

ORIENTADOR: Robson Pederiva
COMISSÃO JULGADORA EM 08/08/1997.
PORTARI
E APROVADA PELA
TESE DEFENDIDA POR MARCUS VINÍCIUS
ESTE EXEMPLAR CORRESPONDE A REDAÇÃO FINAL DA

**UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA MECÂNICA**

Diagnóstico de Defeitos em Equipamentos utilizando Métodos Estatísticos de Reconhecimento de Padrões

Autor: Marcus Vinícius Portari
Orientador: Prof. Dr. Robson Pederiva

08/97

P832d

32317/BC

**UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA MECÂNICA
DEPARTAMENTO DE PROJETO MECÂNICO**

Diagnóstico de Defeitos em Equipamentos utilizando Métodos Estatísticos de Reconhecimento de Padrões

Autor: Marcus Vinícius Portari

Orientador: Prof. Dr. Robson Pederiva

Curso: Engenharia Mecânica

Área de Concentração: Mecânica dos Sólidos e Projeto Mecânico

Dissertação de mestrado apresentada à comissão de Pós Graduação da Faculdade de Engenharia Mecânica, como requisito para a obtenção do título de Mestre em Engenharia Mecânica.

Campinas, 1997

São Paulo - Brasil



688.3889

UNIDADE	BC
N.º CHAMADA	TIUNICAMP
	P832d
V.	Ex.
TOMBO BC	32337
PROC.	284/57
C	☐
D	☒
PREÇO	R\$ 11,00
DATA	28/11/97
N.º CPD	

CM-00103388-1

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

P832d

Portari, Marcus Vinícius

Diagnóstico de defeitos em equipamentos utilizando métodos estatísticos de reconhecimento de padrões / Marcus Vinícius Portari.--Campinas, SP: [s.n.], 1997.

Orientador: Robson Pederiva.

Dissertação (mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Mecânica.

1. Localização de falhas (Engenharia). 2. Reconhecimento de padrões - Métodos estatísticos. 3. Bombas centrífugas - Defeitos. 4. Máquinas - Monitoração. 5. Equipamento industrial - Manutenção e reparos. I. Pederiva, Robson. II. Universidade Estadual de Campinas. Faculdade de Engenharia Mecânica. III. Título.

**UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA MECÂNICA
DEPARTAMENTO DE PROJETO MECÂNICO**

DISSERTAÇÃO DE MESTRADO

**Diagnóstico de Defeitos em Equipamentos
utilizando Métodos Estatísticos de
Reconhecimento de Padrões**

Autor: **Marcus Vinícius Portari**
Orientador: **Prof. Dr. Robson Pederiva**



Prof. Dr. Robson Pederiva
UNICAMP - FEM - Campinas



Prof. Dr. Sérgio Tonini Button
UNICAMP - FEM - Campinas



Prof. Dr. Vicente Lopes Junior
UNESP - FEIS - Ilha Solteira

Campinas, 08 de agosto de 1997.

Dedicatória

Dedico este trabalho aos meus pais:

Ademar Portari

e

Júlia Kiyomi Portari

Agradecimentos

A realização deste trabalho tornou-se possível devido à colaboração de diversas pessoas, às quais agradeço sinceramente.

- ao meu orientador Prof. Dr. Robson Pederiva pela orientação, incentivo e amizade;
- aos professores Milton e Loir do DPM-UNICAMP pelo apoio e amizade;
- aos professores Gilberto, Turra, Vicente e Amarildo da UNESP de Ilha Solteira pela grande amizade;
- aos “*the Dinamikos*” pela grande amizade, incentivo e momentos de lazer proporcionados;
- ao amigo Rivaldo Araújo Vilela pela amizade, sugestões e auxílio na montagem experimental e rotinas computacionais;
- aos técnicos Ferreira, Gastão, Maurício, Tiago e Eli pela amizade e pela colaboração na montagem experimental;
- às bibliotecárias Raquel e Rose pela atenção e gentileza nos serviços prestados;
- à CAPES pelo apoio financeiro;
- enfim, a todos os professores, colegas e funcionários da UNICAMP que auxiliaram, de forma direta ou indireta, na conclusão deste trabalho.

É preciso correr riscos, seguir certos caminhos e abandonar outros.

Nenhuma pessoa é capaz de escolher sem medo.

Paulo Coelho

Sumário

CAPÍTULO 1

INTRODUÇÃO	1
-------------------	----------

CAPÍTULO 2

REVISÃO BIBLIOGRÁFICA	8
------------------------------	----------

2.1 - Introdução	8
-------------------------	----------

2.2 - Métodos baseados em sintomas	9
---	----------

2.2.1 - Sistemas fuzzy	11
------------------------	----

2.2.2 - Árvores de decisão	15
----------------------------	----

2.2.3 - Sistemas especialistas	20
--------------------------------	----

2.2.4 - Redes neurais artificiais	28
-----------------------------------	----

CAPÍTULO 3

MÉTODOS ESTATÍSTICOS	35
-----------------------------	-----------

3.1 - Introdução	35
-------------------------	-----------

3.2 - Conceitos básicos de classificação de padrões	36
--	-----------

3.3 - Métodos estatísticos	39
-----------------------------------	-----------

3.3.1 - Métodos estatísticos paramétricos	41
---	----

3.3.2 - Métodos estatísticos de distribuição livre ou não paramétricos	46
--	----

3.4 - Considerações sobre amostras para treinamento e desempenho do classificador	52
--	-----------

3.4.1 - Considerações sobre as amostras de treinamento	52
--	----

3.4.2 - Considerações sobre o desempenho do classificador	52
---	----

CAPÍTULO 4

MONTAGEM E PROCEDIMENTO EXPERIMENTAL	54
4.1 - Descrição da montagem experimental	54
4.2 - Descrição do programa de aquisição de dados	55
4.3 - Descrição do experimento	58
4.3.1 - Índices de monitoramento utilizados	60
4.3.2 - Definição das classes de funcionamento	61
4.3.3 - Medições	64

CAPÍTULO 5

ANÁLISE DOS RESULTADOS	68
5.1 - Obtenção dos padrões rotulados e das matrizes de covariância	68
5.2 - Variação dos índices de monitoramento	69
5.2.1 - Variação dos índices de monitoramento para a condição operacional 1	69
5.2.2 - Variação dos índices de monitoramento para a condição operacional 2	71
5.2.3 - Variação dos índices de monitoramento para a condição operacional 3	72
5.3 - Desempenho do classificador	74
5.3.1 - Desempenho dos classificadores para a condição operacional 1 ($f_a = 50$ Hz)	75
5.3.2 - Desempenho dos classificadores para a condição operacional 2 ($f_a = 60$ Hz)	77
5.3.3 - Desempenho dos classificadores para a condição operacional 3 ($f_a = 70$ Hz)	80
5.3.4 - Desempenho global dos classificadores	83
5.4 - Desempenho dos métodos com análise independente dos índices de monitoramento	83
5.5 - Análise crítica dos resultados	86

CAPÍTULO 6

CONCLUSÕES E SUGESTÕES	90
6.1 - Conclusões	90
6.2 - Sugestões para trabalhos futuros	93
REFERÊNCIAS BIBLIOGRÁFICAS	95

Lista de Figuras

Figura 1.1 - Diagrama de um processo básico de aquisição.	2
Figura 1.2 - Um processo de monitoramento completo.	3
Figura 1.3 - Etapas principais de um processo de monitoramento.	3
Figura 2.1 - Modelo geral para o monitoramento de máquinas.	9
Figura 2.2 - Funções de pertinência.	12
Figura 2.3 - Ilustração de uma árvore de decisão decomposta no espaço de índices de monitoramento.	17
Figura 2.4 - Árvore de decisão com o atributo cabelo formando a raiz.	18
Figura 2.5 - Ilustração de uma árvore de decisão fuzzy.	20
Figura 2.6 - Arquitetura básica de um sistema especialista.	22
Figura 2.7 - Comparação de um neurônio artificial simples com um neurônio real.	29
Figura 2.8 - Algumas funções de ativação: (a) função linear, (b) função rampa, (c) função tangente hiperbólica e (d) função sigmoidal.	31
Figura 2.9 - Arquiteturas básicas de redes: (a) Redes em camadas, (b) Redes recorrentes, (c) Redes reticuladas.	31
Figura 2.10 - Arquitetura de Multicamadas.	32
Figura 3.1 - Sistema básico de classificação de padrões.	37
Figura 3.2 - Processo de aprendizado e classificação.	37
Figura 3.3 - Representação gráfica dos conceitos de atributos, classes e classificação.	39
Figura 3.4 - Representação da classificação por um diagrama.	40
Figura 3.5 - Representação gráfica de elipses de contorno de densidade de probabilidade constante, usando dois atributos.	45
Figura 3.6 - Processo de classificação por função de discriminação linear para duas classes.	47
Figura 3.7 - Representação gráfica para a classificação de três classes, usando dois atributos.	48

Figura 3.8 - Ilustração de um método de distribuição livre utilizado no reconhecimento de padrões.	49
Figura 4.1 - Esquema da interligação dos equipamentos no processo de medição de dados. ..	55
Figura 4.2 - Disposição física do sistema mecânico estudado.	55
Figura 4.3 - Fluxograma do programa de aquisição de dados desenvolvido.	56
Figura 4.4 - A janela Configuração.	57
Figura 4.5 - A janela Aquisição.	58
Figura 4.6 - Pontos de medição utilizados.	59
Figura 4.7 - Desenho esquemático das válvulas das tubulações de sucção e de recalque.	62
Figura 4.8 - Defeito na pá do rotor $\rightarrow d = 3 \text{ mm}$	63
Figura 4.9 - Defeito na pá do rotor $\rightarrow d = 6 \text{ mm}$	63
Figura 4.10 - Comportamento vibratório do sistema após parada e partida do motor.	64
Figura 4.11 - Processo de desmontagem utilizado na bomba.	65
Figura 4.12 - Comportamento vibratório do sistema após desmontagens da bomba.	66
Figura 5.1 - Variação dos índices de monitoramento X_1^v , X_2^v e X_3^v para $f_a = 50 \text{ Hz}$	70
Figura 5.2 - Variação dos índices de monitoramento X_1^h , X_2^h e X_3^h para $f_a = 50 \text{ Hz}$	70
Figura 5.3 - Variação dos índices de monitoramento X_4^v , X_5^v e X_6^v para $f_a = 60 \text{ Hz}$	71
Figura 5.4 - Variação dos índices de monitoramento X_4^h , X_5^h e X_6^h para $f_a = 60 \text{ Hz}$	72
Figura 5.5 - Variação dos índices de monitoramento X_7^v , X_8^v e X_9^v para $f_a = 70 \text{ Hz}$	73
Figura 5.6 - Variação dos índices de monitoramento X_7^h , X_8^h e X_9^h para $f_a = 70 \text{ Hz}$	73
Figura 5.7 - Procedimento de classificação baseado na distância Euclideana.	74
Figura 5.8 - Procedimento de classificação baseado na distância de <i>Mahalanobis</i>	75
Figura 5.9 - Classificação baseada na distância Euclideana utilizando-se dois índices de monitoramento.	89
Figura 5.10 - Classificação baseada na distância de <i>Mahalanobis</i> utilizando-se dois índices de monitoramento.	89

Lista de Tabelas

Tabela 2.1 - As amostras de aprendizagem.	10
Tabela 2.2 - Alguns domínios de trabalho dos sistemas especialistas.	26
Tabela 4.1 - Configuração usada para a aquisição de dados.	59
Tabela 4.2 - Tabela comumente usada para o diagnóstico de conjuntos moto-bomba.	60
Tabela 4.3 - Índices de monitoramento utilizados para cada condição operacional.....	61
Tabela 4.4 - Número de conjuntos de medições realizados para cada classe de funcionamento	66
Tabela 5.1(a) - Índices de sensibilidade obtidos utilizando o classificador de <i>Mahalanobis</i> ($f_a = 50$ Hz - direções vertical e horizontal).	76
Tabela 5.1(b) - Índices de sensibilidade obtidos utilizando o classificador Euclidiano ($f_a = 50$ Hz - direções vertical e horizontal).	76
Tabela 5.2(a) - Índices de sensibilidade obtidos utilizando o classificador de <i>Mahalanobis</i> ($f_a = 50$ Hz - direção vertical).	76
Tabela 5.2(b) - Índices de sensibilidade obtidos utilizando o classificador Euclidiano ($f_a = 50$ Hz - direção vertical).	77
Tabela 5.3(a) - Índices de sensibilidade obtidos utilizando o classificador de <i>Mahalanobis</i> ($f_a = 50$ Hz - direção horizontal).	77
Tabela 5.3(b) - Índices de sensibilidade obtidos utilizando o classificador Euclidiano ($f_a = 50$ Hz - direção horizontal).	77
Tabela 5.4(a) - Índices de sensibilidade obtidos utilizando o classificador de <i>Mahalanobis</i> ($f_a = 60$ Hz - direções vertical e horizontal).	78
Tabela 5.4(b) - Índices de sensibilidade obtidos utilizando o classificador Euclidiano ($f_a = 60$ Hz - direções vertical e horizontal).	78

Tabela 5.5(a) - Índices de sensibilidade obtidos utilizando o classificador de <i>Mahalanobis</i> ($f_a = 60$ Hz - direção vertical).....	79
Tabela 5.5(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideo ($f_a = 60$ Hz - direção vertical).....	79
Tabela 5.6(a) - Índices de sensibilidade obtidos utilizando o classificador de <i>Mahalanobis</i> ($f_a = 60$ Hz - direção horizontal).....	79
Tabela 5.6(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideo ($f_a = 60$ Hz - direção horizontal).....	80
Tabela 5.7(a) - Índices de sensibilidade obtidos utilizando o classificador de <i>Mahalanobis</i> ($f_a = 70$ Hz - direções vertical e horizontal).....	81
Tabela 5.7(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideo ($f_a = 70$ Hz - direções vertical e horizontal).....	81
Tabela 5.8(a) - Índices de sensibilidade obtidos utilizando o classificador de <i>Mahalanobis</i> ($f_a = 70$ Hz - direção vertical).....	81
Tabela 5.8(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideo ($f_a = 70$ Hz - direção vertical).....	82
Tabela 5.9(a) - Índices de sensibilidade obtidos utilizando o classificador de <i>Mahalanobis</i> ($f_a = 70$ Hz - direção horizontal).....	82
Tabela 5.9(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideo ($f_a = 70$ Hz - direção horizontal).....	82
Tabela 5.10 - Comparação dos índices de sensibilidade globais obtidos utilizando o classificador de <i>Mahalanobis</i> e o Euclideo.....	83
Tabela 5.11(a) - Índices de sensibilidade obtidos utilizando o classificador de <i>Mahalanobis</i> ($f_a = 70$ Hz - direções vertical e horizontal - índices de monitoramento X_9^v e X_9^h).....	84
Tabela 5.11(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideo ($f_a = 70$ Hz - direções vertical e horizontal - índices de monitoramento X_9^v e X_9^h).....	84
Tabela 5.12(a) - Índices de sensibilidade obtidos utilizando o classificador de <i>Mahalanobis</i> ($f_a = 70$ Hz - direção vertical - índice de monitoramento X_9^v).....	85
Tabela 5.12(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideo ($f_a = 70$ Hz - direção vertical - índice de monitoramento X_9^v).....	85

Tabela 5.13(a) - Índices de sensibilidade obtidos utilizando o classificador de <i>Mahalanobis</i> ($f_a = 70$ Hz - direção horizontal - índice de monitoramento X_9^h).	85
Tabela 5.13(b) - Índices de sensibilidade obtidos utilizando o classificador Euclidiano ($f_a = 70$ Hz - direção horizontal - índice de monitoramento X_9^h).	86
Tabela 5.14 - Comparação dos índices de sensibilidade globais obtidos utilizando o classificador de <i>Mahalanobis</i> e o Euclidiano (índices de monitoramento relacionados a 4xRPM).	86

Resumo

PORTARI, Marcus Vinícius, *Diagnóstico de Defeitos em Equipamentos utilizando Métodos Estatísticos de Reconhecimento de Padrões*, Campinas: Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas, 1997. 101 p. Dissertação (Mestrado)

Atualmente, a maioria dos métodos de diagnóstico fundamenta-se na comparação e/ou classificação de padrões. Estes métodos têm recebido uma atenção especial nas últimas duas décadas e as teorias recentes de sistemas fuzzy, árvores de decisão, sistemas especialistas e redes neurais artificiais têm sido muito utilizadas nos processos de comparação/classificação de padrões. No entanto, grande parte destes métodos exige um esforço elevado na obtenção de padrões e inclui procedimentos complexos que inviabilizam seu uso no diagnóstico de equipamentos de menor responsabilidade e custo. Este trabalho contribui neste ponto analisando dois métodos de diagnóstico alternativos, um baseado na distância Euclideana e outro na distância de *Mahalanobis*. Estes métodos não exigiram um grande esforço na fase de treinamento e foram facilmente implementados. Para estudá-los, desenvolveu-se um trabalho experimental em um conjunto moto-bomba simulando-se alguns tipos de defeitos, tais como, desbalanceamento, obstruções nas tubulações de sucção e recalque, e modificações geométricas nas pás do rotor. A análise comparativa dos métodos foi realizada através de seus desempenhos na classificação de amostras não utilizadas na fase de treinamento. O desempenho de cada método foi calculado através de um índice de sensibilidade obtido pelo número de amostras corretamente classificadas dentro de um conjunto. O método de diagnóstico baseado na distância de *Mahalanobis* apresentou, em geral, melhores resultados do que o método baseado na distância Euclideana.

Palavras Chave: bombas, bombas centrífugas, defeitos, diagnóstico, falhas, manutenção, manutenção preditiva, monitoramento, reconhecimento de padrões.

Abstract

PORTARI, Marcus Vinícius, *Diagnóstico de Defeitos em Equipamentos utilizando Métodos Estatísticos de Reconhecimento de Padrões*, Campinas: Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas, 1997. 101 p. Dissertação (Mestrado)

Presently, most diagnostic methods are based on pattern matching or pattern classification. In the past two decades, a large number of pattern matching and pattern classification have been developed. Currently, research has been extended to the applications of expert systems, fuzzy systems, decision trees and artificial neural networks. However most of these methods requires high efforts in learning phase. They include complex procedures that are unsuitable for the diagnostic of simple machines. This work analyzes two alternative diagnostic methods, one based on Euclidean distance and another based on Mahalanobis distance. These methods did not require high efforts in learning phase and were easily implemented. An experimental setup involving a centrifugal pump was designed to analyze these methods. Some faults like unbalance, suction and discharge piping obstructions, and geometrical modifications on rotor blades were studied. Data from healthy and fault conditions were used to evaluate the sensitivity of the two methods. In general, the diagnostic method based on Mahalanobis distance presented better results than the method based on Euclidean distance.

Keywords: centrifugal pumps, diagnostic, failures, faults, maintenance, monitoring, pattern recognition, predictive maintenance, pumps.

Capítulo 1

Introdução

Um conceito tem-se tornado muito comum nos meios produtivos e nunca foi tão perseguido quanto neste final de século: *produtividade*. Busca-se produtividade em todas as etapas do processo produtivo: desenvolvimento de projetos, processo de fabricação, gerenciamento do estoque, ou seja, tudo que agrega valor ao produto final.

Baseado nessa tendência, a manutenção surge como uma parte estratégica da cadeia produtiva, onde o seu papel é ser a mais rápida e a mais eficiente possível a fim de não prejudicar essa mesma cadeia. Porém esse papel da manutenção já é praticado há algum tempo em setores críticos como, por exemplo, o setor de geração e distribuição de energia elétrica, o setor de tratamento e distribuição de água, entre outros setores da infra-estrutura nacional.

Na atividade de manutenção de equipamentos, a manutenção preditiva tem sido objeto de estudo de muitos pesquisadores nas últimas décadas. A manutenção preditiva trabalha com o monitoramento regular do comportamento do equipamento, tentando prever assim, o momento da possível falha. Com isso pode-se programar uma interrupção da produção e executar a manutenção no menor tempo possível. Entretanto, é difícil garantir a eliminação de falhas ocasionais e é necessário investir em sensores, painéis de monitoramento, sistema de tratamento de dados e o treinamento do pessoal da manutenção. Os benefícios da implementação de um programa de manutenção preditiva tem que ser equilibrado com seus custos, que são decorrentes dos investimentos acima citados. A manutenção preditiva depende basicamente do processo de monitoramento.

No entanto, é necessário diferenciar um processo de monitoramento de um processo de aquisição de dados, pois existe uma certa confusão a respeito destes termos. No processo de aquisição, os dados coletados através dos sensores não influenciam o processo produtivo. A Figura 1.1 traz o diagrama de um processo de aquisição básico.

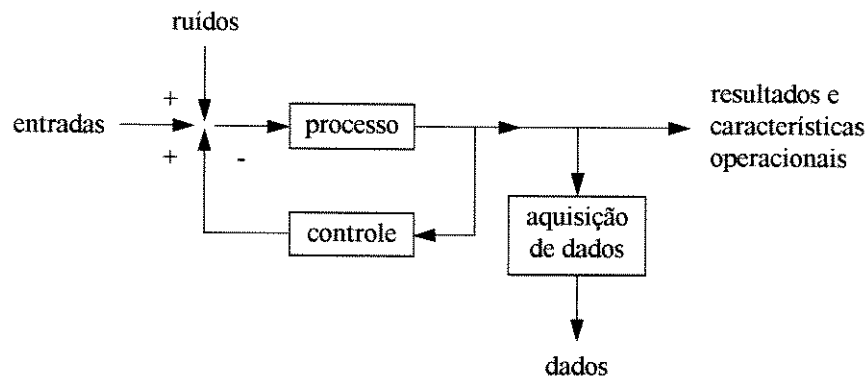


Figura 1.1 - Diagrama de um processo básico de aquisição.

Já o processo de monitoramento deve ser pensado como um fluxo de informações, incluindo uma fonte de informações, a função de supervisão e um usuário da informação (BRAWLEY, 1989). A utilização dos dados coletados é essencial para um monitoramento efetivo. Um processo de monitoramento efetivo deve ser capaz de ordenar e selecionar todas as informações que são relevantes ao processo. A próxima etapa é, com o conhecimento das condições normais de operação, interpretar os dados. A interpretação dos dados fica vinculada a perícia do analista que, dependendo do tipo de interpretação, chega a conclusões das condições de funcionamento da máquina. Após todo este processo, uma decisão para alterar ou não algumas características do processo ou do seu controle deve ser tomada. Esta decisão pode ser a nível de configuração do controle, uma simples manutenção ou até uma complexa mudança no processo. A Figura 1.2 mostra, de uma forma esquemática, um processo de monitoramento completo.

Alternativamente, DU *et al.* (1995a) esquematiza um processo de monitoramento em três etapas: a coleta dos dados, o processamento dos sinais e a análise dos dados já tratados (Figura 1.3). A coleta dos dados refere-se a medição dos sinais provenientes dos transdutores montados nos equipamentos. Os parâmetros medidos devem possuir uma forte correlação com os defeitos mais comuns nos equipamentos. Os parâmetros mais comumente medidos são

vibração, carga, pressão, força, fluxo, temperatura e velocidade. O processamento dos sinais e a análise dos dados são geralmente tratadas como um único processo, recebendo a denominação de métodos de monitoramento.

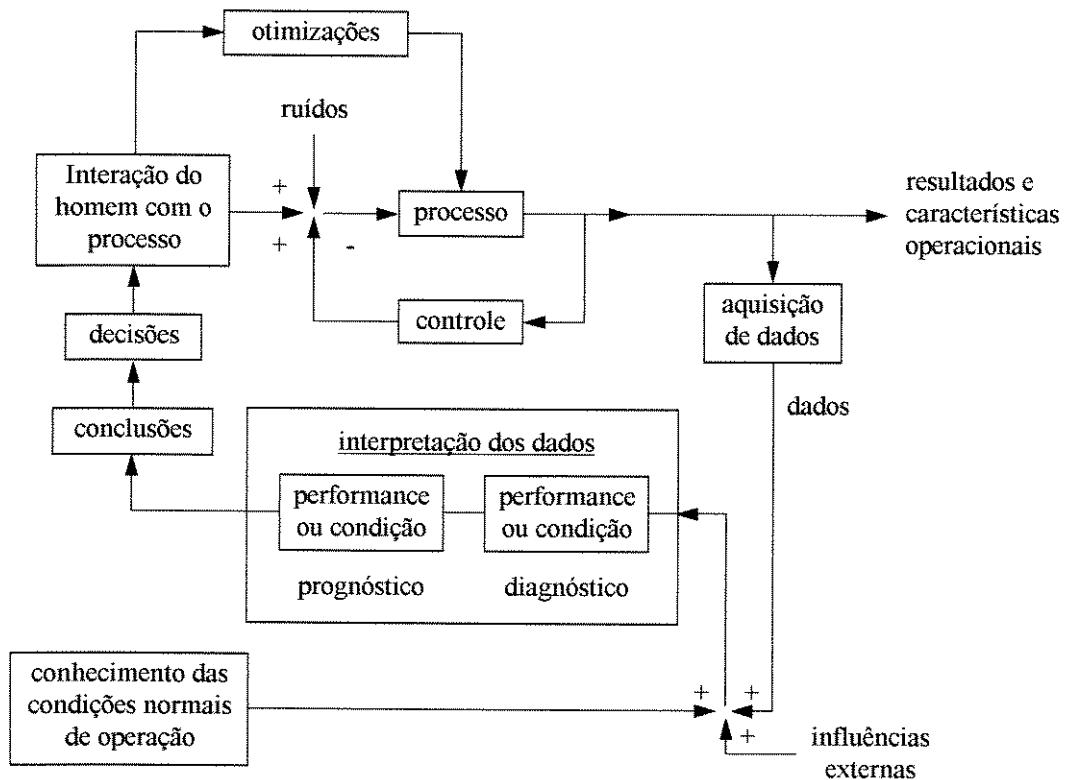


Figura 1.2 - Um processo de monitoramento completo.

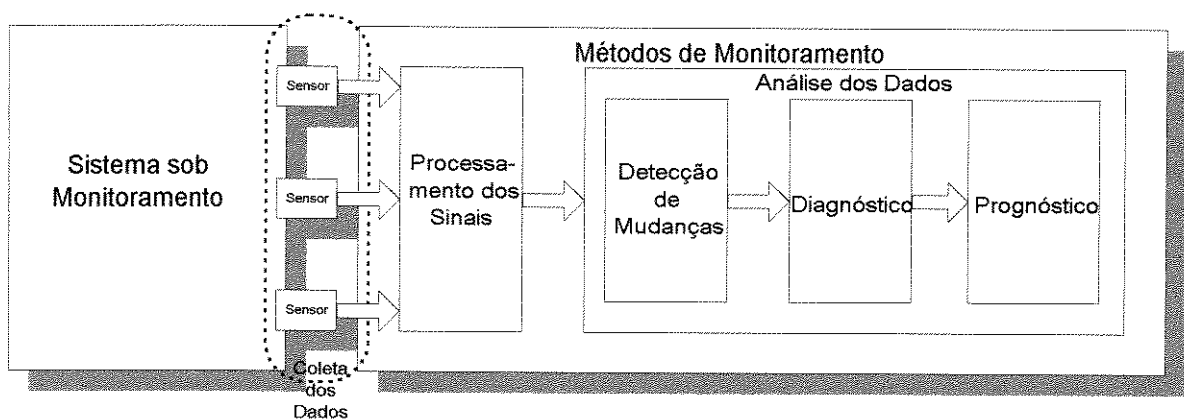


Figura 1.3 - Etapas principais de um processo de monitoramento.

Os métodos de monitoramento podem ser divididos, em geral, em duas categorias: métodos baseados em modelos e métodos baseados em sintomas (DU *et al.*, 1995a). Para

alguns processos, pode-se considerar que os sinais coletados são a saída de um modelo dinâmico. Consequentemente, o processo de monitoramento pode ser conduzido baseado no modelamento do sistema e na avaliação deste modelo. Já os métodos baseados em sintomas utilizam características adequadas dos sinais para identificar as classes de funcionamento do equipamento.

Nos métodos baseados em modelos, o processamento dos sinais tem por objetivo a estimação dos parâmetros do modelo do processo, enquanto que nos métodos baseados em sintomas, o processamento busca extrair, dos sinais coletados, características peculiares a cada tipo de processamento. A estas características e aos parâmetros do modelo do processo, indicadores das classes de funcionamento do sistema monitorado, intitula-se, comumente, índices de monitoramento. Com base na natureza dos sinais, do processo e nos tipos de defeitos a serem detectados, diferentes técnicas podem ser utilizadas, sendo as mais comuns, para os métodos baseados em sintomas, a análise espectral, análise estatística, distribuições tempo-frequência e espectros de ordem superior, enquanto que para os métodos baseados em modelos, os modelos de estado, Auto-Regressivo (AR) e Auto-Regressivo Média Móvel (ARMA).

Processados os dados, a próxima etapa é a análise destes. Nesta nova etapa, avalia-se os índices de monitoramento com o intuito de deduzir os prováveis defeitos, a extensão destes, onde se localizam, como evoluem, possíveis causas e correções. Neste sentido, três fases fazem-se necessárias. A primeira está relacionada à observação de mudanças no sistema. Esta mudança, geralmente, reflete-se na variação dos índices de monitoramento relativamente a valores que representam o funcionamento do equipamento em condições normais, ou seja, um padrão. Para os métodos de monitoramento baseados em modelos, as mudanças do sistema são, geralmente, avaliadas pelas mudanças nos parâmetros do modelo, ou, pela mudança das respostas esperadas, através do cálculo do erro de previsão. Quando os métodos baseados em sintomas utilizam um único índice de monitoramento, geralmente a mudança é detectada quando este ultrapassa um valor limite preestabelecido, enquanto que para casos onde usam-se múltiplos índices, através de técnicas de fusão de sinais.

Uma vez detectada uma mudança é necessária a realização do diagnóstico do defeito, ou seja, a identificação do tipo do defeito, as partes ou componentes comprometidos e, se possível, a extensão do comprometimento. O diagnóstico pode ser realizado de diferentes formas, mas, basicamente, fundamenta-se no conhecimento da relação entre as classes de funcionamento do equipamento e os índices de monitoramento. Para o monitoramento baseado em sintomas, o diagnóstico constitui-se de duas fases: o treinamento e a classificação (LOUKIS *et al.*, 1994). O treinamento é a fase onde busca-se estabelecer a relação entre as classes de funcionamento do equipamento e os índices de monitoramento. A classificação é a fase que identifica os defeitos com base na relação obtida. A terceira e última fase da análise de dados é o prognóstico. O prognóstico tem por objetivo avaliar a evolução do defeito e, conseqüentemente, prever o tempo seguro durante o qual o sistema monitorado pode continuar operando.

