

UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE
PRODUÇÃO

Reconhecimento de Padrões em Energia de Emissões Acústicas na
Frequência Audível em Processos de Torneamento de Aços
Endurecidos

Edielson Prevato Frigieri

Itajubá

Dezembro 2016

UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE
PRODUÇÃO

Edielson Prevato Frigieri

Reconhecimento de Padrões em Energia de Emissões Acústicas na
Frequência Audível em Processos de Torneamento de Aços
Endurecidos

Tese submetida ao Programa de Pós-Graduação em Engenharia de Produção como parte dos requisitos para a obtenção do título de Doutor em Ciências em Engenharia de Produção.

Orientador: Prof. Dr. Anderson Paulo de Paiva

Itajubá
Dezembro 2016

UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE
PRODUÇÃO

Edielson Prevato Frigieri

Reconhecimento de Padrões em Energia de Emissões Acústicas na
Frequência Audível em Processos de Torneamento de Aços
Endurecidos

Tese aprovada por banca examinadora em 7 de Dezembro de 2016, conferindo ao autor o título de Doutor em Ciências em Engenharia de Produção.

Banca Examinadora:

Prof. Dr. Anderson Paulo de Paiva (Orientador)

Prof. Dr. João Roberto Ferreira

Prof. Dr. Carlos Alberto Ynoguti

Prof. Dr. Sebastião Carlos da Costa

Prof. Dr. Guilherme Augusto Barucke Marcondes

Prof. Dr. Rafael Coradi Leme

Itajubá
Dezembro 2016

Aos meus pais, Edi e Suely, à minha esposa Karina e à minha filha Julia, pelo apoio total e incondicional.

AGRADECIMENTOS

Os agradecimentos principais são direcionados aos orientadores e amigos, Anderson P. Paiva e Carlos A. Ynoguti, pelo tempo e paciência dedicados durante meu tempo de aprendizado.

Aos meus colegas de pós-graduação, em especial Paulo Henrique Campos, Rogério Peruchi, Luis Gustavo, Tarcísio e Pedro Papandrea, que trabalharam muito para tornarem o trabalho possível.

Agradeço também à Universidade Federal de Itajubá, que me ofereceu a oportunidade de retomar meus estudos, e pela oportunidade de participar do programa de Doutorado Sanduíche na Northwestern University. Um agradecimento especial aos Prof. Pedro Paulo Balestrassi e David P. Morton pela oportunidade e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo auxílio financeiro através da concessão de minha bolsa de doutorado sanduíche (processo BEX 3203/15-8).

Agradecimentos especiais são direcionados à minha segunda mãe Marli, pelo incentivo e dedicação e pelos cafés e biscoitos quentinhos e deliciosos sempre preparados com muito carinho.

Aos meus pais, Edi e Suely, pelo incentivo, amizade, amor e dedicação, importantes para minha formação como homem. Amo muito vocês.

À minha esposa Karina e à minha filha Julia, pelo apoio total e incondicional, pelas noites mal dormidas, pelo tempo distante da família. Minhas sinceras desculpas pelo tempo que não pude dedicar a vocês.

Agradeço a Deus por ser um guia nas difíceis decisões e um suporte nos momentos mais difíceis.

*“Viva como se fosse morrer amanhã.
Aprenda como se fosse viver para sempre.”
(Mahatma Gandhi)*

RESUMO

O objetivo deste trabalho é avaliar a emissão de sinais acústicos durante o processo de torneamento do aço ABNT 52100 endurecido e identificar quais aspectos do sinal são capazes de representar características do processo, mais precisamente, a rugosidade resultante da peça. Dois parâmetros são avaliados com este intuito: os coeficientes Mel-cepstrais (MFCC) e os escores dos componentes principais obtidos a partir do espectro de potência do sinal acústico. A técnica de análise de correlação é utilizada a fim de determinar o grau de relação entre os parâmetros do sinal acústico e as características do processo, definindo os melhores aspectos do sinal a serem utilizados no sistema de monitoramento e, com base nos resultados, dois sistemas são propostos: um baseado em Modelo de Mistura de Gaussianas (GMM) e, outro, que utiliza Máquina de Vetores de Suporte (SVM). Ambos os sistemas são testados com diferentes combinações de parâmetros e, através de análise fatorial, são definidas as configurações ótimas para cada um. Utilizando as configurações ótimas, ambos os sistemas são utilizados para verificar os parâmetros extraídos dos sinais acústicos a fim de determinar qual parâmetro resulta no melhor desempenho na tarefa de monitoramento. Os resultados obtidos mostram que tanto o MFCC quanto os escores dos componentes principais da energia do sinal apresentam ótimo desempenho na classificação de níveis de rugosidade, atingindo precisões acima de 95%, porém com uma leve vantagem para os escores da energia do sinal, que chegaram a atingir uma precisão de 100% para algumas configurações. Outro resultado importante é a superioridade do sistema de monitoramento baseado em SVM, que apresentou precisões melhores que do sistema baseado em GMM. A partir destes resultados, concluiu-se que o sinal acústico (som), através de parâmetros como os coeficientes mel-cepstrais e dos escores da energia do sinal, pode ser utilizado como forma de identificar variações no comportamento do processo, principalmente a rugosidade resultante da peça, o que permite o monitoramento de sua qualidade de forma indireta através de sistemas baseados em classificadores como GMM e SVM.

Palavras-chaves: Monitoramento; Som; Coeficientes Mel Cepstrais; Análise de Componentes Principais; Modelo de Mistura de Gaussianas; Máquina de Vetores de Suporte.

ABSTRACT

The purpose of this work is to evaluate the acoustic signal emission during the AISI 52100 steel hard turning process and identify which signal parameters are capable of representing the process characteristics, specially the resulting surface roughness of the work-piece. Two features are evaluated for this purpose: the Mel-Frequency Cepstral Coefficients (MFCC) and the scores of the principal components obtained from the energy of the acoustic signal. Correlation analysis is used to determine the degree of relationship between them and characteristics of the process, achieving the best features to be used in the monitoring system. Using the parameters with higher correlation, two monitoring systems are proposed: one based on Gaussian Mixture Model (GMM) and other using Support Vector Machine (SVM). Both systems are tested with different feature combinations, and through factorial analysis, the optimal setup for each system is obtained. Using the optimal setup, both systems are used to verify the parameters extracted from acoustic signals in order to determine which one result in the best performance in the monitoring task. Obtained results show that both the MFCC and the scores of the signal energy present great performance in the classification task, with accuracy over 95%, however slightly outperformed by the scores of the signal energy, which reached an accuracy of 100% for some setups. Another important result is the superiority of SVM-based monitoring system that presented better accuracy than the system based on GMM. From these results, it was concluded that the acoustic signal (sound), represented by parameters such as Mel-Frequency Cepstral Coefficients and the scores of the signal energy, may be used as a means of identifying changes in process behavior, mainly the surface roughness, which allows the quality monitoring of the process in a indirect way through systems based on classifiers such as GMM and SVM.

Key-words: Monitoring; Sound; Mel-frequency Cepstral Coefficientes; Principal Component Analysis; Gaussian Mixture Model; Support Vector Machine.

LISTA DE ILUSTRAÇÕES

Figura 2.1 –Principais etapas de um processo de monitoramento.	8
Figura 2.2 –Distribuição anual das publicações revisadas	10
Figura 2.3 –Processos de fabricação e mecânico estudados nas publicações revisadas	11
Figura 2.4 –Parâmetros de sinais estudados nas publicações revisadas	14
Figura 2.5 –Classificadores estudados nas publicações revisadas	15
Figura 3.1 –Etapas básicas do processo de extração de parâmetros.	17
Figura 3.2 –Exemplo de janelamento	18
Figura 3.3 –Gráfico da janela Hamming para $N_W = 250$	20
Figura 3.4 –Representação do mascaramento auditivo.	22
Figura 3.5 –Banco de filtros na escala <i>mel</i>	23
Figura 3.6 –Exemplo do processo de projeção ortogonal para dados com duas va- riáveis.	24
Figura 4.1 –Exemplo de distribuição formada pela combinação linear de Gaussianas	29
Figura 4.2 –Representação de um modelo de misturas de Gaussianas	30
Figura 4.3 –Exemplo do conceito de margem adotado pelas máquinas de vetores de suporte.	34
Figura 4.4 –Ilustração das variáveis de folga.	38
Figura 5.1 –Exemplo de arranjo composto	46
Figura 5.2 –Arranjo do tipo Box-Behnken para três fatores.	47
Figura 6.1 –Detalhes da configuração do experimento	54
Figura 6.2 –Corpo de prova do aço ABNT 52100.	55
Figura 6.3 –Gráficos de contorno e superfície de resposta para a rugosidade <i>Ra</i> . . .	57
Figura 6.4 –Gráficos de contorno para as rugosidades <i>Ry</i> , <i>Rq</i> , <i>Rz</i> e <i>Rt</i>	58
Figura 6.5 –Ilustração do processo de análise espectral	59
Figura 6.6 –Comparação das características do sinal acústico nos domínios do tempo e da frequência (espectrograma).	60
Figura 6.7 –Som gravado durante usinagem do aço ABNT 52100	61
Figura 6.8 –Exemplo do espectrograma obtido para o som gravado durante usina- gem do aço ABNT 52100	61
Figura 6.9 –Espectrogramas de sinais acústicos gerados por diferentes configurações de usinagem.	62
Figura 6.10 –Gráfico de avaliação do resultado da análise de componentes principais.	64
Figura 6.11 –Gráfico de linha comparando PC_1 e taxa de remoção de material <i>TRM</i>	65
Figura 6.12 –Gráfico de linha comparando PC_3 e velocidade de corte <i>Vc</i>	66

Figura 6.13 –Gráfico de linha comparando PC_1 e as medidas de rugosidade	66
Figura 6.14 –Gráfico de dispersão relacionando PC_1 e a máxima rugosidade Ry	67
Figura 6.15 –Ilustração do processo de extração dos coeficientes Mel-Cepstrais	68
Figura 6.16 –Gráfico de linha comparando c_2 e as medidas de rugosidade	69
Figura 6.17 –Gráfico de linha comparando $-c_2$ e a medidas de rugosidade Rt	70
Figura 6.18 –Gráfico de contorno para as rugosidades Ra e Rt e o coeficiente Mel-Cepstral c_2	71
Figura 7.1 –Acoplamento Ward baseado na Distância Euclideana para Ra e TRM	74
Figura 7.2 –Diagrama geral de um sistema de monitoramento baseado em sinais acústicos	76
Figura 7.3 –Fluxo de treinamento do sistema baseado em GMM	77
Figura 7.4 –Fluxo de diagnóstico do sistema baseado em GMM	78
Figura 7.5 –Gráfico com os efeitos principais dos fatores sobre a precisão do sistema baseado em GMM	80
Figura 7.6 –Fluxo de treinamento do sistema baseado em SVM para o método OAA	82
Figura 7.7 –Fluxo de treinamento do sistema baseado em SVM para o método OAO	82
Figura 7.8 –Fluxo de diagnóstico do sistema baseado em SVM	83
Figura 7.9 –Gráfico com os efeitos principais dos fatores sobre a precisão do sistema baseado em SVM	85
Figura 7.10 –Gráfico de Pareto dos efeitos padronizados para os termos do modelo.	87
Figura 7.11 –Teste de normalidade dos resíduos padronizados.	87
Figura 7.12 –Gráfico com os efeitos principais dos fatores sobre a precisão dos sistema baseados em GMM e SVM	88
Figura 7.13 –Gráfico com a resposta da otimização para o modelo obtido.	88

LISTA DE TABELAS

Tabela 3.1 –Frequências centrais e bandas críticas para o banco de filtros na escala <i>mel</i>	22
Tabela 6.1 –Projeto D-ótimo para o torneamento do aço duro AISI 52100.	55
Tabela 6.2 –Média das medidas de rugosidade para todas as configurações de usinagem.	56
Tabela 6.3 –Modelos Quadráticos Completos para as rugosidades do aço AISI 52100.	56
Tabela 6.4 –Análise dos autovetores da matriz de correlação entre os espectros de potência dos sinais.	63
Tabela 6.5 –Análise de correlação entre escores, configurações de usinagem e medidas de rugosidade.	63
Tabela 6.6 –Análise de correlação entre coeficientes Mel-Cepstrais, configurações de usinagem e medidas de rugosidade.	69
Tabela 6.7 –Coeficientes de regressão e coeficientes de determinação para os modelos de rugosidade e segundo coeficiente <i>mel</i>	72
Tabela 7.1 –Grupos formados a partir da análise hierárquica de <i>cluster</i> - método de Ward.	74
Tabela 7.2 –Conjuntos de treinamento e teste dos métodos de monitoramento.	75
Tabela 7.3 –Resultados do experimento de validação do sistema baseado em GMM.	79
Tabela 7.4 –Resultados do experimento de validação do sistema baseado em SVM.	84
Tabela 7.5 –Resultados do experimento de validação dos sistemas operando com configuração ótima.	86
Tabela A.1 –Medidas de Rugosidade	104
Tabela B.1 –Cinco primeiros escores obtidos pela ACP (passos de 1 ao 4).	106
Tabela B.2 –Cinco primeiros escores obtidos pela ACP (passos de 5 ao 8).	107
Tabela B.3 –Cinco primeiros escores obtidos pela ACP (passos de 9 ao 12).	108
Tabela B.4 –Cinco primeiros escores obtidos pela ACP (passos de 13 ao 15).	109
Tabela C.1 –Seis primeiros coeficientes Mel-Cepstrais (replicas de 1 ao 4).	111
Tabela C.2 –Seis primeiros coeficientes Mel-Cepstrais (replicas de 5 ao 8).	112
Tabela C.3 –Seis primeiros coeficientes Mel-Cepstrais (replicas de 9 ao 12).	113
Tabela C.4 –Seis primeiros coeficientes Mel-Cepstrais (replicas de 13 ao 15).	114

LISTA DE QUADROS

Quadro 2.1 – Exemplo de métodos diretos e indiretos para o monitoramento de processos.	9
Quadro 2.2 – Sumário dos processos específicos estudados nas publicações revisadas.	10
Quadro 2.3 – Sumário dos sinais e seus parâmetros utilizados nas publicações revisadas.	12
Quadro 5.1 – Hipóteses presentes em uma ANOVA para dois fatores.	49
Quadro 5.2 – Fórmulas para ANOVA em dois níveis.	50
Quadro 5.3 – Fórmulas para análise de variância em dois níveis.	50

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
ACP	Análise de Componentes Principais
ANN	<i>Artificial Neural Networks</i>
BW	<i>Bandwidth</i>
CCD	Arranjo Composto Central ou <i>Box-Wilson Central Composite Design</i>
CNC	Comando Numérico Computadorizado
DFT	<i>Discrete Fourier Transform</i>
DSP	<i>Digital Signal Processing</i>
EA	Emissão Acústica
EM	<i>Expectation Maximization</i>
FFT	<i>Fast Fourier Transform</i>
GMM	<i>Gaussian Mixture Model</i>
HMM	<i>Hidden Markov Models</i>
HRC	Rockwell escala C
IMF	<i>Intrinsic Mode Function</i>
LPCC	<i>Linear Predictive Cepstral Coefficient</i>
LR	<i>Logistic Regression</i>
MFCC	<i>Mel-Frequency Cepstral Coefficients</i>
OAA	<i>One-Against-All</i>
OAo	<i>One-Against-One</i>
PCA	<i>Principal Component Analysis</i>
PSD	<i>Power Spectral Density</i>
PSO	<i>Particle Swarm Optimization</i>
RBF	<i>Radial Basis Function</i>

RMS	<i>Root Mean Square</i>
RNA	Redes Neurais Artificiais
SOM	Redes Auto-Organizáveis (<i>Self-Organizing Maps</i>)
SPAM	Mensagem indesejada
SSA	<i>Singular Spectral Analysis</i>
SVM	<i>Support Vector Machine</i>
WPD	<i>Wavelet Packet Decomposition</i>

LISTA DE SÍMBOLOS

Hz	Unidade para frequência (Hertz)
MHz	Mega (10^6) Hertz
kHz	Kilo (10^3) Hertz
ms	Mili (10^{-3}) segundos
kW	Kilo (10^3) Watt
π	Letra grega que representa o valor 3,141516
ln	Logaritmo natural ou neperiano
AP	Profundidade de corte
F	Taxa de avanço
f_s	Frequência de amostragem
G_i	Representa a i -ésimo grupo (cluster) de rugosidade
MS_i	Representa a i -ésima configuração de usinagem
PC_i	Representa o i -ésimo componente principal
c_i	Representa o i -ésimo coeficiente mel-cepstral
VC	Velocidade de corte
R_a	Rugosidade média aritmética
R_q	Rugosidade média quadrática
R_t	Rugosidade total
R_y	Rugosidade máxima
R_z	Rugosidade média
rpm	Rotações por minuto
TRM	Taxa de remoção de material

SUMÁRIO

Agradecimentos	v
Resumo	vii
Abstract	viii
Lista de Ilustrações	viii
Lista de Tabelas	x
Lista de Quadros	xi
Lista de Abreviaturas e Siglas	xiv
Lista de Símbolos	xv
1 Introdução	1
1.1 Considerações Iniciais	1
1.2 Justificativa e Relevância	3
1.3 Objetivos	3
1.3.1 Objetivo Geral	3
1.3.2 Objetivos Específicos	3
1.4 Delimitação do Trabalho	4
1.5 Caráter Inédito	4
1.6 Metodologia	5
1.7 Organização do Texto	6
2 Métodos de Monitoramento Aplicados a Processos de Usinagem	7
2.1 Considerações Iniciais	7
2.2 Métodos de Monitoramento de Processos	7
2.2.1 Principais Sinais Aplicados ao Monitoramento de Processos	10
2.2.2 Principais Parâmetros Utilizados no Monitoramento de Processos	13
2.2.3 Principais Técnicas de Classificação Aplicados ao Monitoramento de Processos	14
2.3 Considerações Finais Sobre o Capítulo	16
3 Análise Espectral e Extração de Parâmetros	17
3.1 Considerações Iniciais	17
3.2 Análise Espectral	17

3.2.1	Janelamento	18
3.2.2	Obtenção do Espectro de Frequência	19
3.3	Parametrização	20
3.3.1	Coefficientes Mel-Cepstrais	21
3.3.2	Análise dos Componentes Principais	23
3.3.2.1	Formulação Baseada na Maximização da Variância	25
3.3.2.2	Aplicação de ACP para Pré-Processamento de Dados	26
3.4	Considerações Finais Sobre o Capítulo	27
4	Algoritmos e Técnicas de Classificação	28
4.1	Considerações Iniciais	28
4.2	Modelo de Mistura de Gaussianas	28
4.2.1	Treinamento de Modelo de Mistura de Gaussianas	30
4.2.1.1	Algoritmo K -means	31
4.2.2	Classificação Utilizando Modelo de Mistura de Gaussianas	33
4.3	Máquina de Vetores de Suporte	33
4.3.1	Máquina de Vetores de Suporte Aplicada a Classes com Distribuições Sobrepostas	37
4.3.2	Máquina de Vetores de Suporte Aplicada a Problemas Multi-Classes	40
4.3.3	Métodos de <i>Kernels</i>	41
4.3.3.1	Funções de Base Radial	42
4.4	Considerações Finais Sobre o Capítulo	43
5	Modelagem Experimental	44
5.1	Considerações Iniciais	44
5.2	Metodologia de Superfície de Resposta	44
5.2.1	Arranjos Experimentais para Superfície de Resposta	46
5.3	Análise de Variância (ANOVA)	47
5.4	Teste Para a Significância do Modelo	51
5.5	Critério de Optimalidade D	51
5.6	Considerações Finais Sobre o Capítulo	52
6	Análise de Correlação e Detecção de Padrões em Sinais Acústicos	53
6.1	Considerações Iniciais	53
6.2	Máquinas, Materiais, Ferramentas e Instrumentos de Medição	53
6.3	Projeto Experimental e Aquisição de Dados	54
6.3.1	Medidas de Rugosidade	55
6.3.2	Aquisição do Som e Análise Espectral	57
6.4	Extração de Parâmetros dos Sinais Acústicos Utilizando ACP	62
6.5	Extração de Parâmetros dos Sinais Acústicos Utilizando MFCC	67

6.6	Considerações Finais Sobre o Capítulo	72
7	Métodos de Monitoramento Indiretos Baseados em Sinais Acústicos	73
7.1	Considerações Iniciais	73
7.2	Definição dos Critérios de Agrupamento	73
7.3	Método de Monitoramento Baseado em Sinais Acústicos	75
7.3.1	Método de Monitoramento Utilizando GMM e Sinais Acústicos	76
7.3.1.1	Procedimento de Treinamento dos Modelos de Mistura de Gaussianas	77
7.3.1.2	Procedimento de Diagnóstico Utilizando Modelos de Mistura de Gaussianas	78
7.3.1.3	Resultados Obtidos para o Sistema Baseado em GMM	78
7.3.2	Método de Monitoramento Utilizando SVM e Sinais Acústicos	81
7.3.2.1	Procedimento de Treinamento das Máquinas de Vetores de Suporte	81
7.3.2.2	Procedimento de Diagnóstico Utilizando Máquina de Vetores de Suporte	83
7.3.2.3	Resultados Obtidos para o Sistema Baseado em SVM	83
7.4	Seleção das Configurações Ótimas do Sistema de Monitoramento Baseado em Sinais Acústicos	85
7.5	Considerações Finais	89
8	Conclusão	90
	Conclusão	90
	Referências	92
	Apêndices	102
	APÊNDICE A Tabela Completa das Medidas de Rugosidade.	103
	APÊNDICE B Escores Obtidos Através da Análise de Componentes Principais	105
	APÊNDICE C Tabela com os Coeficientes Mel-Cepstrais Extraídos dos Sinais	110
	APÊNDICE D Códigos para Implementação da Extração de MFCC	115
	APÊNDICE E Códigos para Implementação do Classificador baseado em GMM	122
	APÊNDICE F Códigos para Implementação do algoritmo K-means	128

Anexos	139
ANEXO A Pattern recognition in audible sound energy emissions of AISI 52100 hardened steel turning: a MFCC-based approach	140
ANEXO B A mel-frequency cepstral coefficient-based approach for surface roughness diagnosis in hard turning using acoustic signals and gaussian mixture model	151

1 INTRODUÇÃO

1.1 CONSIDERAÇÕES INICIAIS

O processo de torneamento de aço duro tem sido amplamente estudado por apresentar mais benefícios que os processos tradicionais. Um de seus principais benefícios é a eficiência na redução do tempo de processo em cada operação, que só pode ser alcançada com valores adequados dos parâmetros do processo e escolha adequada do material da ferramenta e sua geometria (PAIVA *et al.*, 2012). Após determinar o ponto ótimo de operação do processo, material e geometria da ferramenta, o desafio é conseguir manter o equipamento operando com estes parâmetros ou mesmo identificar quando houver alterações nas características de saída como, por exemplo, a rugosidade da peça usinada. O próprio desgaste da ferramenta provoca alterações na rugosidade, o que a torna um parâmetro importante na identificação do momento adequado para a troca da ferramenta e manutenção do ponto ótimo (BONIFACIO; DINIZ, 1994).

Diferentes métodos de monitoramento têm sido propostos com o objetivo de garantir a operação do processo no ponto ótimo (ZHU *et al.*, 2009; MARINESCU; AXINTE, 2009; RAJA *et al.*, 2012). De acordo com Sick (2002), estes métodos podem ser classificados como diretos (ótico, radioativo, elétrico resistivo, etc.) ou indiretos (emissão acústica, corrente do motor, força de corte (SR.; LISTER, 2000), vibração (XIQING; CHUANGWEN, 2009), etc.) dependendo da forma com que os sinais utilizados para monitoramento são adquiridos. Os trabalhos mais recentes têm concentrado os esforços nos métodos indiretos (BHASKARAN *et al.*, 2012; FADARE *et al.*, 2012), que utilizam meios de medição não intrusivos e que não necessitam que o processo pare para que as medições sejam realizadas, o que possibilita ajustes em tempo real (CHEN *et al.*, 2011).

O foco deste trabalho está nos métodos indiretos por apresentarem um real ganho de produtividade ao permitir que a qualidade do processo seja monitorada sem interferência. Dentro dos métodos indiretos, a emissão acústica (EA) tem tido destaque por ser um sinal com faixa de frequência acima dos demais sinais de vibração da máquina e ruídos do ambiente (XI *et al.*, 2010). Podem-se encontrar diversas propostas de métodos que utilizam emissão acústica para monitoramento de processos de usinagem como retificação, fresamento (ZHU *et al.*, 2009), furação (ANDOH *et al.*, 2008), torneamento (SCHEFFER *et al.*, 2003), entre outros.

Entre todas as fontes de informação utilizadas pelos métodos indiretos, o som gerado durante o processo de usinagem tem atraído recentemente a atenção dos pesquisadores (RUBIO; TETI, 2010; MANNAN *et al.*, 2000; AI *et al.*, 2011). Principalmente porque usar o som produzido para medir a informação sobre o processo de usinagem tem

duas grandes vantagens: a facilidade de instalação (RUBIO; TETI, 2010) e a implementação a um custo menor do que outros sensores (RAJA *et al.*, 2012).

Alguns exemplos de investigações sobre o uso de emissões sonoras para determinar o estado de processos e das estruturas incluem, mas não se limitam ao trabalho Mannan *et al.* (2000), no qual os autores concluíram que o desgaste da ferramenta pode ser monitorado combinando dados sensoriais de um microfone e uma câmera com sensor CCD (do inglês *Charge-Coupled Devices*), através da correlação entre padrões de som e desgaste da ferramenta.

Posteriormente, Salgado e Alonso (2007) propuseram um sistema de monitoramento do desgaste da ferramenta através da combinação da corrente do motor de alimentação e do som. As características extraídas do som, através da análise de espectro singular (SSA do inglês *Singular Spectrum Analysis*), foram utilizadas para correlacionar com o desgaste da ferramenta. Experimentos utilizando aço ABNT 1040 mostraram que a técnica proposta é rápida e confiável. Já o método proposto por Boutros e Liang (2011) utiliza os sinais sonoros, aliados ao HMM discreto (do inglês *Hidden Markov Models*), para detectar e diagnosticar falhas mecânicas em processos de usinagem e máquinas rotativa. Esta técnica detectou corretamente três níveis diferentes de desgaste para a ferramenta de corte: afiada, desgastada ou quebrada. Para os testes de rolamentos, o modelo classificou a gravidade das falhas, em dois rolamentos diferentes, com uma taxa de sucesso superior a 95%.

Mais recentemente, Lu e Wan (2012) analisaram os sinais acústicos de alta frequência (faixa entre 20 e 80 kHz), gerados no processo de micro-fresamento. Um método de monitoramento de desgaste de ferramenta foi proposto em que os resultados indicaram que os sinais sonoros normalizados podem ser potencialmente aplicados em métodos de monitoramento, desde que seja utilizada uma seleção adequada da largura de banda das características do sinal. Além disso, Robben *et al.* (2010) encontraram no som propagado pelo ar, uma valiosa fonte de informações sobre o processo de usinagem contínuo para a moagem de corte de concreto. Os autores destacaram que devido à alta emissão sonora do processo de usinagem, em ambiente controlado, não houve influência total do ruído ambiente sobre o método de monitoramento proposto.

Baseando-se nos métodos indiretos e principalmente nos métodos que utilizam emissão acústica, uma nova abordagem é apresentada neste trabalho: utilizar os sinais acústicos (sons) emitidos durante o processo como método indireto de monitoramento, que ainda são pouco explorados mas que vêm ganhando importância em trabalhos recentes (RUBIO; TETI, 2008; DOWNEY *et al.*, 2014). Este tipo de sinal encontra-se na faixa de frequência audível (20 Hz a 20.000 Hz), o que possibilita a utilização de um microfone como meio de captura, tornando o sistema mais simples e com custo menor por não necessitar de sensores especiais, como acelerômetros ou dinamômetros normalmente utilizados para

sinais de emissão acústica (RAJA *et al.*, 2012).

Portanto, a proposta deste trabalho é capturar e analisar o som emitido durante o processo de torneamento do aço ABNT 52100 endurecido, identificar quais parâmetros melhor representam o fenômeno e correlacioná-los com a rugosidade da peça e com os parâmetros de corte. Dois parâmetros são analisados: a energia do sinal acústico, através dos escores de seus componentes principais, e os coeficientes Mel-Cepstrais (MFCC – do inglês *Mel-Frequency Cepstral Coefficients*). A partir dos parâmetros mais representativos, são avaliados dois métodos indiretos de monitoramento da rugosidade, um utilizando Modelos de Mistura de Gaussianas (GMM – do inglês *Gaussian Mixture Models*) e outro baseado em Máquina de Vetores de Suporte (SVM – do inglês *Support Vector Machine*).

1.2 JUSTIFICATIVA E RELEVÂNCIA

As justificativas para a realização deste trabalho provêm das necessidades enfrentadas pela indústria de manufatura e pela frequente busca por métodos de monitoramento que sejam eficazes e de baixo custo. Técnicas de monitoramento baseadas no som ainda não foram extensivamente estudadas para aplicações em sistemas industriais (SICK, 2002; TETI *et al.*, 2010). Porém, o custo benefício dos microfones utilizados para captura destes sinais, associado à facilidade de seu posicionamento, tem motivado trabalhos recentes nesta área (DOWNEY *et al.*, 2014; RAJA *et al.*, 2012; RUBIO; TETI, 2008). Além disso, novas técnicas de processamento de sinais digitais têm ajudado a mitigar deficiências associadas ao uso de sinais na faixa audível como, por exemplo, os ruídos provenientes de outros processos.

Enquanto a maioria das pesquisas relacionadas ao tema focam no monitoramento do desgaste da ferramenta (FADARE *et al.*, 2012), este trabalho apresenta uma nova abordagem capaz de avaliar a qualidade final do processo através do som, identificando diferentes níveis da rugosidade resultante.

1.3 OBJETIVOS

1.3.1 Objetivo Geral

O objetivo principal deste trabalho é validar a utilização do som, através de suas características, como um sinal a ser aplicado de forma eficiente no monitoramento indireto de processos de usinagem.

1.3.2 Objetivos Específicos

Os objetivos específicos deste trabalho incluem:

- identificar características dos sinais audíveis (sons) que possam ser utilizadas em métodos de monitoramento indiretos para processos de usinagem.
- identificar quais características extraídas do som possuem maior correlação com a rugosidade resultante do processo de torneamento, assim como com os parâmetros de usinagem;
- validar a utilização do som emitido no processo como método indireto de monitoramento da rugosidade através de classificadores;
- comparar o desempenho de classificação da rugosidade para os métodos de monitoramento propostos e características do som utilizadas.

1.4 DELIMITAÇÃO DO TRABALHO

Este trabalho avalia o som emitido durante o processo de usinagem como uma fonte de sinal para o monitoramento de processos. Por se tratar de um sinal na faixa de frequência audível, é extremamente susceptível a ruídos. Portanto, este trabalho se limita a avaliar o som em um ambiente controlado, livre de ruídos provenientes de outras fontes, como por exemplo, outras máquinas ou mesmo ruídos de ambiente.

Este trabalho não considera o desgaste da ferramenta e portanto, todos os experimentos são controlados de forma a garantir que as ferramentas tenham um nível de desgaste que não afete os resultados.

Os métodos propostos são validados utilizando um processo de torneamento em aço endurecido. Foi escolhido este processo devido às suas características de usinagem e por ser bastante controlado. A avaliação dos métodos propostos em outros processos de usinagem não são abordadas neste trabalho, ficando como proposta de trabalhos futuros.

Este trabalho utilizará dois algoritmos de classificação: SVM (*Support Vector Machine*) operando com *kernel* RBF (*Radial Basis Function*) e GMM (*Gaussian Mixture Model*) aplicados à classificação da rugosidade da peça em processo de usinagem por torneamento através do som. Os resultados obtidos não serão extrapolados para outros classificadores além dos incluídos nesta tese, deixando estas avaliações para trabalhos futuros.

1.5 CARÁTER INÉDITO

O presente trabalho pretende inovar em relação ao estado atual da pesquisa sobre o tema nos seguintes aspectos:

- identificação de características do som que sejam altamente correlacionadas com os parâmetros de corte e principalmente com a qualidade do processo representada pela rugosidade da superfície da peça;
- utilização de classificadores que, baseados nas características extraídas do som, sejam capazes de identificar e classificar diferentes níveis de qualidade do processo sem que haja necessidade de intervenção no mesmo.

1.6 METODOLOGIA

A análise da literatura revela que a maioria dos estudos sobre sistemas de monitoramento segue uma sequência comum de atividades que inclui a definição do tipo de sinal a ser utilizado, assim como os parâmetros que serão extraídos do mesmo, a definição de uma estratégia experimental, a coleta de dados, a escolha de uma técnica de classificação (classificador), treinamento do classificador e algum tipo de validação dos resultados.

O método experimental de pesquisa aparece como mais adequado, tendo em vista que é possível, pelo uso de ferramentas computacionais, manter todas as variáveis de projeto sob controle. A metodologia do Projeto de Experimentos (DOE) é utilizada com ênfase nas técnicas de Fatoriais Completos. Os fatores do DOE (as variáveis independentes) foram, na fase de coleta de dados, os parâmetros de usinagem do processo de torneamento, em que o planejamento experimental teve por meta construir uma base de dados capaz de mapear as características do processo através do som emitido durante a usinagem. Em uma segunda etapa, a fase de análise de dados, os fatores do DOE foram os parâmetros dos classificadores, com o objetivo de se obter a configuração ótima de cada sistema de forma a atingir o maior nível de precisão possível. Finalmente, um último projeto experimental foi utilizado visando avaliar a importância relativa de cada parâmetro extraído do sinal acústico e sua interação com o classificador utilizado. A variável dependente foi a medida do desempenho dos sistemas na predição da rugosidade. A grandeza utilizada para medição do desempenho foi a precisão dada em termos percentuais.

Para cada etapa, os resultados obtidos foram analisados e interpretados. A validação do método proposto foi realizada por meio de experimentos de usinagem, que deram origem a conjuntos de treinamento e testes. Os sistemas ótimos obtidos pela aplicação do método foram utilizados para classificação da rugosidade da peça para o conjunto de testes e avaliados por seu desempenho nesta tarefa.

Desta forma, pode-se classificar o trabalho realizado, quanto à sua natureza, como pesquisa aplicada. Em relação ao seu objetivo, o trabalho se classifica como sendo Normativo (BERTRAND *et al.*, 2002), com abordagem quantitativa (CRESWELL; CLARK, 2011). O método utilizado foi o experimental (BRYMAN, 1989).

1.7 ORGANIZAÇÃO DO TEXTO

O trabalho está dividido conforme resumo a seguir. O capítulo 2 apresenta a revisão bibliográfica referente aos temas utilizados no desenvolvimento desta pesquisa, apresentando trabalhos científicos já publicados e que sejam relacionados aos assuntos abordados. No capítulo 3 é apresentado o método de processamento de sinais digitais, que envolve a extração de parâmetros do sinal acústico, com detalhamento dos processos de extração do MFCC e dos escores dos componentes principais a partir da energia do sinal. O capítulo 4 aborda a natureza dos classificadores definidos para este trabalho, o GMM e o SVM, detalhando seu funcionamento e mostrando os algoritmos das etapas de treinamento e classificação. Já o capítulo 5 aborda a teoria de planejamento de experimentos e as principais formas de análises, bem como a metodologia de superfície de resposta. O capítulo 6 detalha os equipamentos utilizados assim como a montagem da base de dados utilizada na avaliação do método proposto, além de apresentar como cada parâmetro do sinal acústico se correlaciona com as características do processo. O capítulo 7 apresenta os sistemas de monitoramento propostos bem como um detalhamento das fases de treinamento e teste. O desempenho destes sistemas é avaliado para um conjunto de testes, detalhando o comportamento de cada um de acordo com o tipo de parâmetro extraído do sinal acústico. Por fim, conclusões e análises do trabalho baseados nos resultados obtidos nos capítulos anteriores são discutidos na Conclusão, onde também serão apresentadas sugestões para trabalhos futuros.

2 MÉTODOS DE MONITORAMENTO APLICADOS A PROCESSOS DE USINAGEM

2.1 CONSIDERAÇÕES INICIAIS

Neste capítulo, parte da fundamentação teórica relativa às técnicas de monitoramento aplicadas a diferentes processos de manufatura é apresentada. Iniciando-se com uma visão geral sobre as diferentes técnicas utilizadas, uma importante separação entre métodos *diretos* e *indiretos*, classificados com base no tipo de sinal aplicado, é identificada. Com foco nos métodos indiretos, procura-se também explorar o aspecto da insuficiência de técnicas que utilizam o sinal acústico (som) como base para o monitoramento da qualidade em processos de usinagem. Além disso, discutem-se as principais técnicas de extração de parâmetros e algoritmos de classificação utilizados de acordo com o tipo de sinal, destacando as preferências entre os autores e identificando técnicas ainda não exploradas.

2.2 MÉTODOS DE MONITORAMENTO DE PROCESSOS

Com o aumento da competitividade, as empresas cada vez mais se preocupam com a rápida mudança das tecnologias e formatos de produtos. Estas rápidas mudanças exigem que os processos de fabricação sejam constantemente redefinidos, buscando redução de custos e aumento de produtividade aliados à qualidade. Desta forma, métodos de monitoramento têm sido estudados e desenvolvidos a fim de identificar alterações, falhas ou desgastes no processo.

De uma forma geral, um método de monitoramento pode ser dividido em quatro etapas principais: processo de usinagem, que indica qual é o processo alvo a ser monitorado (por exemplo, torneamento e fresamento); aquisição de sinais, que especifica o tipo de sinal que será capturado durante o processo assim como o sensor a ser utilizado (por exemplo, acelerômetros para capturar sinais de emissão acústica); extração de parâmetros, que detalha a técnica utilizada para extrair informações do sinal que se relacionem com o objeto a ser monitorado (por exemplo, MFCC e LPCC) e classificação, que define a técnica de aprendizado de máquina utilizada para classificação ou diagnóstico (por exemplo, SVM e ANN). Estas etapas são ilustradas na Figura 2.1.

O Quadro 2.1 traz uma lista com os principais métodos de monitoramento, que podem ser classificados como diretos ou indiretos (SICK, 2002). A maioria dos métodos diretos exige que o processo seja interrompido para que sejam efetuadas medidas em alguma parte específica, como por exemplo, o desgaste da ferramenta ou a rugosidade da peça usinada. Por esta razão, o grande destaque têm sido os métodos de monitoramento

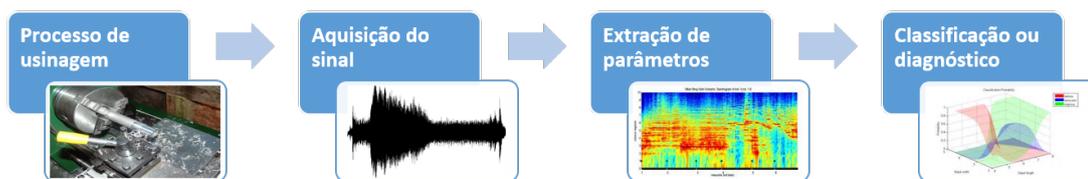


Figura 2.1 – Principais etapas de um processo de monitoramento.

indiretos por não precisarem parar ou interferir no processo, o que aumenta a eficiência do sistema e permite ajustes em tempo real.

A fim de verificar a evolução das pesquisas relacionadas com o monitoramento de processos desde o trabalho de Sick (2002), vários artigos foram selecionados para análise utilizando os seguintes critérios:

- foco no monitoramento de um processo de fabricação ou mecânico;
- utilização de métodos indiretos de monitoramento;
- utilização de alguma técnica para extração de parâmetros de sinais e para modelagem;
- utilização de alguma técnica de aprendizado de máquina ou inteligência artificial para classificação ou diagnóstico.

Com base nos critérios listados, as publicações analisadas podem ser apresentadas de acordo com a distribuição cronológica demonstrada na Figura 2.2. Observa-se que o monitoramento de processos tem atraído o interesse dos pesquisadores, pois é possível notar um aumento considerável no número de publicações sobre o assunto entre 2007 e 2012. De 2012 em diante estão os artigos relacionados à utilização de sinais acústicos na frequência audível como método indireto. A utilização deste tipo de sinal só ganhou destaque recentemente e o pequeno número de trabalhos encontrados demonstra que a utilização do som no monitoramento de processos é uma importante contribuição do presente trabalho.

Os processos específicos investigados ao longo das publicações selecionadas são indicados na Figura 2.3. Dos trabalhos avaliados, treze são baseados no processo de torneamento e nove são relacionados com algum tipo de processo de fresamento. Alguns trabalhos investigam falhas em rolamentos, totalizando quatro artigos, e o restante se dedica ao monitoramento de processos de soldagem, esmerilhamento e engrenagem-rotor. A distribuição dos trabalhos de acordo com os processos pode ser vista no Quadro 2.2. O processo que apareceu por mais vezes entre as publicações revisadas foi o torneamento, responsável por aproximadamente 43% do total. O mesmo foi observado por Sick (2002) e

Quadro 2.1 – Exemplo de métodos diretos e indiretos para o monitoramento de processos.

Desgaste ou parâmetros do processo	Exemplos de procedimentos de medição e transdutores
	<i>Métodos Diretos</i>
Forma ou a posição da aresta de corte ou a área de desgaste	Medição com métodos ópticos (por exemplo, câmara CCD ou de sensores fibra óptica) ou integração de sensores de filme fino no revestimento de uma ferramenta de corte
Perda volumétrica global da ferramenta	Medição do tamanho e concentração de partículas de desgaste no refrigerante (e análise eletroquímica) ou medição da radioatividade (para ferramentas especificamente preparadas)
Alterações da resistência elétrica na junção da ferramenta e da peça	Medição de tensão em um revestimento de ferramenta condutora específica
Alterações nas dimensões da peça	Medição das dimensões através de medidor micrômetro ou óptico, pneumático, ultrassom ou transdutores electromagnéticos
Alterações da distância entre a ferramenta (suporte da ferramenta) e a peça	Medidas das dimensões por meio de medidor micrômetro, pneumático, transdutores de deslocamento (por exemplo, indutiva ou capacitiva), ou sensores ultrassônicos
	<i>Métodos Indiretos</i>
Forças de corte	Medição de força com medidores de tensão ou sensores piezoelétricos na ferramenta ou no porta-ferramentas, força piezoelétrica, medição das placas ou anéis, medição de força nos rolamentos, da medição de torque na rotação principal
Vibrações da ferramenta ou do suporte da ferramenta	Medidas de vibração com acelerômetros
Ruídos (sons) propagados pelo ar ou materiais	Medição da emissão acústica (EA) por meio de transdutores integrados no porta-ferramenta ou acoplado através do líquido de arrefecimento; medição de sinais acústicos (som) com microfones
Corrente elétrica, potência ou energia	Medidas do consumo de corrente ou potência de motores (por exemplo, amperímetro ou dinamômetro)
Temperatura de corte	Medição de temperatura por meio de termopares ou pirômetros, reflectância da superfície da peça ou cavaco
Rugosidade da superfície usinada	Medida com uma caneta mecânica ou métodos ópticos (por exemplo, câmara CCD ou um sensor de fibra óptica)

Fonte: Sick (2002).

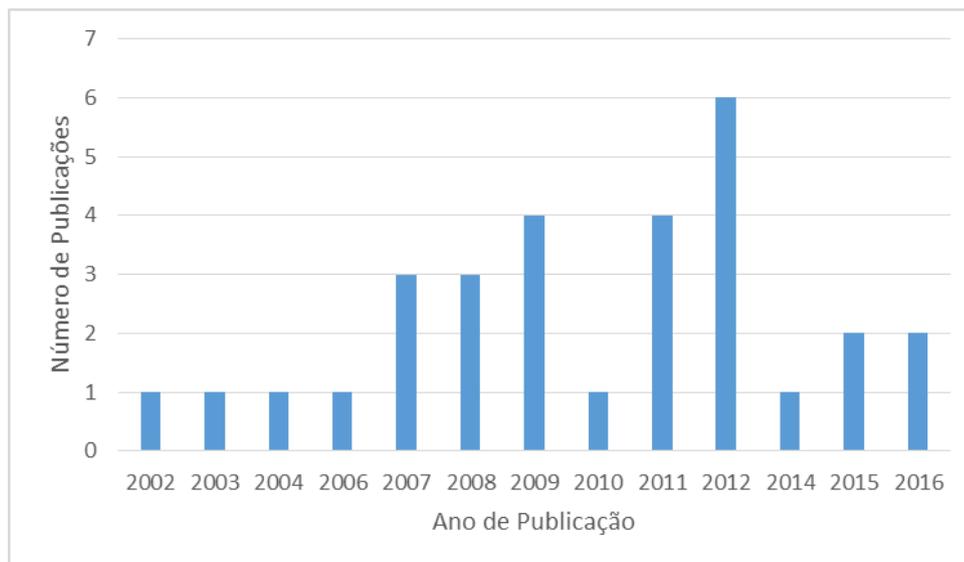


Figura 2.2 – Distribuição anual das publicações revisadas. Adaptado de Frigieri (2013).

Pontes (2011) em seus respectivos trabalhos. Em segundo lugar dentre os processos estudados nos artigos revistos aparece o fresamento, tema de 30% das publicações analisadas.

Quadro 2.2 – Sumário dos processos específicos estudados nas publicações revisadas.

Processo	Referência
Torneamento (corte contínuo)	(BHASKARAN <i>et al.</i> , 2012), (CHEN <i>et al.</i> , 2011), (DOWNEY <i>et al.</i> , 2014), (FADARE <i>et al.</i> , 2012), (GAJATE <i>et al.</i> , 2010), (LIAO <i>et al.</i> , 2016), (LIU; JOLLEY, 2015), (RAJA <i>et al.</i> , 2012), (SALGADO; ALONSO, 2007), (SCHEFFER <i>et al.</i> , 2003), (SHARMA <i>et al.</i> , 2007), (TEKINER; YESILYURT, 2004), (WANG <i>et al.</i> , 2002)
Fresamento (corte intermitente)	(AI <i>et al.</i> , 2011), (BOUTROS; LIANG, 2011), (JEMIELNIAK; ARRAZOLA, 2008), (KHORASANI; YAZDI, 2015), (LU <i>et al.</i> , 2015), (MARINESCU; AXINTE, 2009), (RUBIO; TETI, 2008), (XIQING; CHUANGWEN, 2009), (ZHU <i>et al.</i> , 2009)
Falhas em rolamentos	(BIN <i>et al.</i> , 2012), (HU <i>et al.</i> , 2007), (MARWALA <i>et al.</i> , 2006), (HECKE <i>et al.</i> , 2016)
Soldagem	(CAYO; ALFARO, 2009), (KASBAN <i>et al.</i> , 2011)
Esmerilhamento	(AGUIAR <i>et al.</i> , 2008)
Engrenagem-rotor	(WANG; KANG, 2012)

2.2.1 Principais Sinais Aplicados ao Monitoramento de Processos

Conforme apresentado anteriormente, muitos sinais podem ser utilizados para monitoramento indireto de processos como, por exemplo, os sinais de força de corte (LIAO

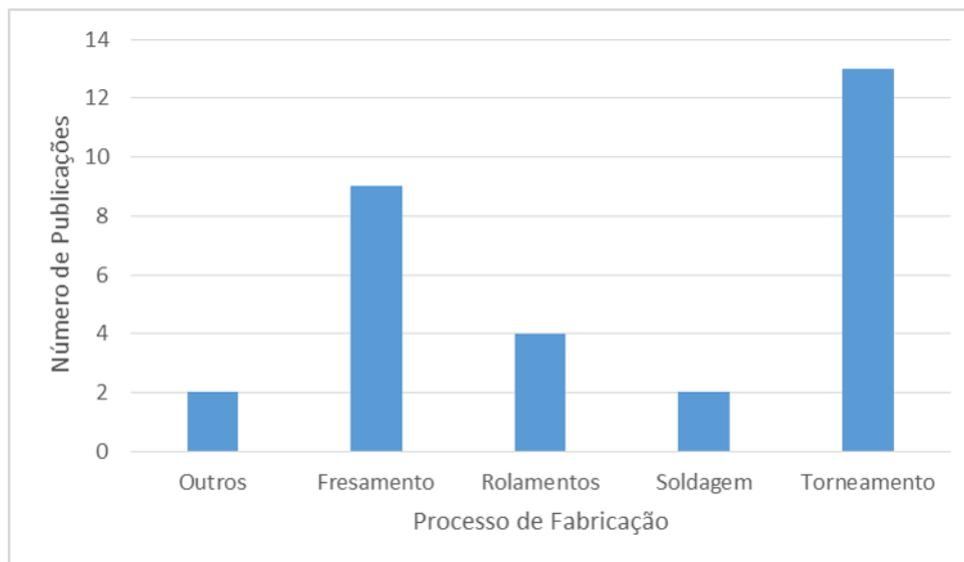


Figura 2.3 – Processos de fabricação e mecânico estudados nas publicações revisadas. Adaptado de Frigieri (2013).

et al., 2016), sinais de vibração mecânica (XIQING; CHUANGWEN, 2009), corrente elétrica, potência, energia, temperatura de corte, imagem (KASBAN *et al.*, 2011), entre outros.

O Quadro 2.3 sintetiza quais sinais e respectivos parâmetros foram adotados pelos autores dos artigos, assim como quais técnicas de classificação foram utilizadas. Foram adotadas algumas legendas para a melhor organização das informações: N/A (não se aplica); EA (Emissão Acústica); DPO (desvio padrão da raiz média quadrática do sinal de emissão acústica, multiplicada pela máxima potência do motor); DPKS (soma da diferença elevada à quarta potência entre a potência de corte e o desvio padrão da potência de corte, multiplicado pelo desvio padrão do sinal de emissão acústica); LPCC (do inglês *Linear Predictive Cepstrum Coefficient*); RMS (representa parâmetros extraídos do sinal no domínio do tempo como média, desvio padrão, assimetria e curtose da distribuição β da raiz média quadrática do sinal); WPD (decomposição de pacotes Wavelet); EMD (do inglês *Empirical Mode Decomposition*); MIV (combinação entre filtragem e K-means que resulta em um vetor de indexação); AIF (do inglês *Acoustic Ignition Frequency*); SPL (nível de pressão do som); WPT (do inglês *Wavelet Package Transform*); MFCC (do inglês *Mel-Frequency Cepstral Coefficients*); MFD (do inglês *Multi-scale Fractal Dimension*); PCA (análise de componentes principais); SSA (análise de espectro singular); CWT (do inglês *Continuous Wavelet Transform*); IMF (do inglês *Intrinsic Mode Function*); DSMA (do inglês *Difference Signal Maximum Amplitude*); PSD (densidade espectral de potência); TFOS (do inglês *Time Frequency Order Spectrum*); RNA (redes neurais artificiais); HMM (cadeias ocultas de Markov); SVM (máquina de vetores de suporte); GMM (modelo de mistura de Gaussianas); HHT (transformada de Hilbert); ANFIS (do inglês *Adaptive Neuro Fuzzy Inference System*); PSO (otimização por colônia de partículas); SOM (redes

auto-organizáveis); LR (regressão logística).

Quadro 2.3 – Sumário dos sinais e seus parâmetros utilizados nas publicações revisadas.

Referência	Sinal	Parâmetros	Classificador
Aguiar <i>et al.</i> (2008)	EA, Potência do Motor	DPO e DPKS	RNA
Ai <i>et al.</i> (2011)	Som	LPCC	N/A
Bhaskaran <i>et al.</i> (2012)	EA	RMS	N/A
Bin <i>et al.</i> (2012)	EA	WPD e EMD	RNA
Boutros e Liang (2011)	Som	MIV	HMM
Cayo e Alfaro (2009)	Som	AIF e SPL	N/A
Chen <i>et al.</i> (2011)	EA, Vibrações	RMS	LR
Downey <i>et al.</i> (2014)	Som	Análise Espectral	N/A
Fadare <i>et al.</i> (2012)	EA	RMS e FFT	N/A
Gajate <i>et al.</i> (2010)	EA, Forças de Corte, Vibração	Normalização	Neuro-Fuzzy
Hu <i>et al.</i> (2007)	EA	WPT	SVM
Jemielniak e Arrazola (2008)	EA, Forças de corte	RMS	N/A
Kasban <i>et al.</i> (2011)	Imagem	MFCC	RNA
Khorasani e Yazdi (2015)	Vibração	RMS	RNA
Liao <i>et al.</i> (2016)	Forças de corte	CWT	HMM
Liu e Jolley (2015)	Forças de corte	RMS	RNA
Lu <i>et al.</i> (2015)	Som	Análise Espectral	HMM
Marinescu e Axinte (2009)	EA	Análise Espectral	N/A
Marwala <i>et al.</i> (2006)	EA	MFD, PCA	HMM, GMM
Raja <i>et al.</i> (2012)	Som	IMF, HHT	N/A
Rubio e Teti (2008)	Som	DSMA	RNA
Salgado e Alonso (2007)	Corrente do motor, Som	SSA	SVM
Scheffer <i>et al.</i> (2003)	Forças de corte, Vibração	PSD	RNA, PSO, SOM
Sharma <i>et al.</i> (2007)	EA, Forças de corte, Vibração	RMS	ANFIS
Tekiner e Yesilyurt (2004)	Som	SPL	N/A
Hecke <i>et al.</i> (2016)	EA	Análise Espectral	N/A
Wang e Kang (2012)	EA	TFOS	RNA
Wang <i>et al.</i> (2002)	EA	WPT	HMM
Xiqing e Chuangwen (2009)	EA	RMS	N/A
Zhu <i>et al.</i> (2009)	Forças de corte	WPD	HMM

A partir do Quadro 2.3 é possível identificar que os sinais acústicos estão entre os mais utilizados. Este tipo de sinal possui diferentes classificações de acordo com a faixa de frequência da vibração. Quando amostrados com frequências na faixa de alguns MHz, são chamados de sinais de emissão acústica (EA) e normalmente são capturados utilizando acelerômetros acoplados a alguma parte do equipamento. A grande vantagem da utilização dos sinais de EA é que a faixa de frequência do sinal está acima das vibrações

da máquina e dos ruídos do ambiente (LI, 2002). A possibilidade de identificar alterações nas características do sinal de emissão acústica a partir de variações nas condições de corte e no desgaste da ferramenta (FADARE *et al.*, 2012) fazem deste sinal uma importante ferramenta para monitorar os parâmetros do processo e identificar possíveis falhas ou alterações na qualidade. Devido a esta característica, este tipo de sinal tem sido muito utilizado para o monitoramento de processos de torneamento.

Já os sinais acústicos na faixa de frequência audível (até 20 kHz) são comumente conhecidos como som e podem ser capturados através de microfones. O trabalho de Grabec *et al.* (1998) destacou que o som também tem sua importância no monitoramento de processos, assim como o sinal de EA, porém apenas recentemente este sinal ganhou importância na comunidade científica. Apesar de ainda pouco explorados, estes sinais apresentam características semelhantes aos sinais de emissão acústica, além de apresentarem como grande vantagem o tipo de sensor (microfone) utilizado para captura do sinal. O principal benefício do microfone está associado com a sua natureza física. Por serem fáceis de montar, podendo ser instalados próximos ao ponto de usinagem, com pouca ou nenhuma interferência na máquina, na ferramenta ou peça a ser usinada (RUBIO; TETI, 2008), apresentam baixo custo quando comparados aos sensores utilizados para aquisição dos sinais de EA (RAJA *et al.*, 2012). Este trabalho utiliza o som como sinal para o método de monitoramento proposto devido à vantagem na aquisição apresentada, além de ser um sinal pouco explorado para monitoramento da qualidade de processos de usinagem.

2.2.2 Principais Parâmetros Utilizados no Monitoramento de Processos

Associada à escolha do sinal está a técnica de processamento de sinais que contempla alguma forma de extração de parâmetros. Entre os principais parâmetros extraídos de sinais no domínio do tempo se destacam: média, variância, curtose, assimetria, etc. (LIU; JOLLEY, 2015; KHORASANI; YAZDI, 2015). Já para sinais no domínio da frequência, destacam-se: análise espectral (MARINESCU; AXINTE, 2009), MFCC (KASBAN *et al.*, 2011), LPCC (AI *et al.*, 2011), entre outras. Outro parâmetro bastante utilizado são os pacotes Wavelet, que estão presentes em pelo menos cinco dos trabalhos revisados (BIN *et al.*, 2012; HU *et al.*, 2007; WANG *et al.*, 2002). A Figura 2.4 mostra a distribuição dos parâmetros nas publicações analisadas. Note que alguns parâmetros foram agrupados dentro de um tópico geral, como é o caso do SSA e do TFOS que foram classificados como *Análise Espectral*, assim como os parâmetros extraídos de sinais no domínio no tempo estão em *RMS*.

Na maioria dos trabalhos analisados que utilizam sinais de vibração, emissão acústica (EA) ou mesmo forças de corte, existe uma predominância das técnicas de extração de parâmetros destes sinais no domínio do tempo. Já nos trabalhos que utilizam o som, existe uma preferência por técnicas de análise espectral, explorando características no

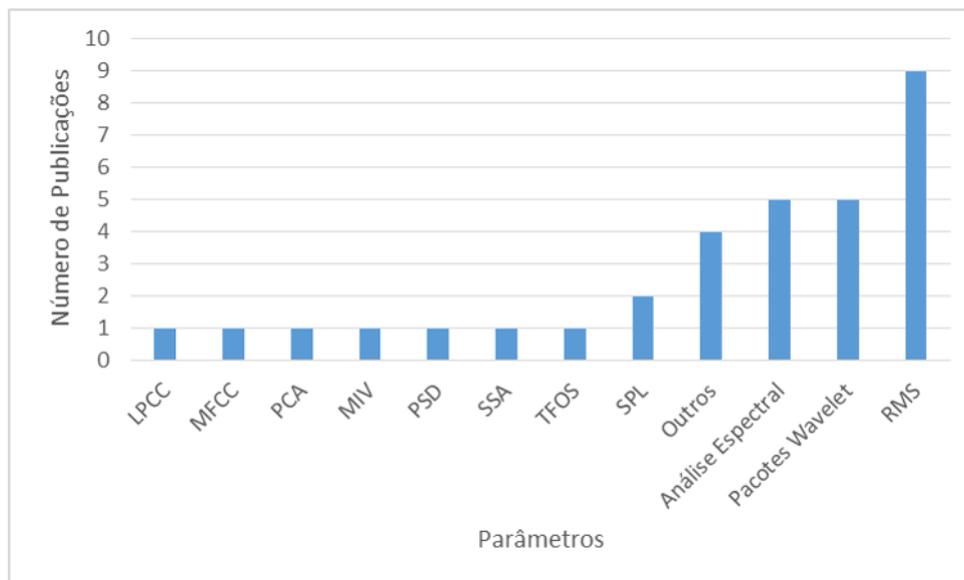


Figura 2.4 – Parâmetros de sinais estudados nas publicações revisadas. Adaptado de Frigieri (2013).

domínio da frequência.

É possível verificar que as técnicas de monitoramento que utilizam o som se assemelham às técnicas de processamento de voz e de sinais de áudio no que tange à extração de parâmetros e modelagem estatística dos sinais (ZHU *et al.*, 2009). Um parâmetro muito utilizado em trabalhos relacionados ao reconhecimento de padrões de voz é o MFCC (do inglês *Mel-Frequency Cepstral Coefficients*) (JUNQIN; JUNJUN, 2011; YUJIN *et al.*, 2010).

O uso do MFCC por Frigieri (2013) apontou fortes indícios de correlação deste parâmetro com a rugosidade resultante do processo de torneamento, mostrando grande potencial para sua utilização para monitoramento de processos de manufatura. Esta forte correlação do MFCC com os parâmetros do processo apresentam indicações de que a própria energia do sinal acústico também tem correlação com o processo, o que motivou sua avaliação neste trabalho através da utilização de Análise de Componentes Principais (ACP).

