de uma borda. Para esta eliminação usa-se a expansão de *Taylor* (equação 2.27) sobre a função espaço escala $D(x, y, \sigma)$, deslocada de modo a ter sua origem no ponto de amostragem.

$$D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \quad (2.27)$$

Onde, D e suas derivadas são calculadas no ponto de amostragem e $x = (x, y, \sigma)^T$ é o deslocamento deste ponto. A localização do extremo $\hat{x}$ é dada pela derivada desta função (2.27) com o respectivo x em 0, resultando em:

$$\hat{x} = -\frac{\partial^2 D^{-1} \partial D}{\partial x^2 \partial x} \quad (2.28)$$

O valor da função no extremo $D(\hat{x})$ é útil para rejeição de extremos instáveis com baixo contraste. Isto pode ser obtido se substituir a equação 2.28 em 2.29, sendo:

$$D(\hat{x}) = D + \frac{1}{2} \frac{\partial D^T}{\partial x} \hat{x} \quad (2.29)$$

A eliminação de *keypoints* próximos da borda é feita usando-se uma matriz Hessiana H (2.30) na localização e escala do *keypoint*:

$$H(x, y) = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (2.30)$$

As derivadas são estimadas pelas diferenças dos pontos de amostragem da vizinhança. Os autovalores de H são proporcionais a principal curvatura de D . A Hessiana representa uma segunda derivada, permitindo mensurar o tamanho das curvaturas de D em seus autovalores. A curvatura próxima a pontos cujo gradiente é paralelo às extremidades possui baixa amplitude, mas em pontos com gradiente perpendicular a essas extremidades a amplitude é elevada.

Sendo λ_1 o autovalor com maior magnitude e λ_2 o de menor magnitude, pode-se computar a soma e o produto destes autovalores:

$$\begin{aligned} Tr(H) &= D_{xx} + D_{yy} = \lambda_2 + \lambda_1, \\ Det(H) &= D_{xx}D_{yy} - (D_{xy})^2 = \lambda_2 \lambda_1. \end{aligned} \quad (2.31)$$

Caso o determinante seja negativo, as curvaturas têm sinais diferentes e o ponto é descartado, não sendo um extremo. Sendo r a razão entre o maior e o menor autovalor ($\lambda_1 = r \lambda_2$). Então,

$$\frac{Tr(H)^2}{Det(H)} = \frac{(\lambda_1 + \lambda_2)^2}{\lambda_1 \lambda_2} = \frac{(r\lambda_2 + \lambda_2)^2}{r\lambda_2^2} = \frac{(r+1)^2}{r} \quad (2.32)$$

A equação (2.32) depende apenas da razão entre os autovalores. A razão $(r+1)^2/r$ oferece uma medida do quanto os autovalores são distintos. Eliminam-se

pontos próximos a extremidades descartando-se pontos abaixo de um determinado limiar (r):

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r} \quad (2.33)$$

Assim, eliminam-se *keypoints* que não são estáveis, apesar de estarem próximos de extremidades.

A escala do *keypoint* é usada para selecionar a Gaussiana da imagem suavizada L , com a escala aproximada, somente então, todos os cálculos são feitos em uma escala invariante. Para cada amostra da imagem $L(x,y)$, em sua escala, a magnitude do gradiente $m(x,y)$, e orientação $\theta(x,y)$ calculados, usando a diferença dos pixels:

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (2.34)$$

$$\theta(x,y) = \tan^{-1}((L(x,y+1) - L(x,y-1)) / (L(x+1,y) - L(x-1,y))) \quad (2.35)$$

Um histograma das orientações é formado a partir das orientações dos gradientes dos pontos amostrados na região vizinha ao *keypoint*. Esse histograma possui 36 regiões, cobrindo os 360° de orientações.

Picos de orientação do histograma correspondem a direções dominantes para os gradientes locais. O maior pico do histograma descoberto e outros picos que correspondam acima de 80% do valor do maior pico serão usados para também criar um *keypoint* com aquela orientação. Então, para locais de múltiplos picos de magnitude semelhante, são criados diferentes *keypoints* na mesma localização, mas com diferentes orientações. Somente aproximadamente 15% dos pontos são apontados para as múltiplas orientações, entretanto eles contribuem significativamente à estabilidade de comparação. Finalmente, para uma boa interpretação dos dados, um gráfico, contendo uma parábola deve ser usado para interpolar os três valores do histograma mais próximos do pico, de forma a se ter uma melhor precisão de sua posição.

Finalmente o descritor do *keypoint* é criado computando-se as magnitudes e orientações dos gradientes que são amostradas ao redor da localização do *keypoint* e, utiliza a escala do *keypoint* para escolher o nível de suavização gaussiana para a imagem. Para alcançar invariância de orientação, as coordenadas e a orientação dos

gradientes do descritor são giradas em relação à orientação do *keypoint*. Para maior eficiência, os gradientes são computados para todos os níveis da pirâmide e este procedimento é mostrado pelas pequenas setas em cada amostra da localização na Figura 2.8.

Figura 2.8 – Gradiente da Imagem. (Fonte: LOWE, 2004)

Uma função de suavização gaussiana σ , igual à metade da largura da janela do descritor é usada para dar pesos à magnitude do gradiente em cada ponto amostrado. Este passo é mostrado como uma janela circular, na Figura 2.8.

Uma vez realizada a suavização dos gradientes, criam-se histogramas de orientação em 4x4 regiões de amostragem. Na figura 2.9 estão representadas oito direções para cada histograma e o tamanho de cada seta representa a magnitude referente a respectiva entrada do histograma.

Figura 2.9 – Descritor do *keypoint*. (Fonte: LOWE, 2004)

Para evitar efeitos de borda na qual o descritor abruptamente muda enquanto a amostra se desloca suavemente indo de um histograma para outro ou de uma orientação para outra, uma interpolação tri linear é usada para distribuir o valor de cada amostra de gradiente para a sua posição no histograma adjacente. O descritor é formado de um vetor contendo os valores de todas as orientações dos histogramas correspondentes aos tamanhos das setas representando um vetor de características com 128 elementos (LOWE, 2004).

O vetor é modificado para reduzir os efeitos de mudança de iluminação. O primeiro passo é normalizar o vetor. Em seguida, tem-se uma mudança no contraste da imagem. Cada valor do pixel será multiplicado por uma constante e os gradientes também serão multiplicados por essa mesma constante, de modo que a alteração será

cancelada pelo vetor normalizado. Uma mudança de brilho na qual uma constante é adicionada a cada pixel da imagem não irá afetar os valores dos gradientes, pois eles são computados através da diferença de pixels. Portanto, o descritor é invariante a mudanças de iluminação.

Entretanto, mudanças de iluminação não-linear podem ocorrer por saturação das câmeras ou por mudanças de iluminação que afetam superfícies 3D. Estes efeitos podem causar uma grande mudança em alguns gradientes, mas afetam menos as orientações dos gradientes. Reduzindo a influência das grandes magnitudes pode-se estabelecer que valores acima de 0,2 sejam ajustados para 0,2 (LOWE, 2004); isto significa que a correspondência para as grandes magnitudes dos gradientes já não são tão importantes, e que a distribuição das orientações tem maior destaque.

Determinando a correspondência dos pontos

Geralmente, os métodos de estimação procuram utilizar o maior número de pontos para uma estimativa inicial e então, ir eliminando os pontos inválidos. O método RANSAC (*Random Sample Consensus*) (FISCHLER; BOLLES, 1981), ao contrário desses métodos, diferencia-se pelo fato de utilizar apenas o número mínimo e suficiente de pontos necessários para uma primeira estimativa, aumentando o conjunto com novos pontos sempre que possível.

Suponha um conjunto de P pontos de amostras para se instanciar um modelo que requer o mínimo de n pontos, tal que $P > n$. O método de RANSAC pode ser descrito por:

1. Seleciona-se aleatoriamente um conjunto S_1 formado por n pontos a partir do conjunto P para uma primeira estimativa do modelo M_1 ;
2. Usa-se o modelo M_1 para determinar o novo conjunto S_1 de pontos de P que se aproximam desse modelo com erros dentro de certa tolerância;
3. Se o número de pontos em S_1 é maior que um certo limite t , que é função da estimativa do nível de erros em P , usa-se S_1 para uma nova e melhor estimativa M_1 do modelo, usando um método de ajuste como por exemplo, Mínimos Quadrados;
4. Se o número de pontos em S_1 é menor que o limite t , seleciona-se aleatoriamente um novo conjunto S_2 de n pontos e repete-se o procedimento descrito;
5. Se depois de certo número de tentativas, não se obtiver um conjunto S_i de t ou mais pontos, pode-se fazer uma estimativa final do modelo com o maior conjunto encontrado, ou terminar o procedimento indicando-se erro.

Esse método requer que três parâmetros sejam especificados:

- A tolerância do erro que é utilizada para definir se um ponto é ou não compatível com o modelo instanciado;
- O número de tentativas do procedimento;
- O número t que representa o número mínimo de pontos que um conjunto S_i deve ter para que o modelo estimado a partir dos seus pontos seja considerado correto.

RANSAC é um procedimento de estimação robusto que usa o conjunto mínimo das correspondências amostradas randomicamente para estimar os parâmetros de transformações das imagens e achar a solução que tem o melhor consenso com os dados. Para o caso dos panoramas devem ser selecionados conjuntos de r correspondências de características e calcular a homográfica H entre eles utilizando o método da transformação linear direta (DLT). Isto deve ser repetido n vezes e após, seleciona-se a solução que tem o maior número de *inliers*.

A probabilidade que uma característica correspondente está de acordo entre um par de imagens correspondentes (a probabilidade *inlier*) é P_i e a probabilidade de achar a transformação correta após n tentativas é:

$$P(H \text{ é correto}) = 1 - (1 - (P_i)^r)^n \quad (2.36)$$

Após uma grande quantidade de tentativas, a probabilidade de achar a correta homográfica é grande.

A matriz homográfica mapeia uma transformação do plano projetivo real para o plano projetivo, fazendo as correções geradas por essa transformação. A matriz homográfica possui oito graus de liberdade no que diz respeito a isometria, transformação de similaridade, transformação afim, transformação projetiva e projeção em perspectiva.

Outra forma de obter a correspondência entre os pontos encontrados é através de simples métricas de distâncias, sendo a euclidiana a mais usada.

2.3 Considerações Finais

A evolução da tecnológica tem propiciado o surgimento de inúmeros algoritmos de detecção de movimento. Foram apresentados neste capítulo os algoritmos estudados para este trabalho, evidenciando que os métodos diretos, como o fluxo ótico, não operam adequadamente em imagens com variações de brilho, os movimentos são detectados sem a necessidade de nenhum outro parâmetro

intermediário ser determinado, apresenta densos campos, mas são mais sensíveis à variações na aparência geral. Muito útil para movimentos em regiões menores que 10 pixels. A técnica SIFT, baseada nos atributos extraídos (textura e cantos), é robusta apesar de muitos campos esparsos e a mistura de Gaussianas, pode funcionar adequadamente na aplicação de detecção de insetos.

Por fim, no próximo capítulo será apresentada uma revisão da técnica Wavelet-Multifractal para extração de atributos de comportamento por som.

Capítulo 3

Wavelet-Multifractal

3.1 Considerações Iniciais

No capítulo anterior foi apresentada uma revisão de literatura para os métodos de detecção de movimento visando o rastreamento dos insetos. Neste capítulo, é feita uma revisão da técnica Wavelet-Multifractal, para posterior aplicação no estudo comportamental dos insetos.

As ferramentas para estudo de sistemas dinâmicos não lineares, como a teoria fractal, a multifractal e a junção da transformada Wavelet com o multifractal são utilizadas para a caracterização, modelagem e extração de características de fenômenos complexos. Será feita neste capítulo uma breve introdução à Transformada Wavelet e um sucinto estudo do formalismo multifractal, tendo como objetivo principal definir um sinal fractal ou multifractal, evidenciando a sua importância na análise de sinais como os sons dos insetos, que representam sistemas não lineares, de caráter estocástico, além de apresentar singularidades nos seus espectros através do tempo (ARNEODO et al., 1995; McATEER et al., 2007; RELJIN et al., 2008).

Uma das formas de se aplicar a técnica de análise multifractal é obter o Módulo Máximo da Transformada Wavelet (MMTW), que fornece o espectro de singularidades de uma série temporal. O espectro de singularidades pode ser interpretado como uma "assinatura" de uma série temporal, sendo muito importante na caracterização da mesma.

Uma série temporal é um conjunto de observações de uma variável dispostas sequencialmente no tempo e pode ser classificada como determinística ou estocástica. Será determinística se a série é estacionária, ou seja, permanece relativamente independente no tempo. Já os sinais não estacionários ou estocásticos trazem variações substanciais e significativas de padrão e comportamento, dependendo do instante de tempo considerado.

A teoria Wavelet-Multifractal permite a caracterização de fenômenos complexos de forma quantitativa, tanto nas variações temporais como espaciais. Os conceitos da técnica multifractal são cada vez mais reconhecidos, sendo considerada a ferramenta mais apropriada e potente para a análise e simulação dos fenômenos complexos em uma grande variedade de escalas (SCHERTZER; LOVEJOY, 1994), onde as aplicações são muito amplas e os resultados efetivos (IVANOV et al., 1999; YU et al., 2001; LATKA et al., 2002; POSADAS et al., 2003, POSADAS et al., 2005; McATEER et al., 2007).

Neste capítulo serão apresentadas a Transformada Wavelet Contínua (CWT) e a Transformada Wavelet Discreta (DWT), o Módulo Máximo da Transformada Wavelet (MMTW) bem como as teorias fractal e multifractal.

3.2 A Transformada *Wavelet* (TW)

A grande maioria de dados em séries temporais, como no caso dos sons, representa sistemas não lineares, de caráter estocástico, apresentando singularidades nos seus espectros através do tempo.

Características típicas do comportamento singular podem ser observadas na Figura 3.1. Os tipos “escalão” (b1) e “chapéu, *tipo delta de Dirac*” (c1), são denominados de *singularidades* do sinal. Na Figura 3.1(b), também se ilustra um fator complicado e muito comum neste tipo de sinal, *singularidades* não isoladas, ou seja, que aparecem muito próximas entre si, fazendo sua caracterização ainda mais complexa.

As ferramentas clássicas como a série de Taylor ou as séries de *Fourier* são incapazes de descrever o comportamento singular. Uma das grandes desvantagens da análise de *Fourier* (espectro) provém do fato de apresentar apenas resolução na frequência e não no tempo. Isto significa que, embora capaz de determinar o conteúdo de frequências presentes em um sinal, não há noção de quando (em que intervalo de tempo) elas ocorrem.

A informação em frequência de um sinal calculado pela transformada clássica de *Fourier* é a média de todo o tempo de duração do sinal. Assim, se existe um sinal transitório local durante algum intervalo de tempo pequeno, o transitório contribuirá para a Transformada de *Fourier* (TF), mas a sua localização no eixo do tempo será

perdida. A análise tradicional de *Fourier* não considera frequências que variam com o tempo, isto é, sinais não estacionários.

Figura 3.1 - Sinais exemplos de séries de tempo: (a) um ruído gerado numericamente; (b) eletrocardiograma, (c) ruído gerado numericamente (Fonte: <<http://www.physionet.org/tutorials/multifractal/behavior.htm>> Acesso dia 20/03/2009).

Com o objetivo de corrigir a deficiência mencionada da TF, objetivando ter uma análise, ao mesmo tempo, temporal e frequencial de sinais não estacionários Dennis Gabor, desenvolveu o Princípio da Incerteza de Heisenberg (tempo x frequência) e adaptou a TF para analisar apenas uma pequena extensão do tempo do sinal, originando a chamada Transformada de *Fourier* de Tempo Curto (*Short-Time Fourier Transform* - STFT), também conhecida como Transformada de *Fourier* Janelada (TFJ) ou Transformada de Gabor (MISITI et al., 2002).

A desvantagem da Gabor é a limitação da largura da janela a qual precisa ser fixada. Com uma janela larga, por exemplo, obtém-se uma boa resolução em frequência, mas uma pobre resolução no tempo, enquanto que com uma janela estreita será obtida uma boa resolução no tempo, mas uma pobre resolução em frequência.

A TW supera as limitações do método de *Fourier* através do emprego de funções que são locais no tempo e em frequência. Em particular, a habilidade das *wavelets* em localizar curtos intervalos de tempo para componentes de alta frequência

e longos intervalos de tempo para componentes de baixa frequência, melhora a análise de sinais com impulsos e oscilações, particularmente na presença de harmônicos fundamentais e de baixa ordem.

A análise *wavelet* apesar dos seus fundamentos matemáticos apontarem para o trabalho de Joseph *Fourier* no século XIX, é relativamente recente. Haar, em 1909, fez o primeiro registro do uso de *wavelets*. As *wavelets* de Haar possuem a característica de terem suportes compactos, porém não são continuamente diferenciáveis, o que limitou o seu uso. Em 1930, alguns autores utilizaram *wavelets* de Haar com funções de base com escala variável. *Wavelets* a partir de 1985, pelo trabalho de Stephane Mallat na área de processamento digital de sinais, tomaram um impulso quando Meyer construiu a primeira *wavelet* não-trivial. As *wavelets* de Meyer são continuamente diferenciáveis e não possuem suporte compacto, ao contrário da *wavelet* de Haar. Mais tarde, Ingrid Daubechies utilizou o trabalho de Stephane Mallat na construção de um conjunto de *wavelets* com função de base ortonormais suaves, com suportes compactos. Os trabalhos de Ingrid Daubechies foram os alicerces das aplicações atuais de *wavelets*. (MISITI et al., 2002).

A Transformada *Wavelet* (TW) consiste em uma técnica de “janelamento” variável. A TW permite o uso de uma janela de tempo maior para analisar informações de baixa frequência, de forma mais precisa, e de uma pequena janela para informações de alta frequência (MISITI ET AL., 2002).

De acordo com Daubechies (1992), define-se a família das funções *wavelets* $\psi_{a,b}$, como na equação 3.1:

$$\psi_{a,b}(t) = |a|^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right) \quad a, b \in \mathfrak{R}, \quad a \neq 0, \quad (3.1)$$

Sendo que:

t representa a variável de tempo ou espaço, gerada a partir das operações de expansão (fator de escala a) e translação (fator b) da mesma função complexa ψ , que é a *wavelet*-mãe. Na Figura 3.2 podem ser observadas diferentes funções *Wavelet*-mães utilizadas.

Figura 3.2 - Exemplos de funções *Wavelet*-mãe

A TW pode ser encontrada em duas versões diferentes, a contínua e a discreta:

a) Transformada *Wavelet* Contínua (TWC): faz o mapeamento de um sinal original unidimensional no domínio do tempo, para uma nova função, em um espaço bidimensional, através dos fatores de escala a e translação b pela TW.

b) Transformada *Wavelet* Discreta (TWD): faz a decomposição de um sinal discretizado em diferentes níveis de resolução e realiza o mapeamento de uma sequência de números em outra sequência de números.

