

**INSTITUTO MILITAR DE ENGENHARIA**

**Ten Cel ANGELO MARCIO CARDOSO RIBEIRO BORZINO**

**PROCESSAMENTO DE SINAIS DE ÁUDIO PROVENIENTE  
DE DISPARO DE ARMA DE FOGO E APLICAÇÕES**

Tese de Doutorado apresentada ao Curso de Doutorado em Engenharia de Defesa do Instituto Militar de Engenharia, como requisito parcial para obtenção do título de Doutor em Ciências em Engenharia de Defesa.

Orientadores: Prof. José Antonio Apolinário Jr. - D.Sc.  
Prof. Marcello Luiz R. de Campos - Ph.D.

Rio de Janeiro  
2016

c2016

INSTITUTO MILITAR DE ENGENHARIA  
Praça General Tibúrcio, 80-Praia Vermelha  
Rio de Janeiro-RJ CEP 22290-270

Este exemplar é de propriedade do Instituto Militar de Engenharia, que poderá incluí-lo em base de dados, armazenar em computador, microfilmar ou adotar qualquer forma de arquivamento.

É permitida a menção, reprodução parcial ou integral e a transmissão entre bibliotecas deste trabalho, sem modificação de seu texto, em qualquer meio que esteja ou venha a ser fixado, para pesquisa acadêmica, comentários e citações, desde que sem finalidade comercial e que seja feita a referência bibliográfica completa.

Os conceitos expressos neste trabalho são de responsabilidade do(s) autor(es) e do(s) orientador(es).

623.3 Borzino, Angelo Marcio Cardoso Ribeiro  
B739p Processamento de sinais de áudio proveniente de disparo de arma de fogo e aplicações / Angelo Marcio Cardoso Ribeiro Borzino; orientado por José Antonio Apolinário Júnior, Marcello Luiz Rodrigues de Campos - Rio de Janeiro: Instituto Militar de Engenharia, 2016.

122 p.: il, graf., tab.

Tese (doutorado) - Instituto Militar de Engenharia- Rio de Janeiro, 2016

1. Curso de Engenharia de Defesa - teses e dissertações. 2. Processamento de sinais. 3. Estimacão de direção de chegada 4. Classificacão de munição. I. Apolinário Júnior, José Antonio. II. Campos, Marcello Luiz Rodrigues de. III. Título. IV. Instituto Militar de Engenharia.

**INSTITUTO MILITAR DE ENGENHARIA**

**TEN CEL ANGELO MARCIO CARDOSO RIBEIRO BORZINO**

**PROCESSAMENTO DE SINAIS DE ÁUDIO PROVENIENTE  
DE DISPARO DE ARMA DE FOGO E APLICAÇÕES**

Tese de Doutorado apresentada ao Curso de Doutorado em Engenharia de Defesa do Instituto Militar de Engenharia, como requisito parcial para obtenção do título de Doutor em Ciências em Engenharia de Defesa.

Orientadores: Prof. José Antonio Apolinário Jr. - D.Sc.

Prof. Marcello Luiz Rodrigues de Campos - Ph.D.

Aprovada em 25 de julho de 2016 pela seguinte Banca Examinadora:

---

Prof. José Antonio Apolinário Jr. - D.Sc. do IME - Presidente

---

Prof. Marcello Luiz Rodrigues de Campos - Ph.D. da UFRJ

---

Prof. Paulo Fernando Ferreira Rosa - Ph.D. do IME

---

Prof. Ernesto Leite Pinto - D.Sc. do IME

---

Prof. Wallace Alves Martins - D.Sc. da UFRJ

Rio de Janeiro  
2016

Dedico esta tese à minha esposa Letícia, aos meus pais Amedeo e Rosane, à minha filha Victória e aos meus filhos(as) que, em breve, farão parte da minha vida, tornando-a mais feliz e prazerosa.

## AGRADECIMENTOS

Agradeço primeiramente a Deus, por me guiar nos caminhos escolhidos e estar sempre presente em minha vida.

Ao Departamento de Ciência e Tecnologia do Exército, pela designação para o curso de doutorado em Engenharia de Defesa no Instituto Militar de Engenharia. Agradeço também em particular ao Gen Bda Pedro Paulo e ao TC Schmeikal por concordarem com minha participação no doutorado sanduíche e pelo empenho em todo o processo administrativo decorrente da viagem.

Ao Instituto Militar de Engenharia, por me dar a oportunidade de realização do curso de doutorado e prover os recursos necessários à pesquisa.

Ao Programa de Pós-Graduação em Engenharia de Defesa, pelo apoio nos procedimentos administrativos e por me acolher no programa.

Aos meus orientadores, Prof. José Antonio Apolinário Jr. e Prof. Marcello Luiz Rodriguez de Campos, por me guiarem na pesquisa, apontando áreas onde eu deveria focar, a fim de obter os resultados necessários para a aprovação no curso. Agradeço também o estímulo para participar de conferências e escrever artigos, assim como o incentivo para a realização do doutorado sanduíche nos Estados Unidos, possibilitando que eu vivenciasse uma experiência única.

Ao Svd Civ Bomfim, do Laboratório de Processamento Digital de Sinais (LPDS), por auxiliar nas gravações de sinais de tiro, enriquecendo assim o conteúdo da tese.

Ao Centro de Avaliações do Exército, em particular aos tenente-coronéis Gatti e Malízia, por cederem espaço, armamentos, munições e pessoal para a realização de gravações de sinais de tiro.

À *Montana State University* (MSU), em particular ao Prof. Robert C. Maher, Chefe do Departamento de Engenharia Elétrica e de Computação da MSU, por me acolher durante quatro meses no doutorado sanduíche, por facilitar minha estadia na cidade de Bozeman e pela parceria na pesquisa a respeito de classificação de armamento e munição, possibilitando a escrita de dois artigos.

À minha esposa Letícia, pela paciência e compreensão sobre o tempo durante o qual tive que ficar “afastado” para a consecução desta pesquisa. Obrigado pelo constante apoio, incentivo e admiração por este trabalho.

À minha mãe, pelo amor, pelos ensinamentos e por me fazer perceber que o estudo

é importante para o sucesso na vida profissional. À minha irmã, também pelo amor e carinho que tem por mim. Ao meu pai (em memória), por possibilitar que eu desse os primeiros passos na vida para me tornar feliz com a profissão escolhida. À minha filha, pelo entendimento da minha ocupação na pesquisa e pelo orgulho que tem de mim.

Aos meus companheiros de laboratório, Manolo, Rigel, Celestino, pelo convívio salutar e descontraído, pelo apoio mútuo e pelos momentos de descontração e conversas informais nos intervalos de nossas pesquisas.

Enfim, a todos que contribuíram de forma direta ou indireta para a conclusão deste trabalho.

*Angelo Marcio Cardoso Ribeiro Borzino*

“No que diz respeito ao empenho, ao compromisso, ao esforço, à dedicação, não existe meio termo: ou você faz uma coisa bem feita ou não faz.”

AYRTON SENNA

## SUMÁRIO

LISTA DE ILUSTRAÇÕES .....	10
LISTA DE TABELAS .....	13
LISTA DE ABREVIATURAS .....	16
<b>1 INTRODUÇÃO .....</b>	<b>19</b>
1.1 Contexto .....	19
1.2 Revisão da literatura .....	19
1.2.1 O sinal de tiro .....	19
1.2.2 Estimação da direção de chegada de sinais de tiro .....	22
1.2.3 Classificação de armamento ou munição .....	27
1.3 Motivação .....	29
1.4 Objetivo .....	29
1.5 Organização da tese .....	30
<b>2 TÉCNICAS DE REDUÇÃO DE RUÍDO .....</b>	<b>31</b>
2.1 Subtração espectral .....	31
2.2 Filtro mediana .....	33
<b>3 A BUSCA EXAUSTIVA .....</b>	<b>38</b>
<b>4 LOOP FUNDAMENTAL CONSISTENTE NA ESTIMAÇÃO DE DOA .....</b>	<b>44</b>
4.1 Motivação .....	44
4.2 Um algoritmo para buscar cFL ( <i>Searching Consistent Fundamental Loop</i> ou SCFL) .....	46
<b>5 O ALGORITMO ES-SCFL .....</b>	<b>50</b>
<b>6 RESULTADOS EXPERIMENTAIS DOS ALGORITMOS DE ESTIMAÇÃO DE DOA .....</b>	<b>56</b>
6.1 Comparando os algoritmos ES-SCFL, ES(5), ES(4) e SCFL .....	56
6.2 Sinais com DOAs simulados .....	57
6.3 Sinais gravados .....	60

<b>7</b>	<b>CLASSIFICAÇÃO DE MUNIÇÃO</b> .....	67
7.1	Métodos .....	67
7.2	Resultados da classificação de munição .....	76
<b>8</b>	<b>A INFLUÊNCIA DA DIREÇÃO DO CANO DA ARMA EM RE- LAÇÃO AO MICROFONE</b> .....	78
8.1	A característica direcional do muzzle blast .....	78
8.2	Análise da primeira base de dados .....	79
8.2.1	Resultados de classificação de munição com as oito armas .....	80
8.3	Inclusão da segunda base de dados .....	85
8.4	Analisando sinais corrompidos .....	94
8.5	Classificação binária .....	95
8.6	Classificação de azimute .....	96
8.7	Outros testes .....	99
8.7.1	Classificação quanto ao porte de arma .....	99
8.7.2	Classificação entre armas de cano curto .....	100
8.7.3	Classificação entre armas de cano longo .....	102
<b>9</b>	<b>CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS</b>	107
9.1	Conclusões .....	107
9.2	Sugestões para trabalhos futuros .....	109
<b>10</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	110
<b>11</b>	<b><u>APÊNDICES</u></b> .....	116
11.1	Apêndice 1: Atributos de textura utilizados neste trabalho .....	117
11.1.1	Primeiro conjunto de atributos .....	117
11.1.2	Segundo conjunto de atributos .....	118
11.1.3	Terceiro conjunto de atributos .....	119
11.2	Apêndice 2: Publicações por capítulo .....	121

## LISTA DE ILUSTRAÇÕES

FIG.1.1	Uma configuração típica para gravação de sinais de tiro (BORZINO, 2016). . . . .	21
FIG.1.2	Geometria da SW para duas velocidades do projétil diferentes. . . . .	21
FIG.1.3	SW e MB originados de um FAL (BORZINO, 2015a). . . . .	22
FIG.1.4	Arranjo de sete microfones utilizado nos experimentos de estimação de DOA e ângulos azimutal e zenital (FREIRE, 2012). . . . .	23
FIG.1.5	Sinais sem ruído forte e reflexões. TDOA é indicado corretamente pelo maior pico no GCC-PHAT $\hat{r}_{x_i x_j}(\tau)$ . . . . .	25
FIG.1.6	Sinais com forte ruído. TDOA não é indicado corretamente pelo maior pico no GCC-PHAT $\hat{r}_{x_i x_j}(\tau)$ . . . . .	26
FIG.1.7	MB de um rifle a aproximadamente 600 metros da posição de gravação. . . . .	30
FIG.2.1	Espectro em frequência do sinal de tiro e ruído. Janela de 1.800 ms para ruído e janelas de 60 ms para SW e MB . . . . .	32
FIG.2.2	Redução de ruído com MF: ambas componentes SW e MB são realçadas. . . . .	34
FIG.2.3	Efeitos de SS e MF na componente MB (BORZINO, 2015b). . . . .	34
FIG.2.4	Comparação de erros de estimação de DOA para SE, SS e MF, com SNR = 0 dB. . . . .	36
FIG.3.1	GCC-PHAT entre sinais que chegam em um par de microfones com diferentes valores de SNR. . . . .	39
FIG.3.2	Ambiguidade em ES(3) (BORZINO, 2015a). . . . .	40
FIG.3.3	Comparando ES(4) para SNR = -7 dB. . . . .	42
FIG.4.1	Histograma da soma de TDOAs para determinação de $\gamma_0$ . . . . .	49
FIG.5.1	O algoritmo proposto ES-SCFL (BORZINO, 2016). . . . .	51
FIG.5.2	Histogramas da função custo para SE e a escolha de $\gamma_1$ tal que a probabilidade de falso alarme seja desprezível com a menor possível probabilidade de perda (BORZINO, 2016). . . . .	52
FIG.5.3	Histogramas da função custo para ES(5) e a escolha de $\gamma_2$ que resulta em probabilidade de falso alarme igual a 5% (BORZINO,	

	2016). . . . .	53
FIG.5.4	Histogramas da função custo para ES(4) e a escolha de $\gamma_3$ que resulta em probabilidade de falso alarme igual a 5% (BORZINO, 2016). . . . .	54
FIG.6.1	Comparando o algoritmo ES-SCFL com os algoritmos: (a) ES(5), onde há 17 “•” e 20 “★” sobrepostos representando erros aproximadamente nulos; (b) ES(4), onde há 18 “•” e 20 “★” sobrepostos; (c) SCFL, onde há 18 “•” and 20 “★” sobrepostos; (d) SE e ILS (BORZINO, 2016). . . . .	57
FIG.6.2	Histogramas de erros angulares para SNR = -8 dB. . . . .	58
FIG.6.3	Histogramas de erros angulares para o algoritmo ES-SCFL com SNR = -8 dB. . . . .	58
FIG.6.4	Local de tiro: o arranjo de microfones foi posicionado na origem, que corresponde a coordenadas UTM (Zona 23 K) 643.826,86 m E e 7.450.217,78 m S. Atiradores S1, S2 e S3 estão a seguintes distâncias do arranjo: 758 m, 929 m e 1.166 m, respectivamente. . . . .	61
FIG.6.5	Desempenho do algoritmo ES-SCFL (frente a seus concorrentes) com sinais gravados a uma distância de 1.166 m, com SNRs média de -10,21 dB e -9,11 dB (BORZINO, 2016). . . . .	61
FIG.6.6	Desempenho do algoritmo ES-SCFL com sinais gravados a uma distância de 929 m, com SNRs média de -8,58 dB e -7,45 dB (BORZINO, 2016). . . . .	63
FIG.6.7	Desempenho do algoritmo ES-SCFL com sinais gravados a uma distância de 758 m, com SNR média de -7,56 dB (5.56 mm) e -6,49 dB (7.62 mm). . . . .	64
FIG.6.8	Distribuição de probabilidade cumulativa para sinais gravados a uma distância de 1.166 m (BORZINO, 2016). . . . .	65
FIG.6.10	Distribuição de probabilidade cumulativa para sinais gravados a uma distância de 758 m. . . . .	65
FIG.6.9	Distribuição de probabilidade cumulativa para sinais gravados a uma distância de 929 m (BORZINO, 2016). . . . .	66
FIG.7.1	Componentes MB de mesmo calibre. . . . .	69
FIG.7.2	Componentes MB de quatro diferentes calibres. . . . .	69

FIG.7.3	Transformadas wavelet do MB de quatro diferentes calibres. ....	70
FIG.7.4	RLM das quatro WT apresentadas anteriormente. ....	71
FIG.7.5	GLCM das quatro WT apresentadas anteriormente. ....	73
FIG.7.6	ADAG utilizada nos experimentos. ....	75
FIG.8.1	Componentes MB de quatro azimutes diferentes da mesma arma e mesmo disparo. ....	79
FIG.8.2	Transformadas wavelet do MB de quatro azimutes diferentes da mesma arma e mesmo disparo. ....	80
FIG.8.3	ADAG para as oito armas/munições. ....	81
FIG.8.4	Primeira árvore de classificação para a Base 2. Nenhuma arma ou munição é repetida. ....	87
FIG.8.5	Segunda árvore de classificação. Uma mesma arma (Ruger SP101) utilizando dois tipos diferentes de munição, que foram separadas em diferentes classes. ....	89
FIG.8.6	Terceira árvore de classificação. Uma mesma munição (357h1 e 357h2) utilizada com dois armamentos diferentes, compreendendo uma única classe. ....	89
FIG.8.7	Quarta árvore de classificação. Diferentes munições (357 Magnum e 38 Special) utilizadas de uma mesma arma (Ruger SP101), compreendendo uma única classe. ....	90
FIG.8.8	Quinta árvore de classificação. O 38h, o 357h1 e o 357h2 foram separados em diferentes classes. ....	90
FIG.8.9	Árvore de classificação com armas de cano curto. ....	101
FIG.8.10	Primeira árvore de classificação com armas de cano longo. ....	102
FIG.8.11	Segunda árvore de classificação com armas de cano longo. ....	104

## LISTA DE TABELAS

TAB.1.1	Pseudo código para o algoritmo ILS ( $M = 7$ ) .....	26
TAB.2.1	Algoritmo para MF .....	33
TAB.2.2	Desempenho ( $E_{dB}$ ) das técnicas de redução de ruído .....	36
TAB.3.1	Pseudo código para o algoritmo ES( $n$ ) .....	39
TAB.3.2	Frequência relativa do menor erro (FRME) para ES( $n$ ) com SNR = -8 dB .....	41
TAB.3.3	Desempenho dos algoritmos SE, ILS, SRP-PHAT e ES(4). .....	43
TAB.4.1	FLs com quatro pares para $M = 7$ .....	46
TAB.4.2	Pseudo código para o algoritmo SCFL .....	47
TAB.5.1	Frequência de cada decisão de ES-SCFL .....	54
TAB.5.2	Porcentagem de cDOA .....	54
TAB.6.1	Porcentagem de resultados com erros menores que 1 grau (para 500 DOAs simulados) .....	59
TAB.6.2	Estatísticas de erros para SNR = -8 dB .....	59
TAB.6.3	Porcentagem de resultados com erros menores que 1 grau (com MF) .....	60
TAB.6.4	Estatísticas de erros com MF para SNR = -8 dB .....	60
TAB.6.5	$E_\theta$ e $E_\phi$ para sinais gravados a uma distância de 1.166 m .....	62
TAB.6.6	$E_\theta$ e $E_\phi$ para sinais gravados a uma distância de 929 m .....	63
TAB.6.7	$E_\theta$ e $E_\phi$ para sinais gravados a uma distância de 758 m .....	64
TAB.7.1	Exemplo de imagem quantizada .....	71
TAB.7.2	Outro exemplo de imagem quantizada .....	72
TAB.7.3	Atributos de textura utilizados nesta tese .....	74
TAB.7.4	Matriz de confusão para a classificação de quatro calibres uti- lizando características espectro-temporais .....	76
TAB.7.5	Matriz de confusão para a classificação de quatro calibres uti- lizando espectrograma quantizado em níveis de cinza .....	76
TAB.7.6	Matriz de confusão para a classificação de quatro calibres uti- lizando WT quantizada em níveis de cinza .....	76

TAB.8.1	Matriz de confusão com azimutes até $95^\circ$ .....	81
TAB.8.2	Matriz de confusão com azimutes até $130^\circ$ .....	82
TAB.8.3	Matriz de confusão com azimutes até $140^\circ$ .....	82
TAB.8.4	Matriz de confusão com azimutes até $171^\circ$ .....	83
TAB.8.5	Matriz de confusão com azimutes até $180^\circ$ .....	84
TAB.8.6	Taxas de acerto utilizando CET, espectrograma e WT .....	84
TAB.8.7	Azimutes aproximados dos microfones para a Base 2 .....	85
TAB.8.8	Armamentos, munições e número de tiros da Base 2 .....	86
TAB.8.9	Taxas de acerto para a primeira árvore de classificação para cada azimute .....	87
TAB.8.10	Conjuntos de azimutes e respectivas taxas de acerto para a primeira árvore de classificação .....	88
TAB.8.11	Taxas de acerto para a segunda até a quinta árvore de classificação para cada azimute .....	91
TAB.8.12	Taxas de acerto para a segunda até a quinta árvore de classificação para cada conjunto de microfones .....	91
TAB.8.13	Número de sinais de tiro de cada base .....	92
TAB.8.14	Azimutes e taxas de acerto para a primeira árvore de classificação com duas bases .....	93
TAB.8.15	Taxas de acerto para a primeira árvore de classificação para con- juntos de microfones com duas bases .....	93
TAB.8.16	Comparando taxas de acerto por azimute com e sem filtro passa- baixas .....	95
TAB.8.17	Comparando taxas de acerto por conjunto de microfones com e sem filtro passa-baixas .....	96
TAB.8.18	Taxas de acerto entre duas classes com a Base 2 .....	97
TAB.8.19	Taxas de acerto entre duas classes com duas bases .....	98
TAB.8.20	Número de sinais para classificação de azimute .....	98
TAB.8.21	Matriz de confusão para classificação de azimute .....	98
TAB.8.22	Número de sinais para classificação de azimute excluindo um dos azimutes .....	99
TAB.8.23	Matriz de confusão para classificação de azimute excluindo o az- imute $115^\circ$ .....	99
TAB.8.24	Número de sinais para classificação de porte de arma .....	100

TAB.8.25	Matriz de confusão para classificação de porte de arma .....	100
TAB.8.26	Taxas de acerto para a árvore de classificação envolvendo somente armas de cano curto, por azimute .....	101
TAB.8.27	Taxas de acerto para a árvore de classificação envolvendo somente armas de cano curto, por conjunto de microfones .....	102
TAB.8.28	Taxas de acerto para a primeira árvore de classificação envolvendo somente armas de cano longo, por azimute .....	103
TAB.8.29	Taxas de acerto para a primeira árvore de classificação envolvendo somente armas de cano longo, por conjunto de microfones .....	103
TAB.8.30	Número de sinais de tiro de cada base por arma de cano longo .....	104
TAB.8.31	Azimutes e taxas de acerto para a segunda árvore de classificação envolvendo somente armas de cano longo com duas bases .....	105
TAB.8.32	Conjuntos de azimutes e taxas de acerto para a segunda árvore de classificação envolvendo somente armas de cano longo com duas bases .....	106

## LISTA DE ABREVIATURAS

cDOA	-	Direção de chegada correta (do inglês <i>correct Direction of Arrival</i> )
CET	-	Características Espectro-Temporais
cFL	-	<i>Loop</i> fundamental consistente (do inglês <i>consistent Fundamental Loop</i> )
DOA	-	Direção de chegada (do inglês <i>Direction of Arrival</i> )
ES	-	Busca exaustiva (do inglês <i>Exhaustive Search</i> )
ES-SCFL	-	Busca exaustiva-Buscando <i>loop</i> fundamental consistente (do inglês <i>Exhaustive Search-Searching Consistent Fundamental Loop</i> )
FL	-	<i>Loop</i> fundamental (do inglês <i>Fundamental Loop</i> )
FRME	-	Frequência Relativa de Menor Erro
GCC	-	Correlação Cruzada Generalizada (do inglês <i>Generalized Cross Correlation</i> )
GLCM	-	Matriz de co-ocorrência de níveis de cinza (do inglês <i>Gray Level Co-occurrence Matrix</i> )
iDOA	-	Direção de chegada incorreta (do inglês <i>incorrect Direction of Arrival</i> )
ILS	-	Mínimos quadrados iterativo (do inglês <i>Iterative Least Squares</i> )
LS	-	Mínimos quadrados (do inglês <i>Least Squares</i> )
MB	-	<i>Muzzle Blast</i>
MF	-	Filtro mediana (do inglês <i>Median Filter</i> )
RLM	-	Matriz de comprimento de corrida (do inglês <i>Run Length Matrix</i> )
SCFL	-	Buscando <i>loop</i> fundamental consistente (do inglês <i>Searching Consistent Fundamental Loop</i> )
SE	-	Estimação padrão (do inglês <i>Standard Estimation</i> )
SRP	-	<i>Steering Response Power</i>
SS	-	Subtração espectral (do inglês <i>Spectral Subtraction</i> )
SVM	-	Máquina de vetores de suporte (do inglês <i>Support Vector Machine</i> )
SW	-	<i>Shock Wave</i>
TDOA	-	Diferença de tempo de chegada (do inglês <i>Time Difference of Arrival</i> )
WT	-	Transformada wavelet (do inglês <i>Wavelet Transform</i> )
ZCS	-	Soma cíclica nula (do inglês <i>Zero Cyclic Sum</i> )

## RESUMO

O estudo de sinais de tiro gravados por um ou mais sensores pode ser de interesse das Forças Armadas, para estimar a localização de um atirador, e de órgãos de segurança pública, para identificar características sobre o tipo de armamento e munição utilizados em uma cena criminal. Esta tese aborda tópicos úteis para esses dois objetivos.

Quanto à localização do atirador, o primeiro passo é a determinação da direção de chegada (DOA) do projétil. No entanto, a estimação de DOA de uma das componentes do sinal de tiro, o *muzzle blast* (MB), torna-se difícil à medida que o atirador se distancia do arranjo de microfones, em um cenário típico de *sniper*. Neste caso, o MB sofre forte atenuação, surgindo assim casos em que a amplitude do MB apresenta valor menor que a amplitude do ruído, degradando a estimativa de DOA.

Na tentativa de se obter melhor desempenho na determinação da direção de chegada de um tiro para o caso de sinais ruidosos, podemos aplicar técnicas de redução de ruído antes da utilização de algoritmos de estimação de DOA. Tais técnicas, porém, são eficientes até certo nível de ruído, além do qual os resultados tornam-se imprecisos para os algoritmos disponíveis na literatura. Para contornar esse problema, é proposto um novo algoritmo de estimação de DOA intrinsecamente robusto a ruídos aditivos. Testes com sinais simulados e com tiros reais mostram que esse novo algoritmo apresenta desempenho melhor que os outros, quando o nível de ruído é alto, e apresenta resultado equivalente quando o ruído é baixo.

Quanto à identificação do armamento ou da munição, esta tese apresenta um método de reconhecimento de padrões, que pode ser aplicado na classificação de munição e de armamento. Esse método utiliza atributos de textura obtidos da Transformada Wavelet do MB, tratada como imagem e quantizada em níveis de cinza. Testes comprovam que tais atributos conseguem distinguir o porte da arma (cano curto ou longo), assim como uma variedade de munições com taxas de acerto acima das obtidas com o estado da arte.

Um estudo sobre a influência da orientação do cano da arma em relação ao microfone na classificação de munição é discutido neste trabalho. Várias árvores de classificação são testadas, inclusive com junção de bases de dados gravadas em locais e dias diferentes, e os resultados apontam que as taxas de acerto se mantêm elevadas quando essa orientação não ultrapassa determinado valor. Além disso, é mostrado que armas de cano longo sofrem menor influência da orientação do cano da arma em relação ao microfone do que armas de cano curto.

## ABSTRACT

Research regarding gunshot audio recorded by one or more sensors is of great value to government agencies. This work covers two important topics: shooter localization as well as weapon and ammunition signature identification.

The first step for shooter localization is to find the direction of arrival (DOA) of the gunshot signal. DOA estimation of one of the components of a gunshot signal, the muzzle blast (MB), becomes difficult as the shooter moves away from the microphone array, in a typical sniper scenario. In this case, the MB is severely attenuated, originating cases in which the MB amplitude is lower than the amplitude of the noise, which degrades DOA estimation.

In an attempt to attain a better performance with respect to DOA estimation of a noisy gunshot signal, we can apply denoising techniques prior to DOA estimation algorithms. These techniques, however, are efficient up to a certain noise level, beyond which the results become inaccurate. To overcome this problem, a new DOA estimation algorithm is proposed, which is robust to additive noise. Experiments with simulated signals and field recorded gunshots show that the new algorithm achieves better results than state-of-the-art algorithms when the noise level is high, and presents similar results when the noise is low.

This work also proposes a pattern recognition method that can be applied to weapon and ammunition classification. This method uses texture features obtained from the Wavelet Transform of the MB, treated as an image and quantized in gray levels. Tests confirm that such features are capable of distinguishing the type of the firearm (short-barrel or long-barrel weapon), as well as a variety of ammunitions with classification rates higher than the ones obtained from competing techniques.

A study regarding how much the direction of the barrel of the firearm with respect to the microphone influences ammunition classification is covered in this work. Several classification trees are tested, joining databases whose signals were recorded in different places, and the results point out that the classification rates remain high when such orientation of the barrel is not higher than a certain value. Besides, it is shown that the direction of the barrel of the firearm with respect to the microphone influences less long-barrel weapons than short-barrel ones.

# 1 INTRODUÇÃO

O estudo de sinais de tiro é de grande utilidade para as Forças Armadas e Órgãos de Segurança Pública, dentre outras instituições. Gravações de áudio desses sinais podem ser utilizadas, com outras evidências forenses, para identificar ou eliminar posições da origem do tiro e para estabelecer alguns atributos do armamento utilizado (FREYTAG, 2005). Neste capítulo, serão abordados os aspectos teóricos, o estado da arte, a motivação deste trabalho, o objetivo e a organização da tese.

## 1.1 CONTEXTO

Duas aplicações de interesse são abordadas nesta tese: estimação da direção de chegada (DOA) e reconhecimento (classificação) de armamento ou munição. A primeira pode ser útil para determinar a localização do atirador, conforme citado em (SALLAI, 2011) e (MAKINEN, 2010); a segunda pode ser um fator decisivo em uma investigação criminal, pois o conhecimento do tipo de arma utilizada pelo atirador pode ajudar a revelar sua identidade ou auxiliar um analista forense a reconstruir a cena de um crime (MAHER, 2008).

Um exemplo de situação prática em que o reconhecimento do armamento pode ser útil, para o Exército em particular, consiste na hipótese de algum armamento de dotação da Força ser roubado: microfones espalhados em áreas suspeitas podem captar algum sinal de tiro e, com isso, o classificador poderá indicar se o armamento de dotação do Exército está sendo usado naquela área em particular.

## 1.2 REVISÃO DA LITERATURA

Nesta seção, serão explicados conceitos a respeito de sinais sonoros produzidos por um tiro e será apresentado o estado da arte das duas aplicações a serem discutidas nesta tese.

### 1.2.1 O SINAL DE TIRO

Um dos primeiros trabalhos envolvendo sinais de tiro, datado de 1946 (DUMOND, 1946), descreveu os experimentos realizados para se determinar as formas de onda e leis

de propagação e dissipação de ondas balísticas a uma distância de 80 jardas (aproximadamente 73 metros) da trajetória do projétil. Assuntos relacionados à física de propagação do som foram inicialmente abordados em (PIERCE, 1971). Desde então, esse tópico tem chamado a atenção de um número crescente de pesquisadores.

Um sinal de tiro típico é impulsivo e consiste principalmente de duas formas de onda características: o estouro no cano (conhecido pela expressão em inglês *MUZZLE BLAST*, MB) e a onda de choque (*SHOCK WAVE*, SW). O primeiro é consequência da explosão da carga no cano da arma, tem duração típica de 3 ms (MAHER, 2006), podendo chegar a 5 ms (dependendo do armamento), e se propaga no ar com velocidade do som; a segunda aparece quando o projétil se move com velocidade supersônica (usual no caso de um fuzil) e é devida à dispersão das moléculas de ar, causada pelo movimento do projétil (MAHER, 2007). A SW dura tipicamente de 0,3 a 0,5 ms (MAHER, 2006). Ainda existem dois sons que não serão levados em consideração, por não influenciarem na estimativa de DOA e no reconhecimento do armamento: o do mecanismo de disparo e o do impacto no alvo.

A FIG. 1.1 mostra as direções das componentes de SW e MB em uma configuração típica para gravação de sinais de tiro. A frente de onda originária do MB se propaga direcionalmente (em uma linha reta), à velocidade do som, entre a posição do atirador e o receptor (microfone ou arranjo de microfones), enquanto que a frente de onda da componente de SW que chega ao receptor, propagando-se também com velocidade do som, provém de um ponto pertencente à trajetória do projétil, conhecido como *Detach(ment) Point*. Portanto, até que a SW chegue ao microfone, deve-se levar em conta o percurso do projétil, com velocidade supersônica, desde a posição do atirador até o *Detach Point* e, a partir daí, o percurso da SW com velocidade sonora. Desta forma, se estimarmos a direção de chegada a partir da SW, obteremos a direção do *Detach Point*, enquanto que se utilizarmos o MB na rotina de estimação de DOA, obteremos a direção de onde o tiro efetivamente se originou.

A velocidade de som no ar é função da temperatura (BOHN, 1988):

$$v_{som} = c_0 \sqrt{1 + \frac{T}{273}} \quad (1.1)$$

onde  $T$  é a temperatura do ar em graus Celsius e  $c_0 = 331$  m/s é a velocidade do som a  $0^\circ$  C.

Na FIG. 1.1, a localização do *Detach Point* depende de  $\theta_M$ , o chamado “ângulo *Mach*” (MAHER, 2006), onde  $\theta_M = \arcsen(\frac{1}{M})$ , onde  $M = (v_p/v_{som})$ , sendo  $v_{som}$  a

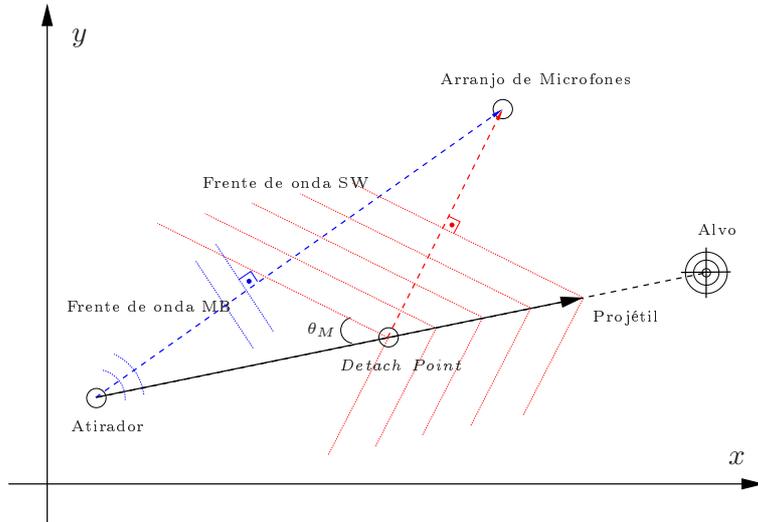


FIG. 1.1: Uma configuração típica para gravação de sinais de tiro (BORZINO, 2016).

velocidade do som e  $v_p$  a velocidade do projétil, que depende do armamento. O *Detach Point* é tal que a reta que o une ao receptor é perpendicular à frente de onda da SW. Note que, para projéteis supersônicos,  $M > 1$ .

A FIG. 1.2 ilustra dois casos de ângulo *Mach*: 1) quando o projétil se desloca bem mais rápido do que a velocidade do som (pequeno ângulo *Mach*), caso em que a SW se propaga quase que perpendicular à trajetória do projétil; 2) quando o projétil se desloca um pouco acima da velocidade do som (ângulo *Mach* perto de  $90^\circ$ ), caso em que a SW se propaga quase que paralelamente à trajetória do projétil.

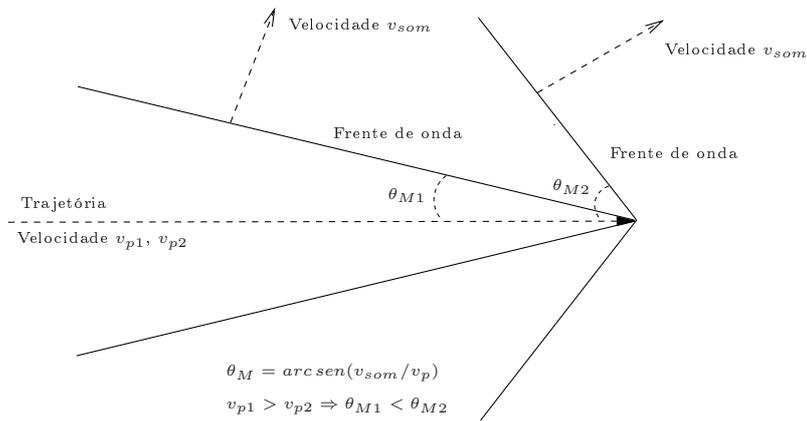


FIG. 1.2: Geometria da SW para duas velocidades do projétil diferentes.

Mesmo quando a velocidade do projétil é supersônica, podem existir situações em que não se consegue captar a SW, quando por exemplo o microfone encontra-se fora do campo de visão dessa componente (na direção oposta à trajetória do tiro ou em uma posição que forma um ângulo maior que  $90^\circ - \theta_M$  com a direção da trajetória do projétil).

A FIG. 1.3 mostra um sinal de tiro proveniente de um M964 Fuzil Automático Leve (FAL), distante de aproximadamente 236 metros da posição de gravação. Os trechos de alta frequência do sinal nos lados esquerdo e direito referem-se, respectivamente, às componentes SW e MB, e suas reflexões. A componente devida ao MB, geralmente, é mais atenuada, sempre que o microfone estiver mais perto da trajetória do projétil do que do atirador, cenário típico de um *sniper* ou caçador (EXÉRCITO, 1998), como é conhecido no Exército Brasileiro (EB). Quando existe SW, quanto maior a velocidade inicial do projétil e quanto maior a distância entre o microfone e o atirador, maior será o intervalo de tempo entre a recepção da SW e do MB, já que o tempo total para a SW atingir o microfone leva em conta a velocidade do projétil.