2.2.3 Principais Técnicas de Classificação Aplicados ao Monitoramento de Processos

Os parâmetros extraídos dos sinais são aplicados no treinamento de modelos estatísticos que, mais tarde, são utilizados para classificação de padrões ou diagnóstico. Dentre os classificadores que se destacam entre os trabalhos avaliados estão as Cadeias Ocultas de Markov (HMM do inglês *Hidden Markov Models*) (BOUTROS; LIANG, 2011; LU *et al.*, 2015) e as Redes Neurais Artificiais (ANN do inglês *Artificial Neural Networks*)

(RUBIO; TETI, 2008; WANG; KANG, 2012), Máquina de Vetores de Suporte (SVM do inglês *Support Vector Machine*) (HU *et al.*, 2007; SALGADO; ALONSO, 2007) e Modelo de Mistura de Gaussianas (GMM do inglês *Gaussian Mixture Models*) (MARWALA *et al.*, 2006), com destaque para os dois primeiros métodos, como pode ser visto na Figura 2.5.

Para sinais em que a informação temporal é importante, a utilização do HMM apresenta bom desempenho na modelagem tanto de sinais dinâmicos como estáticos (BOU-TROS; LIANG, 2011). Por outro lado, quando a informação temporal não é relevante, ganham destaque os classificadores baseados em RNA, SVM ou GMM. No caso do GMM, todo o sinal é tratado com uma entidade singular, representado por uma única propriedade espectral (BROWN; SMARAGDIS, 2009) e desta forma, pode ser visto como um HMM de um único estado. Ao utilizar o GMM ganha-se em tempo de treinamento e classificação além da simplificação do modelo necessário para representar o problema.

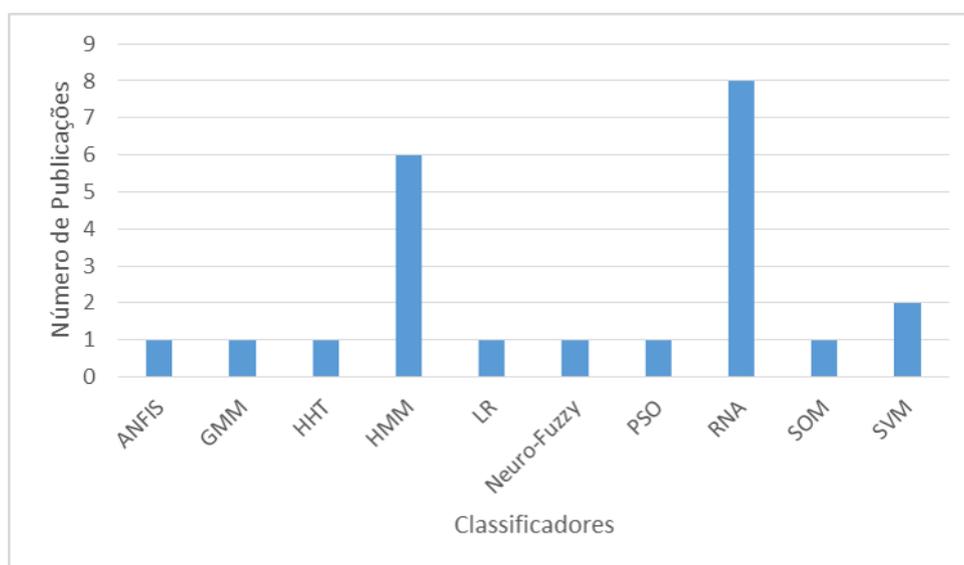


Figura 2.5 – Classificadores estudados nas publicações revisadas. Adaptado de Frigieri (2013).

As redes neurais (RNA) têm sido amplamente utilizadas por apresentar um desempenho excelente para os mais diversos problemas, dos mais simples aos mais complexos. Pode-se encontrar aplicações da RNA para métodos de monitoramento baseados em quase todos os sinais: emissão acústica (BIN *et al.*, 2012), vibração (KHORASANI; YAZDI, 2015), forças de corte (LIU; JOLLEY, 2015), som (RUBIO; TETI, 2008) e até mesmo imagem (KASBAN *et al.*, 2011).

Por outro lado, o SVM é uma ferramenta muito utilizada em outras áreas como sistemas inteligentes para classificação de mensagens indesejáveis (SPAMS) (KAMBLE, 2012) e indexação e classificação de sinais acústicos (DHANALAKSHMI *et al.*, 2009). Por ainda ser pouco aplicado a técnicas de monitoramento de processos de manufatura, sua utilização em novos trabalhos apresenta um forte apelo científico e permite comparações com outros resultados presentes na literatura.

2.3 CONSIDERAÇÕES FINAIS SOBRE O CAPÍTULO

Conforme pode ser constatado na revisão da literatura, a utilização de sinais acústicos (som) para o monitoramento de processos, em especial no monitoramento da qualidade, é ainda pouco explorado. Algumas características que podem ser extraídas deste sinal, como o MFCC e a energia espectral, são muito utilizadas em outras áreas mas ainda não foram aplicadas em monitoramento de processos de usinagem. Análises sobre a utilização destes parâmetros podem contribuir para a proposta de novas técnicas eficientes e com baixo custo de implantação.

3 ANÁLISE ESPECTRAL E EXTRAÇÃO DE PARÂMETROS

3.1 CONSIDERAÇÕES INICIAIS

Neste capítulo, apresenta-se a fundamentação teórica sobre o processo de extração de parâmetros de sinais acústicos, que passa por duas etapas principais: a análise espectral e a parametrização. A análise espectral é responsável pela conversão do sinal do domínio do tempo para o domínio da frequência. Já a etapa de parametrização utiliza técnicas matemáticas para extrair características que sejam capazes de representar o sinal de uma forma compacta e composta apenas de informações significativas que descrevem o processo que o gerou, como são os casos das técnicas MFCC e ACP.

Este capítulo é importante para a compreensão de quais informações estão contidas dentro de um vetor de parâmetros e que podem ajudar no entendimento da correlação entre o sinal e o processo que o gerou.

3.2 ANÁLISE ESPECTRAL

O sinal amostrado e representado no domínio do tempo, muitas vezes, não apresenta informações relevantes capazes de identificar o fenômeno físico que o gerou. Desta forma, faz-se necessário converter o sinal para o domínio da frequência, processo chamado de análise espectral (em inglês: *Spectral Analysis*), a fim de extrair características que consigam maximizar a separação destes sinais em diferentes classes. A análise espectral de sinais acústicos pode ser dividida em duas etapas básicas (KOOLAGUDI *et al.*, 2012): o processo de janelamento, que apresenta uma forma de análise em que o sinal é dividido em pequenos quadros temporais e; a obtenção do espectro de frequência. Uma vez obtida a potência do espectro de cada quadro, eles estão prontos para a etapa de parametrização que é detalhada na Seção 3.3. Uma visão geral deste processo é apresentada na Figura 3.1.

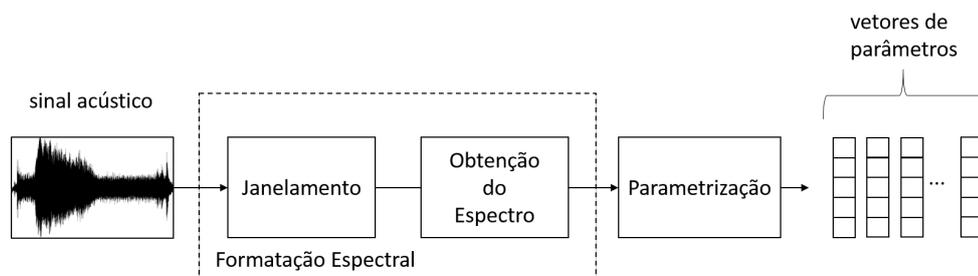


Figura 3.1 – Etapas básicas do processo de extração de parâmetros.

3.2.1 Janelamento

O som proveniente de um processo de usinagem é um sinal dinâmico consequente das variações que ocorrem ao longo do tempo. Estas variações podem acontecer devido a vários fatores, como por exemplo, pequenas diferenças no material usinado, flutuações na velocidade de rotação, entre outros.

Para estas situações, é comum avaliar o sinal considerando pequenos quadros (*frames*) temporais, em que o tamanho do quadro corresponde a N_F amostras e consequentemente a um período de t_F segundos. Este comprimento representa o tempo necessário para que os parâmetros do sinal sejam válidos dentro de um período curto de tempo.

Igualmente importante, no entanto, é o intervalo durante o qual a energia é calculada. O número de amostras usadas para o cálculo é conhecido como N_W e corresponde à duração da janela (em amostras), que normalmente é medido em unidades de tempo (t_W segundos). A duração da janela controla a quantidade média, ou alisamento, usado no cálculo de energia. A duração do quadro e a duração da janela controlam juntas a taxa em que os valores de potência afetam a dinâmica do sinal (RABINER; JUANG, 1993).

De um modo geral, uma duração menor do quadro é usado para capturar dinâmicas rápidas do espectro, assim como a duração da janela deve ser proporcionalmente mais curta de modo que o detalhe no espectro não seja excessivamente alisado. Portanto, escolhe-se o número de amostras para N_W e N_F de tal forma que $N_F \leq N_W$, resultando em uma sobreposição entre quadros adjacentes. A extensão da sobreposição depende de quão rapidamente os parâmetros podem mudar de quadro para quadro, e sua porcentagem correspondente pode ser calculada por:

$$\% \text{ Sobreposição} = \frac{t_w - t_f}{t_w} \times 100 \quad (3.1)$$

No caso de sinais de voz, normalmente utiliza-se uma sobreposição de 50% (PICONE, 1993), conforme exemplificado na Figura 3.2.

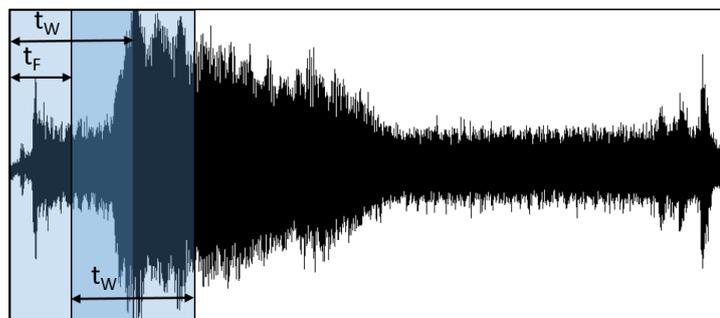


Figura 3.2 – Exemplo de janelamento com sobreposição de 50%.

Cada janela é multiplicada por uma função de janelamento, que tem como objetivo suavizar variações bruscas e aumentar a continuidade entre os quadros adjacentes

(KOOLAGUDI *et al.*, 2012; FAHMY, 2010). Existem muitos tipos de funções de janelamento, incluindo Retangular, Hamming, Hanning, Blackman, Bartlett e Kaiser. Todavia, a função de Hamming é uma das mais utilizadas, pelo seu compromisso entre um lóbulo central estreito e caudas atenuadas (PICONE, 1993) e, por isto, foi escolhida para este trabalho. A função de Hamming é calculada por:

$$w(n) = \frac{\alpha_w - (1 - \alpha_w) \cos(2\pi n / (N_W - 1))}{\beta_w}, 0 \leq n \leq N_W - 1 \quad (3.2)$$

em que N_W representa o número de amostras da janela. O fator α_w é definido como uma janela constante entre o intervalo $[0, 1]$. Normalmente, utiliza-se o valor $\alpha_w = 0,54$ para a Janela de Hamming (DHANALAKSHMI *et al.*, 2009). O fator β_w é uma normalização constante definida de tal forma que a raiz do valor quadrático médio (RMS - *Root Mean Square*) da janela seja unitária. O valor de β_w pode ser obtido da seguinte forma:

$$\beta_w = \sqrt{\frac{1}{N_W} \sum_{n=0}^{N_W-1} w^2(n)}. \quad (3.3)$$

Desta forma, uma janela do sinal de comprimento N_W pode ser ponderado pela janela de Hamming de mesmo comprimento executando uma multiplicação ponto a ponto conforme:

$$\tilde{s}(n) = s(n)w(n), 0 \leq n \leq N_W - 1 \quad (3.4)$$

em que $s(n)$ representa o trecho do sinal sob análise.

O objetivo do janelamento é de favorecer as amostras que se encontram mais ao centro da janela, conforme pode ser visto na Figura 3.3. Esta característica, juntamente com a análise de sobreposição discutida anteriormente, desempenha uma função importante na obtenção das estimativas dos parâmetros com variabilidade suave. É importante que a largura do lobo principal da resposta de frequência da janela seja tão pequena quanto possível, ou o processo de janelamento pode ter um efeito prejudicial sobre a análise espectral subsequente (RABINER; SCHAFER, 2007; RABINER; JUANG, 1993).

3.2.2 Obtenção do Espectro de Frequência

A fim de se obter o espectro de frequência de um sinal $\tilde{s}(n)$, utiliza-se a Transformada Discreta de Fourier (DFT - *Discrete Fourier Transform*), obtida a partir da equação:

$$S(k) = \sum_{n=0}^{N-1} \tilde{s}(n) e^{-j2\pi kn/N}, k = 0, 1, \dots, N - 1 \quad (3.5)$$

na qual N corresponde ao número de amostras. A potência do espectro é obtida calculando-se

$$P(k) = |S(k)|^2, k = 0, 1, \dots, N/2. \quad (3.6)$$

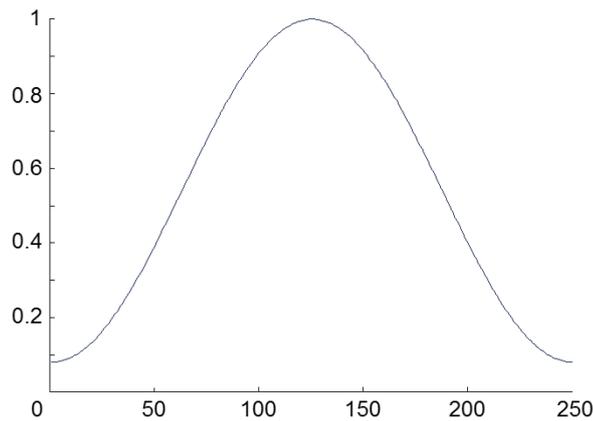


Figura 3.3 – Gráfico da janela Hamming para $N_W = 250$.

Normalmente, consideram-se apenas os termos de 0 a $N/2$, que corresponde à frequência de Nyquist.

A Transformada Rápida de Fourier (do inglês *Fast Fourier Transform* - FFT) (BRIGHAM; MORROW, 1967) também pode ser utilizada como um método alternativo para calcular o espectro do sinal. A FFT é uma implementação da DFT mais eficiente em termos computacionais, sob a restrição de que o espectro deve ser avaliado em um conjunto discreto de frequências, conhecidas como frequências ortogonais, que são múltiplos de f_s/N , em que f_s corresponde à frequência de amostragem. A principal vantagem da FFT é sua eficiência computacional: aproximadamente $N \log N$ adições e $N \log N/2$ multiplicações são requeridas, enquanto a DFT exige um valor proporcional a N^2 operações. A principal desvantagem é que o mapeamento de frequência não é linear e deve ser ajustado para coincidir com as restrições de frequência ortogonal da FFT.

O processo de janelamento associado à extração da potência do espectro de frequência tem como resultado um vetor de valores de potência (ou pares potência-frequência) para cada janela do sinal. Estes são geralmente utilizados para a extração de parâmetros, processo que é discutido na Seção 3.3.

3.3 PARAMETRIZAÇÃO

A parametrização, ou extração de parâmetros, consiste no processo de obter informações de um sinal em um determinado formato, de tal forma que os parâmetros extraídos eliminem a maior quantidade de informação irrelevante possível, mantendo somente informações que sejam significativas.

Dentre as formas de extração de informações (parâmetros), este trabalho tem como foco dois métodos principais: os coeficientes Cepstrais na frequência Mel, também chamados de coeficientes Mel-Cepstrais, que são comumente utilizados para extração de informações de sinais na faixa audível e; a Análise de Componentes Principais, utilizada

para redução da dimensionalidade de vetores e extração de parâmetros (BISHOP, 2006). Ambas as técnicas são detalhadas a seguir.

3.3.1 Coeficientes Mel-Cepstrais

Os coeficientes Mel-Cepstrais (do inglês *Mel-Frequency Cepstral Coefficients* – MFCC) foram propostos inicialmente para a área de reconhecimento de voz (DAVIS; MERMELSTEIN, 1980). Este parâmetro é comumente utilizado em outras áreas de pesquisa mas ainda não foi explorado para aplicação em métodos de monitoramento. Isto faz a utilização deste parâmetro uma grande contribuição deste trabalho. Dentre as principais aplicações que utilizam o MFCC, pode-se destacar o processo de reconhecimento de padrões de voz, que pode ser subdividido em reconhecimento de voz (MAYORGA *et al.*, 2003; POLUR; MILLER, 2006; WANG *et al.*, 2008), reconhecimento de locutor (KAMRUZZAMAN *et al.*, 2010; KUMARI *et al.*, 2012; ZUNJING; ZHIGANG, 2005), identificação de idioma (KOOLAGUDI *et al.*, 2012) e classificação de sinais de áudio (DHANALAKSHMI *et al.*, 2009; DHANALAKSHMI *et al.*, 2011). Também é possível encontrar trabalhos que tratam de diferentes tipos de processos, como identificação de defeitos em soldagem (KASBAN *et al.*, 2011), reconhecimento de palma de mão (FAHMY, 2010) e identificação de sons para fins médicos (CHAUHAN *et al.*, 2008).

Baseado em um banco de filtros, esta técnica pode ser considerada como um modelo dos estágios iniciais de transdução do sistema auditivo humano (PICONE, 1993). As grandes motivações para este tipo de representação são: primeiro, a posição de máximo deslocamento ao longo da membrana basilar por estímulos, tais como tons puros, é proporcional ao logaritmo da frequência do tom; segundo, experimentos sobre percepção humana têm mostrado que as frequências de um som complexo dentro de uma determinada largura de banda, em alguma frequência nominal, não podem ser identificadas individualmente. Em outras palavras, identificações individuais destes componentes só ocorrem quando os mesmos caem fora desta largura de banda, conforme exemplificado na Figura 3.4. Esta largura de banda é conhecida como banda crítica e nominalmente corresponde de 10% a 20% da frequência central.

Baseada nestas informações, a ideia principal desta técnica corresponde ao mapeamento de uma frequência f na escala real (Hz) para uma escala “perceptível”. Dentre as principais funções de mapeamento pode-se destacar a escala Bark, que define o mapeamento através da seguinte equação (PICONE, 1993):

$$f_{Bark} = 13 \arctan \left(\frac{0,76f}{1000} \right) + 3,75 \arctan \left(\frac{f^2}{(7500)^2} \right). \quad (3.7)$$

Outra função popular para este tipo de mapeamento é conhecida como escala *mel*,

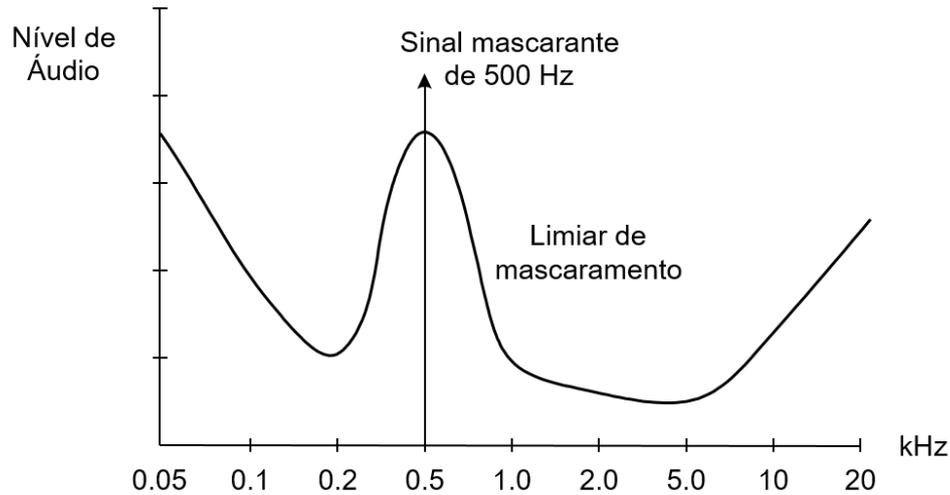


Figura 3.4 – Representação do mascaramento auditivo.

e pode ser obtida através da equação:

$$f_{mel} = 2595 \log_{10} \left(1 + \frac{f}{700} \right). \quad (3.8)$$

Para cada frequência central, a respectiva banda crítica pode ser obtida utilizando a equação:

$$BW_{critica} = 25 + 75 \left(1 + 1,4 \left(\frac{f}{1000} \right)^2 \right)^{0,69}. \quad (3.9)$$

Na escala *mel*, as frequências entre 0 Hz e 1 kHz são aproximadas linearmente e, para frequências acima de 1 kHz, é utilizada uma escala logarítmica em que as frequências são calculadas de acordo com a Equação (3.8). As frequências centrais e respectivas bandas críticas podem ser visualizadas na Tabela 3.1.

Tabela 3.1 – Frequências centrais e bandas críticas para o banco de filtros na escala *mel*.

Indice	Freq. Central (Hz)	BW (Hz)	Indice	Freq. Central (Hz)	BW (Hz)
1	100	100	13	1516	211
2	200	100	14	1741	242
3	300	100	15	2000	278
4	400	100	16	2297	320
5	500	100	17	2639	367
6	600	100	18	3031	422
7	700	100	19	3482	484
8	800	100	20	4000	556
9	900	100	21	4595	639
10	1000	124	22	5278	734
11	1149	160	23	6063	843
12	1320	184	24	6964	969

O banco de filtros na escala Mel pode ser implementado a partir de um conjunto de filtros passa-faixa triangulares, onde o k -ésimo filtro (k corresponde ao índice do filtro) é construído com frequência central $f(k)$, com suas faixas se estendendo da frequência $f(k-1)$ à frequência $f(k+1)$, de modo que $f(k+1)$ não ultrapasse $f_s/2$ (f_s é a frequência de amostragem), conforme ilustrado na Figura 3.5.

Uma vez gerado o banco de filtros, ele pode ser utilizado para a extração dos coeficientes Mel-Cepstrais, de acordo com os seguintes passos. Primeiro, a amplitude do espectro do sinal janelado é escalonado pelo banco de filtros:

$$\tilde{s}(l) = \sum_{n=0}^{N/2} S(n)M_l(n) \quad (3.10)$$

em que M_l representa o l -ésimo filtro na escala *mel* (KASBAN *et al.*, 2011; KUMARI *et al.*, 2012).

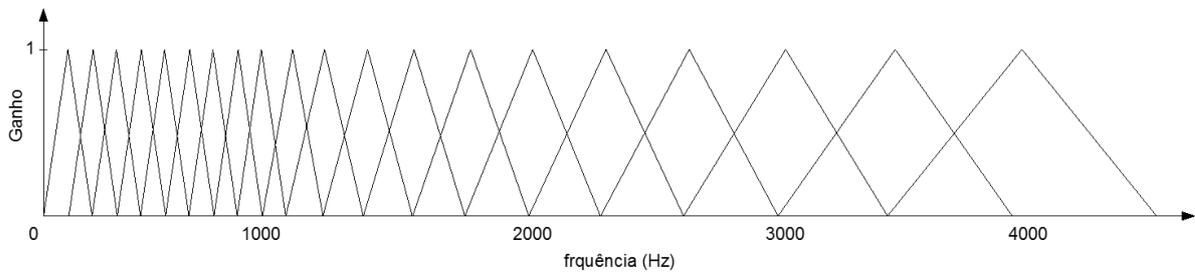


Figura 3.5 – Banco composto por $L = 20$ filtros na escala *mel*.

Na sequência, os espectros escalonados são integrados e compactados através da função logaritmo, para se obter um coeficiente de log-energia para cada banda de frequência. Pode-se dizer que os coeficientes Mel-Cepstrais são as amplitudes resultantes da aplicação da Transformada Inversa de Fourier (do inglês *Inverse Fourier Transform* - IFT) sobre a sequência dos coeficientes de log-energia. No entanto, como o argumento da IFT é uma sequência real e par, o cálculo é geralmente simplificado com a Transformada de Cossenos (VIGNOLO *et al.*, 2011), conforme a equação:

$$c(m) = \sum_{l=0}^{L-1} \log(\tilde{s}(l)) \cos\left(\frac{\pi m}{2L}(2l+1)\right), \forall m = 1, 2, \dots, C \quad (3.11)$$

em que C é o número de coeficientes desejado e L o número total de filtros utilizado.

Como a maior parte da informação do sinal se encontra nos primeiros coeficientes, normalmente se utiliza C entre 12 e 20 (DHANALAKSHMI *et al.*, 2011; FAHMY, 2010).

3.3.2 Análise dos Componentes Principais

A Análise dos Componentes Principais (ACP) é uma técnica estatística largamente utilizada para diferentes tipos de aplicações como, por exemplo, redução de dimensionali-

dade, compressão de dados, extração de parâmetros e visualização de dados (JOLLIFFE, 2002).

Entre alguns trabalhos utilizando esta técnica pode-se destacar o trabalho de Marwala *et al.* (2006) sobre detecção de falhas em rolamentos em que ACP foi utilizada para a redução da dimensão dos parâmetros extraídos do sinal antes de serem aplicados para classificação através de GMM.

Paiva *et al.* (2012) propõem uma abordagem de projeto de parâmetro robusto multivariado para a otimização do processo de torneamento. Nesta proposta, a ACP é utilizada sobre as respostas obtidas através de um projeto experimental para posteriormente serem efetuadas análises sobre os respectivos escores das respostas.

Podem-se encontrar duas definições mais usuais para a ACP, que deram origem ao mesmo algoritmo (BISHOP, 2006). Uma delas foi proposta por Hotelling (1933) e trata a ACP como uma projeção ortogonal dos dados em direção a um espaço linear de menor dimensão, conhecido como *subespaço principal*, de forma que a variância do dado projetado seja maximizada. De forma equivalente, pode-se defini-la como a projeção linear que minimiza o custo médio da projeção, definida como a distância média quadrática entre todos os pontos da projeção (PEARSON, 1901). O processo de projeção ortogonal é ilustrado na Figura 3.6. A linha central, paralela ao indicador u_1 , representa o subespaço principal no qual a projeção (linhas menores perpendiculares) dos pontos para a reta maximizam a variância dos dados. Os pontos sobre esta linha central representam os dados projetados. Outra forma de analisar a mesma figura é através da minimização da soma dos quadrados dos erros das projeções (linhas menores perpendiculares).

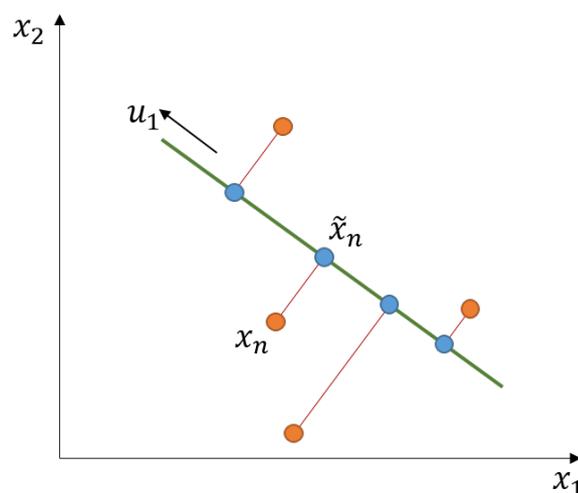


Figura 3.6 – Exemplo do processo de projeção ortogonal para dados com duas variáveis. Fonte: (BISHOP, 2006).

A formulação adotada neste trabalho se baseia na maximização da variância dos dados e será detalhada a seguir.

3.3.2.1 Formulação Baseada na Maximização da Variância

Considere um conjunto de observações $\{\mathbf{x}_n\}$, em que $n = 1, \dots, N$, e \mathbf{x}_n é uma variável Euclidiana com dimensionalidade D . O objetivo é projetar o conjunto de dados na direção de um espaço com dimensionalidade $M < D$ de forma a maximizar a variância dos dados projetados. Para as análises iniciais nesta seção, foi considerado que o valor de M é conhecido e, na sequência, são apresentadas técnicas para sua obtenção.

Em uma análise inicial, pode-se considerar a projeção para um espaço unidimensional $M = 1$. A direção deste espaço pode ser definida através de um vetor D -dimensional denominado \mathbf{u}_1 , que por conveniência (e sem perda de generalidade) pode ser definido como um vetor unitário, de forma que $\mathbf{u}_1^T \mathbf{u}_1 = 1$. Vale ressaltar ainda que o importante é a direção do vetor \mathbf{u}_1 e não sua magnitude. Cada ponto \mathbf{x}_n do conjunto de dados é projetado para um valor escalar $\mathbf{u}_1^T \mathbf{x}_n$. O valor médio de todos os dados projetados pode ser obtido fazendo $\mathbf{u}_1^T \bar{\mathbf{x}}_n$, em que $\bar{\mathbf{x}}_n$ é a média do conjunto de amostras calculado por:

$$\bar{\mathbf{x}}_n = \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n \quad (3.12)$$

e a variância dos dados projetados é dada por:

$$\frac{1}{N} \sum_{n=1}^N \{\mathbf{u}_1^T \mathbf{x}_n - \mathbf{u}_1^T \bar{\mathbf{x}}_n\}^2 = \mathbf{u}_1^T \mathbf{S} \mathbf{u}_1 \quad (3.13)$$

em que \mathbf{S} é a matriz de covariância dos dados obtida através da seguinte equação:

$$\mathbf{S} = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}_n - \bar{\mathbf{x}}_n)(\mathbf{x}_n - \bar{\mathbf{x}}_n)^T. \quad (3.14)$$

Então, basta maximizar a variância projetada $\mathbf{u}_1^T \mathbf{S} \mathbf{u}_1$ em relação a \mathbf{u}_1 . Para evitar que $\|\mathbf{u}_1\| \rightarrow \infty$, deve-se efetuar uma otimização restrita utilizando como restrição a condição de normalização $\mathbf{u}_1^T \mathbf{u}_1 = 1$. É possível reformular o problema utilizando um multiplicador de Lagrange, definido como λ_1 , tornando-o uma otimização sem restrições da forma:

$$\mathbf{u}_1^T \mathbf{S} \mathbf{u}_1 + \lambda_1 (1 - \mathbf{u}_1^T \mathbf{u}_1). \quad (3.15)$$

Fazendo a derivada em relação a \mathbf{u}_1 igual a zero, é possível identificar que ocorrerá um ponto estacionário quando:

$$\mathbf{S} \mathbf{u}_1 = \lambda_1 \mathbf{u}_1 \quad (3.16)$$

o que significa que \mathbf{u}_1 deve ser um autovetor de \mathbf{S} . Multiplicando ambos os lados por \mathbf{u}^T e utilizando a normalização $\mathbf{u}^T \mathbf{u} = 1$, pode-se obter a variância através da equação:

$$\mathbf{u}^T \mathbf{S} \mathbf{u}_1 = \lambda_1 \quad (3.17)$$

e portanto, a variância será máxima quando \mathbf{u}_1 for configurado para ser o autovetor com o maior autovalor λ_1 . Este autovetor é conhecido como o primeiro componente principal.

Componentes principais adicionais podem ser definidos de forma incremental, ou seja, escolhendo cada nova direção de forma que maximize a variância projetada entre todas as possíveis direções ortogonais às direções já definidas. Generalizando para o caso de um espaço de projeção M -dimensional, a projeção linear ótima para a qual a variância é maximizada é definida pelos M autovetores $\mathbf{u}_1, \dots, \mathbf{u}_M$ da matriz de covariância dos dados, \mathbf{S} , correspondendo aos maiores autovalores $\lambda_1, \dots, \lambda_M$.

Resumindo, a análise dos componentes principais envolve a avaliação da média $\bar{\mathbf{x}}_n$ e da matriz de covariância \mathbf{S} do conjunto de dados e, então, a determinar os M autovetores de \mathbf{S} que correspondem aos M maiores autovalores. Algoritmos para obtenção de autovetores e autovalores, assim como teoremas sobre a decomposição de autovetores, podem ser encontrados em [Golub e Van Loan \(1996\)](#). Pode-se dizer que o custo computacional pra se realizar a decomposição de todos os autovetores de uma matriz de tamanho $D \times D$ é $O(D^3)$. Por outro lado, para achar a projeção dos dados sobre apenas os primeiros M componentes principais, é necessário apenas calcular os primeiros M autovalores e autovetores, que pode ser obtido com técnicas mais eficientes resultando em um tempo computacional $O(MD^2)$. De forma geral, as etapas necessárias para o análise ACP pode ser definida como:

1. calcular a matriz de covariância do conjunto de dados;
2. calcular os autovetores e os autovalores da matriz de covariância;
3. reter os componentes principais no qual a explicação acumulada da estrutura variância-covariância seja pelo menos 90%;
4. projetar os dados originais sobre os autovetores reduzidos obtendo os respectivos escores, e conseqüentemente reduzindo a dimensão dos dados.

3.3.2.2 Aplicação de ACP para Pré-Processamento de Dados

Em algumas aplicações o objetivo principal não está apenas na redução da dimensionalidade dos dados mas também em extrair propriedades significativas do conjunto de dados. Isto pode ajudar na aplicação de algoritmos de reconhecimento de padrões (classificação) de forma mais satisfatória. Normalmente, utiliza-se esta técnica quando as variáveis são medidas em diferentes unidades ou possuem variabilidade significativamente diferentes ([BISHOP, 2006](#)).

Em casos como este, pode-se redimensionar as variáveis individualmente, de forma que cada uma tenha média zero e variância unitária. Este processo é conhecido como normalização (do inglês: *standardizing*) no qual a matriz de covariância para dados nor-

malizados tem componentes do tipo

$$\rho_{ij} = \frac{1}{N} \sum_{n=1}^N \frac{(x_{ni} - \bar{x}_i)}{\sigma_i} \frac{(x_{nj} - \bar{x}_j)}{\sigma_j} \quad (3.18)$$

em que σ_i é o desvio padrão de x_i e σ_j é o desvio padrão de x_j . Conhecida como matriz de *correlação* dos dados originais, tem como propriedade apresentar $\rho_{ij} = 1$ quando duas variáveis x_i e x_j são perfeitamente correlacionadas, e $\rho_{ij} = 0$ caso sejam descorrelacionadas.

A normalização de dados pode ser feita de uma forma mais adequada utilizando ACP, resultando em dados com média zero e covariância unitária de forma que as diferentes variáveis se tornem descorrelacionadas. Para isto é preciso escrever a Equação (3.16) para os autovetores na forma:

$$\mathbf{S}\mathbf{U} = \mathbf{U}\mathbf{L} \quad (3.19)$$

no qual \mathbf{L} é uma matriz diagonal com os elementos λ_i e dimensão $D \times D$, e \mathbf{U} é uma matriz ortogonal com colunas formadas pelos elementos \mathbf{u}_i e também de dimensão $D \times D$. Portanto, pode-se definir para cada elemento do conjunto de dados, um valor transformado dado por:

$$\mathbf{y}_n = \mathbf{L}^{-1/2} \mathbf{U}^T (\mathbf{x}_n - \bar{\mathbf{x}}) \quad (3.20)$$

em que $\bar{\mathbf{x}}$ é a média das amostras obtida pela Equação (3.12). Claramente, o conjunto $\{\mathbf{y}_n\}$ possui média 0 e sua matriz de covariância é representada pela matriz identidade, o que pode ser demonstrado através da seguinte equação:

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^N \mathbf{y}_n \mathbf{y}_n^T &= \frac{1}{N} \sum_{n=1}^N \mathbf{L}^{-1/2} \mathbf{U}^T (\mathbf{x}_n - \bar{\mathbf{x}}) (\mathbf{x}_n - \bar{\mathbf{x}})^T \mathbf{U} \mathbf{L}^{-1/2} \\ &= \mathbf{L}^{-1/2} \mathbf{U}^T \mathbf{S} \mathbf{U} \mathbf{L}^{-1/2} = \mathbf{L}^{-1/2} \mathbf{L} \mathbf{L}^{-1/2} = \mathbf{I}. \end{aligned} \quad (3.21)$$

Esta operação é conhecida como *whitening* (BISHOP, 2006).

3.4 CONSIDERAÇÕES FINAIS SOBRE O CAPÍTULO

Neste capítulo, discutiram-se diferentes maneiras de extrair parâmetros de sinais acústicos. Foram apresentadas as etapas necessárias a fim de preparar o sinal para a fase de parametrização, efetuando sua conversão do domínio do tempo para o domínio da frequência. Posteriormente, foram detalhadas as principais técnicas de parametrização apresentando as deduções matemáticas e características de cada um dos métodos propostos: o MFCC e a ACP.

4 ALGORITMOS E TÉCNICAS DE CLASSIFICAÇÃO

4.1 CONSIDERAÇÕES INICIAIS

Neste capítulo, apresenta-se a fundamentação teórica de duas técnicas de aprendizado de máquinas utilizadas para a classificação ou diagnóstico, o GMM e o SVM. Cada técnica é detalhada se destacando os passos necessários para o treinamento (estimação do modelo) e para classificação (diagnóstico).

O intuito principal deste capítulo está no entendimento de como funciona o processo de treinamento destes classificadores, destacando a importância de fornecer informações relevantes sobre o sistema a ser modelado através dos parâmetros de entrada.

4.2 MODELO DE MISTURA DE GAUSSIANAS

O Modelo de Mistura de Gaussianas (do inglês *Gaussian Mixture Model* - GMM) é um método comumente utilizado para estimar a função densidade de probabilidade para sistemas de classificação baseados em estatística. Classificadores baseados em GMM são conhecidos por sua capacidade de modelar distribuições arbitrariamente complexas e por apresentarem grande eficácia para muitas tarefas (BROWN; SMARAGDIS, 2009). Apesar de esta técnica ter sido inicialmente utilizada em pesquisas relacionadas à fala e, principalmente, no reconhecimento de locutor (REYNOLDS; ROSE, 1995), atualmente tem sido aplicada a diferentes sistemas que se baseiam em características extraídas de sinais acústicos (OLIVEIRA *et al.*, 2015; LU; WU, 2013).

A fim de modelar dados provenientes do mundo real, a utilização de somente uma distribuição Gaussiana em geral não é suficiente, fazendo-se necessárias estruturas mais complexas. Por exemplo, a Figura 4.1 apresenta uma distribuição formada pela soma ponderada de três Gaussianas que, de certa forma, consegue modelar os dados de forma mais precisa do que se fosse utilizada apenas uma única Gaussiana. Estas superposições, formadas pela combinação linear de distribuições básicas, neste caso Gaussianas, podem ser formuladas como modelos probabilísticos conhecidos como distribuições de misturas (MCLACHLAN; BASFORD, 1988). Logo, utilizando um número suficiente de Gaussianas e ajustando suas médias e covariâncias assim como os coeficientes da combinação linear, é possível aproximar qualquer densidade contínua com razoável precisão.

Na abordagem utilizando o modelo de mistura de Gaussianas, cada classe formada pelas características do sinal é identificado e modelado por uma função densidade de probabilidade. De acordo com Reynolds e Rose (1995), dado um vetor \mathbf{x} com dimensão

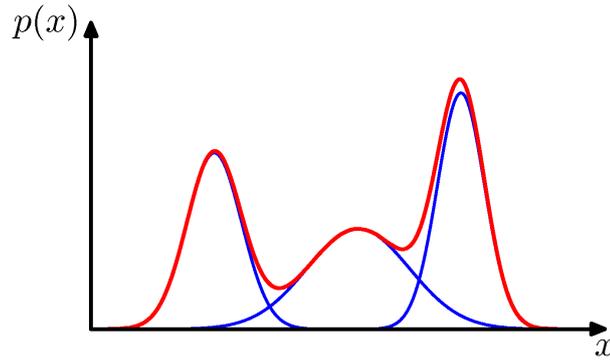


Figura 4.1 – Exemplo de distribuição formada pela combinação linear de três Gaussianas. Fonte: (BISHOP, 2006).

D , a densidade de misturas de Gaussiana pode ser definida conforme a equação

$$p(\mathbf{x} | \lambda) = \sum_{m=1}^M w_m b_m(\mathbf{x}) \quad (4.1)$$

no qual a densidade $p(\mathbf{x}|\lambda)$ é uma combinação linear ponderada de $m = 1, \dots, M$ densidades de Gaussianas, representadas por $b_m(\mathbf{x})$, em que cada densidade é parametrizada por um vetor de média $\boldsymbol{\mu}_m$, uma matriz de covariância $\boldsymbol{\Sigma}_m$ e pesos de mistura w_m , que satisfazem a restrição $\sum_{m=1}^M w_m = 1$ (KUMARI *et al.*, 2012; DHANALAKSHMI *et al.*, 2011). Logo, cada componente $b_m(\mathbf{x})$ da densidade é uma função Gaussiana D -variada na forma

$$b_m(\mathbf{x}) = \frac{1}{(2\pi)^{D/2} |\boldsymbol{\Sigma}_m|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_m)^T \boldsymbol{\Sigma}_m^{-1} (\mathbf{x} - \boldsymbol{\mu}_m) \right\} \quad (4.2)$$

Em geral, a densidade de mistura de Gaussianas completa pode ser parametrizada por $\lambda = (w_m, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m)$, para $m = 1, \dots, M$, em que o número de componentes da mistura, M , é escolhido empiricamente de acordo com o conjunto de dados a ser modelado. Um exemplo desta representação pode ser visto na Figura 4.2.

Existem diversas variantes do GMM, definido como $\lambda = (w_m, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m)$. A matriz de covariância, $\boldsymbol{\Sigma}_m$, pode ser utilizada completa ou restrita pela diagonal principal. Além disso, os parâmetros podem ser compartilhados (*shared*) ou amarrados (*tied*) entre os componentes (Gaussianas), ou seja, pode haver uma matriz de covariância comum para todos os componentes. A escolha do modelo de configuração (número de componentes, matrizes completas ou diagonal) é muitas vezes determinada pela quantidade de dados disponíveis para estimar os parâmetros do GMM e como o modelo é utilizado em uma aplicação particular. É também importante notar que, como os componentes (Gaussianas) atuam em conjunto para modelar a função de densidade global, as matrizes de covariância completas não são necessárias, mesmo que as características não sejam estatisticamente independentes.

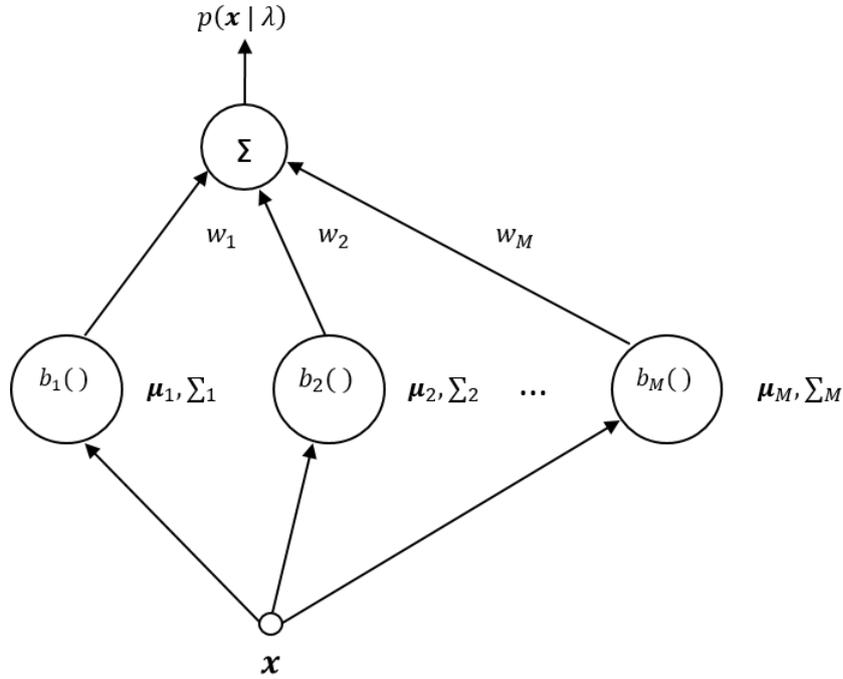


Figura 4.2 – Representação de um modelo de misturas com M Gaussianas.

4.2.1 Treinamento de Modelo de Mistura de Gaussianas

Com o intuito de gerar resultados confiáveis e precisos, classificadores baseados em GMM, assim como acontece com outras técnicas (Redes Neurais, HMM etc.), precisam passar por um procedimento de treinamento, que consiste em modelar um padrão representativo para uma dada classe, antes de serem finalmente utilizados para classificação.

Para estimar os parâmetros do modelo (w_m, μ_m, Σ_m) , pode-se utilizar um método poderoso conhecido como algoritmo *EM* (do inglês: *expectation-maximization*) (DEMPSTER *et al.*, 1977), que tem como objetivo encontrar as soluções que maximizam a verossimilhança dos modelos. Para uma sequência de T vetores de treinamento, $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_T)$, a verossimilhança do GMM pode ser calculada através da equação

$$p(\mathbf{X}|\lambda) = \prod_{t=1}^T p(\mathbf{x}_t|\lambda). \quad (4.3)$$

Conseqüentemente, os parâmetros são ajustados aumentando a verossimilhança do modelo estimado de forma iterativa, ou seja, para iterações k e $k + 1$, $p(\mathbf{X}|\lambda^{k+1}) > p(\mathbf{X}|\lambda^k)$. Em cada iteração, os pesos da mistura, as médias e variâncias são calculadas através das respectivas equações:

$$\bar{w}_m = \frac{1}{T} \sum_{t=1}^T p(m|\mathbf{x}_t, \lambda) \quad (4.4)$$

$$\bar{\boldsymbol{\mu}}_m = \frac{\sum_{t=1}^T p(m|\mathbf{x}_t, \lambda) \mathbf{x}_t}{\sum_{t=1}^T p(m|\mathbf{x}_t, \lambda)} \quad (4.5)$$

$$\bar{\boldsymbol{\sigma}}_m^2 = \frac{\sum_{t=1}^T p(m|\mathbf{x}_t, \lambda) \mathbf{x}_t \mathbf{x}_t^T}{\sum_{t=1}^T p(m|\mathbf{x}_t, \lambda)} - \bar{\boldsymbol{\mu}}_m^2. \quad (4.6)$$

Por fim, a probabilidade de uma determinada classe m pertencer ao modelo λ , após a atualização dos parâmetros, pode ser obtida através da equação:

$$p(m|\mathbf{x}_t, \lambda) = \frac{w_m b_m(\mathbf{x}_t)}{\sum_{k=1}^M w_k b_k(\mathbf{x}_t)}. \quad (4.7)$$

Dois pontos críticos no treinamento de modelos de mistura de gaussianas são a escolha do número M de Gaussianas na mistura e a inicialização dos parâmetros do modelo para iniciar o algoritmo *EM*. O método utilizado para a inicialização dos parâmetros neste trabalho foi o algoritmo *K-means*.

4.2.1.1 Algoritmo *K-means*

O algoritmo *K-means* (LLOYD, 1982) é uma técnica não-probabilística utilizada em problemas que envolvem encontrar agrupamentos (*clusters*) em um conjunto de dados. É importante destacar que esta técnica não utiliza nenhum conhecimento sobre os dados em análise (sem rótulos), o que a classifica como uma abordagem de aprendizado não-supervisionada.

Considerando um problema em que se deseja identificar grupos em conjunto de dados em um espaço multidimensional, para um conjunto de dados $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ composto por N observações de uma variável aleatória Euclidiana \mathbf{x} com dimensão D , o objetivo é particionar o dado em um determinado número de grupos K (inicialmente se considera que K é conhecido).

Intuitivamente, pode-se pensar em um *cluster* como um grupo de amostras no qual a distância entre elas é menor que quando comparada com a distância para amostras que não fazem parte do *cluster*. Pode-se formalizar esta notação primeiro introduzindo um conjunto de vetores $\boldsymbol{\mu}_k$ com dimensão D , em que $k = 1, \dots, K$, no qual $\boldsymbol{\mu}_k$ é um protótipo associado ao k -ésimo *cluster*, ou melhor, uma representação do centro do respectivo *cluster*. Portanto, o objetivo é encontrar uma atribuição de amostras para *clusters*, ou ainda um conjunto de vetores $\{\boldsymbol{\mu}_k\}$, de forma que a soma do quadrado das distâncias de cada amostra para o seu vetor $\boldsymbol{\mu}_k$ mais próximo seja mínimo.

Para representar a atribuição de um ponto (ou amostra) a um determinado *cluster*, pode-se definir para cada amostra \mathbf{x}_n uma variável binária de indicação $r_{nk} \in \{0, 1\}$, em que $k = 1, \dots, K$ descreve a qual dos K *clusters* está atribuída a amostra \mathbf{x}_n . Ou seja, se \mathbf{x}_n está atribuída ao *cluster* k , então $r_{nk} = 1$ e $r_{nj} = 0$ para todo $j \neq k$. Este esquema é conhecido como codificação 1-para- K . Logo, a função objetivo, também conhecida como *medida de distorção*, pode ser definida como

$$J = \sum_{n=1}^N \sum_{k=1}^K r_{nk} \|\mathbf{x}_n - \boldsymbol{\mu}_k\|^2 \quad (4.8)$$

que representa a soma dos quadrados das distâncias entre cada amostra e o vetor $\boldsymbol{\mu}_k$ ao qual está atribuído. O objetivo é encontrar valores para $\{r_{nk}\}$ e $\{\boldsymbol{\mu}_k\}$ que minimizem o valor de J . Isto pode ser feito iterativamente – cada iteração envolve duas etapas sucessivas que correspondem a sucessivas otimizações em relação a r_{nk} e $\boldsymbol{\mu}_k$. Este procedimento pode ser descrito da seguinte forma:

1. Escolher valores iniciais para $\boldsymbol{\mu}_k$;
2. Minimizar J em relação a r_{nk} , mantendo $\boldsymbol{\mu}_k$ fixo;
3. Minimizar J em relação a $\boldsymbol{\mu}_k$, mantendo r_{nk} fixo;
4. Repetir os dois estágios de otimização (etapas 2 e 3) até que haja convergência.

Estes dois estágios responsáveis por atualizar r_{nk} e $\boldsymbol{\mu}_k$ correspondem respectivamente às etapas **E** (do inglês *Expectation*) e **M** (do inglês *Maximization*), do algoritmo **EM**. Para enfatizar esta ideia, os termos passo **E** e passo **M** são utilizados na descrição do algoritmo K -means (BISHOP, 2006).

Pode-se iniciar pela determinação de r_{nk} . Como J na Equação (4.8) é uma função linear com relação a r_{nk} , esta otimização pode ser realizada facilmente para dar uma solução de forma fechada. Os termos envolvendo diferentes valores de n são independentes e portanto podem ser otimizados separadamente escolhendo r_{nk} igual a 1 para qualquer valor de k que minimiza o valor de $\|\mathbf{x}_n - \boldsymbol{\mu}_k\|^2$. Em outras palavras, corresponde a atribuir a mostra n ao *cluster* mais próximo. Formalmente, pode ser expresso por

$$r_{nk} = \begin{cases} 1 & \text{se } k = \arg \min_j \|\mathbf{x}_n - \boldsymbol{\mu}_j\|^2 \\ 0 & \text{caso contrário.} \end{cases} \quad (4.9)$$

O próximo passo é a otimização de $\boldsymbol{\mu}_k$ com r_{nk} fixo. A função objetivo J é uma função quadrática de $\boldsymbol{\mu}_k$, e pode-se minimizá-la fazendo sua derivada em relação $\boldsymbol{\mu}_k$ igual a zero, resultando

$$2 \sum_{n=1}^N r_{nk} (\mathbf{x}_n - \boldsymbol{\mu}_k) = 0 \quad (4.10)$$

que facilmente pode ser resolvido para $\boldsymbol{\mu}_k$

$$\boldsymbol{\mu}_k = \frac{\sum_n r_{nk} \mathbf{x}_n}{\sum_n r_{nk}}. \quad (4.11)$$

O denominador desta expressão representa o número de pontos atribuídos ao *cluster* k e portanto se pode interpretar este resultado como a média de todos os pontos \mathbf{x}_n atribuídos ao *cluster* k . Por isto o nome do algoritmo é K -means (em que *means* significa médias).

O algoritmo K -means é baseado no uso de distâncias Euclidianas ao quadrado como medida de dissimilaridade entre as amostras e os vetores centrais. Isto não apenas limita o tipo de dado a ser considerado (pode ser inapropriado quando algumas ou todas as variáveis representam rótulos de categorias), mas também pode gerar as médias dos *clusters* de forma que não sejam robustas a exceções. Desta forma, outras abordagens baseadas em probabilidade se tornam mais apropriadas, como é o caso dos Modelos de Mistura de Gaussianas (GMM). Neste caso, o algoritmo K -means é utilizado para a inicialização dos parâmetros do GMM antes de aplicar o algoritmo **EM**.

4.2.2 Classificação Utilizando Modelo de Mistura de Gaussianas

Considerando um grupo de S classes diferentes, representadas por $\mathcal{S} = \{1, 2, \dots, S\}$, cujos modelos GMM, $\lambda_1, \lambda_2, \dots, \lambda_S$, foram obtidos através do processo de treinamento baseado no algoritmo *EM*, a classificação é o processo de encontrar o modelo λ com a maior verossimilhança, calculado através da equação:

$$\hat{\mathcal{S}} = \arg \max_{1 \leq k \leq S} p(\mathbf{X} | \lambda_k) \quad (4.12)$$

na qual $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T)$ representa a sequência de observações em análise. Assumindo-se que os vetores (observações) de \mathbf{X} são independentes, pode-se aplicar o logaritmo para substituir o produto entre termos por um somatório a fim de evitar que o produto entre valores muito pequenos resulte em números extremamente pequenos ou nulos. Desta forma, a equação pode ser definida como:

$$\hat{\mathcal{S}} = \arg \max_{1 \leq k \leq S} \sum_{t=1}^T \log p(\mathbf{x}_t | \lambda_k) \quad (4.13)$$

na qual $p(\mathbf{x}_t | \lambda_k)$ é obtido pela Equação (4.1).

4.3 MÁQUINA DE VETORES DE SUPORTE

A Máquina de Vetores de Suporte (do inglês *Support Vector Machine* - SMV) (VAPNIK, 1999) é uma técnica de aprendizado de máquina bastante utilizada no reconhecimento de padrões, e que tem sido aplicada com sucesso em áreas com ênfase em classificação utilizando sinais acústicos (BARKANA; UZKENT, 2011; DHANALAKSHMI

et al., 2009). Baseada na minimização do risco estrutural, tem como objetivo encontrar o classificador que minimiza o erro esperado, maximizando a margem de separação entre as classes (WIDODO; YANG, 2007).

Para entender esta técnica, é preciso recorrer aos problemas de classificação que usam modelos lineares para identificação de duas classes, que podem ser formulados da seguinte maneira:

$$y(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b \quad (4.14)$$

em que $\phi(\mathbf{x})$ representa uma transformação fixa do vetor de características, e b representa o fator de tendenciamento de forma explícita. O conjunto de dados de treinamento é composto por N vetores de entrada, $\mathbf{x}_1, \dots, \mathbf{x}_N$, com os respectivos valores alvo t_1, \dots, t_N sendo $t_n \in \{-1, 1\}$. Para este caso, novos dados \mathbf{x} podem ser classificados de acordo com o sinal de $y(\mathbf{x})$, ou seja, $y(\mathbf{x}) > 0$ para os pontos com $t_n = +1$ e $y(\mathbf{x}) < 0$ para os pontos com $t_n = -1$. Portanto, pode-se dizer que $t_n y(x_n) > 0$ para todos os pontos do conjunto de treinamento.

Vários conceitos podem ser utilizados para encontrar a solução que separe as duas classes corretamente. No caso das máquinas de vetores de suporte, este problema é abordado através do conceito de *margem*, que corresponde à menor distância entre a fronteira de decisão e qualquer uma das amostras. Este conceito é ilustrado na Figura 4.3, em que a margem é definida como a distância perpendicular entre a fronteira de decisão ($y(\mathbf{x}) = 0$) e o ponto mais próximo do conjunto de dados. Maximizar a margem conduz a uma escolha particular da fronteira de decisão, determinada por um subconjunto de dados conhecidos como *vetores de suporte*, ilustrados pelas amostras com círculos na Figura 4.3.

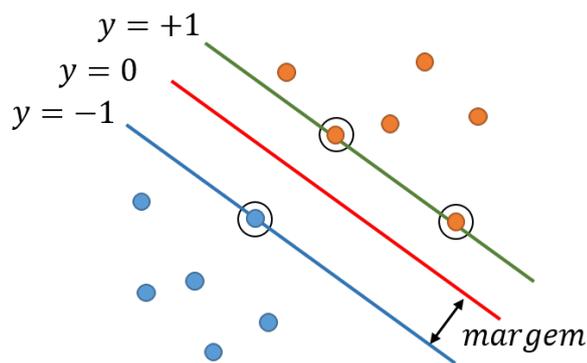


Figura 4.3 – Exemplo do conceito de margem adotado pelas máquinas de vetores de suporte.

A distância perpendicular de um ponto \mathbf{x} em relação a um hiperplano definido por $y(\mathbf{x}) = 0$, obtido através da Equação (4.14), é obtida por $|y(\mathbf{x})|/\|\mathbf{w}\|$ (BISHOP, 2006). Considerando que interessam apenas as soluções em que todos os dados são classificados corretamente, ou seja, $t_n y(x_n) > 0$ para todo n , a solução com a máxima margem pode

ser encontrada resolvendo:

$$\arg \max_{\mathbf{w}, b} \left\{ \frac{1}{\|\mathbf{w}\|} \min_n [t_n(\mathbf{w}^T \phi(x_n) + b)] \right\}. \quad (4.15)$$

O termo $1/\|\mathbf{w}\|$ é deixado fora da otimização porque \mathbf{w} não depende de n . É possível obter soluções diretas para este problema de otimização, porém com alto custo de complexidade, de forma que a melhor opção é convertê-lo para um problema equivalente mais fácil de resolver. Uma forma possível é escalonar $\mathbf{w} \rightarrow k\mathbf{w}$ e $b \rightarrow kb$, de forma que a distância entre qualquer ponto x_n e a superfície de separação $t_n y(x_n)/\|\mathbf{w}\|$ não se altere. Esta alteração permite fazer

$$t_n(\mathbf{w}^T \phi(x_n) + b) = 1 \quad (4.16)$$

para o ponto que está mais próximo à superfície e, neste caso, todos os pontos vão atender às restrições

$$t_n(\mathbf{w}^T \phi(x_n) + b) \geq 1, n = 1, \dots, N. \quad (4.17)$$

Esta modificação é conhecida como representação canônica do hiperplano de decisão. Para os pontos nos quais a igualdade é válida, diz-se que as restrições estão ativas, e para os casos restantes, inativas. Por definição, sempre existirá pelo menos uma restrição ativa porque sempre haverá um ponto mais próximo, e uma vez que a margem seja maximizada, haverá pelo menos duas restrições ativas. Desta forma, o problema de otimização se restringe a maximizar o termo $\|\mathbf{w}\|^{-1}$, que é equivalente a minimizar $\|\mathbf{w}\|^2$, gerando o seguinte problema de otimização:

$$\arg \min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 \quad (4.18)$$

sujeito às restrições em (4.17). O termo $1/2$ é adicionado por conveniência em tratamentos posteriores. Esta nova formulação representa um problema de programação quadrática onde se busca minimizar uma função quadrática com uma série de restrições baseadas em desigualdades lineares. O parâmetro b está implícito na formulação uma vez que alterações em $\|\mathbf{w}\|$ precisam ser compensadas por alterações em b .