3.2.1 A Transformada *Wavelet* Contínua (TWC)

A análise *wavelet* emprega um protótipo de função chamado *wavelet*-mãe que tem média zero e parte central oscilante, ou seja, decai para zero em ambos os lados de sua trajetória. Matematicamente, a TWC de um dado sinal $x(t)$ em relação à *wavelet*-mãe $\psi(t)$ é definida por:

$$TWC(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt \quad (3.2)$$

sendo:

a a dilatação ou fator de escala, b o fator de translação, e essas variáveis são contínuas.

A equação (3.2) mostra que o sinal unidimensional original no domínio do tempo $x(t)$ é mapeado para uma nova função, em um espaço bidimensional, por meio dos fatores de escala a e translação b pela TW. Um coeficiente da TW, em uma

determinada escala e translação – $TWC(a,b)$, representa quão bem o sinal original $x(t)$ e a *wavelet*-mãe escalada e transladada se combinam. Dessa forma, o conjunto de todos os coeficientes $TWC(a,b)$ associados a um sinal particular são a representação *wavelet* do sinal original $x(t)$ em relação à *wavelet*-mãe $\psi(t)$.

Pode-se visualizar a *wavelet*-mãe como uma função janela. O fator de escala a e o tamanho da função janela são interdependentes, desse fato decorre que menores escalas implicam em menores janelas. Consequentemente, podem-se analisar componentes de bandas estreitas de frequência de um sinal com um pequeno fator de escala, e componentes de bandas largas de frequência com fatores de escala maiores, o que permite capturar todas as características de um sinal particular (ADDISON, 2002).

Escalamento

A análise *wavelet* produz um escalamento de um sinal no tempo.

Escalar uma *wavelet* significa simplesmente expandí-la ou comprimí-la. O fator de escala, usualmente denotado pela letra a , é utilizado para representar esses efeitos. Por exemplo, o efeito do fator de escala é facilmente observado, como é ilustrado na Figura 3.3.

Figura 1.3 - Escalamento de uma função *Wavelet*-mãe

A Figura 3.3 mostra que o fator de escala a está relacionado de maneira inversa à frequência do sinal em radianos w . O que corresponde a:

- baixa escala $a \rightarrow$ *wavelet* comprimida $\rightarrow$ detalhes que mudam rapidamente $\rightarrow$ alta frequência w .
- alta escala $a \rightarrow$ *wavelet* expandida $\rightarrow$ detalhes que mudam vagarosamente $\rightarrow$ baixa frequência w .

Translação

Transladar uma *wavelet* significa simplesmente atrasá-la ou adiantá-la em relação ao eixo das abscissas. Matematicamente, atrasar uma função $f(x)$ de um fator b , consiste na expressão $f(x-b)$, conforme mostrado na Figura 3.4.

Figura 3.4 - Translação aplicada a uma *wavelet*

Cálculo da Transformada *Wavelet* Contínua

A TWC é a integral em todo o intervalo de definição do sinal, multiplicada pelas versões escaladas e transladadas da *wavelet*-mãe. Esse processo produz coeficientes *wavelet* que são função da escala e posição. A seguir são ilustrados cinco passos para criar a TWC (MISITI et al., 2002):

1. alinhar uma *wavelet*-mãe e compará-la com uma seção inicial do sinal original.
2. calcular o coeficiente C , o qual representa quão intimamente está relacionada a *wavelet*-mãe com a seção do sinal, em que $C = TWC(a,b)$. Entre os coeficientes calculados, os maiores valores de C significam uma maior similaridade do sinal com a *wavelet*. Mais precisamente, se a energia do sinal e a energia da *wavelet* são iguais a um, o coeficiente C pode ser interpretado como um coeficiente de correlação. Note que o resultado dependerá da forma da *wavelet*-mãe escolhida, como na Figura 3.5.

Figura 3.5 - Cálculo do fator C de relação entre a *wavelet*-mãe e o sinal

3. transladar a *wavelet*-mãe para a direita, como mostrado na Figura 3.6, e repetir os passos 1 e 2 até que se tenha coberto todo o sinal.

Figura 3.6 - Translação da função *wavelet*-mãe

4. escalar (expandir) a *wavelet*-mãe, como ilustrado na Figura 3.7, e repetir os passos de 1 a 3.

Figura 3.7 - Escalamento da função wavelet-mãe

5. repetir os passos de 1 a 4 para todas as escalas.

Quando o processo terminar, têm-se os coeficientes da TWC produzidos em diferentes escalas, por diferentes seções do sinal (MISITI et al., 2002).

Em vez de interpretar os parâmetros nos domínios tempo e frequência ($f \times t$), costuma-se utilizar os domínios escala e deslocamento ($a \times b$). Assim, o resultado da transformada *Wavelet* gera uma imagem tipo “mapa”, também denominado de escalograma, conforme apresentado na Figura 3.8.

Um exemplo deste mapa gerado para o sinal do inseto macho pode ser visto na Figura 3.9. Na Figura 3.9(a) é observado o sinal ao longo do tempo e na Figura 3.9(b) a respectiva TWC aplicada com uma *Wavelet*-mãe tipo Paul (ADDISON, 2002) (Figura 3.10).

Figura 3.8 - Representação do espectro *Wavelet*, ou escalograma, resultante com todos os coeficientes calculados. (Fonte: ADDISON, 2002)

Figura 3.9 - (a) Sinal de som de um inseto macho e (b) o escalograma da TWC do sinal (a) em aplicando a *Wavelet*-mãe da Figura 3.10, obtidos pelo programa desenvolvido neste trabalho.

Figura 3.10 - Função *Wavelet*-mãe de Morlet (MISITI et al., 2002)

Propriedades Transformada *Wavelet* Contínua

As propriedades mais importantes das *wavelets* são as condições de admissibilidade e oscilatória (ADDISON, 2002).

Funções que atendem à condição de admissibilidade podem ser utilizadas para analisar e reconstruir um sinal sem a perda de informação. Ou seja:

$$C_{\Psi} = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < +\infty \quad (3.3)$$

Esta condição implica que TF de $\Psi(t)$, $\hat{\Psi}(\omega)$, se anula quando a frequência é zero, ou seja:

$$|\hat{\Psi}(\omega)|^2 \Big|_{\omega=0} = 0 \quad (3.4)$$

A Transformada *Fourier* nula em $\omega = 0$ implica que, no tempo, o valor médio na *wavelet* é nulo, ou melhor, é oscilatória, logo:

$$\int_{-\infty}^{\infty} \Psi(t) dt = 0 \quad (3.5)$$

3.2.2 A Transformada *Wavelet* Discreta

A TF pode ser encontrada em duas versões diferentes, a contínua (TCF) e a discreta (TDF). De maneira análoga, a TWC tem uma versão digitalmente

implementável, chamada de Transformada *Wavelet* Discreta (TWD) que é definida na equação (3.6):

$$TWD(m, k) = \frac{1}{\sqrt{a_0^m}} \sum_n x(n) \psi\left(\frac{k - nb_0 a_0^m}{a_0^m}\right) \quad (3.6)$$

sendo:

$\psi(\cdot)$ a *wavelet*-mãe e os parâmetros de escala e de translação a e b são funções de um parâmetro inteiro m , isto é, $a = a_0^m$ e $b = nb_0 a_0^m$, que permite uma expansão da família originada pela *wavelet*-mãe, gerando as *wavelets* filhas. Na equação (3.6), k é uma variável inteira que se refere a um número particular de amostras de um determinado sinal de entrada. O parâmetro de escala permite o aumento da escala geométrica, isto é, $1, 1/a_0, 1/a_0^2, \dots$

A saída da TWD pode ser representada em um espaço bidimensional de maneira semelhante à da Transformada Discreta de *Fourier* Janelada (TDFJ), mas com divisões muito diferentes no tempo e na frequência. A análise da TWD produz bandas de frequência de tempo retangulares, as quais são estreitas nas componentes de alta frequência e largas nas componentes de baixa frequência (GRAPS, 1995).

Filtragem e Análise Multirresolução

Filtrar consiste em permitir passar algumas componentes de frequência de um sinal e rejeitar outras. Um exemplo disso é a análise da voz humana. Se forem retiradas as componentes de alta frequência, a voz terá um som diferente, mas é possível entender o que está sendo dito. Entretanto, se forem retiradas as componentes de baixa frequência, será emitido um palavreado sem sentido.

Na análise *wavelet*, fala-se usualmente em aproximações e detalhes. As aproximações são as altas escalas, ou seja, as componentes de baixa frequência do sinal. Os detalhes são as baixas escalas, isto é, as componentes de alta frequência (MISITI et al., 2002).

O processo de filtragem, em seu nível mais básico, está ilustrado na Figura 3.11.

Figura 3.11 - Processo de filtragem de um sinal

O sinal original S passa através de dois filtros complementares que fornecem dois sinais como saída.

Em um sinal digital real, obtém-se duas vezes mais a quantidade de dados em relação aos dados iniciais. Por exemplo, considere um sinal original S contendo 1000 amostras de dados. Sendo assim, as versões aproximadas (A) e detalhadas (D) do sinal original terão 1000 amostras cada uma, totalizando 2000 amostras.

Para corrigir esse problema, foi introduzido um operador que reduz o número de amostras que é chamado de operador *downsampling* (Figura 3.12). Esse operador considera dados intercalados, ou seja, leva em consideração o primeiro dado e rejeita o segundo, e assim por diante. Com isso, é introduzido um fenômeno *aliasing*, ou superposição de espectro, nas componentes do sinal, o que deverá ser levado em conta em tratamentos posteriores.

O processo na direita da Figura 3.12 ilustra a atuação do operador *downsampling*, o qual produz os coeficientes da TWD. O sinal é uma senóide pura adicionada com ruído de alta frequência.

Figura 3.12 - Exemplo de filtragem com *downsampling* de um sinal senoidal ruidoso

Observa-se na Figura 3.12 que os coeficientes de detalhe cD são menores e consistem principalmente do ruído de alta frequência, enquanto que os coeficientes cA contêm menos ruído que o sinal original.

As famílias *Wavelets* frequentemente mais usadas para processamento de sinais são as *Wavelets* Daubechies, Morlets, Coiflets e Symlets. Essas *Wavelets* exibem diferentes atributos e critérios de performance quando empregadas em aplicações específicas, tais como: detecção de transitórios, compressão de sinais e filtragem de ruído. Ainda que não exista um critério definido para a seleção das *Wavelets*, a melhor escolha é uma *Wavelet* que melhor caracterize o fenômeno ou o problema a ser estudado (ADDISON, 2002).

3.2.3 Módulo Máximo da Transformada *Wavelet* (MMTW)

O Módulo Máximo é encontrado a partir da TWC original usando um simples algoritmo que percorre o escalograma Wavelet gerado, escala por escala e identifica os máximos e mínimos locais. Detalhes sobre implementações do MMTW pode ser visto em Lin e Sharif (2007), Fachfour e Kinsner (2005).

O método utilizado na caracterização multifractal por *Wavelet* usa o Método do MMTW (MALLAT; HWANG, 1992). Este método foi desenvolvido baseando-se na propriedade apresentada pela transformada *Wavelet* de ser capaz de detectar, no tempo, regiões em que uma função sofre mudanças abruptas de comportamento, além de verificar para quais escalas estas mudanças permanecem relevantes. O MMTW pode ser visto sobreposto ao escalograma da Figura 3.9(b), representado por traços finos em preto.

Na Figura 3.13 pode ser observada como é representado o MMTW para um sinal qualquer.

Figura 3.13 - Representação e formação do MMTW de um sinal qualquer. (Fonte: http://cas.ensmp.fr/~chaplais/wavetour_presentation/Regularite/Detection_of_singularities.html, acessado em 20/06/2009).

3.3 Teoria fractal

A palavra “fractal” vem da junção das palavras Latinas *fractus* que significa “irregular” e *frangere* que significa “quebrar” e a pronuncia correta é “frac’tal”.

A geometria fractal é um ramo relativamente novo na matemática que vem sendo estudado desde a sua descoberta nos anos 60 por MANDELBROT (1982).

A geometria fractal possibilita estudar melhor as formas da natureza, onde a geometria euclidiana não tem grande êxito, pois não descreve perfeitamente as estruturas encontradas com facilidade na natureza. Na Figura 3.14, D_T , d , D_s representam dimensões topológicas, euclidiana e fractal, respectivamente. A dimensão euclidiana apresenta uma dimensão inteira, enquanto que na geometria fractal tem dimensão fracionada e propriedade de auto-similaridade (pedaço de objeto que se assemelham ao objeto todo), como sugere a Figura 3.14.

Um objeto fractal tem como principal característica suas propriedades de medida, como o comprimento ou área, como sendo funções da escala de medida. Um

exemplo clássico para ilustrar esta propriedade é a medida da costa (MANDELBROT, 1982) que quando medida numa escala δ , o comprimento da costa $L(\delta)$ é estimado como um conjunto de N segmentos de retas de comprimento δ .

Figura 3.14 - Comparação entre a geometria fractal e a geometria euclidiana. (Fonte: JORGE et al., 2008)

3.3.1 Auto-similaridade

Segundo descrito por Jorge et al. (2008), a auto-similaridade, ou auto-semelhança é uma idéia antiga, contudo, apesar de ser uma propriedade física simples, apenas no início da década de 70 o homem percebeu a sua existência na natureza, verificando que há uma infinita gama de objetos com estruturas geométricas complexas. Um fractal possui um número imenso de pequenas cópias dele próprio, o que define a auto-semelhança. Este conceito está relacionado com a dimensão fracionária, o que diferencia da dimensão convencional (números inteiros) e serve para caracterizar o fractal.

A auto-similaridade significa que, independente da complexidade da forma de um objeto, observando-se mais detalhadamente a estrutura, um observador pode encontrar as mesmas formas (ou semelhantes) em escalas diferentes (MANDELBROT, 1982; POSADAS et al., 2005).

Na natureza existem vários exemplos claros de fractal, como demonstrado na Figura 3.15.

Figura 3.15 - Modelos de fácil entendimento, (a) um brócolis, (b) é o detalhe das estruturas do brócolis que demonstram a auto-similaridade e (c) é uma folha de samambaia, demonstrando o mesmo. (Fonte: <<http://pt.wikipedia.org/wiki/Fractal>> acessado em 20/05/2008).

Existem vários tipos de fractais, como os estocásticos e os determinísticos. Para um fractal determinístico, representado pela Figura 3.16, (a) tem uma estrutura fractal de crescimento, usando um procedimento iterativo, (b) representa uma estrutura análoga, só que desta vez, a estrutura foi obtida através de sucessivas divisões do quadrado original. Quando $k \rightarrow \infty$, ambas as estruturas tem dimensão fractal $D = 1,465$ (JORGE et al., 2008; VICSEK, 1992).

Figura 3.16 - Exemplo de uma construção de um fractal determinístico. (Fonte: JORGE et al., 2008)

No caso de um sinal 1D pode ser observada a auto-similaridade como na Figura 3.17, ou seja, pequenas partes do sinal possuem mesmo formato que o sinal inteiro.

Figura 3.17 - Auto-similaridade em um sinal (Fonte: GRAPS, 1995).

3.3.2 Determinação da Dimensão Fractal

Um conjunto fractal foi definido por Mandelbrot (1982) como um conjunto para o qual a dimensão de Hausdoff-Besicovitch (D_h) é maior que a dimensão topológica (D_T). A dimensão de Hausdoff-Besicovitch (D_h) é definida como a razão logarítmica entre o número N de objetos internos e o recíproco raio r numa escala de observação δ , ou seja:

$$D_h = \frac{\ln(N)}{\ln(1/r)} \quad (3.7)$$

Neste caso, a equação (3.7) irá ser constituída de N padrões nos quais o tamanho foi reduzido de fatores r . A relação entre a escala δ e o comprimento L obtido na escala será dada por:

$$L(\delta) = k \cdot \delta^{(1 - D_h)} \quad (3.8)$$

Onde k é uma constante e D_h a dimensão fractal.

Existem vários métodos de cálculo da dimensão fractal (BETROUNI; LOPES, 2009). O método mais difundido para a determinação da dimensão fractal é o *Box-counting*, definido por Russel et al. (1980). Utilizando caixas de comprimento r cobrindo toda a estrutura ou o sinal, a dimensão fractal D é calculada por:

$$D = - \lim_{r \rightarrow 0} \frac{\log N(r)}{\log r} \quad (3.9)$$

Onde: $N(r)$ é o número de caixas necessárias para cobrir o sinal.

A Figura 3.18 ilustra a utilização das caixas de diferentes comprimentos τ sobre o sinal. Estes diferentes comprimentos são as escalas de análise.

Figura 3.18 - Utilização de caixas no método “Box-counting” para o cálculo da dimensão fractal, $D_h = 1,39$.(Fonte: POSADAS, 2001).

3.3.3 Multifractal

Uma série temporal $f(t)$ pode possuir não somente uma dimensão fractal, mas um conjunto de dimensões fractais. Neste caso, é chamada de multifractal. Os multifractais são séries temporais, em qualquer dimensão, que são singulares em muitos pontos. Neste caso, uma única dimensão fractal é insuficiente para caracterizar sua geometria, assim como qualquer propriedade física inerente. Com a finalidade de obter informações de forma generalizada foram definidos os multifractais.

Uma das desvantagens do fractal é a incapacidade de identificar as singularidades de um sinal, e por este motivo o multifractal foi introduzido para possibilitar uma descrição estatística de singularidades em termos de funções termodinâmicas tais como a dimensão fractal generalizada e o espectro de singularidades. O fato de não ser considerada a probabilidade de ocorrência da estrutura dentro de cada caixa e sim apenas o número de caixas que contém a

estrutura, a simples análise fractal não é capaz de identificar as singularidades do sistema. Por exemplo, o método não faz distinção entre as diferentes densidades encontradas nos quadros da Figura 3.19, apesar da notável diferença no preenchimento de cada caixa.

Figura 3.19 - Uma aplicação do método de análise fractal em uma estrutura 2D de um sistema poroso em solos. Os poros representam-se com áreas negras, e δ_i é o tamanho dos quadrados, ou escala. Os quadrados que estão sendo especificados na figura mostram os extremos contidos em um quadrado. (Fonte: POSADAS et al., 2001).