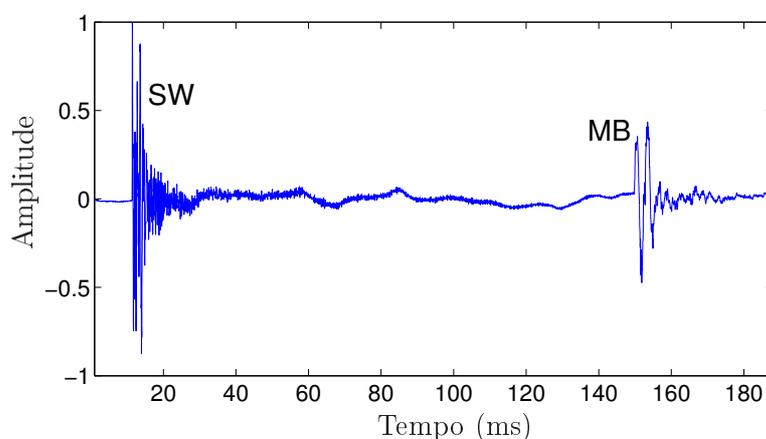


FIG. 1.3: SW e MB originados de um FAL (BORZINO, 2015a).

### 1.2.2 ESTIMAÇÃO DA DIREÇÃO DE CHEGADA DE SINAIS DE TIRO

A estimação de DOA de sinais de tiro é um importante passo para a localização do atirador (SALLAI, 2011). Em (XUE, 2014), foi proposta uma técnica de estimação de DOA para sinais de voz com ruído Gaussiano, utilizando estatísticas de ordem elevada, não sendo então apropriada para sinais de tiro, por serem impulsivos por natureza.

Um sinal de tiro é tipicamente banda larga e, conseqüentemente, algoritmos de processamento de sinais banda estreita, como Capon e MUSIC (VAN TREES, 2002), não são apropriados, em seus formatos originais, para estimar sua direção de chegada. Essas técnicas de banda estreita podem ser utilizadas com sinais de banda larga empregando-se decomposição em subbandas. Além disso, a conformação de feixe conhecida pela expressão em inglês *delay-and-sum beamforming* (RAMOS, 2011) pode ser utilizada em conjunto com a técnica *Steering Response Power* (SRP) (DIBIASE, 2001) para estimação

de DOA de sinais banda larga; contudo, (BORZINO, 2015a) conclui que, para sinais muito ruidosos, essa técnica apresenta pior desempenho que os métodos que empregam a Correlação Cruzada Generalizada (GCC) (KNAPP, 1976).

A estimação de DOA pode ser realizada utilizando diferenças de tempo de chegada (TDOAs) (ELKAMCHOUCI, 2005) entre sinais que chegam em pares de microfones de um arranjo. O GCC, por ser aplicável em sinais banda larga, pode ser utilizado na estimação desses TDOAs de sinais de tiro. A subclasse do GCC chamada *Phase Transform* (PHAT) (KNAPP, 1976) é utilizada neste trabalho, e a resultante estimação de DOA é descrita neste capítulo.

A FIG. 1.4 ilustra o arranjo de sete microfones utilizado nos experimentos desenvolvidos nesta tese, assim como os dois ângulos que caracterizam a direção de chegada (VAN TREES, 2002): o azimute  $\phi$  e o ângulo zenital  $\theta$  (ou complemento da elevação). O azimute varia de  $0^\circ$  a  $360^\circ$ , e o ângulo zenital, de  $0^\circ$  a  $180^\circ$ . Nas gravações feitas, as distâncias entre o atirador e o arranjo sempre foram muito maiores que as distâncias entre os microfones, caracterizando a origem da onda sonora como campo distante.

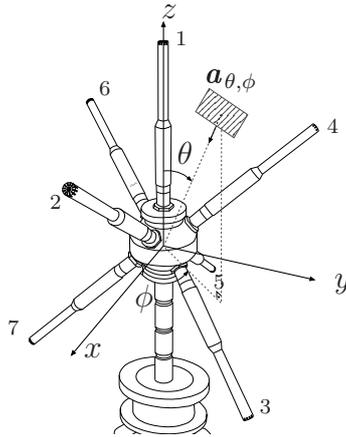


FIG. 1.4: Arranjo de sete microfones utilizado nos experimentos de estimação de DOA e ângulos azimutal e zenital (FREIRE, 2012).

Na FIG. 1.4,  $\mathbf{a}_{\theta, \phi}$  é o vetor unitário na direção de propagação da onda e é dado por

$$\mathbf{a}_{\theta, \phi} = [-\sin \theta \cos \phi \quad -\sin \theta \sin \phi \quad -\cos \theta]^\top. \quad (1.2)$$

Um arranjo de  $M$  microfones define  $N = \binom{M}{2} = \frac{M(M-1)}{2}$  diferentes pares de microfones e um número igual de correlações cruzadas entre seus sinais, que são utilizadas na estimação de DOA baseada no GCC-PHAT. A diferença de tempo de chegada (TDOA), em número de amostras, entre os microfones  $i$  e  $k$ , pode ser estimada levando-se

em consideração a posição dos picos da correlação cruzada entre seus sinais, como em

$$\hat{\tau}_{ik} = \arg \max_{\tau} \hat{r}_{x_i x_k}(\tau), \quad (1.3)$$

onde  $x_i(n)$  e  $x_k(n)$  são os sinais que chegam aos microfones  $i$  e  $k$ , respectivamente. A correlação cruzada, definida usualmente por  $r_{x_i x_k}(\tau) = E[x_i(n)x_k(n-\tau)]$  para sinais estacionários, pode ser estimada, assumindo-se ergodicidade, como

$$\hat{r}_{x_i x_k}(\tau) = \frac{1}{L} \sum_{n=0}^{L-1} x_i(n)x_k(n-\tau), \quad (1.4)$$

para  $x_i(n)$  de tamanho  $L$ . Apesar de sinais de tiro não serem considerados estacionários, a expressão em (1.4) foi utilizada na presente aplicação, pois a posição do seu pico corresponderá ao valor de TDOA. Para se obter resultados mais precisos, pode-se interpolar as correlações em torno do pico (FREIRE, 2011).

A correlação cruzada de (1.4) corresponde a  $\hat{r}_{x_i x_k}(\tau) = \frac{1}{L} [x_i(\tau) * x_k(-\tau)]$ , onde  $*$  denota o operador de convolução, e é usualmente computada como a transformada inversa de Fourier (IFT) da densidade espectral de potência cruzada (CPSD)  $X_i(e^{j\omega})X_k(e^{-j\omega})$ . A correlação cruzada generalizada é definida como (KNAPP, 1976)

$$\hat{r}_{x_i x_k}^G(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Psi(\omega) X_i(e^{j\omega}) X_k(e^{-j\omega}) e^{j\omega\tau} d\omega, \quad (1.5)$$

onde o sobrescrito  $G$  identifica a GCC e  $\Psi(\omega)$  é chamada função de ponderação.

No caso do GCC PHAT, a função de ponderação é dada por

$$\Psi(\omega) = \frac{1}{|X_i(e^{j\omega})X_k(e^{-j\omega})|}. \quad (1.6)$$

Definimos a estimativa de TDOA, em unidade de tempo, como  $\bar{\tau}_{ik} = \frac{\hat{\tau}_{ik}}{f_s}$ ,  $f_s$  sendo a frequência de amostragem.  $\bar{\tau}_{ik}$  corresponde ao tempo que o som (vindo na direção de  $\mathbf{a}_{\theta,\phi}$ ) leva do microfone  $i$  ao microfone  $k$ , uma distância dada por

$$d_{ik} = \mathbf{a}_{\theta,\phi}^T (\mathbf{p}_i - \mathbf{p}_k), \quad (1.7)$$

onde  $\mathbf{p}_i$  e  $\mathbf{p}_k$  são as coordenadas dos microfones. Assim, é esperado que

$$\bar{\tau}_{ik} \approx \frac{d_{ik}}{v_{som}} = \mathbf{a}_{\theta,\phi}^T \Delta \bar{\mathbf{p}}_{ik}, \quad (1.8)$$

onde  $v_{som}$  é a velocidade de propagação do som no ar, e  $\Delta \bar{\mathbf{p}}_{ik} = \frac{\mathbf{p}_i - \mathbf{p}_k}{v_{som}}$ .

Se definirmos uma função custo de mínimos quadrados (LS) contendo todos os  $N = \frac{M(M-1)}{2}$  erros de estimação de TDOA como

$$\xi = \frac{(\bar{\tau}_{12} - \Delta\bar{\mathbf{p}}_{12}^T \mathbf{a}_{\text{DOA}})^2 + \dots + (\bar{\tau}_{(M-1)M} - \Delta\bar{\mathbf{p}}_{(M-1)M}^T \mathbf{a}_{\text{DOA}})^2}{N}, \quad (1.9)$$

sendo  $\mathbf{a}_{\text{DOA}}$  o vetor unitário  $\mathbf{a}_{\theta,\phi}$  a ser determinado na função, e se tomarmos o gradiente dessa função em relação a  $\mathbf{a}_{\text{DOA}}$  e igualarmos o resultado a zero, obtemos uma solução de mínimos quadrados dada por

$$\hat{\mathbf{a}}_{\text{DOA}} = \mathbf{R}^{-1} \mathbf{p}, \quad (1.10)$$

onde

$$\mathbf{R} = \Delta\bar{\mathbf{p}}_{12} \Delta\bar{\mathbf{p}}_{12}^T + \dots + \Delta\bar{\mathbf{p}}_{(M-1)M} \Delta\bar{\mathbf{p}}_{(M-1)M}^T, \quad (1.11)$$

e

$$\mathbf{p} = \bar{\tau}_{12} \Delta\bar{\mathbf{p}}_{12} + \dots + \bar{\tau}_{(M-1)M} \Delta\bar{\mathbf{p}}_{(M-1)M}. \quad (1.12)$$

Dos elementos de  $\hat{\mathbf{a}}_{\text{DOA}} = [\hat{a}_x \ \hat{a}_y \ \hat{a}_z]^T$  computados por (1.10) e de sua definição em (1.2), observamos que o ângulo horizontal (azimute) estimado é dado por  $\hat{\phi} = \tan^{-1} \frac{\hat{a}_y}{\hat{a}_x}$ , e o ângulo vertical (zênite) estimado, por  $\hat{\theta} = \cos^{-1}(-\hat{a}_z)$ . Esse método de estimação de TDOA baseado em mínimos quadrados foi proposto em (CHAN, 1978) e utilizado em (SMITH, 1987) para avaliar diferentes técnicas de localização. A essa estimação clássica baseada no GCC-PHAT utilizando todos os  $N$  pares possíveis (neste caso,  $N = 21$ ), denominamos nesta tese “Estimação Padrão” (SE, do inglês *Standard Estimation*).

Quando a amplitude do sinal de tiro é maior que a amplitude do ruído e quando não existem reflexões significativas, haverá somente um pico dominante em cada correlação cruzada. A Fig. 1.5 ilustra este caso para um par de microfones. Note que TDOA, em número de amostras, entre os picos dos dois sinais de tiro é 74, e o maior pico da correlação cruzada está localizado exatamente no *lag* 74, indicando uma correta estimação de TDOA.

Por outro lado, a Fig. 1.6 ilustra os mesmos sinais da Fig. 1.5 com elevado nível de ruído. O ruído adicionado é real (gravação feita entre dois sinais de tiro). Note que o maior pico da correlação cruzada, realçado pelo círculo vermelho, não está localizado no *lag* 74, que é o TDOA correto entre os dois sinais. O maior pico é devido ao ruído, e não ao sinal de tiro. Quando isso acontece, o maior pico está incorreto, e o TDOA estimado para este par de sinais leva a uma estimativa de DOA errada. Na Fig. 1.6, o *lag* 74 corresponde ao quarto maior pico da correlação cruzada.

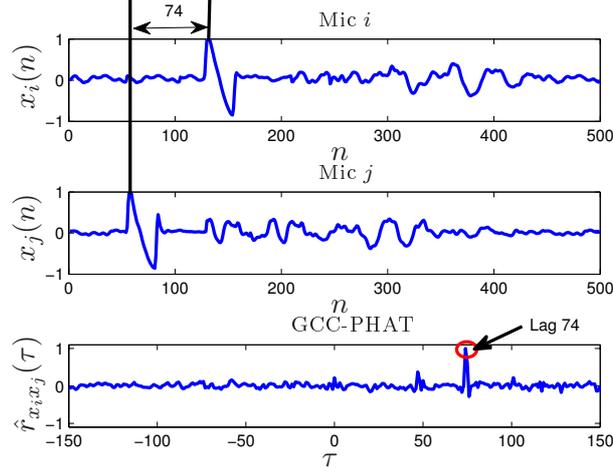


FIG. 1.5: Sinais sem ruído forte e reflexões. TDOA é indicado corretamente pelo maior pico no GCC-PHAT  $\hat{r}_{x_i x_j}(\tau)$ .

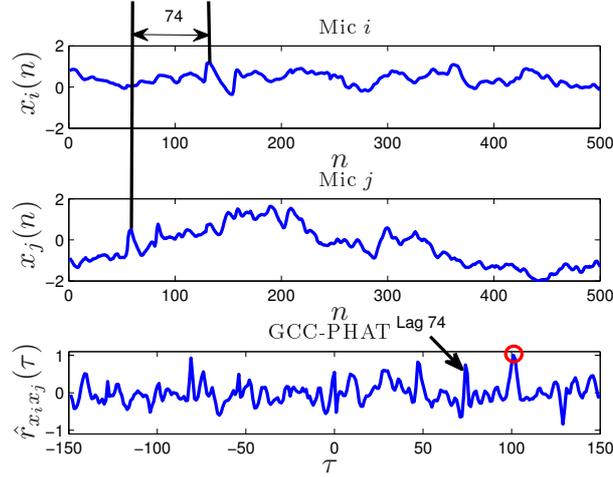


FIG. 1.6: Sinais com forte ruído. TDOA não é indicado corretamente pelo maior pico no GCC-PHAT  $\hat{r}_{x_i x_j}(\tau)$ .

Para resolver o problema de estimação incorreta de DOA nos casos em que nem todos TDOAs são corretamente estimados (como na Fig. 1.6), o algoritmo *Iterative Least Squares* (ILS) (FREIRE, 2014) tenta eliminar os efeitos deletérios dos picos incorretos nas correlações cruzadas entre os sinais que chegam em um par de microfones. O algoritmo ILS elimina a correlação (dentre  $\frac{M(M-1)}{2} = 21$  possibilidades, para  $M = 7$ ) que menos contribui para a minimização da função custo em (1.9). Em outras palavras, o par  $(i, j)$  de microfones que origina o maior valor  $(\bar{\tau}_{ij} - \Delta \bar{\mathbf{p}}_{ij}^T \hat{\mathbf{a}}_{\text{DOA}})^2$  é removido e a função custo é calculada novamente, normalizada pelo número de termos. Esse procedimento é repetido até que 6 ou 5 termos (nos experimentos, utilizou-se 6) permaneçam na função (FREIRE, 2012, 2014). A TAB. 1.1 detalha o algoritmo ILS.

TAB. 1.1: Pseudo código para o algoritmo ILS ( $M = 7$ )

**Faça**  $N \leftarrow 21$   
**Se**  $N \geq 6$   
    **Estime**  $\hat{\mathbf{a}}_{\text{DOA}}$  em (1.10) utilizando os  $N$  TDOAs, obtendo  $\hat{\theta}$  e  $\hat{\phi}$   
    **Calcule** as parcelas  $(\bar{\tau}_{ij} - \Delta \bar{\mathbf{p}}_{ij}^T \hat{\mathbf{a}}_{\text{DOA}})^2$  da função custo  
    em (1.9) para cada par  $(i, j)$  de microfones  
    **Elimine** o par  $(i, j)$  que gera a maior parcela em (1.9)  
    **Faça**  $N \leftarrow N - 1$   
**Fim**  
**Retorne**  $\hat{\theta}$  e  $\hat{\phi}$  do último  $\hat{\mathbf{a}}_{\text{DOA}}$  obtido

Em (BORZINO, 2014a), o algoritmo ILS foi aplicado após a técnica de redução de ruído conhecida como Subtração Espectral (BOLL, 1979), possibilitando melhoria de desempenho na estimação de DOA de sinais de tiro com razão sinal-ruído (SNR) entre -3 dB e 2 dB. O próximo capítulo abordará a aplicação de técnicas de redução de ruído nos sinais de tiro.

Com relação a reflexões, (BORZINO, 2014b) propôs um esquema de deconvolução, a fim de eliminá-las do sinal de tiro. Porém, essa técnica só se mostrou útil para SNR entre 3 dB e 12 dB. Além disso, em (PAREDES, 2015), uma melhoria do DOA estimado foi conseguida removendo-se microfones não alinhados com uma primeira direção estimada, mas esse método só é aplicável para valores de SNR maiores que 10 dB.

### 1.2.3 CLASSIFICAÇÃO DE ARMAMENTO OU MUNIÇÃO

O reconhecimento de padrões é um ramo de aprendizagem de máquina que envolve a classificação de informações em imagens, sons e outras representações digitais. Aplicações incluem reconhecimento de voz (CHOU, 2003), reconhecimento automático de imagens de faces humanas (UWECHUE, 2006), classificação de documentos (WILLET, 1988), bioinformática (DING, 2003), aplicações médicas (HUNG, 2012) e muitos outros exemplos. Este trabalho utiliza reconhecimento de padrões para identificar que tipo de munição ou armamento foi utilizado no disparo de um projétil.

A tarefa de classificação está relacionada com a procura de regularidades estruturais em conjuntos de dados, para que seja possível organizar esses dados em diferentes categorias ou classes (BISHOP, 2006). Para isto, vários algoritmos podem ser aplicados, tais como *Support Vector Machine* (SVM) (CORTES, 1995), Redes Neurais (HAYKIN, 1999), *Linear Discriminant Analysis* (LDA) (JAMES, 2013), dentre outros.

A identificação do tipo de armamento ou do seu calibre pode ser um fator decisivo

em uma investigação criminal, ter um alto valor tático ou ser um elemento de inteligência importante nas operações policiais ou militares. O conhecimento do tipo de arma utilizada pelo atirador pode ajudar a revelar sua identidade, auxiliar um analista forense a solucionar um crime, ajudar a encontrar um armamento roubado de algum órgão do governo, ou alertar uma tropa operacional de uma ameaça durante uma ação real no campo de combate ou em uma operação policial. Sabe-se que seres humanos bem treinados são capazes reconhecer o tipo de arma (ou munição) apenas escutando o som do tiro. Entretanto, para não dependermos dessa habilidade de poucos indivíduos, podemos utilizar alguma técnica de aprendizado de máquina para classificação de munição.

Há alguns trabalhos na literatura cujo objetivo é identificar (detectar) um sinal de tiro dentro de uma gravação com ruído ou outros sinais impulsivos, independente do tipo de munição. Uma recente pesquisa que explora esse assunto pode ser encontrada em (AHMED, 2013), que apresenta um sistema de detecção de tiro com alta taxa de detecção em dois estágios: o primeiro detecta um sinal impulsivo (evento acústico), como tiro, palmas, batida de porta, estouro de balão, etc. O segundo estágio tenta reconhecer um sinal de tiro dentre os sinais impulsivos previamente detectados. Apenas dois armamentos foram utilizados em (AHMED, 2013): Heckler & Koch G3 e Heckler & Koch MP5, que usam calibres 7,62 mm e 9 mm, respectivamente.

Em (KHAN, 2010), são apresentados método e dispositivo utilizados naquele trabalho para identificar padrões acústicos, como os de um sinal de tiro. As armas são classificadas segundo uma hierarquia: um primeiro nível classifica os sinais de acordo com o porte da arma (pequeno, médio ou grande porte). Um segundo nível compreende a classificação em armas específicas. Os tiros foram gravados a aproximadamente 10 metros do atirador, e três condições de captura foram simuladas, chamadas de *Room reverb.*, *Concert Reverb.* e *Doppler Effect.* A técnica apresentada em (KHAN, 2010) funciona bem quando se pretende classificar o sinal de tiro entre rifle e *handgun* sob as condições analisadas no artigo, mas pode falhar em classificar um calibre específico (taxa de classificação menor que 70%). Os testes de classificação de armamento/munição realizados nesta tese compreendem somente gravações reais, em vez de condições simuladas.

Em (SALLAI, 2011), é proposta uma estimacão de calibre baseada na duracão  $T$  da SW; logo, só é aplicável quando a velocidade do projétil é supersônica. Porém, para se determinar  $T$ , precisamos saber a trajetória do projétil, e isto somente é viável se existirem arranjos de microfones no campo de visão da SW. Além disso, o método utilizado em (SALLAI, 2011) para estimacão de calibre depende da correta determinacão

da duração da SW. A técnica de reconhecimento de munição apresentada nesta tese requer apenas um único microfone (e não arranjos) e é baseada na componente MB, que é consequência da explosão da carga no cano da arma (MAHER, 2006) sendo, então, influenciada pelas características da munição.

Em (MAYS, 2001), foi desenvolvida uma classificação entre SW e MB, utilizando *Short Time Fourier Transform* (STFT), distribuição em tempo-frequência de Wigner-Ville e Transformada Wavelet (WT); a última apresentou melhor desempenho para classificar as duas componentes do sinal de tiro. Além disso, (TIAN, 2013) e (BORZINO, 2014b) quantizaram o espectrograma em níveis de cinza e o utilizaram em tarefas de reconhecimento: o primeiro, em aplicações de radar e o último, em reconhecimento de armamento após um esquema de deconvolução. Esta tese combina as ideias de WT e quantização em níveis de cinza na classificação de munição.

Em (KIKTOVA-VOZARIKOVA, 2015), são propostos atributos temporais e espectrais para detecção de eventos acústicos. São eles: taxa de cruzamento com zero, fluxo espectral, *Log-frequency filter bank coefficients*, *Audio Spectrum Envelope* (ASE), *Audio Spectrum Flatness* (ASF), *Audio Spectrum Centroid* (ASC), *Audio Spectrum Spread* (ASS) e *Spectral Roll-off Frequency*. Tais atributos são comparados com a proposta desta tese para classificação de munição.

### 1.3 MOTIVAÇÃO

Dois grandes problemas que influenciam nas duas aplicações de interesse são as reflexões e o ruído. No tocante ao primeiro problema citado, reflexões na superfície da terra ou em outros obstáculos são situações comuns encontradas nas componentes de SW e MB, como na FIG. 1.3. Isto degrada, tanto a estimativa de DOA quanto o reconhecimento de armamento, pois os sinais que chegam ao arranjo de microfones podem ser modelados como resultado da convolução do sinal de tiro limpo com a resposta ao impulso do caminho acústico, adicionado ao ruído ambiente.

Com relação à influência do ruído na estimação de DOA, quando o atirador está longe da posição de gravação (mais de 400 metros), o MB torna-se muito atenuado, e a SNR diminui, principalmente se houver vento forte. Neste caso, técnicas clássicas de estimação de DOA não são suficientes para se obter um resultado preciso. O método mais utilizado para estimativa de DOA baseia-se na correlação cruzada de sinais gravados em pares de microfones. Assim, picos de ruído podem ser confundidos com picos de MB, ocasionando uma estimação errônea. No próximo capítulo, será ilustrado e explicado com

mais detalhes o problema do ruído. A FIG. 1.7 ilustra um sinal de tiro (apenas o MB) originado de um rifle, a aproximadamente 600 metros do arranjo de microfones. O MB, realçado com um retângulo vermelho, possui amplitudes comparáveis aos picos de ruído, o que degrada consideravelmente a estimação de DOA.

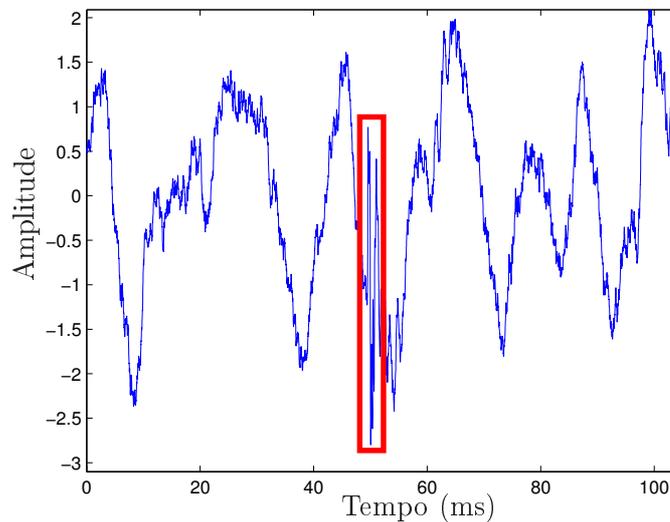


FIG. 1.7: MB de um rifle a aproximadamente 600 metros da posição de gravação.

#### 1.4 OBJETIVO

Os objetivos desta tese compreendem formular técnicas de estimação de DOA que possam ser empregadas inclusive em sinais com SNR muito baixa (abaixo de -3 dB) e com alguma reflexão, em um cenário típico de *sniper*, com precisão de  $1^\circ$ , além de criar um método de classificação de munição com maior taxa de acerto do que a encontrada na literatura. A precisão de  $1^\circ$  será detalhada no Capítulo 5. A SNR muito baixa é considerada, neste trabalho, menor que -3 dB, porque os algoritmos para estimação de DOA existentes na literatura possuem desempenho satisfatório até esse valor.

#### 1.5 ORGANIZAÇÃO DA TESE

A tese está organizada da seguinte forma: o Capítulo 2 aborda técnicas de redução de ruído que podem ser empregadas antes da estimação de DOA propriamente dita; o Capítulo 3 explica a técnica de busca exaustiva; o Capítulo 4 apresenta o conceito de *loop* fundamental consistente aplicado à estimativa de DOA; o Capítulo 5 propõe o algoritmo ES-SCFL, que pode ser empregado para estimar o DOA de sinais com qualquer SNR; o

Capítulo 6 apresenta os resultados de estimação de DOA utilizando direções simuladas e sinais reais gravados; o Capítulo 7 aborda a classificação de munição utilizando a WT quantizada em níveis de cinza; o Capítulo 8 analisa os aspectos direcionais das armas de fogo e a influência na classificação de munição, e o Capítulo 9 conclui a tese e sugere tópicos para trabalhos futuros.

## 2 TÉCNICAS DE REDUÇÃO DE RUÍDO

Como o ruído é um dos problemas para as aplicações de interesse deste trabalho, este capítulo analisa algumas técnicas de redução de ruído (*denoising*), a fim de verificar qual delas contribui para uma melhor estimação de DOA. Na literatura, existem diversas técnicas de *denoising* que podem ser aplicadas a sinais impulsivos, tais como *Wavelet denoising* (MEDINA, 2003; ISA, 2011; RAJ, 2011; DONOHO, 1994), Subtração espectral (BOLL, 1979; RAMOS, 2013, 2012b,a) e Filtro mediana (FITZGERALD, 2010; CANADAS-QUESADA, 2014; BORZINO, 2015b). Foram feitos testes com sinais de tiro, e as duas últimas técnicas citadas apresentaram melhores resultados e serão discutidas neste capítulo.

### 2.1 SUBTRAÇÃO ESPECTRAL

A Subtração espectral (SS, do inglês *Spectral Subtraction*) foi introduzida em (BOLL, 1979), no contexto de redução do ruído em sinais de voz, e sua aplicação em sinais de tiro foi abordada em (RAMOS, 2013, 2012a). Para este método, considera-se que a informação necessária para estimar o espectro do ruído esteja disponível (RAMOS, 2012b), o que é coerente assumir, pois os sensores captam o sinal antes do tiro ser realizado. Dada a natureza impulsiva dos sinais de tiro, o ruído ambiente pode ser considerado estacionário durante o curto período em que o tiro aconteceu (RAMOS, 2012b).

Sejam  $X(e^{j\omega})$ ,  $S(e^{j\omega})$  e  $R(e^{j\omega})$  as transformadas de Fourier do sinal ruidoso que chega a um microfone, do sinal limpo correspondente e do ruído aditivo (tomado, por exemplo, antes do sinal de tiro), respectivamente. Então  $X(e^{j\omega}) = S(e^{j\omega}) + R(e^{j\omega})$ . Uma estimativa do espectro do sinal de interesse,  $\hat{S}(e^{j\omega})$ , é dada por (BOLL, 1979),

$$\hat{S}(e^{j\omega}) = [ |X(e^{j\omega})| - \overline{R}(e^{j\omega}) ] e^{j\angle X(e^{j\omega})}, \quad (2.1)$$

onde  $\overline{R}(e^{j\omega})$  é o valor absoluto médio de  $R(e^{j\omega})$ , estimado da porção de ruído sem sinal imediatamente antes do evento de interesse. A Equação (2.1) pode ser reescrita como

$$\hat{S}(e^{j\omega}) = \left[ 1 - \alpha \frac{\overline{R}(e^{j\omega})}{|X(e^{j\omega})|} \right] X(e^{j\omega}) = H(e^{j\omega})X(e^{j\omega}), \quad (2.2)$$

onde  $\alpha$  é real positivo e tem a função de regular as componentes frequenciais que se tornam nulas,  $H(e^{j\omega})$  é real e usualmente designado filtro de subtração espectral. Sempre que  $H(e^{j\omega}) < 0$  em (2.2),  $H(e^{j\omega})$  é forçado a ser zero (BOLL, 1979).

A respeito do parâmetro  $\alpha$ , de acordo com experimentos em (BORZINO, 2014a), se o aumentarmos (acima de 10, por exemplo) para maior redução de ruído, o processamento altamente não linear distorce severamente os picos de MB (da onda direta e das reflexões), influenciando negativamente na estimação de DOA. Experimentos indicam que valores de  $\alpha$  de 5 a 10 geram resultados satisfatórios para a aplicação em questão (estimação de DOA com MB). Em (BORZINO, 2014a), a aplicação do algoritmo ILS após SS melhorou o desempenho da estimação de DOA para sinais de tiro com SNR entre 2 e -3 dB.

A FIG. 2.1 mostra um sinal de tiro real completo (com SW e MB) e a resposta em frequência dessas componentes e do ruído. Para cada componente, a resposta em frequência foi determinada a partir de uma janela de 60 ms de sinal contendo a SW ou o MB. Para o ruído, a janela foi de 1.800 ms antes da SW.

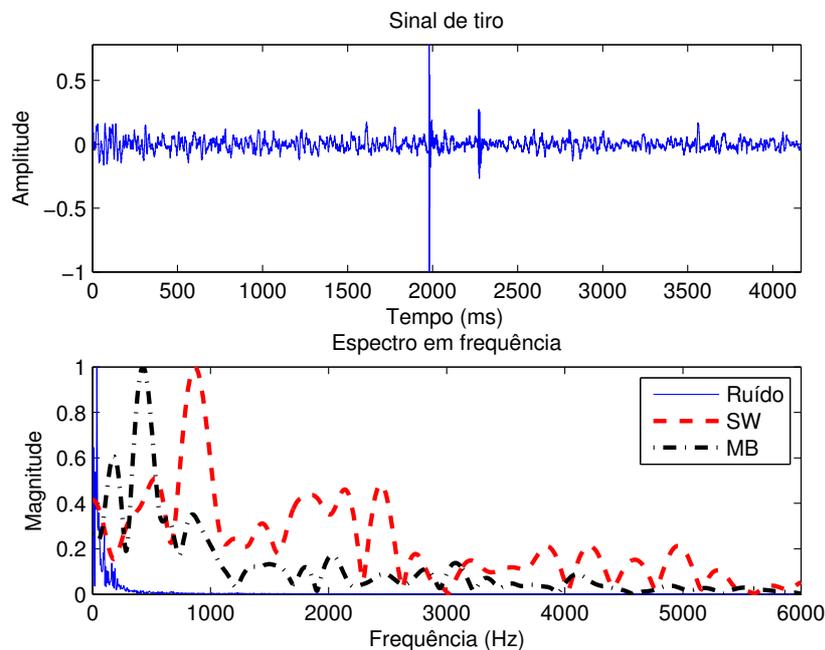


FIG. 2.1: Espectro em frequência do sinal de tiro e ruído. Janela de 1.800 ms para ruído e janelas de 60 ms para SW e MB

Da FIG. 2.1, observa-se que as principais componentes frequenciais do ruído são diferentes das do sinal de tiro. O ruído tem predominância nas frequências mais baixas, o que contribui para a aplicação da subtração espectral, que se comportará, nesse caso, como um filtro passa-altas, de modo a cortar as baixas frequências do ruído. Ainda, analisando as componentes do sinal de tiro, como esperado, a SW possui maiores frequências que o MB.

## 2.2 FILTRO MEDIANA

Em (FITZGERALD, 2010), foi apresentado o uso do filtro mediana (MF, do inglês *Median Filter*) em sinais musicais com o objetivo de separar sons percussivos e harmônicos (CANADAS-QUESADA, 2014). Sinais impulsivos, como tiros em ambiente ruidoso, possui espectro semelhante ao de um sinal musical, ambos possuindo componentes de banda larga, representadas por linhas verticais no espectrograma, significando a presença de várias componentes frequenciais. Em (DUFAUX, 2000) e (CHACÓN-RODRÍGUEZ, 2011), o MF foi utilizado para detectar som impulsivo em ambiente ruidoso.

O MF simplesmente substitui cada amostra da sequência de entrada  $x(n)$  pela mediana (valor central de uma lista ordenada) da janela limitada por  $\pm\Delta$  amostras em torno da  $n$ -ésima amostra. Com uma escolha adequada de  $\Delta$ , esse procedimento se torna uma estratégia eficiente para eliminar *outliers* em uma sequência.

Em (FITZGERALD, 2010), o objetivo é reproduzir os sons separados, o que requer processamento adicional. Nesta aplicação, entretanto, deseja-se realçar a componente MB para melhorar a estimação de DOA. Se o sinal de tiro for modelado por  $x(n) = s(n) + r(n)$ , onde  $s(n)$  é o sinal de tiro limpo (impulsivo) e  $r(n)$  é o resíduo indesejável, as componentes de tiro, impulsivas, serão os *outliers* que o MF é capaz de eliminar. Assim, após filtrar  $x(n)$ , obtemos uma estimativa  $\hat{r}(n)$  do resíduo. O sinal de tiro realçado pode ser facilmente computado como  $\hat{s}(n) = x(n) - \hat{r}(n)$ . O algoritmo resultante, detalhado na TAB. 2.1, é simples e só requer que algumas amostras sejam armazenadas (buffer), devido à natureza não causal do MF. Na TAB. 2.1,  $x(n)$  é o sinal ruidoso com  $N$  amostras.

TAB. 2.1: Algoritmo para MF

1: <b>Inicialize</b> $\hat{r}(n)$ e $\hat{s}(n)$ com $N$ amostras nulas
2: <b>Faça</b> $\Delta = (L - 1)/2$ ( $L$ , um número ímpar)
3: <b>Para</b> $n = (1 + \Delta) : (N - \Delta)$
4: $\hat{r}(n) \leftarrow \text{mediana}(x(n - \Delta : n + \Delta))$
5: $\hat{s}(n) \leftarrow x(n) - \hat{r}(n)$
6: <b>Fim</b>

Na frequência de amostragem ( $f_s$ ) de 96 kHz, por exemplo, os experimentos produziram bons resultados para  $\Delta = 150$  amostras quando se utiliza tanto a SW como o MB. Uma motivação para essa escolha vem do fato de  $\Delta$ , convertido em unidade de tempo, ser maior que metade da duração típica do MB. A FIG. 2.2 ilustra um sinal ruidoso gravado contendo 4 tiros, a saída do MF (resíduo) e o sinal realçado. Nota-se que tal

esquema de redução de ruído realçou ambas as componentes (SW e MB) dos sinais de tiro, permitindo, em particular, uma clara visualização das componentes MB.

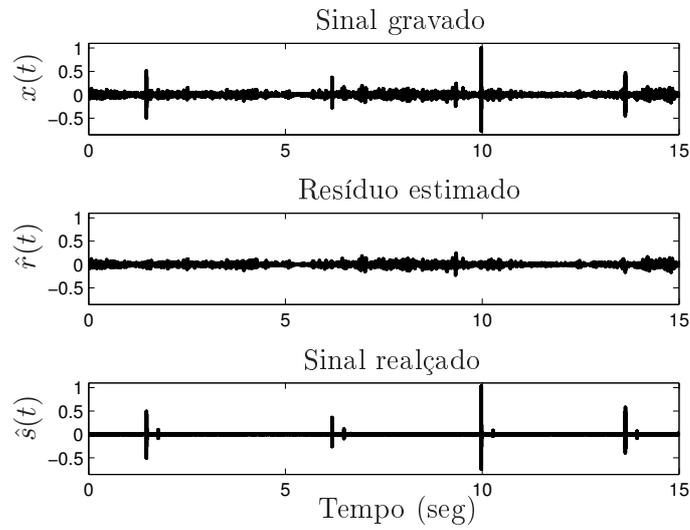


FIG. 2.2: Redução de ruído com MF: ambas componentes SW e MB são realçadas.