O objetivo deste trabalho baseia-se no diagnóstico de defeitos em equipamentos. Os métodos de monitoramento baseados em modelos, também referenciados na literatura como métodos de detecção de falhas baseados em modelos, têm sido utilizados em várias aplicações (QI *et al.*, 1990; DANAI e ULSOY, 1987; CHEN e WU, 1987). Um estudo mais detalhado destes métodos pode ser encontrado em ISERMANN (1993; 1995); MARTIN (1994); BASSEVILLE (1988). No entanto, os métodos baseados em modelos têm 3 limitações significativas (DU *et al.*, 1995a). A primeira limitação é que a maioria dos processos é constituída de sistemas não lineares variantes no tempo. Um exemplo típico é dado por TLUSTY e ISMAIL (1981) em um processo de fabricação, onde a não linearidade é causada pela interação entre a estrutura vibratória e as forças de corte. Ainda que projetos de adaptação possam ser usados, os métodos baseados em modelos são, geralmente, menos sensíveis às mudanças estruturais do processo. A segunda limitação é que as medições são dependentes das classes de funcionamento do processo. É difícil identificar se a mudança nas medições é devido a uma mudança da classe de funcionamento do equipamento ou a uma deterioração do processo. A terceira limitação refere-se ao custo para a obtenção do modelo do equipamento. Geralmente, este custo é elevado e o modelo obtido, às vezes, não representa de maneira confiável o equipamento real. Com base nas limitações acima citadas, os métodos baseados em sintomas foram escolhidos como nosso objeto de estudo.

Atualmente, a maioria dos métodos baseados em sintomas, utilizados para o diagnóstico de equipamentos, fundamenta-se na comparação e/ou classificação de padrões (CHEN *et al.*, 1995). Estes métodos têm recebido uma atenção especial nas últimas duas décadas e teorias recentes têm sido utilizadas nos processos de comparação/classificação de padrões, tais como, sistemas fuzzy, árvores de decisão (*decision trees*), sistemas especialistas e redes neurais artificiais. No entanto, grande parte destes métodos exige um esforço elevado na obtenção de padrões e inclui procedimentos complexos que inviabilizam seu uso no diagnóstico de equipamentos de menor responsabilidade e custo. Além disso, estes métodos demandam de uma carga computacional elevada provocando, consequentemente, um aumento do custo de sua implementação (TROEDER *et al.*, 1987). Este trabalho vem contribuir neste ponto apresentando dois métodos alternativos, um baseado na distância Euclideana e outro na distância de *Mahalanobis*, que exigem um menor esforço na fase de treinamento, são de fácil implementação e com resultados satisfatórios no diagnóstico do equipamento estudado.

O trabalho tem como objetivo principal a análise destes dois métodos de diagnóstico (método baseado na distância Euclideana e método baseado na distância de *Mahalanobis*). Para isso, desenvolveu-se um trabalho experimental em um conjunto moto-bomba em três diferentes condições operacionais, representadas por diferentes rotações do motor elétrico. Simulou-se experimentalmente alguns tipos de defeitos, tais como, desbalanceamento, obstruções nas tubulações de sucção e recalque, e modificações geométricas nas pás do rotor. A análise dos métodos foi realizada através dos seus desempenhos na classificação de amostras não utilizadas na fase de treinamento.

Na busca por métodos de diagnóstico alternativos e de fácil implementação foi inevitável uma revisão bibliográfica para analisar, baseando-se na crescente complexidade e custo de implementação, como as teorias recentes de sistemas fuzzy, árvores de decisão, sistemas especialistas e redes neurais têm sido utilizadas nos processos de comparação/classificação de padrões. No capítulo 2 apresenta-se este estudo enfocando-se a maneira como sistemas fuzzy, árvores de decisão, sistemas especialistas e redes neurais têm sido utilizadas nestes processos.

Já no capítulo 3, antes de apresentar os métodos de diagnóstico que foram implementados, uma revisão bibliográfica é realizada sobre os conceitos básicos da teoria de

classificação de padrões e as principais divisões das metodologias de reconhecimento de padrões (*pattern recognition*). Dentre essas divisões, uma ênfase maior é dada aos métodos estatísticos de reconhecimento de padrões mais utilizados. Finalmente, os métodos de diagnóstico nos quais foram baseados o desenvolvimento deste trabalho são apresentados.

O capítulo 4 descreve a montagem experimental, a utilização do *software* de aquisição de dados desenvolvido, seu fluxograma e o experimento.

No capítulo 5 mostram-se os resultados experimentais e realiza-se uma análise dos métodos de diagnóstico utilizados.

Finalmente, as conclusões obtidas e propostas para trabalhos futuros são apresentadas no capítulo 6.

Capítulo 2

Revisão Bibliográfica

2.1 - Introdução

Para um cardiologista, a linha ziguezague de um eletrocardiograma é uma fonte valiosa de informações. O padrão das amplitudes e frequências dos batimentos cardíacos podem revelar uma série de doenças. Se o diagnóstico não é imediatamente óbvio, o cardiologista pode direcionar seus estudos destacando áreas irregulares para análises posteriores. Analogamente, engenheiros interessados com a “saúde” de seus pacientes mecânicos podem monitorar e diagnosticar problemas usando métodos baseados em análise de vibrações. Como os médicos, os engenheiros sabem que a prevenção é melhor do que os tratamentos caros e arriscados (SARGEANT, 1995).

De acordo com a literatura, um grande número de métodos tem sido desenvolvido para o monitoramento e diagnóstico de equipamentos. A maioria destes métodos é baseada em sintomas e fundamenta-se na comparação e/ou classificação de padrões (CHEN *et al.*, 1995). Este capítulo procura mostrar as idéias básicas que fundamentam a utilização de sistemas fuzzy, árvores de decisão, sistemas especialistas e redes neurais artificiais nos processos de comparação/classificação de padrões. Dentro do contexto deste trabalho, este capítulo tem um caráter exploratório do nível de complexidade necessário para a implementação de certos métodos, salientando a importância de se explorar o potencial de métodos menos sofisticados.

2.2 - Métodos baseados em sintomas

De uma forma geral, os métodos de monitoramento baseados em sintomas utilizam características peculiares dos sinais coletados para identificar as classes de funcionamento do equipamento (DU *et al.*, 1995a). Estas características peculiares, chamadas de índices de monitoramento, podem ser as características do sinal no domínio do tempo e/ou da frequência, tais como, média, variância, curtose e potência em uma banda de frequência específica. Eventos discretos (por exemplo, o sinal lógico “ligado” e “desligado”) e as condições operacionais do processo (por exemplo, a pressão na saída de uma bomba centrífuga) também podem ser usadas como índices de monitoramento.

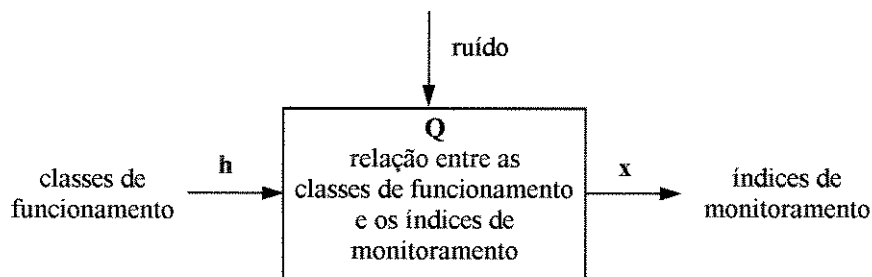


Figura 2.1 - Modelo geral para o monitoramento de máquinas.

Os processos de monitoramento podem ser diferentes de aplicação para aplicação. Entretanto, as formas básicas destes processos são notavelmente similares. Assumindo que as classes de funcionamento do equipamento, $\mathbf{h} = \{h_1, h_2, \dots, h_n\}$, estejam definidas e que os índices de monitoramento, $\mathbf{x} = \{x_1, x_2, \dots, x_m\}$, tenham sido definidos e calculados para cada classe, então existe uma relação $\mathbf{Q}$ entre as classes de funcionamento do equipamento e os índices de monitoramento como mostrado na Figura 2.1. Pode-se representar $\mathbf{Q}$, da forma:

$$\mathbf{Q}: \mathbf{h} \rightarrow \mathbf{x} \quad (2.1)$$

Deve ser notado que esta relação pode ser obtida por diferentes maneiras utilizando-se, por exemplo, uma função analítica, um padrão (*pattern*), um sistema fuzzy, uma árvore de decisão, um sistema especialista ou uma rede neural. Para o monitoramento baseado em sintomas, o diagnóstico constitui-se de duas fases: a aprendizagem e a classificação (LOUKIS *et al.*,

1994). A aprendizagem, também chamada de treinamento, consiste em estabelecer a relação $\mathbf{Q}$ entre $\mathbf{h}$ e $\mathbf{x}$. Existem dois tipos de métodos de aprendizagem: aprendizagem por amostras e aprendizagem por instruções. Uma discussão mais detalhada sobre aprendizagem pode ser encontrada em MICHALSKI *et al.* (1990). Para o monitoramento da maioria das aplicações, a aprendizagem por amostras é geralmente mais efetiva já que instruções precisas são, na maioria das vezes, indisponíveis ou muito limitadas. Um resumo das tecnologias de aprendizagem por amostras pode ser encontrado em GAMS *et al.* (1991). Em geral, as amostras de aprendizagem podem ser organizadas como mostra a Tabela 2.1: onde, m é o número de índices de monitoramento, n é o número de classes de funcionamento do equipamento e N é o número de amostras de aprendizagem disponíveis. Também, $\mathbf{x}_k = [x(1, k), x(2, k), \dots, x(m, k)]$ denota a k -ésima amostra de aprendizagem e $h(\mathbf{x}_k) \in \{h_1, h_2, \dots, h_n\}$ implica que a k -ésima amostra de aprendizagem é obtida de uma das classes conhecidas: $h_1, h_2, \dots, h_n$. Baseada nas amostras de aprendizagem, a relação $\mathbf{Q}$ pode ser estabelecida como discutido nas seções subsequentes.

Amostras	Índices de monitoramento						Classes de funcionamento
	X_1	X_2		X_i		X_m	
$\mathbf{x}_1$	$x(1, 1)$	$x(2, 1)$		$x(i, 1)$		$x(m, 1)$	$h(\mathbf{x}_1) \in (h_1, h_2, \dots, h_n)$
$\mathbf{x}_2$	$x(1, 2)$	$x(2, 2)$		$x(i, 2)$		$x(m, 2)$	$h(\mathbf{x}_2) \in (h_1, h_2, \dots, h_n)$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
$\mathbf{x}_N$	$x(1, N)$	$x(2, N)$		$x(i, N)$		$x(m, N)$	$h(\mathbf{x}_N) \in (h_1, h_2, \dots, h_n)$

Tabela 2.1 - As amostras de aprendizagem.

Na fase de classificação, baseada na relação $\mathbf{Q}$, e na nova amostra $\mathbf{x}$, a classe de funcionamento estimada do equipamento é identificada pela operação inversa da equação (2.1):

$$\mathbf{Q}^{-1}: \mathbf{x} \rightarrow \mathbf{h} \quad (2.2)$$

O diagnóstico pode ser realizado de diferentes formas, mas, basicamente, fundamenta-se no conhecimento da relação entre as classes de funcionamento do equipamento e os índices de monitoramento. Como mencionado anteriormente, estas relações podem ser obtidas utilizando-

se, por exemplo, sistemas fuzzy, árvores de decisão, sistemas especialistas ou redes neurais. Alguns processos de obtenção destas relações a partir destas teorias são apresentados a seguir.

2.2.1 - Sistemas fuzzy

A formulação da lógica fuzzy começou a tomar corpo quando ZADEH (1965) publicou o trabalho denominado “*Fuzzy Sets*”.

A lógica fuzzy é a lógica baseada nos conjuntos fuzzy (GOMIDE e GUDWIN, 1994), diferindo dos sistemas lógicos tradicionais em suas características e detalhes. Ela suporta os modos de raciocínio que são aproximados ao invés de exatos. O raciocínio exato corresponde a um caso limite do aproximado, sendo interpretado como um processo de composição de relações fuzzy.

Na lógica fuzzy, o valor verdade de uma proposição pode ser um subconjunto fuzzy de qualquer conjunto parcialmente ordenado, ao contrário dos sistemas lógicos binários, onde o valor verdade só pode assumir dois valores: verdadeiro (1) ou falso (0). Nos sistemas lógicos multivalores, o valor verdade de uma proposição pode ser um elemento de um conjunto finito num intervalo, ou uma álgebra booleana. Na lógica fuzzy, os valores verdade são expressos linguisticamente, ou seja: verdade, muito verdade, pouco verdade, falso, dentre outros. Cada termo linguístico é interpretado como um subconjunto fuzzy de intervalo unitário.

Em sistemas lógicos binários, os predicados são exatos (por exemplo, par, ímpar, maior que, menor que) e o modificador de predicado mais utilizado é a negação, ao passo que, na lógica fuzzy, os predicados são nebulosos (por exemplo, alto, baixo, grande) e há uma grande variedade de modificadores de predicado (por exemplo, muito, pouco, mais ou menos) que proporcionam uma grande variedade de variáveis linguísticas.

Como mencionado anteriormente, na teoria clássica de conjuntos, um elemento ou pertence a um conjunto ou não. Dado um universo U e um elemento particular $x \in U$, o grau de pertinência $\mu_A(x)$ com respeito a um conjunto $A \subseteq U$ é dado por:

$$\mu_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases} \quad (2.3)$$

A função $\mu_A(x)$, representada pela equação (2.3), é chamada de função característica na teoria clássica de conjuntos. Frequentemente, uma generalização desta idéia é usada, por exemplo, para a manipulação de dados com erros limitados. Todos os números dentro de um erro percentual terão um fator de pertinência 1, tendo os demais, um fator de pertinência 0. Isto pode ser observado na Figura 2.2 (a).

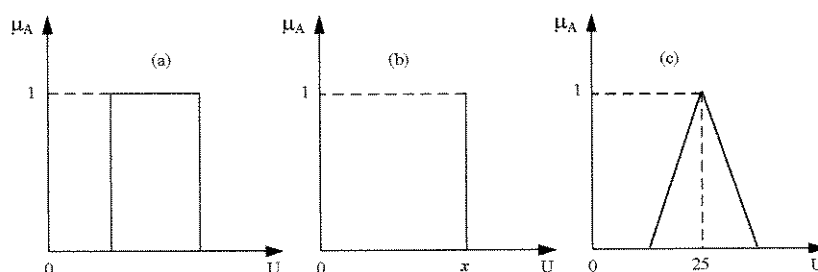


Figura 2.2 - Funções de pertinência.

Para o caso preciso, o fator de pertinência é 1 somente no número exato, sendo 0 para os demais valores, Figura 2.2 (b). Zadeh em 1965 propôs uma caracterização mais ampla, na medida em que sugere que alguns elementos são mais membros de um conjunto que outros. O fator de pertinência ou degrau fuzzy pode então assumir qualquer valor entre 0 e 1, sendo que o valor 0 indica uma completa exclusão e um valor 1, completa pertinência. Esta generalização aumenta o poder de expressão da função característica. Por exemplo, para expressar a idéia de que uma velocidade tem seu valor por volta de 25 m/s, pode-se utilizar a função de pertinência triangular, com pico em 25, para sugerir a idéia de que quanto mais perto o número de 25, mais ele se identifica com o conceito representado (Figura 2.2 (c)).

Formalmente, sendo U uma coleção de objetos denominados genericamente por $\{u\}$, U é chamado de universo de discurso, podendo ser contínuo ou discreto. Um conjunto fuzzy A em um universo de discurso U pode ser descrito por:

$$A = \{x \mid \mu_A(x)\} \quad (2.4)$$

onde, $x \in U$ é o valor de A e $\mu_A(x)$ é o degrau fuzzy. O degrau fuzzy, $\mu_A(x)$, é uma função monótona, $0 \leq \mu_A \leq 1$ e representa a imprecisão do evento.

Uma discussão mais detalhada sobre sistemas fuzzy pode ser encontrada em KLIR e FOLGER (1988) e GOMIDE e GUDWIN (1994).

Baseado em sistemas fuzzy, vários métodos de classificação têm sido desenvolvidos (RUSPINI, 1981), mas somente alguns têm sido utilizados para o monitoramento de equipamentos. Estes incluem o método média-C fuzzy (LI e WU, 1988; WANG *et al.*, 1985), o método da equação de relação fuzzy (TROEDER *et al.*, 1987; PEEKEN *et al.*, 1985), o método da matriz equivalente fuzzy (WANG *et al.*, 1985) e o método da equação linear fuzzy (DU *et al.*, 1992). Dentre eles, o método média-C fuzzy envolve conceitos simples e será explicado a seguir.

O método média-C fuzzy foi proposto pela primeira vez por BEZDEK (1981). Ele utiliza a função custo definida abaixo:

$$J(\mathbf{U}, \mathbf{V}, \mathbf{X}) = \sum_{k=1}^N \sum_{j=1}^n u_{jk}^b \|\mathbf{x}_k - \mathbf{v}_j\|^2 \quad (2.5)$$

onde $\mathbf{V} = \{\mathbf{v}_j\}$ é o agrupamento central fuzzy com $\mathbf{v}_j = [v_{1j}, v_{2j}, \dots, v_{mj}]$, $\mathbf{X} = \{\mathbf{x}_k\}$ são as amostras $\mathbf{x}_k = [x_{1k}, x_{2k}, \dots, x_{mk}]$, $\mathbf{U} = \{u_{jk}\}$ é o degrau fuzzy com $u_{jk} \in [0, 1]$ representando o degrau fuzzy do k -ésimo conjunto de dados com relação a j -ésima classe de funcionamento do equipamento, b é um número inteiro positivo que controla o formato do degrau fuzzy (geralmente, $b = 2$) e $\|\cdot\|$ denota a norma.

Segundo LI e WU (1988), o agrupamento central fuzzy é uma matriz que relaciona as classes de funcionamento com os índices de monitoramento. Na fase de aprendizagem, este agrupamento central é determinado pelo seguinte procedimento de otimização:

$$\min_{(\mathbf{U}, \mathbf{V}) \in \mathbf{M}} \{J(\mathbf{U}, \mathbf{V}, \mathbf{X})\} \quad (2.6)$$

onde

$$\mathbf{M} = \left\{ (u, v) \left| \sum_{j=1}^n u_{jk} = 1, \quad \forall k = 1, 2, \dots, N \right. \right\}$$

BEZDEK (1981) mostra que, quando $b > 1$ e $\mathbf{x}_k \neq \mathbf{v}_j$, a condição necessária para $(\mathbf{U}, \mathbf{V})$ ser uma solução ótima da equação (2.6) é:

$$u_{jk} = \frac{1}{\sum_{\alpha=1}^n \left(\frac{\|\mathbf{x}_k - \mathbf{v}_j\|}{\|\mathbf{x}_k - \mathbf{v}_\alpha\|} \right)^{b/(b-1)}} \quad (2.7)$$

$$\mathbf{v}_{ij} = \frac{\sum_{k=1}^N u_{jk}^b \cdot x_{ik}}{\sum_{k=1}^N u_{jk}^b} \quad (2.8)$$

com $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$; $\alpha = 1, 2, \dots, n$ e $k = 1, 2, \dots, N$. As notações m , n e N representam, respectivamente, o número de índices de monitoramento, o número de classes de funcionamento e o número de amostras utilizadas na fase de treinamento.

Note que as equações (2.7) e (2.8) não podem ser resolvidas analiticamente. Entretanto, existe um algoritmo iterativo que pode fornecer uma solução aproximada para o problema (BEZDEK, 1981; WANG *et al.*, 1985; LI e WU, 1988). Para a aplicação deste algoritmo iterativo, a matriz inicial $\mathbf{U}$ não pode ser definida de uma forma arbitrária. As condições para a definição da matriz inicial $\mathbf{U}$ são encontradas em WANG *et al.* (1985).

Na fase de classificação, baseado na amostra a ser classificada $\mathbf{x}_c$ e no agrupamento central $\mathbf{V}$ obtido na fase de treinamento, calcula-se:

$$u_j = \frac{1}{\sum_{\alpha=1}^n \left(\frac{\|\mathbf{x}_c - \mathbf{v}_j\|}{\|\mathbf{x}_c - \mathbf{v}_\alpha\|} \right)^{b/(b-1)}} \quad (2.9)$$

onde u_j é o degrau fuzzy da amostra a ser classificada sendo atribuído a j -ésima classe de funcionamento do equipamento. Consequentemente, a classe de funcionamento do equipamento é identificada pelo degrau fuzzy máximo:

$$h_j^* = \max\{u_j\} \quad (2.10)$$

Os métodos fuzzy são todos baseados em algum modo de ponderação dos índices de monitoramento. Por exemplo, no método da equação linear fuzzy, os índices de monitoramento são igualmente ponderados em subintervalos igualmente distribuídos. Na prática, entretanto, um índice de monitoramento deve ser particularmente sensível a uma certa classe de funcionamento do equipamento. Isto indica que os índices de monitoramento podem ser usados de uma forma mais efetiva pela decomposição do espaço de índice de monitoramento em várias regiões (DU *et al.*, 1995a). O conceito de decomposição conduz aos métodos de árvore de decisão discutidos abaixo.

2.2.2 - Árvores de decisão

Classificação é o processo de atribuir a uma amostra, a partir de um conjunto de regras, a classe à qual ela pertence. As classes que o procedimento de classificação escolhe podem ser descritas de várias maneiras. Se a definição das classes for feita em termos estruturais, a técnica de árvores de decisão pode ser utilizada (RICH e KNIGHT, 1994).

Em geral, as árvores de decisão podem ser representadas da seguinte forma (QUINLAN, 1987):

$$\begin{array}{c} C_1: D_1 \\ C_2: D_2 \\ \vdots \\ C_M: D_M \end{array} \quad (2.11)$$

onde, C_α , $\alpha = 1, 2, \dots, M$, são condições lógicas mutuamente exclusivas e completas; e D_α é uma regra de classificação que indica indicando uma classe de funcionamento do equipamento (por exemplo, classe de funcionamento = h_i). A condição lógica, C_α , representa uma regra de classificação e normalmente envolve um dos índices de monitoramento. Para um índice de monitoramento contínuo, X_i , a condição lógica pode ser escrita como:

$$C_\alpha := X_i \leq t_{xi} \text{ (ou } X_i > t_{xi}) \quad (2.12)$$

onde t_{xi} é um valor limite. Para um índice de monitoramento discreto, a condição lógica é:

$$C_\alpha := X_i = d \text{ (ou } X_i \in D_\beta) \quad (2.13)$$

onde $d \in \mathcal{D}$ ($\mathcal{D}$ é o conjunto de todos os eventos discretos possíveis), e $\mathcal{D}_\beta$ é um subconjunto de $\mathcal{D}$.

A Figura 2.3 ilustra uma árvore de decisão decomposta no espaço dos índices de monitoramento.

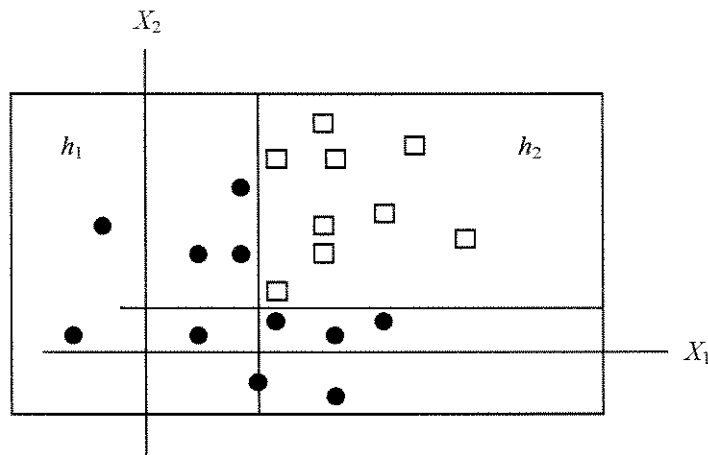


Figura 2.3 - Ilustração de uma árvore de decisão decomposta no espaço de índices de monitoramento.

O processo de construção das árvores de decisão é chamado de aprendizagem ou treinamento. Existem dois tipos de métodos de aprendizagem: aprendizagem por instruções e aprendizagem por amostras. Segundo DU *et al.* (1995a), a aprendizagem por amostras é geralmente preferida, pois elas podem ser obtidas, por exemplo, através da realização de experimentos.

A idéia básica da aprendizagem por amostras é o particionamento. Assumindo que as amostras são dadas como mostrado na Tabela 2.1, uma partição trivial é usar cada amostra de aprendizagem como um galho da árvore de decisão, ou seja, como uma regra de classificação. Isto resultará em uma árvore de decisão extensa que dificilmente irá perceber as variações das novas amostras (DU *et al.*, 1995a). O melhor caminho é agrupar as amostras similares para formar os galhos da árvore de decisão. Obviamente, suponha que se possa encontrar todos os agrupamentos ou partições possíveis, então a árvore de decisão ótima pode ser encontrada. Isto requer uma carga computacional exponencial e pode nem sempre ser possível se existir um grande número de amostras de aprendizagem (GAREY e JOHNSON, 1979). Vários algoritmos foram desenvolvidos para construir árvores de decisão sub-ótimas. Entre estes algoritmos, o algoritmo ID3 (*Iterating Dichotomizer Three*) é o mais utilizado.

O ID3 representa conceitos em forma de árvore (Figura 2.4). Para classificar uma determinada entrada, começamos na raiz da árvore e respondemos a perguntas até alcançar uma folha, onde a classificação é armazenada. O algoritmo ID3 é um algoritmo recursivo, usado na fase de aprendizagem para construir a árvore de decisão. Ele dá preferência às

árvores simples sobre as complexas, com base na teoria de que as árvores simples classificam com maior precisão entradas futuras. O ID3 começa escolhendo um subconjunto aleatório de exemplos de treinamento. Esse subconjunto é chamado de janela. O algoritmo constrói uma árvore de decisão que classifica corretamente todos os exemplos da janela. A árvore é então testada com base nos exemplos de treinamento fora da janela. Se todos os exemplos forem classificados corretamente, o algoritmo pára. Caso contrário, ele acrescenta outros exemplos de treinamento à janela e repete o processo. Evidências empíricas indicam que a estratégia iterativa é mais eficiente do que a análise de todo o conjunto de treinamento de uma só vez (RICH e KNIGHT, 1994).

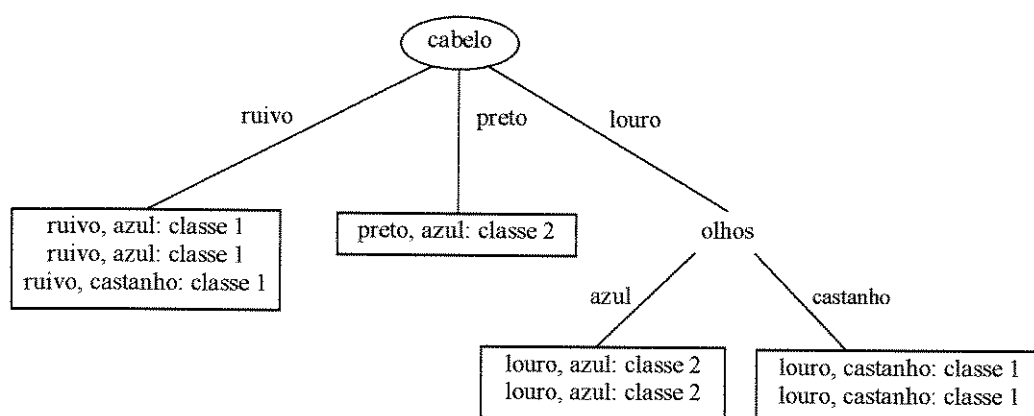


Figura 2.4: Árvore de decisão com o atributo cabelo formando a raiz.

No algoritmo ID3, a construção de um nó significa escolher algum atributo (índice de monitoramento) para ser testado. Em um determinado ponto da árvore, alguns atributos produzem mais informações que outros. Por exemplo, testar o atributo temperatura é inútil se o comportamento vibratório do equipamento não ajudar a classificá-lo corretamente. O ideal seria que um atributo separasse as amostras de treinamento em classes que tenham uma marcação comum. Nesse caso, a ramificação termina e os nós-folha são marcados.

Há muitas variações neste algoritmo básico. Por exemplo, quando acrescenta-se um teste que tenha mais de duas ramificações, é possível que uma ramificação não tenha amostras de treinamento correspondentes. Nesse caso, pode-se deixar o nó sem marcação, ou pode-se tentar adivinhar com base nas propriedades estatísticas do conjunto de amostras que está sendo testado naquele ponto da árvore (RICH e KNIGHT, 1994). Os ruídos são um outro problema. Uma maneira de lidar com eles é evitar criar novas ramificações se o ganho de informação for

pequeno. Em outras palavras, não se quer complicar a árvore para que ela leve em conta amostras inesperadas. Uma outra fonte de incerteza é que os valores dos atributos podem ser desconhecidos. Por exemplo, o registro médico de um paciente pode estar incompleto. Uma solução é adivinhar a ramificação correta a ser seguida. O algoritmo completo e detalhado pode ser encontrado em DU *et al.* (1995a), QUINLAN (1990); QUINLAN (1986).

Segundo QUINLAN (1990), as vantagens do uso do algoritmo ID3 no processo de construção das árvores de decisão são:

- o método converge rapidamente; geralmente 4 iterações são requeridas para encontrar a árvore de decisão correta.
- o processo não é muito sensível aos parâmetros, tais como, o tamanho do conjunto das amostras de treinamento.
- o tempo para obter uma árvore de decisão confiável aumenta linearmente com a dificuldade do problema.

Estas particularidades, principalmente a última, têm permitido ao algoritmo ID3 encontrar árvores de decisão confiáveis para vários problemas de classificação.

Uma desvantagem do algoritmo ID3 é que pode ser difícil para os seres humanos compreenderem árvores de decisão grandes e complexas e, por isso, o sistema de decisão poderá se complicar, tentando explicar os motivos de suas classificações (RICH e KNIGHT, 1994).