Para resolver este problema de otimização com restrições, faz-se necessário utilizar novamente de multiplicadores de Lagrange, utilizados anteriormente na Seção 3.3.2. Neste caso, será introduzido um multiplicador de Lagrange $a_n \geq 0$ para cada restrição em (4.17), resultando na seguinte função Lagrangiana

$$L(\mathbf{w}, b, \mathbf{a}) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{n=1}^N a_n \{t_n(\mathbf{w}^T \phi(x_n) + b) - 1\} \quad (4.19)$$

na qual $\mathbf{a} = (a_1, \dots, a_N)^T$. O sinal de subtração aplicado aos multiplicadores de Lagrange destacam que o problema é minimizado em relação a \mathbf{w} e b porém maximizado em relação a \mathbf{a} . Igualando a zero as derivadas de $L(\mathbf{w}, b, \mathbf{a})$ em relação a \mathbf{w} e b , obtêm-se as seguintes condições

$$\sum_{n=1}^N a_n t_n \phi(x_n) = \mathbf{w} \quad (4.20)$$

$$\sum_{n=1}^N a_n t_n = 0. \quad (4.21)$$

Utilizando estas duas condições, é possível eliminar \mathbf{w} e b de $L(\mathbf{w}, b, \mathbf{a})$, obtendo a representação dual do problema, no qual se maximiza

$$\tilde{L}(\mathbf{a}) = \sum_{n=1}^N a_n - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N a_n a_m t_n t_m k(\mathbf{x}_n, \mathbf{x}_m) \quad (4.22)$$

sujeito às seguintes restrições

$$\sum_{n=1}^N a_n t_n = 0, \quad (4.23)$$

$$a_n \geq 0, n = 1, \dots, N. \quad (4.24)$$

Nesta formulação, é introduzida a ideia de função *kernel*, representada por $k(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{x})^T \phi(\mathbf{x}')$.

As funções de *kernel* permitem estender a utilização de máquinas de vetores de suporte para problemas de classificações não-lineares (XU *et al.*, 2012). De acordo com Vapnik (1995), qualquer função que satisfaça o teorema de Mercer pode ser utilizada como uma função de *kernel* para calcular o produto vetorial no espaço de características. Dentre os vários tipos de funções, podem-se destacar a linear, a polinomial e a Gaussianas RBF (do inglês *Radial Basis Function*).

Finalmente, a fim de classificar novos dados utilizando o modelo treinado, basta avaliar o sinal da seguinte função

$$y(\mathbf{x}) = \sum_{n=1}^N a_n t_n k(\mathbf{x}, \mathbf{x}_n) + b \quad (4.25)$$

gerada ao substituir (4.20) em (4.14).

Avaliando a Equação (4.25), percebe-se que sua aplicação prática fica inviável dependendo do número de pontos N utilizados no processo de treinamento. A fim de viabilizar a utilização prática das máquinas de vetores de suporte, deve-se recorrer às condições de *Karush-Kuhn-Tucker* (KKT), que se aplicam a este tipo de otimização, exigindo que as três seguintes propriedades sejam atendidas

$$a_n \geq 0 \quad (4.26)$$

$$t_n y(x_n) - 1 \geq 0 \quad (4.27)$$

$$a_n \{t_n y(x_n) - 1\} = 0. \quad (4.28)$$

Logo, considerando todas as amostras, ou $a_n = 0$ ou $t_n y(x_n) = 1$. Toda amostra que tenha $a_n = 0$ não irá aparecer na soma da Equação (4.25) e portanto não interfere na predição de novas amostras. Os pontos remanescentes são conhecidos como *vetores de suporte*, que satisfazem a condição $t_n y(x_n) = 1$ e estão situados sobre o hiperplano de margem máxima no espaço de características. Esta propriedade é de extrema importância a fim de permitir a aplicação prática desta técnica. Uma vez que o modelo é treinado, pode-se descartar uma proporção significativa dos dados e manter somente os vetores de suporte.

Uma vez resolvido o problema de programação quadrática e encontrando o valor de \mathbf{a} , pode-se encontrar o valor do parâmetro b através da observação de que cada vetor suporte x_n satisfaz $t_n y(x_n) = 1$. Usando a Equação (4.25) obtém-se

$$t_n \left(\sum_{m \in \mathcal{S}} a_m t_m k(\mathbf{x}_n, \mathbf{x}_m) + b \right) = 1 \quad (4.29)$$

em que \mathcal{S} representa o conjunto de índices dos vetores de suporte. Multiplicando esta equação por t_n , e usando $t_n^2 = 1$, e em seguida obtendo a média em relação a todos os vetores de suporte e resolvendo em relação a b , obtém-se

$$b = \frac{1}{N_{\mathcal{S}}} \sum_{n \in \mathcal{S}} \left(t_n - \sum_{m \in \mathcal{S}} a_m t_m k(\mathbf{x}_n, \mathbf{x}_m) \right) \quad (4.30)$$

em que $N_{\mathcal{S}}$ é o número total de vetores de suporte.

4.3.1 Máquina de Vetores de Suporte Aplicada a Classes com Distribuições Sobrepostas

Considerou-se até agora que a separação entre as classes eram exatas, não havendo sobreposição entre elas. Isto não ocorre nos problemas reais e desta forma, faz-se necessário modificar a formulação do problema para permitir que algumas amostras sejam posicionadas no lado oposto à margem de separação, porém com uma penalização que aumenta de acordo com a distância em relação à margem. Para incluir esta penalidade, podem-se introduzir *variáveis de folga* $\xi_n \geq 0$ para cada amostra do conjunto de treinamento $n = 1, \dots, N$ (VAPNIK, 1995). Estas variáveis podem ser definidas como $\xi_n = 0$ para as amostras que estão sobre ou dentro da margem de separação e $\xi_n = |t_n - y(x_n)|$ para as demais. Ou seja, uma amostra que esteja em cima do limiar de decisão $y(x_n) = 0$ terá $\xi_n = 1$ e amostras com $\xi_n \geq 1$ serão classificados incorretamente, conforme ilustrado na Figura 4.4. Logo, a restrição descrita na Equação (4.17) pode ser reescrita na forma

$$t_n y(x_n) \geq 1 - \xi_n, n = 1, \dots, N. \quad (4.31)$$

no qual cada variável de folga deve satisfazer a condição $\xi_n \geq 0$. Note que apesar das variáveis de folga permitirem sobreposição entre as classes, esta abordagem é sensível a

valores discrepantes (*outliers*) porque a penalidade imposta, devido à classificação incorreta, aumenta linearmente com ξ .

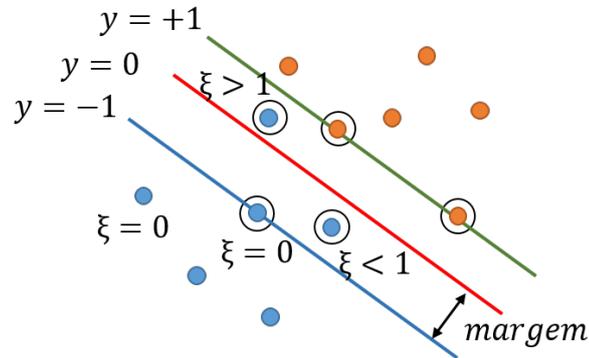


Figura 4.4 – Ilustração das variáveis de folga.

O objetivo agora é maximizar a margem enquanto penalizações suaves são aplicadas aos pontos que estão no lado oposto do limiar da margem. Logo, a equação pode ser escrita como

$$C \sum_{n=1}^N \xi_n + \frac{1}{2} \|\mathbf{w}\|^2 \quad (4.32)$$

em que o parâmetro $C > 0$ controla a troca entre a penalização da variável de folga e a margem. No limite $C \rightarrow \infty$, recorre-se ao problema anterior com dados totalmente separáveis. Como cada ponto classificado incorretamente apresenta $\xi_n > 1$, pode-se dizer que $\sum_n \xi_n$ é um limite superior ao número de pontos classificados incorretamente.

O problema agora é minimizar a Equação (4.32) com as restrições em (4.31) juntamente com $\xi_n \geq 0$. A função Lagrangiana correspondente fica

$$L(\mathbf{w}, b, \mathbf{a}, \boldsymbol{\xi}, \boldsymbol{\mu}) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{n=1}^N \xi_n - \sum_{n=1}^N a_n \{t_n y(x_n) - 1 + \xi_n\} - \sum_{n=1}^N \mu_n \xi_n \quad (4.33)$$

em que $\{a_n \geq 0\}$ e $\{\mu_n \geq 0\}$ são multiplicadores de Lagrange. As correspondentes condições KKT são dadas por

$$a_n \geq 0 \quad (4.34)$$

$$t_n y(x_n) - 1 + \xi_n \geq 0 \quad (4.35)$$

$$a_n \{t_n y(x_n) - 1 + \xi_n\} = 0. \quad (4.36)$$

$$\mu_n \geq 0 \quad (4.37)$$

$$\xi_n \geq 0 \quad (4.38)$$

$$\mu_n \xi_n = 0 \quad (4.39)$$

nas quais $n = 1, \dots, N$.

Podem-se otimizar os termos \mathbf{w} , b e $\{\xi_n\}$ utilizando a definição em (4.14) para $y(x_n)$, sendo:

$$\frac{\partial L}{\partial \mathbf{w}} = 0 \Rightarrow \sum_{n=1}^N a_n t_n \phi(x_n) = \mathbf{w} \quad (4.40)$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{n=1}^N a_n t_n = 0 \quad (4.41)$$

$$\frac{\partial L}{\partial \xi_n} = 0 \Rightarrow a_n = C - \mu_n. \quad (4.42)$$

Utilizando estes resultados, é possível eliminar os termos \mathbf{w} , b e $\{\xi_n\}$ da Lagrangiana, obtendo

$$\tilde{L}(\mathbf{a}) = \sum_{n=1}^N a_n - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N a_n a_m t_n t_m k(\mathbf{x}_n, \mathbf{x}_m) \quad (4.43)$$

que é a mesma função obtida para o caso com separação total, com exceção de algumas restrições. O problema agora é maximizar a Equação (4.43) com relação às variáveis $\{a_n\}$, sujeito às seguintes restrições:

$$0 \leq a_n \leq C \quad (4.44)$$

$$\sum_{n=1}^N a_n t_n = 0 \quad (4.45)$$

para $n = 1, \dots, N$. Assim como anteriormente, esta formulação representa um problema de programação quadrática e substituindo a Equação (4.40) em (4.14), predições para novas amostras podem ser efetuadas utilizando a Equação (4.25). Da mesma forma, um conjunto de pontos deve apresentar $a_n = 0$, não contribuindo para o modelo de predição. Os pontos remanescentes constituem os vetores de suporte tendo $a_n > 0$ que, a partir da restrição (4.36), devem satisfazer

$$t_n y(x_n) = 1 - \xi_n. \quad (4.46)$$

Se $a_n < C$, então $\mu_n > 0$ em (4.42) e $\xi_n = 0$ em (4.39), significando que estes pontos estão sobre a margem. Para os casos em que $a_n = C$, os pontos estão dentro da margem e podem ser classificados corretamente se $\xi_n \leq 1$ ou incorretamente para $\xi_n > 1$.

O parâmetro b pode ser determinado para este caso considerando que os vetores de suporte possuem $0 \leq a_n \leq C$ e $\xi_n \leq 1$ de forma que $t_n y(x_n) = 1$ e portanto irá satisfazer

$$t_n \left(\sum_{m \in \mathcal{S}} a_m t_m k(\mathbf{x}_n, \mathbf{x}_m) + b \right) = 1. \quad (4.47)$$

Novamente pode-se obter uma solução numérica estável através da média, obtendo

$$b = \frac{1}{N_{\mathcal{M}}} \sum_{n \in \mathcal{M}} \left(t_n - \sum_{m \in \mathcal{S}} a_m t_m k(\mathbf{x}_n, \mathbf{x}_m) \right) \quad (4.48)$$

em que \mathcal{M} é conjunto de índices de amostras que possuem $0 \leq a_n \leq C$.

4.3.2 Máquina de Vetores de Suporte Aplicada a Problemas Multi-Classes

A Máquina de Vetores de Suporte foi inicialmente desenvolvida para problemas de classificação envolvendo duas classes (classificação binária). No entanto, problemas do mundo real, que normalmente possuem mais do que duas classes, contribuíram para o desenvolvimento de estratégias para classificação multi-classes. Entre as técnicas mais utilizadas encontram-se a *um-contra-todas* (do inglês *one-against-all* - OAA) e *um-contra-um* (do inglês *one-against-one* - OAO) (WIDODO; YANG, 2007).

No primeiro método, OAA, um modelo SVM é obtido para cada classe $i = 1, \dots, I$. O i -ésimo modelo SVM é treinado fazendo todas as amostras da i -ésima classe rotuladas como $+1$, e todas as demais amostras rotuladas como -1 . Logo, para cada amostra de treinamento x_n , sendo $n = 1, \dots, N$, existe uma classe correspondente $t_n \in \{1, \dots, I\}$. Portanto, para treinar o modelo para a i -ésima classe, basta resolver o seguinte problema de otimização:

$$\frac{1}{2} \|\mathbf{w}^i\|^2 + C \sum_{n=1}^N \xi_n^i \quad (4.49)$$

no qual C corresponde ao parâmetro de penalização. Este problema deve ser minimizado considerando as seguintes restrições:

$$t_n \{(\mathbf{w}^i)^T \phi(x_n) + b^i\} \geq 1 - \xi_n^i \quad \text{if } t_n = i, \quad (4.50)$$

$$t_n \{(\mathbf{w}^i)^T \phi(x_n) + b^i\} \geq -1 + \xi_n^i \quad \text{if } t_n \neq i, \quad (4.51)$$

$$\xi_n^i \geq 0, n = 1, \dots, N. \quad (4.52)$$

Já no caso do método OAO, $I(I-1)/2$ modelos são gerados utilizando dados de treinamento que correspondem a duas classes diferentes. Logo, para treinar cada modelo de ambas i -ésima e j -ésima classes, recorre-se ao seguinte problema de classificação binária:

$$\frac{1}{2} \|\mathbf{w}^{ij}\|^2 + C \sum_{n=1}^N \xi_n^{ij} \quad (4.53)$$

no qual C corresponde ao parâmetro de penalização. Este problema deve ser minimizado considerando as seguintes restrições:

$$t_n \{(\mathbf{w}^{ij})^T \phi(x_n) + b^{ij}\} \geq 1 - \xi_n^{ij} \quad \text{if } t_n = i, \quad (4.54)$$

$$t_n \{(\mathbf{w}^{ij})^T \phi(x_n) + b^{ij}\} \geq -1 + \xi_n^{ij} \quad \text{if } t_n \neq i, \quad (4.55)$$

$$\xi_n^{ij} \geq 0, n = 1, \dots, N. \quad (4.56)$$

Para executar a classificação usando todos os $I(I-1)/2$ classificadores, as decisões podem ser tomadas usando a seguinte estratégia: se $\text{signal}((\mathbf{w}^{ij})^T \phi(\mathbf{x}) + b^{ij})$ indicar que \mathbf{x} pertence à classe i , então um voto é computado para esta classe. Caso contrário, à classe j é incrementada de um. Finalmente, a classe à qual pertence a amostra \mathbf{x} é a que obtiver o maior número de votos. Esta abordagem baseada em votação é chamada de estratégia *Max Win* (WIDODO; YANG, 2007).

4.3.3 Métodos de *Kernels*

O conceito de *kernel* foi introduzido inicialmente por Aizerman *et al.* (1964) em um contexto de método de funções potenciais, assim chamadas devido à analogia com eletrostática. Após alguns anos, foi reintroduzido em aprendizado de máquinas por Boser *et al.* (1992), no contexto de classificadores baseados em margens, dando origem à técnica de Máquinas de Vetores de Suporte (SVM).

Como apresentado pelo método de SVM, vários modelos de parametrização lineares podem ser redefinidos por uma ‘representação-dual’ equivalente na qual as previsões são baseadas em uma combinação linear de uma função de *kernel* avaliada pelos pontos do conjunto de treinamento. Conforme demonstrado na Seção 4.3, modelos baseados em um mapeamento $\phi(\mathbf{x})$ de um espaço de características fixo e não-linear apresentam uma função de *kernel* definida pela seguinte relação

$$k(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{x})^T \phi(\mathbf{x}'). \quad (4.57)$$

Através desta definição, é possível verificar que o *kernel* é uma função simétrica de seus argumentos de forma que $k(\mathbf{x}, \mathbf{x}') = k(\mathbf{x}', \mathbf{x})$.

O exemplo mais simples de uma função de *kernel* é obtido considerando a identidade de mapeamento do espaço de características da forma $\phi(\mathbf{x}) = \mathbf{x}$, que resulta em $k(\mathbf{x}, \mathbf{x}') = \mathbf{x}^T \mathbf{x}'$, conhecido como *kernel* linear. Existem vários outros formatos de funções de *kernel*. Dentre os mais utilizados encontra-se o *kernel* estacionário, que leva em consideração apenas a diferença entre os argumentos, de forma que $k(\mathbf{x}, \mathbf{x}') = k(\mathbf{x} - \mathbf{x}')$. Outra importante especialização são os *kernels* homogêneos, também conhecidos como funções de bases radiais, que dependem apenas da magnitude da distância (tipicamente Euclidiana) entre os argumentos, de forma que $k(\mathbf{x}, \mathbf{x}') = k(\|\mathbf{x} - \mathbf{x}'\|)$.

4.3.3.1 Funções de Base Radial

Como mencionado anteriormente, as funções de bases radiais têm sido largamente aplicadas como função de *kernel*. Inicialmente introduzida para a interpolação de funções exatas (POWELL, 1987), esta função tem a propriedade de cada função base depender apenas da distância radial (tipicamente Euclidiana) de um centro $\boldsymbol{\mu}_j$, de forma que $\phi_j(\mathbf{x}) = h(\|\mathbf{x} - \boldsymbol{\mu}_j\|)$.

Dado um conjunto de dados de vetores de entrada $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ com os respectivos valores alvo $\{t_1, \dots, t_N\}$, o objetivo é encontrar uma função $f(\mathbf{x})$ que ajusta de maneira exata todos os valores alvo, de forma que $f(\mathbf{x}_n) = t_n$, para $n = 1, \dots, N$. Isto pode ser alcançado expressando $f(\mathbf{x})$ como uma combinação linear de funções de base radial, em que cada uma é centrada em cada ponto do conjunto de dados, resultando na expressão

$$f(\mathbf{x}) = \sum_{n=1}^N w_n h(\|\mathbf{x} - \mathbf{x}_n\|). \quad (4.58)$$

Os coeficientes w_n podem ser obtidos através do método dos mínimos quadrados, e por existir o mesmo número de coeficientes e restrições, o resultado é uma função que ajusta todos os valores alvos de forma exata. Porém, em aplicações que envolvem o reconhecimento de padrões, os valores alvos são geralmente ruidosos e a interpolação exata torna-se indesejável uma vez que leva a uma solução *overfitted*.

Outra motivação para as funções de base radial vem da consideração do problema de interpolação quando as variáveis de entrada (em vez dos alvos) são ruidosos (BISHOP, 1995). Se o ruído das variáveis de entrada \mathbf{x} é descrito por uma variável $\boldsymbol{\xi}$ que tem uma distribuição $\nu(\boldsymbol{\xi})$, então a função de erro da soma-de-quadrados pode ser descrita como

$$E = \frac{1}{2} \sum_{n=1}^N \int \{y(\mathbf{x}_n + \boldsymbol{\xi}) - t_n\}^2 \nu(\boldsymbol{\xi}) d\boldsymbol{\xi}. \quad (4.59)$$

Usando o cálculo de variações, pode-se otimizar com relação à função $y(\mathbf{x})$ para obter

$$y(\mathbf{x}) = \sum_{n=1}^N t_n h(\mathbf{x} - \mathbf{x}_n) \quad (4.60)$$

na qual as funções base são dadas por

$$h(\mathbf{x} - \mathbf{x}_n) = \frac{\nu(\mathbf{x} - \mathbf{x}_n)}{\sum_{n=1}^N \nu(\mathbf{x} - \mathbf{x}_n)}. \quad (4.61)$$

É possível notar que existe uma função base para cada ponto do conjunto de dados. Este modelo é conhecido como o modelo de *Nadaraya-Watson*. Quando a distribuição do ruído $\nu(\boldsymbol{\xi})$ é isotrópica, a função depende apenas de $\|\boldsymbol{\xi}\|$ e sua base é radial.

4.4 CONSIDERAÇÕES FINAIS SOBRE O CAPÍTULO

Neste capítulo, discutiram-se detalhes sobre duas diferentes técnicas de classificação: o GMM e o SVM. Foram detalhadas as deduções matemáticas e procedimentos necessários para estimar os modelos usados para representar o sistema de classificação, assim como o procedimento de diagnóstico baseado nestes modelos.

5 MODELAGEM EXPERIMENTAL

5.1 CONSIDERAÇÕES INICIAIS

Neste capítulo, apresenta-se a fundamentação teórica relativa à Metodologia de Projeto de Experimentos, com foco na Metodologia de Superfície de Resposta.

A finalidade deste capítulo é fornecer uma visão geral sobre esta técnica que possibilita a criação e análise de projetos de experimentos capazes de dar consistência estatística aos resultados experimentais.

5.2 METODOLOGIA DE SUPERFÍCIE DE RESPOSTA

De acordo com [MYERS e MONTGOMERY \(1995\)](#), a Metodologia de Superfície de Resposta (MSR) é uma coleção de técnicas matemáticas e estatísticas para modelar e analisar problemas em que a resposta de interesse é influenciada por muitas variáveis. Nestes problemas, a variável de resposta y deve alcançar um valor ótimo, mas sua forma de relacionamento com as variáveis independentes é desconhecida. Portanto, o primeiro passo dentro da metodologia MSR é encontrar uma aproximação razoável para o verdadeiro relacionamento entre y e o conjunto de variáveis, o que usualmente é feito empregando-se um polinômio de baixa ordem.

Supondo que a resposta esperada $E(Y)$ seja função de k variáveis preditoras codificadas x_1, x_2, \dots, x_k , o relacionamento entre y e as variáveis preditoras pode ser expresso segundo uma expansão em Série de Taylor ([BOX; DRAPER, 1987](#)), tal que:

$$E(Y) = \eta = \eta_0 + \sum_{i=1}^k \left[\frac{\partial \eta}{\partial x_i} \right]_0 + \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k \left[\frac{\partial^2 \eta}{\partial x_i \partial x_j} \right] + \dots \quad (5.1)$$

na qual o subscrito zero indica a avaliação na origem.

Se a resposta puder ser modelada de forma satisfatória por uma função linear das variáveis independentes, então a função de aproximação será um modelo de primeira ordem do tipo

$$\eta = \beta_0 + \sum_{i=1}^k \beta_i x_i + \epsilon. \quad (5.2)$$

Se houver curvatura no processo, então um polinômio de ordem mais alta deve ser utilizado, tal como um modelo de segunda ordem, e a expansão pode ser escrita da seguinte forma

$$\eta = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j + \epsilon. \quad (5.3)$$

Segundo [Box e Draper \(1987\)](#), quase todos os problemas de superfície de resposta utilizam um desses dois modelos. Contudo, é improvável que um modelo polinomial seja uma aproximação razoável do modelo real em todo o espaço experimental, mas terá resultado satisfatório para uma determinada região ([MONTGOMERY, 2013](#)). Para estimar os parâmetros β do modelo polinomial, emprega-se o método dos mínimos quadrados, que também pode ser escrito em forma matricial.

A análise de uma superfície de resposta é feita em termos de uma superfície ajustada. Se tal superfície é adequada, sua análise será aproximadamente equivalente à análise da superfície real. Evidentemente, usando-se um arranjo experimental apropriado para coletar os dados de resposta y , os parâmetros do modelo serão estimados de forma eficaz. Geralmente, quando se está em um ponto da região experimental distante do ponto de ótimo, a curvatura do sistema é bem pequena, o que torna um modelo de primeira ordem adequado para modelar a resposta. O objetivo experimental deve ser, portanto, caminhar em direção à região de ótimo. Uma vez encontrada tal região, um modelo quadrático deve ser utilizado. Em algumas circunstâncias, um modelo envolvendo apenas efeitos principais e algumas interações pode ser apropriado para descrever a superfície de resposta, principalmente quando a análise dos resultados revelar a inexistência de curvatura. Em outras condições, a completa descrição do comportamento do processo poderá requerer um modelo quadrático ou cúbico. Se os limites para os fatores forem definidos apropriadamente, será improvável a necessidade de modelos de terceira ordem.

Na maioria dos experimentos, assume-se inicialmente que o modelo linear é adequado. Entretanto, para se confirmar se existe ou não falta de ajuste, através da existência de curvatura, devem-se utilizar pontos experimentais extras denominados de *Pontos Centrais*. De acordo com [Box et al. \(2005\)](#), tal procedimento consiste em adicionar ao fatorial completo, pontos que sejam intermediários aos níveis dos fatores. Serão adicionados tantos pontos centrais quantos forem os k fatores do experimento. Se a resposta média obtida com os n_f pontos fatoriais for significativamente diferente da média formada pelos n_c pontos centrais, então se pode caracterizar uma região de curvatura significativa. Segundo [Montgomery \(2013\)](#), a soma de quadrados para curvatura é dada por

$$SS_{QP} = \frac{(n_f n_c x)(\bar{y}_f - \bar{y}_c)^2}{n_f + n_c}. \quad (5.4)$$

O erro médio quadrático e o teste de hipótese são respectivamente:

$$MS_e = \frac{SS_e}{n_c - 1} = \frac{\sum_{i=1}^{n_c} (y_i - \bar{y})^2}{n_c - 1} \quad (5.5)$$

$$H_0 : \sum_{j=1}^k \beta_{jj} = 0, H_1 : \beta_{jj} \neq 0. \quad (5.6)$$

Se SS_{QP} for muito maior do que MS_e , haverá evidência suficiente para rejeitar a hipótese nula de que os coeficientes quadráticos são nulos.

5.2.1 Arranjos Experimentais para Superfície de Resposta

Existem dois tipos principais de arranjos experimentais para a coleta de dados na Metodologia de Superfície de Resposta: o arranjo central composto e o Box-Behnken.

O arranjo central composto (Box-Wilson *Central Composite Design* - CCD) é uma matriz formada por três grupos distintos de elementos experimentais: um fatorial completo ou fracionado; um conjunto de pontos centrais e, adicionalmente, um grupo de níveis extras denominados *Pontos Axiais*. Se a distância entre o centro do arranjo e ponto fatorial (+1; -1) for aproximadamente 1 (em módulo), a distância do centro ao ponto axial será maior que a unidade. Esta distância, comumente representada por α , depende de certas propriedades desejadas para o experimento e do número de fatores envolvidos (MONTGOMERY, 2013), tal como ilustra a Figura 5.1. O número de pontos axiais em um CCD é igual ao dobro do número de fatores e representam seus valores extremos. Em função de sua localização, podem ser circunscritos, inscritos ou de face centrada.

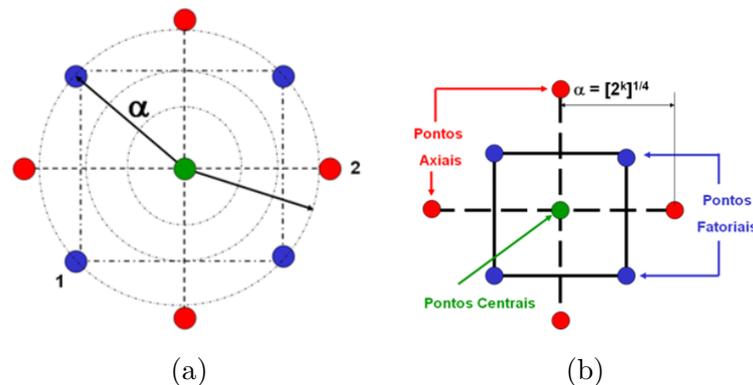


Figura 5.1 – (a) Rotacionalidade; (b) Arranjo central composto para 2 fatores.

Dentre os tipos de arranjos pode-se destacar o Circunscrito (CCC - do inglês *Central Composite Circumscribed*), que corresponde ao CCD original. Nele, os pontos axiais estão a uma distância α do centro, baseado nas propriedades desejadas do projeto. Este arranjo requer cinco níveis para cada fator. Já o Arranjo inscrito (CCI - do inglês *Central Composite Inscribed*) é adequado às situações nas quais os limites especificados não podem ser extrapolados, quer por medida de segurança, quer por incapacidade física de realização. Neste caso, o CCI utiliza os níveis dos fatores como pontos axiais e cria um fatorial completo ou fracionado dentro desses limites, o que requer cinco níveis. O arranjo de Face Centrada (CCF - do inglês *Central Composite Face Centered*) caracteriza-se por dispor os pontos axiais sobre o centro de cada face do espaço fatorial, ou seja, $\alpha = +1$ ou -1 , e requer três níveis para cada fator.

Nos arranjos centrais compostos, a variância predita é constante ao longo da esfera de raio α , ou seja, $V[\hat{y}(x)]_{(1)} = V[\hat{y}(x)]_{(2)} = \dots = V[\hat{y}(x)]_{(i)}$, e é definida como

$$V[\hat{y}(x)] = \sigma^2 x^T (X^T X)^{-1} x. \quad (5.7)$$

Esta propriedade se denomina rotacionalidade. Segundo [Box e Draper \(1987\)](#), para manter a rotacionalidade de um arranjo, o valor de α depende do número de experimentos k da porção fatorial do CCD, tal que $\alpha = (2^k)^{1/4}$. Quando for necessário se empregar um arranjo em blocos, α é dado por:

$$\alpha = \left[k \frac{\left(1 + \frac{n_{s0}}{n_s}\right)}{\left(1 + \frac{n_{c0}}{n_c}\right)} \right]^{\frac{1}{2}} \quad (5.8)$$

em que n_{s0} é o número de pontos centrais presentes na porção axial do arranjo; n_s é a quantidade restante de pontos da porção axial; n_{c0} é o número de pontos centrais da porção cúbica do arranjo (fatorial completo) e n_c é a quantidade restante de pontos da porção cúbica ([BOX; DRAPER, 1987](#)).

Um arranjo do tipo Box-Behnken ([BOX; BEHNKEN, 1960](#)) é um modelo quadrático independente que não contém fatoriais completos ou fracionários embutidos. Neste tipo de arranjo, as combinações de fatores são os pontos médios das arestas das faces do espaço experimental Ω , além dos pontos centrais, como pode ser visto na Figura 5.2. Estes arranjos são rotacionáveis e requerem três níveis para cada fator.

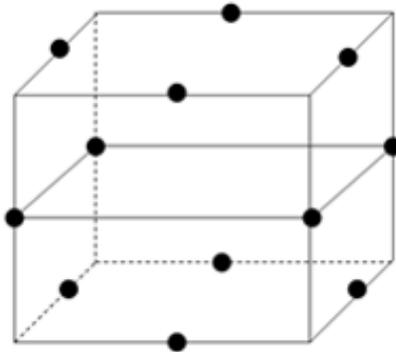


Figura 5.2 – Arranjo do tipo Box-Behnken para três fatores.

5.3 ANÁLISE DE VARIÂNCIA (ANOVA)

Em planejamento de experimentos, para testar se a mudança de níveis de um fator ou interação é significativa se emprega um teste de hipótese para média chamada de Análise de Variância (ANOVA) ([MONTGOMERY; RUNGER, 2011](#)).

Quando dois níveis de um fator geram respostas médias estatisticamente iguais, assume-se que o fator não influencia a resposta de interesse. Por outro lado, quando se

detecta uma diferença significativa, o fator é importante. Para um fatorial completo com de dois fatores, A e B , em dois níveis (+1, -1), pode-se escrever o seguinte modelo:

$$y_{ijk} = \mu + A_i + B_j + AB_{ij} + \varepsilon_{ijk} \quad (5.9)$$

no qual i é número de níveis de A , j é número de níveis de B e k , o número de réplicas. O termo y_{ijk} representa a ijk -ésima observação obtida no experimento, μ é a média geral, A_i é o efeito do i -ésimo tratamento do Fator A , B_j é o efeito do j -ésimo tratamento do Fator B , AB_{ij} é o efeito da ij -ésima interação AB entre os fatores e ε_{ijk} é um componente do erro aleatório.

Para verificar a igualdade entre estas respostas médias, a ANOVA utiliza a relação existente entre a variação *dentro* de um tratamento e a variação *entre* tratamentos. A variação *dentro* é resultado dos repetidos ensaios executados com cada tratamento, calculado através da média das réplicas. Quanto mais as réplicas diferirem desta média, maior será a variação *dentro* deste tratamento. Dividindo-se a variação *entre* pela variação *dentro* se obtém a estatística de teste F , que quando comparada a um valor de $F_{\text{crítico}}$, definido segundo o nível de significância e o número de graus de liberdade da variação *dentro*, possibilita a aceitação ou rejeição da hipótese nula de igualdade entre as médias dos níveis dos fatores (MONTGOMERY; RUNGER, 2011). Ou seja, $F > F_{\text{crítico}}$ comprova a adequação do modelo.

Sempre que se faz uma inferência sobre uma população a partir de dados amostrais, corre-se o risco de cometer algum tipo de erro. Considerando que uma hipótese pode ser definida como uma afirmação que se deseja testar, tendo como fonte de provas os dados provenientes da amostra, existem dois tipos de erro (mutuamente exclusivos) que podem ser cometidos em um teste de hipóteses (MONTGOMERY; RUNGER, 2011). O primeiro é o Erro Tipo I, que consiste em se rejeitar a hipótese nula quando ela é verdadeira. A probabilidade de se cometer este tipo de erro é α , denominado nível de significância do teste. O segundo é o Erro Tipo II, que consiste em se aceitar a hipótese nula quando ela for falsa. A probabilidade de se cometer este tipo de erro é β .

A probabilidade de aceitar a hipótese nula quando ela é verdadeira é $1 - \alpha$, e é denominada nível de confiança do teste. Da mesma forma que a probabilidade de rejeitar a hipótese nula quando ela for falsa é $1 - \beta$, chamada de Poder do Teste (MONTGOMERY; RUNGER, 2011). Dada uma amostra, pode-se calcular a estatística de teste que, quando comparada ao valor crítico determinado por α , conduz à aceitação ou rejeição da hipótese nula. Ao conjunto de todos os valores da estatística de teste que levam à rejeição da hipótese nula dá-se o nome de Região Crítica (MONTGOMERY, 2013).

Para avaliar a significância do efeito dos níveis de um tratamento A , bem como de um tratamento B , faz-se necessário testar a hipótese da igualdade entre as respostas médias obtidas com os níveis dos dois fatores. É importante determinar também se existe

interação entre os dois tratamentos. As hipóteses necessárias neste caso estão representadas no Quadro 5.1.

Quadro 5.1 – Hipóteses presentes em uma ANOVA para dois fatores.

Hipóteses	Efeitos Principais	Interações
Hipótese Nula (H_0)	$H_0 : A_1 = A_2$	$H_0 : AB_{ij} = 0$
Hipótese Alternativa (H_1)	$H_1 : A_1 \neq A_2$	$H_1 : AB_{ij} \neq 0$

Para os dois tratamentos, a aceitação de H_0 significa que os efeitos na resposta obtidos com os dois níveis dos fatores A e B são iguais; a hipótese alternativa é que os tratamentos são diferentes. Na interação, a aceitação de H_0 indica que a interação não é significativa. Para se calcularem os elementos de uma ANOVA para dois fatores, sendo o Fator A com a níveis e Fator B com b níveis, para n réplicas, pode-se empregar as fórmulas do Quadro 5.2. Utilizando as somas definidas, a Soma de Quadrados Total (SS_T) pode ser calculada como

$$\begin{aligned} \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_t)^2 &= \\ &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n [(\bar{y}_l - \bar{y}_t) + (\bar{y}_c - \bar{y}_t) + (\bar{y}_{cel} - \bar{y}_l - \bar{y}_c - \bar{y}_t) + (y_{ijk} - \bar{y}_t)]^2 \\ &= bn \sum_{i=1}^a (\bar{y}_l - \bar{y}_t)^2 + an \sum_{j=1}^b (\bar{y}_c - \bar{y}_t)^2 + n \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{cel} - \bar{y}_l - \bar{y}_c - \bar{y}_t)^2 + \\ &\quad + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_t)^2 \end{aligned}$$

Dessa maneira, a Soma de Quadrados Total do arranjo pode ser dividida em somas de quadrados devido aos tratamentos de linha (Fator A) SS_A , aos tratamentos de coluna (Fator B) SS_B , a soma de quadrados devido à interação entre A e B , SS_{AB} , e a soma de quadrados devido ao erro experimental SS_E . Desse modo, pode-se escrever que

$$SS_T = SS_A + SS_B + SS_{AB} + SS_E. \tag{5.10}$$

As equações das somas de quadrados podem ainda ser escritas como:

$$SS_T = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}^2 - \frac{y_{st}^2}{abn} \tag{5.11}$$

$$SS_A = \frac{1}{bn} \sum_{i=1}^a y_{sl}^2 - \frac{y_{st}^2}{abn} \tag{5.12}$$

$$SS_B = \frac{1}{an} \sum_{j=1}^b y_{sc}^2 - \frac{y_{st}^2}{abn}. \tag{5.13}$$

Estas equações podem ser resumidas conforme o Quadro 5.3.

Quadro 5.2 – Fórmulas para ANOVA em dois níveis.

Soma de cada linha:	$y_{sl} = \sum_{j=1}^b \sum_{k=1}^n y_{ijk}$
Soma de cada coluna:	$y_{sc} = \sum_{i=1}^a \sum_{k=1}^n y_{ijk}$
Soma dentro de cada célula:	$y_{scel} = \sum_{k=1}^n y_{ijk}$
Soma Geral:	$y_{st} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}$
Média de cada linha:	$\bar{y}_l = \frac{1}{bn} \sum_{j=1}^b \sum_{k=1}^n y_{ijk}$
Média de cada coluna:	$\bar{y}_c = \frac{1}{an} \sum_{i=1}^a \sum_{k=1}^n y_{ijk}$
Média dentro de cada célula:	$\bar{y}_{cel} = \frac{1}{n} \sum_{k=1}^n y_{ijk}$
Média Geral:	$\bar{y}_t = \frac{1}{abn} \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}$

Fonte: Adaptado de [Montgomery e Runger \(2011\)](#).

Quadro 5.3 – Fórmulas para análise de variância em dois níveis.

Fonte de Variação	Soma de Quadrados	Graus de Liberdade	Média Quadrática	F_0
Fator A	SS_A	$a - 1$	$MS_A = \frac{SS_A}{a - 1}$	$\frac{MS_A}{MS_E}$
Fator B	SS_B	$b - 1$	$MS_B = \frac{SS_B}{b - 1}$	$\frac{MS_B}{MS_E}$
Interação AB	SS_{AB}	$(a - 1)(b - 1)$	$MS_{AB} = \frac{SS_{AB}}{(a - 1)(b - 1)}$	$\frac{MS_{AB}}{MS_E}$
Erro	SS_E	$ab(n - 1)$	$MS_E = \frac{SS_E}{ab(n - 1)}$	
Total	SS_T	$abn - 1$		

Para se obter uma simplificação da soma de quadrados referentes à interação é aconselhável se calcular primeiramente a soma de quadrados parciais, tal que

$$SS_P = \frac{1}{n} \sum_{i=1}^a \sum_{j=1}^b y_{scel}^2 - \frac{y_{st}^2}{abn}. \quad (5.14)$$

Em seguida, subtrai-se SS_A e SS_B de SS_P , de modo que:

$$SS_{AB} = SS_P - SS_A - SS_B. \quad (5.15)$$

Assim, pode-se encontrar a parcela de variação devido ao erro, tal que:

$$SS_E = SS_T - SS_{AB} - SS_A - SS_B. \quad (5.16)$$

O quadrado médio é o quociente entre a soma de quadrados e os graus de liberdade associados a cada fonte de variação. F_0 representa a estatística de teste para a análise de dois grupos de dados.

5.4 TESTE PARA A SIGNIFICÂNCIA DO MODELO

A fim de se avaliar um modelo e testar se a mudança de níveis de um fator ou interação é significativa se emprega um teste de hipótese para média, que no caso do planejamento de experimentos é a Análise de Variância (ANOVA) (MONTGOMERY; RUNGER, 2011).

Este teste é realizado calculando a razão entre a média quadrática dos termos de regressão e a média quadrática do erro, encontrando-se a estatística F . Comparando-se esta estatística com o valor crítico de F obtido a partir do nível de significância desejado, conclui-se sobre a significância do modelo. Se $F > F_{\text{crítico}}$ (o que equivale a um valor de P menor que o nível de significância), então o modelo é adequado (MONTGOMERY; RUNGER, 2011).

O teste individual de significância de cada coeficiente pode indicar se ele poderá ser adicionado ou removido do modelo, o que pode conduzir à sua otimização. Se o valor de P do teste individual para os termos for inferior ao nível de significância, então o termo é adequado ao modelo e deve, portanto, ser mantido. Caso contrário, o termo deve ser excluído se tal procedimento conduzir a um aumento do coeficiente de determinação R^2 conjuntamente com a diminuição do erro residual S e o valor de P , referente à falta de ajuste do modelo, for superior ao nível de significância. Além disso, a retirada de qualquer termo deve obedecer ao princípio da Hierarquia, ou seja, quando um termo de ordem alta é mantido no modelo, o de ordem baixa que o compõe também deve ser conservado, mesmo não sendo estatisticamente significativo.

5.5 CRITÉRIO DE OPTIMALIDADE D

Utilizando como exemplo um arranjo de superfície de resposta para $k = 3$ fatores, haveria oito pontos fatoriais, seis pontos axiais e cinco pontos centrais, resultando em um total de 19 combinações experimentais. Dependendo do objeto de estudo, são necessárias várias réplicas para cada combinação experimental e portanto a adoção do CCD completo inviabiliza a condução do estudo, uma vez que as réplicas são importantes para obter consistência na geração dos padrões. Assim, neste trabalho, optou-se pela adoção de uma fração do CCD completo escolhida com base no critério de optimalidade D.

De acordo com Montgomery (2013), um modelo pode ser considerado *D-optimal* se $|(X^T X)^{-1}|$ for minimizado. Desta forma, um modelo *D-optimal* minimiza o volume do intervalo de confiança do vetor de coeficientes de regressão. A equação que mede o

desempenho relativo entre o modelo original (1) e o modelo otimizado (2), baseado no critério de optimalidade D é dada por

$$D_e = \left(\frac{|(X_2^T X_2)^{-1}|}{|(X_1^T X_1)^{-1}|} \right)^{1/p} \quad (5.17)$$

na qual X_1 e X_2 são as matrizes para os dois modelos e p é o número de parâmetros.

5.6 CONSIDERAÇÕES FINAIS SOBRE O CAPÍTULO

Neste capítulo, discutiu-se a maneira como diferentes métodos estatísticos são utilizados para resolver problemas de modelagem de respostas correlacionadas. Foram apresentadas as deduções matemáticas, as principais estatísticas de teste, as vantagens e as desvantagens de cada um dos métodos, além dos principais arranjos experimentais disponíveis.

6 ANÁLISE DE CORRELAÇÃO E DETECÇÃO DE PADRÕES EM SINAIS ACÚSTICOS

6.1 CONSIDERAÇÕES INICIAIS

Nos capítulos anteriores, foram discutidos os conceitos fundamentais e necessários para a criação de um método de monitoramento indireto a partir de sinais coletados do processo. O som foi escolhido como sinal a ser estudado neste trabalho devido a suas vantagens como facilidade de instalação e custo dos microfones, destacadas na Seção 2.2.1. Parâmetros extraídos do sinal acústico (som) emitido durante o processo de torneamento são analisados a fim de identificar padrões que possam ser utilizados para o monitoramento das configurações do processo ou até mesmo da qualidade.

Dois parâmetros do sinal acústico são alvos de análises para identificação de padrões que possam ser utilizados como fonte de informação sobre o processo: o MFCC e o espectro de potência médio do sinal, destacados na Seção 2.2.2. Detalhes sobre o processo de extração e a análise de correlação deste parâmetros com as configurações do processo são detalhados neste capítulo.

A fim de avaliar a utilização do som e seus parâmetros para esta finalidade, foi necessário recorrer a um projeto experimental e coletar dados emitidos durante um processo de usinagem. O processo escolhido para esta finalidade foi o de torneamento em aço ABNT 52100 em virtude de sua grande aplicação no mercado industrial e por se tratar de um material com alta dureza. Este tipo de material pode trazer benefícios ao processo como: redução do tempo de corte, baixa rugosidade da peça, aumento da remoção de material, precisão dimensional e usinagem sem fluido de corte. Vale ressaltar que estes benefícios só podem ser atingidos caso sejam utilizadas configurações de usinagem adequadas (PAIVA *et al.*, 2012).

6.2 MÁQUINAS, MATERIAIS, FERRAMENTAS E INSTRUMENTOS DE MEDIÇÃO

No processo de torneamento foi utilizado o torno CNC Nardini Logic 175, com velocidade de rotação máxima de 4.000 rpm e potência de corte de 5,5 kW.

As peças utilizadas para usinagem possuíam dimensões de ϕ 49 mm \times 50 mm. Todas as peças foram previamente temperadas e revenidas, resultando em uma dureza entre 49 e 52 HRC, até uma profundidade de três milímetros abaixo da superfície.

As ferramentas de corte utilizadas foram do tipo cerâmica mista alisadora (Al_2O_3+

TiC), geometria ISO CNGA 120408 S01525WH, revestida com uma camada de nitreto de titânio (TiN). Diferentes ferramentas foram utilizadas a fim de realizar todos os passos com ferramentas não desgastadas.

Os parâmetros de rugosidade foram medidos com um rugosímetro portátil Mitutoyo modelo Surftest SJ-201P, fixo a um comprimento de amostragem de 0,25 mm.

Para a captura do som, foi utilizado um microfone acoplado à placa de som de um laptop Dell Vostro. As gravações foram realizadas através do software Audacity (DANNENBERG, 2013) configurado com uma frequência de amostragem de 44,1 kHz e resolução de 16 bits.

A configuração dos equipamentos para a execução dos experimentos pode ser vista na Figura 6.1.

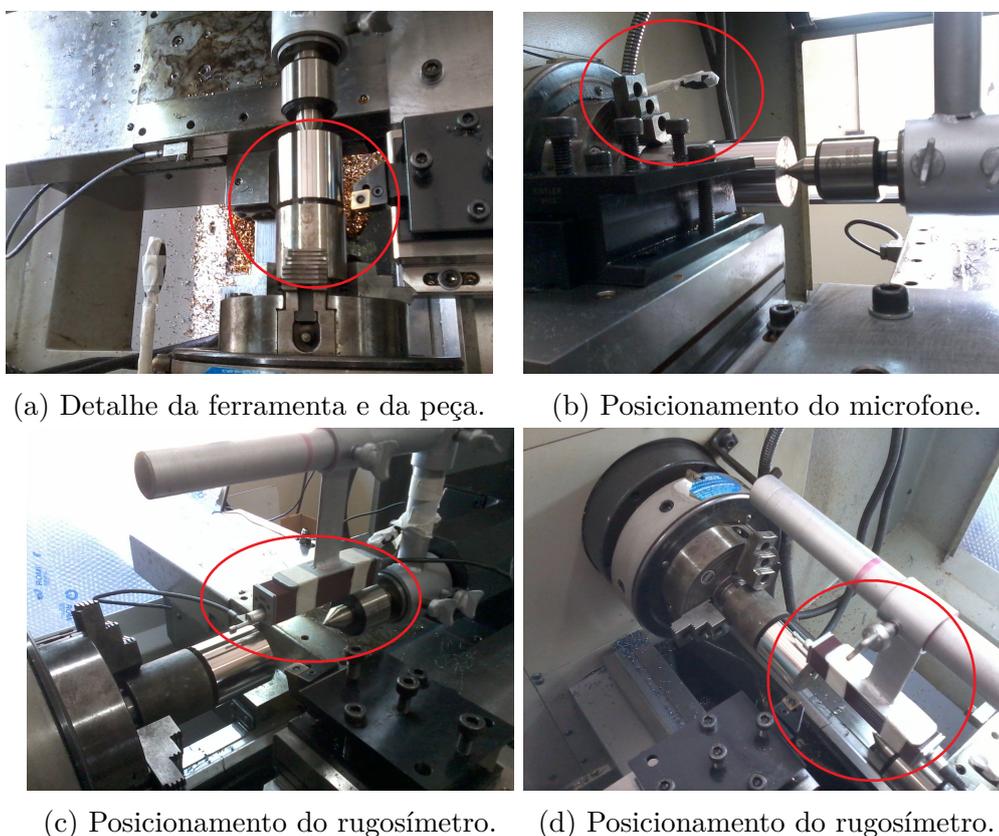


Figura 6.1 – Detalhes da configuração do experimento.

6.3 PROJETO EXPERIMENTAL E AQUISIÇÃO DE DADOS

A fim de gerar um número de experimentos capaz de mapear diferentes configurações de usinagem para análise, elaborou-se um conjunto sequencial de execuções experimentais usando um arranjo composto central (CCD) construído de acordo com Paiva *et al.* (2012). Devido ao número de ferramentas disponíveis para o experimento, 10 conjuntos de parâmetros de usinagem foram selecionados usando um critério D-ótimo (do inglês

D-optimal) (MONTGOMERY; JOHNSON, 2009), resultando na sequência experimental apresentada na Tabela 6.1, na qual Vc representa a velocidade de corte (m/min), f representa o avanço (mm/rev), ap é a profundidade de corte (mm) e TRM , a taxa de remoção de material, calculado através da equação $TRM = Vc \times f \times ap$.

Tabela 6.1 – Projeto D-ótimo para o torneamento do aço duro AISI 52100.

Configuração	Vc (m/min)	f (mm/rev)	ap (mm)	TRM (cm ³ /min)
MS_1	200,00	0,10	0,10	2,00
MS_2	240,00	0,10	0,10	2,40
MS_3	200,00	0,20	0,10	4,00
MS_4	200,00	0,10	0,20	4,00
MS_5	240,00	0,10	0,20	4,80
MS_6	240,00	0,20	0,20	9,60
MS_7	186,36	0,15	0,15	4,19
MS_8	220,00	0,23	0,15	7,72
MS_9	220,00	0,15	0,23	7,72
MS_{10}	220,00	0,15	0,15	4,95

Cada configuração de usinagem foi executada 15 vezes para formar uma base de dados de treinamento e teste capaz de garantir a confiabilidade estatística do método de monitoramento. Durante o experimento, garantiu-se que outros processos de usinagem não estivessem em operação nas imediações, o que poderia acrescentar componentes acústicos que pudessem comprometer os sinais a serem avaliados.

6.3.1 Medidas de Rugosidade

A cada passe, foram medidos cinco parâmetros de rugosidades: média aritmética da rugosidade (Ra); rugosidade máxima (Ry); média quadrática da rugosidade (Rq); rugosidade média (Rz); rugosidade máxima pico/vale (Rt); e largura média da rugosidade máxima pico/vale (Rsm). A rugosidade foi medida quatro vezes em cada ponto (A, B e C), conforme ilustrado na Figura 6.2. As médias das quatro medidas podem ser vistas na Tabela 6.2 (todas as medidas estão disponíveis na Tabela A.1, no Apêndice A).

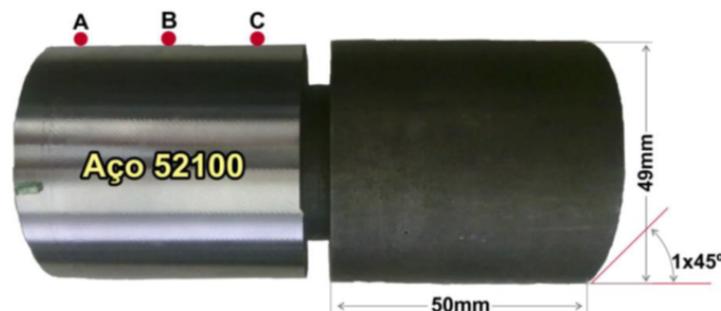


Figura 6.2 – Corpo de prova do aço ABNT 52100. Adaptado de Paiva *et al.* (2012).

Tabela 6.2 – Média das medidas de rugosidade para todas as configurações de usinagem.

Configuração de usinagem	Rugosidade					
	Ra	Ry	Rz	Rq	Rt	Rsm
MS_1	0,15	1,05	0,75	0,18	1,44	85,60
MS_2	0,16	1,17	0,81	0,20	1,56	91,30
MS_3	0,44	2,39	1,76	0,53	3,09	169,72
MS_4	0,19	1,36	1,00	0,24	1,80	58,91
MS_5	0,18	1,32	0,89	0,23	1,89	100,67
MS_6	0,52	2,90	2,24	0,63	3,41	195,46
MS_7	0,26	1,87	1,25	0,33	2,71	104,31
MS_8	0,50	2,62	2,06	0,60	3,10	212,34
MS_9	0,26	1,90	1,29	0,33	2,70	73,04
MS_{10}	0,19	1,42	0,97	0,24	2,10	88,14

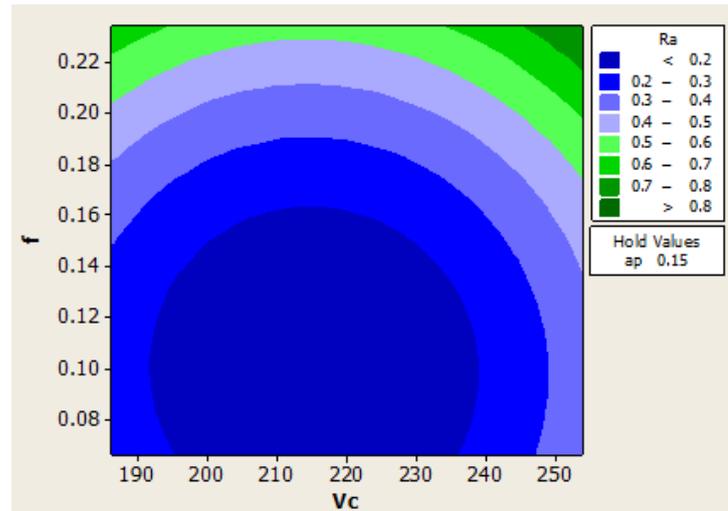
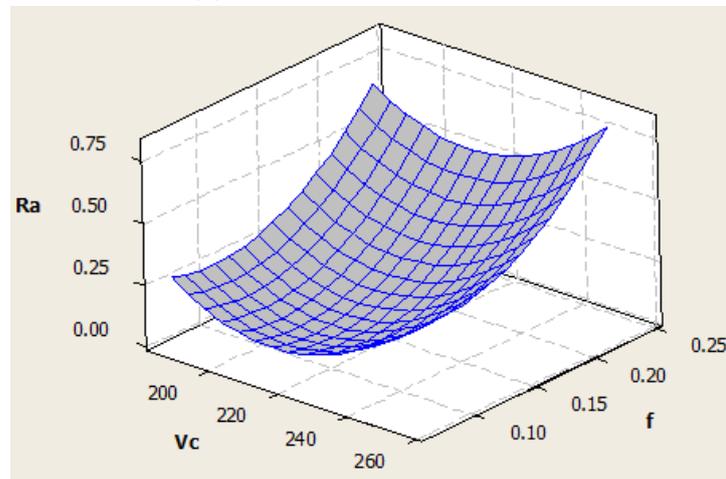
A Tabela 6.3 traz os modelos quadráticos completos para as rugosidades do aço ABNT 52100. A partir destes modelos é possível confirmar que as rugosidades podem ser explicadas através do avanço (f) e da velocidade de corte (Vc), sendo que (f) apresenta a maior correlação conforme previamente demonstrado por outros trabalhos (PAIVA *et al.*, 2012; TEKINER; YESILYURT, 2004).

A superfície de resposta e o gráfico de contorno para a rugosidade Ra é representada na Figura 6.3. Estes gráficos apresentam a relação entre os parâmetros de corte (velocidade de corte e avanço) com a rugosidade resultante. De acordo com as figuras, existe um conjunto de parâmetros que minimiza o valor da rugosidade Ra . A Figura 6.4 apresenta os gráficos de contorno para as demais medidas de rugosidade (Ry , Rq , Rz , e Rt).

Tabela 6.3 – Modelos Quadráticos Completos para as rugosidades do aço AISI 52100.

Modelo	Ra	Ry	Rz	Rq	Rt
Constante	0,16791 ¹	1,2248	0,8254	0,212	1,7701
Vc	0,0365	0,1991	0,1494	0,0475	0,4784
f	0,1207	0,5041	0,4061	0,1376	0,6354
ap	-0,0055	0,0205	0,0217	-0,0072	-0,232
Vc^2	0,0695	0,4065	0,2883	0,0849	0,6477
f^2	0,059	0,2545	0,2434	0,0677	0,1219
ap^2	0,0125	0,0282	0,0319	0,0144	0,0938
$Vc \times f$	0,0029	-0,0648	-0,004	0,0018	0,0698
$Vc \times ap$	0,0022	-0,0329	-0,0081	0,0022	-0,1545
$f \times ap$	0,007	0,0988	0,0693	0,0087	-0,1543
$R^2 adj_{completo}$	91,87%	85,92%	87,61%	91,13%	87,34%
$R^2 adj_{reduzido}$	93,61%	88,73%	90,08%	93,09%	84,28%

¹ valores em negrito correspondem aos termos estatisticamente significantes (P-value < 5%).

(a) Gráfico de contorno de Ra .(b) Superfície de Resposta para Ra .Figura 6.3 – Gráficos de contorno e superfície de resposta para a rugosidade Ra .

A correlação entre os parâmetros de corte f e Vc e a rugosidade resultante mostra que, caso seja possível identificar variações no som de acordo com as variações nos parâmetros de usinagem, talvez seja possível também identificar alterações na rugosidade e, conseqüentemente, monitorar a qualidade do processo. Portanto, ao se obter os parâmetros ótimos do processo que minimizam a rugosidade da resposta (Figura 6.3), pode-se utilizar o som para monitorar este ponto de operação e garantir a sua qualidade.

6.3.2 Aquisição do Som e Análise Espectral

Para a gravação do áudio, foi utilizado o esquema proposto por Fadare *et al.* (2012). Porém, para este trabalho, foi utilizado um microfone acoplado a um computador no lugar do acelerômetro (sensor de emissão acústica). Para cada configuração de usinagem da Tabela 6.1 foi gerado um total de 15 gravações totalizando 150 arquivos de áudio para análise.

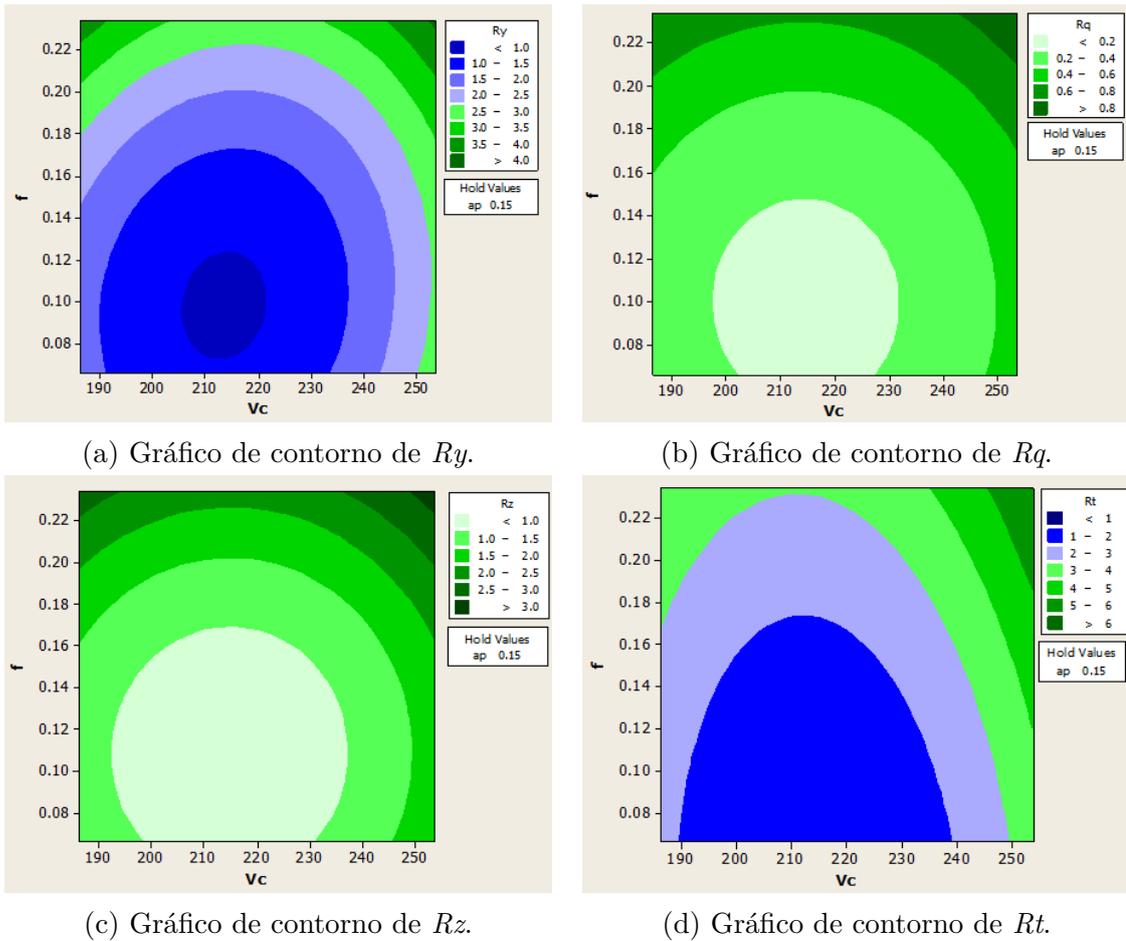


Figura 6.4 – Gráficos de contorno para as rugosidades R_y , R_q , R_z e R_t .