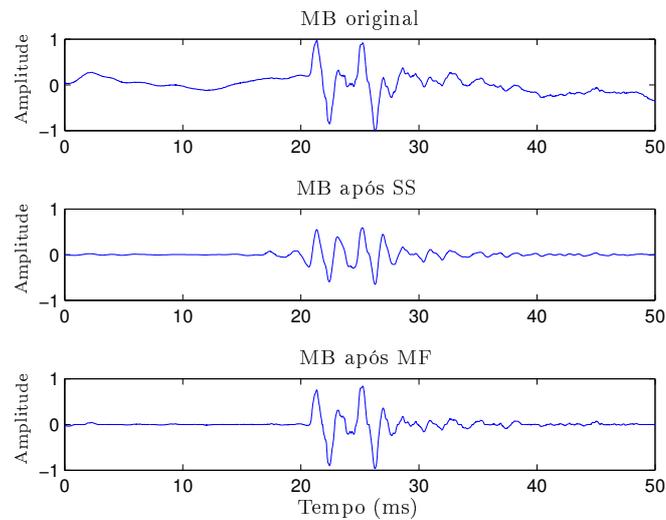


FIG. 2.3: Efeitos de SS e MF na componente MB (BORZINO, 2015b).

Para verificar os efeitos na componente MB (picos da onda direta e reflexões) das duas técnicas de redução de ruído detalhadas nesta seção, a FIG. 2.3 mostra os resultados do pré-processamento do sinal gravado com SS (com  $\alpha = 10$ ) e MF. Nota-se que este último preserva mais a forma de onda do MB do que o primeiro, o que contribui para melhor estimativa de DOA pois, na SS, a energia das componentes frequenciais eliminadas

é transferida para as componentes que permaneceram no sinal original. Já, com o MF, isso não ocorre.

Com o objetivo de avaliar o desempenho da aplicação da SS e do MF antes da estimação de DOA, foi conduzido o experimento detalhado a seguir: são simulados DOAs de sinais de tiro utilizando como referência “limpa” uma componente MB de um rifle gravada a uma distância de 236 metros da posição de tiro (SNR maior que 20 dB). Escolhem-se, então, direções aleatórias (azimutes e zênites), e são dados os atrasos correspondentes no sinal de referência, para simular os sinais chegando em cada microfone do arranjo ilustrado na FIG. 1.4. Depois disso, é adicionado ruído gravado pelo arranjo (somente ruído existente entre tiros sucessivos), de modo que seja possível controlar a SNR. Note que somente a direção é simulada, já que tanto o sinal de tiro como o ruído são reais. Para cada direção simulada, são adicionados diferentes trechos de ruído. Para o cálculo da SNR, denota-se a variância de um sinal limpo contendo a componente MB de  $\sigma_s^2$ , calculada a partir de uma janela de 7,5 ms, e a variância do ruído de  $\sigma_n^2$ ; a SNR (em dB) é então calculada como  $10 \log(\sigma_s^2/\sigma_n^2)$ .

Os experimentos mostraram que, para SNR maior que 1 dB, a aplicação de técnica de *denoising* não prejudica, mas também não resulta em significativa melhora na estimação de DOA. Foram simulados 200 DOAs aleatórios para cada SNR de -3 dB a 3 dB (com incremento de 1 dB) e comparados os resultados para três casos:

- SE (somente a estimação padrão GCC-PHAT);
- SS (o melhor resultado obtido com  $\alpha$  inteiro de 5 a 10) + SE;
- MF ( $\Delta = 150$ ) + SE.

Para quantificar os resultados, foram definidos, para cada sinal de tiro,  $e_\theta = \theta - \hat{\theta}$ , o erro entre os ângulos zenitais simulado ( $\theta$ ) e estimado ( $\hat{\theta}$ ), e  $e_\phi = \phi - \hat{\phi}$ , o erro entre os ângulos azimutais simulado ( $\phi$ ) e estimado ( $\hat{\phi}$ ). Com isso, foram calculados os erros, utilizando RMSE (do inglês *Root Mean Squared Error*), para cada SNR, assim definidos:

$$E_\theta = \sqrt{\sum_{k=1}^K \frac{(e_{\theta,k})^2}{K}}, \quad e \quad (2.3)$$

$$E_\phi = \sqrt{\sum_{k=1}^K \frac{(e_{\phi,k})^2}{K}}, \quad (2.4)$$

onde “K” é o número de tiros simulados, e o índice “k” refere-se ao  $k$ -ésimo tiro.

A TAB. 2.2 mostra o resultado de  $E_{\text{dB}} = 10 \log_{10} (E_{\theta} + E_{\phi})$ , utilizando 200 sinais simulados ( $K = 200$ ) para cada SNR, para os três casos citados (SE, SS + SE e MF + SE).

TAB. 2.2: Desempenho ( $E_{\text{dB}}$ ) das técnicas de redução de ruído

SNR	3 dB	2 dB	1 dB	0 dB	-1 dB	-2 dB	-3 dB
SE	-14,01	-13,40	0,58	11,95	13,41	13,52	14,67
SS	-16,01	-13,98	-2,33	6,95	12,89	13,48	14,58
MF	<b>-17,53</b>	<b>-17,58</b>	<b>-15,37</b>	<b>-13,28</b>	<b>-11,84</b>	<b>2,66</b>	<b>8,84</b>

Da TAB. 2.2, observa-se que a aplicação do filtro mediana antes da estimação padrão melhora o desempenho, consideravelmente, para SNR de -2 dB a 1 dB. Para SNR acima dessa faixa, o desempenho com MF é apenas um pouco melhor ou não faz diferença. Para SNR abaixo de -2 dB, o erro quadrático médio, em dB, fica muito elevado. Nota-se também que a aplicação do MF apresenta melhores resultados que a SS para todos os valores de SNR analisados na TAB. 2.2.

A Fig. 2.4 compara  $|e_{\phi}|$  e  $|e_{\theta}|$  nos casos de SE, SS + SE e MF + SE para 25 tiros com DOAs aleatórios com SNR igual a 0 dB. Dessa figura, observa-se que MF produz erros muito próximos de zero, demonstrando o desempenho superior do MF comparado com SS.

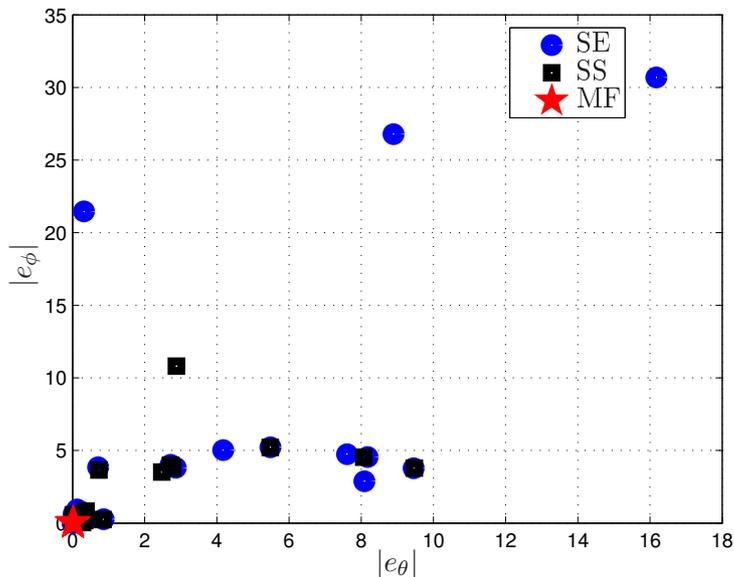


FIG. 2.4: Comparação de erros de estimação de DOA para SE, SS e MF, com SNR = 0 dB.

Neste capítulo, foram detalhadas as técnicas de SS e MF para redução de ruído e foram comparados os desempenhos desses métodos antes da estimação padrão de DOA. Pelo fato da SS ser suscetível a alterações nos picos do MB, o MF apresentou melhores resultados na estimação de DOA.

### 3 A BUSCA EXAUSTIVA

A função custo resultante do algoritmo ILS possui seis termos (pares de microfones selecionados pelo algoritmo da TAB. 1.1); porém, não se pode garantir que esses termos originam o menor valor da função custo, considerando-se todas as possíveis combinações de seis pares de microfones. Assim, pode-se dizer que o algoritmo ILS resulta em uma solução subótima, do tipo “vizinho mais próximo”. Um modo de se obter uma melhor solução é avaliando-se a função custo para todas as possíveis combinações e selecionando-se a combinação que produza o menor valor da função. Este procedimento foi chamado de “Busca Exaustiva” (ES do inglês *Exhaustive Search*).

O arranjo utilizado possui sete microfones e, então,  $N = 21$  pares de microfones. Se aplicarmos ES considerando-se  $n$  termos próximo de  $\frac{N}{2}$ , levaria muito tempo para testar todas as possíveis combinações, não tolerado em aplicações em tempo real. Por outro lado, se  $n$  não for próximo de  $\frac{N}{2}$ , um curto período de tempo é suficiente, causando um retardo aceitável. Define-se  $ES(n)$  como a busca exaustiva de  $n$  pares. Note que, neste caso, são necessárias  $\binom{N}{n}$  avaliações da função custo, esta possuindo somente  $n$  termos. Um modo de encontrar todas as possíveis combinações de  $n$  pares, de um total de  $N$  possibilidades, é realizando incrementos em um contador binário de 1 a  $2^N$  e separando aquelas cuja soma dos bits seja igual a  $n$ , com a posição de cada bit correspondendo a um par  $(i, j)$  específico de microfones (por exemplo, o primeiro bit sempre corresponde ao par  $(1, 2)$ ; o segundo bit, ao par  $(1, 3)$ , ..., o último bit, ao par  $(6, 7)$ ). Um pseudo-código para  $ES(n)$  para um arranjo de  $M$  microfones é descrito na TAB. 3.1 (BORZINO, 2015a).

Para sinais muito ruidosos (abaixo de -3 dB), o efeito na correlação cruzada é semelhante ao apresentado na FIG. 1.6, em que o maior pico não está localizado no *lag* correto, ou seja, o maior pico da correlação cruzada pode ser devido ao ruído, e não ao sinal de tiro. A FIG. 3.1 ilustra as correlações cruzadas entre sinais que chegam em dois microfones, com SNR igual a 30 dB (figura da esquerda) e -8 dB (figura da direita).

Da FIG. 3.1, observamos que, quando a SNR é alta, existe somente um único pico na correlação cruzada, e esse pico é certamente devido ao sinal de tiro. Por outro lado, quando a SNR é baixa, pode haver mais de um pico com amplitudes similares, e não é possível inferir qual pico corresponde ao sinal de tiro. Tal figura sugere que, para valores de SNR muito baixos, dentre todas as  $N = 21$  possíveis correlações cruzadas,

TAB. 3.1: Pseudo código para o algoritmo ES( $n$ )

```

1: Faça (contador binário)  $BC = 1$ ,  $N = \frac{M(M-1)}{2}$ ,  $k = 0$  e  $n$ 
2: Para  $BC = 1$  a  $2^N$ 
3:   Se soma dos bits de  $BC = n$ 
4:     Faça  $k \leftarrow k + 1$ 
5:     Ache  $\hat{\mathbf{a}}_{DOA}(k)$  em (1.10)
6:     Avalie a função custo  $\xi(k)$ , em (1.9), com  $n$  (em vez de  $N$ ) pares
7:     Armazene  $\xi(k)$ 
8:   Fim
9: Fim
10: Ache  $k$  com o menor  $\xi(k)$  calculado no passo 6
11: Retorne  $\hat{\mathbf{a}}_{DOA}(k)$ ,  $k$  obtido no passo anterior.

```

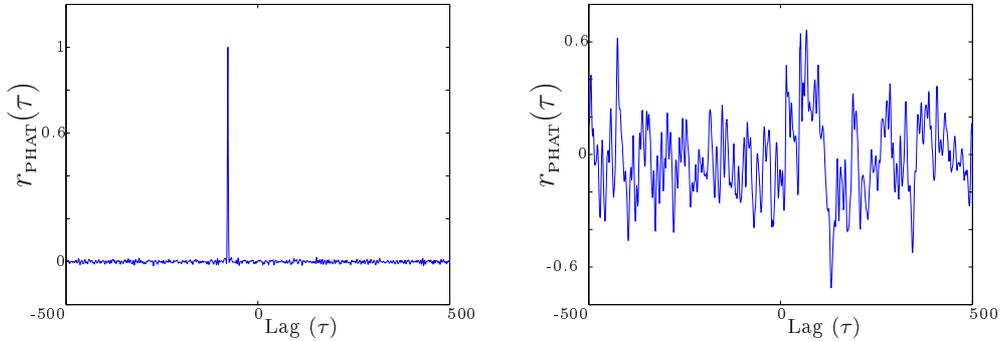


FIG. 3.1: GCC-PHAT entre sinais que chegam em um par de microfones com diferentes valores de SNR.

pode haver casos em que somente poucas delas possuem o maior pico devido ao sinal (e não ao ruído). Nestes casos, um pequeno valor de  $n$  apresenta os melhores resultados. Entretanto, ES(3), ou  $n = 3$  (o menor valor de  $n$  para estimar o vetor de três elementos  $\mathbf{a}_{DOA}$ ) nem sempre gera resultados precisos, visto que algumas combinações de  $n = 3$  pares de microfones geram ambiguidade na estimação de DOA. A ambiguidade (matriz  $\mathbf{R}$  é singular e (1.10) não é aplicável) ocorre quando os três pares são formados por apenas três microfones que, como consequência, definem um plano (BORZINO, 2015a).

Na FIG. 3.2, é mostrado um exemplo em que os pares (2,4), (4,6) e (6,2), formados pelos microfones 2, 4 e 6, definem um plano. Na FIG. 3.2, observa-se que os vetores unitários  $\mathbf{a}_1$  e  $\mathbf{a}_2$  possuem a mesma projeção horizontal (no plano  $x'y'$ ) e, então, as mesmas componentes  $\mathbf{a}_{1x} = \mathbf{a}_{2x}$  e  $\mathbf{a}_{1y} = \mathbf{a}_{2y}$ , causando o mesmo atraso. Além disso, suas componentes verticais têm a mesma magnitude, mas direções opostas,  $\mathbf{a}_{1z} = -\mathbf{a}_{2z}$ . Como essas projeções verticais são perpendiculares ao plano  $x'y'$ , elas não influenciam os

retardos. Então, os azimutes dos sinais que chegam das direções dos vetores unitários  $\mathbf{a}_1$  e  $\mathbf{a}_2$  são os mesmos ( $\phi_1 = \phi_2$ ), e seus ângulos zenitais são simétricos ( $\theta_1 = -\theta_2$ ). Visto que o atraso causado por esses dois vetores quando os sinais adquiridos por esses três microfones é o mesmo, não é possível determinar, a partir desses três TDOAs, se o DOA correto é  $\mathbf{a}_1$  ou  $\mathbf{a}_2$ . Em um caso como esse (quando  $n = 3$  pares formam um plano), a matriz  $\mathbf{R}$  em (1.10) é singular. Por causa dessa ambiguidade, sempre que 3 pares definem um plano, não há uma única solução para ES(3).

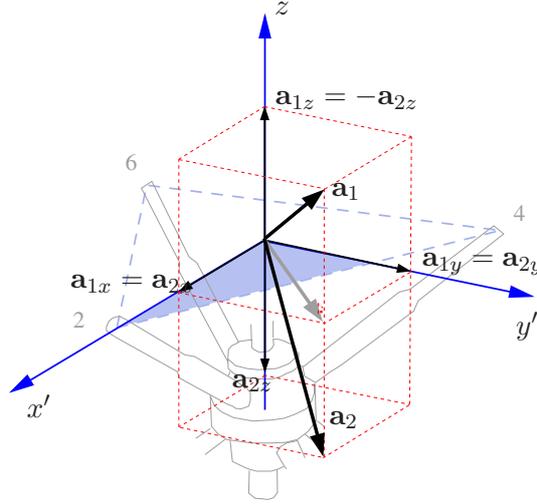


FIG. 3.2: Ambiguidade em ES(3) (BORZINO, 2015a).

Com o objetivo de avaliar o desempenho de ES( $n$ ) para diferentes valores de  $n$ , foram simulados DOAs, conforme explicado no Capítulo 2. Para cada direção simulada, variou-se  $n$  em ES( $n$ ) de 3 a 21 e foi observado qual  $n$  em ES( $n$ ) originou o melhor resultado, para SNR de -3 dB a -8 dB, com decremento unitário. Para SNR acima de -3 dB, ES(4) garante resultados comparáveis a qualquer valor de  $n$  em ES( $n$ ).

A TAB. 3.2 (BORZINO, 2015a) mostra a frequência relativa de menor erro (FRME) em função de  $n$ ,  $3 \leq n \leq 7$ , onde FRME( $n$ ) significa quantas vezes ES( $n$ ) produziu o melhor resultado dividido pelo número de simulações realizadas (no caso, 200), com SNR de -8 dB. Para cada direção simulada, o melhor resultado corresponde ao ES( $n$ ),  $n$  variando de 3 a 7, que origina o menor erro definido como  $e_\theta^2 + e_\phi^2$ , onde  $e_\theta = \theta - \hat{\theta}$  é o erro entre os ângulos zenitais simulado e estimado, enquanto que  $e_\phi = \phi - \hat{\phi}$  é o erro entre os ângulos azimutais simulado e estimado. Por exemplo, na TAB. 3.2, FRME(4) = 0,36 significa que, em 36% do conjunto de simulações, o ES(4) resultou no menor erro. Para  $n = 3$ , a busca exaustiva foi feita somente no subconjunto das combinações possíveis que não configuravam um plano. Note, nesse caso, que  $\sum_{n=3}^7 \text{FRME}(n) = 1$ .

TAB. 3.2: Frequência relativa do menor erro (FRME) para ES( $n$ ) com SNR = -8 dB

$n$	3	4	5	6	7
FRME	0,10	<b>0,36</b>	0,26	0,14	0,14

Da TAB. 3.2, observa-se que ES(4) originou o menor erro (na média) para SNR = -8 dB, pois possui a maior FRME. Experimentos mostraram que o mesmo acontece para diferentes SNR de -8 dB a -3 dB. Todos os valores de  $n$  de 8 a  $N = 21$  também foram testados com sinais nessa faixa de SNR, mas não tiveram melhores resultados que aqueles obtidos com menores valores de  $n$ .

Com o intuito de avaliar o desempenho do ES(4) com os algoritmos existentes na literatura, variou-se a SNR, de -3 dB para baixo, e os resultados foram comparados. Quatro algoritmos foram analisados: o ES(4), o ILS, o SE, além do *Steering Response Power - Phase Transform* (SRP-PHAT) (DIBIASE, 2001). O algoritmo SRP baseia-se na estimação espectral espacial (DMOCHOWSKI, 2007), fazendo uso da matriz de correlação espacial. Em (RAMOS, 2011), são descritas as equações que governam um arranjo 3-D, e a equação que corresponde à estimativa de DOA é apresentada como

$$\hat{\Theta}(\theta, \phi) = \arg \max_{(\theta, \phi)} \mathbf{1}^T \hat{\mathbf{R}}_{\theta, \phi} \mathbf{1}, \quad (3.1)$$

onde  $\hat{\mathbf{R}}_{\theta, \phi}$  denota uma estimativa da matriz de correlação espacial, obtida por meio da média das amostras (do sinal de tamanho  $L$ ) que chegam ao arranjo, ou seja

$$\hat{\mathbf{R}}_{\theta, \phi} = \frac{1}{L} \sum_{l=0}^{L-1} \mathbf{x}_{\theta, \phi}(n-l) \mathbf{x}_{\theta, \phi}^T(n-l), \quad (3.2)$$

onde

$$\mathbf{x}_{\theta, \phi}(n) = [x_1(n + \tau_1(\theta, \phi)) \quad x_2(n + \tau_2(\theta, \phi)) \quad \dots \quad x_M(n + \tau_M(\theta, \phi))]^T. \quad (3.3)$$

onde  $\tau_i(\theta, \phi) = \bar{\tau}_i(\theta, \phi) \times f_s$ , em que  $\bar{\tau}_i(\theta, \phi) = \frac{\mathbf{a}_{\theta, \phi}^T \mathbf{p}_i}{V_{som}}$  é a diferença de tempo de chegada do sinal ao microfone  $i$  de coordenadas  $\mathbf{p}_i$ , com relação a uma origem arbitrária, de um arranjo com  $M$  microfones. Em outras palavras, SRP procura (em um conjunto de valores possíveis) pelo par de ângulos  $(\theta, \phi)$  que maximiza a potência de saída do arranjo.

Como a correlação cruzada é usualmente computada como a transformada inversa de Fourier da densidade espectral de potência cruzada  $X_{i, \theta, \phi}(e^{j\omega})X_{k, \theta, \phi}(e^{-j\omega})$  entre os sinais que chegam aos microfones  $i$  e  $k$ , a (3.1) pode ser escrita como

$$\hat{\Theta}(\theta, \phi) = \arg \max_{(\theta, \phi)} \sum_{i=1}^M \sum_{k=1}^M \int_{-\infty}^{\infty} X_{i, \theta, \phi}(e^{j\omega}) X_{k, \theta, \phi}(e^{-j\omega}) e^{j\omega \tau_{ik}(\theta, \phi)} d\omega, \quad (3.4)$$

onde  $\tau_{ik}(\theta, \phi) = \tau_i(\theta, \phi) - \tau_k(\theta, \phi)$ .

Pode-se perceber que o SRP e o GCC possuem expressões quase idênticas; porém, no SRP, somam-se as densidades espectrais de potência cruzadas em todos os microfones. O SRP-PHAT (semelhante ao GCC-PHAT) é uma subclasse do SRP, em que este é normalizado pelo valor absoluto da densidade espectral de potência cruzada entre os sinais que chegam a um par de microfones. Assim, a expressão para o SRP-PHAT é definida como

$$\hat{\Theta}(\theta, \phi) = \arg \max_{(\theta, \phi)} \sum_{i=1}^M \sum_{k=1}^M \int_{-\infty}^{\infty} \frac{1}{|X_{i,\theta,\phi}(e^{j\omega})X_{k,\theta,\phi}(e^{-j\omega})|} X_{i,\theta,\phi}(e^{j\omega})X_{k,\theta,\phi}(e^{-j\omega})e^{j\omega\tau_{ik}(\theta,\phi)} d\omega. \quad (3.5)$$

Em todos os testes, o algoritmo ES(4) produziu resultados melhores (ou, no mínimo, similares) do que os outros.

A FIG. 3.3 compara SE, ILS, SRP-PHAT e ES(4) em 25 sinais simulados, com SNR = -7 dB. Observa-se que ES(4) possui desempenho muito melhor do que os outros algoritmos do atual estado da arte.

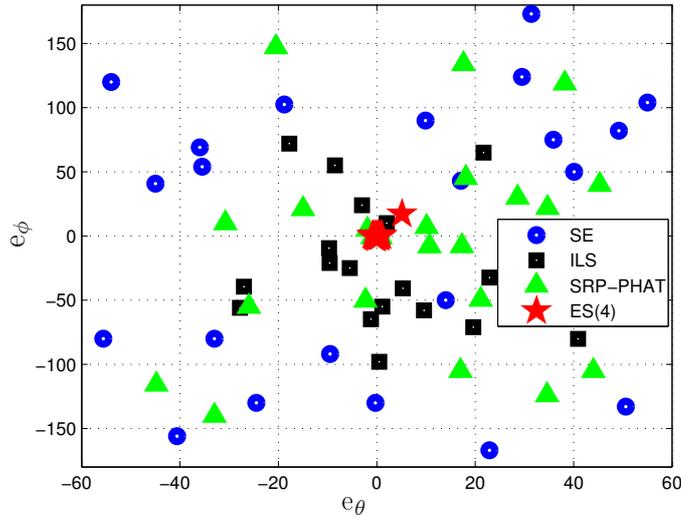


FIG. 3.3: Comparando ES(4) para SNR = -7 dB.

Com a finalidade de comparar o desempenho do ES(4) com os outros algoritmos em uma faixa de SNR (de -4 a -8 dB), foram calculados os erros, utilizando RMSE, conforme definidos em (2.3) e (2.4), para cada SNR.

A TAB. 3.3 mostra o resultado de  $10 \log_{10}(E_\theta + E_\phi)$ , utilizando 500 sinais simulados para cada SNR, para os quatro algoritmos em análise neste capítulo, e observa-se que o ES(4) apresenta melhor desempenho.

TAB. 3.3: Desempenho dos algoritmos SE, ILS, SRP-PHAT e ES(4).

SNR	SE	ILS	SRP-PHAT	ES(4)
-4 dB	13,71	6,83	8,44	<b>-1,98</b>
-5 dB	16,15	10,69	11,58	<b>0,63</b>
-6 dB	18,48	15,61	16,59	<b>2,71</b>
-7 dB	19,79	17,50	18,58	<b>7,18</b>
-8 dB	20,96	18,73	20,05	<b>14,51</b>

Neste capítulo, foi apresentado o algoritmo  $ES(n)$ , útil para estimar DOA de sinais ruidosos, em que nem todas as correlações cruzadas possuem o maior pico devido ao MB. Também foi explicado o problema da ambiguidade no caso do  $ES(3)$ , motivo pelo qual foi dada ênfase ao  $ES(4)$ , que gerou bons resultados comparados com SE, ILS e SRP-PHAT.

## 4 LOOP FUNDAMENTAL CONSISTENTE NA ESTIMAÇÃO DE DOA

Devido à interferência ou ao ruído, o maior pico de uma correlação cruzada pode não corresponder ao pico correto para a estimação de DOA. A fim de melhorar a robustez em cenários de baixa SNR, podemos realizar uma busca exaustiva, ES(4), considerando-se um número pré-definido de picos, ao invés de se utilizar apenas o maior deles. Com base em observações, percebeu-se que, nos casos em que o ES( $n$ ) falha, o pico de uma correlação cruzada devido ao sinal de tiro usualmente aparece entre os 10 maiores picos. A escolha dos 10 picos mais altos, considerando uma busca num menor conjunto de combinações, é um bom compromisso entre o custo computacional e a probabilidade de incluir o pico correto em cenários de baixa SNR. Entretanto, testar a função custo  $5,985 \times 10^7$  vezes – que resulta da combinação dos 10 maiores picos e 4 correlações cruzadas a partir de 21 pares de microfones – seria ineficaz para aplicações em tempo real. Além disso, nos experimentos com sinais muito ruidosos, observou-se que, considerando somente a combinação de pares com o menor valor da função custo não é suficiente para se obter uma estimativa de DOA precisa. Como uma forma de superar esse problema, pode-se recorrer ao conceito de *loop fundamental consistente* (cFL, do inglês *consistent Fundamental Loop*) (SCHEUING, 2008) e restringir a pesquisa a um conjunto muito menor de combinações (350.000).

### 4.1 MOTIVAÇÃO

Um *loop fundamental* (FL) é formado quando um conjunto de pares de microfones (quatro no exemplo seguinte) completa um caminho cíclico, isto é, atrasos no tempo são calculados do microfone  $i$  ao  $j$ ,  $\bar{\tau}_{ij}$ , depois do  $j$  ao  $k$ ,  $\bar{\tau}_{jk}$ , do  $k$  ao  $l$ ,  $\bar{\tau}_{kl}$ , e do  $l$  de volta ao microfone  $i$ ,  $\bar{\tau}_{li}$ . O FL é dito consistente (chamado cFL) quando a soma dos TDOAs no loop é zero:

$$\bar{\tau}_{ij} + \bar{\tau}_{jk} + \bar{\tau}_{kl} + \bar{\tau}_{li} = 0. \quad (4.1)$$

Isso é conhecido como condição de “soma cíclica nula” (ZCS, do inglês *Zero Cyclic Sum*) (SCHEUING, 2008). Apesar de não ser suficiente para verificar a correta estimação dos 4 TDOAs, a condição ZCS é necessária. Portanto, podemos utilizá-la para selecionar, dentre todas as combinações possíveis de 4 pares de microfones formando um FL, somente aquelas que satisfazem a condição ZCS.

Considere que quatro pares de microfones formem um FL e que os correspondentes TDOAs estimados (em unidade de tempo) sejam  $\bar{\tau}_{ij}$ ,  $\bar{\tau}_{jk}$ ,  $\bar{\tau}_{kl}$  e  $\bar{\tau}_{li}$ . A função custo LS é dada por:

$$\xi = \frac{(\bar{\tau}_{ij} - \Delta\bar{\mathbf{p}}_{ij}^T \mathbf{a}_{\text{DOA}})^2 + \dots + (\bar{\tau}_{li} - \Delta\bar{\mathbf{p}}_{li}^T \mathbf{a}_{\text{DOA}})^2}{4}, \quad (4.2)$$

que também é dada por  $\mathbf{e}^T \mathbf{e}$ , vetor de erros  $\mathbf{e}$  definido como

$$\mathbf{e} = \underbrace{\begin{bmatrix} \bar{\tau}_{ij} \\ \bar{\tau}_{jk} \\ \bar{\tau}_{kl} \\ \bar{\tau}_{li} \end{bmatrix}}_{\bar{\boldsymbol{\tau}}} - \underbrace{\begin{bmatrix} \Delta\bar{\mathbf{p}}_{ij}^T \\ \Delta\bar{\mathbf{p}}_{jk}^T \\ \Delta\bar{\mathbf{p}}_{kl}^T \\ \Delta\bar{\mathbf{p}}_{li}^T \end{bmatrix}}_{\Delta\bar{\mathbf{P}}^T} \mathbf{a}_{\text{DOA}} = \bar{\boldsymbol{\tau}} - \Delta\bar{\mathbf{P}}^T \mathbf{a}_{\text{DOA}}. \quad (4.3)$$

A solução LS de (4.2),  $\hat{\mathbf{a}}_{\text{DOA}} = \mathbf{R}^{-1} \mathbf{p}$ , torna-se

$$\hat{\mathbf{a}}_{\text{DOA}} = \left( \underbrace{\Delta\bar{\mathbf{P}} \Delta\bar{\mathbf{P}}^T}_{\mathbf{R}} \right)^{-1} \underbrace{\Delta\bar{\mathbf{P}} \bar{\boldsymbol{\tau}}}_{\mathbf{p}}, \quad (4.4)$$

tal que o vetor de erros é dado por

$$\mathbf{e} = \bar{\boldsymbol{\tau}} - \Delta\bar{\mathbf{P}}^T \left( \Delta\bar{\mathbf{P}} \Delta\bar{\mathbf{P}}^T \right)^{-1} \Delta\bar{\mathbf{P}} \bar{\boldsymbol{\tau}}. \quad (4.5)$$

Se todos os TDOAs estimados em um FL estiverem corretos, pode-se observar de (1.8) que  $\bar{\boldsymbol{\tau}} = \Delta\bar{\mathbf{P}}^T \mathbf{a}_{\theta, \phi}$  e, de (4.3),  $\mathbf{e} = \mathbf{0}$ . Consequentemente, a função custo em (4.2) é nula e, assumindo  $\Delta\bar{\mathbf{p}}_{ij} + \Delta\bar{\mathbf{p}}_{jk} + \Delta\bar{\mathbf{p}}_{kl} + \Delta\bar{\mathbf{p}}_{li} = \mathbf{0}$  (arranjo calibrado), a condição ZCS é satisfeita. Por outro lado, se nem todos os TDOAs estimados em um FL forem corretos, independente da condição ZCS ser satisfeita ou não, podemos dizer que

$$\bar{\boldsymbol{\tau}} = \Delta\bar{\mathbf{P}}^T \mathbf{a}_{\theta, \phi} + \mathbf{n}, \quad (4.6)$$

onde  $\mathbf{n}$  é definido como o vetor de erro de estimação de TDOA. Na prática, até mesmo em cenários de alta SNR, podemos esperar erros na estimação de TDOA, devido a erros de quantização e de calibração. A condição para se obter ZCS sem que todos os TDOAs estimados estejam corretos é  $\mathbf{n}^T [1 \ 1 \ 1 \ 1]^T = 0$ , o que implica compensação de erros, tal que sua soma seja nula. Uma vez que isso é improvável de acontecer, não devemos esperar que a condição ZCS seja exatamente satisfeita. No entanto, pode-se afirmar que uma condição quase-ZCS com  $\xi = \mathbf{e}^T \mathbf{e}$  também próximo de zero é uma melhor indicação de que o DOA estimado está correto do que quando temos uma única condição.

Substituindo (4.6) em (4.5) e rearranjando a equação resultante, temos

$$\mathbf{e} = \left[ \mathbf{I} - \Delta \bar{\mathbf{P}}^T \left( \Delta \bar{\mathbf{P}} \Delta \bar{\mathbf{P}}^T \right)^{-1} \Delta \bar{\mathbf{P}} \right] \mathbf{n}. \quad (4.7)$$

De (4.7), observa-se que o vetor  $\mathbf{e}$  torna-se nulo ou quando  $\mathbf{n}$  for zero, ou quando pertence ao espaço gerado pelas colunas de  $\Delta \bar{\mathbf{P}}^T$ , o que é pouco provável. Assim, é proposto um algoritmo que seleciona uma combinação adequada de pares de microfones que atendam tanto a condição de quase-ZCS quanto o menor valor de  $\xi$ .

## 4.2 UM ALGORITMO PARA BUSCAR CFL (*SEARCHING CONSISTENT FUNDAMENTAL LOOP* OU SCFL)

Sejam P(1) a P(21) os 21 pares de microfones. Por exemplo, P(1) corresponde ao par (1,2), P(2) corresponde ao par (1,3), e assim sucessivamente. Armazenemos as posições dos 10 maiores picos da correlação cruzada de cada par de microfones, denotadas por HP[ $n, b$ ], onde  $n$  é o índice do par (de 1 a 21) e  $b$  é a ordem do pico (de 1 – o maior – até 10). Por exemplo, HP[2,3] significa o terceiro maior pico de P(2), o segundo par (1,3). A TAB. 4.1 resalta alguns FLs com quatro pares.

TAB. 4.1: FLs com quatro pares para  $M = 7$

FL	Par 1	Par 2	Par 3	Par 4
FL(1)	P(1)=(1,2)	P(7)=(2,3)	P(12)=(3,4)	P(3)=(4,1)
FL(2)	P(1)=(1,2)	P(7)=(2,3)	P(13)=(3,5)	P(4)=(5,1)
FL(3)	P(1)=(1,2)	P(7)=(2,3)	P(14)=(3,6)	P(5)=(6,1)
FL(4)	P(1)=(1,2)	P(7)=(2,3)	P(15)=(3,7)	P(6)=(7,1)
FL(5)	P(1)=(1,2)	P(8)=(2,4)	P(16)=(4,5)	P(4)=(5,1)
...	...	...	...	...
FL(35)	P(16)=(4,5)	P(19)=(5,6)	P(21)=(6,7)	P(18)=(7,4)

Note que o mesmo par de microfones  $i$  e  $j$  podem ser escritos como  $(i, j)$ , que corresponde ao TDOA  $\bar{\tau}_{ij}$ , ou como  $(j, i)$ , que corresponde ao TDOA  $\bar{\tau}_{ji}$ , onde  $\bar{\tau}_{ij} = -\bar{\tau}_{ji}$ . Na TAB. 4.1, “Par 1”, “Par 2” e “Par 3” correspondem a  $\bar{\tau}_{ij}$ , e “Par 4” corresponde a  $\bar{\tau}_{ji}$ , de modo que um FL seja realmente formado. Por exemplo, para FL(1), os TDOAs para os quatro pares são  $\bar{\tau}_{12}$ ,  $\bar{\tau}_{23}$ ,  $\bar{\tau}_{34}$  e  $\bar{\tau}_{41}$ .

De acordo com a TAB. 4.1, os possíveis valores de  $n$  em HP[ $n, b$ ], no caso de FL(1), são HP[1,  $b$ ], HP[7,  $b$ ], HP[12,  $b$ ] e HP[3,  $b$ ], já que P(1), P(7), P(12) e P(3) formam FL(1). Das 5.985 combinações possíveis de quatro pares em um arranjo de sete microfones, somente as 35 combinações que formam um FL são consideradas. Para cada uma dessas

35 combinações, a existência de loops fundamentais consistentes (neste caso, soma dos TDOAs próxima de zero, caracterizando a condição de quase-ZCS) precisa ser testada para os 10 maiores picos da correlação cruzada de cada par de microfones, totalizando  $10^4$  testes. A combinação que gera o menor valor da função de custo é utilizada para a estimativa de DOA. A TAB. 4.2 descreve um pseudo-código para o algoritmo SCFL, que estima o DOA utilizando os melhores quatro pares de microfones que satisfazem a condição de quase-ZCS e o menor valor de  $\xi$ . A condição de quase-ZCS é dada por  $20\log \left| \sum_{ij} \bar{\tau}_{ij} \right| < \gamma_{0dB}$ .