As regras de classificação nas árvores de decisão são caracterizadas por declarações de condição *sharp-cut* tal como “se $x_i > t_{xi}$ então h_j ”. Com o fim de minimizar o efeito deste *sharp-cut* na fase de classificação, pode-se usar a árvore de decisão de probabilidades (CARTER *et al.*, 1988) ou a árvore de decisão fuzzy (LI *et al.*, 1992). Na árvore de decisão fuzzy, os nós são determinados por um degrau fuzzy e/ou por um agrupamento central fuzzy. Na fase de aprendizagem, o degrau fuzzy pode ser determinado usando o método da equação linear fuzzy. O agrupamento central pode ser determinado usando o método média-C fuzzy, que por sua vez determina o degrau fuzzy. O degrau fuzzy e o agrupamento central representam um *soft-cut* como mostrado na Figura 2.5. Na fase de classificação, toda a árvore

de decisão fuzzy tem que ser pesquisada para encontrar o degrau fuzzy de cada nova amostra para todas as classes de funcionamento do equipamento. Consequentemente, a partir de uma nova amostra, a classe de funcionamento do equipamento pode ser estimada com base na equação (2.10).

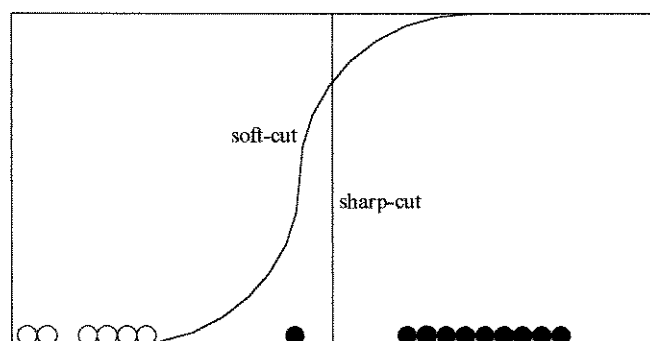


Figura 2.5 - Ilustração de uma árvore de decisão fuzzy.

2.2.3 - Sistemas especialistas

A Inteligência Artificial (IA) e o seu segmento de aplicação denominado Sistema Especialista constituem-se em instrumentos adequados para o tratamento de problemas que são decorrentes da experiência e das preferências pessoais do especialista (TEIXEIRA JR. *et al.*, 1989).

Segundo SCHILDT (1989), sistemas especialistas são programas que imitam o comportamento de especialistas humanos. Os sistemas especialistas usam informações que o usuário fornece para emitir uma opinião sobre determinado assunto.

Segundo PASSOS (1987), sistema especialista é um programa de computador destinado a solucionar problemas em um campo específico do conhecimento, que tem para isso uma base de conhecimento desse domínio restrito. Usa um raciocínio inferencial para executar tarefas e tem desempenho comparável aos especialistas humanos. Na prática, uma das mais importantes características de um sistema especialista é a capacidade de explanação. Eles surgiram porque as estratégias de resolução de propósitos gerais da IA eram muito ineficientes na solução de problemas complexos. Na década de 70, ficou claro que para os pesquisadores de IA conseguirem que seus sistemas resolvessem satisfatoriamente problemas reais, era necessário

incorporar-lhes grandes quantidades de conhecimentos sobre o problema. Este fato demandou a criação do campo da Engenharia do Conhecimento que procura formas de usar conhecimentos de especialistas na solução de problemas complexos.

A verdadeira definição de sistema especialista (BARBORAK *et al.*, 1991) vive constantemente em debate, mas pode ser genericamente descrito como um sistema projetado para simular o conhecimento e raciocínio de um especialista humano fazendo com que o conhecimento fique disponível para outras pessoas de forma útil. Os termos Sistemas de Conhecimento e Sistemas Baseados em Conhecimento são frequentemente usados como sinônimos de sistemas especialistas.

Os benefícios que um sistema especialista pode trazer são basicamente (SMART, 1987):

- sintetizar e conservar conhecimento coletivo;
- treinamento;
- determinação de especificações funcionais;
- assegurar consistência e perfeição em decisões repetitivas;
- execução de tarefas por leigos;
- realizar as tomadas de decisão em profundidade;
- implementação de sistemas fáceis de serem atualizados.

Os sistemas especialistas utilizam-se de fatos, informações específicas sobre a área de interesse, e heurísticas, que são representadas através de regras práticas que o especialista humano utiliza para chegar a conclusões relevantes em sua área de atuação.

A estrutura de um sistema especialista caracteriza-se pela separação entre as estratégias de solução do problema e os algoritmos de controle do processamento da informação. Uma arquitetura típica é composta dos seguintes elementos (KOWALSKI *et al.*, 1993):

- base de conhecimento: onde reside o conjunto de informações que codificam a estratégia de solução e representação do conhecimento.
- base de dados: contém as características e parâmetros relativos ao problema a ser resolvido.

No encadeamento para a frente as regras são exaustivamente executadas baseadas nos dados conhecidos e suas principais características são (TRINZIC, 1994):

- é orientado a dados;
- analisa as regras a partir de sua premissa;
- o processo é continuado até que todas as possíveis ocorrências sejam inferidas.

O encadeamento para a frente é disparado para a frente quando novos dados tornam-se conhecidos, ou seja, novos valores são atribuídos aos dados. Durante o processo de encadeamento para a frente, a máquina de inferência processa as regras na medida em que suas premissas tornam-se verdadeiras. A execução das regras, por sua vez, inferem outros novos dados. E assim o processo de encadeamento continua até que todo o possível conhecimento tenha sido inferido e todas as possíveis conclusões sejam alcançadas.

A execução do encadeamento para a frente dá-se a partir do ciclo:

1. entrada de novos dados;
2. disparo de regras quando suas premissas são verdadeiras;
3. ações das regras atribuem novos valores aos dados;
4. execute o passo (2).

O mecanismo de inferência seleciona para a resolução todas as regras da base de conhecimento cujas premissas são verdadeiras e o processo continua enquanto houver regras que possam ser disparadas. Dessa forma, ao final do processamento, todos os dados que poderiam ser inferidos em função dos dados conhecidos estarão presentes na base de dados.

No encadeamento para trás a máquina de inferência tem um objetivo previamente estabelecido a ser determinado, e trabalha para trás através das regras para determinar valores que possam resolver o objetivo. As principais características do encadeamento para trás são (TRINZIC, 1994):

- é orientado ao objetivo;
- analisa as regras a partir de sua ação;
- somente as informações necessárias são acessadas da base de conhecimento.

O ciclo de execução da máquina de inferência no encadeamento para trás é:

1. identifica o objetivo;
2. procura pela regra que resolva o objetivo na sua cláusula de ação;
3. identifica objetivos intermediários nas premissas das regras;
4. para cada objetivo, um resultado é obtido: seja o objetivo resolvido ou não. Faça perguntas e/ou acione regras adicionais (passo 2) necessárias para tentar resolver cada objetivo.

Em muitos casos a escolha de qual direção o sistema deverá executar é feita reproduzindo a maneira utilizada por uma pessoa para resolver o problema.

Um sistema especialista tem como características:

- flexibilidade: um sistema especialista pode crescer ativamente pela adição de mais conhecimento na forma de regras. Além disso, estas regras podem ser adicionadas pelo especialista sem nenhum conhecimento de computação.
- adaptabilidade: como as questões podem ser formuladas pela mudança do sistema especialista, novas regras devem ser adicionadas. Com esta habilidade para adaptar novas situações, a vida do sistema especialista pode ser estendida indefinidamente, sendo que os custos de desenvolvimento após o estágio inicial podem ser reduzidos.
- desempenho e eficiência: bom desempenho e eficiência na busca das soluções.
- explicação do raciocínio: a fim de tornar o programa auto explicativo, no que diz respeito ao raciocínio empregado, são respondidas algumas perguntas elucidativas. Na implementação do sistema, perguntas fundamentais podem ser feitas, tais como: “Como a máquina de inferência chegou a determinada conclusão?”; “Por que é necessário fornecermos novos dados?”; “Por que não foi usada determinada regra?”.
- interação com o usuário: os sistemas tem boa interação com o usuário, pois têm características que auxiliam o usuário amplamente. Alguns, por exemplo, questionam o usuário, a fim de chegar à meta mais rapidamente e com um grau de certeza maior.
- incerteza: muitos sistemas especialistas podem trabalhar com a incerteza. Pode existir uma resposta com uma determinada chance percentual de ser verdadeira, então eles expressam implicitamente o grau de confiança que se tem no fato de uma resposta ser

verdadeira. Para isso, cada variável de um sistema especialista deve possuir associada a ela um fator de certeza.

O processo de construção de um sistema especialista é frequentemente chamado de Engenharia do Conhecimento (*Knowledge Engineering*). Tipicamente ele envolve uma forma especial de interação entre o construtor do sistema especialista, chamado Engenheiro do Conhecimento, e um ou mais especialistas humanos de alguma área.

O Engenheiro do Conhecimento obtém dos especialistas humanos seus procedimentos, estratégias e regras práticas para a solução dos problemas construindo este conhecimento num sistema especialista. O Engenheiro do Conhecimento pratica a arte de buscar os princípios e ferramentas que requerem o conhecimento de especialistas para suas soluções.

As questões técnicas na aquisição desse conhecimento, na representação, e na utilização adequada para a construção e explicação das linhas de raciocínio são problemas importantes no projeto de Sistemas Baseados em Conhecimento.

Os sistemas especialistas têm sido construídos para resolver diferentes tipos de problemas, mas suas atividades básicas podem ser agrupadas em categorias. Segundo WATERMAN (1986), estas categorias são: interpretação, previsão ou prognóstico, diagnose, projeto, planejamento, monitoramento, depuramento, reparo, instrução e controle.

Como a maioria dos sistemas especialistas realizam mais do que uma atividade (por exemplo, a diagnose frequentemente ocorre com depuramento, monitoramento com controle e planejamento com projeto), é mais conveniente agrupá-los pelos tipos de problemas que eles resolvem. A Tabela 2.2 (WATERMAN, 1986) mostra alguns dos domínios em que os sistemas especialistas estão trabalhando. Destas áreas, a área médica é a mais popular.

Agricultura	Fabricação
Química	Matemática
Sistemas de Computação	Medicina
Eletrônica	Meteorologia
Engenharia	Ciência Militar
Geologia	Física
Administração de Informação	Controle de Processos
Advocacia	Tecnologia Espacial

Tabela 2.2 - Alguns domínios de trabalho dos sistemas especialistas.

O sistema especialista MYCIN, desenvolvido em meados dos anos 70, foi um dos primeiros a ser utilizado com sucesso (SHORTLIFFE, 1976). Este sistema diagnostica doenças infecciosas acessando suas regras através de encadeamentos para trás. Desde então, centenas de sistemas especialistas têm sido desenvolvidos para várias aplicações. Para o monitoramento e diagnóstico de máquinas, o método de árvore de decisão discutido na seção anterior é recomendado com uma combinação de outras informações e conhecimentos incluindo regras empíricas ou semi-empíricas, simulações computacionais e testes experimentais. Por exemplo, regras do tipo “se $x_1 \geq T$ então verificar sensor 1” podem ser usadas para verificar se o sensor está funcionando satisfatoriamente. Alguns pacotes comerciais (*Ex-Tran*, *RuleMaster* e *1st-Class*) expressam as regras na forma de árvores de decisão.

Dentro da área de engenharia, temos, por exemplo, os seguintes sistemas especialistas:

- DELTA (Diesel-Electric Locomotive Troubleshooting Aid) - Este sistema especialista ajuda o pessoal da manutenção a identificar e corrigir mau funcionamento em locomotivas aplicando estratégias de diagnóstico para a manutenção de locomotivas (MARCUS, 1983). O sistema orienta o usuário através de um procedimento de reparo, apresentando desenhos das partes e subsistemas, sequências de reparo na forma de filmes e instruções específicas de reparo uma vez que o mau funcionamento é identificado. DELTA é um sistema com a representação do conhecimento baseado em regras desenvolvido em linguagem LISP. DELTA acessa suas regras através de encadeamentos para a frente e para trás e usa fatores de certeza para lidar com premissas de regras incertas.

- **REACTOR** - Este sistema especialista auxilia os operadores na detecção e localização de defeitos e/ou falhas em reatores através das medições feitas pelos instrumentos de monitoramento (por exemplo, fluxo da água de alimentação e nível de radiação), procurando por desvios nas condições normais de operação (NELSON, 1982). Quando o sistema detecta um desvio, ele estima a situação e recomenda uma ação apropriada. Para isso, utiliza de conhecimentos sobre a configuração do reator e as relações funcionais de componentes aliados com os conhecimentos sobre o comportamento esperado do reator sobre condições anormais conhecidas. REACTOR foi implementado em linguagem LISP usando métodos baseados em regras com encadeamentos para a frente e para trás.
- **STEAMER** - Este sistema especialista instrui estudantes de Engenharia Naval na operação da planta de propulsão de vapor da fragata classe 1078 (HOLLAN, 1984). O sistema monitora o estudante reconhecendo ações apropriadas e corrigindo as inapropriadas. O sistema trabalha ligando uma simulação matemática da planta de propulsão com um sofisticado programa de interface gráfica que mostra diagramas dos subsistemas da planta. O estudante pode manipular componentes, tais como válvulas e bombas, e observar os efeitos nos parâmetros da planta, tais como mudanças nas pressões, temperaturas e vazões. STEAMER foi implementado em ZETALISP.
- **NPPC (Nuclear Power Plant Consultant)** - Este sistema especialista ajuda operadores de usinas nucleares a determinar a causa de alguns eventos anormais pela aplicação de regras em conjunção com o modelo de operação da usina (UNDERWOOD, 1982). O sistema usa um modelo primário do sistema de refrigeração, incluindo bombas, reator, geração de vapor e sistema de refrigeração do núcleo de emergência, para diagnosticar a causa de anormalidades ou acidentes, e então sugere procedimentos para a correção do problema.
- **SPERIL-I** - Este sistema especialista realiza uma avaliação estrutural de estruturas que estão sujeitas a terremotos (ISHIZUKA, 1981). O sistema determina o quanto as estruturas estão danificadas a partir dos dados de inspeção e das medições de aceleração. A perícia deste sistema consiste em conhecimentos coletados de engenheiros civis experientes e inclui relações entre os fatores tais como amortecimento estrutural e rigidez. Estes conhecimentos são representados como regras acessadas através de encadeamento para a frente. O sistema usa fatores de certeza combinados com Lógica Fuzzy para calcular o quanto a estrutura está danificada. SPERIL-I foi escrito em linguagem C.

2.2.4 - Redes neurais artificiais

Ao contrário dos sistemas especialistas, que processam expressões simbólicas de uma forma sequencial, as redes neurais processam informações em paralelo (LIPPMANN, 1987).

Segundo PECK e BURROWS (1994), uma rede neural é um sistema computacional constituído por elementos de processamento (neurônios) que estão conectados formando uma rede de neurônios. Cada neurônio é uma unidade de processamento matemático simples sem inteligência própria. Entretanto, quando vários neurônios são interconectados, uma estrutura extremamente complexa é produzida, e uma vez treinada pode fornecer a funcionalidade necessária para um processamento inteligente.

A origem das redes neurais decorreu do contraste existente entre o processamento e a estrutura de computadores e o cérebro humano. Deste modo vem se buscando imitar o funcionamento do cérebro a partir do conhecimento de suas estruturas. O cérebro humano é o mais notável e o mais estudado dos sistemas neuro-biológicos e é constituído por cerca de 100 bilhões de neurônios que se comunicam através de uma rede de conexões de axônios e sinapses. Os neurônios possuem um grande número de vias de entrada (os dendritos), um corpo celular (soma) e uma única via de saída (o axônio), que pode estar conectada a inúmeras outras entradas de diferentes neurônios. Um modelo da estrutura e do funcionamento de um neurônio foi primeiramente apresentado por McCULLOCH e PITTS (1943), e deu origem às pesquisas sobre redes neurais artificiais.

A Figura 2.7 mostra um modelo de neurônio artificial em comparação à um neurônio real. Neste modelo simplificado, proposto por ROSENBLATT (1959), a entrada consiste em um conjunto de valores (x_i) que suprem o neurônio artificial com informações. Os pesos (w_{ij}) formam um conjunto de valores numéricos diretamente responsáveis pela ativação ou não do neurônio artificial, pois ditam o quanto de influência uma entrada tem sobre o elemento de processamento. A função de agregação corresponde à regra de propagação da rede, ou seja, como os pesos se combinam às entradas. A função de ativação é uma função matemática que determina o novo valor do elemento de processamento a partir da interpretação do resultado da função de agregação, definindo assim o valor de saída do neurônio artificial.

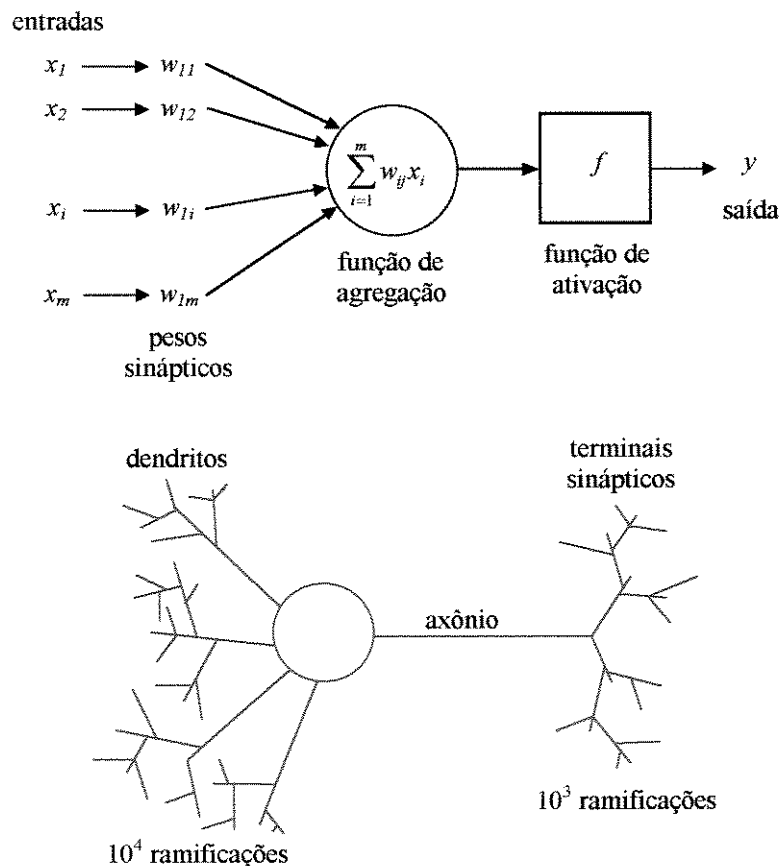


Figura 2.7 - Comparação de um neurônio artificial simples com um neurônio real.

Uma das principais características das redes neurais artificiais é sua capacidade de aprender. A habilidade de armazenar informações e de ser inteligente de uma rede neural ocorre devido à sua concepção paralela, distribuída e interativa. O aprendizado de uma rede artificial se dá através da alteração dos pesos entre os elementos de processamento e é um processo mais direto quando comparado ao aprendizado humano (LIPPMANN, 1987).

Para atingir seus objetivos uma rede neural deve se auto-arrumar em termos de possuir as conexões adequadas a partir das Regras de Aprendizado. O neurônio artificial é um elemento adaptativo cujos pesos são modificados dependendo dos valores de entrada recebidos e das saídas requeridas (JONES e HOSKINS, 1987).

Um conjunto bem definido de regras para a minimização do erro, ou obtenção dos ajustes dos pesos, é denominado algoritmo de aprendizagem e a maneira pela qual o ambiente influencia a rede em seu aprendizado define o paradigma de aprendizagem para o sistema

completo. Segundo TURBAN (1992), como exemplos de algoritmos de aprendizagem, têm-se a Lei de Hebb, algoritmos competitivos e algoritmo *backpropagation*. Como exemplos de paradigmas de aprendizagem destacam-se (KARNA e BREEN, 1989):

- aprendizado supervisionado: pares de padrões de entrada/saída relacionados são apresentados à rede, que então utiliza um algoritmo de modificação dos pesos das conexões para aprender o mapeamento desejado entre estes padrões.
- aprendizado de reforço: a rede não fornece a saída correta mas indica se a saída produzida é boa ou não.
- aprendizado não supervisionado: são utilizados algoritmos competitivos no qual a rede desenvolve suas próprias regras de classificação a partir de exemplos de entradas.

Estando a rede devidamente treinada por uma regra de aprendizagem adequada, pode-se apresentar a ela entradas variadas que serão então analisadas. Este treinamento tem a função de distribuir os pesos nas camadas da rede em resposta a exemplos e à experiência adquirida.

No processamento de uma rede neural, a regra de propagação é na maioria dos casos uma função somatório do tipo:

$$NET_j = \sum_{i=1}^m w_{ij} x_i \quad (2.14)$$

onde m é o número de entradas do elemento de processamento.

As funções de ativação são elaboradas de modo que seja possível a obtenção da saída desejada na aplicação de entradas, por exemplo, no caso onde deseja-se saída 0 ou 1 para quaisquer valores de NET. A Figura 2.8 ilustra graficamente quatro funções de ativação diferentes, são elas: a função linear, a função rampa, a função tangente hiperbólica e a função sigmoidal.

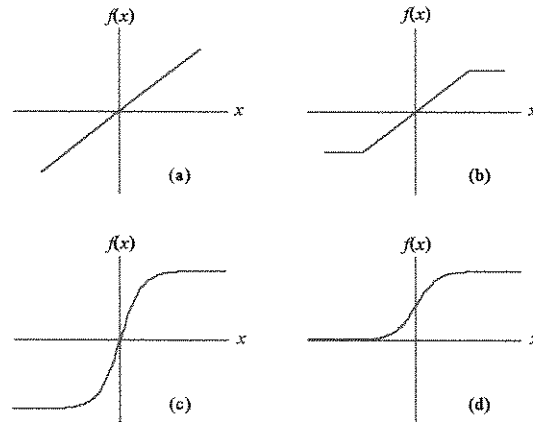


Figura 2.8 - Algumas funções de ativação: (a) função linear, (b) função rampa, (c) função tangente hiperbólica e (d) função sigmoïdal.

Uma rede neural é composta por um conjunto de neurônios artificiais agrupados em camadas. De modo análogo às redes biológicas, uma rede artificial pode ser organizada de diferentes maneiras, isto é, os neurônios podem ser interconectados de diferentes modos, mas há três arquiteturas básicas como mostra a Figura 2.9 (TURBAN, 1992).

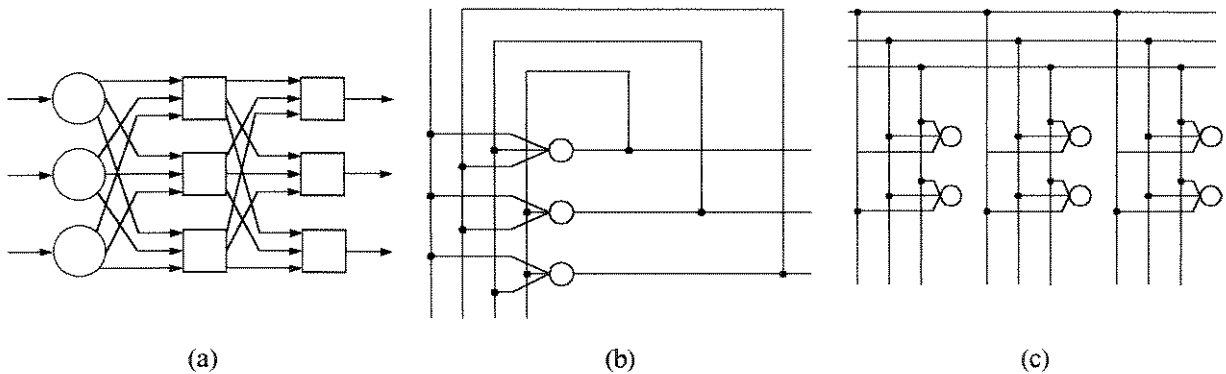


Figura 2.9 - Arquiteturas básicas de redes: (a) Redes em camadas, (b) Redes recorrentes, (c) Redes reticuladas.

Os algoritmos de aprendizagem descrevem métodos para mudar o comportamento individual dos neurônios almejando um comportamento desejado para a rede como um todo. O par formado pela arquitetura Multicamadas e pelo algoritmo de aprendizagem *backpropagation* constitui-se em um dos pares arquitetura-algoritmo de aprendizagem mais utilizados na área de monitoramento e diagnóstico de máquinas.

A Figura 2.10 ilustra a arquitetura de multicamadas onde o algoritmo *backpropagation* pode ser aplicado. O padrão de entrada é colocado na camada de entrada para estimular a

rede. Os neurônios na camada de entrada geralmente não realizam nenhuma operação sobre os padrões de entrada e podem ser vistos como um *buffer*. O número de neurônios nesta camada é igual ao tamanho do padrão de entrada. A segunda e demais camadas são compostas de neurônios do tipo *McCulloch-Pitts*, por exemplo. A última camada produz a saída da rede, assim ela é também conhecida como camada de saída. Existe uma outra camada onde a saída desejada é colocada. Esta camada simplesmente fornece a saída desejada à camada de saída a fim de ser computado o erro ocorrido. Estes erros são retropropagados da camada de saída em direção a camada de entrada para ajustar os pesos das conexões nas camadas anteriores. Daí a origem do termo *backpropagation*. As camadas entre a camada de entrada e a camada de saída são chamadas de camadas escondidas, *hidden layers*, porque a saída de seus neurônios não é vista ou acessada por procedimentos externos. Os sinais propagam-se unidirecionalmente da camada de entrada para a camada de saída através das camadas escondidas. Consequentemente, este tipo de rede é do tipo *feedforward*.

Os neurônios destas multicamadas têm uma ativação não linear na saída, isto é, uma função de transferência não linear. Isto habilita o seu uso em aplicações que exigem um mapeamento padrão, isto é, dada uma entrada padrão, a rede produz uma saída associada. Uma vez aprendido o mapeamento padrão a partir dos conjuntos de treinamento entrada-saída, para qualquer nova entrada a rede será capaz de produzir uma saída padrão baseada no conhecimento conseguido a partir dos conjuntos de treinamento entrada-saída.

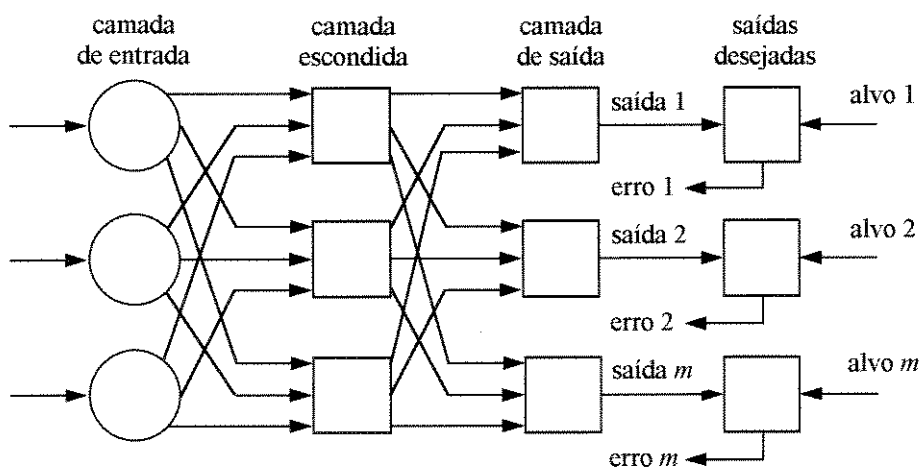


Figura 2.10 - Arquitetura de Multicamadas.

O algoritmo *backpropagation* é um método de aprendizado supervisionado, gerando um controlador de erro entre o que se pretende como resposta e o que se obtém durante o processamento. Este algoritmo não garante chegar sempre ao erro global mínimo de padrões linearmente dependentes. Contudo, na grande maioria das vezes, ele consegue realizar a associação desejada. Este problema é referenciado como o problema do erro de mínimo local.

O algoritmo *backpropagation* e a arquitetura de multicamadas são, sem dúvida, o paradigma de redes neurais mais utilizado nas pesquisas em redes neurais. O algoritmo *backpropagation* tem sido muito utilizado para o monitoramento e diagnóstico de equipamentos (LIAO e CHEN, 1994; PHAM e BAYRO-CORROCHANO, 1995; KIM *et al.*, 1991; MARZI e MARTIN, 1991; MARTIN e MARZI, 1992). O algoritmo *backpropagation* pode ser encontrado em várias referências, como por exemplo, LIPPMANN (1987); LETA e SOARES (1996).

O algoritmo *backpropagation* pode ser aplicado em uma arquitetura com qualquer número de camadas. Na fase de treinamento, uma modificação bastante conhecida é a inclusão da constante de *momento* que varia de 0 a 1. Segundo PHAM e BAYRO-CORROCHANO (1995), esta modificação faz com que o algoritmo tenha o tempo de convergência mais rápido, pois não modifica os pesos de uma forma brusca. Na grande maioria dos casos, o algoritmo *backpropagation* é utilizado com a presença de momento, bem como com outras técnicas que aceleram a velocidade da fase de aprendizagem (SILVA e ALMEIDA, 1990; MOLLER, 1993).

Existem várias outras classes de modelos de redes neurais artificiais, cada uma com suas respectivas arquiteturas e algoritmos de aprendizagem, tais como, a teoria da adaptação ressonante (ART - *Adaptive Resonance Theory*) proposta por GROSSBERG (1976), ART1 (CARPENTER e GROSSBERG, 1987), ART2 (TANSEL e McLAUGHLIN, 1993), ART3 (CARPENTER e GROSSBERG, 1990), redes RAM-based (ALEKSANDER e MANDANI, 1968) e o modelo de Kohonen (redes auto-organizáveis de Kohonen) (KOHONEN *et al.*, 1984; KOHONEN, 1982; TANAKA *et al.*, 1995).

As principais atrações oferecidas pela abordagem de redes neurais para resolver problemas computacionais diversos são as seguintes:

- *Aprendizagem por Exemplos.* Um sistema de computação que utiliza uma aprendizagem por exemplos é desejado quando a complexidade computacional do problema a ser resolvido é alta, no sentido de que o problema não possui todas as suas variáveis conhecidas.
- *Independência do Problema.* Uma rede neural antes da aprendizagem não possui nenhum conhecimento sobre os problemas que se pretende resolver, assim uma mesma rede neural pode ser aplicada a problemas diferentes sem necessidade de qualquer modificação de sua estrutura.
- *Obtenção de Resultados Desejados.* Caso uma rede neural não esteja fornecendo resultados desejados aceitáveis, a sua arquitetura pode ser modificada em busca de otimizações.
- *Processamento Distribuído, Paralelo e Local.* A própria arquitetura da rede e a natureza dos neurônios propiciam um processamento distribuído, paralelo e local. O paralelismo se dá nos níveis de módulos de redes neurais, camadas, neurônios e conexões.
- *Implementação Realística.* Diversos modelos de redes neurais têm sido realizados em hardware, a fim de otimizar as implementações e aumentar a sua faixa de utilização.