A análise espectral de cada sinal foi realizada utilizando um tempo de quadro t_F de 10ms e um tempo de janela t_W de 20ms, resultando em uma sobreposição de 50 %, obtida através da Equação 3.1. Estes são valores comumente utilizados para a análise espectral de sinais aplicados a reconhecimento de padrões de voz (PICONE, 1993). Cada janela foi, então, multiplicada pela função de Hamming (Equação 3.2), usando $\alpha_w = 0.54$ e $\beta_w = 1$ e, finalmente, obteve-se sua Transformada de Fourier. Um vetor de espectro \vec{v} composto de 129 valores de potência foi obtido para cada janela, através do algoritmo da FFT utilizando $N = 256$ pontos (veja Equações 3.5 e 3.6). O número de pontos da FFT foi escolhido de forma a garantir a resolução necessária do espectro de potência, sem aumentar o tempo computacional necessário para o processamento do sinal. O processo de análise espectral é ilustrado na Figura 6.5.

Mapeando cada vetor de espectro de potência \vec{v} em uma escala de cores, na qual tons mais escuros correspondem a maiores amplitudes, pode-se representar o sinal em uma escala de tempo \times frequência conhecido como espectrograma. Um exemplo desta representação pode ser visto na Figura 6.6, para o som gravado durante a usinagem do aço ABNT 52100, utilizando os parâmetros $Vc = 200$, $f = 0,10$ e $ap = 0,10$. A partir desta representação é possível identificar como a potência do espectro se comporta ao

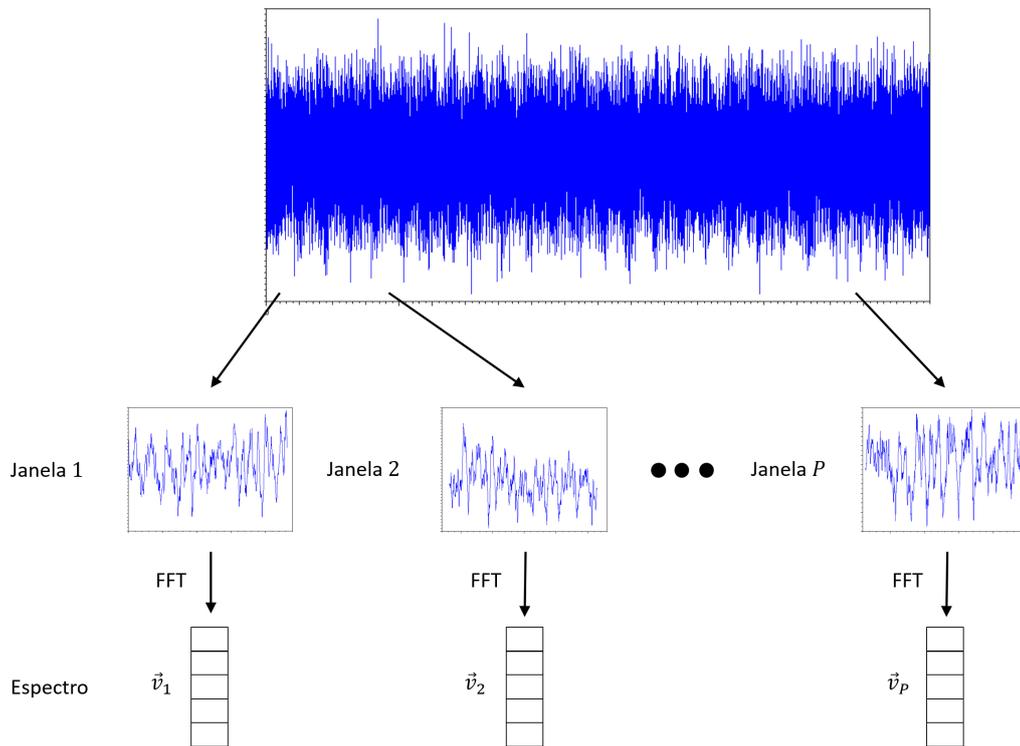
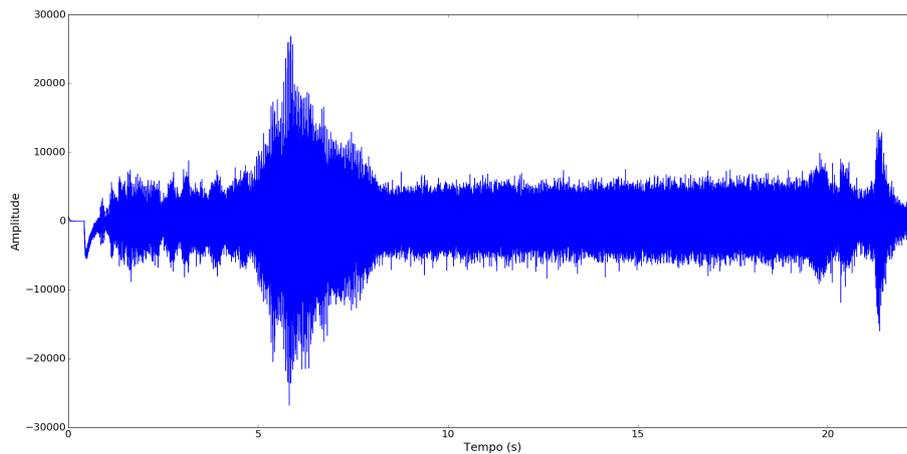


Figura 6.5 – Ilustração do processo de análise espectral.

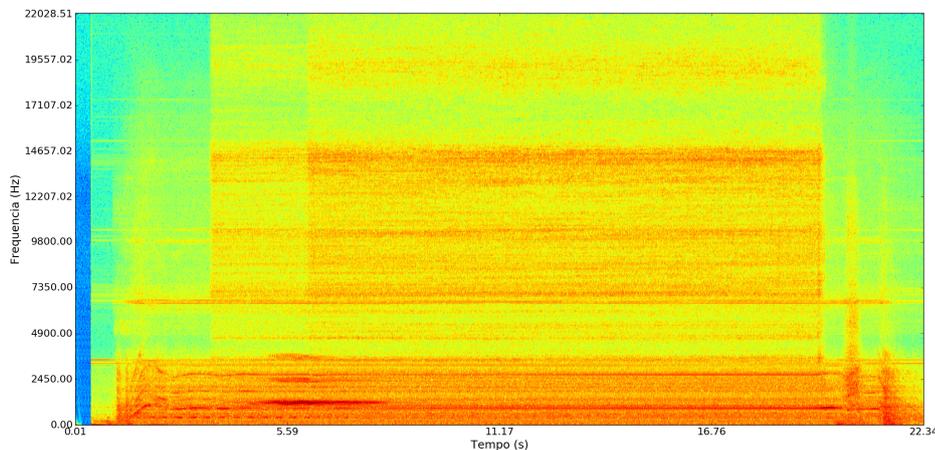
longo do tempo.

Como é possível observar, o som emitido pelo processo apresenta um período de transitório assim que a ferramenta entra em contato com a peça e, também, durante sua saída. Esta variação resulta em valores de potência elevados para os componentes de frequência abaixo de 2,45 kHz e potências mais baixas acima desta frequência. Durante o período de usinagem, existe uma forte concentração de potência na faixa de 14 kHz, que é mais suave durante os transitórios. Esta diferença de distribuição de potência em relação ao período de usinagem é ainda mais perceptível para os períodos em que a ferramenta de corte não está em contato com a peça, como pode ser notado nos períodos antes de 5 segundos e depois de 20 segundos, aproximadamente. A fim de selecionar apenas a parte estacionária de cada sinal, foram extraídos intervalos de tempo em torno de 7 a 10 segundos, evitando os momentos em que a ferramenta entra e sai da peça, conforme ilustrado na Figura 6.7.

Logo, o processo de análise espectral descrito acima foi aplicado apenas à parte selecionada, resultando em uma faixa de sinal com um comportamento estacionário, conforme pode ser visto na Figura 6.8. Utilizando esta análise para avaliar os sinais obtidos durante os experimentos, foi possível identificar que todos apresentavam um comportamento estacionário, ou seja, uma distribuição de potência estável ao longo do tempo, conforme pode ser visto pelos exemplos apresentados na Figura 6.9.



(a) Gráfico do sinal no domínio do tempo.



(b) Gráfico do espectrograma do sinal.

Figura 6.6 – Comparação das características do sinal acústico nos domínios do tempo e da frequência (espectrograma).

Esta avaliação foi importante porque trouxe uma série de benefícios para os passos posteriores em relação à extração de informações destes sinais. O primeiro grande benefício foi mostrar que não existem informações relevantes no tempo do sinal, acarretando a redução do número de amostras necessárias para sua representação. Em outras palavras, pode-se utilizar apenas uma média de todas as observações fazendo

$$\vec{v}_{avg} = \frac{1}{P} \sum_{i=1}^P \vec{v}_i \quad (6.1)$$

em que P representa o número total de vetores de potência. Pode-se dizer que esta média representa a potência média do espectro do sinal ao longo do tempo ou a sua energia dividida pelo intervalo de tempo em análise (HAYKIN; VEEN, 2003). Outro impacto resultante desta análise diz respeito ao tipo de classificador que pode ser utilizado na fase de diagnóstico do sistema de monitoramento. Por se tratar de um sinal sem informações

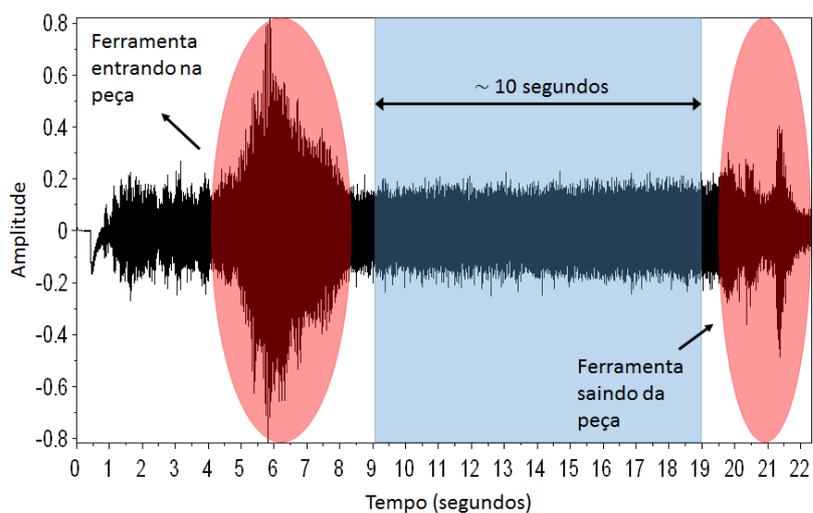


Figura 6.7 – Som gravado durante usinagem do aço ABNT 52100 ($V_c = 200$, $f = 0,10$ e $ap = 0,10$).

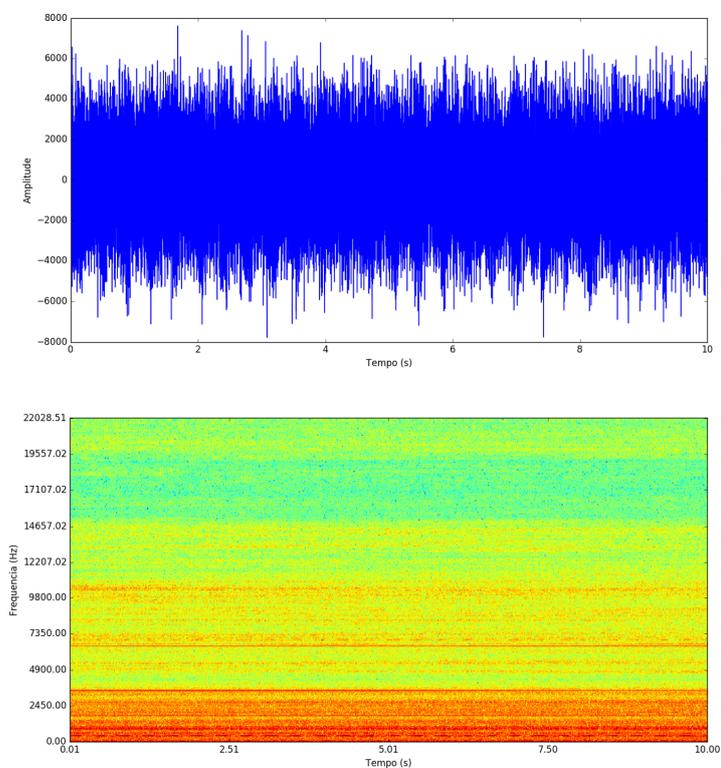
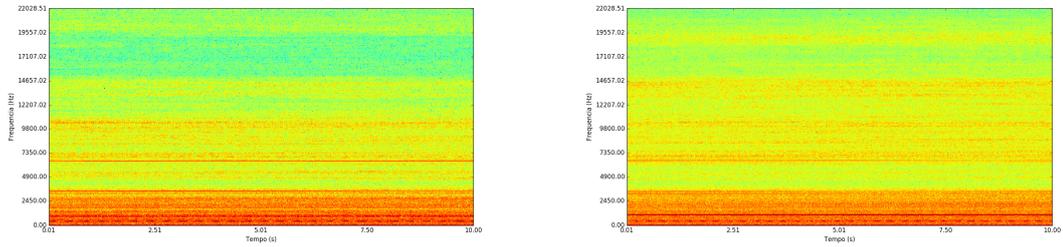
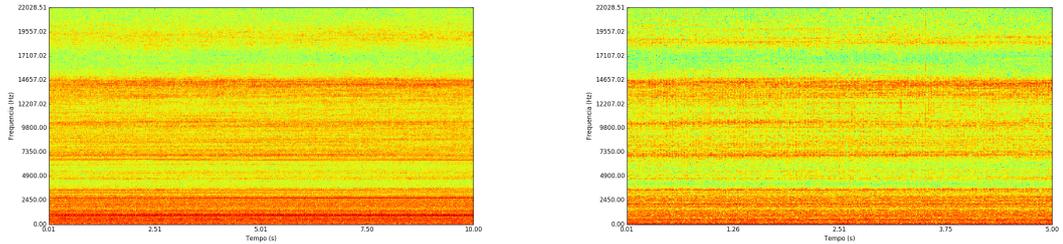


Figura 6.8 – Exemplo do espectrograma obtido para o som gravado durante usinagem do aço ABNT 52100 ($V_c = 200$, $f = 0,10$ e $ap = 0,10$).

temporais, pode-se utilizar classificadores baseados em GMM ou SVM, por exemplo, em vez de usar técnicas baseadas em HMM (ver Capítulo 2.2.3 para mais detalhes).



(a) MS_1 ($Vc = 200$, $f = 0, 10$ e $ap = 0, 10$). (b) MS_2 ($Vc = 240$, $f = 0, 10$ e $ap = 0, 10$).



(c) MS_7 ($Vc = 186$, $f = 0, 15$ e $ap = 0, 15$). (d) MS_8 ($Vc = 220$, $f = 0, 23$ e $ap = 0, 15$).

Figura 6.9 – Espectrogramas de sinais acústicos gerados por diferentes configurações de usinagem.

Ainda com base na Figura 6.9, é possível identificar que diferentes parâmetros de usinagem resultam em diferentes espectros de potência, se comparados MS_1 com MS_8 , por exemplo. O espectrograma resultante de MS_1 apresenta uma maior concentração de energia nas frequências mais baixas (abaixo de 7 kHz), enquanto o MS_8 produz altos valores de potência também para as frequências na faixa entre 7 kHz e 15 kHz.

Por outro lado, esta diferença não é tão significativa se forem comparados MS_1 e MS_2 , o que de certa forma dá indícios de que alguns sinais podem ser correlacionados e que esta correlação pode ser identificada por características do sinal, neste caso, seu espectro de potência médio. Com base nestas informações, a próxima etapa é extrair os parâmetros dos sinais e verificar se tal correlação existe e se pode ser utilizada para o monitoramento deste processo.

6.4 EXTRAÇÃO DE PARÂMETROS DOS SINAIS ACÚSTICOS UTILIZANDO ACP

Como detalhado na Seção 6.3.2, o processo de análise espectral gerou um vetor \vec{v}_{avg} com a potência média do espectro para cada sinal acústico. Devido ao número de pontos utilizados para o cálculo da FFT, este vetor tem um número grande de componentes para uma análise de correlação (neste caso 129).

A fim de extrair deste vetor apenas as informações mais relevantes e ainda reduzir a sua dimensão, pode-se utilizar a técnica de Análise de Componentes Principais (ACP), descrita na Seção 3.3.2. Aplicando ACP sobre o conjunto de vetores de potência, obtiveram-se os escores apresentados no Apêndice B. É importante ressaltar que para esta análise, o algoritmo para a obtenção dos componentes principais utilizou a matriz de covariância completa do conjunto de dados. É possível verificar através da Figura 6.10 e da Tabela 6.4 que os oito primeiros componentes representam a maior parte da variabilidade dos dados. Assim, uma redução na dimensão é obtida ao projetar os dados dos vetores originais sobre os autovetores reduzidos.

Tabela 6.4 – Análise dos autovetores da matriz de correlação entre os espectros de potência dos sinais.

Componente	PC_1	PC_2	PC_3	PC_4	PC_5	PC_6	PC_7	PC_8	PC_9
Autovalor	87,331	14,722	5,174	2,846	2,544	1,884	1,543	1,253	1,101
Proporção	0,677	0,114	0,04	0,022	0,02	0,015	0,012	0,01	0,009
Cumulativo	0,677	0,791	0,831	0,853	0,873	0,888	0,9	0,909	0,918

Porém, utilizando a análise de correlação foi possível identificar que apenas os quatro primeiros componentes principais apresentam correlação estatisticamente significativa com os parâmetros de usinagem e rugosidade de resposta. O respectivos níveis de correlação dos componentes principais com estes parâmetros podem ser vistos na Tabela 6.5. O escore correspondente ao primeiro componente principal (PC_1) apresentou alta correlação de Pearson¹ com todos os parâmetros de corte e medidas de rugosidade, destacando-se f , TRM , Ra e Ry . O segundo componente (PC_2) apresentou correlação com a velocidade de corte (Vc), assim como o componente PC_3 , porém o último se correlaciona com maior intensidade. Já o componente PC_4 apresentou correlação significativa com a profundidade de corte (ap) e a rugosidade Rsm .

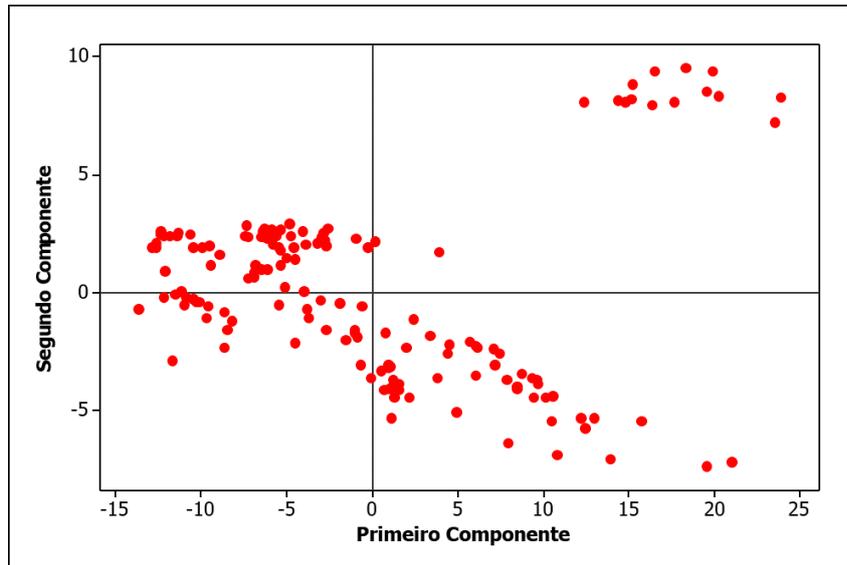
Tabela 6.5 – Análise de correlação entre escores, configurações de usinagem e medidas de rugosidade.

Escore	Vc	f	ap	TRM	Ra	Ry	Rz	Rq	Rt	Rsm
PC_1	0,208^{a,b}	0,676	0,532	0,879	0,667	0,738	0,704	0,680	0,731	0,523
PC_2	0,213	-0,087	0,037	0,072	0,175	0,139	0,183	0,168	0,038	0,120
PC_3	0,797	-0,256	0,345	0,225	-0,219	-0,192	-0,178	-0,216	-0,218	-0,184
PC_4	-0,266	-0,217	0,415	0,021	-0,234	-0,167	-0,172	-0,226	-0,106	-0,474

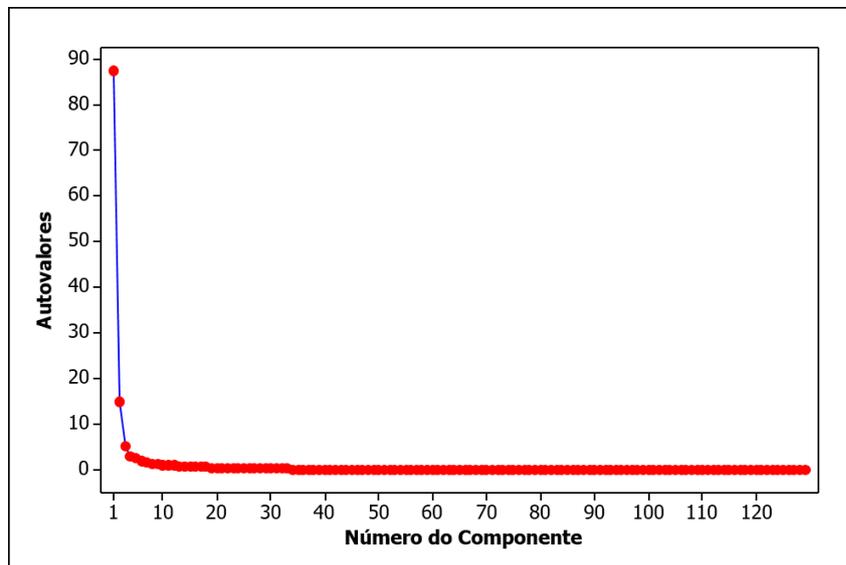
^a Valores em negrito representam os valores estatisticamente significativos (P-Value < 5%).

^b Correlação de Pearson.

¹ A correlação produto-momento de Pearson é utilizada para avaliar a força e a direção da associação entre duas variáveis contínuas que são linearmente relacionadas. Seu coeficiente, r , indica a força e a direção dessa relação e pode variar de -1, para uma relação linear negativa perfeita, a +1 para uma para uma relação linear positiva perfeita. Um valor igual a 0 (zero) indica que não há nenhuma relação entre as duas variáveis.



(a) Gráfico dos escores correspondentes ao primeiro e segundo componentes principais.



(b) Gráfico com a distribuição de significância dos autovalores.

Figura 6.10 – Gráfico de avaliação do resultado da análise de componentes principais.

Outra maneira de demonstrar tais correlações é através de um gráfico de linha das variáveis padronizadas, no qual é possível verificar como as variações dos parâmetros impactam os respectivos escores aos quais são correlacionados. A forte correlação entre a TRM e o escore PC_1 é enfatizada na Figura 6.11, na qual é possível verificar que PC_1 segue, aproximadamente, as variações da TRM . Este comportamento apoia a ideia de que as emissões acústicas audíveis transportam informações sobre o processo de usinagem, já que a energia do som (para certas frequências) aumenta de acordo com o aumento da taxa de remoção de material.

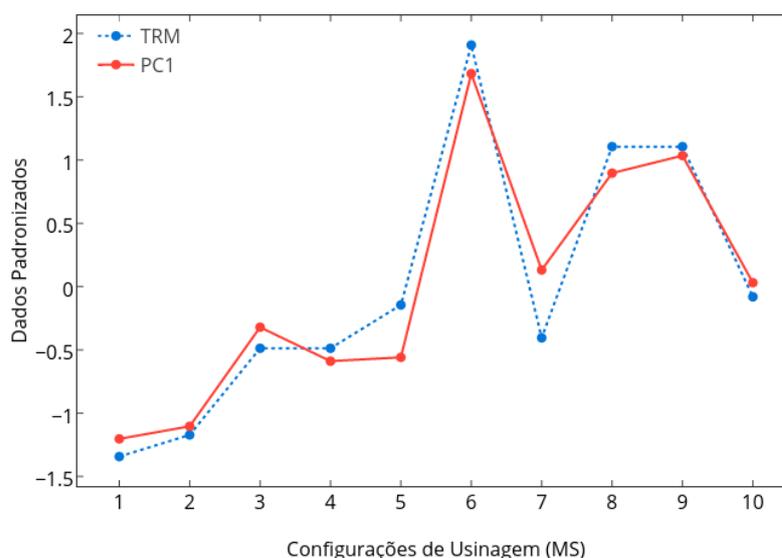


Figura 6.11 – Gráfico de linha comparando PC_1 e taxa de remoção de material TRM .

O escore PC_3 também apresentou alta correlação com o parâmetro velocidade de corte (Vc), como pode ser visto na Figura 6.12. O aumento da velocidade de corte reflete diretamente no escore PC_3 , que também aumenta. Como PC_1 se correlaciona com $TRM = Vc \times f \times ap$, conhecendo a velocidade de corte (Vc) de comportamento através PC_3 , é possível combinar os dois escores para ter alguma ideia sobre o avanço (f) e profundidade de corte (ap), mapeando quase todos os parâmetros de usinagem.

Mais importante do que detectar variações nos parâmetros de usinagem usando a energia do som, é medir a qualidade do processo por meio da rugosidade da peça, o que permite utilizar as emissões sonoras audíveis para técnicas de monitoramento indiretos. A Figura 6.13 mostra que o escore PC_1 segue as tendências da maioria das medidas de rugosidade, apresentando maior correlação com Ry e Rt . É importante destacar que as configurações de usinagem 3 e 4 resultaram o mesmo $TRM = 4$, mas com valor de rugosidade diferente, conforme apresentado na Figura 6.13. Conforme apresentado por Tekiner e Yesilyurt (2004), a rugosidade da superfície é fortemente correlacionada com o avanço (f), o que explica uma maior rugosidade de superfície para $f = 0,20$ na configuração de usinagem 3, em comparação com $f = 0,10$, da configuração de usinagem

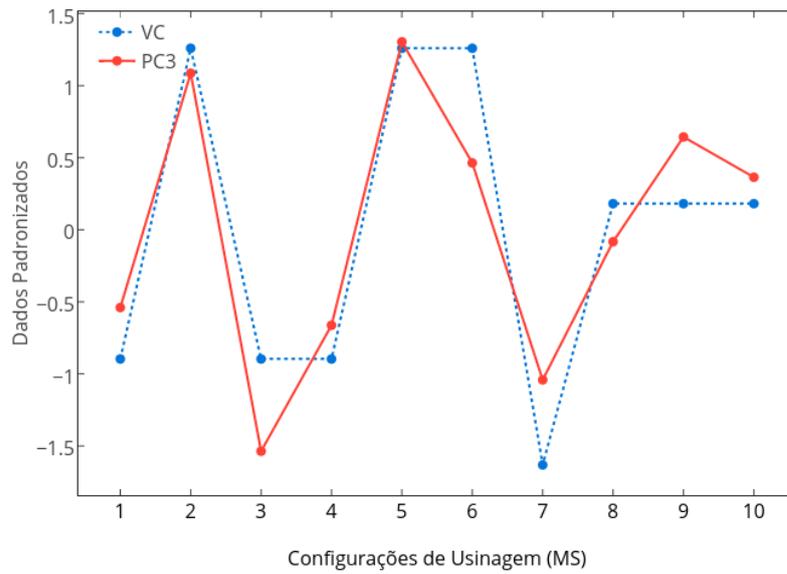


Figura 6.12 – Gráfico de linha comparando PC_3 e velocidade de corte V_c .

4. O mesmo comportamento pode ser observado para as configurações de usinagem 8 e 9, que apresentam os mesmos valores para TRM , mas para diferentes taxas de avanço ($f = 0,23$ e $f = 0,15$ respectivamente). Essa diferença no avanço impacta diretamente a rugosidade da superfície, mas não é detectada pelo escore PC_1 .

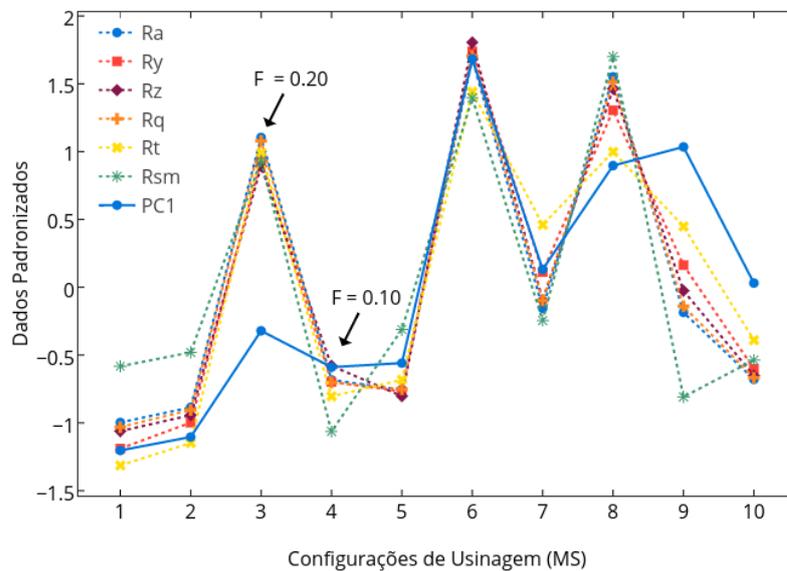


Figura 6.13 – Gráfico de linha comparando PC_1 e as medidas de rugosidade.

No entanto, o escore PC_1 tem características que o tornam um parâmetro interessante a ser aplicado em monitoramento de processos. A Figura 6.14 mostra um gráfico de dispersão cruzando Ry e PC_1 . É possível visualizar pelo menos três grupos separáveis, nos quais PC_1 pode detectar o nível de rugosidade da superfície. O impacto do avanço (f) sobre a medida de rugosidade da superfície também é destacado na Figura 6.14, mostrando

que altos valores de f resultam em rugosidades de superfície elevadas.

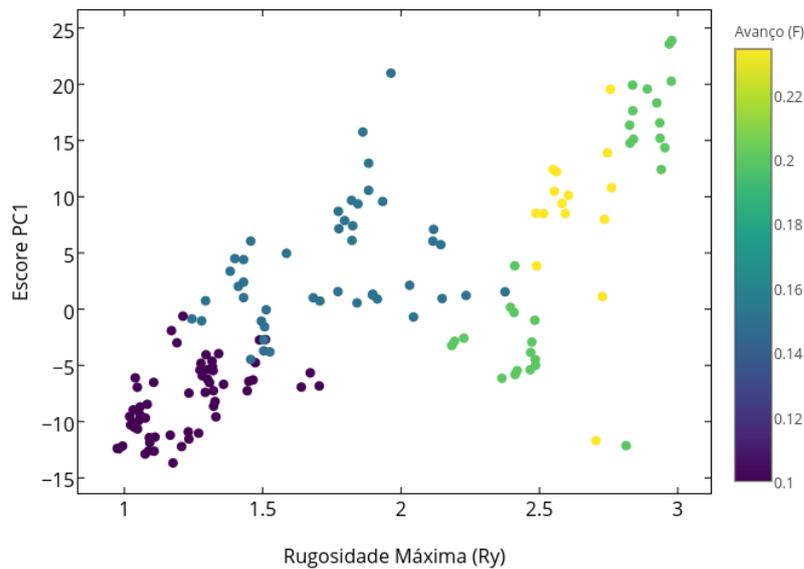


Figura 6.14 – Gráfico de dispersão relacionando PC_1 e a máxima rugosidade Ry , mapeadas pelo valor do avanço f .

6.5 EXTRAÇÃO DE PARÂMETROS DOS SINAIS ACÚSTICOS UTILIZANDO MFCC

Para a extração dos parâmetros Mel-Cepstrais conforme as etapas descritas na Seção 3.3.1, foi necessária sua codificação utilizando a linguagem de programação C (veja códigos no Apêndice D). A partir da implementação, foram extraídos $C = 12$ coeficientes de cada janela do sinal. Foram utilizados $L = 31$ filtros para o cálculo dos coeficientes de forma que $f(31 + 1) = 21.112$ seja menor que $44.100/2 = 22.050$, em que 44.100 Hz corresponde à frequência de amostragem utilizada.

De acordo com a revisão apresentada na Seção 3.3.1, os coeficientes Mel-Cepstrais são calculados a partir dos vetores de espectro de potência, de forma que se obtém um vetor de coeficientes mel \vec{c}_i para cada vetor de espectro de potência \vec{v}_i , para $i = 1, 2, \dots, P$, sendo P o número total de janelas (veja ilustração na Figura 6.15).

Conforme discutido anteriormente, os vetores de espectro de potência não apresentaram variação significativa ao longo do tempo, o que resultou na utilização de um único vetor contendo a média de todos os vetores. Desta forma, a mesma abordagem pode ser utilizada para o MFCC, na qual se pode adotar um vetor de coeficientes Mel-Cepstrais composto pela média de todas as observações, ou seja,

$$\vec{c}_{avg} = \frac{1}{P} \sum_{i=1}^P \vec{c}_i. \quad (6.2)$$

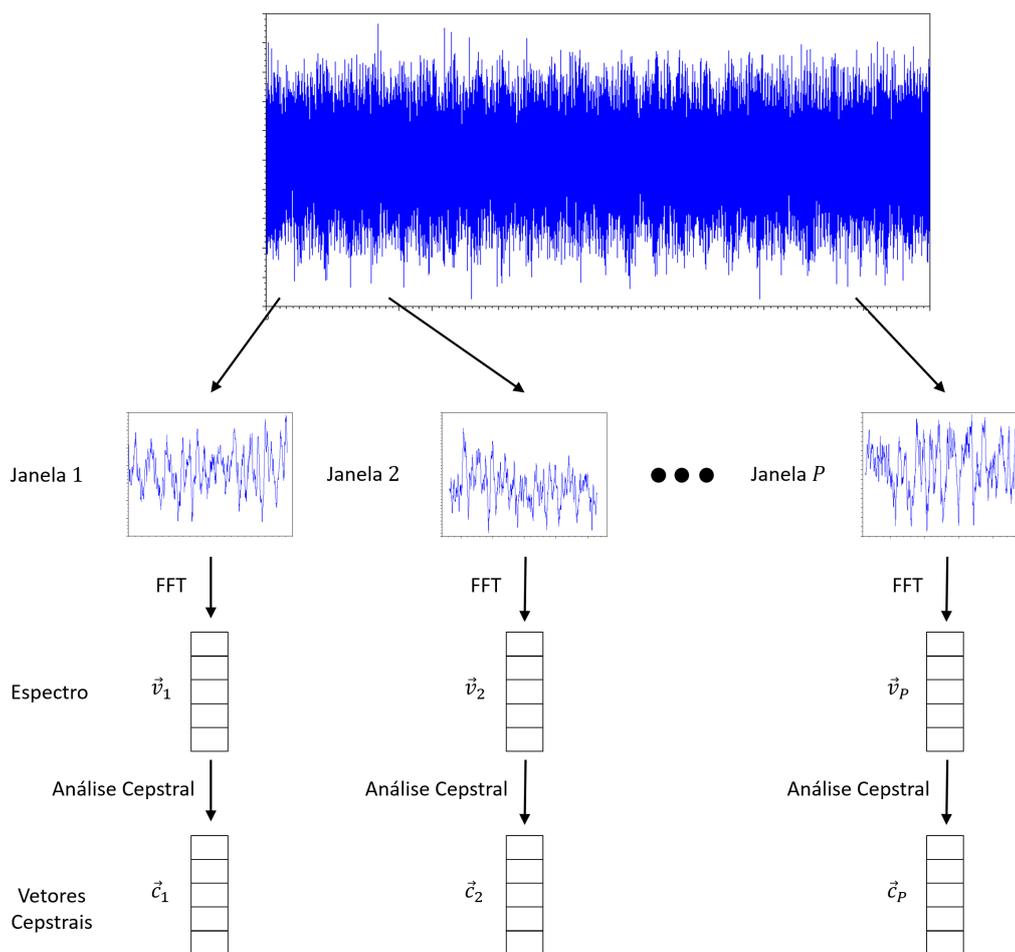


Figura 6.15 – Ilustração do processo de extração dos coeficientes Mel-Cepstrais.

Assim como acontece ao utilizar ACP sobre os vetores de espectro de potência, a extração dos coeficientes mel resulta em uma redução na dimensão deste vetor, já que são obtidos como resultado 12 coeficientes. As tabelas com os seis primeiros coeficientes Mel-Cepstrais, obtidos para todos os sinais, são apresentadas no Apêndice C. Entretanto, ainda é necessário verificar se estes coeficientes se correlacionam com os parâmetros do processo de usinagem.

Através da análise de correlação, foi possível identificar quais coeficientes Mel-Cepstrais são fortemente correlacionados com parâmetros de usinagem e rugosidade da superfície (veja Tabela 6.6). Cada coeficiente apresentou maior correlação de Pearson com diferentes parâmetros de usinagem, como por exemplo, os coeficientes $c_5, c_6, c_8, c_9, c_{11}$ e c_{12} , que possuem correlação mais forte com a velocidade de corte (Vc).

Já os coeficientes c_1 e c_2 são fortemente correlacionados com a taxa de remoção de material (TRM) e também com todas as medidas de rugosidade (Ra, Ry, Rz, Rq, Rt e Rsm), nas quais Rt (o pico máximo de rugosidade de superfície) apresentou a maior correlação entre as medidas de rugosidade. A Figura 6.16 ilustra como o segundo coeficiente

(c_2) é negativamente correlacionado com as medidas de rugosidade da superfície. Quando o nível de rugosidade aumenta, c_2 tem o mesmo comportamento, mas no sentido oposto, o que explica a correlação negativa. A Figura 6.17 ilustra a forte correlação entre a medida de rugosidade Rt e o coeficiente c_2 , mostrando o coeficiente com o sinal invertido.

Tabela 6.6 – Análise de correlação entre coeficientes Mel-Cepstrais, configurações de usinagem e medidas de rugosidade.

MFCC	V_c	f	ap	TRM	Ra	Ry	Rz	Rq	Rt	Rsm
c_1	0,061 ^a	0,594 ¹	0,563	0,781	0,595	0,665	0,629	0,607	0,685	0,419
c_2	-0,069	-0,703	-0,506	-0,802	-0,649	-0,722	-0,671	-0,662	-0,752	-0,507
c_3	0,123	0,430	0,210	0,414	0,228	0,265	0,220	0,235	0,318	0,285
c_4	0,440	-0,240	0,334	0,204	-0,040	-0,053	-0,003	-0,048	-0,158	-0,097
c_5	0,673	0,029	0,437	0,432	0,003	0,037	0,037	0,006	0,018	0,021
c_6	0,566	0,001	-0,098	0,058	-0,052	-0,126	-0,066	-0,063	-0,216	0,093
c_7	-0,069	0,029	0,077	0,106	-0,081	-0,058	-0,039	-0,072	-0,037	-0,135
c_8	-0,607	0,207	-0,228	-0,173	0,063	0,076	0,051	0,068	0,165	-0,031
c_9	-0,562	-0,021	-0,361	-0,296	-0,032	-0,066	-0,069	-0,036	-0,043	0,010
c_{10}	0,013	0,239	-0,032	0,110	0,291	0,237	0,237	0,280	0,206	0,346
c_{11}	0,616	-0,037	0,271	0,324	0,062	0,062	0,065	0,059	0,014	0,159
c_{12}	0,614	0,052	0,334	0,363	-0,020	0,025	0,012	-0,013	0,024	0,006

¹ Valores em negrito representam os valores estatisticamente significativos (P-Value < 5%).

^a Correlação de Pearson.

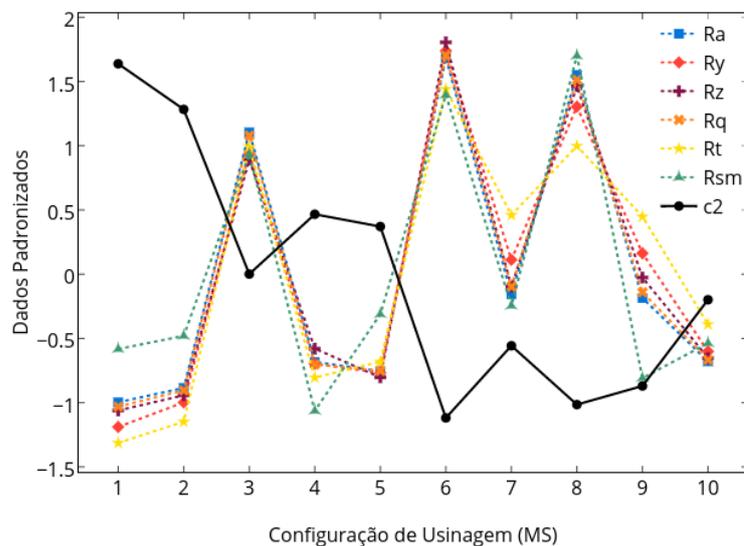


Figura 6.16 – Gráfico de linha comparando c_2 e as medidas de rugosidade.

Com o objetivo de destacar tal correlação, utilizou-se a análise de superfície de resposta em duas réplicas selecionadas aleatoriamente para o coeficiente c_2 , que apresenta a maior correlação com as medidas de rugosidade, resultando nos coeficientes de regressão apresentados na Tabela 6.7. Os resultados não só apresentaram coeficientes de determinação adequados (acima de 95%, exceção para R_t que apresentou um coeficiente de determinação de 86,13%), mas também evidências suficientes para afirmar que os resíduos

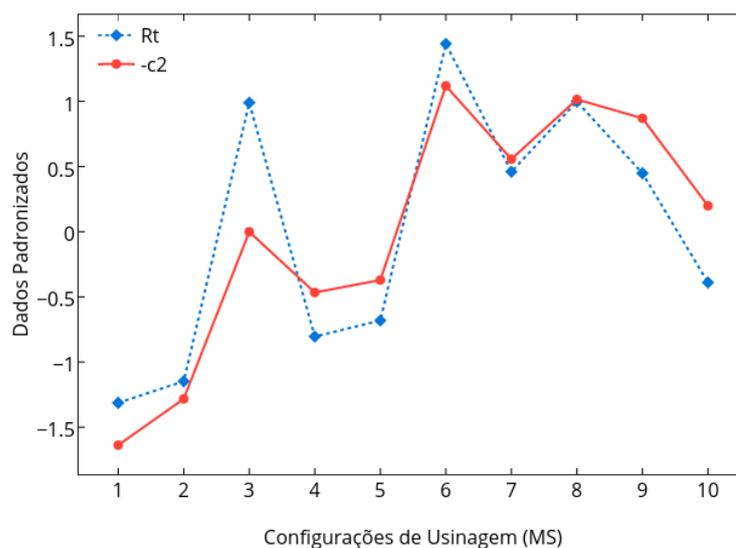
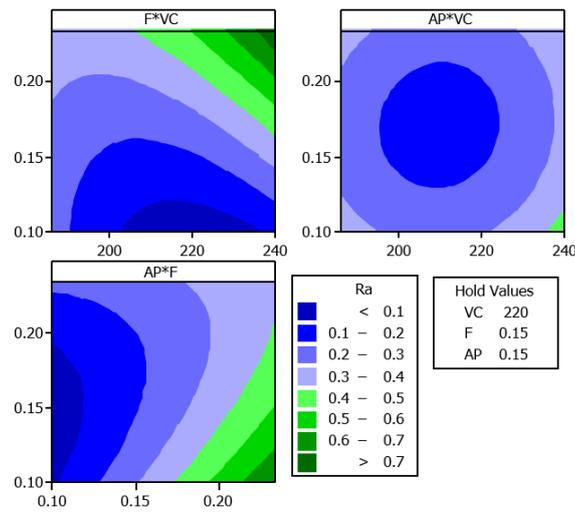


Figura 6.17 – Gráfico de linha comparando $-c_2$ e a medidas de rugosidade Rt .

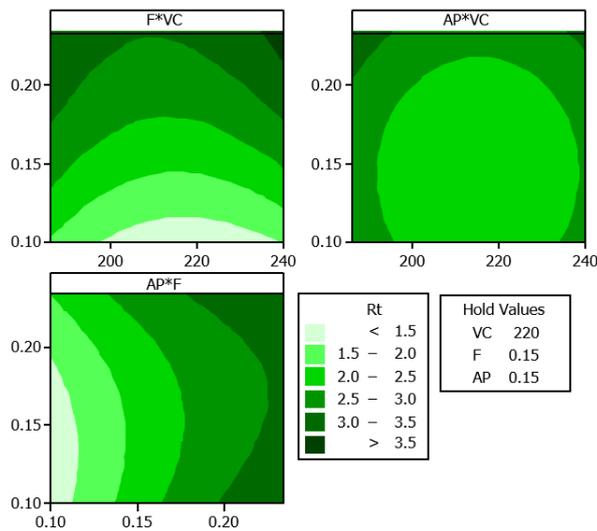
seguem uma distribuição normal. Como pode ser observado na Tabela 6.7, o avanço (f) foi o fator mais significativo para explicar o comportamento médio das medidas de rugosidade e também foi um dos fatores mais significativos para c_2 . Pode-se evocar novamente a afirmação de Tekiner e Yesilyurt (2004) sobre a rugosidade ser fortemente correlacionada com o avanço (f) para destacar que esta pode ser uma das razões pelas quais existe forte correlação entre as medidas de rugosidade e os coeficientes Mel-Cepstrais. Por outro lado, a profundidade de corte (ap) teve impacto mais elevado sobre o coeficiente Mel do que sobre as medidas de rugosidade. Por isso, os profissionais devem dedicar muita atenção a esta variável de controle, a fim de garantir a confiabilidade do processo de tomada de decisão no sistema de monitoramento. Provavelmente, esta é a principal razão pela qual a estrutura de correlação não foi ainda mais forte.

Finalmente, a Figura 6.18 representa graficamente os modelos de superfície de resposta, comparando dois parâmetros de rugosidade (Ra e Rt) e o segundo coeficiente Mel-Cepstral c_2 . As relações negativas são claramente visualizadas, principalmente devido aos gráficos de contorno para “ $f \times Vc$ ” e “ $ap \times f$ ”.

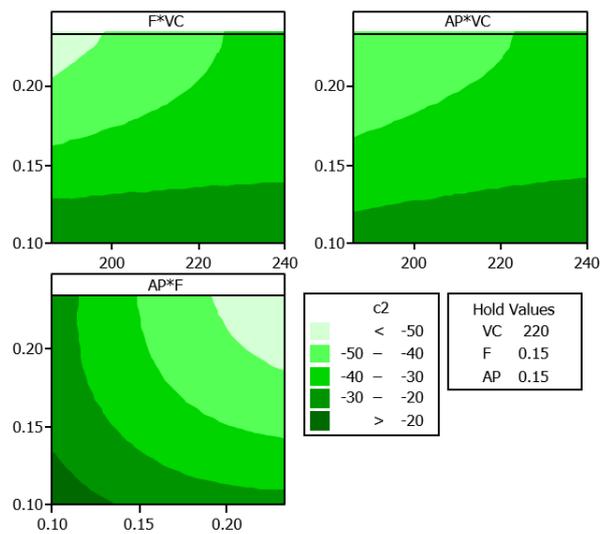
A análise apresentada mostrou que a energia específica de corte (relacionada com o TRM) aplicada ao sistema é provavelmente transformada em outros tipos de energia, como vibração, calor (ferramenta, peça, cavaco e ar), som e outras. Neste caso, a parte da energia transferida para o som audível pode ser identificada por MFCC, comprovando que este parâmetro pode ser aplicado em monitoramento de processos.



(a) Gráfico de contorno para a rugosidade Ra .



(b) Gráfico de contorno para a rugosidade Rt .



(c) Gráfico de contorno para o coeficiente Mel-Cepstral c_2 .

Figura 6.18 – Gráfico de contorno para as rugosidades Ra e Rt e o coeficiente Mel-Cepstral c_2 .

Tabela 6.7 – Coeficientes de regressão e coeficientes de determinação para os modelos de rugosidade e segundo coeficiente mel.

	R_a	R_y	R_z	R_q	R_t	c_2
Constant	0,185¹	1,387	0,958	0,236	2,133	-32,234
V_c	0,074	0,202	0,24	0,079	0,074	2,765
f	0,169	0,746	0,616	0,194	0,773	-9,278
ap	-0,053	-0,035	-0,111	-0,051	0,075	-10,073
V_c^2	0,075	0,33	0,267	0,084	0,201	0,091
f^2	0,014	0,002	0,044	0,015	-0,123	0,78
ap^2	0,057	0,198	0,172	0,062	0,162	3,079
$V_c \times f$	0,077	0,197	0,266	0,081	0,043	5,743
$V_c \times ap$	-0,01	-0,051	-0,052	-0,011	-0,009	0,743
$f \times ap$	-0,063	-0,081	-0,159	-0,063	0,046	-5,759
S	0,013	0,135	0,085	0,018	0,262	1,283
R-Sq(adj)	99,27%	95,79%	97,62%	98,80%	86,13%	98,70%
Teste de normalidade	0,433 ^a	0,268	0,117	0,118	0,328	0,382
para resíduos	0,273 ^b	0,646	0,988	0,988	0,492	0,365

¹ Valores em negrito identificam os termos significativos do modelo (P-Value < 5%).

^a Teste estatístico Anderson-Darling.

^b P-Value.

6.6 CONSIDERAÇÕES FINAIS SOBRE O CAPÍTULO

O presente capítulo teve como objetivo investigar se os parâmetros extraídos do som – vetores de espectro de potência através do PCA e MFCC – tinham correlação com os parâmetros de usinagem e com as medidas de rugosidade, a fim de serem aplicados no monitoramento de processos.

Através de técnicas estatísticas, como a metodologia de superfície de resposta e a análise de variância, foi possível identificar diversos fatores que se correlacionam com os parâmetros do processo e que demonstram grande potencial para o monitoramento deste processo de usinagem.

Uma parte das conclusões obtidas neste capítulo resultou em um artigo com o título “*Pattern recognition in audible sound energy emissions of AISI 52100 hardened steel turning: a MFCC-based approach*” publicado na revista *International Journal of Advanced Manufacturing Technology*, páginas 1-10, em 2016 (FRIGIERI *et al.*, 2016a). O artigo pode ser visualizado no Anexo A.

O próximo capítulo se dedica à aplicação destes fatores em algumas técnicas de monitoramento para validação das observações relatadas, a fim de consolidar o som como um sinal em potencial para este tipo de aplicação.

7 MÉTODOS DE MONITORAMENTO INDIRETOS BASEADOS EM SINAIS ACÚSTICOS

7.1 CONSIDERAÇÕES INICIAIS

O presente capítulo tem como objetivo avaliar a utilização dos parâmetros extraídos do som que apresentaram alta correlação com os parâmetros do processo a serem monitorados. Desta forma pretende-se consolidar a utilização do som como um meio eficaz de monitoramento de processos de usinagem.

Para esta finalidade, dois métodos de monitoramento são propostos e avaliados na classificação de níveis de rugosidade no processo de torneamento: um baseado em GMM e outro em SVM. Detalhes sobre o processo de treinamento e classificação de cada método são discutidos neste capítulo.

7.2 DEFINIÇÃO DOS CRITÉRIOS DE AGRUPAMENTO

Para reconhecer os padrões dos sinais acústicos e correlacioná-los com os valores obtidos de rugosidade, é necessário avaliar-se em quantas categorias distintas os valores de rugosidade podem ser separados.

Utilizando a técnica estatística de Análise Hierárquica de Cluster – Método de Ward (JOHNSON; WICHERN, 2007), observou-se que algumas configurações experimentais não apresentam valores de rugosidade suficientemente diferentes para suportar a hipótese de que seja possível separar em 10 conjuntos de parâmetros de usinagem. Desse modo, algumas configurações experimentais diferentes podem conduzir aos mesmos valores médios de rugosidade. A Figura 7.1 mostra como estes experimentos se agrupam assim como, em alguns casos, eles são perfeitamente separáveis formando três grupos distintos.

Nesta análise, avaliou-se a semelhança entre os experimentos no que tange ao comportamento da rugosidade média (Ra) e à taxa de remoção de material (TRM). Algumas configurações de usinagem resultaram em valores de rugosidade muito similares como, por exemplo, as configurações de usinagem MS_1 e MS_2 , que acabaram sendo classificadas dentro do mesmo grupo ($G1$). O mesmo ocorre com as configurações MS_3 , MS_4 , MS_5 , MS_7 e MS_{10} que formam um segundo grupo ($G2$), e MS_6 , MS_8 e MS_9 que formam o grupo 3 ($G3$). A Tabela 7.1 resume a separação de grupos formada a partir de suas similaridades. A formação do grupo $G1$, por exemplo, pode indicar uma dificuldade em separar a rugosidade resultante das usinagem com parâmetros $MS_1 = \{Vc = 200, f = 0,10 \text{ e } ap = 0,10\}$ e $MS_2 = \{Vc = 240, f = 0,10 \text{ e } ap = 0,10\}$. A partir da Tabela 6.2 é possível notar que os valores de rugosidade pra estas duas configurações são realmente

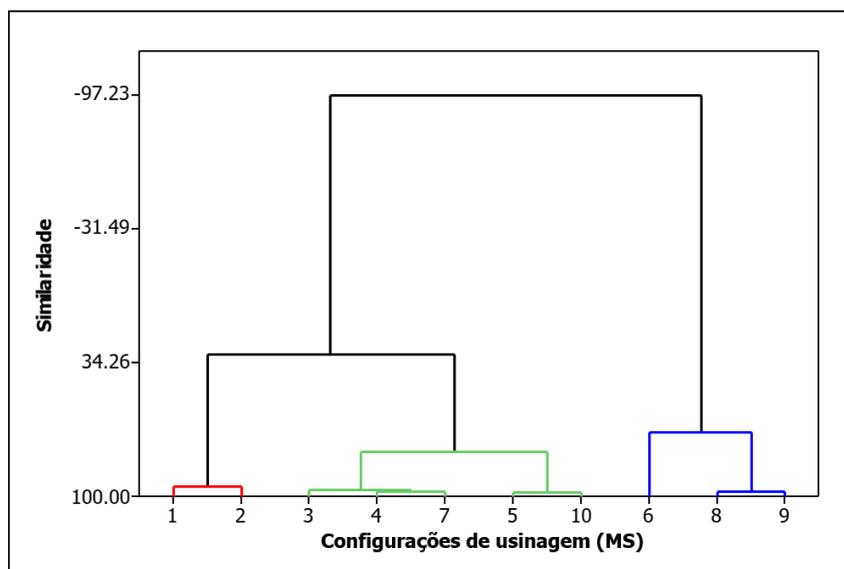


Figura 7.1 – Acoplamento Ward baseado na Distância Euclideana para Ra e TRM .

muito próximas, o que dificultaria a separação das duas condições. Esta mesma observação é válida para os demais grupos formados.

Tabela 7.1 – Grupos formados a partir da análise hierárquica de *cluster* - método de Ward.

Configurações	MS_1	MS_2	MS_3	MS_4	MS_5	MS_6	MS_7	MS_8	MS_9	MS_{10}
Grupo	1	1	2	2	2	3	2	3	3	2

Partindo dos grupos obtidos, pode-se definir três grupos de rugosidade distintos e portanto gerar um conjunto de dados de treinamento e teste a fim de validar os métodos de monitoramento que são apresentados a seguir. Definiu-se pela utilização de um segundo grupo de treinamento e de teste com o objetivo de gerar um experimento de confirmação e garantir a veracidade estatística dos dados. Ambos os conjuntos gerados a partir da aleatorização das amostras podem ser observados na Tabela 7.2.

Cada número da tabela representa o índice da réplica do experimento do qual os arquivos com os sinais acústicos foram selecionados. Vale lembrar que cada replica é composta por 10 parâmetros de usinagem diferentes (veja Apêndice A). Por fim, obteve-se um conjunto de oito arquivos para treinamento e sete arquivos para teste, para cada configuração de usinagem. Como o grupo $G1$ é formado por dois conjuntos de configurações de usinagem, resultou um total de dezesseis arquivos de treinamento e quatorze para teste. No caso do grupo $G2$, por possuir cinco conjuntos de configuração de usinagem, o número de arquivos de treinamento total é quarenta e de teste trinta e cinco. Para o grupo $G3$, o número de arquivos de treinamento é vinte e quatro e de teste, vinte e um.

Os processos de treinamento e teste de cada sistema proposto estão detalhados a seguir.

Tabela 7.2 – Conjuntos de treinamento e teste dos métodos de monitoramento.

Treinamento 1	Teste 1	Treinamento 2	Teste 2
1 ^a	2	1	3
3	4	2	4
5	6	5	7
7	8	6	8
9	10	9	11
11	12	10	12
13	14	13	15
15		14	

^a Valores da tabela indicam o número da replica com os respectivos sinais acústicos a serem utilizados.

7.3 MÉTODO DE MONITORAMENTO BASEADO EM SINAIS ACÚSTICOS

De forma geral, um método de monitoramento baseado em sinais acústicos deve seguir as etapas apresentadas na Figura 7.2. Conforme detalhado no Capítulo 3, é necessário extrair informações do sinal acústico que consigam representar o fenômeno físico que o gerou. Portanto o sinal passa primeiramente por uma etapa de formatação espectral e posteriormente os parâmetros são extraídos. Neste trabalho foram apresentados dois parâmetros que mostraram forte correlação com o processo de usinagem alvo, que foram o espectro de potência, através da utilização de seus coeficientes principais, e os coeficientes Mel-Cepstrais (veja Seções 6.4 e 6.5).

Também na Figura 7.2 é possível identificar que a fase de extração de parâmetros deve ocorrer para ambas as etapas que envolvem um sistema de classificação, que são as etapas de treinamento e de teste. Cada uma destas etapas depende diretamente do algoritmo de classificação utilizado mas de modo geral pode-se dizer que a etapa de treinamento deve ocorrer antes do sistema entrar em operação. Nesta etapa utiliza-se um conjunto de dados razoavelmente grande, chamado de conjunto de treinamento, capaz de representar a população global, ao qual o classificador deverá ser aplicado. As amostras de treinamento são apresentadas ao classificador para que o mesmo possa estimar modelos matemáticos e/ou estatísticos capazes de aprender sobre a natureza das amostras e a forma correta de separá-las.

Uma vez que os modelos estão estimados, pode-se dizer que o sistema está treinado e pronto para ser utilizado. Para concluir sobre o grau de aprendizado do sistema, verifica-se sua taxa de acerto para os dados de treinamento e muitas vezes para outro conjunto de dados chamado conjunto de dados de validação. Uma alta taxa de acerto na classificação destes dados dá indícios de que o sistema terá bom desempenho com dados ainda não

apresentados a ele.