TAB. 4.2: Pseudo código para o algoritmo SCFL

**Para**  $l = 1$  a 35  
**Indique** os pares que constituem FL( $l$ ) por  $q(1)$  a  $q(4)$   
**Indique** os 10 maiores picos desses pares por  
 $HP[q(1), b]$  a  $HP[q(4), b]$ ,  $b$  de 1 a 10.  
**Selecione** um (dentre 10) pico de cada um dos 4 pares,  
um total de  $10^4$  combinações possíveis  
**Faça**  $k = 1$   
**Para**  $s = 1$  a  $10^4$   
**Calcule**  $\hat{\mathbf{a}}_{\text{DOA}}(l, s)$  utilizando (4.4) e guarde seu valor  
**Avalie** a função custo  $\xi$ , utilizando (4.2),  
denotando o resultado como  $\xi(l, s)$   
**Se** o FL for consistente  $\left( 20\log \left| \sum_{ij} \bar{\tau}_{ij} \right| < \gamma_{0dB} \right)$   
**Faça**  $pos(l, k) = s$   
**Faça**  $k = k + 1$   
**Fim**  
**Fim**  
**Fim**  
**Ache**  $(l, pos(l, k))$  com o menor  $\xi(l, pos(l, k))$  e utilize o resultado  
para estimar DOA como  $\hat{\mathbf{a}}_{\text{DOA}}(l, pos(l, k))$

Na TAB. 4.2, para  $l = 1$ , temos  $q(1) = 1$ ,  $q(2) = 7$ ,  $q(3) = 12$  e  $q(4) = 3$ , pois P(1), P(7), P(12) e P(3) formam FL(1), de acordo com a TAB. 4.1. Para  $l = 2$ , temos  $q(1) = 1$ ,  $q(2) = 7$ ,  $q(3) = 13$  e  $q(4) = 4$ , pois P(1), P(7), P(13) e P(4) formam FL(2), e assim sucessivamente.

O limiar  $\gamma_0$  aparece na TAB. 4.2 porque um cFL raramente apresenta soma de TDOAs igual a zero, devido a erros de quantificação e calibração. A fim de estabelecer o valor de  $\gamma_0$ , primeiramente definiu-se o máximo erro angular tolerável, para a aplicação de interesse (estimação de DOA de tiros de *sniper*), igual a  $1^\circ$ . A uma distância de 1.000 m, que requer grandes habilidades de um atirador treinado, o erro angular azimutal

correspondente a  $1^\circ$  é de aproximadamente 17,5 m, que pode ser facilmente coberto pelas forças de defesa. O limiar  $\gamma_0$  é escolhido de forma a conter erros na soma de TDOAs, em FLs, correspondentes a erros de até  $1^\circ$  na estimativa de DOA.

A fim de se determinar um valor apropriado para  $\gamma_0$ , foram simulados muitos cenários diferentes, utilizando dados reais de sinais de tiro e ruído ambiente gravados. Como referência de um sinal limpo (SNR > 20 dB), foi utilizado um sinal de MB de um rifle localizado a aproximadamente 200 m do microfone. Cada cenário correspondeu a um DOA selecionado aleatoriamente. Para simular o sinal de saída de cada microfone do arranjo para um dado DOA, o sinal limpo de MB foi apropriadamente deslocado no tempo e, em seguida, misturado com o ruído gravado, cuja amplitude foi dimensionada para produzir valores de SNR entre -8 dB e -5 dB. Essa faixa de SNR resultou em estimativas precisas de DOA com a soma de TDOAs não muito próxima de zero, para que o limiar  $\gamma_0$  pudesse ser ajustado adequadamente.

Para cada direção simulada aleatoriamente, ambos TDOAs e DOAs foram estimados considerando-se somente os maiores picos das correlações cruzadas e os 35 FLs obtidos das combinações de quatro pares. Se a diferença entre os DOAs estimado e simulado for menor que  $1^\circ$  (para azimute e zênite), o DOA é classificado como correto (cDOA), e armazena-se a correspondente soma de TDOAs em um vetor chamado “soma de TDOAs para cDOAs”. Caso contrário, o DOA é classificado como incorreto (iDOA), e armazena-se a correspondente soma de TDOAs em um vetor chamado “soma de TDOAs para iDOAs”.

De acordo com testes, alguns casos de iDOA possuem soma de TDOAs em um FL próximo de zero, talvez porque os erros dos TDOAs estimados se cancelem, ou porque pode haver um forte sinal interferente proveniente de uma direção diferente da desejada (do sinal de tiro). Após obter 4.000 cDOAs e 4.000 iDOAs, histogramas dos valores absolutos dos vetores “soma de TDOAs para cDOAs” e “soma de TDOAs para iDOAs” foram representados graficamente, conforme a FIG. 4.1. O limiar  $\gamma_{0dB}$  escolhido produz iguais probabilidades de falso alarme e de perda (EER - do inglês *Equal Error Rate*). O valor encontrado foi -97.64 dB ( $1,3125 \times 10^{-5}$  s), representado pela linha tracejada na FIG. 4.1, que corresponde a um EER igual a 5,42%. Portanto, um FL será considerado consistente se o valor absoluto da soma de suas estimativas de TDOA for menor que  $1,3125 \times 10^{-5}$  s. Para outras geometrias de arranjo, outro limiar deve ser calculado, seguindo o mesmo procedimento explicado nesta seção.

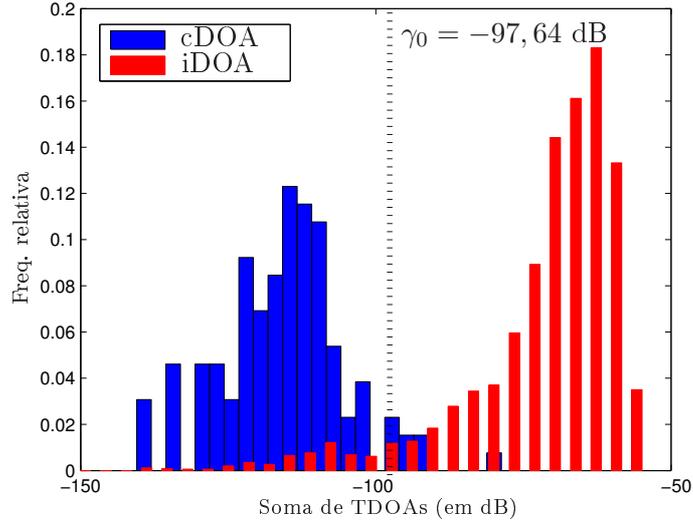


FIG. 4.1: Histograma da soma de TDOAs para determinação de  $\gamma_0$ .

Neste capítulo, foi introduzido o conceito de *loop* fundamental consistente na estimação de DOA, possibilitando que outros picos, além dos dominantes, das correlações cruzadas sejam avaliados. Esse procedimento é útil nos casos de sinais muito ruidosos em que o maior pico da correlação cruzada é devido ao ruído. Com a utilização desse conceito, limita-se a busca dentre todas as combinações de quatro pares, impedindo assim que a estimação de DOA seja ineficaz para aplicações em tempo real. O próximo capítulo propõe o emprego da técnica SCFL, junto com  $ES(n)$ , na concepção de um novo algoritmo para sinais muito ruidosos (menor que -7 dB).

## 5 O ALGORITMO ES-SCFL

Apesar dos algoritmos ES e SCFL serem adequados para sinais ruidosos, resultados experimentais com valores muito baixos de SNR (menor que -7 dB) têm mostrado que existem casos em que um falha enquanto o outro apresenta bom desempenho, e vice-versa em outros casos.

Motivado por essa observação experimental, foi proposto um algoritmo, que é uma combinação das técnicas SE, ES e SCFL. Tal algoritmo, apresentado na FIG. 5.1 e referido como ES-SCFL (BORZINO, 2016), utiliza uma árvore de decisão e compara a função custo com limiares cuidadosamente selecionados. Do capítulo anterior, pode-se assumir que um pequeno valor da função custo LS é um bom indicador de que o vetor DOA  $\mathbf{a}_{\text{DOA}}$  foi corretamente estimado. Depois de calcular  $\hat{\mathbf{a}}_{\text{DOA}}$  com  $n$  pares de microfones, o menor valor da função custo, com a notação empregada no capítulo anterior para  $n$  pares, é calculado como em  $\xi_n = \mathbf{e}^T \mathbf{e} / n$ , onde  $\mathbf{e} = \bar{\boldsymbol{\tau}} - \Delta \bar{\mathbf{P}}^T \hat{\mathbf{a}}_{\text{DOA}}$ . Note que, quando  $n = N$ , temos  $\xi_N = \xi_{\text{SE}}$ , o menor valor da função custo para SE. Da mesma forma, o menor valor da função custo obtido com uma busca exaustiva de cinco pares é representado como  $\xi_5$ . Seguindo esse raciocínio, “Computa ES(4)” na FIG. 5.1 significa “estime o vetor unitário  $\mathbf{a}_{\text{DOA}}$  com uma busca exaustiva de  $n = 4$  pares”, enquanto “função custo em ES (4)” significa o correspondente mínimo valor da função custo  $\xi_4$ .

A estimação padrão (SE) foi incluída para ser aplicada a sinais com alta SNR, pois essa técnica apresenta bons resultados na estimação de DOA no caso de sinais pouco ruidosos, não sendo necessário computar ES ou SCFL.

O algoritmo ES-SCFL utiliza o mínimo da função de custo na árvore de decisão, a partir da solução mais simples (SE) para a mais complexa (SCFL). Os limiares, designados como  $\gamma_1$ ,  $\gamma_2$  e  $\gamma_3$  na FIG. 5.1, são o ponto principal desse algoritmo e foram determinados conforme explicado a seguir. A configuração utilizada para a obtenção dos limiares é a mesma do Capítulo 4 no que diz respeito à simulação de DOAs aleatórios de sinais de tiro. Os valores de SNR utilizados foram: -1 dB para computar  $\gamma_1$ , -9 dB para computar  $\gamma_2$  e -10 dB para computar  $\gamma_3$ . Esses valores de SNR foram escolhidos, para cada limiar, com base no conhecimento prévio de qual faixa de SNR produziria uma quantidade aproximadamente igual de DOAs corretos e incorretos para o respectivo algoritmo. Se maiores valores de SNR fossem utilizados, teríamos mais cDOAs. Por

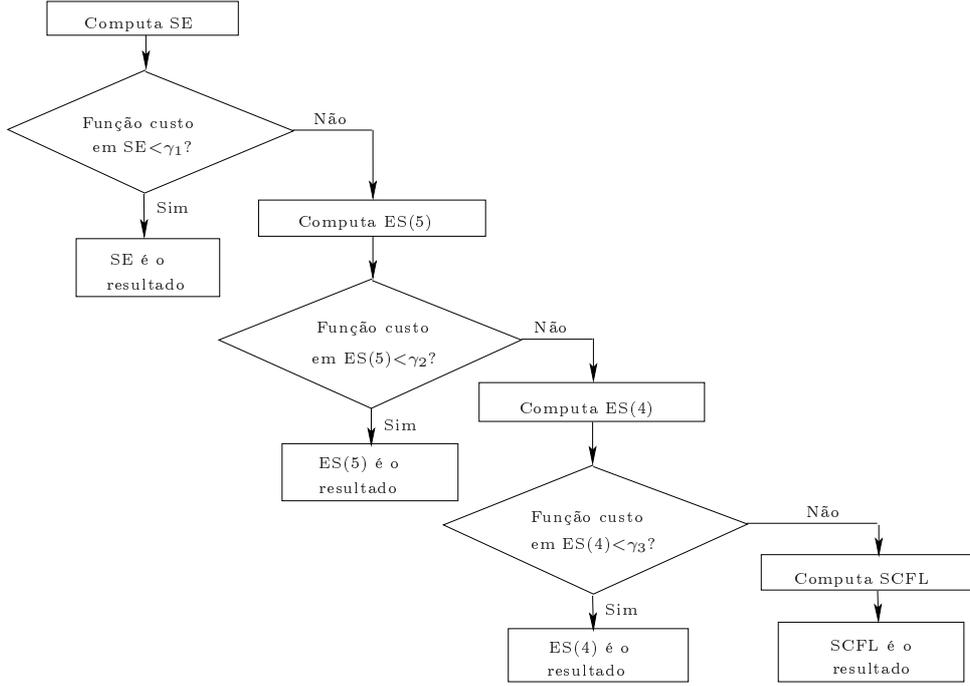


FIG. 5.1: O algoritmo proposto ES-SCFL (BORZINO, 2016).

outro lado, se menores valores de SNR fossem utilizados, teríamos mais iDOAs. Foram medidos erros angulares  $e_\phi = \phi - \hat{\phi}$  e  $e_\theta = \theta - \hat{\theta}$ . Quando  $|e_\theta|$  e  $|e_\phi|$  são ambos inferiores a  $1^\circ$ , classifica-se a estimativa como cDOA; caso contrário, classifica-se como iDOA. Após a estimação dos DOAs e o cálculo das funções custo correspondentes (em dB), dois histogramas foram gerados (um para 1.000 cDOA e outro para 1.000 iDOA) para SE, ES(5) e ES(4).

A FIG. 5.2 mostra os histogramas da função custo, em dB, para cDOA e iDOA com SE e também inclui uma curva de função densidade de probabilidade (PDF) que melhor modela visualmente o caso de iDOA: distribuição Logística (PAPOULIS, 2001) com média -76,6 dB e desvio padrão 2,9 dB. Como a intenção é utilizar a estimativa de DOA gerada por SE somente se tivermos razoável certeza que está correta, o limiar  $\gamma_1$ , representado pela linha tracejada, foi estabelecido como o maior valor da função custo que garante probabilidade de falso alarme desprezível. Nesse caso,  $\gamma_1 = -90,24$  dB, que corresponde a  $9,47 \times 10^{-10}$ . Então, se a função custo com todos os 21 termos (para o caso de SE) for menor que  $\gamma_1 = 9,47 \times 10^{-10}$ , considera-se que a estimativa de DOA obtida de SE é o resultado correto. Caso contrário, computa-se ES(5).

A FIG. 5.3 mostra os histogramas da função custo, em dB, para cDOA e iDOA com ES(5), assim como PDFs que melhor modelam visualmente cDOA (distribuição Gaussiana com média -121,3 dB e desvio padrão 8,4 dB) e iDOA (distribuição Gaussiana

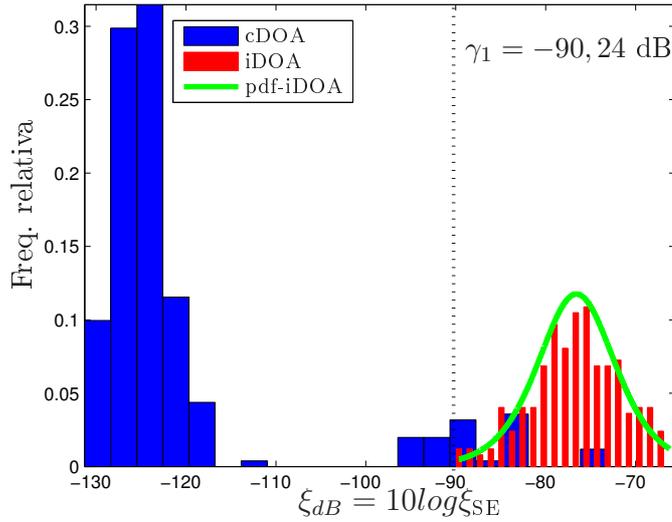


FIG. 5.2: Histogramas da função custo para SE e a escolha de  $\gamma_1$  tal que a probabilidade de falso alarme seja desprezível com a menor possível probabilidade de perda (BORZINO, 2016).

com média -103,3 dB e desvio padrão 4,8 dB). Note que, neste caso, a área de interseção dos histogramas é maior que a da FIG. 5.2, o que impede o uso da estratégia anterior (limiar tal que a probabilidade de falso alarme seja desprezível). Em vez disso, a partir das distribuições Gaussianas aproximadas, aplica-se o critério de Neyman-Pearson (TREES, 2004), maximizando a probabilidade de detecção e não permitindo que a probabilidade de falso alarme seja maior que 5%. Como resultado, o limiar  $\gamma_2$ , representado pela linha tracejada, é igual a  $-110,86$  dB, que corresponde a  $8,2 \times 10^{-12}$ .

A Fig. 5.4 mostra os histogramas da função custo, em dB, para cDOA e iDOA com ES(4) e PDFs aproximadas para cDOA (distribuição Logística com média -125,7 dB e desvio padrão 5,4 dB) e iDOA (distribuição de Valor Extremo com média -109,5 dB e desvio padrão 5,1 dB). Como resultado da aplicação do critério de Neyman-Pearson com probabilidade de falso alarme não maior que 5%, o limiar  $\gamma_3$ , representado pela linha tracejada, é igual a  $-126$  dB, que corresponde a  $2,5 \times 10^{-13}$ .

Pode-se notar que todos os limiares deste capítulo são muito baixos. Portanto, deve-se ter cuidado ao implementar esse algoritmo em um ambiente de precisão limitada. A fim de mostrar que os valores da função custo podem assumir baixas ordens de magnitude, realizemos a seguinte análise: a taxa de amostragem utilizada é de 96.000 amostras por segundo; interpolam-se as correlações cruzadas em torno do pico (FREIRE, 2011) por um fator de 100, a fim de se obter resultados mais precisos. Se a diferença entre os TDOAs estimado e calculado for de apenas uma amostra, isso corresponderá a aproximadamente

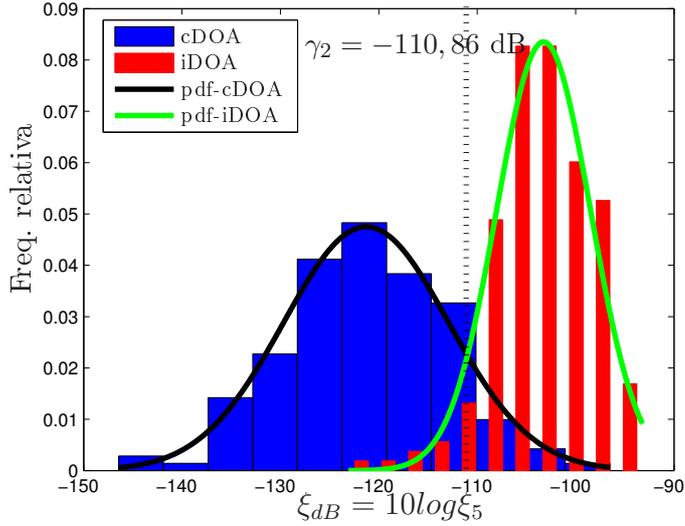


FIG. 5.3: Histogramas da função custo para ES(5) e a escolha de  $\gamma_2$  que resulta em probabilidade de falso alarme igual a 5% (BORZINO, 2016).

a  $1,04 \times 10^{-7}$  s. Como a função custo é composta dos quadrados das diferenças entre os TDOAs estimado e calculado, uma ordem de magnitude inferior a  $10^{-13}$  é possível de acontecer. Na prática, o intervalo possível da TDOAs vai de zero ao atraso máximo, que ocorre quando o som percorre a distância entre os dois microfones mais distantes (cerca de 1,62 ms no arranjo utilizado, assumindo velocidade de propagação do som igual a 340 m/s).

Todos esses limiares dependem do espaçamento inter-sensor e da quantidade de microfones. Se, pelo menos, um espaçamento inter-sensor for modificado, a distância que a frente de onda percorrerá também vai mudar e, como consequência, o TDOA será diferente e a função custo terá seu valor alterado, afetando os limiares. Além disso, se simplesmente um dos microfones do arranjo for excluído, o número de termos na função custo mudará, cujo valor também será diferente, alterando assim os limiares. No entanto, pode-se adotar os procedimentos descritos nesta tese para determinar quaisquer novos limiares.

A fim de avaliar a frequência de decisão do algoritmo ES-SCFL entre as técnicas SE, ES(5), ES(4) e SCFL, foram realizadas simulações com 500 DOAs aleatórios para cada SNR utilizada para determinar os limiares. A TAB. 5.1 mostra os resultados. Por exemplo, para  $SNR = -1$  dB, o algoritmo decide por SE e ES(5), respectivamente, 43,2% e 56,8% dos 500 DOAs aleatórios.

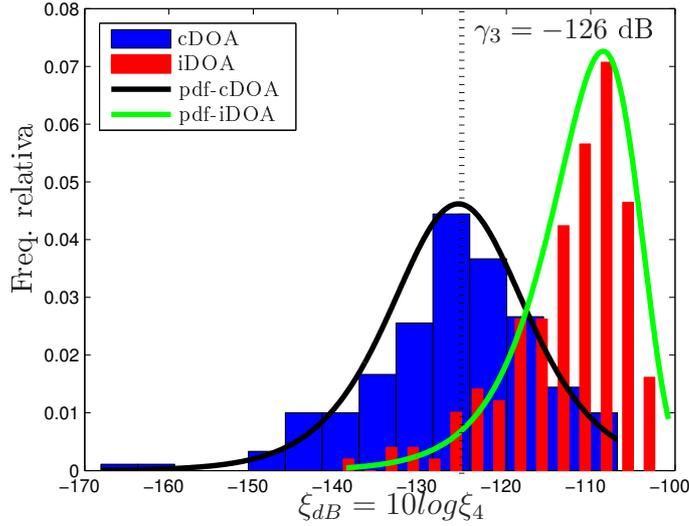


FIG. 5.4: Histogramas da função custo para ES(4) e a escolha de  $\gamma_3$  que resulta em probabilidade de falso alarme igual a 5% (BORZINO, 2016).

TAB. 5.1: Frequência de cada decisão de ES-SCFL

SNR	SE	ES(5)	ES(4)	SCFL
-1 dB	43,2%	56,8%	0%	0%
-9 dB	0%	40,4%	15,4%	44,2%
-10 dB	0%	16,2%	9,6%	74,2%

A TAB. 5.2 apresenta a porcentagem condicional de cDOA, dado que o algoritmo ES-SCFL decidiu por SE, ES(5), ES(4) ou SCFL, assim como a porcentagem total de cDOA (última coluna). Note que, exceto pela última coluna, todos os outros valores nesta tabela são porcentagens *condicionais*. Por exemplo, para SNR = -9 dB, 94,6% das decisões feitas por ES(5) produziram DOAs corretos.

TAB. 5.2: Porcentagem de cDOA

SNR	SE	ES(5)	ES(4)	SCFL	ES-SCFL
-1 dB	100%	100%	-	-	100%
-9 dB	-	94,6%	88,3%	82,4%	88,2%
-10 dB	-	90,0%	83,3%	65,2%	71,0%

Foi medido o tempo de execução de cada ramo do algoritmo ES-SCFL, a fim de inferir suas complexidades computacionais. Para isso, foi utilizando um notebook com 4 GB de memória RAM e processador Intel Core i5. Todos os experimentos foram realizados utilizando MATLAB<sup>®</sup>. Para o algoritmo SE, o tempo necessário é de aproximadamente

0,19 s, incluindo a rotina de correlações cruzadas, que também é utilizada pelos outros ramos. Para os algoritmos ES(5), ES(4) e SCFL, os tempos computacionais médios são 0,58 s, 0,17 s e 2,36 s, respectivamente. Para o algoritmo ES-SCFL como um todo, o tempo requerido para estimação de DOA depende da árvore de decisão da Fig. 5.1 que, por sua vez, depende da SNR do sinal. Por exemplo, para  $SNR = -9\text{dB}$ , considerando os resultados da TAB. 5.1, o tempo médio necessário para computar uma estimativa de DOA é de aproximadamente 1,5 s, utilizando MATLAB<sup>®</sup> e o notebook mencionado.

Como esta tese aborda sinais extremamente ruidosos, técnicas de redução de ruído, como SS (RAMOS, 2013; BOLL, 1979) ou MF (DUFAUX, 2000; CHACÓN-RODRÍGUEZ, 2011), exploradas no Capítulo 2, podem ser aplicadas para reduzir o ruído aditivo que contamina os sinais de tiro, antes do emprego de um esquema eficiente de estimação de DOA, como o algoritmo ES-SCFL (sem alterar os limiares). Particularmente, o uso do MF tem se mostrado efetivo para estimação de DOA (BORZINO, 2015b). Vale ressaltar que, para sinais com SNR abaixo de -2 dB, somente a aplicação de MF antes da estimação padrão não melhora tanto os resultados, conforme apresentado no Capítulo 2.

Neste capítulo, foi proposto um algoritmo de estimação de DOA, combinando a estimação padrão e os conceitos de busca exaustiva e *loop* fundamental consistente, e que pode ser empregado para sinais com qualquer valor de SNR.

## 6 RESULTADOS EXPERIMENTAIS DOS ALGORITMOS DE ESTIMAÇÃO DE DOA

A fim de avaliar o desempenho do algoritmo ES-SCFL, este foi primeiramente comparado com as técnicas ES(5), ES(4) e SCFL isoladamente, assim como com os algoritmos SE e ILS. Em seguida, o algoritmo ES-SCFL foi testado com sinais com vários DOAs simulados e com sinais de campo gravados.

### 6.1 COMPARANDO OS ALGORITMOS ES-SCFL, ES(5), ES(4) E SCFL

Com o intuito de mostrar que o algoritmo ES-SCFL pode melhorar os resultados de ES(5), ES(4) e SCFL, empregados isoladamente, foram simulados 20 DOAs aleatórios de sinais de tiro com SNR = -8 dB e medidos os erros angulares  $e_\phi = \phi - \hat{\phi}$  e  $e_\theta = \theta - \hat{\theta}$ . As FIGs. 6.1(a), (b) e (c) comparam o desempenho do algoritmo ES-SCFL com os algoritmos ES(5), ES(4) e SCFL, respectivamente. Além disso, a FIG. 6.1(d) compara o desempenho dos algoritmos ES-SCFL, SE e ILS. Observa-se que ES-SCFL produz os melhores resultados dentre todos os outros algoritmos.

Observa-se, na FIG. 6.1(a), que há três estimativas incorretas do algoritmo ES(5), que são corretamente estimadas pelo algoritmo ES-SCFL proposto, descrito na FIG. 5.1. Como a função custo é maior que  $\gamma_2$  para esses três sinais de tiro, o algoritmo calcula novas estimativas, de acordo com o algoritmo ES(4). Em um desses três casos, a estimativa obtida de ES(4) está correta, enquanto que, para os outros dois, a função custo é maior que  $\gamma_3$ , e o algoritmo calcula então a solução SCFL, produzindo estimativas corretas de DOA. Na FIG. 6.1(c), duas estimativas obtidas do algoritmo SCFL são incorretas, mas não correspondem aos mesmos sinais de tiro incorretamente estimados pelo algoritmo ES(5) e mostrados na FIG. 6.1(a). As duas estimativas incorretas na FIG. 6.1(c) podem ter ocorrido por causa de um sinal interferente suficientemente forte capaz de gerar um cFL, assim como o menor valor de função custo para uma dada combinação dos 10 maiores picos das correlações cruzadas.

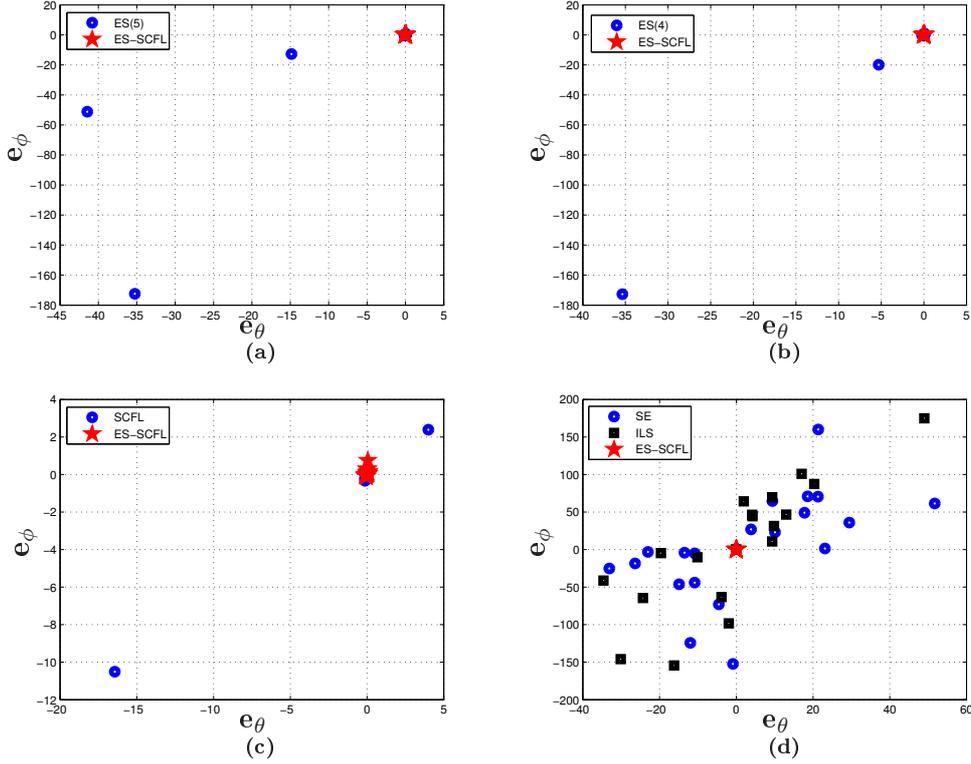


FIG. 6.1: Comparando o algoritmo ES-SCFL com os algoritmos: (a) ES(5), onde há 17 “•” e 20 “\*” sobrepostos representando erros aproximadamente nulos; (b) ES(4), onde há 18 “•” e 20 “\*” sobrepostos; (c) SCFL, onde há 18 “•” and 20 “\*” sobrepostos; (d) SE e ILS (BORZINO, 2016).

## 6.2 SINAIS COM DOAS SIMULADOS

Antes de testes com sinais reais, foram simulados DOAs de sinais de tiro, conforme detalhado no Capítulo 2. Já que ES(4) produz resultados satisfatórios para SNR menor que -3 dB (BORZINO, 2015a), variou-se a SNR de -5 dB para baixo, com intervalos de -1 dB, e foram comparados os resultados dos algoritmos SE, ILS e ES-SCFL. Foram simulados 500 DOAs para cada SNR. A FIG. 6.2 ilustra histogramas de erros azimutais e zenitais. Nesse caso, a SNR é de -8 dB, que é próxima à SNR estimada nos sinais gravados a uma distância de 929 metros da posição do atirador, com o arranjo de microfones fora do campo de visão da SW. Portanto, a SNR de -8 dB é considerada um caso típico para o cenário de um *sniper*. Neste caso de DOAs simulados, somente ruído foi introduzido (e não reflexões).

Da FIG. 6.2, observa-se que, para o algoritmo ES-SCFL, a grande maioria dos erros está próxima de zero. A FIG. 6.3 realça (para o ES-SCFL) a área próxima de zero. Então, pode-se verificar que a maioria dos erros possui valores absolutos menores que 1

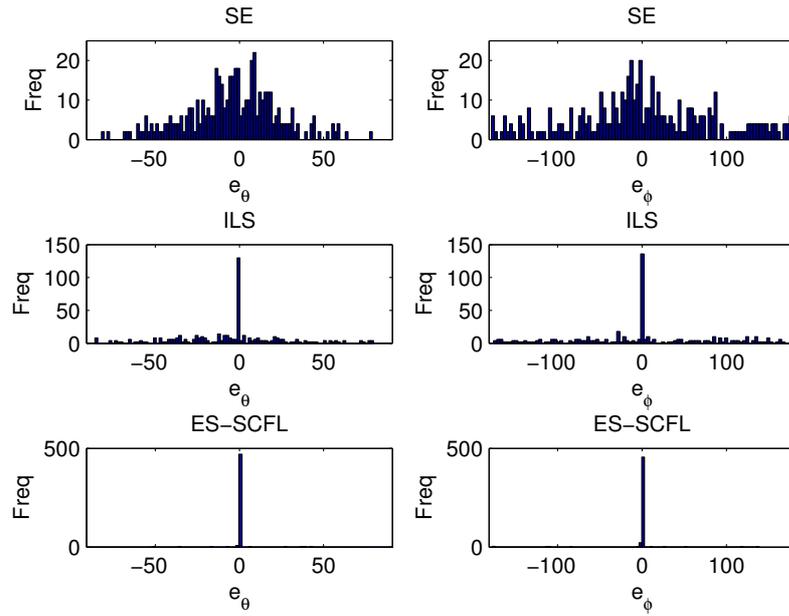


FIG. 6.2: Histogramas de erros angulares para  $\text{SNR} = -8$  dB.

grau. Para quantificar o desempenho dos resultados de estimação de DOA na FIG. 6.2, a TAB. 6.1 mostra, para vários valores de SNR, os resultados de  $\frac{\#(|e_\theta| \text{ e } |e_\phi|) < 1}{500} \times 100$ , para os três algoritmos abordados na FIG. 6.2: SE, ILS e ES-SCFL. Note que  $\#(|e_\theta| \text{ e } |e_\phi|) < 1$  significa o número de casos (dos 500 DOAs simulados) em que os valores absolutos dos erros em  $\theta$  e em  $\phi$  são ambos menores que um grau. Da TAB. 6.1, pode-se observar que o algoritmo ES-SCFL sempre produz a maior porcentagem de erros menor que um grau, quando comparado com os algoritmos SE e ILS.

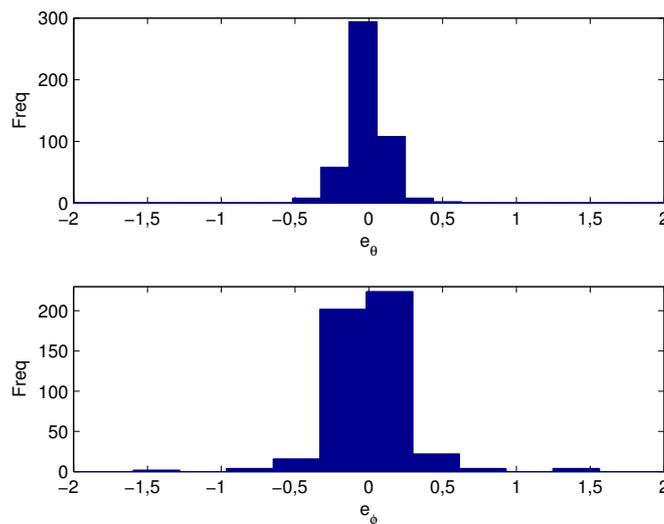


FIG. 6.3: Histogramas de erros angulares para o algoritmo ES-SCFL com  $\text{SNR} = -8$  dB.

TAB. 6.1: Porcentagem de resultados com erros menores que 1 grau  
(para 500 DOAs simulados)

SNR	SE	ILS	ES-SCFL
-5 dB	2,8%	81,6%	<b>100%</b>
-6 dB	0,6%	43,2%	<b>98,6%</b>
-7 dB	0,2%	30,6%	<b>97,0%</b>
-8 dB	0%	24,2%	<b>94,4%</b>
-9 dB	0%	13,4%	<b>88,2%</b>
-10 dB	0%	4,0%	<b>71,0%</b>

A TAB. 6.2 apresenta a média dos erros em  $\theta$ , uma estimativa de  $m_\theta \triangleq E[\theta - \hat{\theta}]$ , a média dos erros em  $\phi$ , uma estimativa de  $m_\phi \triangleq E[\phi - \hat{\phi}]$ , o desvio padrão do erro em  $\theta$ , uma estimativa de  $\sigma_\theta \triangleq \sqrt{E[(e_\theta - m_\theta)^2]}$  e o desvio padrão do erro em  $\phi$ , uma estimativa de  $\sigma_\phi \triangleq \sqrt{E[(e_\phi - m_\phi)^2]}$ , para os mesmos 500 sinais analisados nos histogramas da FIG. 6.2, onde SNR = -8 dB.

TAB. 6.2: Estatísticas de erros para SNR = -8 dB

Algoritmo	$\hat{m}_\theta$	$\hat{\sigma}_\theta$	$\hat{m}_\phi$	$\hat{\sigma}_\phi$
SE	-4,0	27,4	1,3	83,7
ILS	-5,3	12,6	1,1	44,9
ES-SCFL	0,13	5,4	-0,34	17,4

Da TAB. 6.2, pode-se concluir que a média dos erros do ES-SCFL é próxima de zero, e o desvio padrão do erro desse algoritmo é menor que o desvio padrão do SE e ILS. Apesar de apresentar o menor valor, o desvio padrão do erro do algoritmo ES-SCFL não é muito próximo de zero: isso acontece porque há alguns poucos erros próximos de  $50^\circ$  ou  $100^\circ$  (*outliers*) nos últimos dois gráficos (referentes ao ES-SCFL) da FIG. 6.2. Se desconsiderarmos os *outliers* no caso do algoritmo ES-SCFL, verifica-se que  $\hat{m}_\theta = 0,0326$ ,  $\hat{\sigma}_\theta = 0,33$ ,  $\hat{m}_\phi = -0,0652$  e  $\hat{\sigma}_\phi = 0,40$ .