As principais desvantagens da utilização de redes neurais são: primeiro, a falta e a dificuldade de um formalismo na especificação e na análise de modelos de redes neurais, uma vez que, para se compreender os mecanismos fundamentais da rede é necessário realizar simulações que, na maioria dos casos, são tarefas árduas e distantes da realidade do modelo; segundo, o estado da arte das implementações ainda limita certas aplicações, isto é, certos problemas exigem poderes computacionais específicos das implementações que ainda não foram atingidos.

Capítulo 3

Métodos Estatísticos

3.1 - Introdução

Mostrou-se no capítulo anterior como alguns métodos de diagnóstico, fundamentados em teorias recentes, têm sido utilizados nos processos de classificação de padrões. Estes métodos exigem um esforço elevado na fase de treinamento e incluem procedimentos complexos que inviabilizam seu uso no diagnóstico de equipamentos de menor responsabilidade e custo. Além disso, estes métodos demandam uma carga computacional elevada provocando, conseqüentemente, um aumento do custo de sua implementação. Entretanto, para equipamentos sofisticados e de grande importância no processo produtivo, onde diferentes defeitos podem exibir sintomas similares e estes podem ocorrer simultaneamente, é mais vantajoso empregar técnicas de diagnóstico baseadas em técnicas mais complexas, uma vez que a relação **Q** entre as classes de funcionamento e os índices de monitoramento pode não ser determinada de forma única e neste caso, o custo de sua implementação se justifica (TROEDER *et al.*, 1987).

Já para o diagnóstico de equipamentos de baixo custo ou de menor responsabilidade dentro do processo produtivo, onde a relação entre a importância do equipamento e o custo de implementação de técnicas avançadas é baixa, a utilização das teorias acima citadas não se justifica. Também, espera-se que a maioria destes equipamentos sejam simples, possibilitando, conseqüentemente, a definição da relação **Q** entre as classes de funcionamento e os índices de monitoramento de forma única. Segundo TROEDER *et al.* (1987), se a relação **Q** estiver bem definida, métodos de classificação baseados em distâncias estatísticas podem ser empregados

com sucesso. Este capítulo apresenta dois métodos alternativos de classificação de padrões, um baseado na distância Euclideana e outro na distância de *Mahalanobis*, que exigem um menor esforço na fase de treinamento, são de fácil implementação e com resultados satisfatórios no diagnóstico de equipamentos.

Apresenta-se inicialmente uma revisão bibliográfica sobre os conceitos básicos da teoria de classificação de padrões e as principais divisões das metodologias de reconhecimento de padrões (*pattern recognition*). Dentre essas divisões, uma ênfase maior é dada aos métodos estatísticos de reconhecimento de padrões mais utilizados.

3.2 - Conceitos básicos de classificação de padrões

As propriedades de um padrão são denominadas atributos, no nosso caso usaremos índices de monitoramento, e o conjunto de m -atributos formam um vetor $\mathbf{x}$ que representa o padrão.

$$\mathbf{x} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{Bmatrix} \quad (3.1)$$

Um conjunto de padrões similares representam uma classe e o conjunto dessas classes é representado por h_j , $j = 1, 2, \dots, n$. Assim, por exemplo, no reconhecimento de números decimais, um conjunto será formado pelos símbolos 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, num total de 10 classes. Então, tem-se:

- número de classes: 10, $n = 10$
- número de atributos: m

Portanto, o problema de classificação consiste em determinar o conjunto de medidas que descrevem o padrão e que permitam a associação do padrão a uma das classes da aplicação.

A Figura 3.1 mostra um sistema básico de classificação.

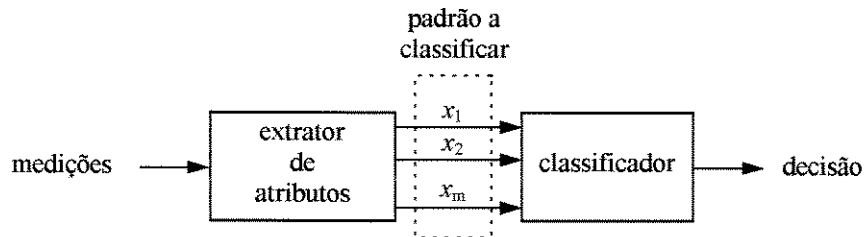


Figura 3.1 - Sistema básico de classificação de padrões.

O processo de classificação é feito pela formulação de regras ou funções de decisão. Assim temos que, nos seres humanos, o processo de classificação é precedido por um processo de aprendizado onde são geradas as regras de decisão. Estas regras de decisão são constantemente melhoradas através de experiências passadas ou treinamento.

Se não é possível a classificação de um padrão, isto significa que as medidas obtidas na etapa de extração não são suficientes para sua definição a partir das regras existentes, ou as regras existentes não definem o padrão.

Deste modo, um sistema de reconhecimento de padrões deverá incluir uma etapa de aprendizado e uma etapa de classificação; em alguns casos, o processo de aprendizado pode ser iterativo.

A Figura 3.2 mostra um diagrama de blocos que define as etapas do processo de aprendizado e classificação, num sistema de reconhecimento de padrões.

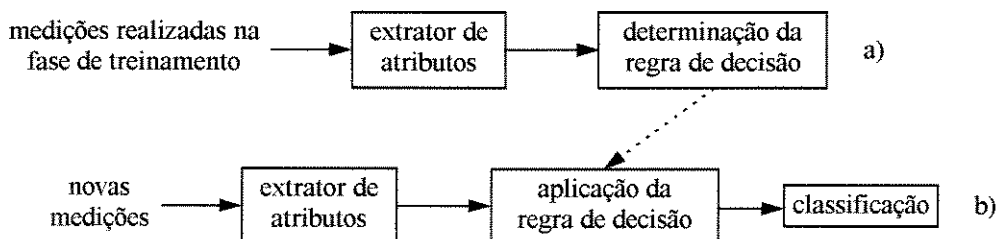


Figura 3.2 - Processo de aprendizado e classificação.

No processo de aprendizado (Figura 3.2a) são examinados padrões conhecidos obtendo-se, assim, valores para os atributos que caracterizam estes padrões.

O conjunto de atributos dos padrões da mesma classe do processo determinam uma regra de decisão. Obtém-se deste modo um conjunto de regras que definem unicamente uma das classes.

No processo de classificação (Figura 3.2b), com as regras já definidas, são apresentados ao sistema um conjunto de padrões de teste: estes padrões são processados e classificados com as regras de decisão obtidas no processo de aprendizado.

Se o resultado da classificação não for o desejado, isto significa que as regras obtidas não são corretas e que a partição do espaço de atributos nas diferentes classes não está bem definida, sendo possivelmente necessário considerar-se uma maior quantidade de atributos para uma melhor separação entre as classes.

O processo descrito pode ser sistematizado do seguinte modo:

- definir um espaço padrão de dimensão m (m : número de atributos), no qual estão representadas as n possíveis classes, $h_1, h_2, \dots, h_n$.
- um padrão da classe h_j é descrito pelo vetor $\mathbf{x}$, no qual estão contidos os m -atributos.
- realizar o processo de ajuste, denominado de aprendizado, que se define um método de classificação: este método otimiza a separação das classes ou a determinação da fronteira entre as regiões. O processo de aprendizado é realizado a partir de um conjunto de padrões rotulados, gerando um procedimento de classificação o qual é denominado regras de classificação. Portanto, o procedimento de aprendizado consiste na divisão do espaço padrão em n regiões, cada uma associada com uma classe. No caso de incorporação de novas classes estas regras são geralmente modificadas.
- o processo de reconhecimento é, basicamente, o uso do procedimento de classificação para a associação do padrão desconhecido a uma das classes.

A Figura 3.3 mostra graficamente os conceitos de atributos, classes e classificação definidos anteriormente.

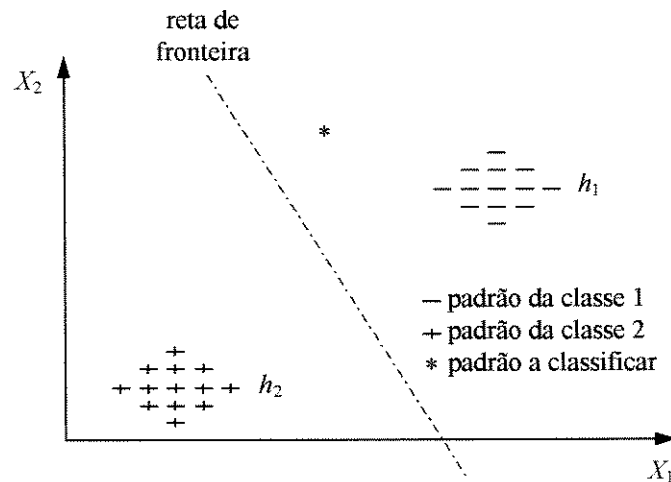


Figura 3.3 - Representação gráfica dos conceitos de atributos, classes e classificação.

Na Figura 3.3 estão representadas duas classes h_1 e h_2 formadas por padrões que possuem como atributos X_1 e X_2 . O padrão representado por *, poderá ser classificado como pertencente a classe h_1 , por existir uma maior proximidade da classe ou pelo fato do padrão encontrar-se acima da reta de fronteira que separa as duas classes.

A maioria dos classificadores de padrões baseiam-se numa das seguintes abordagens (SING-TZE BOW, 1984; FU, 1982):

- estatística (decisão teórica);
- sintática (estrutural).

3.3 - Métodos estatísticos

Os métodos estatísticos de reconhecimento de padrões têm sido utilizados no monitoramento e diagnóstico de equipamentos (MONOSTORI, 1986; HOUSHMAND e KANNATEY-ASIBU, 1989; ELBESTAWI *et al.*, 1989; DU *et al.*, 1992; LOUKIS *et al.*, 1992; LOUKIS *et al.*, 1994; DU *et al.*, 1995a; DU *et al.*, 1995b).

Os métodos estatísticos de reconhecimento de padrões podem ser divididos em dois grupos (SING-TZE BOW, 1984; FU, 1982):

- métodos paramétricos;
- métodos de distribuição livre ou não paramétricos.

Os métodos estatísticos são baseados na determinação de regras de decisão usando propriedades que permitem uma separação dos padrões em classes. Seja $\mathbf{x} = \{x_1, x_2, \dots, x_m\}$ um vetor que representa um padrão. Para uma aplicação com classes $h_1, h_2, \dots, h_n$, o problema básico consiste na geração de n funções de discriminação $d_1(\mathbf{x}), d_2(\mathbf{x}), \dots, d_n(\mathbf{x})$, que, para qualquer padrão $\mathbf{x} \in h_i$ satisfaça a equação:

$$d_i(\mathbf{x}) > d_j(\mathbf{x}) \quad \forall j \neq i \text{ para } j = 1, 2, \dots, n \quad (3.2)$$

Deste modo, o classificador pode ser representado pelo diagrama da Figura 3.4, onde cada bloco da esquerda representa o cálculo do valor $d_i(\mathbf{x})$ para $i = 1, 2, \dots, n$, e o bloco da direita representa a seleção do maior valor de d e a associação à classe i . Os critérios para a classificação são conectados em paralelo, de forma a tornar as combinações dos valores de d suficientes para distinguir as n classes.

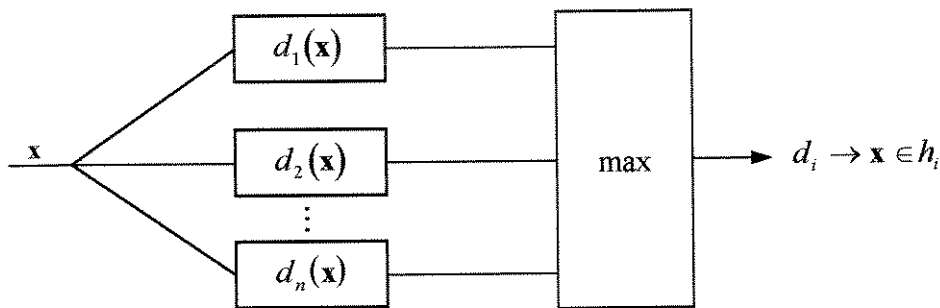


Figura 3.4 - Representação da classificação por um diagrama.

3.3.1 - Métodos estatísticos paramétricos

Os métodos paramétricos geram funções de discriminação estatísticas. Segundo SING-TZE BOW (1984), o uso de funções de discriminação estatísticas é vantajoso principalmente porque (1) o considerável conhecimento já existente em teorias de detecção, teorias de decisão, dentre outras, pode ser aplicado diretamente em problemas de reconhecimento de padrões; e (2) a formulação estatística é particularmente adequada, já que muitos processos de reconhecimento de padrões são modelados estatisticamente.

Os métodos paramétricos são baseados em modelos de distribuição de probabilidades, ou seja, a função de decisão baseia-se em uma função de densidade de probabilidade, uma probabilidade de ocorrência definida *a priori* e uma função de perda ou custo.

Antes de estabelecer as funções de perda, fazem-se necessárias algumas hipóteses:

- a probabilidade *a priori* $p(h_i)$ de ocorrer um padrão $\mathbf{x}$ de uma das classes h_i é conhecida ou pode ser estimada;
- a função de densidade de probabilidade condicional $p(\mathbf{x}|h_i)$ é conhecida ou pode ser estimada diretamente a partir do conjunto de treinamento;
- a probabilidade *a posteriori* $p(h_i|\mathbf{x})$ de $\mathbf{x}$ pertencer à classe h_i geralmente não é conhecida.

A função de perda L_{ik} pode ser definida como a perda ou o custo de decidir que $\mathbf{x} \in h_k$ quando, de fato, $\mathbf{x} \in h_i$. Analogamente, a perda média condicional $r_k(\mathbf{x})$ pode ser definida como (SING-TZE BOW, 1984):

$$r_k(\mathbf{x}) = \sum_{i=1}^n L_{ik} p(h_i | \mathbf{x}) \quad (3.3)$$

que é a perda média de classificar erroneamente $\mathbf{x} \in h_k$, quando de fato, $\mathbf{x}$ pertence a outra classe h_i , $i = 1, 2, \dots, n$ e $i \neq k$.

Pela regra de Bayes, define-se a probabilidade *a posteriori* $p(h_i | \mathbf{x})$ como (ANDREWS, 1972; SING-TZE BOW, 1984):

$$p(h_i | \mathbf{x}) = \frac{p(\mathbf{x} | h_i) p(h_i)}{p(\mathbf{x})} \quad (3.4)$$

onde $p(\mathbf{x}) = \sum_{i=1}^n p(\mathbf{x} | h_i) p(h_i)$.

Substituindo a equação (3.4) em (3.3) obtém-se a perda média de uma decisão sob todas as classes, que é expressa como:

$$r_k(\mathbf{x}) = \frac{1}{p(\mathbf{x})} \sum_{i=1}^n L_{ik} p(\mathbf{x} | h_i) p(h_i) \quad (3.5)$$

Como $p(\mathbf{x})$ da equação (3.5) é comum para todos $r_j(\mathbf{x})$, $j = 1, 2, \dots, n$, ele pode ser eliminado e então procura-se somente pelo mínimo:

$$\min_k r_k(\mathbf{x}) = \min_k \sum_{i=1}^n L_{ik} p(\mathbf{x} | h_i) p(h_i) \quad (3.6)$$

para se obter a melhor decisão entre todas as possíveis. Alternativamente, pode-se dizer que:

$$d_k(\mathbf{x}) = -r_k(\mathbf{x}) \quad (3.7)$$

onde $d_k(\mathbf{x})$ é a função de discriminação de Bayes. O classificador baseado nessa minimização é chamado de classificador de Bayes, o qual consegue um desempenho ótimo do ponto de vista estatístico, porém esta relação é de difícil aplicação, tendo em vista os seguintes aspectos:

- a probabilidade *a priori* $p(h_i)$ de ocorrer um padrão $\mathbf{x}$ de uma das classes h_i nem sempre é conhecida;
- a função de perda L_{ik} é difícil de ser avaliada.

Assim para uma implementação prática do classificador de Bayes assume-se que não existe perda associada com uma classificação correta e uma perda igual para qualquer classificação errada, ou seja:

$$L_{ik} = \begin{cases} 0 & \text{para } i = k \\ 1 & \text{para } i \neq k \end{cases} \quad (3.8)$$

Então o classificador de Bayes resulta na seleção da classe k que satisfaz:

$$p(\mathbf{x} | h_k) p(h_k) \geq p(\mathbf{x} | h_i) p(h_i) \quad (3.9)$$

Então a função de discriminação é agora:

$$d_k(\mathbf{x}) = p(\mathbf{x} | h_k) p(h_k) \quad (3.10)$$

Uma forma alternativa da função de discriminação acima é:

$$d'_k(\mathbf{x}) = \log p(\mathbf{x} | h_k) + \log p(h_k) \quad (3.11)$$

Assumindo, na equação (3.9), que $p(h_i) = p(h_k) \quad \forall i, k$, é possível obter a seguinte função de discriminação:

$$d_k(\mathbf{x}) = p(\mathbf{x} | h_k) \quad (3.12)$$

A associação do padrão à classe k é determinada pelo máximo valor da função $d_k(\mathbf{x})$, obtido a partir da densidade de probabilidade condicional. Segundo SING-TZE BOW (1984), o classificador que se baseia em algumas das três equações acima (3.10, 3.11 ou 3.12) é denominado classificador por máxima probabilidade ou máxima verossimilhança (*maximum likelihood classifier*).

Como um exemplo ilustrativo, suponha que as funções de densidade de probabilidade condicional $p(\mathbf{x}|h_k)$, $k = 1, 2, \dots, n$, sejam normais (gaussianas) com vetor média μ_k e matriz de covariância Σ_k conhecidas. A função densidade de probabilidade de uma distribuição normal multivariada para n classes é definida como (SING-TZE BOW, 1984; DUDA e HART, 1973):

$$p(\mathbf{x}|h_k) = \frac{1}{(2\pi)^{g/2} |\Sigma_k|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x} - \mu_k)' \Sigma_k^{-1} (\mathbf{x} - \mu_k) \right] \quad (3.13)$$

onde $\mathbf{x}$ é um vetor coluna, μ_k é o vetor média, Σ_k é a matriz de covariância, $|\Sigma_k|$ é o determinante de Σ_k e g é a dimensão do vetor padrão.

Substituindo-se a equação da densidade de probabilidade de uma distribuição normal multivariada (equação 3.13) na equação (3.11), a função de discriminação torna-se:

$$d_k''(\mathbf{x}) = -\frac{1}{2} (\mathbf{x} - \mu_k)' \Sigma_k^{-1} (\mathbf{x} - \mu_k) + \left[\log p(h_k) - \frac{1}{2} \log |\Sigma_k| \right] \quad (3.14)$$

Esta equação representa uma função de discriminação quadrática e pode ser colocada de uma forma mais compacta como:

$$d_k''(\mathbf{x}) = -\frac{1}{2} r^2 + f(k) \text{ para } \mathbf{x} \in h_k \quad (3.15)$$

onde $f(k) = \log p(h_k) - (1/2) \log |\Sigma_k|$ e $r^2 = (\mathbf{x} - \mu_k)' \Sigma_k^{-1} (\mathbf{x} - \mu_k)$.

Segundo DUDA e HART (1973), amostras retiradas de uma população normal formam um agrupamento simples, com centro determinado pelo vetor média obtido das amostras e formato determinado pela matriz de covariância. Segue-se da equação (3.13) que a localização dos pontos que possuem densidade de probabilidade constante são hiperelipsóides para as quais a forma quadrática $(\mathbf{x} - \mu)' \Sigma^{-1} (\mathbf{x} - \mu)$ é constante. Os eixos principais desses

hiperelipsóides são dados pelos auto-vetores de Σ , enquanto os auto-valores determinam os comprimentos desses eixos. A quantidade:

$$r^2 = (\mathbf{x} - \boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \quad (3.16)$$

é chamada de distância de *Mahalanobis* de $\mathbf{x}$ até $\boldsymbol{\mu}$. Então, os contornos que possuem densidade de probabilidade constante são hiperelipsóides com distâncias de *Mahalanobis* constantes até $\boldsymbol{\mu}$. A Figura 3.5 mostra uma representação gráfica de elipses com contornos de densidade de probabilidade constante, usando-se dois atributos (X_1 e X_2).

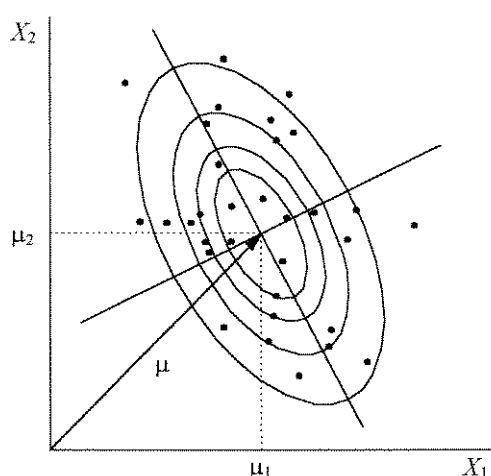


Figura 3.5 - Representação gráfica de elipses de contorno de densidade de probabilidade constante, usando dois atributos.

A performance dos métodos estatísticos paramétricos depende da distribuição de probabilidade $p(\mathbf{x}|h_i)$. Se as distribuições de probabilidade aproximam-se de modelos de distribuição de probabilidades conhecidos, a estimativa de Bayes é um ótimo classificador. Entretanto, para a maioria das aplicações, as distribuições de probabilidade não são próximas de distribuições de probabilidade conhecidas. Então, os métodos estatísticos de distribuição livre são, geralmente, preferidos.

3.3.2 - Métodos estatísticos de distribuição livre ou não paramétricos

Os métodos de distribuição livre baseiam-se na partição do espaço de atributos em regiões mutuamente exclusivas, com cada região correspondendo a uma classe. Superfícies de decisão (*decision surfaces*) podem ser usadas para separar os padrões conhecidos em suas respectivas classes e também para classificar padrões desconhecidos. A função de discriminação é uma função $d(x)$ que define a superfície de decisão. Existem vários caminhos para definir as funções de discriminação. Isto resulta em vários métodos de distribuição livre. Alguns métodos utilizados para a geração das funções de discriminação são apresentados a seguir.

3.3.2.1 - Classificador por função de discriminação linear

O classificador por função de discriminação linear consiste na geração de funções que são uma combinação linear dos atributos medidos das amostras de aprendizagem.

A função de discriminação para um classificador linear é representada na forma (ANDREWS, 1972; FU, 1982; ELBESTAWI *et al.*, 1989):

$$d_i(\mathbf{x}) = \sum_{k=1}^m a_{ik} x_k + a_{i, m+1}, \quad i = 1, \dots, n \quad (3.17)$$

A filosofia do classificador é definir uma superfície que separa os padrões pertencentes a uma classe dos pertencentes às demais classes. A equação desta superfície é definida abaixo (FU, 1982):

$$d_i(\mathbf{x}) - d_j(\mathbf{x}) = \sum_{k=1}^m a_k x_k + a_{m+1} = 0 \quad (3.18)$$

com $a_k = a_{ik} - a_{jk}$ e $a_{m+1} = a_{i, m+1} - a_{j, m+1}$.

A função de discriminação linear é determinada na fase de aprendizado a partir das amostras de aprendizagem.

Uma representação gráfica do classificador para o caso $n = 2$ é apresentada na Figura 3.6. O gráfico desta figura é obtido a partir das amostras de treinamento.

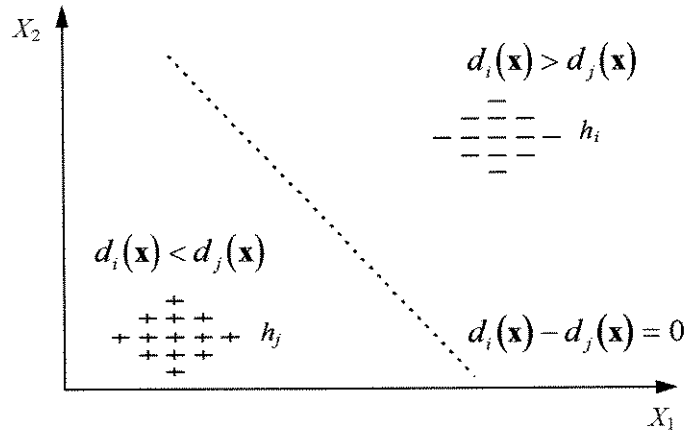


Figura 3.6 - Processo de classificação por função de discriminação linear para duas classes.

Na fase de classificação de um padrão, os valores dos atributos de $\mathbf{x}$ são substituídos na função de discriminação linear e se o resultado é positivo o padrão é associado à classe h_i , em caso contrário é associado à classe h_j .

$$\mathbf{x} \in h_i \text{ se } d_i(\mathbf{x}) - d_j(\mathbf{x}) > 0 \quad (3.19)$$

$$\mathbf{x} \in h_j \text{ se } d_i(\mathbf{x}) - d_j(\mathbf{x}) < 0 \quad (3.20)$$

No caso de duas classes ($n = 2$) é possível visualizar graficamente uma função de discriminação linear aproximada. Para uma aplicação com classes $h_1, h_2, \dots, h_n$, a determinação dos coeficientes a_k deve ser feita analiticamente, obtendo-se n funções de discriminação $d_1(\mathbf{x}), d_2(\mathbf{x}), \dots, d_n(\mathbf{x})$, que, para qualquer padrão $\mathbf{x} \in h_i$ satisfaçam a equação:

$$d_i(\mathbf{x}) > d_j(\mathbf{x}) \quad \forall j \neq i \text{ para } j = 1, 2, \dots, n \quad (3.21)$$

Deste modo, o classificador pode ser representado pelo diagrama da Figura 3.4.

Uma representação gráfica para o caso de 3 classes é apresentado na Figura 3.7.

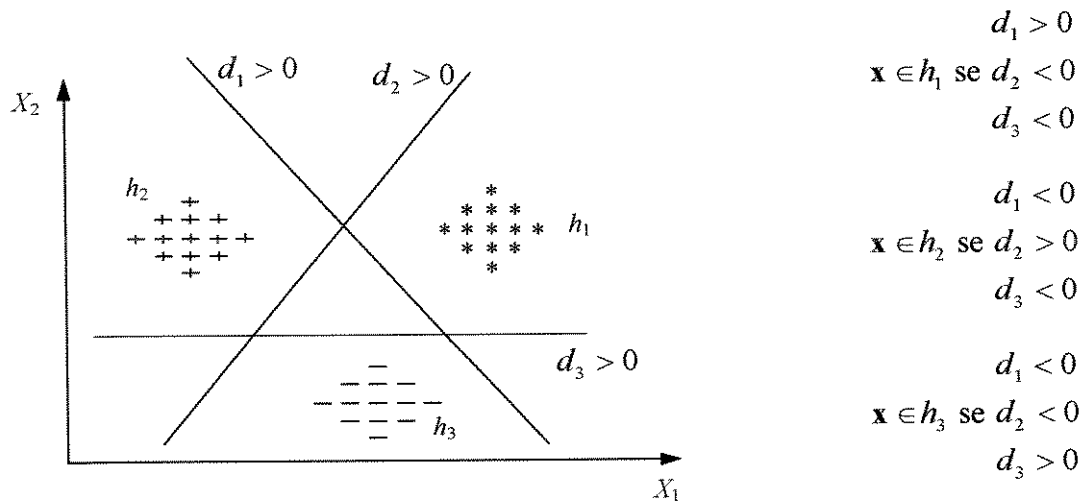


Figura 3.7 - Representação gráfica para a classificação de três classes, usando dois atributos.

A determinação dos coeficientes α_k só é possível se as amostras das diferentes classes forem linearmente separáveis, ou seja, puderem ser separadas por uma superfície de decisão.

Para a obtenção dos valores α_k são utilizados algoritmos que geram uma função de discriminação linear para cada classe (ANDREWS, 1972; FU 1982). A determinação dos valores α_k é feita de forma iterativa até que se obtenha uma classificação para as amostras de treinamento e consequentemente a obtenção das funções de discriminação linear para todas as classes.

Uma aplicação deste classificador pode ser encontrada em ELBESTAWI *et al.* (1989). Nesta aplicação, as amostras utilizadas na fase de classificação foram as mesmas da fase de treinamento, ocasionando, consequentemente, em altas taxas de reconhecimento das classes (85 a 100%).

3.3.2.2 - Classificador por distância mínima

O classificador por distância mínima é muito utilizado em reconhecimento estatístico de padrões. Usa como critério de classificação a distância entre o vetor definido pelos valores dos atributos extraídos do padrão a ser reconhecido e o conjunto de vetores obtidos na fase de aprendizagem. Do ponto de vista geométrico, os índices de monitoramento cobrem um espaço

m -dimensional $\mathcal{S}$. Neste espaço, cada classe, h_j , é caracterizada por um vetor padrão $\mathbf{p}_j = [p_{1j}, p_{2j}, \dots, p_{mj}]$. Por outro lado, a amostra $\mathbf{x}$ também é um vetor no espaço. A similaridade entre a amostra e o padrão pode ser medida pela distância entre os dois vetores. A distância mínima é então usada como critério para classificar a amostra. A Figura 3.8 mostra um simples exemplo, onde os índices de monitoramento são X_1 , X_2 e X_3 , e as três classes são descritas pelos padrões $\mathbf{p}_1$, $\mathbf{p}_2$ e $\mathbf{p}_3$. Dada uma amostra $\mathbf{x}$, as distâncias entre a amostra e os padrões são D_1 , D_2 e D_3 , respectivamente. Se a menor distância é D_1 , estima-se que a amostra pertence a classe h_1 .

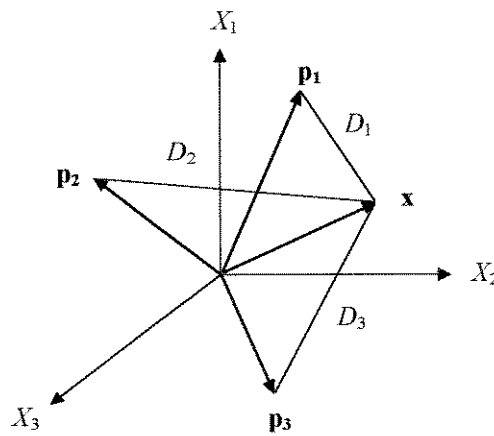


Figura 3.8 - Ilustração de um método de distribuição livre utilizado no reconhecimento de padrões.