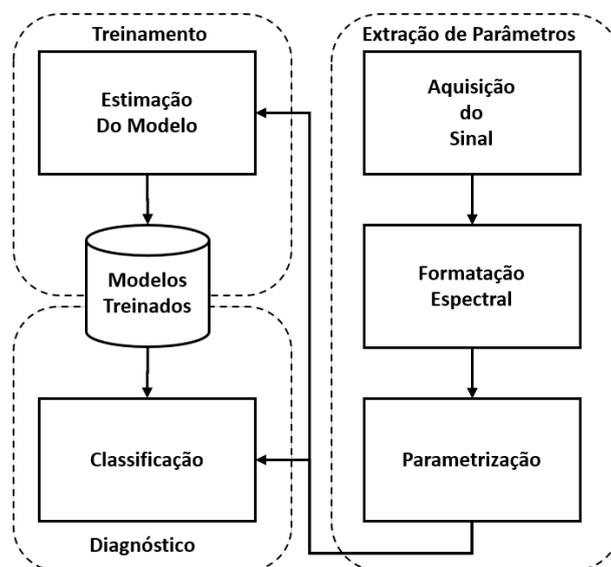


Figura 7.2 – Diagrama geral de um sistema de monitoramento baseado em sinais acústicos.

Dois algoritmos de classificação foram escolhidos para este trabalho, que são o GMM e o SVM. Detalhes sobre o algoritmo de treinamento e como utilizar os modelos treinado para classificação podem ser acessados no Capítulo 4. É importante ressaltar sobre a natureza destes dois classificadores já que o GMM é baseado em decisão probabilística enquanto o SVM trabalha com decisões baseadas em distâncias Euclidianas¹. Esta diferença é importante a fim de validar a utilização do sinal acústico para diferentes sistemas de monitoramento.

A seguir são apresentados os detalhes sobre cada sistema proposto utilizando GMM e SVM, os parâmetros utilizados para os classificadores assim como os resultados obtidos.

7.3.1 Método de Monitoramento Utilizando GMM e Sinais Acústicos

Este método foi primeiramente proposto por Frigieri (2013) e os códigos das implementações para o GMM e o algoritmo utilizado na inicialização, K-means, podem ser acessados nos Apêndices E e F, respectivamente.

No trabalho proposto por Frigieri (2013) foram testadas duas configurações para o GMM: uma utilizando uma mistura de quatro Gaussianas e outra de oito. Porém neste trabalho utilizou-se na mistura apenas o vetor de variâncias, que corresponde à diagonal principal da matriz de covariância. A fim de melhor avaliar o desempenho do classificador GMM para os parâmetros propostos, duas modalidades para a matriz de covariância foram testadas, a matriz completa, denominada ‘full’, e a matriz achatada, denominada ‘tied’

¹ É importante ressaltar que classificadores baseados em SVM podem utilizar decisões probabilísticas, porém, esta abordagem não faz parte do escopo deste trabalho.

(veja detalhes na Seção 4.2). O número de Gaussianas na mistura foi avaliado para estes dois casos de matriz de covariância.

7.3.1.1 Procedimento de Treinamento dos Modelos de Mistura de Gaussianas

Conforme foi apresentado na Seção 7.2, o conjunto de dados foi separado em dois conjuntos distintos, um conjunto de treinamento e outro de teste, conforme Tabela 7.2. Neste caso, foram criados dois conjuntos de treinamento a fim de validar estatisticamente os resultados.

Desta forma, para treinar os modelos GMM para cada conjunto de treinamento, foi utilizado o procedimento apresentado na Figura 7.3. Para cada um dos G_i grupos obtidos na Tabela 7.1, extrai-se os parâmetros de todos os sinais. A partir dos parâmetros obtidos, cada amostra do conjunto de treinamento é apresentada sequencialmente ao algoritmo *EM* a fim de estimar o modelo de mistura de Gaussianas que representará o respectivo grupo (detalhes do funcionamento do algoritmo *EM* podem ser vistos na Seção 4.2). Ao fim da sequência de treinamento, o modelo está pronto para a fase de diagnósticos e seus parâmetros são armazenados: a média $\bar{\mu}_m$, variância $\bar{\sigma}_m^2$ e peso \bar{w}_m para cada Gaussiana m da mistura. Este procedimento se repete para cada um dos grupos ao qual se deseja estimar.

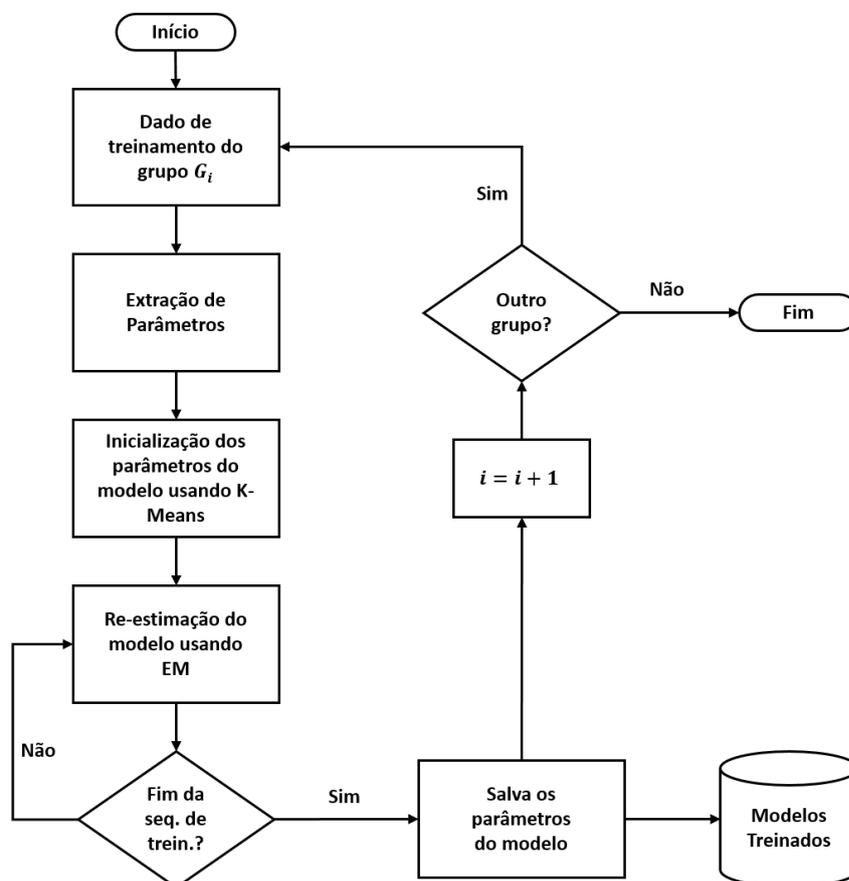


Figura 7.3 – Fluxo de treinamento do sistema baseado em GMM.

7.3.1.2 Procedimento de Diagnóstico Utilizando Modelos de Mistura de Gaussianas

Uma vez que os modelos foram estimados para cada grupo, o sistema está apto para a fase de operação ou diagnóstico. No caso do sistema baseado em GMM, a fim de se estimar a qual grupo uma determinada amostra pertence, deve-se seguir o procedimento apresentado na Figura 7.4. Para uma determinada amostra, verifica-se a probabilidade dele pertencer a cada um dos modelos previamente treinados, através do cálculo do logaritmo da probabilidade a *priori* obtida através da Equação (4.13), conforme apresentado na Seção 4.2.2. O maior valor de probabilidade resultante indica o modelo ao qual a amostra tem a maior chance de pertencer.

A fim de validar o desempenho do sistema, um conjunto de teste foi definido para cada conjunto de treinamento, de acordo com a Tabela 7.2. Cada amostra do conjunto de teste foi diagnosticada pelo sistema treinado e o grupo identificado foi comparado ao verdadeiro grupo ao qual a mesma pertencia. Finalmente calculou-se o número de amostras corretamente classificadas dividido pelo número total de amostras testadas, resultando na métrica de desempenho definida como *Precisão*. A equação para obtenção desta métrica pode ser formalmente definida como

$$\text{Precisão \%} = \frac{\text{número de amostras classificadas corretamente}}{\text{número total de amostras verificadas}} \quad (7.1)$$

Portanto, quanto maior o valor da *Precisão*, melhor é o desempenho do sistema na tarefa de monitoramento.

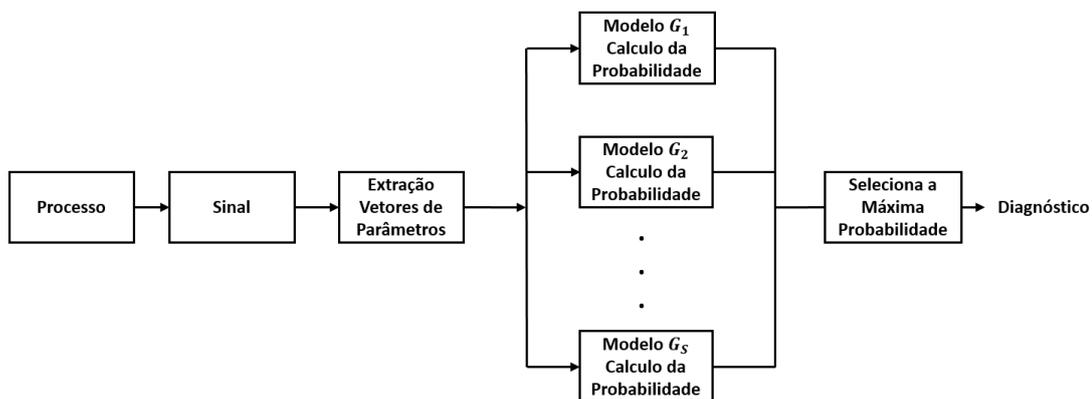


Figura 7.4 – Fluxo de diagnóstico do sistema baseado em GMM.

7.3.1.3 Resultados Obtidos para o Sistema Baseado em GMM

Para avaliar o sistema proposto baseado em GMM, uma matriz experimental foi gerada considerando os seguintes fatores: tipo de matriz de covariância (*'full'*, *'tied'*); número de Gaussianas por modelo (quatro ou oito); parâmetros do sinal (*'mfcc'*, *'pca'*, *'mix'*). Em relação à matriz de covariância, o fator *'full'* representa a matriz de covariância completa, enquanto o fator *'tied'* representa a matriz de covariância amarrada ao

número de componentes do modelo. O fator ‘*pca*’ corresponde aos três primeiros escores que apresentaram a maior correlação com os parâmetros de usinagem e rugosidade, conforme Tabela 6.5 apresentada na Seção 6.4. O fator ‘*mfcc*’ corresponde aos dois primeiros coeficientes mel que apresentaram a maior correlação com os parâmetros de usinagem e rugosidade, conforme Tabela 6.6 apresentada na Seção 6.5. Já o fator ‘*mix*’ corresponde à combinação dos parâmetros ‘*mfcc*’ e ‘*pca*’. A Tabela 7.3 apresenta o desempenho médio de dez execuções para cada uma das combinações dos fatores do experimento.

Tabela 7.3 – Resultados do experimento de validação do sistema baseado em GMM.

Replica	Num. Gaus. (<i>Ga</i>)	Matriz de Cov. (<i>Cv</i>)	Parâmetros (<i>Pa</i>)	Precisão (%)
1	4	full	pca	92,85
1	4	full	mfcc	97,14
1	4	full	mix	92,85
1	4	tied	pca	97,28
1	4	tied	mfcc	97,14
1	4	tied	mix	97,28
1	8	full	pca	81,85
1	8	full	mfcc	88,85
1	8	full	mix	81,85
1	8	tied	pca	95,71
1	8	tied	mfcc	95,57
1	8	tied	mix	95,71
2	4	full	pca	87,14
2	4	full	mfcc	92,85
2	4	full	mix	94,28
2	4	tied	pca	95,71
2	4	tied	mfcc	95,14
2	4	tied	mix	94,28
2	8	full	pca	87,14
2	8	full	mfcc	73,00
2	8	full	mix	86,42
2	8	tied	pca	94,28
2	8	tied	mfcc	93,00
2	8	tied	mix	95,00

Utilizando a análise de variância sobre os resultados obtidos, é possível identificar que os fatores *Ga* e *Cv* são estatisticamente significativos enquanto *Pa* não contribuiu de forma efetiva no desempenho do sistema (P-value > 5%). Tanto os coeficientes Mel-Cepstrais como o escores dos componentes principais, ou mesmo a combinação de ambos (*mix*), resultam em média no mesmo desempenho na classificação dos grupos de rugosidade. Este efeito é ilustrado pela Figura 7.5. Nenhuma das interações entre os fatores se mostraram estatisticamente significativas e portanto não foram avaliadas.

Utilizando análise dos efeitos principais também é possível identificar que utilizar quatro Gaussianas na mistura tem um efeito positivo no desempenho do sistema, aumen-

tando sua taxa de acerto. O mesmo acontece quando a matriz de covariância dos modelos GMM é definida como ‘tied’. Embora se esperaria que utilizar a matriz de covariância completa resultaria um melhor desempenho, em geral para pequenos conjuntos de dados, existe a possibilidade de *overfitting*² resultando em uma má generalização para os dados de teste.

Apesar dos parâmetros extraídos dos sinais não serem estatisticamente significativos para o aumento de desempenho do sistema, é possível notar que os coeficientes Mel-Cepstrais apresentam uma leve superioridade quando comparados aos escores dos componentes principais. Seguindo esta linha, o desempenho resultante da combinação de todos os parâmetros se sobressai sobre a utilização dos mesmos de forma isolada. Portanto, é possível identificar a configuração ótima de operação deste sistema com os modelos de mistura formados por quatro Gaussianas, treinados a partir da matriz de covariância definida como ‘tied’ e utilizando todos os parâmetros para treinamento e diagnóstico.

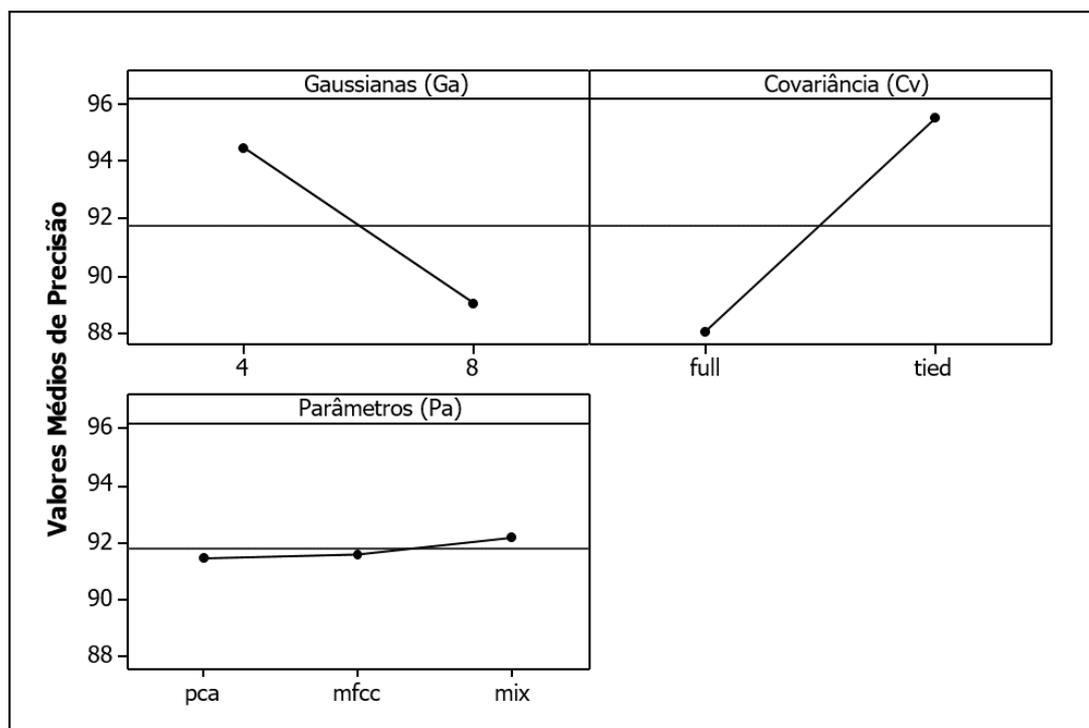


Figura 7.5 – Gráfico com os efeitos principais dos fatores sobre a precisão do sistema baseado em GMM.

² Na estatística e na aprendizagem de máquina, uma das tarefas mais comuns consiste em ajustar um "modelo" a um conjunto de dados de treinamento, de modo que se possa fazer previsões confiáveis sobre dados não treinados de forma generalizada. Caso ocorra um treinamento excessivo deste modelo, pequenas flutuações nos dados de treinamento geram uma reação exagerada do modelo, resultando em uma má generalização para os dados de teste e conseqüentemente um mau desempenho preditivo. Para este treinamento excessivo dá-se o nome de *overfitting*.

7.3.2 Método de Monitoramento Utilizando SVM e Sinais Acústicos

Um método de monitoramento baseado em Máquinas de Vetores de Suporte é apresentado neste trabalho com o objetivo de validar a utilização dos parâmetros extraídos dos sinais acústicos. Outro ponto importante a ser avaliado com este novo método é a natureza do classificador que melhor se adapta ao problema de monitoramento da rugosidade.

O algoritmo utilizado para o treinamento dos modelos SVM assim como os algoritmos de diagnósticos utilizados foram baseados na libSVM (CHANG; LIN, 2011). Detalhes em relação ao procedimento de treinamento e diagnóstico para este sistema são apresentados a seguir.

7.3.2.1 Procedimento de Treinamento das Máquinas de Vetores de Suporte

Os mesmos conjuntos de treinamento e teste da Tabela 7.2 foram utilizados para a avaliação do sistema baseado em SVM. Como o problema de diagnóstico do grupo de rugosidade contempla três grupos distintos, é necessária a utilização de técnicas de decisão para problemas multi classes. Desta forma, ambas as técnicas *OAA* e *OAO* apresentadas na Seção 4.3.2 foram avaliadas.

Para treinar o sistema SVM a fim de operar no modo *OAA*, deve-se efetuar o procedimento de treinamento apresentado na Figura 7.6. Para cada um dos G_i grupos obtidos na Tabela 7.1, extraem-se os parâmetros de todos os sinais. O treinamento do SVM utiliza todas as amostras para estimar os modelos que representarão cada grupo. Porém, para estimar cada modelo i rotulam-se todas as amostras deste grupo como +1 e as amostras dos demais grupos como -1, resultando um modelo capaz de separar este grupo dos demais. Este procedimento resulta um número total de modelos igual ao total de grupos a serem separados (detalhes do funcionamento do algoritmo podem ser vistos na Seção 4.3). Ao fim da sequência de treinamento, os modelos estão prontos para a fase de diagnósticos e seus parâmetros, que são os vetores de suporte, são armazenados.

Outro método de treinamento considerado neste trabalho é o *OAO*. Nesta técnica cria-se um modelo SVM para separar cada combinação de dois grupos de um total de I , resultando $I(I - 1)/2$ modelos diferentes. Esta técnica tem como deficiência a criação de um número grande de modelos, o que pode impactar no tempo de processamento durante a fase de diagnóstico. Como este trabalho não tem como escopo avaliar os classificadores em relação ao tempo de classificação, esta técnica será utilizada durante a avaliação do sistema baseado em SVM. O procedimento necessário para o treinamento utilizando a técnica *OAO* é apresentado na Figura 7.7.

As máquinas de vetores de suporte foram criadas inicialmente para a separação linear de grupos. Porém, a evolução da técnica desenvolveu outras funções de separação

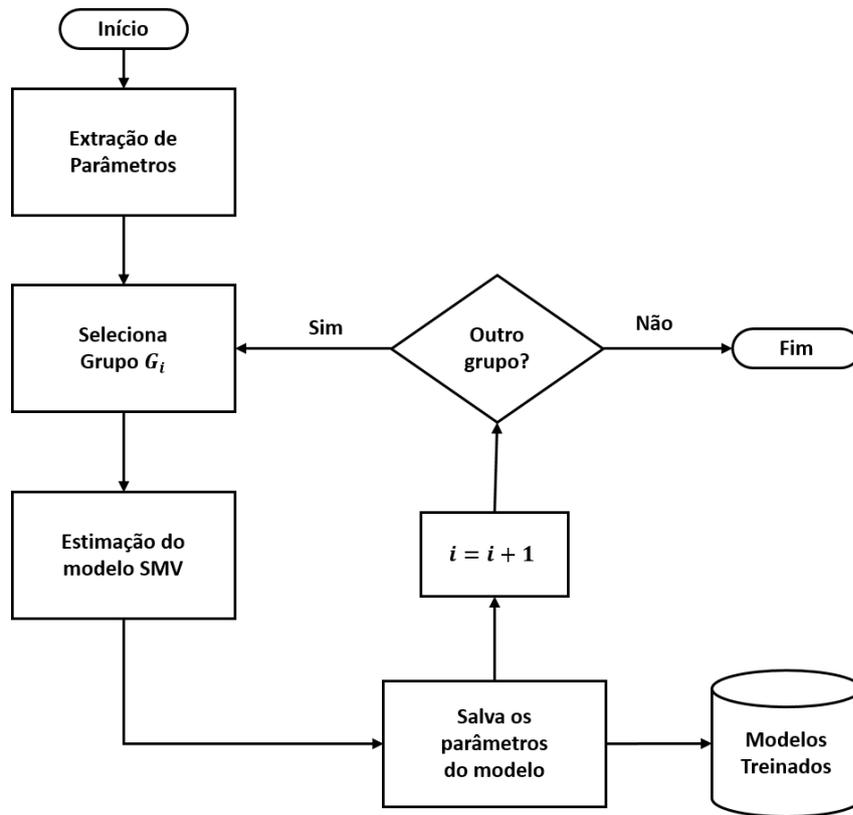


Figura 7.6 – Fluxo de treinamento do sistema baseado em SVM para o método *OAA*.

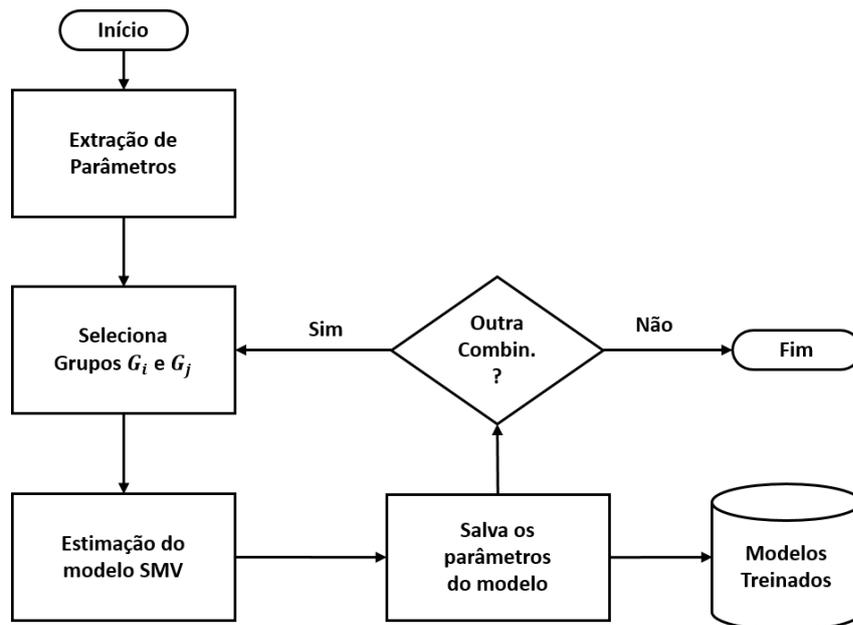


Figura 7.7 – Fluxo de treinamento do sistema baseado em SVM para o método *OAO*.

dando mais poderes de classificação ao SVM. A estas funções deu-se o nome de *kernels*. A fim de validar o ganho de desempenho destas novas técnicas, um *kernel* baseado em funções de base radiais (do inglês *Radial Basis Function* – RBF) foi avaliado neste trabalho e comparado ao *kernel* linear.

7.3.2.2 Procedimento de Diagnóstico Utilizando Máquina de Vetores de Suporte

Mais uma vez, com os modelos estimados para cada grupo o sistema está apto para a fase de operação ou diagnóstico. No caso do sistema com SVM baseado na técnica *OAA*, a ideia de classificação é mais direta, já que existe um modelo para cada grupo, ou seja, a fim de se estimar a qual grupo uma determinada amostra pertence, deve-se verificar qual dos modelos resultará +1, já que sinaliza a classificação de um grupo em relação a todos os demais. No caso do sistema que utiliza *OAO*, o procedimento é um pouco mais complicado e, a fim de diagnosticar uma determinada amostra, utiliza-se o procedimento apresentado na Figura 7.8. Neste procedimento computa-se um voto para cada modelo que identificar o grupo i . Ao final da avaliação de todos os $I(I - 1)/2$ modelos, aquele que obteve o maior número de votos será o grupo ao qual a amostra será atribuída.

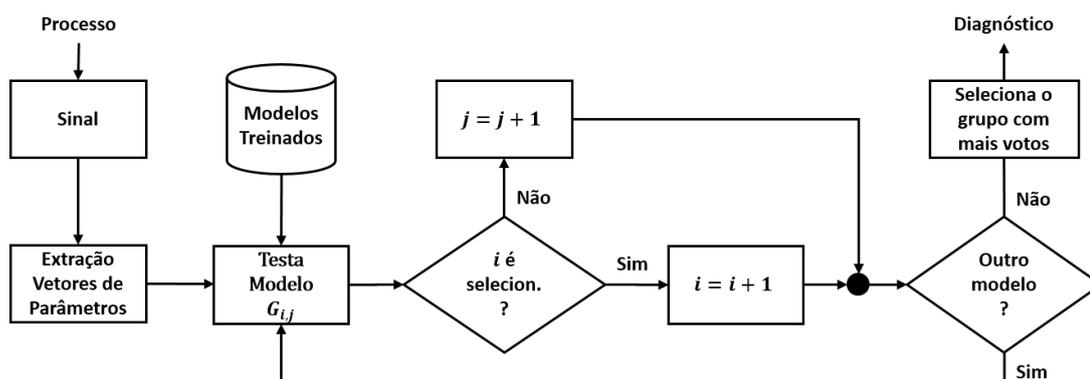


Figura 7.8 – Fluxo de diagnóstico do sistema baseado em SVM.

A fim de validar o desempenho do sistema, utilizou-se os conjuntos de teste definidos na Tabela 7.2 e o mesmo procedimento adotado para a avaliação do sistema baseado em GMM. Ou seja, para cada amostra do conjunto de teste, o modelo previsto pelo sistema treinado foi verificado e comparado com o verdadeiro grupo ao qual pertencia a amostra. O número de amostras corretamente classificadas dividido pelo número total de amostras testadas resultou na *Precisão* do sistema, métrica de desempenho definida anteriormente na Equação 7.1.

7.3.2.3 Resultados Obtidos para o Sistema Baseado em SVM

Para avaliar o sistema proposto baseado em SVM, uma matriz experimental foi gerada considerando os seguintes fatores: tipo de *kernel* (*linear*, *rbf*); técnica multi-classe (*oao*, *oaa*); parâmetros do sinal (*mfcc*, *pca*, *mix*). Assim como foi definido

para o GMM, o fator ‘*pca*’ corresponde ao três primeiros escores, o fator ‘*mfcc*’ corresponde ao dois primeiros coeficientes mel e o fator ‘*mix*’ corresponde à combinação dos parâmetros ‘*mfcc*’ e ‘*pca*’. A Tabela 7.4 apresenta o desempenho médio de dez execuções para cada uma das combinações dos fatores do experimento.

Tabela 7.4 – Resultados do experimento de validação do sistema baseado em SVM.

Replica	Kernel (<i>Kn</i>)	Téc. Multi-classe (<i>Mc</i>)	Parâmetros (<i>Pa</i>)	Precisão
1	linear	oao	pca	97,14
1	linear	oao	mfcc	95,71
1	linear	oao	mix	97,14
1	linear	oaa	pca	97,14
1	linear	oaa	mfcc	95,71
1	linear	oaa	mix	97,14
1	rbf	oao	pca	100,00
1	rbf	oao	mfcc	98,57
1	rbf	oao	mix	100,00
1	rbf	oaa	pca	100,00
1	rbf	oaa	mfcc	98,57
1	rbf	oaa	mix	100,00
2	linear	oao	pca	97,14
2	linear	oao	mfcc	97,14
2	linear	oao	mix	98,57
2	linear	oaa	pca	97,14
2	linear	oaa	mfcc	97,14
2	linear	oaa	mix	98,57
2	rbf	oao	pca	100,00
2	rbf	oao	mfcc	100,00
2	rbf	oao	mix	100,00
2	rbf	oaa	pca	100,00
2	rbf	oaa	mfcc	100,00
2	rbf	oaa	mix	100,00

Utilizando a análise de variância sobre os resultados obtidos, foi possível identificar que os fatores *Kn* e *Pa* são estatisticamente significativos enquanto a técnica multi-classe *Mc* e as interações entre os fatores não contribuíram de forma efetiva para o desempenho do sistema, ou seja, não são estatisticamente significativos (P-value > 5%). A efetividade do *kernel* RBF foi comprovada para este problema devido aos níveis de precisão resultantes serem muito maiores do que os obtidos pelo sistema baseado no *kernel* linear. Vale ressaltar também que os parâmetros extraídos do sinal acústico resultaram diferentes níveis de desempenho ao sistema operando com SVM, diferentemente do comportamento ocorrido com o GMM. Os coeficientes Mel-Cepstrais apresentaram os menores níveis de desempenho para este sistema, quando utilizados sozinhos. Porém, ao serem utilizados em conjunto com os escores dos componentes principais, a um aumento considerável na precisão do sistema. Estes efeitos são ilustrados pela Figura 7.9. Logo, a configuração ótima

de operação deste sistema utiliza *kernel* RBF, e é alimentado por todos os parâmetros, os coeficientes Mel-Cepstrais e os escores dos componentes principais (*mix*), para treinamento e diagnóstico. Para a configuração ótima deste sistema pode-se optar por qualquer uma das técnicas multi-classe: *OAO* ou *OAA*.

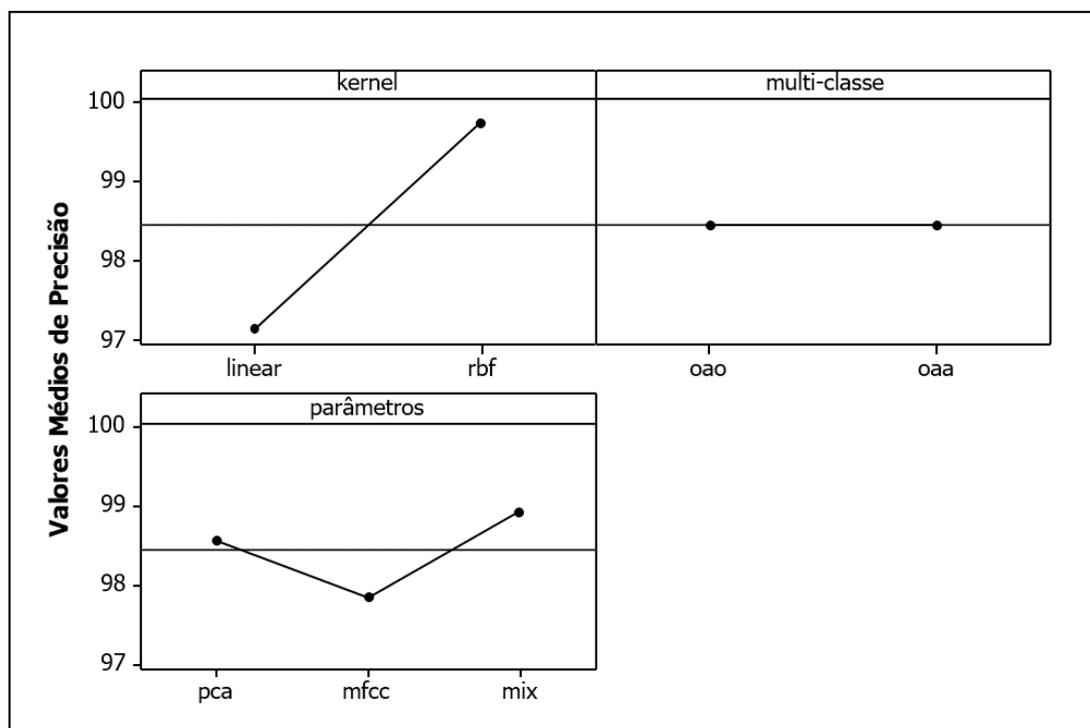


Figura 7.9 – Gráfico com os efeitos principais dos fatores sobre a precisão do sistema baseado em SVM.

7.4 SELEÇÃO DAS CONFIGURAÇÕES ÓTIMAS DO SISTEMA DE MONITORAMENTO BASEADO EM SINAIS ACÚSTICOS

No Capítulo 7.3 foram apresentadas duas propostas de sistemas de monitoramento baseados em sinais acústicos: um sistema baseado em classificador GMM e outro em SVM. Como pôde ser visto, cada sistema proposto é composto por diversos parâmetros de configuração, o que demandou a utilização da análise fatorial para a determinação da configuração ótima de cada sistema.

Com base nos experimentos realizados, foi possível notar que ambos os sistemas apresentaram desempenho satisfatório, nos quais algumas configurações chegaram a resultar uma precisão de 100%. A fim de se obter qual o sistema ótimo para o monitoramento da rugosidade através dos sinais acústicos, um novo projeto experimental foi delineado a partir das configurações ótimas obtidas para cada um dos sistemas propostos. Considerando duas réplicas e os conjuntos de treinamento e teste apresentados na Tabela 7.2, o projeto experimental pode ser definido conforme Tabela 7.5, em que *gmm-opt* representa o

sistema baseado em GMM com a seguinte configuração ótima: quatro gaussianas, matriz de covariância ‘tied’; e *svm-opt* representa o sistema baseado em SVM com a seguinte configuração ótima: *kernel* RBF, técnica multi classe *OAA*. Apesar da utilização de ambos os parâmetros MFCC e PCA (*mix*) terem resultado os maiores valores de precisão para ambos os sistemas, foram utilizados para esta análise os parâmetros separados com o intuito de avaliar qual o impacto de cada parâmetro no desempenho dos sistemas ótimos.

Tabela 7.5 – Resultados do experimento de validação dos sistemas operando com configuração ótima.

Replica	Sistema	Parâmetro	Precisão (%)
1	gmm-opt	pca	97,14
1	svm-opt	pca	100,00
1	gmm-opt	mfcc	97,14
1	svm-opt	mfcc	98,57
2	gmm-opt	pca	95,71
2	svm-opt	pca	100,00
2	gmm-opt	mfcc	95,28
2	svm-opt	mfcc	100,00

Através da análise fatorial foi possível identificar que apenas o fator que representa o tipo de sistema se apresentou estatisticamente significativo (P-value < 5%). Tanto o fator que representa o parâmetro como a interação entre os fatores não são estatisticamente significativos, conforme ilustrado pela Figura 7.10. O modelo resultante apresentou um valor de ajuste R-Sq(adj) igual a 75,03% que, apesar de não representar um valor de coeficiente de determinação elevado, apresenta valores para o teste de normalidade dos resíduos (Anderson-Darling e P-value > 5%) com evidências suficientes para afirmar que os resíduos seguem uma distribuição normal, conforme apresentado pela Figura 7.11.

A análise dos efeitos principais, apresentada na Figura 7.12, demonstra que o sistema baseado em SVM se sobressai em comparação com o sistema baseado em GMM, ou seja, resulta maiores valores de precisão. Outro resultado importante é a eficiência do PCA sobre o MFCC, pois ambos os sistemas utilizando PCA apresentaram precisões maiores que os sistemas baseados em MFCC, apesar da diferença não ser significativa. Estas observações podem ser comprovadas através da otimização do modelo obtido, resultando o sistema *svm-opt*, operando com o parâmetro *pca*, conforme ilustrado pela Figura 7.13.

O melhor desempenho do PCA sobre o MFCC pode significar que o aparelho auditivo humano filtra características importantes sobre o processo, mais precisamente sobre a rugosidade resultante da usinagem, contidas no som. Desta forma, apesar de o MFCC resultar um bom desempenho de classificação no método de monitoramento proposto, esta técnica de extração de características baseada na modelagem do aparelho auditivo humano não é a melhor escolha para o monitoramento de processos.

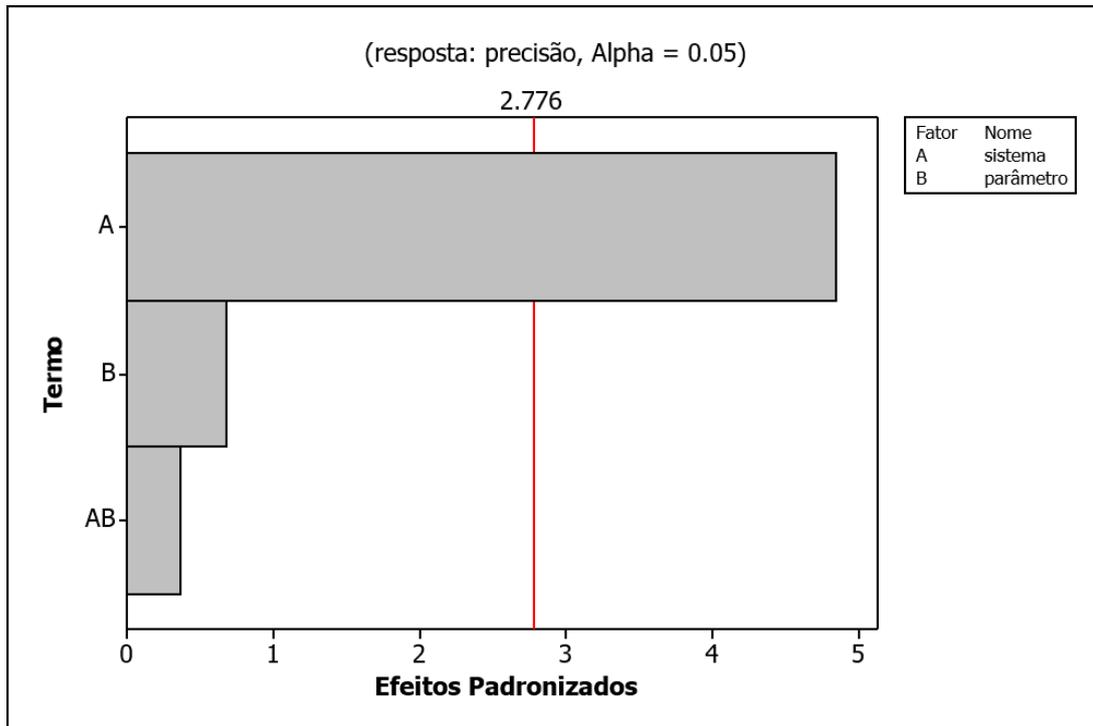


Figura 7.10 – Gráfico de Pareto dos efeitos padronizados para os termos do modelo.

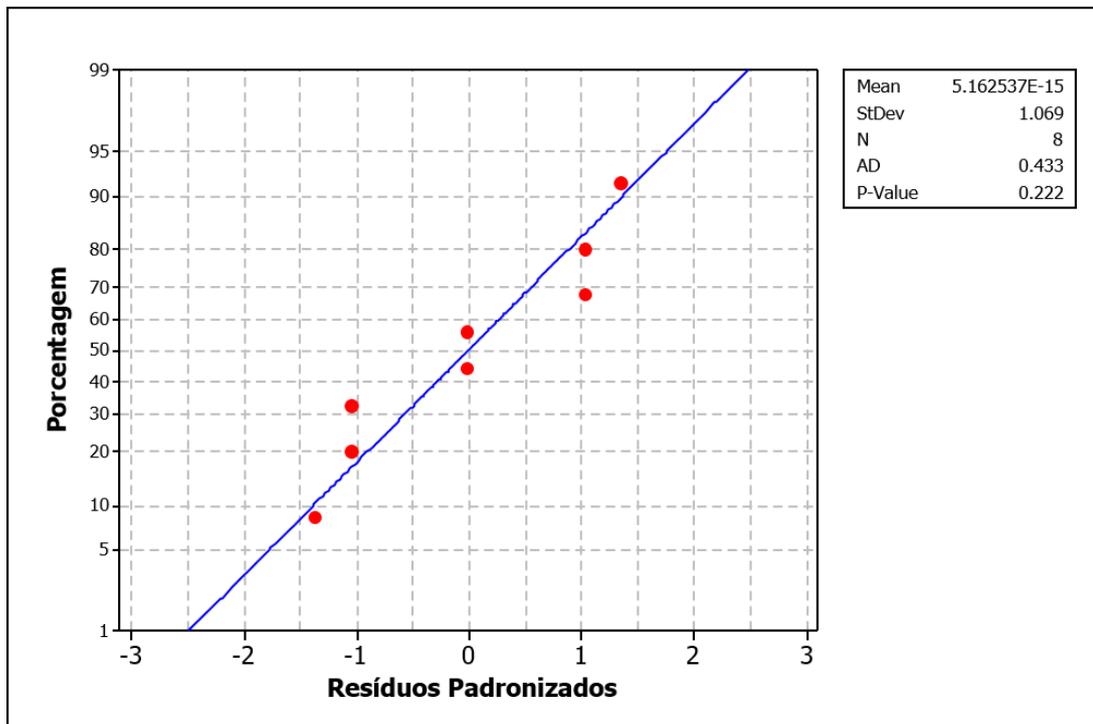


Figura 7.11 – Teste de normalidade dos resíduos padronizados.

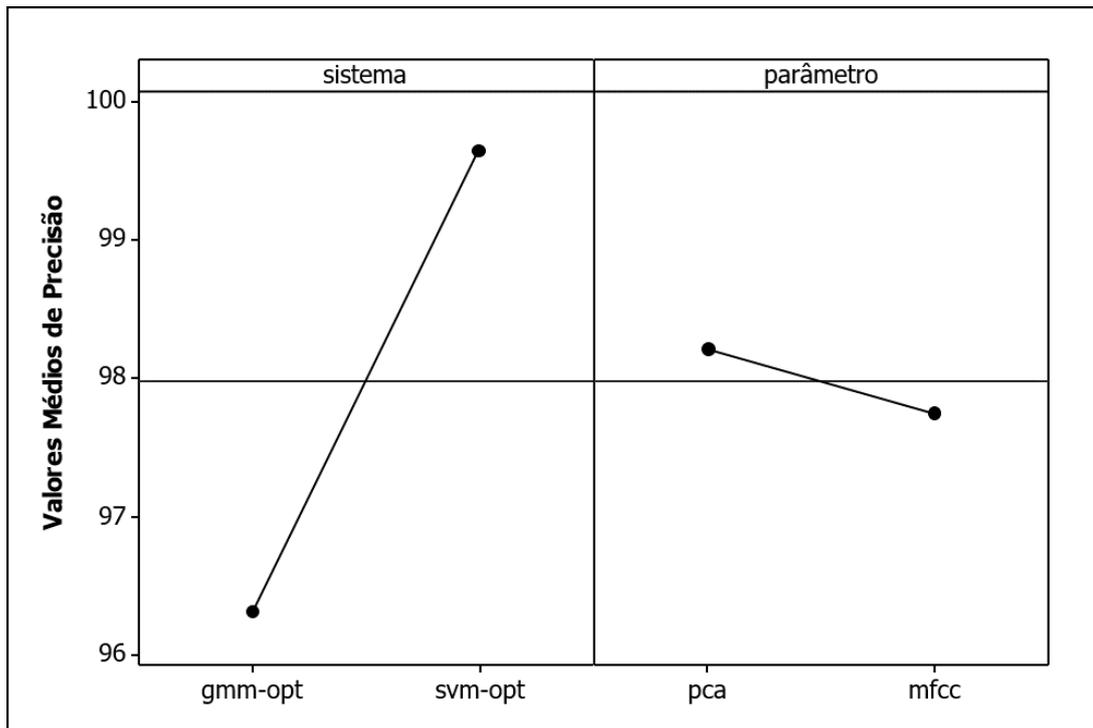


Figura 7.12 – Gráfico com os efeitos principais dos fatores sobre a precisão dos sistemas baseados em GMM e SVM.

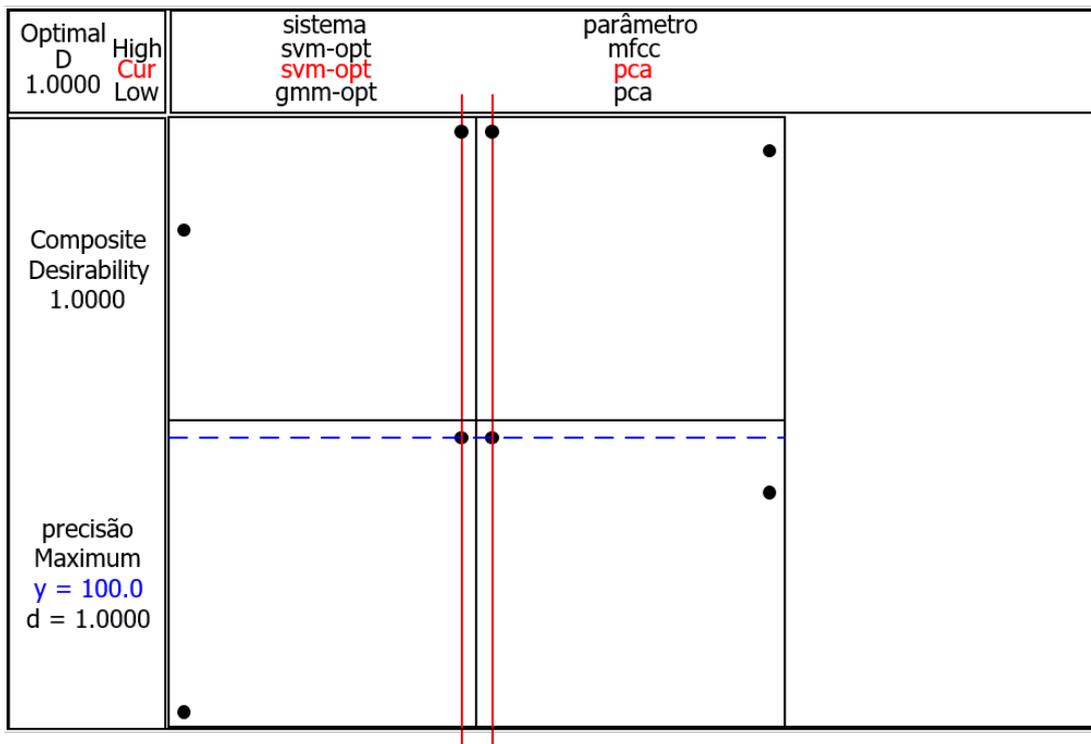


Figura 7.13 – Gráfico com a resposta da otimização para o modelo obtido.

7.5 CONSIDERAÇÕES FINAIS

O presente capítulo teve como objetivo avaliar a utilização dos parâmetros extraídos do som que apresentaram alta correlação com os parâmetros do processo a serem monitorados, através de sistemas baseados em classificadores GMM e SVM.

A partir desta análise foi possível confirmar a eficácia de ambos os parâmetros (PCA e MFCC) no monitoramento da rugosidade resultante do processo de torneamento, obtendo precisões de classificação acima de 95%, podendo atingir 100% operando sob configuração ótima. Outra resultando importante deste capítulo foi a comparação de dois classificadores distintos sendo que o SVM apresentou, de forma geral, desempenho superior ao GMM nesta tarefa de classificação.

Uma parte das conclusões obtidas neste capítulo resultou em um artigo com o título “A mel-frequency cepstral coefficient-based approach for surface roughness diagnosis in hard turning using acoustic signals and gaussian mixture model”, publicado na revista *Applied Acoustics*, volume 113, páginas 230-237, em 2016 ([FRIGIERI *et al.*, 2016b](#)). O artigo pode ser visualizado no Anexo B.

8 CONCLUSÃO

Este trabalho teve como objetivo investigar os sinais acústicos na frequência audível (som) como uma fonte capaz de identificar características do processo de usinagem, em tempo-real, para sua aplicação em novos métodos de monitoramento. A fim de validar esta ideia, dois parâmetros extraídos do sinal acústico resultante de um processo de torneamento foram investigados: os coeficientes Mel-cepstrais (MFCC) e os escores dos componentes principais da energia do sinal.

Através da análise de correlação foi possível identificar quais parâmetros extraídos do sinal conseguem representar características do processo, como por exemplo, os coeficientes mel c_1 , c_2 e c_5 , que apresentaram alta correlação com todas as rugosidades resultantes e com a velocidade de corte, respectivamente. Os escores dos componentes principais PC_1 e PC_3 também apresentaram forte correlação com as rugosidades e a velocidade de corte. Outro resultado importante desta análise foi a forte correlação de ambos os parâmetros com a taxa de remoção de material (TRM), o que dá indícios de que a energia específica de corte é provavelmente transformada em outras formas de energia, neste caso o som audível emitido pelo processo, comprovando que o sinal acústico pode ser utilizado para o monitoramento de processos.

A fim de validar os resultados obtidos pela análise de correlação dos parâmetros dos sinais acústicos, foram propostos dois sistemas de monitoramento da rugosidade do processo de torneamento: um sistema baseado no classificador GMM e outro em classificador SVM. Através da análise fatorial ambos os sistemas foram testados para diferentes configurações e combinações de parâmetros dos sinais acústicos de forma que foi possível identificar a configuração ótima para cada um. No caso do sistema baseado em GMM, sua configuração ótima é um modelo de mistura composto de quatro gaussianas que utiliza a matriz de covariância adaptada para utilizar a mesma matriz para todos os componentes ('tied'). Para o sistema baseado em SVM, a configuração ótima obtida utiliza um *kernel* de base radial com método de classificação OAA (*one-against-all*). Para os dois sistemas a combinação de ambos os parâmetros, MFCC a PCA, resultaram nos maiores níveis de precisão.

A partir das configurações ótimas de cada sistema de monitoramento proposto, foi realizado um novo experimento a fim de avaliar qual parâmetro do sinal acústico e qual classificador apresentavam o melhor desempenho na classificação da rugosidade. Através da análise fatorial concluiu-se que, de forma geral, ambos os parâmetros resultaram em ótimo desempenho de classificação, porém, os escores dos componentes principais apresentando uma leve vantagem sobre o MFCC. Este fato apresenta fortes indícios de que a forma como ouvimos o som filtra informações que são importantes para a caracteriza-

ção do processo, e que o mesmo não é limitado pelo PCA. Esta observação caracteriza uma das maiores contribuições deste trabalho em relação aos resultados apresentados em Frigieri (2013).

Também podemos destacar que a utilização do SVM como técnica de classificação tornou o sistema de monitoramento mais preciso, chegando a alcançar uma taxa de certo de 100% operando na configuração ótima. Estes resultados corroboram com o princípio de que os parâmetros dos sinais acústicos podem ser utilizados para o monitoramento de processos de torneamento duro.

Vale ainda ressaltar que este trabalho apresenta uma proposta de método de monitoramento da qualidade do processo de torneamento, através da identificação de níveis de rugosidade da peça, diferente da maioria dos trabalhos encontrados na literatura, que têm como foco o monitoramento do desgaste da ferramenta. Os resultados apresentados comprovam que o sinal acústico na faixa audível pode ser utilizado diretamente para o monitoramento da qualidade, destacando outra importante contribuição.

Algumas investigações podem ser realizadas em trabalhos futuros a fim de consolidar ainda mais a utilização dos sinais acústicos em monitoramento de processos:

- Avaliar como o desgaste da ferramenta pode influenciar no método proposto;
- Avaliar a influência de ruído proveniente de outras máquinas, ou mesmo do ambiente, no método proposto;
- Testar o método em outros processos de manufatura, tais como, soldagem, micro-fresamento, retificação, furação etc.;
- Testar o método proposto para o monitoramento do desgaste da ferramenta;
- Avaliar outras técnicas de classificação como, por exemplo, as Redes Neurais, Modelos Ocultos de Markov (HMM), entre outras;
- Avaliar outros parâmetros do sinal acústico;
- Avaliar outras combinações com diferentes materiais, equipamentos ou ferramentas;

REFERÊNCIAS

- AGUIAR, P. R.; CRUZ, C. E. D.; PAULA, W. C. F.; BIANCHI, E. C. Predicting surface roughness in grinding using neural networks. *Advances in Robotics, Automation and Control*, n. October, p. 33–44, 2008. Citado 2 vezes nas páginas 10 e 12.
- AI, C. S.; SUN, Y. J.; HE, G. W.; ZE, X. B.; LI, W.; MAO, K. The milling tool wear monitoring using the acoustic spectrum. *The International Journal of Advanced Manufacturing Technology*, v. 61, n. 5-8, p. 457–463, nov 2011. ISSN 0268-3768. Citado 4 vezes nas páginas 1, 10, 12 e 13.
- AIZERMAN, M. A.; BRAVERMAN, E. A.; ROZONOER, L. Theoretical foundations of the potential function method in pattern recognition learning. *Automation and Remote Control*, v. 25, p. 821–837, 1964. Citado na página 41.
- ANDOH, P. Y.; DAVIS, F.; ANTONIO, J. Effects of cutting parameters on acoustic emission signal response during drilling of laminated composites using factorial design method. *Journal of Science and Technology (Ghana)*, v. 27, n. 3, p. 98–106, apr 2008. ISSN 0855-0395. Disponível em: <<http://www.ajol.info/index.php/just/article/view/33062>>. Citado na página 1.
- BARKANA, B. D.; UZKENT, B. Environmental noise classifier using a new set of feature parameters based on pitch range. *Applied Acoustics*, Elsevier Ltd, v. 72, n. 11, p. 841–848, 2011. ISSN 0003682X. Disponível em: <<http://dx.doi.org/10.1016/j.apacoust.2011.05.008>>. Citado 2 vezes nas páginas 33 e 34.
- BERTRAND, W. M. J.; FRANSOO, J. C.; BERTRAND, J. W. M.; FRANSOO, J. C. Operations management research methodologies using quantitative modeling. *International Journal of Operations and Production Management*, v. 22, n. 2, p. 241–264, 2002. ISSN 0144-3577. Disponível em: <<http://www.emeraldinsight.com/10.1108/01443570210414338>>. Citado na página 5.
- BHASKARAN, J.; MURUGAN, M.; BALASHANMUGAM, N.; CHELLAMALAI, M. Monitoring of hard turning using acoustic emission signal. *Journal of Mechanical Science and Technology*, v. 26, n. 2, p. 609–615, apr 2012. ISSN 1738-494X. Disponível em: <<http://www.springerlink.com/index/10.1007/s12206-011-1036-1>>. Citado 3 vezes nas páginas 1, 10 e 12.
- BIN, G.; GAO, J.; LI, X.; DHILLON, B. Early fault diagnosis of rotating machinery based on wavelet packets - empirical mode decomposition feature extraction and neural network. *Mechanical Systems and Signal Processing*, v. 27, p. 696–711, feb 2012. ISSN 08883270. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0888327011003207>>. Citado 4 vezes nas páginas 10, 12, 13 e 15.
- BISHOP, C. *Neural Networks for Pattern Recognition*. [S.l.]: Oxford University Press., 1995. 504 p. ISBN 9780198538646. Citado na página 42.
- BISHOP, C. M. *Pattern Recognition and Machine Learning*. [s.n.], 2006. 738 p. ISSN 10179909. ISBN 9780387310732. Disponível em: <<http://www.library.wisc->

[.edu/selectedtocs/bg0137.pdf](#)>. Citado 7 vezes nas páginas 21, 24, 26, 27, 29, 32 e 34.

BONIFACIO, M.; DINIZ, A. Correlating tool wear, tool life, surface roughness and tool vibration in finish turning with coated carbide tools. *Wear*, v. 173, n. 1-2, p. 137–144, apr 1994. ISSN 00431648. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/0043164894902666>>. Citado na página 1.

BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. A Training Algorithm for Optimal Margin Classifiers. *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory*, p. 144–152, 1992. ISSN 0-89791-497-X. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.21.3818>>. Citado na página 41.

BOUTROS, T.; LIANG, M. Detection and diagnosis of bearing and cutting tool faults using hidden markov models. *Mechanical Systems and Signal Processing*, Elsevier, v. 25, n. 6, p. 2102–2124, aug 2011. ISSN 08883270. Citado 5 vezes nas páginas 2, 10, 12, 14 e 15.

BOX, G. E. P.; BEHNKEN, D. W. Some new three-level designs for the study of quantitative variables. *Technometrics*, v. 2, p. 455–475, 1960. Citado na página 47.

BOX, G. E. P.; DRAPER, N. R. *Empirical Model-Building and Response Surfaces*, John Wiley & Sons. 1. ed. [S.l.]: John Wiley & Sons, 1987. 650 p. Citado 3 vezes nas páginas 44, 45 e 47.

BOX, G. E. P.; HUNTER, W. G.; HUNTER, J. S. *Statistics for Experimenters: Design, Innovation, and Discovery*. 2. ed. [S.l.]: John Wiley & Sons, 2005. 664 p. ISBN 978-0-471-71813-0. Citado na página 45.

BRIGHAM, E. O.; MORROW, R. E. The fast Fourier transform. *Spectrum, IEEE*, v. 4, n. 12, p. 63–70, 1967. ISSN 0018-9235. Citado na página 20.

BROWN, J. C.; SMARAGDIS, P. Hidden markov and gaussian mixture models for automatic call classification. *The Journal of the Acoustical Society of America*, v. 125, n. 6, p. EL221–4, jun 2009. ISSN 1520-8524. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19507925>>. Citado 2 vezes nas páginas 15 e 28.

BRYMAN, A. *Research Methods and Organization Studies*. [S.l.: s.n.], 1989. 304 p. ISBN 0415084040. Citado na página 5.

CAYO, E. H.; ALFARO, S. C. A. A Non-Intrusive GMA Welding Process Quality Monitoring System Using Acoustic Sensing. *Sensors (Basel, Switzerland)*, v. 9, n. 9, p. 7150–66, jan 2009. ISSN 1424-8220. Citado 2 vezes nas páginas 10 e 12.

CHANG, C.-C.; LIN, C.-J. LIBSVM. *ACM Transactions on Intelligent Systems and Technology*, v. 2, n. 3, p. 1–27, apr 2011. ISSN 21576904. Disponível em: <<http://dl.acm.org/citation.cfm?doid=1961189.1961199> <http://www.csie.ntu.edu.tw/~cjlin/libsvm>>. Citado na página 81.

CHAUHAN, S.; WANG, P.; Sing Lim, C.; ANANTHARAMAN, V. A computer-aided MFCC-based HMM system for automatic auscultation. *Computers in Biology and Medicine*, v. 38, n. 2, p. 221–233, 2008. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18045582>>. Citado na página 21.