Após isso, foram simulados novamente 500 DOAs aleatórios (para cada SNR), aplicando-se o “filtro mediana” (MF) explicado no Capítulo 2, e foram computados  $\frac{\#(|e_\theta| \text{ and } |e_\phi|) < 1}{500} \times 100$  para os algoritmos SE, ILS e ES-SCFL. A TAB. 6.3 apresenta os resultados com MF.

Da TAB. 6.3, notam-se melhores resultados com a aplicação de MF, principalmente para os algoritmos ILS e ES-SCFL. Se comparadas as TABs. 6.1 e 6.3 para o ES-SCFL, observa-se que os resultados obtidos sem MF a uma dada SNR<sub>dB</sub> são aproximadamente equivalentes àqueles com MF, mas com SNR<sub>dB</sub> - 5 dB.

TAB. 6.3: Porcentagem de resultados com erros menores que 1 grau (com MF)

SNR	SE	ILS	ES-SCFL
-8 dB	6,4%	91,2%	<b>100%</b>
-9 dB	2,2%	82,2%	<b>100%</b>
-10 dB	0,4%	64,6%	<b>99,6%</b>
-11 dB	0%	40,0%	<b>98,2%</b>
-12 dB	0%	21,8%	<b>96,4%</b>
-13 dB	0%	10,4%	<b>93,8%</b>

A TAB. 6.4 apresenta estimativas da média dos erros em  $\theta$  ( $m_\theta$ ), da média dos erros em  $\phi$  ( $m_\phi$ ), do desvio padrão do erro em  $\theta$  ( $\sigma_\theta$ ) e do desvio padrão do erro em  $\phi$  ( $\sigma_\phi$ ), para os mesmos 500 sinais com SNR = -8 dB, computados na TAB. 6.2, mas pré-processados com MF.

TAB. 6.4: Estatísticas de erros com MF para SNR = -8 dB

Algoritmo	$\hat{m}_\theta$	$\hat{\sigma}_\theta$	$\hat{m}_\phi$	$\hat{\sigma}_\phi$
SE	-2,36	10,2	-4,75	1.642,1
ILS	-0,99	8,85	-3,53	599,9
ES-SCFL	0,004	0,17	0,006	0,26

Da TAB. 6.4, observa-se que, para o cenário crítico de SNR = -8 dB, a aplicação de MF faz com que a média dos erros do algoritmo ES-SCFL fique mais próxima de zero e com que o desvio padrão do erro desse algoritmo sejam muito menor que o desvio padrão sem MF.

### 6.3 SINAIS GRAVADOS

A fim de avaliar o desempenho do algoritmo proposto em sinais de tiro gravados (reais), foram testados todos os algoritmos SE, ILS, ES(5), ES(4), SCFL e ES-SCFL em 60 tiros com três diferentes posições. As gravações foram feitas no Centro de Avaliações do Exército (CAEx), localizado na Restinga da Marambaia, Rio de Janeiro. O arranjo foi posicionado em um local com visada direta entre os microfones e o atirador, mas fora do campo de visão da SW. O azimute e o ângulo vertical corretos para cada posição foram medidos por uma estação topográfica “TOPCON”, modelo CTS3000, com precisão linear de 3 mm e precisão angular de 5”. As armas utilizadas nesse experimento foram um Fuzil Automático Leve 7,62 mm M964 (762) e um rifle de assalto Tavor 5,56 mm (556). A FIG. 6.4 ilustra o local de tiro.

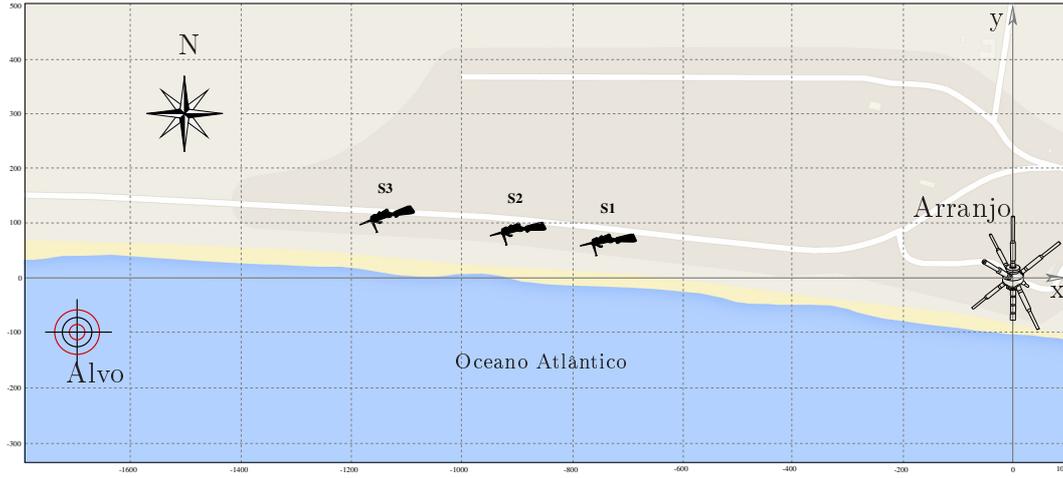


FIG. 6.4: Local de tiro: o arranjo de microfones foi posicionado na origem, que corresponde a coordenadas UTM (Zona 23 K) 643.826,86 m E e 7.450.217,78 m S. Atiradores S1, S2 e S3 estão a seguintes distâncias do arranjo: 758 m, 929 m e 1.166 m, respectivamente.

Inicialmente, o algoritmo ES-SCFL foi aplicado em 20 sinais de tiro (10 de cada arma), em que a distância entre o atirador e o arranjo era de 1.166 m (a maior distância entre as três posições), e os resultados foram comparados com os algoritmos SE e ILS. Nesse caso, a SNR média é igual a  $-10,21$  dB e  $-9,11$  dB, referentes aos rifles 5,56 mm e 7,62 mm, respectivamente. A FIG. 6.5 mostra o erro azimutal ( $e_\phi$ ) e o erro zenital ( $e_\theta$ ) quando comparados os três algoritmos para esses 20 sinais de tiro.

Na FIG. 6.5, observa-se que o ES-SCFL produz menores erros que os outros algoritmos. Para quantificar o erro de cada algoritmo, para cada arma, foram calculados os erros, utilizando RMSE, assim definidos:

$$E_\theta = \sqrt{\frac{1}{10} \sum_{k=1}^{10} (e_{\theta,k})^2}, \quad e \quad (6.1)$$

$$E_\phi = \sqrt{\frac{1}{10} \sum_{k=1}^{10} (e_{\phi,k})^2}, \quad (6.2)$$

onde o índice “ $k$ ” refere-se ao  $k$ -ésimo tiro.

A TAB. 6.5 mostra os resultados das Eqs. (6.1) e (6.2) aplicadas a esses tiros disparados a 1.166 m do arranjo de microfones.

Da TAB. 6.5, observa-se que o algoritmo ES-SCFL produz os menores erros, tanto em  $\theta$  como em  $\phi$ . Além disso, os erros originados do rifle de assalto 5,56 mm são maiores que os erros originados do fuzil 7,62 mm. Isso é justificado pelo fato de que a SNR é

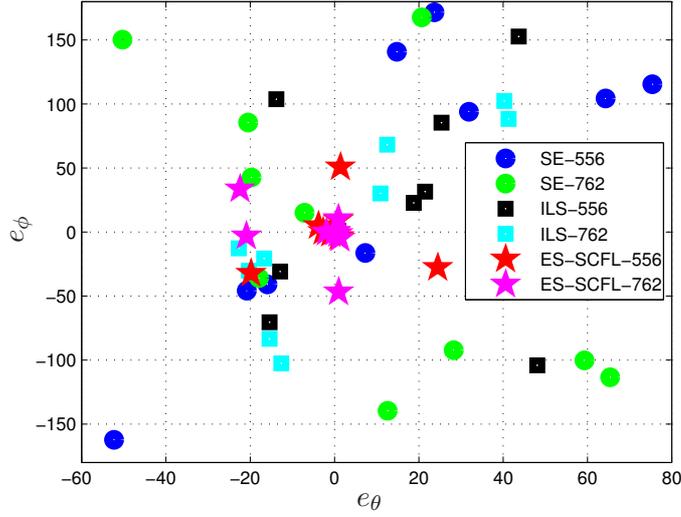


FIG. 6.5: Desempenho do algoritmo ES-SCFL (frente a seus concorrentes) com sinais gravados a uma distância de 1.166 m, com SNRs média de  $-10, 21$  dB e  $-9, 11$  dB (BORZINO, 2016).

TAB. 6.5:  $E_\theta$  e  $E_\phi$  para sinais gravados a uma distância de 1.166 m

Arma	Erro	SE	ILS	ES-SCFL
5,56 mm	$E_\theta$	43,8	39,2	10,2
7,62 mm		31,0	25,4	9,6
5,56 mm	$E_\phi$	105,2	82,6	21,1
7,62 mm		102,4	66,9	18,5

menor quando são utilizados cartuchos 5,56 mm, do que quando são utilizados cartuchos 7,62 mm. Verifica-se que os erros (RMSE) obtidos na TAB. 6.5 são maiores que o desvio padrão dos erros da TAB. 6.2, o que é explicado pelo fato dos valores de SNR dos tiros da TAB. 6.5 serem maiores que a SNR dos sinais da TAB. 6.2.

Na sequência, os algoritmos ES-SCFL, ILS e SE foram aplicados em outros 20 sinais de tiro (10 com o fuzil 7,62 mm e 10 com o fuzil 5,56 mm) cuja distância entre o atirador e o arranjo era de 929 m. Nesse caso, a SNR média é igual a  $-8.58$  dB e  $-7.45$  dB, referentes aos rifles 5,56 mm e 7,62 mm, respectivamente. A FIG. 6.6 mostra o erro azimutal ( $e_\phi$ ) e o erro zenital ( $e_\theta$ ) quando comparados os três algoritmos para esses 20 sinais de tiro.

Na FIG. 6.6, observa-se que o ES-SCFL também produz erros menores que os outros algoritmos e que esses erros são menores que os da FIG. 6.5, visto que a SNR aqui é maior. A TAB. 6.6 mostra os resultados dos erros, conforme as Eqs. (6.1) e (6.2) aplicadas a esses tiros disparados a 929 m do arranjo de microfones. Da TAB. 6.6, observa-se que

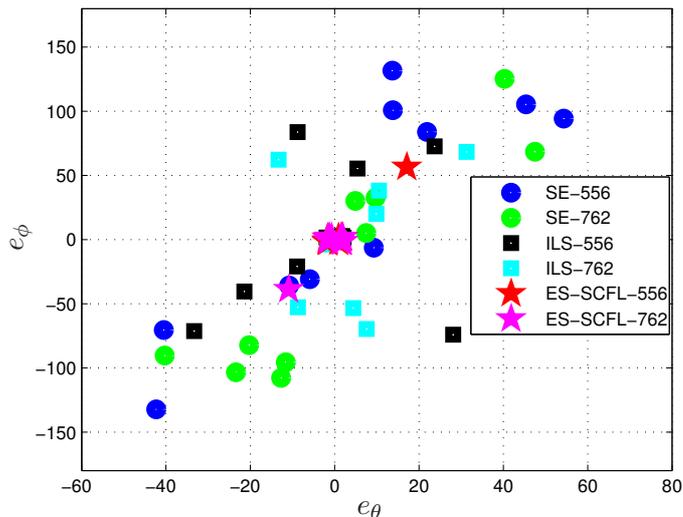


FIG. 6.6: Desempenho do algoritmo ES-SCFL com sinais gravados a uma distância de 929 m, com SNRs média de  $-8,58$  dB e  $-7,45$  dB (BORZINO, 2016).

o algoritmo ES-SCFL também produz os menores erros, tanto em  $\theta$  como em  $\phi$ . Além disso, pode-se verificar novamente que os erros originados do rifle de assalto 5,56 mm são maiores que os erros originados do fuzil 7,62 mm.

TAB. 6.6:  $E_\theta$  e  $E_\phi$  para sinais gravados a uma distância de 929 m

Arma	Erro	SE	ILS	ES-SCFL
5,56 mm	$E_\theta$	30,9	17,6	5,5
7,62 mm		26,3	12,4	3,7
5,56 mm	$E_\phi$	89,8	53,7	17,9
7,62 mm		82,8	45,2	12,4

Finalmente, os algoritmos sob investigação foram aplicados a outros 20 sinais de tiro (metade deles de cada rifle) cuja distância entre o atirador e o arranjo era de 758 m. Nesse caso, a SNR média é igual a  $-7.56$  dB e  $-6.49$  dB, referentes aos rifles 5,56 mm e 7,62 mm, respectivamente. A FIG. 6.7 mostra o erro azimutal ( $e_\phi$ ) e o erro zenital ( $e_\theta$ ) quando comparados os três algoritmos para esses 20 sinais de tiro.

Na FIG. 6.7, observa-se novamente que o algoritmo ES-SCFL produz erros menores que o SE e o ILS e que esses erros são mais próximos de zero do que as duas figuras anteriores.

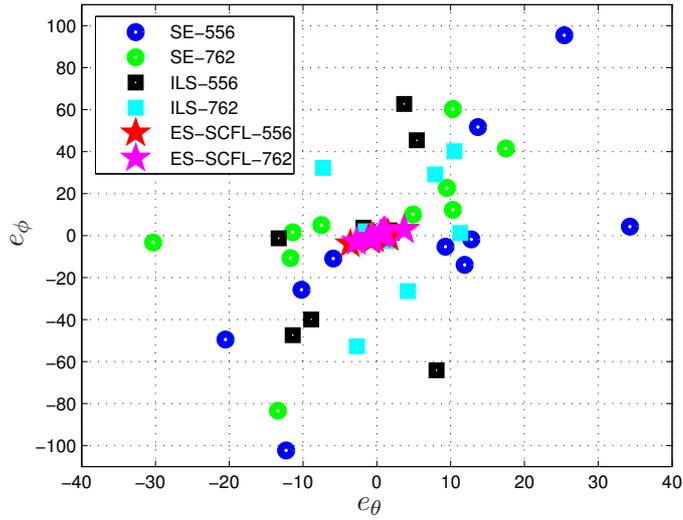


FIG. 6.7: Desempenho do algoritmo ES-SCFL com sinais gravados a uma distância de 758 m, com SNR média de  $-7,56$  dB (5.56 mm) e  $-6,49$  dB (7.62 mm).

A TAB. 6.7 mostra os resultados de  $E_\theta$  and  $E_\phi$  aplicados a esses tiros disparados a 758 m do arranjo de microfones. Da TAB. 6.7, observa-se que o algoritmo ES-SCFL também produz os menores erros, tanto em  $\theta$  como em  $\phi$ , para distâncias menores, indicando que ele pode ser utilizado em casos com SNR mais alta. Além disso, podemos observar, mais uma vez, que os erros originados do rifle 5,56 mm são maiores que os originados do rifle 7,62 mm.

TAB. 6.7:  $E_\theta$  e  $E_\phi$  para sinais gravados a uma distância de 758 m

Arma	Erro	SE	ILS	ES-SCFL
5,56 mm	$E_\theta$	17,6	7,1	1,9
7,62 mm		14,3	6,2	1,6
5,56 mm	$E_\phi$	50,7	37,0	2,0
7,62 mm		36,4	26,5	1,7

Foi utilizada outra métrica para quantificar a acurácia de todos os algoritmos: a função distribuição acumulada (cdf, do inglês *cumulative distribution function*)  $F(x)$ , denotando a probabilidade do erro ser menor que  $x$ . Neste trabalho, erro engloba tanto azimute quanto zênite. Por exemplo, para  $x = 1$ ,  $F(x)$  denota a probabilidade de ambos os erros azimutal e zenital serem menores que  $1^\circ$ . As FIGs. 6.8, 6.9 e 6.10 ilustram as funções distribuição acumulada para os sinais gravados a distâncias de, respectivamente, 1.166 m, 929 m e 758 m.

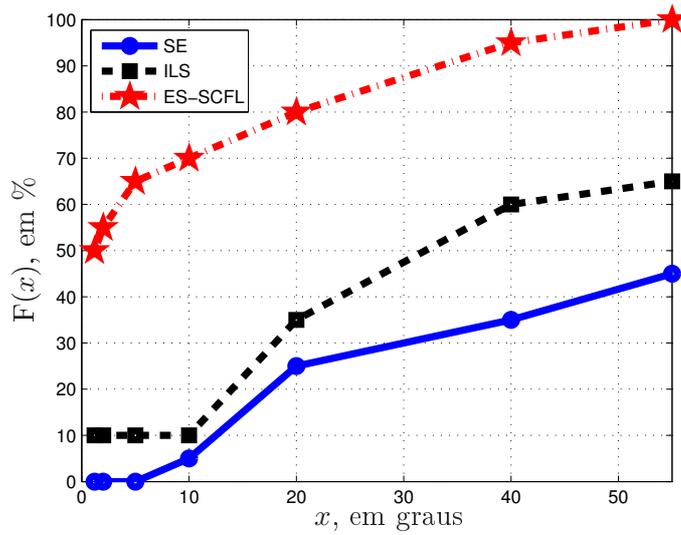


FIG. 6.8: Distribuição de probabilidade cumulativa para sinais gravados a uma distância de 1.166 m (BORZINO, 2016).

Das FIGs. 6.8, 6.9 e 6.10, observa-se que o algoritmo ES-SCFL possui desempenho melhor que os algoritmos SE e ILS, consistente com os resultados obtidos com direções simuladas. Particularmente, da FIG. 6.9, pode-se verificar que a cdf para o algoritmo ES-SCFL aumenta abruptamente para pequenos ângulos (em torno de  $1^\circ$ ). Já na FIG. 6.10, referente à menor das três distâncias analisadas, a cdf para o algoritmo ES-SCFL já atinge seu máximo quando  $x = 4^\circ$ .

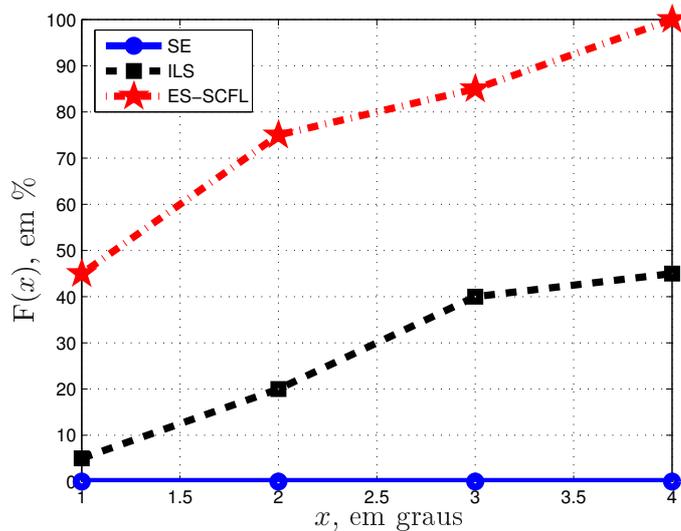


FIG. 6.10: Distribuição de probabilidade cumulativa para sinais gravados a uma distância de 758 m.

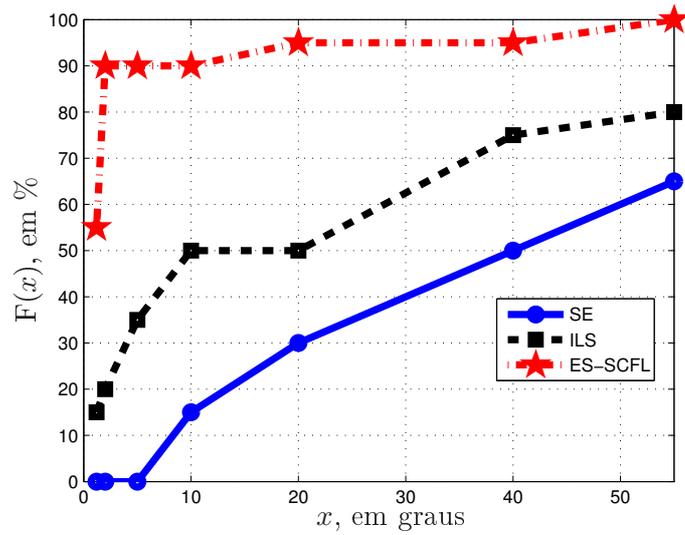


FIG. 6.9: Distribuição de probabilidade cumulativa para sinais gravados a uma distância de 929 m (BORZINO, 2016).

Neste capítulo, foi avaliado o desempenho do algoritmo ES-SCFL e comparado com os outros algoritmos da estado da arte. Foram feitos testes com DOAs simulados e com sinais reais gravados no Centro de Avaliações do Exército. Em todos os casos, o algoritmo ES-SCFL produziu os melhores resultados.

## 7 CLASSIFICAÇÃO DE MUNIÇÃO

Neste capítulo, será abordada a segunda aplicação de interesse deste trabalho: classificação de munição ou armamento, útil na área de análise forense. O termo “classificação” é utilizado quando assume-se que o sinal que queremos analisar (sinal de teste) pertence a um dado (conhecido) conjunto de categorias ou classes (SCHEIRER, 2013). Neste caso, dizemos que estamos lidando com classificação em conjunto fechado (*closed set*). Por outro lado, o termo “reconhecimento” refere-se a casos em que o sinal de teste pode não pertencer às classes consideradas. Neste caso, dizemos que estamos lidando com reconhecimento em conjunto aberto (*open set*). Esta tese focou em aplicações *closed set*.

### 7.1 MÉTODOS

Em reconhecimento de padrões, atributos são propriedades mensuráveis individuais obtidas dos dados (BISHOP, 2006), sendo a seleção de atributos um passo crucial para alcançar um bom desempenho em tarefas de classificação. É desejável escolher atributos que não mudam muito entre os dados de uma mesma classe, mas que variam significativamente quando comparamos duas ou mais classes distintas. Trabalhos recentes publicados quantizaram o espectrograma de um sinal em níveis de cinza (tratado como imagem), e utilizaram esses atributos em aplicações de radar (TIAN, 2013) e em classificação de armamento após um pré-processamento que realiza uma deconvolução (BORZINO, 2014b). Em (MAYS, 2001), classificação de sinais de tiro entre suas componentes (SW e MB) apresentou bom desempenho com a utilização da WT. Esta tese combina as ideias de WT e quantização em níveis de cinza na classificação de munição e estende o método descrito em (TIAN, 2013), que considera somente atributos derivados dos valores quantizados da intensidade de nível de cinza de cada pixel do espectrograma, não considerando a relação entre pixels vizinhos. Neste trabalho, leva-se em conta esta relação.

A WT de um sinal é computada como (DAUBECHIES, 1992)

$$C(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi \left( \frac{t - b}{a} \right) dt, \quad (7.1)$$

onde  $x(t)$  é o sinal analisado,  $\psi(t)$  é a wavelet, que translada de acordo com o deslocamento (salto)  $b$ , e  $a$  é o fator de escala, que estende ou comprime a wavelet, que é finita por definição. Como estamos lidando com sinais discretos, a integral em (7.1) torna-se

uma soma das amostras envolvidas para os valores particulares de  $a$  e  $b$ . Em outras palavras, o número  $C(a, b)$  representa o quanto a wavelet se correlaciona ou se assemelha com uma seção específica do sinal. Se o valor de  $b$  for alterado, a wavelet será deslocada e comparada com outras seções do sinal. Quanto maior  $C$ , maior a semelhança. Os resultados dependem da forma da wavelet escolhida. Note que  $C$  é bidimensional e pode ser visto como uma matriz.

Para a quantização da WT em  $N$  níveis de cinza (nos testes,  $N = 256$  apresentou bons resultados), depois que  $C$  é obtido para o intervalo de valores desejados de  $a$  e  $b$ , normaliza-se  $C$ , de modo que seus elementos variem entre zero e um. Esta gama de valores normalizados é então dividida em  $N$  bins (faixas) de larguras iguais. Valores em um bin são mapeados em um único nível de cinza. Em seguida, os elementos da matriz  $C$  são transformados em pixels, gerando uma imagem, cujo eixo horizontal é o deslocamento (número de amostras), e o eixo vertical representa a escala (utilizou-se 64 escalas). Quanto maior  $C$ , maior a intensidade de nível de cinza e mais claro o pixel.

Para cada sinal de tiro, primeiramente toma-se a WT da componente MB, quantiza-a em níveis de cinza, utilizando-a como imagem, e extraem-se atributos de textura (HARALICK, 1973), (GALLOWAY, 1975) em um típico cenário de reconhecimento de padrões. Os atributos servem então como entrada de um classificador *Support Vector Machine* (SVM) (CORTES, 1995), (CRISTIANINI, 2000) para possibilitar reconhecimento de munição. Foram testadas três wavelets, pelas suas semelhanças com o MB: *daubechies2*, *sym7* e *biorthogonal1.5* (MALLAT, 1999), e esta última apresentou os melhores resultados. Por esta razão, serão apresentados apenas os resultados com a wavelet *biorthogonal1.5*.

Indicam-se, nas FIGs. 7.1 e 7.2, sinais no domínio do tempo relativos às componentes MB de mesmo calibre e de quatro diferentes calibres, respectivamente. Na FIG. 7.1, a munição utilizada foi 12.7 mm (.50 BMG), enquanto que, na FIG. 7.2, as munições foram 5.56 mm, 7.62 mm, 9 mm e 12.7 mm (.50 BMG).

Da FIG. 7.1, observa-se que as componentes MB do mesmo calibre no domínio do tempo são parecidas e, da FIG. 7.2, que as quatro componentes MB são subjetivamente diferentes umas das outras, características essas desejáveis na tarefa de classificação. Como a WT produz uma representação objetiva em tempo-frequência de um sinal (DEBNATH, 2003), é razoável supor que as WT das componentes MB na FIG. 7.2 também serão diferentes. A FIG. 7.3 corrobora esta afirmação, uma vez que mostra as WT (em 256 níveis de cinza) dos quatro sinais de tiro, respectivamente. Quanto maior o coeficiente da WT,

mais claro é o pixel.

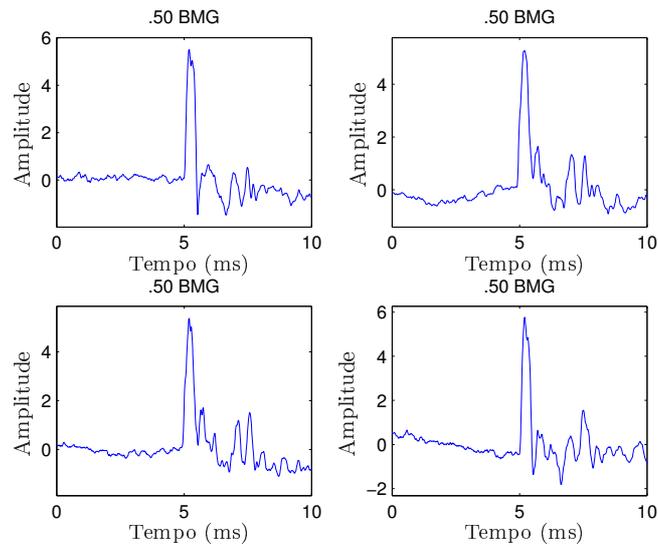


FIG. 7.1: Componentes MB de mesmo calibre.

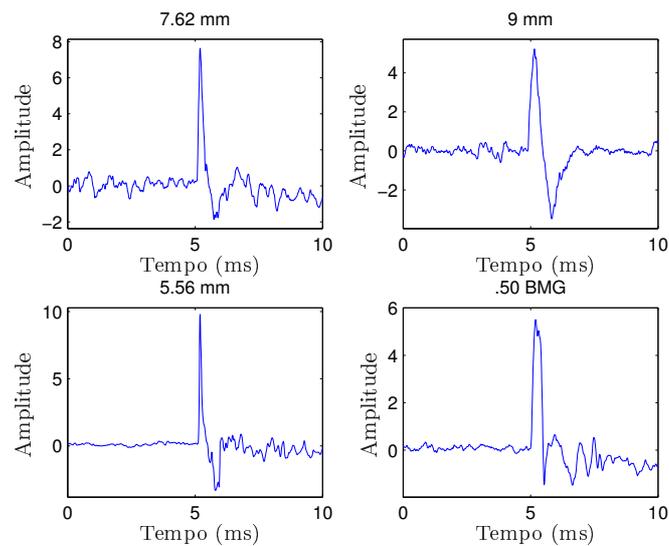


FIG. 7.2: Componentes MB de quatro diferentes calibres.

Da FIG. 7.3, se olharmos cuidadosamente, podemos notar pequenas diferenças entre seus quatro gráficos. Entretanto, quando os atributos de textura são calculados, mais diferenças relevantes aparecem, sugerindo que atributos obtidos da WT tratada como imagem são potenciais discriminantes interclasse. Mais adiante neste capítulo, quando forem apresentadas duas matrizes utilizadas para gerar alguns atributos de textura, essa afirmação será corroborada.

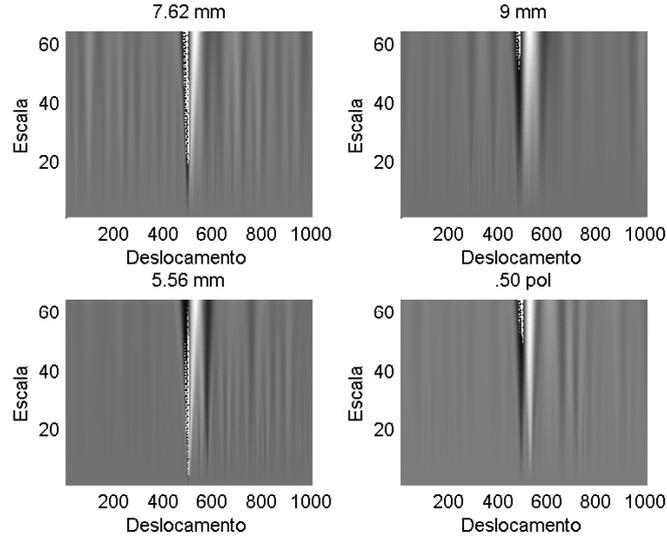


FIG. 7.3: Transformadas wavelet do MB de quatro diferentes calibres.

Vários métodos podem ser utilizados para caracterizar uma imagem (SELVARAJAH, 2011). O primeiro leva em conta somente o valor quantizado da intensidade de nível de cinza de cada pixel da imagem da WT, sem considerar as relações com pixels vizinhos. Atributos derivados dessa abordagem são explicados no Apêndice e incluem energia, entropia, correlação, contraste, momento diferencial inverso, parâmetro R, *skewness*, *kurtosis* (TIAN, 2013), (SELVARAJAH, 2011), todos utilizados nos experimentos.

O segundo método caracteriza imagens de textura baseado nos “comprimentos de corrida” de níveis de cinza. Pixels consecutivos de mesmo valor (nível) de cinza da imagem da WT, numa dada direção, constituem uma corrida. Assim, podemos formar a chamada Matriz de Comprimento de Corrida (RLM, do inglês *Run Length Matrix*) (SELVARAJAH, 2011), (CHU, 1990), cujo elemento  $p(i, j)$  refere-se à quantidade de corridas de pixels de nível de cinza  $i$  e comprimento de corrida  $j$ , numa dada direção (horizontal, vertical ou diagonal). Nos experimentos, foram consideradas todas as direções. Por exemplo,  $p(3, 5) = 2$  significa que 5 pixels consecutivos de nível de cinza 3 aparecem duas vezes na imagem inteira da WT. A partir da RLM, extraem-se atributos desse segundo método de caracterização de imagem. Para ilustrar melhor esse conceito de comprimento de corrida, considere, por exemplo, que uma imagem quadrada possua 25 pixels quantizados em 8 níveis de cinza (de 1 a 8), conforme a TAB. 7.1.

Da TAB. 7.1, ao considerarmos todas as direções, podemos inferir, por exemplo, que:

- $p(3, 4) = 2$ , ou seja, o elemento da terceira linha e quarta coluna da RLM vale 2;

TAB. 7.1: Exemplo de imagem quantizada

1	3	3	3	3
8	5	7	3	1
8	5	5	3	3
8	4	6	3	5
4	1	6	5	5

- $p(5, 2) = 6$ , ou seja, o elemento da quinta linha e segunda coluna da RLM vale 6;
  - $p(3, 2) = 3$ , ou seja, o elemento da terceira linha e segunda coluna da RLM vale 3;
- etc.

A FIG. 7.4 mostra a RLM (também normalizada e quantizada em 256 níveis de cinza) das quatro WT da FIG. 7.3, respectivamente. Para ratificar a afirmação de que os atributos de textura são capazes de discriminar os quatro calibres citados, observe na FIG. 7.4 que existem diferenças visíveis entre os gráficos. Nesta figura, um dos eixos deveria variar de 1 a 256 (quantidade de níveis de cinza utilizados), e o outro eixo deveria variar de 1 até o número máximo de pixels consecutivos com o mesmo nível de cinza; porém, somente a parte da figura (eixos) que realça as diferenças entre os calibres é mostrada.

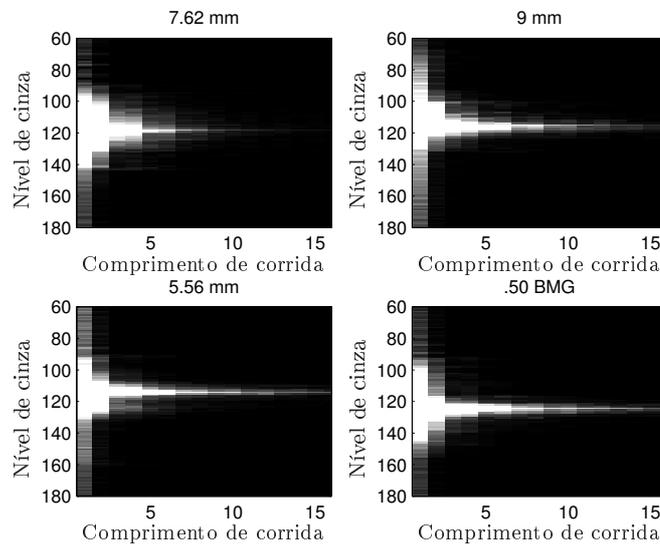


FIG. 7.4: RLM das quatro WT apresentadas anteriormente.

Atributos derivados da RLM, detalhados no Apêndice, são (SELVARAJAH, 2011): *Short Run Emphasis (SRE)*, *Long Run Emphasis (LRE)*, *Gray-Level Nonuniformity (GLN)*, *Run Percentage (RP)*, *Low Gray-Level Run Emphasis (LGRE)*, *High Gray-Level*

*Run Emphasis* (HGRE), *Short Run Low Gray-Level Emphasis* (SRLGE), *Short Run High Gray-Level Emphasis* (SRHGE), *Long Run Low Gray-Level Emphasis* (LRLGE) e *Long Run High Gray-Level Emphasis* (LRHGE).

O terceiro e último método para caracterizar uma imagem baseia-se na “matriz de co-ocorrência de níveis de cinza” (GLCM, do inglês *Gray Level Co-occurrence Matrix*) (HARALICK, 1973), SELVARAJAH (2011), cujo elemento  $p_d(i, j)$  refere-se à quantidade de vezes que um pixel de nível de cinza  $i$  é adjacente a um pixel de nível de cinza  $j$  da imagem da WT. A adjacência pode ser horizontal (para esquerda ou direita), vertical (para cima ou para baixo) ou diagonal (direções NE, SE, NW ou SW). Nos experimentos, foram computadas todas as oito direções. Por exemplo,  $p_d(3, 5) = 2$  significa que pixels de níveis de cinza 3 e 5 são adjacentes em duas ocasiões na imagem inteira da WT. Note que  $p_d(i, j)$  pode ser igual ou diferente de  $p_d(j, i)$ , dependendo das direções consideradas para computar a GLCM. A partir da GLCM, extraem-se atributos desse terceiro método de caracterização de imagem. Para melhor entendimento, considere, por exemplo, que uma imagem quadrada possua 25 pixels quantizados em 8 níveis de cinza (de 1 a 8), conforme a TAB. 7.2.

TAB. 7.2: Outro exemplo de imagem quantizada

1	3	3	3	3
8	5	7	3	1
8	5	5	3	3
8	4	6	3	5
4	1	6	5	5

Da TAB. 7.2, ao considerarmos todas as direções, podemos inferir, por exemplo, que:

- $p_d(3, 1) = p_d(1, 3) = 6$ , ou seja, o elemento da terceira linha e primeira coluna e o da primeira linha e terceira coluna da GLCM valem 6;
- $p_d(5, 3) = p_d(3, 5) = 10$ , ou seja, o elemento da quinta linha e terceira coluna e o da terceira linha e quinta coluna da GLCM valem 10;
- $p_d(8, 5) = p_d(5, 8) = 5$ , ou seja, o elemento da oitava linha e quinta coluna e o da quinta linha e oitava coluna da GLCM valem 5; etc.

