

MINISTÉRIO DA DEFESA
EXÉRCITO BRASILEIRO
SECRETARIA DE CIÊNCIA E TECNOLOGIA
INSTITUTO MILITAR DE ENGENHARIA
CURSO DE MESTRADO EM SISTEMAS E COMPUTAÇÃO

RAFAEL LIMA DE CARVALHO

SISTEMA DE IDENTIFICAÇÃO PARA A CASA INTELIGENTE
UTILIZANDO SOM

Rio de Janeiro
2008

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

INSTITUTO MILITAR DE ENGENHARIA

RAFAEL LIMA DE CARVALHO

SISTEMA DE IDENTIFICAÇÃO PARA A CASA INTELIGENTE
UTILIZANDO SOM

Dissertação de Mestrado apresentada ao Curso de Mestrado em Sistemas e Computação do Instituto Militar de Engenharia, como requisito parcial para obtenção do título de Mestre em Sistemas e Computação.

Orientador: Prof. Paulo Fernando Ferreira Rosa - Ph.D.

Rio de Janeiro
2008

c2008

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

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

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

Os conceitos expressos neste trabalho são de responsabilidade do autor e do orientador.

C331s Carvalho, R. L.
Sistema de Identificação para a Casa Inteligente Utilizando Som/ Rafael Lima de Carvalho.
– Rio de Janeiro: Instituto Militar de Engenharia, 2008.
146 p.: il., tab.

Dissertação (mestrado) – Instituto Militar de Engenharia – Rio de Janeiro, 2008.

1. Domótica (Robótica). 2. Casa Inteligente (Robótica). I. Carvalho, Rafael Lima de. II. Instituto Militar de Engenharia.

CDD 629.892

INSTITUTO MILITAR DE ENGENHARIA

RAFAEL LIMA DE CARVALHO

**SISTEMA DE IDENTIFICAÇÃO PARA A CASA INTELIGENTE
UTILIZANDO SOM**

Dissertação de Mestrado apresentada ao Curso de Mestrado em Sistemas e Computação do Instituto Militar de Engenharia, como requisito parcial para obtenção do título de Mestre em Sistemas e Computação.

Orientador: Prof. Paulo Fernando Ferreira Rosa - Ph.D.

Aprovada em 08 de maio de 2008 pela seguinte Banca Examinadora:

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

Prof. Cláudia Marcella Justel - D.Sc. do IME

Prof. Paulo Renato da Costa Pereira - D.Sc. do IME

Prof. Ronaldo Ribeiro Goldschmidt - D.Sc. do IST/FAETEC

Rio de Janeiro
2008

Em memória de Maria de São Pedro Ferreira Rosa

AGRADECIMENTOS

A Deus pelo fornecimento de força, resistência e fé em todo o curso.

A minha mãe, Maria Iones Pereira Lima, simplesmente por tudo.

Ao Exército Brasileiro, que através do Instituto Militar de Engenharia, concedeu-me essa oportunidade.

Ao meu orientador Prof. Paulo Rosa, por todo apoio, dedicação e orientações passados durante o período em que estive no IME.

A minha noiva Débora Sabatini, pelo carinho e compreensão.

Agradeço a todas as pessoas que contribuíram com o desenvolvimento desta dissertação, tenha sido por meio de críticas, idéias, apoio, incentivo ou qualquer outra forma de auxílio. Em especial, desejo agradecer às pessoas citadas a seguir: Fernanda, Seimou, Daniel, Alexandre Rossini, Fábio Vidal, Liliana, Marco Antônio, Monael, Wagner, Cap. Sandro, Lu, Silvano, Lucas, Ivan, Warley.

Quero agradecer de forma especial ao Cel. Pedro Eduardo e sua esposa Rosely, pelo fundamental apoio no final deste trabalho, me abrigando em sua casa. Nada do que eu possa fazer pagaria tal caridade.

A Vovó Salete, por sempre abençoar meu dia antes que eu o começasse.

Aos alunos da graduação e demais participantes do experimento realizado.

Por fim, a todos os professores e funcionários do Departamento de Engenharia de Sistemas (SE/8) do Instituto Militar de Engenharia.

Rafael Lima de Carvalho

A vida e a morte são determinadas demais, por demais implacáveis para que sejam puramente acidentais.

(Charles Chaplin)

SUMÁRIO

LISTA DE ILUSTRAÇÕES	9
LISTA DE TABELAS	13
LISTA DE ABREVIATURAS E SÍMBOLOS	17
1 INTRODUÇÃO	20
1.1 Considerações iniciais	22
1.2 Objetivo	25
1.2.1 Objetivo Geral	25
1.2.2 Objetivos Específicos	25
1.3 Motivação	26
1.4 Estrutura da Dissertação	27
2 ESTADO DA ARTE E REVISÃO DA LITERATURA	28
2.1 Ambientes Inteligentes	28
2.2 Revisão da Literatura	36
3 FUNDAMENTAÇÃO TEÓRICA	42
3.1 Conceitos Básicos	42
3.2 Características do Som	48
3.2.1 Parâmetros Objetivos	48
3.2.2 Parâmetros Subjetivos	56
3.3 Métodos de Seleção de Atributos	68
3.4 Algoritmos de Classificação	70
4 IDENTIFICANDO INDIVÍDUOS ATRAVÉS DO SOM DOS PAS-	
SOS	73
4.1 Análise da Marcha	73
4.2 Acústica dos Passos	77
4.3 Identificando Indivíduos a Partir do Som dos Passos	81
5 ORGANIZAÇÃO DOS EXPERIMENTOS	83
5.1 Base de Dados	83

5.2	Metodologia	86
5.2.1	Pré-processamento	88
5.2.2	Extração e Seleção das Características	91
5.2.3	Identificação	101
6	CONSIDERAÇÕES FINAIS	126
6.1	Discussões Gerais	126
6.2	Conclusões	129
6.3	Trabalhos Futuros	130
6.4	Agradecimentos	131
7	REFERÊNCIAS BIBLIOGRÁFICAS	132
8	APÊNDICES	138
8.1	APÊNDICE 1: Resultados complementares	139
8.1.1	K-Means aplicado a todos os coeficientes mel-cepestrais	139
8.1.2	K-NN aplicado a todos os coeficientes mel-cepestrais	144

LISTA DE ILUSTRAÇÕES

FIG.1.1	Pegadas na malha de sensores (adaptado de (NASCIMENTO, 2002)). 23	
FIG.1.2	Esquema do SMA da Casa Inteligente do IME (BOTELHO, 2005).	24
FIG.1.3	Circuito do sensor de passos proposto em (LIMA, 2005). (a) Corte da malha normal. (b) Sensor deformado.	24
FIG.2.1	Imagens do HomeLab (DERUYTER, 2005). (a) Imagem obtida de uma das 34 câmeras instaladas no laboratório. (b) Cabine do técnico responsável por identificar e gravar as ações dos moradores.	29
FIG.2.2	Imagens do PlaceLab. (a) PlaceLab. (b) Software mostrando as leituras dos sensores.	32
FIG.2.3	Foto do interior da Adaptive House e o programa que monitora os sensores.	34
FIG.3.1	Processo de amostragem de um sinal analógico.	44
FIG.3.2	Funções janelas.	47
FIG.3.3	Identificação dos PMPs (Pico Máximo do Passo) em uma sequência de passos.	50
FIG.3.4	Sistema linear do mecanismo de produção de voz.	52
FIG.3.5	Coeficientes cepstrais obtidos pelo método LPC. A sequência utilizada foi digitalizada a 44.1 KHz e contém o som de quatro passos.	53
FIG.3.6	Resultado da estimação do envelope espectral. (a) Envelope usando o método de Welch. (b) Envelope usando o método do periodograma.	56
FIG.3.7	Corte longitudinal do ouvido, com suas principais estruturas (BARBEDO, 2004).	58
FIG.3.8	Cóclea, membrana basilar e vibração da membrana em duas diferentes frequências (adaptado de (BARBEDO, 2004)).	59
FIG.3.9	Banco de filtros triangulares espaçados segundo a escala <i>mel</i> (DAVIS, 1980).	62
FIG.3.10	Diagrama em blocos para a extração dos coeficientes mel-cepstrais (REYNOLDS, 1995; DELIMA, 2001) (O filtro de pré-ênfase é uti-	

	lizado em sinais de voz).	63
FIG.3.11	Curvas isofônicas de Fletcher-Munson.	64
FIG.3.12	Relação entre a escala bark e hertz.	66
FIG.4.1	Diferenças entre passo e passada.	74
FIG.4.2	Períodos da fase de apoio.	75
FIG.4.3	Períodos da fase de balanço.	76
FIG.4.4	Componentes da força de reação do solo(adaptado de (CATTIN, 2002)).	77
FIG.5.1	Configuração do experimento.	83
FIG.5.2	Palco de madeira onde se sucederam os experimentos.	84
FIG.5.3	Tipos de calçados utilizados no experimento. (a) ST1: Coturno. (b) ST2: Tênis. (c) ST3: Sapato. (d) ST4: Sapato de borracha. (e) ST5: Sapatilha. (f) ST6: Sapatênis.	86
FIG.5.4	Padrão de nomenclatura utilizado no armazenamento das amostras do experimento realizado.	86
FIG.5.5	Visão geral do sistema.	88
FIG.5.6	Exemplo de seleção de passos em uma amostra completa.	89
FIG.5.7	ZCR normalizado com janela de 16 ms e 32 ms. (a) ZCR normal- izado, com janela de 16ms (706 amostras). (b) ZCR normalizado, com janela de 32ms (1412 amostras).	90
FIG.5.8	SEL normalizada com janela de 16ms e 32ms. (a) SEL normalizado, com janela de 16 ms (706 amostras). (b) SEL normalizado, com janela de 32 ms (1412 amostras).	90
FIG.5.9	Período médio dos passos extraído dos indivíduos participantes dos grupos COTURNO (a) e MESMOCALCADO (b).	92
FIG.5.10	Período médio dos passos extraído dos indivíduos participantes dos grupos SAPATO e UMPORCALCADO.	93
FIG.5.12	Análise 1 do Envelope Espectral obtido através dos algoritmos Welch e Periodograma. São exibidos os envelopes da amostra 13 do conjunto MESMOCALCADO, obtidos através do método de Welch (a) e Periodograma (b).	93

FIG.5.11	Período médio dos passos extraído dos indivíduos participantes dos grupos TENIS e IGNORECALCADO.	94
FIG.5.13	Análise 2 do Envelope Espectral obtido através dos algoritmos Welch e Periodograma. São exibidos os envelopes do indivíduo 13 usando coturno (ST1), estimados pelo método baseado no Welch (a) e Periodograma (b).	94
FIG.5.14	Análise 1 dos coeficientes mel-cepestrais. Em (a) e (b) são exibidos os dois primeiros coeficientes obtidos usando as amostras dos participantes do conjunto COTURNO.	95
FIG.5.15	Análise 1 dos coeficientes mel-cepestrais. Em (a) e (b) são exibidos o 12º e 16º coeficientes, respectivamente, obtidos usando as amostras dos participantes do conjunto COTURNO.	96
FIG.5.16	Análise 2 dos coeficientes mel-cepestrais. Em (a) o primeiro coeficiente de todas as amostras do indivíduo ID5 calçando tênis (ST2). Em (b) é mostrado o primeiro coeficiente de todas as amostras do indivíduo ID2 utilizando sapato (ST3).	97
FIG.5.17	Análise 1 do sonoridade segundo o modelo de <i>Zwicker</i> . Em (a) é mostrado a sonoridade específica em torno de 24 bandas críticas, para os indivíduos participantes do conjunto UMPORCALCADO. Em (b) a sonoridade específico em torno de 24 bandas críticas para os indivíduos participantes do conjunto SAPATO.	98
FIG.5.18	Análise 2 do <i>sonoridade</i> segundo o modelo de <i>Zwicker</i> . Em (a) é mostrado a <i>sonoridade</i> específica em torno de 24 bandas críticas, para as amostras do indivíduo IGNORECALCADO. Em (b) a <i>sonoridade</i> específica em torno de 24 bandas críticas para o indivíduo ID3 utilizando o tipo de calçado TENIS.	98
FIG.5.19	Análise 1 dos coeficientes cepestrais. Em (a) os coeficientes da amostra 6 do conjunto UMPORCALCADO. Em (b) são reunidos os coeficientes cepestrais da amostra 7 de todos os indivíduos participantes do conjunto COTURNO.	99
FIG.5.20	Análise 2 dos coeficientes cepestrais. Em (a), os coeficientes cepestrais de 10 amostras do indivíduo ID1 utilizando o tipo de calçado coturno. Em (b) os coeficientes cepestrais das amostras do in-	

divíduo ID5 calçando tênis. 100

LISTA DE TABELAS

TAB.3.1	Bandas críticas (em Bark).	61
TAB.5.1	Dados físicos dos indivíduos participantes do experimento.	85
TAB.5.2	Resultado da classificação, avaliando a frequência dos passos através do K-NN aplicado a todos os conjuntos.	102
TAB.5.3	Resultado da classificação, avaliando a frequência dos passos através do K-Means aplicado a todos os conjuntos.	102
TAB.5.4	Resultado da classificação, avaliando a sonoridade total através do K-NN aplicado a todos os conjuntos.	103
TAB.5.5	Resultado da classificação, avaliando a frequência dos passos através do K-Means aplicado a todos os conjuntos.	103
TAB.5.6	Resultado da classificação, avaliando a sonoridade específica através do K-Means aplicado a todos os conjuntos.	104
TAB.5.7	Resultado da classificação, avaliando a sonoridade específica através do K-Means aplicado a todos os conjuntos.	104
TAB.5.8	Resumo dos resultados ao aplicar o K-NN aos coeficientes mel-cepestrais, para todos os conjuntos. A tabela reúne apenas os melhores resultados; e estes ocorreram quando k foi igual a 1.	105
TAB.5.9	Resumo com os melhores resultados ao aplicar o K-Means nos coeficientes mel-cepestrais, para todos os conjuntos, com $K = 5$	106
TAB.5.10	Resumo com os melhores resultados ao aplicar o K-Means nos coeficientes mel-cepestrais, para todos os conjuntos, com $K = 10$	107
TAB.5.11	Resultados da aplicação do K-NN aos envelopes espectrais com extração baseada nos métodos: Periodograma e Welch.	107
TAB.5.12	Resultados da aplicação do K-Means aos envelopes espectrais com extração baseada nos métodos: Periodograma e Welch.	108
TAB.5.13	Resultados da aplicação do K-NN aos coeficientes cepestrais baseados no LPC.	109
TAB.5.14	Resultados da aplicação do K-Means aos coeficientes cepestrais baseados no LPC.	109
TAB.5.15	Resumo com os melhores resultados ao se aplicar o KNN em todas as características e grupos. A primeira coluna mostra a carac-	

	terística, as siglas significam: FP - Frequência dos passos, ST - Sonoridade Total, SE - Sonoridade Específica, EP - Envelope espectral usando Periodograma, EW - Envelope espectral usando o método de Welch e CC - Coeficientes Cepstrais.	110
TAB.5.16	Resumo com os melhores resultados ao se aplicar o K-Means em todas as características e grupos. A primeira coluna mostra a característica, as siglas significam: FP - Frequência dos passos, ST - Sonoridade Total, SE - Sonoridade Específica, EP - Envelope espectral usando Periodograma, EW - Envelope espectral usando o método de Welch e CC - Coeficientes Cepstrais.	110
TAB.5.17	Características combinadas para o conjunto COTURNO, avaliando o K-NN.	111
TAB.5.18	Resultado do algoritmo K-NN após as combinações para o conjunto COTURNO.	111
TAB.5.19	Características combinadas para o conjunto COTURNO, avaliando o K-Means com $K = 5$	112
TAB.5.20	Resultado do algoritmo K-Means com $K = 5$, após as combinações para o conjunto COTURNO.	112
TAB.5.21	Características combinadas para o conjunto COTURNO, avaliando o K-Means com $K = 10$	112
TAB.5.22	Resultado do algoritmo K-Means com $K = 10$, após as combinações para o conjunto COTURNO.	113
TAB.5.23	Características combinadas para o conjunto TENIS, avaliando o K-NN.	113
TAB.5.24	Resultado do algoritmo K-NN após as combinações para o conjunto TENIS.	113
TAB.5.25	Características combinadas para o conjunto TENIS, avaliando o K-Means com $K = 5$	114
TAB.5.26	Resultado do algoritmo K-Means com $K = 5$, após as combinações para o conjunto TENIS.	114
TAB.5.27	Características combinadas para o conjunto TENIS, avaliando o K-Means com $K = 10$	114
TAB.5.28	Resultado do algoritmo K-Means com $K = 10$, após as combinações	

	para o conjunto TENIS.	115
TAB.5.29	Características combinadas para o conjunto SAPATO, avaliando o K-NN.	115
TAB.5.30	Resultado do algoritmo K-NN após as combinações para o conjunto SAPATO.	115
TAB.5.31	Características combinadas para o conjunto SAPATO, avaliando o K-Means com $K = 5$	116
TAB.5.32	Resultado do algoritmo K-Means com $K = 5$, após as combinações para o conjunto SAPATO.	116
TAB.5.33	Características combinadas para o conjunto SAPATO, avaliando o K-Means com $K = 10$	117
TAB.5.34	Resultado do algoritmo K-Means com $K = 10$, após as combinações para o conjunto SAPATO.	117
TAB.5.35	Características combinadas para o conjunto UMPORCALCADO, avaliando o K-NN.	118
TAB.5.36	Resultado do algoritmo K-NN após as combinações para o conjunto UMPORCALCADO.	118
TAB.5.37	Características combinadas para o conjunto UMPORCALCADO, avaliando o K-Means com $K = 5$	118
TAB.5.38	Resultado do algoritmo K-Means com $K = 5$, após as combinações para o conjunto UMPORCALCADO.	119
TAB.5.39	Características combinadas para o conjunto UMPORCALCADO, avaliando o K-Means com $K = 10$	119
TAB.5.40	Resultado do algoritmo K-Means com $K = 10$, após as combinações para o conjunto UMPORCALCADO.	120
TAB.5.41	Características combinadas para o conjunto MESMOCALCADO, avaliando o K-NN.	120
TAB.5.42	Resultado do algoritmo K-NN após as combinações para o conjunto MESMOCALCADO.	120
TAB.5.43	Características combinadas para o conjunto MESMOCALCADO, avaliando o K-Means com $K = 5$	121
TAB.5.44	Resultado do algoritmo K-Means com $K = 5$, após as combinações para o conjunto MESMOCALCADO.	121

TAB.5.45	Características combinadas para o conjunto MESMOCALCADO, avaliando o K-Means com $K = 10$	122
TAB.5.46	Resultado do algoritmo K-Means com $K = 10$, após as combinações para o conjunto MESMOCALCADO.	122
TAB.5.47	Características combinadas para o conjunto IGNORECALCADO, avaliando o K-NN.	122
TAB.5.48	Resultado do algoritmo K-NN após as combinações para o conjunto IGNORECALCADO.	123
TAB.5.49	Características combinadas para o conjunto IGNORECALCADO, avaliando o K-Means com $K = 5$	123
TAB.5.50	Resultado do algoritmo K-Means com $K = 5$, após as combinações para o conjunto IGNORECALCADO.	124
TAB.5.51	Características combinadas para o conjunto IGNORECALCADO, avaliando o K-Means com $K = 10$	124
TAB.5.52	Resultado do algoritmo K-Means com $K = 10$, após as combinações para o conjunto IGNORECALCADO.	125
TAB.8.1	K-Means aplicado aos coeficientes 01 ao 04.	139
TAB.8.2	K-Means aplicado aos coeficientes 05 ao 08.	140
TAB.8.3	K-Means aplicado aos coeficientes 09 ao 12.	141
TAB.8.4	K-Means aplicado aos coeficientes 13 ao 16.	142
TAB.8.5	K-Means aplicado aos coeficientes 17 ao 20.	143
TAB.8.6	KNN aplicado aos coeficientes mel-cepestrais 01 ao 08.	144
TAB.8.7	KNN aplicado aos coeficientes mel-cepestrais 09 ao 16.	145
TAB.8.8	KNN aplicado aos coeficientes mel-cepestrais 17 ao 20.	146

LISTA DE ABREVIATURAS E SÍMBOLOS

ABREVIATURAS

A/D	-	<i>Analógico-digital</i>
ANSI	-	<i>American National Standard Institute</i>
fdp	-	<i>Função de distribuição de probabilidade</i>
FFT	-	<i>Fast Fourier Transform</i>
IDE	-	<i>Integrated Development Environment</i>
IME	-	<i>Instituto Militar de Engenharia</i>
LPC	-	<i>Linear Predictive Coding</i>
MaSE	-	<i>Multiagent Software Engineering</i>
SMA	-	<i>Sistema Multiagentes</i>
S/N	-	<i>Taxa de Sinal por Ruído</i>
K-NN	-	<i>K Nearest Neighbors</i>
ART	-	<i>Art Resonance Theory</i>
MIT	-	<i>Massachusetts Institute of Technology</i>
ACHE	-	<i>Adaptive Control of Home Environments</i>
PMP	-	<i>Pico Máximo do Passo</i>
MCC	-	<i>Coeficiente Mel Cepstral</i>
PCA	-	<i>Análise de Componentes Principais</i>
FLDA	-	<i>Discriminante Linear de Fisher</i>
FRS	-	<i>Força de Reação do Solo</i>
SEL	-	<i>Nível de Energia do Sinal</i>
ZCR	-	<i>Taxa de Cruzamento por Zero</i>
FP	-	<i>Frequência dos passos</i>
ST	-	<i>Sonoridade Total</i>
SE	-	<i>Sonoridade específica</i>
EP	-	<i>Envelope espectral usando o Periodograma</i>
EW	-	<i>Envelope espectral usando o método de Welch</i>
CC	-	<i>Coeficientes Cepstrais</i>

RESUMO

O presente trabalho lida com o problema de identificação de indivíduos a partir do som dos passos. A abordagem utilizada foi estudar características aplicadas a outros sistemas de identificação e realizar uma avaliação de suas propriedades discriminatórias, quando aplicadas ao sinal acústico do caminhar.

As características foram divididas em dois grupos: objetivas e subjetivas. As objetivas se baseiam nos aspectos físicos da onda. As subjetivas são parâmetros processados levando-se em consideração o modelo acústico do sistema auditivo humano, ou seja, como o som é sentido pelo ouvido. Os parâmetros considerados geraram um volume considerável de dados, fazendo-se necessária a aplicação de métodos para eleição de atributos que melhor representem o padrão do caminhar.

A fase de identificação é composta pela avaliação de dois algoritmos: uma versão modificada do K-Means e k vizinhos mais próximos - KNN, utilizando validação cruzada com k conjuntos e *leave-one-out*, como metodologias de partição de dados, respectivamente. O critério de Fisher foi utilizado como método de seleção de atributos. Em ambos, foram testadas diversas configurações utilizando o critério de seleção, com variadas proporções de dados.

Para avaliar o sistema, foi realizado um experimento real através de gravações do caminhar de um conjunto de indivíduos, os quais utilizaram seis tipos de calçados diferentes no total. Uma análise foi feita através da separação do conjunto total em seis subgrupos característicos: COTURNO, TENIS, SAPATO, UMPORCALCADO, MESMOCALCADO e IGNORECALCADO.

A metodologia de testes consistiu em aplicar cada característica, singularmente, nos algoritmos de classificação. Em seguida, as características são selecionadas e combinadas de acordo com a quantidade de acertos obtidos na etapa anterior. Essa metodologia resultou em taxas de identificação que vão de 33.8 % a 97.5 %.

ABSTRACT

The current work approaches the project of a personal identification system based on the sound of the footsteps.

Psychoacoustics parameters were applied to the footsteps sounds. The considered features generated a high amount of data. This implicates the application of methods to ranking the best features to choose the best representation of the gait pattern.

The identification phase was composed of two classification algorithms performances: k-means and K-NN (k nearest neighbors) . The data partition methodology was k cross validation (k-means performance) and leave-one-out (K-NN performance). The parameter selection used was the Fisher's criterion.

In order to evaluate the proposed system, a real experiment was done in with a set of individuals walking on a wooden platform using six different types of shoes. One analysis was driven by database segregation in six specific subsets: COTURNO, TENIS, SAPATO, UMPORCALCADO, MESMOCALCADO E IGNORECALCADO.

The methodology of tests was consolidated in running each feature, peculiarly, on the considered classifiers. Furthermore, the better features were selected and combined in accordance with the level of classification accuracy. This methodology resulted on a range of 33.8 % to 97.5 % identification rates.

1 INTRODUÇÃO

A casa inteligente pode ser definida, de maneira sistemática, como uma casa equipada com objetos inteligentes interligados por uma rede doméstica, capaz de transmitir informações entre esses objetos e um mecanismo para conectar o ambiente com meios de comunicação externos (internet).

Vigilância, segurança e economia de recursos, como água, energia e gás, são alguns dos serviços básicos que uma casa inteligente pode oferecer. No projeto casa inteligente proposto por (LIMA, 2005), o serviço de iluminação adaptativa, que aproveita a iluminação exterior, mostrou uma economia de 32 % de energia elétrica. O conforto e serviços personalizados da casa vão muito além disso. É possível, por exemplo, adaptar a casa para o cuidado de pessoas idosas ou enfermas, sendo configurados serviços como: monitoramento de torneiras, chuveiros, fogões, além da detecção de alguns perigos como quedas.

Há um vasto campo de aplicações para a casa inteligente. Para exemplificar, imagine que uma pessoa seja submetida a uma cirurgia e more distante do recurso hospitalar. Para o acompanhamento do médico na recuperação do paciente, os gastos que essa pessoa teria para permanecer internada seriam consideravelmente altos. O ambiente onde ela reside pode ser capaz de obter informações vitais e enviá-las a um especialista para acompanhamento remoto. Com exceção de algum problema detectado, o indivíduo não necessitaria sair de casa e teria acesso ao acompanhamento médico. Isso implicaria no aumento da capacidade de atendimento do hospital, além do aumento de conforto dos pacientes.

Antes de imaginar e projetar aplicações para a casa, é necessário identificar seus habitantes para que os serviços possam ser configurados de forma personalizada e de acordo com as suas necessidades. Há algum tempo, pesquisadores vêm se esforçando na busca de novas formas de identificação. Os sistemas tradicionais de identificação, tais como senhas e crachás, podem ser esquecidos ou extraviados. Os sistemas biométricos de identificação levam vantagem em relação aos convencionais por estarem presentes no indivíduo sem o problema de esquecimento ou extravio (no caso de senhas e crachás), porém os mais conhecidos (íris, impressão digital, reconhecimento de face) apesar de apresentarem métodos que geram bons resultados, são invasivos por necessitarem a atenção do indivíduo. Além

de preciso, o sistema de identificação da casa deve ser transparente, ou seja, o indivíduo deve ser identificado através de uma interação implícita ou passiva. Analisando tais necessidades, um sistema de identificação baseado no padrão de caminhar foi considerado em algumas aplicações de domótica, pois caminhar é um ato primordial e natural na locomoção do ser humano.

Na literatura há alguns trabalhos (ADDLESEE, 1997; ORR, 2000; LIMA, 2005; SUTALA, 2005; SAVIO, 2007) sobre pisos inteligentes, os quais colhem dados dos passos dos indivíduos que caminham sobre os mesmos e então os submetem a um algoritmo de classificação. Apesar de apresentarem bons resultados, a implementação dessas soluções alteraria a infraestrutura da residência, uma vez que o piso teria que ser total ou quase totalmente substituído. Além disso, pelas pesquisas feitas até o momento, ainda não foram encontrados trabalhos onde seja possível identificar indivíduos locomovendo-se em vários modos (correndo, caminhando em velocidade normal, vagarosamente) como também identificar mais de um indivíduo simultaneamente.

Este trabalho se propõe ao estudo de um sistema de identificação que se baseia em características extraíveis de sinais de som, e no caso particular desse estudo, o som dos passos. Dentre as características exploradas em (LIMA, 2005), a frequência dos passos foi uma das determinantes para a identificação do indivíduo. No som também é possível a extração do período dos passos. Para isso, foi desenvolvido um algoritmo que se baseia no valor máximo produzido pelo impacto do calcanhar com o solo. No decorrer do estudo, todos os parâmetros pesquisados foram organizados em dois grupos: objetivos e subjetivos. O primeiro grupo reúne as características objetivas do som, ou seja, seus cálculos se baseiam no aspecto físico da onda. Enquanto que o segundo grupo, leva em consideração os aspectos subjetivos, ou seja, como o sinal sonoro é percebido pelo ouvido humano.

A dimensão dos dados obtidos na tarefa de extração das características estudadas em uma sequência sonora de passos, em geral é razoavelmente grande. Portanto, métodos de redução de atributos constituem uma das fases do projeto. Para essa tarefa, foi utilizado o discriminante de Fisher (vide Seção 3.3), o qual ordena as características segundo sua importância mediante ao cálculo da razão entre a variância entre-classes por intra-classes.

Após extrair as características, estas foram submetidas a dois algoritmos de classificação avaliados nesta dissertação, a saber: K vizinhos mais próximos - K-NN e o algoritmo de agrupamento K -Means. As metodologias de partição de dados utilizadas na avaliação dos métodos K-NN e K-Means foram respectivamente *leave-one-out*, que con-

siste em deixar uma amostra para teste e as demais para treinamento até que todas as amostras sejam testadas, e validação cruzada com k conjuntos, que consiste em dividir a base em k subconjuntos disjuntos, com aproximadamente o mesmo número de elementos, onde cada um dos k subconjuntos é utilizado como conjunto de teste e os demais para treinamento (GOLDSCHMIDT, 2005).

Um experimento real foi feito para avaliação do sistema. O conjunto foi particionado em seis subgrupos característicos. No intuito de obter as melhores taxas de reconhecimento, a seguinte metodologia de testes foi utilizada. Inicialmente, os algoritmos de classificação foram testados, para cada subgrupo, utilizando-se cada característica singularmente. Após isso, as melhores características são combinadas gerando resultados ainda melhores. A próxima seção introduz a casa inteligente do IME, comentando-se os trabalhos predecessores pertencentes ao projeto.

1.1 CONSIDERAÇÕES INICIAIS

O trabalho encontrado em (NASCIMENTO, 2002), foi marcado o início do projeto da Casa Inteligente do IME. Intitulado inicialmente como Sistema Dinâmico de Automação Residencial - SDAR, nessa primeira versão da casa, foram levantados os seguintes quesitos: protocolos de automação, sistema de identificação através de malha sensora com o desenvolvimento dos algoritmos de passo e considerado um sistema de iluminação objetivando economia. Os algoritmos de passo foram desenvolvidos para extraírem quatro parâmetros, dados dois pontos na malha feita de piezoelétrico ou células de carga, a saber: ângulos dos pés direito (δ_d) e esquerdo (δ_e), (f) frequência do passo, (p) peso e (s) comprimento do passo. A FIG. 1.1 mostra os parâmetros na malha propostos em (NASCIMENTO, 2002). Para a tarefa de identificação foi proposto o uso da rede neural do tipo *Art Resonance Theory* - ART1.

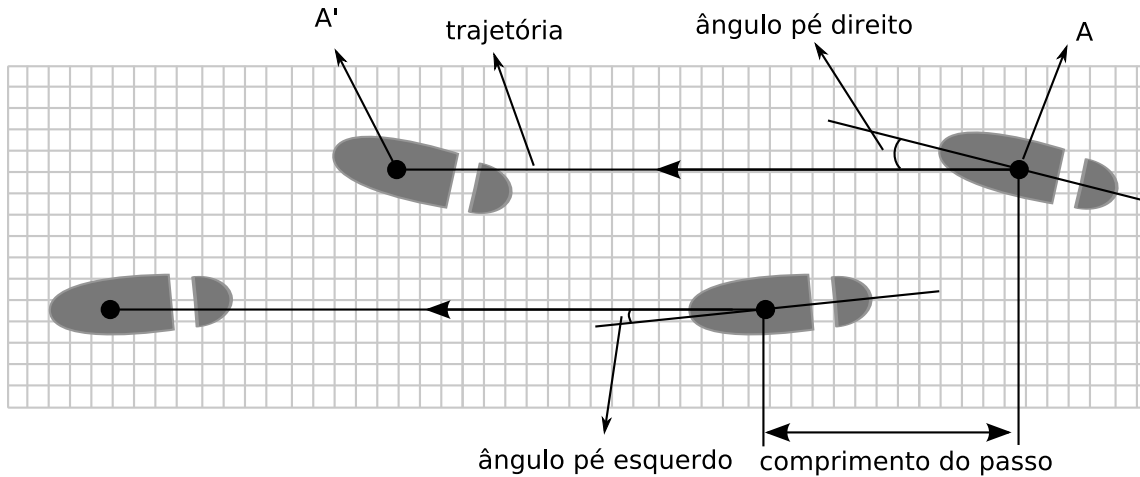


FIG. 1.1: Pegadas na malha de sensores (adaptado de (NASCIMENTO, 2002)).

Os estudos de Fitzner (2002) foram prosseguidos nos trabalhos de Botelho (BOTELHO, 2005) e Lima (LIMA, 2005). No trabalho feito por Botelho (BOTELHO, 2005), foi apresentada uma arquitetura de um Sistema Multiagentes - SMA para a casa inteligente. A arquitetura proposta contém duas camadas de agentes. Na primeira camada ficam os agentes interfaces, os quais são responsáveis por monitorar e controlar os eletrodomésticos da casa (luz, ar condicionado, sensor de passos). Na outra camada, existem os agentes: Rede Neural, Temperatura, Iluminação, Identificação, Consumo de Energia e Externo, com as respectivas principais funções: classificar o padrão do indivíduo, ajustar a temperatura, ajustar a iluminação, carregar preferências e julgar padrão advindo do agente Rede Neural, calcular o consumo de energia e representar a firma de segurança ou polícia. Tal arquitetura se destaca por sua flexibilidade, uma vez que o sistema pode ser usado em quaisquer ambientes (escolas, igrejas, comércio e casas), como também podem ser adicionadas novas funcionalidades sem alteração da mesma. A arquitetura SMA da casa foi testada via simulação e o sistema se comportou corretamente para o que foi projetado. A FIG. 1.2 mostra o modelo SMA para um cômodo (quarto) da casa.

No trabalho descrito em (LIMA, 2005), um sistema adaptativo para a casa inteligente foi modelado segundo a orientação a objetos, foi produzido um sensor de passos genuíno onde foi possível validar e melhorar fisicamente o trabalho predecessor. Além disso, o sistema de iluminação foi implementado e validado através de simulações, nas quais os resultados apresentaram uma economia de aproximadamente 40 %.

O sensor de passos é formado por pontos de pressão distribuídos de forma equidistante, para que o número de dispositivos detectores de contato excitados durante o caminhar

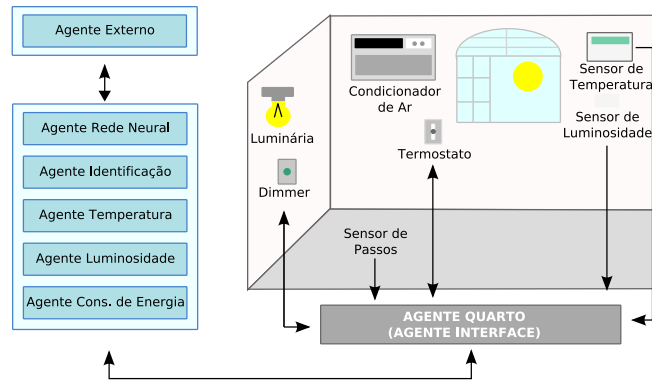


FIG. 1.2: Esquema do SMA da Casa Inteligente do IME (BOTELHO, 2005).

do indivíduo independa da direção de deslocamento. Basicamente, o sensor é constituído por duas camadas, superior e inferior. A camada superior é formada pela borracha *Eva* (etileno/vinil acetato), um material resistente porém que apresenta deformação quando submetido à pressão exercida pelo pé de um indivíduo. Abaixo dessa camada, encontram-se trilhas de alumínio, por onde a corrente circula no momento que uma pessoa caminhar sobre esta região do sensor. Em cada um dos lados das trilhas, fontes de tensão contínua de mesmo valor estão ligadas. Acondicionadas em orifícios feitos na camada amortecedora, encontram-se molas as quais juntamente com as trilhas de alumínio, formam o circuito por onde circula a corrente determinando a coordenada (x,y) , referente a área dos pés do caminhante. Para notar as diferenças de voltagem, resistores conectam molas vizinhas. A FIG. 1.3 mostra o corte da malha normal e deformado.