Se $\mathbf{x} = \{x_1, x_2, \dots, x_m\}$ é um padrão desconhecido e $\{h_1, h_2, \dots, h_n\}$ são as classes com padrões de referências $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n\}$, então $\mathbf{x}$ é tido como classificado na classe h_j se:

$$D(\mathbf{x}, \mathbf{p}_j) < D(\mathbf{x}, \mathbf{p}_i) \text{ para } i \neq j; i = 1, 2, \dots, n \quad (3.22)$$

onde $D(\mathbf{x}, \mathbf{p}_j)$ é a distância entre $\mathbf{x}$ e $\mathbf{p}_j$.

A distância entre o vetor definido pelos valores dos atributos extraídos do padrão a ser reconhecido e o conjunto de vetores obtidos na fase de aprendizagem pode ser calculada de várias maneiras. As medidas de distâncias mais utilizadas são a distância Euclideana e a distância de *Mahalanobis*.

A distância Euclideana é a mais simples e a mais utilizada. Os padrões são as médias das amostras de aprendizagem (isto é, $\mathbf{p}_j = \boldsymbol{\mu}_j$, $j = 1, 2, \dots, n$) e a distância entre $\mathbf{x}$ e $\boldsymbol{\mu}_j$ é definida como:

$$D^2(\mathbf{x}, \boldsymbol{\mu}_j) = (\mathbf{x} - \boldsymbol{\mu}_j)'(\mathbf{x} - \boldsymbol{\mu}_j) \quad (3.23)$$

onde

$$\boldsymbol{\mu}_j = \frac{1}{N_j} \sum_{k=1}^{N_j} \mathbf{x}_k \quad (3.24)$$

representa o padrão de referência da classe j calculado com base nas amostras de aprendizagem $\mathbf{x}_k$ e N_j é o número de amostras que pertencem a j -ésima classe.

Na fase de classificação, para a nova amostra $\mathbf{x}$, as distâncias $D^2(\mathbf{x}, \boldsymbol{\mu}_j)$, $j = 1, \dots, n$ são calculadas. Então, a classe estimada é determinada por:

$$\mathbf{x} \in h_j \text{ se } D^2(\mathbf{x}, \boldsymbol{\mu}_j) < D^2(\mathbf{x}, \boldsymbol{\mu}_i) \text{ para } i \neq j; i = 1, \dots, n \quad (3.25)$$

A classe também pode ser determinada a partir da função de discriminação para o classificador baseado na mínima distância Euclideana (SING-TZE BOW, 1984; ANDREWS, 1972). Para que a amostra $\mathbf{x}$ seja classificada na classe j , a seguinte equação tem que ser satisfeita (SING-TZE BOW, 1984):

$$\mathbf{x} \in h_j \text{ se } \mathbf{x}'\boldsymbol{\mu}_j - \frac{1}{2}\boldsymbol{\mu}_j'\boldsymbol{\mu}_j > \mathbf{x}'\boldsymbol{\mu}_i - \frac{1}{2}\boldsymbol{\mu}_i'\boldsymbol{\mu}_i \text{ para } j \neq i; i = 1, 2, \dots, n \quad (3.26)$$

Na distância de *Mahalanobis*, os padrões também são as médias das amostras de aprendizagem e a distância é definida como:

$$D^2(\mathbf{x}, \mu_j) = (\mathbf{x} - \mu_j)^t \Sigma_j^{-1} (\mathbf{x} - \mu_j) \quad (3.27)$$

Na equação acima, μ_j representa o padrão de referência da classe j que é calculado através da equação (3.24) e,

$$\Sigma_j = \frac{1}{N_j - 1} \sum_{k=1}^{N_j} (\mathbf{x}_k - \mu_j)(\mathbf{x}_k - \mu_j)^t \quad (3.28)$$

representa a matriz de covariância da classe j calculada também com base nas amostras de aprendizagem $\mathbf{x}_k \in h_j$. Novamente, N_j é o número de amostras que pertencem a j -ésima classe.

Analogamente ao caso anterior, na fase de classificação, para a nova amostra $\mathbf{x}$, as distâncias $D^2(\mathbf{x}, \mu_j)$, $j = 1, \dots, n$ são calculadas. Então, a classe estimada é determinada por:

$$\mathbf{x} \in h_j \text{ se } D^2(\mathbf{x}, \mu_j) < D^2(\mathbf{x}, \mu_i) \text{ para } i \neq j; i = 1, \dots, n \quad (3.29)$$

A classe também pode ser determinada a partir da função de discriminação para o classificador baseado na mínima distância de *Mahalanobis*. Para que a amostra $\mathbf{x}$ seja classificada na classe j , a seguinte equação tem que ser satisfeita para todos $j \neq i$; $i = 1, 2, \dots, n$:

$$2\mathbf{x}^t \Sigma_j^{-1} \mu_j - \mathbf{x}^t \Sigma_j^{-1} \mathbf{x} - \mu_j^t \Sigma_j^{-1} \mu_j > 2\mathbf{x}^t \Sigma_i^{-1} \mu_i - \mathbf{x}^t \Sigma_i^{-1} \mathbf{x} - \mu_i^t \Sigma_i^{-1} \mu_i \quad (3.30)$$

Os classificadores por distância mínima, baseados nas distâncias Euclideana e de *Mahalanobis*, foram utilizados neste trabalho para a classificação das amostras não utilizadas na fase de treinamento. Estes dois classificadores não exigem um grande esforço na fase de

treinamento, são de fácil implementação e apresentam bons resultados. O classificador de *Mahalanobis*, ao contrário do classificador por mínima distância Euclideana, considera as propriedades estatísticas μ e Σ das amostras utilizadas na fase de treinamento.

3.4 - Considerações sobre amostras para treinamento e desempenho do classificador

Como observado nos itens anteriores, é necessária uma fase de aprendizado na qual são geradas as regras de classificação a partir de amostras rotuladas.

A seguir apresentam-se algumas considerações sobre as amostras de treinamento e a análise do desempenho do classificador obtido.

3.4.1 - Considerações sobre as amostras de treinamento

As amostras de treinamento usadas para definir o classificador ou regras de decisão devem satisfazer os seguintes requisitos:

- as amostras de treinamento devem ser representativas da aplicação, uma vez que a inclusão de amostras distorcidas levam à geração de regras de classificação sub-ótimas, com conseqüente diminuição do desempenho do classificador;
- o conjunto de amostras de treinamento deve ser o maior possível com a finalidade de prover a cobertura máxima de padrões da mesma classe;
- das amostras de treinamento deve ser possível a extração dos atributos mais significativos.

3.4.2 - Considerações sobre o desempenho do classificador

A análise do desempenho do classificador, do mesmo modo que a geração das regras de decisão é feita a partir de amostras rotuladas. As condições para a análise do desempenho do classificador são:

- não se deve usar para a validação o mesmo conjunto de padrões rotulados que foram utilizados na fase de aprendizado, neste caso o resultado obtido no desempenho do classificador leva a altíssimas taxas de reconhecimento, resultando num valor incorreto

para a medida do desempenho. Deve-se então separar o conjunto de padrões rotulados em amostras para gerar a regra de classificação e amostras para a análise do desempenho do classificador.

- se o desempenho do classificador não é o desejado, deverão ser geradas novas regras de decisão considerando-se atributos mais representativos.

Capítulo 4

Montagem e Procedimento Experimental

4.1 - Descrição da montagem experimental

No procedimento de aquisição e processamento de dados foram utilizados os seguintes equipamentos:

- [1] - Inversor FUJI, modelo FVR040E7S-2EX, 3 ϕ , 6.5 kVA, 17 A;
- [2] - Conjunto moto-bomba
 - bomba DARKA, modelo CM2
 - rotor CAÇULA, 4 pás, $\phi_{\text{ext}} = 106$ mm, fabricado em alumínio
 - eixo do rotor CAÇULA
 - motor WEG, modelo 56B0877, 0.75 CV;
- [3] - Acelerômetros ENDEVCO, modelo 272, 15.0 pC/g;
- [4] - Condicionadores/Amplificadores UNHOLTZ-DICKIE, modelo D22;
- [5] - Filtro analógico;
- [6] - Computador PC-AT 486 DX2 66 MHz;
- [7] - Placa de aquisição e conversão de sinais analógicos/digitais NATIONAL INSTRUMENTS, modelo AT-MIO-16F-5.

O esquema que mostra a interligação desses equipamentos no processo de medição de dados é mostrado na Figura 4.1, enquanto a Figura 4.2 mostra a disposição física do sistema estudado (conjunto moto-bomba, tubulações, válvulas, conexões e reservatório de água).

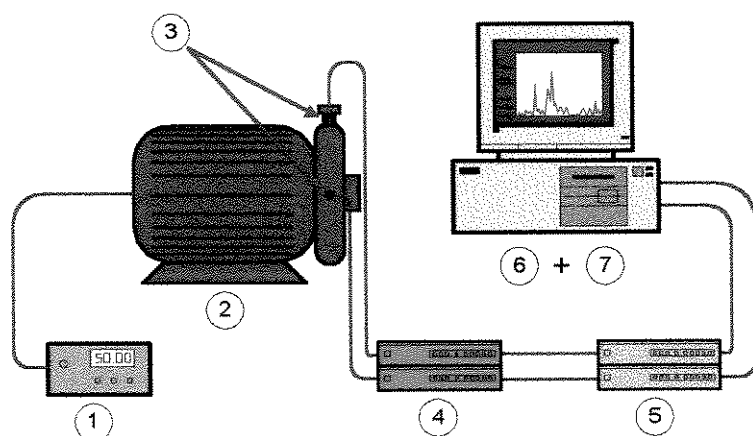


Figura 4.1 - Esquema da interligação dos equipamentos no processo de medição de dados.

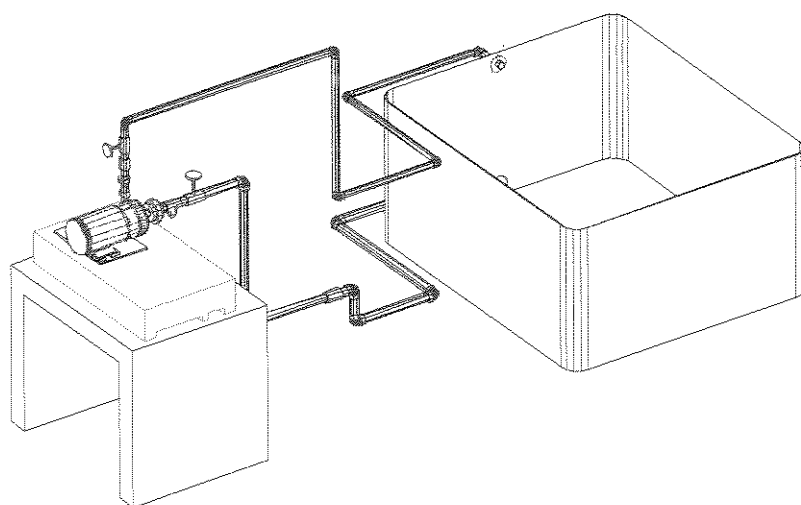


Figura 4.2 - Disposição física do sistema mecânico estudado.

4.2 - Descrição do programa de aquisição de dados

Com o objetivo de tornar a aquisição de dados on-line e feita de forma automática em intervalos pré-determinados, desenvolveu-se o programa Aquis 1.0. O sistema de aquisição de dados utilizado é constituído de uma placa de conversão conectada a um *slot* de expansão do microcomputador. Neste caso, a idéia principal foi a de aproveitar toda a estrutura de *hardware* e *software* do microcomputador de uma maneira simples e econômica.

O programa de aquisição de dados Aquis 1.0 foi desenvolvido utilizando-se o *software* BORLAND DELPHI® versão 1.0 para Windows 3.1x. O *software* DELPHI® foi escolhido devido a sua ótima interface gráfica e a sua versátil linguagem de programação (Pascal). O fluxograma do programa está esquematizado na Figura 4.3.

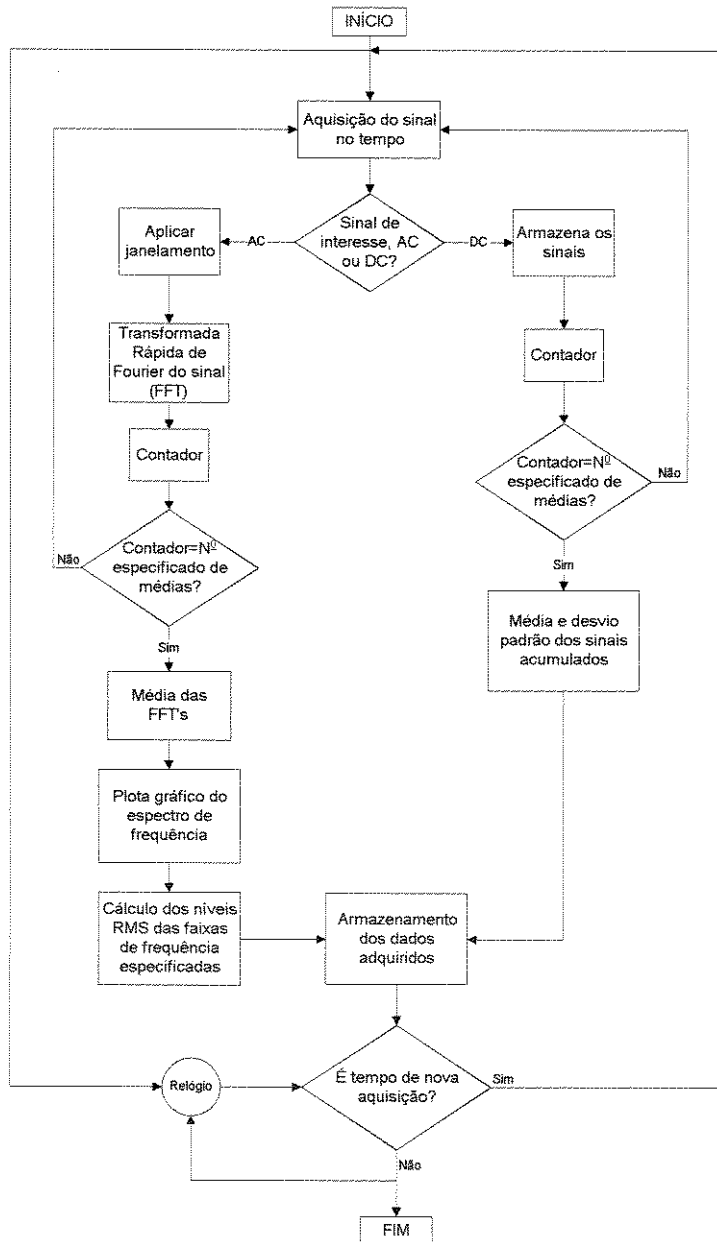


Figura 4.3 - Fluxograma do programa de aquisição de dados desenvolvido.

No programa de aquisição desenvolvido toda a interação com o usuário é executada através de janelas. O uso desta interface é bastante simples, onde todas as funções são apresentadas em menus, botões e caixas de texto. A configuração da aquisição de dados e a

visualização dos resultados gráficos são facilmente alteradas utilizando o mouse e/ou o teclado. O programa consiste de duas janelas, cada uma com uma função específica, sendo elas: janela Aquisição e janela Configuração.

Na janela Configuração, mostrada na Figura 4.4, os nomes dos arquivos para o banco de dados e as bandas de frequência utilizadas para o cálculo dos valores RMS são definidas. Nesta janela define-se também a configuração da aquisição de dados, ou seja, canais habilitados, sinal AC ou DC, sensibilidade, taxa de amostragem (amostras por segundo), frequência de análise, intervalo entre amostragens, número de amostras, tipo de janelamento e número de médias.

Análise de Dados

Configuração da aquisição

Canal 1 ☒ habilitado mV/Un: 1000 ☒ AC ☐ DC

Canal 2 ☒ habilitado mV/Un: 1000 ☒ AC ☐ DC

Canal 3 ☐ habilitado mV/Un: 1000 ☒ AC ☐ DC

Canal 4 ☐ habilitado mV/Un: 1000 ☒ AC ☐ DC

Canal 5 ☐ habilitado mV/Un: 1000 ☒ AC ☐ DC

Canal 6 ☐ habilitado mV/Un: 1000 ☒ AC ☐ DC

Canal 7 ☐ habilitado mV/Un: 1000 ☒ AC ☐ DC

Canal 8 ☐ habilitado mV/Un: 1000 ☒ AC ☐ DC

Dispositivo: 1

Taxa de amostragem: 4000

Intervalo de amostragem [s]: 600

Nº de amostras: 1024

Nº de médias: 50

Tipo de janela: Hanning

Frequência de análise: 1000

Arquivo de dados

Nome: C:\dados\dados25.dbf

Bandas de análise:

	Frequência [Hz]	Faixa [Hz]
1	58.59	5
2	119.14	5

Aquisição Configuração

Sair

Figura 4.4 - A janela Configuração.

Na janela Aquisição, mostrada na Figura 4.5, o gráfico da FFT (caso AC) ou o valor da média e do desvio padrão do sinal (caso DC) são mostrados. A janela gráfica mostra o resultado de um canal por vez. Para visualizar um resultado de um outro canal, basta clicar no número do canal desejado. Nesta janela temos também duas caixas de texto onde é possível visualizar os valores das coordenadas X e Y do ponto clicado na janela gráfica (caso AC). É ainda nesta janela que opta-se por salvar os dados em disco ou não e onde encontra-se o botão de controle do início e fim da aquisição dos dados.

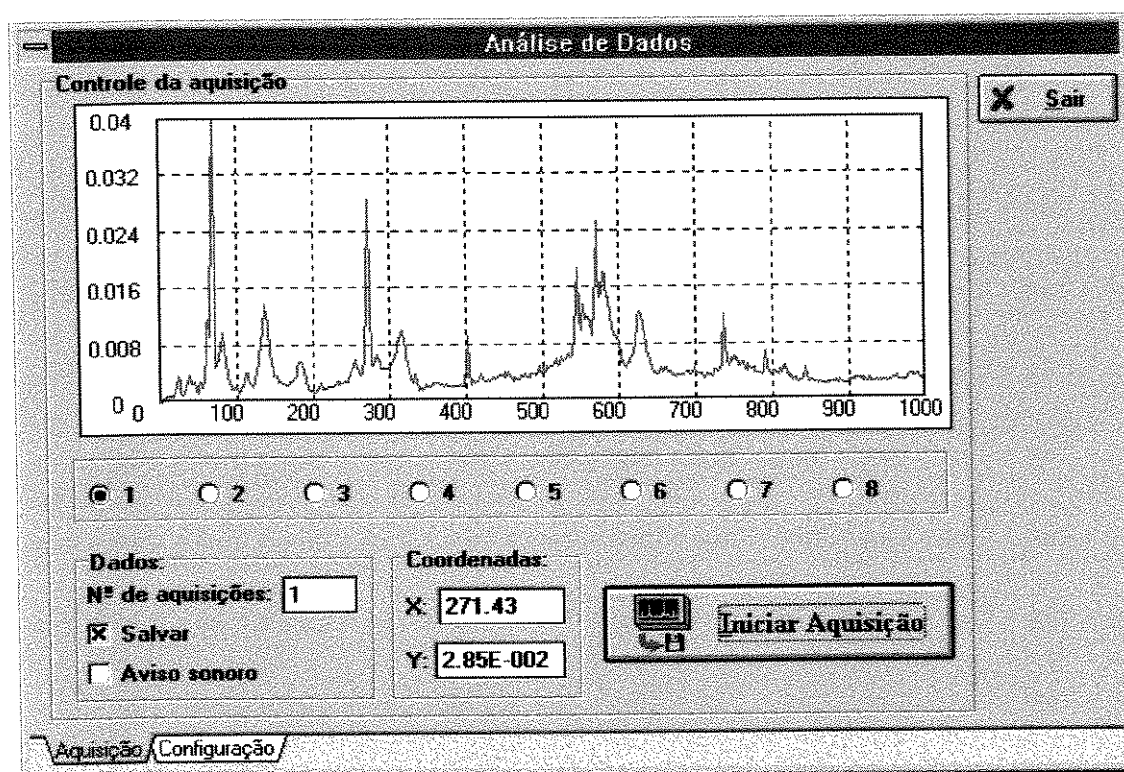


Figura 4.5 - A janela Aquisição.

4.3 - Descrição do experimento

Os experimentos foram conduzidos em três condições operacionais diferentes: 1) motor elétrico com frequência de alimentação $f_a = 50$ Hz, 2) motor elétrico com frequência de alimentação $f_a = 60$ Hz e 3) motor elétrico com frequência de alimentação $f_a = 70$ Hz. A frequência de alimentação foi modificada através de um inversor instalado entre a rede elétrica e o motor. Foram definidos dois pontos de medição, como mostra a Figura 4.6. Os sinais de vibração, aceleração vertical (canal 1) e aceleração horizontal (canal 2), foram medidos usando-se acelerômetros aparafusados em bases de fibra que estavam coladas na bomba.

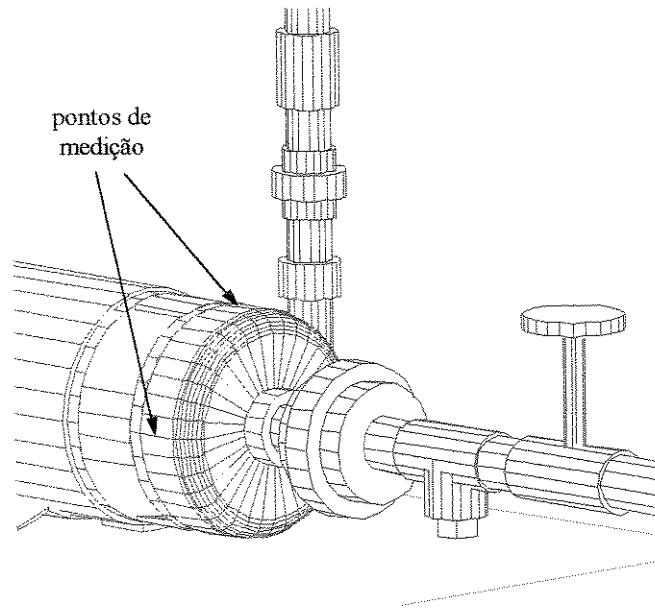


Figura 4.6 - Pontos de medição utilizados.

Durante os experimentos, os sinais foram condicionados e filtrados (frequência de corte: 1000 Hz); e então amostrados pelo conjunto computador+placa de aquisição a uma frequência de 2000 Hz por canal. Cada conjunto de dados consistia de 1024 pontos. A janela *Hanning* foi aplicada sobre cada conjunto de dados com o objetivo de minimizar o efeito de truncamento (*leakage*). A Transformada Rápida de Hartley (FHT) foi usada para o cálculo da FFT (SORENSEN *et al.*, 1985). Cada FFT obtida era resultado da média de valores obtidos em 50 diferentes aquisições. Finalmente, os valores RMS das acelerações foram obtidos a partir de bandas de frequência específicas. Todo esse processo foi repetido a cada 10 minutos. A Tabela 4.1 traz, de forma resumida, a configuração usada para a aquisição de dados.

INTERVALO ENTRE AMOSTRAGENS = 600 s = 10 min
NÚMERO DE MÉDIAS = 50
NÚMERO DE AMOSTRAS = 1024
JANELAMENTO = HANNING
FREQUÊNCIA DE ANÁLISE = 1000 Hz
NÚMERO DE CANAIS HABILITADOS = 2
TAXA DE AMOSTRAGEM POR CANAL = 2000 Hz

Tabela 4.1 - Configuração usada para a aquisição de dados.

4.3.1 - Índices de monitoramento utilizados

Cada tipo de defeito afeta o comportamento vibratório dos equipamentos de forma peculiar. Para a detecção e identificação destes defeitos uma metodologia bastante aplicada é a análise do espectro de frequência em banda estreita. Nesta análise procura-se relacionar aquelas frequências que tem amplitude modificada com o surgimento de um determinado defeito. Para os defeitos mais comumente encontrados em equipamentos estas relações são bem conhecidas e facilmente encontradas na literatura. A Tabela 4.2 (TRANTER, 1989) é um exemplo destas.

Frequência	Causa possível
1xRPM	desbalanceamento
	desalinhamento ou empenamento
	ressonância elétrica - diminui sensivelmente com alguma variação de velocidade e geralmente apresenta bandas em 2x a frequência
	perda de tensão (<i>Strain Looseness</i>) - causada por distorções na estrutura ou fundação ou de estruturas acopladas (p. e. encanamentos)
2xRPM	desalinhamento ou empenamento
NxRPM	N = nº aletas ou pás As harmônicas geralmente indicam algum problema
ressonância	muitas fontes, incluindo eixo, carcaça, fundação ou estruturas acopladas. Testes de partida e análise modal são úteis nesta área.

Tabela 4.2 - Tabela comumente usada para o diagnóstico de conjuntos moto-bomba.

Os principais tipos de defeitos que ocorrem em moto-bombas são desbalanceamento, empenamento, desalinhamento, vazamentos no selo mecânico e variações no fluxo provocados por obstruções e/ou fenômenos hidráulicos. Os três primeiros defeitos, pela Tabela 4.2, estão relacionados a 1x e 2xRPM. As variações no fluxo geralmente não estão relacionadas a frequências específicas, mas reflete-se na frequência de passagem de pás, que no conjunto moto-bomba estudado é igual a 4xRPM (4 pás).

Com base nesta análise escolheu-se como índices de monitoramento os níveis de aceleração RMS nas bandas de frequência em torno de 1x, 2x e 4xRPM. A Tabela 4.3 apresenta, para cada rotação do motor, os índices de monitoramento utilizados. A notação X_i^v representa o i -ésimo índice de monitoramento da direção vertical e X_i^h representa o i -ésimo índice de monitoramento da direção horizontal.

frequência de alimentação do motor [Hz]	índices de monitoramento	descrição dos índices de monitoramento
50	X_1^v e X_1^h	aceleração RMS na banda [44,92 - 52,74] Hz
	X_2^v e X_2^h	aceleração RMS na banda [95,7 - 103,52] Hz
	X_3^v e X_3^h	aceleração RMS na banda [191,4 - 199,22] Hz
60	X_4^v e X_4^h	aceleração RMS na banda [54,69 - 62,51] Hz
	X_5^v e X_5^h	aceleração RMS na banda [115,23 - 123,05] Hz
	X_6^v e X_6^h	aceleração RMS na banda [230,47 - 238,29] Hz
70	X_7^v e X_7^h	aceleração RMS na banda [64,45 - 72,27] Hz
	X_8^v e X_8^h	aceleração RMS na banda [136,72 - 144,54] Hz
	X_9^v e X_9^h	aceleração RMS na banda [267,57 - 275,39] Hz

Tabela 4.3 - Índices de monitoramento utilizados para cada condição operacional.

4.3.2 - Definição das classes de funcionamento

Dentro de cada condição operacional definiu-se várias classes de funcionamento. A classe de funcionamento 1 refere-se à condição normal de operação do conjunto moto-bomba. As classes de funcionamento 2 a 5 simulam obstruções na tubulação de sucção ou de recalque. Estas obstruções foram provocadas através do fechamento das válvulas globo esquematizadas na Figura 4.7. As obstruções são definidas como percentagens do número de voltas necessárias para o fechamento total da válvula, a partir da abertura total. Já as classes 6 a 8 simulam modificações geométricas em uma ou em quatro pás do rotor. Estas classes são explicadas mais detalhadamente a seguir.

- **Classe de funcionamento 1** - Condição normal de operação do conjunto moto-bomba: válvulas globo da tubulação de sucção ($\phi = 1''$) e da tubulação de recalque ($\phi = 3/4''$) totalmente abertas. As válvulas das tubulações de sucção e de recalque são mostradas na Figura 4.7.

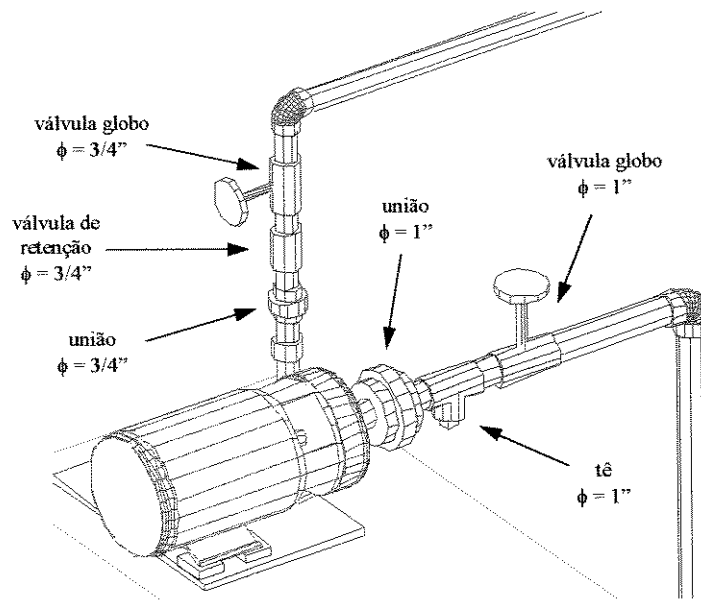


Figura 4.7 - Desenho esquemático das válvulas das tubulações de sucção e de recalque.

- **Classe de funcionamento 2** - Simulação de obstrução 1 na tubulação de recalque: válvula globo da tubulação de sucção totalmente aberta e válvula globo da tubulação de recalque com 45% de obstrução, ou seja, 45% fechada.
- **Classe de funcionamento 3** - Simulação de obstrução 2 na tubulação de recalque: válvula globo da tubulação de sucção totalmente aberta e válvula globo da tubulação de recalque com 83% de obstrução.
- **Classe de funcionamento 4** - Simulação de obstrução 1 na tubulação de sucção: válvula globo da tubulação de recalque totalmente aberta e válvula globo da tubulação de sucção 26% de obstrução.
- **Classe de funcionamento 5** - Simulação de obstrução 2 na tubulação de sucção: válvula globo da tubulação de recalque totalmente aberta e válvula globo da tubulação de sucção 61% de obstrução.
- **Classe de funcionamento 6** - Simulação de defeito 1 em uma pá do rotor: através de uma lima desbastou-se uma pá do rotor em 3 mm, como mostra a Figura 4.8.