Em particular, se for considerado o suporte da medida μ do sinal em análise na caixa de tamanho δ , e definindo-se $P_i(\delta)$ como a probabilidade na i - *ésima* caixa dada, define-se um expoente α como sendo a singularidade Hölder pela relação (3.7):

$$P_i(\delta) \sim \delta^\alpha \quad (3.10)$$

Se forem contadas as caixas $N_\delta(\alpha)$, na escala δ , com probabilidade P_i , que possuem singularidade num intervalo de α e $\alpha + d\alpha$, então $f(\alpha)$ pode ser definido como sendo a dimensão fractal do conjunto de caixas com singularidade α pela equação (3.11) :

$$N_\delta(\alpha) \sim \delta^{-f(\alpha)} \quad (3.11)$$

Sendo $f(\alpha)$ também o espectro de singularidades, ou dimensão de Hausdorff do conjunto de caixas com singularidades α .

Dimensão fractal generalizada

A dimensão fractal generalizada D_q , que corresponde ao espaço dos q momentos da medida μ fornecem também as medidas singulares ou multifractais.

Definindo-se uma função de partição (3.12):

$$Z_q(\delta) = \sum_{i=1}^{N(\delta)} \mu_i^q(\delta) \quad (3.12)$$

Onde q é o momento de distribuição estatística. A dimensão fractal no espaço dos momentos q é D_q . A função de partição, Z_q , está relacionada com a escala δ através de (3.13):

$$Z_q(\delta) \sim \delta^{-\tau(q)} \quad (3.13)$$

Segundo Arneodo et al. (1995):

$$D_q = \tau(q)/(q - 1) \quad (3.14)$$

$$\alpha(q) = \frac{d\tau(q)}{dq} \quad (3.15)$$

$$f(\alpha(q)) = \tau(q) - \alpha(q)q \quad (3.16)$$

Assim, os espectros de singularidades $f(\alpha)$ são avaliados numericamente pela transformada de Legendre de $\tau(q)$.

No entanto, para evitar algumas distorções originadas na transformada de Legendre, Chhabra et al. (1989), propõe uso do método canônico, baseado nas equações (3.17) e (3.18) para cálculos das singularidades $f(\alpha)$ e α .

$$f(\alpha(q)) = \lim_{\delta \rightarrow 0} \frac{\sum_{i=1}^{N(\delta)} \mu_i^q \log \mu_i^q}{\log \delta} \quad (3.17)$$

$$\alpha(q) = \lim_{\delta \rightarrow 0} \frac{\sum_{i=1}^{N(\delta)} p_i^q \log \mu_i^q}{\log \delta} \quad (3.18)$$

Originalmente, pode ser calculada a dimensão fractal generalizada como na equação (3.19):

$$D_q = \frac{1}{q-1} \lim_{\delta \rightarrow 0} \frac{\log \mu(q, \delta)}{\log \delta} = \frac{1}{q-1} \lim_{\delta \rightarrow 0} \frac{\log \sum_{i=1}^{N(\delta)} \mu_i^q}{\log \delta} \quad (3.19)$$

A dimensão generalizada D_q é uma função monótona decrescente para todos os valores reais de q dentro do intervalo $[-\infty, \infty]$. Quando $q < 0$, μ acentua regiões na distribuição com menos concentração de uma medida, e o contrário é verdadeiro, para $q > 0$.

O espectro multifractal $f(\alpha)$ e a dimensão generalizada D_q contêm informações semelhantes, ambas caracterizam os arranjos estatísticos dos fractais de dimensão $f(\alpha_i)$. Em cada i -ésima fractal, é observado escalas δ_i com expoente Lipschitz-Hölder, α_i . Ao contrário do D_q que é associado à geometria do sistema, o espectro multifractal $f(\alpha)$ associa-se a parâmetros termodinâmicos do sistema, sendo f e α a entropia e a energia interna do sistema, respectivamente. Assim, o espectro multifractal $f(\alpha)$ tem a capacidade de descrever fisicamente qualquer sistema estudado.

A dimensão fractal $D(q=0)$ ou D_0 é a dimensão do “box-counting”, a $D(q=1)$ ou D_1 é a dimensão da informação e $D(q=2)$ ou D_2 é a dimensão de correlação, dadas por (3.20) a (3.22), respectivamente.

$$D_0 = \lim_{\delta \rightarrow 0} \frac{\log(N(\delta))}{\log \delta} \quad (3.20)$$

$$D_1 = \lim_{\delta \rightarrow 0} \frac{\sum_{i=1}^{N(\delta)} \mu_{ii}(\delta) \log(\mu(\delta_i))}{\log \delta} \quad (3.21)$$

$$D_2 = \lim_{\delta \rightarrow 0} \frac{\log(C(\delta))}{\log \delta} \quad (3.22)$$

Sendo $C(\delta)$ a função de correlação.

D_0 é a dimensão de capacidade independente de q e fornece uma informação global (ou média) do sistema. A D_1 é relacionada à informação de entropia de Shannon e quantifica o grau de desordem presente em uma distribuição, medindo a forma que a entropia de Shannon seja, por exemplo, $\sum \mu(\delta) \log[\mu(\delta)]$ quando δ tende a 0. Para uma medida $\mu \in [0,1]$, o valor de D_1 está no intervalo $0 < D_1 < 1$. Um valor de

D_1 próximo a 1,0 caracteriza um sistema uniforme distribuído através de todas as escalas, enquanto que um valor de D_1 perto de 0 reflete um subconjunto de escalas na qual as irregularidades estão concentradas. O D_2 associa-se matematicamente à função de correlação e computa a correlação de medidas contidas em intervalos de tamanho δ . A relação entre D_0 , D_1 e D_2 é definida em (3.23).

$$D_0 \leq D_1 \leq D_2 \quad (3.23)$$

onde a igualdade $D_0 = D_1 = D_2$ ocorre só se o fractal for estatisticamente auto-similar e homogêneo.

Wavelet-Multifractal

O expoente de Hölder numa função $x(t)$ é também a sua singularidade α . É definida como o maior expoente de um polinômio $P_n(t)$ de ordem n satisfazendo a seguinte condição:

$$|x(t) - P_n(t - t_0)| \leq C|t - t_0|^\alpha \quad (3.24)$$

Para todo t na vizinhança de t_0 . O polinômio $P_n(t)$ corresponde à série de Taylor de x .

O expoente mede a irregularidade de uma função $x(t)$ num ponto t_0 . Quanto maior o expoente de $\alpha(t_0)$, mais regular é a função $x(t)$.

Utilizando localmente a TWC definida na equação (3.2), ou seja, com $b = t_0$, pode-se escrever (3.25):

$$TWC(t_0, a) \sim a^{\alpha(t_0)} \quad (3.25)$$

Onde $\alpha(t_0)$ é o expoente de singularidade ao redor de t_0 .

Utilizando-se ao invés da TWC e sim o MMTW, uma vez que se tem um conjunto de valores discretos, assim a função de partição é definida em (3.26):

$$Z_q(a) = \sum_{i=1}^{N(\delta)} |MMTW(a, t_i(a))|^q \sim a^{\tau(q)} \quad (3.26)$$

Pelo método canônico (CHHABRA et al., 1989), pode ser utilizada a medida normalizada μ pela equação (3.27):

$$\mu(q, a, t_0) = \frac{|MMTW(a, t_0)|^q}{\sum |MMTW(a, t_i)|^q} \quad (3.27)$$

sendo MMTW na escala a .

E por fim o espectro de singularidade pode ser obtido por (MA et al., 2005):

$$\alpha(q) = \lim_{a \rightarrow 0} \frac{\sum_{i=1}^{t_i(a)} \mu_{ti}^q \log |MMTW(a, t_i)|}{\log a} \quad (3.28)$$

$$f(\alpha(q)) = \lim_{a \rightarrow 0} \frac{\sum_{i=1}^{t_i(a)} \mu_{ti}^q \log \mu_{ti}^q}{\log a} \quad (3.29)$$

Espectro Multifractal

A Figura 3.20 representa um espectro multifractal típico, onde os momentos q estão relacionados à singularidade do espectro: para valores $q > 1$, as medidas são fortemente reforçadas e para $q < 1$, as menores áreas singulares são enfatizadas. O ponto em que $f(\alpha)$ é igual a α e corresponde ao ponto $q = 1$, a média do sistema, o ponto $q = 0$ corresponde ao valor máximo e o ponto $q = 2$ corresponde à correlação. Para monofractais, o espectro é constante, enquanto que para multifractais são representados por uma função monótona decrescente. (HURST, 1951 e STRUZI, 2000).

De acordo com Posadas et al. (2005), um espectro multifractal, $f(\alpha)$ por α (Figura 3.20), $f(\alpha)$ está associada à entropia generalizada e α é a energia interna do sistema. O espectro caracteriza quantitativamente a variabilidade da medida. A largura do espectro é a indicação da variabilidade total. Para gráficos com a distância de $\Delta\alpha$ maior, sinaliza que o sistema é mais heterogêneo, conseqüentemente mais entrópico, e se a distância é menor, indica que o sistema é menos entrópico, ou seja, mais homogêneo.

Figura 3.20 - Espectro multifractal.

Quando o espectro multifractal apresenta apenas um ponto, o sinal é monofractal, quando se encontra valores distintos, este é multifractal.

3.4 Extração de características em sinais sonoros

Processamento de sinais normalmente tem sido realizado no domínio da frequência usando-se a Transformada discreta de *Fourier* (DFT), ou mesmo a transformada Z. No domínio da frequência as filtragens se tornam simples multiplicações e, como isto, é computacionalmente mais rápido. Normalmente, frequências de corte, de rejeição, máxima largura de banda de transição, máxima variação na banda de passagem e a máxima variação na banda de atenuação são os principais parâmetros avaliados. A DFT fornece a informação das frequências que compõe o sinal, mas não o relaciona com o tempo.

Da mesma forma acontece para o processamento de áudio e voz.

No processamento digital de voz, devido o sistema bio-gerador de voz, trata-se dois tipos de sinais: *voiced speech* (basicamente vogais) e *unvoiced speech* (sons como “s”, “f”, “th”, etc). Para sinais “vozeados”, a frequência fundamental é a frequência de *pitch*, em torno de 80 a 100Hz, para mulheres pode alcançar até 300Hz, e para crianças até mesmo cerca de 400 ou 500Hz. O período associado, igual ao inverso da frequência é o período de *pitch*.

As análises das características importantes do sinal de voz são obtidas da análise conjunta tempo-frequência. Os espectrogramas representam o sinal de voz visualmente e permite análise tempo-frequência-potência. O espectro de voz *cepstrum*, dilatado apropriadamente, cria uma escala chamada escala Mel, muito usada para reconhecimento de voz. Devidamente filtrado, são gerados 20 sinais que são usados para comparação e reconhecimento. Assim, o reconhecimento de fala é realizado por árvores de decisão, separando, vozeado de não vozeado, alta frequência de baixa frequência, silêncio de som, vogal de consoante, parâmetros MEL, usando no final *template-matching* ou *knowledge-based*. (Arabi et al., 2006).

Na literatura, a maioria dos trabalhos apresenta o *Dynamic Time Warping* (DTW) como uma das técnicas de *template-matching*, seguidas daquelas baseadas em correlação e dimensão fractal. No caso de *Knowledge-based*, diversos parâmetros são extraídos e um sistema mais complexo, baseado em aprendizado é utilizado. Tais sistemas envolvem redes neurais, modelos ocultos de Markov (*Hidden Markov Models* – HMM) e modelos de misturas Gaussianas (*Gaussian Mixture Models* – GMM). (DENG; O'SHAUGHNESSY, 2003).

No reconhecimento de voz, algumas técnicas visam identificar o locutor, dentre outros sinais. Verificação do locutor autorizado para uma dada entrada é uma das aplicações que mais se destaca e se baseia em um modelo para autorizado e outro para não autorizado. O algoritmo de reconhecimento de locutor tenta segmentar o sinal encontrando regiões de silêncio, isolar cada sub-sinal anterior e, utilizando preferencialmente as partes nasais, extrair características, e por fim, fazer o reconhecimento baseado num modelo previamente treinado. Os modelos mais usados são obtidos por redes neurais, GMM e HMM, fundamentalmente do tipo *Knowledge-based*.

Existem também técnicas para síntese de fala, que a partir de um dicionário de tempo de duração de cada fonema e características específicas (voiced/unvoiced, formantes, pitch, etc), estabelece um modelo (HMM).

Algumas patologias pelo funcionamento indevido de alguns órgãos da fala causam alterações no sinal da fala. Entre as anormalidades pode-se citar o edema de Renke, nódulos, entre outros, causando ruídos de alta frequência desvio do período de *pitch*, entre outras anormalidades. As principais técnicas e parâmetros utilizam Redes Neurais Artificiais, fractais, e *Wavelet*, analisando o período de *pitch* e seus desvios, frequências formantes, *low-to-high energy ratio*, *noise-to-harmonics ratio*, *harmonics-*

to-noise ratio. (HADDJITODOROV et al., 2000; ROSA et al., 2000); FONSECA et al., 2007).

A TW também tem sido utilizada para detecção do período de *pitch* de sinais de fala, identificando locutores, voz com doenças etc.

No caso de insetos, alguns trabalhos descrevem os registros e tipos de cantos e sons emitidos, baseados apenas em espectrogramas (VIRANT-DOBERLET; COKL, 2004; MORAES et al., 2005; LAUMANN et al., 2007). No reconhecimento, destaca-se Ganchev et al. (2007), que utiliza GMM, HMM e Redes Neurais, com parâmetros do tipo harmônicas dominantes, duração de pulsos e picos, energia e dispersão ao redor do pico máximo.

O uso de fractal para processamento de sinais sonoros tem recebido maior destaque nos últimos anos, destacando-se o estudo de turbulência (MUZY et al., 1991). Um bom exemplo de uso da Wavelet-Multifractal para processamento de sinais pode ser visto em Soemintapura e Langi (1998).

3.5 Considerações Finais

A evolução tecnológica tanto de hardware como software tem proporcionado processamentos mais pesados e técnicas mais avançadas, podendo treinar sistemas com grande quantidade de dados.

Neste capítulo foram apresentadas técnicas de análise de som que utilizam grande capacidade de processamento e tempo de máquina na fase de treinamento. Várias são as características que podem ser extraídas tanto de sons como das imagens. Por este motivo, no próximo capítulo se descreve os métodos de seleção das características sonoras de interesse bem como das levantadas no rastreamento para o reconhecimento do comportamento animal, que neste caso está sendo os percevejos que atacam a lavoura brasileira.

Capítulo 4

Fusão

4.1 Considerações Iniciais

A fusão de informações utiliza diferentes técnicas para trabalhar dados provenientes de múltiplas fontes: sensores, bancos de dados, sistemas inteligentes, seres humanos, etc., com o principal objetivo de obter o melhor entendimento de algum fenômeno para introduzir inteligência e funções de decisão. O próprio cérebro humano é um exemplo, pois recebe vários sinais provenientes da visão, odores, sensação de tato, sons ou mesmo sabor, fazendo a sua devida interpretação e fusão de informações para tomada de decisão. Múltiplos sensores normalmente melhoram o desempenho de um sistema de decisão na detecção, o rastreamento e identificação, além de interpretação da situação. A redundância de sensores torna os sistemas mais robustos e melhora o desempenho dos mesmos diante da falha de algum sensor. Além disso, informações espaciais e temporais podem ser cobertas com múltiplos sensores de forma mais completa.

Por outro lado, diferentes sensores podem tornar mais complexos os algoritmos de decisão. A grande quantidade e tipos de sensores, o volume de dados, os requisitos de sistema em tempo real e dificuldades de propagação de sinais são alguns dos problemas que devem ser tratados no caso de técnicas de fusão. A natureza dos sensores e o tipo de informação fornecida, a localização e a capacidade computacional nos sensores devem ser observadas com cuidado para se definir a arquitetura do sistema de fusão. A topologia, a comunicação e os recursos computacionais definem muitas vezes o nível da fusão, que pode ser realizada diretamente nos dados, nas características extraídas dos dados, ou mesmo na fase de tomada de decisão.

O tipo do dado, os atributos e localização, comportamento, relacionamento de cada entidade, forma de acesso e desempenho na avaliação e disponibilidade de recursos definem o nível da fusão. Dependendo do nível, diferentes técnicas são

utilizadas, partindo de uma simples detecção do sinal, uma estimativa ou filtragem, passando por sistemas baseados em redes neurais, clusterização, até mesmo definições usando lógica nebulosa ou sistemas baseados em conhecimento até o controle e algoritmos de otimização. Em geral a solução de fusões de sistemas complexos requer a multidisciplinaridade com a integração de diversos algoritmos e técnicas.

A coleta de grandes quantidades de dados em diversas áreas tem sido cada vez mais uma realidade. Neste trabalho foram arquivados vídeos de várias horas de monitoramento do comportamento de insetos, bem como sons de todo o período. Esses dados foram posteriormente usados para descrever e caracterizar o comportamento dos insetos com base no movimento e no sinal sonoro dos insetos em condições diferenciadas. Sobre esses dados, foram extraídas características e estas submetidas a algoritmos de aprendizado supervisionado foram utilizadas para induzir um modelo para prever os rótulos de novos exemplos. O sucesso na construção de um modelo de previsão adequado, mais especificamente um classificador, depende da relevância das características do conjunto de dados. Características irrelevantes interferem na construção desses modelos e podem acarretar baixa acurácia de predição.

A seleção de características relevantes é um tema do Aprendizado de Máquina (AM). Embora a maioria dos algoritmos de AM busque, intrinsecamente, selecionar características ou atribuir-lhes graus de importância, estudos indicam que a presença de muitos atributos irrelevantes deteriora o desempenho desses algoritmos na construção das hipóteses (FERREIRA; JORGE, 2007). Por isso, métodos de seleção de características representam uma alternativa para aumentar a acurácia das hipóteses geradas, reduzir custos de armazenamento e viabilizar a interpretação de atributos e dos modelos gerados. Nesse contexto, o objetivo deste capítulo é apresentar as abordagens de seleção de características, como uma alternativa de técnica de fusão de atributos relevantes de som e de imagem para o estudo comportamental dos insetos.

4.2 Níveis de Fusão

Na Figura 4.1 podem ser observados os três níveis de fusão (HALL; LLINAS, 2001).

O nível mais baixo de fusão ocorre na fusão dos dados e utiliza os dados obtidos de todos os sensores fundidos diretamente. Erros de registro de dados pode gerar o maior gasto computacional. Neste nível, decisões são baseadas nas informações de todos os sensores. A quantidade de dados para processar neste nível é bem superior, exigindo mais tempo de processamento.

No nível intermediário, a fusão é realizada nas características, combinando as características extraídas dos dados, ou seja, cada sensor possui um domínio individual no qual são extraídas as características de acordo com a aplicação. Este nível pressupõe capacidade de detecção de cada sensor independente. As características de cada sensor são independentemente extraídas em cada domínio, e um espaço de características é utilizado posteriormente para a classificação. A fusão efetuada neste nível reduz a quantidade de dados, uma vez que são extraídas características relevantes de cada sensor. As características são mescladas em um espaço de decisão comum.