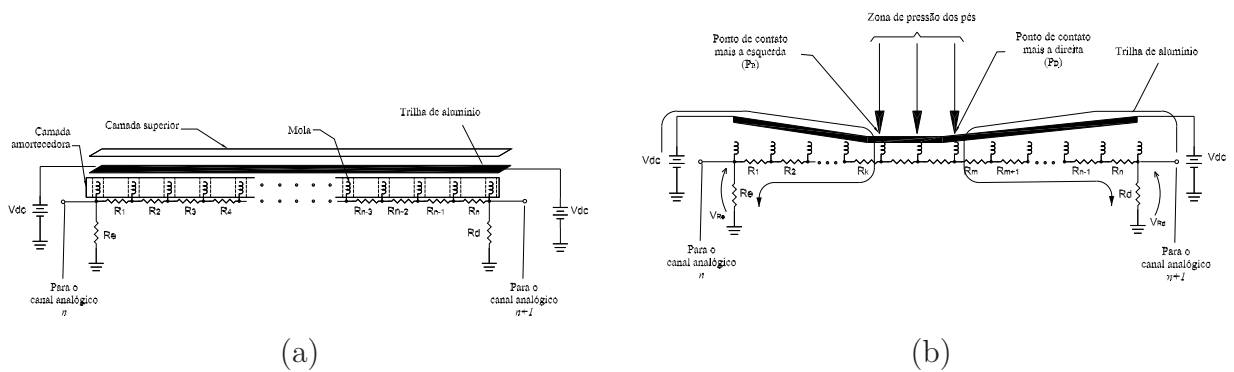


FIG. 1.3: Circuito do sensor de passos proposto em (LIMA, 2005). (a) Corte da malha normal. (b) Sensor deformado.

Através do conjunto (X,Y) formado por todos os pares excitados no sensor durante o caminhar, são extraídas três características, a saber: frequência, abertura e distância da

passada do indivíduo. Esses três parâmetros são introduzidos em uma rede neural ART2, onde além do ângulo formado entre os vetores, foi considerado o módulo do vetor, que com um conjunto de 10 amostras de 20 alunos do departamento, obteve, no geral, uma precisão de 87 % de acertos. Esse algoritmo de identificação foi modificado posteriormente no trabalho reunido em (ROSA, 2006), com a função de ativação da rede sendo a distância euclidiana do raio do neurônio, aumentando o índice de acertos para 90,2 %.

O trabalho atual, dá continuidade ao projeto da Casa Inteligente do IME, através da investigação de um sensor de passos baseado no som. É importante ressaltar que no sensor de passos proposto por Lima (2005), a influência do ruído ambiente é praticamente nula. O que no caso do som, é um problema complexo de se eliminar. No som, não é possível obter a abertura e a distância da passada, pois mesmo as soluções mais robustas em sistemas de localização de fonte sonora, não é possível alcançar a mesma precisão da malha sensora. Além disso, o som dos passos é muito mais sensível aos tipos de piso e calçados, do que a malha. Dos parâmetros explorados no trabalho de Lima (2005), somente a frequência dos passos pode ser identificada na onda sonora, sendo um dos parâmetros explorados nesse trabalho. As demais características apresentadas nessa pesquisa, são provenientes de sistemas de identificação de locutores e outras que se mostraram eficazes, no âmbito dos sinais dos passos.

1.2 OBJETIVO

Nesta seção são apresentados os principais objetivos da presente pesquisa. Para isso, eles serão divididos em *Objetivos gerais* e *Objetivos específicos*, destacando-se as finalidades globais e locais esperadas.

1.2.1 OBJETIVO GERAL

O objetivo geral desse trabalho é avaliar a capacidade de discriminação de um grupo de características objetivas e subjetivas do som, quando aplicadas ao som dos passos, e também avaliar sua aplicação em algoritmos de classificação, no intuito de identificar os moradores da casa inteligente.

1.2.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos são citados da seguinte forma:

- Reunir características aplicadas ao sinal sonoro e prosseguir com uma investigação a respeito do poder discriminatório, ao ser aplicado ao sinal dos passos;
- Aplicar métodos para eleição das melhores características, no quesito qualidade de classificação;
- Comparar algoritmos de classificação para a tarefa de identificação.
- Reunir o sistema composto pelos três tópicos anteriores e testá-lo em um experimento real, de magnitude maior que os trabalhos encontrados na literatura.

1.3 MOTIVAÇÃO

Apesar do conjunto de teste ser bastante limitado, alguns trabalhos na literatura (SHOJI, 2004, 2005; ITAI, 2006) mostraram bons resultados na identificação através do som das passadas. Porém, o conjunto utilizado em tais trabalhos é muito restrito, tendo somente cinco indivíduos caminhando com um tipo de calçado. Além disso o problema de reconhecer mais de um habitante no mesmo local e simultaneamente, ainda é considerado um problema difícil e sem solução até a presente pesquisa. O presente trabalho faz um estudo na identificação de indivíduos utilizando o som dos passos, em um ambiente doméstico e com um conjunto de testes mais complexo do que nos trabalhos encontrados.

Na literatura encontra-se muita documentação a respeito do funcionamento e tratamento do sinal de voz (FECHINE, 2000; JR., 1993; PICONE, 1993; SILVA, 2002; DELIMA, 2001). Infelizmente, não existe algo semelhante para o sinal produzido pelos passos. Este trabalho não propõe novas formulações matemáticas para representar os estímulos do som provocado pelos passos. Porém, são exploradas características utilizadas em outros trabalhos de classificação de locutor, com exceção à frequência dos passos, a qual já foi provada ser uma importante característica na identificação de pessoas a partir do caminhar (LIMA, 2005), que é um parâmetro particular ao ato de caminhar.

Na tarefa de classificação, a qualidade dos parâmetros é essencial. Neste trabalho são estudadas e implementadas dois grupos de características: objetivas e subjetivas. As características objetivas são aquelas em que se julga apenas o formato da onda, como os componentes de frequência que a formam até parâmetros mais complexos como os coeficientes mel-Cepstrais. Já os parâmetros subjetivos ou psicoacústicos são aqueles modelados de acordo com o sistema auditivo humano, ou seja, modelos matemáticos de

como o ouvido “percebe” o som. O algoritmo de classificação utilizado nos trabalhos anteriores da casa foi adaptado neste trabalho para classificar os indivíduos a partir das características estudadas. No decorrer do trabalho também é feita uma análise dos problemas encontrados durante o estudo do som dos passos.

1.4 ESTRUTURA DA DISSERTAÇÃO

Este trabalho está estruturado em sete capítulos. O presente capítulo teve por objetivo permitir ao leitor ter o conceito de casa inteligente, os benefícios trazidos por ela, como também mostrar os trabalhos do projeto da casa inteligente o qual o presente trabalho dá continuidade. Além disso, esta seção apresenta uma breve visualização dos demais capítulos deste documento, mostrando nos parágrafos a seguir uma descrição sucinta desses capítulos.

O Capítulo 2 reúne três dos principais projetos de ambientes inteligentes na literatura e comenta os trabalhos mais relacionados ao tema abordado por esse trabalho. Os projetos que formam a seção do estado da arte, foram escolhidos por serem casas reais e os sistemas inteligentes de automação testados com moradores humanos.

O Capítulo 3 reúne informações básicas sobre a área de processamento de sinais, relaciona as características do som avaliadas nesse trabalho, comenta sobre soluções a respeito de eleição dos melhores parâmetros, finalizando com os algoritmos de classificação examinados na presente pesquisa.

No Capítulo 4 a estrutura biométrica do caminhar, bem como as forças geradas por esse ato são congregadas. Além disso, é discutida a ausência de um modelo explicativo, tal como no sistema de identificação de locutores que o sinal de fala pode ser aproximado por um sistema linear. Neste capítulo também são levantados os requisitos de um sistema ideal de identificação através do som dos passos.

O experimento, metodologia e resultados são apresentados no Capítulo 5. Em seguida são apresentadas as conclusões no Capítulo 6, encerrando o texto com as referências bibliográficas.

2 ESTADO DA ARTE E REVISÃO DA LITERATURA

O presente capítulo reúne recentes pesquisas em ambientes inteligentes, e também os principais trabalhos relacionados ao problema de classificar indivíduos através do som dos passos.

2.1 AMBIENTES INTELIGENTES

Essa seção tem por finalidade apresentar projetos de ambientes inteligentes onde, de modo geral, o foco é estudar maneiras onde as pessoas possam ser inseridas à ambientes de alta tecnologia, sem ter que aprender a manipular sistemas de controle complexos. Especialmente, três ambientes foram reunidos: o *HomeLab*, projeto de pesquisa da empresa *Koninklijke Philips Electronics*; o *PlaceLab*, projeto do *Massachusetts Institute of Technology - MIT* em conjunto com o laboratório da *TIAX/LLC*¹; e o *Adaptive House*, projeto de pesquisa do Departamento de Ciência da Computação da Universidade do Colorado. Tais projetos foram escolhidos por representarem o estado da arte na área de ambientes inteligentes.

O *HomeLab* é um projeto da empresa *Philips*, criado no intuito de estudar o comportamento de indivíduos, mediante a apresentação de tecnologias domésticas voltadas para o lazer. Nesse sentido, as seguintes áreas iniciais foram consideradas: a experiência por imersão através de dispositivos de visualização, que estendem-se além dos limites físicos das telas, cobrindo aspectos como a luminosidade do ambiente; e a experiência de presença social e conectividade com amigos e familiares remotos (DERUYTER, 2005).

Em aspectos físicos, o *HomeLab* se parece com uma casa comum, com uma mobília moderna (vide FIG. 2.1a). Ela possui uma sala de estar, uma cozinha, dois quartos, um banheiro e um escritório. A casa também é equipada com 34 câmeras de vídeo e microfones posicionados no teto dos cômodos. A finalidade desses equipamentos é capturar a reação dos indivíduos durante o teste de interatividade com as tecnologias apresentadas. Portanto, técnicos monitoram a casa em período integral, analisando e armazenando o áudio e vídeo das interações dos indivíduos (ver a FIG. 2.1b). O *HomeLab*

¹TIAX Tecnology <http://www.tiaxllc.com/>

também conta com um sistema de controle de energia possibilitando que as luzes sejam controladas remotamente. Contudo, os moradores ainda têm a opção de ligar e desligar as luzes manualmente. Além disso, o laboratório tem suporte a redes sem fio e com fio, permitindo à casa acessar conteúdos exteriores que necessitar (RUYTER, 2003).



FIG. 2.1: Imagens do HomeLab (DERUYTER, 2005). (a) Imagem obtida de uma das 34 câmeras instaladas no laboratório. (b) Cabine do técnico responsável por identificar e gravar as ações dos moradores.

Para fazer uso das funcionalidades fornecidas por dispositivos móveis (tocadores de música, celulares e PDAs), os pesquisadores do *HomeLab* estão desenvolvendo uma tecnologia chamada *Context Aware Messaging Platform - CAMP*, para alimentar os dispositivos com informações condizentes ao ambiente, no qual estão inseridos. A idéia é conectar “sensores de orientação” dispostos em toda a casa, via *bluetooth*, ao dispositivo móvel, fornecendo-o informações sobre o contexto, aumentando ainda mais sua capacidade de interatividade. Por exemplo, se o morador está na cozinha, seu telefone celular poderia receber informações a respeito de receitas, levando em consideração seus níveis de colesterol.

Outra tecnologia que está atualmente sendo viabilizada pelos laboratórios da *Philips*, juntamente com uma empresa de materiais esportivos, é a de bio-sensores embutidos em roupas, com a finalidade de detectar e transmitir parâmetros como taxas de batidas do coração e respiração, como também níveis de oxigênio no sangue. Espera-se que com

tal tecnologia, seja possível monitorar as condições vitais de um indivíduo e realizar o seu acompanhamento em atividades físicas. Para tanto, a *Philips* está desenvolvendo sensores com maior poder de processamento, os quais poderão ser integrados em circuitos com baixo consumo de energia.

Atualmente, quatro tecnologias estão sendo testadas no *HomeLab*, a saber: brinquedos interativos, *iCat*, *dreamscreen* e espelho interativo. As novidades citadas são descritas a seguir.

A tecnologia chamada **brinquedos interativos** consiste na aplicação do ambiente inteligente a brinquedos inteligentes, preservando sua facilidade natural de uso e criando princípios de interação com o usuário. No *HomeLab*, um ambiente de *StoryTelling* com tratamento de áudio e uma interface palpável foi implementada utilizando redes de sensores. O protótipo consiste em uma pequena fazenda onde os animais são nós da rede de sensores (FONTIJN, 2005).

O **iCat** é um robô de assistência pessoal que tem a aparência de um gato. Ele possui 38cm de altura com 13 servos que controlam diferentes partes de sua face, tais como sobrancelhas, olhos, pálpebras, boca e posições da cabeça; possibilitando gerar diferentes expressões faciais (VANBREEMEN, 2005b,a). O Robô é equipado com uma câmera e um microfone, possuindo software para reconhecer objetos e faces, como também capacidade de sintetizar voz e entender algumas frases. Além disso, sensores de toque e LEDs multicoloridos estão instalados em suas orelhas e pés, os quais codificam os estados do robô (cansado, dormindo, etc.), através das cores dos LEDs e reage ao sentir algum toque. Esse aparato robótico é controlado por módulos que rodam em uma plataforma chamada *Open Platform for Personal Robots - OPPr*. O OPPr oferece um simulador que possibilita desenvolver módulos para o *iCat* em computadores pessoais - PCs. Utilizando o OPPr, um módulo de jogo de xadrez foi desenvolvido em (LEITE, 2007). Finalmente, dentre as funções sociais testadas atualmente no robô que habita o *HomeLab*, destacam-se: (a) informações sobre a agenda do indivíduo; e (b) capacidade de fornecer sugestões sobre tipos de pratos e atividades, levando em consideração as preferências pessoais do morador.

No projeto tecnológico **dreamscreen** é estudado como o conjunto “som e imagem” será usado no futuro para criar experiências de imersão. Por exemplo, ao transformar janelas em superfícies será possível uma série de novos serviços e aplicações, incluindo um verdadeiro *home theater* e substituição de cenas reais por cenas virtualmente mais atrativas. Os elementos necessários para criar tais experiências são: telas com grandes

dimensões, som direcional e soluções de interação cientes do contexto². Para aprender o que as pessoas farão no futuro, quando tais possibilidades existirem, um teste interativo da *DreamScreen* foi criado no *HomeLab*. Na largura das janelas da sala, com imagens, vídeos e som direcional. Por fim, soluções de interação intuitivas estão sendo exploradas, usando a tecnologia de carpete inteligente para localizar e identificar usuários.

O **espelho interativo** versa um espelho que possui funções de uma (tela sensível ao toque), enquanto adiciona interatividade por aceitar entradas do usuário. Para explorar as possibilidades oferecidas por esse equipamento, foi montado um protótipo no banheiro do *HomeLab*, com alguns serviços como: fornecimento de informações sobre as condições do tempo e trânsito; examinar peso; mostrar opções de penteado; e controle de iluminação do ambiente. Para ativar os controles do espelho interativo, dois conceitos de interação foram desenvolvidos: a borda e por proximidade. Esses conceitos combinam as vantagens de apontamento direto e, ao mesmo tempo, não requer toques no espelho, evitando marcas de digitais.

Os pesquisadores da *Philips* acreditam que o futuro dos ambientes inteligentes não será regido pelas tecnologias, mas pelas pessoas, através de decisões em relação às mudanças em suas vidas (RUYTER, 2003). Por essa razão, acreditam na validade de seu projeto como meio de alcançar essa aproximação entre sua tecnologia e os desejos de seus consumidores.

Outro projeto de ambiente inteligente em destaque é o *PlaceLab*. Localizado em Cambridge, Massachusetts, o *PlaceLab* é um apartamento totalmente funcional e altamente flexível, construído pelas instituições TIAX/LLC e MIT. Para habilitar o estudo das reações comportamentais dos indivíduos, mediante tecnologias domésticas, alguns voluntários ocupam o laboratório periodicamente e suas ações são gravadas para uma análise posterior. O interior do apartamento é formado por 15 móveis pré-fabricados e pré-configurados. Cada um deles contém um micro-controlador, um sistema multimídia e uma rede de 25 a 30 sensores (INTILLE, 2005). Novos sensores podem ser facilmente adicionados à rede, sendo que os atuais gravam dados audiovisuais das atividades dos ocupantes. Todos os dispositivos de sensoriamento são discretamente integrados aos móveis, ferramentas e utensílios da casa.

Para estudar a interação dos indivíduos com os objetos do ambiente e medir a ace-

²A computação ciente de contexto estuda aplicações que adaptem seu comportamento com base em informações (ou contexto) capturadas de um ambiente físico com algum suporte computacional (KUDO, 2004).

lação em diferentes partes do corpo humano, foram desenvolvidos sensores ambientais chamados *MITes* (*MIT Environmental Sensors*). Com dimensões aproximadamente de 3,81 x 2,54 cm, um *MITes* é composto por cinco componentes principais (TAPIA, 2004): (1) um Chip nRF24E1³, com conexão sem fio operando em 2.4 GHz; (2) uma EEPROM de 4 KB; (3) acelerômetros ADXL202/210; (4) uma antena de 50 ohms; e (5) uma bateria CR2032. Para o uso desses dispositivos no *PlaceLab*, o *MITes* foi posteriormente modificado, incluindo versões com: (a) acelerômetros de 2 e 3 eixos para captar movimentos em objetos e no corpo humano, respectivamente; (b) chaves magnéticas para detectar eventos ligado-desligado e aberto-fechado de gavetas, portas, etc.; (c) sensores de temperatura, luminosidade para controle dos níveis ambientais e pressão arterial e frequência cardíaca para monitoramento dos dados vitais de um indivíduo; e (d) leitor RFID, como proposta de identificação do habitante. Adicionando ao *MITes* a flexibilidade de embutir qualquer outro sensor, mantendo os padrões de conectividade com a rede e transmissão dos dados.



FIG. 2.2: Imagens do PlaceLab. (a) PlaceLab. (b) Software mostrando as leituras dos sensores.

A estrutura de sensores do *PlaceLab* hospeda 80 (oitenta) pequenos sensores com fio que detectam eventos “ligado-desligado” e “aberto-fechado”, tais como ligar e desligar do forno e abrir e fechar de portas. As condições ambientais do *PlaceLab* são obtidas usando 34 sensores de temperatura, 10 de umidade, 5 de iluminação e 1 barômetro, distribuídos pelos cômodos. Além disso, nove câmeras coloridas, nove câmeras infravermelhas e 18 mi-

³Nordic VLSI Semiconductors.

crofones estão instalados nos móveis distribuídos na casa e sobre superfícies de trabalho, tais como escritório e cozinha. Vinte computadores usam algoritmos de processamento de imagens para selecionar 4 *streams* de vídeo e 1 *stream* de áudio, que melhor captura o comportamento do ocupante. Esses dados são sincronizados com os dados dos demais sensores, do ambiente ocupado pelo indivíduo, e são armazenados para uma análise posterior. O objetivo desse procedimento é testar algoritmos de detecção de contexto e fabricar protótipos de aplicações orientadas ao contexto para dispositivos móveis e computadores pessoais.

Além da infra-estrutura de sensores, o *PlaceLab* possui: controle da iluminação (a intensidade e temperatura da cor da luz podem ser dinamicamente controladas em cada cômodo); e controle de temperatura, além de um sofisticado filtro de ar permitindo controle dinâmico das qualidades ambientais da casa.

De posse do conhecimento da infra-estrutura do *PlaceLab*, as áreas de estudo atualmente desenvolvidas são: (a) cuidados médicos pró-ativos e fornecimento de informações *just-in-time*; (b) monitoramento de atividades do dia-a-dia; (c) monitoramento biométrico; (d) qualidade do ar dentro do ambiente; e (d) energia e eficiência.

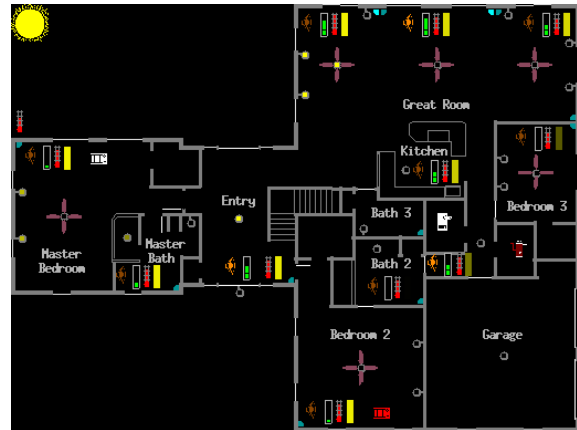
O projeto *Adaptive House* é um ambiente inteligente implementado em uma casa real, com o foco de se adaptar e se atualizar de acordo com as mudanças nos hábitos de seus moradores. O objetivo principal do projeto é desenvolver uma casa que essencialmente aprenda os padrões ambientais de seus ocupantes e programe-se para regular as condições do ambiente automaticamente, segundo o estilo de vida e os desejos de seus habitantes. A casa é equipada com sensores e atuadores que monitoram e regulam as condições ambientais da casa. Mais especificamente, os sensores e atuadores da casa podem: medir e regular os níveis de iluminação, temperatura da água, velocidade dos ventiladores, intensidade sonora e temperatura do ambiente; medir a temperatura e luminosidade do lado de fora da casa; ligar e desligar os aquecedores de água e ar; ligar e desligar o gás do forno; saber o status de janelas e portas (aberta ou fechada); medir gastos de gás, energia elétrica e água; e detectar movimentos. A FIG. 2.3a mostra a planta baixa da *Adaptive House* indicando a localização dos sensores e atuadores.

Algumas funcionalidades desse ambiente incluem:

- Prever quando os ocupantes irão retornar à casa e determinar quando iniciar o sistema de aquecimento da casa para que, ao chegar, os ocupantes encontrem a casa em uma temperatura confortável;



(a) O interior da Adaptive House.



(b) Planta baixa da *Adaptive House* indicando a localização dos sensores e atuadores (MOZER, 1998).

FIG. 2.3: Foto do interior da Adaptive House e o programa que monitora os sensores.

- Detectar padrões estatísticos do uso de água, de maneira que se a água quente é usada ao meio dia nos dias da semana, o aquecedor permaneça desligado nas demais horas do dia;
- Inferir a localização do ocupante e em quais atividades o mesmo está envolvido e controlar os padrões de iluminação;
- Acionar o sistema de iluminação antes de algum ocupante entrar no cômodo.

O sistema de controle desenvolvido para o projeto *Adaptive House* é chamado de *Adaptive Control of Home Environments - ACHE*. Ele é responsável por controlar sistemas básicos de conforto residencial - temperatura da água, iluminação ventilação e temperatura do ambiente. O ACHE foi configurado para atingir dois objetivos: (1) Preparar as condições ambientais da casa, segundo as preferências dos habitantes, antecipadamente. Dessa forma, ao adaptar as condições ambientais de algum cômodo e algum habitante alterar manualmente algum sistema, o ACHE encara como sinal de treinamento; e (2) Conservação de energia. Além disso, o ACHE é capaz de controlar todos os sensores e atuadores da casa (MOZER, 1998).

A arquitetura do sistema ACHE é formada por cinco blocos principais: *state transformation*, *occupancy model*, *predictors*, *setpoint generator* e *device regulator*. O módulo *state transformation* computa os dados oriundos dos sensores no instante atual, e os envia aos módulos *state transformation* e *occupancy model*. No primeiro, levando em conta os

dados de tempos passados, são computados as médias, mínimos, máximos e variâncias, produzindo mais informações sobre o ambiente do que os valores instantâneos. No segundo, é determinado para cada cômodo, os estados de “ocupado” e “não ocupado”. Além disso, são geradas regras que dizem essencialmente: “um determinado cômodo continua com o status *ocupado*, mesmo que não exista movimento, até que o cômodo adjacente seja desocupado”. Os três outros módulos são responsáveis pelo comportamento adaptável do ambiente. O módulo “predictors” se atenta ao estado atual e aos estados futuros. Exemplos de predição incluem: padrões ocupacionais esperados na casa nas próximas cinco horas; ou o uso de água quente esperado. O módulo de predição é implementado com redes neurais *feed forward* treinadas com o algoritmo *back propagation*. Dado as predições dos estados futuros, as decisões de controle necessitam ser feitas de acordo com o gasto de energia dos dispositivos da casa. O processo de decisão é dividido em dois estágios. O módulo “setpoint generator” determina o ponto de ajuste, especificando o valor-alvo de alguma variável ambiental (níveis de luminosidade, temperatura do ar, etc.) sob uma janela de tempo. O módulo “device regulator” controla os dispositivos físicos para alcançar o ponto de ajuste, determinado pelo módulo anterior. A razão para dividir o controle entre os módulos “setpoint generator” e o “device regulator” é encapsular o conhecimento. O “setpoint generator” requer conhecimento sobre as preferências dos habitantes, enquanto que o “device regulator” tem o conhecimento sobre as condições físicas do ambiente e o gasto dos dispositivos controlados. Se os habitantes e/ou suas preferências mudarem no tempo, somente o “setpoint generator” necessita se atualizar.

Para avaliar o ACHE, testes de simulação utilizando o sistema de temperatura, usando dados reais dos ocupantes e perfis de condições de temperatura exteriores à casa, foi possível testar diversas políticas de controle. Dentre elas, o ACHE melhor executou robustamente três políticas alternativas, mostrando um custo total mais baixo (desconforto + energia), através de uma taxa de valores para o custo relativo de desconforto e o grau de não-determinismo nos padrões dos habitantes.

Os pesquisadores da *Adaptive House* concluem chamando a atenção para o cuidado ao avaliar o ACHE. Devido ao fato da maioria das pessoas afirmarem que seus horários não seguem um padrão, esse comportamento sutil também é útil ao ACHE. Pois acredita-se que tais padrões possam ser explorados de forma útil, no controle adaptativo de ambientes domésticos.

2.2 REVISÃO DA LITERATURA

Nessa seção serão discutidos os trabalhos de (LI, 1991), (SHOJI, 2004), (SHOJI, 2005) e (ITAI, 2006). Tais trabalhos foram escolhidos, por tratarem do problema de classificação, de gênero e de indivíduos, a partir do som produzido pelo caminhar. Cada um deles utilizou um conjunto de características, filtros e algoritmos de classificação. O método utilizado por cada um, será detalhado a seguir.

O trabalho congregado em (LI, 1991), consistiu em avaliar a viabilidade da classificação do gênero, a partir do som dos passos. Para tanto, foram realizados quatro experimentos, sendo os experimentos 1, 3 e 4 de caráter comportamental⁴ e o experimento 2 de análise estatística espectral do sinal sonoro proveniente dos passos. Para a execução dos experimentos, 16 indivíduos (8 homens, rotulados de M1 a M8 e 8 mulheres, rotuladas de F1 a F8) foram convidados a participar do processo de gravação do som de seus passos. Para avaliar a influência do calçado no julgamento do gênero, foi solicitado que os indivíduos M4 e M8 utilizassem o calçado do indivíduo M3, os demais participantes utilizaram seus próprios calçados. Por 3 vezes, cada indivíduo deu 8 passos em um palco de madeira, com 7,3 m de comprimento por 1 m de largura, sendo esses passos gravados a uma taxa de amostragem de 10 KHz com 12 *bits* de quantização. Para evitar possíveis propriedades irregulares associadas com o início e término da caminhada, somente os 4 passos intermediários da sequência foram utilizados nos experimentos, sendo descartados os 2 passos iniciais e finais.

O experimento 1 consistiu em selecionar um grupo de indivíduos para ouvir as sequências gravadas e para julgar se o som produzido pela respectiva sequência, foi produzido por um homem ou uma mulher. Nesse primeiro experimento, 69 % dos homens e 25.5 % das mulheres foram julgados(as) como homens.

O experimento 2 do trabalho de (LI, 1991) se divide em duas etapas: (1) análise do ritmo de passada e proporções temporais das fases de apoio e balanço; e (2) análise das seguintes medidas estatísticas: média; moda; inclinações de alta e baixa frequências; assimetria e curtose; e quatro momentos (todas as medidas estatísticas se basearam na distribuição de frequências obtidas pela transformada rápida de Fourier⁵). Nas duas etapas, são verificados seus relacionamentos com o julgamento obtido no experimento 1.

⁴No sentido de avaliar se os indivíduos eram capazes de identificar o gênero de outros indivíduos, ouvindo-se o som dos passos, original e intencionalmente modificado

⁵A transformada rápida de Fourier será abordada no Capítulo 3.

Na primeira etapa do experimento 2 (dois) do trabalho de (LI, 1991), um teste de hipótese foi conduzido nos dados oriundos da média do ritmo de passada e da porcentagem de duração das fases de apoio e balanço. O teste mostrou que indivíduos do sexo feminino têm duração da fase de apoio maior que os do sexo masculino. Para examinar se o ritmo de passo e a relativa diferença temporal nas fases de apoio e balanço entre as duas classes afetam o julgamento do gênero, coeficientes de correlação de Pearson foram calculados entre o resultado do experimento 1 e estes dados. O teste mostrou que o sexo dos indivíduos, com um ritmo de passada mais rápido, tende a ser julgado como feminino, apesar do ritmo da passada em média não diferenciar os gêneros significativamente.

Na segunda etapa do experimento 2 (dois) do trabalho de (LI, 1991), foram utilizadas sete características estatísticas de distribuições espectrais obtidas através da transformada de Fourier. A partir dos dados oriundos da extração dessas características, uma análise de componentes principais foi realizada e dimensões hipotéticas extraídas. Com o critério de componentes com autovalores maiores que 1, dois componentes, PC1 e PC2, foram selecionados para explicar o relacionamento dos dados. Aplicando-se então uma função discriminante linear, técnica de classificação e previsão utilizada em estatística, obteve-se o seguinte resultado para a classificação: 3 dos 8 homens foram classificados como mulheres; e 1 mulher classificada como homem. Utilizando esse resultado, foi realizada uma correlação de Pearson com os resultados do experimento 1 para identificar os atributos físicos usados pelos ouvintes. Como conclusão desse procedimento, verificou-se que tendem a ser julgados como mulheres, os indivíduos que apresentam estímulos com: (a) altos valores para média e moda espectrais; (b) valores significativos para energia em altas frequências. Em contrapartida, tendem a ser julgados como homens, os indivíduos que apresentam: (a) baixos valores para a média e moda espectral; (b) altos valores de assimetria e curtose; e (c) baixo crescimento e queda espectrais.

O experimento 3 foi conduzido para avaliar os efeitos na percepção do gênero ao se efetuar manipulações no espectro. Através de um equalizador de frequências, a moda espectral dos estímulos dos indivíduos M6 e F7 (no experimento 1, apresentaram porcentagem relativamente neutra em relação ao julgamento do gênero), foi configurada com os valores 65, 125, 250, 500 e 1000 Hz. Dez sujeitos ouviram os estímulos modificados e classificaram o gênero. Em resumo, observou-se que quando a moda foi: (1) movida para baixas frequências, o julgamento de masculinidade cresceu de 50% para 90%; e (2) movida para 1000 Hz, o julgamento de masculinidade caiu para 10%. Com esse experimento, foi

possível concluir que: (1) manipulações no espectro podem alterar o julgamento do gênero; e (2) a moda espectral tem um papel importante na classificação do gênero.

Por fim, o experimento 4 consistiu no estudo da influência do tipo de calçado no julgamento do gênero. Para tanto, 11 sequências de passos foram gravadas, utilizando-se 1 homem, M7 calçando dois tipos de calçados (A e B), e três mulheres, F2, F6 e F8 calçando seus próprios calçados e os de M7. Por 3 vezes, as 11 sequências foram ouvidas e julgadas por um grupo de 8 pessoas. A partir dos dados oriundos desse teste, as probabilidades dos indivíduos F2, F6 e F8 serem julgadas como homem, respectivamente, foram as seguintes: (1) com o sapato A: 0.70, 0.71 e 0.92; (2) com o sapato B: 0.44, 0.32 e 0.39; e (3) com seus próprios sapatos: 0.06, 0.3 e 0.39. A probabilidade do indivíduo M7 ser julgado como homem foi de 0.83, usando o sapato A, e 0.40, utilizando o sapato B. Esse simples experimento mostrou que o tipo de calçado influencia o julgamento do gênero.

Nos trabalhos (SHOJI, 2004), (SHOJI, 2005) e (ITAI, 2006), foram utilizados uma mesma base de dados. Essa base foi composta da gravação dos passos de cinco estudantes do sexo masculino, calçando um único tipo de calçado (chinelos). O ambiente por onde caminharam foi um corredor de uma universidade. Cada indivíduo caminhou dez vezes um percurso de cinco metros até o microfone, onde foram gravadas suas passadas em uma taxa de amostragem de 44.1 KHz e o microfone utilizado foi um SOKKI LA-5120, com uma taxa de frequência de 20 a 12.5 KHz. Os três trabalhos, organizaram-se basicamente em três fases: (1) pré-processamento; (2) extração das características; e (3) classificação. Cada uma dessas fases será detalhada a seguir, levando-se em conta a particularidade de cada um.

Em (SHOJI, 2004) se encontra um dos primeiros trabalhos encontrado na literatura tendo em vista a classificação de indivíduos, baseando-se na extração de características do som dos passos. Como pré-processamento, utilizou-se a *transformada wavelet* como filtro do ruído ambiente. Obtido gravando-se o som de fundo do ambiente, um limiar de 0.017 foi passado como parâmetro à transformada. Na fase de extração das características, três parâmetros foram extraídos dos sinais: a análise mel-cepestral, o ritmo de caminhada e o grau de similaridade do envelope espectral.

A primeira característica considerada em (SHOJI, 2004) foi a análise mel-cepestral. Esta consiste em uma técnica muito utilizada em síntese de voz e reconhecimento de locutor (maiores detalhes no Capítulo 3) e no trabalho de (SHOJI, 2004), foi conside-

rado um filtro MLSA (*Mel Log Spectrum Approximation*) para obtenção dos coeficientes mel-cepestrais. Uma amostra de cada um dos indivíduos foi escolhida aleatoriamente e submetida ao filtro. Como resultado desse procedimento, foi notificada uma diferença entre os indivíduos, ao se observar os primeiros picos de frequência. Dessa maneira, os 6 (seis) primeiros picos de frequência foram selecionados para compor um vetor de características $F = \{f_1, f_2, \dots, f_6\}$. A reunião dos vetores F , calculados para cada amostra da base de dados, formou o conjunto $A = \{F_1, F_2, \dots, F_M\}$, onde $M = 10$ é o número de medidas feitas para cada sujeito. O conjunto formado pelos vetores A_i de cada sujeito i foi chamado de $\alpha = \{A_1, A_2, \dots, A_N\}$, onde N é o número de indivíduos.

A segunda característica explorada no trabalho de (SHOJI, 2004) foi o ritmo de caminhada. Para calculá-lo, foi assumido que o ritmo de caminhada é equivalente ao período fundamental da onda. Um algoritmo para extração da frequência fundamental utilizando uma função de autocorrelação, foi utilizado. Para cada amostra i , um período médio p_i foi calculado e um conjunto $B_i = \{p_1, p_2, \dots, p_M\}$ foi formado para cada sujeito i . O conjunto de todos os B_i formaram o vetor de características β , definido por $\beta = \{B_1, B_2, \dots, B_N\}$, armazenando os valores dos intervalos entre os passos.

A terceira característica explorada no trabalho de (SHOJI, 2004) foi o grau de similaridade do envelope espectral (EE). O EE é a representação da forma de onda no domínio da frequência. Para obter um efeito de “suavização”, o EE utilizado nesse trabalho foi submetido a um filtro de média móvel. Dado um sinal de entrada $y(n)$, o processo de média móvel calcula a média para os N pontos passados do valor $y(n)$ presente e o sinal $\bar{y}(n)$ é obtido novamente. Ao se calcular os valores dos gradientes dos EEs, em amostras de diferentes indivíduos, notou-se diferenças entre eles. Portanto, o seguinte procedimento foi tomado: (1) calcular o EE de cada amostra; (2) escolher arbitrariamente um EE como base; e (3) calcular o grau de similaridade, definido pela distância euclidiana entre o EE base e os demais EEs. Chamando k o valor obtido por esse último cálculo, definiu-se $C = \{k_1, k_2, \dots, k_M\}$, o conjunto dos graus de similaridade das M amostras de cada sujeito. Por fim, um vetor $\gamma = \{C_1, C_2, \dots, C_N\}$ formado pelos N sujeitos foi definido como o vetor de características representando o grau de similaridade.