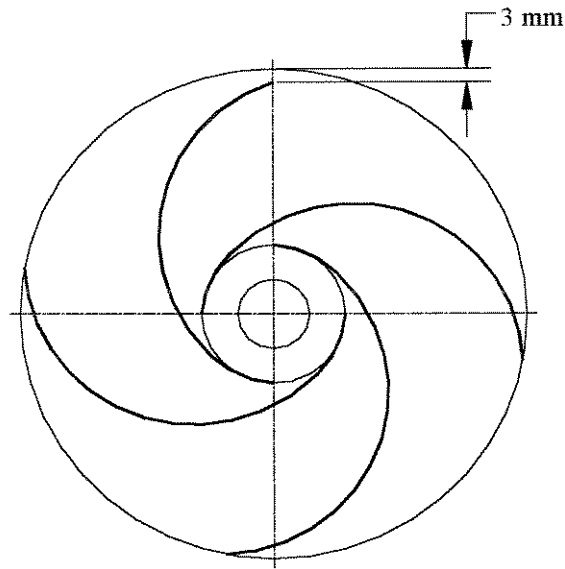


Figura 4.8 - Defeito na pá do rotor $\rightarrow d = 3 \text{ mm}$.

- **Classe de funcionamento 7** - Simulação de defeito 2 em uma pá do rotor: a mesma pá do rotor foi desbastada até a profundidade de 6 mm (Figura 4.9).

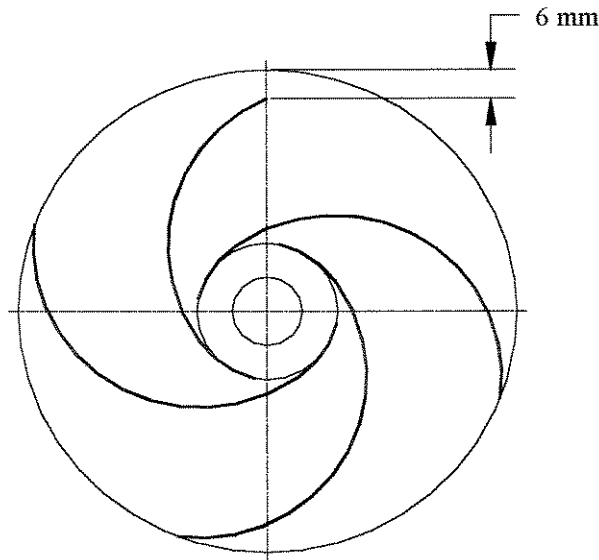


Figura 4.9 - Defeito na pá do rotor $\rightarrow d = 6 \text{ mm}$.

- **Classe de funcionamento 8** - Simulação de defeito nas 4 pás do rotor: as 4 pás do rotor foram desbastadas até a profundidade de 6 mm.

4.3.3 - Medições

Antes do início das medições, o conjunto moto-bomba foi submetido a uma operação de manutenção. Nesta operação de manutenção substituiu-se o rotor, o eixo do rotor e o selo mecânico.

Numa fase preliminar mediu-se a aceleração em dois pontos do conjunto moto-bomba durante 24 horas ininterruptamente. Depois, o equipamento ficou 24 horas parado para ser submetido, no dia seguinte, a um novo período de medições de 5 horas. Pela análise da Figura 4.10 percebe-se inicialmente um período de acomodação dos componentes decorrente da manutenção realizada. Depois deste período de acomodação os valores de aceleração medidos na direção vertical apresentaram característica aproximadamente estacionária com pequenas variações periódicas relacionadas a condições climáticas. Após uma parada de 24 horas e consequente partida do motor, os sinais apresentaram características similares de média e variância. Com base nesta análise, adotou-se um período de 12 horas de medições para representar as classes de funcionamento, pois, se permite que os efeitos climáticos sejam atenuados sem aumentar em demasia o tempo de medição.

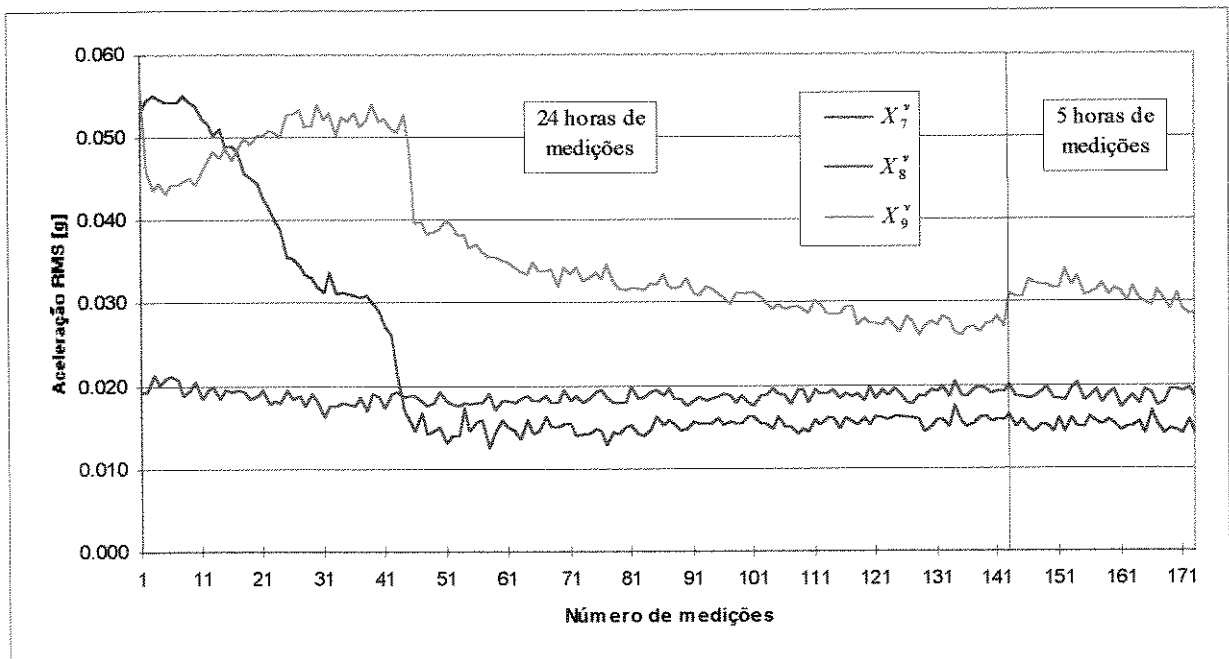


Figura 4.10 - Comportamento vibratório do sistema após parada e partida do motor.

Para inserir os defeitos representativos das classes de funcionamento 6, 7 e 8 foi necessária a desmontagem da bomba. No entanto, procurou-se efetuar esta desmontagem afetando o menor número possível de componentes conforme mostra a Figura 4.11, ou seja, somente a carcaça da bomba foi retirada, sem alterações no alinhamento, selo mecânico e posicionamento do rotor. Nesta fase, para avaliar a influência da operação de desmontagem, realizou-se o seguinte procedimento. Inicialmente realizou-se 18 horas de medições. A seguir, o conjunto foi submetido ao processo de desmontagem descrito acima e as acelerações vertical e horizontal foram medidas durante um período de 35 horas. Ainda, um novo período de medições de 12 horas foi realizado após uma segunda desmontagem na qual desbastou-se uma pá do rotor em 3 mm. Os dados obtidos são mostrados na Figura 4.12. Analisando esta figura percebe-se que a mudança inserida pela desmontagem não compromete a análise de diagnóstico, pois a variação dos índices de monitoramento apresentada não foi significativa em relação a variação dos índices obtidos a partir da simulação do defeito.

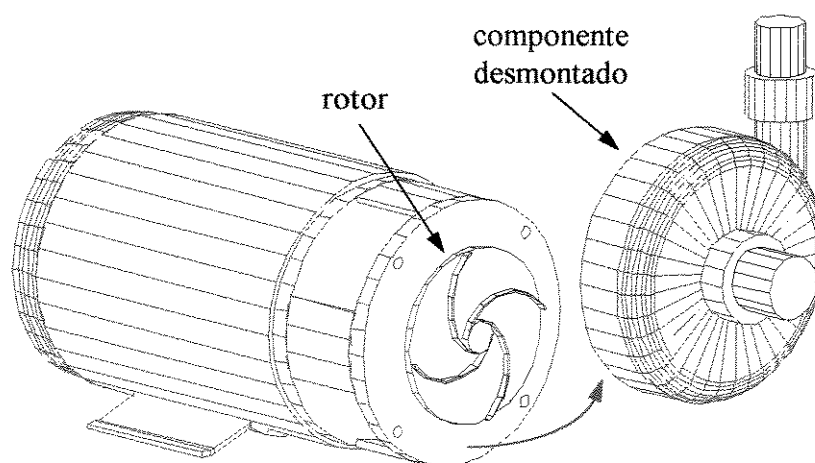


Figura 4.11 - Processo de desmontagem utilizado na bomba.

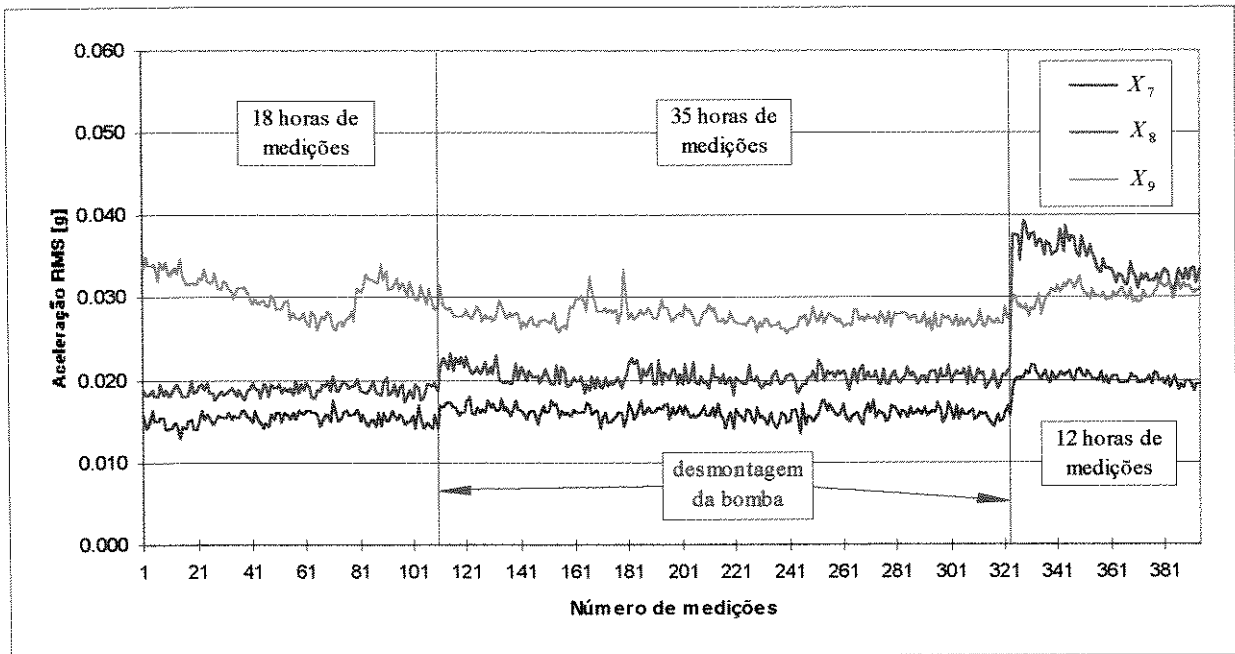


Figura 4.12 - Comportamento vibratório do sistema após desmontagens da bomba.

Após a solução destes problemas, partiu-se para o processo de medição propriamente dito. Para cada classe, obteve-se um determinado número de conjunto de medições, onde cada conjunto compreende as medições realizadas durante um período total de 36 horas sem o desligamento do motor, sendo 12 horas sob a frequência de alimentação de 50 Hz, 12 horas sob 60 Hz e as outras 12 horas sob 70 Hz. Em outras palavras, para uma mesma classe de funcionamento, um conjunto é constituído de três períodos de medições de 12 horas, cada um correspondendo a uma condição operacional diferente. A Tabela 4.4 mostra o número de conjuntos de medições realizados para cada classe de funcionamento.

Classes de funcionamento	Conjuntos de medições
Classe 1	3
Classe 2	2
Classe 3	2
Classe 4	2
Classe 5	2
Classe 6	2
Classe 7	3
Classe 8	3

Tabela 4.4 - Número de conjuntos de medições realizados para cada classe de funcionamento

Para se verificar o efeito de uma operação de manutenção no conjunto moto-bomba, os componentes rotor, eixo do rotor e selo mecânico foram novamente trocados. Então, após a manutenção do conjunto moto-bomba, definiu-se a classe de funcionamento 9: condição normal de operação do conjunto moto-bomba com válvulas globo da tubulação de sucção e de recalque totalmente abertas. Dois novos conjuntos de medição foram obtidos para a classe 9 após um tempo de estabilização de 36 horas. O objetivo é verificar se o método de diagnóstico conseguirá classificar corretamente estes novos padrões da classe 9 em relação aos padrões da classe 1 obtidos anteriormente. Seria interessante que o resultado desta classificação fosse satisfatório, para que os padrões anteriores das classes 2 a 8 também pudessem servir como referências para as novas medições.

Capítulo 5

Análise dos Resultados

5.1 - Obtenção dos padrões rotulados e das matrizes de covariância

Antes de se analisar o desempenho do classificador baseado na distância de *Mahalanobis* e do classificador baseado na distância Euclideana, faz-se necessária a obtenção dos padrões e das matrizes de covariância para a geração das regras de classificação. Os padrões, que representam as classes de funcionamento, e as matrizes de covariância foram obtidas a partir de um conjunto de amostras de treinamento usando-se as equações:

$$\mu_j = \frac{1}{N_j} \sum_{k=1}^{N_j} \mathbf{x}_k \quad (5.1)$$

$$\Sigma_j = \frac{1}{N_j - 1} \sum_{k=1}^{N_j} (\mathbf{x}_k - \mu_j)(\mathbf{x}_k - \mu_j)^t \quad (5.2)$$

onde μ_j é o padrão de referência da classe j calculado com base nas amostras de aprendizagem $\mathbf{x}_k$. Σ_k é a matriz de covariância da classe j também calculada com base nas amostras de aprendizagem $\mathbf{x}_k$ e N_j é o número de amostras que pertencem a j -ésima classe. Em todos os casos, o primeiro conjunto de medições obtido, para cada classe de funcionamento, foi considerado como o conjunto de amostras de treinamento com N_j em torno de 72 amostras. É útil lembrar a definição da expressão “conjunto de medições”. Para cada classe de funcionamento, obteve-se um determinado número de “conjunto de medições”, onde cada conjunto compreende as medições realizadas durante um período total de 36 horas

sem o desligamento do motor, sendo 12 horas sob a frequência de alimentação de 50 Hz, 12 horas sob 60 Hz e as outras 12 horas sob 70 Hz. A Tabela 4.4 do capítulo anterior mostra o número de conjuntos de medições que foram obtidos para cada classe de funcionamento.

5.2 - Variação dos índices de monitoramento

As figuras 5.1 a 5.6 mostram como foram as variações das médias dos índices de monitoramento em relação às diferentes classes de funcionamento do conjunto moto-bomba. Os pontos apresentados nestas figuras representam as médias dos índices de monitoramento calculadas de acordo com a equação (5.1) com N_j , em torno de 72 amostras.

Na Figura 5.1, por exemplo, os primeiros pontos de cada classe, referentes aos índices de monitoramento X_1^v , X_2^v e X_3^v , representam as componentes do vetor padrão considerando apenas a direção vertical. Os demais, também obtidos pelo uso da equação (5.1), representam as médias das amostras a serem classificadas. Nas figuras 5.1 a 5.6, embora estejam apresentadas as médias das amostras a serem classificadas, a classificação foi feita considerando todas as amostras e não somente a média destas.

5.2.1 - Variação dos índices de monitoramento para a condição operacional 1

A Figura 5.1 mostra, para a condição operacional 1 ($f_a = 50$ Hz), como foi a variação dos índices de monitoramento da direção vertical (X_1^v , X_2^v e X_3^v) em relação às classes de funcionamento do conjunto moto-bomba.

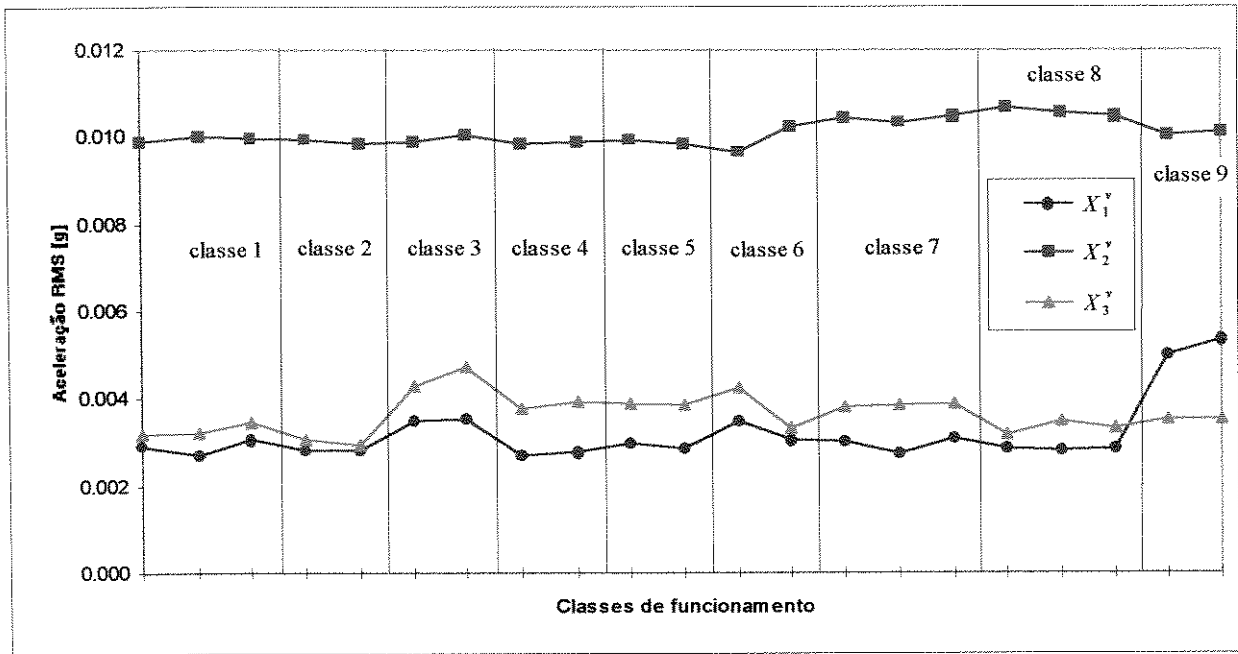


Figura 5.1 - Variação dos índices de monitoramento X_1^v , X_2^v e X_3^v para $f_a = 50$ Hz.

A Figura 5.2 mostra, para a condição operacional 1 ($f_a = 50$ Hz), como foi a variação dos índices de monitoramento da direção horizontal (X_1^h , X_2^h e X_3^h) em relação às classes de funcionamento do conjunto moto-bomba.

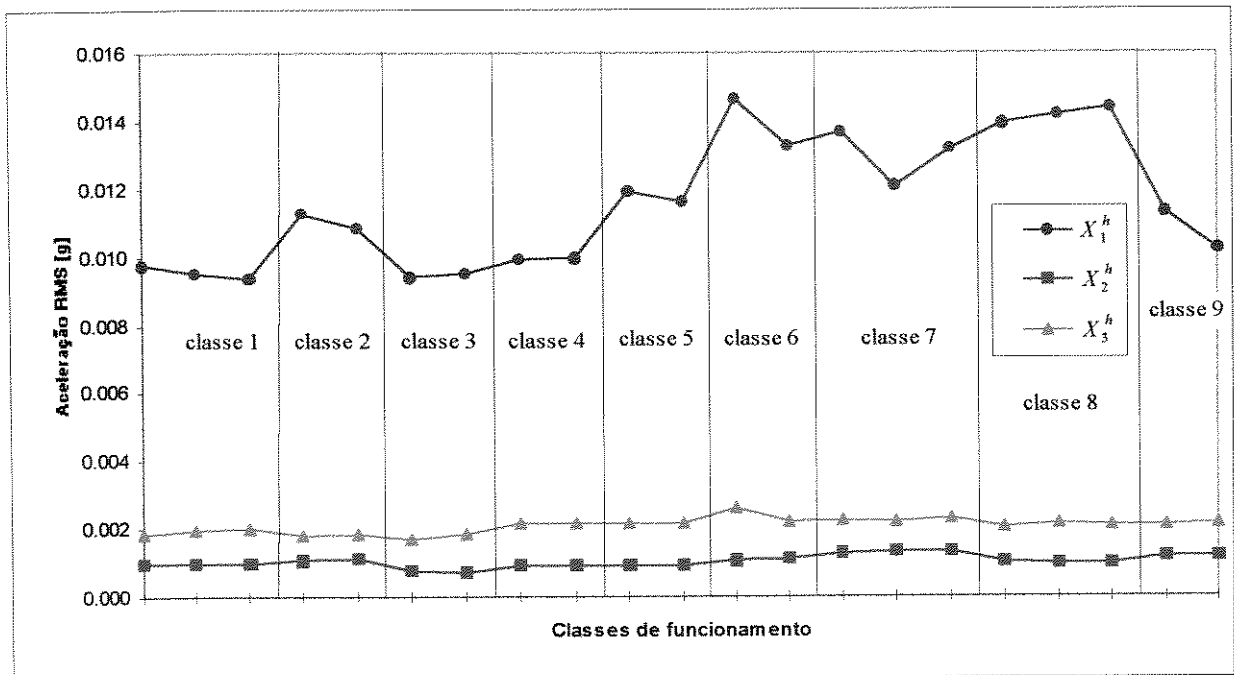


Figura 5.2 - Variação dos índices de monitoramento X_1^h , X_2^h e X_3^h para $f_a = 50$ Hz.

5.2.2 - Variação dos índices de monitoramento para a condição operacional 2

A Figura 5.3 mostra, para a condição operacional 2 ($f_a = 60$ Hz), como foi a variação dos índices de monitoramento da direção vertical (X_4^v , X_5^v e X_6^v) em relação às classes de funcionamento do conjunto moto-bomba.

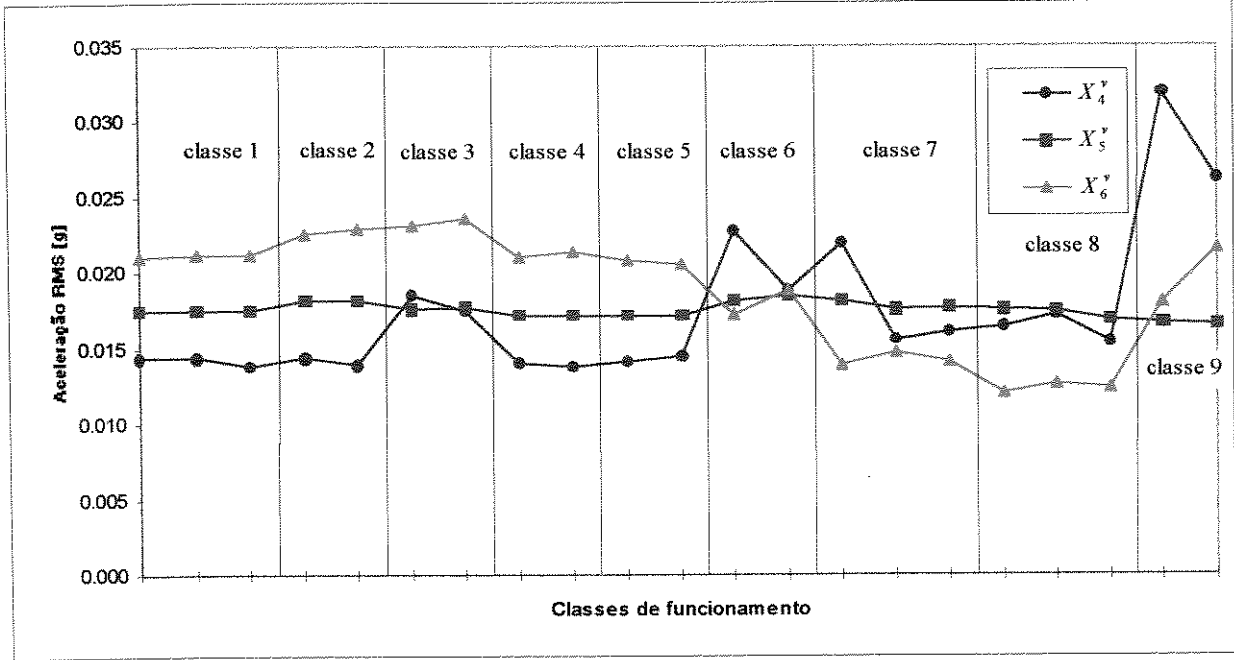


Figura 5.3 - Variação dos índices de monitoramento X_4^v , X_5^v e X_6^v para $f_a = 60$ Hz.

A Figura 5.4 mostra, para a condição operacional 2 ($f_a = 60$ Hz), como foi a variação dos índices de monitoramento da direção horizontal (X_4^h , X_5^h e X_6^h) em relação às classes de funcionamento do conjunto moto-bomba.

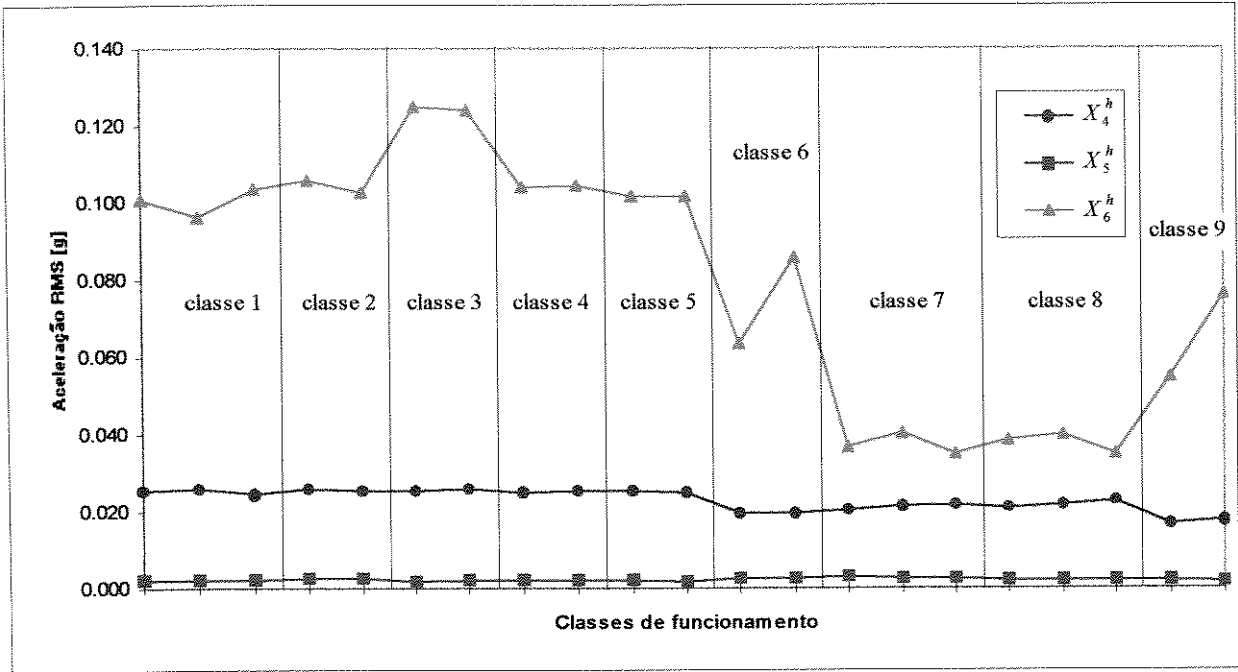


Figura 5.4 - Variação dos índices de monitoramento X_4^h , X_5^h e X_6^h para $f_a = 60$ Hz.

5.2.3 - Variação dos índices de monitoramento para a condição operacional 3

A Figura 5.5 mostra, para a condição operacional 3 ($f_a = 70$ Hz), como foi a variação dos índices de monitoramento da direção vertical (X_7^v , X_8^v e X_9^v) em relação às classes de funcionamento do conjunto moto-bomba.

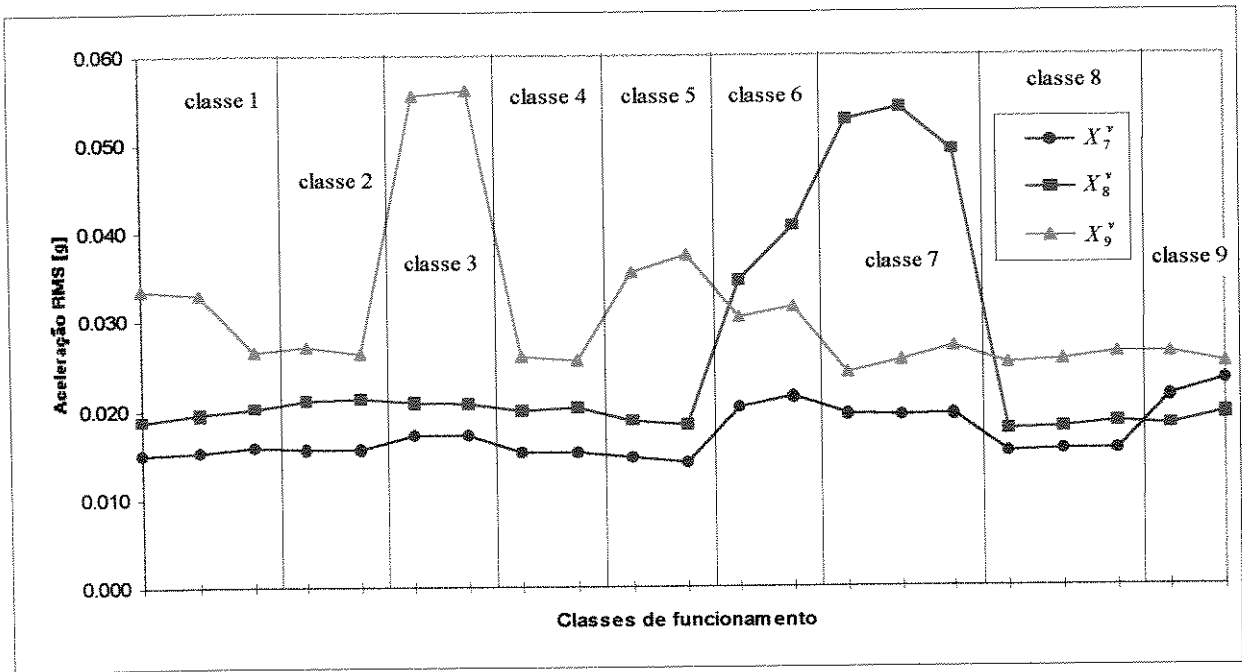


Figura 5.5 - Variação dos índices de monitoramento X_7^v , X_8^v e X_9^v para $f_a = 70$ Hz.