Um exemplo de fusão nível básico e nível intermediário, ou seja, fusão nível de dados e fusão nível de características, pode ser observado em Bebis et al. (2006). Neste trabalho utilizado um algoritmo genético para encontrar a melhor estratégia de fusão aplicado no reconhecimento de faces.

Fusão no nível de decisão, também chamado de pós-decisão, combina as decisões independente do sensor usando uma heurística, atribuindo mais ou menos pesos num ou noutro sensor. Na fusão em nível de decisão, cada sensor deve permitir de forma independente realizar a classificação, e posteriormente cada decisão tomada será ponderada para formar a decisão final. Um exemplo de fusão no nível de decisão, usando algoritmos de aprendizado de máquina foi apresentado por Gigli et al. (2007).

Neste trabalho, foi adotada a opção de nível intermediário, ou seja, a extração de características, pois o sinal sonoro e o vídeo são dados complexos e ao longo do tempo aumentam significativamente. Já é feita uma redução no espaço de dados quando se extrai as características. A fusão adotada adotou a abordagem por características, fazendo a seleção das características mais relevantes para a aplicação, e gerando um algoritmo classificador. A seleção de subconjuntos de características por técnicas de mineração de dados foi então a abordagem deste trabalho.

4.3 Seleção de Subconjunto de Características

A relevância das características consideradas no conjunto de exemplos de treinamento T define a qualidade do algoritmo indutor. Em problemas de classificação, em geral, o conjunto de exemplos T é composto por m instâncias (A, C) , sendo o vetor A um elemento do produto cartesiano $a_1 \times a_2 \times \dots \times a_i$; C é a classe que A pertence; a_i é o domínio do i -ésimo atributo e i é o número de características utilizadas para descrever o exemplo em questão. A relevância das características é fundamental para o algoritmo de AM no processo de aprendizado. Quanto maior a quantidade de características irrelevantes, maior a necessidade de exemplos de treinamento para alcançar uma dada acurácia (FERREIRA; JORGE, 2007).

Figura 4.1 - Níveis de fusão: (a) básico ou por dados; (b) intermediário ou por características e (c) pós-decisão ou por decisão.

A tarefa de aprendizado de conceitos pode ser dividida em (1) decidir que características utilizar na descrição do conceito e (2) decidir como combiná-las (BARANAUSKAS, 2001). Os métodos de seleção de subconjunto de características (FSS - *Feature Subset Selection*) têm o objetivo de encontrar uma combinação

adequada de características relevantes para o aprendizado de um conceito. A relevância de características é um conceito de definição complexa. Características individualmente irrelevantes podem ser relevantes quando combinadas. A relevância depende também das peculiaridades do modelo de aprendizado (algoritmo indutor) adotado. Assim, um subconjunto relevante para um determinado modelo, pode não ser similar para outro modelo de aprendizado distinto. Diversas definições para relevância de características podem ser encontradas em Blum e Langley (1997).

Um processo de FSS é caracterizado pela redução da dimensão do vetor de características. A partir de um conjunto U , um subconjunto de características R é selecionado usando algum critério, tal que $(U-R) \neq \emptyset$. Em outras palavras, o conjunto original U é analisado e processado para produzir um subconjunto R com um número menor de atributos. Várias razões justificam e motivam a aplicação de FSS (LEE et al., 1999). A primeira razão é que alguns algoritmos de AM computacionalmente viáveis não trabalham adequadamente na presença de um grande número de características, principalmente se houver características irrelevantes. Isso significa que a FSS é um instrumento importante para aumentar a acurácia dos classificadores gerados por esses algoritmos. Outra razão é que a FSS pode auxiliar na compreensão dos dados pois, pode indicar características importantes para o processo de predição. A terceira razão é que FSS pode ajudar a reduzir os custos envolvidos na coleta de dados de alguns domínios. Finalmente, a redução de características por FSS pode reduzir significativamente o custo de processamento de grandes quantidades de dados.

O processo de FSS pode ser descrito como uma busca em um espaço de estados, onde cada estado representa um subconjunto de características que é mensurável por meio de critérios pré-estabelecidos para prover uma estimativa da relevância daquele subconjunto. Operadores de adição ou remoção de características são comuns e possibilitam a travessia ou caminhamento no espaço de estados. Na Figura 4.2 é mostrado um grafo que representa o espaço de busca para quatro características, onde cada nó corresponde a um estado.

Figura 4.2 - Espaço de busca (estados) (BLUM; LANGLEY, 1997).

Os círculos em preto representam a presença das características e os círculos em branco a ausência delas. Cada estado no espaço dos subconjuntos de características especifica quais são os atributos candidatos à indução de um classificador. Os estados no espaço de busca são estruturados (ordenados) por número de características. Em cada nível, caminhando da esquerda para direita, uma característica é acrescentada. Os extremos representam a ausência e a presença total das características do problema em questão.

Percorrer todo o espaço de busca, busca exaustiva, para encontrar o melhor subconjunto de características é uma tarefa computacionalmente inviável na grande maioria dos casos, pois a complexidade computacional é da ordem de 2^n , onde n é o número de características do conjunto U . Para exemplificar, suponha uma imagem descrita por 50 características ($n = 50$) representando as cores, por exemplo, indicando diferentes padrões de cobertura do solo. Se a estimativa de precisão de um estado do espaço levasse um milésimo de segundo para ser computada e se a avaliação de todos os possíveis estados fosse requerida, o melhor resultado levaria cerca de 35700 anos para ser emitido. Portanto, a busca exaustiva garante a melhor solução, mas é impraticável na maioria dos casos. Se fosse conseguido paralelizar a busca e usar arquitetura paralela, se torna viável esta busca.

Para tornar a FSS viável, diversas soluções para seleção das características relevantes são encontradas na literatura. Para geração de novos candidatos há três abordagens:

- *completa* - é a busca exaustiva por todos os estados no espaço de busca. Conforme descrito anteriormente é impraticável quando o número de características é relativamente grande. Sua complexidade é $O(2^n)$;
- *heurística* - em cada iteração, todas as características ainda não selecionadas (ou selecionadas) são consideradas e avaliadas para seleção (ou remoção). A utilização de heurísticas pode diminuir drasticamente a complexidade da busca. Sua complexidade é $O(n^2)$;
- *aleatória* - é uma abordagem relativamente nova comparada as outras duas categorias. Em geral, os subconjuntos candidatos são aleatoriamente amostrados no espaço de busca. Embora sua complexidade seja $O(2^n)$, os métodos dessa categoria efetuam, tipicamente, uma busca em um número limitado de subconjuntos de características. Algoritmos genéticos aplicados a FSS são excelentes alternativas desta categoria.

Um paradigma conveniente na seleção de características é a busca heurística, apresentado a seguir.

4.4 Seleção de Características como Busca Heurística

A visualização das abordagens de seleção de características como uma busca heurística permite que qualquer método de FSS seja caracterizado em relação a quatro pontos básicos:

- ponto de partida;
- organização da busca;
- estratégia da busca;
- critério de parada.

4.4.1 Ponto de partida

O ponto de partida determina qual é o estado no espaço onde a busca deve iniciar. A determinação do ponto de partida influencia na direção da busca e nos operadores que serão utilizados para geração dos estados sucessores. Pode-se observar no grafo da Figura 4.2 a existência de um ordenamento lógico entre os estados, pois o estado adjacente à direita tem uma característica adicional e o estado adjacente à esquerda tem uma característica subtraída. Se o ponto inicial é determinado pelo estado sem nenhuma característica (primeiro estado à esquerda) e o operador adiciona características gradativamente, então a abordagem é conhecida como seleção *forward*. Se o ponto inicial é o estado com todas as características (último estado à direita) e o operador remove características gradativamente, então a abordagem é conhecida como seleção *backward*. Além disso, pode ser empregada uma variação de ambas, escolhendo-se como ponto de partida qualquer estado do espaço e movendo-se a partir desse ponto – seleção *outward*.

4.4.2 Organização da busca

Conforme descrito anteriormente, a busca exaustiva por todos os estados no espaço de busca é impraticável para muitos atributos. Uma abordagem mais prática e viável é a utilização de métodos gulosos² para travessia do espaço de busca: em cada estado do espaço consideram-se alterações locais sobre o atual subconjunto de características, seleciona-se uma nova característica e realiza-se uma nova iteração. A abordagem *hill-climbing*, por exemplo, conhecida por seleção ou eliminação

2 Métodos que selecionam sempre o estado com a melhor estimativa a cada iteração.

stepwise, considera tanto a adição quanto a remoção de características em cada ponto (estado) de decisão, além de permitir que um estado anterior possa ser reconsiderado durante a busca.

4.4.3 Estratégia de Avaliação

Um dos pontos fundamentais na busca é a estratégia utilizada para avaliar o subconjunto de características. Uma métrica geralmente empregada é a capacidade que uma característica tem para discriminar as classes em um conjunto de treinamento. Diversos algoritmos empregam um critério baseado na Teoria da Informação, enquanto outros medem diretamente a acurácia do classificador sobre o conjunto de treinamento ou sobre um conjunto de teste.

A estratégia de avaliação está relacionada à maneira com que o algoritmo de FSS lida com o algoritmo básico de AM. A independência de avaliação com respeito ao algoritmo de AM é uma característica fundamental da abordagem *filtro*. A dependência do algoritmo de AM caracteriza a abordagem conhecida como *wrapper*.

4.4.4 Critério de parada

A dimensão do espaço de estados pode tornar o processo de seleção de características computacionalmente custoso, mesmo na busca heurística. Assim, algum critério para parar a busca deve ser definido. O critério de parada pré-estabelece as condições em que a busca deve cessar. Alguns critérios utilizados são:

- parar de adicionar ou remover características quando nenhuma das alternativas melhora a acurácia estimada do classificador;
- parar quando um dos extremos do espaço de busca for alcançado e selecionar a melhor solução;
- parar quando um número pré-definido de características é alcançado;
- parar após um número pré-definido de iterações.

4.5 Abordagens para Seleção de Características

Segundo Blum e Langley (1997), os métodos de FSS podem ser agrupados em três classes/abordagens: embutida, filtros e *wrapper*. Cada uma dessas abordagens é discutida nas seções subseqüentes.

4.5.1 Embutida

Na indução de hipóteses, alguns algoritmos de aprendizado podem realizar uma seleção dinâmica de características. Nesses casos, a FSS é um componente intrínseco do algoritmo de aprendizado.

Os métodos de particionamento recursivo, tais como as árvores de decisão, efetuam uma busca gulosa através do espaço de estados/características. A cada passo, uma função de avaliação determina qual característica tem maior capacidade de discriminação das classes. Com base na característica escolhida, o conjunto de treinamento é particionado e o processo repetido, estendendo a árvore até que nenhuma discriminação adicional seja possível. Esse tipo de abordagem é usada pelo algoritmo C4.5 (QUINLAN, 1993). Na Figura 4.3 é apresentada uma ilustração da abordagem de FSS embutida para um algoritmo de construção de uma árvore de decisão. Observa-se que a árvore de decisão final (hipótese) utiliza apenas uma parcela das características (representadas como f_2 , f_6 e f_7) do conjunto total.

Figura 4.3 - Abordagem embutida (FERREIRA ; JORGE, 2007).

4.5.2 Filtro

Essa abordagem de seleção de características é caracterizada pela independência de seleção em relação ao algoritmo de AM utilizado para gerar o modelo final. Os filtros atuam na fase de pré-processamento dos dados e fazem uma seleção de características, recebendo como entrada o conjunto de treinamento, com todas as características, avaliando as propriedades intrínsecas dos dados e produzindo um subconjunto de maior relevância segundo algum critério estatístico. Na Figura 4.4 é ilustrada a abordagem filtro.

Um dos esquemas mais simples de filtragem é a avaliação de cada característica individualmente, baseada na sua correlação com a classe (C), escolhendo as r características com melhor valor.

Figura 4.4 - Abordagem Filtro (FERREIRA; JORGE, 2007) .

A aplicação de algoritmos de filtragem oferece vantagens, tais como: eficiência para lidar com grande quantidade de dados, baixo custo de processamento computacional e independência do algoritmo de AM. A principal desvantagem dessa abordagem é que os filtros ignoram totalmente os efeitos do subconjunto de características no desempenho do algoritmo de AM ao qual os dados reduzidos serão submetidos (BARANAUSKAS, 2001). Por isso, Kohavi e John (1997) propõem uma abordagem, denominada *wrapper*, que considera o próprio algoritmo de AM no processo de seleção.

4.5.3 Wrapper

Os algoritmos baseados nessa abordagem geram um subconjunto de características como candidatas, executam o algoritmo de AM com os dados de treinamento desse subconjunto e medem a acurácia do modelo gerado. Uma medida de acurácia do modelo é utilizada para avaliar o subconjunto de características em questão. Esse processo é repetido a cada novo subconjunto gerado até que algum critério de parada seja satisfeito. Todo processo envolvido na abordagem *wrapper* é apresentado na Figura 4.5.

Figura 4.5 - Abordagem Wrapper (KOHAVI; JOHN, 1997).

A principal desvantagem dessa abordagem é o custo computacional. Para cada subconjunto de características, um modelo deve ser gerado e sua acurácia estimada.

Esse fator encarece computacionalmente todo o processo de busca pelo subconjunto adequado. Entretanto, a principal vantagem é que o subconjunto de características selecionado é otimizado para o modelo final. Por essa razão, *wrapper* geralmente supera a abordagem filtro (JORGE et al., 2007a; JORGE et al., 2007b).

Em Sun et al. (2004) foi utilizada a abordagem *wrapper* com um algoritmo genético, porém combinando com a abordagem filtro para então elaborar um classificador de padrões. Neste caso, o filtro foi um algoritmo de análise das componentes principais para reduzir o espaço de atributos e dentro deste espaço das componentes principais foi aplicado o *wrapper* com o algoritmo genético para então selecionar o subconjunto ótimo para o classificador implementado com uma máquina de vetor de suporte, que guia o algoritmo genético para obter a melhor acurácia.

Quando o espaço de busca é demasiadamente grande, torna-se impraticável o uso de técnicas de busca tradicionais em *wrapper* como, por exemplo, *hill-climbing* com algoritmos de AM de RNAs (Redes Neurais Artificiais). Assim, algoritmos genéticos são, geralmente, mais eficazes para uma rápida busca global em grandes espaços de busca, além de oferecerem um atrativo para uma otimização multi-objetivo (YANG; HONAVAR, 1998).

4.6 Seleção de Características como Fusão de Dados

A utilização de seleção de características em um vetor de atributos permite buscar as características relevantes para a aplicação, sejam elas provenientes ou não de diferentes fontes. Caso os atributos de um dos sensores, imagem ou som, não sejam incluídos nos algoritmos indutores gerados, significa que não são relevantes para o processo de aprendizado.

Na prática métodos de seleção de características são empregados para aumentar o desempenho do processo de aprendizado de máquina. Uma das consequências é a redução de dimensionalidade do domínio dos dados visando aumentar o desempenho dos algoritmos de mineração e reduzindo a carga computacional. Na abordagem *wrapper* adotada neste trabalho, o algoritmo de mineração é o elemento-chave na avaliação do conjunto de atributos. Na aprendizagem supervisionada, em geral, estamos tentando minimizar o erro de classificação ou alguma função de perda com base em análises *recall x precision*. Assim, a fusão em si é simplesmente executada por esta operação de seleção de características e confere a melhor fusão no nível de características.

4.7 Considerações Finais

Neste capítulo foram apresentados os níveis de fusão geralmente utilizados e os conceitos de aprendizado de máquina utilizados para seleção dos atributos relevantes no vetor de características. Este vetor é composto de características extraídas do vídeo e do som emitidos pelos insetos.

No próximo capítulo serão apresentados os algoritmos indutores implementados e a descrição dos atributos extraídos das imagens e sons e os resultados encontrados.

Capítulo 5

Resultados e Discussões

5.1 Considerações Iniciais

Nos capítulos anteriores foram apresentadas algumas das técnicas utilizadas para estudo do comportamento de insetos por som e por imagem, bem como a seleção de características por técnicas de mineração de dados. Neste capítulo serão apresentadas as metodologias, os resultados e as discussões com as técnicas empregadas.

Na Figura 5.1, é apresentado um fluxograma com as etapas do trabalho. Observa-se que, no mesmo experimento, no mesmo instante de tempo, foram coletados vídeos e sons emitidos por insetos. Escolheu-se simular ou induzir alguns comportamentos já conhecidos para facilitar o processo de aquisição de dados e também minimizar o tempo de captura de dados. Assim mesmo, foram realizados dois períodos de uma semana para obtenção de dados, um no início dos estudos, para ter material para iniciar os testes em dezembro de 2006 e outra captura de dados em julho de 2007. Os vídeos coletados foram digitalizados e processados pelos diferentes algoritmos de rastreamento selecionados. A trilha gerada em cada um dos segmentos de vídeo foi analisada, e posteriormente, foram definidos e extraídos os atributos mais adequados para representar o comportamento descrito pelo inseto. Ao mesmo tempo em que foram coletados os dados, os diferentes padrões de comportamento foram devidamente identificados e registrados por um especialista em entomologia, criando assim a referência ou o *ground-truth*. Foram utilizados os algoritmos de rastreamento de vídeo para gerar as trilhas percorridas pelos insetos. Das trilhas foram extraídos os atributos que seriam somados aos demais para aplicar o algoritmo de fusão.

Da mesma forma, em cada período correspondente aos vídeos foi capturado o som emitido pelo inseto. Este sinal sonoro foi processado pelas técnicas Wavelet-Multifractal visando extrair os atributos para a posterior fusão com os atributos de movimentação do inseto.

Figura 5.1 - Etapas de desenvolvimento da fusão de som e imagem

5.2 Materiais

Para a realização deste estudo foram utilizados percevejos (*Hemiptera: Pentatomidas*), de colônias de laboratório da Embrapa Recursos Biológicos e Biotecnologia, em Brasília, DF, mantidas em salas climatizadas por volta de 26°C e 65,0% de umidade relativa. Todos insetos sexualmente maduros e criados em salas diferentes, até o momento de serem utilizados nos experimentos.

Para estabelecer as categorias comportamentais foram escolhidas todas as fases do acasalamento ou cópula das espécies. Os insetos foram isolados em arenas de estudo com fundo de placas de petri de 9 cm de diâmetro, que possui um dos extremos abertos. Este extremo colocado invertido sobre a membrana formada por papel de filtro de 12 cm de diâmetro, sobre um autofalante invertido para operar como microfone e um amplificador (Sonifex Redbox, RB-MA1), um sistema de digitalização do som (Aardvark-Direct Pro 24/96) e um computador. No caso das imagens, foi fixada uma filmadora Sony HandCam MiniDV logo acima da arena, com a qual eram disparadas as capturas das imagens. Também uma placa digitalizadora Pinnacle foi utilizada para captura dos *frames* e vídeos.