Para avaliar a capacidade de identificação de indivíduos através dos três conjuntos de características descritos anteriormente, o algoritmo *K-Means* foi escolhido a tarefa de classificação. Em síntese, o algoritmo K-Means é um algoritmo de agrupamento (clusterização) em que um de seus pontos principais é a calibração de um elemento chamado

centróide, o qual é ajustado a cada interação. O algoritmo pára quando não há mais atualizações no elemento centróide. Cada centróide representa um grupo (nesse caso, o indivíduo) e um elemento (uma amostra daquele determinado indivíduo) pertence a um determinado grupo, se a distância entre ele e o centróide do grupo for menor que determinado limiar. Durante o teste de classificação, foi submetido ao algoritmo combinações de conjuntos para avaliar a eficácia das combinações. Primeiramente, o algoritmo rodou tendo como entrada somente os conjuntos α , β , γ , com 60 % (30/50), 54 % (27/50) e 52 % (26/50) de acertos, respectivamente. Por último, as combinações dos conjuntos identificadas por $P1\{\alpha, \beta\}$, $P2\{\alpha, \gamma\}$, $P3\{\beta, \gamma\}$, $P4\{\alpha, \beta, \gamma\}$ ao serem submetidas ao algoritmo K-Means, a acurácia foi aumentada para 100 %, 84 %, 86 % e 100 %, respectivamente.

O trabalho publicado em (SHOJI, 2005) se diferencia do anterior ((SHOJI, 2004)) em dois aspectos principais: (1) na etapa de pré-processamento, como técnica de extração do ruído ambiente, foi utilizado um método de subtração espectral; e (2) um algoritmo de quantização vetorial (QV) foi utilizado como método de classificação. As amostras foram submetidas à etapa de extração de características antes e após a extração do ruído. Considerando a mesma organização dos conjuntos de características definidos em (SHOJI, 2004), para os conjuntos α , β , γ , as taxas de classificação foram 60 %, 54 % e 52 %, com ruído ambiente e 64 %, 54 % e 58 %, após a extração do ruído, respectivamente. Ao realizar as mesmas combinações: $P1\{\alpha, \beta\}$, $P2\{\alpha, \gamma\}$, $P3\{\beta, \gamma\}$, $P4\{\alpha, \beta, \gamma\}$, a acurácia na classificação cresceu para 80 %, 84 %, 80 % e 100 %, com ruído ambiente e 82 %, 88 %, 84 % e 100 %, após a extração do ruído, respectivamente.

Em (ITAI, 2006), não houve pré-processamento. Na fase de extração de características foram considerados quatro parâmetros psicoacústicos: *sonoridade*, *sharpness*, *fluctuation strength* e *roughness*. Os parâmetros psicoacústicos são medidas baseadas em aproximações do sentido do ouvido, ou seja, são parâmetros modelados levando em conta as reações do sistema auditivo humano. O cálculo desses parâmetros dependem dos modelos utilizados para tal. O modelo utilizado no trabalho de (ITAI, 2006) foi o de *E. Zwicker*. Este modelo produz a densidade de *sonoridade* em 24 bandas críticas, onde no trabalho de (ITAI, 2006) foram consideradas 20 bandas críticas. O nível de *sonoridade* total é obtido somando a resposta em cada uma das bandas. Os demais parâmetros são escalares obtidos através da densidade de *sonoridade*. Para avaliar a qualidade dos parâmetros psicoacústicos, dois testes foram conduzidos utilizando-se o algoritmo *K-Means* para a classificação. O primeiro consistiu em juntar os quatro parâmetros calculados de cada

amostra, como um vetor de características e submetê-los ao algoritmo. Como resultado do primeiro teste, 3 dos 5 indivíduos foram agrupados corretamente. O segundo teste, consistiu em: (1) definir outro vetor de características utilizando somente a *sonoridade* específica em torno de 20 bandas críticas, para cada amostra; (2) para cada sujeito, separar dois conjuntos: um de teste contendo 1 vetor e um de treinamento contendo 9 vetores; e (3) associar o vetor de teste ao centróide do grupo que produzir o menor erro absoluto, obtido pela soma das distâncias entre o elemento teste e o centróide. Como resultado do segundo teste, 44 amostras foram corretamente identificados.

Em resumo, o trabalho apresentado em (LI, 1991) mostrou a qualidade das características estatísticas na classificação do gênero, a partir dos passos dos indivíduos, como também fez um estudo no relacionamento destas características com a percepção dos ouvintes, no julgamento do gênero. O trabalho apresentado nessa dissertação se diferencia por utilizar parâmetros objetivos e subjetivos, no intuito de identificar habitantes na casa inteligente a partir do som de seus passos. Os trabalhos reunidos em (SHOJI, 2004), (SHOJI, 2005) e (ITAI, 2006) mostraram a possibilidade de identificar os indivíduos através do sinal sonoro provocado pelo caminhar. Entretanto, o conjunto de teste foi bastante restrito, limitando-se a 5 indivíduos com um tipo de calçado, como também somente um algoritmo de classificação foi avaliado. Uma das etapas do trabalho atual consistiu em realizar um experimento com 15 (quinze) indivíduos calçando 6 (seis) tipos distintos de calçados (vide Capítulo 5). Dois algoritmos para a tarefa de classificação foram implementados utilizando técnicas distintas de partição de dados. Além disso, foi implementado um critério de seleção de atributos, para diminuição da dimensionalidade dos dados.

Após a apresentação das pesquisas mais recentes na área de ambientes inteligentes e a discussão dos trabalhos relacionados diretamente com o foco do presente trabalho, é necessário apresentar os fundamentos teóricos envolvidos nessa obra, que estão reunidos no próximo capítulo.

3 FUNDAMENTAÇÃO TEÓRICA

No capítulo anterior foram reunidas as recentes pesquisas em ambientes inteligentes, como também foram abordados trabalhos que tratam da identificação de indivíduos usando a análise do sinal sonoro dos passos. O atual capítulo fornece subsídios para o entendimento da análise acústica de sinais (Seção 3.1), como também as características e seus respectivos métodos de extração (Seção 3.2), para então serem posteriormente processadas no intuito de encontrar padrões no caminhar. Devido à magnitude dos dados provenientes dos parâmetros estudados, fez-se necessária a busca por técnicas de seleção de atributos, com o objetivo de contribuir na qualidade de classificação e consequentemente no desempenho do sistema. As técnicas de seleção de atributos estão reunidas na Seção 3.3. Por fim, a Seção 3.4 encerra a fundamentação teórica com a descrição dos algoritmos de classificação avaliados.

3.1 CONCEITOS BÁSICOS

Em todo ciclo de vida, o ser humano é rodeado de sinais que carregam consigo informações ajudando-o a tomar decisões. Do ponto de vista da engenharia, sinais são funções ou sequências que servem para transportar informação de uma origem para um destino. O som é um sinal que propaga no ar através de um movimento ordenado das partículas constituintes, provocando variações na pressão do ar. Essas variações são percebidas pelo sistema auditivo humano, constituindo a sensação de audição. Assim como os demais sinais na natureza, o som em sua forma natural é um sinal contínuo. Devido às limitações dos sistemas digitais, faz-se necessário trabalhá-lo com sua versão discreta. Fazendo-se necessário ainda, transformar sua forma analógica para digital, para então poder ser processada como informação. Esse processo, chamado de conversão analógico-digital A/D, é realizado em duas etapas: (1) transformação da variação de pressão sonora em sinais elétricos; e (2) amostragem e codificação (também conhecida como quantização) dos sinais oriundos da primeira etapa.

Um dispositivo chamado transdutor acústico-elétrico (microfone) é o responsável por transformar a variação de pressão sonora do ar, em variação de tensão elétrica. Este dispositivo é conectado a uma interface especial chamada *placa de som*. Esta por sua vez,

realiza as etapas de amostragem e codificação. Cada amostra é medida em uma determinada frequência, denominada *taxa de amostragem* (tempo de amostragem e frequência de amostragem também são termos utilizados), e codificada em uma quantidade definida de *bits*, conhecida como *taxa de quantização*. Para obter a representação da onda analógica em forma digital sem perda significativa de informação, é preciso escolher uma taxa de amostragem e quantidade de *bits* de quantização adequados. Se a taxa de amostragem for escolhida acima da maior frequência contida no sinal, haverá informação redundante no sinal amostrado. Por outro lado, se for escolhida uma taxa de amostragem menor que a maior frequência, haverá perda de informação. Nesse sentido, para encontrar uma relação entre as frequências contidas em um sinal e a devida taxa de amostragem necessária para digitalizá-lo, Shannon⁶ (SHANNON, 1949) desenvolveu o *Teorema da Amostragem* que diz, essencialmente:

“Um sinal analógico contendo componentes acima de alguma frequência máxima f_1 Hz pode ser completamente representado por amostras regularmente espaçadas, desde que a taxa de amostragem seja pelo menos $2f_1$ amostras por segundo.”

Durante o processo de conversão A/D, existe ainda um problema chamado falseamento (do inglês *aliasing*). Esse problema ocorre quando o sinal a ser digitalizado possui frequências mais altas que a frequência de *Nyquist*. Quando isso ocorre, as frequências acima de $fs/2$ são introduzidas no sinal amostrado como um ruído de frequências baixas. Para resolver tal problema, os conversores A/D aplicam um filtro conhecido como *anti-aliasing filter* antes da amostragem, o qual permite a passagem somente de frequências abaixo da metade da frequência de amostragem. A conversão digital para analógico D/A faz o processo inverso, lendo as amostras e gerando os pulsos elétricos correspondentes. O processo de conversão analógico digital pode ser visto na FIG. 3.1.

Dois conceitos comumente utilizados são a energia e a potência de um sinal $x(n)$. A potência $p_x(n)$ é definida por

$$p_x(n) = |x(n)|^2 \tag{3.1}$$

onde n representa o número da amostra analisada. E a energia pela seguinte fórmula

$$E_x = \sum_{n=-\infty}^{\infty} p_x(n) \tag{3.2}$$

⁶A frequência máxima contida em um sinal amostrado de acordo com o teorema da amostragem, $fs/2$, é conhecida como frequência de *Nyquist*.

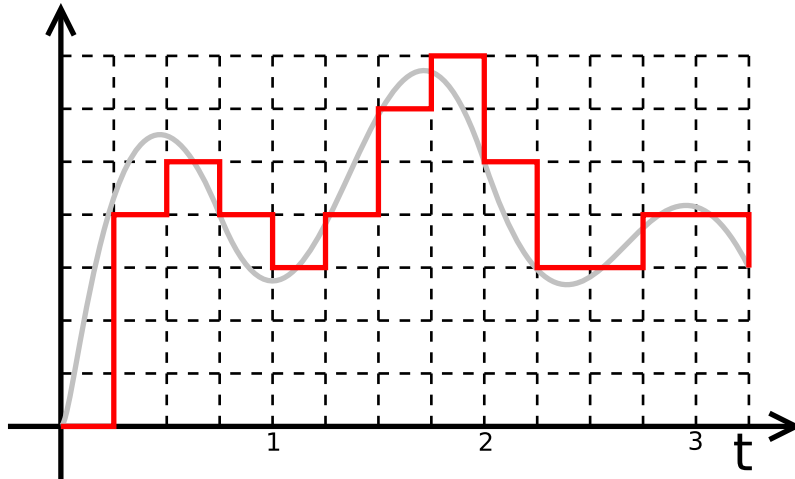


FIG. 3.1: Processo de amostragem de um sinal analógico.

Dessa maneira, tais valores não significam muito, porém são utilizados para comparar dois sinais. Para uma melhor representação, foi criado a unidade decibel (dB) definida por

$$R = 10 \log_{10} \frac{E_1}{E_2} \quad (3.3)$$

onde R é a relação em dB de dois sinais com energias E_1 e E_2 , respectivamente. O decibel geralmente é escolhido para uso prático, devido a taxa dinâmica do ouvido humano ser muito larga, os valores em decibel permanecem em uma taxa tratável.

Após a conversão digital, o sinal $x(n)$ é representado no domínio do tempo como uma sequência de números. De acordo com a idéia inicial de Fourier, qualquer função, mesmo as não periódicas, podem ser representadas como uma soma de senóides (periódicas). No domínio dos números complexos, uma senóide pode ser expressa pelo Teorema de Euler:

$$e^{j\omega n} = \cos \omega n + j \sin \omega n \quad (3.4)$$

com a frequência angular ω dada por

$$\omega = \frac{2\pi}{N} k \quad (3.5)$$

para alguma frequência discreta k e N bandas de frequência, ou, relacionado a frequências em Hertz (Hz), como

$$\omega = \frac{2\pi}{f_s} f \quad (3.6)$$

Portanto, para revelar a estrutura de frequência de x , este pode ser convertido em seu domínio da frequência X pela transformada discreta de Fourier (DFT):

$$X(k) \triangleq \sum_{n=0}^{N-1} e^{j\frac{2\pi}{N}nk} x(n) \quad (3.7)$$

X é chamado de espectro de frequências do sinal digital x de tamanho N . Para transformar X para o domínio do tempo, usa-se a transformada discreta inversa de Fourier:

$$x(n) \triangleq \frac{1}{N} \sum_{k=0}^{N-1} e^{-j\frac{2\pi}{N}nk} X(k) \quad (3.8)$$

O resultado da transformada de Fourier (eq. 3.7) são N números complexos $X(k)$ os quais definem o espectro de magnitude $M(k)$ e o espectro de fase $\phi(k)$ para a sequência discreta k :

$$M(k) = |X(k)| \quad (3.9)$$

$$\phi(k) = \arctan(X(k)) = \arctan \left| \frac{reX(k)}{imX(k)} \right| \quad (3.10)$$

onde $reX(k)$ e $imX(k)$ representam a parte real e imaginária de $X(k)$.

O espectro de magnitude $M(k)$ dá a intensidade de uma senóide na frequência k , enquanto que o espectro de fase mostra seu deslocamento no tempo. Contudo, o espectro de fase frequentemente é ignorado e o termo espectro de Fourier ou espectro é usado para o espectro de magnitude.

Existe uma generalização da transformada de Fourier⁷, chamada transformada Z , o qual simplifica as representações no domínio da frequência. A transformada Z de uma sequência discreta $x(n)$, onde z é uma variável complexa, é definido como:

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n} \quad (3.11)$$

A complexidade computacional da DFT, quando calculada diretamente, é $O(N^2)$, onde N é o número de amostras constituintes do sinal de entrada. Fazendo uso da periodicidade da senóide em análise $e^{-j\frac{2\pi n}{N}k}$ e aplicando o princípio *dividir para conquistar*

⁷Da mesma maneira que, para o caso contínuo, a transformada de Laplace é a generalização da transformada contínua de Fourier

particionando o problema em subproblemas menores, uma variedade de algoritmos mais eficientes de complexidade $O(N \log N)$ foram desenvolvidos. Tais algoritmos são conhecidos coletivamente como transformada rápida de Fourier (FFT do inglês *Fast Fourier Transform*). Para sua aplicação, geralmente é necessário que N seja múltiplo de dois. Como a DFT e sua inversa diferem somente no sinal do expoente e no fator de escala, os mesmos princípios levam à versão rápida da FFT.

Para o uso de técnicas convencionais aplicadas à sinais do mundo real, é necessário trabalhar com pequenos intervalos do sinal, supostos estacionários. Supondo que o som dos passos é um processo estocástico, em geral não estacionário, muitas partes da onda acústica podem ser supostas estacionárias num curto intervalo de tempo. Este intervalo caracteriza o tamanho da janela de análise a ser utilizada, em cuja duração, de 10 a 40 ms, o som dos passos pode ser considerado como um processo estacionário. Uma janela de comprimento longo tende a produzir uma melhor representação espectral do sinal, desde que este esteja na região de estacionariedade. Em contrapartida, uma janela de comprimento curto tende a ser melhor em análises no domínio do tempo. Objetivando a atenuação no efeito de *Gibbs* (*ripple* em amplitude na resposta em frequência da janela retangular) devido ao truncamento do sinal de análise no domínio do tempo, deve-se utilizar janelas que possuam, no domínio da frequência, um lóbulo principal o mais estreito possível e uma grande diferença de amplitude entre o lóbulo principal e o primeiro lóbulo lateral. As janelas mais utilizadas na prática e que procuram atender tais condições são as seguintes: Hamming, Hanning, Retangular, Bartlett (triangular), Blackman e Kaiser (JR., 1993). Cada janela representa uma parte do sinal do qual serão extraídas as características desejadas, uma visualização de cada função janela citada anteriormente se encontra na FIG. 3.2. Após esse processamento, conhecido como janelamento, cada parte do sinal é chamada de quadro.

Para compensar a suavização da janela temporal, uma sobreposição entre as janelas é efetuada aumentando a correlação entre janelas adjacentes. A porcentagem da sobreposição é dada por:

$$Sobr = \frac{T_{jan} - T_{inc}}{T_{jan}} \times 100\% \quad (3.12)$$

onde T_{jan} e T_{inc} são o tamanho da janela e o tamanho do incremento para uma nova janela, respectivamente, ambos em unidade de tempo ou em número de amostras.

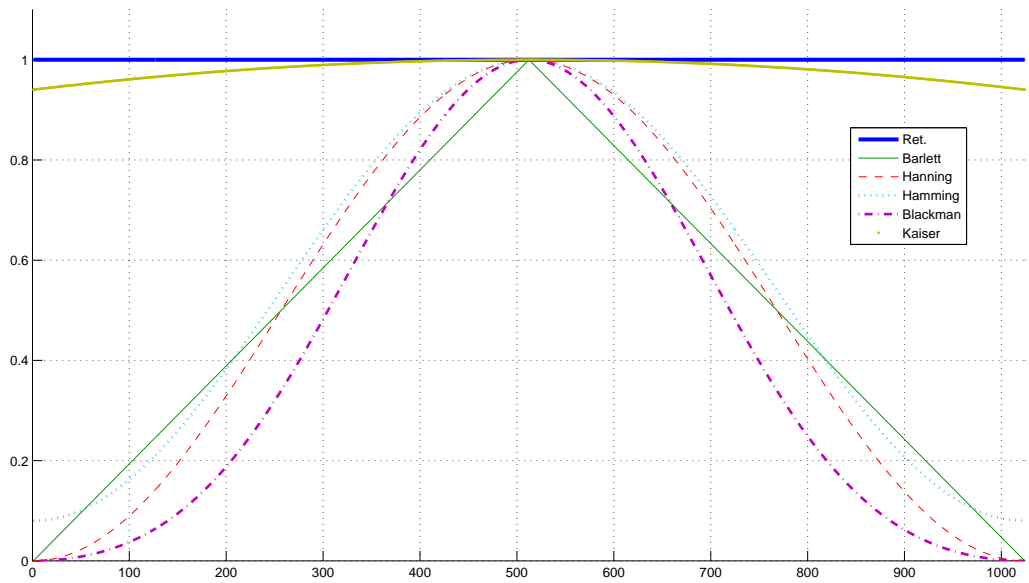


FIG. 3.2: Funções janelas.

Sobreposições típicas estão na ordem de 50% ou mais (JR., 1993).

MÉTRICAS DE DISTÂNCIA

O conceito de distância é formalizado como uma função $D : E \times E \rightarrow R$ (a cada par de pontos associa um valor real) que atente às seguintes restrições (LIMA, 1970):

- $D(x, x) = 0$;
- $D(x, y) = D(y, x)$;
- $D(x, y) \leq D(x, z) + D(z, y)$.

Algumas distâncias conhecidas na literatura são mostradas a seguir.

- Euclidiana: $d(X, Y) = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2}$;
- Hamming: $d(X, Y) = \sum_{i=1}^n |X_i - Y_i|$;
- Minkowsky: $d(X, Y) = (\sum_{i=1}^n |X_i - Y_i|^p)^{\frac{1}{p}}$;
- Bhattacharyya: $B_{k,i} = \frac{1}{8}(\mu_k - \mu_i)^t \left[\frac{\sum_k + \sum_i}{2} \right]^{-1} (\mu_k - \mu_i) + \frac{1}{2} \ln \frac{|\frac{\sum_k + \sum_i}{2}|}{\sqrt{|\sum_k| |\sum_i|}}$
onde \sum_k e \sum_i são as matrizes de covariância das classes k e i , μ_k e μ_i são os respectivos vetores de média .

- Distância média de Jefferys-Motusita (JM):

$$JM_{k,i} = \sqrt{2(1 - e^{-B_{k,i}})} \quad JM \in [0, \sqrt{2}] \quad (3.13)$$

onde $B_{k,i}$ é a distância de Bhattacharyya.

De posse dos conhecimentos básicos, serão apresentadas na próxima seção quais as características consideradas para o problema de classificação em questão.

3.2 CARACTERÍSTICAS DO SOM

A fase de extração e seleção de características é tida como a mais crítica em sistemas de classificação. E como uma das partes desse trabalho envolve essa tarefa, se fez necessário uma busca pelas características utilizadas em outros sistemas de classificação, que utilizam o som como fonte de informação, tais como reconhecimento de locutor, como os dos trabalhos relacionados na Seção 2.2. Além disso, são propostos alguns algoritmos próprios para o contexto, tais como a frequência dos passos e a diferença entre a amplitude dos passos esquerdo e direito.

Todos os métodos reunidos nessa seção, compartilham a mesma entrada: o sinal oriundo dos passos dos indivíduos. Entretanto, alguns parâmetros são processados levando-se em consideração o modelo acústico do sistema auditivo humano, devido a estudos na área de psicoacústica. Em virtude disso, esta seção congrega os parâmetros estudados em dois grupos: parâmetros objetivos (Seção 3.2.1) e parâmetros psicoacústicos (Seção 3.2.2).

3.2.1 PARÂMETROS OBJETIVOS

Essa seção apresenta as características baseadas nos aspectos físicos da onda acústica produzida pelos passos.

ENERGIA E TAXA DE CRUZAMENTOS POR ZERO

A energia é uma das informações mais simples extraídas de um sinal, calculada a cada N amostras, como:

$$E = \sum_{n=1}^N x^2(n) \quad (3.14)$$

Costuma-se utilizar o logaritmo da energia, para suavizar grandes variações de magnitude e ressaltar as pequenas. Esta característica é conhecida como Log-energia.

A taxa de cruzamentos por zero (TCZ) é um método simples de análise no domínio do tempo que baseia-se na quantidade de vezes que a onda cruza o eixo das abcissas (GERHARD, 2003), ou seja, quando as amplitudes de amostras temporais sucessivas possuem sinais diferentes. Pode ser definida pela fórmula abaixo:

$$zcr_i = \frac{1}{2} \sum_{n=1}^N |sgn[x_i(n)] - sgn[x_i(n-1)]| \quad (3.15)$$

Onde x_i representa as amostras temporais do quadro i , N é o número de amostras por quadro e sgn é a função de avaliação do sinal da amostra, definida por

$$sgn(x) = \begin{cases} -1 : x < 0 \\ 0 : x = 0 \\ 1 : x > 0 \end{cases} \quad (3.16)$$

A energia e o TCZ foram utilizados no projeto de um algoritmo que marca o início e o fim dos passos. Na literatura de processamento de voz, essas marcas são conhecidas como *endpoints*. O algoritmo e seus resultados são mostrados no próximo capítulo, por pertencerem a etapa de pré-processamento.

FREQUÊNCIA DOS PASSOS

Foi mostrado em trabalhos pretéritos (LIMA, 2005; SHOJI, 2004, 2005) que a frequência dos passos foi uma das principais características com capacidade de distinção de indivíduos. Analisando a forma de onda dos passos obtida nos experimentos, notou-se que a cada impacto do pé com o solo, é gerada uma onda caracterizada por um rápido crescimento até um pico máximo, e posteriormente um decaimento em menor velocidade. A partir desse argumento, pode-se afirmar que toda onda⁸ gerada pelo ligeiro impacto entre o pé e o solo possui um valor máximo, que será chamado de pico máximo do passo - PMP.

Para explorar esse fato, foi desenvolvido um algoritmo simples para extrair os PMPs de cada sequência gravada e posteriormente calcular o tempo médio, chamado período

⁸Toda onda obtida do experimento realizado, descrito no Capítulo 5.

dos passos. Ao inverter o período dos passos, obtém-se a frequência de caminhada. Os passos do algoritmo batizado de *maxValueFootstepId* são dados a seguir:

- a) Calcule o número total de amostras de uma determinada sequência de passos;
- b) Considerando N o número de estímulos de uma forma de onda, divida as amostras em N partes;
- c) Calcule o valor máximo absoluto das N partes e armazene os N índices dos respectivos valores máximos;
- d) Calcule as diferenças entre os índices obtidos no passo c) e obtenha a média (período dos passos).
- e) Divida a média pela frequência de amostragem, transformando a média de amostras para segundos, por fim inverta-a para transformá-la em Hz.

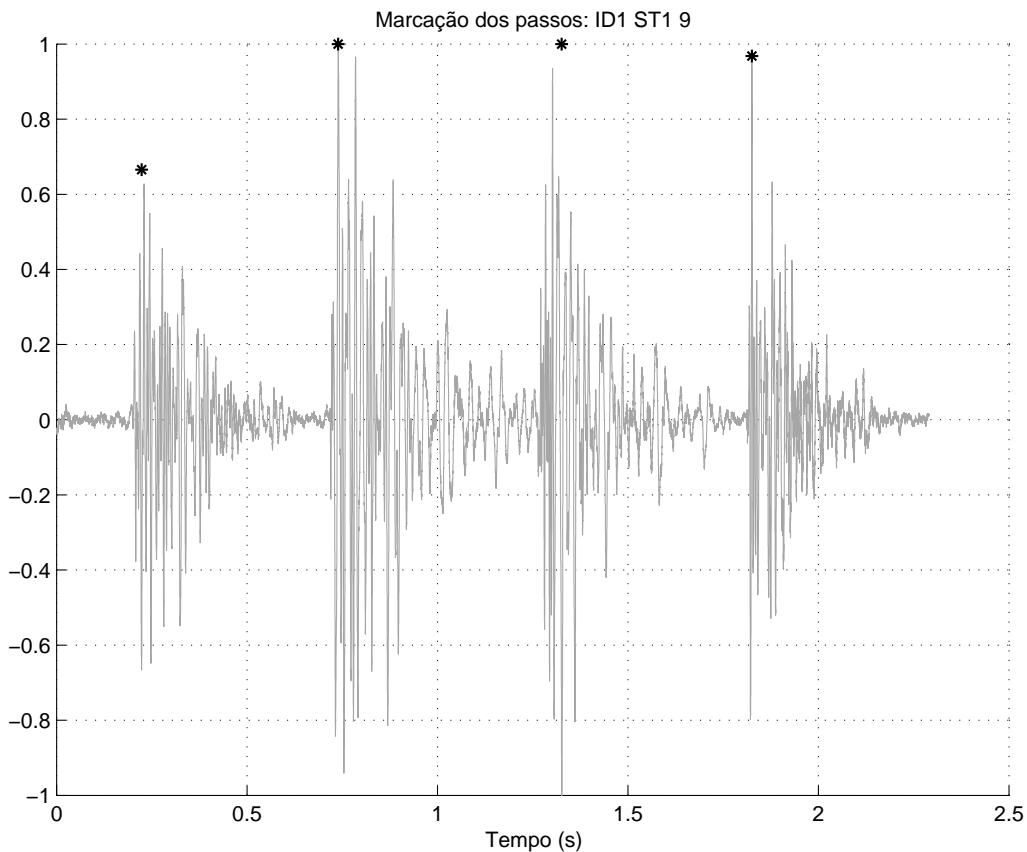


FIG. 3.3: Identificação dos PMPs (Pico Máximo do Passo) em uma sequência de passos.

A FIG. 3.3 mostra o resultado do algoritmo proposto para uma amostra.

Esse algoritmo foi elaborado baseado em amostras de passadas de indivíduos, obtidas experimentalmente (o experimento é descrito no Capítulo 5) e sem levar em consideração o ruído ambiente. Acredita-se que caso o algoritmo seja testado com amostras gravadas em ambientes com ruídos semelhantes ao som de um passo, tais como queda de objetos, seu desempenho poderá diminuir consideravelmente (por se basear em picos máximos). Entretanto, para o tratamento de ruídos é necessário um estudo especializado nesse tópico, no intuito de encontrar um conjunto de filtros munidos da capacidade de recuperar o sinal o mais próximo do tratado no atual trabalho.

COEFICIENTES CEPESTRAS

O cepestro é uma técnica bastante utilizada em sistemas de processamento de voz. A voz é um sinal produzido como resultado de várias transformações que ocorrem em diferentes níveis: semântico, lingüístico, articulatório e acústico (FECHINE, 2000). Tais transformações refletem diretamente nas propriedades acústicas do sinal de voz. Em reconhecimento de locutor, por exemplo, as diferenças relacionadas às diferenças anatômicas inerentes ao trato vocal⁹ e seu movimento dinâmico (forma como a pessoa fala) podem ser utilizadas para discriminar locutores (FECHINE, 2000; REYNOLDS, 1995; RABINER, 1993).

Os sinais gerados pela modulação da corrente de ar dos pulmões pelo abrir e fechar das cordas vocais na glote ou laringe, são chamados de sinais vozeados (pronúncia de consoantes como /m/, /n/). Enquanto que os sinais gerados pela baixa pressão do trato vocal provocando fluxos de ar turbulentos, os quais resultam em ruído, são chamados de sinais não-vozeados. Os dois tipos de sinais podem ser aplicados simultaneamente, por exemplo, em fricativas como /z/ ou /v/.

Do ponto de vista de processamento de sinais, este comportamento pode ser modelado por um sistema linear (FIG. 3.4). O sinal de excitação $e(n)$ é modelado pela contribuição da componente proveniente dos sinais vozeados (trem de pulsos) juntamente com a dos sinais não-vozeados (aleatório). O sinal de voz $s(n)$ é o resultado da convolução entre o sinal de excitação $e(n)$ e o modelo do trato vocal $h(n)$.

⁹Nome genérico dado ao conjunto de cavidades e estruturas que participam diretamente da produção sonora de voz.

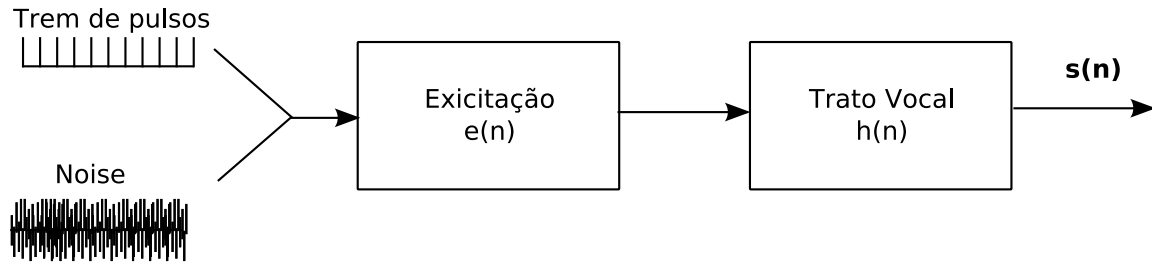


FIG. 3.4: Sistema linear do mecanismo de produção de voz.

Para o reconhecimento de locutores, é desejável separar (deconvoluir) o sinal $s(n)$, de forma a obter a informação do trato vocal $h(n)$. No cepstro os representantes dos sinais componentes serão separados e serão *linearmente combinados* (JR., 1993).

Existem duas formas para a obtenção dos coeficientes cepestrais, a saber: coeficientes cepestrais FFT e coeficientes cepestrais LPC.

Na análise cepestral FFT é aplicada, diretamente ao sinal, uma transformada inversa rápida de Fourier. O i -ésimo cepestro, $ce_i(n)$, é calculado por (JR., 1993):

$$ce_i = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log_{10} |X_i(e^{j\omega})| e^{j\omega n} d\omega, \quad (3.17)$$

em que $-\infty < n < \infty$ e X_i representa o i -ésimo bloco do espectro de potência do sinal a curtos intervalos de tempo.

Na análise cepestral LPC, a transformada- z é aplicada no sinal de voz modelado pela análise LPC. Os coeficientes Cepestrais, do espectro obtido na análise LPC, podem ser calculados recursivamente, a partir dos coeficientes LPC, c_i , por (MAMMONE, 1996):

$$ce_i(1) = c_i(1), \quad (3.18)$$

$$ce_i(n) = c_i(n) + \sum_{j=1}^{n-1} 1 - \frac{j}{n} c_i(j) ce_i(n-j) \quad , 1 < n \leq K, \quad (3.19)$$

onde n é o índice do coeficiente e i o índice do bloco de amostras.

O uso da relação recursiva leva a uma computação eficiente dos coeficientes cepestrais, $ce_i(n)$, e evita a fatoração polinomial. Uma vez que $ce_i(n)$ tem duração infinita, o vetor de características, de dimensão K , é constituído das componentes $ce_i(1)$ a $ce_i(K)$, as quais são as mais significativas devido ao “decaimento” da sequência com o aumento de n (FECHINE, 2000).

O método LPC é comumente utilizado para definir as características do locutor, pois modela seu trato vocal. Entretanto, a FFT modela a forma de onda, necessitando de outras técnicas para auxiliar a extração das características do locutor (FARRELL, 1994).

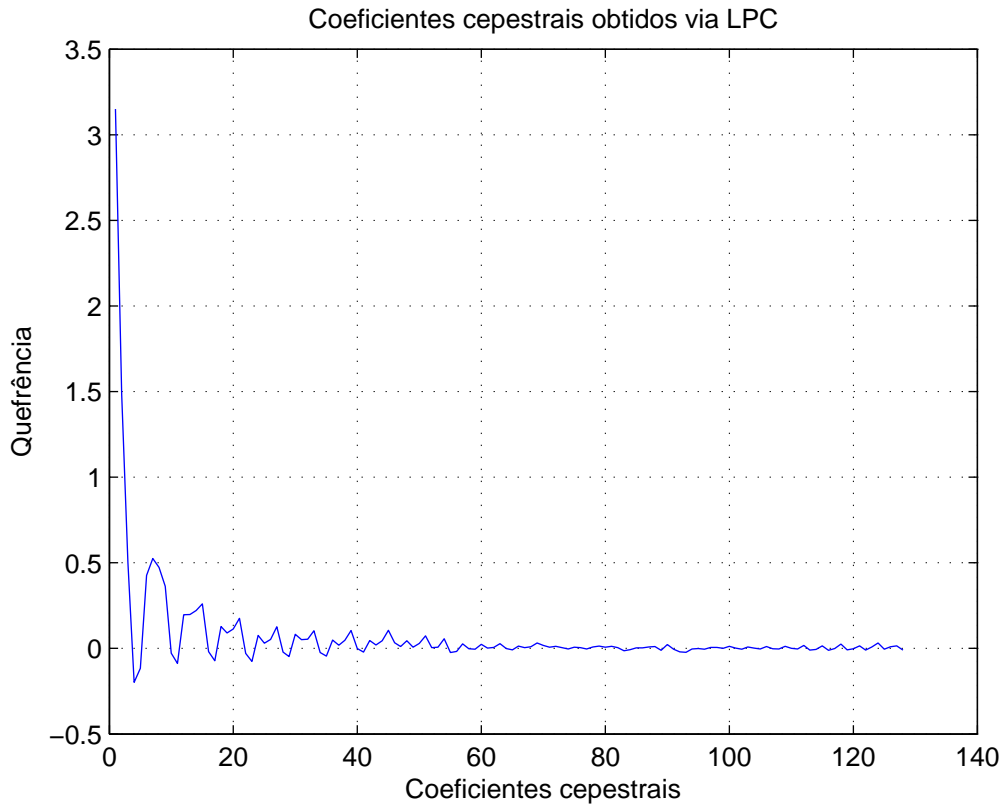


FIG. 3.5: Coeficientes cepstrais obtidos pelo método LPC. A sequência utilizada foi digitalizada a 44.1 KHz e contém o som de quatro passos.