- CHEN, B.; CHEN, X.; LI, B.; HE, Z.; CAO, H.; CAI, G. Reliability estimation for cutting tools based on logistic regression model using vibration signals. *Mechanical Systems and Signal Processing*, Elsevier, v. 25, n. 7, p. 2526–2537, oct 2011. ISSN 08883270. Citado 3 vezes nas páginas 1, 10 e 12.
- CRESWELL, J.; CLARK, V. *Designing and conducting mixed-methods research*. [S.l.: s.n.], 2011. 488 p. Citado na página 5.
- DANNENBERG, R. *Audacity*. 2013. Disponível em: <<http://audacity.sourceforge.net/>>. Citado na página 54.
- DAVIS, S.; MERMELSTEIN, P. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics Speech and Signal Processing*, IEEE, v. 28, n. 4, p. 357–366, 1980. ISSN 00963518. Citado na página 21.
- DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. Maximum likelihood from incompatible data via EM algorithm. *Journal of the Royal Statistical Society*, v. 39, n. 1, p. 1–38, 1977. Citado na página 30.
- DHANALAKSHMI, P.; PALANIVEL, S.; RAMALINGAM, V. Classification of audio signals using SVM and RBFNN. *Expert Systems with Applications*, Elsevier Ltd, v. 36, n. 3, p. 6069–6075, apr 2009. ISSN 09574174. Citado 5 vezes nas páginas 15, 19, 21, 33 e 34.
- DHANALAKSHMI, P.; PALANIVEL, S.; RAMALINGAM, V. Classification of audio signals using AANN and GMM. *Applied Soft Computing*, v. 11, n. 1, p. 716–723, jan 2011. ISSN 15684946. Citado 3 vezes nas páginas 21, 23 e 29.
- DOWNEY, J.; O'LEARY, P.; RAGHAVENDRA, R. Comparison and analysis of audible sound energy emissions during single point machining of HSTS with PVD TiCN cutter insert across full tool life. *Wear*, Elsevier, v. 313, n. 1-2, p. 53–62, may 2014. ISSN 00431648. Citado 4 vezes nas páginas 2, 3, 10 e 12.
- FADARE, D. A.; SALES, W. F.; BONNEY, J.; EZUGWU, E. O. Influence of cutting parameters and tool wear on acoustic emission signal in high-speed turning of Ti-6Al-4V alloy. *Journal of Emerging Trends in Engineering and Applied Sciences*, v. 3, n. 3, p. 547–555, 2012. Citado 6 vezes nas páginas 1, 3, 10, 12, 13 e 57.
- FAHMY, M. M. Palmprint recognition based on mel frequency cepstral coefficients feature extraction. *Ain Shams Engineering Journal*, Faculty of Engineering, Ain Shams University, v. 1, n. 1, p. 39–47, sep 2010. ISSN 20904479. Citado 3 vezes nas páginas 19, 21 e 23.
- FRIGIERI, E. P. *Monitoramento da Rugosidade no Processo de Torneamento Duro Utilizando Coeficientes Mel-Cepstrais de Sinais Acústicos e Modelos de Misturas de Gaussianas* Monitoramento da Rugosidade no Processo de Torneamento Duro Utilizando Coeficientes Mel-Cepstrais de. 122 p. Tese (Master thesis) — Universidade Federal de Itajuba (UNIFEI), 2013. Citado 6 vezes nas páginas 10, 11, 14, 15, 76 e 91.
- FRIGIERI, E. P.; BRITO, T. G.; YNOGUTI, C. A.; PAIVA, A. P.; FERREIRA, J. R.; BALESTRASSI, P. P. Pattern recognition in audible sound energy emissions of AISI 52100 hardened steel turning: a MFCC-based approach. *The International Journal*

of *Advanced Manufacturing Technology*, may 2016. ISSN 0268-3768. Disponível em: <<http://link.springer.com/10.1007/s00170-016-8748-4>>. Citado na página 72.

FRIGIERI, E. P.; CAMPOS, P. H. S.; PAIVA, A. P.; BALESTRASSI, P. P. A mel-frequency cepstral coefficient-based approach for surface roughness diagnosis in hard turning using acoustic signals and gaussian mixture models. *Applied Acoustics*, v. 113, p. 230–237, dec 2016. ISSN 0003682X. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0003682X16301840>>. Citado na página 89.

GAJATE, A.; HABER, R.; TORO, R.; VEGA, P.; BUSTILLO, A. Tool wear monitoring using neuro-fuzzy techniques: a comparative study in a turning process. *Journal of Intelligent Manufacturing*, v. 23, n. 3, p. 869–882, aug 2010. ISSN 0956-5515. Disponível em: <<http://link.springer.com/10.1007/s10845-010-0443-y>>. Citado 2 vezes nas páginas 10 e 12.

GOLUB, G. H.; Van Loan, C. F. *Matrix Computations*. 1996. 48 p. Citado na página 26.

GRABEC, I.; GOVEKAR, E.; SUSIČ, E.; ANTOLOVIČ, B. Monitoring manufacturing processes by utilizing empirical modeling. *Ultrasonics*, v. 36, n. 1-5, p. 263–271, feb 1998. ISSN 0041624X. Disponível em: <[http://dx.doi.org/10.1016/S0041-624X\(97\)00051-6](http://dx.doi.org/10.1016/S0041-624X(97)00051-6)>. Citado na página 13.

HAYKIN, S.; VEEN, B. V. *Signals and Systems*. 2nd. ed. [S.l.]: John Wiley & Sons, 2003. 816 p. ISBN 0471378518. Citado na página 60.

HECKE, B. V.; YOON, J.; HE, D. Low speed bearing fault diagnosis using acoustic emission sensors. *Applied Acoustics*, Elsevier Ltd, v. 105, p. 35–44, 2016. ISSN 1872910X. Disponível em: <<http://dx.doi.org/10.1016/j.apacoust.2015.10.028>>. Citado 2 vezes nas páginas 10 e 12.

HOTELLING, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, v. 24, n. 6, p. 417–441, 1933. ISSN 0022-0663. Citado na página 24.

HU, Q.; HE, Z.; ZHANG, Z.; ZI, Y. Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble. *Mechanical Systems and Signal Processing*, v. 21, n. 2, p. 688–705, feb 2007. ISSN 08883270. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0888327006000306>>. Citado 4 vezes nas páginas 10, 12, 13 e 15.

JEMIELNIAK, K.; ARRAZOLA, P. Application of AE and cutting force signals in tool condition monitoring in micro-milling. *CIRP Journal of Manufacturing Science and Technology*, v. 1, n. 2, p. 97–102, jan 2008. ISSN 17555817. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1755581708000400>>. Citado 2 vezes nas páginas 10 e 12.

JOHNSON, R. A.; WICHERN, D. W. *Applied Multivariate Statistical Analysis*. 5. ed. New Jersey: Prentice-Hall Inc, 2007. 1–767 p. ISBN 9780131877153. Citado na página 73.

- JOLLIFFE, I. T. Principal Component Analysis, Second Edition. *Encyclopedia of Statistics in Behavioral Science*, v. 30, n. 3, p. 487, 2002. ISSN 00401706. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1002/0470013192.bsa501/full>>. Citado na página 24.
- JUNQIN, W.; JUNJUN, Y. An improved arithmetic of MFCC in speech recognition system. In: FACULTY OF INFORMATION ENGINEERING, JIANGXI UNIVERSITY OF SCIENCE AND TECHNOLOGY, GANZHOU, CHINA. *2011 International Conference on Electronics Communications and Control ICECC*. [S.l.]: IEEE, 2011. p. 719–722. ISBN 9781457703195. Citado na página 14.
- KAMBLE, M. Detecting Image Spam Using Principal Component Analysis & SVM Classifier. *International Journal of Computer Science and Information Technology & Security*, v. 2, n. 6, p. 1217–1220, 2012. Citado na página 15.
- KAMRUZZAMAN, S. M.; KARIM, A. N. M. R.; ISLAM, M. S.; HAQUE, M. E. Speaker identification using MFCC-domain support vector machine. *Computer*, I, n. 10, p. 5, 2010. Disponível em: <<http://arxiv.org/abs/1009.4972>>. Citado na página 21.
- KASBAN, H.; ZAHRAN, O.; ARAFA, H.; EL-KORDY, M.; ELARABY, S.; Abd El-Samie, F. Welding defect detection from radiography images with a cepstral approach. *NDT & E International*, Elsevier, v. 44, n. 2, p. 226–231, mar 2011. ISSN 09638695. Citado 7 vezes nas páginas 10, 11, 12, 13, 15, 21 e 23.
- KHORASANI, A.; YAZDI, M. R. S. Development of a dynamic surface roughness monitoring system based on artificial neural networks (ANN) in milling operation. *The International Journal of Advanced Manufacturing Technology*, oct 2015. ISSN 0268-3768. Citado 4 vezes nas páginas 10, 12, 13 e 15.
- KOOLAGUDI, S. G.; RASTOGI, D.; RAO, K. S. Identification of Language using Mel-Frequency Cepstral Coefficients (MFCC). *Procedia Engineering*, v. 38, p. 3391–3398, jan 2012. ISSN 18777058. Citado 3 vezes nas páginas 17, 19 e 21.
- KUMARI, R. S.; NIDHYANANTHAN, S. S.; G, A. Fused mel feature sets based text-independent speaker identification using gaussian mixture model. *Procedia Engineering*, v. 30, p. 319–326, jan 2012. ISSN 18777058. Citado 3 vezes nas páginas 21, 23 e 29.
- LI, X. A brief review: acoustic emission method for tool wear monitoring during turning. *International Journal of Machine Tools and Manufacture*, v. 42, n. 2, p. 157–165, jan 2002. ISSN 08906955. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0890695501001080>>. Citado na página 13.
- LIAO, Z.; GAO, D.; LU, Y.; LV, Z. Multi-scale hybrid HMM for tool wear condition monitoring. *The International Journal of Advanced Manufacturing Technology*, v. 84, n. 9-12, p. 2437–2448, jun 2016. ISSN 0268-3768. Disponível em: <<http://link.springer.com/10.1007/s00170-015-7895-3>>. Citado 3 vezes nas páginas 10, 11 e 12.
- LIU, T.-I.; JOLLEY, B. Tool condition monitoring (TCM) using neural networks. *The International Journal of Advanced Manufacturing Technology*, v. 78, n. 9-12, p. 1999–2007, 2015. ISSN 0268-3768. Disponível em: <<http://link.springer.com/10.1007/s00170-014-6738-y>>. Citado 4 vezes nas páginas 10, 12, 13 e 15.

LLOYD, S. P. Least Squares Quantization in PCM. *IEEE Transactions on Information Theory*, v. 28, n. 2, p. 129–137, 1982. ISSN 15579654. Citado na página 31.

LU, M.-C.; WAN, B.-S. Study of high-frequency sound signals for tool wear monitoring in micromilling. *The International Journal of Advanced Manufacturing Technology*, p. 1785–1792, aug 2012. ISSN 0268-3768. Citado na página 2.

LU, X.; WANG, P.; NIYATO, D.; KIM, D. I.; HAN, Z. Wireless Networks With RF Energy Harvesting: A Contemporary Survey. *IEEE Communications Surveys & Tutorials*, v. 17, n. 2, p. 757–789, jan 2015. ISSN 1553-877X. Citado 3 vezes nas páginas 10, 12 e 14.

LU, Y.; WU, Z. Maximum Likelihood Normalization for Robust Speech Recognition. *Applied Acoustics*, v. 74, n. 5, p. 640–646, 2013. ISSN 0003-682X. Disponível em: <<http://www.ee.ust.hk/~eemsiu/Paper/eurospeech2003-lai.p>>. Citado na página 28.

MANNAN, M.; KASSIM, A. a.; JING, M. Application of image and sound analysis techniques to monitor the condition of cutting tools. *Pattern Recognition Letters*, v. 21, n. 11, p. 969–979, oct 2000. ISSN 01678655. Citado 2 vezes nas páginas 1 e 2.

MARINESCU, I.; AXINTE, D. A time-frequency acoustic emission-based monitoring technique to identify workpiece surface malfunctions in milling with multiple teeth cutting simultaneously. *International Journal of Machine Tools and Manufacture*, v. 49, n. 1, p. 53–65, 2009. ISSN 08906955. Citado 4 vezes nas páginas 1, 10, 12 e 13.

MARWALA, T.; MAHOLA, U.; NELWAMONDO, F. Hidden markov models and gaussian mixture models for bearing fault detection using fractals. In: *The 2006 IEEE International Joint Conference on Neural Network Proceedings*. [S.l.]: IEEE, 2006. p. 3237–3242. ISBN 0-7803-9490-9. Citado 4 vezes nas páginas 10, 12, 15 e 24.

MAYORGA, P.; BESACIER, L.; LAMY, R. Audio packet loss over IP and speech recognition. *Automatic Speech Recognition*, 2003. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1318>. Citado na página 21.

MCLACHLAN, G. J.; BASFORD, K. E. Mixture Models: Inference and Applications to Clustering. *Statistics Textbooks and Monographs*, v. 1, p. 272, 1988. ISSN 0035-9254. Citado na página 28.

MONTGOMERY, D. C. *Design and Analysis of Experiments*. 8. ed. Hoboken: John Wiley and Sons, 2013. 752 p. ISBN 9781118146927. Citado 4 vezes nas páginas 45, 46, 48 e 51.

MONTGOMERY, D. C.; JOHNSON, R. T. Choice of second-order response surface designs for logistic and Poisson regression models. *International Journal of Experimental Design and Process Optimisation*, v. 1, n. 1, p. 2, 2009. ISSN 2040-2252. Disponível em: <<http://www.inderscience.com/link.php?id=28954>>. Citado na página 55.

MONTGOMERY, D. C.; RUNGER, G. C. *Applied Statistics and Probability for Engineers*. [S.l.: s.n.], 2011. 792 p. ISBN 9780470053041. Citado 4 vezes nas páginas 47, 48, 50 e 51.

MYERS, R. H.; MONTGOMERY, D. C. *Response Surface Methodology: Process and Product Optimization Using Design of Experiments*. 2. ed. New York: Wiley – Interscience, 1995. 700 p. ISBN 978-0470174463. Citado na página 44.

- OLIVEIRA, A. G. D.; VENTURA, T. M.; GANCHEV, T. D.; De Figueiredo, J. M.; JAHN, O.; MARQUES, M. I.; SCHUCHMANN, K. L. Bird acoustic activity detection based on morphological filtering of the spectrogram. *Applied Acoustics*, Elsevier Ltd, v. 98, p. 34–42, 2015. ISSN 1872910X. Disponível em: <<http://dx.doi.org/10.1016/j.apacoust.2015.04.014>>. Citado na página 28.
- PAIVA, A.; CAMPOS, P.; FERREIRA, J.; LOPES, L.; PAIVA, E.; BALESTRASSI, P. A multivariate robust parameter design approach for optimization of AISI 52100 hardened steel turning with wiper mixed ceramic tool. *International Journal of Refractory Metals and Hard Materials*, v. 30, n. 1, p. 152–163, jan 2012. ISSN 02634368. Citado 6 vezes nas páginas 1, 24, 53, 54, 55 e 56.
- PEARSON, K. LIII. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine Series 6*, v. 2, n. 11, p. 559–572, nov 1901. ISSN 1941-5982. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/14786440109462720>>. Citado na página 24.
- PICONE, J. Signal modeling techniques in speech recognition. *Proceedings of the IEEE*, v. 81, n. 9, p. 1215–1247, 1993. ISSN 00189219. Citado 4 vezes nas páginas 18, 19, 21 e 58.
- POLUR, P. D.; MILLER, G. E. Investigation of an HMM/ANN hybrid structure in pattern recognition application using cepstral analysis of dysarthric (distorted) speech signals. *Medical engineering & physics*, v. 28, n. 8, p. 741–8, oct 2006. ISSN 1350-4533. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16359906>>. Citado na página 21.
- PONTES, F. J. *Projeto Otimizado de Redes Neurais Artificiais para Predição da Rugosidade em Processos de Usinagem com a Utilização da Metodologia de Projeto de Experimentos*. 259 p. Tese (Doutorado) — Universidade Paulista, 2011. Citado na página 10.
- POWELL, M. J. D. Radial Basis Functions for Multivariable Interpolation: A Review. In: MASON, J. C.; COX, M. G. (Ed.). *Algorithms for Approximation*. New York, NY: Clarendon Press, 1987. p. 143–167. ISBN 0-19-853612-7. Disponível em: <<http://dl.acm.org/citation.cfm?id=48424.48433>>. Citado na página 42.
- RABINER, L.; JUANG, B.-H. *Fundamentals of Speech Recognition*. [S.l.: s.n.], 1993. 224–225 p. ISBN 0130151572. Citado 2 vezes nas páginas 18 e 19.
- RABINER, L. R.; SCHAFER, R. W. Introduction to digital speech processing. *Foundations and Trends in Signal Processing*, v. 1, n. 1, p. 1–194, 2007. ISSN 1932-8346. Disponível em: <<http://dx.doi.org/10.1561/2000000001>>. Citado na página 19.
- RAJA, J. E.; LIM, W. S.; VENKATASESHIAH, C. Emitted sound analysis for tool flank wear monitoring using hilbert huang transform. *International Journal of Computer and Electrical Engineering*, v. 4, n. 2, 2012. ISSN 17938163. Citado 6 vezes nas páginas 1, 2, 3, 10, 12 e 13.
- REYNOLDS, D. a.; ROSE, R. C. Robust text-independent speaker identification using gaussian mixture speaker models. *IEEE Transactions On Speech And Audio Processing*, p. 72–83, 1995. Citado na página 28.

- ROBBEN, L.; RAHMAN, S.; BUHL, J. C.; DENKENA, B.; KONOPATZKI, B. Airborne sound emission as a process monitoring tool in the cut-off grinding of concrete. *Applied Acoustics*, Elsevier Ltd, v. 71, n. 1, p. 52–60, 2010. ISSN 0003682X. Citado na página 2.
- RUBIO, E. M.; TETI, R. Cutting parameters analysis for the development of a milling process monitoring system based on audible energy sound. *Journal of Intelligent Manufacturing*, v. 20, n. 1, p. 43–54, apr 2008. ISSN 0956-5515. Disponível em: <<http://link.springer.com/10.1007/s10845-008-0102-8>>. Citado 6 vezes nas páginas 2, 3, 10, 12, 13 e 15.
- RUBIO, E. M.; TETI, R. Process Monitoring Systems for Machining Using Audible Sound Energy Sensors. In: AIZED, T. (Ed.). *Future Manufacturing Systems*. [S.l.]: Sciyo, 2010. p. 217–235. ISBN 978-953-307-128-2. Citado 2 vezes nas páginas 1 e 2.
- SALGADO, D. R.; ALONSO, F. J. An approach based on current and sound signals for in-process tool wear monitoring. *International Journal of Machine Tools and Manufacture*, v. 47, n. 14, p. 2140–2152, 2007. ISSN 08906955. Citado 4 vezes nas páginas 2, 10, 12 e 15.
- SCHEFFER, C.; KRATZ, H.; HEYNS, P.; KLOCKE, F. Development of a tool wear-monitoring system for hard turning. *International Journal of Machine Tools and Manufacture*, v. 43, n. 10, p. 973–985, aug 2003. ISSN 08906955. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S089069550300110X>>. Citado 3 vezes nas páginas 1, 10 e 12.
- SHARMA, V. S.; SHARMA, S. K.; SHARMA, A. K. Cutting tool wear estimation for turning. *Journal of Intelligent Manufacturing*, v. 19, n. 1, p. 99–108, jul 2007. ISSN 0956-5515. Disponível em: <<http://www.springerlink.com/index/10.1007/s10845-007-0048-2>>. Citado 2 vezes nas páginas 10 e 12.
- SICK, B. On-line and indirect tool wear monitoring in turning with artificial neural networks: a review of more than a decade of research. *Mechanical Systems and Signal Processing*, v. 16, n. 4, p. 487–546, jul 2002. ISSN 08883270. Citado 5 vezes nas páginas 1, 3, 7, 8 e 9.
- SR., D. E. D.; LISTER, P. M. On-line metal cutting tool condition monitoring . I : force and vibration analyses. *International Journal of Machine Tools and Manufacture*, v. 40, p. 739–768, 2000. Citado na página 1.
- TEKINER, Z.; YESILYURT, S. Investigation of the cutting parameters depending on process sound during turning of AISI 304 austenitic stainless steel. *Materials & Design*, v. 25, n. 6, p. 507–513, sep 2004. ISSN 02613069. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0261306903002632>>. Citado 5 vezes nas páginas 10, 12, 56, 65 e 70.
- TETI, R.; JEMIELNIAK, K.; O'DONNELL, G.; DORNFELD, D. Advanced monitoring of machining operations. *CIRP Annals - Manufacturing Technology*, v. 59, n. 2, p. 717–739, jan 2010. ISSN 00078506. Citado na página 3.
- VAPNIK, V. An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, v. 10, n. 5, p. 988–999, 1999. ISSN 10459227. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=788640>>. Citado na página 33.

- VAPNIK, V. N. The Nature of Statistical Learning Theory. *Springer*, v. 8, p. 187, 1995. ISSN 10459227. Disponível em: <<http://portal.acm.org/citation.cfm?id=211359>>. Citado 2 vezes nas páginas 36 e 37.
- VIGNOLO, L. D.; RUFINER, H. L.; MILONE, D. H.; GODDARD, J. C. Evolutionary cepstral coefficients. *Applied Soft Computing*, Elsevier B.V., v. 11, n. 4, p. 3419–3428, jun 2011. ISSN 15684946. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1568494611000226>>. Citado na página 23.
- WANG, C.; MIAO, Z.; MENG, X. Differential MFCC and vector quantization used for real-time speaker recognition system. In: *2008 Congress on Image and Signal Processing*. Ieee, 2008. v. 5, p. 319–323. ISBN 9780769531199. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4566841>>. Citado na página 21.
- WANG, C.-C.; KANG, Y. Feature Extraction Techniques of Non-Stationary Signals for Fault Diagnosis in Machinery Systems. *Journal of Signal and Information Processing*, v. 03, n. 01, p. 16–25, 2012. ISSN 2159-4465. Citado 3 vezes nas páginas 10, 12 e 15.
- WANG, L.; MEHRABI, M. G.; KANNATEY-ASIBU, E. Hidden markov model-based tool wear monitoring in turning. *Journal of Manufacturing Science and Engineering*, v. 124, n. 3, p. 651, 2002. ISSN 10871357. Citado 3 vezes nas páginas 10, 12 e 13.
- WIDODO, A.; YANG, B.-S. Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical Systems and Signal Processing*, v. 21, n. 6, p. 2560–2574, aug 2007. ISSN 08883270. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0888327007000027>>. Citado 3 vezes nas páginas 34, 40 e 41.
- XI, J.; HAN, W.; LIU, Y. Relationship analysis between chaotic characteristics of acoustic emission signal and tool wear condition. In: *Third International Workshop on Advanced Computational Intelligence*. IEEE, 2010. p. 612–617. ISBN 978-1-4244-6334-3. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5585164>>. Citado na página 1.
- XIQING, M.; CHUANGWEN, X. Tool Wear Monitoring of Acoustic Emission Signals from Milling Processes. In: *2009 First International Workshop on Education Technology and Computer Science*. IEEE, 2009. p. 431–435. ISBN 978-0-7695-3557-9. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4958808>>. Citado 4 vezes nas páginas 1, 10, 11 e 12.
- XU, Q.; XIANG, E. W.; YANG, Q.; DU, J.; ZHONG, J. SMS Spam Detection Using Noncontent Features. *IEEE Intelligent Systems*, v. 27, n. 6, p. 44–51, nov 2012. ISSN 1541-1672. Citado na página 36.
- YUJIN, Y. Y. Y.; PEIHUA, Z. P. Z.; QUN, Z. Q. Z. Research of speaker recognition based on combination of LPCC and MFCC. In: *Intelligent Computing and Intelligent Systems ICIS 2010 IEEE International Conference on*. [S.l.: s.n.], 2010. v. 3, p. 765–767. ISBN 9781424465828. Citado na página 14.
- ZHU, K.; WONG, Y. S.; HONG, G. S. Multi-category micro-milling tool wear monitoring with continuous hidden markov models. *Mechanical Systems and Signal Processing*,

v. 23, n. 2, p. 547–560, feb 2009. ISSN 08883270. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0888327008001088>>. Citado 4 vezes nas páginas 1, 10, 12 e 14.

ZUNJING, W. U.; ZHIGANG, C. A. O. Improved MFCC-based feature for robust speaker identification. *Tsinghua Science and Technology*, v. 10, n. 2, p. 158–161, 2005. ISSN 10070214. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1007021405700481>>. Citado na página 21.

Apêndices

APÊNDICE A – TABELA COMPLETA DAS MEDIDAS DE RUGOSIDADE.

Tabela A.1 – Medidas de Rugosidade

Config.	Rugosidade	Passos														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
MS1	Ra	0.15	0.15	0.15	0.14	0.14	0.14	0.15	0.15	0.15	0.13	0.14	0.13	0.16	0.16	0.15
	Ry	1.06	1.03	1.02	1.05	1.04	1.05	1.09	1.11	1.09	0.99	0.97	0.98	1.11	1.09	1.08
	Rz	0.71	0.71	0.70	0.76	0.76	0.75	0.76	0.77	0.77	0.73	0.74	0.73	0.82	0.81	0.80
	Rq	0.19	0.19	0.19	0.18	0.18	0.18	0.18	0.19	0.18	0.17	0.17	0.17	0.20	0.19	0.19
	Rt	1.44	1.44	1.43	1.46	1.43	1.54	1.53	1.55	1.56	1.31	1.27	1.30	1.49	1.43	1.41
	Rsm	110.50	111.17	106.25	79.00	75.50	77.17	87.00	95.33	89.50	82.75	81.00	78.17	71.75	68.75	70.17
	Média	2503.50	2370.10	1687.40	1352.10	978.72	971.56	984.58	1063.40	1129.00	965.24	1210.40	1109.70	1063.40	972.77	971.30
	D. Padrão	337.27	324.34	328.25	368.55	54.04	32.54	29.28	140.25	223.96	45.01	241.67	155.61	123.89	12.14	32.48
MS2	Ra	0.16	0.16	0.16	0.17	0.17	0.17	0.14	0.14	0.14	0.18	0.18	0.17	0.16	0.16	0.16
	Ry	1.06	1.08	1.08	1.33	1.32	1.33	1.03	1.02	1.04	1.27	1.23	1.21	1.23	1.17	1.18
	Rz	0.71	0.72	0.72	0.85	0.87	0.86	0.75	0.74	0.74	0.89	0.88	0.88	0.89	0.86	0.86
	Rq	0.20	0.20	0.20	0.22	0.22	0.22	0.18	0.18	0.18	0.22	0.22	0.21	0.20	0.20	0.20
	Rt	1.41	1.51	1.51	1.88	1.81	1.88	1.34	1.29	1.34	1.61	1.59	1.53	1.65	1.52	1.49
	Rsm	129.50	116.33	129.00	93.42	92.50	87.67	81.33	78.58	76.00	85.92	84.08	88.50	78.67	74.42	73.58
	Média	1101.80	1549.40	1339.60	2003.10	1418.80	1845.90	1505.20	2101.60	1687.40	2001.90	1824.90	1411.10	1865.90	1880.50	1605.70
	D. Padrão	10.24	342.67	275.94	341.29	317.38	337.68	339.12	360.94	352.41	244.05	318.07	334.69	269.72	303.10	331.35
MS3	Ra	0.40	0.40	0.40	0.44	0.43	0.44	0.47	0.47	0.47	0.45	0.45	0.45	0.45	0.45	0.44
	Ry	2.41	2.41	2.40	2.23	2.19	2.18	2.47	2.48	2.47	2.49	2.49	2.47	2.42	2.41	2.36
	Rz	1.61	1.60	1.59	1.66	1.64	1.65	1.89	1.89	1.92	1.89	1.87	1.88	1.79	1.79	1.78
	Rq	0.49	0.49	0.49	0.51	0.51	0.51	0.56	0.56	0.56	0.54	0.54	0.54	0.53	0.53	0.53
	Rt	3.55	3.54	3.55	2.77	2.73	2.74	3.06	3.03	2.99	3.25	3.26	3.21	2.95	2.91	2.80
	Rsm	190.00	189.25	187.67	196.08	194.75	193.00	194.42	190.33	190.00	164.83	161.00	166.00	111.83	107.75	108.92
	Média	4366.40	2146.30	1442.10	1313.20	2023.50	1987.70	1437.90	3586.30	2407.50	1453.00	1319.00	1004.80	1373.40	2882.70	2209.00
	D. Padrão	1788.20	445.48	381.22	379.97	595.79	821.98	659.18	1538.20	707.53	409.12	271.92	94.56	363.83	45.03	326.93
MS4	Ra	0.20	0.20	0.20	0.19	0.19	0.19	0.18	0.19	0.18	0.19	0.19	0.19	0.19	0.19	0.18
	Ry	1.51	1.49	1.47	1.30	1.31	1.29	1.23	1.32	1.30	1.28	1.27	1.28	1.47	1.45	1.44
	Rz	1.01	1.00	1.00	0.98	0.99	0.99	0.92	0.95	0.93	0.94	0.94	0.93	1.15	1.15	1.15
	Rq	0.25	0.25	0.25	0.23	0.23	0.23	0.22	0.23	0.23	0.23	0.23	0.23	0.24	0.23	0.23
	Rt	2.21	2.18	2.15	1.74	1.70	1.66	1.56	1.81	1.77	1.61	1.63	1.66	1.81	1.79	1.73
	Rsm	68.08	65.92	61.58	49.92	50.08	48.67	64.00	69.50	66.58	65.42	68.50	69.25	46.25	45.17	44.67
	Média	3023.60	2947.70	2557.10	1054.30	2310.00	1770.30	1005.70	1020.90	2322.50	2807.60	2552.30	2131.00	2545.20	2496.70	2884.60
	D. Padrão	490.33	331.95	285.12	126.02	462.01	584.37	66.27	79.91	426.18	238.65	396.92	304.97	382.17	416.29	277.55
MS5	Ra	0.18	0.18	0.18	0.17	0.17	0.17	0.15	0.15	0.15	0.18	0.18	0.18	0.22	0.23	0.23
	Ry	1.21	1.17	1.19	1.34	1.32	1.30	1.04	1.11	1.05	1.36	1.32	1.32	1.64	1.70	1.67
	Rz	0.78	0.76	0.77	0.85	0.84	0.84	0.75	0.76	0.75	0.95	0.94	0.93	1.12	1.14	1.12
	Rq	0.22	0.22	0.22	0.22	0.22	0.22	0.18	0.19	0.18	0.23	0.22	0.22	0.29	0.29	0.29
	Rt	1.68	1.61	1.63	2.04	1.95	1.98	1.32	1.63	1.37	2.06	1.99	2.01	2.25	2.25	2.36
	Rsm	152.83	163.83	162.58	107.08	99.75	105.33	82.75	82.00	83.92	74.33	73.00	73.67	81.67	83.17	84.08
	Média	2958.10	2678.20	1008.10	1517.20	1768.80	1969.80	1759.50	1808.50	1957.10	2538.10	323.13	1042.30	3107.20	3445.20	3786.90
	D. Padrão	675.09	498.51	561.57	530.83	464.39	422.46	441.33	470.53	402.09	254.29	309.92	563.87	540.73	226.39	861.19
MS6	Ra	0.53	0.53	0.52	0.53	0.52	0.53	0.52	0.52	0.51	0.52	0.52	0.52	0.52	0.52	0.52
	Ry	2.97	2.98	2.98	2.84	2.83	2.84	2.95	2.94	2.89	2.81	2.83	2.84	2.92	2.94	2.94
	Rz	2.28	2.27	2.26	2.24	2.24	2.25	2.24	2.24	2.19	2.22	2.22	2.23	2.24	2.24	2.23
	Rq	0.64	0.64	0.64	0.63	0.63	0.63	0.62	0.62	0.62	0.63	0.63	0.63	0.63	0.63	0.62
	Rt	3.46	3.48	3.50	3.24	3.21	3.23	3.52	3.54	3.42	3.25	3.32	3.33	3.49	3.59	3.62
	Rsm	197.83	194.92	195.92	195.50	196.42	195.92	194.50	193.25	194.58	198.83	197.92	198.92	193.33	191.83	192.25
	Média	4550.80	9131.70	9961.00	9988.50	9626.90	8776.70	6588.60	7424.90	5288.50	1028.40	8214.40	5076.70	5561.80	7136.00	3609.60
	D. Padrão	1766.00	956.95	550.39	511.51	658.48	961.65	1777.30	1518.60	1885.30	423.95	971.58	1760.40	1775.50	1253.70	1433.40
MS7	Ra	0.29	0.33	0.28	0.31	0.25	0.28	0.27	0.25	0.26	0.24	0.25	0.26	0.22	0.23	0.25
	Ry	2.03	2.38	2.24	2.15	1.68	1.91	2.05	1.59	1.90	1.77	1.84	1.90	1.51	1.43	1.71
	Rz	1.33	1.53	1.39	1.41	1.18	1.28	1.34	1.07	1.19	1.27	1.18	1.22	1.05	1.09	1.26
	Rq	0.37	0.43	0.37	0.40	0.31	0.34	0.34	0.31	0.33	0.32	0.32	0.33	0.27	0.29	0.32
	Rt	2.68	3.64	3.11	3.08	2.31	2.36	2.79	2.13	2.65	3.49	2.83	2.88	2.17	1.99	2.55
	Rsm	160.83	141.33	135.33	105.42	92.67	80.92	75.25	139.92	123.42	119.58	92.17	85.92	76.08	72.33	63.50
	Média	10189.00	4582.90	2669.00	2251.30	3098.80	8524.20	4040.70	2711.90	6480.70	6895.30	6097.90	2354.50	4427.40	3508.40	3628.30
	D. Padrão	1617.20	1830.30	363.50	629.65	2244.00	2045.40	1617.40	580.46	2004.60	2400.00	2380.90	547.05	1687.00	1382.80	1571.30
MS8	Ra	0.51	0.51	0.51	0.50	0.50	0.50	0.51	0.51	0.51	0.50	0.50	0.49	0.49	0.49	0.49
	Ry	2.76	2.73	2.70	2.76	2.75	2.74	2.56	2.55	2.55	2.61	2.59	2.58	2.52	2.49	2.49
	Rz	2.15	2.12	2.13	2.22	2.22	2.22	2.02	2.01	2.02	2.05	2.01	2.02	1.92	1.90	1.89
	Rq	0.62	0.62	0.61	0.61	0.61	0.61	0.60	0.60	0.60	0.59	0.59	0.58	0.58	0.58	0.57
	Rt	3.39	3.35	3.25	3.24	3.26	3.27	2.90	2.92	2.92	3.04	3.00	3.00	3.01	2.95	2.94
	Rsm	219.25	220.75	220.17	224.33	219.00	220.67	211.92	209.92	211.92	208.08	207.58	212.33	196.58	204.00	198.67
	Média	13486.00	1228.00	1047.00	4683.30	9695.90	9977.80	13951.00	12900.00	13648.00	11824.00	11666.00	8593.70	13374.00	2763.80	6780.30
	D. Padrão	1266.40	271.49	21917.00	1842.60	2527.20	2860.60	425.65	1599.50	941.52	2056.00	1935.20	3483.20	720.46	1883.40	2967.50
MS9	Ra	0.25	0.25	0.25	0.24	0.25	0.25	0.27	0.27	0.27	0.25	0.25	0.25	0.28	0.28	0.28
	Ry	1.96	1.86	1.88	1.82	1.88	1.93	1.84	1.77	1.80	1.78	1.82	1.83	2.14	2.12	2.12
	Rz	1.29	1.25	1.25	1.21	1.23	1.25	1.27	1.24	1.24	1.29	1.30	1.30	1.42	1.41	1.41
	Rq	0.32	0.32	0.32	0.31	0.32	0.32	0.33	0.33	0.33	0.32	0.32	0.32	0.36	0.36	0.36
	Rt	2.85	2.58	2.65	2.69	2.92	2.91	2.53	2.40	2.50	2.47	2.67	2.63	2.90	2.95	2.90
	Rsm	67.50	67.83	69.67	73.75	76.50	77.08	79.58	78.08	78.83	71.00	70.92	70.67	71.33	72.08	70.83
	Média	13070.00	9203.00	9486.30	3277.40	4353.60	3463.60	6623.00	5606.50	2541.10	4439.40	4800.50	5978.90	3747.00	2913.80	3583.20
	D. Padrão	948.38	2037.60	2046.60	660.68	1581.60	2043.50	2233.70	2063.40	975.11	1781.30	1844.20	2125.30	790.02	2713.90	796.22
MS10	Ra	0.19	0.20	0.19	0.18	0.18	0.18	0.19	0.19	0.19	0.18	0.18	0.17	0.22	0.21	0.21
	Ry	1.46	1.43	1.41	1.29	1.2										

APÊNDICE B – ESCORES OBTIDOS ATRAVÉS DA ANÁLISE DE COMPONENTES PRINCIPAIS

Tabela B.1 – Cinco primeiros escores obtidos pela ACP (passos de 1 ao 4).

Replica	Configuração	PC_1	PC_2	PC_3	PC_4	PC_5
1	MS1	-9.40709	1.144496	1.17623	0.215581	2.047965
1	MS2	-8.68914	-2.34808	-0.51338	0.389215	2.550003
1	MS3	3.857607	1.752847	4.657514	1.330931	-0.59141
1	MS4	-2.69929	1.975467	2.439561	-1.21317	-0.29424
1	MS5	-0.62022	-0.5643	-2.67867	1.202403	-0.62213
1	MS6	23.57674	7.206892	0.652583	1.009246	4.743238
1	MS7	2.12315	-4.42434	2.978939	0.606172	-0.86813
1	MS8	19.54606	-7.35613	0.172635	1.357134	-1.97615
1	MS9	20.9901	-7.15394	0.284492	-2.03205	-0.10797
1	MS10	6.055453	-3.53236	1.110368	-2.55468	0.012877
2	MS1	-8.95681	1.583808	1.054509	0.41955	1.481251
2	MS2	-9.68146	-1.06229	-1.78604	1.393979	1.784174
2	MS3	-0.29575	1.901593	3.955877	2.346172	-0.417
2	MS4	-2.73785	2.258503	2.834283	-0.42221	-1.44323
2	MS5	-1.90983	-0.45271	-2.71347	0.719838	-1.20025
2	MS6	23.88522	8.290601	0.254908	1.562296	2.489401
2	MS7	1.544493	-3.90004	3.105584	1.179189	-0.85812
2	MS8	1.120174	-5.32605	0.451424	0.722586	2.596413
2	MS9	15.75463	-5.41297	-0.07057	-2.13654	0.285304
2	MS10	2.4119	-1.13039	-1.03216	1.96251	1.204079
3	MS1	-9.54313	1.961445	1.354035	1.740091	0.487334
3	MS2	-8.45993	-1.55555	-2.06795	1.687031	2.928687
3	MS3	0.163719	2.187978	4.99248	2.486847	-1.46789
3	MS4	-4.75594	2.391703	3.585083	0.44283	-0.68455
3	MS5	-2.98989	-0.30968	-3.67639	0.393915	-1.27949
3	MS6	20.2652	8.34567	0.163586	2.861294	3.078789
3	MS7	1.219945	-3.66302	3.691983	1.467446	-1.29918
3	MS8	-11.6927	-2.86304	-1.28161	1.136323	3.479627
3	MS9	12.98336	-5.29655	-0.75846	-2.2313	0.758015
3	MS10	2.031348	-2.3176	-0.09121	-1.54584	0.307962
4	MS1	-9.92036	1.9099	1.92332	2.111679	0.522161
4	MS2	-8.21029	-1.20519	-2.70008	2.16274	2.289913
4	MS3	-2.57436	2.700363	4.886421	2.610634	-1.38958
4	MS4	-6.23344	2.638379	3.786447	-0.67552	-1.62288
4	MS5	-3.95563	0.044655	-3.6978	0.027728	-2.55443
4	MS6	17.64478	8.075194	-0.66605	2.880332	2.661084
4	MS7	0.943779	-3.08842	4.82792	1.379603	-2.70948
4	MS8	10.80632	-6.85728	1.895559	0.988617	1.654682
4	MS9	9.675796	-3.87024	-0.58058	-2.31809	1.607245
4	MS10	0.749096	-1.71139	-1.2637	-2.08399	1.543085

Tabela B.2 – Cinco primeiros escores obtidos pela ACP (passos de 5 ao 8).

Replica	Configuração	PC_1	PC_2	PC_3	PC_4	PC_5
5	MS1	-10.4382	1.918289	1.7852	1.805376	0.806149
5	MS2	-8.63264	-0.83583	-2.3877	1.753472	2.498644
5	MS3	-2.85233	2.545021	4.275021	1.826566	-0.9424
5	MS4	-6.52986	2.375017	1.697055	-1.05675	-0.06059
5	MS5	-5.12204	0.251481	-2.48223	0.341644	-1.77897
5	MS6	16.36432	7.952551	-0.13523	0.823499	1.212688
5	MS7	1.015059	-3.10953	4.055351	0.458854	-1.71553
5	MS8	13.90062	-7.0692	1.989225	1.756101	1.084454
5	MS9	10.56402	-4.36684	-0.78131	-2.55477	1.254504
5	MS10	-0.86453	-1.8612	-1.0949	-1.2467	1.028619
6	MS1	-10.6561	2.483806	2.249363	2.356297	0.04291
6	MS2	-9.57686	-0.57083	-1.7232	1.277734	2.015165
6	MS3	-3.23558	2.090061	3.723381	1.276852	-0.79818
6	MS4	-7.39839	2.845961	1.884617	-1.31028	-0.59197
6	MS5	-5.36495	1.198013	-2.96405	0.688343	-3.30575
6	MS6	15.11866	8.229985	-0.33762	0.350885	0.650297
6	MS7	0.909571	-3.19686	3.032518	-0.44969	-1.63209
6	MS8	7.994081	-6.37122	1.440511	1.53797	0.635477
6	MS9	9.572064	-3.66245	-1.07881	-3.14769	2.61456
6	MS10	-1.04281	-1.71484	-1.75594	-2.25625	1.732166
7	MS1	-11.42	2.447033	2.885431	2.054358	-0.51038
7	MS2	-10.1207	-0.40913	-2.14939	0.812883	2.159708
7	MS3	-2.92334	2.3744	3.652633	0.912539	-0.64059
7	MS4	-7.45919	2.422988	1.260318	-2.00997	-0.73479
7	MS5	-6.1109	0.974847	-2.15956	0.80688	-2.62594
7	MS6	14.35848	8.17219	-0.96647	0.21322	0.570197
7	MS7	-0.69539	-3.08761	2.68906	-0.24765	-1.69876
7	MS8	12.22548	-5.31998	0.705498	2.534143	-0.11523
7	MS9	9.360786	-3.6312	-0.56844	-1.28809	1.377856
7	MS10	-1.05894	-1.59878	-1.61405	-1.66804	1.094998
8	MS1	-11.3762	2.561691	1.176391	1.333719	-0.22194
8	MS2	-10.2961	-0.38342	-2.92769	1.530574	1.631283
8	MS3	-0.9825	2.297248	3.838673	0.266871	-0.61904
8	MS4	-4.6139	1.9298	1.454193	-2.2138	0.593266
8	MS5	-6.5067	0.954056	-2.26285	0.696556	-2.56666
8	MS6	12.41594	8.064295	-1.43296	0.330208	0.002157
8	MS7	4.969151	-5.08377	1.787103	0.302566	-1.12856
8	MS8	10.47594	-5.43094	0.820561	1.444475	-0.62922
8	MS9	8.699377	-3.47243	-0.40745	-0.77744	0.52862
8	MS10	-1.56798	-1.97909	-0.87022	-1.63979	1.798647

Tabela B.3 – Cinco primeiros escores obtidos pela ACP (passos de 9 ao 12).

Replica	Configuração	PC_1	PC_2	PC_3	PC_4	PC_5
9	MS1	-11.8623	2.408616	0.985751	0.666054	0.31313
9	MS2	-10.5119	-0.23797	-2.80185	1.474818	1.05784
9	MS3	-3.85214	2.028756	3.611173	-0.10997	-1.09266
9	MS4	-4.05558	2.598638	1.652331	-4.48057	0.521306
9	MS5	-6.94489	0.673627	-1.20528	-0.50803	-2.05657
9	MS6	19.58213	8.527651	-1.27629	-0.6423	-0.1259
9	MS7	1.334868	-4.45399	2.121906	-0.42554	-0.28854
9	MS8	12.42244	-5.75058	0.365571	1.922424	-0.40092
9	MS9	7.878886	-3.67739	-0.8452	-0.32559	1.740923
9	MS10	-2.71844	-1.55323	-0.59337	-1.39816	2.044489
10	MS1	-12.1693	2.436856	1.624777	1.145978	-0.13783
10	MS2	-11.029	-0.50505	-2.47847	1.008308	0.74948
10	MS3	-4.48279	1.413897	2.818432	0.16327	0.18577
10	MS4	-4.80564	2.938026	1.920646	-6.56651	0.576625
10	MS5	-6.68411	1.014449	-1.54624	-0.21859	-3.0705
10	MS6	-12.135	0.943146	-3.4224	1.030473	0.581268
10	MS7	1.556632	-4.12051	1.206192	-1.32525	-1.07254
10	MS8	10.12393	-4.44283	-0.78219	1.356416	-1.73818
10	MS9	7.152814	-3.08908	-1.16354	0.603217	0.456102
10	MS10	-3.70962	-1.06013	-1.05403	-1.12247	1.971765
11	MS1	-12.3813	2.453037	0.624312	0.67175	0.079263
11	MS2	-10.9071	-0.20064	-3.08496	1.306426	0.810085
11	MS3	-4.99753	1.486572	2.537329	0.090441	0.436152
11	MS4	-5.40866	2.694649	1.660894	-5.26731	1.193923
11	MS5	-5.46837	-0.53098	-3.40034	0.911689	-0.77049
11	MS6	14.76721	8.111444	-1.79019	0.009305	-0.69211
11	MS7	0.558202	-3.29041	1.343452	-2.90937	-1.28629
11	MS8	8.508437	-3.98356	-0.41091	0.901103	-1.8547
11	MS9	6.110096	-2.3308	-1.98993	1.050454	0.355344
11	MS10	-3.80117	-0.72183	-1.25712	-0.07465	1.358171
12	MS1	-12.4115	2.631246	0.832214	-0.07654	0.078492
12	MS2	-12.2156	-0.22961	-2.8679	1.044576	0.532608
12	MS3	-5.38877	1.773939	2.46984	-0.19173	-0.13517
12	MS4	-5.92327	2.668019	0.656634	-4.81917	0.520465
12	MS5	-7.24943	0.588011	-2.90186	0.481792	-1.08217
12	MS6	19.92615	9.429266	-1.55908	-0.26047	-0.71321
12	MS7	1.244254	-3.93605	1.78235	-2.23587	-0.27579
12	MS8	9.40289	-4.46223	-0.7163	1.160006	-1.07247
12	MS9	7.412198	-2.59614	-2.14187	0.054053	-0.76566
12	MS10	-4.46912	-2.11757	-0.99899	-0.1474	2.576439

Tabela B.4 – Cinco primeiros escores obtidos pela ACP (passos de 13 ao 15).

Replica	Configuração	PC_1	PC_2	PC_3	PC_4	PC_5
13	MS1	-12.6264	2.098927	0.709121	-0.66586	0.107974
13	MS2	-11.54	-0.09655	-3.34168	1.299291	0.275902
13	MS3	-5.46469	1.919132	2.481312	-0.74318	-0.01466
13	MS4	-6.29544	2.713039	-0.08053	-3.71305	0.556788
13	MS5	-6.92527	0.835123	-3.42063	-0.10613	-2.81185
13	MS6	18.34234	9.508976	-1.8747	-1.01638	-1.80889
13	MS7	-0.05113	-3.60876	1.436426	-2.02495	-0.58232
13	MS8	8.493135	-3.99554	-0.42421	0.669703	-1.70945
13	MS9	5.739438	-2.05511	-3.21468	0.278717	-1.42065
13	MS10	4.40698	-2.56635	-0.32241	1.88389	1.372004
14	MS1	-12.6054	1.926408	0.119125	-0.58153	0.502015
14	MS2	-11.2045	0.02919	-3.58182	1.190304	-0.19044
14	MS3	-5.8164	2.03846	2.628514	-1.36624	0.241739
14	MS4	-6.42164	2.611554	-0.73485	-3.69992	1.01529
14	MS5	-6.82915	1.194752	-3.96908	-0.71549	-3.35927
14	MS6	16.5648	9.42633	-2.05413	-1.23161	-2.42042
14	MS7	1.024051	-4.07338	1.21634	-1.80509	-0.01875
14	MS8	8.516689	-4.05004	-0.61586	0.472044	-1.82735
14	MS9	7.096617	-2.36672	-4.05046	-0.82161	-2.87553
14	MS10	4.49951	-2.17356	-0.68648	2.196119	1.094122
15	MS1	-12.8811	1.908609	-0.08711	-0.63373	0.550176
15	MS2	-13.6723	-0.72795	-2.66616	1.139835	1.122068
15	MS3	-6.14488	2.322414	1.908927	-1.39656	-0.18404
15	MS4	-7.25473	2.359612	-1.40191	-3.19783	1.042074
15	MS5	-5.66205	2.396157	-5.41018	-0.13446	-5.9499
15	MS6	15.20263	8.845336	-1.3928	-0.76751	-1.89899
15	MS7	0.728578	-4.15234	0.265298	-1.52966	-0.27512
15	MS8	3.84735	-3.64191	-0.78797	0.496576	-1.0283
15	MS9	6.047963	-2.25029	-4.61194	-0.30481	-2.7634
15	MS10	3.373128	-1.80633	-0.92057	1.675873	1.608291

APÊNDICE C – TABELA COM OS COEFICIENTES MEL-
CEPSTRAIS EXTRAÍDOS DOS SINAIS

Tabela C.1 – Seis primeiros coeficientes Mel-Cepstrais (replicas de 1 ao 4).

Replica	Configuração	c_1	c_2	c_3	c_4	c_5	c_6
1	MS1	-16.9210	-22.2110	-7.8412	-3.6083	-4.3933	4.8022
1	MS2	-13.5969	-20.1315	-3.0378	-13.6358	-3.2735	8.8081
1	MS3	11.8948	-37.6071	-3.7767	-8.5160	-7.8161	3.7991
1	MS4	0.3633	-31.9064	-2.9841	0.3731	-1.1789	6.7274
1	MS5	2.5963	-32.4457	-2.0114	-6.1207	-0.1864	7.2536
1	MS6	33.7071	-46.1875	-7.2668	-5.3614	-4.6531	6.7638
1	MS7	0.8152	-40.1748	0.2365	-12.1428	-4.2424	3.2383
1	MS8	18.7614	-48.2844	4.7162	-6.2940	3.1280	6.1223
1	MS9	20.2758	-49.8485	0.8796	-6.3546	-0.1454	3.5187
1	MS10	11.3284	-38.0102	2.5430	-6.1665	1.0502	7.8652
2	MS1	-15.0687	-21.3327	-6.8181	-2.1135	-3.8131	5.3948
2	MS2	-18.6072	-17.9317	-1.7641	-7.2820	-0.7087	7.4327
2	MS3	7.2900	-32.8473	-5.7888	-10.3472	-8.8569	3.8766
2	MS4	2.7852	-30.5463	-5.6902	-2.8672	-2.9460	3.0962
2	MS5	0.3161	-30.9036	-2.4469	-6.5102	-0.1085	7.6369
2	MS6	32.5116	-48.6721	-6.2777	-1.5472	-1.5755	6.3765
2	MS7	4.5191	-37.3456	-1.2367	-13.6916	-5.1998	3.0779
2	MS8	2.5589	-34.1922	2.8156	-12.8424	-0.5056	9.1048
2	MS9	19.3956	-45.0326	0.9192	-7.1387	-0.2976	5.0517
2	MS10	4.1226	-36.6015	-0.7174	-6.1385	-0.9783	7.6527
3	MS1	-10.5173	-17.2538	-8.0797	-3.7597	-5.1134	4.1765
3	MS2	-16.3720	-20.8712	-1.5112	-7.3369	0.2805	8.6923
3	MS3	10.7701	-32.6398	-7.5302	-12.1909	-10.1674	1.6276
3	MS4	3.1310	-26.2280	-8.9185	-6.8602	-6.8853	2.0427
3	MS5	0.0929	-28.6101	-1.0289	-4.5489	2.5474	9.1061
3	MS6	26.3998	-48.8055	-6.2084	-0.9043	-1.1229	5.6447
3	MS7	5.3962	-36.1304	-1.5552	-15.4776	-7.4714	1.5047
3	MS8	-25.1120	-14.6833	-3.1409	-11.2208	0.5387	11.3424
3	MS9	16.7649	-44.2453	-0.1215	-7.6841	-0.4423	4.3349
3	MS10	6.1463	-34.8973	1.2185	-5.4774	1.3074	8.1782
4	MS1	-9.7322	-15.7328	-10.5301	-7.1053	-8.0925	2.3868
4	MS2	-15.8223	-21.3021	-1.9578	-7.1834	-0.9174	7.1275
4	MS3	5.4118	-29.7337	-7.3238	-11.6804	-10.3834	2.8389
4	MS4	-2.6508	-22.1969	-8.0792	-7.2980	-9.5031	2.2425
4	MS5	-1.5600	-27.4515	-2.1430	-4.8783	2.8530	8.8203
4	MS6	24.4210	-47.5931	-6.0993	-0.4743	-1.1620	5.2316
4	MS7	7.1264	-34.9991	-1.9049	-14.7333	-7.2102	1.3310
4	MS8	13.2129	-44.0848	-0.3398	-14.1201	-3.2697	6.5225
4	MS9	18.9244	-39.8620	-1.9772	-9.5580	-1.2980	5.4038
4	MS10	9.1386	-31.7807	-1.1116	-8.1031	-1.2409	7.9702

Tabela C.2 – Seis primeiros coeficientes Mel-Cepstrais (replicas de 5 ao 8).

Replica	Configuração	c_1	c_2	c_3	c_4	c_5	c_6
5	MS1	-11.9421	-13.1406	-8.4057	-6.1663	-8.4227	4.0226
5	MS2	-17.0412	-20.8324	-1.7394	-5.5080	0.7109	9.1671
5	MS3	3.4509	-30.1338	-7.4457	-9.5641	-8.4961	3.0278
5	MS4	-3.6027	-24.3130	-5.4245	-0.4848	-2.3402	5.4258
5	MS5	-4.7593	-26.5511	-2.8838	-4.8167	1.0654	7.4834
5	MS6	23.9068	-46.6802	-5.6221	-0.0104	-1.3332	5.2341
5	MS7	6.1657	-35.5207	-1.8019	-12.8861	-6.5224	1.7746
5	MS8	14.8227	-46.0221	2.0508	-12.8654	-2.0009	6.5210
5	MS9	17.9615	-40.3955	-0.6367	-9.5853	-2.6839	5.0291
5	MS10	1.3823	-33.6482	-0.8722	-4.7506	2.3801	8.2122
6	MS1	-9.7303	-11.7349	-11.3331	-7.7615	-9.6625	2.1264
6	MS2	-18.0641	-19.2097	-4.0541	-7.5300	-1.3604	8.3284
6	MS3	3.2423	-30.0420	-6.6638	-9.2673	-8.3156	2.9372
6	MS4	1.4235	-21.2058	-8.2784	-3.2464	-4.3515	5.7143
6	MS5	-4.3774	-26.2047	-4.1745	-4.1798	1.6922	8.5152
6	MS6	24.7719	-44.5704	-5.6722	0.0487	-1.9681	4.9550
6	MS7	1.5736	-37.7722	-0.3885	-9.8149	-2.9552	4.1400
6	MS8	6.2770	-43.7332	1.8717	-11.0823	-0.8957	7.4338
6	MS9	17.9903	-39.8202	-1.2667	-9.6314	-3.0534	5.2744
6	MS10	2.7354	-33.1010	-3.1915	-7.1329	0.7023	6.9787
7	MS1	-11.8263	-10.2071	-11.2329	-6.1170	-9.6992	1.9954
7	MS2	-18.7543	-17.6174	-4.4306	-6.2296	-0.6810	8.8537
7	MS3	4.2473	-28.9478	-6.4993	-9.8819	-8.8955	4.3072
7	MS4	-1.1401	-20.9603	-9.9366	-6.6912	-8.6177	3.4456
7	MS5	-5.3308	-24.9036	-2.7669	-3.0816	2.3091	8.5150
7	MS6	24.6007	-43.7177	-4.5080	1.6137	0.1747	6.8676
7	MS7	2.3444	-35.4784	-1.7722	-11.9664	-4.4470	4.2991
7	MS8	10.6714	-47.5334	2.2208	-9.1246	-0.1594	6.2152
7	MS9	14.8821	-41.1090	-0.3614	-7.4469	-0.9080	5.3978
7	MS10	4.2254	-31.3672	-0.9902	-6.5392	0.7972	7.3558
8	MS1	-13.4468	-10.8135	-10.5611	-3.4899	-8.0129	3.6233
8	MS2	-18.9915	-17.4762	-5.6427	-5.0894	0.5727	7.1586
8	MS3	3.2358	-35.1032	-5.5146	-7.4808	-5.4512	5.7301
8	MS4	3.7756	-25.3748	-5.1065	-4.7649	-6.5690	6.1526
8	MS5	-6.5522	-24.7613	-4.0559	-4.8149	1.0242	7.6234
8	MS6	22.6829	-43.0835	-5.3475	0.6422	-0.6429	6.9005
8	MS7	8.8688	-38.2947	0.3244	-14.0533	-4.9916	2.5975
8	MS8	9.7460	-45.8632	3.7344	-7.1755	1.9812	7.1935
8	MS9	14.3209	-40.7242	-0.6898	-7.6808	-0.5965	6.3673
8	MS10	4.1964	-30.2596	-1.9024	-9.1189	-1.5273	7.1490

Tabela C.3 – Seis primeiros coeficientes Mel-Cepstrais (replicas de 9 ao 12).

Replica	Configuração	c_1	c_2	c_3	c_4	c_5	c_6
9	MS1	-17.1459	-10.4438	-10.9983	-1.6876	-6.7098	3.7215
9	MS2	-18.9693	-15.6748	-5.1014	-5.9805	-0.6218	8.5145
9	MS3	-0.7379	-30.5417	-6.3242	-10.3518	-7.6267	5.0413
9	MS4	4.7547	-27.9553	-9.4673	-5.3562	-6.2707	3.3710
9	MS5	-6.8876	-23.7559	-4.1940	-3.2784	2.1088	7.5219
9	MS6	-28.3313	-45.1980	-4.6231	1.9478	-0.4092	6.4719
9	MS7	1.4579	-39.3480	-0.2713	-10.0977	-2.1024	3.6238
9	MS8	10.7255	-47.5347	3.2760	-6.1924	2.0637	6.2217
9	MS9	15.8530	-39.1943	-2.9773	-11.4904	-3.7634	5.2658
9	MS10	2.9303	-29.0558	-2.0200	-8.0559	-0.6044	9.1931
10	MS1	-14.4091	-8.6640	-11.5628	-1.7120	-6.6277	4.2166
10	MS2	-20.8816	-14.9481	-6.5642	-7.5005	-1.1739	7.3419
10	MS3	-0.6628	-29.8209	-6.5619	-10.7365	-8.6366	5.1048
10	MS4	4.2380	-26.3783	-9.5602	-3.3724	-4.4164	5.3063
10	MS5	-6.6229	-24.5485	-4.9045	-4.1500	1.6145	6.4149
10	MS6	23.9708	-11.6824	-9.7503	-2.1341	1.5898	8.9603
10	MS7	3.0881	-38.7887	-1.8868	-11.8110	-2.9126	3.2719
10	MS8	16.0028	-42.1274	3.2923	-7.9236	0.3785	6.9105
10	MS9	13.1835	-39.4901	-1.1704	-7.6766	-0.6969	5.8365
10	MS10	2.2032	-27.4066	-3.0064	-7.6202	-0.7357	8.9515
11	MS1	-19.0695	-10.5302	-10.6064	1.5037	-1.9071	7.3037
11	MS2	-19.5580	-14.4230	-5.9707	-5.3513	0.8038	8.0238
11	MS3	-2.7398	-29.7830	-7.7086	-10.9963	-9.2615	5.0370
11	MS4	0.1897	-27.3284	-9.3235	-3.9568	-6.1735	5.6127
11	MS5	-5.1048	-25.5446	-1.0251	-1.6564	5.2295	8.1465
11	MS6	22.6721	-44.3268	-6.1472	0.4785	-1.6030	5.9442
11	MS7	3.7777	-37.9710	-4.0686	-10.4699	0.6012	3.5997
11	MS8	10.9358	-43.6591	1.7872	-6.3209	1.6221	5.9598
11	MS9	10.4859	-38.9722	-0.4067	-5.8638	0.5327	5.9890
11	MS10	-0.7934	-28.7173	-3.2473	-7.5076	-0.7092	7.7047
12	MS1	-18.8311	-9.6135	-11.2203	-0.7794	-4.5053	6.2671
12	MS2	-24.2597	-10.5003	-5.8406	-4.5137	0.8191	9.1253
12	MS3	0.9822	-27.4155	-7.3570	-11.6354	-10.6576	5.4203
12	MS4	1.6178	-24.2614	-7.8468	-2.3698	-4.1485	5.9001
12	MS5	-7.4930	-23.6635	-4.8945	-1.1803	3.3759	8.0027
12	MS6	28.3117	-46.1040	-4.6835	3.8024	1.3424	6.8124
12	MS7	7.6021	-35.0916	-1.6740	-12.9388	-3.4746	4.2488
12	MS8	10.9779	-43.9637	3.7383	-4.8912	3.1377	7.7119
12	MS9	14.8423	-42.4215	-4.3906	-2.2164	2.6985	1.1103
12	MS10	-1.3970	-27.1140	-4.4737	-11.4430	-4.5558	7.6040

Tabela C.4 – Seis primeiros coeficientes Mel-Cepstrais (replicas de 13 ao 15).