A Figura 5.6 mostra, para a condição operacional 3 ($f_a = 70$ Hz), como foi a variação dos índices de monitoramento da direção horizontal (X_7^h , X_8^h e X_9^h) em relação às classes de funcionamento do conjunto moto-bomba.

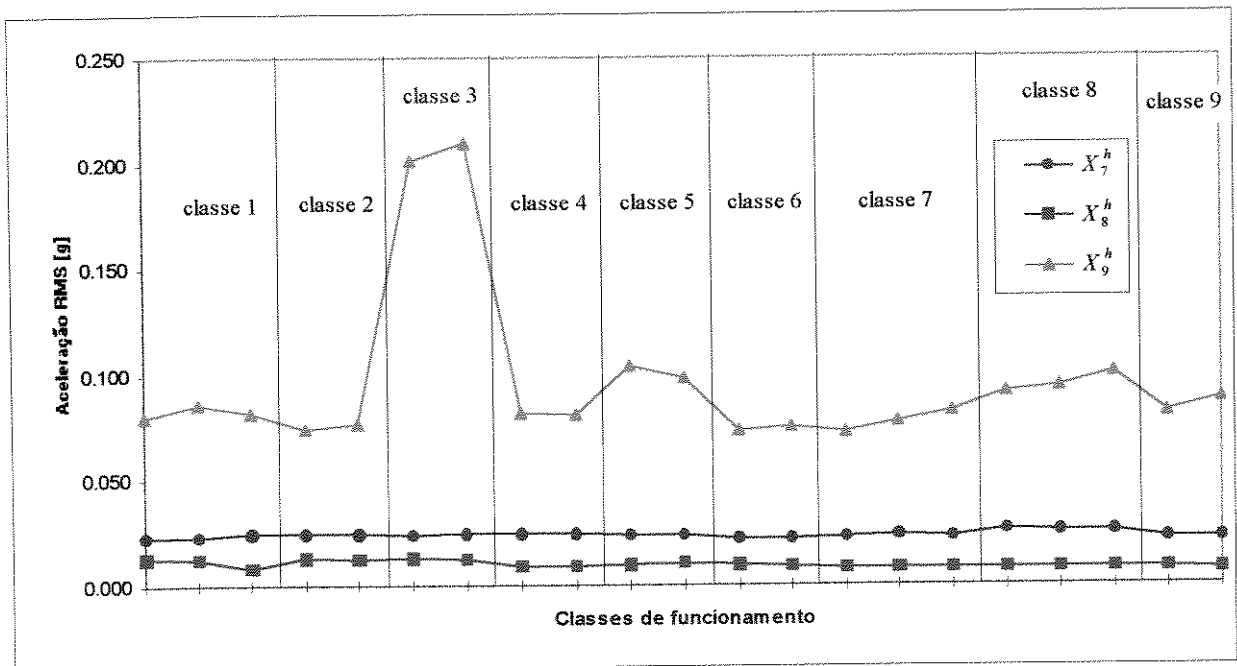


Figura 5.6 - Variação dos índices de monitoramento X_7^h , X_8^h e X_9^h para $f_a = 70$ Hz.

5.3 - Desempenho do classificador

A análise do desempenho do classificador foi realizada através do índice de sensibilidade **ST** (*sensitivity*). Segundo DU *et al.* (1995a), o índice **ST** é definido como:

$$ST = \frac{(ACC)_j}{(NTA)_j} \quad (5.3)$$

onde $(ACC)_j$ é o número de amostras corretamente classificadas na classe j , e $(NTA)_j$ é o número total de amostras da classe j submetidas à classificação. A maioria dos autores usa o índice **SR** (*success rate*) para analisar o desempenho do classificador. O índice **SR** também é definido pela equação (5.3), mas utiliza as mesmas amostras para o treinamento e para a classificação, ou seja, é a taxa de sucesso em um teste de auto-classificação. Neste trabalho, o índice **ST** foi utilizado, pois considera que as amostras usadas na fase de classificação não pertencem ao conjunto das amostras de treinamento, o que torna o processo de análise mais eficaz. Em todos os casos, o primeiro conjunto de medições obtido para cada classe de funcionamento foi considerado como o conjunto de amostras de treinamento e os demais conjuntos foram utilizados na fase de classificação. As Figuras 5.7 e 5.8 mostram diagramas de blocos que esquematizam o procedimento utilizado na classificação das novas amostras. Na primeira figura, a classificação baseia-se na distância Euclideana e na segunda, na distância de *Mahalanobis*.

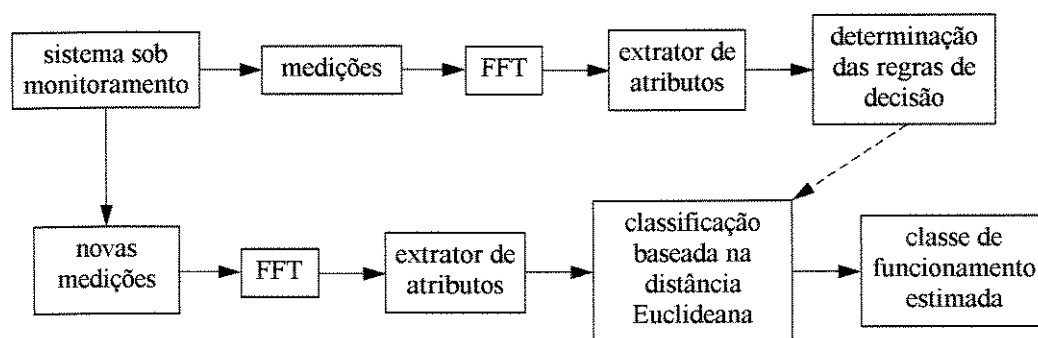


Figura 5.7 - Procedimento de classificação baseado na distância Euclideana.

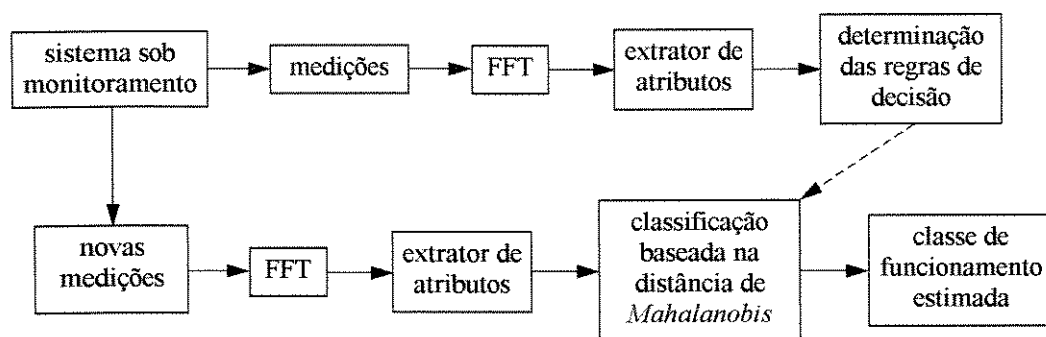


Figura 5.8 - Procedimento de classificação baseado na distância de *Mahalanobis*.

A seguir mostram-se os índices de sensibilidade obtidos nos testes de classificação para as três condições operacionais, ou seja, condições com frequências de alimentação $f_a = 50$ Hz, $f_a = 60$ Hz e $f_a = 70$ Hz, utilizando-se o classificador baseado na distância de *Mahalanobis* e o classificador baseado na distância Euclideana. Dentro de cada condição operacional, os resultados foram apresentados para três casos distintos: a) considerando todos os índices de monitoramento das direções vertical e horizontal (canais 1 e 2); b) considerando todos os índices de monitoramento da direção vertical; e c) considerando todos os índices de monitoramento da direção horizontal. Esta separação em casos b) e c) leva em consideração a redução de equipamentos de medição (condicionadores, filtros e acelerômetros), pois se os desempenhos dos classificadores de *Mahalanobis* e Euclideano forem satisfatórios, o processo de monitoramento poderia ser conduzido baseado somente nas medições realizadas por um único acelerômetro.

5.3.1 - Desempenho dos classificadores para a condição operacional 1 ($f_a = 50$ Hz)

A apresentação dos resultados foi dividida em etapas. Primeiro, nas Tabelas 5.1(a) e 5.1(b) apresentam-se os índices de sensibilidade que foram obtidos, considerando todos os índices de monitoramento (X_1^v , X_1^h , X_2^v , X_2^h , X_3^v e X_3^h). A seguir, as Tabelas 5.2(a) e 5.2(b) apresentam os índices de sensibilidade obtidos considerando os índices de monitoramento da direção vertical (X_1^v , X_2^v e X_3^v) e, as Tabelas 5.3(a) e 5.3(b) os índices de sensibilidade obtidos considerando os índices de monitoramento da direção horizontal (X_1^h , X_2^h e X_3^h). Os índices de sensibilidade apresentados nas Tabelas 5.1(a), 5.2(a) e 5.3(a) foram obtidos com a utilização do classificador baseado na distância de *Mahalanobis*, e os índices de sensibilidade

apresentados nas Tabelas 5.1(b), 5.2(b) e 5.3(b) foram obtidos com a utilização do classificador baseado na distância Euclideana. Nestas tabelas, os índices de sensibilidade estão representados em percentagem e os espaços em branco indicam percentagem nula.

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	90,63		1,56	7,81				
classe 2	34,29	65,71						
classe 3			100,00					
classe 4				77,14	22,86			
classe 5				30,55	66,67	2,78		
classe 6	5,63			2,82		1,41	84,51	5,63
classe 7	1,32		0,66	1,99			96,03	
classe 8							0,68	99,32
classe 9	98,61		1,39					

Tabela 5.1(a) - Índices de sensibilidade obtidos utilizando o classificador de *Mahalanobis* ($f_a = 50$ Hz - direções vertical e horizontal).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	88,28	3,12	4,69	3,91				
classe 2	11,43	88,57						
classe 3			100,00					
classe 4				77,14	22,86			
classe 5		11,11		8,33	80,56			
classe 6					18,31	0,00	50,70	30,99
classe 7		1,99			48,34		49,67	
classe 8							12,84	87,16
classe 9		10,42	53,47		36,11			

Tabela 5.1(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideano ($f_a = 50$ Hz - direções vertical e horizontal).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	46,88	15,63	5,47	1,56	10,94		19,52	
classe 2	20,00	80,00						
classe 3			97,22				2,78	
classe 4				45,71	40,00		14,29	
classe 5	5,56		2,78	36,11	41,67	2,78	11,10	
classe 6	5,63	8,45	1,41		2,82	0,00	71,83	9,86
classe 7	1,32		0,66	0,66	2,66		94,04	0,66
classe 8		0,68					49,32	50,00
classe 9			1,39			98,61		

Tabela 5.2(a) - Índices de sensibilidade obtidos utilizando o classificador de *Mahalanobis* ($f_a = 50$ Hz - direção vertical).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	52,34	23,44	1,56	5,47	10,94		3,91	2,34
classe 2	8,57	91,43						
classe 3			100,00					
classe 4				45,71	54,29			
classe 5			2,78	52,78	44,44			
classe 6	38,03	9,86	1,41		8,45	0,00	22,53	19,72
classe 7	1,99		5,30	8,61	9,27		70,86	3,97
classe 8	0,68	0,68			0,68		33,78	64,18
classe 9			52,08			47,92		

Tabela 5.2(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideano ($f_a = 50$ Hz - direção vertical).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	61,72			38,28				
classe 2	42,86	57,14						
classe 3	5,56		86,11	8,33				
classe 4	2,86			74,28	22,86			
classe 5	5,56			30,56	63,88			
classe 6	2,82			1,41		18,31	61,97	15,49
classe 7				1,32			98,68	
classe 8						3,38	10,14	86,48
classe 9	45,83	18,06		27,78	5,55		2,78	

Tabela 5.3(a) - Índices de sensibilidade obtidos utilizando o classificador de Mahalanobis ($f_a = 50$ Hz - direção horizontal).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	21,88	1,56	67,97	8,59				
classe 2	2,86	88,57		8,57				
classe 3	16,67		72,22	11,11				
classe 4			42,86	34,28	22,86			
classe 5		25,00		5,56	69,44			
classe 6					16,90	2,82	71,83	8,45
classe 7		4,64			47,02		48,34	
classe 8						7,43		92,57
classe 9		38,89		50,69	10,42			

Tabela 5.3(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideano ($f_a = 50$ Hz - direção horizontal).

5.3.2 - Desempenho dos classificadores para a condição operacional 2 ($f_a = 60$ Hz)

Para esta condição operacional, a apresentação dos resultados também foi dividida em etapas. Primeiro, nas Tabelas 5.4(a) e 5.4(b) apresentam-se os índices de sensibilidade que

foram obtidos, considerando todos os índices de monitoramento (X_4^v , X_4^h , X_5^v , X_5^h , X_6^v e X_6^h). A seguir, as Tabelas 5.5(a) e 5.5(b) apresentam os índices de sensibilidade obtidos considerando os índices de monitoramento da direção vertical (X_4^v , X_5^v e X_6^v) e, as Tabelas 5.6(a) e 5.6(b) os índices de sensibilidade obtidos considerando os índices de monitoramento da direção horizontal (X_4^h , X_5^h e X_6^h). Os índices de sensibilidade apresentados nas Tabelas 5.4(a), 5.5(a) e 5.6(a) foram obtidos com a utilização do classificador baseado na distância de *Mahalanobis*, e os índices de sensibilidade apresentados nas Tabelas 5.4(b), 5.5(b) e 5.6(b) foram obtidos com a utilização do classificador baseado na distância Euclideana. Novamente, nas tabelas a seguir, os índices de sensibilidade estão representados em percentagem e os espaços em branco indicam percentagem nula.

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	65,49			30,28	4,23			
classe 2		100,00						
classe 3			100,00					
classe 4	37,84			48,65	13,51			
classe 5	2,86			22,86	74,28			
classe 6						100,00		
classe 7						0,70	95,07	4,23
classe 8								100,00
classe 9						100,00		

Tabela 5.4(a) - Índices de sensibilidade obtidos utilizando o classificador de *Mahalanobis* ($f_a = 60$ Hz - direções vertical e horizontal).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	48,58	7,04		30,28	14,08			
classe 2	30,56	25,00		25,00	19,44			
classe 3			100,00					
classe 4		18,92		70,27	10,81			
classe 5	28,57			17,14	54,29			
classe 6	95,77					4,23		
classe 7						0,70	95,07	99,30
classe 8								100,00
classe 9	0,70					96,48	2,82	

Tabela 5.4(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideano ($f_a = 60$ Hz - direções vertical e horizontal).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	71,83	0,70		23,24	4,23			
classe 2	2,78	97,22						
classe 3			100,00					
classe 4	27,03			70,27	2,70			
classe 5	37,14			8,57	54,29			
classe 6	29,58					70,42		
classe 7	1,41						97,89	0,70
classe 8							20,14	79,86
classe 9			18,31			76,06	5,63	

Tabela 5.5(a) - Índices de sensibilidade obtidos utilizando o classificador de Mahalanobis ($f_a = 60$ Hz - direção vertical).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	34,51	5,63		46,48	13,38			
classe 2		100,00						
classe 3			100,00					
classe 4	13,52			83,78	2,70			
classe 5	34,28			2,86	62,86			
classe 6	25,35		22,54			52,11		
classe 7					1,41		0,00	98,59
classe 8								100,00
classe 9						100,00		

Tabela 5.5(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideano ($f_a = 60$ Hz - direção vertical).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	52,82			22,54	24,64			
classe 2		88,89			11,11			
classe 3			100,00					
classe 4	32,43			51,35	16,22			
classe 5	5,71			20,00	74,29			
classe 6						100,00		
classe 7						1,41	23,24	75,35
classe 8							1,39	98,61
classe 9						64,79		35,21

Tabela 5.6(a) - Índices de sensibilidade obtidos utilizando o classificador de Mahalanobis ($f_a = 60$ Hz - direção horizontal).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	45,77	9,86		26,76	17,61			
classe 2	2,78	8,33		22,22	66,67			
classe 3			100,00					
classe 4		29,73		59,46	10,81			
classe 5	25,71			20,00	54,29			
classe 6	95,77					4,23		
classe 7						0,70	48,60	50,70
classe 8							48,61	51,39
classe 9	2,11					85,92		11,97

Tabela 5.6(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideano ($f_a = 60$ Hz - direção horizontal).

5.3.3 - Desempenho dos classificadores para a condição operacional 3 ($f_a = 70$ Hz)

Para esta condição operacional, a apresentação dos resultados também foi dividida em etapas. Primeiro, nas Tabelas 5.7(a) e 5.7(b) apresentam-se os índices de sensibilidade que foram obtidos, considerando todos os índices de monitoramento (X_7^v , X_7^h , X_8^v , X_8^h , X_9^v e X_9^h). A seguir, as Tabelas 5.8(a) e 5.8(b) apresentam os índices de sensibilidade obtidos considerando os índices de monitoramento da direção vertical (X_7^v , X_8^v e X_9^v) e, as Tabelas 5.9(a) e 5.9(b) os índices de sensibilidade obtidos considerando os índices de monitoramento da direção horizontal (X_7^h , X_8^h e X_9^h). Os índices de sensibilidade apresentados nas Tabelas 5.7(a), 5.8(a) e 5.9(a) foram obtidos com a utilização do classificador baseado na distância de *Mahalanobis*, e os índices de sensibilidade apresentados nas Tabelas 5.7(b), 5.8(b) e 5.9(b) foram obtidos com a utilização do classificador baseado na distância Euclideana. Novamente, nas tabelas a seguir, os índices de sensibilidade estão representados em percentagem e os espaços em branco indicam percentagem nula.

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	76,52			23,48				
classe 2		100,00						
classe 3			100,00					
classe 4				100,00				
classe 5	5,71				94,29			
classe 6						100,00		
classe 7							100,00	
classe 8	0,70							99,30
classe 9	100,00							

Tabela 5.7(a) - Índices de sensibilidade obtidos utilizando o classificador de Mahalanobis ($f_a = 70$ Hz - direções vertical e horizontal).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	43,94			53,03				3,03
classe 2		75,00		25,00				
classe 3			100,00					
classe 4				100,00				
classe 5					100,00			
classe 6						100,00		
classe 7							100,00	
classe 8				0,70	25,17			74,13
classe 9				64,08				35,92

Tabela 5.7(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideano ($f_a = 70$ Hz - direções vertical e horizontal).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	46,21	27,27		22,73				3,79
classe 2		80,56		19,44				
classe 3			100,00					
classe 4		27,03		67,57				5,40
classe 5	22,86				77,14			
classe 6						100,00		
classe 7						0,68	99,32	
classe 8	0,70	2,80		13,98				82,52
classe 9	3,52	7,75		2,11		0,70	85,92	

Tabela 5.8(a) - Índices de sensibilidade obtidos utilizando o classificador de Mahalanobis ($f_a = 70$ Hz - direção vertical).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	46,21	15,15		34,85				3,79
classe 2		77,78		22,22				
classe 3			100,00					
classe 4		13,51		78,38				8,11
classe 5					100,00			
classe 6						100,00		
classe 7							100,00	
classe 8		1,40		27,27				71,33
classe 9		38,74		23,94				37,32

Tabela 5.8(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideano ($f_a = 70$ Hz - direção vertical).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	46,21			23,48		0,76	28,79	0,76
classe 2		100,00						
classe 3			100,00					
classe 4				100,00				
classe 5					100,00			
classe 6						76,39	23,61	
classe 7						50,68	49,32	
classe 8					0,70	0,70		98,60
classe 9						100,00		

Tabela 5.9(a) - Índices de sensibilidade obtidos utilizando o classificador de Mahalanobis ($f_a = 70$ Hz - direção horizontal).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	12,12			76,52		0,76	2,27	8,33
classe 2	44,44	55,56						
classe 3			100,00					
classe 4				97,30			2,70	
classe 5					74,29			25,71
classe 6				33,33		19,45	47,22	
classe 7				76,03		0,68	22,61	0,68
classe 8				0,70	45,45			53,85
classe 9				60,57		1,41	0,70	37,32

Tabela 5.9(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideano ($f_a = 70$ Hz - direção horizontal).

5.3.4 - Desempenho global dos classificadores

A Tabela 5.10 mostra, de uma forma comparativa, os índices de sensibilidade globais obtidos utilizando-se o classificador baseado na distância de *Mahalanobis* e o classificador baseado na distância Euclideana.

Condição operacional 1 ($f_a = 50$ Hz)		
índices de monitoramento utilizados	ST (%) classificador <i>Mahalanobis</i>	ST (%) classificador Euclideano
$X_1^v, X_1^h, X_2^v, X_2^h, X_3^v$ e X_3^h	81,1	68,8
X_1^v, X_2^v e X_3^v	57,8	57,7
X_1^h, X_2^h e X_3^h	73,3	52,2
Condição operacional 2 ($f_a = 60$ Hz)		
índices de monitoramento utilizados	ST (%) classificador <i>Mahalanobis</i>	ST (%) classificador Euclideano
$X_4^v, X_4^h, X_5^v, X_5^h, X_6^v$ e X_6^h	87	47,7
X_4^v, X_5^v e X_6^v	81,2	55,3
X_4^h, X_5^h e X_6^h	67,5	45,3
Condição operacional 3 ($f_a = 70$ Hz)		
índices de monitoramento utilizados	ST (%) classificador <i>Mahalanobis</i>	ST (%) classificador Euclideano
$X_7^v, X_7^h, X_8^v, X_8^h, X_9^v$ e X_9^h	94,7	81,1
X_7^v, X_8^v e X_9^v	80,6	80
X_7^h, X_8^h e X_9^h	74,3	40,6

Tabela 5.10 - Comparação dos índices de sensibilidade globais obtidos utilizando o classificador de *Mahalanobis* e o Euclideano.

5.4 - Desempenho dos métodos com análise independente dos índices de monitoramento

A recente facilidade de adquirir dados e informações dos processos fez com que um grande número de fabricantes desenvolvessem sistemas de auxílio à manutenção (SCUDELLER Jr., 1993; APPLICATION..., s. d.; VIBROCAM..., s. d.). Estes *softwares* comerciais de diagnóstico analisam os índices de monitoramento de uma forma independente, ou seja, associam os índices de monitoramento relacionados, por exemplo, a 1xRPM, 2xRPM ou n xRPM, a um determinado tipo de defeito. Vários autores também realizam o monitoramento e o diagnóstico de equipamentos baseado somente na variação de índices de monitoramento relacionados a 1xRPM, 2xRPM ou n xRPM (GLATT, 1995; DIANA *et al.*, 1994; ERICSSON, 1985; ZIMMER e BENTLY, 1985; BROWN e JORGENSEN, s. d.).

Para a verificação do comportamento dos métodos de diagnóstico apresentados no capítulo 3, quando analisa-se somente os índices de monitoramento relacionados a 4xRPM, foram calculados seus desempenhos para a condição operacional 3 ($f_a = 70$ Hz). Esta condição operacional foi escolhida, pois foi a que apresentou índices de sensibilidade mais elevados, como pode ser observado na Tabela 5.10. As Tabelas 5.11(a) e 5.11(b) apresentam os índices de sensibilidade que foram obtidos, considerando os índices de monitoramento X_g^v e X_g^h . As Tabelas 5.12(a) e 5.12(b) apresentam os índices de sensibilidade obtidos considerando o índice de monitoramento da direção vertical X_g^v e, as Tabelas 5.13(a) e 5.13(b) os índices de sensibilidade obtidos considerando o índice de monitoramento da direção horizontal X_g^h . Os índices de sensibilidade apresentados nas Tabelas 5.11(a), 5.12(a) e 5.13(a) foram obtidos com a utilização do classificador baseado na distância de *Mahalanobis*, e os índices de sensibilidade apresentados nas Tabelas 5.11(b), 5.12(b) e 5.13(b) foram obtidos com a utilização do classificador baseado na distância Euclideana. Nestas tabelas, os índices de sensibilidade estão representados em percentagem e os espaços em branco indicam percentagem nula.

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	43,94	5,30		47,73		3,03		
classe 2		55,56		30,56			13,88	
classe 3			100,00					
classe 4		2,70		94,60			2,70	
classe 5					100,00			
classe 6	43,06					56,94		
classe 7	2,05	15,07		71,92		0,69	10,27	
classe 8								100,00

Tabela 5.11(a) - Índices de sensibilidade obtidos utilizando o classificador de *Mahalanobis* ($f_a = 70$ Hz - direções vertical e horizontal - índices de monitoramento X_g^v e X_g^h).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	42,42	6,82		46,21		0,76		3,79
classe 2		55,56		36,11			8,33	
classe 3			100,00					
classe 4		2,70		97,30				
classe 5					100,00			
classe 6	37,50					62,50		
classe 7		26,72		69,18		0,68	2,74	0,68
classe 8				0,70	25,87			73,43

Tabela 5.11(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideano ($f_a = 70$ Hz - direções vertical e horizontal - índices de monitoramento X_g^v e X_g^h).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	44,70	27,27		21,96		1,52	0,76	3,79
classe 2		50,00		30,56			2,77	16,67
classe 3			100,00					
classe 4		16,22		40,54			8,11	35,13
classe 5	11,43				88,57			
classe 6	54,17					45,83		
classe 7		44,52		18,49		2,05	3,43	31,51
classe 8		34,27		27,97			9,79	27,97

Tabela 5.12(a) - Índices de sensibilidade obtidos utilizando o classificador de Mahalanobis ($f_a = 70$ Hz - direção vertical - índice de monitoramento X_9^v).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	42,42	23,48		27,27	0,76	3,03	0,76	2,28
classe 2		44,44		41,67			2,78	11,11
classe 3			100,00					
classe 4		13,51		45,95			8,11	32,43
classe 5					100,00			
classe 6	40,28					59,72		
classe 7		44,52		23,29		2,05	3,42	26,72
classe 8		32,87		30,77			9,79	26,57

Tabela 5.12(b) - Índices de sensibilidade obtidos utilizando o classificador Euclidiano ($f_a = 70$ Hz - direção vertical - índice de monitoramento X_9^v).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	19,70			48,48		3,79		28,03
classe 2	36,10	30,56		5,56		25,00	2,78	
classe 3			100,00					
classe 4	59,46			37,84		2,70		
classe 5					37,14			62,86
classe 6	29,17	15,28		4,17		50,00	1,38	
classe 7	41,10	4,11		26,03		21,23	0,68	6,85
classe 8					35,66			64,34

Tabela 5.13(a) - Índices de sensibilidade obtidos utilizando o classificador de Mahalanobis ($f_a = 70$ Hz - direção horizontal - índice de monitoramento X_9^h).

amostras a classificar	Padrões obtidos na fase de treinamento							
	classe 1	classe 2	classe 3	classe 4	classe 5	classe 6	classe 7	classe 8
classe 1	19,70	2,27		62,88			0,76	14,39
classe 2	44,44	47,22		5,56			2,78	
classe 3			100,00					
classe 4	54,06	2,70		43,24				
classe 5					62,86			37,14
classe 6	27,78	22,22		5,56		5,56	38,88	
classe 7	46,58	19,18		32,19			1,37	0,68
classe 8				0,70	47,55			51,75

Tabela 5.13(b) - Índices de sensibilidade obtidos utilizando o classificador Euclideano ($f_a = 70$ Hz - direção horizontal - índice de monitoramento X_9^h).

A Tabela 5.14 mostra, de uma forma comparativa, os índices de sensibilidade globais obtidos utilizando-se o classificador de *Mahalanobis* e o Euclideano considerando apenas os índices de monitoramento relacionados a 4xRPM.

Condição operacional 3 ($f_a = 70$ Hz)		
índices de monitoramento	ST (%) <i>Mahalanobis</i>	ST (%) Euclideano
X_9^v e X_9^h	60,2	53
X_9^v	37,3	38,7
X_9^h	36,1	31

Tabela 5.14 - Comparação dos índices de sensibilidade globais obtidos utilizando o classificador de *Mahalanobis* e o Euclideano (índices de monitoramento relacionados a 4xRPM).

5.5 - Análise crítica dos resultados

Analisando as Tabelas 5.1(a) a 5.9(b), em alguns casos, nota-se que os índices de sensibilidade obtidos foram baixos. Tomando, como exemplo, os resultados obtidos na classificação das amostras pertencentes à classe 6 mostrados nas Tabelas 5.3(a) e 5.3(b). Estas amostras foram, em sua maioria, classificadas como pertencentes à classe 7. Isto aconteceu, provavelmente, devido a uma má definição dos padrões da classe 6, como pode ser visto na Figura 5.2. Nesta mesma figura, percebe-se também que a média das amostras da classe 6 a classificar aproxima-se dos padrões da classe 7, fazendo com que o resultado desta classificação ficasse prejudicado. Pelos resultados, nota-se uma má definição dos padrões das seguintes classes: classe 6 na condição operacional $f_a = 50$ Hz; classes 6 e 7 na condição operacional $f_a = 60$ Hz e, classes 1, 6 e 7 na condição operacional $f_a = 70$ Hz.

Tomando agora, como outro exemplo, os resultados obtidos na classificação das amostras pertencentes à classe 1 mostrados nas Tabelas 5.4(a) e 5.4(b). Algumas amostras pertencentes à classe 1 foram classificadas como pertencentes à classe 4. Isto ocorreu devido a semelhança existente entre as classes 1 e 4, como pode ser visto nas Figuras 5.3 e 5.4. Na classe de funcionamento 1, as válvulas globo das tubulações de sucção e recalque estão totalmente abertas. Já na classe de funcionamento 4, a válvula da tubulação de recalque está totalmente aberta e a de sucção está com 26% de obstrução. Este pequeno fechamento da válvula da tubulação de sucção acarreta em uma pequena variação na pressão de entrada da bomba, variação esta pouco suficiente para alterar o comportamento vibratório do conjunto moto-bomba. Esta semelhança entre as classes 1 e 4 acarretou em uma diminuição do desempenho dos métodos, como mostrado nas seguintes tabelas: 5.3(a), 5.3(b), 5.5(a), 5.5(b), 5.6(a), 5.6(b), 5.7(a), 5.7(b), 5.8(a), 5.8(b), 5.9(a) e 5.9(b).

Um outro exemplo do baixo desempenho dos classificadores pode ser apresentado com os resultados obtidos na classificação do segundo conjunto de amostras pertencentes à classe 7 mostrado na Tabela 5.3(b). Nota-se que a não repetibilidade das medições deste conjunto de amostras fez com que ele fosse classificado na classe 5 (Figura 5.2). Esta classificação errônea devido a esta pequena diferença entre os dados medidos prejudicou o desempenho do método. Já o terceiro conjunto de amostras da classe 7 foi classificado corretamente em sua classe.