ENVELOPE ESPECTRAL

O espectro de um sinal mostra a quantidade de energia distribuída em uma determinada faixa de frequências. Em geral, o espectro de sinais sonoros reais possui muitas variações (SCHWARZ, 1998). Entretanto, a curva traçada no plano do espectro do som dos passos possui um comportamento mais ameno, podendo ser utilizada como característica para distinguir indivíduos (SHOJI, 2004, 2005). Segundo (ABU-SHIKHAH, 2007), a curva no plano amplitude-frequência, derivado de um espectro, é chamado envelope espectral. Estatisticamente, a descrição geral do envelope espectral pode ser vista como uma técnica de componentes principais, baseada nas frequências, aplicadas a séries temporais multivariadas (STOFFER, 2000). Após calculado o envelope, o grau de similaridade entre os

envelopes das amostras de um mesmo indivíduo é obtido utilizando-se alguma métrica de distância (vide Seção 3.1).

De acordo com (SCHWARZ, 1998), todo método para estimar o envelope espectral deve levar em consideração as seguintes propriedades:

Precisão - a curva estimada deve descrever um envelope do espectro, isto é, deve cobrir estreitamente em torno da magnitude do espectro, ligando os picos¹⁰;

Robustez - o método de estimação deve ser aplicável a um vasto número de sinais com diversas características, de sons com altas harmônicas fundamentais com várias parciais espaçadas a sons ruidosos ou misturas de sons harmônicos e ruidosos;

Regularidade - uma certa suavidade ou regularidade é requerida. Isto significa que o envelope espectral não deve oscilar demasiadamente, contudo deve manter a idéia geral da distribuição da energia do sinal sobre a frequência.

Para estimar o envelope espectral é necessário primeiramente estimar o espectro. Os métodos de Welch e Periodograma foram eleitos por apresentarem bons resultados quando se trabalha com sinais quase aleatórios.

O periodograma de uma seqüência x_1, \dots, x_N é dado pela seguinte fórmula:

$$P(e^{j\omega}) = \frac{1}{N} \left| \sum_{l=1}^N x_l e^{0j\omega l} \right|^2 \quad (3.20)$$

Entretanto para um espectro mais suave, foi aplicado a função janela de Hamming multiplicada pela forma de onda de uma seqüência. Supondo que a seqüência w_1, \dots, w_N seja a seqüência gerada pela função janela, então o periodograma fica definido da seguinte maneira:

$$P(e^{j\omega}) = \frac{\frac{1}{N} \left| \sum_{l=1}^N x_l \cdot w_l e^{0j\omega l} \right|^2}{\frac{1}{N} \sum_{l=1}^N |w_l|^2} \quad (3.21)$$

Também conhecido como periodograma modificado, o método de Welch se baseia na idéia de dividir o sinal em sucessivos blocos e calcular a média da magnitude ao

¹⁰Algumas vezes não será possível ligar cada pico. Por exemplo, quando a análise encontra um grupo de picos próximos uns aos outros em uma faixa de frequência alta, o envelope espectral deve encontrar um caminho intermediário. Uma possível solução para este caso, seria pôr o centro de gravidade de cada fatia de frequência do grupo.

quadrado, obtida através da transformada discreta de Fourier aplicada a cada um dos blocos (SMITH, 2003).

O algoritmo para estimar o espectro segundo o método de Welch pode ser descrito através dos seguintes passos (WELCH, Jun 1967):

- O sinal x de entrada é segmentado em k seções de acordo com a função janela utilizada e a porcentagem de sobreposição escolhida;
- Cada segmento é multiplicado pela função janela escolhida do mesmo tamanho do segmento;
- O periodograma de cada segmento janelado é computado;
- Toma-se a média dos periodogramas para formar o espectro estimado $S(e^{j\omega})$.

Após estimar o espectro de frequências, seu valor é transformado para a escala logarítmica dB, através da seguinte fórmula

$$\hat{P} = 10 * \log_{10}(P) \tag{3.22}$$

onde P é o espectro estimado através dos métodos discutidos anteriormente.

Após ser transformado para a escala logarítmica, os máximos locais são calculados e um novo sinal, o envelope espectral, é estimado através da interpolação dos máximos locais. O método de interpolação utilizado foi o “splines” cúbico, por suavizar a forma de onda resultante.

Os resultados da interpolação do espectro obtido pelo periodograma e pelo método de Welch, são mostrados nas FIG. 3.6a e FIG. 3.6b.

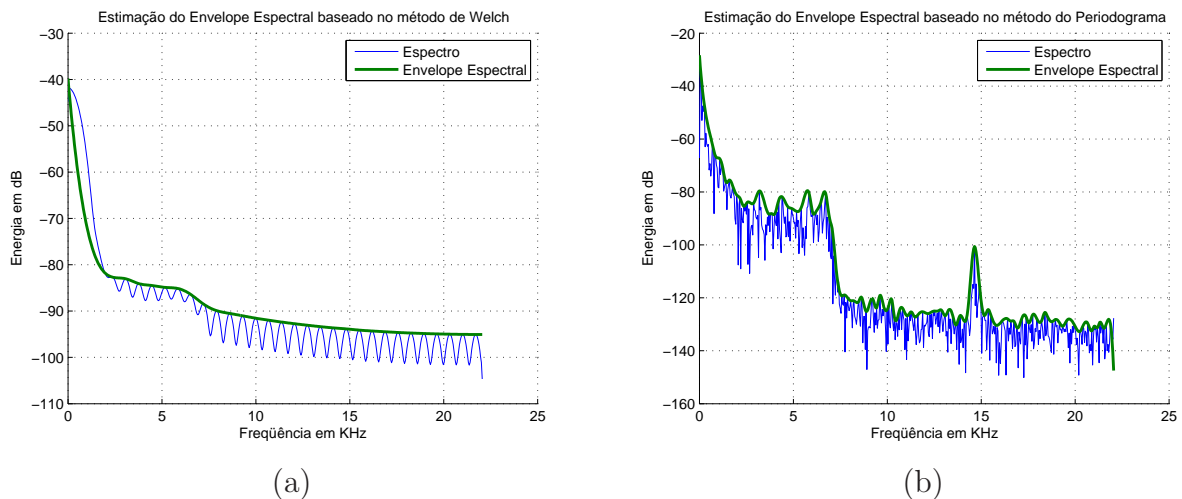


FIG. 3.6: Resultado da estimação do envelope espectral. (a) Envelope usando o método de Welch. (b) Envelope usando o método do periodograma.

3.2.2 PARÂMETROS SUBJETIVOS

A percepção do som pelo sistema auditivo e sua interpretação pelo cérebro caracteriza o termo aspecto subjetivo. A percepção do som é estudada pela psicoacústica e a interpretação pelas ciências cognitivas. A psicoacústica define o som, sob o aspecto subjetivo, como a sensação auditiva produzida pelo ouvido ocasionada pela alteração em pressão, deslocamento ou movimentação de partículas, que se propaga em um meio elástico (OLSEN, 1967) *apud* (JÚNIOR, 2003).

Ao notar diferenças de pressão sonora, o sistema auditivo capta as oscilações dessas diferenças, convertendo-as de vibrações mecânicas a impulsos elétricos, posteriormente transportados pelo nervo auditivo ao cérebro que então as interpreta como a sensação fisiológica conhecida como audição. O motor auditivo do ser humano privilegia sons importantes para sua sobrevivência e relacionamento com outros humanos. A exemplo disso, a banda de frequência mais sensível pelo ouvido é de 1 a 5 KHz, faixa em que se encontra o sinal de voz.

Do mesmo modo que para outros sentidos da percepção humana, a audição apresenta limites de percepção. Escutamos os sons que ocorrem dentro de uma faixa de intensidade, frequência e tempo. O limite da percepção de intensidade sonora é dado pelo nível mínimo de percepção sonora, onde o ouvido percebe a existência do som, até o limiar da dor, onde a intensidade sonora é tão grande que provoca sensação de desconforto ou dor no ouvinte.

A percepção da intensidade está relacionada a frequência das componentes do som. Para sons simples, com apenas uma componente sonora, a percepção da intensidade sonora varia aproximadamente entre 0 dB para o limiar da percepção até 120 dB para o limiar da dor.

No domínio do tempo, a percepção do som ocorre em dois níveis perceptuais: macroscópico e microscópico. A divisão entre estes se dá pela definição de um intervalo de tempo conhecido como persistência auditiva. Eventos sonoros que ocorram separados no tempo por intervalos menores que o da persistência auditiva são percebidos pela audição como se ocorressem simultaneamente. O intervalo da persistência auditiva médio é 30 ms.

A percepção macroscópica leva em conta a organização temporal ou rítmica do som. É neste nível de percepção que a audição reconhece ritmos, melodias, sílabas e palavras. A percepção macroscópica não leva em conta o timbre do instrumento ou da voz, que é definido adiante, desse modo é possível para a audição reconhecer sílabas e palavras pronunciadas por diferentes indivíduos, com timbres de voz diferentes. Já no nível de percepção microscópico, a audição reconhece o timbre, ou seja, as características estruturais do som, como seu ataque e a sua composição espectral. Através da percepção microscópica a audição reconhece, por exemplo, a diferença das vozes de indivíduos ou a diferença entre o som de instrumentos musicais, mesmo soando a mesma frase ou nota musical.

A psicoacústica é a ciência que estuda a percepção do som pela audição humana, levando em consideração seus limites e não-linearidades. A percepção das grandezas acústicas é estudada pela psicoacústica de modo a fornecer um mapeamento de cada grandeza em relação à sua percepção subjetiva. Deste mapeamento surgem as grandezas psicoacústicas. Para a percepção da intensidade sonora, tem-se a *sonoridade*. Para a percepção da frequência, tem-se o *pitch*. Para a percepção das componentes em frequência o tem-se a distribuição espectral. Além destas grandezas perceptuais, o sistema auditivo é composto pelos dois ouvidos, percepção também chamada de bi-audição. Esta permite reconhecer a localização espacial de uma fonte sonora, pela diferença de tempo de chegada do som a cada ouvido, bem como por outros detalhes, como ecos, reverberações e reflexos na estrutura da orelha e no ombro do ouvinte (JÚNIOR, 2003).

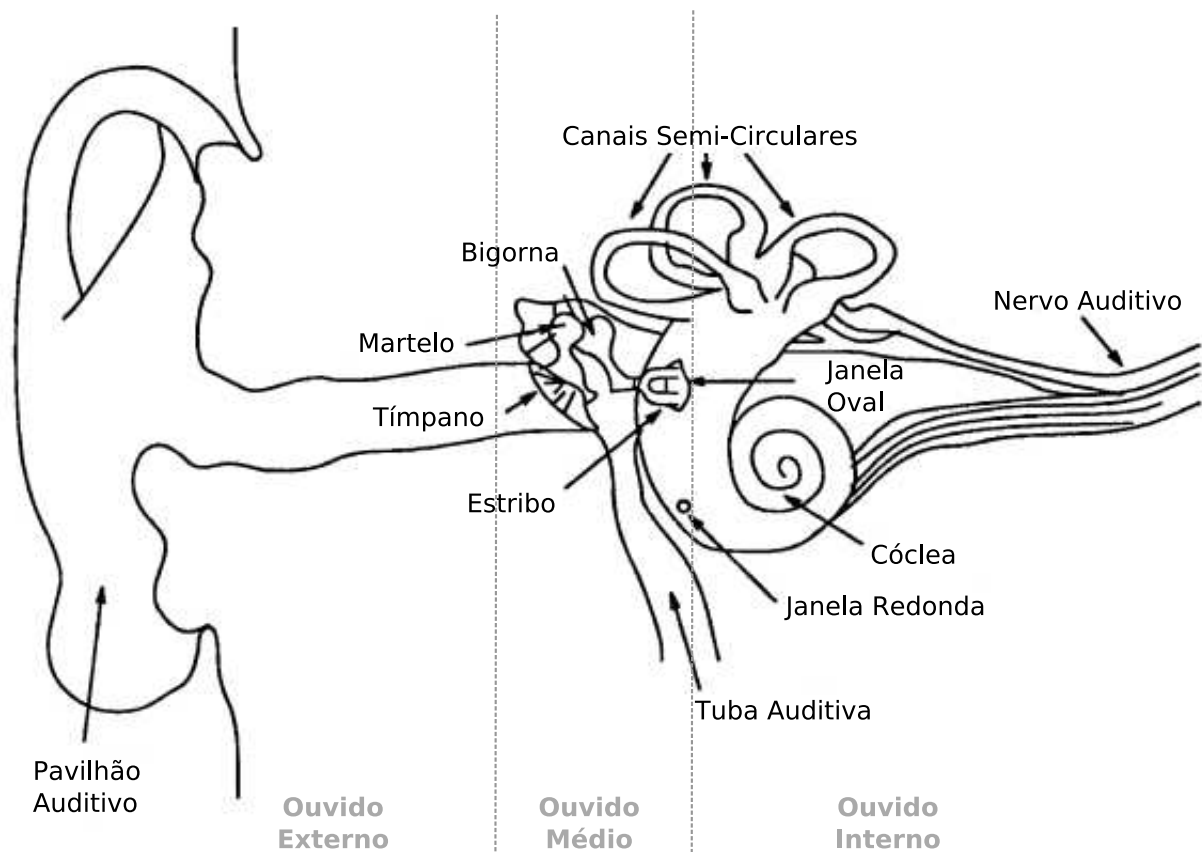


FIG. 3.7: Corte longitudinal do ouvido, com suas principais estruturas (BARBEDO, 2004).

Como pode ser visto na FIG. 3.7, o ouvido é formado por três partes: ouvido externo, médio e interno. O ouvido externo e a cabeça são componentes de um complexo sistema de recepção acústica, fazendo a ligação entre o tímpano e o campo sonoro externo. Além de proteger o tímpano de danos mecânicos, melhora o acoplamento entre este e o campo sonoro além de ter a propriedade de filtrar o som de modo a realçar as frequências mais importantes para o reconhecimento da voz humana. Além disso, contribui no processo de localização da posição da fonte sonora no espaço.

O ouvido médio é composto pela membrana timpânica, a cadeia ossicular com os respectivos ligamentos e músculos e a cavidade preenchida com ar, onde se encontram os ossículos. Sua principal função é melhorar a transmissão sonora entre os demais ouvidos (interno e externo), por possuir a capacidade de reduzir a reflexão ocorrida quando uma onda sonora incide em uma superfície fluida. Pode-se considerá-lo um transformador de impedância que reduz a alta impedância do fluido coclear (descrito mais adiante) para um valor semelhante ao do ar. O tímpano marca a divisão entre os ouvidos médio e externo;

as janelas oval e redonda da cóclea são os limites entre os ouvidos médio e interno. Os ossículos (*martelo, bigorna e estribo*) fazem a transmissão do som que é recebido pelo tímpano diretamente à janela oval, além de protegerem a janela redonda do som, o qual chega a ela com menor amplitude. É interessante destacar que se isto não ocorresse, o som chegaria ao mesmo tempo às duas janelas com a mesma magnitude e na mesma fase sonora, pois a distância entre elas é muito pequena, e isto não promoveria a movimentação do fluido do ouvido interno, tornando a audição impossível (BARBEDO, 2004).

Os ossos do ouvido médio estimulam a cóclea através da janela oval, fazendo com que seu líquido interno se movimente. A cóclea pode ser modelada como um tubo de aproximadamente 30 mm com duas câmaras separadas por uma estrutura conhecida como membrana basilar (FIG. 3.8). A membrana basilar apresenta uma resistência mecânica que varia ao longo de sua extensão, ressoando em frequências mais altas no extremo próximo à janela oval, enquanto que no seu final (ápice) ressoa em frequências mais baixas. As ondas geradas pelo estribo, em resposta a um sinal senoidal, viajam ao longo da cóclea, fazendo vibrar a membrana basilar na mesma frequência do sinal de entrada.

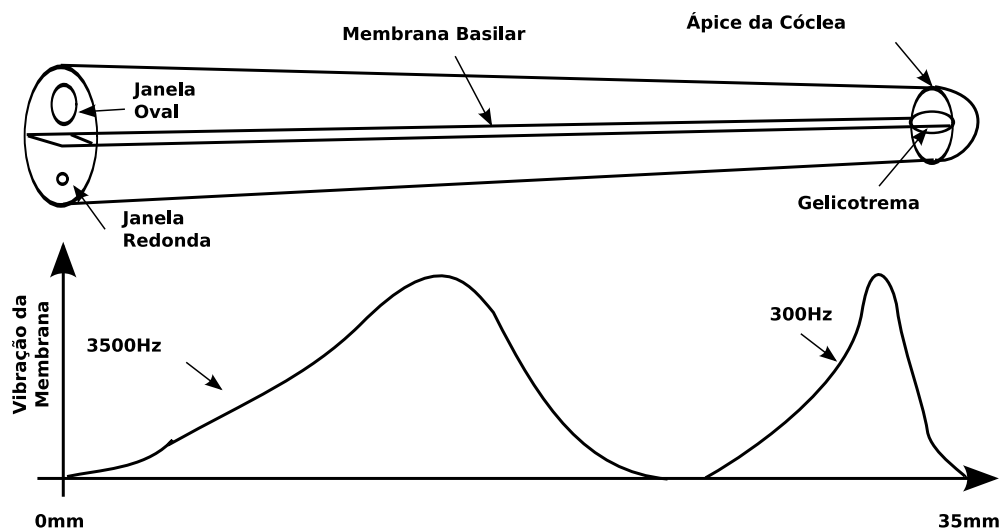


FIG. 3.8: Cóclea, membrana basilar e vibração da membrana em duas diferentes frequências (adaptado de (BARBEDO, 2004)).

Cada ponto da membrana basilar é mais sensível a uma determinada frequência, chamada de frequência característica. As fibras basilares localizadas na região de altas frequências características, respondem em uma maior faixa de frequências do que as fibras na região de baixas frequências características. O gráfico abaixo da membrana basilar na FIG. 3.8, mostra a curva de resposta ao longo da membrana basilar para um

tom numa frequência específica. Para cada frequência, há um ponto da membrana basilar em que a vibração é máxima. A posição desse ponto, medida a partir do helicotrema, é aproximadamente proporcional ao logaritmo da frequência do som. Ao redor desse ponto haverá uma faixa, de cerca de 1,5 mm, onde a vibração estará presente, atenuando-se conforme se afasta do ponto (BARBEDO, 2004). Essa faixa determina o conceito de bandas críticas.

Segundo (RABINER, 1993) *apud* (SILVA, 2002), uma banda crítica é uma função de frequência que quantifica a faixa de passagem do filtro coclear. Em outras palavras, é um critério subjetivo do conteúdo de frequência de um sinal que se refere a faixa de passagem para o qual respostas subjetivas tornam-se significativamente diferentes. A sonoridade (definida mais abaixo) de uma banda de ruído para uma pressão sonora constante, permanece constante quando a faixa de ruído aumenta até a largura da banda crítica; após ultrapassar o limite percebe-se mudança na sonoridade. Uma banda crítica define uma faixa em torno de uma frequência central, a qual está associada a um ponto da membrana basilar, de modo que a cada ponto é possível definir uma banda crítica. Quando dois sinais se situam dentro de uma abanda crítica, o de maior energia poderá dominar a percepção e mascarar o outro estímulo sonoro. Este fenômeno é conhecido como mascaramento simultâneo. Assim, dependendo dos níveis, dois tons distintos só serão perceptualmente distinguidos quando estiverem em bandas críticas diferentes.

As bandas críticas podem ser definidas em torno de qualquer frequência central. A largura de faixa das bandas corresponde a um espaçamento uniforme de 1,5 mm ao longo da membrana basilar, correspondendo a aproximadamente 100 Hz para frequências abaixo de 500 Hz e de aproximadamente 20 % da frequência central, para frequências acima de 1000 Hz (em direção à janela oval). Apesar de existir uma banda crítica em torno de cada frequência, convencionou-se (com pequenas variações) a adoção dos valores mostrados na TAB. 3.1. Os valores para a coluna de banda crítica são dados na escala Bark.

Banda Crítica	Frequência (Hz)		
	Inferior	Superior	Faixa
0	0	100	100
1	100	200	100
2	200	300	100
3	300	400	100
4	400	510	110
5	510	630	120
6	630	770	140
7	770	920	150
8	920	1080	160
9	1080	1270	190
10	1270	1480	210
11	1480	1720	240
12	1720	2000	280

Banda Crítica	Frequência (Hz)		
	Inferior	Superior	Faixa
13	2000	2320	320
14	2320	2700	380
15	2700	3150	450
16	3150	3700	550
17	3700	4400	700
18	4400	5300	900
19	5300	6400	1100
20	6400	7700	1300
21	7700	9500	1800
22	9500	12000	2500
23	12000	15500	3500
24	15500	22050	6550

TAB. 3.1: Bandas críticas (em Bark).

COEFICIENTES MEL CEPESTRAIS

Para explanação dos coeficientes mel cepestrais, é mister um breve estudo no campo físico-acústico o qual examina a percepção auditiva humana. Tais estudos mostram que a escala de frequências da percepção da voz humana é não-linear. Para cada tom com uma frequência de percepção medida em Hertz, há uma relação com uma frequência de percepção em uma escala por nome *mel*. Stevens e Volkman (STEVENS, 1940) foram os pioneiros neste tipo de trabalho, elaborando um experimento que consistiu em: arbitrariamente escolher a frequência de 1000 Hz a 1000 *mels* em um tom de 30 db acima do limiar mínimo da percepção auditiva humana. Em seguida, realizaram testes com pessoas aumentando e diminuindo a frequência até que as pessoas percebessem. Ilustrado pela FIG. 3.12, a conclusão desse experimento foi de que o mapeamento entre a frequência em *Hz* e em *mels* é aproximadamente linear abaixo dos 1000 Hz e logarítmica acima. Um mapeamento da frequência percebida (*mel*) versus a frequência real resulta na escala *mel*, expressa por (PICONE, 1993):

$$mel = 2595 \log_{10}\left(1 + \frac{f}{700}\right) \quad (3.23)$$

A percepção de uma frequência particular para o sistema auditivo humano é influenciada pela energia dentro de uma banda crítica centrada em torno da frequência em questão (DELIMA, 2001). Devido a esse fato, é comum a utilização de filtros de banda crítica (filtros passa faixas) para calcular os coeficientes mel-cepestrais. Alguns pesquisadores sugerem a utilização da log-energia total encontrada dentro das bandas críticas em torno de cada frequência, em vez de utilizar a log-magnitude (JR., 1993). Além disso, a largura da banda dos filtros varia com a frequência, começando por volta de 100 Hz para frequências abaixo de 1 KHz, e aumentando logarithmicamente acima de 1 KHz. Para a obtenção dos coeficientes mel-cepestrais, costuma-se utilizar 20 (vinte) filtros passa-banda triangulares (JR., 1993). Cada filtro é centrado em uma frequência *mel*, que determina a largura de banda crítica do filtro. A FIG. 3.9 mostra os filtros para o cálculo dos coeficientes mel-cepestrais para um sinal amostrado a 8 KHz.

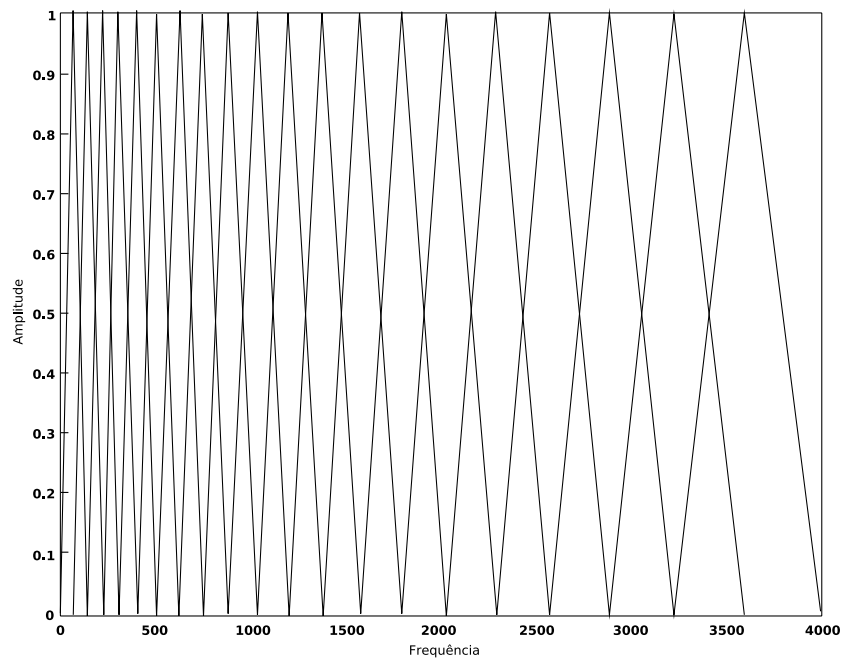


FIG. 3.9: Banco de filtros triangulares espaçados segundo a escala *mel* (DAVIS, 1980).

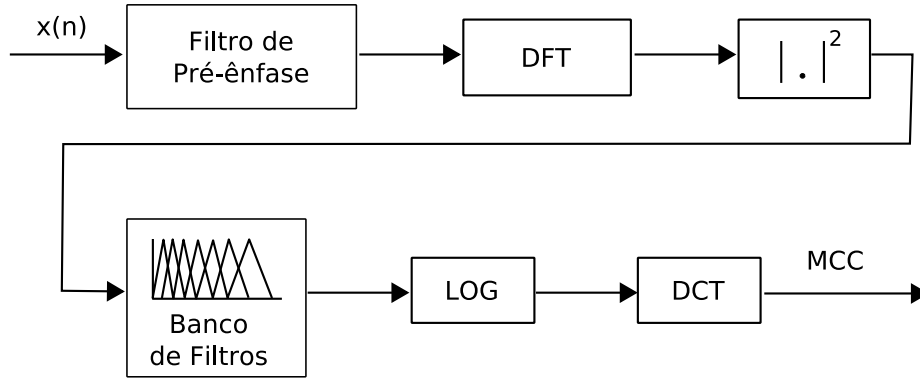


FIG. 3.10: Diagrama em blocos para a extração dos coeficientes mel-cepestrais (REYNOLDS, 1995; DELIMA, 2001) (O filtro de pré-ênfase é utilizado em sinais de voz).

De acordo com (DAVIS, 1980), os coeficientes mel-cepestrais, baseados num banco de filtros de banda crítica, podem ser calculados como:

$$MCC_i = \sum_{k=1}^N X_k \cos \left[i \left(k - \frac{1}{2} \right) \frac{\pi}{N} \right] \quad i = 1, 2, \dots, M \quad (3.24)$$

onde M é o número de coeficientes mel-cepestrais, e X_k , $k = 1, 2, \dots, N$, representa a energia logarítmica do k -ésimo filtro e N é o número de filtros do banco de filtros. A FIG. 3.10 mostra o diagrama de blocos da sequência para o cálculo dos coeficientes mel-cepestrais.

Para obter os coeficientes mel-cepestrais em sinais de voz, inicialmente o sinal de voz digitalizado passa por um filtro de pré-ênfase, posteriormente extrai-se o espectro do sinal por meio da transformada discreta de Fourier (*Discrete Fourier Transform - DFT*) ou dos coeficientes de predição linear. Neste trabalho, o espectro foi obtido através da DFT, por ser mais robusto em relação ao ruído. Logo depois, o espectro de energia é calculado, sendo então filtrado através da multiplicação por uma série de filtros triangulares espaçados segundo a escala *mel* (escala projetada para simular a resposta de frequências do ouvido humano). A energia resultante da filtragem é aplicada a uma função logarítmica, e finalmente, é utilizada a transformada cosseno discreta (*Discrete Cosine Transform - DCT*) para se obter os coeficientes no domínio cepestro de frequência (quefrência).

A implementação do modelo acima descrito se encontra em (ELLIS, 2005). Nesta *toolbox* gratuita para Matlab, plataforma utilizada na confecção do sistema, é possível configurar o número de filtros utilizados, a faixa de frequência do sinal de entrada e o

número de coeficientes desejados. Por essas razões, esta ferramenta foi utilizada nessa dissertação.

SONORIDADE

A grandeza da percepção da intensidade sonora é chamada de *sonoridade*. Experimentos realizados por Fletcher e Munson, (FLETCHER, 1933) demonstraram que para sons senoidais, ou seja, com apenas uma componente sonora, a percepção da intensidade é dependente da frequência da componente. A unidade de *sonoridade* é chamada de *phon* e as curvas cuja *sonoridade* se mantêm constante são as curvas isofônicas. Estes experimentos foram realizados dentro dos limites de percepção de intensidade e frequência, ou seja, sons senoidais de intensidades variando entre 0 e 120 dB e frequência entre 20 e 20000 Hz. A partir desses dados empíricos estabeleceu-se o que é conhecido hoje como curvas isofônicas de Fletcher e Munson, vistas na FIG. 3.11.

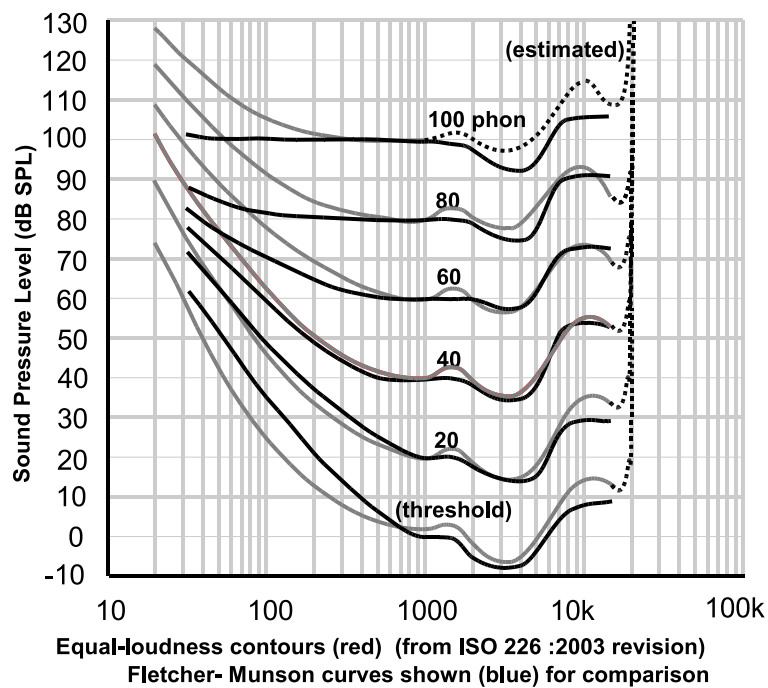


FIG. 3.11: Curvas isofônicas de Fletcher-Munson.

Como pode ser verificado na FIG. 3.11, o ouvido é mais sensível para intensidades sonoras com frequências medianas, frequências onde se encontra a fala humana. É importante realçar que tais experimentos foram realizados para sons senoidais, que possuem um único componente em frequência. Na realidade a quase totalidade dos sons que escutamos

são sons complexos, compostos por muitas componentes sonoras que variam dinamicamente ao longo do tempo. Assume-se assim que os formatos das curvas isofônicas devem vir a se modificar de acordo com a composição do espectro de frequência de cada som complexo.

O limite da percepção da frequência sonora é relacionado ao formato em cunha do órgão de Corti, dentro da cóclea, que é sensível às frequências sonoras aproximadamente entre 20 Hz e 20.000 Hz. Para efeito de comparação, as frequências fundamentais das notas do piano, um dos instrumentos com maior extensão de escala musical, vão de 27,5 Hz para a primeira nota, o A_0 , até 4.186 Hz, para a última nota, o C_8 . A voz humana varia a frequência fundamental entre 80 Hz para baixos até 1.000 Hz para sopranos.

A percepção da frequência sonora se reduz com a idade do indivíduo. Entre indivíduos de audição normal, crianças podem escutar até acima de 20 KHz, adolescentes e jovens adultos até 16 KHz, pessoas muito idosas, consideradas com audição normal, podem apresentar esta percepção diminuída para o máximo de 5 KHz (LAUNER, 1995).

Para representar a percepção auditiva da variação de frequência sonora foi criada a escala Bark de frequência. Ao invés da escala linear de frequência em Hertz, a escala Bark apresenta maior resolução para baixas frequências e menor resolução a medida que a frequência aumenta. O gráfico da relação entre Bark e Hertz é dada pela Figura 3.12:

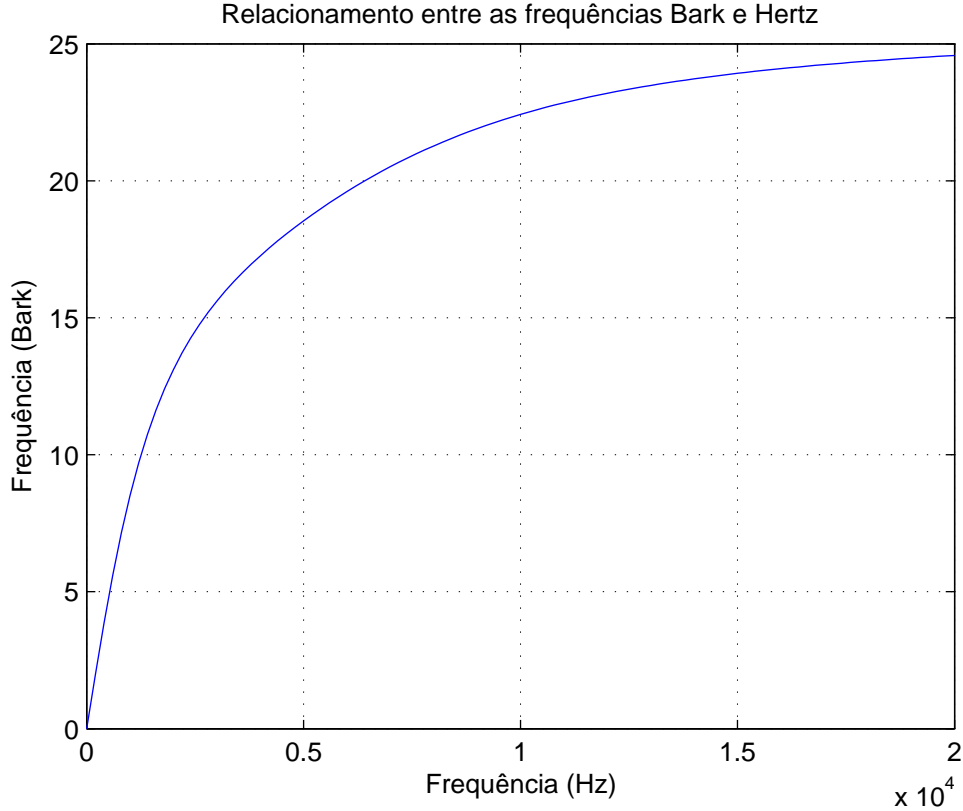


FIG. 3.12: Relação entre a escala bark e hertz.

O modelo para cálculo de sonoridade proposto por *Zwicker* inclui a maior parte dos passos utilizados no cálculo das medidas perceptuais. O algoritmo que implementa o modelo utilizado nessa dissertação é uma versão simplificada que modela a resposta em frequência do ouvido usando filtros com um formato determinado pelas terças-oitavas, tem sido amplamente utilizada para a estimação da sonoridade no campo da prevenção de ruído, tornando-se parte de um padrão internacional (ISO532). Para estimar a sonoridade, segundo o modelo de *Zwicker*, o primeiro passo é transformar o sinal de entrada para o domínio da frequência e agrupá-lo em bandas críticas. Esta operação pode ser expressa como uma função da densidade de energia na escala Bark, dada por

$$A(z) = \int_{z-0.5}^{z+0.5} \frac{dI}{dz'} \cdot dz' \quad (3.25)$$

Onde z é o valor da banda crítica, em Bark, $\frac{dI}{dz'}$ é a densidade de energia da banda crítica e $A(z)$ denota a energia total na banda crítica z . A operação pode também ser

expressa como uma função da densidade espectral de energia, na forma

$$A(z) = \int_{f(z-0.5)}^{f(z+0.5)} \frac{dI}{df} \cdot df \quad (3.26)$$

onde dI/df é a densidade espectral de energia para a faixa de frequência em Hz correspondente à banda crítica em questão. Segundo (BARBEDO, 2004), os dados utilizados no modelo de *Zwicker* foram obtidos através de experimentos fisiológicos os quais consistiram, basicamente, da observação do comportamento da membrana basilar mediante a apresentação de determinadas excitações e na contagem dos neurônios envolvidos em cada situação. Devido a complexidade matemática envolvida no processo de desenvolvimento do modelo de *Zwicker*, o próximo parágrafo reúne um resumo do processo dado por (BARBEDO, 2004). Maiores informações consultar (ZWICKER, 1990; FASTL, 2006).