Replica	Configuração	c_1	c_2	c_3	c_4	c_5	c_6
13	MS1	-19.1644	-8.8981	-11.7011	-1.5292	-6.1652	6.4002
13	MS2	-21.3722	-12.7664	-5.5399	-4.8315	0.8016	8.0549
13	MS3	-2.0471	-28.5770	-8.6010	-10.6944	-8.9347	5.0717
13	MS4	-3.0250	-25.2659	-7.3130	0.0004	-2.6369	5.7040
13	MS5	-3.1635	-25.6014	-9.4851	-0.2260	3.7574	3.2569
13	MS6	27.7570	-45.1321	-5.7770	2.2806	0.1793	6.4464
13	MS7	2.3457	-36.1197	-0.8586	-10.6394	-2.7274	5.3034
13	MS8	8.9021	-45.0470	0.3368	-7.2781	0.5574	5.5704
13	MS9	17.6880	-39.2699	-3.6705	-1.4534	3.2526	1.2771
13	MS10	8.4834	-36.7224	-0.6207	-9.9307	-2.1828	9.1886
14	MS1	-15.8963	-7.4062	-11.6110	-1.3042	-5.1868	6.2510
14	MS2	-19.9608	-13.4489	-6.5883	-5.6547	0.4254	8.6341
14	MS3	-2.4099	-29.6892	-9.7989	-8.2093	-5.5072	4.8864
14	MS4	0.0450	-23.6634	-10.1365	-3.7464	-4.9606	4.5772
14	MS5	-1.5620	-25.9874	-10.7400	1.7076	2.1431	1.6376
14	MS6	25.9279	-44.3587	-6.5709	2.0684	-0.2583	5.8294
14	MS7	3.8915	-35.7845	0.8034	-9.9004	-2.1109	6.2229
14	MS8	10.9158	-43.6156	2.4596	-4.7202	3.5838	7.5719
14	MS9	-27.0757	-27.7868	-4.1080	0.2005	2.1715	3.2658
14	MS10	5.6835	-38.6120	0.2216	-6.0853	-0.3802	8.3731
15	MS1	-17.9545	-6.5734	-11.3052	-1.2849	-5.5220	7.2170
15	MS2	-32.7439	-7.2041	-8.1267	-4.9537	-0.2068	10.4097
15	MS3	-2.8434	-27.5076	-8.1888	-8.8287	-6.5735	6.7961
15	MS4	-4.8914	-24.3135	-10.4229	-0.7927	-1.0933	5.9808
15	MS5	5.0386	-27.1116	-12.9176	5.0582	1.5795	1.0471
15	MS6	25.5180	-42.9364	-3.9792	4.3085	1.8564	7.1223
15	MS7	2.9666	-35.5119	2.2121	-8.2514	-0.9076	6.4374
15	MS8	3.9787	-41.2239	1.0672	-5.3898	2.6259	6.0591
15	MS9	17.5008	-38.9532	-2.2367	-0.4532	4.2534	2.9753
15	MS10	5.7274	-36.7760	0.7718	-5.8969	-0.1174	8.5262

APÊNDICE D – CÓDIGOS PARA IMPLEMENTAÇÃO DA EXTRAÇÃO DE MFCC

Listing D.1 – Função para o calculo da FFT

```

void fft(double *temp_DATA, double *fft_DATA, int M, int fft_SIZE, int
    window_SIZE)
{
    double Tr, Ti, Ur, Ui, Wr, Wi, b; // var auxiliares para o calculo da FFT
    int ip, n1, n2, n3;           // var aux para o calculo da FFT
    register int i, j, k;        // var auxiliares para loops
    double *Xr,*Xi;             // partes real e imaginaria da FFT

    //Aloca memoria para dados float
    Xr = (double*) malloc (sizeof(double) * (fft_SIZE));
    Xi = (double *) malloc (sizeof(double) * (fft_SIZE));

    Tr = Ti = Ur = Ui = Wr = Wi = b = 0.0;
    ip = n1 = n2 = n3 = 0;

    /* Calculo da FFT de N pontos do vetor temp_DATA. Usa-se o algoritmo
       proposto no Prob 6.4, pagina 333, do livro "Digital Signal Processing"
       de Oppenheim & Schaffer. Nomenclatura usada no livro e aqui: LE= n1,
       LE1= n2, NV2=n3.
    */
    for(k=0; k<fft_SIZE; k++) // inicializacao
    {
        Xr[k] = 0.0;
        Xi[k] = 0.0;
    }
    for(k=0; k<window_SIZE; k++) Xr[k] = temp_DATA[k];

    // Embaralhamento da sequencia de entrada
    n3 = fft_SIZE / 2;
    j = 0;
    for(i=0; i<fft_SIZE-1; i++)
    {
        if(i < j)
        {

```

```
    Tr = Xr[j];
    Xr[j] = Xr[i];
    Xr[i] = Tr;
    /* Ti = Xi[j];    // para seq.complexas
    Xi[j] = Xi[i];
    Xi[i] = Ti; */
}
k = n3;
while(k <= j)
{
    j -= k;
    k /= 2;
}
j += k;
}for(k=0; k<fft_SIZE; k++) // inicializacao
{
    Xr[k] = 0.0;
    Xi[k] = 0.0;
}
for(k=0; k<window_SIZE; k++)
    Xr[k] = temp_DATA[k];

// Embaralhamento da sequencia de entrada
n3 = fft_SIZE / 2;
j = 0;
for(i=0; i<fft_SIZE-1; i++) // comeco do "for"
{
    if(i < j)
    {
        Tr = Xr[j];
        Xr[j] = Xr[i];
        Xr[i] = Tr;
        /* Ti = Xi[j];    // para seq.complexas
        Xi[j] = Xi[i];
        Xi[i] = Ti; */
    }
}
k = n3;
while(k <= j)
{
    j -= k;
    k /= 2;
}
```

```

    j += k;
}    // fim do "for" (e do embaralhamento)

// Calculo da FFT
for(k=1; k<=M; k++) // "for 1"
{
    n1 = 1;
    for(i=1; i<=k; i++) // equivale a fazer 2**k
        n1 *= 2;
    n2 = n1 / 2;
    Ur = 1.0;
    Ui = 0.0;
    b = M_PI / n2; // M_PI=3.1415... (definido em <math.h>)
    Wr = cos(b);
    Wi = -sin(b);

    for(j=0; j<n2; j++) // "for 2"
    {
        for(i=j; i<fft_SIZE; i+=n1) // "for 3"
        {
            ip = i + n2;
            Tr = Xr[ip] * Ur - Xi[ip] * Ui;
            Ti = Xr[ip] * Ui + Xi[ip] * Ur;
            Xr[ip] = Xr[i] - Tr;
            Xi[ip] = Xi[i] - Ti;
            Xr[i] = Xr[i] + Tr;
            Xi[i] = Xi[i] + Ti;
        } // fim "for 3"
        Tr = Ur;
        Ur = Ur * Wr - Ui * Wi;
        Ui = Tr * Wi + Ui * Wr;
    } // fim "for 2"
} // Fim do calculo da FFT. A seq. obtida esta na ordem sequencial

// Calculo do quadrado do modulo da FFT (equivale a energia)
for(k=0; k<fft_SIZE/2; k++)
    fft_DATA[k] = Xr[k] * Xr[k] + Xi[k] * Xi[k];
// Desalocando ponteiros
free(Xr);
free(Xi);
}

```

Listing D.2 – Função para o calculo da Pré-ênfase.

```
void pre_ênfase(double *fDATA, double *eDATA, long eSIZE)
{
    int i; //contador

    eDATA[0]=fDATA[0];
    for (i=1;i<eSIZE;i++)
    {
        eDATA[i]=fDATA[i]-0.95*fDATA[i-1];
    }
}
```

Listing D.3 – Função para a janela de Hamming

```
void hamming(double* pdHamming, long N)
{
    long i;

    for(i=0;i<N;i++)
    {
        pdHamming[i] = 0.54 - 0.46*cos((2*M_PI*i)/(N-1));
    }
}
```

Listing D.4 – Função para interpolação linear.

```
double InterpLinear(double x1,double y1, double x2, double y2, double y)
{
    // x1 componente x do ponto 1
    // y1 componente y do ponto 1
    // x2 componente x do ponto 2
    // y2 componente y do ponto 2
    // y componente y do ponto intermediario

    double a; // coeficiente angular da reta que passa pelos pontos 1 e 2
    double b; // termo independente da reta que passa pelos pontos 1 e 2
    double aux_x, aux_y; // variaveis auxiliares
    double x; // valor de retorno da funcao

    aux_x = x1 - x2;
    aux_y = y1 - y2;
```

```

a = aux_y / aux_x;
b = y1-a*x1;

x = (y - b)/a;

return (x);
}

```

Listing D.5 – Função para gerar banco de filtros mel.

```

void MelParameters_melFilters(double *mel_DATA, double *fft_DATA, int
    fft_SIZE, int filtros_SIZE, int iSampleFreq)
{
    double* pdCenterFreq; //frequencias Centrais dos Filtros Mel
    double f1,f2,f3; //frequencias para calculo da equacao da reta
    double aC, bC; //coeficientes da equacao da reta crescente
    double aD, bD; //coeficientes da equacao da reta decrescente
    double Hz_DATA; //converte dado para Hz
    double dfft_SIZE; //tamanho da fft em formato double
    int i,j; //contadores

    //Aloca memoria para dados double
    pdCenterFreq = (double *) malloc (sizeof(double) * (filtros_SIZE+1));

    //Converte fft_SIZE para double
    dfft_SIZE=fft_SIZE;

    //Inicializa frequencias centrais
    pdCenterFreq[0]=100;
    for(i=1;i<filtros_SIZE+1;i++)
    {
        if(i<10) pdCenterFreq[i]=pdCenterFreq[i-1]+100;
        else pdCenterFreq[i]=MelParameters_roundx(pdCenterFreq[i-1]*pow(2,0.2));
    }

    //Calcula os filtros
    for (i=0;i<filtros_SIZE;i++)
    {
        //Coeficientes equacao da reta
        if(i==0)
        {
            f1=0;

```

```

        f2=pdCenterFreq[i];
        f3=pdCenterFreq[i+1];
    }
    else
    {
        f1=pdCenterFreq[i-1];
        f2=pdCenterFreq[i];
        f3=pdCenterFreq[i+1];
    }

    aC=1/(f2-f1);
    bC=-f1/(f2-f1);
    aD=1/(f3-f2);
    bD=f3/(f3-f2);

    //Zera mel_DATA
    mel_DATA[i]=0.0;

    for(j=0;j<(fft_SIZE/2);j++)
    {
        Hz_DATA=j*((iSampleFreq/2)/((dfft_SIZE/2)-1));

        if(f1<=Hz_DATA && Hz_DATA<=f2)
            mel_DATA[i]=mel_DATA[i]+fft_DATA[j]*(aC*Hz_DATA+bC);
        else if(f2<Hz_DATA && Hz_DATA<=f3)
            mel_DATA[i]=mel_DATA[i]+fft_DATA[j]*(-aD*Hz_DATA+bD);
    }
}
free(pdCenterFreq);
}

```

Listing D.6 – Função para calcular coeficientes mel-cepstrais.

```

void MelParameters_melCepstrals(double* mel_DATA, double* mel_ceps_DATA, int
    P, int filtros_SIZE)
{
    int i,m; //contador
    double aux; //variavel auxiliar para calculo do mel ceps

    //Zera mel_ceps_DATA
    for(i=0;i<P;i++)
    {

```

```
    mel_ceps_DATA[i]=0.0;
}

//Calculo do Mel Cepstrais
for(m=0;m<P;m++)
{
    aux=0.0;

    for(i=0;i<filtros_SIZE;i++)
    {
        aux=aux+log(mel_DATA[i])*cos(((M_PI*m)/(2*filtros_SIZE))*(2*i+1));
    }
    mel_ceps_DATA[m]=aux;
}
}
```

APÊNDICE E – CÓDIGOS PARA IMPLEMENTAÇÃO DO CLASSIFICADOR BASEADO EM GMM

Listing E.1 – Função para o calculo da probabilidade posteriori.

```
double mixture
(
    double *x,          // feature vector
    int i,              // id of Gaussian under analysis
    struct mixture *l, // GMM models
    int nGaussians,    // number of Gaussians in the mixture
    int dim             // dimension of the acoustic vectors
)
{
    double num, den; // aux variables
    int j;           // counter

    num = l[i].p*b(x,l[i].m,l[i].s,dim);

    den = 0.0;
    for (j=0;j<nGaussians;j++)
    {
        den += l[j].p * b(x,l[j].m,l[j].s,dim);
    }
    return num/den;
}
```

Listing E.2 – Função para o atualização dos coeficientes da mistura (variância).

```
void update_s
(
    double **x,          // sequence of feature vectors x[t][i]
    int dim,             // dimension of feature vectors
    int nFrames,        // number of feature vectors in x
    struct mixture *l, // GMMs
    int nGaussians,    // number of Gaussians in the mixture
    double **m,        // gaussian means m[nGaussians][dim]
    double ***s1       // gaussian variances s[nGaussians][dim]
)
```

```
)
{
    double **s;      // Gaussian variances
    double *num, den, aux; // aux variables
    int i,j,t;      // counters
    double threshold = 0.001; // minimum values allowed for the variances

    s = *s1;

    // Allocating memory
    num = malloc(sizeof(double)*dim);

    // Updating Gaussian means
    for (i=0;i<nGaussians;i++)
    {
        den = 0.0;
        for (j=0;j<dim;j++)
            num[j] = 0.0;

        for (t=0;t<nFrames;t++)
        {
            aux = mixture(x[t],i,l,nGaussians,dim);
            den += aux;

            for (j=0;j<dim;j++) num[j] += aux*x[t][j]*x[t][j];
        }
        for (j=0;j<dim;j++) s[i][j] = num[j]/den-m[i][j]*m[i][j];
    }

    // Limiting the values of the variances
    for(i=0;i<nGaussians;i++)
        for(j=0;j<dim;j++)
            if(s[i][j] < threshold) s[i][j] = threshold;

    // Freeing memory
    free (num);

    // Returning updated variances
    *s1 = s;
}
```

Listing E.3 – Função para o atualização dos coeficientes da mistura (peso).

```
void update_p
(
    double **x,      // sequence of feature vectors x[t][i]
    int dim,         // dimension of feature vectors
    int nFrames,    // number of feature vectors in x
    struct mixture *l, // GMMs
    int nGaussians, // number of Gaussians in the mixture
    double *p
)
{
    int i,t; // counters

    // Updating mixture coefficients
    for (i=0;i<nGaussians;i++)
        p[i] = 0;

    for (i=0;i<nGaussians;i++)
    {
        for (t=0;t<nFrames;t++)
            p[i] += mixture(x[t],i,l,nGaussians,dim);

        p[i] /= (double)nFrames;
    }
}
```

Listing E.4 – Função para o atualização dos coeficientes da mistura (média).

```
void update_m
(
    double **x,      // sequence of feature vectors x[t][i]
    int dim,         // dimension of feature vectors
    int nFrames,    // number of feature vectors in x
    struct mixture *l, // GMMs
    int nGaussians, // number of Gaussians in the mixture
    double ***m1     // gaussian means m[nGaussians][dim]
)
{
    double **m;      // Gaussian means
    double *num, den, aux; // aux variables
```

```
int i,j,t;          // counters

m = *m1;

// Allocating memory
num = malloc(sizeof(double)*dim);

// Updating Gaussian means
for (i=0;i<nGaussians;i++)
{
    den = 0.0;
    for (j=0;j<dim;j++) num[j] = 0.0;

    for (t=0;t<nFrames;t++)
    {
        aux = mixture(x[t],i,l,nGaussians,dim);
        den += aux;

        for (j=0;j<dim;j++)
        {
            num[j] += aux*x[t][j];
        }
    }
    for (j=0;j<dim;j++)
    {
        m[i][j] = num[j]/den;
    }
}
// Freeing memory
free (num);

// Returning updated mean
*m1 = m;
}
```

Listing E.5 – Função que implementa uma fdp gaussiana multidimensional.

```
double b
(
    double *x,      // vetor de parametros
    double *media, // vetor com as medias
```

```
double *var, // vetor com as variancias
int ordem // ordem do vetor de parametros
)
{
    // Declaracao das variaveis locais
    double aux=0.0,aux1,aux2 = 0.0; // variaveis auxiliares
    double det; // determinante da matriz de covariancia
    double dif; // var aux p/ o calculo da gaussiana
    int i; // contador
    double prob=0.0; // probabilidade do vetor dada a distribuicao

    aux1 = 2.0*PI;
    aux2 = ordem / 2.0;
    aux1 = pow(aux1,aux2);

    // Calculando determinante da matriz de covariancia
    det = 1.0;
    for (i=0;i<ordem;i++)
        det *= var[i];

    if (det != 0)
    {
        aux2 = fabs(det);
        aux2 = pow(aux2,0.5);

        for (i=0;i<ordem;i++)
        {
            dif = x[i] - media[i];
            aux += (dif*dif)/var[i];
        }
        aux *= (-0.5);
        aux = exp(aux);

        prob = aux/(aux1*aux2);
    }
    return(prob);
}
```

Listing E.6 – Função que calcula a probabilidade posteriori para uma amostra dado o modelo GMM.

```
double testgmm
(
    double **x, // observation sequence
    int dim, //dimension of feature vectors
    int nFrames, // number of vectors in x
    struct mixture *l, // GMM
    int nGaussians
)
{
    int i,t; // counters
    double p = 0.0; // a posteriori probability
    double aux;

    for (t=0;t<nFrames;t++)
    {
        aux = 0.0;
        for (i=0;i<nGaussians;i++)
            aux += l[i].p*b(x[t],l[i].m,l[i].s,dim);
        p += log(aux);
    }
    return p;
}
```

APÊNDICE F – CÓDIGOS PARA IMPLEMENTAÇÃO DO ALGORITMO K-MEANS

Listing F.1 – Funcao que calcula n codebooks a partir de m vetores.

```

void lbg
(
    double ***codebook1, // armazena o codebook
    int n_codevectors, // numero de codebooks desejado para o quantizador
    int n_vetores, // numero de vetores exemplo para o calculo dos codebooks
    int ordem, // ordem dos vetores exemplo
    double **x // vetores exemplo para o calculo dos codebooks
)
{
    // Declaracao de variaveis locais
    double **codebook; // armazena o codebook
    double **codebook2; // codebook gerado com 2^n vetores
    double epsilon=0.005; // perturbacao (ver demonstracao do algoritmo)
    int expoente; // var aux para calcular num_vet
    int register i,j; // contadores
    int *numero; // armazena o numero de vetores em cada particao
    int num_vet; // menor numero maior ou igual a n_codevectors que e tambem
        potencia de 2.
    int n_vet; // tamanho atual do codebook
    double *p1,*p2; // variaveis auxiliares no calculo de distancias euclidianas
    double **sommas; // variavel auxiliar

    codebook = *codebook1;

    // Determinando menor numero maior ou igual a n_codevectors, e que seja
        potencia
    // de 2 (pois na versao splitting, o LBG calcula apenas codebooks com numero
    // de vetores que seja potencia de 2)
    expoente = 0;
    num_vet = pow(2,expoente);
    while (num_vet < n_codevectors)
    {
        expoente++;
    }
}

```

```
    num_vet = pow(2,expoente);
}

// Alocando memoria p/ os ponteiros

// Codebook
codebook2 = malloc(sizeof(double *)*num_vet);
for (i=0;i<num_vet;i++)
    codebook2[i] = malloc(sizeof(double)*ordem);

// Matriz auxiliar no calculo do codebook
somas = malloc(sizeof(double *)*num_vet);
for (i=0;i<num_vet;i++)
    somas[i] = malloc(sizeof(double)*ordem);

// Var's auxiliares no calculo da dist. Euclidiana
p1 = malloc(sizeof(double)*ordem);
p2 = malloc(sizeof(double)*ordem);

// Numero de vetores em cada particao
numero = malloc(sizeof(int)*num_vet);

// Calculando centroide de todos os vetores (codebook inicial)
calcula_centroide(&codebook2[0],n_vetores,ordem,x);
n_vet = 1;

printf("Number of vectors: %d\n",n_vet);
// Calculando demais codebooks
while (n_vet < num_vet)
{

    // Splitting matriz codebook
    for (i=0; i < n_vet; i++)
        for (j=0; j < ordem; j++)
            somas[i][j] = codebook2[i][j];

    for (i=1; i<=n_vet; i++)
        for (j=0; j < ordem; j++)
        {
            codebook2[2*(n_vet-i)][j] = somas[n_vet-i][j] - epsilon;
            codebook2[2*(n_vet-i)+1][j] = somas[n_vet-i][j] + epsilon;
```

```
    }

    // Proximo codebook
    n_vet *= 2;

    printf("Number of vectors: %d\n",n_vet);

    // Calculando novo codebook
    novo_codebook(&codebook2,&numero,n_vet,n_vetores,ordem,&somas,x);
}

//-----
// Temos agora um codebook otimizado com num_vet vetores codigo
// Se (n_codevectors < num_vet), devemos retirar alguns vetores codigo do
// codebook
// O criterio de retirada e o numero de vetores de treinamento associados a
// cada
// um dos vetores codigo. Aqueles com menor numero de vetores de treinamento
// associados sao eliminados primeiro.

if (n_codevectors < num_vet)
{
    // Eliminando vetores codigo de 'codebook2' para gerar 'codebook' com
    // 'n_codevectors'
    // vetores codigo
    ordena_elimina(codebook2,&codebook,n_codevectors,numero,num_vet,ordem);

    // Atualizando codebook
    novo_codebook(&codebook,&numero,n_codevectors,n_vetores,ordem,&somas,x);
}
else
    for (i=0;i<n_codevectors;i++)
        for (j=0;j<ordem;j++)
            codebook[i][j] = codebook2[i][j];

// Desalocando ponteiros
free(p1);
free(p2);

for (i=0;i<num_vet;i++)
{
    free(somas[i]);
}
```

```
    free(codebook2[i]);
}
free(somas);
free(codebook2);

free(numero);

// Retornando codebook calculado
*codebook1 = codebook;
}
```

Listing F.2 – Funcao que calcula a distância Euclideana.

```
double deuc1
(
    double *p1, // vetor 1
    double *p2, // vetor 2
    int ordem // ordem dos vetores
)
{
    int i;
    double soma;

    soma = 0.0;
    for (i=0; i<ordem; i++)
        soma += (p1[i]-p2[i])*(p1[i]-p2[i]);

    return sqrt(soma);
}
```

Listing F.3 – Funcao que retorna o centroide de um conjunto de vetores.

```
void calcula_centroide
(
    double **centroide1, // centroide
    int n_vetores,       // numero de vetores exemplo para o calculo dos
                        // codebooks
    int ordem,          // ordem dos vetores exemplo
    double **x          // vetores exemplo para o calculo dos codebooks
)
```

```

{
    // Declaracao das variaveis locais
    double *centroide; // centroide dos vetores exemplo
    int i,j; // contadores

    centroide = *centroide1;

    // Inicializando variavel que ira conter o centroide
    for (i=0;i<ordem;i++)
        centroide[i]=0.0;

    // Somando os quadros
    for (i=0;i<n_vetores;i++)
        for (j=0;j<ordem;j++)
            centroide[j] += x[i][j];

    for (i=0;i<ordem;i++)
        centroide[i] /= (float)n_vetores;

    *centroide1 = centroide;
}

```

Listing F.4 – Funcao que ordena os vetores do codebook em ordem decrescente de acordo com o numero de vetores de treinamento associado a cada um dos vetores codigo.

```

void ordena_elimina
(
    double **codebook_in, // codebook a ser reordenado
    double ***codebook_out1, // codebook reordenado com apenas 'n_codevectors'
        vetores codigo
    int n_codevectors, // numero de vetores codigo desejado
    int *numero, // numero de vetores de treinamento associados a cada vetor
        codigo
    int num_vet, // numero de vetores codigo no codebook atual
    int ordem // ordem dos vetores exemplo
)
{
    double **codebook_out; // codebook reordenado com apenas 'n_codevectors'
        vetores codigo
    int done = 0; // flag auxilliar para fim de processamento

```

```
register int i,j,k; // contadores
int *index; // indices dos vetores codigo do codebook (var aux para
ordenacao)

codebook_out = *codebook_out1;

// Alocando memoria
index = malloc(sizeof(int)*num_vet);

for (i=0;i<num_vet;i++)
    index[i] = i;

while (!done)
{
    done = 1;
    for (i=0;i<(num_vet-1);i++)
    {
        j = index[i];
        k = index[i+1];
        if (numero[j] < numero[k])
        {
            index[i] = k;
            index[i+1] = j;
            done = 0;
        }
    }
}

// Salvando 'n_codevectors' em 'codebook_out'
for (i=0;i<n_codevectors;i++)
    for (j=0;j<ordem;j++)
        codebook_out[i][j] = codebook_in[index[i]][j];

// Desalocando ponteiro
free(index);

// Codebook reduzido
*codebook_out1 = codebook_out;
}
```

Listing F.5 – Função que realiza operações para o cálculo de um novo codebook depois do splitting.

```
void novo_codebook
(
    double ***codebook1, // armazena o codebook a ser processado
    int **numero1, // armazena o numero de vetores em cada particao
    int n_vet, // tamanho atual do codebook
    int n_vetores, // numero de vetores exemplo para o calculo dos codebooks
    int ordem, // ordem dos vetores exemplo
    double ***somas1, // variavel auxiliar
    double **x // vetores exemplo para o calculo dos codebooks
)
{
    double **codebook; // // armazena o codebook a ser processado
    double distorcao; // verifica se o algoritmo convergiu
    double d_media[2]; // distancia media ao realizar a quantizacao com o
        codebook
    int register i,j,n; // contadores
    int *numero; // armazena o numero de vetores em cada particao
    double **somas; // variavel auxiliar

    numero = *numero1;
    somas = *somas1;
    codebook = *codebook1;

    // Inicializando vetor que armazena a distancia media
    d_media[0] = d_media[1] = 1e20;

    // Inicializando a variavel numero (conta quantos vetores em cada particao)
    for (i=0;i<n_vet;i++)
        numero[i] = 0;

    // Inicializando a variavel somas (var aux p/ calcular o centroide de cada
        particao)
    for (i=0;i<ordem;i++)
        for (j=0;j<n_vet;j++)
            somas[j][i] = 0.0;

    // Associando vetores a particoes
    associa(codebook,&d_media[1],&numero,n_vet,n_vetores,ordem,&somas,x);
}
```

```
distorcao = 1;
n=0;
while (distorcao > 0.005)
{
    // Calculando o centroide de cada particao
    for (i=0;i<n_vet;i++)
        for (j=0;j<ordem;j++)
            codebook[i][j] = somas[i][j]/numero[i];

    // Inicializando a variavel numero (conta quantos vetores em cada particao)
    for (i=0;i<n_vet;i++)
        numero[i] = 0;

    // Inicializando a variavel somas (var aux p/ calcular o centroide de cada
    // particao)
    for (i=0;i<ordem;i++)
        for (j=0;j<n_vet;j++)
            somas[j][i] = 0.0;

    // Calculando a distancia media com este codebook
    d_media[0] = d_media[1];
    d_media[1] = 0.0;

    // Associando vetores a particoes
    associa(codebook,&d_media[1],&numero,n_vet,n_vetores,ordem,&somas,x);

    d_media[1] /= (float)n_vetores;

    // Calculando a distorcao
    if (d_media[1] != 0)
        distorcao = (d_media[0]-d_media[1])/d_media[1];
    else
        distorcao = 0; // se a distancia media e zero, temos um quantizador otimo
    if (distorcao < 0.0)
        distorcao=1;
    n++;
}

// Valores de retorno da funcao
*codebook1 = codebook;
*numero1 = numero;
```

```

    *somas1 = somas;
}

```

Listing F.6 – Função que associa vetores de treinamento ao vetor código mais próximo.

```

void associa
(
    double **codebook, // codebook
    double *d_medial, // distancia media com este codebook
    int **numero1, // armazena o numero de vetores em cada particao
    int n_vet, // tamanho atual do codebook
    int n_vetores, // numero de vetores exemplo para o calculo dos codebooks
    int ordem, // ordem dos vetores exemplo
    double ***somas1, // variavel auxiliar
    double **x // vetores exemplo para o calculo dos codebooks
)
{
    //double distancia; // distancia euclidiana entre dois vetores
    double d_media; // distancia media com este codebook
    int register i,l; // contadores
    double minimo; // menor distancia
    int *numero; // armazena o numero de vetores em cada particao
    int qual; // vetor codigo associado ao vetor de treinamento atual
    double **somas; // variavel auxiliar

    somas = *somas1;
    numero = *numero1;
    d_media = *d_medial;

    // Calculando distancias dos parametros aos centroides
    for (i=0;i<n_vetores;i++)
    {
        // Verificando vetor codigo mais proximo de x[i]
        qual = mais_proximo(codebook,&minimo,n_vet,ordem,x[i]);

        // Atualizando contagens dos clusters
        for (l=0;l<ordem;l++)
            somas[qual][l] += x[i][l];
        numero[qual]++;

        // Atualizando d_media
    }
}

```

```

    d_media += minimo;
}

// Valores de retorno da funcao
*d_media1 = d_media;
*somas1 = somas;
*numero1 = numero;
}

```

Listing F.7 – Função que verifica qual vetor mais próximo de x e uma lista de vetores.

```

int mais_proximo
(
    double **lista, // lista de vetores com a qual o vetor 'x' ira ser analisado
    double *minimo1, // distancia euclidiana entre o vetor exemplo e o vetor
        codigo mais proximo
    int n_codevectors, // numero de vetores no codebook
    int ordem, // ordem dos vetores
    double *x // vetor sob analise
)
{
    double distancia; // distancia euclidiana entre dois vetores
    int register l; //contadores
    double minimo; // distancia euclidiana entre o vetor exemplo e o vetor
        codigo mais proximo
    int qual; // vetor codigo mais proximo do vetor exemplo

    minimo = *minimo1;

    minimo = 1e20;

    for (l=0;l<n_codevectors;l++)
    {
        distancia = deucl(x,lista[l],ordem);
        if (distancia < minimo)
        {
            minimo = distancia;
            qual = l;
        }
    }
}

```

```
*minimo1 = minimo;  
return qual;  
}
```

Anexos

ANEXO A – PATTERN RECOGNITION IN AUDIBLE SOUND
ENERGY EMISSIONS OF AISI 52100 HARDENED
STEEL TURNING: A MFCC-BASED APPROACH

Pattern recognition in audible sound energy emissions of AISI 52100 hardened steel turning: a MFCC-based approach

Edielson P. Frigieri¹ · Tarcisio G. Brito¹ · Carlos A. Ynoguti² · Anderson P. Paiva¹ · João R. Ferreira¹ · Pedro P. Balestrassi¹

Received: 17 January 2016 / Accepted: 5 April 2016
© Springer-Verlag London 2016

Abstract The main objective in machining processes is to produce a high-quality surface finish which, however, can be measured only at the end of the machining cycle. A more preferable method would be to monitor the quality during the cycle, what result a real-time, low-cost, and accurate monitoring method that can dynamically adjust the machining parameters and keep the target surface finish. Motivated by this premise, results of investigation on the relationship between emitted sound signal and surface finish during turning process are reported in this paper. Through experiments with AISI 52100 hardened steel, this work shows that such a correlation does exist presenting strong evidences that Mel-Frequency Cepstral Coefficients, extracted from sound energy, can detect different surface roughness levels, what makes it a promising feature for real-time process quality monitoring methods.

Keywords Sound · Machining · Monitoring · Mel-frequency cepstral coefficients

1 Introduction

Customers who need parts machined place a premium on the quality of surface finish [1]. Thus, to identify changes, failures, or tears in the machining process, researchers have studied and developed monitoring methods for many years, as can be seen in the overview presented by Teti et al. [2]. Sick [3] showed that monitoring methods can be classified as either direct or indirect, where direct methods analyze measurements of tool wear or surface roughness but carrying out of such methods must either interrupt the machining process or wait until its completion. Indirect methods however make use of other sources of information, such as, cutting forces [4], vibration signals [5], electric current, image [6], and many other sources. The great advantage of indirect methods is that they can be carried out without stopping or interfering with the machining process, thereby increasing efficiency and allowing online adjustments [7].

One indirect method that has recently attracted researchers attention concerns is the sound generated during the machining process [8–10] because using the produced sound to gauge information about machining process has two major advantages: ease of installation [8] and implemented at a lower cost than other sensors [11].

Some examples of investigations into the use of acoustic emissions to determine the status of processes and structures include, but are not limited to, Boutros and Liang [12], where a method based on sound signals allied to discrete HMM was proposed to detect and diagnose mechanical faults in machining processes and rotating machinery. In the case of cutting tool, the proposed method detected correctly three different states (i.e., sharp, worn, or broken). For bearing tests, the model classified the severity of the faults

✉ Edielson P. Frigieri
edielsonpf@unifei.edu.br

Carlos A. Ynoguti
ynoguti@inatel.br

¹ Industrial Engineering Institute, Federal University of Itajuba, 1303 BPS Avenue, Itajuba MG 37500-903, Brazil

² Electrical Engineering Department, National Institute of Telecommunication, 510 João de Camargo Avenue, Santa Rita do Sapucaí MG 37540-000, Brazil

seeded in two different engine bearings with a success rate greater than 95 %.

In another work based on sound signals collected in the milling process, Ai et al. [10] showed a relationship between each order component of LPCC and the flank wear, mainly concentrated in the sixth-, seventh-, and eighth-order components of LPCC. Correlating sound patterns and tool wear was carried out in another study by Mannan et al. [9] where authors concluded that tool condition can be monitored by combining sensory data from a microphone and a central composite design (CCD) camera.

Airborne sound was found by Robben et al. [13] to be a valuable source of information in an ongoing machining process for the cut-off grinding of concrete. The authors highlighted that because of the very high sound emission of the machining process in a controlled environment, there was no total influence of environmental noise on the proposed monitoring method.

Salgado and Alonso [14] proposed a tool condition monitoring system based on feed motor current and sound signal where Singular Spectrum Analysis (SSA) was used to extract information correlated with tool wear from the sound signal. Experiments using AISI 1040 steel showed that the proposed technique is fast and reliable for tool condition monitoring.

Lu and Wang [15] analyzed the high-frequency sound signals (range between 20 and 80 kHz) generated in micro-milling process and proposed a wear monitoring method. Results indicated that the normalized sound signals can be potentially applied in monitoring methods with the proper selection of feature bandwidth and other parameters.

Downey et al. [16] examined the feasibility of audible sound emissions in the frequency range spectrum of human hearing (between 0 and 20 kHz). They observed that it was possible to correlate the sound energy with tool wear in machining operations. Furthermore, audible acoustic spectra highlighted the possibility of identifying discrete phases in the cutting interface performance.

Most of presented works have focus on monitoring tool wear, and analysis of data that is offered by the machining process from the sensor configurations. As the objective of a machining process is the final quality of the work-piece, actually many efforts are directed to find efficient ways to monitor the quality of the machining process. Surface roughness raised as an important parameter for quality monitoring [17], as presented by [18], which proposed a dynamic surface roughness monitoring system based on artificial neural networks (ANN) achieving a success rate of 99 %.

Following this idea, the fundamental aim of this work is to determine if, instead of detecting tool wear, it is possible to identify differences in audible acoustic emissions for different finishing surface roughness levels and also if

those differences are correlated with the machining cutting parameters, enabling its utilization in quality monitoring of machining process. Also, Mel-Frequency Cepstral Coefficients (MFCC) extracted from sound energy are used as acoustic spectrum features and their correlation with surface roughness are analyzed. MFCC is commonly used in other research areas, such as speech recognition [19, 20], but has not been explored for monitoring machining process, so this is the major contribution of this work.

The rest of the paper is organized as follows. The process of extracting acoustic features is presented in Section 2. In Section 3, the experiment is outlined showing all the steps for database creation. In Section 4, the obtained results are shown and also detailed in Section 5. Finally, conclusions and future work are given in Section 6.

2 Signal processing and feature extraction

2.1 Short-time fast Fourier transform

The sound that comes from the machining process is a dynamic signal: it varies with time due to small differences in the machined material, vibration, fluctuations in rotating speed, etc. Therefore, it is reasonable to consider it as a non-stationary signal and the use of classic FFT for analysis can lose the dynamic behaviors of all frequency components along the time. Such type of signal requires analysis in both time and frequency domain which can be achieved by the use of short-time Fourier transform (STFT), or spectrogram (power of STFT) [21, 22]. The STFT is based on the assumption that, during short time intervals, the signal can be considered stationary [23]. Therefore, the entire signal is divided into small successive overlapped frames multiplied by a non-zero window function, on which the FFT is applied [24].

According to [23], the STFT for a signal $s(t)$ can be expressed as

$$STFT(t, \omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} s(\tau)h(\tau - t)e^{-j\tau\omega} d\tau \quad (1)$$

where $h(t)$ denotes a window function at time t . Window functions are used to avoid signal distortion and ensure a smooth transition from frame to frame of the estimated parameters [25]. A Hamming window was selected to be used in this paper.

Thus, the energy density at instant t can be calculated as

$$SPE(t, \omega) = |STFT(t, \omega)|^2 \quad (2)$$

where $STFT(t, \omega)$ is obtained from Eq. 1.

As recorded data is a discrete signal, time instants are updated on a frame by frame basis where a frame consists of N_S samples of signal representing a time frame T_F . This

is the length of time (in seconds) over which a set of parameters is valid. Then the analysis moves T_W samples forward for a new frame k , and a new frame energy is calculated. It is important to notice that there is a superposition between adjacent frames. The amount of overlap to some extent controls how quickly parameters can change from frame to frame, which normally corresponds to an overlap of 50 % [25]. Finally, the result for each frame $k = 1, 2, \dots, K$, is a spectral energy vector \vec{S}_k containing $N/2 + 1$ values of energy, where N corresponds to the number of FFT points. The total number of frames K depends on the selected frame interval (T_F) and time shift (T_W) besides the total signal length.

2.2 Mel-frequency cepstral coefficients

Depending on the size of N , a correlation analysis of all $N/2 + 1$ energy values with each machining parameter and also finishing surface roughness may not be satisfactory since many energy components tend to be auto-correlated. Also, this representation should remove as much irrelevant information as possible, for instance intensity, background noise, and retain only the information relevant for the signal classification. Thus, the acoustic vectors should attend the following requirements:

1. be of low dimensionality to allow a reliable estimation of the free parameters of the recognition system,
2. be independent of the recording environment, and
3. be characteristic for each machine setup, to allow an optimal discrimination between the different acoustic models.

One such parameter that has achieved a great success in this task is the MFCC feature, first proposed for the speech recognition problems [26]. It consists on a filter bank-based approach, which can be regarded as a crude model of the

initial stages of transduction in the human auditory system [25].

There are two main motivations for the filter bank representation: first, the position of maximum displacement along the basilar membrane for stimuli such as pure tones is proportional to the logarithm of the frequency of the tone; second, experiments in human perception have shown that frequencies of a complex sound within a certain bandwidth of some nominal frequency cannot be individually identified. When one of the components of this sound falls outside this bandwidth, it can be individually distinguished. We refer to this bandwidth as the critical bandwidth [25]. A critical bandwidth is nominally 10 to 20 % of the center frequency of the sound.

The *Mel* scale attempts to map the perceived frequency of a tone onto a linear scale. It is often approximated as a linear scale from 0 to 1 kHz, and then a logarithmic scale beyond 1 kHz, and it is possible to define a mapping from the actual frequency f to a perceptual frequency scale by

$$f_{mel} = 2595 \log_{10} \left(1 + \frac{f}{700} \right). \tag{3}$$

The process for extracting the MFCC using this idea is outlined as follow. The first step is to calculate the energy spectrum of each frame k [19, 27] that is given by Eq. 2. Next, each energy frame S_k is filtered using a triangular filter bank, with the center frequencies (see Table 1) of each filter calculated by Eq. 3, resulting \tilde{s}_k . The filtering process can be represented by

$$\tilde{s}_k(l) = \sum_{j=0}^{N/2} S_k(j) M_l(j), \forall l = 1, 2, \dots, L \tag{4}$$

where N is the number of FFT points and L is the number of *Mel* filters [6, 27].

Table 1 Central frequencies for a filter bank in *Mel* scale

Filter index	Center Freq. (Hz)	Filter index	Center Freq. (Hz)	Filter index	Center Freq. (Hz)
1	100	13	1516	25	8000
2	200	14	1741	26	9190
3	300	15	2000	27	10,556
4	400	16	2297	28	12,126
5	500	17	2639	29	13,929
6	600	18	3031	30	16,000
7	700	19	3482	31	18,379
8	800	20	4000	32	21,112
9	900	21	4595	33	24,251
10	1000	22	5278	34	27,858
11	1149	23	6063	35	32,000
12	1320	24	6964	36	36,758

Finally, a discrete cosine transform (DCT) is applied to the natural logarithm of the *Mel* spectrum, resulting in the mel-frequency Cepstral coefficients, as follows

$$c(m) = \sum_{i=0}^{L-1} \ln(\tilde{s}_k(i)) \cos\left(\frac{\pi m}{2L}(2i+1)\right) \quad (5)$$

where $m = 1, 2, \dots, C$ and C is the number of desired coefficients [28, 29]. Since most of the signal information is represented by the first coefficients, normally C is chosen between 12 and 20.

3 Experimental outline

3.1 Experimental setup

The experimental setup is based on a CNC lathe with a maximum rotational speed of 4000 rpm and power of 5.5 kW and using Wiper mixed ceramic (Al₂O₃+TiC) inserts (ISO code CNGA 120408S01525WH) coated with a very thin layer of titanium nitride (TiN; Sandvik-Coromant GC 6050). The workpieces were made of AISI 52100 steel with dimensions of 49–50 mm and a hardness between 49 and 52 HRC, up to a depth of 3 mm below the surface. Figure 1 detail the actual experimental setup, the location of microphone in relation to cutting interface and the tool holding configuration.

Surface roughness was measured using a Mitutoyo portable roughness meter model Surftest SJ 201, fixed to a cut-off length of 0.25 mm. The sound emissions were recorded using an audio microphone connected to the soundcard of a Dell Vostro laptop, using the software

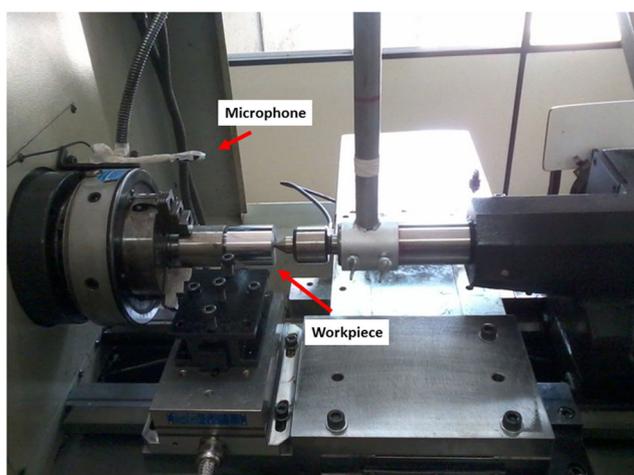


Fig. 1 Machining setup

Audacity [30], configured for a sample frequency of 44.1 kHz with a resolution of 16 bits.

3.2 Experimental methodology

CCD was used to determine the machining operational parameter values and to create a sequential set of experimental runs [31]. Due to the number of tools available for the experiment, 10 sets of machining parameters were selected based on D-optimality criteria [32]. Table 2 shows the resulted values for each experimental run considering cutting speed (CS m/min), feed rate (F mm/rev), depth of cut (D mm) and material removal rate (MRR).

During the experiment, it was guaranteed that there were no other machining operations taking place in the vicinity, which might have added interfering acoustic components and compromise the signals being detected. Each row of Table 2 was replicated 15 times to guarantee the statistical reliability of the experiment, resulting a total of 150 experimental runs. For each execution, sound signal was recorded and the following surface roughness parameters were measured: arithmetic average surface roughness (Ra), maximum surface roughness (Ry), root mean square roughness (Rq), 10-point height (Rz), and maximum peak to valley (Rt), measured three times at four positions in the work-piece middle, as proposed by [31]. From these results, the mean and variance were calculated for each surface roughness obtained. For each experimental cycle, the only change was the replacement of the cutting tool insert with an identical in order to guarantee the same wear condition.

4 Results

As outlined measurements of the surface finish of the work-piece were taken regularly during the experimental cycles to

Table 2 D-Optimal design for hard turning of AISI 52100 steel

Machining Setup	CS	F	D	MRR
Ms1	200.00	0.10	0.10	2.00
Ms2	240.00	0.10	0.10	2.40
Ms3	200.00	0.20	0.10	4.00
Ms4	200.00	0.10	0.20	4.00
Ms5	240.00	0.10	0.20	4.80
Ms6	240.00	0.20	0.20	9.60
Ms7	186.36	0.15	0.15	4.19
Ms8	220.00	0.23	0.15	7.72
Ms9	220.00	0.15	0.23	7.72
Ms10	220.00	0.15	0.15	4.95

Table 3 Surface roughness average for all machining setups

Machining setup	Surface roughness					
	<i>Ra</i>	<i>Ry</i>	<i>Rz</i>	<i>Rq</i>	<i>Rt</i>	<i>Rsm</i>
Ms1	0.15	1.05	0.75	0.18	1.44	85.60
Ms2	0.16	1.17	0.81	0.20	1.56	91.30
Ms3	0.44	2.39	1.76	0.53	3.09	169.72
Ms4	0.19	1.36	1.00	0.24	1.80	58.91
Ms5	0.18	1.32	0.89	0.23	1.89	100.67
Ms6	0.52	2.90	2.24	0.63	3.41	195.46
Ms7	0.26	1.87	1.25	0.33	2.71	104.31
Ms8	0.50	2.62	2.06	0.60	3.10	212.34
Ms9	0.26	1.90	1.29	0.33	2.70	73.04
Ms10	0.19	1.42	0.97	0.24	2.10	88.14

evaluate the performance of the cutting operation and also correlate with sound energy. Table 3 shows the average of 15 replicas for all surface roughness parameters resulted for each experimental run.

Samples in the range of 7 to 10 seconds duration were taken from the audio data for all experimental trials. The sample time range was chosen considering just the stable cutting period avoiding moments when cutting tool enters and exits of the work-piece [16], as illustrated by Fig. 2.

Energy frames was obtained from audio data using $N = 1024$ FFT points, a frame length of 20 ms and frame shift of 10 ms, that are common values used in speech analysis [27] and also have been applied in machining monitoring methods [10]. For each frame, $C = 12$ Mel-frequency Cepstral Coefficients were calculated using $L = 31$ filters which resulted a total of $K = 900$ energy vectors (on average).

Differently then expected, MFCC level distribution presented a stochastic behavior along time what reflected directly to the MFCC behavior. The same behavior was noted for all audio data, as can be seen in some examples on Fig. 3. Based on that, instead of using all $K = 900$ frames, it was calculated a mean and a standard deviation vector to represent each audio data. Therefore, the MFCC of each machining setup can be represented by $\vec{c}_m = (\vec{\mu}_m, \vec{\sigma}_m)$, where $\vec{\mu}_m$ and $\vec{\sigma}_m$ are the mean and standard deviation of all K frames for each coefficient $m = 1, 2, \dots, C$, respectively, composed by $C = 12$ MFCC levels each.

5 Discussion of results

Correlation analysis was applied to identify which Mel-Frequency Cepstral Coefficient have strongly correlation

Fig. 2 Full audio recorded during machining ($CS = 200$, $F = 0.10$ and $D = 0.10$)

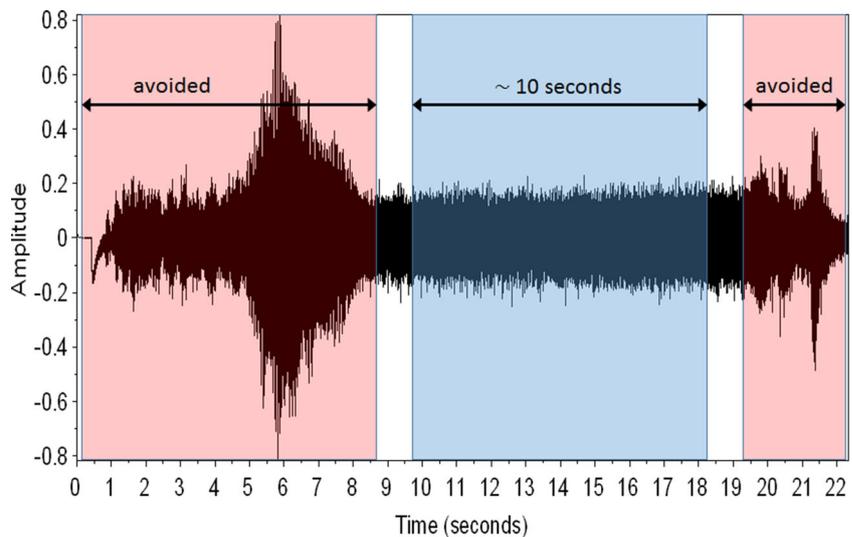
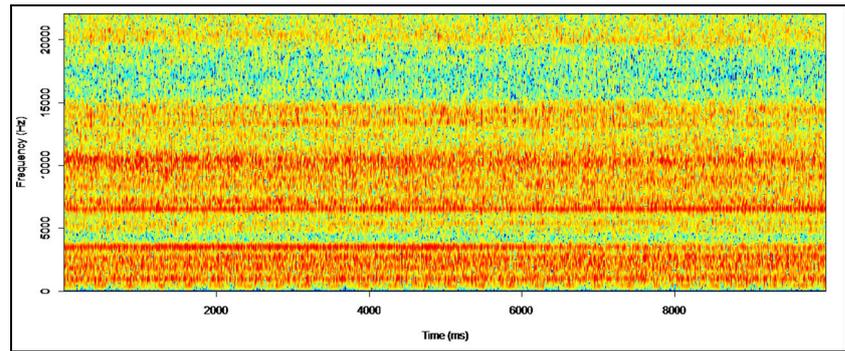
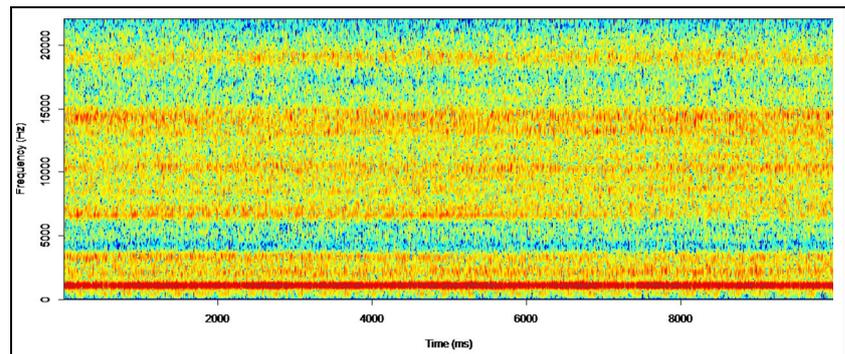


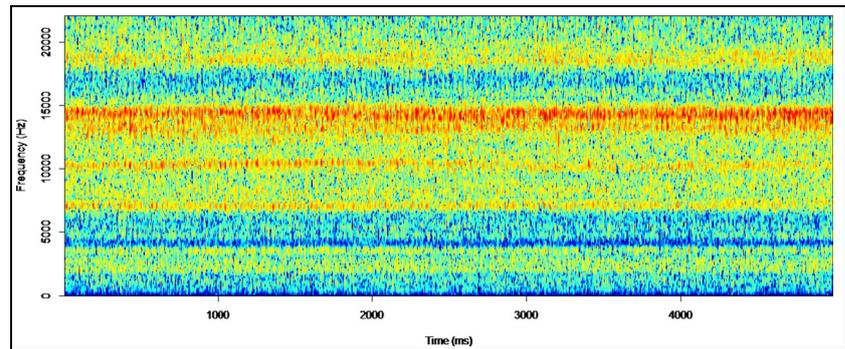
Fig. 3 Resulted spectrogram for different machining setups



(a) Machining Setup 1 (*Ms1*).



(b) Machining Setup 2 (*Ms2*).



(c) Machining Setup 8 (*Ms8*)

with machining parameters and surface roughness. Table 4 shows that each coefficient presents higher Pearson¹ correlation with different machining parameter, for example, c_1 and c_2 coefficient correlates with material remove

¹The Pearson product-moment correlation is used to assess the strength and direction of association between two continuous variables that are linearly related. Its coefficient, r , indicates the strength and direction of this relationship and can range from -1 for a perfect negative linear relationship to +1 for a perfect positive linear relationship. A value of 0 (zero) indicates that there is no relationship between the two variables.

rate (*MRR*), and also with all finishing surface roughness parameter (R_a , R_y , R_z , R_q , R_t , and R_{sm}) presenting higher correlation with maximum peak to valley surface roughness (R_t). Already c_5 coefficient, stronger correlation occurs with cutting speed (*CS*) machining parameter. Each coefficient c_m presented in Table 4 is represented by its mean μ_m since standard deviation σ_m presented no strong correlation with any machining parameter or surface roughness, and hence it was omitted.

This correlation presents that the cutting specific energy (related to the *MRR*) applied to the system is probably transformed in other types of energy like heat (tool edge, work-piece, chip, and air), sound, vibration, and others. In

Table 4 Correlation analysis between scores, machining setup parameter and surface roughness

MFCC	CS	F	D	MRR	Ra	Ry	Rz	Rq	Rt	Rsm
c ₁	0.061 ^b	<i>0.594^a</i>	<i>0.563</i>	<i>0.781</i>	<i>0.595</i>	<i>0.665</i>	<i>0.629</i>	<i>0.607</i>	<i>0.685</i>	<i>0.419</i>
	0.457 ^c	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
c ₂	-0.069	<i>-0.703</i>	<i>-0.506</i>	<i>-0.802</i>	<i>-0.649</i>	<i>-0.722</i>	<i>-0.671</i>	<i>-0.662</i>	<i>-0.752</i>	<i>-0.507</i>
	0.403	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
c ₃	0.123	<i>0.430</i>	0.210	<i>0.414</i>	0.228	0.265	0.220	0.235	0.318	0.285
	0.135	0.000	0.010	0.000	0.005	0.001	0.007	0.004	0.000	0.000
c ₄	0.440	-0.240	0.334	0.204	-0.040	-0.053	-0.003	-0.048	-0.158	-0.097
	0.000	0.003	0.000	0.012	0.629	0.523	0.974	0.558	0.054	0.238
c ₅	<i>0.673</i>	0.029	0.437	0.432	0.003	0.037	0.037	0.006	0.018	0.021
	0.000	0.722	0.000	0.000	0.974	0.657	0.656	0.938	0.831	0.803
c ₆	<i>0.566</i>	0.001	-0.098	0.058	-0.052	-0.126	-0.066	-0.063	-0.216	0.093
	0.000	0.995	0.233	0.483	0.528	0.125	0.422	0.445	0.008	0.260
c ₇	-0.069	0.029	0.077	0.106	-0.081	-0.058	-0.039	-0.072	-0.037	-0.135
	0.403	0.720	0.349	0.196	0.327	0.477	0.637	0.378	0.650	0.099
c ₈	<i>-0.607</i>	0.207	-0.228	-0.173	0.063	0.076	0.051	0.068	0.165	-0.031
	0.000	0.011	0.005	0.034	0.447	0.358	0.534	0.411	0.043	0.707
c ₉	<i>-0.562</i>	-0.021	-0.361	-0.296	-0.032	-0.066	-0.069	-0.036	-0.043	0.010
	0.000	0.799	0.000	0.000	0.694	0.424	0.403	0.666	0.602	0.900
c ₁₀	0.013	0.239	-0.032	0.110	0.291	0.237	0.237	0.280	0.206	0.346
	0.874	0.003	0.696	0.179	0.000	0.003	0.004	0.001	0.011	0.000
c ₁₁	<i>0.616</i>	-0.037	0.271	0.324	0.062	0.062	0.065	0.059	0.014	0.159
	0.000	0.653	0.001	0.000	0.453	0.448	0.433	0.472	0.867	0.052
c ₁₂	<i>0.614</i>	0.052	0.334	0.363	-0.020	0.025	0.012	-0.013	0.024	0.006
	0.000	0.529	0.000	0.000	0.808	0.758	0.880	0.872	0.772	0.946

^aItalic values represent higher Pearson correlations that are statistically significant (P -value < 5 %)

^bPearson correlation

^c P -value

this case, the part of the energy transferred to the acoustic signal can be identified by MFCC, what makes this feature a good component for applying in monitoring methods.

Figure 4 illustrates the experimental results for 15 replicas as well as determines how strongly negative were the correlations between second MFCC (c_2) and surface

Fig. 4 Line plots for standardized surface roughness parameters and Mel-frequency cepstral coefficient (c_2)

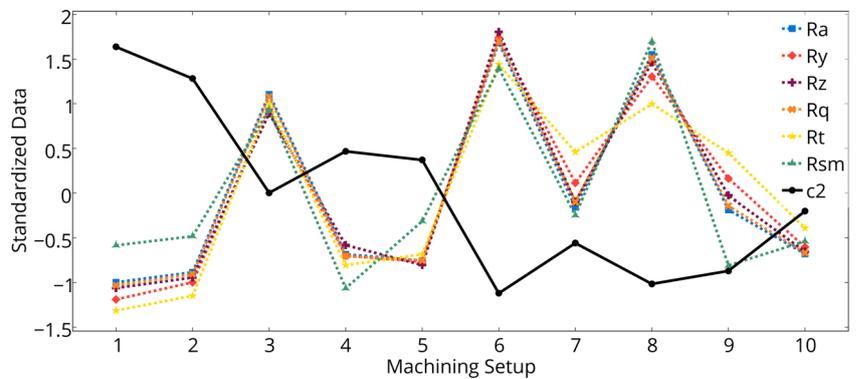
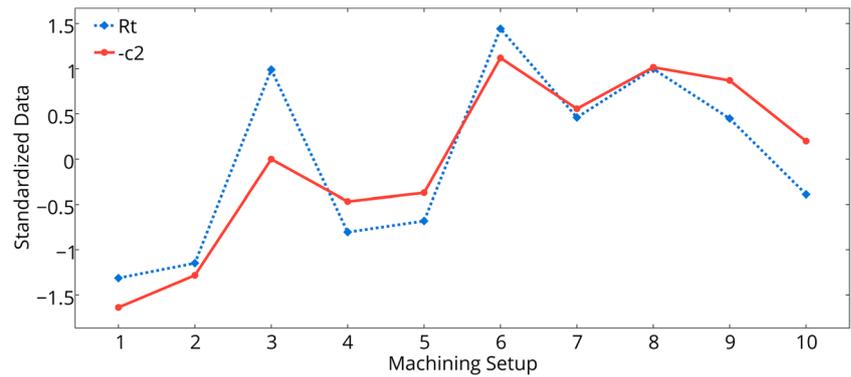


Fig. 5 Line plots for standardized surface roughness (R_t) and Mel-frequency cepstral coefficient ($-c_2$)



roughness parameters. It highlights this correlation showing that for all machining setups, when the surface roughness increases, the MFCC has the same behavior but in the opposite way, presenting a negative correlation. Also, Fig. 5 highlights that c_2 level follows surface roughness R_t with a higher correlation, as presented before on Table 4.