Na Figura 5.2 pode ser observada a vista geral do local do experimento, mostrando a câmara isolada de som onde foram colocadas as arenas da Figura 5.3.

O processo de registro do som foi realizado segundo critérios utilizados pelos especialistas. Um etograma foi elaborado para determinar os pontos de gravação do som e da imagem. No primeiro momento, foram gravados trechos de no máximo 2 minutos em cada fase descrita no etograma da Figura 5.4.

(a)

(b)

Figura 5.2 - Vista geral do local do experimento: (a) distribuição de microfones e suporte da câmera de vídeo (b) autofalante invertido, sendo usado como microfone.

(a)

(b)

Figura 5.3 - Tomada de uma arena do tipo circular utilizada com 2 insetos. (a) vista lateral sob um auto-falante, e (b) vista superior com microfone de eletreto e superfície de polímero excelente resistência.

No etograma da figura 5.4 podem ser observados os pontos de captura de sinais sonoros de Machos (CM-1 a CM-4), de fêmeas (CF-1 a CF-2). Os nomes utilizados são CM para insetos controle machos e CF para controle fêmeas. Estes sinais sonoros foram caracterizados por Moraes et al. (2005), conforme demonstrado na Figura 5.5. Na Figura 5.5(a) pode ser observado um canto de fêmea com contato visual com o macho e na Figura 5.5(b) o mesmo canto sem contato visual. No caso do macho, o mesmo pode ser observado nas Figuras 5.5(c) e (d). Neste experimento foi possível constatar os mesmos padrões sonoros encontrados por Moraes et al. (2005), porém vale ressaltar que mesmo em condições controladas de isolamento, os níveis de ruído nos sinais obtidos nos dias dos experimentos foram significativos. Além disso, foi utilizado um novo sistema de captura de sons, especialmente desenvolvido pela Embrapa Instrumentação para este experimento de tese. Este foi projetado com microfones de eletreto e películas de polímero resistente, o que também tornou o sistema mais sensível.

Os números em parênteses no Etograma da Figura 5.4 demonstram o número de casais de insetos e o respectivo comportamento. Neste caso foram 25 casais de insetos. Além disso, foram utilizados mais alguns para simular a presença de mais machos para se observar a rivalidade entre os mesmos. Na Figura 5.6 podem ser observados exemplos típicos dos padrões sonoros emitidos pelos machos, pelas fêmeas e na rivalidade de machos, obtidos durante experimento e usados para treinamento nas técnicas estudadas. Observa-se também o espectrograma de cada um dos sinais determinados pelo software. Na Figura 5.7 pode ser observada uma típica de sequência de vídeo obtida durante os monitoramentos de até 2 min. Foram digitalizados vídeos com taxas médias de 5 quadros por segundo. Para insetos

menores, principalmente os predadores dos percevejos, foram capturados testes com taxa máxima de 15 quadros por segundo, uma vez que os insetos se deslocavam com velocidade maior.

Figura 5.4 - Etograma dos comportamentos reprodutivos de machos e fêmeas .
(Fonte: Embrapa Recursos Genéticos).

De cima para baixo da esquerda para a direita foram capturados quadros com pouco mais de 2 segundos, conforme Figura 5.7. Verifica-se que a movimentação do inseto neste período é de pequenos deslocamentos, mas as antenas ficam mudando de posição o tempo todo e movimentos com o corpo com mudanças de direção são possíveis de observar. Em geral, para cobrir todo o deslocamento de 2 minutos, e para não ficar extremamente lento, o sistema não armazena as imagens e sim faz o processamento em tempo real de tal forma que a única informação armazenada seja a trilha gerada pelo inseto. Um detalhe a ser observado nas imagens da Figura 5.7 é um fio de arame colocado na arena. É através deste fio que é transmitido sons previamente gravados de cantos de machos ou fêmeas para estimular o canto.

Figura 5.5 - Sons do *Euschistus heros* notificados por Moraes et al. (2005) cantos (a) e (b) de fêmea, (c) e (d) de macho, (e) de rivalidade dos machos e (f) macho após recusa de cópula pela fêmea. (Fonte: Moraes et al.,2005)

Figura 5.6 - Exemplos dos sons capturados no experimento e os respectivos espectrogramas calculados.

Figura 5.7 - Sequência de vídeo típica capturada nos experimentos.

5.3 Detecção de movimento

Como foi descrito no capítulo 2 foram implementados e testados diferentes métodos de detecção de movimento. Os quadros sucessivos foram estudados inicialmente pela simples subtração de frames (JORGE et al. 2006; SPINK et al., 2001; WAJNBERG; COLAZZA, 1998), o detector de HARRIS (JORGE et al., 2011), bem como uma mistura de GMM com quadros sucessivos.

Na figura 5.8, pode ser observada uma sequência subtrações dos frames da Figura 5.7, sendo os quatro primeiros os resultados da simples diferença de frames consecutivos da Figura 5.7. Neste caso, apenas as movimentações dos insetos ficaram evidenciadas. Ao contrário deste resultado, as duas últimas subtrações apresentam os insetos duplicados, bem como outros artefatos resultantes não apenas dos insetos se movimentando. Estes foram resultados da subtração não de frames imediatamente consecutivos, mas sim a cada 9 a 10 frames. Ou seja, percebe-se uma grande alteração com a mudança da posição do inseto, bem como a interferência no

background com a ação de retirada do arame presente na arena, o que mostra o quanto é sensível à mudanças no fundo, necessitando mais cuidados.

Figura 5.8 - Imagens de diferenças de quadros consecutivos do vídeo da Figura 5.7, bem como a cada 9 a 10 quadros.

No caso do algoritmo Lucas e Kanade (1981) do fluxo ótico, foram realizados alguns testes para identificar a detecção de movimento nas imagens. Na Figura 5.9 pode ser observado que mais pontos na imagem evidenciaram movimento, não só os insetos. Neste caso, o *background* não se movimentou. Foram efeitos de variações das fontes de luz fluorescente utilizadas durante o experimento. As pequenas variações ficaram mais visíveis. Verificando a diferença de variação na direção vertical ou horizontal separadamente, conforme Figura 5.10, percebe-se que os efeitos de iluminação estão evidentes nas bordas da arena.

Figura 5.9 - Imagem mostrando resultado de movimento utilizando o Fluxo ótico Lucas-Kanade com média de 8 frames consecutivos.

(a) (b)

Figura 5.10 - Utilizando a diferença entre dois quadros, (a) na vertical e (b) na horizontal, por Lucas-kanade, (1981).

Uma vez que o fundo não se altera, e as pequenas modificações são provenientes das pequenas variações na iluminação, foram realizados testes de aplicar a mistura de Gaussianas de alguns quadros consecutivos, de tal forma que o fundo fosse atualizado pela distribuição de luz presente no experimento. Sendo assim, a diferença entre os quadros resultaria em uma detecção com muito menos artefatos provenientes da iluminação, como pode ser observado na Figura 5.11.

Alguns trabalhos mostram alternativas similares para detecção do animal em movimento detectando formas em movimento (PISTORI et al., 2010); outros, inclusive

um dos produtos comerciais mais usados (NOLDUS et al. 2002), não evidencia claramente o algoritmo utilizado, mas percebe-se que é uma variação da subtração de quadros em vídeos e atualização do fundo ao longo do tempo, indicando uso de GMM.

No entanto, se por um lado corrige-se o problema da iluminação, acrescenta-se outro problema. Se o inseto fica parado ele acaba sendo incorporado no fundo, sumindo a referência de onde estava o inseto. Nestes casos, o algoritmo criou algumas condições de contorno para tratar estes casos, tentando não perder o inseto em questão. Isto se torna um problema mais complicado quando se tem muitos insetos na imagem, sendo uns parados e outros em movimento.

Figura 5.11 - Detecção de movimento com GMM adaptado

Finalmente, foram realizados testes com as técnicas (LOWE, 2004; HARRIS; STEPHENS, 1988) que se destacaram mais recentemente no rastreamento de objetos em vídeos, também utilizadas por Fisher e Xiao (2010). As técnicas que utilizam pontos salientes ou *Keypoints* têm sido utilizadas em aplicações de monitoramento humano ou de segurança (YILMAZ et al., 2006). Nas figuras 5.12 (a) e (b) podem ser vistos os *Keypoints* gerados nas técnicas SIFT e detector de Harris, ambos apresentam pontos semelhantes e foram utilizados de forma complementar, sendo o detector de Harris mais rápido no processamento. O detector de Harris evidencia os cantos e a SIFT as bordas. Na Figura 5.13 é apresentado o resultado do algoritmo de casamento de pontos usando o RANSAC para ligação dos pontos e detecção de movimento.

Figura 5.12 - Keypoints nos quadros nos instantes de tempo (a) t e (b) $t+1$ determinados pela transformada SIFT

Tanto o SIFT como o detector de Harris são invariantes em relação à escala, rotação e translação, no entanto também são sensíveis com as variações de iluminação. Para se minimizar este efeito, é necessário também acrescentar a informação do tamanho do inseto de tal forma definir a sensibilidade do método. Ou seja, optou-se por trabalhar em uma área cujos pontos salientes sejam mais fortemente presentes e representem um inseto. No caso do aplicativo implementado, foram feitos vários testes na condição de luz e o tipo de inseto para então definir o padrão de processamento.

Percebe-se na Figura 5.13 que o casamento de pontos pelo método RANSAC já elimina alguns pontos menos importantes e diminui o número. Além deste, é necessário calcular o agrupamentos numa forma do inseto em estudo. Além disso, foi introduzido o parâmetro de limitação de área onde o inseto poderia estar na arena, eliminando os efeitos de borda da arena.

Figura 5.13 - Pontos característicos obtidos na transformada SIFT e Harris, interligados pelo algoritmo RANSAC.

Foram basicamente escolhidos os algoritmos de subtração e mistura Gaussiana bem como os detectores de Harris para a análise de 30 padrões de sinais de cada comportamento.

Com base nos parâmetros analisados em cada trabalho de comportamento animal estudado (NOLDUS et al., 2002a; NOLDUS et al., 2002b; YE; BELL, 1991; WAJNBERG; COLAZZA, 2009; EILAM, 2003; SPINK et al., 2001, WAWERLA et al., 2009), bem como pela experiência dos entomologistas que acompanharam o experimento, depois dos vários testes, definiu-se os parâmetros necessários a serem determinados em cada trilha capturada, sendo eles:

- velocidade angular (w);
- velocidade linear (v);
- deslocamento angular (dw);
- deslocamento linear (dx);
- tortuosidade (T).

A tortuosidade foi definida por (5.1):

$$T = 1 - \text{lin} / \text{curv} \quad (5.1)$$

sendo *lin* a distância linear entre o início e o fim do deslocamento e *curv*, a soma de todos os pequenos deslocamentos ao longo do *tracking*, como em (5.2)

$$\text{curv} = \sum \|p_i - p_{i+1}\| \quad (5.2)$$

para p_i a posição inicial. Assim, cada trilha detectada para cada inseto gera um vetor de características do deslocamento representado por (5.3)

$$V_{C_{\text{video}}} = [v, dx, w, dw, T] \quad (5.3)$$

Nos testes onde foram inseridos mais insetos, utilizou-se os valores médios das características para os insetos presentes. Em geral se utilizou um macho e uma fêmea para o acasalamento e mais machos para estudar rivalidade.

A maior dificuldade foram os seguintes pontos:

- perda do rastreio quando os insetos não se movimentam;
- tratar as oclusões quando um inseto se sobrepõe ao outro;
- movimentos falsos provocados por movimento nas patas ou asas;
- variações de iluminação, apesar do fundo se manter relativamente padrão.

Foi inserido um parâmetro de sensibilidade no software que leva em conta a distribuição de tons na janela de detecção proporcional ao tamanho do inseto. Foram estabelecidos níveis baixo, médio e alto. Dependendo da seleção, são detectados mais ou menos pontos de movimento. O programa foi elaborado de tal forma a detectar movimento apenas em regiões de tamanho previamente definido na sensibilidade do algoritmo. Os movimentos que não se encontram dentro da área de vizinhança que define um inseto é considerado outro inseto.

Os parâmetros foram inicialmente obtidos com apenas um inseto. Depois de bem entendido o comportamento, adotou-se valores médios de todos os insetos na imagem, pois estes valores facilitam um entendimento médio do comportamento.

No caso das oclusões causadas pelos próprios insetos, estas foram desconsideradas uma vez que se optou pelo valor médio. A iluminação foi contornada mantendo-se uma fonte padrão de infra-vermelho, servindo inclusive para o experimento noturno. Este foi um ponto importante a ser considerado uma vez que os insetos escolhidos, percevejos que atacam as lavouras brasileiras, são de hábitos noturnos.

Na Figura 5.14 pode ser observado o relatório típico que é armazenado para posterior análise. Além dos parâmetros para compor o vetor de características, são determinadas algumas informações adicionais, tais como tempo de residência em cada área de interesse, número de vezes que entrou ou saiu de cada área bem como o tempo do experimento. Estas áreas são definidas pelo usuário antes do experimento iniciar.

Um resumo dos processamentos executados pelo software SACAM desenvolvido está descrito na Figura 5.15. Neste, cria-se uma pasta para cada experimento, onde serão armazenados os dados de rastreo, são configurados os parâmetros de ajuste de cada algoritmo de detecção, a escala da análise em unidades reais, as áreas de interesse, o tamanho do inseto e sensibilidade do algoritmo e posteriormente pode-se disparar um processamento ou programar um processamento, por exemplo, para o período noturno.

5.4 Detecção de Som

Basicamente foi implementado o algoritmo Wavelet-Multifractal com base no artigo de Chhabra e Jensen (1989), de Muzy et al.(1991) e fundamentalmente o trabalho de Jordan e Miksad (1997) que foi utilizado para implementar a transformada Wavelet contínua, a partir da convolução das TF.

Figura 5.14 - Exemplo típico de um relatório gerado para cada trilha do experimento.

Figura 5.15 - Fluxograma de ações no software SACAM para cada experimento

O fluxograma da Figura 5.16 apresenta a forma implementada neste trabalho para o cálculo da transformada Wavelet. A partir da Wavelet mãe e do sinal Wavelet, em potência de dois (2^n), aplica-se a escala, e então, a transformada rápida de *Fourier* (FFT – *Fast Fourier Transform*) deve ser aplicada tanto na Wavelet mãe como no sinal que se pretende analisar. Efetua-se a convolução e então, aplica-se a transformada de *Fourier* inversa. Caso a escala não seja a máxima, o processamento é encerrado e a TWC é concluída, caso contrário aplica-se uma nova escala, sendo esta potência de dois.

Figura 5.16 - Algoritmo da Transformada Wavelet contínua

De acordo com Posadas et al. (2005), um espectro multifractal $f(\alpha) \times \alpha$ caracteriza quantitativamente a variabilidade da medida, sendo $f(\alpha)$ associado à entropia generalizada e α é a energia interna do sistema. O comprimento do espectro é a indicação da variabilidade total. O comprimento de $\Delta\alpha$ maior sinaliza que o sistema é mais heterogêneo e consequentemente, mais entrópico. Já se o comprimento é menor indica que o sistema é menos entrópico, ou seja, mais homogêneo.

Foram então selecionados sons típicos de macho e fêmea nos seguintes comportamentos: chamada pelo parceiro, cômte ou cópula, e rivalidade. O espectrograma de cada som foi estudado, gerado o escalograma a partir da TWC, posteriormente determinado o MMWT e os respectivos espectros multifractais representando cada comportamento, sendo 30 amostras de cada comportamento. Foram testadas diferentes *Wavelet*-mães: *Paul*, *Morlet*, *Gauss*, Chapéu Mexicano, e de *Shannon*) (MISITI, et al., 2002). O espectro de singularidades foi resumido pelos seguintes parâmetros Δf , $\Delta f_{\text{máx}}$, Δf_{min} para representar a simetria da curva, o coeficiente de Hölder α , as dimensões fractais D_0 , D_1 e D_2 , relacionadas à auto similaridade, entropia e correlação, respectivamente.

Assim o vetor de características do som ($V_{c_{som}}$) foi definido por (5.4):

$$V_{c_{som}} = [\Delta q; \Delta f; \Delta f_{\max}; \Delta f_{\min}; \alpha_{\max}; \alpha_{\min}; \Delta \alpha; f(0) - f(-1); f(0) - f(1); \alpha(0) - \alpha(-1); D_0; D_1; D_2] \quad (5.4)$$

Para se determinar o espectro multifractal, deve-se primeiro definir a função *Wavelet*-mãe a ser utilizada na análise.

Nas Figuras 5.17 e 5.18 são apresentados os espectros multifractal obtidos com as diferentes *Wavelet*-mães escolhidas para cantos de fêmea e macho respectivamente. Alguns espectros multifractais, como nas Figuras 5.17(c) e 5.18(c), não apresentam respostas típicas multifractais, indicando que estas *Wavelet*-mães selecionadas não são adequadas ou a faixa de sinal não é adequada.

Além do formato do espectro multifractal deve-se observar o espectro da TW para cada uma das funções *Wavelet*-mães, verificando se ela percorre toda a faixa de escalas. Se concentrar apenas nas primeiras escalas, isto indica que não é adequada para análise multifractal dos sinais estudados. Nas Figuras 5.17 e 5.18, itens (b), (d) e (f) podem ser observados os espectros da TW.

Para as faixas de momento utilizadas, variando de $-2 < q < 2$, com deslocamentos 0,1, mais de um tipo de *Wavelet*-mãe pode ser utilizado na análise. Foram escolhida as *Wavelets*-mães *Morlet* e *Paul* por possuírem o espectro multifractal mais completo para $q < 0$ e $q > 0$.

Na Tabela 5.1 são apresentados os espectros multifractais encontrados, bem como os escalogramas a partir da TWC de *Morlet* e os respectivos MMTW, para cantos típicos de chamamento de macho, de fêmea, de acasalamento e de rivalidade entre machos obtidos para um conjunto de insetos. Os valores dos espectros estão apresentados na Tabela 5.2, sendo os parâmetros de Δf , $\Delta f_{\max}$ e $\Delta f_{\min}$ meramente para ver a simetria da curva, α é o coeficiente de Hölder, o D_0 é a dimensão fractal propriamente dita, D_1 e D_2 são as dimensões de entropia e de correlação, respectivamente.

Pode-se observar que um único macho apresentou padrões de emissões heterogêneas em relação às frequências. Isso se deve ao comportamento do espectro multifractal $f(\alpha) \times \alpha$, nos quais quanto mais pontual for o gráfico, mais homogêneo é o sistema analisado e quanto mais afastado os pontos negativos estiverem dos positivos, determina-se que o sistema seja heterogêneo, ou seja, mais entrópico.