De acordo com (BARBEDO, 2004), em (ZWICKER, 1990), o efeito do mascaramento simultâneo (domínio da frequência) é interpretado como o resultado de um espalhamento das excitações dos neurônios presentes na área basilar, que corresponde à faixa de frequência do estímulo sonoro, para áreas que na verdade não respondem ao estímulo sonoro; em outras palavras, a excitação de uma determinada região da membrana basilar irá disparar não apenas os neurônios ligados a tal região, mas também as bandas críticas adjacentes. Assim, um sinal de grande amplitude irá disparar um grande número de neurônios na região da membrana basilar correspondente à sua frequência, e também certo número de neurônios nas regiões adjacentes; se outro sinal for executado simultaneamente ao primeiro, este só será audível se for capaz de disparar um maior número de neurônios que o primeiro, na região correspondente à sua frequência na membrana basilar. Portanto, o número de neurônios excitados é outra maneira de se explicar as curvas de mascaramento geradas por determinado tom.

Portanto, as excitações designadas para as bandas críticas adjacentes são determinadas pela forma das curvas de mascaramento (onde o mascaramento é mais intenso, mais neurônios são envolvidos). De acordo ao modelo padronizado pela ISO, a excitação é determinada pelo maior valor entre as excitações parciais devidas a cada componente, ou seja, as excitações são espalhadas de maneira apropriada e, a cada banda, apenas o maior componente presente determinará a excitação resultante. No modelo de *Zwicker*, utilizado nesse trabalho, os padrões de excitação são transformados em uma função densidade $N'(z)$. Tal função é conhecida como sonoridade específica. Destarte, a transformação do padrão de excitação $E(z)$ para o padrão de excitação específico $N'(z)$, dado

em sôns/Bark, é dada por uma função de compressão não-linear, a saber (ZWICKER, 1990) *apud* (APPELL, 2002):

$$N'(z) = k \cdot \left(\frac{1}{s} \cdot \frac{E_l(z)}{E_0} \right)^\gamma \cdot \left[\left(1 - s + s \cdot \frac{E(z)}{E_l(z)} \right)^\gamma - 1 \right] \quad (3.27)$$

onde k é o fator de escala, γ uma constante, E é a excitação, E_l excitação correspondente ao limiar absoluto de audibilidade, E_0 excitação correspondente a um nível de pressão sonora de 40 dB (fator de escala), s fator de limiar.

A equação de sonoridade utilizada neste trabalho foi a seguinte (TIMONEY, 2004):

$$N' = 0.0064 \cdot 10^{0.025 L_{EQT}} \left[\left(11 + \frac{1}{4} 10^{0.1(L_G - a_0 - L_{ETQ})} \right) - 1 \right] \quad (3.28)$$

onde os valores de L_G , L_{ETQ} e a_0 se encontram em (ZWICKER, 1984) *apud* (TIMONEY, 2004).

A sonoridade total do sinal é dada por

$$N = \int_0^{24} N'(z) dz \quad (3.29)$$

o intervalo de integração cobre as bandas críticas (em Bark) consideradas.

Existem outros modelos para o cálculo do sonoridade, baseados no modelo de *Zwicker*. Modelo de Moore e Glasberg (MOORE, 1996), modelo de Launer (LAUNER, 1995) e modelo de Marzinik (MARZINIK, 1996) são alguns exemplos. Maiores detalhes dos modelos mencionados encontram-se em (APPELL, 2002).

3.3 MÉTODOS DE SELEÇÃO DE ATRIBUTOS

No corrente estudo, algumas das características estudadas possuem dimensionalidades consideravelmente grandes. Uma grande dimensionalidade do espaço de atributos pode causar degradação na classificação e um alto custo computacional. Dessa forma, é necessário aplicar um procedimento de redução dessa propriedade sem alterar a representatividade da característica. Além disso, é preciso mensurar a qualidade dos atributos de forma a eleger os mais discriminantes, para então serem utilizados no algoritmo de classificação.

A seleção de atributos pode ser vista como um processo de busca onde o algoritmo usado deve encontrar o menor subconjunto de atributos que possa proporcionar a melhor

acurácia de classificação (PAPPA, 2002). O objetivo de algoritmos de seleção de atributos é escolher o menor subconjunto que ofereça a melhor classificação em conformidade com custos computacionais razoáveis. A redução de atributos é um problema de otimização que busca criar novas características através de uma combinação do conjunto original de características mantendo a eficiência no processo de classificação. Nesse intuito, será posta adiante a técnica avaliada nessa dissertação para tal finalidade.

DISCRIMINANTE LINEAR DE FISHER

O discriminante linear de Fisher - FLDA, utiliza informações das categorias associadas a cada padrão para extrair linearmente as características mais discriminantes. Através da análise discriminante de Fisher é possível realizar a segregação entre classes, através de processos não supervisionados, quando não se tem um padrão conhecido. A FLDA é um método supervisionado na concepção da estatística e deve ser empregado quando algumas condições forem atendidas, tais como:

- a) As classes sob investigação são mutuamente exclusivas;
- b) Cada classe é obtida de uma população normal multivariada;
- c) Duas medidas não podem ser perfeitamente correlacionadas, entre outras.

A técnica consiste na computação de uma combinação linear de m variáveis quantitativas, que mais eficientemente separam grupos de amostras em um espaço m -dimensional. Para isso a razão da variância entre-grupos e inter-grupos é maximizada. A separação intra-classe é realizada através da seguinte fórmula estabelecida por Fisher, a saber.

- a) Espalhamento intra-classes:

$$S_{\omega} = \sum_{j=1}^c \sum_{i=1}^{T_j} (x_i^j - u_j) \cdot (x_i^j - u_j)^t, \quad (3.30)$$

em que x_i^j é o i -ésimo exemplo da classe j , u_j é a média da classe j , T_j é o número de amostras da classe j , c é a quantidade de classes e t representa a operação de transposição de matrizes.

- b) Espalhamento inter-classes:

$$S_b = \sum_{j=1}^c (u_j - u)(u_j - u)^t, \quad (3.31)$$

em que u é a média de todas as classes.

A partir do cálculo de espalhamento intra-classe e inter-classe de um conjunto de amostras, é possível seguir o critério de Fisher, que é maximizar a medida $S_f = S_b/S_\omega$. Essa fórmula mostra que se S_ω é uma matriz não singular, então essa taxa é maximizada quando os vetores colunas da matriz S_f são os autovetores de $S_\omega^{-1}.S_b$ (CAMPOS, 2001).

A alta dimensionalidade do espaço de atributos pode resultar em problemas como alto custo de processamento e a geração do fenômeno conhecido como maldição da dimensionalidade dos dados. A maldição da dimensionalidade, pode ser caracterizada como uma degradação nos resultados de classificação, devido ao aumento da dimensionabilidade dos dados.

Em análise discriminante de Fisher a redução de atributos é realizada a partir de um conjunto de amostras de n classes, tendo p variáveis, com o objetivo de reduzir para m atributos. Para a redução de atributos por Fisher segue-se o seguinte algoritmo:

- a) Calcular o espalhamento S_ω e S_b para n classes;
- b) Maximizar a medida inter-classes e minimizar a medida intra-classes S_f a partir de $S_\omega^{-1}.S_b$.

A partir de S_f é possível a redução de atributos com base nos auto-valores, em que os atributos selecionados serão os auto-vetores referentes aos seus m maiores auto-valores.

3.4 ALGORITMOS DE CLASSIFICAÇÃO

Esta seção põe adiante os algoritmos de classificação avaliados neste projeto.

K VIZINHOS MAIS PRÓXIMOS - K-NN

O método K-NN (do inglês *K-Nearest Neighbors*) é muito utilizado em aplicações envolvendo a tarefa de classificação. Trata-se de um método de fácil entendimento e implementação e que não requer treinamento prévio para ser aplicado. O funcionamento do K-NN encontra-se descrito a seguir.

Considerando uma base de dados (base de referência) de um problema envolvendo a tarefa de classificação (que contém um atributo cujos valores são rótulos de classes predefinidas) e cada novo registro a ser classificado (registro da base de teste), os seguintes passos são executados (GOLDSCHMIDT, 2005).

- a) Cálculo da distância do novo registro a cada um dos registros existentes na base de referência;
- b) Identificação dos k registros da base de referência que apresentaram menor distância em relação ao novo registro (mais similares);
- c) Apuração da classe mais frequente entre os k registros identificados no passo anterior;
- d) Comparação da classe apurada com a classe real, computando erro ou acerto do algoritmo. Este último passo só deve ser utilizado quando as classes dos novos registros são conhecidas e deseja-se avaliar o desempenho do método K-NN na base de dados em questão. Nas demais situações, este algoritmo se encerra no passo anterior.

O algoritmo KNN já foi aplicado à classificação de indivíduos a partir do padrão do caminhar em (CARVALHO, 2007) e em (CAMPOS, 2001) no reconhecimento de faces.

K-MEANS

O método *K-Means* é um método de particionamento que trata as observações dos dados como objetos que possuem posicionamento em uma base de características, onde é possível aplicar alguma métrica de distância entre os elementos. Ele particiona os objetos em k grupos mutuamente exclusivos, de forma que os objetos dentro de cada grupo são tão próximos uns dos outros quanto possível, e ao mesmo tempo tão distantes quanto possíveis de objetos em outros grupos. Cada grupo é identificado por um centróide ou ponto central.

Segundo (HAN, 2006), o *K-Means* é um dos mais simples algoritmos de aprendizagem não-supervisionada objetivando resolver o problema de agrupamento. O procedimento segue uma maneira simples e fácil de identificar subconjuntos de uma base de dados especificando-se o número de grupos (assume-se k grupos) definido *a priori*. A idéia principal é definir k centróides, um para cada grupo. Inicialmente, tais centróides devem ser colocados de maneira inteligente, pois diferentes locais iniciais causam diferentes resultados. Portanto, uma maneira bastante utilizada é posicioná-los o mais distante possível uns dos outros. Posteriormente, cada elemento do conjunto é associado ao centróide mais próximo, levando-se em consideração alguma métrica de distância. Quando nenhum ponto está pendente, o primeiro passo é completado e um agrupamento inicial é feito.

Portanto, é necessário recalcular k novos centróides como baricentros dos grupos resultantes do passo anterior. Após obter os novos k centróides, a associação dos dados aos respectivos centróides mais próximos deve ser refeita. Essa ação é executada até que não haja modificações a serem feitas, ou seja, a posição dos centróides não se modificam mais.

O algoritmo *K-Means* de agrupamento objetiva a minimização da seguinte função

$$J = \sum_{j=1}^n \sum_{i=1}^k \|x_i^{(j)} - c_j\|^2 \quad (3.32)$$

onde $\|x_i^{(j)} - c_j\|^2$ é a medida de distância escolhida entre o dado $x_i^{(j)}$ e o centro do grupo c_j indicando a distância dos n pontos de dados de seus respectivos centros.

O algoritmo é composto dos seguintes passos:

- a) Escolha k pontos no espaço representado pelos objetos que serão agrupados. Esses pontos representam os centróides iniciais.
- b) Associe cada objeto ao grupo cuja distância entre ele e o centróide seja mínima.
- c) Quando todos os objetos forem agregados a um grupo, recalcule as posições dos k centróides.
- d) Repita os passos (b) e (c) até que não haja mais atualizações dos centróides.

Apresentado os fundamentos necessários para o entendimento do restante do trabalho, o próximo capítulo aborda o problema de identificar indivíduos a partir dos passos.

4 IDENTIFICANDO INDIVÍDUOS ATRAVÉS DO SOM DOS PASSOS

Esse trabalho trata o problema de classificar indivíduos através do som dos passos, no objetivo da aplicação em um ambiente doméstico inteligente, como meio de identificação para prover serviços personalizados. Para tanto, é preciso estudar inicialmente as propriedades do tipo de caminhar buscando entender como o som da marcha é gerado, para então ser tratado. A próxima seção (Seção 4.1) diz respeito aos aspectos fundamentais do movimento humano através dos pés. Em seguida (Seção 4.2), é realizado um apanhado de estudos sobre a acústica dos passos. Por fim, a Seção 4.3 conclui o capítulo com uma discussão sobre o problema em questão.

4.1 ANÁLISE DA MARCHA

O corpo humano pode ser definido fisicamente como um complexo sistema de segmentos articulados em equilíbrio estático ou dinâmico, onde o movimento é causado por forças internas atuando fora do eixo articular, provocando deslocamentos angulares dos segmentos, e por forças externas ao corpo. A ciência que descreve, analisa e modela os sistemas biológicos é chamada Biomecânica. O ramo da Biomecânica que busca explicar como as formas de movimento dos corpos de seres vivos acontecem na natureza, a partir de parâmetros cinemáticos e dinâmicos é chamada *biomecânica do movimento* (ZERNICKE, 1981).

Dentre as funções básicas do corpo humano está a locomoção. Ela é atingida através de movimentos coordenados dos segmentos corporais numa interação dinâmica das forças internas (muscular e articular) e forças externas (inercial, gravitacional, friccional, etc.) (ZERNICKE, 1981). Dentre as formas de locomoção do ser humano, a marcha é a maneira mais comum. Assim, se tratando da marcha, a biomecânica do movimento estuda como modelar seus aspectos físicos como forças, torques, força de reação do solo, e acelerações, por exemplo.

O estudo fotográfico de Marey e Muybridge nos anos de 1870 marcou o início da contribuição científica para a análise da marcha (BAKER, 2007). Tradicionalmente, um ciclo completo do andar é determinado por dois contatos consecutivos do mesmo calcanhar no solo e o intervalo de tempo entre estes eventos é a duração do ciclo do andar (ALLARD,

1995). A distância percorrida nesse intervalo é chamada de passada. O início do ciclo de um passo é determinado pelo contato do calcanhar de uma perna e se finaliza com o mesmo movimento, porém com a outra perna. A FIG. 4.1 mostra as diferenças entre passo e passada.

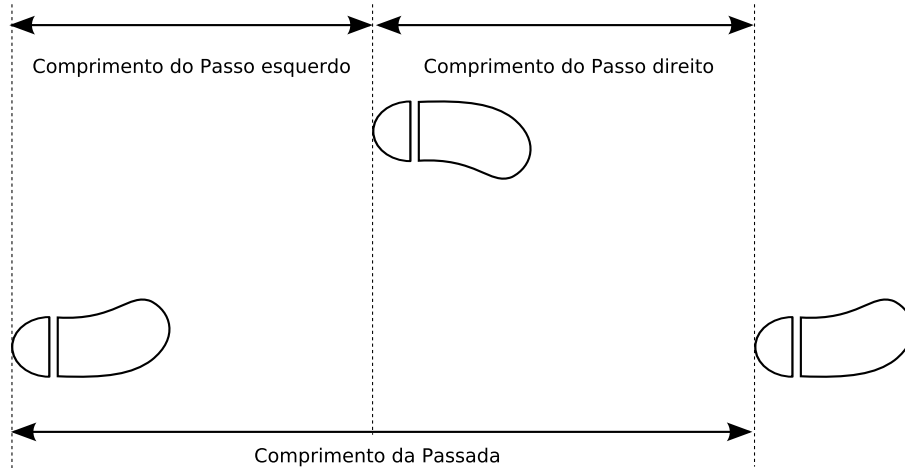


FIG. 4.1: Diferenças entre passo e passada.

Segundo (CHAN, 1981), uma passada é composta de um ciclo de duas fases, chamadas *Fase de Apoio* (*stance phase*) e *Fase de Balanço* (*swing phase*). A fase de apoio corresponde ao período em que o pé está em contato com o solo e pode ser dividida nas seguintes etapas: contato do calcanhar, pé aplainado, apoio médio e impulso final. A fase de balanço, em contrapartida, diz respeito ao período em que o pé não está em contato com o solo e divide-se em: balanço inicial, médio e terminal. Cada uma dessas etapas será abordada a seguir.

Períodos da fase de apoio (conforme ilustrado na FIG. 4.2):

- a) **Contato do Calcanhar** Nesse movimento, a perna de apoio se alonga totalmente, formando um ângulo de 30 graus em relação à perna de balanço, aproximadamente. Então o calcanhar entra em contato com o chão e o corpo está pronto para receber a força de impulsão do outro pé, marcando o início da fase de apoio. (Mais adiante será destacada a importância dessa fase no estudo do som dos passos, devido a esse período ser o responsável pela maior contribuição acústica da passada)
- b) **Pé Aplainado** Nesse período, o fêmur da perna de balanço e o tronco se movem sobre a perna de suporte, devido à força de propulsão provocada pela fase de impulso final, do ciclo anterior. A perna de balanço fica a 15 graus do fêmur da perna de

apoio. Esse movimento provoca uma força de extensão no joelho. Os músculos da panturrilha se contraem formando um ângulo próximo de 15 graus, entre o eixo do fêmur e a parte inferior da perna. Ao fim dessa etapa, o pé de apoio está pronto para receber o peso do corpo.

- c) **Apoio Médio** Nesse estágio, o joelho continua se estendendo até a perna ficar totalmente alongada. Em seguida, a perna de balanço passa pela perna de suporte, se preparando para o início de sua fase de apoio. Quando o pé da perna de suporte está totalmente apoiado no chão, os músculos da coxa se contraem estabilizando o corpo devido a influência de seu peso.
- d) **Impulso Final** Essa ação movimentava o fêmur em 10 graus em relação ao eixo do tronco e o joelho se flexiona em 35 graus em relação ao eixo da parte superior da perna. Ao fim desta ação, os músculos da panturrilha e coxa usam sua energia empurrando o tronco para frente, dando início ao processo de transferência do peso do corpo para a outra perna. Isto marca o fim da fase de apoio de uma perna e aponta seu início para a outra perna.

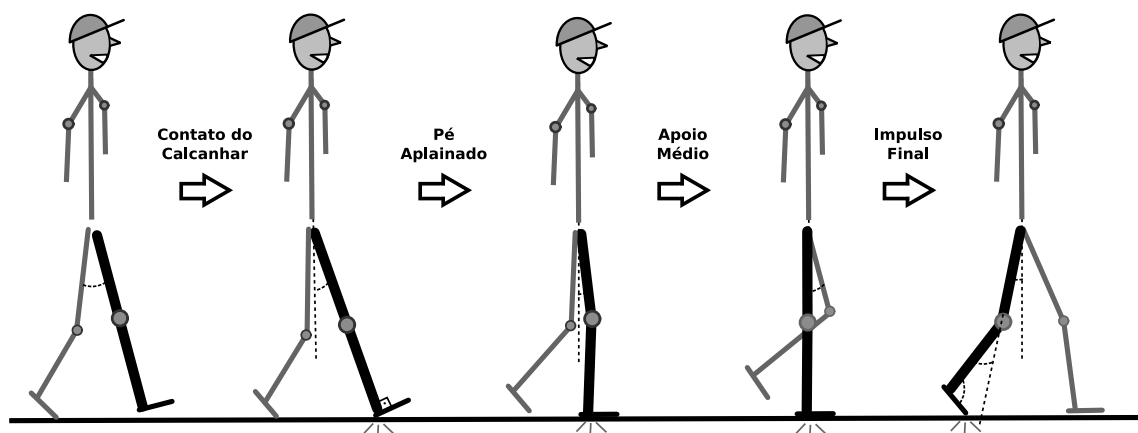


FIG. 4.2: Períodos da fase de apoio.

Períodos da fase de balanço são (conforme ilustrado na Figura 4.3):

- a) **Balanço Inicial** Marcando a fase de balanço, este período se inicia com a parte superior da perna sendo flexionada de 10 para 20 graus em relação ao eixo do tronco e o joelho para 70 graus, em relação ao eixo do fêmur. Esse movimento é produzido pelo trabalho concêntrico do grupo muscular correspondente.

- b) **Balanço Médio** Nesse estágio, o fêmur da perna de balanço se move adiante da perna de apoio, formando um ângulo de aproximadamente 10 graus entre as partes superiores das duas pernas e o ângulo entre a parte inferior da perna e o eixo do fêmur diminui para aproximadamente 30 graus, devido à influência da força gravitacional.
- c) **Balanço Terminal** Por fim, a fase de balanço é findada com o movimento do fêmur para frente, formando um ângulo próximo de 30 graus em relação ao eixo do tronco e o joelho é flexionando para 0 graus, em relação ao eixo do fêmur, deixando a perna totalmente alongada e preparada para o início da fase de apoio.

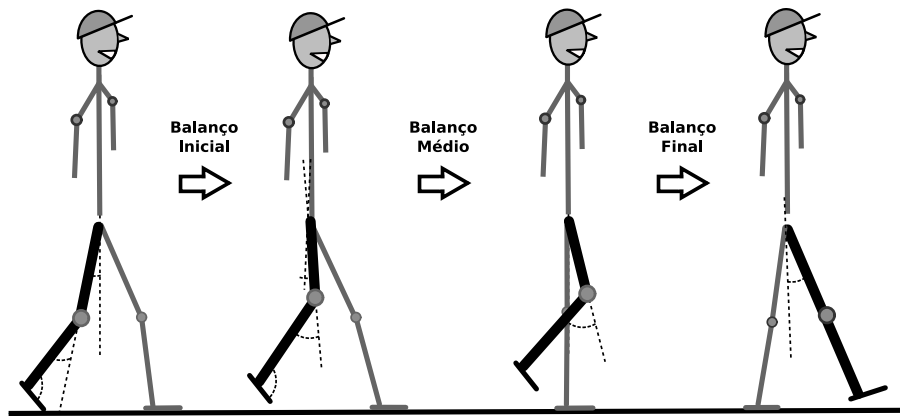


FIG. 4.3: Períodos da fase de balanço.

Enquanto caminha, o pé do indivíduo contribui para a geração de duas forças chamadas: Força de Reação do Solo (FRS) e Força de atrito. Segundo (WINTER, 1984), a FRS representa uma das mais importantes grandezas para análises biomecânicas do movimento, por ser soma algébrica da aceleração da massa de todos os segmentos corporais, ou seja, o total líquido de todas as forças musculares e gravitacionais atuando em cada instante durante a fase de apoio. Como pode ser visto na FIG. 4.4, a FRS é um vetor tridimensional e, segundo (CATTIN, 2002), as três componentes vetor FRS F , conhecidas como anterior/posterior (F_x), vertical (F_y) e lateral/medial (F_z), podem ser medidas com células de carga. A força de atrito é gerada pelo atrito entre o pé do indivíduo e o chão, no ato de caminhar. A força de atrito pode ser obtida utilizando os coeficientes de atrito do chão e da sola do calçado do indivíduo. Serão vistas na próxima seção as vibrações geradas por cada uma dessas forças e a existência de ondas sonoras em frequências ultra-sônicas.

Conhecida a estrutura do caminhar e as forças que são geradas através desse ato, será visto a seguir o tratamento acústico dos passos.

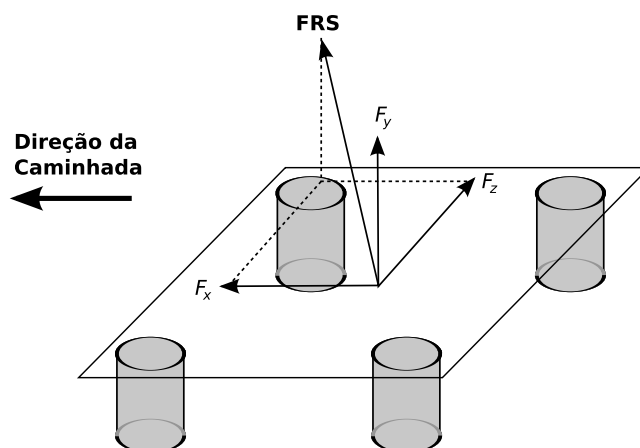


FIG. 4.4: Componentes da força de reação do solo(adaptado de (CATTIN, 2002)).

4.2 ACÚSTICA DOS PASSOS

Como foi visto na seção anterior, ao entrar em contato com o solo, o pé do indivíduo caminhante provoca vibrações causadas pela força de atrito e de reação do solo. Obviamente, essas forças são geradas durante a fase de apoio. Durante o contato do pé com o solo, além das forças geradas, a pressão do ar é alterada durante esse impacto, suscitando um conjunto de ondas sonoras no ar e vibrações no solo caracterizando a acústica do passo.

Em (WATTERS, 1965), foi medida a força de impacto de um único tipo de sapato de salto sólido de um indivíduo, caminhando sobre vários tipos de piso (madeira, concreto e linóleo¹¹ colado ao piso de madeira) em um edifício. Para medir a força, foi utilizado um sensor de força anexado ao salto do calçado de teste. Uma análise de bandas de oitava¹² da força média para um número de passos mostrou um pico máximo na primeira oitava (com o centro da frequência em 31.5Hz). As amplitudes espectrais da força tiveram um crescimento negativo em uma taxa de 30 a 50 dB da primeira oitava à mais alta oitava (centro de frequência em 4 KHz) dependendo do tipo de piso.

Em (GALBRAITH, 1970) foi realizado um experimento para obter o tempo de carga dos passos. O procedimento de teste consistiu em usar uma plataforma com duas placas, uma placa feita de material sólido e a outra coberta com areia, para medir a força dos

¹¹Tecido impermeável, feito de juta e untado com óleo de linhaça e cortiça em pó, usado para tapete(Dicionário Aurélio).

¹²A análise de bandas de oitava consiste em segmentar as frequências do sinal em larguras proporcionais (bandas de oitava) e posteriormente analisá-las. Nesta análise, cada banda é representada por um centro definido.

passos. Com sapatos de modelos idênticos e com meias, três indivíduos caminharam em velocidades diferentes sobre a plataforma de força. Os parâmetros extraídos foram o ritmo e a força peso dos passos. Os resultados desse experimento mostraram que os tipos de superfície e calçado foram relativamente desprezíveis ao que o teste se propôs e foi encontrada uma boa correlação entre o ritmo de passada e o crescimento da carga do passo dividido pelo peso do indivíduo.

Em (CRESS, 1978) foi investigada a velocidade de vibração de partículas provocada por indivíduos se movimentando de formas variadas (rastejante, caminhando em velocidade regular e correndo), em uma área de análise de 6 metros em tipos de solo ao ar livre (*outdoor*). Tal experimento revelou que as mesmas atividades (rastejando, caminhando e correndo) obtiveram diferentes espectros de vibração, em diferentes solos devido as características distintas dos terrenos. Entretanto, diferentes modalidades de locomoção, em um mesmo tipo de solo, resultou em respostas espectrais diferentes apenas por um constante fator de escala da amplitude. Nesse trabalho também foi desenvolvido um modelo para explicar e prever o espectro de vibração do solo, provocado pelo impacto dos passos. De acordo com tal modelo, a resposta da velocidade sísmica (assinatura de vibração) $\widehat{V}(\omega, R, t)$, na frequência $f = (\omega = 2\pi f)$ e distância R da atividade, resulta da multiplicação entre a força do passo $\widehat{F}(\omega, t)$ e a função de transferência do solo $\widehat{T}(\omega, R)$, denotada por:

$$\widehat{V}(\omega, R, t) = \widehat{F}(\omega, t) \times \widehat{T}(\omega, R) \quad (4.1)$$

O sinal ($\widehat{\cdot}$) indica que a função é complexa e t é o tempo. A função de transferência do solo, $T(\omega, R)$, descreve a propagação da vibração entre o local origem da força e a posição do detector. Em seu trabalho, Cress (CRESS, 1978) também demonstrou que a função de transferência $\widehat{T} = (\omega, R)$ é específica ao solo e tem um comportamento semelhante a um filtro passa-faixa, com faixa dentro da banda de frequência de 20 a 90 Hz.

Com base nos estudos de Cress (CRESS, 1978), no trabalho apresentado em (EKIMOV, 2006), foi conduzido um estudo das vibrações acústicas do som e das partículas sísmicas, no piso sólido de uma construção, que foi revestido com linóleo (com uma parte coberta com areia e outra parte coberta com um carpete), onde 6 (seis) indivíduos caminharam em diferentes modos (regular, suave e cauteloso). Em seu teste, (EKIMOV, 2006) mostrou que existem duas bandas de frequência nas ondas sísmicas e sonoras dos

passos. A primeira banda é gerada pela força normal da superfície de contato (FRS) e está concentrada em uma extensão de baixas frequências abaixo de 500 Hz. A segunda banda de frequências é gerada pela força tangencial (de atrito) e está localizada em uma extensão de alta frequência, 1 KHz acima do limiar de frequências ultra-sônicas. No trabalho de (EKIMOV, 2006) foi estabelecida uma relação entre a velocidade de vibração das partículas (V) e o nível de pressão sonora (P). Essa relação se estabelece devido a P ser proporcional à V no ponto de contato entre o pé e o piso sendo representada por

$$\widehat{P}_2(\omega, R) = \widehat{Z}_{12}(\omega, R) \times \widehat{V}_1(\omega) \quad (4.2)$$

onde $\widehat{P}_2(\omega, R)$ é o nível de pressão sonora medida pelo microfone, $\widehat{V}_1(\omega)$ é a velocidade de vibração de partículas no ponto de contato entre o pé e o chão, e $\widehat{Z}_{12}(\omega, R)$ é a função de transferência (impedância) entre a localização do microfone e o ponto de contato. A distância entre esses dois locais utilizada no teste foi $R = 1m$.

Posteriormente, no trabalho apresentado em (EKIMOV, 2007) foi conduzido outro teste para avaliar a onda ultra-sônica gerada pelos passos no trabalho anterior [(EKIMOV, 2006)], porém em um ambiente ao ar livre (outdoor). O ambiente de teste foi uma faixa circular de 1 m de diâmetro, onde um homem caminhou em velocidade constante em três modos diferentes (regular, suave e cauteloso). Para captar as ondas ultra-sônicas, um sensor cerâmico (25OSR), com o ângulo de captação de 60 graus, foi instalado a 1.2 m de altura e 6 m de distância do centro do círculo(essa distância foi necessária para a captação do sensor cobrir toda a faixa). Para captar as vibrações do solo, um acelerômetro de três eixos (PCB 356B21) foi colocado no topo de uma estaca introduzida no meio do círculo. Com o teste, foi detectado uma máxima, na resposta das vibrações do solo, em uma banda de frequência próxima a 19 Hz e com significativas diferenças nas amplitudes dos modos de caminhar(em relação ao modo regular, as magnitudes decresceram -7.5 dB e -12.5 dB para os modos suave e cauteloso, respectivamente). O som dos passos foi analisado numa banda estreita, entre 25 e 26 KHz, do espectro de frequências ultra-sônicas e os níveis de pressão sonora em 25.5 KHz sofreram diferenças em torno de 1 dB para os 3 (três) modos de caminhar. A análise do som, mostrou que a magnitude dos sinais ultra-sônicos, no ar, não possuem uma dependência tão forte com o modo de caminhar, ao contrário das vibrações sísmicas.

Em (EKIMOV, 2007), foi estabelecido que a pressão sonora resultante do som dos

passos, $P(\omega, R, \alpha_s, T)$, a uma distância R do caminhante pode ser representada por

$$P(\omega, R, \alpha_s, T) = P_0(\omega, t) \times A_{\alpha_s}(R, \alpha_s) \times A_s(R) \quad (4.3)$$

onde $P_0(\omega, t)$ é a assinatura sonora do passo no local do impacto do pé, $\omega = 2\pi f$, f é a frequência, t é o tempo, $A_{\alpha_s}(R, \alpha_s)$ é a função de absorção sonora, α_s é o coeficiente de absorção do ar e $A_s(R)$ é a função de atenuação devido a dilatação geométrica do som.

A absorção atmosférica do som é causado pelas perdas de atrito no meio de transmissão (EVANS, 1972). O resultado da atenuação sonora $A_{\alpha_s}(R, \alpha_s)$ pela absorção atmosférica é apresentada como uma função da distância R e do coeficiente de absorção α_s em dB/m:

$$A_{\alpha_s}(R, \alpha_s) = 10^{-(1/20) \times \alpha_s R} \quad (4.4)$$

Segundo Evans e Bass *apud* (EKIMOV, 2007), o valor máximo de absorção α_s (em dB/m) no ar em uma temperatura ambiente sob todas as umidades e para frequências acima de 50 kHz é dado por:

$$\alpha_s(f) = 0.000033 \times f \quad (4.5)$$

onde f é a frequência do som em Hz.

Com os trabalhos que examinaram os sinais provenientes da marcha humana, foi possível observar uma evolução dos estudos no intuito de encontrar um modelo para a origem esses sinais. Isso se deve ao fato do avanço em equipamentos de medição e hardwares com maior poder de processamento. Porém, mesmo com essas atualizações, não se chegou a um modelo explicando a combinação (linear, não-linear, convolução, etc) de sinais, mostrando as variáveis envolvidas (tipo de piso, forma de pisada, tipo de calçado, peso do indivíduo, etc), que formam a onda sonora do passo.

Alguns trabalhos reportaram a existência de sinais sonoros em frequências ultrassônicas, geradas pela força de atrito existente no contato do calçado com o piso. Todavia, para tratar tais sinais é necessário um sensor especial, tornando sua aplicação prática inacessível. Além disso, não foi mostrado que através de tais sinais é possível diferenciar indivíduos. Contudo, alguns trabalhos mostraram possibilidades na classificação de indivíduos a partir de características extraídas do som dos passos (vide Seção 2.2).

Diante do exposto, foi visto que não existe um modelo para explicar a geração da onda sonora dos passos. Encontrar tal modelo, foge ao escopo desse trabalho. Entretanto, nesta

dissertação, o problema é abordado no ponto de vista da engenharia, onde serão extraídas características estatísticas, psicoacústicas e pertinentes aos sistemas de identificação de locutor. Posteriormente, as características serão submetidas a um algoritmo classificador para identificar os indivíduos, atendendo os requisitos necessários para o funcionamento do sistema inteligente de controle da casa.

A próxima seção aborda o referido problema de identificação do ponto de vista da engenharia, identifica os pontos principais para um sistema ideal e por fim limita o escopo envolvido por essa dissertação.

4.3 IDENTIFICANDO INDIVÍDUOS A PARTIR DO SOM DOS PASSOS

Em sistemas de identificação de locutor, o sinal de voz pode ser representado por um sistema linear variante no tempo e possui propriedades que variam lentamente no tempo. Porém, o princípio básico de análise da voz diz que curtos segmentos desse sinal podem ser efetivamente modelados como a resposta de um sistema linear invariante no tempo (LIT) excitado por um trem de impulsos quase-periódicos (FECHINE, 2000; JR., 1993; PICONE, 1993; SILVA, 2002; DELIMA, 2001). O sinal $s(n)$ de saída de um LIT é combinado entre a convolução do sinal de excitação $e(n)$ e a resposta ao impulso $\Theta(n)$, representado por:

$$s(n) = e(n) \otimes \Theta(n) \tag{4.6}$$

Aplicando o processo de *deconvolução* no sinal $s(n)$, obtém-se a envoltória do $\Theta(n)$, isto é, a informação do trato vocal. Para a classificação do indivíduo a partir do som de sua voz, algumas características são extraídas do sinal representativo do trato vocal, as quais possibilitam identificar o indivíduo.

Conforme visto na seção precedente e diferente do sinal da fala, ainda não foi estabelecido um modelo para o sinal sonoro gerado pelos passos. Contudo, o problema pode ser abordado através da aplicação de algoritmos que extraiam características da onda gerada pelos passos. Do ponto de vista da engenharia, identificar indivíduos a partir do som dos passos pode ser tratado como a tarefa de encontrar um conjunto de características, extraídas do sinal sonoro provocado pela marcha, onde seja possível aplicar um algoritmo de classificação, possibilitando discriminar os padrões de indivíduos diferentes como também assimilar o padrão de um mesmo indivíduo.