A FIG. 7.5 mostra a GLCM (também quantizada em 256 níveis de cinza) das quatro WT da FIG. 7.3, respectivamente. Novamente, para confirmar que os atributos de textura são potenciais discriminantes interclasse (para a tarefa de classificação de munição),

observe na FIG. 7.5 que há diferenças visíveis entre os gráficos. Nesta figura, os dois eixos deveriam variar de 1 a 256 (quantidade de níveis de cinza utilizados), mas somente a parte da figura que realça as diferenças entre os calibres é mostrada.

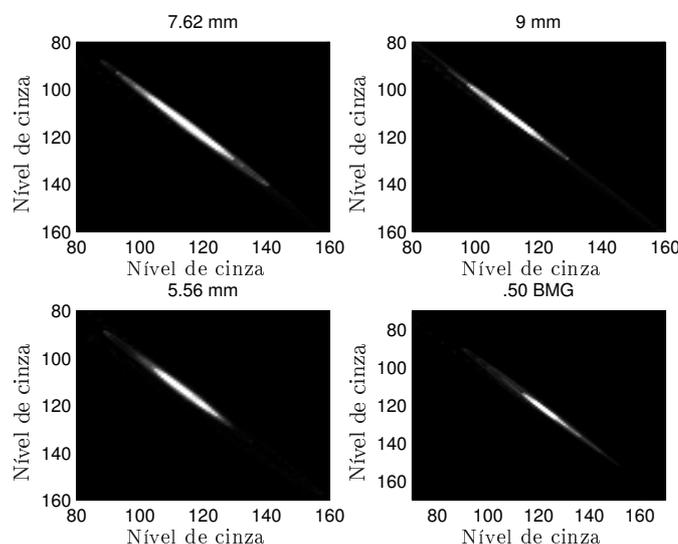


FIG. 7.5: GLCM das quatro WT apresentadas anteriormente.

Atributos derivados da GLCM são apresentados no Apêndice e incluem: energia, entropia, correlação, contraste, momento diferencial inverso, variância, dissimilaridade, homogeneidade (HARALICK, 1973), (SELVARAJAH, 2011), todos utilizados nos experimentos.

Note que alguns atributos (energia, entropia, correlação, contraste e momento diferencial inverso) são definidos tanto na primeira abordagem (que leva em conta somente os níveis de cinza de cada pixel) (TIAN, 2013), (SELVARAJAH, 2011), como para a GLCM (HARALICK, 1973), (SELVARAJAH, 2011).

A TAB. 7.3 enumera todos os 26 atributos citados, onde os 8 primeiros são derivados diretamente da imagem (WT) quantizada em níveis de cinza; os 10 atributos seguintes são derivados da RLM, e os 8 últimos, da GLCM.

Com a intenção de verificar se a tarefa de classificação proposta sofre pouca influência do ambiente, foram utilizados, nos experimentos iniciais, sinais de tiro de diferentes posições, incluindo diferentes países: Brasil e Finlândia (agradecimentos a TTY-säätiö, particularmente a Toni Mäkinen e Pasi Pertilä, por fornecer base de dados contendo sinais de tiro gravados no sul da Finlândia), com quatro tipos de munição, da seguinte maneira:

TAB. 7.3: Atributos de textura utilizados nesta tese

Atributos obtidos diretamente da imagem	Correlação
	Contraste
	Energia
	Entropia
	Momento diferencial inverso
	Parâmetro R
	<i>Skewness</i>
	<i>Kurtosis</i>
Atributos obtidos da RLM	<i>Short Run Emphasis</i>
	<i>Long Run Emphasis</i>
	<i>Gray-Level Nonuniformity</i>
	<i>Run Percentage</i>
	<i>Low Gray-Level Run Emphasis</i>
	<i>High Gray-Level Run Emphasis</i>
	<i>Short Run Low Gray-Level Emphasis</i>
	<i>Short Run High Gray-Level Emphasis</i>
	<i>Long Run Low Gray-Level Emphasis</i>
<i>Long Run High Gray-Level Emphasis</i>	
Atributos obtidos da GLCM	Correlação
	Contraste
	Energia
	Entropia
	Momento diferencial inverso
	Variância
	Dissimilaridade
	Homogeneidade

- 7.62 mm: 30 tiros de rifle brasileiro (2 posições) e 20 tiros de rifle finlandês (1 posição);
- 5.56 mm: 50 tiros de rifle brasileiro (3 posições);
- 9 mm: 25 tiros de pistola brasileira (1 posição) e 25 tiros de pistola finlandesa (2 posições);
- .50 BMG: 50 tiros de metralhadora brasileira (3 posições).

Note que a base contém 50 tiros de cada calibre, totalizando 200 tiros. Para os calibres 5.56 mm e .50 BMG, somente existiam tiros no Brasil (base do IME). Note também que há 3 posições diferentes para cada tipo de munição. No caso da base do IME, os tiros foram realizados no CAEx (Centro de Avaliações do Exército), inclusive em

dias diferentes, e as distâncias entre o atirador e o microfone são de, aproximadamente, 236, 540 e 860 m. No caso da base finlandesa, as distâncias são de 50 e 75 m.

Todos os 26 atributos mencionados anteriormente (8 oriundos do valor quantizado da intensidade de nível de cinza dos pixels, 10 da RLM e 8 da GLCM) servem como entrada de um classificador. Foram testados três classificadores: SVM, linear e redes neurais *feedforward* com duas camadas (foram testados de 10 a 40 neurônios na camada intermediária, em intervalos de 5, com o melhor desempenho obtido com 20 neurônios). O SVM apresentou melhores resultados e, como este classificador é binário (é capaz de separar somente duas classes), ele foi combinado em um multiclasse (HSU, 2002), utilizando a estrutura chamada *Adaptive Directed Acyclic Graph* (ADAG) (KIJSIRIKUL, 2002). A FIG. 7.6 mostra a ADAG utilizada neste trabalho: para classificar um sinal de tiro (teste), primeiramente aplica-se o SVM com 7.62 mm e 9 mm; o classificador indica qual munição mais se assemelha com o sinal de teste, e o resultado é denominado A1. Então, aplica-se o SVM com 5.56 mm e .50 BMG, e o resultado é chamado de A2. Após, aplica-se SVM com A1 e A2, e o resultado final é a classe (calibre) que a estrutura classifica o sinal de tiro testado.

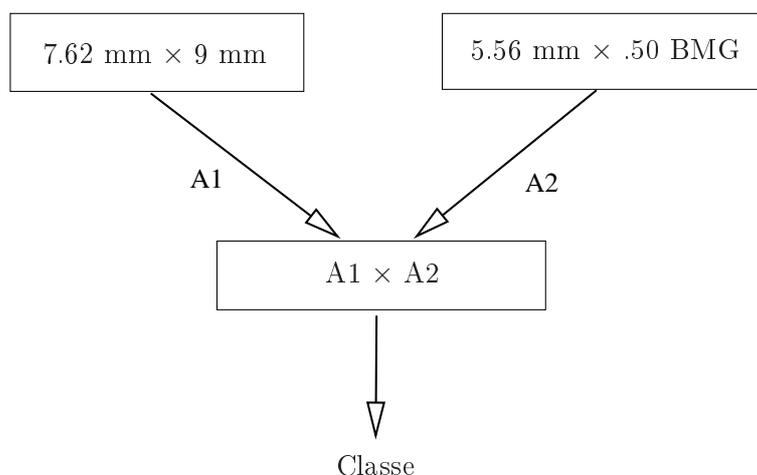


FIG. 7.6: ADAG utilizada nos experimentos.

Na FIG. 7.6, quando o sinal a ser testado não pertence às duas classes de um classificador binário, utilizam-se todos os 50 tiros de cada classe. Caso contrário, adota-se a validação cruzada *leave-one-out* (CAWLEY, 2006). Por exemplo, quando um tiro de 7.62 mm for o sinal de teste, utilizam-se 99 sinais quando é aplicado o SVM com 7.62 mm e 9 mm, enquanto que são utilizados 100 sinais quando é aplicado o SVM com 5.56 mm e .50 BMG.

## 7.2 RESULTADOS DA CLASSIFICAÇÃO DE MUNIÇÃO

Nesta seção, serão comparados os resultados da tarefa de classificação utilizando os atributos propostos em (KIKTOVA-VOZARIKOVA, 2015) (chamados aqui de características espectro-temporais), além dos 26 atributos da seção anterior, obtidos a partir do espectrograma e da WT quantizados em níveis de cinza. As TABs. 7.4, 7.5 e 7.6 mostram as matrizes de confusão para a classificação de munição realizada com as 200 componentes MB descritas na seção anterior, a partir das características espectro-temporais e dos atributos do espectrograma e da WT, respectivamente.

TAB. 7.4: Matriz de confusão para a classificação de quatro calibres utilizando características espectro-temporais

ARMA (CLASSE)	7.62 mm	5.56 mm	9 mm	.50BMG
7.62 mm	<b>37</b>	8	1	0
5.56 mm	8	<b>40</b>	0	0
9 mm	4	2	<b>49</b>	1
.50BMG	1	0	0	<b>49</b>

TAB. 7.5: Matriz de confusão para a classificação de quatro calibres utilizando espectrograma quantizado em níveis de cinza

ARMA (CLASSE)	7.62 mm	5.56 mm	9 mm	.50BMG
7.62 mm	<b>41</b>	7	1	5
5.56 mm	6	<b>40</b>	3	3
9 mm	1	2	<b>45</b>	1
.50BMG	2	1	1	<b>41</b>

TAB. 7.6: Matriz de confusão para a classificação de quatro calibres utilizando WT quantizada em níveis de cinza

ARMA (CLASSE)	7.62 mm	5.56 mm	9 mm	.50BMG
7.62 mm	<b>46</b>	5	1	1
5.56 mm	3	<b>44</b>	1	1
9 mm	1	1	<b>48</b>	0
.50BMG	0	0	0	<b>48</b>

Das TABs. 7.4, 7.5 e 7.6, podemos calcular as taxas de acerto: 87.5% com as características espectro-temporais, 83.5% quando utiliza-se o espectrograma quantizado em níveis de cinza, e 93.0% no caso de se utilizar WT, o que sugere melhor desempenho da WT quando comparada com os outros dois. Apesar das características espectro-temporais classificarem quase que 100% as munições 9 mm e .50, não conseguem identificar com igual desempenho os calibres 7.62 mm e 5.56 mm. Além disso, todos os atributos utilizados neste capítulo serão analisados em outro experimento, em que ficará mais evidente a superioridade de desempenho, na classificação de munição, dos 26 atributos oriundos da WT quantizada em níveis de cinza.

Neste capítulo, passou a ser abordada a segunda aplicação de interesse deste trabalho: a classificação de munição. Foram considerados sinais oriundos de quatro munições, gravados em diferentes ambientes. Foi visto que atributos de textura obtidos da WT quantizada em níveis de cinza, tratada como imagem, podem ser considerados bons separadores entre as classes aqui analisadas: 7.62 mm, 5.56 mm, 9 mm e .50 BMG, apresentando melhor desempenho que o espectrograma quantizado em níveis de cinza e as características espectro-temporais citadas em (KIKTOVA-VOZARIKOVA, 2015).

Em todos os 200 tiros analisados neste capítulo, a direção do cano da arma fazia um ângulo menor que  $45^\circ$  em relação à reta que une a ponta do cano ao microfone. Em (MAHER, 2010), é realizado um estudo da forma de onda da componente MB, considerando a orientação relativa do cano da arma e o microfone, de onde se conclui que o MB muda significativamente quando a arma está apontada na direção contrária à posição do microfone. Assim, é provável que a taxa de acerto da classificação de munição dependerá da direção do cano da arma em relação ao microfone. O próximo capítulo analisa esse assunto.

## 8 A INFLUÊNCIA DA DIREÇÃO DO CANO DA ARMA EM RELAÇÃO AO MICROFONE

Este capítulo explora os aspectos direcionais das gravações da componente MB. Na Seção 8.1, serão mostrados os experimentos realizados na tentativa de ilustrar que existem diferenças no domínio do tempo e, conseqüentemente, no domínio da WT, se o sinal de tiro que chega ao microfone provém de uma arma cujo cano estava em sua direção ou em direção oposta. A análise dos resultados da classificação de munição será feita em duas etapas: a primeira, abordada na Seção 8.2, faz uso de uma base mais antiga (de 2010); a segunda, analisada na Seção 8.3, utiliza uma base mais recente (de 2015), separadamente e em conjunto com a base de 2010. Nas duas bases, que serão descritas nas Seções 8.2 e 8.3, há tiros de várias direções entre o cano da arma e o microfone. Neste capítulo, será considerado que o termo azimute significa o ângulo entre a direção do cano da arma e a reta que une a ponta do cano ao microfone. A WT utilizada neste capítulo também é a *biorthogonal*1.5.

### 8.1 A CARACTERÍSTICA DIRECIONAL DO MUZZLE BLAST

O MB possui características direcionais que dependem do armamento e da munição utilizados (MAHER, 2010), ou seja, o modo como a orientação do cano da arma afeta o sinal recebido depende de especificações do armamento e do projétil, indicando uma possível degradação da taxa de acerto na tarefa de classificação de munição quando sinais de diferentes azimutes são considerados. A FIG. 8.1 mostra quatro componentes MB de um mesmo armamento (ruger sp101 com munição .38 Special), do mesmo disparo, mas com azimutes diferentes ( $16^\circ$ ,  $98^\circ$ ,  $130^\circ$  e  $164^\circ$ ).

Da FIG. 8.1, pode-se observar que as componentes MB no domínio do tempo são ligeiramente diferentes quando o azimute muda de  $16^\circ$  para  $98^\circ$ ; porém, para os azimutes de  $130^\circ$  e  $164^\circ$ , as diferenças são maiores quando comparadas com os outros dois. Como os coeficientes da WT indicam o quanto a wavelet se assemelha com uma seção do sinal, ao mudar a forma de onda no domínio do tempo (como no caso da FIG. 8.1), os valores dos coeficientes da WT também serão diferentes, ocasionando possíveis erros na classificação de munição. A FIG. 8.2 ilustra as WT das quatro componentes MB da FIG. 8.1.

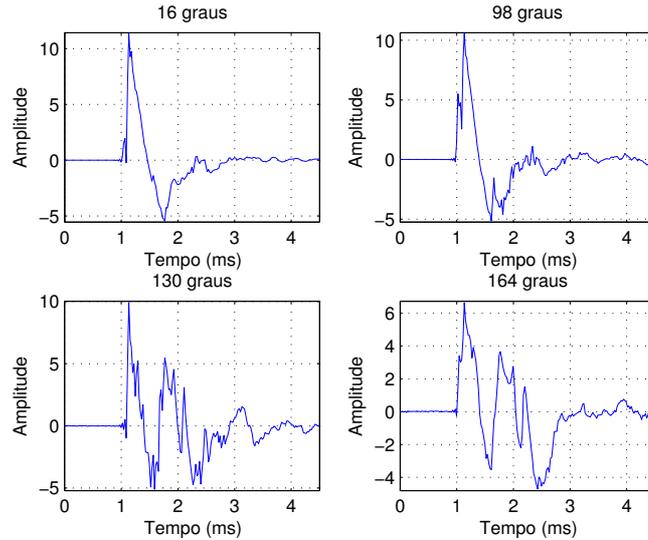


FIG. 8.1: Componentes MB de quatro azimutes diferentes da mesma arma e mesmo disparo.

Da FIG. 8.2, observamos que, para os azimutes menores que  $100^\circ$ , as WT do MB são pouco diferentes. Por outro lado, para os dois azimutes mais próximos de  $180^\circ$ , as diferenças para os azimutes mais baixos são mais expressivas. Desta análise, é razoável supor que as taxas de acerto da classificação de munição terão resultados diferentes se forem considerados somente sinais com azimutes até próximos de  $100^\circ$  ou se forem incluídos sinais com azimutes próximos de  $180^\circ$ .

## 8.2 ANÁLISE DA PRIMEIRA BASE DE DADOS

Nesta primeira análise a respeito da influência do azimuth na classificação de munição, foi utilizada uma base (Base 1) com sinais de tiro gravados na cidade de Bozeman, estado de Montana, nos Estados Unidos da América (agradecimentos ao Prof. Robert C. Maher, da Universidade Estadual de Montana, por disponibilizar os sinais dessa base), de oito armamentos/munições diferentes (MAHER, 2010), dos quais quatro são rifles ou espingardas, denominados armas de cano longo (308 Winchester, 223 Remington, 12 gauge shotgun, 22 long rifle) e quatro são pistolas ou revólveres, denominados armas de cano curto ou *handguns* (45 ACP, 357 Magnum,  $9 \times 19$  mm, 38 Special handgun). Serão utilizadas as seguintes nomenclaturas para identificar as oito munições citadas: 308r, 223r, 12s, 22r, 45h, 357h, 9h, 38h, onde “r” significa rifle, “h”, *handgun*, e “s”, espingarda (do inglês *shotgun*). Os microfones foram colocados a um raio de 3 m da posição de tiro, englobando uma ampla faixa de azimutes, com  $0^\circ$  sendo à frente do atirador, com uma

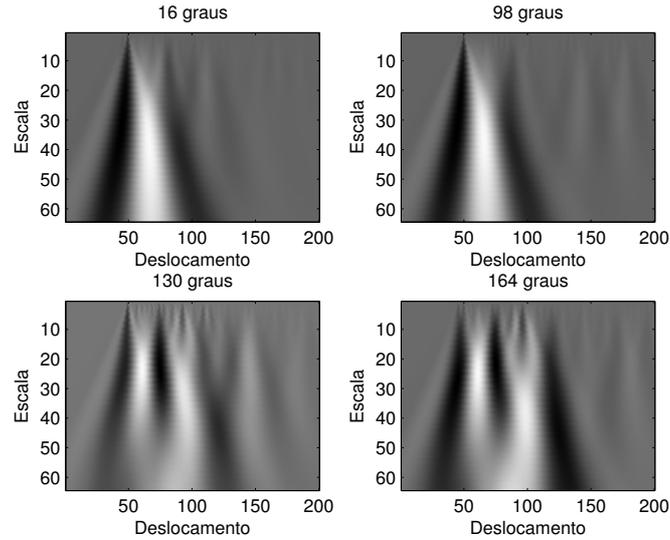


FIG. 8.2: Transformadas wavelet do MB de quatro azimutes diferentes da mesma arma e mesmo disparo.

elevação de 3 m, para que a reflexão no solo fosse atrasada de aproximadamente 10 ms, se compararmos com a onda direta da origem do tiro ao microfone.

Para cada munição, há gravações com 10 diferentes azimutes ( $0^\circ$ ,  $9^\circ$ ,  $40^\circ$ ,  $50^\circ$ ,  $85^\circ$ ,  $95^\circ$ ,  $130^\circ$ ,  $140^\circ$ ,  $171^\circ$  e  $180^\circ$ ), sendo 1 tiro de cada arma para cada azimuth, totalizando 80 tiros (a cada tiro, o microfone teve sua posição alterada). Para evitar que se atrasasse no microfone, o azimuth de zero grau foi ligeiramente deslocado.

### 8.2.1 RESULTADOS DE CLASSIFICAÇÃO DE MUNIÇÃO COM AS OITO ARMAS

A fim de verificar a influência de azimutes maiores que  $100^\circ$  na classificação de munição, considerando as oito armas citadas e utilizando os 26 atributos oriundos da WT quantizada em níveis de cinza, foram realizados testes da seguinte forma:

- Primeiro teste: azimutes até  $95^\circ$ , com 6 tiros de cada arma, totalizando 48 tiros;
- Segundo teste: azimutes até  $130^\circ$ , com 7 tiros de cada arma, totalizando 56 tiros;
- Terceiro teste: azimutes até  $140^\circ$ , com 8 tiros de cada arma, totalizando 64 tiros;
- Quarto teste: azimutes até  $171^\circ$ , com 9 tiros de cada arma, totalizando 72 tiros;
- Quinto teste: azimutes até  $180^\circ$ , com 10 tiros de cada arma, totalizando 80 tiros.

Assim como no Capítulo 7, foi utilizado o classificador SVM e a estrutura ADAG com procedimentos semelhantes. A FIG. 8.3 mostra a ADAG implementada contendo as oito

armas em estudo neste capítulo. Para classificar um sinal de tiro (teste), primeiramente aplica-se o SVM quatro vezes: entre 9h e 22r, 38h e 12s, 357h e 223r, 45h e 308r, e os resultados são denominados A1, A2, A3 e A4, respectivamente. Depois, no segundo nível, aplica-se SVM duas vezes: entre A1 e A2, A3 e A4, e os resultados são A5 e A6. E, finalmente, aplica-se SVM entre A5 e A6, e o resultado é a classe (calibre) que a estrutura classifica o sinal de tiro testado. Note que, no primeiro nível, foram escolhidas combinações entre armas de cano curto e de cano longo, visto que suas diferenças na componente MB são presumidamente maiores que entre duas armas de cano curto ou entre duas armas de cano longo.

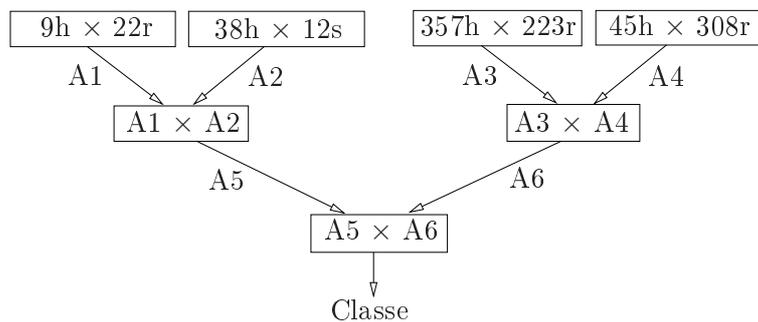


FIG. 8.3: ADAG para as oito armas/munições.

Na FIG. 8.3, quando o sinal a ser testado não pertence às duas classes de um classificador binário, utilizam-se todos os tiros de cada classe. Caso contrário, adota-se a validação cruzada *leave-one-out*, que foi utilizada em todas as classificações desta tese.

A TAB. 8.1 mostra a matriz de confusão para a classificação de munição relativa ao primeiro teste (azimutes até  $95^\circ$ ).

TAB. 8.1: Matriz de confusão com azimutes até  $95^\circ$

ARMA (CLASSE)	9h	22r	12s	38h	223r	357h	308r	45h
9h	<b>6</b>	0	0	0	0	0	0	0
22r	0	<b>6</b>	0	0	0	0	0	0
12s	0	0	<b>6</b>	0	0	0	0	0
38h	0	0	0	<b>5</b>	0	0	0	0
223r	0	0	0	0	<b>5</b>	0	0	0
357h	0	0	0	0	0	<b>6</b>	0	0
308r	0	0	0	0	1	0	<b>6</b>	0
45h	0	0	0	1	0	0	0	<b>6</b>

Da TAB. 8.1, observa-se que, dos 48 sinais, apenas 2 tiveram erro na classificação,

sendo um erro entre as armas de cano curto 45h e 38h e outro entre os rifles 223r e 308r, gerando uma taxa de acerto igual a **95,83%**.

A TAB. 8.2 mostra a matriz de confusão para a classificação de munição relativa ao segundo teste (azimutes até 130°).

TAB. 8.2: Matriz de confusão com azimutes até 130°

ARMA (CLASSE)	9h	22r	12s	38h	223r	357h	308r	45h
9h	<b>6</b>	0	0	0	0	0	0	0
22r	0	<b>7</b>	0	0	0	0	0	0
12s	1	0	<b>7</b>	2	0	0	0	0
38h	0	0	0	<b>5</b>	0	0	0	0
223r	0	0	0	0	<b>6</b>	0	1	0
357h	0	0	0	0	1	<b>6</b>	0	2
308r	0	0	0	0	0	0	<b>6</b>	1
45h	0	0	0	0	0	1	0	<b>4</b>

Da TAB. 8.2, pode-se observar que há 9 classificações erradas entre 56 sinais, fazendo com que a taxa de acerto diminua para **83,93%**. Dos 9 erros, 3 ocorrem entre armas de cano curto, 1 entre rifles, e aparecem 5 erros entre armas de cano curto e de cano longo, indicando que, para azimutes maiores que 100°, pode haver confusão não só entre munições de armas de mesmo porte, mas também entre armas de calibres bem diferentes.

A TAB. 8.3 mostra a matriz de confusão para a classificação de munição relativa ao terceiro teste (azimutes até 140°).

TAB. 8.3: Matriz de confusão com azimutes até 140°

ARMA (CLASSE)	9h	22r	12s	38h	223r	357h	308r	45h
9h	<b>8</b>	0	0	0	1	0	0	1
22r	0	<b>8</b>	0	0	0	0	0	0
12s	0	0	<b>8</b>	1	0	1	0	1
38h	0	0	0	<b>7</b>	0	0	0	0
223r	0	0	0	0	<b>4</b>	0	1	0
357h	0	0	0	0	1	<b>7</b>	0	1
308r	0	0	0	0	2	0	<b>7</b>	0
45h	0	0	0	0	0	0	0	<b>5</b>

Da TAB. 8.3, pode-se observar que há 10 classificações erradas entre 64 sinais, gerando uma taxa de acerto de **84,38%**. Novamente, há vários erros entre armas de

cano curto e de cano longo, confirmando a confusão entre munições de armas de calibres bem diferentes quando são considerados azimutes grandes.

Das TABs. 8.1, 8.2 e 8.3, nota-se que há uma diminuição considerável da taxa de acerto quando sinais com azimute de  $130^\circ$  (TAB 8.2) são adicionados aos sinais com azimutes de até  $95^\circ$  (TAB 8.1); porém, há somente um ligeiro aumento dessa taxa quando sinais com azimutes de  $140^\circ$  (TAB 8.3) são adicionados a esse conjunto. Uma explicação para isso é a diferença de azimute entre  $130^\circ$  e  $95^\circ$  e entre  $140^\circ$  e  $130^\circ$ . Como esta última diferença é pequena, um sinal com azimute de  $140^\circ$  tende a ser parecido com um sinal com azimute de  $130^\circ$ , o que não deve ocorrer entre sinais com azimutes de  $130^\circ$  e  $95^\circ$ . Assim, a inclusão de sinais parecidos no treinamento (da TAB 8.2 para a TAB 8.3) deve facilitar a tarefa de classificação.

A TAB. 8.4 mostra a matriz de confusão para a classificação de munição relativa ao quarto teste (azimutes até  $171^\circ$ ).

TAB. 8.4: Matriz de confusão com azimutes até  $171^\circ$

ARMA (CLASSE)	9p	22r	12r	38p	223r	357p	308r	45p
9p	8	1	1	0	1	0	0	0
22r	0	8	0	0	0	0	0	0
12r	1	0	5	1	0	2	2	0
38p	0	0	0	7	0	0	0	2
223r	0	0	0	0	3	1	0	0
357p	0	0	1	0	1	6	0	2
308r	0	0	2	0	4	0	7	0
45p	0	0	0	1	0	0	0	5

Da TAB. 8.4, pode-se observar que existem 23 classificações erradas entre 72 sinais, fazendo com que a taxa de acerto diminua para **68,06%**. Neste caso, há 10 erros entre armas de cano curto e de cano longo. Note que a taxa de acerto diminuiu consideravelmente da TAB. 8.3 para a TAB. 8.4, visto que foi incluído um azimute próximo a  $180^\circ$  no conjunto cujo maior azimute era de  $140^\circ$ .

A TAB. 8.5 mostra a matriz de confusão para a classificação de munição relativa ao quinto teste (azimutes até  $180^\circ$ ).

Da TAB. 8.5, pode-se observar que há 24 classificações erradas entre 80 sinais, gerando uma taxa de acerto de **70%**, com 8 erros entre armas de cano curto e de cano longo.

Note que, da mesma forma entre as TABs. 8.2 e 8.3, a taxa de acerto sofreu um

TAB. 8.5: Matriz de confusão com azimutes até  $180^\circ$

ARMA (CLASSE)	9h	22r	12s	38h	223r	357h	308r	45h
9h	<b>9</b>	1	1	0	0	0	0	0
22r	0	<b>9</b>	0	0	0	0	0	0
12s	0	0	<b>7</b>	1	1	1	0	0
38h	0	0	0	<b>7</b>	0	1	0	0
223r	0	0	0	0	<b>4</b>	0	3	0
357h	0	0	1	0	0	<b>6</b>	0	2
308r	1	0	1	0	5	1	<b>7</b>	1
45h	0	0	0	2	0	1	0	<b>7</b>

pequeno aumento da TAB. 8.4 para a TAB. 8.5, pois o azimute incluído ( $180^\circ$ ) na última tabela é próximo do maior azimute da tabela anterior. Com isso, maior é a probabilidade de haver sinais parecidos no treinamento para cada arma, o que explica a pequena melhora na classificação de munição.

Foi visto que a classificação utilizando WT quantizada em níveis de cinza se deteriorou com a inserção de azimutes próximos de  $180^\circ$ . Da mesma forma que no capítulo anterior, também foram testados o espectrograma quantizado em níveis de cinza e as características espectro-temporais, a fim de verificar seus desempenhos perante azimutes elevados. A TAB. 8.6 compara as taxas de acerto quando se utilizam WT e espectrograma, ambos quantizados em níveis de cinza, e as características espectro-temporais (CET) propostas em (KIKTOVA-VOZARIKOVA, 2015).

TAB. 8.6: Taxas de acerto utilizando CET, espectrograma e WT

Azimute	CET	Espectrograma	WT
Até $95^\circ$	72,92%	52,08%	<b>95,83%</b>
Até $130^\circ$	58,93%	53,57%	<b>83,93%</b>
Até $140^\circ$	60,94%	56,25%	<b>84,38%</b>
Até $171^\circ$	52,78%	51,39%	<b>68,06%</b>
Até $180^\circ$	58,75%	42,50%	<b>70%</b>

Da TAB. 8.6, observa-se que atributos extraídos da WT quantizada em níveis de cinza continuam a apresentar melhores resultados na tarefa de classificação de munição, em comparação com os outros dois, desta vez com uma diferença de desempenho superior à do capítulo anterior.

### 8.3 INCLUSÃO DA SEGUNDA BASE DE DADOS

A gravação da segunda base (Base 2) foi realizada em 2015 no estado de Montana, nos Estados Unidos da América, em campo aberto, e contou com a participação deste autor. Neste caso, 12 microfones foram posicionados simultaneamente a um raio de 3 m do atirador e igualmente espaçados ao redor de  $180^\circ$  de azimute, com  $180^\circ$  sendo atrás do atirador. A TAB. 8.7 enumera os azimutes aproximados dos microfones, denominados “M1” a “M12”.

TAB. 8.7: Azimutes aproximados dos microfones para a Base 2

Microfone	Azimute
M1	$0^\circ$
M2	$16^\circ$
M3	$33^\circ$
M4	$49^\circ$
M5	$65^\circ$
M6	$82^\circ$
M7	$98^\circ$
M8	$115^\circ$
M9	$131^\circ$
M10	$147^\circ$
M11	$164^\circ$
M12	$180^\circ$

Para as gravações, foram utilizados conjuntos de microfones G.R.A.S. tipo 46DP, consistindo de diafragma tipo 40DP de  $1/8''$ , pré-amplificadores tipo 26TC de  $1/4''$  e módulos de potência tipo 12AA e 12 AG. Os microfones operam na faixa de frequência de 6,5 Hz a 140 kHz, com sensibilidade de 0,7 mV/Pa.

O sistema de gravação de áudio consiste de um conjunto NI PXIe-1071, da *National Instruments*, composto por um processador NI PXIe-8840 e por um módulo de aquisição de dados (A/D) NI PXIe-6358. O conversor analógico-digital possui 16 entradas simultâneas, com resolução de 16 bits. A frequência de amostragem foi de 500.000 amostras por segundo por canal.

Durante os testes, o armamento foi orientado horizontalmente e afastado aproximadamente 3 m do solo, da mesma forma que a Base 1. Todos os microfones foram calibrados, de modo a evitar a clipagem dos sinais. Para evitar que se atirasse no microfone, o azimute de zero grau foi ligeiramente deslocado.

Depois que o aparato foi colocado na posição correta e o atirador posicionou-se no

topo de uma escada, cada tiro foi captado e gravado simultaneamente pelos 12 microfones. Este procedimento foi repetido para cada tiro. A TAB. 8.8 apresenta os armamentos utilizados e suas respectivas munições, bem como as abreviações adotadas e o número de tiros executados.

TAB. 8.8: Armamentos, munições e número de tiros da Base 2

Armamento	Munição	Abrev.	Tiros
Colt 1911	45 ACP	45h	10
Glock 19	9 × 19	9h	10
Sig239	357 Sig	357h1	10
Ruger SP101	38 Special	38h	9
Ruger SP101	357 Magnum	357h2	10
CZ 452	22 LR	22r	11
Surgeon / AI	308 Winchester	308r	10
Stag Arms AR 15	5.56 NATO	15r	10
Remington 870	12 gauge	12s	3

Da TAB. 8.8, note que o Ruger SP101 foi utilizado com dois tipos de munição (38 Special e 357 Magnum), e que a mesma munição (357) foi utilizada com duas armas diferentes (Ruger SP101 e Sig239). Tal procedimento é útil para mostrar o bom desempenho do classificador com relação à classificação de munição e de armamento.

Quando da análise da Base 1 na seção anterior, foi constatado que o método de classificação baseado nos atributos de textura obtidos da WT da componente MB de um sinal de tiro teve desempenho melhor que outras técnicas, com taxa de acerto maior que 95% ao se considerar azimutes menores que 100° (aqui denominados azimutes pequenos). Foram utilizados, novamente, os 26 atributos de textura, com o classificador SVM com kernel linear, com a mesma estrutura ADAG, em que, no primeiro nível, somente há combinações entre armas de cano curto e de cano longo.

Para a Base 2, foram adotadas várias árvores de classificação (conjuntos de armamentos e munições em uma árvore de classificadores binários, aqui chamados de classificador multiclasse). Quando o sinal a ser testado não pertence às duas classes de um classificador binário, utilizam-se todos os tiros de cada classe. Caso contrário, adota-se a validação cruzada *leave-one-out*.

A FIG. 8.4 mostra a primeira árvore de classificação para a Base 2, em que nenhuma arma ou munição é repetida, isto é, o armamento Ruger SP101 é considerado apenas com a munição 38 Special, e a munição 357 só é utilizada com o armamento Sig239. Para esta análise, temos então 73 sinais: 11 de 22r, 9 de 38h, 3 de 12s e 10 de cada um dos

diferentes armamentos restantes, ou seja, excluindo os tiros denominados 357h2.

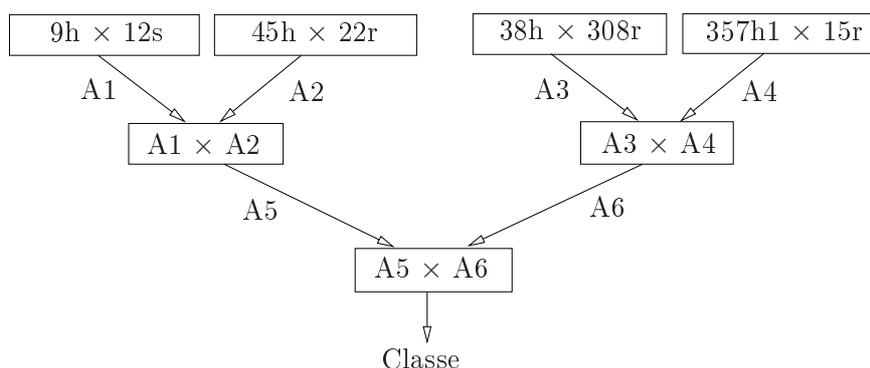


FIG. 8.4: Primeira árvore de classificação para a Base 2. Nenhuma arma ou munição é repetida.

A TAB. 8.9 apresenta as taxas de acerto dos 73 tiros para cada um dos 12 azimutes da TAB. 8.7, no caso da primeira árvore de classificação.

TAB. 8.9: Taxas de acerto para a primeira árvore de classificação para cada azimute

Microfone	Azimute	Taxa de acerto
M1	0°	100%
M2	16°	98,63%
M3	33°	97,26%
M4	49°	98,63%
M5	65°	98,63%
M6	82°	94,52%
M7	98°	97,26%
M8	115°	87,67%
M9	130°	95,89%
M10	147°	95,89%
M11	164°	86,30%
M12	180°	94,52%

Da TAB. 8.9, pode-se observar que o azimute próximo de 0° é o melhor caso e que as taxas de acerto são maiores que 94% quando consideramos azimutes pequenos (até M7). Além disso, a menor taxa de acerto é 86.30%.