Two replicates were randomly selected for applying the response surface analysis in the surface roughness parameters R_a , R_y , R_z , R_q , R_t , and the Mel-frequency cepstral coefficient c_2 . Table 5 shows regression coefficients, R-Sq (adj.) and Anderson-Darling normality test for residuals of each response. The results have presented not only adequate coefficients of determination (above 95 %, exception for $R_t = 86.13$ %), but also enough evidence to affirm that residuals are following normal distributions. As can be observed in Table 5, feed rate F was the most important factor explaining the average behavior of surface roughness

parameters and was also one of the most important factors explaining the second MFCC. Reference [33] showed that the surface roughness is strongly correlated to feed rate (F) so this similarity might be one of the reasons why strong correlations between surface roughness parameters and MFCC were observed. On the other hand, the depth of cut D had higher impact on MFCC coefficient than on mechanical properties. Hence, practitioners must dedicate close attention to this control variable in order to guarantee the reliability of decision-makings for process monitoring. Probably, this is the main reason why the correlation structure was not even stronger.

Finally, Fig. 6 graphically represents the response surface models comparing two surface roughness parameters (R_y and R_t) to the second Mel-Frequency Cepstral Coefficient c_2 and their negative relationship are clearly visualized, mainly due to the contour plots for $F \times CS$ and $D \times F$.

Table 5 Regression coefficients and coefficients of determination

	R_a	R_y	R_z	R_q	R_t	c_2
Constant	0.185 ^a	1.387	0.958	0.236	2.133	-32.234
CS	0.074	0.202	0.24	0.079	0.074	2.765
F	0.169	0.746	0.616	0.194	0.773	-9.278
D	-0.053	-0.035	-0.111	-0.051	0.075	-10.073
CS^2	0.075	0.33	0.267	0.084	0.201	0.091
F^2	0.014	0.002	0.044	0.015	-0.123	0.78
D^2	0.057	0.198	0.172	0.062	0.162	3.079
$CS \times F$	0.077	0.197	0.266	0.081	0.043	5.743
$CS \times D$	-0.01	-0.051	-0.052	-0.011	-0.009	0.743
$F \times D$	-0.063	-0.081	-0.159	-0.063	0.046	-5.759
S	0.013	0.135	0.085	0.018	0.262	1.283
R-Sq(adj)	99.27 %	95.79 %	97.62 %	98.80 %	86.13 %	98.70 %
Normality test for residuals	0.433 ^b	0.268	0.117	0.118	0.328	0.382
	0.273 ^c	0.646	0.988	0.988	0.492	0.365

¹Italic values represent the significant terms in the models (P -value < 5 %)

^aAnderson-Darling test statistic

^b P -value

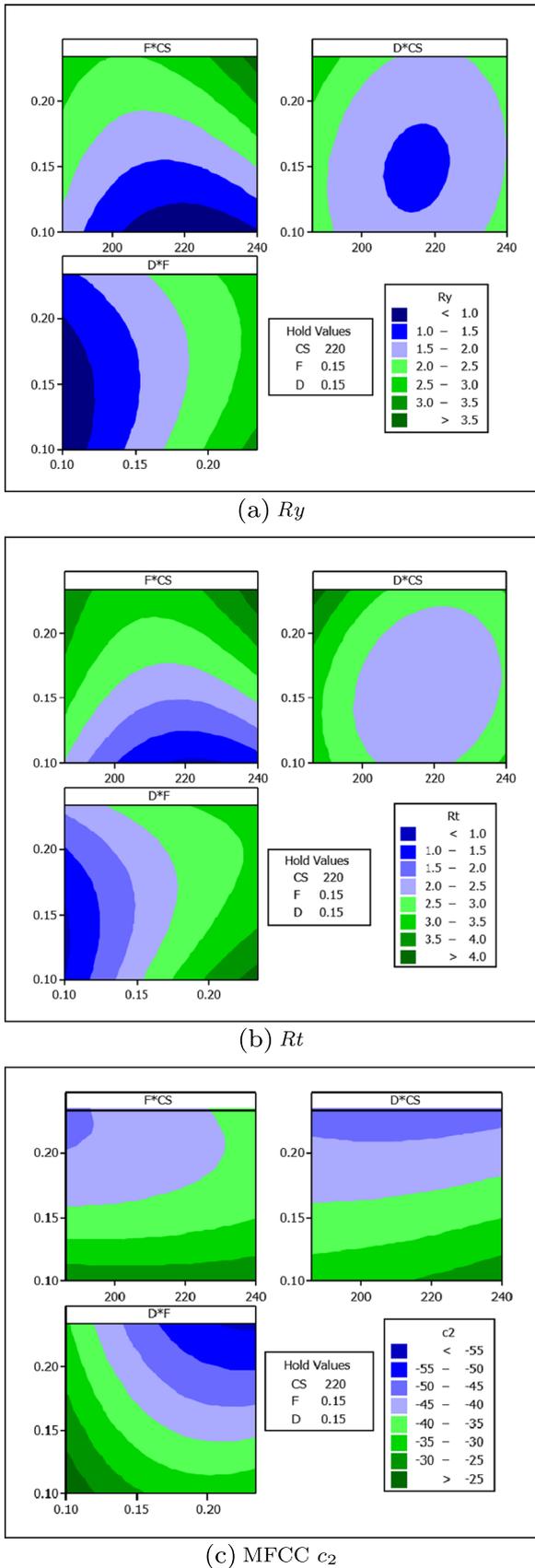


Fig. 6 Contour plots for for surface roughness and Mel coefficient

6 Conclusions

During the last years, notable efforts have been made to develop reliable and industrially applicable machining monitoring systems based on different types of sensors, especially indirect methods that does not required to interrupt the machining process. The main focus of this work is the establishment of correlation between audible sound energy and surface roughness in order to allow the effective use of such feature in quality monitoring methods.

In order to characterize the audible sound energy signals emitted by different cutting conditions during turning of AISI-5210 hardened steel, machining parameters were varied based on a CCD and the corresponding acoustic signals were detected and processed in the frequency domain by calculating the Mel-Frequency Cespstral Coefficients.

Pearson correlation analysis was performed to evaluate the linear relationship between MFCC levels and machining parameters where it was possible to detect strong correlations such as c_1 and c_2 with material removal rate (MRR) and all surface roughness, and also c_5 with cutting speed (CS). To better understand such relationship, graphical analysis was used where the linear relationship was highlighted since machining parameters change resulted proportional changes in the MFCC levels.

Also, a response surface analysis was applied in the surface roughness parameters R_a , R_y , R_z , R_q , R_t and the Mel-frequency cepstral coefficient c_2 and showed that feed rate F was the most important factor explaining the average behavior of surface roughness parameters and was also one of the most import factors explaining the second Mel coefficient, which explains the strong correlations between surface roughness parameters and MFCC. Those observations lead us to conclude that features obtained from sound energy can be applied in quality monitoring process methods.

Further works can be performed in order to test this approach for different machining process, such as milling and drilling. Classification methods based on ANN and also Gaussian Mixture Models (GMM) can be tested in order to propose monitoring methods using MFCC extracted from sound energy.

Acknowledgments This research was supported by CAPES under process number BEX 3203/15-8. Thanks also to the Brazilian governmental agencies of CNPq and FAPEMIG.

References

- Özel T, Karpat Y (2005) Predictive modeling of surface roughness and tool wear in hard turning using regression and neural networks. Int J Mach Tools Manuf 45(4–5):467–479, 4. doi:10.1016/j.ijmachtools.2004.09.007

2. Teti R, Jemielniak K, O'Donnell G, Dornfeld D (2010) Advanced monitoring of machining operations. *CIRP Ann Manuf Technol* 59(2):717–739, 1. doi:[10.1016/j.cirp.2010.05.010](https://doi.org/10.1016/j.cirp.2010.05.010)
3. Bernhard S (2002) On-line and indirect tool wear monitoring in turning with artificial neural networks: a review of more than a decade of research. *Mech Syst Signal Process* 16(4):487–546, 7. doi:[10.1006/mssp.2001.1460](https://doi.org/10.1006/mssp.2001.1460)
4. Dimla DE Sr., Lister PM (2000) On-line metal cutting tool condition monitoring. i : force and vibration analyses. *Int J Mach Tools Manuf* 40:739–768
5. Xiqing Mu, Chuangwen Xu (2009) Tool wear monitoring of acoustic emission signals from milling processes. In: 2009 First international workshop on education technology and computer science, pp 431–435. IEEE. doi:[10.1109/ETCS.2009.105](https://doi.org/10.1109/ETCS.2009.105)
6. Kasban H, Zahran O, Arafa H, El-Kordy M, Elaraby SMS, Abd El-Samie FE (2011) Welding defect detection from radiography images with a cepstral approach. *NDT & E International* 44(2):226–231, 3. doi:[10.1016/j.ndteint.2010.10.005](https://doi.org/10.1016/j.ndteint.2010.10.005)
7. Chen B, Chen X, Li B, He Z, Cao H, Cai G (2011) Reliability estimation for cutting tools based on logistic regression model using vibration signals. *Mech Syst Signal Process* 25(7):2526–2537, 10. doi:[10.1016/j.ymsp.2011.03.001](https://doi.org/10.1016/j.ymsp.2011.03.001)
8. Rubio EM, Teti R (2010) Process monitoring systems for machining using audible sound energy sensors. In: Aized T (ed) *Future Manufacturing Systems*, pp 217–235. Sciyo. doi:[10.5772/55601](https://doi.org/10.5772/55601), (to appear in print)
9. Mannan MA, Kassim AA, Ma J (2000) Application of image and sound analysis techniques to monitor the condition of cutting tools. *Pattern Recogn Lett* 21(11):969–979, 10. doi:[10.1016/S0167-8655\(00\)00050-7](https://doi.org/10.1016/S0167-8655(00)00050-7)
10. Ai CS, Sun YJ, He GW, Ze XB, Li W, Mao K (2011) The milling tool wear monitoring using the acoustic spectrum. *Int J Adv Manuf Technol* 61(5–8):457–463, 11. doi:[10.1007/s00170-011-3738-z](https://doi.org/10.1007/s00170-011-3738-z)
11. Emerson Raja J, Lim WS, Venkateshaiah C (2012) Emitted sound analysis for tool flank wear monitoring using hilbert huang transform. *International Journal of Computer and Electrical Engineering* 4(2). doi:[10.7763/IJCEE.2012.V4.460](https://doi.org/10.7763/IJCEE.2012.V4.460)
12. Boutros T, Liang M (2011) Detection and diagnosis of bearing and cutting tool faults using hidden markov models. *Mech Syst Signal Process* 25(6):2102–2124, 8. doi:[10.1016/j.ymsp.2011.01.013](https://doi.org/10.1016/j.ymsp.2011.01.013)
13. Robben L, Rahman S, Buhl JC, Denkena B, Konopatzki B (2010) Airborne sound emission as a process monitoring tool in the cut-off grinding of concrete. *Appl Acoust* 71(1):52–60. doi:[10.1016/j.apacoust.2009.07.004](https://doi.org/10.1016/j.apacoust.2009.07.004)
14. Salgado DR, Alonso FJ (2007) An approach based on current and sound signals for in-process tool wear monitoring. *Int J Mach Tools Manuf* 47(14):2140–2152. doi:[10.1016/j.ijmactools.2007.04.013](https://doi.org/10.1016/j.ijmactools.2007.04.013)
15. Lu M-C, Wan B-S (2012) Study of high-frequency sound signals for tool wear monitoring in micromilling. *Int J Adv Manuf Technol*:1785–1792, 8. doi:[10.1007/s00170-012-4458-8](https://doi.org/10.1007/s00170-012-4458-8)
16. Downey J, O'Leary P, Raghavendra R (2014) Comparison and analysis of audible sound energy emissions during single point machining of hsts with pvd ticn cutter insert across full tool life. *Wear* 313(1–2):53–62, 5. doi:[10.1016/j.wear.2014.02.004](https://doi.org/10.1016/j.wear.2014.02.004)
17. Lee SS, Chen JC (2003) On-line surface roughness recognition system using artificial neural networks system in turning operations. *Int J Adv Manuf Technol* 22(7–8):498–509, 11. doi:[10.1007/s00170-002-1511-z](https://doi.org/10.1007/s00170-002-1511-z)
18. Khorasani AM, Yazdi MRS (2015) Development of a dynamic surface roughness monitoring system based on artificial neural networks (ann) in milling operation. *Int J Adv Manuf Technol*:10. doi:[10.1007/s00170-015-7922-4](https://doi.org/10.1007/s00170-015-7922-4)
19. Koolagudi SG, Rastogi D, Sreenivasa Rao K (2012) Identification of language using mel-frequency cepstral coefficients (mfcc). *Procedia Engineering* 38:3391–3398, 1. doi:[10.1016/j.proeng.2012.06.392](https://doi.org/10.1016/j.proeng.2012.06.392)
20. Junqin W, Junjun Y (2011) An improved arithmetic of mfcc in speech recognition system. In: 2011 International conference on electronics communications and control ICECC, pp 719–722. IEEE, DOI doi:[10.1109/ICECC.2011.6066676](https://doi.org/10.1109/ICECC.2011.6066676), (to appear in print)
21. Jardine AKS, Lin D, Banjevic D (2006) A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech Syst Signal Process* 20(7):1483–1510, 10. doi:[10.1016/j.ymsp.2005.09.012](https://doi.org/10.1016/j.ymsp.2005.09.012)
22. Wang C-C, Kang Y (2012) Feature extraction techniques of non-stationary signals for fault diagnosis in machinery systems. *Journal of Signal and Information Processing* 03(01):16–25. doi:[10.4236/jsip.2012.31002](https://doi.org/10.4236/jsip.2012.31002)
23. Cohen L (1994) *Time Frequency Analysis: Theory and Applications*, 1st edn. Prentice Hall
24. Marinescu I, Axinte D (2009) A time-frequency acoustic emission-based monitoring technique to identify work-piece surface malfunctions in milling with multiple teeth cutting simultaneously. *Int J Mach Tools Manuf* 49(1):53–65. doi:[10.1016/j.ijmactools.2008.08.002](https://doi.org/10.1016/j.ijmactools.2008.08.002)
25. Picone JW (1993) Signal modeling techniques in speech recognition. *Proc IEEE* 81(9):1215–1247. doi:[10.1109/5.237532](https://doi.org/10.1109/5.237532)
26. Davis S, Mermelstein P (1980) Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences
27. Shantha Selva Kumari R, Selva Nidhyananthan S, Anand G (2012) Fused mel feature sets based text-independent speaker identification using gaussian mixture model. *Procedia Engineering* 30:319–326, 1. doi:[10.1016/j.proeng.2012.01.867](https://doi.org/10.1016/j.proeng.2012.01.867)
28. Dhanalakshmi P, Palanivel S, Ramalingam V (2011) Classification of audio signals using aann and gmm. *Appl Soft Comput* 11(1):716–723, 1. doi:[10.1016/j.asoc.2009.12.033](https://doi.org/10.1016/j.asoc.2009.12.033)
29. Fahmy MMM (2010) Palmprint recognition based on mel frequency cepstral coefficients feature extraction. *Ain Shams Engineering Journal* 1(1):39–47, 9. doi:[10.1016/j.asej.2010.09.005](https://doi.org/10.1016/j.asej.2010.09.005)
30. Dannenberg R (2013) Audacity. <http://audacity.sourceforge.net/>
31. Paiva AP, Campos PH, Ferreira JR, Lopes LGD, Paiva EJ, Balestrassi PP (2012) A multivariate robust parameter design approach for optimization of aisi 52100 hardened steel turning with wiper mixed ceramic tool. *Int J Refract Met Hard Mater* 30(1):152–163, 1. doi:[10.1016/j.ijrmhm.2011.08.001](https://doi.org/10.1016/j.ijrmhm.2011.08.001)
32. Montgomery DC (2013) *Design and Analysis of Experiments*, 8edn. Wiley
33. Tekner Z, Yeşlyurt S (2004) Investigation of the cutting parameters depending on process sound during turning of aisi 304 austenitic stainless steel. *Mater Des* 25(6):507–513, 9. doi:[10.1016/j.matdes.2003.12.011](https://doi.org/10.1016/j.matdes.2003.12.011). <http://linkinghub.elsevier.com/retrieve/pii/S0261306903002632>

ANEXO B – A MEL-FREQUENCY CEPSTRAL COEFFICIENT-
BASED APPROACH FOR SURFACE ROUGHNESS
DIAGNOSIS IN HARD TURNING USING ACOUS-
TIC SIGNALS AND GAUSSIAN MIXTURE MODEL



Technical note

A mel-frequency cepstral coefficient-based approach for surface roughness diagnosis in hard turning using acoustic signals and gaussian mixture models



Edielson P. Frigieri^{b,*}, Paulo H. S. Campos^a, Anderson P. Paiva^a, Pedro P. Balestrassi^a, João Roberto Ferreira^a, Carlos A. Ynoguti^b

^a Industrial Engineering Institute, Federal University of Itajuba, Minas Gerais, Brazil

^b Department of Computer Engineering, National Institute of Telecommunication, Minas Gerais, Brazil

ARTICLE INFO

Article history:

Received 10 May 2016

Received in revised form 28 June 2016

Accepted 29 June 2016

Keywords:

Hard turning

Mel-frequency cepstral coefficients

Gaussian mixture models

Sound

Monitoring

ABSTRACT

During the last years, notable efforts have been made to develop reliable and industrially applicable machining monitoring systems based on different types of sensors, especially indirect methods that does not required to interrupt the machining process. As the main objective in machining processes is to produce a high-quality surface finish which, however, can be measured only at the end of the machining cycle, a more preferable method would be to monitor the quality during the cycle. Motivated by this premise, results of investigation on the relationship between audible sound emitted during process and the resulted surface finish are reported in this paper. Through experiments with AISI 52100 hardened steel, this work shows that such a correlation does exist between the surface roughness and the Mel-Frequency Cepstral Coefficients (MFCCs) and based on that correlation, a new quality monitoring method is proposed using Gaussian Mixture Models (GMM). Obtained results show that this method can identify three different levels of surface roughness with an average accuracy of 98.125%.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Customers who need parts machined place a premium on the quality of surface finish [1]. Thus to identify changes, failures, or tears in the machining process, researchers have studied and developed monitoring methods for many years, as can be seen in the overview presented by Teti et al. [2]. Monitoring methods can be classified as either direct or indirect, where direct methods analyze measurements of tool wear or surface roughness but, carrying out of such methods must either interrupt the machining process or wait until its completion [3]. Indirect methods however, can be carried out without stopping or interfering with the machining process, thereby increasing efficiency and allowing online adjustments [4]. To achieve this advantage it is necessary to make use of other sources of information, such as, cutting forces

[5], vibration signals [6], acoustic emission [7], electric current, image [8], and many other sources.

Among all source of information used indirect methods, the sound generated during machining process has recently attracted researchers attention [9–11]. Mainly because using the produced sound to gauge information about machining process has two major advantages: ease of installation [9] and implementation at a lower cost than other sensors [12].

Some examples of investigations into the use of sound emissions to determine the status of processes and structures include, but are not limited to, Lu and Wan [13] that analyzed the high-frequency sound signals (range between 20 and 80 kHz) generated in micro-milling process and proposed a tool wear monitoring method. Results indicated that the normalized sound signals can be potentially applied in monitoring methods with the proper selection of feature bandwidth and other parameters.

Also, airborne sound was found by Robben et al. [14] to be a valuable source of information in an ongoing machining process for the cut-off grinding of concrete. The authors highlighted that because of the very high sound emission of the machining process in a controlled environment, there was no total influence of environmental noise on the proposed monitoring method.

* Corresponding author at: Department of Computer Engineering, National Institute of Telecommunication, 510 João de Camargo Avenue, Santa Rita do Sapucaí, MG 37540-000, Brazil.

E-mail addresses: edielson@inatel.br (E.P. Frigieri), paulo.campos@unifei.edu.br (P.H.S. Campos), andersonpaiva@unifei.edu.br (A.P. Paiva), pedro@unifei.edu.br (P.P. Balestrassi), jorofe@unifei.edu.br (J.R. Ferreira), ynoguti@inatel.br (C.A. Ynoguti).

In the case of audible sound emissions in the frequency range spectrum of human hearing (between 0 and 20 kHz), Downey et al. [15] observed that it was possible to correlate the sound energy with tool wear in machining operations. Furthermore, audible acoustic spectra highlighted the possibility of identifying discrete phases in the cutting interface performance. This work was based in the premise that some experienced machinists are indeed capable of discerning between a good and a degraded machining process through variations in the audible sound emissions from the process.

Following this idea, the fundamental aim of this work is to determine if, besides detecting tool wear, it is possible to identify differences in audible acoustic emissions for different finishing surface roughness levels enabling its utilization in quality monitoring of machining process. One feature that can successfully model human auditory system is the Mel-Frequency Cepstral Coefficients (MFCC) [16], extensively used for speech recognition [17,18]. This feature has not been explored for monitoring machining process, what makes it the major contribution of this work.

Other important point to highlight is that most of presented work have focus on monitoring tool wear and analysis of data that is offered by the machining process from the sensor configurations. As the objective of a machining process is the final quality of the work-piece, actually many efforts are directed to find efficient ways to monitor the quality of the machining process, where surface roughness raised as an important parameter for quality monitoring [19,20]. Using the MFCC extracted from sound energy as acoustic spectrum features, a new surface roughness diagnosis approach based on Gaussian Mixture Models (GMM) is proposed.

The rest of the paper is organized as follows. Section 2 provides a detailed explanation of the methodology. In Section 3, the experiment is outlined showing all the steps for database creation. In Section 4, the obtained results are presented. Finally, conclusions and future work are given in Section 5.

2. The methodology

Fig. 1 presents an overview of the presented methodology. First, a windowing approach is used over the sound signal providing a frame-by-frame basis. Second, mel-frequency cepstral coefficients are extracted from the windowed sound signal. Next, extracted

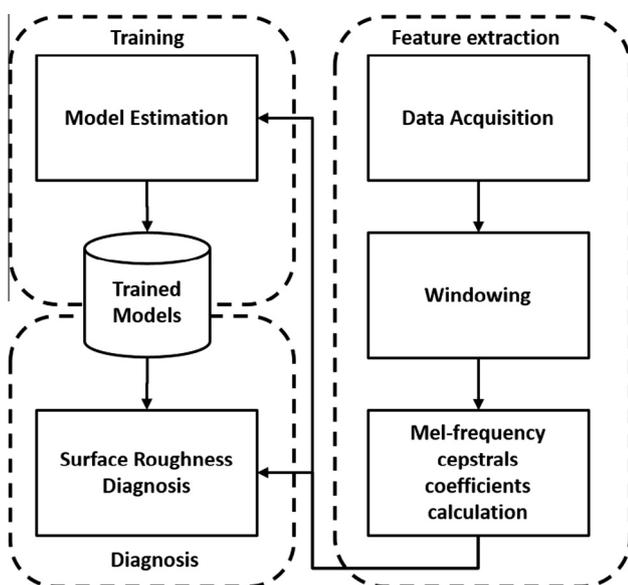


Fig. 1. Overview of the methodology.

coefficients can be used for training and estimate GMM models or, if all models are already estimated, used for surface roughness diagnosis. The methodology will be presented in 4 sections. Section 2.1 discusses how sound signal is divided in frames and the power spectral density is obtained for each frame. Section 2.2 shows the feature extraction technique used to calculate the mel-frequency cepstral coefficients. This is an important step since it is used in both system training and surface roughness diagnosis, as detailed in the next sections. Then, in Section 2.3 is presented a brief review on Gaussian Mixture Models and also the models estimation procedure is detailed. Finally, Section 2.4 presents the diagnosis procedure.

2.1. Signal processing and windowing in sound signals

The sound that comes from the machining process is a dynamic signal. It varies with time due to small differences in the machined material, vibration, fluctuations in rotating speed, and so forth. Therefore, it is reasonable to update the acoustic information on a frame-by-frame basis where a frame consists of F samples of the signal corresponding to a period of t_F seconds. This corresponds to the length of time (in seconds) over which a set of parameters is valid. Then the analysis moves t_W samples forward for a new frame, where $W \leq F$, resulting in superposition between adjacent frames. The amount of overlap to some extent controls how quickly parameters can change from frame to frame, which normally corresponds to an overlap of 50% [16].

Each frame is multiplied by a windowing function in order to smooth quick variations and increase continuity between frames [18,21]. For this work, it was used a Hamming function that can be calculated by

$$w(n) = 0.54 - 0.46 \cos(2\pi n / (F - 1)), 0 \leq n \leq F - 1$$

where F represents the frame size. The windowed frame can be expressed as

$$\tilde{x}_k(n) = x(n)w(n), 0 \leq n \leq F - 1$$

where F is the frame size. The total number of frames depends on the selected frame interval F and time shift W besides the total signal length.

Finally, for each k windowed frame is calculated the spectral energy $S_k = |X_k|^2$, where X_k is the Discrete Fourier Transform (DFT) of the windowed frame \tilde{x}_k , calculated through the FFT algorithm by

$$X_k = \sum_{n=0}^{N-1} \tilde{x}_k(n) e^{2\pi i k n / N}.$$

S_k is composed by $N/2 + 1$ values of energy where N corresponds to the number of FFT points.

2.2. Mel-frequency cepstral coefficients extraction

Depending on the number of FFT points N , a correlation analysis of all $N/2 + 1$ energy values with each machining parameter and also finishing surface roughness may not be satisfactory since many components tend to be correlated. Also, for successful classification, features extracted from signal must eliminate as much irrelevant information as possible and retain only the significant information. Thus the feature vectors should attend the following requirements: first, be of low dimensionality to allow a reliable estimation of the free parameters of the recognition system; second, be independent of the recording environment, and finally be characteristic for each machining setup, to allow an optimal discrimination between the different acoustic models.

One such parameter that has achieved a great success in this task is the Mel-Frequency Cepstral Coefficients (MFCC), first proposed for the speech recognition problems [22]. Based on filter bank, can be regarded as a crude model of the initial stages of transduction in the human auditory system [16]. The main motivations for this kind of representation are: first, the position of maximum displacement along the basilar membrane for stimuli, such as pure tones, is proportional to the logarithm of the frequency of the tone; second, experiments in human perception have shown that frequencies of a complex sound within a certain bandwidth of some nominal frequency cannot be individually identified. When one of the components of this sound falls outside this bandwidth, it can be individually distinguished. We refer to this bandwidth as the critical bandwidth, which is nominally 10–20% of the center frequency of the sound [16].

The main idea behind Mel-scale is therefore a mapping between the real frequency scale (Hz) and the perceived frequency scale (Mels). Frequencies between 0 Hz and 1 kHz are linearly approximated and then a logarithmic scale is used for frequencies beyond 1 kHz. It is possible to define a mapping from the actual frequency f to a perceptual frequency scale by

$$f_{mel} = 2595 \log_{10} \left(1 + \frac{f}{700} \right). \quad (1)$$

The process of extracting the MFCC can be outlined as follow. The first step is to separate the signal into K frames of size F , and calculate the energy spectrum S_k for each frame $k = 1, \dots, K$ [18]. This first step was detailed in Section 2.1. Next, each energy frame is filtered using a triangular filter bank, which center frequencies can be calculated by (1), resulting

$$\tilde{s}_k(l) = \sum_{n=0}^{N/2} S_k(n) M_l(n)$$

where N is the number of FFT points, L is the number of Mel-scale filters and M_l is the l th-filter [8,23]. Finally, a Discrete Cosine Transform (DCT) is applied to the natural logarithm of the Mel spectrum, resulting in the Mel-frequency Cepstral Coefficients, as follows

$$c_k(m) = \sum_{l=0}^{L-1} \log(\tilde{s}_k(l)) \cos \left(\frac{\pi m}{2L} (2l+1) \right), \quad \forall k = 1, \dots, K \quad (2)$$

where $m = 1, 2, \dots, C$ and C is the number of desired coefficients. Since most of the signal information is represented by the first coefficients, normally C is chosen between 12 and 20 [24,21].

2.3. Estimation of surface roughness models

Gaussian mixture model (GMM) was adopted in order to perform surface roughness diagnosis, as it provides a probabilistic model of the underlying sound. The use of GMM is present in many applications where features extracted from audible sound are used for some kind of classification [25,17]. According to [26], for a D dimensional feature vector \vec{x} , a Gaussian mixture density is given by

$$p(\vec{x} | \lambda) = \sum_i^M w_i b_i(\vec{x}) \quad (3)$$

The density $p(\vec{x} | \lambda)$ is a weighted linear combination of $i = 1, \dots, M$ component densities $b_i(\vec{x})$, each parameterized by a $D \times 1$ mean vector, $\vec{\mu}_i$, a $D \times D$ co-variance matrix, Σ_i , and D -mixture weights w_i that satisfy the constraint that $\sum_{i=1}^M w_i = 1$ [23,24]. Each component density $b_i(\vec{x})$ is a D -variate Gaussian function of the form

$$b_i(\vec{x}) = \frac{1}{(2\pi)^{D/2} |\Sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2} (\vec{x} - \vec{\mu}_i)' \Sigma_i^{-1} (\vec{x} - \vec{\mu}_i) \right\}$$

In summary, the complete Gaussian mixture density λ can be parameterized by $\lambda = (w_i, \vec{\mu}_i, \Sigma_i)$. The number of mixture components is empirically chosen for a given data set.

In order to generate reliable and accurate results, the GMM-based pattern recognition technique, as with other techniques (neural network, HMM, etc.), needs a training procedure prior to diagnosis.

After extracting features from signals (see details in Sections 2.1 and 2.2), it is necessary to find the best cluster separation that can accurately represent the different desired classes. Once defining such clusters, models parameters $(w_i, \vec{\mu}_i, \Sigma_i)$ can be estimated using the expectation-maximization (EM) algorithm [27]. For a sequence of T training vectors $X = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_T)$, the GMM likelihood can be calculated as $p(X | \lambda) = \prod_{t=1}^T p(\vec{x}_t | \lambda)$. So, the parameters are adjusted on each iteration, increasing the likelihood of the estimated model, that is, for iterations k and $k+1$, $p(X | \lambda^{k+1}) > p(X | \lambda^k)$. On each iteration, the mixture weights, means, and variances are re-estimated using the respective formulas

$$\bar{w}_i = \frac{1}{T} \sum_{t=1}^T p(i | \vec{x}_t, \lambda)$$

$$\bar{\mu}_i = \frac{\sum_{t=1}^T p(i | \vec{x}_t, \lambda) \vec{x}_t}{\sum_{t=1}^T p(i | \vec{x}_t, \lambda)}$$

$$\bar{\sigma}_i^2 = \frac{\sum_{t=1}^T p(i | \vec{x}_t, \lambda) \vec{x}_t^2}{\sum_{t=1}^T p(i | \vec{x}_t, \lambda)} - \bar{\mu}_i^2$$

The *a posteriori* probability for a class i can be calculated as

$$p(i | \vec{x}_t, \lambda) = \frac{w_i b_i(\vec{x}_t)}{\sum_{k=1}^M w_k b_k(\vec{x}_t)}$$

2.4. Surface roughness diagnosis

Supposing that a group of S surface roughness clusters obtained in the training procedure are represented by the GMM models $\lambda_1, \lambda_2, \dots, \lambda_S$. Surface roughness diagnosis in the process of finding the model λ with the maximum likelihood representation, calculating

$$\hat{S} = \arg \max_{1 \leq k \leq S} p(X | \lambda_k)$$

where $X = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_T)$ is the feature vector sequence under analysis. Assuming that the feature vectors of X are independent, the log-likelihood of a model λ can be computed by

$$\hat{S} = \arg \max_{1 \leq k \leq S} \sum_{t=1}^T \log p(\vec{x}_t | \lambda_k)$$

where $p(\vec{x}_t | \lambda_k)$ is calculated by (3).

3. Experimental outline

3.1. Experimental setup

The experimental setup is based on a CNC lathe with a maximum rotational speed of 4000 rpm and power of 5.5 kW and using Wiper mixed ceramic (Al₂O₃ + TiC) inserts (ISO code CNGA 120408S01525WH) coated with a very thin layer of titanium

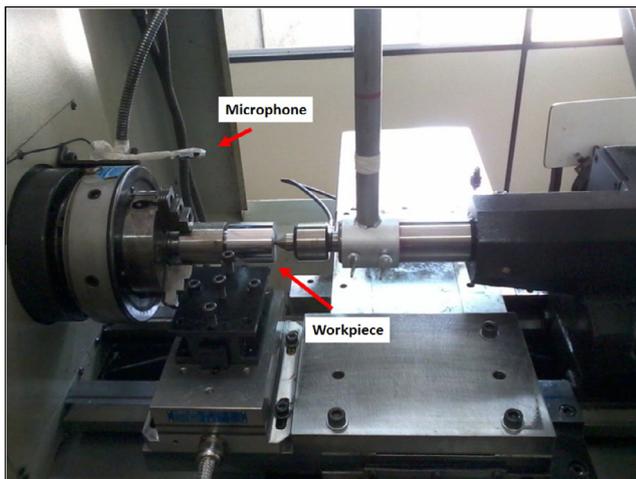


Fig. 2. Machining setup.

nitride (TiN; Sandvik-Coromant GC 6050). The work-pieces were made of AISI 52100 steel with dimensions of 49 mm 50 mm and a hardness between 49 and 52 HRC, up to a depth of 3 mm below the surface. Details related to the experimental setup such as the location of microphone in relation to cutting interface and the tool holding configuration can be seen in Fig. 2.

Sound emissions were collected using an audio microphone connected to the sound card of a Dell Vostro laptop, recorded through the software Audacity [28] using a sample frequency of 44.1 kHz and 16 bits resolution. Respective surface roughness was measured using a Mitutoyo portable roughness meter model SurfTest SJ 201, fixed to a cut-off length of 0.25 mm.

3.2. Experimental methodology

In order to determine the machining setup parameters, a Central Composite Design (CCD) was used to create a sequential set of experimental runs [29]. The number of tools available for the experiment led us to choose 10 sets of machining setups which were selected based on D-optimality criteria [30]. Table 1 shows the resulted parameter values for each experimental run considering cutting speed (CS m/min), feed rate (F mm/rev), depth of cut (D mm) and material removal rate (MRR).

Each experimental run was replicated 15 times to guarantee the statistical reliability of the experiment, resulting a total of 150 experimental runs. During the experiment, it was guaranteed that there were no other machining operations taking place in the vicinity, which might have added interfering acoustic components and compromise the signals being detected. For each execution, sound signal was recorded and the following surface roughness

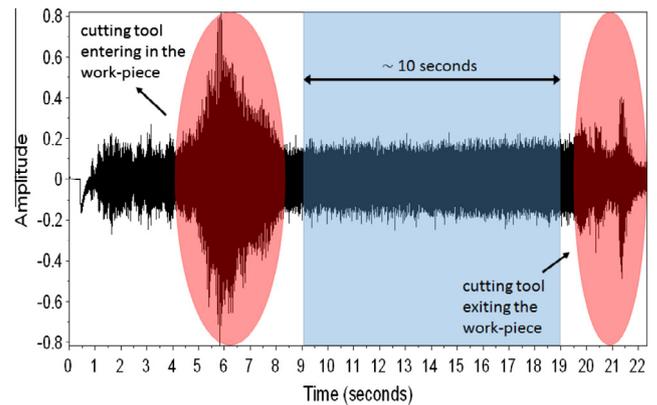


Fig. 3. Full audio recorded during machining ($CS = 200$, $F = 0.10$ and $D = 0.10$).

parameters were measured: arithmetic average surface roughness (R_a), maximum surface roughness (R_y), root mean square roughness (R_q), 10-point height (R_z), and maximum peak to valley (R_t), measured three times at four positions in the work-piece middle, as proposed by Paiva et al. [29]. Finally, mean and variance were calculated for each surface roughness obtained resulting the values presented in Table 1. For each experimental cycle, the only change was the replacement of the cutting tool insert in order to guarantee the same tool wear condition.

3.3. Feature extraction

As outlined, measurements of the surface finish of the work-piece were taken regularly during the experimental runs. Samples in the range of 7–10 s duration were taken from each recorded audio data, for all experimental runs. The sample time range was chosen considering just the stable cutting period avoiding moments when cutting tool enters and exits of the work-piece [15], as illustrated by Fig. 3. Feature vectors were extracted, according to Section 2.2, using $N = 1024$ FFT points, a frame length $t_f = 20$ ms and frame shift $t_w = 10$ ms, that are common values used in speech analysis [23] and also applied in machining monitoring methods [11]. On average, a total of $K = 900$ frames was obtained from each data file, each one composed by $C = 12$ Mel-frequency Cepstral Coefficients, calculated using $L = 31$ filters.

4. Results

4.1. Correlation analysis

Fig. 4 represents the spectrogram of two different machining setups (Ms_1 and Ms_7), where it is possible to note that power spectral density has a stochastic behavior along time. As the MFCC feature

Table 1
Surface roughness average for all machining setups.

Machining setup	Parameters				Surface roughness					
	CS	F	D	MRR	R_a	R_y	R_z	R_q	R_t	R_{sm}
Ms_1	200.00	0.10	0.10	2.00	0.15	1.05	0.75	0.18	1.44	85.60
Ms_2	240.00	0.10	0.10	2.40	0.16	1.17	0.81	0.20	1.56	91.30
Ms_3	200.00	0.20	0.10	4.00	0.44	2.39	1.76	0.53	3.09	169.72
Ms_4	200.00	0.10	0.20	4.00	0.19	1.36	1.00	0.24	1.80	58.91
Ms_5	240.00	0.10	0.20	4.80	0.18	1.32	0.89	0.23	1.89	100.67
Ms_6	240.00	0.20	0.20	9.60	0.52	2.90	2.24	0.63	3.41	195.46
Ms_7	186.36	0.15	0.15	4.19	0.26	1.87	1.25	0.33	2.71	104.31
Ms_8	220.00	0.23	0.15	7.72	0.50	2.62	2.06	0.60	3.10	212.34
Ms_9	220.00	0.15	0.23	7.72	0.26	1.90	1.29	0.33	2.70	73.04
Ms_{10}	220.00	0.15	0.15	4.95	0.19	1.42	0.97	0.24	2.10	88.14

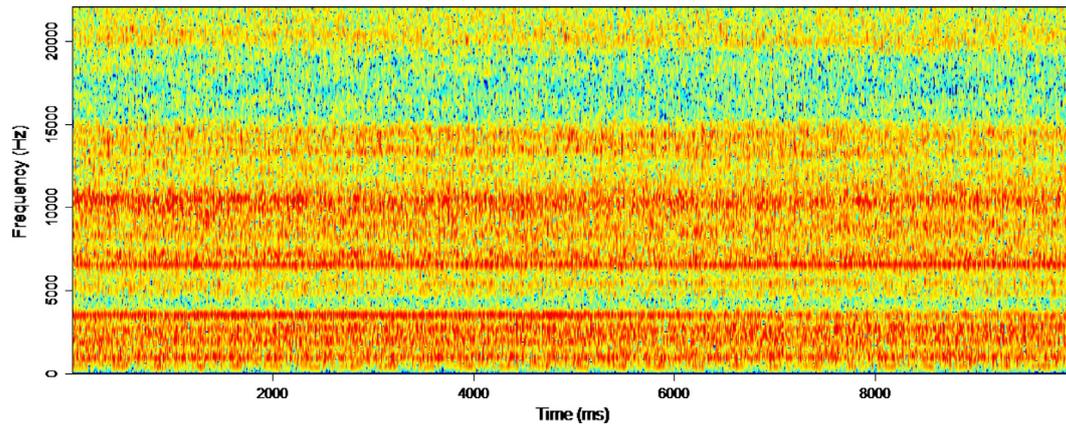
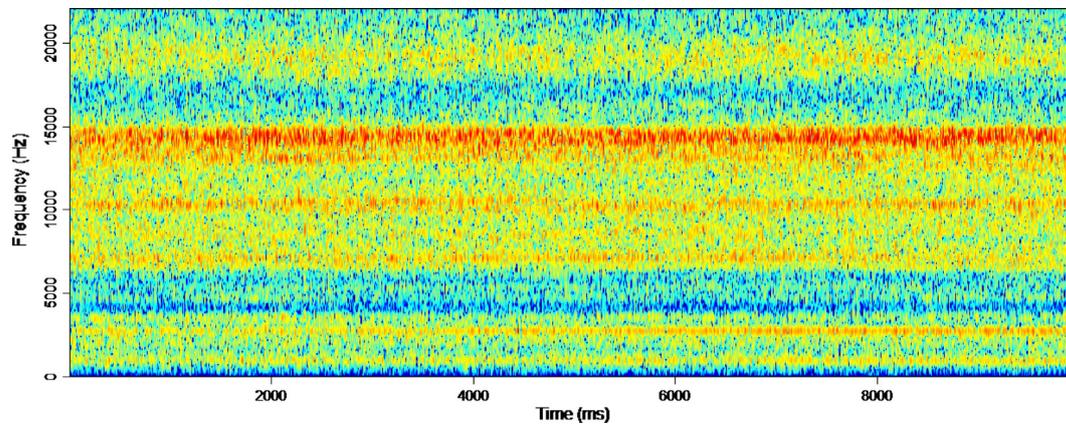
(a) Machining setup 1 (Ms_1).(b) Machining setup 7 (Ms_7).

Fig. 4. Resulted spectrogram for different machining setups.

Table 2

Correlation analysis between mel-frequency cepstral coefficients, machining setup parameter and surface roughness.

MFCC	CS	F	D	MRR	Ra	Ry	Rz	Rq	Rt	Rsm
c_1	0.061 ^a	0.594^A	0.563	0.781	0.595	0.665	0.629	0.607	0.685	0.419
c_2	-0.069	-0.703	-0.506	-0.802	-0.649	-0.722	-0.671	-0.662	-0.752	-0.507
c_3	0.123	0.430	0.210	0.414	0.228	0.265	0.220	0.235	0.318	0.285
c_4	0.440	-0.240	0.334	0.204	-0.040	-0.053	-0.003	-0.048	-0.158	-0.097
c_5	0.673	0.029	0.437	0.432	0.003	0.037	0.037	0.006	0.018	0.021
c_6	0.566	0.001	-0.098	0.058	-0.052	-0.126	-0.066	-0.063	-0.216	0.093
c_7	-0.069	0.029	0.077	0.106	-0.081	-0.058	-0.039	-0.072	-0.037	-0.135
c_8	-0.607	0.207	-0.228	-0.173	0.063	0.076	0.051	0.068	0.165	-0.031
c_9	-0.562	-0.021	-0.361	-0.296	-0.032	-0.066	-0.069	-0.036	-0.043	0.010
c_{10}	0.013	0.239	-0.032	0.110	0.291	0.237	0.237	0.280	0.206	0.346
c_{11}	0.616	-0.037	0.271	0.324	0.062	0.062	0.065	0.059	0.014	0.159
c_{12}	0.614	0.052	0.334	0.363	-0.020	0.025	0.012	-0.013	0.024	0.006

^A Bold values represent correlations that are statistically significant (P -Value < 5%).^a Pearson correlation.

vector describes the power spectral envelope of a single frame, this stochastic behavior means that there is no information in the trajectories of the MFCC coefficients over time (dynamics), and time based features as delta or delta-delta are not applicable. Based on that information, the MFCC of all K frames can be substituted by one vector containing their average, i.e., the MFCC of each machining setup can be represented by $c_m = 1/K \sum_{k=1}^K c_k(m)$, $\forall m = 1, 2, \dots, C$ and $c_k(m)$ obtained by (2) for each frame k .

Through correlation analysis was possible to identify which Mel-Frequency Cepstral Coefficients were strongly correlated with machining parameters and surface roughness (see Table 2). Each Mel coefficient presented higher correlation with different machining parameter, i.e., coefficient c_2 is correlated with material remove rate (MRR) and also with all finishing surface roughness parameter (Ra , Ry , Rz , Rq , Rt and Rsm), where maximum peak to valley surface roughness (Rt) presented highest correlation. Already c_5 coefficient

presented stronger correlation with cutting speed (CS) machining parameter. Fig. 5 illustrates how second coefficient (c_2) is strongly negative correlated with surface roughness measures. When surface roughness level increases, c_2 has the same behavior but in the opposite way, which explains the negative correlation.

With the aim of highlight such a correlation, response surface analysis was applied in two replicas randomly selected resulting the regression coefficients, R-Sq (adj.) and Anderson–Darling normality test for residuals of each response, presented in Table 3. Results have presented not only adequate coefficients of determination (above 95%, exception for $R_t = 86.13\%$), but also enough evidence to affirm that residuals are following normal distributions. As can be observed in Table 3, feed rate F was the most important factor explaining the average behavior of surface roughness parameters and was also one of the most import factors explaining c_2 . According to Tekner and Yeşlyurt [31], surface roughness is strongly correlated with feed rate (F) so this similarity might be one of the reasons why strong correlations between surface roughness parameters and MFCC were observed.

The presented analysis showed that the cutting specific energy (related to the MRR) applied to the system is probably transformed in other types of energy like vibration, heat (tool edge, work-piece, chip and air), sound and others. In this case, the part of the energy transferred to the audible sound can be identified by MFCC, proving that this feature can be applied in monitoring methods.

4.2. Cluster criteria definition

To recognize the patterns of the acoustic signals and correlate them with the surface roughness measured, it is necessary to evaluate in how many distinct clusters the measured surface roughness can be separated. Using the statistical technique of cluster hierarchical analysis based on the Ward method [32], the machining setups were separated into three different clusters: $S_1 = \{Ms_1, Ms_2\}$, $S_2 = \{Ms_3, Ms_4, Ms_5, Ms_7, Ms_{10}\}$ and $S_3 = \{Ms_6, Ms_8, Ms_9\}$. The similarity among the machining setups was evaluated regarding the behavior of the average surface roughness Ra and the material removal rate MRR . As can be observed, some machining setups resulted a surface roughness Ra not sufficiently different to support the hypothesis that it is possible to separate the 10 machining setups. In this way, some different machining setups can result similar surface roughness levels.

4.3. Surface roughness diagnosis

Based on the obtained clusters, two training and testing sets were separated using the 150 total sound data files. A second

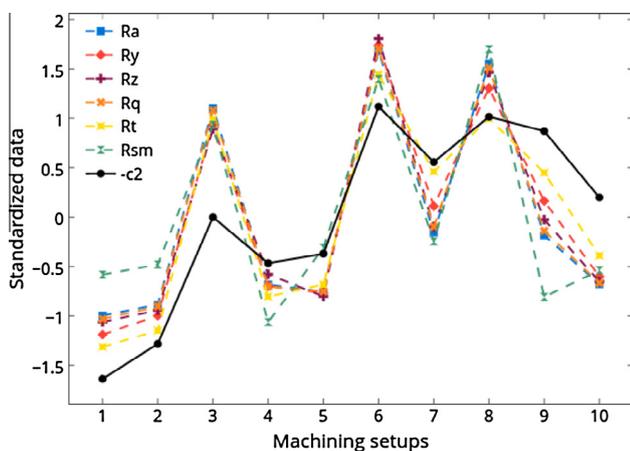


Fig. 5. Line plots for standardized surface roughness parameters and Mel-frequency cepstral coefficient ($-c_2$).

Table 3
Regression coefficients and coefficients of determination.

	Ra	Ry	Rz	Rq	Rt	c_2
Constant	0.185^A	1.387	0.958	0.236	2.133	-32.234
CS	0.074	0.202	0.24	0.079	0.074	2.765
F	0.169	0.746	0.616	0.194	0.773	-9.278
D	-0.053	-0.035	-0.111	-0.051	0.075	-10.073
CS^2	0.075	0.33	0.267	0.084	0.201	0.091
F^2	0.014	0.002	0.044	0.015	-0.123	0.78
D^2	0.057	0.198	0.172	0.062	0.162	3.079
$CS \times F$	0.077	0.197	0.266	0.081	0.043	5.743
$CS \times D$	-0.01	-0.051	-0.052	-0.011	-0.009	0.743
$F \times D$	-0.063	-0.081	-0.159	-0.063	0.046	-5.759
S	0.013	0.135	0.085	0.018	0.262	1.283
R-Sq(adj)	99.27%	95.79%	97.62%	98.80%	86.13%	98.70%
Normality test	0.433 ^a	0.268	0.117	0.118	0.328	0.382
For residuals	0.273 ^b	0.646	0.988	0.988	0.492	0.365

^A Bold values represent the significant terms in the models (P -Value < 5%).

^a Anderson–Darling statistic test.

^b P -Value.

Table 4
Gaussian mixture model for surface roughness cluster S_1 .

Density	w	μ_{c1}	μ_{c2}	σ_{c1}	σ_{c2}
1	0.12	-18.42	-17	24.56	27.1
2	0.12	-8.45	-8.81	36.79	17.75
3	0.13	-9.44	0.43	5.71	5.2
4	0.11	-6.92	-5.04	19.67	6.2
5	0.21	-6.04	4.42	5.08	5.56
6	0.14	-1.52	8.8	8.73	6.93
7	0.13	2.24	-4.22	19.27	7.5
8	0.04	11.11	-1.28	6.67	5.43

Table 5
Gaussian mixture model for surface roughness cluster S_2 .

Density	w	μ_{c1}	μ_{c2}	σ_{c1}	σ_{c2}
1	0.05	-0.85	-31.66	25.84	43.28
2	0.06	6.6	-36.51	29.29	14.68
3	0.12	-2.28	-8.48	12.27	14.09
4	0.06	-0.87	-26.41	33.89	15.11
5	0.13	-10.75	-1.44	7.32	17.53
6	0.17	3.81	-2.14	11.12	9.21
7	0.26	-6.73	4.45	7.42	13.8
8	0.14	-1.49	7.99	13.17	10.26

Table 6
Gaussian mixture model for surface roughness cluster S_3 .

Density	w	μ_{c1}	μ_{c2}	σ_{c1}	σ_{c2}
1	0.01	3.13	-35.85	17.85	14.03
2	0.04	17.39	-38.88	18.77	10.02
3	0.06	15.61	-45.36	46.75	18.78
4	0.05	27.25	-45.45	39.84	15.5
5	0.1	-0.19	-9.03	8.48	10.13
6	0.13	4.09	-3.99	6.11	7.09
7	0.01	-17.38	-10.67	83.47	17.05
8	0.61	-3.69	3.87	22.1	16.3

training and testing set was used as a confirmation experiment to ensure the statistical accuracy of the proposed system. For each machining setup, eight of the 15 sound data files were separated for training and seven for testing. According to the number of machining setups in the cluster, the number of training and testing data files can vary from 16 to 40 and from 14 to 35, respectively, considering that each machining setup has 15 different replicas.

Table 7
Surface roughness diagnosis accuracy.

Cluster No.	Cutting speed, feed rate, depth of cut	Testing set	No. of replicas	Replicas	Accuracy (%)
2	186 m/min, 0.15 mm/rev, 0.15 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
3	220 m/min, 0.23 mm/rev, 0.15 mm	1	7	2, 4, 6, 8, 10, 12, 14	85.7
3	220 m/min, 0.15 mm/rev, 0.23 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
1	240 m/min, 0.10 mm/rev, 0.10 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
2	240 m/min, 0.10 mm/rev, 0.20 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
2	220 m/min, 0.15 mm/rev, 0.15 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
3	240 m/min, 0.20 mm/rev, 0.20 mm	1	7	2, 4, 6, 8, 10, 12, 14	85.7
1	200 m/min, 0.10 mm/rev, 0.10 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
2	200 m/min, 0.20 mm/rev, 0.10 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
2	200 m/min, 0.10 mm/rev, 0.20 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
2	186 m/min, 0.15 mm/rev, 0.15 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
3	220 m/min, 0.23 mm/rev, 0.15 mm	2	7	3, 4, 7, 8, 11, 12, 15	85.7
3	220 m/min, 0.15 mm/rev, 0.23 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
1	240 m/min, 0.10 mm/rev, 0.10 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
2	240 m/min, 0.10 mm/rev, 0.20 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
2	220 m/min, 0.15 mm/rev, 0.15 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
3	240 m/min, 0.20 mm/rev, 0.20 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
1	200 m/min, 0.10 mm/rev, 0.10 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
2	200 m/min, 0.20 mm/rev, 0.10 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
2	200 m/min, 0.10 mm/rev, 0.20 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
Average					98.125

GMM model estimation algorithm was implemented in C programming language according to the method described in Section 2.3. Each surface roughness model, identified as S_1, S_2, S_3 , was estimated using $M = 8$ component densities. Observations sequence used during model estimation were composed only by the first and second Mel-frequency cepstral coefficients, as they presented the highest correlation with surface roughness parameters (see Section 4.1). The resulted GMM models for each surface roughness cluster can be seen on Tables 4–6.

Surface roughness diagnosis procedure was also developed using C programming language according to the method described in Section 2.4. Observations sequence, composed by the first and second mel-frequency cepstral coefficients, were used to find the model with the maximum likelihood. The testing set was used to access the accuracy of the proposed system calculating the number of sound signals correctly classified divided by the total number of sound signals tested. The proposed method achieved an average accuracy rate of 98.125%, as presented in Table 7. As the testing set for each surface roughness cluster is composed by different machining setups, this percentage indicates that the method is very robust and is not sensitive to cutting speed, feed rate, or depth-of-cut parameters.

5. Conclusions

Machining parameters were varied based on a Central Composite Design (CCD) in order to characterize the audible sound energy signals emitted by different cutting conditions during turning of AISI-5210 hardened steel. The corresponding acoustic signals were detected and processed in the frequency domain by extracting the Mel-Frequency Cepstral Coefficients.

The linear relationship between MFCC levels and machining parameters were evaluated through Pearson correlation analysis where it was possible to detect strong correlations such as c_1 and c_2 with material removal rate (MRR) and all surface roughness, and also c_5 with cutting speed (CS). To better understand such relationship, graphical analysis was used to demonstrate the linear relationship since changes in the machining parameters resulted in proportional changes in the MFCC levels.

Also, a response surface analysis was applied in the surface roughness parameters and MFCC. Feed rate (F) was the most

important factor explaining the average behavior of surface roughness parameters and was also one of the most important factors explaining second Mel coefficient (c_2), which highlights the strong correlations between surface roughness parameters and MFCC. Finally, using the MFCC as sound feature, this paper proposed a quality monitoring method based on Gaussian Mixture Models. Presented results validated the method and the effectiveness was evidenced by a diagnosis accuracy of 98.125%.

For future study, the proposed method could be analyzed using other machining processes like milling, drilling, welding and others. Also for future work, the MFCC applied with ANN or even SVM could be tested.

Acknowledgment

This research was supported by CAPES under process number BEX 3203/15-8. Thanks also to the Brazilian governmental agencies of CNPq and FAPEMIG.

References

- Özel T, Karpat Y. Predictive modeling of surface roughness and tool wear in hard turning using regression and neural networks. *Int J Mach Tools Manuf* 2005;45:467–79.
- Teti R, Jemielniak K, O'Donnell G, Dornfeld D. Advanced monitoring of machining operations. *CIRP Ann – Manuf Technol* 2010;59:717–39.
- Sick B. On-line and indirect tool wear monitoring in turning with artificial neural networks: a review of more than a decade of research. *Mech Syst Signal Process* 2002;16:487–546.
- Chen B, Chen X, Li B, He Z, Cao H, Cai G. Reliability estimation for cutting tools based on logistic regression model using vibration signals. *Mech Syst Signal Process* 2011;25:2526–37.
- Dimla Sr DE, Lister PM. On-line metal cutting tool condition monitoring. I: force and vibration analyses. *Int J Mach Tools Manuf* 2000;40:739–68.
- Xiqing M, Chuangwen X. Tool wear monitoring of acoustic emission signals from milling processes. In: 2009 First international workshop on education technology and computer science. IEEE; 2009. p. 431–5.
- Van Hecke B, Yoon J, He D. Low speed bearing fault diagnosis using acoustic emission sensors. *Appl Acoust* 2016;105:35–44.
- Kasban H, Zahran O, Arafa H, El-Kordy M, Elaraby S, Abd El-Samie F. Welding defect detection from radiography images with a cepstral approach. *NDT & E Int* 2011;44:226–31.
- Rubio EM, Teti R. Process monitoring systems for machining using audible sound energy sensors. In: Aized T, editor. *Future manufacturing systems*. Sciyo; 2010. p. 217–35.
- Mannan M, Kassim Aa, Jing M. Application of image and sound analysis techniques to monitor the condition of cutting tools. *Pattern Recogn Lett* 2000;21:969–79.

- [11] Ai CS, Sun YJ, He GW, Ze XB, Li W, Mao K. The milling tool wear monitoring using the acoustic spectrum. *Int J Adv Manuf Technol* 2011;61:457–63.
- [12] Raja JE, Lim WS, Venkateshaiah C. Emitted sound analysis for tool flank wear monitoring using hilbert huang transform. *Int J Comput Electr Eng* 2012;4.
- [13] Lu M-C, Wan B-S. Study of high-frequency sound signals for tool wear monitoring in micromilling. *Int J Adv Manuf Technol* 2012;1785–92.
- [14] Robben L, Rahman S, Buhl JC, Denkena B, Konopatzki B. Airborne sound emission as a process monitoring tool in the cut-off grinding of concrete. *Appl Acoust* 2010;71:52–60.
- [15] Downey J, O'Leary P, Raghavendra R. Comparison and analysis of audible sound energy emissions during single point machining of HSTS with PVD TiCN cutter insert across full tool life. *Wear* 2014;313:53–62.
- [16] Picone J. Signal modeling techniques in speech recognition. *Proc IEEE* 1993;81:1215–47.
- [17] Lu Y, Wu Z. Maximum likelihood normalization for robust speech recognition. *Appl Acoust* 2013;74:640–6.
- [18] Koolagudi SG, Rastogi D, Rao KS. Identification of Language using Mel-Frequency Cepstral Coefficients (MFCC). *Procedia Eng* 2012;38:3391–8.
- [19] Lee SS, Chen JC. On-line surface roughness recognition system using artificial neural networks system in turning operations. *Int J Adv Manuf Technol* 2003;22:498–509.
- [20] Khorasani A, Yazdi MRS. Development of a dynamic surface roughness monitoring system based on artificial neural networks (ANN) in milling operation. *Int J Adv Manuf Technol* 2015.
- [21] Fahmy MM. Palmprint recognition based on mel frequency cepstral coefficients feature extraction. *Ain Shams Eng J* 2010;1:39–47.
- [22] Davis S, Mermelstein P. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Trans Acoust Speech Signal Process* 1980;28:357–66.
- [23] Kumari RS, Nidhyananthan SS, G A. Fused mel feature sets based text-independent speaker identification using gaussian mixture model. *Procedia Eng* 2012;30:319–26.
- [24] Dhanalakshmi P, Palanivel S, Ramalingam V. Classification of audio signals using AANN and GMM. *Appl Soft Comput* 2011;11:716–23.
- [25] De Oliveira AG, Ventura TM, Ganchev TD, De Figueiredo JM, Jahn O, Marques MI, et al. Bird acoustic activity detection based on morphological filtering of the spectrogram. *Appl Acoust* 2015;98:34–42.
- [26] Reynolds DA, Rose RC. Robust text-independent speaker identification using gaussian mixture speaker models. *IEEE Trans Speech Audio Process* 1995:72–83.
- [27] Dempster AP, Laird NM, Rubin D. Maximum likelihood from incompatible data via EM algorithm. *J Roy Stat Soc* 1977;39:1–38.
- [28] Dannenberg R. *Audacity*; 2013.
- [29] Paiva A, Campos P, Ferreira J, Lopes L, Paiva E, Balestrassi P. A multivariate robust parameter design approach for optimization of AISI 52100 hardened steel turning with wiper mixed ceramic tool. *Int J Refract Met Hard Mater* 2012;30:152–63.
- [30] Montgomery DC. *Design and analysis of experiments*. 8 ed. Hoboken: John Wiley and Sons; 2013.
- [31] Tekner Z, Yeşilyurt S. Investigation of the cutting parameters depending on process sound during turning of AISI 304 austenitic stainless steel. *Mater Des* 2004;25:507–13.
- [32] Johnson RA, Wichern DW. *Applied multivariate statistical analysis*, 5 ed., vol. 6. New Jersey: Prentice-Hall Inc.; 2007.