Num sistema de identificação ideal para o problema em questão, é desejada robustez suficiente para superar as dificuldades provocadas pelas variabilidades, as quais incluem principalmente:

- Canal de gravação (microfone utilizado);
- Ruído ambiente e provocado por eventos que geram som (queda de objetos, por exemplo);
- Modo de andar (correndo, andando, etc.);
- Tipo de piso;
- Tipo de calçado (ou descalço);
- Identificar todos os habitantes da casa simultaneamente;
- Versatilidade na amplitude da onda sonora (devido a aumento do peso do indivíduo, por exemplo).

Mesmo que o objetivo, a longo prazo, seja alcançar o sistema ideal de identificação apresentado anteriormente, o escopo da presente pesquisa limita-se em identificar um habitante através de uma sequência sonora do seu caminhar, dentro de um conjunto distinto formado por sequências do caminhar de outros habitantes, usando um conjunto variado de calçados, em um tipo especial de piso (madeira) e em um ambiente com pouco ruído.

Portanto, essa dissertação se propõe em aplicar as características congregadas no Capítulo 3 em amostras obtidas de um experimento real, selecionando as melhores representantes no quesito de classificação, e por fim submetê-las aos algoritmos classificadores reunidos na Seção 3.4. O experimento, a metodologia do sistema proposto e seus resultados serão vistos a seguir, no próximo capítulo.

5 ORGANIZAÇÃO DOS EXPERIMENTOS

O capítulo anterior apresentou as características físicas do passo, suas fases e sub-fases, como também discutiu os principais problemas encontrados ao se utilizar o som como sinal de entrada para classificação dos indivíduos a partir dos passos. O capítulo atual apresenta o processo de aquisição dos dados (sons da passada), metodologia utilizada na condução dos experimentos, o resultado dos algoritmos para extração dos parâmetros, finalizando com a implementação e resultados dos algoritmos classificadores.

5.1 BASE DE DADOS

Para a realização dos testes de atuação do sistema, um experimento foi realizado com 15 (quinze) indivíduos utilizando diversos tipos de calçados. Os sons das passadas dos indivíduos foram adquiridos a uma taxa de amostragem de 44.1 KHz com 16 *bits* de quantização. O microfone utilizado foi o de um fone de ouvido com microfone da *Leadership*¹³ conectado a uma placa de som ADI AD1986A *SoundMAX 6-channel* de alta definição (ANALOG DEVICES, 2007). A escolha desse microfone se deve ao fato de, em experimentos, ter apresentado um nível mediano de sensibilidade quando conectado à placa de som descrita anteriormente e por apresentar um baixo valor comercial.

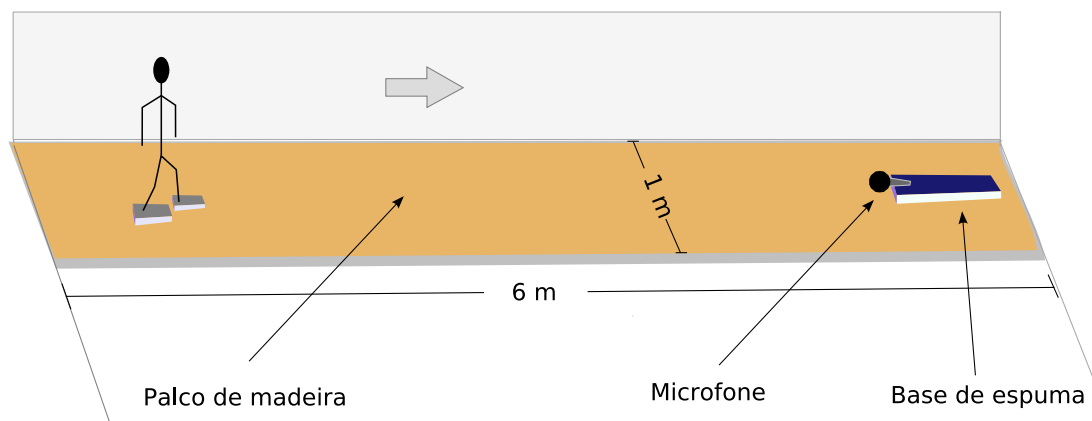


FIG. 5.1: Configuração do experimento.

¹³<http://www.leadership.com.br/>



FIG. 5.2: Palco de madeira onde se sucederam os experimentos.

Durante o experimento foram captadas 15 amostras de cada um dos 15 indivíduos que caminharam sobre um palco de madeira (FIG. 5.2) com dimensão 6x1m (comprimento por largura). De uma extremidade do palco, os voluntários caminharam rumo ao microfone, localizado na outra extremidade (FIG. 5.1). O microfone foi colocado em uma base de espuma para que a influência das vibrações causadas na madeira não implicassem ruído durante o experimento. Foram utilizados 6 (seis) tipos de calçados nesse teste. A FIG. 5.3 mostra os tipos de calçados utilizados e é importante ressaltar que nessa figura aparece somente um tipo de tênis, porém cada indivíduo caminhou com seu próprio calçado, aumentando ainda mais a aleatoriedade da base de dados. Para encontrar a influência entre os dados físicos dos indivíduos e a resposta do sistema, os participantes informaram algumas de suas características físicas as quais estão reunidas na TAB. 5.1. Para avaliar o comportamento do sistema mediante a apresentação de um mesmo indivíduo usando mais de um calçado, foi solicitado a seis indivíduos (ID2 - ID7), realizar o trajeto calçando o tipo de calçado exibido na FIG. 5.3b.

O software para a digitalização das variações em pressão sonora provocada pelos passos, foi utilizado o *Audacity* versão 1.2.6 (OETZMANN, 2008), um editor sonoro de código livre. Durante o experimento, o indivíduo ficou parado no início do caminho, iniciando sua caminhada ao perceber um sinal visual do operador. Devido a esse fato, as amostras de um mesmo indivíduo se diferenciaram no total de milissegundos, fazendo-se necessária

Id.	Peso	Altura	Sexo	Calçado (s)
ID2	67	1,79	M	Tenis/Sapato
ID3	68	1,80	M	Tenis/Sapato
ID4	87	1,85	M	Tenis/Sapato
ID5	70	1,84	M	Tenis/Sapato
ID6	63	1,60	M	Tenis/Sapato
ID7	65	1,82	M	Tenis/Sapato
ID8	70	1,82	M	Tenis
ID9	80	1,92	M	Sapato de Borracha
ID10	57	1,54	F	Sapatilha
ID11	85	1,70	M	Sapatenis
ID12	95	1,90	M	Tenis
ID13	70	1,59	M	Coturno
ID14	60	1,70	M	Coturno
ID15	53	1,63	M	Coturno
ID1	63	1,63	F	Coturno

TAB. 5.1: Dados físicos dos indivíduos participantes do experimento.

a padronização das mesmas, assunto abordado com mais abrangência na Seção 5.2.1.

Para uma identificação direta a respeito da amostra, os nomes dos arquivos adquiridos também seguem um padrão. A FIG. 5.4 mostra a metodologia da organização utilizada. Esse padrão de nomes foi adotado para facilitar a identificação das amostras nos gráficos dos resultados.

Apesar de ser comum caminhar em ambientes domésticos calçando-se chinelos, pantufas ou mesmo sem calçado algum, tais situações não foram abordadas no experimento devido às limitações do equipamento de aquisição sonora utilizado. Infelizmente, os sons captados usando-se duas (chinelos e sem calçado), das três situações discutidas, ficaram com uma amplitude muito baixa, quase imperceptíveis, impossibilitando a separação entre o sinal útil do ruído ambiente. Porém, acredita-se que a capacidade de identificação do sistema também cobrirá tais situações mediante o uso de um microfone mais sensível e a aplicação de filtros específicos para aumentar a taxa de sinal ruído (S/R), bem como eliminar¹⁴ o ruído ambiente.

¹⁴Sabe-se que não existe filtro ideal, porém acredita-se que é possível projetar um conjunto de filtros adaptativos capazes de realizar essa separação.



FIG. 5.3: Tipos de calçados utilizados no experimento. (a) ST1: Coturno. (b) ST2: Tênis. (c) ST3: Sapato. (d) ST4: Sapato de borracha. (e) ST5: Sapatilha. (f) ST6: Sapatênis.

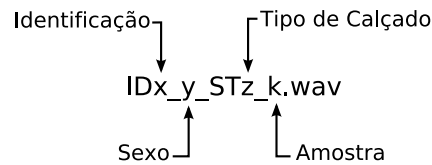


FIG. 5.4: Padrão de nomenclatura utilizado no armazenamento das amostras do experimento realizado.

5.2 METODOLOGIA

Para analisar o comportamento do sistema mediante a variedade de calçados no experimento explicado na seção anterior, foi proposta uma seleção de 6 (seis) subconjuntos derivados do experimento. De maneira induzida os conjuntos foram selecionados da seguinte forma:

- 1 **COTURNO (ST1)** Constituído por 4 (quatro) integrantes do conjunto que utilizaram o tipo de calçado coturno (ID1, ID13, ID14 e ID15);
- 2 **TENIS (ST2)** Composto pelas amostras de 8 (oito) indivíduos (ID2, ID3, ID4,

- ID5, ID6, ID7, ID8 E ID12) que caminharam utilizando o tipo de calçado tênis;
- 3 **SAPATO (ST3)** Formado por 6 (seis) indivíduos (ID2, ID3, ID4, ID5, ID6, ID7) que calçaram sapato;
 - 4 **UMPORCALCADO** Formado por 6 (seis) indivíduos (ID1, ID8, ID2, ID9, ID10, ID11);
 - 5 **MESMOCALCADO** Contém os indivíduos que compartilharam do mesmo tipo de calçado (ID2, ID3, ID4, ID5, ID6, ID7)¹⁵ no experimento;
 - 6 **IGNORECALCADO** Constituído por todas as amostras de todos os indivíduos utilizando todos os tipos de calçados.

A análise proposta neste trabalho consiste de três fases: 1. Pré-processamento; 2. Extração e seleção das características; e 3. Identificação. A FIG. 5.5 mostra a visão geral do sistema separado em quatro partes (os blocos A a D). No bloco A da figura, está concentrado o banco de amostras (vide Seção 5.1), ou seja, os arquivos contendo a gravação do som dos passos. No bloco B, encontra-se a etapa de pré-processamento. É nessa fase que os arquivos serão filtrados (bloco B.3 da figura), haverá a detecção de estímulos (a duração do som provocado pelo impacto do pé ao solo, mostrado no bloco B.1) e a seleção dos estímulos a serem utilizados (bloco B.2 da figura). Esta última etapa seleciona os passos do meio do arquivo. Isso é necessário para obter o estado mais natural do caminhar, pois no início da caminhada a pessoa está acelerando e no término está desacelerando. A Seção 5.2.1 mostra os detalhes de implementação do pré-processamento.

Ainda observando a FIG. 5.5, o bloco C (extração e seleção de parâmetros) se divide em 3 sub-blocos. No sub-bloco onde estão contidos os quadros C.1 a C.4, estão reunidos os parâmetros físicos e estatísticos da onda sonora. E o sub-bloco com os quadros C.5 a C.7 estão concentrados os parâmetros psicoacústicos. O terceiro sub-bloco (C.8 na FIG. 5.5) é a fase onde as melhores características são selecionadas para então serem submetidas ao algoritmo de classificação (bloco D). As duas setas ligando os blocos B e C significam a entrada de dados filtrados e não filtrados, isto é necessário para um comparativo da importância do filtro dos dados (a tarefa filtro neste trabalho se relaciona a diminuir ao máximo o ruído ambiente). A fase de extração e seleção de parâmetros será descrita na Seção 5.2.2.

¹⁵Note que são os mesmos indivíduos do grupo SAPATO, porém com amostras de mais de um calçado.

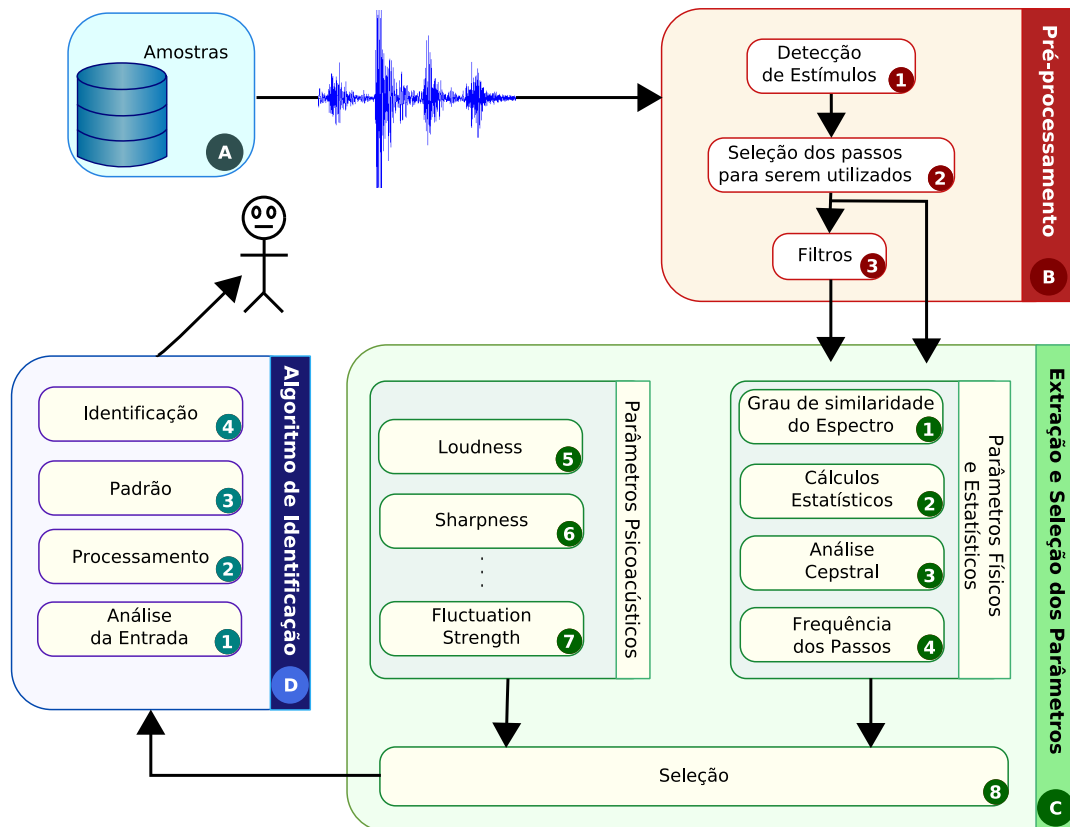


FIG. 5.5: Visão geral do sistema.

Por fim, o bloco D da figura FIG. 5.5 recebe as características selecionadas e trabalha no objetivo de identificar o indivíduo. As etapas contidas nos quadros D.1 a D.4 foram colocadas como passos de um algoritmo genérico de classificação, entretanto foram testados dois algoritmos de classificação. A fase de identificação é abordada em maiores detalhes na Seção 5.2.3.

5.2.1 PRÉ-PROCESSAMENTO

Como na maioria dos sistemas de identificação, a fase de pré-processamento é uma das etapas mais importantes do sistema. Nessa fase, os sinais serão preparados para se submeterem aos algoritmos de extração de características. Fazendo referência à FIG. 5.5, neste trabalho identificou-se a necessidade de realizar três tarefas nesta fase: B.1) Detecção de estímulos; B.2) Seleção dos passos; e B.3) Filtros.

Durante a realização dos experimentos, o tamanho das amostras captadas é variável. Isso ocorre devido ao tempo que o operador do computador tem para habilitar a função de gravação e avisar, através de sinais visuais, o indivíduo, somado ao tempo que este tem

para reagir e começar a andar. Além disso, como já foi discutido no capítulo anterior, o tempo de caminhada é diferente entre os indivíduos. Devido a esses problemas, criou-se a necessidade de uma forma de identificar os passos na onda. Fazendo uma conexão com a literatura de identificação de locutor, seria o problema de identificar o início e fim das palavras.

Na literatura de tratamento do sinal sonoro produzido pela voz, é possível classificar os sons em três classes: sonoros, surdos e explosivos. O problema então é separar os segmentos de voz nessas classes. Para efetuar essa classificação, o som da voz é separado em pequenos segmentos chamados janelas. Em cada janela é então aplicado um cálculo de alguma métrica. As métricas mais comuns são a taxa de cruzamento por zero (ZCR) e o nível de energia do sinal (SEL). Se determinada janela tem um SEL ou ZCR em determinado nível então recebe uma classificação. No problema dos passos seria necessário somente identificar onde existe estímulo provocado pelo contato do pé com o solo e quando não tem.

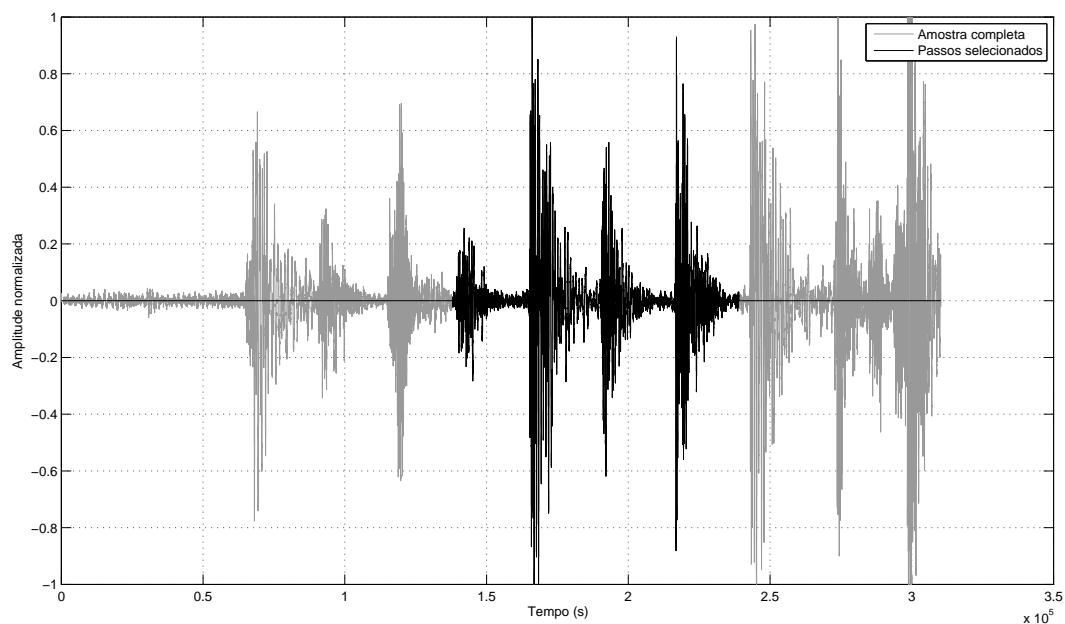


FIG. 5.6: Exemplo de seleção de passos em uma amostra completa.

A FIG. 5.6 mostra uma amostra de um dos experimentos. Ao se observar a FIG. 5.6, fica fácil separar os estímulos provocados pelo contato do pé do indivíduo e o solo. Utilizando os mesmos métodos usados na classificação de sons sonoros e não-sonoros da literatura de reconhecimento de locutor, não foi possível chegar a um nível padrão ou

a um comportamento computacionalmente explicável dos níveis ZCR e SEL para um julgamento eficaz de quais segmentos são estímulos e quais não são. Porém elaborar um algoritmo para realizar essa tarefa automaticamente não foi possível.

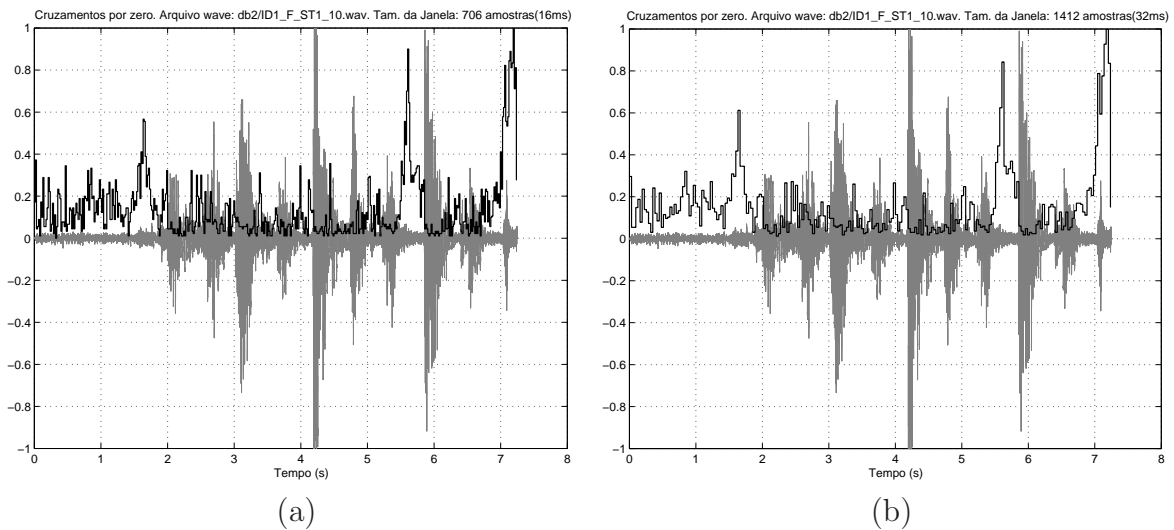


FIG. 5.7: ZCR normalizado com janela de 16 ms e 32 ms. (a) ZCR normalizado, com janela de 16ms (706 amostras). (b) ZCR normalizado, com janela de 32ms (1412 amostras).

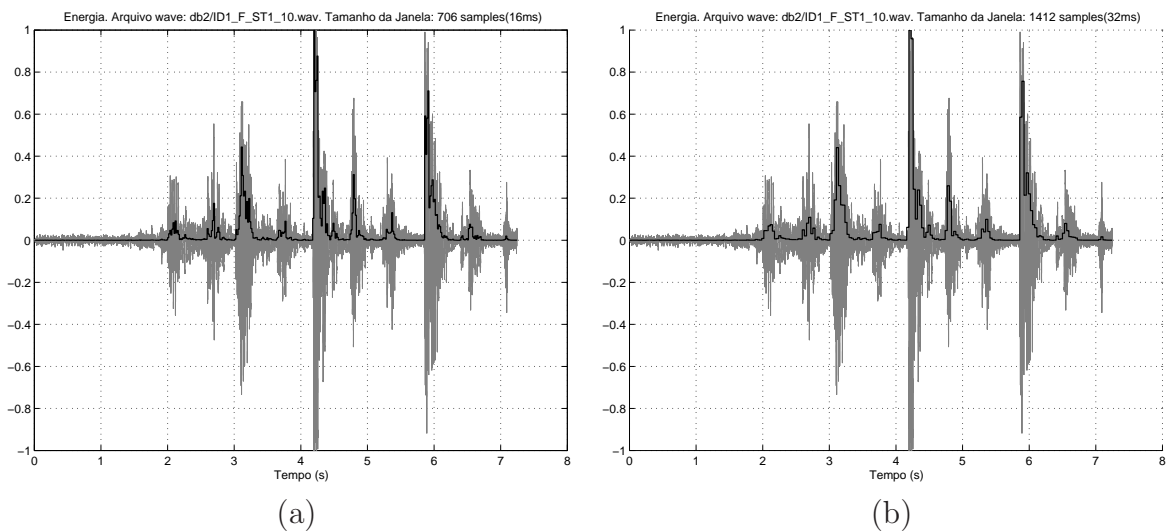


FIG. 5.8: SEL normalizada com janela de 16ms e 32ms. (a) SEL normalizado, com janela de 16 ms (706 amostras). (b) SEL normalizado, com janela de 32 ms (1412 amostras).

Para chegar a tal conclusão, algumas amostras do experimento foram selecionadas aleatoriamente e feito o cálculo de SEL e ZCR utilizando dois tamanhos diferentes de

janela. De acordo com alguns trabalhos de reconhecimento de locutor, valores típicos de janela são 16 ms e 32 ms (FECHINE, 2000), portanto esses dois valores foram utilizados na tentativa de encontrar um nível de corte para tentar classificar as janelas com e sem estímulo. As figuras FIG. 5.7a e FIG. 5.8a, e FIG. 5.7b e FIG. 5.8b mostram os níveis de ZCR e SEL calculados para janelas de 16 ms e 32 ms, respectivamente. Ao observar a FIG. 5.7, percebe-se que não há um padrão possibilitando a identificação dos sinais com estímulo. Na FIG. 5.8, nota-se que a energia tende a crescer na presença de estímulos, entretanto alguns limiares foram testados porém não se obteve êxito. Diferentemente dos sinais de voz, os quais na literatura se encontram algoritmos para separação dos sinais vozeados, que se baseiam nos valores de SEL e ZCR, para a onda dos passos é requerido um algoritmo mais robusto.

Para o sistema é importante receber os passos da forma mais natural possível. Em cada amostra do experimento, os indivíduos passam por três fases: 1) fase de aceleração; 2) fase de estabilidade; e 3) fase de desaceleração. Para o sistema é importante trabalhar com a forma mais natural do andar e essa forma ocorre na fase de estabilidade. Neste trabalho, assumiu-se que os quatro passos do meio do arquivo pertencem à fase de estabilidade. Como o processo de identificação de estímulos não pode ser automatizado, a identificação e seleção dos passos foi feita manualmente. Para cada arquivo *.wav*, quatro estímulos foram selecionados (vide FIG. 5.6) sendo recortada 101000 (2.29 s, aproximadamente) amostras de cada uma delas. Os processos acima descritos fizeram parte do pré-processamento das amostras obtidas nos experimentos para a próxima fase do sistema, extração e seleção das características, a qual será descrita na Seção 5.2.2.

5.2.2 EXTRAÇÃO E SELEÇÃO DAS CARACTERÍSTICAS

Após receber o sinal tratado na etapa predecessora, cada amostra é enviada aos algoritmos de extração de características. De maneira a avaliar os resultados de cada um dos métodos reunidos na Seção 3.2, a seguir é feita uma análise da decorrência dos algoritmos extra-tores de parâmetros, para os grupos COTURNO, TENIS, SAPATO, UMPORCALCADO, MESMOCALCADO e IGNORECALCADO definidos na seção anterior.

FREQUÊNCIA DOS PASSOS

A FIG. 5.9 mostra o período médio dos passos obtido pelo algoritmo *maxValueFootstepId* para todos os grupos. Cada indivíduo foi separado distribuindo os períodos médios de

suas amostras por múltiplos de 0.5 no eixo das ordenadas. Nessa forma de representação, é possível observar as semelhanças entre a média dos passos. De modo geral, o período médio dos passos não mostrou ser uma característica tão discriminante. Pois, para o grupo COTURNO (FIG. 5.9a), por exemplo, a maioria das amostras não podem ser separadas. Na FIG. 5.9b, existe o dobro de amostras que nos demais grupos, isso se deve ao fato de estar reunidas amostras de dois tipos de calçado: tênis (ST2) e sapato (ST3). Como pode ser visto na FIG. 5.11b, o indivíduo ID12 obteve um período de caminhada praticamente constante.

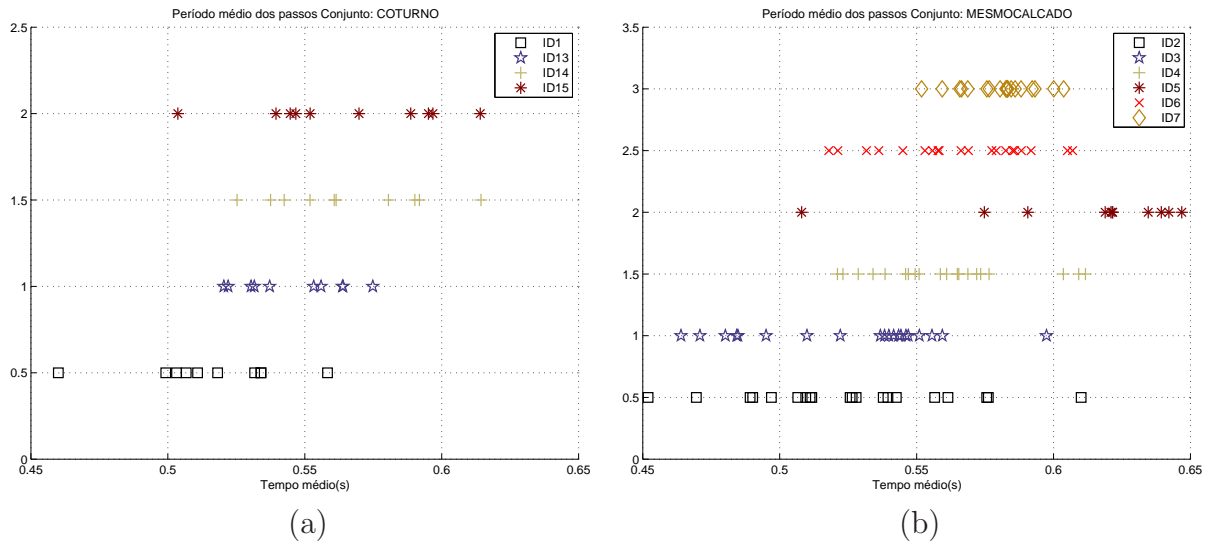


FIG. 5.9: Período médio dos passos extraído dos indivíduos participantes dos grupos COTURNO (a) e MESMOCALCADO (b).

ENVELOPE ESPECTRAL

Duas análises visuais foram feitas para avaliar a qualidade dos dois algoritmos desenvolvidos para estimação do envelope espectral. A análise 1 consistiu em mostrar os dados do envelope de cada amostra, variando os indivíduos. E a análise 2 versou em mostrar, para cada indivíduo, os envelopes de todas as amostras. A primeira análise visa destacar as diferenças, enquanto que a segunda mostra a similaridade das amostras de um mesmo indivíduo.

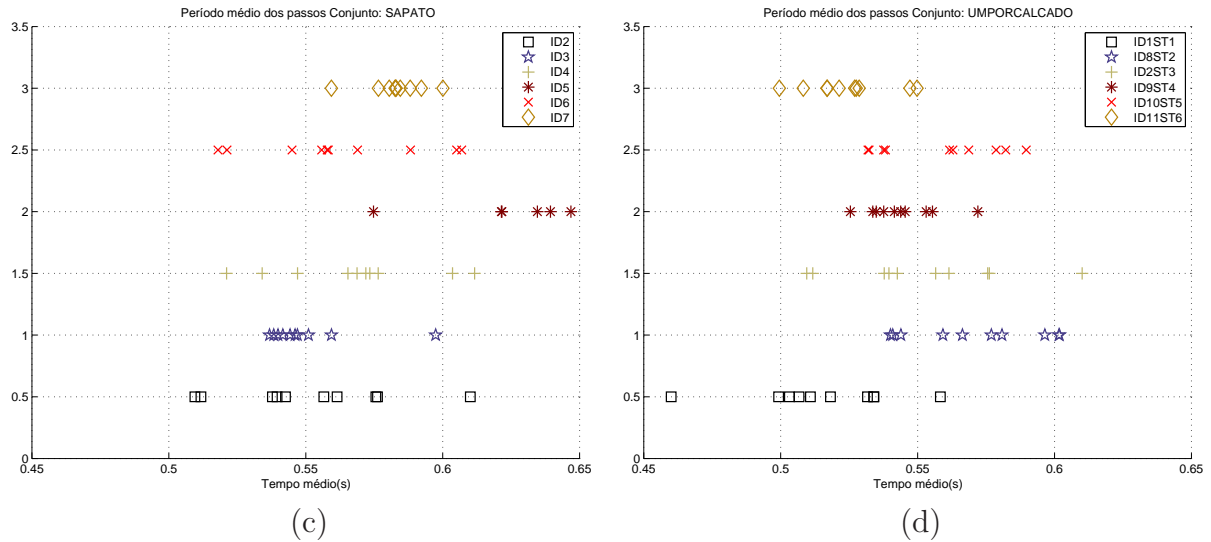


FIG. 5.10: Período médio dos passos extraído dos indivíduos participantes dos grupos SAPATO e UMPORCALCADO.

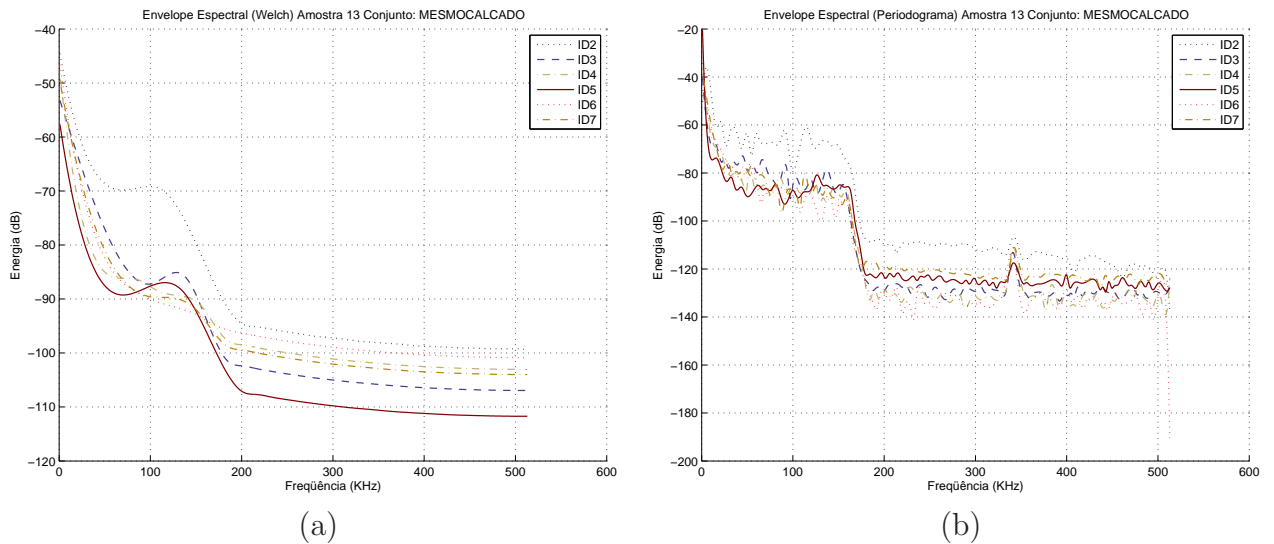


FIG. 5.12: Análise 1 do Envelope Espectral obtido através dos algoritmos Welch e Periodograma. São exibidos os envelopes da amostra 13 do conjunto MESMOCALCADO, obtidos através do método de Welch (a) e Periodograma (b).

Na FIG. 5.12 estão reunidos os envelopes de uma amostra de cada indivíduo participante do grupo MESMOCALCADO, utilizando os dois métodos (Welch e Periodograma). Observando o resultado dos dois algoritmos, para a amostra considerada, é possível diferir, visualmente, os indivíduos a partir de seus envelopes espectrais.

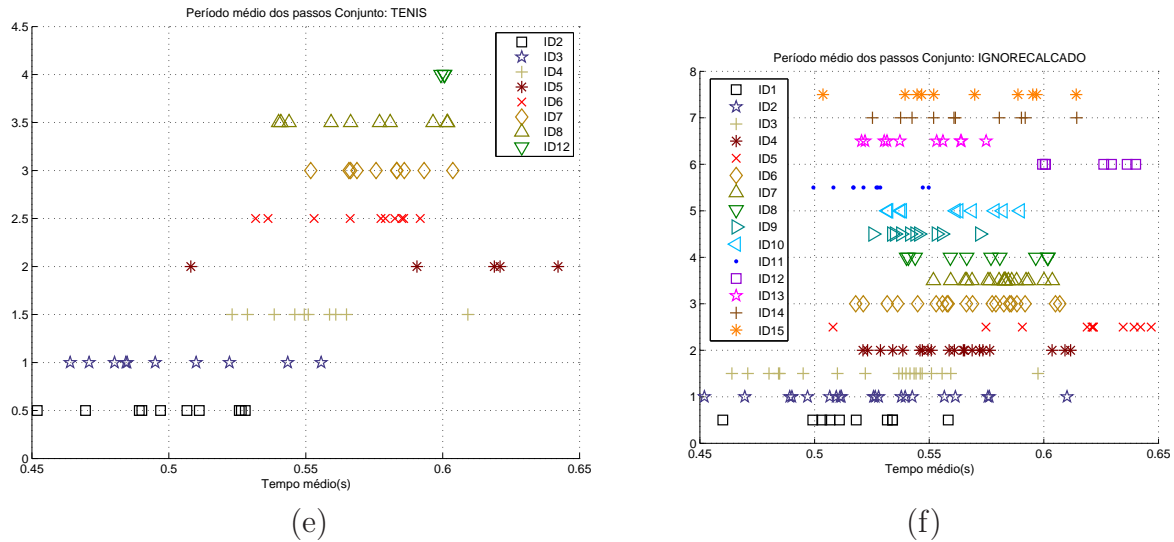


FIG. 5.11: Período médio dos passos extraído dos indivíduos participantes dos grupos TENIS e IGNORECALCADO.