Em seguida, a fim de incluir uma variabilidade de azimutes na mesma análise, foram criados 10 conjuntos de microfones, conforme ilustrado na TAB. 8.10. Os primeiros 6 conjuntos compreendem apenas pequenos azimutes, e os outros conjuntos incluem azimutes maiores que 100° (aqui chamados grandes azimutes). Em particular, o nono conjunto compreende apenas grandes azimutes, e o último conjunto é composto por quase todas

as orientações. Para cada um dos 73 tiros incluídos na primeira árvore de classificação, apenas um microfone foi considerado, de modo que não mais que uma gravação do mesmo tiro faça parte da classificação apresentada na FIG. 8.4, assegurando, assim, que qualquer sinal de teste não seja muito semelhante a qualquer sinal que pertença ao conjunto de treinamento. Também houve cuidado no sentido de que, para cada conjunto, considerou-se que o número de sinais de tiro gravados por um microfone seja igual ao número de sinais dos outros microfones do conjunto ou, na impossibilidade dessa igualdade, que a quantidade de tiros de cada microfone seja diferente de apenas um tiro por armamento. Por exemplo, no primeiro conjunto, para o “45h”, foram utilizados 5 tiros gravados de M1 e 5 tiros de M2. Por outro lado, para o “38h”, foram utilizados 5 tiros de M1 e 4 tiros de M2, pois havia apenas 9 tiros com esta arma+munição. A TAB. 8.10 também mostra as taxas de acerto para cada conjunto de microfones para a primeira árvore de classificação.

TAB. 8.10: Conjuntos de azimutes e respectivas taxas de acerto para a primeira árvore de classificação

<b>Ordem do cj</b>	<b>Mics envolvidos</b>	<b>Taxa acerto</b>
Cj 01	1, 2	98,63%
Cj 02	1, 2, 3	100%
Cj 03	1, 2, 3, 4	97,26%
Cj 04	1, 2, 3, 4, 5	95,89%
Cj 05	1, 2, 3, 4, 5, 6	93,15%
Cj 06	2, 3, 4, 5, 6, 7	94,52%
Cj 07	6, 7, 8, 9, 10	90,41%
Cj 08	7, 8, 9, 10, 11	78,08%
Cj 09	8, 9, 10, 11, 12	76,71%
Cj 10	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	78,08%

Da TAB. 8.10, conclui-se que quando consideram-se apenas azimutes pequenos, a taxa de acerto é maior que 93%. Por outro lado, ao se considerar (ou incluir) azimutes grandes, a taxa de acerto diminui, e seu menor valor é 76,71%, que corresponde ao caso em que há somente azimutes grandes.

O mesmo algoritmo de classificação foi testado em quatro outras árvores de classificação com a Base 2, e os resultados foram analisados da mesma forma que a TAB. 8.9 (por azimuth) e a TAB. 8.10 (por conjunto de microfones). As FIGs. 8.5, 8.6, 8.7 e 8.8 apresentam, respectivamente, as segunda, terceira, quarta e quinta árvores de classificação.

Na FIG. 8.5, observa-se que a única diferença para a FIG. 8.4 é a substituição de

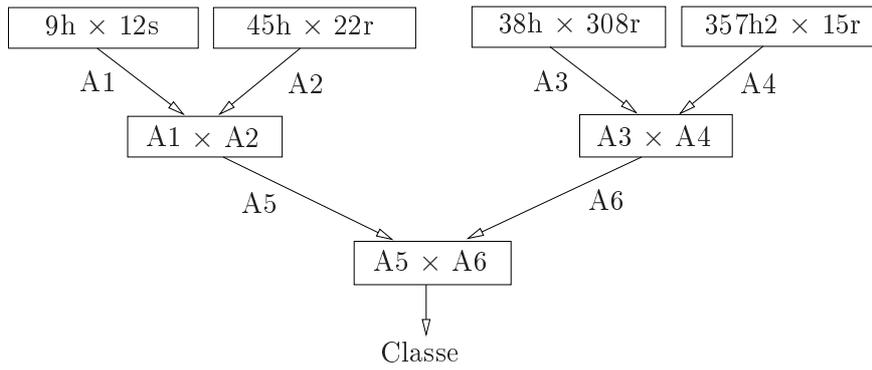


FIG. 8.5: Segunda árvore de classificação. Uma mesma arma (Ruger SP101) utilizando dois tipos diferentes de munição, que foram separadas em diferentes classes.

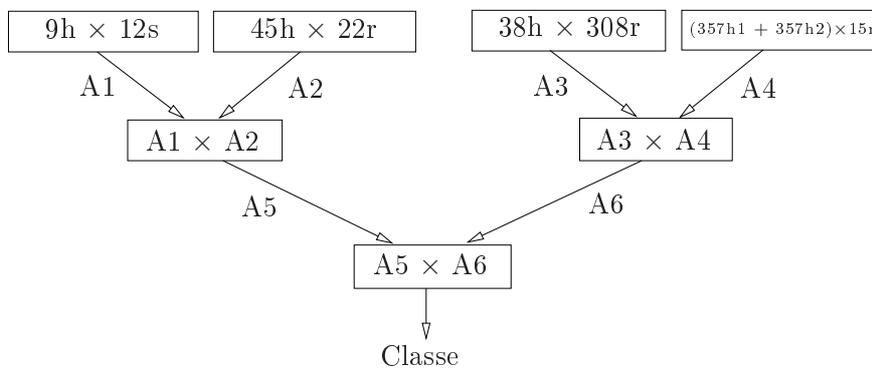


FIG. 8.6: Terceira árvore de classificação. Uma mesma munição (357h1 e 357h2) utilizada com dois armamentos diferentes, compreendendo uma única classe.

357h1 por 357h2 no primeiro estágio do classificador; porém, a mesma arma (Ruger SP101) utilizou munições diferentes (357 Sig e 38 Special) em duas classes distintas. Esta árvore de classificação testa a robustez do classificador quando a mesma arma é utilizada com dois tipos de munição representando classes diferentes. O número total de tiros também é igual a 73 (11 de 22r, 9 de 38h, 3 de 12s e 10 de cada um dos pares de arma+munição restantes).

Na FIG. 8.6, o mesmo tipo de munição (357h1 e 357h2) usada com dois diferentes armamentos foi considerado como uma única classe. A árvore desta figura funciona como uma classificação de munição, ao juntar armamentos diferentes com mesma munição na mesma classe. O número total de tiros agora é igual a 83 (11 de 22r, 9 de 38h, 3 de 12s e 10 de cada um dos pares de arma+munição restantes). Note que 10 tiros de 357h1 e 10 tiros de 357h2 constituem uma única classe “357”, que contém 20 sinais.

Na FIG. 8.7, diferentes munições (357 Magnum e 38 Special) deflagradas da mesma

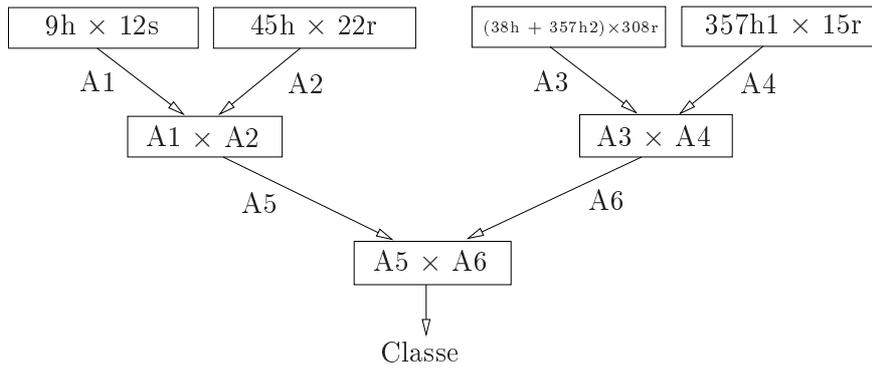


FIG. 8.7: Quarta árvore de classificação. Diferentes munições (357 Magnum e 38 Special) utilizadas de uma mesma arma (Ruger SP101), compreendendo uma única classe.

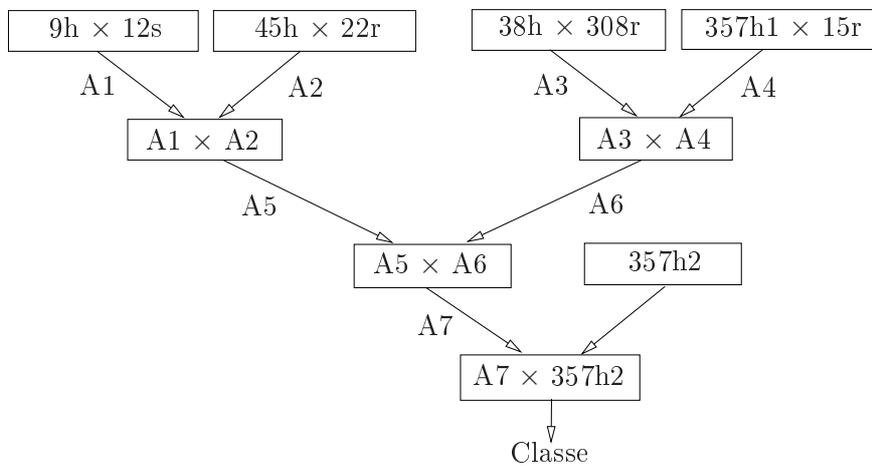


FIG. 8.8: Quinta árvore de classificação. O 38h, o 357h1 e o 357h2 foram separados em diferentes classes.

arma (Ruger SP101) foram juntadas em uma única classe. A árvore desta figura funciona como uma classificação de armamento. O número total de tiros também é igual a 83 (11 de 22r, 9 de 38h, 3 de 12s e 10 de cada um dos pares de arma+munição restantes). Note que 10 tiros de 357h2 e 9 tiros de 38h constituem uma única classe “Ruger SP101”, que contém 19 sinais.

Na FIG. 8.8, o 38h, o 357h1 e o 357h2 foram separados em três diferentes classes. Dessa forma, pode-se verificar o desempenho do algoritmo com relação à classificação de armamento e munição em uma única árvore de classificação. Neste caso, o número total de tiros também é igual a 83 (11 de 22r, 9 de 38h, 3 de 12s e 10 de cada um dos pares de arma+munição restantes), mas têm-se 9 classes.

De modo semelhante à TAB. 8.9, a TAB. 8.11 apresenta as taxas de acerto para a segunda até a quinta árvore de classificação para cada azimute. Da TAB. 8.11, observa-se

que a taxa de acerto é sempre maior que 95% quando são considerados somente azimutes pequenos. Para azimutes grandes, a taxa de acerto usualmente decresce, com menor valor igual a 84.93%.

TAB. 8.11: Taxas de acerto para a segunda até a quinta árvore de classificação para cada azimute

Mic	Árvore de classificação			
	2 <sup>a</sup>	3 <sup>a</sup>	4 <sup>a</sup>	5 <sup>a</sup>
M1	95,89%	97,59%	97,59%	96,39%
M2	98,63%	98,80%	98,80%	98,80%
M3	97,26%	96,39%	96,39%	97,59%
M4	98,63%	98,80%	98,80%	98,80%
M5	98,63%	98,80%	98,80%	98,80%
M6	98,63%	96,39%	95,18%	95,18%
M7	95,89%	96,39%	97,59%	96,39%
M8	86,30%	87,95%	89,16%	89,16%
M9	94,52%	96,39%	96,39%	95,18%
M10	91,78%	87,95%	96,39%	90,36%
M11	90,41%	87,95%	86,75%	86,75%
M12	84,93%	85,54%	93,98%	85,54%

A TAB. 8.12 apresenta as taxas de acerto para a segunda até a quinta árvore de classificação para cada conjunto de microfones, semelhante à TAB. 8.10.

TAB. 8.12: Taxas de acerto para a segunda até a quinta árvore de classificação para cada conjunto de microfones

Conjunto	Árvore de classificação			
	2 <sup>a</sup>	3 <sup>a</sup>	4 <sup>a</sup>	5 <sup>a</sup>
01	97,26%	97,59%	95,18%	95,18%
02	95,89%	98,80%	90,36%	91,57%
03	95,89%	98,80%	93,98%	91,57%
04	97,26%	97,59%	89,16%	92,77%
05	91,78%	92,77%	90,36%	91,57%
06	94,52%	95,18%	89,16%	93,98%
07	86,30%	73,49%	84,34%	83,13%
08	79,45%	69,88%	79,52%	75,90%
09	78,08%	71,08%	79,52%	71,08%
10	75,34%	72,29%	71,08%	68,67%

Da TAB. 8.12, pode-se verificar que, quando considera-se um conjunto de microfones somente com azimutes pequenos, a taxa de acerto é maior que 89%. Com azimutes

grandes, essa taxa decresce, como esperado.

Das TABs. 8.9, 8.10, 8.11 e 8.12, pode-se concluir que o algoritmo pode ser utilizado indistintamente para classificação de munição e de armamento, particularmente para azimutes pequenos.

A fim de checar se o método de classificação apresentado é robusto a mudanças ambientais, pelo menos para azimutes pequenos, foram juntados sinais das Bases 1 e 2, utilizando a primeira árvore de classificação descrita na FIG. 8.4. A TAB. 8.13 mostra o número de sinais de tiro de cada base por arma+munição (A+M) utilizado nesta primeira árvore.

TAB. 8.13: Número de sinais de tiro de cada base

A+M	Número de sinais		
	Base 1	Base 2	Bases 1 e 2
9h	6	10	16
45h	6	10	16
38h	6	9	15
357h1	6	10	16
22r	6	11	17
308r	6	10	16
12s	6	3	9
15r	0	10	10
TOTAL	42	73	115

Da TAB. 8.13, note que, apesar de não haver tiros de 15r na Base 1, a quantidade total de sinais desta base corresponde a mais de  $\frac{1}{3}$  de todos os sinais utilizados, não sendo uma porcentagem desprezível do total. Observe também que a quantidade de tiros de cada arma da Base 1, quando não nulo, é igual a 6, correspondente aos 6 azimutes pequenos daquela base.

Como a Base 1 contém somente um tiro por azimute por armamento, não é possível analisar cada azimute separadamente (conforme realizado na TAB. 8.9). Assim, a TAB. 8.14 apresenta como as duas bases foram combinadas em termos de azimute e mostra também as taxas de acerto. Note que, em todos os casos (linhas da TAB. 8.14), foi considerada a faixa de azimutes de  $0^\circ$  a  $95^\circ$  para a Base 1.

Da TAB. 8.14, observa-se que a taxa de classificação permanece maior que 90% quando somente pequenos azimutes são considerados, mesmo juntando sinais de bases diferentes, sugerindo que, para pequenas distâncias e campo aberto, o ambiente não influencia tanto os resultados.

TAB. 8.14: Azimutes e taxas de acerto para a primeira árvore de classificação com duas bases

Caso	Azimutes		Taxa de acerto
	Base 1	Base 2	
1	$0^\circ - 95^\circ$	$0^\circ$	94,78%
2	$0^\circ - 95^\circ$	$16^\circ$	92,17%
3	$0^\circ - 95^\circ$	$33^\circ$	93,91%
4	$0^\circ - 95^\circ$	$49^\circ$	98,26%
5	$0^\circ - 95^\circ$	$65^\circ$	96,52%
6	$0^\circ - 95^\circ$	$82^\circ$	95,65%
7	$0^\circ - 95^\circ$	$98^\circ$	91,31%

A fim de desenvolver testes similares aos apresentados na TAB. 8.10, em que foram criados conjuntos de microfones para a Base 2, as duas bases foram combinadas, conforme ilustrado na TAB. 8.15. As taxas de acerto também são apresentadas nesta tabela.

TAB. 8.15: Taxas de acerto para a primeira árvore de classificação para conjuntos de microfones com duas bases

Conjunto	Azimutes		Taxa de acerto
	Base 1	Base 2	
1	$0^\circ - 95^\circ$	$0^\circ - 16^\circ$	93,04%
2	$0^\circ - 95^\circ$	$0^\circ - 33^\circ$	90,43%
3	$0^\circ - 95^\circ$	$0^\circ - 49^\circ$	93,91%
4	$0^\circ - 95^\circ$	$0^\circ - 65^\circ$	91,30%
5	$0^\circ - 95^\circ$	$0^\circ - 82^\circ$	90,43%
6	$0^\circ - 95^\circ$	$16^\circ - 98^\circ$	89,57%

Da TAB. 8.15, observa-se que a taxa de acerto é maior que 89%, mesmo com a combinação de sinais com azimutes pequenos.

Analisando os resultados obtidos até agora neste capítulo, pode-se afirmar que o algoritmo de classificação apresentado possui bom desempenho se os azimutes forem menores que  $100^\circ$ . Por outro lado, se incluirmos ou só considerarmos azimutes grandes, a taxa de acerto se reduz. Então, se fôssemos capazes de confirmar que um tiro fora originado de um azimute pequeno, o resultado de classificação seria mais confiável. Desta seção, conclui-se também que os atributos de textura utilizados são capazes de classificar tanto a munição quanto o armamento.

## 8.4 ANALISANDO SINAIS CORROMPIDOS

Até a seção anterior, foi analisado o desempenho do classificador considerando sinais não corrompidos gravados por microfones profissionais, que causam distorções imperceptíveis. No entanto, em situações reais, o gravador pode degradar o sinal de interesse de várias maneiras (por exemplo, distorções em amplitude e frequência e clipagem).

Como a Transformada Wavelet normaliza o sinal, é esperado que, caso somente a amplitude do sinal de teste seja alterada, os resultados de classificação sejam os mesmos que os apresentados na seção anterior. Foram realizados experimentos considerando a primeira árvore de classificação da FIG. 8.4, mudando a amplitude de cada sinal de teste aleatoriamente por meio do aumento ou diminuição do sinal original em 10%, 20%, 30%, 40% ou 50%. As taxas de acerto obtidas permaneceram as mesmas. Neste experimento, para cada sinal de teste que foi classificado, o conjunto de treinamento correspondente não foi alterado.

Foram realizados experimentos clipando-se o sinal de teste, mas esta ação não produziu resultados satisfatórios. Assim, é importante que o microfone possua uma elevada *Sound Pressure Level* (SPL), além do cuidado para que o amplificador não sature o sinal.

Em seguida, visto que a maioria dos gravadores comerciais cortam altas frequências, foi aplicado, nos sinais de teste, um filtro passa-baixas *equiripple* em 8 kHz (frequências limite da banda passante e da banda de rejeição iguais a 7,95 kHz e 8,05 kHz, respectivamente), com 0,01 dB de *ripple* na banda passante e atenuação superior a 60 dB na banda de rejeição. As TABs. 8.16 e 8.17 comparam as taxas de acerto, respectivamente, por azimuth e por conjunto de microfones, considerando a primeira árvore de classificação, ao aplicar ou não o filtro. As colunas “Sinais originais” dessas tabelas são as últimas colunas das TABs. 8.9 e 8.10, respectivamente.

Das TABs. 8.16 e 8.17, note que a aplicação de um filtro passa-baixas a 8 kHz não altera substancialmente os resultados, principalmente se considerarmos azimues pequenos. Além disso, em alguns casos, esse filtro até ocasiona uma melhora na taxa de acerto, talvez pelo fato da informação de interesse residir nas baixas frequências (inferiores a 8 kHz). Do exposto, pode-se afirmar que os atributos de textura utilizados nos experimentos são robustos a filtragem passa-baixas.

TAB. 8.16: Comparando taxas de acerto por azimuth com e sem filtro passa-baixas

Microfone	Taxas de acerto	
	Sinais originais	Com filtro
M1	100%	100%
M2	98,63%	100%
M3	97,26%	98,63%
M4	98,63%	98,63%
M5	98,63%	97,26%
M6	94,52%	95,89%
M7	97,26%	94,52%
M8	87,67%	87,67%
M9	95,89%	94,52%
M10	95,89%	90,41%
M11	86,30%	86,30%
M12	94,52%	93,15%

## 8.5 CLASSIFICAÇÃO BINÁRIA

Pode haver situações em que não é necessário utilizar um classificador multiclasse. Por exemplo, suponha que uma gravação com dois tiros de diferentes armamentos ou munições esteja disponível e que saibamos as duas armas usadas, talvez por terem sido capturadas pelos investigadores. Suponha ainda que um dos tiros atinja uma pessoa e que foi possível identificar que a vítima gritou somente quando o segundo tiro fora disparado. Neste caso, precisa-se determinar qual das duas armas atirou por último e, então, a classificação pode ser realizada entre duas classes apenas. Assim, esse é um exemplo de emprego de classificação binária.

Para os testes binários, foram utilizadas as nove armas/munições da Base 2: 9h, 45h, 38h, 357h1, 357h2, 22r, 308r, 12s e 15r. Como o sexto conjunto da TAB. 8.10 contém sinais de vários microfones (com azimuths pequenos), foram realizados experimentos considerando esse conjunto. É de se esperar que classificações binárias entre armas do mesmo tipo ou porte (cano curto ou cano longo) apresentem piores resultados que classificações envolvendo uma arma de cada porte, o que foi comprovado nos testes. Todas as classificações binárias entre armas de cano curto e de cano longo (de portes diferentes) tiveram taxa de acerto igual a 100%, o que não ocorreu para classificações dentro do mesmo tipo de arma, conforme apresentado na TAB. 8.18, que mostra as taxas de acerto, assim como o número de tiros considerados em cada caso (par de armamentos de mesmo porte).

Da TAB. 8.18, pode-se verificar que parece ser mais fácil distinguir entre duas armas

TAB. 8.17: Comparando taxas de acerto por conjunto de microfones com e sem filtro passa-baixas

Ordem do cj	Taxas de acerto	
	Sinais originais	Com filtro
Cj 01	98,63%	98,63%
Cj 02	100%	98,63%
Cj 03	97,26%	95,89%
Cj 04	95,89%	95,89%
Cj 05	93,15%	90,41%
Cj 06	94,52%	91,78%
Cj 07	90,41%	79,45%
Cj 08	78,08%	73,97%
Cj 09	76,71%	76,71%
Cj 10	78,08%	79,45%

de cano longo do que entre duas armas de cano curto, pois a taxa de acerto média entre aqueles é maior do que entre duas armas de cano curto. Além disso, a menor taxa de acerto é próxima de 90%, que ocorre quando comparamos “9h” e “38h”.

Da mesma forma que na Seção 8.3, também foram realizadas classificações binárias juntando-se as duas bases. De novo, foram feitos testes considerando o sexto conjunto da TAB. 8.10. Quase todas as classificações binárias entre armas de cano curto e de cano longo (de portes diferentes) tiveram taxa de acerto igual a 100%; apenas as classificações entre “12s” e “9h”, “12s” e “38h” e “12s” e “45h” tiveram um erro cada. A TAB. 8.19 mostra as taxas de acerto, assim como o número de tiros considerados em cada caso (par de armamentos) para as classificações dentro do mesmo tipo de arma.

Da TAB. 8.19, observa-se que continua sendo mais fácil distinguir entre duas armas de cano longo do que entre duas armas de cano curto. Além disso, a menor taxa de acerto é próxima de 90%, que ocorre quando comparamos “9h” e “38h”, ou “9h” e “45h”, ou “45h” e “38h”.

## 8.6 CLASSIFICAÇÃO DE AZIMUTE

A partir dos resultados deste capítulo até agora, concluímos que a classificação de munição ou armamento produz resultados melhores quando consideramos somente azimutes pequenos. Este resultado tende a ser mais confiável se houver alguma evidência de que o tiro foi executado com azimute menor que  $100^\circ$ , utilizando, portanto, para o conjunto de treinamento, somente sinais com azimutes pequenos. Mas se não houver

TAB. 8.18: Taxas de acerto entre duas classes com a Base 2

Par	Tiros	Taxa de acerto
9h × 45h	20	95%
9h × 38h	19	89,47%
9h × 357h1	20	100%
9h × 357h2	20	100%
45h × 38h	19	100%
45h × 357h1	20	90%
45h × 357h2	20	100%
38h × 357h1	19	100%
38h × 357h2	19	94,74%
357h1 × 357h2	20	95%
22r × 12s	14	100%
22r × 308r	21	100%
22r × 15r	21	100%
308r × 15r	20	95%
308r × 12s	13	100%
12s × 15p	13	100%

certeza a respeito da orientação do cano da arma, uma classificação binária entre duas classes (abaixo ou acima de  $100^\circ$ , isto é, azimutes pequenos ou grandes) seria útil antes da classificação de munição.

O mesmo método (atributos de textura) utilizado na classificação de munição foi também aplicado para classificar o azimute. Para isso, foram incluídos todos os sinais das duas bases citadas neste trabalho. Os azimutes da Base 1 são  $0^\circ$ ,  $9^\circ$ ,  $40^\circ$ ,  $50^\circ$ ,  $85^\circ$ ,  $95^\circ$ ,  $130^\circ$ ,  $140^\circ$ ,  $171^\circ$  e  $180^\circ$ , e os da Base 2 estão descritos na TAB. 8.9. A TAB. 8.20 mostra a quantidade de sinais de cada base por classe, totalizando 1.076 sinais.

Foi adotada a validação cruzada *leave-one-out*, da mesma forma que na classificação de munição. A TAB. 8.21 apresenta a matriz de confusão.

Da TAB. 8.21, conclui-se que a taxa de acerto é 88,66%. Note que foram considerados todos os azimutes. Porém, se alguma(s) orientação(ões) próxima(s) de  $100^\circ$  (limiar de separação das classes) for(em) excluída(s), talvez o resultado melhore, já que sinais oriundos de azimutes perto do limiar podem ser classificados incorretamente com mais frequência do que sinais distantes do limiar.

No intuito de analisar a influência de azimutes próximos do limiar de  $100^\circ$ , foram realizados testes com a exclusão de alguns azimutes, concluindo-se que a exclusão de azimutes ligeiramente maiores que o limiar (por exemplo,  $115^\circ$  da Base 2) geram melhores resultados que a exclusão de azimutes ligeiramente menores (por exemplo,  $95^\circ$  da Base

TAB. 8.19: Taxas de acerto entre duas classes com duas bases

Par	Tiros	Taxa de acerto
9h × 45h	32	90,63%
9h × 38h	31	90,32%
9h × 357h1	32	100%
9h × 357h2	26	96,15%
45h × 38h	31	90,38%
45h × 357h1	32	96,88%
45h × 357h2	26	92,31%
38h × 357h1	31	96,77%
38h × 357h2	25	92%
357h1 × 357h2	26	92,31%
22r × 12s	26	100%
22r × 308r	33	100%
22r × 15r	27	100%
308r × 15r	26	92,31%
308r × 12s	25	92%
12s × 15p	19	100%

TAB. 8.20: Número de sinais para classificação de azimute

AZIMUTE	Base 1	Base 2	TOTAL
< 100°	48	581	629
> 100°	32	415	447

1). Portanto, os melhores resultados foram obtidos com a exclusão dos sinais da Base 2 gravados pelo microfone 8 (azimute de 115°) e mantendo todos os sinais da Base 1, visto que seus azimutes grandes não são próximos de 115° (o azimute menos distante vale 130°). A TAB. 8.22 mostra a quantidade de sinais de cada base por classe, totalizando 993 sinais, e a TAB. 8.23 apresenta a matriz de confusão correspondente.

Da TAB. 8.23, conclui-se que a taxa de acerto é 95,07%, que é aproximadamente 6.5% maior que o resultado quando sinais de todos os azimutes disponíveis são incluídos. Além disso, se estivermos interessados que o tiro possua azimute pequeno (chamando assim a

TAB. 8.21: Matriz de confusão para classificação de azimute

CLASSE	AZIMUTE	
	< 100°	> 100°
< 100°	<b>560</b>	53
> 100°	69	<b>394</b>

TAB. 8.22: Número de sinais para classificação de azimute excluindo um dos azimutes

AZIMUTE	Base 1	Base 2	TOTAL
< 100°	48	581	629
> 100°	32	332	364

TAB. 8.23: Matriz de confusão para classificação de azimute excluindo o azimute 115°

CLASSE	AZIMUTE	
	< 100°	> 100°
< 100°	<b>592</b>	12
> 100°	37	<b>352</b>

classe “menor que 100 graus” como a verdadeira), para que a classificação de munição se torne mais confiável, chega-se à conclusão de que a taxa de falso alarme é de apenas 1,99%. Isto significa que se o algoritmo classifica um sinal como “azimute pequeno”, há uma alta probabilidade de que esta classificação esteja correta. Note que somente uma reduzida porção (em torno de 115°) da faixa inteira de azimutes foi excluída.

Então, dado um sinal de teste, poder-se-ia primeiramente classificar o azimute. Se o resultado for “azimute pequeno”, aplica-se a classificação de munição (como descrito neste capítulo), considerando apenas sinais com azimutes pequenos no conjunto de treinamento, estando certo de que o resultado será mais confiável do que no caso do primeiro classificador rotular o sinal de teste como “azimute grande”.

## 8.7 OUTROS TESTES

As árvores de classificação apresentadas na Seção 8.3 compreendem munições e armamentos de diferentes portes. Com o intuito de verificar se a influência do azimute nas taxas de acerto se comporta de modo distinto quando consideramos somente armas de cano curto ou de cano longo, foi testada primeiramente classificação binária entre portes de armas (cano curto e cano longo) para depois verificar como o azimute influencia cada classe de porte de arma.

### 8.7.1 CLASSIFICAÇÃO QUANTO AO PORTE DE ARMA

O mesmo método (atributos de textura) utilizado na classificação de munição foi também aplicado para classificar o porte de arma. Para isso, foram incluídos todos os sinais das duas bases citadas neste trabalho. Ao todo, são 5 armas de cano curto ou pequeno porte (9h, 45h, 38h, 357h1 e 357h2) e 5 armas de cano longo ou grande porte

(22r, 223r, 308r, 15r e 12s). A TAB. 8.24 mostra a quantidade de sinais de cada base por classe, totalizando 1.076 sinais.

TAB. 8.24: Número de sinais para classificação de porte de arma

Porte da arma	Base 1	Base 2	TOTAL
Cano curto	40	588	628
Cano longo	40	408	448

Foi adotada a validação cruzada *leave-one-out*, da mesma forma que na classificação de munição. A TAB. 8.25 apresenta a matriz de confusão.

TAB. 8.25: Matriz de confusão para classificação de porte de arma

CLASSE	PORTE DA ARMA	
	Cano curto	Cano longo
Cano curto	<b>593</b>	30
Cano longo	35	<b>418</b>

Da TAB. 8.25, pode-se verificar que a taxa de acerto é igual a 93,96%, concluindo que os atributos de textura também podem ser utilizados para classificar o porte de arma. Após isso, foi analisada a influência do azimute em cada tipo de arma (cano curto e longo).

## 8.7.2 CLASSIFICAÇÃO ENTRE ARMAS DE CANO CURTO

Foram feitos testes com as quatro armas de cano curto utilizadas na árvore de classificação da FIG. 8.4, nos mesmos moldes das TABs. 8.9 e 8.10 (por azimute e por conjunto de microfones, respectivamente). A FIG. 8.9 ilustra a árvore de classificação utilizada somente com armas de cano curto, com quantidade total de tiros igual a 39.

A TAB. 8.26 apresenta as taxas de acerto da FIG. 8.9 para cada um dos 12 azimutes da TAB. 8.9.

Comparando as TABs. 8.9 e 8.26, conclui-se que, quando são utilizadas somente as quatro armas de cano curto da FIG. 8.9, as taxas de acerto são ligeiramente menores, na média, do que quando são utilizadas as oito armas da FIG. 8.4. Ainda assim, a menor taxa de acerto para azimutes pequenos ficou próxima de 90%.

A TAB. 8.27 apresenta as taxas de acerto da FIG. 8.9 para cada conjunto de microfones da TAB. 8.10.

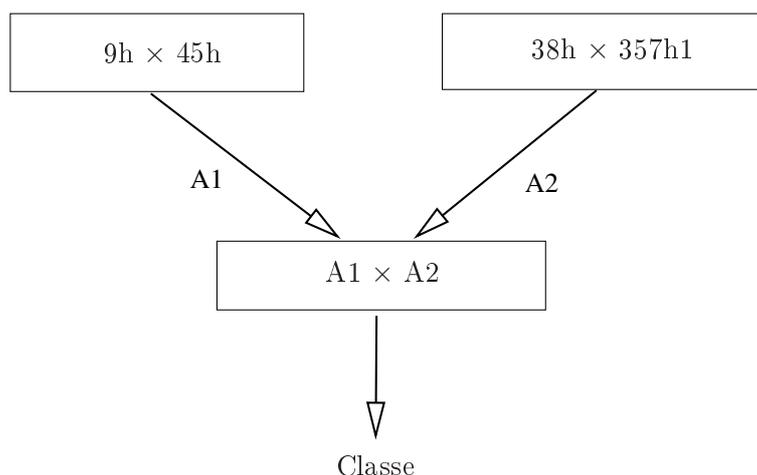


FIG. 8.9: Árvore de classificação com armas de cano curto.

TAB. 8.26: Taxas de acerto para a árvore de classificação envolvendo somente armas de cano curto, por azimute

Microfone	Azimute	Taxa de acerto
M1	0°	100%
M2	16°	97,44%
M3	33°	94,87%
M4	49°	97,44%
M5	65°	100%
M6	82°	89,74%
M7	98°	97,44%
M8	115°	84,62%
M9	130°	94,87%
M10	147°	92,31%
M11	164°	84,62%
M12	180°	100%

Comparando as TABs. 8.10 e 8.27, conclui-se que, quando são utilizadas somente as quatro armas de cano curto da FIG. 8.9, nenhuma taxa de acerto é maior do que quando são utilizadas as oito armas da FIG. 8.4. Além disso, dois dos conjuntos de azimutes pequenos tiveram taxa de acerto menor que 90%.

Mais testes foram realizados entre as armas de cano curto: com as quatro munições da FIG. 8.5 e com os cinco armamentos/munição da FIG. 8.8. Entretanto, as taxas de acerto foram, na média, (ligeiramente) menores do que quando são testadas todas as armas (canos curto e longo).

TAB. 8.27: Taxas de acerto para a árvore de classificação envolvendo somente armas de cano curto, por conjunto de microfones

Ordem do cj	Mics envolvidos	Taxa acerto
Cj 01	1, 2	97,44%
Cj 02	1, 2, 3	100%
Cj 03	1, 2, 3, 4	94,87%
Cj 04	1, 2, 3, 4, 5	92,31%
Cj 05	1, 2, 3, 4, 5, 6	87,18%
Cj 06	2, 3, 4, 5, 6, 7	92,31%
Cj 07	6, 7, 8, 9, 10	87,18%
Cj 08	7, 8, 9, 10, 11	71,79%
Cj 09	8, 9, 10, 11, 12	71,79%
Cj 10	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	74,36%

### 8.7.3 CLASSIFICAÇÃO ENTRE ARMAS DE CANO LONGO

Agora, veremos a influência da direção do cano da arma em relação ao microfone quando são consideradas somente armas de grande porte. Inicialmente, foram feitos testes com os quatro rifles/espingarda utilizados nas árvores de classificação das FIGs. 8.4 até 8.8, nos mesmos moldes das TABs. 8.9 e 8.10 (por azimute e por conjunto de microfones, respectivamente). A FIG. 8.10 ilustra a primeira árvore de classificação utilizada somente com armas de cano longo, com quantidade total de tiros igual a 34.

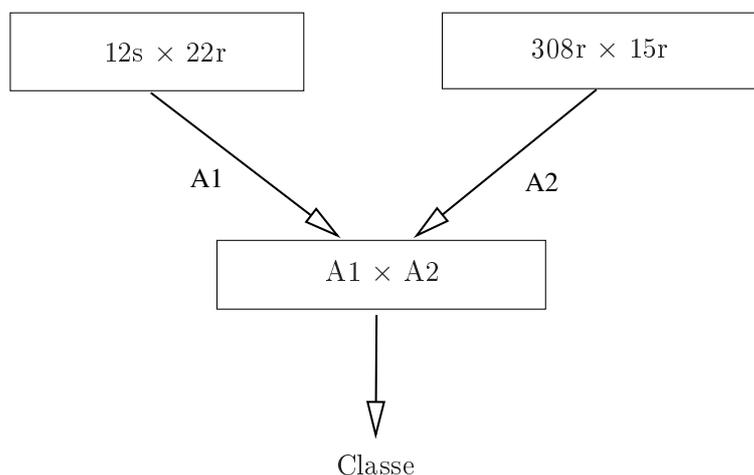


FIG. 8.10: Primeira árvore de classificação com armas de cano longo.

A TAB. 8.28 apresenta as taxas de acerto da FIG. 8.10 para cada um dos 12 azimutes da TAB. 8.9.

Da TAB. 8.28, percebe-se que as taxas de acerto são sempre maiores que 90% para

TAB. 8.28: Taxas de acerto para a primeira árvore de classificação envolvendo somente armas de cano longo, por azimute

Microfone	Azimute	Taxa de acerto
M1	0°	100%
M2	16°	100%
M3	33°	100%
M4	49°	100%
M5	65°	97,06%
M6	82°	100%
M7	98°	97,06%
M8	115°	94,12%
M9	130°	97,06%
M10	147°	100%
M11	164°	91,18%
M12	180°	100%

qualquer azimute, diferentemente das taxas de acerto referentes às FIGs. 8.4 a 8.8, em que bons resultados não são alcançados por todos os azimutes grandes.