(a)

(b)

(c)

(d)

(e)

(f)

Figura 5.17 - Wavelet-mães para cantos de fêmea: (a) Morlet (c) Chapéu Mexicano (e) Shanon, com os respectivos espectros TW

(a)

(b)

(c)

(d)

(e)

(f)

Figura 5.18 - Wavelet-mães para cantos de macho: (a) Morlet (c) Chapéu Mexicano (e) quarta derivada Gaussiana, com os respectivos espectros TW

TABELA 5.1 – Espectros Wavelet com módulo máximo e espectros fractais para exemplos de canto de macho, canto de fêmea, acasalamento e rivalidade.

TABELA 5.2 - Dados das emissões de sons de insetos do programa, onde temos os diferentes tipos de comportamento. Os parâmetros de Δf , $\Delta f_{\text{máx}}$ e Δf_{min} são meramente para ver a simetria da curva, α é o coeficiente de Hölder, o D_0 é a dimensão fractal propriamente dita, D_1 e D_2 são as dimensões de entropia e de correlação, respectivamente. [] números negativos; - dado não existente; q=0, 0.

Tipo	Análise	Δq	Δf	$\Delta f_{\text{máx}}$	Δf_{min}	$\alpha_{\text{máx}}$	α_{min}	$\Delta \alpha$	$f(0) - f(-1)$	$f(0) - f(1)$	$\alpha(0) - \alpha(-1)$	$\alpha(0) - \alpha(1)$	D_0	D_1	D_2
	Todo	0	0	0	0	0	0	0	0	0	0	0	1,213208	1,306172	1,3790525
	Frequência 1	2,8	0,0842	0,0274	0,1116	1,3903	1,039	0,3513	0,1083	0,0323	[0,2434]	0,1072	1,084902	1,052563	1,0786003
	Frequência 2	0	0	0	0	0	0	0	0	0	0	0	1,431979	1,5095	1,5342198
	Frequência 3	0	0	0	0	0	0	0	0	0	0	0	1,45302	1,518696	1,523292
Fêmea Eh c1 8	Todo	2	0,1297	0,0326	0,1623	1,459	0,7538	0,7052	-	0,0428	-	0,1688	0,164949	0,122063	0,1571888
	Frequência 1	0	0	0	0	0	0	0	0	0	0	0	0,623331	0,781395	0,8796509
	Frequência 2	0	0	0	0	0	0	0	0	0	0	0	0,61492	0,67546	0,7021639
	Frequência 3	0	0	0	0	0	0	0	0	0	0	0	0,724888	0,810397	0,8181167
Acasalamento Eh7	Todo	1,4	0,4346	0	0,4346	1,177	0,8018	0,3752	0,0957	-	[0,0982]	-	0,472333	0,555331	0,7465601
	Frequência 1	0	0	0	0	0	0	0	0	0	0	0	1,20546	1,265009	1,2690264
	Frequência 2	4	0,0653	0,0061	0,0714	1,1641	1,1034	0,0607	0,0083	0,0016	[0,0136]	0,0038	1,089246	1,087646	1,0830576
	Frequência 3	0	0	0	0	0	0	0	0	0	0	0	1,178569	1,183632	1,184007
	Frequência 4	4	0,031	0,0121	0,0431	1,1447	1,0941	0,0506	0,0073	0,0038	[0,0136]	0,0083	1,122606	1,118765	1,1105093
registro 28--mFêmea+somMacho	Todo	1,3	0,6621	[0,0015]	0,6606	1,9739	0,9237	1,0502	-	-	-	-	0,890552	0,931166	1,0506457
	Frequência 1	4	[0,0608]	0,1009	0,0401	1,4499	1,2703	0,1796	0,0234	0,0326	[0,0557]	0,0678	0,950657	0,918124	0,8497744
	Frequência 2	0	0	0	0	0	0	0	0	0	0	0	1,593459	1,597098	1,6152158
registro 30 50m Macho+somFêmea	Todo	2,4	0,0953	0,0645	0,1598	1,5424	0,5798	0,9626	-	0,0757	-	0,2857	0,215536	0,139796	0,150992
	Frequência 1	2,7	[0,0402]	0,0333	[0,0069]	1,4041	1,37	0,0341	-	0,0135	-	0,0281	1,307535	1,293991	1,2741994
	Frequência 2	2,9	[0,0647]	0,0583	[0,0064]	1,2003	1,1307	0,0696	-	0,0185	-	0,0385	0,761268	0,742824	0,7029826
registro 49 110 F+3m rivalidade	Todo	1,1	0,5885	[0,0014]	0,5871	1,3932	0,5631	0,8301	-	-	-	-	0,719563	0,78602	0,9978195
	Frequência 1	3,3	0,0583	0,0541	0,1124	1,353	1,1227	0,2303	0,0602	0,0232	[0,1095]	0,0537	0,127026	0,103842	0,0729199
	Frequência 2	4	0,1211	0,3543	0,4754	2,0085	0,9454	1,0631	0,139	0,1667	[0,341]	0,3767	0,68784	0,521064	0,3335333
	Frequência 3	3	0,0333	0,0895	0,1228	1,3832	1,02	0,3632	0,1228	0,0437	[0,2255]	0,1048	0,133539	0,089762	0,0439876
	Frequência 4	2,3	[0,7816]	0,7815	[0,0001]	1,7109	0,8613	0,8496	-	0,2634	-	0,4712	1,323314	1,059871	0,5417992
registro 31 58m Macho+FêmeaAposTentativaacasalam	Todo	2	[0,0211]	0	[0,0211]	1,3182	1,326	[0,0078]	0,0047	-	[0,0057]	-	0,956203	0,970044	1,0195955
	Frequência 1	0	0	0	0	0	0	0	0	0	0	0	0,90084	0,911671	0,9276506
	Frequência 2	2,3	0,1264	[0,0005]	0,1259	1,3958	1,3138	0,082	0,0099	-	[0,0121]	-	0,872975	0,882602	0,9141077

Para a fêmea, os espectros das frequências se apresentaram todos pontuais, sendo que, de maneira geral, são mais homogêneos em relação ao do macho, o que pode ser um indicativo dos diferentes “sons” emitidos pelos machos durante o chamamento reprodutivo, a cópula e mesmo a função de territorialidade.

Também, nota-se predominância da heterogeneidade dos pulsos emitidos no acasalamento, possivelmente pelos comportamentos variados envolvendo mais de um indivíduo, além das sexualidades diferentes. Ressalta-se a variabilidade dos espectros de territorialidade, pois mesmo composta de indivíduos todos do sexo masculino apresentou heterogeneidade em todas as frequências capturadas.

A Figura 5.19 demonstra os resultados preliminares dos dados dos insetos para a variação da entropia nos diferentes comportamentos. Nota-se uma variabilidade significativa ao longo do tempo, sendo os maiores picos determinados pelas maiores entropias e, conseqüentemente, maior heterogeneidade, já os menores valores determinam menores entropias e maior homogeneidade dos sons

Figura 5.19 - Gráfico da variação da entropia do comportamento dos insetos ao longo do tempo. Linhas: vermelha: macho, azul: fêmea, preta: acasalamento, verde: fêmea com som de macho induzido, roxa: macho com som de fêmea induzido, amarela: rivalidade, cinza: pós-tentativa de acasalamento.

Pela análise das respostas e dos sinais encontrados, pode-se tirar algumas conclusões do comportamento, mas os valores do espectro multifractal foram utilizados e fundidos com os atributos do rastreamento obtidos dos vídeos, conforme descrito. Em seguida, foram aplicadas as técnicas de fusão para selecionar os parâmetros mais relevantes.

O próximo passo foi gerar um classificador os parâmetros selecionados ou fundidos e treinar um classificador de padrões para posterior processamento do

comportamento do inseto. Após o treinamento, foram aplicados a um conjunto de teste para se obter os resultados.

5.5 Fusão por seleção de características

Neste trabalho, o principal objetivo foi selecionar as características que melhor caracterizassem o comportamento, seja por som, ou por imagem, formando assim um vetor de características mesclando ambos sinais, ou seja (5.5):

$$Vc = Vc_{som} + Vc_{video} \quad (5.5)$$

Este vetor de características resultante incorpora todos os dados extraídos de som e de imagem para os insetos monitorados em diferentes situações.

A escolha dos atributos é de extrema importância para o resultado final. Sendo os atributos provenientes de diferentes sensores, a fusão poderia ser desde uma simples soma como apresentado na equação (5.5) até outra técnica de fusão nos diferentes níveis já definidos. Os sinais brutos, de vídeo ou som, seriam praticamente impossíveis de serem analisados automaticamente, sem a extração prévia de atributos adequados. Por este motivo foram extraídos os atributos já descritos. No entanto, estes atributos podem ser irrelevantes ou podem apresentar correlação entre eles piorando o processo de classificação. Sendo assim, foi proposta neste trabalho a técnica de fusão por seleção de características, procurando um subconjunto ótimo de componentes e suas características mais representativas do comportamento, o que é conseguido através do balanço adequado entre informações de sons emitidos e movimentações realizadas.

Baseando-se na noção de separabilidade de classes, foi proposto um método para selecionar um subconjunto de atributos visando conseguir a melhor seleção de componentes utilizando as técnicas descritas.

A obtenção de bons resultados também depende do algoritmo utilizado na extração e seleção de características relevantes, pois é a partir desta fase que é criado o conjunto de atributos que será utilizado na etapa de treinamento. Pode-se pensar que quanto maior o número de atributos para representar um padrão (classe), maior o poder discriminatório de um classificador, mas nem sempre isso é verdade. Na prática, o que acontece é uma degradação na acurácia dos resultados da classificação com o aumento da dimensionalidade dos dados mantendo-se constante o número de amostras de treinamento.

Depois de feita a seleção e extração das características, inicia-se o processo de decisão onde serão definidas quais as características mais representativas para o comportamento do inseto.

A API (*Application Programming Interface*) do *toolbox* WEKA (WEKA, 2010) foi utilizada, pois ela possui vários algoritmos disponíveis, que permitiram a extração das principais características e a geração de um classificador. Além disso, o código fonte está disponível e pode ser alterado para integrar na ferramenta desenvolvida.

Para a extração do subconjunto de características relevantes é necessária a utilização da combinação de um algoritmo de seleção de características, responsável por selecionar as características mais relevantes, e de um algoritmo de busca, que é responsável por realizar a busca no subconjunto de características gerado pela execução dos algoritmos de seleção de características.

Seleção de características baseados na correlação de características ou *Correlation-based Feature Selection* (CFS) e busca gulosa (*Greedy Stepwise*). O algoritmo CFS utilizado para seleção de características, avalia o valor de um subconjunto de características, considerando a habilidade preditiva de cada característica permitindo um certo grau de redundância entre estas. Esse algoritmo escolhe os subconjuntos de características que são correlacionados com a classe e que possuam correlação interna. O algoritmo *Greedy Stepwise* é utilizado para realizar a busca de características, e faz uma busca gulosa nos subconjuntos de cada espaço de atributos. Essa busca pode ser iniciada em todas as características ou pode ser iniciada arbitrariamente em qualquer ponto do espaço de características (WITTEN; FRANK, 2005).

Para a classificação foi avaliado o algoritmo C4.5 que é baseado em uma árvore de decisão que representa as informações do algoritmo de aprendizagem de máquina e expressa essa estrutura dos dados de maneira rápida e eficiente (QUINLAN, 1993). Árvores de Decisão são algoritmos recursivos de partição de dados de acordo com um critério de divisão retirado do algoritmo indutor, independentes das variáveis dos dados. Os dados são recursivamente particionados de forma a diminuir os erros com relação à variável dependente para os subconjuntos. Consequentemente, quanto mais baixo se está na estrutura em árvore, mais homogêneos são os subconjuntos. Árvores de Decisão são ferramentas úteis para os dados porque permitem redução de dados e uma descrição compacta dos dados; classificação, pois permitem a descoberta de hipótese com base em classes separadas, que pode ser explicado pelas suas variáveis indutivas, e generalização,

porque eles descobrem mapeamento de variáveis independentes para variáveis dependentes. Apesar da superfície de separação ser considerada ruim para funções complexas, a idéia foi considerar esta como sendo o pior caso, sabendo que o uso de máquinas de vetores de suporte, redes neurais, algoritmos genéticos, dentre outros, poderia ser melhorar a performance do classificador.

A árvore de decisão é um classificador na forma de um grafo acíclico direcionado. Os tipos mais comuns de Árvores de Decisão são as binárias. Eles têm zero ou um valor de arestas de saída em qualquer nó da árvore. Um nó com zero arestas de saída é chamado de nó folha, todos os outros nós são chamados internos. Nós folhas são rotuladas com rótulos de classe: Uma categoria é uma variável dependente. O atributo preditor do nó tem uma divisão predicado associado a ele. Um nó interno produz uma divisão da base de dados em dois subconjuntos disjuntos. A parte dos dados originais que se encaixa no critério será considerado pela sub-árvore esquerda da Árvore de Decisão e da parte dos dados que não satisfazem o predicado divisão será considerada pelo lado direito da sub-árvore. Agora um caminho da raiz até uma folha representa uma regra de classificação. Para cada sub-árvore à esquerda tomadas o predicado do nó é unido ao predicado anterior e para cada sub-árvore direita é tomada a negação do predicado.

A seleção de características foi dividida em dois experimentos: o primeiro faz a fusão dos 5 atributos do vetor de características de vídeo (equação 5.3) com o 14 atributos do vetor de características de som (equação 5.4) conforme descrito anteriormente; e o segundo, ao invés de usar o vetor de atributos de som extraídos do espectro multifractal, vai utilizar o próprio espectro, ou seja, os valores 41 atributos de $f(\alpha)$ e de α , para os intervalos de $-2 < q < 2$ com passos de 0,1. Neste caso, estudou-se o uso do espectro de singularidades para todos os valores de q , identificando quais parâmetros descrevem as auto-similaridades do sinal. Uma vez identificadas as variáveis mais relevantes, poderia ser simplificado o espectro de singularidades usando estes valores ao invés dos definidos na equação 5.4. A fusão foi realizada pela seleção de características utilizando o *Wrapper* e CFS, com busca exaustiva, quando possível, caso contrário com busca gulosa.

No primeiro experimento, os atributos selecionados foram: Δf , $\Delta f_{\max}$, $\Delta f_{\min}$, $f(0)-f(-1)$, $\alpha(0)-\alpha(-1)$. No segundo, selecionou as seguintes características: $\alpha(q=-1.6)$, $\alpha(q=1.7)$, $f(\alpha(q=-2))$ e $f(q=1.8)$. Nota-se, neste caso, que os atributos foram reduzidos drasticamente e os atributos de vídeo não foram selecionados, utilizando o *Wrapper*. Verifica-se que existiu um desbalanceamento no tamanho do vetor de características

de vídeo quando comparado com vetor de características de som, ou seja, o número de atributos de som é bem maior que o número de atributos de vídeo. No processo de seleção de características, recomenda-se as classes distintas devem ter aproximadamente o mesmo número amostras, pois caso contrário podem inferir erro no classificador ou apresentar tendência de classificação para a classe com maior distribuição. O mesmo pode acontecer se forem considerados atributos de diferentes fontes de forma desbalanceada. No entanto, o que se esperava era que os descritores de som e vídeo, apesar de número diferentes, seriam os mais representativos do comportamento, o que bastaria para a seleção de características.

Usando o CFS o atributo Tortuosidade dos atributos do rastreamento de vídeo foi selecionado como sendo relevante. No entanto, a precisão das classes foi bem inferior.

Após a execução do algoritmo de seleção de características, a melhor solução de classificação com o algoritmo C4.5 resultou 99,2% de precisão. Este classificador foi gerado usando a validação cruzada com 10 *folds*.

A Tabela 5.3 apresenta os resultados da análise de atributos para os dois experimentos.

TABELA 5.3: Resultados da análise dos experimentos com e sem o Wrapper

	<i>Experimentos</i>			
	<i>Primeiro</i>	<i>Segundo</i>	<i>Primeiro wrapper</i>	<i>Segundo wrapper</i>
<i>Instancia classificada correta</i>	67.37%	39.07%	80,99%	97.61%
<i>Erro médio</i>	0.25%	0.47%	0.15%	0.01%

A classificação correta não resultou em uma boa precisão para todas as classes. A Tabela 5.4 apresenta os valores de precisão obtidos para cada classe de análise, ou seja, situações onde o macho cantava, apresentava rivalidade, a fêmea cantava e o acasalamento. A rivalidade é o momento cujo espectro multifractal melhor descreveu os resultados e a precisão foi melhor.

TABELA 5.4: Precisão de cada classe analisada

Primeiro Experimento				
	Canto Macho	Canto fêmea	Rivalidade	Acasalamento
<i>Com Wrapper</i>	59.9%	45.1%	67.1%	27.9%
<i>Sem Wrapper</i>	68.3%	55.2%	87.8%	37.1%
Segundo Experimento				
<i>Com wrapper</i>	88.43%	78.23%	99.1%	98.3%
<i>Sem wrapper</i>	40.1%	65.2%	75.2%	65.7%

Observando-se os testes realizados com os vetores de atributos obtidos, e os resultados da seleção de características, verificou-se que a partir dos gráficos $f(\alpha) \times \alpha$, pode-se simplificar a análise dos atributos multifractais. Esta simplificação proposta foi, ao invés do espectro multifractal completo ou dos atributos Vc_{som} , uma aproximação do espectro multifractal por um triângulo. Os pontos do triângulo foram determinados pelos valores de $(-q \text{ max}, f(-q(\text{max})))$ e $(q \text{ máx}, f(q(\text{máx})))$, bem como o valor de $(q=0, f(q(0)))$. Foram incluídos no projeto estes valores e executado novamente o processo de classificação. Na Figura 5.20 é possível visualizar o triângulo determinado. A seleção de características utilizou os seguintes atributos utilizados: Vc_{video} + *lados do triângulo com vértices em $f(-q_{max})$, $f(q_{max})$ e $f(q=0)$.*

Na seleção de características foram selecionadas as características $\alpha(q=-1,5)$, $\alpha(q=1,6)$, $f(q=-1,8)$, $f(q=1,5)$ e dw e T , somente com a abordagem *wrapper*. Nas Tabelas 5.5 e 5.6 podem ser observados os resultados de precisão encontrados.