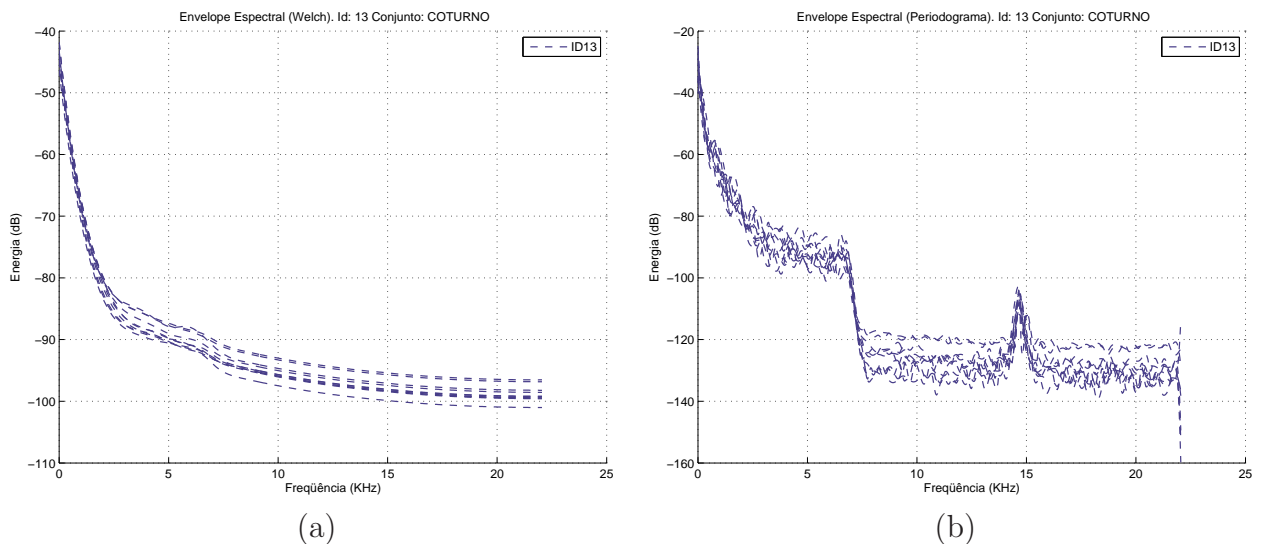


FIG. 5.13: Análise 2 do Envelope Espectral obtido através dos algoritmos Welch e Periodograma. São exibidos os envelopes do indivíduo 13 usando coturno (ST1), estimados pelo método baseado no Welch (a) e Periodograma (b).

A FIG. 5.13 mostra os envelopes calculados para as 10 amostras do indivíduo ID13, obtido pelos dois métodos. Nota-se através da FIG. 5.13a um grau de similaridade entre os envelopes calculados por meio do algoritmo baseado no método de Welch. O mesmo ocorre na FIG. 5.13b os envelopes reunidos de um mesmo indivíduo, quando o cálculo do envelope baseou-se no método baseado no Periodograma.

Pela análise visual, os métodos avaliados se mostraram discriminantes. Entretanto, uma avaliação mais completa será realizada nas seções posteriores.

COEFICIENTES MEL-CEPESTRAIS

Os coeficientes mel-cepestrais geram uma massa de dados consideravelmente grande, em relação às demais características. Cada coeficiente consiste em um vetor de 226 posições, que são as saídas dos filtros triangulares utilizados. O número de coeficientes escolhidos varia de aplicação para aplicação. Vinte coeficientes são geralmente utilizados em sistemas de reconhecimento de palavras e de locutores (DELIMA, 2001; FECHINE, 2000). Para esta aplicação, vinte coeficientes foram escolhidos e serão analisados segundo seu poder de representatividade.

Para avaliar os coeficientes utilizados, o mesmo processo aplicado às demais características foi realizado, porém com a seguinte diferença: cada coeficiente foi tratado como uma característica. Por serem muitos dados, serão mostrados alguns coeficientes aplicados a alguns calçados e indivíduos.

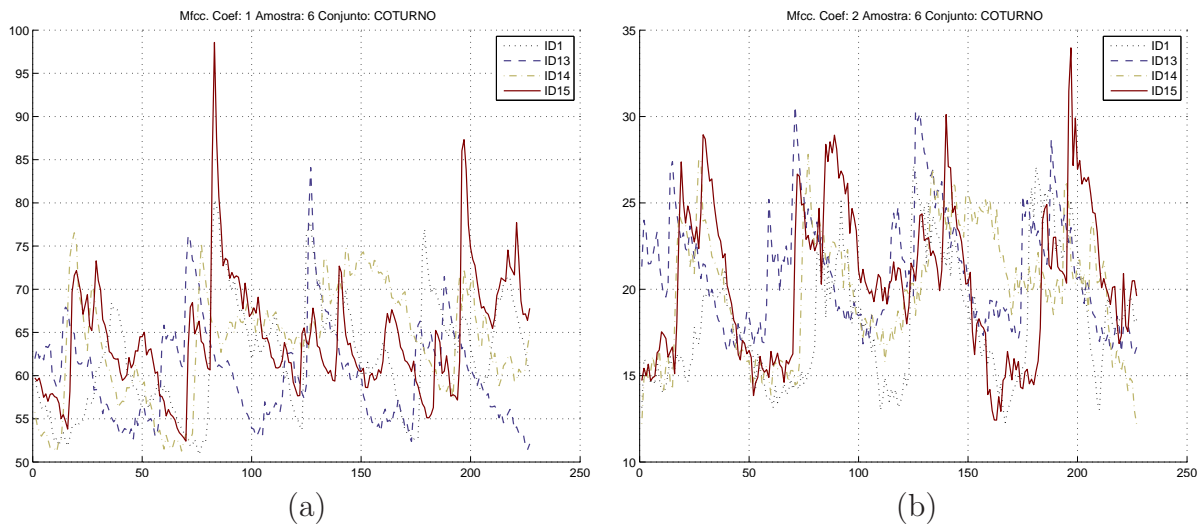


FIG. 5.14: Análise 1 dos coeficientes mel-cepestrais. Em (a) e (b) são exibidos os dois primeiros coeficientes obtidos usando as amostras dos participantes do conjunto COTURNO.

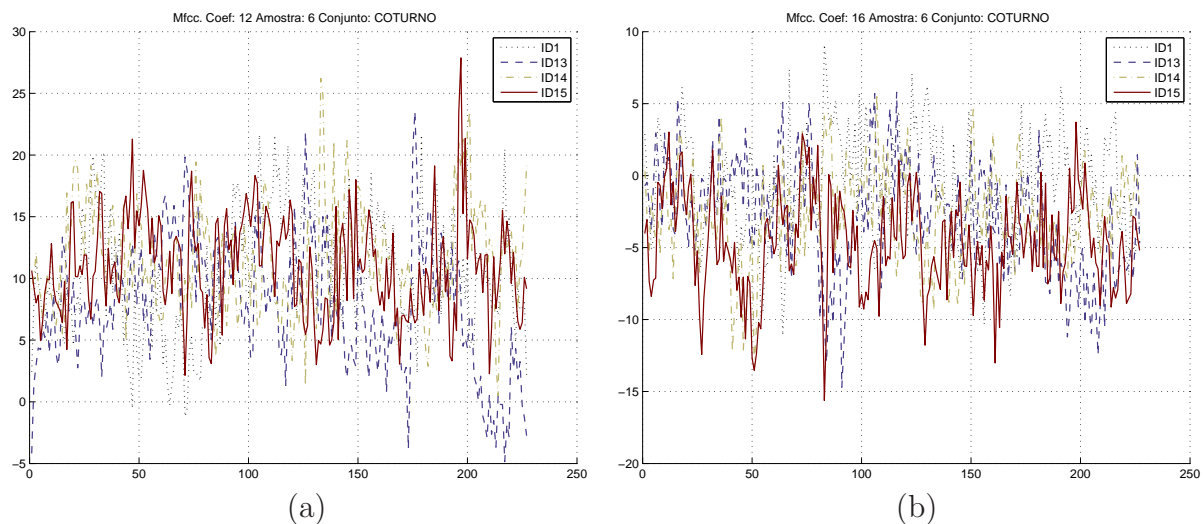


FIG. 5.15: Análise 1 dos coeficientes mel-cepestrais. Em (a) e (b) são exibidos o 12º e 16º coeficientes, respectivamente, obtidos usando as amostras dos participantes do conjunto COTURNO.

Nas FIG. 5.14a e FIG. 5.14b são mostrados os dois primeiros coeficientes mel-cepestrais. Como pode ser visto, é possível distinguir entre os indivíduos, partindo-se desses dois primeiros coeficientes. Entretanto no 12º e 16º coeficientes (ilustrado nas FIG. 5.16a e FIG. 5.16b), não é possível realizar essa separação visualmente. Uma possível justificativa para esse ocorrido é devido ao sinal dos passos apresentar energia em frequências abaixo de 1 KHz. Contudo para a confirmação dessa hipótese é necessário uma análise mais minuciosa, utilizando-se analisadores sofisticados.

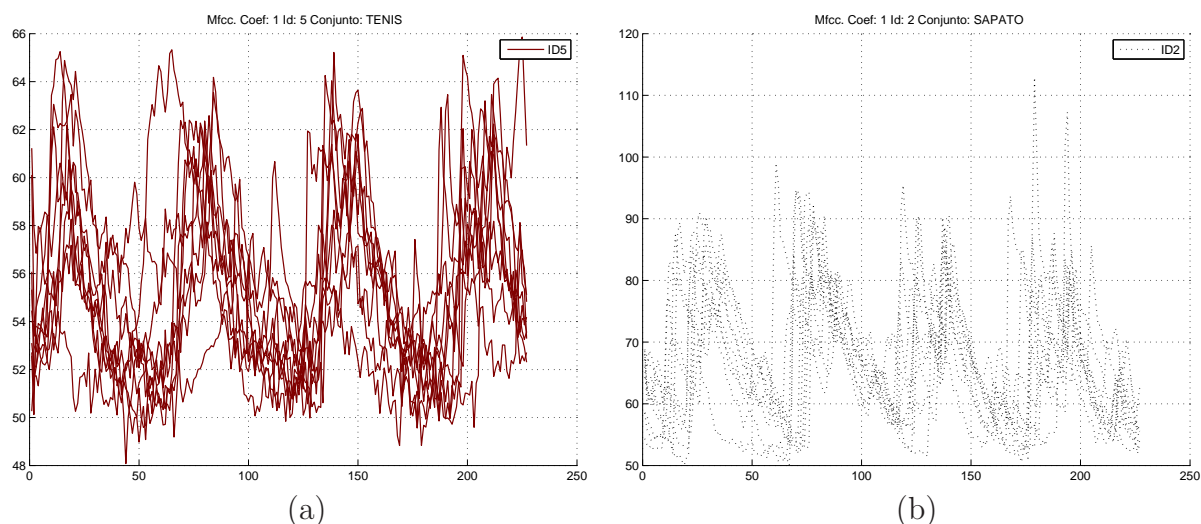


FIG. 5.16: Análise 2 dos coeficientes mel-cepestrais. Em (a) o primeiro coeficiente de todas as amostras do indivíduo ID5 calçando tênis (ST2). Em (b) é mostrado o primeiro coeficiente de todas as amostras do indivíduo ID2 utilizando sapato (ST3).

Na análise 2 dos coeficientes mel-cepestrais, a FIG. 5.16a exibe o primeiro coeficiente para o indivíduo ID5 utilizando tênis e a FIG. 5.16b mostra o primeiro coeficiente para o indivíduo ID2 utilizando sapato. Em ambas figuras, pode ser visto que os coeficientes de um mesmo indivíduo estão praticamente relacionados por um fator escalar. Todavia, não é possível afirmar que na presença dos primeiros coeficientes de outros indivíduos é possível discriminá-los. Alguns dos demais coeficientes se comportam de maneira aleatória, entretanto não se pode descartá-los por essa análise visual. Os métodos de seleção de características julgarão quais são mais discriminativas, justificativa suficiente para a existência dessa etapa.

SONORIDADE ESPECÍFICA

Como pode ser observado nas FIG. 5.17a e FIG. 5.17b, a partir da medida de sonoridade específica distribuída em torno de 24 bandas críticas, é possível visualmente separar indivíduos distintos. Além disso, em FIG. 5.18a e FIG. 5.18b mostra um comportamento similar entre as amostras de um mesmo indivíduo.

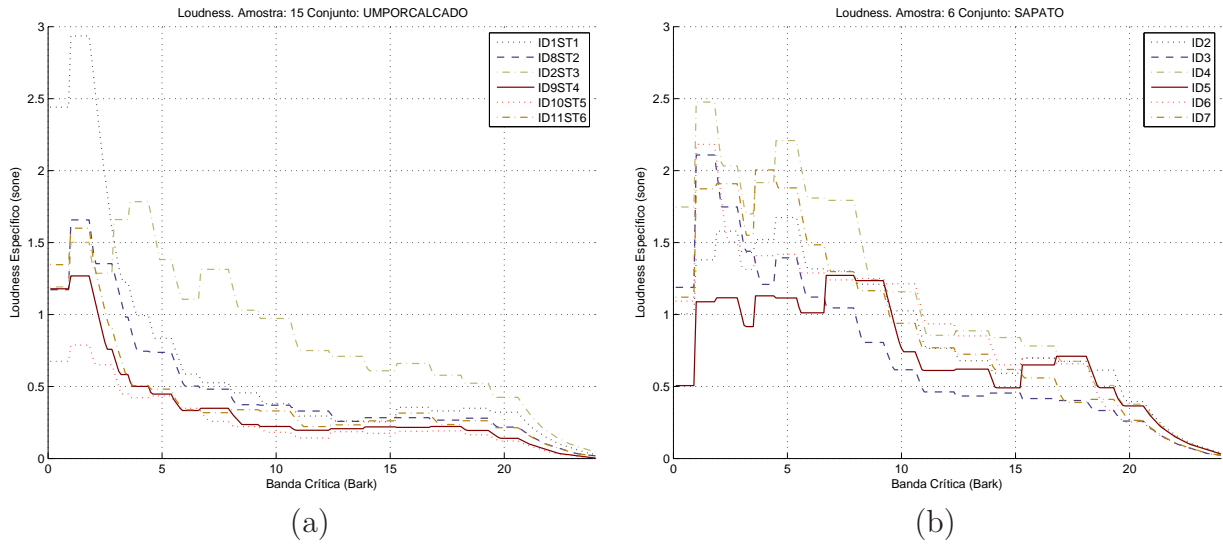


FIG. 5.17: Análise 1 do *sonoridade* segundo o modelo de *Zwicker*. Em (a) é mostrado a *sonoridade* específica em torno de 24 bandas críticas, para os indivíduos participantes do conjunto UMPORCALCADO. Em (b) a *sonoridade* específico em torno de 24 bandas críticas para os indivíduos participantes do conjunto SAPATO.

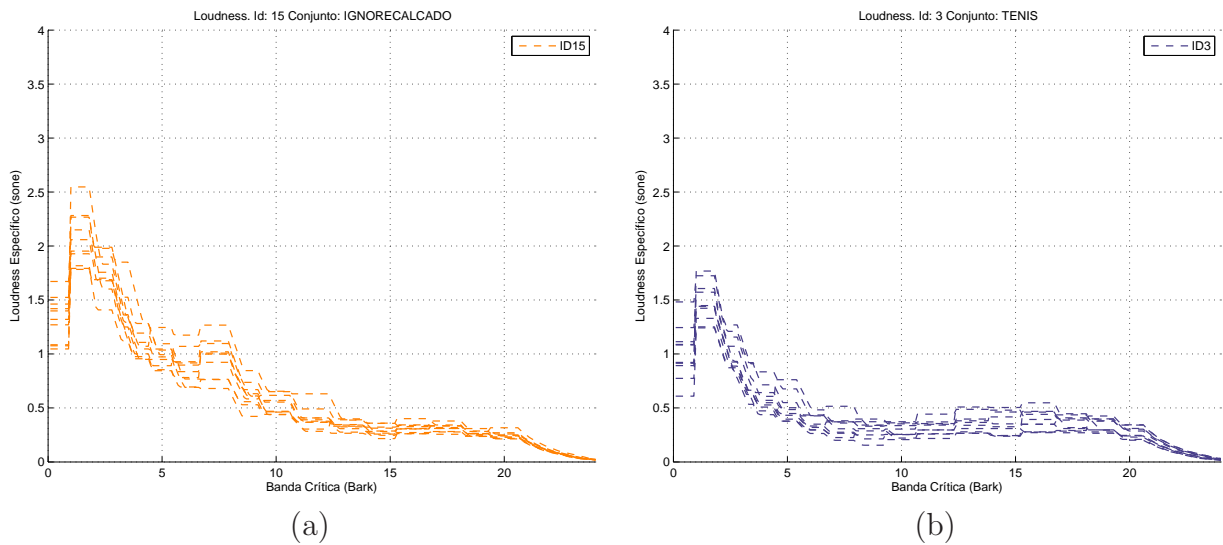


FIG. 5.18: Análise 2 do *sonoridade* segundo o modelo de *Zwicker*. Em (a) é mostrado a *sonoridade* específica em torno de 24 bandas críticas, para as amostras do indivíduo IGNORECALCADO. Em (b) a *sonoridade* específica em torno de 24 bandas críticas para o indivíduo ID3 utilizando o tipo de calçado TENIS.

COEFICIENTES CEPESTRAIS

Como pode ser observado nas FIG. 5.19a e FIG. 5.19b, é possível notar uma pequena diferença visual entre os coeficientes de indivíduos distintos. Nas FIG. 5.20a e FIG. 5.20b nota-se similaridade entre os coeficientes cepestrais de um mesmo indivíduo. Não obstante visualmente as diferenças não serem tão notáveis quanto as outras, não é possível julgar os coeficientes cepestrais. Essa tarefa se dará através da avaliação descrita nas seções posteriores.

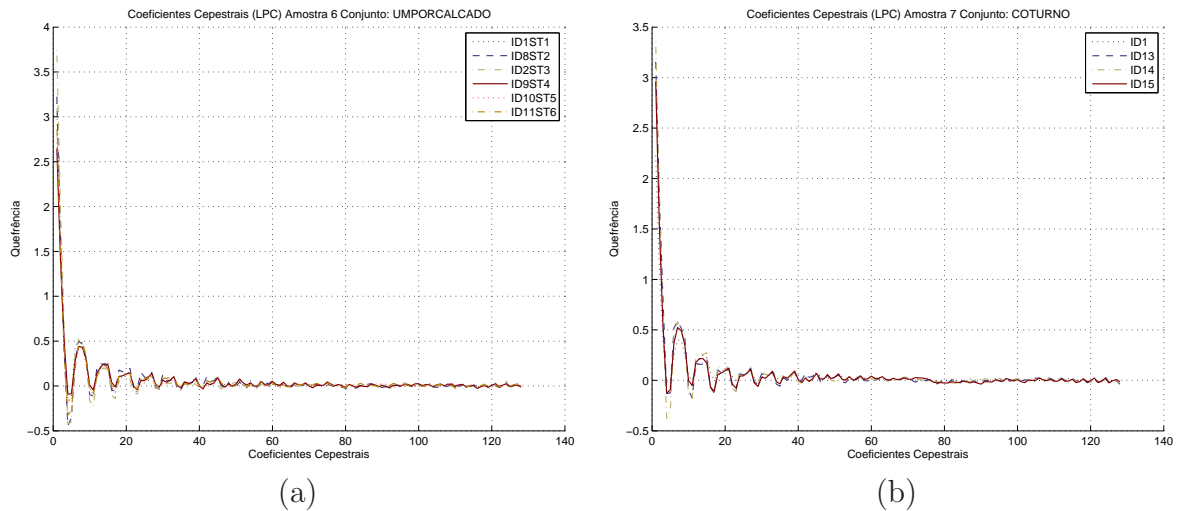


FIG. 5.19: Análise 1 dos coeficientes cepestrais. Em (a) os coeficientes da amostra 6 do conjunto UMPORCALCADO. Em (b) são reunidos os coeficientes cepestrais da amostra 7 de todos os indivíduos participantes do conjunto COTURNO.

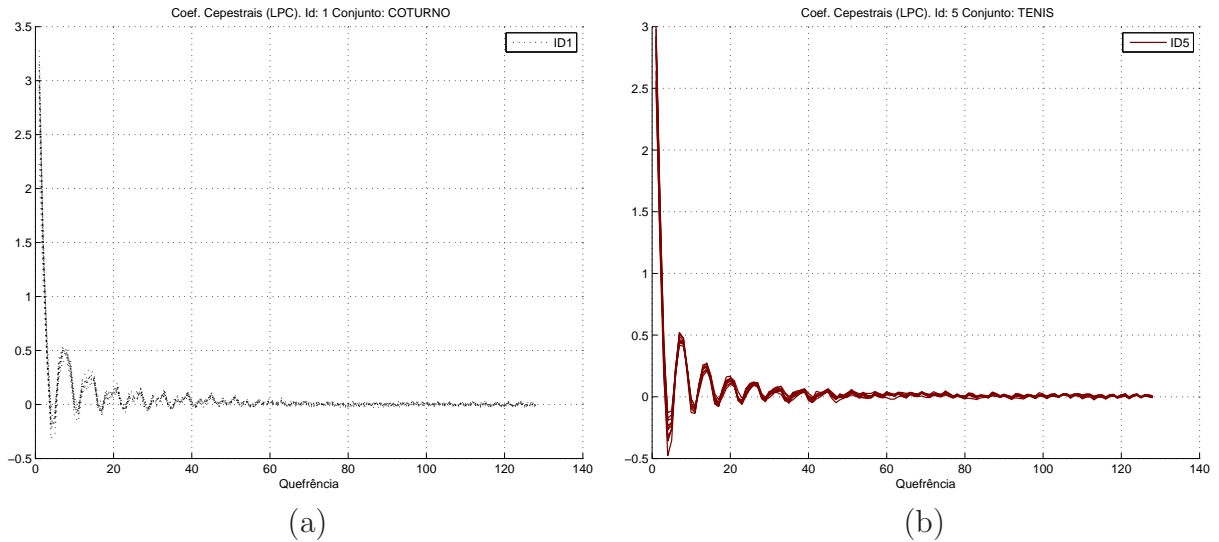


FIG. 5.20: Análise 2 dos coeficientes cepstrais. Em (a), os coeficientes cepstrais de 10 amostras do indivíduo ID1 utilizando o tipo de calçado coturno. Em (b) os coeficientes cepstrais das amostras do indivíduo ID5 calçando tênis.

SELEÇÃO

Para a execução do processo de seleção de características, foi implementada a seleção de atributos usando o critério de Fisher (vide Seção 3.3). Tal critério foi aplicado a cada coluna dos atributos multi-coluna (densidade específica de sonoridade, envelope espectral, coeficientes cepstrais e mel cepstrais). Posteriormente, o conjunto (para cada coluna obtém-se um valor) dos valores obtidos pela aplicação do critério de seleção foi ordenado de forma decrescente, para que os “melhores”, segundo o critério de Fisher, apareçam no início do conjunto. Então, uma porcentagem P dos primeiros elementos do conjunto é escolhida, elegendo assim o conjunto representante dos dados. Em seguida, esse conjunto reduzido é submetido aos algoritmos de classificação, onde é avaliado segundo a quantidade de acertos.

O critério de seleção foi avaliado de quatro maneiras: sem seleção, e com seleção atribuindo os valores de 0.05, 0.1 e 0.2 para P . Quando determinada característica é submetida aos métodos de classificação sem passar pela seleção de atributos, o valor de P nas tabelas de resultados aparecerá como “-”; e para os demais casos, o próprio valor de P . Os resultados da seleção são mostrados na próxima seção.

5.2.3 IDENTIFICAÇÃO

Para avaliar os parâmetros considerados, os dois algoritmos reunidos na Seção 3.4, foram implementados para a tarefa de classificação. No método *K-NN*, a metodologia de partição de dados foi a *leave-one-out*, na qual retira-se um indivíduo para teste e designa-se os demais para servirem de base, repetindo esta ação até que todos os indivíduos da base tenham sido testados. Na implementação do *K-Means*, a metodologia de partição da base de dados foi a validação cruzada com K conjuntos, que consiste em dividir aleatoriamente o conjunto de dados K subconjuntos disjuntos, com aproximadamente o mesmo número de elementos, onde cada um dos K subconjuntos é utilizado como conjunto de teste e os demais para treinamento (GOLDSCHMIDT, 2005).

Em ambos os métodos classificatórios, cada subconjunto definido na Seção 5.2 foi utilizado como base de dados. Inicialmente, cada característica considerada na fase de extração de características será avaliada independentemente. Para os atributos multi-colunas será mostrado o resultado da classificação antes e após a aplicação do método de seleção de atributos.

No método *K-NN* foi utilizada uma faixa de valores para k variando de 1 a 20. E para o *k-means*, o número de *clusters* considerados foi o número de classes e os valores do parâmetro K da metodologia de partição da base de dados foram 5 e 10. Esse valor foi adotado devido ao número de amostras de cada classe, nos grupos eleitos, apresentar um valor máximo de 20 e mínimo de 10. Portanto, os valores de $K = 5$ e $K = 10$ escolhidos representam 20% e 40%, nas classes com 20 amostras, e 10% e 20%, nas classes com 10 amostras, respectivamente.

No intuito de designar o *K-Means* para classificação (uma vez que seu intuito é para a tarefa de agrupamento), a seguinte modificação foi efetuada:

- Inicialmente os indivíduos são identificados segundo segundo a classe;
- Cada conjunto formado pelos indivíduos de determinada classe é dividido em K subconjuntos, que serão testados um a um;
- Cada um dos conjuntos formados pelo passo anterior é submetido para teste, avaliando a distância de seus indivíduos com a média dos outros conjuntos;
- A distância mínima entre a média e o indivíduo determina rotula a classe do indivíduo do grupo de teste;

- Avalia se foi rotulado corretamente acumulando os acertos.

RESULTADOS DA CLASSIFICAÇÃO DE CADA CARACTERÍSTICA

As TAB. 5.2 e TAB. 5.3 reúnem os resultados dos algoritmos K-NN e K-Means, respectivamente, quando a variável em avaliação é a frequência média dos passos. Observando tais tabelas, nota-se uma diferença razoável na porcentagem de acertos. Essas tabelas confirmam a análise visual feita na seção anterior, que através do gráfico da média dos passos, não era possível distinguir os indivíduos. O melhor resultado (42.5%), utilizando somente a frequência média dos passos, foi obtido durante a avaliação do conjunto COTURNO, quando submetido ao K-NN.

K-NN - Frequência dos passos (maxvalueid)					
COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
K/ (%)	K/ (%)	K/ (%)	K/ (%)	K/ (%)	K/ (%)
1/42.5	1/36.3	1/36.7	1/30	1/30	1/19

TAB. 5.2: Resultado da classificação, avaliando a frequência dos passos através do K-NN aplicado a todos os conjuntos.

K-Means - Frequência dos passos (maxvalueid)					
COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
K/ (%)	K/ (%)	K/ (%)	K/ (%)	K/ (%)	K/ (%)
10/25	10/15	10/15	10/18.33	10/17.50	10/5.71
05/25	05/11.25	05/15	05/16.67	05/15.83	05/5.71

TAB. 5.3: Resultado da classificação, avaliando a frequência dos passos através do K-Means aplicado a todos os conjuntos.

Nas tabelas TAB. 5.4 e TAB. 5.5 estão reunidos os resultados dos algoritmos K-NN e K-Means, respectivamente, quando o parâmetro considerado é a sonoridade total. Utilizando a sonoridade total, o melhor resultado (60%) de classificação ocorreu durante a avaliação do conjunto UMPORCALCADO, utilizando o K-Means. Neste mesmo conjunto, o K-NN classificou corretamente 56.67% da base, quando $k = 1$. Mediante os resultados dos algoritmos avaliados, notou-se que a frequência média e a sonoridade total mostraram-se uma característica muito discriminante.

K-NN - Sonoridade Total (<i>Zwicker</i>)					
COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
K/ (%)	K/ (%)	K/ (%)	K/ (%)	K/ (%)	K/ (%)
1/30	1/28.75	1/25	1/56.67	1/30	1/21.9

TAB. 5.4: Resultado da classificação, avaliando a sonoridade total através do K-NN aplicado a todos os conjuntos.

K-Means - Sonoridade Total (<i>Zwicker</i>)					
COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
K/ (%)	K/ (%)	K/ (%)	K/ (%)	K/ (%)	K/ (%)
10/27.5	10/32.5	10/36.67	10/60	10/18.33	10/10
05/27.5	05/23.75	05/26.67	05/35	05/21.67	05/4.76

TAB. 5.5: Resultado da classificação, avaliando a frequência dos passos através do K-Means aplicado a todos os conjuntos.

As tabelas TAB. 5.6 e TAB. 5.7 reúnem os resultados dos algoritmos K-NN e K-Means, respectivamente, quando a variável em avaliação é a sonoridade específica. Observando a TAB. 5.6, a queda na qualidade de classificação é notável, quando aplicado o critério de Fisher para redução dos dados. Isso mostra que todos os pontos da sonoridade específica foram essenciais na classificação, segundo o K-NN. Além disso, nesta mesma tabela verifica-se a qualidade desse parâmetro; com exceção ao grupo SAPATO, para todos os demais grupos a classificação superou 50%, com o melhor resultado no grupo UMPORCALCADO com 81.67%, com $k = 1$.

Analisando a TAB. 5.7, percebe-se que a precisão dos acertos decaiu para todos os grupos (no grupo SAPATO quando $K=10$; e MESMOCALCADO quando $K=5$) e cresceu para os grupos SAPATO (quando $K=5$) e MESMOCALCADO (quando $K=10$), quando aplicado o critério de Fisher. Isto mostra a sensibilidade da sonoridade específica, na utilização do K-Means para classificação, mediante as diferenças dos tipos de calçados. Cabe ainda observar que o melhor resultado continuou no grupo UMPORCALCADO, quando relacionado ao algoritmo K-NN (com 81.67 %, na TAB. 5.6), com 75 % de acertos, quando $K = 10$. Este ocorrido mostra que a intensidade sonora subjetiva é mais saliente, diante da variação provocada pelos diferentes tipos de calçados.

K-NN - Sonoridade Específica						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)
-	5/62.5	1/65	1/46.67	1/81.67	1/54.17	1/54.76
p=0.05	3/42.5	1/36.25	1/21.67	1/40	1/33.33	1/23.81
p=0.1	3/42.5	1/42.5	1/30	3/41.67	1/38.33	1/29.52
p=0.2	1/57.5	1/50	1/33.33	3/48.33	1/38.33	1/33.33

TAB. 5.6: Resultado da classificação, avaliando a sonoridade específica através do K-Means aplicado a todos os conjuntos.

K-Means - Sonoridade Específica						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)
-	10/55	10/58.75	10/40	10/75	10/29.17	10/26.67
-	5/47.5	5/43.75	5/36.67	5/53.33	5/30.83	5/18.1
p=0.05	10/55	10/33.75	10/38.33	10/56.67	10/27.5	10/21.9
p=0.05	5/37.5	5/33.75	5/35	5/45	5/27.5	5/16.19
p=0.1	10/55	10/37.5	10/35	10/50	10/31.67	10/18.1
p=0.1	5/35	5/33.75	5/38.33	5/45	5/28.33	5/12.86
p=0.2	10/47.5	10/48.75	10/36.67	10/53.33	10/35	10/22.86
p=0.2	5/37.5	5/41.25	5/41.67	5/45	5/25	5/15.71

TAB. 5.7: Resultado da classificação, avaliando a sonoridade específica através do K-Means aplicado a todos os conjuntos.

Devido ao grande número de dados, os resultados dos coeficientes foram resumidos nas tabelas TAB. 5.8, TAB. 5.9 e TAB. 5.10. Os resultados completos se encontram no Anexo 8.1.

A tabela TAB. 5.8 abriga um resumo com os melhores resultados, ao aplicar o algoritmo K-NN nos 20 coeficientes mel-cepestrais. O parâmetro k não é exibido na tabela TAB. 5.8, pois em todos os melhores resultados k foi igual a 1. Para cada conjunto e coeficiente, são mostrados o parâmetro P indicando a proporção usada no processo de seleção e a porcentagem de acertos. Para o conjunto COTURNO, todos os coeficientes em geral obtiveram bons resultados. Os coeficientes 1 ao 3 foram os que apresentaram as maiores taxas de classificação para o grupo IGNORECALCADO. É importante ressaltar que para este grupo, a maioria das melhores taxas foram alcançadas sem passar pela seleção (“-”). Isso implica em duas observações, ou a quantidade P não foi escolhida adequadamente (o máximo de 20 % não foi suficiente), ou todas as 227 colunas foram realmente necessárias para alcançar essa taxa.

K-NN (MFCC) Resumo de resultados (k=1)						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
C	P/(%)	P/(%)	P/(%)	P/(%)	P/(%)	P/(%)
1	0.05/70.0	- /68.8	- /61.7	- /78.3	- /65.8	- /60.0
2	0.20/85.0	0.10/66.3	- /56.7	- /51.7	- /65.0	- /52.9
3	- /70.0	- /68.8	0.20/51.7	- /80.0	- /53.3	- /54.8
4	0.05/90.0	- /50.0	0.10/48.3	0.10/58.3	- /40.0	- /35.7
5	0.20/72.5	0.20/45.0	0.20/55.0	0.10/43.3	- /39.2	- /31.4
6	0.20/65.0	- /40.0	0.20/41.7	0.20/63.3	0.10/34.2	0.20/27.1
7	0.10/67.5	0.10/52.5	0.20/46.7	- /73.3	0.10/43.3	- /32.9
8	0.05/70.0	0.20/46.3	0.20/45.0	- /55.0	0.10/37.5	- /31.4
9	0.20/50.0	0.05/48.8	0.10/50.0	0.20/48.3	0.20/33.3	0.20/19.0
10	0.20/77.5	0.05/38.8	- /48.3	0.20/45.0	0.20/45.8	- /27.6
11	0.20/67.5	0.20/43.8	0.20/55.0	- /53.3	- /34.2	0.20/25.7
12	0.10/75.0	0.20/40.0	0.05/41.7	0.10/55.0	- /29.2	- /20.5
13	0.20/57.5	0.20/48.8	0.20/56.7	0.20/55.0	0.20/42.5	- /25.7
14	0.20/65.0	0.10/37.5	0.20/45.0	0.05/50.0	0.20/34.2	- /21.4
15	0.10/65.0	0.20/32.5	- /36.7	0.20/41.7	- /25.8	- /18.6
16	0.20/62.5	0.20/35.0	0.10/41.7	0.10/53.3	0.05/37.5	0.10/16.2
17	0.20/57.5	0.20/30.0	0.20/45.0	0.20/45.0	0.10/35.0	0.10/19.5
18	0.10/65.0	0.10/37.5	0.05/41.7	0.20/41.7	0.10/32.5	0.10/16.2
19	0.10/70.0	0.05/28.8	0.20/56.7	0.05/35.0	0.10/30.8	- /17.6
20	0.10/70.0	0.10/33.8	0.10/50.0	0.10/45.0	0.10/36.7	0.20/20.0

TAB. 5.8: Resumo dos resultados ao aplicar o K-NN aos coeficientes mel-cepestrais, para todos os conjuntos. A tabela reúne apenas os melhores resultados; e estes ocorreram quando k foi igual a 1.