A TAB. 8.29 apresenta as taxas de acerto da FIG. 8.10 para cada conjunto de microfones da TAB. 8.10.

TAB. 8.29: Taxas de acerto para a primeira árvore de classificação envolvendo somente armas de cano longo, por conjunto de microfones

Ordem do cj	Mics envolvidos	Taxa acerto
Cj 01	1, 2	100%
Cj 02	1, 2, 3	100%
Cj 03	1, 2, 3, 4	100%
Cj 04	1, 2, 3, 4, 5	100%
Cj 05	1, 2, 3, 4, 5, 6	94,12%
Cj 06	2, 3, 4, 5, 6, 7	97,06%
Cj 07	6, 7, 8, 9, 10	94,12%
Cj 08	7, 8, 9, 10, 11	94,12%
Cj 09	8, 9, 10, 11, 12	94,12%
Cj 10	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	94,12%

Da TAB. 8.29, verifica-se que todas as taxas de acerto são maiores que 90%, sugerindo que uma classificação entre armas de cano longo apresenta bons resultados não só com azimutes pequenos.

Da mesma forma que na Seção 8.3, foram juntados sinais de rifles/espingarda das Bases 1 e 2, totalizando 5 armas. A FIG. 8.11 ilustra a segunda árvore de classificação

somente com armas de cano longo.

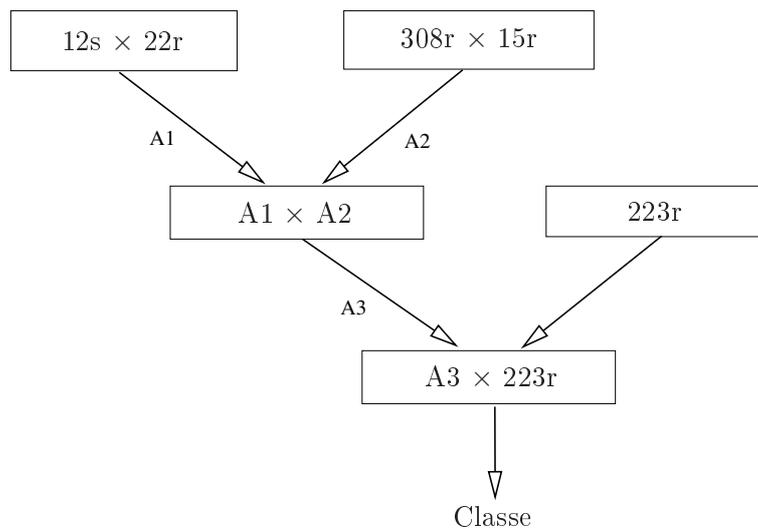


FIG. 8.11: Segunda árvore de classificação com armas de cano longo.

A TAB. 8.30 mostra o número de sinais de tiro de cada base por arma de cano longo para a FIG. 8.11. Observe que a quantidade de tiros de cada arma da Base 1, quando não nulo, é igual a 6, correspondente aos 6 azimutes pequenos daquela base, de modo semelhante à TAB. 8.13.

TAB. 8.30: Número de sinais de tiro de cada base por arma de cano longo

A+M	Número de sinais		
	Base 1	Base 2	Bases 1 e 2
22r	6	11	17
308r	6	10	16
12s	6	3	9
15r	0	10	10
223r	6	0	6
TOTAL	24	34	58

A TAB. 8.31 apresenta como as duas bases foram combinadas em termos de azimute e mostra também as taxas de acerto. Note que, em todos os casos (linhas da TAB. 8.31), foi considerada a faixa de azimutes de  $0^\circ$  a  $95^\circ$  para a Base 1.

Da TAB. 8.31, observa-se que as taxas de acerto continuam acima de 90% para todos os azimutes, mesmo com a junção das duas bases.

Também foram realizados testes com os conjuntos de azimutes da TAB. 8.10, de modo semelhante à TAB. 8.15. A TAB. 8.32 apresenta os azimutes para cada base, assim como as taxas de acerto.

TAB. 8.31: Azimutes e taxas de acerto para a segunda árvore de classificação envolvendo somente armas de cano longo com duas bases

Caso	Azimutes		Taxa de acerto
	Base 1	Base 2	
1	$0^\circ - 95^\circ$	$0^\circ$	96,55%
2	$0^\circ - 95^\circ$	$16^\circ$	96,55%
3	$0^\circ - 95^\circ$	$33^\circ$	96,55%
4	$0^\circ - 95^\circ$	$49^\circ$	98,28%
5	$0^\circ - 95^\circ$	$65^\circ$	98,28%
6	$0^\circ - 95^\circ$	$82^\circ$	98,28%
7	$0^\circ - 95^\circ$	$98^\circ$	98,28%
8	$0^\circ - 95^\circ$	$115^\circ$	94,83%
9	$0^\circ - 95^\circ$	$130^\circ$	96,55%
10	$0^\circ - 95^\circ$	$147^\circ$	94,83%
10	$0^\circ - 95^\circ$	$164^\circ$	91,38%
12	$0^\circ - 95^\circ$	$180^\circ$	98,28%

Da TAB. 8.32, verifica-se que a menor taxa de acerto é próxima de 90%, ratificando o bom desempenho do classificador de munição/armamento para todos os azimutes quando consideram-se apenas armas de cano longo.

Neste capítulo, foram estudadas as características direcionais das armas de fogo e sua influência na classificação de munição. Foi visto que os atributos obtidos da WT quantizada em níveis de cinza continuam com bons resultados e que o classificador proposto sempre apresenta taxa de acerto maior que 89% quando o cano da arma faz um ângulo menor que  $100^\circ$  em relação ao microfone. Foram utilizadas duas bases com sinais gravados em locais diferentes, de 9 armamentos, e foram testadas várias árvores de classificação, concluindo que o método apresentado pode ser empregado para classificar munição e armamento. Foi proposta também a classificação de azimute, de modo que o resultado da classificação de munição seja mais confiável quando o azimute for menor que  $100^\circ$ . Esses atributos também apresentaram bons resultados para classificar o porte da arma (cano curto ou longo), e foi verificado que a classificação entre armas de cano longo apresenta taxas de acerto maiores que aproximadamente 90%, independente de azimute, o que não ocorre entre armas de cano curto. Assim, infere-se que armas de grande porte sofrem menos influência da direção do cano da arma em relação ao microfone do que as armas de pequeno porte.

TAB. 8.32: Conjuntos de azimutes e taxas de acerto para a segunda árvore de classificação envolvendo somente armas de cano longo com duas bases

Conjunto	Azimutes		Taxa de acerto
	Base 1	Base 2	
1	$0^{\circ} - 95^{\circ}$	$0^{\circ} - 16^{\circ}$	96,55%
2	$0^{\circ} - 95^{\circ}$	$0^{\circ} - 33^{\circ}$	94,83%
3	$0^{\circ} - 95^{\circ}$	$0^{\circ} - 49^{\circ}$	94,83%
4	$0^{\circ} - 95^{\circ}$	$0^{\circ} - 65^{\circ}$	94,83%
5	$0^{\circ} - 95^{\circ}$	$0^{\circ} - 82^{\circ}$	91,38%
6	$0^{\circ} - 95^{\circ}$	$16^{\circ} - 98^{\circ}$	94,83%
7	$0^{\circ} - 95^{\circ}$	$82^{\circ} - 147^{\circ}$	89,66%
8	$0^{\circ} - 95^{\circ}$	$98^{\circ} - 164^{\circ}$	91,38%
9	$0^{\circ} - 95^{\circ}$	$115^{\circ} - 180^{\circ}$	91,38%
10	$0^{\circ} - 95^{\circ}$	$0^{\circ} - 147^{\circ}$	91,38%

## 9 CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS

Este capítulo resume as principais contribuições deste trabalho e sugere tópicos que podem ser explorados em futuras pesquisas.

### 9.1 CONCLUSÕES

Esta tese abordou duas aplicações de interesse relacionadas a sinais de tiro: a estimativa da direção de chegada (DOA) de um projétil e a classificação de munição ou armamento. A primeira aplicação consiste em um passo preliminar para a localização do atirador, enquanto que a segunda é útil em investigações forenses. Em todo o trabalho, é analisado o sinal sonoro do tiro captado por um microfone ou arranjo de microfones.

Dentre as duas principais componentes de um sinal de tiro, o MB é que possibilita estimar a direção do atirador. Porém, a grandes distâncias entre o ponto de disparo do tiro e o arranjo de microfones, cenário típico de um *sniper*, o MB aparece fortemente atenuado e, conseqüentemente, com baixa SNR, o que, juntamente com possíveis reflexões, degrada a estimativa de DOA obtida pela estimação padrão (SE).

O arranjo de microfones utilizado nos experimentos de estimativa de DOA consistiu de sete microfones. Como os algoritmos de estimação de DOA apresentados nesta tese utilizam os picos das correlações cruzadas entre sinais que chegam em dois microfones, há 21 combinações possíveis. Esses picos das correlações cruzadas estão relacionados com a diferença de tempo de chegada (TDOA) estimada dos sinais no arranjo. Formulou-se uma função custo de mínimos quadrados, composta de tantas parcelas quantos forem os pares de microfones analisados (21 no caso de SE), contendo os TDOAs estimados. A direção de chegada estimada é aquela que minimiza a referida função custo.

Foram testadas técnicas de redução de ruído antes da aplicação de algoritmo de estimativa de DOA. A subtração espectral (SS) e o filtro mediana (MF) apresentaram melhores resultados, e foram explicados neste trabalho. Foi constatado que o MF produziu melhor desempenho; porém, para SNR abaixo de -2 dB, sua simples aplicação antes do algoritmo SE não produziu resultados satisfatórios, o que ensejou a busca de outras técnicas de estimativa de DOA.

Foi apresentado, no Capítulo 3, o algoritmo denominado Busca Exaustiva,  $ES(n)$ , em que avalia-se a função custo para todas as possíveis combinações de pares de microfones,

selecionando-se a combinação que produza o menor valor da função. Para sinais muito ruidosos, em que nem todas as correlações cruzadas possuem o maior pico devido ao MB (e sim devido ao ruído), este algoritmo obteve bom desempenho, mais precisamente para sinais com SNR até o mínimo de -7 dB.

Nos casos em que os maiores picos de (quase) todas as correlações cruzadas não corresponderem ao pico correto para a estimação de DOA, faz-se necessária a busca por outros picos, ao invés de se utilizar apenas o maior deles. Nesse sentido, aplicou-se ES(4) entre os 10 maiores picos das correlações cruzadas. A fim de não utilizar o custo computacional de testar a função custo  $5,985 \times 10^7$  vezes – que resulta da combinação dos 10 maiores picos e 4 correlações cruzadas a partir de 21 pares de microfones – recorreu-se ao conceito de *loop fundamental consistente* (cFL), restringindo a pesquisa a um conjunto muito menor de combinações (350.000). O algoritmo de estimação de DOA resultante dessa ideia foi chamado de *Searching Consistent Fundamental Loop* (SCFL).

Em experimentos com sinais muito ruidosos (SNR abaixo de -2 dB), ao comparar ES e SCFL, foi verificado que há casos em que um falha enquanto o outro apresenta bom desempenho, e vice-versa em outros casos. Foi então proposto um algoritmo, referido como ES-SCFL, que é uma combinação das técnicas SE, ES e SCFL. Tal algoritmo utiliza uma árvore de decisão, compara a função custo com limiares cuidadosamente selecionados e pode ser empregado para sinais com qualquer valor de SNR. Foi comparado o desempenho do algoritmo ES-SCFL com os algoritmos ES(5), ES(4), SCFL, ILS e SE, utilizando sinais com DOAs simulados e sinais de campo gravados no Centro de Avaliações do Exército. Em todos os casos, o algoritmo ES-SCFL produziu os melhores resultados.

Para a aplicação de classificação de munição, atributos de textura extraídos da WT quantizada em níveis de cinza, tratada como imagem, mostraram melhores resultados que outras técnicas apresentadas neste trabalho. Foram utilizadas várias bases com diferentes armamentos, com sinais gravados em diversos locais: Brasil, Finlândia e Estados Unidos da América. Foi empregado o SVM com kernel linear com o uso de estrutura ADAG para transformar classificadores binários em classificador multiclasse. O método proposto mostrou-se útil para classificar munição e armamento, principalmente quando o ângulo entre a cano da arma e o microfone (azimute) é menor que  $100^\circ$ , visto que, para esses azimutes, as Transformadas Wavelet do muzzle blast são pouco diferentes, ao contrário da de um sinal oriundo de um azimute grande. Foi verificado também que o emprego de um filtro passa-baixas com frequência de corte em 8 kHz não altera substancialmente o desempenho do classificador. Foi proposta a classificação de azimute (maior ou menor que

100°), utilizando os mesmos atributos de textura, de modo que o resultado da classificação de munição seja mais confiável quando o azimute for menor que 100°. Por último, foi constatado que a direcionalidade do cano da arma em relação ao microfone tem menos influência quando o classificador consiste exclusivamente de armas de cano longo do que de armas de cano curto.

## 9.2 SUGESTÕES PARA TRABALHOS FUTUROS

Com relação à estimativa de DOA, poder-se-ia empregar o algoritmo ES-SCFL em microfones localizados em uma quadricóptero ou VANT, a fim de verificar seu desempenho junto à fonte de ruído do motor do veículo. Um bom resultado ensejaria o emprego do VANT, com um arranjo de microfones embarcado, em áreas de risco, para localizar atiradores.

Com relação à classificação de munição, um tópico para trabalho futuro pode ser a identificação dos principais atributos de textura dentre os 26 citados nesta tese, a fim de determinar quais atributos mais contribuem para a alta taxa de acerto na classificação de munição. Outra possibilidade é estender o método para casos em que o sinal de teste pode pertencer a nenhuma das classes consideradas, ou seja, aplicações de reconhecimento *open set*. Os sinais analisados no Capítulo 8 para fins de estudo da influência do azimute na classificação de munição/armamento foram gravados por microfones profissionais a curta distância. Para que o método proposto se torne operacional, com previsão de emprego em situações táticas reais, faz-se necessária uma pesquisa futura com sinais de tiro obtidos de gravadores não profissionais ou mais distantes do atirador.

## 10 REFERÊNCIAS BIBLIOGRÁFICAS

- AHMED, T., UPPAL, M. e MUHAMMAD, A. **Improving efficiency and reliability of gunshot detection systems.** Em *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, págs. 513–517, 2013.
- BISHOP, C. *Pattern recognition and Machine Learning.* Springer Science + Business Media, LLC, 2006.
- BOHN, D. **Environmental Effects on the Speed of Sound.** *Journal of the Audio Engineering Society*, 36(4):223–231, 1988.
- BOLL, S. **Suppression of Acoustic Noise in Speech using Spectral Subtraction.** *IEEE Transactions on Acoustics, Speech, and Signal Processing*, volume 27, número 2, págs. 113-120, Abr. 1979. ISSN 0096-3518.
- BORZINO, A., APOLINÁRIO JR., J. e CAMPOS, M. **Estimating direction of arrival of long range gunshot signals.** Em *International Telecommunications Symposium (ITS)*, págs. 1–4, Ago. 2014a.
- BORZINO, A., APOLINÁRIO JR., J. e CAMPOS, M. **Robust DoA Estimation of Heavily Noisy Gunshot Signals.** Em *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Abr. 2015a.
- BORZINO, A., APOLINÁRIO JR., J. e CAMPOS, M. **Consistent DOA Estimation of Heavily Noisy Gunshot Signals Using a Microphone Array.** *Institution of Engineering and Technology (IET) - Radar, Sonar & Navigation*, Jun. 2016.
- BORZINO, A., APOLINÁRIO JR., J., CAMPOS, M. e PAGLIARI, C. **Gunshot Signal Enhancement for DoA estimation and Weapon Recognition.** Em *European Signal Processing Conference (EUSIPCO)*, págs. 1–4, Set. 2014b.
- BORZINO, A., APOLINÁRIO JR., J., DE CAMPOS, M. L. e BISCAINHO, L. **Signal Enhancement for Gunshot DOA Estimation with Median Filters.** Em *Latin American Symposium on Circuits and Systems (LASCAS)*, Fev. 2015b.
- CANADAS-QUESADA, F., VERA-CANDEAS, P., RUIZ-REYES, N., CARABIAS-ORTI, J. e CABANAS-MOLERO, P. **Percussive/harmonic sound separation by non-negative matrix factorization with smoothness/sparseness constraints.** *EURASIP Journal on Audio, Speech, and Music Processing*, volume 26, págs. 1-17, Jul. 2014.
- CAWLEY, G. **Leave-one out cross-validation based model selection criteria for weighted ls-svm.** Em *International Joint Conference on Neural Networks*, págs. 1661–1668, 2006.

- CHACÓN-RODRÍGUEZ, A., JULIÁN, P. e CASTRO, L. **Evaluation of Gunshot Detection Algorithms.** Em *IEEE Transactions on Circuits and Systems*, volume 58(2), págs. 363–373, Fev. 2011.
- CHAN, Y., HATTIN, R. e PLANT, J. **The least squares estimation of time delay and its use in signal detection.** *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(3), págs. 217-222, Jun. 1978.
- CHOU, W. e JUANG, B. *Pattern Recognition in Speech and Language Processing.* CRC Press, 2003.
- CHU, A., SEHGAL, C. M. e GREENLEAF, J. F. **Use of gray value distribution of run lengths for texture analysis.** Em *Pattern Recognition Letters*, volume 11, págs. 415–420, 1990.
- CORTES, C. e VAPNIK, V. **Support vector networks.** *AT&T Labs-Research, USA*, 20:273–296, 1995.
- CRISTIANINI, N. e SHAW-TAYLOR, J. *An introduction to support vector machines and other kernel-based learning methods.* Cambridge University Press, 2000.
- DAUBECHIES, I. *Ten lectures on wavelets.* Society for Industrial and Applied Mathematics, 1992.
- DEBNATH, L. *Applied and Numerical Harmonic Analysis: Wavelets and Signal Processing.* Springer Science + Business Media New York, 2003.
- DIBIASE, J., SILVERMAN, H. e BRANDSTEIN, M. *Microphone Arrays: signal processing techniques and applications, chapter 8: Robust localization in reverberant rooms.* 2001.
- DING, C. e PENG, H. **Minimum redundancy feature selection from microarray gene expression data.** Em *Proceedings of the Computational Systems Bioinformatics*, págs. 523–528, 2003.
- DMOCHOWSKI, J., BENESTY, J. e AFFES, S. **Direction of arrival estimation using the parameterized spatial correlation matrix.** Em *IEEE Transactions on Audio, Speech, and Language Processing*, volume 15(4), págs. 1327–1339, Maio 2007.
- DONOHU, D. e JOHNSTONE, I. **Threshold selection for wavelet shrinkage of noisy data.** Em *Conference of the IEEE Engineering in Medicine and biology society*, págs. 24a–25a, 1994.
- DUFAUX, A., BESACIER, L., ANSORGE, M. e PELLANDINI, F. **Automatic Sound Detection and Recognition for Noisy Environment.** Em *European Signal processing Conference (EUSIPCO)*, págs. 1–4, Set. 2000.
- DUMOND, J., COHEN, E., PANOFKY, W. e DEEDS, E. **A determination of the wave forms and laws of propagation and dissipation of ballistic shock waves.** *The Journal of the Acoustical Society of America*, 18(1):93–118, 1946.

- ELKAMCHOUCI, H. e MOFFED, M. A. E. **Direction-of-Arrival Methods (DOA) and Time Difference of Arrival (TDOA) Position Location Technique.** Em *Twenty Second National Radio Science Conference (NRSC)*, págs. 173–182, Mar. 2005.
- EXÉRCITO, E. M. *O caçador - Instruções Provisórias.* Ministério do Exército, 1998.
- FITZGERALD, D. **Harmonic/percussive separation using median filtering.** Em *International Conference on Digital Audio Effects (DAFx-10)*, Set. 2010.
- FREIRE, I. e APOLINÁRIO JR., J. **DOA of gunshot signals in a spatial microphone array: performance of the interpolated generalized cross-correlation method.** Em *Argentine School of Micro-Nanoelectronics, Technology, and Applications (EAMTA)*, Ago. 2011.
- FREIRE, I., PRANDEL, P. e APOLINÁRIO JR., J. **Sobre a escolha de Sinais em Arranjos de Microfones estimando DoA com GCC-PhaT.** Em *Simpósio Brasileiro de Telecomunicações (SBrT)*, págs. 1–5, Set. 2012.
- FREIRE, I. **Robust Direction-of-Arrival by matched-lags, applied to gunshots.** *Acoustical Society of America*, volume 135, número 6, págs. 246-251, Maio 2014.
- FREYTAG, J. e BRUSTAD, B. **A survey of audio forensic gunshot investigations.** Em *International Conference: Audio Forensics in Digital Age*, págs. 131–134, Jul. 2005.
- GALLOWAY, M. **Texture analysis using gray level run lengths.** Em *Computer Graphics and Image Processing*, volume 4, págs. 172–179, 1975.
- HARALICK, R. M., SHANMUGAM, K. e DINSTEN, I. **Texture Features for Image Classification.** Em *IEEE Transactions on Systems, Man and Cybernetics*, volume 6, págs. 610–621, 1973.
- HAYKIN, S. *Neural Networks: A Comprehensive Foundation.* Prentice Hall, 1999.
- HSU, C. e LIN, C. **A comparison of methods for multiclass support vector machines.** Em *IEEE Transactions on Neural Networks*, volume 13(2), págs. 415–425, Mar. 2002.
- HUNG, K. **Medical Pattern Recognition: Applying an Improved Intuitionistic Fuzzy Cross-Entropy Approach.** *Advances in Fuzzy Systems*, 2012:1–6, 2012.
- ISA, S., NOVIYANTO, A. e ARYMURTHY, A. **Optimal Selection of Wavelet Thresholding Algorithm for ECG Signal Denoising.** Em *International Conference on Advanced Computer Science and Information Systems*, págs. 365–370, 2011.
- JAMES, G., WITTEN, D., HASTIE, T. e TIBSHIRANI, R. *An Introduction to Statistical Learning with Applications in R.* Springer Science + Business Media New York, 2013.

- KHAN, S., DIVAKARAN, A. e SAWHNEY, H. **Weapon identification using acoustic signatures across varying capture conditions.** *United States Patent Application Publication*, Out. 2010.
- KIJSIRIKUL, B. e USSIVAKUL, N. **Multiclass support vector machines using adaptive directed acyclic graph.** Em *International Joint Conference on Neural Networks*, págs. 980–985, 2002.
- KIKTOVA-VOZARIKOVA, E., JUHAR, J. e CIZMAR, A. **Feature selection for acoustic events detection.** *Multimed Tools Appl.*, 74:4213–4233, 2015.
- KNAPP, C. e CARTER, C. **The generalized correlation method for estimation of time delay.** *IEEE Transactions on Acoustics, Speech, and Signal Processing*, volume 24, número 4, págs. 320–327, Ago. 1976.
- MAHER, R. **Modeling and Signal Processing of Acoustic Gunshot Recordings.** Em *Signal Processing Education Workshop*, págs. 257–261, Set. 2006.
- MAHER, R. **Acoustical Characterization of Gunshots.** Em *IEEE Workshop on Signal Processing Applications for Public Security and Forensics (SAFE'07)*, págs. 1–5, Abr. 2007.
- MAHER, R. e SHAW, S. **Deciphering Gunshot Recordings.** Em *Audio Engineering Society International Conference*, Jun. 2008.
- MAHER, R. e SHAW, S. **Directional aspects of forensic gunshot recordings.** Em *Audio Engineering Society International Conference*, págs. 1–6, Jun. 2010.
- MAKINEN, T. e PERTILA, P. **Shooter localization and bullet trajectory, caliber, and speed estimation based on detected firing sounds.** *Applied Acoustics*, volume 71, págs. 902–913, Jun. 2010.
- MALLAT, S. *A wavelet tour of signal processing.* CA: Academic Press, San Diego, 1999.
- MAYS, B. **Shockwave and Muzzle Blast Classification via Joint Time Frequency and Wavelet Analysis.** *U.S. Army Research Laboratory*, Set. 2001.
- MEDINA, C., ALCAIM, A. e APOLINÁRIO JR., J. **Wavelet denoising of speech using neural networks for threshold selection.** *Electronics Letters*, 39, 2003.
- PAPOULIS, A. e PILLAI, S. *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, 2001.
- PAREDES, D. e APOLINÁRIO JR., J. **Shooter localization based on DoA Estimation of Gunshot Signals and Digital Map Information.** *IEEE Latin America Transactions*, 13:441–447, Fev. 2015.
- PIERCE, A. D. **Statistical theory of atmospheric turbulence effects on sonic-boom rise times.** 49(3B):906–924, 1971.

- RAJ, K., MURUGAN, S., NATARAJAN, V. e RADHA, S. **Denoising Algorithm using Wavelet for Underwater Signal Affected by Wind Driven Ambient Noise.** Em *International Conference on Recent Trends in Information Technology*, págs. 943–946, 2011.
- RAMOS, A., HOLM, S., GUDVANGEN, S. e OTTERLEI, R. **A Modified Spectral Subtraction algorithm for Real-Time noise reduction applied to Gunshot Acoustics.** Em *International Conference on Signals and Electronic Systems (ICSES)*, págs. 1–5, Set. 2012a.
- RAMOS, A., HOLM, S., GUDVANGEN, S. e OTTERLEI, R. **Real-time vehicle noise cancellation techniques for Gunshot Acoustics.** Em *SPIE*, volume 8359, págs. 835917–835917–9, 2012b. URL <http://dx.doi.org/10.1117/12.920275>.
- RAMOS, A., HOLM, S., GUDVANGEN, S. e OTTERLEI, R. **A Spectral Subtraction Based Algorithm for Real-Time Noise Cancellation with Application to gunshot Acoustics.** *International Journal of Electronics and Telecommunications*, volume 59, número 1, págs. 93-98, Maio 2013.
- RAMOS, A. L. L., HOLM, S., GUDVANGEN, S. e OTTERLEI, R. **Delay-and-sum Beamforming for Direction of Arrival Estimation Applied to Gunshot Acoustics.** Em *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense X*, volume 8019, Maio 2011.
- SALLAI, J., HEDGECOCK, W., VOLGYESI, P., NADAS, A., BALOGH, G. e LEDECZI, A. **Weapon classification and shooter localization using distributed multichannel acoustic sensors.** *Journal of Systems Architecture*, 57: 869–885, Nov. 2011.
- SCHEIRER, W., ROCHA, A., SAPKOTA, A. e BOULT, T. **Towards Open Set Recognition.** Em *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 35(7), págs. 1757–1772, Jul. 2013.
- SCHEUING, J. e YANG, B. **Disambiguation of TDOA Estimation for Multiple Sources in Reverberant Environments.** Em *IEEE Transactions on Audio, Speech, and Language Processing*, volume 16(8), págs. 1479–1489, Nov. 2008.
- SELVARAJAH, S. e KODITUWAKKU, S. **Analysis and comparison of texture features for content based image retrieval.** *International Journal of Latest Trends in Computing*, 2(1), 2011.
- SMITH, J. e ABEL, J. **Closed-Form Least-Squares Source Location Estimation from Range-Difference Measurements.** *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(12), págs. 1661-1669, Dez. 1987.
- TIAN, X. e TANG, B. **Spectrum Texture Features Based Radar Deception Jamming Recognition Using Joint Frequency-slow Time Processing.** *Journal of Computational Information Systems*, págs. 5181–5188, 2013.

- TREES, H. *Detection, Estimation, and Modulation Theory, Part I*. WILEY-ITERSCIENCE, 2004.
- UWECHUE, O. e PANDYA, A. *Human face recognition using third-order synthetic neural networks*. Springer Science + Business Media, LLC, 2006.
- VAN TREES, H. *Optimum Array Processing*, volume IV. WILEY-ITERSCIENCE, New York, 2002.
- WILLETT, P. **Recent trends in hierarchic document clustering: a critical review**. *Information Processing & Management*, 24:577–597, 1988.
- XUE, W., LIANG, S. e LIU, W. **DOA estimation of speech source in noisy environments with weighted spatial bispectrum correlation matrix**. Em *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, págs. 2282–2286, May 2014.

## 11 APÊNDICES

## 11.1 APÊNDICE 1: ATRIBUTOS DE TEXTURA UTILIZADOS NESTE TRABALHO

### 11.1.1 PRIMEIRO CONJUNTO DE ATRIBUTOS

Os oito primeiros atributos citados no Capítulo 7 são derivados diretamente do valor quantizado da intensidade de nível de cinza de cada pixel da imagem da WT. As quatro imagens da FIG. 7.3 são exemplos de figuras de onde extraem-se esses oito atributos. Para cada imagem, nas expressões a seguir,  $i$  refere-se à ordem do pixel no eixo horizontal (deslocamento),  $j$  é a ordem do pixel no eixo vertical (escala),  $I(i, j)$  é a intensidade de nível de cinza do pixel de coordenadas  $(i, j)$ ,  $M$  é o número de deslocamentos,  $N$  é o número de escalas,  $\mu$  é a média da intensidade de níveis de cinza e  $\sigma$  é o desvio padrão. Note que a quantidade de pixels da imagem da WT é  $M \times N$ .

- Correlação

$$COR = - \sum_i^M \sum_j^N \frac{(i - \mu_x)(j - \mu_y) I(i, j)}{\sigma_x \sigma_y}, \text{ onde}$$

$$\mu_x = \sum_i^M i \sum_j^N I(i, j),$$

$$\mu_y = \sum_j^N j \sum_i^M I(i, j),$$

$$\sigma_x = \sum_i^M (i - \mu_x)^2 \sum_j^N I(i, j),$$

$$\sigma_y = \sum_j^N (j - \mu_y)^2 \sum_i^M I(i, j).$$

- Contraste

$$CON = \sum_i^M \sum_j^N (i - j)^2 I(i, j).$$

- Energia

$$ENE = \frac{\sum_i^M \sum_j^N I^2(i, j)}{MN}.$$

- Entropia

$$ENT = \frac{- \sum_i^M \sum_j^N I(i, j) \log(I(i, j))}{MN}.$$

- Momento diferencial inverso

$$IDM = \sum_i^M \sum_j^N \frac{1}{1 + (i - j)^2} I(i, j).$$

- Parâmetro R

$$R = \frac{\sigma^2}{\mu^2}.$$

- *Skewness*

$$SKE = \frac{\sum_i^M \sum_j^N (I(i, j) - \mu)^3}{MN\sigma^2}.$$

- *Kurtosis*

$$KUR = \frac{\sum_i^M \sum_j^N (I(i, j) - \mu)^4}{MN\sigma^4} - 3.$$

### 11.1.2 SEGUNDO CONJUNTO DE ATRIBUTOS

Os dez atributos seguintes citados no Capítulo 7 são derivados da RLM. As quatro imagens da FIG. 7.4 são exemplos de figuras de onde extraem-se esses dez atributos. Para cada imagem, nas expressões a seguir,  $p(i, j)$  é o elemento da RLM que corresponde à quantidade de corridas de pixels de nível de cinza  $i$  com comprimento de corrida  $j$  da imagem da WT em todas as direções (horizontal, vertical e diagonal);  $M$  é a quantidade de níveis de cinza (256);  $N$  é o maior comprimento de corrida encontrado na imagem, ou seja, é o número máximo de pixels consecutivos com o mesmo nível de cinza, e  $n_r$  é a quantidade total de corridas, ou seja, é a soma dos elementos da RLM.

- *Short Run Emphasis*

$$SRE = \frac{1}{n_r} \sum_i^M \sum_j^N \frac{p(i, j)}{j^2}.$$

- *Long Run Emphasis*

$$LRE = \frac{1}{n_r} \sum_i^M \sum_j^N p(i, j)j^2.$$

- *Gray-Level Nonuniformity*

$$GLN = \frac{1}{n_r} \sum_i^M \left( \sum_j^N p(i, j) \right)^2.$$

- *Run Percentage*

$$RP = \frac{n_r}{MN}.$$

- *Low Gray-Level Run Emphasis*

$$LGRE = \frac{1}{n_r} \sum_i^M \sum_j^N \frac{p(i, j)}{i^2}.$$

- *High Gray-Level Run Emphasis*

$$HGRE = \frac{1}{n_r} \sum_i^M \sum_j^N p(i, j) i^2.$$

- *Short Run Low Gray-Level Emphasis*

$$SRLGE = \frac{1}{n_r} \sum_i^M \sum_j^N \frac{p(i, j)}{i^2 j^2}.$$

- *Short Run High Gray-Level Emphasis*

$$SRHGE = \frac{1}{n_r} \sum_i^M \sum_j^N \frac{p(i, j) i^2}{j^2}.$$

- *Long Run Low Gray-Level Emphasis*

$$LRLGE = \frac{1}{n_r} \sum_i^M \sum_j^N \frac{p(i, j) j^2}{i^2}.$$

- *Long Run High Gray-Level Emphasis*

$$LRHGE = \frac{1}{n_r} \sum_i^M \sum_j^N p(i, j) i^2 j^2.$$

### 11.1.3 TERCEIRO CONJUNTO DE ATRIBUTOS

Os oito atributos restantes citados no Capítulo 7 são derivados da GLCM. As quatro imagens da FIG. 7.5 são exemplos de figuras de onde extraem-se esses oito atributos. Para cada imagem, nas expressões a seguir,  $p_d(i, j)$  é o elemento da GLCM que corresponde à quantidade de vezes que um pixel de nível de cinza  $i$  é adjacente a um pixel de nível de cinza  $j$  da imagem da WT em todas as direções (horizontal, vertical e diagonal);  $M$  é a quantidade de níveis de cinza (256), e  $\mu_d$  é a média de todos os elementos da GLCM.

- Correlação

$$COR = - \sum_i^M \sum_j^M \frac{(i - \mu_x)(j - \mu_y) I(i, j)}{\sigma_x \sigma_y}, \text{ onde}$$

$$\mu_x = \sum_i^M i \sum_j^M p_d(i, j),$$

$$\mu_y = \sum_j^M j \sum_i^M p_d(i, j),$$

$$\sigma_x = \sum_i^M (i - \mu_x)^2 \sum_j^M p_d(i, j),$$

$$\sigma_y = \sum_j^M (j - \mu_y)^2 \sum_i^M p_d(i, j).$$

- Contraste

$$CON = \sum_i^M \sum_j^M (i - j)^2 p_d(i, j).$$

- Energia

$$ENE = \frac{\sum_i^M \sum_j^M p_d^2(i, j)}{M^2}.$$

- Entropia

$$ENT = \frac{- \sum_i^M \sum_j^M p_d(i, j) \log(p_d(i, j))}{M^2}.$$

- Momento diferencial inverso

$$IDM = \sum_i^M \sum_j^M \frac{1}{1 + (i - j)^2} p_d(i, j).$$

- Variância

$$VAR = \frac{\sum_i^M \sum_j^M (p_d(i, j) - \mu_d)^2}{M^2}.$$

- Dissimilaridade

$$DIS = \sum_i^M \sum_j^M p_d(i, j) |i - j|.$$

- Homogeneidade

$$HOM = \sum_i^M \sum_j^M \frac{1}{1 + |i - j|} p_d(i, j).$$

## 11.2 APÊNDICE 2: PUBLICAÇÕES POR CAPÍTULO

A seguir, são listadas as publicações decorrentes deste trabalho com seus respectivos capítulos.

1) Borzino, A.M.C.R. and Apolinário JR., J.A. and Campos, M.L., **Estimating direction of arrival of long range gunshot signals**, International Telecommunications Symposium (ITS), 2014 ⇒ Capítulo 2.

2) Borzino, A.M.C.R. and Apolinário JR., J.A. and Campos, M.L. and Pagliari, C.L., **Gunshot Signal Enhancement for DoA estimation and Weapon Recognition**, European Signal Processing Conference (EUSIPCO), 2014 ⇒ Capítulo 2.

3) Borzino, A.M.C.R. and Apolinário Jr., J.A. and de Campos, Marcello L.R. and Biscainho, L.W.P., **Signal Enhancement for Gunshot DOA Estimation with Median Filters**, Latin American Symposium on Circuits and Systems (LASCAS), 2015 ⇒ Capítulo 2.

4) Borzino, A.M.C.R. and Apolinário JR., J.A. and Campos, M.L., **Robust DoA Estimation of Heavily Noisy Gunshot Signals**, IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2015 ⇒ Capítulo 3.

5) Borzino, A.M.C.R. and Apolinário JR., J.A. and Campos, M.L., **Consistent DOA Estimation of Heavily Noisy Gunshot Signals Using a Microphone Array**, Institution of Engineering and Technology (IET) - Radar, Sonar & Navigation ⇒ Capítulos 4, 5 e 6.