Figura 5.20 – Exemplo do Triângulo multifractal

TABELA 5.5 - Resultados da análise dois experimentos com e sem o Wrapper

Terceiro Experimento com Wrapper

<i>Instancia classificada correta</i>	79,89%
<i>Erro médio</i>	0.12%

TABELA 5.6 - Precisão de cada classe analisada no terceiro experimento

	Classes			
	Canto Macho	Canto fêmea	Rivalidade	Acasalamento
<i>Com Wrapper</i>	79.9%	42.1%	67.1%	47.9%

5.5 Considerações Finais

Neste capítulo foram apresentados os resultados obtidos, com os softwares implementados. A quantidade de dados extraídos formaram uma excelente base de dados que explorada pelos métodos de mineração de dados puderam chegar a resultados de classificação muito eficiente, apesar da literatura apontar para métodos classificadores mais precisos. O objetivo foi avaliar um classificador cujas superfícies de separação não fossem tão complexas, e supostamente menos precisos, para depois avançar para outros classificadores. No entanto, pelos dados apresentados percebe-se que os atributos de som foram relevantes e, possivelmente, determinantes nos resultados do classificador, não necessitando explorar outro algoritmo classificador. Em artigos que trabalham com apenas som (POTAMATIS, 2006) foi possível fazer o reconhecimento de padrões de sons de insetos de forma automática, porém com taxas de acertos inferiores aos obtidos neste trabalho. A técnica multifractal a partir da transformada Wavelet se mostrou bastante interessante na extração dos atributos pois foi possível caracterizar claramente os comportamentos dos insetos. Os atributos de vídeo provavelmente não foram relevantes pois alguns comportamentos, só poderiam ser observados pela emissão de som, uma vez que não havia movimentação dos insetos.

A quantidade de dados para processar e analisar exigiu muito tempo. Para uso automático, ainda é necessário automatizar alguns passos, como por exemplo a escolha dos parâmetros fractais. Foram 5 fitas de vídeo com mais de 2h de filmagem, juntamente com inúmeros trechos de sons, ambos com os devidos registros de comportamento interpretados por um especialista, formando o “*ground-truth*” do trabalho. Deste conjunto, foram selecionados os exemplos representativos para aplicação das técnicas propostas. Para análise automática, será necessário posteriormente identificar automaticamente os trechos padrões de sinais, como forma de trabalho futuro.

Percebeu-se que os resultados são promissores e que deverão apresentar uma solução para o monitoramento automático do comportamento de insetos. Os pacotes de software gerados estão em fase de pedido de patente. Além dos resultados já obtidos e submetidos para publicação outros tipos de dados estão sendo processados e deverão gerar também artigos inéditos.

No próximo capítulo são apresentadas conclusões deste trabalho, bem como a proposta de novos trabalhos.

Capítulo 6

Conclusões e Trabalhos Futuros

6.1 Conclusões

No presente trabalho de doutoramento foi desenvolvida uma metodologia inédita de fusão de som e vídeo, incluindo técnicas de processamento de vídeo digital, processamento de som, reconhecimento de padrões e técnicas de mineração de dados, além das técnicas de fusão de sensores ou de sinais para estudo do comportamento de insetos.

A metodologia desenvolvida é de grande importância para entender o comportamento de insetos e avaliar procedimentos para o controle de pragas na lavoura. A melhor compreensão da reprodução de insetos e dos sinais envolvidos, por exemplo, dos percevejos presentes nas culturas de soja e milho, pode auxiliar no desenvolvimento de novos procedimentos de manejo de pragas através da utilização de sinais químicos (feromônios) e físicos (sons vibracionais) para monitoramento e/ou controle.

Este trabalho aplicou técnicas inovadoras para seleção de características no espaço de atributos, como a Wavelet-Multifractal e detector de Harris, e a fusão dos mesmos através da seleção de características mais relevantes para caracterizar o comportamento dos insetos, fazendo com que os acertos fossem superiores a 90%.

Foram contribuições inéditas neste trabalho a forma de fusão “inteligente” dos parâmetros obtidos por diferentes fontes, som e vídeo. O uso de seleção de características utilizando técnicas de mineração de dados permitiu obter características que contribuem significativamente para se gerar um classificador automático, evitando que o número excessivo de atributos irrelevantes ou mesmo de atributos correlacionados prejudique no processo de classificação. Apesar dos atributos de vídeo terem apresentado pouca expressão no resultado final, uma vez que foram praticamente excluídos dos atributos depois de aplicada a técnica de seleção de características, a forma de detecção de movimento foi bastante rápida e eficiente, permitindo identificar os pontos de movimentação do inseto de forma prática.

Os atributos de vídeo provavelmente influenciaram menos por duas razões, por apresentarem um número menor no vetor de atributos final ou por não apresentar movimentação ou deslocamento significativo nos comportamentos observados. No caso do som, a técnica de extração de atributos por meio da identificação de auto similaridade nos sinais foi determinante para obtenção dos resultados com altas taxas de acerto. Provavelmente, nos comportamentos observados no trabalho, o som foi a forma determinante no comportamento.

Outra contribuição deste trabalho foi a aplicação da técnica Wavelet-Multifractal de forma inédita para sinais sonoros de insetos, mostrando enorme potencial discriminatório. Tradicionalmente a análise de sinais sonoros avalia principalmente a distribuição de frequências, sem levar em conta o momento no tempo onde a frequência ocorre. A identificação de padrões de auto similaridade, permite identificar uma assinatura típica do comportamento de forma precisa e permite interpretar o sinal como um processo físico modelado dinamicamente. A interpretação do sinal multifractal com a simplificação do triângulo proposto permite que a análise possa ser automatizada com menor custo computacional. No caso dos insetos, comportamentos do tipo perturbadores, como a rivalidade entre machos, disputas no acasalamento, etc. apresentam mudanças na estabilidade do processo físico, representando uma alteração significativa nos valores de entropia do sistema.

Por fim, este trabalho mostra como os sons emitidos por insetos podem ser modelados como processos físicos com características bem definidas ao longo do tempo. Também o modelo de simplificação do espectro multifractal por uma análise de triângulos, reduz enormemente a complexidade de análise automática do espectro multifractal e é outra contribuição marcante do trabalho.

Alguns pontos ainda merecem mais atenção, ficando como sugestões para trabalhos futuros.

6.2 Propostas de trabalhos futuros

Depois de estudar as diferentes técnicas e as respectivas limitações e soluções abordadas, considera-se como propostas futuras os seguintes pontos:

- Aplicar as mesmas técnicas em diferentes animais e comportamentos;
- Automatizar o processo de seleção dos momentos estatísticos utilizados na técnica multifractal;

- Testar diferentes algoritmos de aprendizado de máquina para a seleção das características e fusão;
- Incrementar também a fusão no nível de decisão, para tornar a análise não supervisionada.
- Desenvolver versão paralela dos vários algoritmos visando diminuir o tempo de processamento e viabilizar um número maior de análises.

Referências Bibliográficas

- AGGARVAL, J. K.; CAI, Q. Human motion analysis: A review. **Computer Vision and Image Understanding**, v. 73, n. 3, p. 428-440, 1999.
- ADDISON, P. S. **The illustrated wavelet transform handbook**. 1. ed. Institute of Physics Publishing, 2002. 351 p.
- ARABI, P.; SHI, G.; SHANECHI, M. M.; RABI, S. A. **Phase-based speech processing**. 1. ed. Singapore: B & JO Enterprise, 2006. 141p.
- ARNEODO, A.; BACRY, E.; MUZY, J. F. The thermodynamics of fractals revisited with *wavelets*. **Physica A**, v. 213, p. 232-275, 1995.
- ARNEODO, A.; BACRY, E.; JAFFARD, S.; MUZY, J. F. Oscillating singularities on cantor sets: A grand-canonical multifractal formalism. **Journal of Statistical Physics**, v. 87, n. 1-2, p. 179-209, 1997.
- ARNEODO, A.; DECOSTER, N.; DESTENER, P.; ROUX, S. G. A wavelet-based method for multifractal image analysis: from theoretical concepts to experimental applications. **Advances in Imaging and Electron Physics**, v. 126, p. 1-92, 2003.
- BARANAUSKAS, J. A. **Extração automática de conhecimento por múltiplos indutores**. 2001. 181 f. Tese (Doutorado em Ciências da Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2001.
- BARRON, J. L.; FLEET, D. J.; BEAUCHEMIN, S. S. Performance of optical flow techniques. **International Journal of Computer Vision**, v. 12, n. 1, p. 43-77, 1994.
- BATTIATO, S.; FARINELLA, G. M.; GUARNERA, G. C.; PUGLISI, G.; AZZARO, G.; CACCAMO, M.; LICITRA, G.; FERGUSON, J. D. Estimation of cow's body condition score from images In: VISUAL OBSERVATION AND ANALYSIS OF ANIMAL AND INSECT BEHAVIOR, VAIB 2010, Istanbul, Turkey, **Proceedings...** Isanbul: Istanbul Convention & Exhibition Centre, 2010. (unpaged).
- BETROUNI, N.; LOPES, R. Fractal and multifractal analysis: A review. **Medical Image Analysis**, v. 13, p. 634-649, 2009.
- BLUM, A. L.; LANGLEY, P. Selection of relevant features and examples in machine learning. **Artificial Intelligence**, Amsterdam, p. 245-271, 1997.
- BUGEAU, A.; PÉREZ, P. Detection and segmentation of moving objects in complex scenes. **Journal Computer Vision and Image Understanding**, New York, v. 113, n. 4, p. 459-476, 2009.
- CATCHPOLE, C. K.; SLATER, P. J. B. **Bird Song: Biological Themes and Variation**. Cambridge: University Press, 1995.
- CHHABRA, A. B.; JENSEN, R. V. Direct determination of the $f(\alpha)$ singularity spectrum.

Physical Review Letters, v. 62, n. 12, p. 1327-1330, 1989.

CHHABRA, A. B.; MENEVEAU, C.; JENSEN, R. V.; SREENIVASAN, K. R. Direct determination of the $f(\alpha)$ singularity spectrum and its application to fully developed turbulence. **Physical Review A**, v. 40, n. 9, p. 5284-5294, 1989.

CHEN, Z.; MAHER, R. C. Semi-automatic classification of bird vocalizations using spectral peak tracks. **The Journal of the Acoustical Society of America**, v. 120, n. 5, p. 2974-2984, 2006.

CLINE, D. E.; EDGINGTON, D. R. A detection, tracking and classification system for underwater images. In: INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, 20, 2010, Istanbul, Turkey. **Proceedings...**Istanbul: University Turkey, 2010. (unpaged).

COLAZZA, S.; FUCCARINO, A.; PERI, E.; SALERNO, G.; CONTI, E.; BIN, F. A video tracking and motion analysis system for Linux. In: WORKSHOP INTERNATIONAL ENTOMOPHAGOUS INSECT, 12, 1999, Asilomar, CA, U.S.A. **Proceedings...**Asilomar, CA: 1999. p. 26-30.

DAUBECHIES, I. **Ten lectures on wavelets**. 1. ed. Philadelphia: Society for Industrial and Applied Mathematics, 1992, 357 p.

DAVIS, S. B.; MERMELSTEIN, P. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. **IEEE Transactions on Acoustics, Speech and Signal Processing**, v. 28, n. 4, p. 357-366, 1980.

DENG, L.; O'SHAUGHNESSY, D. **Speech Processing**. Marcel Dekker, v. 17, 626 p. 2003

DICKINSON, P.; QING, C.; LAWSON, S.; FREEMAN, R. Automated visual monitoring of nesting seabirds In: VISUAL OBSERVATION AND ANALYSIS OF ANIMAL AND INSECT BEHAVIOR, VAIB, 2010, Istanbul, Turkey, **Proceedings...**Istanbul: Istanbul Convention & Exhibition Centre, 2010. (unpaged).

EILAM, D. Open-field behavior withstands drastic changes in arena size. **Behavioural Brain Research**, v. 142 p. 53-62, 2003.

EVANS, W.R.. *Applications of Acoustic Bird Monitoring for the Wind Power Industry*. Proceedings of the National Avian-Wind Power Planning Meeting III, San Diego, USA, 2000.

ERIKSSON, T.; SJOBERG, J. Dynamic bit allocation in CELP excitation coding. In: INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, **Proceedings...**1993, v.2, p. 171-174.

EVANS, W.R.. Applications of Acoustic Bird Monitoring for the Wind Power Industry. In: PROCEEDINGS OF THE NATIONAL AVIAN-WIND POWER PLANNING MEETING, 3, 2000, San Diego, USA, **Proceedings...**San Diego: 2000. (unpaged).

FAGERLUND, S.; HÄRMÄ, A. Parametrization of inharmonic bird sounds for automatic recognition. In: EUROPEAN SIGNAL PROCESSING CONFERENCE – EUSIPCO, 13., 2005, Turquia, **Proceeding...** Turquia: 2005. (unpaged).

FAGHFOURI, A.; KINSNER, W. Local and global analysis of multifractal singularity spectrum through *wavelets*. In: CONFERENCE ON ELECTRICEL AND COMPUTER ENGINEERING, 18, 2005, Saskatchewan, Canadá. **Proceeding...**Saskatchewan: 2005. p. 2163-2169.

FANDIÑO-MARIÑO, H.; VIELLIARD, J. M. E. Complex communication Signals: The case of the blue-black grassquit volatinia jacarina (Aves, Emverizidae) Song. Part I – A Structural Analysis. **Anais da Academia Brasileira de Ciências**, v. 76, n. 2, p. 325-334, 2004.

FERREIRA, E. J.; JORGE, L. A. C. **Seleção de características aplicada ao processamento de imagens digitais.**, 2007. 21 p. (Embrapa Instrumentação Agropecuária. Documentos, 33).

FISHER, R.; XIAO, Y. Bat echolocation behavior from highspeed 3D video In: VISUAL OBSERVATION AND ANALYSIS OF ANIMAL AND INSECT BEHAVIOR, VAIB 2010, Istanbul. Turkey, **Anais...** Isanbul: Istanbul Convention & Exhibition Centre, 2010. (Não paginado).

FONSECA, E. S.; GUIDO, R. C.; SCALASSARA, P. R.; MACIEL, C. D.; PEREIRA, J. C. *Wavelet* time-frequency analysis and least squares support vector machines for the identification of voice disorders. **Computers in Biology and Medicine**, v. 37, n. 4, p. 571-578, 2007.

FRY, S. S. N.; BICHSEL, M.; MULLER, P.; ROBERT, D. Tracking of flying insects using pan-tilt cameras. **Jornal Neuroscience Methods**, v. 101, p. 59-67, 2000.

GALVIN, B.; MCCANE, B.; NOVINS, K.; MASON, D.; MILL, S. Recovering motion fields: an evaluation of eight optical flow algorithms In: BRITISH MACHINE VISION CONFERENCE. PROCEEDINGS, 1998, **Proceeding...** 1988. p. 195-204.

GANCHEV, T.; POTAMITIS, I.; FAKOTAKIS, N. Acoustic monitoring of singing insects. In: INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, 4., 2007, Hawai, USA. **Proceeding...**Hawai:, 2007. p. 721-724.

GIGLI, G.; BOSSÉ, E.; LAMPROPOULOS, G.A. An optimized architecture for classification combining data fusion and data mining. **Information Fusion**. v.8, n.4, 2007

GONÇALVES, C.; PEREIRA, L. A.; PIZOLATO JR, J. C.; GONZAGA, A. Reconhecimento de tipos de movimento humano baseados no método GMM. In: Workshop de Visão Computacional, **Anais...** 2006 . (CD-ROM).

GONÇALVES, W. N.; SILVA, J. A.; MACHADO, B. B.; PISTORI, H.; SOUZA, A. S. Modelos ocultos de Markov aplicados na identificação de comportamento de serpentes In: WORKSHOP DE VISÃO COMPUTACIONAL, 3, 2007, São José do Rio Preto, SP. **Anais...**São José do Rio Preto:UNESP, 2007. p. 324-329.

GRAPS, A. An introduction to wavelets. **IEEE Computational Science and Engineering Summer**, Los Alamitos, v. 2, n. 2, p. 1-18, 1995.

GREEN, J.; FRENCH, A. P.; BRIGNELL, C.; HUXLEY, J.; PRIDMORE, T. P. Tools and resources for cow behavior monitoring. In: VISUAL OBSERVATION AND ANALYSIS

OF ANIMAL AND INSECT BEHAVIOR, VAIB 2010, Istanbul, Turkey, **Proceeding...** Isanbul: Istanbul Convention & Exhibition Centre, 2010. (unpaged).

GYORY, G.; RANKOV, V.; GORDON, G.; PERKON, I.; MITCHINSON, B.; GRANT, R.; PRESCOTT, T. An algorithm for automatic tracking of rat whiskers In: VISUAL OBSERVATION AND ANALYSIS OF ANIMAL AND INSECT BEHAVIOR, VAIB, 2010, Istanbul, Turkey, **Proceedings...** Isanbul: Istanbul Convention & Exhibition Centre, 2010. (unpaged).

HADDJITODOROV, S.; BOYANOV, B.; TESTON, B. Laryngeal pathology detection by means of class-specific neural maps. **IEEE Trans. on Information Technology in Biomedicine**, v. 4, n. 1, p. 89-93, 2000.

HALL, D. L.; LLINAS, J. **Hand book of multisensory datafusion** 1. ed. New York: CRC Press, 2001. 501 p.

HARRIS, C.; STEPHENS, M. A. Combined corner and edge detector. In: PROC. ALVEY VISION CONF., 4, 1988. **Proceeding...** 1988. v. 15, p. 147–151.

HELLER, J. R. **Automatic Recognition of Harmonic Bird Sounds**. 2005. 102 p. Tese (Doutorado) - Stony Brook University, EUA, 2005.

HERMANSKY, H. Perceptual linear predictive analysis for speech recognition. **Journal of Acoustical Society of America**, v. 87, n. 4, p. 1738-1752, 1990

HORN, B. K. P.; SCHUNCK, B. G. Determining optical flow. **Artificial Intelligence**, v. 17, p. 185-203, 1981.

HOY, J. B.; KOEHLER, P. G.; PATTERSON, R. S. A microcomputer-based system for real-time analysis of animal movement. **J. of Neuroscience Methods**, v. 64, p. 157-161, 1996.

HURST, H. E. Long-term storage capacity of reservoirs. **Trans. Am. Soc. Civ. Eng.**, v. 116, p. 770-808, 1951.

HUANG, X.; ACERO, A.; HON, H. **Spoken language processing: A guide to theory, algorithm, and system development**. Prentice Hall, 2001. 980p.

IVANOV, P. C.; AMARAL, L.A.N.; GOLDBERGER, A.L.; HAVLIN, S.; ROSENBLUM, M.G.; STRUZIK, Z.R.; STANLEY, H.E.. Multifractality in healthy heartbeat dynamics. **Nature**, v. 399, p. 461-465, 1999.

JHUANG, H.; GARROTE, E.; POGGIO, T.; STEELE, A. D; SERRE, T. Vision-based automated recognition of mice home-cage behaviors In: VISUAL OBSERVATION AND ANALYSIS OF ANIMAL AND INSECT BEHAVIOR, VAIB, 2010, Istanbul, Turkey, **Proceeding...** Isanbul: Istanbul Convention & Exhibition Centre, 2010. (unpaged).