Outro ponto importante a ser analisado é a classificação correta ou não dos padrões da classe 9 em relação aos da classe 1. Somente dois casos conseguiram classificar corretamente os padrões da classe 9 (Tabelas 5.1(a) e 5.7(a)). Apesar desta classificação correta, nota-se nas Figuras 5.1, 5.2, 5.5 e 5.6 que estes padrões da classe 9 não apresentam uma grande semelhança com os padrões da classe 1. Desta forma, para os casos estudados, os métodos de diagnóstico não conseguiram classificar corretamente os padrões da classe 9 em relação aos da classe 1 e quando conseguiram os padrões não eram suficientemente representativos.

O processo de classificação de novas amostras poderia ser conduzido somente com base nas medições realizadas por um único acelerômetro, tendo em vista que os índices de sensibilidade globais obtidos, utilizando a distância de *Mahalanobis*, apresentaram valores satisfatórios. Considerando um processo onde o motor elétrico trabalha em uma única

condição operacional ($f_a = 50$ Hz, $f_a = 60$ Hz ou $f_a = 70$ Hz), a classificação pode ser conduzida da seguinte forma: 1) para a condição operacional $f_a = 50$ Hz: um acelerômetro na direção horizontal, 2) para a condição operacional $f_a = 60$ Hz: um acelerômetro na direção vertical e 3) para a condição operacional $f_a = 70$ Hz: um acelerômetro na direção vertical. Se o motor elétrico trabalha com uma variação constante destas condições operacionais (rotações), então existe a necessidade do uso de acelerômetros nas duas direções, pois a direção que apresentou índices de sensibilidade mais elevados não foi a mesma para todas as condições operacionais.

Neste trabalho os métodos de diagnóstico utilizaram todo o conjunto dos índices de monitoramento para o cálculo das distâncias. Os índices de sensibilidade obtidos, utilizando-se índices de monitoramento relacionados a 1xRPM, 2xRPM e 4xRPM de uma forma conjunta, chegaram até aproximadamente 95% para, por exemplo, a condição operacional 3. Como esta condição operacional foi a que apresentou os índices de sensibilidade mais elevados, ela foi escolhida para verificar o comportamento dos métodos de diagnóstico apresentados, quando se analisou somente os índices de monitoramento relacionados a 4xRPM. Neste caso, os índices de sensibilidade globais obtidos para esta condição operacional não foram superiores a 61% (Tabela 5.14). Isto mostra a importância de se analisar conjuntamente os índices de monitoramento relacionados às diferentes frequências de interesse.

Para exemplificar uma situação em que a classificação de uma nova amostra fosse realizada de forma geométrica ou visual, escolheu-se a condição operacional 2 ($f_a = 60$ Hz) e dois índices de monitoramento X_4^h e X_6^h . Observando a Figura 5.9 nota-se que se a classificação tivesse sido conduzida visualmente ou baseada na distância Euclideana, o resultado seria incorreto. Já a Figura 5.10 mostra que o método baseado na distância de *Mahalanobis* conseguiu classificar corretamente a amostra. Isto demonstra que o procedimento geométrico ou visual, às vezes, não apresenta bons resultados.

O método baseado na distância de *Mahalanobis*, ao contrário do baseado na distância Euclideana, por considerar as propriedades estatísticas μ e Σ das amostras utilizadas na fase de treinamento apresentou, de uma forma geral, melhores resultados que o outro método.

Entretanto, o método baseado na distância Euclidiana apresentou, em alguns casos, melhores resultados em relação ao método baseado na distância de *Mahalanobis*. Estes resultados foram obtidos nos testes de classificação classe a classe realizados com os índices de monitoramento da direção vertical.

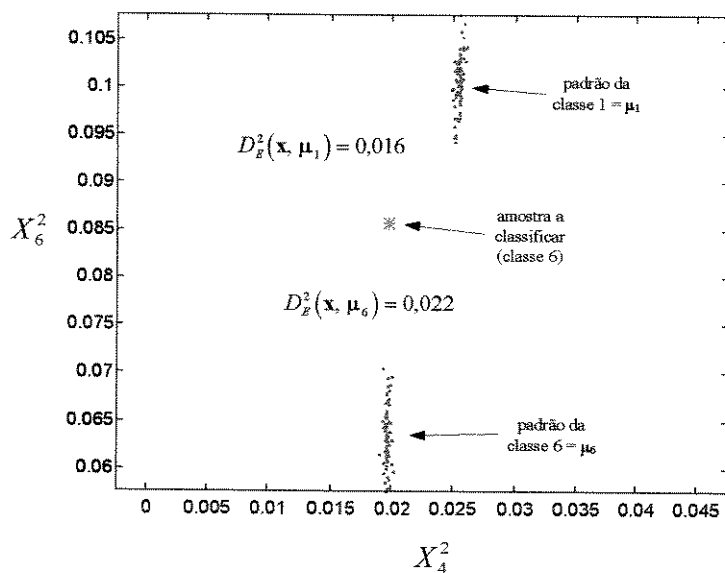


Figura 5.9 - Classificação baseada na distância Euclidiana utilizando-se dois índices de monitoramento.

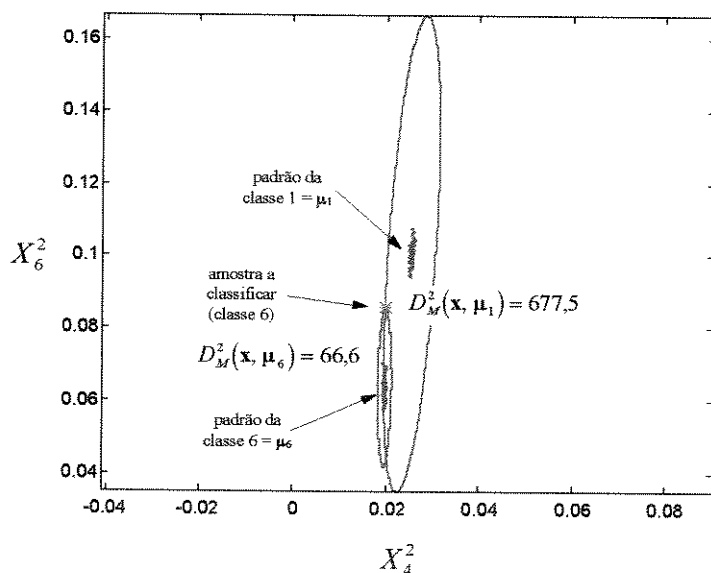


Figura 5.10 - Classificação baseada na distância de *Mahalanobis* utilizando-se dois índices de monitoramento.

Capítulo 6

Conclusões e Sugestões

6.1 - Conclusões

Vários métodos de diagnóstico baseados em sintomas, fundamentados no processo de comparação e/ou classificação de padrões, foram apresentados no capítulo 2. Estes métodos utilizam as teorias recentes de sistemas fuzzy, árvores de decisão, sistemas especialistas e redes neurais neste processo de comparação/classificação de padrões. Estes métodos demandam uma carga computacional elevada, exigem um esforço elevado na fase de treinamento e incluem procedimentos complexos que inviabilizam seu uso no diagnóstico de equipamentos de menor responsabilidade e custo. Entretanto, para equipamentos sofisticados e de grande importância no processo produtivo, onde diferentes defeitos podem exibir sintomas similares e estes podem ocorrer simultaneamente, é mais vantajoso empregar técnicas de diagnóstico baseadas, por exemplo, em lógica fuzzy, uma vez que a relação Q entre as classes de funcionamento e os índices de monitoramento pode não ser determinada de forma única e neste caso, o custo de sua implementação se justifica.

Para o diagnóstico de equipamentos de baixo custo ou de menor responsabilidade dentro do processo produtivo, onde a relação entre a importância do equipamento e o custo de implementação de técnicas avançadas é baixa, a utilização das teorias acima citadas não se justifica. Também, a relação Q entre as classes de funcionamento e os índices de monitoramento de ser determinada de forma única, pois espera-se que a maioria destes equipamentos sejam simples. Se esta relação Q estiver bem definida, métodos de classificação baseados em distâncias estatísticas são empregados com sucesso. O presente trabalho

contribuiu neste ponto analisando dois métodos de classificação de padrões alternativos, um baseado na distância Euclideana e outro na distância de *Mahalanobis*, que exigiram um menor esforço na fase de treinamento, foram facilmente implementados e apresentaram resultados satisfatórios no diagnóstico do equipamento estudado. A descrição destes métodos foi feita no capítulo 3.

O objetivo principal do trabalho foi o de analisar os dois métodos de diagnóstico acima citados. Para isso, desenvolveu-se um trabalho experimental em um conjunto moto-bomba em três diferentes condições operacionais simulando-se alguns tipos de defeitos, tais como, desbalanceamento, obstruções nas tubulações de sucção e recalque, e modificações geométricas nas pás do rotor. A montagem e o procedimento experimental foram descritos no capítulo 4. A análise dos métodos foi realizada através de seus desempenhos na classificação de amostras não utilizadas na fase de treinamento. O cálculo do desempenho de cada método foi realizado através do índice de sensibilidade *ST* (*sensitivity*).

Em alguns casos notou-se que os índices de sensibilidade obtidos foram pouco satisfatórios devido, provavelmente, a três fatores: 1) má definição dos padrões das classes de funcionamento; 2) semelhança existente entre as classes de funcionamento e 3) não repetibilidade das medições para alguns casos isolados. Estes três fatores foram, com grande chance, uma consequência do fato das amostras de treinamento não satisfazerem os seguintes requisitos: 1) as amostras de treinamento deviam ser representativas da aplicação, uma vez que a inclusão de amostras distorcidas levam à geração de regras de classificação sub-ótimas e 2) o conjunto de amostras de treinamento deveria ser o maior possível com a finalidade de prover a cobertura máxima de padrões da mesma classe.

A classificação correta dos padrões da classe 9 em relação aos da classe 1, que correspondia à mesma condição após uma operação de manutenção, foi observada somente em dois casos. Apesar desta classificação correta, notou-se que estes padrões da classe 9 não apresentaram uma grande semelhança com os padrões da classe 1. Dessa maneira, pode-se concluir que o equipamento sofreu pequenas alterações no seu comportamento dinâmico e assim, a utilização dos mesmos padrões de defeitos poderá ser comprometida.

Ao considerar um processo onde o motor elétrico opera em uma única condição operacional ($f_a = 50$ Hz, $f_a = 60$ Hz ou $f_a = 70$ Hz), o processo de monitoramento poderia ser conduzido somente com base nas medições realizadas por um único acelerômetro, uma vez que os índices de sensibilidade obtidos, utilizando a distância de *Mahalanobis*, foram superiores a 70%. A classificação de novas amostras poderia ser conduzida da seguinte forma: 1) para a condição operacional $f_a = 50$ Hz: um acelerômetro na direção horizontal, 2) para a condição operacional $f_a = 60$ Hz: um acelerômetro na direção vertical e 3) para a condição operacional $f_a = 70$ Hz: um acelerômetro na direção vertical. Entretanto, se estas três condições operacionais variarem constantemente, então existe a necessidade do uso de acelerômetros nas duas direções, pois a direção escolhida para a realização das medições não foi a mesma para todas as condições operacionais.

Na verificação do comportamento dos métodos, quando se analisou somente os índices de monitoramento relacionados a 4xRPM, notou-se que os índices de sensibilidade globais obtidos para a condição operacional 3 ($f_a = 70$ Hz) não foram superiores a 61%. Neste trabalho os métodos de diagnóstico utilizaram todo o conjunto dos índices de monitoramento para o cálculo das distâncias entre os padrões. Os índices de sensibilidade obtidos, utilizando-se índices de monitoramento relacionados a 1xRPM, 2xRPM e 4xRPM de uma forma conjunta, chegaram até aproximadamente 95% para a condição operacional 3. Isto mostrou a importância de se analisar conjuntamente os índices de monitoramento relacionados às frequências de interesse.

O método de diagnóstico baseado na distância de *Mahalanobis* apresentou, de uma forma geral, melhores resultados do que o método baseado na distância Euclideana. Estes resultados foram melhores, pois, ao contrário do método baseado na distância Euclideana, o método de *Mahalanobis* considera as propriedades estatísticas μ e Σ das amostras utilizadas na fase de treinamento. Entretanto, o método baseado na distância Euclideana apresentou, para alguns casos, melhores resultados em relação ao método baseado na distância de *Mahalanobis*. Estes resultados foram obtidos nos testes de classificação classe a classe realizados com os índices de monitoramento da direção vertical. Isto justifica o fato de que, apesar de simples, a distância Euclideana ainda é a medida de distância mais utilizada.

Através dos dados coletados, para algumas classes de funcionamento específicas do conjunto moto-bomba, foi possível a obtenção de padrões de funcionamento deste sistema. Como o procedimento é experimental, se o sistema estiver operando em uma classe de funcionamento diferente das classes padrões analisadas, o resultado da classificação fica altamente prejudicado. Em outras palavras, se o sistema estiver operando em uma classe de funcionamento diferente em relação àquelas que já foram vivenciadas (padrões), esta nova classe de funcionamento será estimada com base nos padrões obtidos anteriormente, ou seja, o resultado poderá não ser o correto.

O processamento dos sinais deve ser direcionado à obtenção dos índices de monitoramento, sendo esta obtenção decisiva para o sucesso do diagnóstico. Portanto, os índices devem representar e distinguir, da melhor maneira possível, cada classe de funcionamento do equipamento.

6.2 - Sugestões para trabalhos futuros

Os métodos analisados neste trabalho apresentaram dificuldades na classificação de amostras pertencentes a classes de funcionamento com comportamento vibratório semelhantes. Para estes casos, o diagnóstico não é apenas em função da condição do processo, mas também em função do degrau que o processo está nesta condição. Isto pode ser melhor respondido por sistemas fuzzy. A vantagem do processamento fuzzy é que ele carrega a incerteza das variáveis junto com os cálculos para a determinação das classes de funcionamento mais prováveis. Ainda, quando a relação entre as classes de funcionamento e os índices de monitoramento não pode ser determinada de forma única, as técnicas fuzzy simulam a lógica humana multi-valores de pensamento, uma vez que nós relacionamos simples sintomas com várias causas.

No nosso ensaio experimental estudamos apenas alguns defeitos relacionados com a bomba. Se o estudo fosse realizado considerando, por exemplo, defeitos na bomba, no motor elétrico e nos rolamentos, teríamos que utilizar um número maior de índices de monitoramento para englobarmos toda esta gama de defeitos. Neste caso, um índice de monitoramento deve ser particularmente sensível a uma certo defeito do equipamento. Isto indica que os índices de

monitoramento podem ser melhores utilizados através da representação das classes de funcionamento na forma de árvores de decisão.

Referências Bibliográficas

- ALEKSANDER, I., MANDANI, E. H. Microcircuit learning net: improved recognition by means of a pattern feedback. *Electronics Letters*, vol.4, n.20, p.425-426, 1968.
- ANDREWS, H. C. *Introduction to mathematical techniques in pattern recognition*. USA: John Wiley & Sons, 1972, 242p.
- APPLICATION software for machine condition monitoring with the vibration analyzer type 2515. Denmark: Brüel & Kjær, s. d., 4p. (Product data type 7616).
- BARBORAK, D. M. *et al.* PC - Based expert system and their application to welding. *Welding*, p.29-38, Jan 1991.
- BASSEVILLE, M. Detecting changes in signals and systems - a survey. *Automatica*, vol.24, n.3, p.309-326, 1988.
- BEZDEK, J. C. *Pattern recognition with fuzzy objective function algorithm*. New York: Plenum Press, 1981.
- BITTERMANN, HANS-JÜRGEN Avoid downtime in centrifugal pumps. *PACE Process and Control Engineering*, vol.47, n.1, p.20-22, 1994.
- BRAWLEY, G. H. Diagnostic health condition performance monitoring - does this make sense?. In: *Proceedings of the First International Machinery Monitoring & Diagnostics Conference*, p.394-401, Las Vegas, Nevada, USA, Setembro 1989.
- BROWN, D. N., JORGENSEN, J. C. *Machine-condition monitoring using vibration analysis: a case study from a petrochemical plant*. Denmark: Brüel & Kjær, s. d., 12p. (Application notes).
- CARPENTER, G. A., GROSSBERG, S. A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics and Image Processing*, vol.37, p.54-115, 1987.
- CARPENTER, G. A., GROSSBERG, S. ART3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures. *Neural Networks*, vol.3, n.2, p.129-152, 1990.

- CARTER, C., CATLETT, J., BUDA, R. Testing continuous attributes in class probability trees. *Artificial Intelligence Developments and Applications*, GERO, J. S., STANTON, R., eds., Elsevier Science Publishers, B. V., p.291-300, 1988.
- CHEN, Y., WU, S. M., Machinery condition monitoring by prediction error analysis. *Intelligent and Integrated Manufacturing Analysis and Synthesis - ASME PED*, vol.25, p.129-140, 1987.
- CHEN, Y. D., DU, R., QU, L. S. Fault rotating machinery and diagnosis using sensor fusion. *Journal of Sound and Vibration*, vol.118, n.2, p.227-242, 1995.
- DANAI, K., ULSOY, A. G. An adaptive observer for on-line tool wear estimation in turning - Part I: Theory. *Mechanical Systems and Signal Processing*, vol.1, n.2, p.211-225, 1987.
- DIANA, G., VANIA, A., VALLINI, A., ZANETTA, G. A. Diagnostic techniques in condition monitoring. In: Bachschmid, N. *Rotordynamics and Diagnostics of Rotating Machinery*. Campinas: Unicamp, 1994, p.A1-A13. (Apostila de curso)
- DU, R., ELBESTAWI, M. A., LI, S. Tool condition monitoring in turning using fuzzy set theory. *International Journal of Machine Tools Manufacturing*, vol.32, n.6, p.781-796, 1992.
- DU, R., ELBESTAWI, M. A., WU, S. M. Automated monitoring of manufacturing processes, Part 1 : Monitoring methods. *ASME Journal of Engineering for Industry*, vol.117, n.2, p.121-132, 1995a.
- DU, R., ELBESTAWI, M. A., WU, S. M. Automated monitoring of manufacturing processes, Part 2: Applications. *ASME Journal of Engineering for Industry*, vol.117, n.2, p.133-141, 1995b.
- DUDA, R. O., HART, P. E. *Pattern classification and scene analysis*. Reading: John Wiley & Sons, 1973. Cap.2: Bayes Decision Theory, p.10-43.
- ELBESTAWI, M. A., MARKS, J., PAPAZAFIRIOU, T. Process monitoring in milling by pattern recognition. *Mechanical Systems and Signal Processing*, vol.3, n.3, p.305-315, 1989.
- ERICSSON, U. Vibration monitoring - the state of the art - with case studies. In: *VDI Conference on Schwingungsüberwachung*, Hamburg, Germany, VDI Berichte 568, p.289-304, 1985.
- FU, K. S. Introduction to pattern recognition. In: FU, K. S., ICHIKAWA, T. *Special computer architectures for pattern processing*, CRC Press Inc., p.36-63, 1982.
- GAMS, M., DROBNIC, M., PETKOVSEK, M. Learning from examples - a uniform view. *International Journal of Man-Machine Studies*, vol.34, p.49-68, 1991.

- GAREY, M. R., JOHNSON, D. S. *Computers and Intractability - a guide to the theory of NP-completeness*. Freeman and Company, 1979.
- GLATT, P. Two-level vibration strategy for maintenance. *Noise & Vibration Worldwide*, p.7-9, October, 1995.
- GOMIDE, F. A. C., GUDWIN, R. R. Modelagem, controle, sistemas e lógica fuzzy. *Controle & Automação*, vol.4, n.3, p.97-115, 1994.
- GROSSBERG, S. Adaptive pattern classification and universal recording: feedback, expectation, olfaction, illusions. *Biological Cybernetics*, vol.23, p.187-202, 1976.
- HOLLAN, J. D., HUTCHINS, E. L., WEITZMAN, L. STEAMER: an interactive inspectable simulation-based training system. *The AI Magazine*, vol.5, n.2, 1984.
- HOUSHMAND, A. A., KANNATEY-ASIBU, Jr. E. Statistical process control of acoustic emission for cutting tool monitoring. *Mechanical Systems and Signal Processing*, vol.3, n.4, p.405-424, 1989.
- ISERMANN, R. Fault diagnosis of machines via parameter estimation and knowledge processing - Tutorial paper. *Automatica*, vol.29, n.4, p.815-835, 1993.
- ISERMANN, R. Model based fault detection and diagnosis methods. In: *Proceedings of the American Control Conference*, p.1605-1609, Seattle, Washington, June, 1995.
- ISHIZUKA, M., FU, K. S., YAO, J. T. P. Inexact inference for rule-based damage assessment of existing structures. *Seventh International Joint Conference on Artificial Intelligence*, Vancouver, August, 1981.
- JONES, W. P., HOSKINS, J. Backpropagation - a generalized delta learning rule. *BYTE*, October, 1987.
- KARNA, K. N., BREEN, D. M. An artificial neural networks tutorial - Part1: basics. *Neural Networks*, vol.1, n.1, 1989.
- KIM, D. S., SHIN, Y. S., CARLSON, D. K. Machinery diagnostics for rotating machinery using backpropagation neural networks. *Proceedings of the Third Conference IMMDC*, Las Vegas, USA, 1991.
- KLIR, J. G., FOLGER, A. T. *Fuzzy Sets, Uncertainty, and Information*. Prentice-Hall, 1988.
- KOHONEN, T., MAKISARA, M., SARAMAKI, T., Phonotopic maps - insightful representations of phonological for speech recognition. *Proceedings of the 7th IEEE International Conference on Pattern Recognition*, August 1984.
- KOHONEN, T., Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, p.59-69, 1982.

- KOWALSKI, A. *et al.* PITCH Expert - A problem-solving system for kraft mills. *The AI Magazine*, vol.14, n.3, p.81-99, 1993.
- LETA, F. R., SOARES, A. N. D. Uso de redes neurais para o reconhecimento de perfis. *Anais do IV Congresso de Engenharia Mecânica Norte-Nordeste*, vol.1, p.505-510, Recife, Junho, 1996.
- LI, P. G., WU, S. M. Monitoring of drill wear states using fuzzy pattern recognition technique. *ASME Journal of Engineering for Industry*, vol.110, n3, p.297-300, 1988.
- LI, S., ELBESTAWI, M. A., DU, R. A fuzzy logic approach for multi-sensor process monitoring in machining. *ASME PED vol.55, Sensor and Signal Processing for Manufacturing*, p.1-16, 1992.
- LIAO, T. W., CHEN, L. J. A neural network approach for grinding processes: modelling and optimization. *International Journal of Machine Tools Manufacturing*, vol.34, n.7, p.919-937, 1994.
- LIPPMANN, R. P. An introduction of computing with neural nets. *IEEE ASSP Magazine*, vol.4, p.4-22, 1987.
- LOUKIS, E., MATHIOUDAKIS, K. & PAPILIOU, K. A procedure for automated gas turbine blade fault identification based on spectral pattern analysis. *ASME Journal of Engineering for Gas Turbines and Power*, vol.113, p.201-208, 1992.
- LOUKIS, E., MATHIOUDAKIS, K., PAPILIOU, K. Optimizing automated gas turbine fault detection using statistical pattern recognition. *Journal of Engineering for Gas Turbines and Power*, vol.116, p.165-171, 1994.
- MARCUS, S. J. Computer systems applying expertise. *The New York Times*, August 29, 1983.
- MARTIN, K. F. A review by discussion of condition monitoring and fault diagnosis in machine tools. *International Journal of Machine Tools Manufacturing*, vol.34, n.4, p.527-551, 1994.
- MARTIN, K. F., MARZI, M. H. Neural network solution to coolant system diagnostics. *Fourth International Conference on Profitable Condition Monitoring*, p.217-227, Kluwer Academic, 1992.
- MARZI, M. H., MARTIN, K. F. Artificial neural network condition monitoring and fault diagnosis. *ASME First International Conference on Neural Networks*, San Diego, May, 1991.
- McCULLOCH, W. S., PITTS, W. H. A logical calculus of the ideas imminent in nervous activity. *Bulletin of Mathematical Biophysics*, vol.5, p.115-133, 1943.
- MICHALSKI, R. S., CARBONELL, J. G., MITCHELL, T. M. *Machine learning - an artificial intelligence approach*. Morgan Kaufmann Publishers Inc., vol. I, 1990.

- MOLLER, M. F. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, vol.6, n.4, p.525-533, 1993.
- MONOSTORI, L. Learning procedures in machine tool manufacturing. *Computers in Industry*, vol.7, p.53-64, 1986.
- NELSON, W. R. REACTOR: an expert system for diagnosis and treatment of nuclear reactor accidents. *Proceedings AAAI-82*, p.296-301, 1982.
- PASSOS, E. L. *Inteligência Artificial e Sistemas Especialistas*. LTC Editora Ltda, 1987.
- PECK, J. P., BURROWS, J. On-line condition monitoring of rotating equipment using neural networks. *ISA Transactions*, vol.33, n.2, p.159-164, 1994.
- PEEKEN, H., TROEDER, Ch., MUSCHAWECK, F., XU, J. W. Fuzzy-Logik - eine neue Methode zur technischen Fehlerfrühdiagnose und Betriebsüberwachung an Maschinenanlagen. In: *VDI Conference on Schwingungsüberwachung*, Hamburg, Germany, VDI Berichte 568, p.335-351, 1985.
- PHAM, D. T., BAYRO-CORROCHANO, E. J. Neural networks for classifying surface defects on automotive valve stem seals. *International Journal of Machine Tools Manufacturing*, vol.35, n.8, p.1115-1124, 1995.
- QI, Z., LU, Y., YANG, S. Non-stationary modeling of vibration signals for monitoring the condition of machinery. *Mechanical Systems and Signal Processing*, vol.4, n.5, p.355-365, 1990.
- QUINLAN, J. R. Induction of decision trees. *Machine Learning*, vol.1, p.81-106, 1986.
- QUINLAN, J. R. Simplifying decision trees. *International Journal of Man-Machine Studies*, vol.27, p.221-234, 1987.
- QUINLAN, J. R. Learning efficient classification procedures and their application to chess end games. In: MICHALSKI, R. S., CARBONELL, J. G., MITCHELL, T. M. *Machine learning - an artificial intelligence approach*, Morgan Kaufmann Publishers Inc., vol. I, p.463-482, 1990.
- RICH, E., KNIGHT, K. *Artificial Intelligence*. Makron Books, 1994.
- ROSENBLATT, R. *Principles of Neurodynamics*. New York: Spartan Books, 1959.
- RUSPINI, E. Numerical methods for fuzzy clustering. *Inf. Sci.*, vol.6, p.273-284, 1981.
- SARGEANT, T. Frequency flyers. *Process Engineering*, vol.76, n.9, p.41-42,44, September, 1995.
- SCHILDT, H. *Inteligência Artificial utilizando linguagem C*. McGraw Hill, 1989.

- SCUDELLER Jr., W. *Estudo e implementação de critérios para monitoramento contínuo de máquinas hidroelétricas*. São Paulo: Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas, 1993. 106p. Dissertação (Mestrado).
- SHORTLIFFE, E. H. *Computer-Based Medical Consultation - MYCIN*. Amsterdam: Elsevier Scientific Pub. Co., 1976.
- SILVA, F. M., ALMEIDA, L. B. Speeding up backpropagation. *Advanced Neural Computers*, R. Eckmiller, ed., North-Holland, p.151-160, 1990.
- SING-TZE BOW, *Pattern recognition - applications to a large data-set problems*. New York and Bessel: Marcel Dekker Inc., 1984, 323p.
- SMART, S. B. Planejamento e Arquitetura do "CIM" (Computer Integrated Manufacturing). *7º SCNB - Seminário de Comando Numérico no Brasil e 3ª Jornada Internacional de Automação Industrial*, São Paulo, 1987.
- SORENSEN, H. V., JONES, D. L., BURRUS, S., HEIDEMAN, M. T. On computing the discrete Hartley transform. *IEEE Transactions on Acoustic, Speech and Signal Processing*, ASSP-33 (n4):1231-1238, 1985.
- TANAKA, M., SAKAWA, M., SHIROMARU, I., MATSUMOTO, T. Application of Kohonen's self-organizing network to the diagnosis system for rotating machinery. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 1995. vol.5, p.4039-4044.
- TANSEL, I. N., McLAUGHLIN, C. Detection of tool breakage in milling operations-II, the neural network approach. *Internation Journal of Machine Tools Manufacturing*, vol.33, n.4, p.545-558, 1993.
- TEIXEIRA JR, A. *et al.* Inteligência Artificial aplicada ao Planejamento de Processos de Usinagem. *1º Encontro Regional de Automação e Instrumentação*, p.IA14-IA19, 1989.
- TLUSTY, J., ISMAIL, F. Basic non-linearity in machining chatter. *Annals of the CIRP*, vol.30, p.289-304, 1981.
- TRANter, J. The fundamentals of, and the application of computers to, condition monitoring and predictive maintenance. In: *Proceedings of the First International Machinery Monitoring & Diagnostics Conference*, p.394-401, Las Vegas, Nevada, USA, Setembro 1989.
- TRINZIC *Developing applications with the aion development system - Version 1.0 DC0319*. Trinzic Corporation, 1994.
- TROEDER, Ch., MUSCHAWECK, F., XU, J. W. Machine diagnosis with classification systems. In: *IFTToMM World Congress on the Theory of Machines and Mechanisms*, p.679-683, Sevilha, Espanha, 1987.

- TURBAN, E. *Expert systems and applied artificial intelligence*. MacMillan Inc., 1992.
- UNDERWOOD, W. E. A CSA model-based nuclear power plant consultant. *Proceedings AAAI-82*, p.302-305, 1982.
- VIBROCAM 5000 - A system for computer controlled diagnostic monitoring of turbine sets. Germany: SCHENCK, s. d., 11p. (Product data).
- WANG, M., ZHU, J. Y., ZHANG, Y. Z. Fuzzy pattern recognition of the metal cutting states. *Annals of the CIRP*, vol.34, n1, p.133-136, 1985.
- WATERMAN, D. A. *A guide to expert systems*. Addison-Wesley Publishing Company, 1986.
- ZADEH, L. A. Fuzzy sets. *Information and Control*, 8, p.338-353, 1965.
- ZIMMER, S., BENTLY, D. E. Predictive maintenance programs for rotating machinery using computerized vibration monitoring systems. In: *VDI Conference on Schwingungsüberwachung*, Hamburg, Germany, VDI Berichte 568, p.265-287, 1985.