JORDAN, D.; MIKSAD, R. W. Implementation of the continuous wavelet transform for digital time series analysis. **Rev. Sci. Instrum.**, v.68, n.3, p.1484-1489. 1997.

JORGE, L. A. C.; CRUZ, R. A.; MORAES, M. C. B.; LAUMANN, R.; PALHARES, L.; BORGES, M. Software para análise comportamental de insetos em movimento. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 5, SIMPÓSIO BRASILEIRO DE TECNOLOGIA DA INFORMAÇÃO NO AGRONEGÓCIO COOPERATIVO-

AGRONEGÓCIO, TECNOLOGIA E INOVAÇÃO-SBI-AGRO, 2005, Londrina, PR.
Anais... não paginado – 6f.

JORGE, L. A. C.; RODA, O. V.; MORAES, M. C. B.; LAUMANN, R.; BORGES, M.; IRBER, L. C.; MILARE, R. N. Sistema para análise comportamental de insetos baseado na fusão de sensores: som e imagem: primeiros resultados. In: WORKSHOP DE VISÃO COMPUTACIONAL, WVC, 2., 2006, São Carlos, SP. **Anais...** São Carlos: EESC/USP, 2006. v. CD ROM p. 178-183.

JORGE, L. A. C.; RUIZ, H. S.; FERREIRA, E. J. GONZAGA, A. Wrapper Approach to Select a Subset of Color Components for Image Segmentation with Photometric Variations. In: Brazilian Symposium on Computer Graphics and Image Processing, SIBGRAPI 2007. Belo Horizonte, MG. **Proceedings...** Los Amigos: IEEE Computer Society, 2007. p. 245-252. ISSN: 1530-1834. ISBN: 978-0-7695-2996-7

JORGE, L. A. C.; POSADAS, D. A. N.; NAKAMURA, R.; GUIMARÃES, M. F.; RODA, V. O.; CURMI, P. **Aplicação da Técnica Multifractal para Caracterização de Manejo de Solo** 2008. 35 p. (Embrapa Instrumentação Agropecuária. Documentos, 41).

KOGAN, J. A.; MARGOLIASH, D. Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden markov models: a comparative study. **The Journal of the Acoustical Society of America**, v. 103, n. 4, p. 2185-2196, 1998.

KOHAVER, R.; JOHN, G. Wrappers for feature subset selection. **Artificial Intelligence**, Amsterdam, v. 97, n. 1-2, p. 273-324, 1997.

KO, T.; SOATTO, S.; ESTRIN, D.; CENEDESE, A. Cataloging birds in their natural habitat In: VISUAL OBSERVATION AND ANALYSIS OF ANIMAL AND INSECT BEHAVIOR, VAIB, 2010, Istanbul, Turkey, **Proceedings...** Istanbul: Istanbul Convention & Exhibition Centre, 2010. (unpaged).

KOTTELY, B. **The Art and Business of Speech Recognition: Creating the Noble Voice**. Addison Wesley, 2003. 208 p.

KREBS, J. R.; KROODSMA, D. E. Repertoires and geographical variation in bird song. **Journal of Advances in the Study of Behavior**, v. 11, p. 143-177, 1980.

KUMAR, N. R.; JANAKIRAMAN, T. N.; THIAGARAJAN, H.; SUBAHARAN, K. Automated motion tracking of insects using invariant moments in image sequence. In: INTERNATIONAL CONFERENCE ON COMPUTATIONAL INTELLIGENCE AND MULTIMEDIA APPLICATIONS. **Proceeding...** 2007, p. 220-224.

KUMAR, R.; MARTIN, V.; MOISAN, S. Robust insect classification applied to real time greenhouse infestation monitoring. In: INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, 20, 2010, Istanbul, Turkey. **Proceedings...** Istanbul: University Turkey, 2010. (unpaged).

KWAN, C.; HO, K.; MEI, G.; LI, Y.; REN, Z.; XU, R.; ZHAO, G.; STEVENSON, M.; STANFORD, V.; ROCHET, C. An automated acoustic system to monitor and classify birds. **Journal on Applied Signal Processing**, p. 1-19, 2006.

LATKA, M., GLAUBIC-LATKA, M.; LATKA, D.; WEST, B. The loss of multifractality in migraines. 2002. Disponível em:

<http://arxiv.org/PS_cache/physics/pdf/0204/0204010v1.pdf>. Acesso em: 15 fev. 2009.

LAUMANN, R. A.; MORAES, M. C. B.; COKL, A.; BORGES, M. **Metodologias para estudo da comunicação vibracional de insetos e sua aplicação ao manejo de pragas** Brasília: Embrapa Recursos Genéticos e Biotecnologia, 2005. 30p. (Embrapa Recursos Genéticos e Biotecnologia. Documentos, 152).

LAUMANN, R. A.; MORAES, M. C. B.; COKL, A.; BORGES, M. Eavesdropping on sexual vibratory signals of stink bugs (Hemiptera: Pentatomidae) by the egg parasitoid *Telenomus podisi*. **Animal Behaviour**, v. 73 n. 4, p. 637-649, 2007.

LEE, H. D.; MONARD, M. C.; BARANAUSKAS, J. A. **Empirical comparison of wrapper and filter approaches for feature subset selection**. 1999. 46 p. (USP-ICMC. Relatórios técnicos, 94).

LÉVESQUE, O.; BERGEVIN, R. Detection and identification of animals using stereo vision In: VISUAL OBSERVATION AND ANALYSIS OF ANIMAL AND INSECT BEHAVIOR, VAIB, 2010, Istanbul, Turkey, **Proceedings...** Isanbul: Istanbul Convention & Exhibition Centre, 2010. (unpaged).

LIN, D. C.; SHARIF, A. Wavelet transform modulus maxima based fractal correlation analysis. **Eur. Phys. J. B.**, v. 60, p. 483-491, 2007.

LOPES, R.; BETROUNI, N. Fractal and multifractal analysis: A review. **Medica Image Analysis**, v.13, p. 634-649, 2009.

LOWE, D. G. Distinctive image features from scale-invariant keypoints. **Int. J. Computer Vision**, v. 60, n. 2, p. 91-110, 2004.

LUCAS, B. D.; KANADE, T. An iterative image registration technique with application to stereo vision. In: PROC. OF IMAGE UNDERSTANDING WORKSHOP, 1981, **Proceeding...** 1981. p. 121-130.

MA, L.; HUANG, W.; SHI, Y.; YU, H.; ZHU, J. Multifractal analysis of particle-fluid system in a circulating fluidized bed **China Particology**, v. 3, n. 1-2, p. 80-83, 2005.

MALLAT, S.; HWANG, W. Singularity detection and processing with wavelet. **IEEE Transactions on Information Theory**, v. 38, n. 2, p. 617-643, 1992.

MANDELBROT, B. **The Fractal Geometry of Nature**. 1. ed. New York: W. H. Freeman, 1982. 480 p.

MAURER, M.; RÜTHER, M.; BISCHOF, H. Rigid body reconstruction for motion analysis of giant honey bees using stereo vision. In: INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, 20, 2010, Istambul, Turkey. **Proceedings...** Istambul: University Turkey, 2010. (unpaged).

MCATEER, J. R. T.; YOUNG, C. A.; IRELAND, J.; GALLAGHER, P. T. The bursty nature of solar flare x-ray emission. **The Astrophysical Journal**, v. 662, p. 691-700, 2007.

MILLER, D. C.; LANG, W. H.; GREAVES, J. O. B.; WILSON, R. S. Investigation in aquatic behavioral toxicology using a computerized video quantification system. In: J.G. Pearson, R.B. Foster and W.E. Bishop (Eds.) *Aquatic Toxicology and Hazard*

Assesment: Fifth Conference, ASTM STP 766, **American Society for Testing and Materials**, p. 106-220, 1982.

MISITI, M.; MISITI, Y.; OPPENHEIM, G.; POGGI, J.-M. **Wavelet Toolbox: User's Guide**. Natick, MA: The MathWorks, 934 p., 2002.

MORAES, M. C. B.; LAUMANN, R. A.; ČOKL, A.; BORGES, M. Vibratory signals of four Neotropical stink bug species. **Physiological Entomology**, Oxford, v.30, n.2, p. 175-188, 2005.

MULLER, P.; ROBERT, D. A shot in the dark: the silent quest of a free-flying phonotactic fly. **The Journal of experimental biology**, v. 204, p. 1039-1052, 2001.

MUZY, J. F.; BACRY, E.; ARNEODO, A. *Wavelets* and multifractal formalism for singular signals: Application to turbulence data. **Physical Review Letters**, v. 67 n. 25, p. 3515-3518, 1991.

NICKERSON, C. M.; BLOOMFIELD, L. L.; DAWSON, M. R. W. Artificial neural network discrimination of black-capped chickadee (*Poecile atricapillus*) call notes. **The Journal of the Acoustical Society of America**, v. 120, n. 2, p. 1111-1117, 2006.

NOLDUS, L. P. J. J.; SPINK, A. J. ; TEGELENBOSCH, R. A. J. Computerised video tracking, movement analysis and behavior recognition in insects. **Computers and Electronics in Agriculture**, v. 35, p. 201-227, 2002a.

NOLDUS, L. P. J. J.; TRIENES, R. J. H.; HENDRIKSEN, A. H. M.; JANSEN, H.; JANSEN, R. G. The Observer Video-Pro: New software for the collection, management, and presentation of time-structured data from videotapes and digital media files. **Behavior Research Methods Instruments & Computers**, v. 32, n. 1, p. 197-206, 2002b.

PANDORFI, H. **Avaliação do comportamento de leitões em diferentes sistemas de aquecimento por meio da análise de imagem e identificação eletrônica**. 2002. 54 f. Dissertação de Mestrado, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo – Piracicaba, 2002.

PISTORI, H.; ODAKURA, V. V. V. A. ; MONTEIRO, J. B. O.; GONÇALVES, W. N.; ROEL, A. R.; SILVA, J. A.; MACHADO, B. B. Mice and larvae tracking using a particle filter with an auto-adjustable observation model. **Pattern Recognition Lett**, v. 31, p. 337-346, 2010.

POSADAS, A. N. D.; GIMÉNEZ, D.; BITTELLI, M.; VAZ, C. M. P.; FLURY, M. Multifractal characterization of soil particle-size distributions. **Soil Science Society of America Journal**, v. 65, p. 1361-1367, 2001.

POSADAS, A. N. D.; GIMÉNEZ, D.; QUIROZ, R.; PROTZ, R. Multifractal characterization of soil pore systems. **Soil Science Society of America Journal**, v. 67, n. 5, manuscript: S02-0294, 2003.

POSADAS, A. N. D.; QUIROZ, R.; ZOROGASTÚA, P. E.; LEÓN-VELARDE, C. Multifractal characterization of the spatial distribution of ulexite in Bolivian salt flat. **International Journal of Remote Sensing**, v. 26, n. 3, p. 615-627, 2005.

POTAMITIS, I.; GANCHEV, T.; FAKOTAKIS, N. *Automatic Acoustic Identification of Insects Inspired by the Speaker Recognition Paradigm*,. In: Inter- Speech – ICSLP,

2006, Pittsburgh, PA, USA, **Proceeding...** Pittsburgh: 2006, p. 2126-2129.
BEHAVIOR RESEARCH METHODS INSTRUMENTS & COMPUTERS 32 (1): 197-206.

QUINLAN, J. R. **C4.5: Programs for machine learning**. 1. ed. San Mateo: Morgan Kaufmann, 1993. 302 p.

RELJIN, I. S.; RELJIN, D.; AVRAMOV-IVIC, M. L.; JOVANOVIĆ, D. V.; PLAVEC, G. I.; PETROVIC, S. D.; BOGDANOVIC, G. M. Multifractal analysis of the UV/VIS spectra of malignant ascites: Confirmation of the diagnostic validity of a clinically evaluated spectral analysis. **Physica A**, v. 387, p. 3563-3573, 2008.

ROSA, M. O.; PEREIRA, J. C.; GRELLET, M. Adaptive estimation of residue signal for voice pathology diagnosis. **IEEE Trans. on Biomedical Engineering**, v. 47, p. 96-104, 2000.

ROS-SÁNCHEZ, G.; GARCÍA-MATEOS, G.; VERA, L. M.; SÁNCHEZ-VÁZQUEZ, F. J. A new taxonomy and graphical representation for visual fish analysis with a case study. In: VISUAL OBSERVATION AND ANALYSIS OF ANIMAL AND INSECT BEHAVIOR, VAIB, 2010, Istanbul, Turkey, **Proceedings...** Istanbul: Istanbul Convention & Exhibition Centre, 2010. (unpaged).

RUSSEL, D. A.; HANSON, J. E.; OTT, E. Dimensionality and Lyapunov numbers of strange attractors. **Fis. Rev. Lett.** V. 45, p. 1175-1778, 1980.

SCHERTZER, D.; LOVEJOY, S. Multifractal generation of self-organized criticality, In: NOVAK, M. M. **Fractals in the natural and applied sciences**. Ed. Elsevier, North-Holland, 1994. p. 325-339.

SELIN, A.; TURUNEN, J.; TANTTU, J. T. Wavelets in recognition of bird sounds. **EURASIP Journal on Advances in Signal Processing**, 2007.

SELOUANI, S. A.; KARDOUCHI, M.; HERVET, E.; ROY, D. Automatic birdsong recognition based on autoregressive time-delay neural networks. In: CONFERENCE ON COMPUTATIONAL INTELLIGENCE: METHODS AND APPLICATIONS - IEEE ICSC, 2005, **Proceeding...** 2005. p. 356-369.

SIMAKOV, D.; FROLOVA, D. Matching with invariant features. 2004. Disponível em: <http://www.wisdom.weizmann.ac.il/~dariaf/InvariantFeatures.ppt>. Acesso em: 20 out 2007.

SOEMINTAPURA, K.; LANGI, A. Z. R. Wavelet-multifractal processing of signals. In: IEEE ASIA-PACIFIC CONFERENCE ON CIRCUITS AND SYSTEMS, 1988, Chiangmai, Thailand, **Proceeding...** Chiangmai, 1988. p. 65-68.

SPINK, A. J.; TEGELENBOSCH, R. A. J.; BUMA, M. O. S.; NOLDUS, L. P. J. J. The etho vision video tracking system – A tool for behavioral phenotyping of transgenic mice. **Physiology & Behavior**, v. 73, p. 731-744, 2001.

STAUFFER, C.; GRIMSON, W. E. L. Learning patterns of activity using real-time tracking. **IEEE Transactions on Pattern Analysis & Machine Intelligence**, v. 22, n. 8, p. 747-757, 2000.

STAVENS, D.; THRUN, S. Unsupervised learning of invariant features using video. In: COMPUTER VISION AND PATTERN RECOGNITION – CVPR, 2010, San Francisco, CA, USA, **Proceedings...** San Francisco: 2010 p.1649-1656.

STRUZIK, Z. R. Determining local singularity strengths and their spectra with the wavelet transform. **Fractals**, Singapore, v. 8, p. 163-179, 2000.

SVEN, A. E.; DAVE, A. S.; MARBOLIAH, D. Template-based automatic recognition of birdsong syllables from continuous recordings. **The Journal of the Acoustical Society of America**, v. 100, n. 2, p. 1209-1219, 1996.

THERIAULT, D.; WU, Z.; HRISTOV, N.; SWARTZ, S. M.; BREUER, K. S.; KUNZ, T.; BETKE, M. Reconstruction and analysis of 3D trajectories of brazilian free-tailed bats in flight In: VISUAL OBSERVATION AND ANALYSIS OF ANIMAL AND INSECT BEHAVIOR, VAIB, 2010, Istanbul, Turkey, **Proceedings...** Isanbul: Istanbul Convention & Exhibition Centre, 2010. (unpaged).

VEERARAGHAVAN, A.; CHELLAPPA, R.; SRINIVASAN, M. Shape-and-behavior-encoded tracking of bee dances. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 30 n. 3, p. 463-476, 2008.

VICSEK, T. **Fractal growth phenomena**. 2nd ed. Singapore: World Scientific Publishing Co., 1992.

VIGNEAULT, C.; PANNETON, B.; RAHAVAN, G. S. V. Real-time digitizing system applied to air bubble generator characterization. **ASAE**, St. Joseph, p. 90-3535, 1990.

VIRANT-DOBERLET, M.; ČOKL, A. Vibrational communication in insects. **Neotropical Entomology**, Vacaria, v. 33, n. 2, p. 121-134, 2004.

WAJNBERG, E., COAZZA, S. 1998. Genetic variability in the area searched by a parasitic wasp: analysis from automatic video tracking of the walking path. **Jornal of Insect physiology**, 44:437-444.

WAWERLA, J.; MARSHALL, S.; MORI, G.; ROTHLEY, K.; SABZMEYDANI, P. BearCam: automated wildlife monitoring at the arctic circle. **Machine Vision and Applications**, v. 20, p. 303-317, 2009.

WEIJER, J.; GEVERS, T.; GEUSEBROEK, J. Edge and corner detection by photometric quasi-invariants. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 27, n. 4, p. 625-630, 2005.

XIAO, Y.; FISHER, R. Bat echolocation behavior from high speed 3D video In: INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, 20, 2010, Istambul, Turkey. **Proceedings...**Istambul: University Turkey, 2010. (unpaged).

YAMAGUTI, M.; PRADO, C. P. C. A direct calculation of the spectrum of singularities $f(\alpha)$ of multifractals. **Physics Letters A**, v. 206, p. 318-322, 1995.

YANG, J.; HONAVAR, V. Feature subset selection using a genetic algorithm. **IEEE Intelligent Systems**, v. 13, n. 2, p. 44-49, 1998.

YE, S.; BELL, W. J. A simple video position-digitizer for studying animal movement patterns. **Journal of Veuroscience Methods**, v. 37, p. 215-225, 1991.

YILMAZ, A.; JAVED, O.; SHAH, M. Object tracking: A survey. ACM. **Comput. Surv.**, V.38, n.4, 45 pages, 2006.

YU, Z. G.; ANH, V. V.; LAU, K. S. Measure representation and multifractal analysis of complete genome. **Phys. Rev.**, v. 64, n. 3, 9 pages, 2001.

ZHOU, D.; ZHANG, H. 2D Shape measurement of multiple moving objects by GMM background modeling and optical flow. **Image Analysis and Recognition Lecture Notes in Computer Science**, v. 3656, p. 789-795, 2005

ZEEUW, P. M.; PAUWELS, E. J. ; RANGUELOVA, E. B.; BUONANTONY, D. M.; ECKERT, S. A. Computer assisted photo identification of *Dermochelys coriacea*. In: VISUAL OBSERVATION AND ANALYSIS OF ANIMAL AND INSECT BEHAVIOR, VAIB, 2010, Istanbul, Turkey, **Proceeding...** Isanbul: Istanbul Convention & Exhibition Centre, 2010. (unpaged).