As tabelas TAB. 5.9 e TAB. 5.10 agregam os melhores resultados de classificação utilizando o algoritmo K-Means com $K = 5$ e $K = 10$, respectivamente. Na tabela TAB. 5.9, o maior resultado de classificação entre os coeficientes (33.8 % quando C=8), para o grupo IGNORECALCADO ocorreu nos dados sem o critério de Fisher. Entretanto, ao observar os demais grupos, em ambas tabelas a maioria dos melhores resultados ocorreu com a aplicação do critério de Fisher.

K-Means - MFCC (Resumo de resultados K=5)						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
C	P/(%)	P/(%)	P/(%)	P/(%)	P/(%)	P/(%)
1	- /30.0	- /25.0	0.20/38.3	- /16.7	- /26.7	- /13.8
2	0.10/45.0	0.20/33.8	0.10/41.7	0.05/20.0	0.10/45.8	0.20/16.7
3	0.20/67.5	- /35.0	0.20/45.0	0.10/51.7	- /30.0	0.20/28.1
4	0.10/80.0	0.20/31.3	0.05/38.3	0.10/36.7	- /33.3	- /18.6
5	0.20/80.0	0.10/48.8	0.05/53.3	0.20/48.3	0.20/38.3	- /23.3
6	0.10/62.5	- /51.3	0.05/41.7	0.20/60.0	0.05/40.0	- /28.6
7	0.05/32.5	0.10/37.5	0.05/38.3	0.05/36.7	0.20/32.5	- /16.2
8	0.20/72.5	0.20/56.3	0.20/50.0	0.20/58.3	0.20/45.8	- /33.8
9	0.10/70.0	0.10/60.0	0.10/56.7	0.20/70.0	0.05/35.8	0.10/28.1
10	0.10/62.5	0.05/45.0	0.10/41.7	0.05/50.0	0.10/40.0	0.05/20.0
11	0.20/75.0	0.10/41.3	0.20/53.3	0.10/41.7	0.20/39.2	- /21.4
12	0.20/75.0	0.20/53.8	0.05/50.0	0.05/55.0	0.05/42.5	0.05/19.0
13	0.20/85.0	0.20/58.8	0.20/43.3	0.10/63.3	0.05/39.2	0.20/22.4
14	0.05/67.5	0.20/47.5	0.20/55.0	0.20/60.0	0.10/35.8	0.20/24.3
15	0.10/62.5	0.20/40.0	0.20/61.7	0.20/48.3	0.20/40.8	- /21.0
16	0.20/82.5	0.20/53.8	0.20/51.7	0.20/71.7	0.10/43.3	0.20/25.2
17	0.20/77.5	0.20/47.5	0.20/61.7	0.20/56.7	0.20/36.7	0.10/22.4
18	0.05/70.0	0.10/57.5	0.20/53.3	0.10/50.0	0.20/45.8	0.20/21.9
19	0.20/77.5	0.05/45.0	0.05/65.0	0.20/51.7	0.20/52.5	0.20/21.4
20	0.20/77.5	0.20/46.3	0.20/70.0	0.05/55.0	0.20/49.2	0.20/24.8

TAB. 5.9: Resumo com os melhores resultados ao aplicar o K-Means nos coeficientes mel-cepestrais, para todos os conjuntos, com $K = 5$.

K-Means - MFCC (Resumo de resultados K=10)						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
C	P/(%)	P/(%)	P/(%)	P/(%)	P/(%)	P/(%)
1	- /47.5	- /33.8	0.20/40.0	0.05/26.7	- /41.7	- /21.4
2	0.10/72.5	0.20/52.5	0.20/53.3	0.05/43.3	0.20/51.7	- /25.7
3	0.20/70.0	- /61.3	0.20/58.3	- /68.3	- /50.0	- /44.8
4	0.10/90.0	0.20/55.0	- /46.7	0.10/56.7	0.10/41.7	- /28.1
5	0.10/77.5	0.10/58.8	0.20/55.0	0.10/53.3	0.20/40.0	- /29.5
6	0.20/72.5	0.20/60.0	0.10/53.3	0.10/65.0	0.20/44.2	- /33.3
7	0.20/52.5	0.10/52.5	0.20/55.0	0.20/66.7	0.10/40.0	- /28.1
8	0.20/82.5	- /60.0	0.20/63.3	0.20/63.3	0.20/51.7	- /37.6
9	0.10/72.5	0.10/61.3	0.20/60.0	0.10/70.0	0.20/45.8	0.20/31.9
10	0.20/80.0	0.10/53.8	0.10/51.7	0.05/60.0	0.20/48.3	- /30.5
11	0.20/80.0	0.10/51.3	0.20/61.7	0.20/60.0	0.20/44.2	- /26.7
12	0.20/80.0	0.20/58.8	0.05/55.0	0.05/60.0	0.10/46.7	- /26.2
13	0.05/82.5	0.20/57.5	0.10/46.7	0.10/60.0	0.20/40.0	0.05/21.9
14	0.05/67.5	0.20/48.8	0.20/55.0	0.10/65.0	0.20/40.0	0.20/26.7
15	0.20/72.5	0.20/46.3	0.20/60.0	0.20/60.0	0.20/38.3	- /24.8
16	0.20/75.0	0.20/51.3	0.20/53.3	0.20/68.3	0.10/44.2	0.20/29.0
17	0.20/82.5	0.20/43.8	0.20/63.3	0.20/56.7	0.10/39.2	0.10/25.2
18	0.20/75.0	0.20/60.0	0.20/58.3	0.20/55.0	0.20/52.5	0.20/29.0
19	0.10/82.5	0.05/50.0	0.10/66.7	0.20/55.0	0.20/52.5	- /24.8
20	0.05/77.5	0.20/51.3	0.20/70.0	0.05/58.3	0.10/47.5	0.20/26.2

TAB. 5.10: Resumo com os melhores resultados ao aplicar o K-Means nos coeficientes mel-cepestrais, para todos os conjuntos, com $K = 10$.

K-NN - Envelope Espectral (Periodograma)						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)
-	1/62.50	1/31.25	1/38.33	1/71.67	1/40.83	1/32.86
0.05	1/57.50	1/33.75	1/36.67	1/51.67	3/34.17	1/19.52
0.10	1/62.50	1/43.75	1/46.67	1/61.67	1/37.50	1/24.76
0.20	1/75.00	1/47.50	1/41.67	1/73.33	1/35.83	1/30.95
K-NN - Envelope Espectral (Welch)						
-	1/55.00	1/43.75	1/58.33	1/91.67	1/50.00	1/46.19
0.05	1/65.00	1/36.25	1/48.33	1/81.67	1/31.67	1/30.00
0.10	1/70.00	1/36.25	1/46.67	1/80.00	1/33.33	1/27.62
0.20	1/67.50	1/35.00	1/46.67	1/81.67	1/32.50	1/25.71

TAB. 5.11: Resultados da aplicação do K-NN aos envelopes espectrais com extração baseada nos métodos: Periodograma e Welch.

A TAB. 5.11 congrega os resultados do algoritmo K-NN aplicado aos envelopes espec-

trais obtidos pelos métodos: Periodograma e Welch. Para o grupo UMPORCALCADO o envelope espectral usando Welch, atingiu um valor máximo de 91.67 %, se comportando como uma característica discriminante mediante a variedade de calçados. Enquanto que utilizando o envelope obtido pelo periodograma, o K-NN obteve uma acurácia razoável somente no grupo COTURNO (75 % com $P = 0.2$). Observando os resultados para o método Periodograma, nota-se que o critério de seleção contribuiu no resultado para os grupos: COTURNO (75 % quando $P = 0.2$ e $K = 1$), TENIS (47.5 % quando $P = 0.2$ e $K = 1$), SAPATO (46.67 % quando $P = 0.1$ e $K = 1$) e UMPORCALCADO (73.33 %).

Os resultados do K-Means para os envelopes espectrais baseados nos métodos Periodograma e Welch estão reunidos na TAB. 5.12. Ao observar os resultados para o método de Welch, nota-se que eles permaneceram praticamente inalteráveis, independente do valor de K e P . Em geral, os resultados foram muito baixos quando considerou-se somente o envelope espectral.

K-Means - Envelope Espectral (Periodograma)						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)
-	5/27.50	5/12.50	5/21.67	5/16.67	5/16.67	5/4.76
0.05	5/25.00	5/16.25	5/16.67	5/16.67	5/16.67	5/4.76
0.10	5/25.00	5/13.75	5/18.33	5/16.67	5/16.67	5/4.76
0.20	5/25.00	5/16.25	5/16.67	5/16.67	5/18.33	5/4.76
-	10/32.50	10/16.25	10/21.67	10/21.67	10/20.83	10/7.14
0.05	10/25.00	10/15.00	10/11.67	10/21.67	10/19.17	10/8.57
0.10	10/25.00	10/17.50	10/18.33	10/21.67	10/20.00	10/8.10
0.20	10/25.00	10/22.50	10/21.67	10/21.67	10/19.17	10/9.05
K-Means - Envelope Espectral (Welch)						
-	5/25.00	5/12.50	5/16.67	5/16.67	5/16.67	5/5.24
0.05	5/25.00	5/12.50	5/16.67	5/16.67	5/16.67	5/4.76
0.10	5/25.00	5/12.50	5/16.67	5/16.67	5/16.67	5/4.76
0.20	5/32.50	5/12.50	5/16.67	5/16.67	5/16.67	5/4.76
-	10/25.00	10/12.50	10/16.67	10/21.67	10/16.67	10/7.62
0.05	10/25.00	10/12.50	10/16.67	10/16.67	10/16.67	10/4.76
0.10	10/25.00	10/12.50	10/16.67	10/16.67	10/16.67	10/4.76
0.20	10/32.50	10/12.50	10/16.67	10/16.67	10/16.67	10/4.76

TAB. 5.12: Resultados da aplicação do K-Means aos envelopes espectrais com extração baseada nos métodos: Periodograma e Welch.

Os resultados da aplicação do K-NN e K-Means aos coeficientes cepstrais são mostrados em TAB. 5.13 e TAB. 5.14, respectivamente. Observando a TAB. 5.13 intui-se que o

processo de seleção colaborou para o aumento dos acertos nos subgrupos COTURNO, TENIS, SAPATO E IGNORECALCADO. O processo de seleção também se mostrou eficaz nos resultados do algoritmo K-Means (TAB. 5.14), em todos os subgrupos, aumentando a acurácia da classificação.

K-NN - Coeficientes Cepstrais (LPC)						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)
-	1/67.50	1/58.75	1/48.33	1/75.00	1/54.17	1/53.33
0.05	3/80.00	3/53.75	1/50.00	1/55.00	1/35.83	1/29.52
0.10	1/77.50	1/57.50	1/58.33	1/68.33	1/40.83	1/45.24
0.20	1/77.50	1/68.75	1/50.00	1/71.67	1/50.00	1/58.57

TAB. 5.13: Resultados da aplicação do K-NN aos coeficientes cepstrais baseados no LPC.

K-Means - Coeficientes Cepstrais (LPC)						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)
-	5/45.00	5/23.75	5/23.33	5/18.33	5/25.83	5/18.10
0.05	5/47.50	5/27.50	5/25.00	5/45.00	5/30.83	5/18.10
0.10	5/57.50	5/31.25	5/20.00	5/46.67	5/34.17	5/17.14
0.20	5/60.00	5/38.75	5/21.67	5/55.00	5/30.83	5/16.67
-	10/55.00	10/36.25	10/38.33	10/40.00	10/30.00	10/27.14
0.05	10/72.50	10/43.75	10/43.33	10/60.00	10/27.50	10/30.48
0.10	10/67.50	10/43.75	10/33.33	10/66.67	10/36.67	10/30.48
0.20	10/70.00	10/50.00	10/33.33	10/75.00	10/30.83	10/29.52

TAB. 5.14: Resultados da aplicação do K-Means aos coeficientes cepstrais baseados no LPC.

Por fim, os melhores resultados de todas as características, sob as condições de teste descritas no início dessa seção, são reunidos em TAB. 5.15, para o algoritmo K-NN, e TAB. 5.16, para o algoritmo K-Means. Nota-se que os coeficientes mel-cepstrais não constam nas tabelas citadas, pois um resumo de seus resultados foram reunidos nas tabelas TAB. 5.8, TAB. 5.9 e TAB. 5.10.

K-NN - Resumo com todas as características												
C	COTURNO		TENIS		SAPATO		UMPORCALCADO		MESMOCALCADO		IGNORECALCADO	
	P	k/%	P	k/%	P	k/%	P	k/%	P	k/%	P	k/%
FP	-	1/42.5	-	1/36.3	-	1/36.7	-	1/30	-	1/30	-	1/19
S.T	-	1/30	-	1/28.75	-	1/25	-	1/56.67	-	1/30	-	1/21.9
S.E	-	5/62.5	-	1/65	-	1/46.67	-	1/81.67	-	1/54.17	-	1/54.76
EP	0.2	1/75	0.2	1/47.5	0.1	1/46.67	0.2	1/73.33	-	1/40.83	-	1/32.86
EW	0.1	1/70	-	1/43.75	-	1/58.33	-	1/91.67	-	1/50	-	1/46.19
CC	0.05	3/80	0.2	1/68.75	0.1	1/58.33	-	1/75	-	1/54.17	0.2	1/58.57

TAB. 5.15: Resumo com os melhores resultados ao se aplicar o KNN em todas as características e grupos. A primeira coluna mostra a característica, as siglas significam: FP - Frequência dos passos, ST - Sonoridade Total, SE - Sonoridade Específica, EP - Envelope espectral usando Periodograma, EW - Envelope espectral usando o método de Welch e CC - Coeficientes Cepstrais.

K-Means - Resumo com todas as características													
C	K	COTURNO		TENIS		SAPATO		UMPORCALCADO		MESMOCALCADO		IGNORECALCADO	
		P	%	P	%	P	%	P	%	P	%	P	%
FP	5	-	25	-	11.2	-	15	-	16.7	-	15.8	-	5.7
ST	5	-	27.5	-	23.7	-	26.7	-	35	-	21.7	-	4.8
SE	5	-	47.5	-	43.7	0.2	41.7	-	53.3	-	30.8	-	18.1
EP	5	-	27.5	0.05	16.2	-	21.7	0.05	16.7	0.2	18.3	0.05	4.8
EW	5	0.2	32.5	-	12.5	-	16.7	-	16.7	0.05	16.7	-	5.2
CC	5	0.2	60	0.2	38.7	0.05	25	0.2	55	0.1	34.2	0.05	18.1
FP	10	-	25	-	15	-	15	-	18.3	-	17.5	-	5.7
ST	10	-	27.5	-	32.5	-	36.7	-	60	-	18.33	-	10
SE	10	-	55	-	58.7	-	40	-	75	0.2	35	-	26.7
EP	10	-	32.5	0.2	22.5	0.2	16.7	0.05	16.7	-	16.7	0.2	5.2
EW	10	0.2	32.5	0.05	12.5	0.05	16.7	-	21.7	0.05	16.7	-	7.6
CC	10	0.05	72.5	0.2	50	0.05	43.3	0.2	75	0.1	36.7	0.05	30.4

TAB. 5.16: Resumo com os melhores resultados ao se aplicar o K-Means em todas as características e grupos. A primeira coluna mostra a característica, as siglas significam: FP - Frequência dos passos, ST - Sonoridade Total, SE - Sonoridade Específica, EP - Envelope espectral usando Periodograma, EW - Envelope espectral usando o método de Welch e CC - Coeficientes Cepstrais.

RESULTADOS DA COMBINAÇÃO DE CARACTERÍSTICAS

Nesta seção são mostrados os resultados dos algoritmos de classificação avaliados nesse projeto, ao ter como base de dados combinações de características. A metodologia consistiu em combinar as características e porcentagem utilizada, na seleção dos melhores resultados em cada conjunto. Devido às diferenças de domínio, antes da fusão dos dados para o envio ao algoritmo classificador, cada parâmetro é normalizado linearmente. As características são representadas por símbolos, para facilitar a organização dos resultados em tabelas. No teste de cada grupo, será mostrada uma tabela com as características participantes das combinações. O símbolo de “+” foi utilizado para representar a fusão dos dados de cada característica. A seguir, será mostrado os resultados aplicados a cada

conjunto.

COTURNO

Seis características, reunidas na TAB. 5.17, foram selecionadas para avaliar o K-NN. Devido ao grande número de combinações, são mostrados na TAB. 5.18 somente os resultados maiores que 80 %.

Característica	Símbolo Representativo	P	% Acertos
MFCC 4	γ_4	0.05	90
MFCC 2	γ_2	0.2	85
Coef. Cepstrais	φ	0.05	80
Env. Periodograma	ϖ_p	0.2	75
Env. Welch	ϖ_w	0.1	70
Sonoridade Específica	λ_e	-	62.2

TAB. 5.17: Características combinadas para o conjunto COTURNO, avaliando o K-NN.

Combinação	% Acertos
$\gamma_4 + \gamma_2$	97.5
$\gamma_4 + \gamma_2 + \varphi$	97.5
$\gamma_4 + \gamma_2 + \varphi + \lambda_e$	92.5
$\gamma_2 + \varphi$	92.5
$\gamma_2 + \varphi$	92.5
$\gamma_2 + \lambda_e$	85
$\gamma_4 + \lambda_e$	82.5
$\varpi_p + \varpi_w$	80
$\gamma_4 + \gamma_2 + \varphi + \lambda_e + \varpi_p + \varpi_w$	80
$\gamma_4 + \gamma_2 + \varphi + \lambda_e + \varpi_p + \varpi_w + \lambda_e$	80

TAB. 5.18: Resultado do algoritmo K-NN após as combinações para o conjunto COTURNO.

A TAB. 5.19 mostra as características eleitas para avaliação do K-Means, com $K = 5$. Os resultados das combinações com os melhores resultados se encontram na TAB. 5.20.

Característica	Símbolo Representativo	P	% Acertos
MFCC 13	γ_{13}	0.2	85
MFCC 16	γ_{16}	0.2	82.5
MFCC 4	γ_4	0.1	80
MFCC 5	γ_5	0.2	80
Coef. Cepstrais	φ	0.2	60
Sonoridade Específica	λ_e	-	47

TAB. 5.19: Características combinadas para o conjunto COTURNO, avaliando o K-Means com $K = 5$.

Combinação	% Acertos
$\gamma_{13} + \gamma_4 + \varphi$	97.50
$\gamma_{13} + \gamma_{16}$	95.00
$\gamma_{13} + \gamma_4$	95.00
$\gamma_{13} + \gamma_{16} + \gamma_4$	95.00
$\gamma_{13} + \gamma_{16} + \gamma_4 + \gamma_5 + \varphi$	95.00
$\gamma_{13} + \gamma_{16} + \gamma_4 + \gamma_5 + \varphi + \lambda_e$	95.00
$\gamma_{13} + \gamma_{16} + \gamma_4 + \gamma_5$	92.50
$\gamma_{16} + \gamma_4 + \gamma_5 + \varphi$	92.50
$\gamma_{16} + \gamma_4$	92.50
$\gamma_{13} + \gamma_5$	82.50
$\gamma_{16} + \gamma_5$	87.50

TAB. 5.20: Resultado do algoritmo K-Means com $K = 5$, após as combinações para o conjunto COTURNO.

A TAB. 5.21 mostra as características eleitas para avaliação do K-Means, com $K = 10$. Os resultados das combinações com os melhores resultados se encontram na TAB. 5.22.

Característica	Símbolo Representativo	P	% Acertos
MFCC 4	γ_4	0.1	90
MFCC 13	γ_{13}	0.05	82.5
MFCC 17	γ_{17}	0.2	82.5
MFCC 19	γ_{19}	0.1	82.5
Coef. Cepstrais	φ	0.05	72.5
Sonoridade Específica	λ_e	0.05	55

TAB. 5.21: Características combinadas para o conjunto COTURNO, avaliando o K-Means com $K = 10$.

Combinação	% Acertos
$\gamma_4 + \gamma_{13} + \gamma_{17} + \gamma_{19}$	95.00
$\gamma_4 + \gamma_{13} + \gamma_{17} + \gamma_{19} + \gamma_8$	95.00
$\gamma_4 + \gamma_{13} + \gamma_{17} + \gamma_{19} + \gamma_8 + \varphi$	95.00
$\gamma_4 + \gamma_{13}$	92.50
$\gamma_4 + \gamma_{17}$	92.50
$\gamma_4 + \gamma_{13} + \gamma_{17}$	92.50
$\gamma_{17} + \gamma_{19} + \gamma_8 + \varphi$	90.00
$\gamma_4 + \gamma_{13} + \gamma_{17} + \gamma_{19} + \gamma_8 + \varphi + \lambda_e$	90.00
$\gamma_{13} + \gamma_{19} + \gamma_8$	87.50
$\gamma_{13} + \gamma_{17}$	85.00
$\gamma_8 + \varphi + \lambda_e$	85.00
$\gamma_{19} + \gamma_8 + \varphi$	85.00

TAB. 5.22: Resultado do algoritmo K-Means com $K = 10$, após as combinações para o conjunto COTURNO.

TENIS

As características a serem combinadas e submetidas ao teste com o algoritmo K-NN para o grupo TENIS, estão reunidas na TAB. 5.23. Devido ao grande número de combinações, são mostrados na TAB. 5.24 somente os resultados maiores que 60 %.

Característica	Símbolo Representativo	P	% Acertos
MFCC 1	γ_1	-	68.8
MFCC 2	γ_2	0.1	66.3
MFCC 3	γ_3	-	68.8
Coef. Cepstrais	φ	0.2	68.75
Sonoridade Específica	λ_e	-	65

TAB. 5.23: Características combinadas para o conjunto TENIS, avaliando o K-NN.

Combinação	% Acertos
$\gamma_1 + \gamma_2 + \gamma_3$	86.25
$\gamma_1 + \gamma_2 + \gamma_3 + \varphi$	86.25
$\gamma_1 + \gamma_2 + \gamma_3 + \varphi + \lambda_e$	86.25
$\gamma_1 + \gamma_3$	85
$\gamma_2 + \gamma_3$	75
$\gamma_2 + \lambda_e$	71.25
$\gamma_2 + \lambda_e$	71.25

TAB. 5.24: Resultado do algoritmo K-NN após as combinações para o conjunto TENIS.

A TAB. 5.25 mostra as características eleitas para avaliação do K-Means, com $K = 5$. Os resultados das combinações com os melhores resultados se encontram na TAB. 5.26.

Característica	Símbolo Representativo	P	% Acertos
MFCC 9	γ_9	0.1	60
MFCC 13	γ_{13}	0.2	58.8
MFCC 18	γ_{18}	0.1	57.5
MFCC 8	γ_8	0.2	56.
Sonoridade Específica	λ_e	-	43.7
Coef. Cepstrais	φ	0.2	38.7

TAB. 5.25: Características combinadas para o conjunto TENIS, avaliando o K-Means com $K = 5$.

Combinação	% Acertos
$\gamma_9 + \gamma_{13} + \gamma_{18}$	81.25
$\gamma_9 + \gamma_{13} + \gamma_8 + \gamma_{18} + \lambda_e$	80.00
$\gamma_9 + \gamma_{13} + \gamma_8 + \gamma_{18} + \varphi + \lambda_e$	80.00
$\gamma_9 + \gamma_{13} + \lambda_e$	76.25
$\gamma_9 + \gamma_{13} + \gamma_8 + \lambda_e$	77.50
$\gamma_9 + \gamma_{18}$	75.00
$\gamma_9 + \gamma_{13} + \gamma_8 + \gamma_{18}$	72.50
$\gamma_9 + \gamma_{13} + \gamma_8 + \gamma_{18} + \varphi$	72.50
$\gamma_9 + \gamma_{13}$	71.25
$\gamma_9 + \gamma_{13} + \varphi$	71.25
$\gamma_9 + \gamma_{13} + \gamma_8$	70.00

TAB. 5.26: Resultado do algoritmo K-Means com $K = 5$, após as combinações para o conjunto TENIS.

A TAB. 5.27 mostra as características eleitas para avaliação do K-Means, com $K = 10$. Os resultados das combinações com os melhores resultados se encontram na TAB. 5.28.

Característica	Símbolo Representativo	P	% Acertos
MFCC 9	γ_9	0.1	60
MFCC 13	γ_{13}	0.2	58.8
MFCC 18	γ_{18}	0.1	57.5
MFCC 8	γ_8	0.2	56.
Sonoridade Específica	λ_e	-	43.7
Coef. Cepstrais	φ	0.2	38.7

TAB. 5.27: Características combinadas para o conjunto TENIS, avaliando o K-Means com $K = 10$.

Combinação	% Acertos
$\gamma_3 + \gamma_9 + \gamma_{18} + \gamma_8 + \gamma_6 + \lambda_e$	75.00
$\gamma_3 + \gamma_9 + \gamma_{18} + \gamma_8 + \gamma_6 + \lambda_e + \varphi$	75.00
$\gamma_3 + \gamma_8 + \gamma_6 + \lambda_e + \varphi$	75.00
$\gamma_8 + \gamma_6 + \lambda_e + \varphi$	75.00
$\gamma_3 + \gamma_9 + \gamma_{18} + \gamma_8 + \gamma_6$	73.75
$\gamma_3 + \gamma_6 + \lambda_e$	72.50
$\gamma_3 + \gamma_9 + \gamma_6$	72.50
$\gamma_3 + \gamma_9 + \gamma_{18} + \gamma_6$	72.50
$\gamma_3 + \gamma_6$	71.25

TAB. 5.28: Resultado do algoritmo K-Means com $K = 10$, após as combinações para o conjunto TENIS.

SAPATO

As características a serem combinadas e submetidas ao teste com o algoritmo K-NN para o grupo SAPATO, estão reunidas na TAB. 5.29. Devido ao grande número de combinações, são mostrados na TAB. 5.30 somente os resultados maiores que 60 %.

Característica	Símbolo Representativo	P	% Acertos
MFCC 1	γ_1	-	61.7
MFCC 13	γ_{13}	0.2	56.7
MFCC 19	γ_{19}	0.2	56.7
Coef. Cepstrais	φ	0.2	58.33
Env. Welch	ϖ_w	-	58.33
Sonoridade Específica	λ_e	-	46.67

TAB. 5.29: Características combinadas para o conjunto SAPATO, avaliando o K-NN.

Combinação	% Acertos
$\gamma_1 + \gamma_{13} + \gamma_{19}$	70
$\gamma_1 + \gamma_{13}$	68.33
$\gamma_1 + \lambda_e$	68.33
$\gamma_1 + \gamma_{19}$	65
$\gamma_{10} + \lambda_e$	65
$\gamma_{13} + \lambda_e$	63.33

TAB. 5.30: Resultado do algoritmo K-NN após as combinações para o conjunto SAPATO.

A TAB. 5.31 mostra as características eleitas para avaliação do K-Means, com $K = 5$. Os resultados das combinações com os melhores resultados se encontram na TAB. 5.32.

Característica	Símbolo Representativo	P	% Acertos
MFCC 20	γ_{20}	0.2	70
MFCC 19	γ_{19}	0.05	65
MFCC 17	γ_{17}	0.2	61.7
MFCC 15	γ_{15}	0.2	61.7
Sonoridade Específica	λ_e	0.2	41.7
Sonoridade Total	λ_t	-	26.7

TAB. 5.31: Características combinadas para o conjunto SAPATO, avaliando o K-Means com $K = 5$.

Combinação	% Acertos
$\gamma_{20} + \gamma_{19} + \gamma_{17} + \gamma_{15}$	88.33
$\gamma_{20} + \gamma_{19} + \gamma_{17} + \gamma_{15} + \lambda_t$	88.33
$\gamma_{20} + \gamma_{19} + \gamma_{17}$	85.00
$\gamma_{20} + \gamma_{19} + \gamma_{17} + \lambda_t$	85.00
$\gamma_{20} + \gamma_{19} + \gamma_{15}$	83.33
$\gamma_{19} + \gamma_{17}$	81.67
$\gamma_{20} + \gamma_{19} + \gamma_{17} + \gamma_{15} + \lambda_e$	81.67
$\gamma_{20} + \gamma_{19} + \gamma_{17} + \gamma_{15} + \lambda_e + \lambda_t$	81.67
$\gamma_{20} + \gamma_{17}$	78.33
$\gamma_{20} + \gamma_{19} + \gamma_{17} + \lambda_e$	78.33
$\gamma_{20} + \gamma_{15}$	73.33
$\gamma_{17} + \gamma_{15}$	73.33
$\gamma_{20} + \gamma_{19}$	71.67
$\gamma_{20} + \lambda_t$	71.67
$\gamma_{20} + \gamma_{19} + \lambda_t$	71.67

TAB. 5.32: Resultado do algoritmo K-Means com $K = 5$, após as combinações para o conjunto SAPATO.

A TAB. 5.33 mostra as características eleitas para avaliação do K-Means, com $K = 10$. Os resultados das combinações com os melhores resultados se encontram na TAB. 5.34.

Característica	Símbolo Representativo	P	% Acertos
MFCC 20	γ_{20}	0.2	70
MFCC 19	γ_{19}	0.1	66.7
MFCC 8	γ_8	0.2	63.3
MFCC 17	γ_{17}	0.2	63.3
MFCC 11	γ_{11}	0.2	61.7
Coef. Cepstrais	φ	0.05	43.3
Sonoridade Específica	λ_e	-	40
Sonoridade Total	λ_t	-	36.7

TAB. 5.33: Características combinadas para o conjunto SAPATO, avaliando o K-Means com $K = 10$.

Combinação	% Acertos
$\gamma_8 + \gamma_{11} + \gamma_{17} + \gamma_{20}$	90.00
$\gamma_8 + \gamma_{11} + \gamma_{17} + \gamma_{19} + \gamma_{20}$	90.00
$\gamma_8 + \gamma_{11} + \gamma_{17} + \gamma_{19} + \gamma_{20} + \lambda_e$	90.00
$\gamma_8 + \gamma_{11} + \gamma_{20}$	85.00
$\gamma_8 + \gamma_{11} + \gamma_{17} + \gamma_{19} + \gamma_{20} + \lambda_t$	90.00
$\gamma_{11} + \gamma_{17} + \gamma_{20}$	86.67
$\gamma_8 + \gamma_{11} + \gamma_{17} + \gamma_{19} + \lambda_e$	85.00
$\gamma_8 + \gamma_{11} + \gamma_{17} + \gamma_{19}$	85.00
$\gamma_8 + \gamma_{11} + \gamma_{17} + \gamma_{19} + \lambda_t$	85.00
$\gamma_8 + \gamma_{11} + \gamma_{17} + \gamma_{19} + \varphi$	85.00
$\gamma_8 + \gamma_{11} + \gamma_{17}$	83.33
$\gamma_8 + \gamma_{11} + \gamma_{17} + \lambda_e$	83.33
$\gamma_8 + \gamma_{11} + \gamma_{17} + \lambda_t$	83.33
$\gamma_8 + \gamma_{11} + \gamma_{17} + \varphi$	83.33
$\gamma_{19} + \gamma_{20}$	83.33
$\gamma_8 + \gamma_{11} + \lambda_e$	81.67
$\gamma_8 + \gamma_{11}$	80.00
$\gamma_8 + \gamma_{11} + \lambda_t$	80.00
$\gamma_8 + \gamma_{11} + \varphi$	80.00
$\gamma_{17} + \gamma_{20}$	80.00
$\gamma_8 + \gamma_{20}$	78.33
$\gamma_{17} + \gamma_{19}$	78.33
$\gamma_{11} + \gamma_{20}$	76.67
$\gamma_8 + \gamma_{11} + \gamma_{19}$	76.67
$\gamma_8 + \gamma_{19}$	75.00

TAB. 5.34: Resultado do algoritmo K-Means com $K = 10$, após as combinações para o conjunto SAPATO.

UMPORCALCADO

As características a serem combinadas e submetidas ao teste com o algoritmo K-NN para o grupo UMPORCALCADO, estão reunidas na TAB. 5.35. Devido ao número de combinações, são mostrados na TAB. 5.36 somente os resultados maiores que 85 %.

Característica	Símbolo Representativo	P	% Acertos
Env. Welch	ϖ_w	-	91.66
Sonoridade Específica	λ_e	-	81.66
MFCC 3	γ_3	-	80
MFCC 1	γ_1	-	78.3
Coef. Cepstrais	φ	-	75
MFCC 7	γ_{19}	-	73.3
Env. Periodograma	ϖ_p	0.2	73.33

TAB. 5.35: Características combinadas para o conjunto UMPORCALCADO, avaliando o K-NN.

Combinação	% Acertos
$\varpi_w + \lambda_e$	91.67
$\varpi_w + \varpi_e$	90.00
$\gamma_3 + \gamma_1 + \gamma_7$	90.00
$\lambda_e + \gamma_1$	88.33
$\lambda_e + \gamma_7$	85.00

TAB. 5.36: Resultado do algoritmo K-NN após as combinações para o conjunto UMPORCALCADO.

A TAB. 5.37 mostra as características eleitas para avaliação do K-Means, com $K = 5$. Os resultados das combinações com os melhores resultados se encontram na TAB. 5.38.

Característica	Símbolo Representativo	P	% Acertos
MFCC 6	γ_6	0.2	60
MFCC 9	γ_9	0.2	70
MFCC 13	γ_{13}	0.1	63.3
MFCC 14	γ_{14}	0.2	60
MFCC 16	γ_{16}	0.2	71.7
Sonoridade Específica	λ_e	-	53.3
Coef. Cepstrais	φ	0.2	55

TAB. 5.37: Características combinadas para o conjunto UMPORCALCADO, avaliando o K-Means com $K = 5$.

Combinação	% Acertos
$\gamma_6 + \gamma_9 + \gamma_{13} + \gamma_{14}$	88.33
$\gamma_6 + \gamma_9 + \gamma_{13} + \gamma_{14} + \gamma_{16}$	88.33
$\gamma_6 + \gamma_9 + \gamma_{13} + \gamma_{16}$	88.33
$\gamma_6 + \gamma_9 + \gamma_{13} + \gamma_{14} + \gamma_{16} + \lambda_e$	88.33
$\gamma_9 + \gamma_{16}$	86.67
$\gamma_6 + \gamma_9 + \gamma_{13} + \gamma_{14} + \gamma_{16} + \varphi$	88.33
$\gamma_6 + \gamma_9 + \gamma_{13}$	83.33
$\gamma_6 + \gamma_9 + \gamma_{14}$	83.33
$\gamma_6 + \gamma_9 + \gamma_{16}$	83.33
$\gamma_6 + \gamma_9 + \gamma_{13} + \varphi$	83.33
$\gamma_6 + \gamma_9$	81.67
$\gamma_9 + \gamma_{14}$	81.67
$\gamma_6 + \gamma_9 + \gamma_{13} + \lambda_e$	81.67
$\gamma_{14} + \lambda_e$	81.67
$\gamma_{16} + \lambda_e$	81.67
$\gamma_6 + \gamma_9 + \varphi$	81.67
$\gamma_9 + \gamma_{13}$	80.00
$\gamma_{13} + \gamma_{16}$	80.00

TAB. 5.38: Resultado do algoritmo K-Means com $K = 5$, após as combinações para o conjunto UMPORCALCADO.

A TAB. 5.39 mostra as características eleitas para avaliação do K-Means, com $K = 10$. Os resultados das combinações com os melhores resultados se encontram na TAB. 5.40.

Característica	Símbolo Representativo	P	% Acertos
MFCC 3	γ_{33}	-	68.3
MFCC 7	γ_7	0.2	66.7
MFCC 9	γ_9	0.1	70
MFCC 14	γ_{14}	0.1	65
MFCC 16	γ_{16}	0.2	68.3
Sonoridade Específica	λ_e	-	75
Coef. Cepstrais	φ	0.2	75

TAB. 5.39: Características combinadas para o conjunto UMPORCALCADO, avaliando o K-Means com $K = 10$.

Combinação	% Acertos
$\gamma_9 + \lambda_e$	85.00
$\gamma_{16} + \lambda_e$	85.00
$\gamma_7 + \gamma_{14}$	83.33
$\gamma_9 + \gamma_{16}$	83.33
$\gamma_7 + \gamma_9$	80.00
$\gamma_9 + \gamma_{14}$	80.00
$\gamma_7 + \gamma_{16}$	78.33
$\gamma_7 + \lambda_e$	78.33
$\gamma_{14} + \lambda_e$	78.33
$\gamma_{14} + \gamma_{16}$	75.00
$\gamma_3 + \gamma_7 + \gamma_9 + \gamma_{14} + \gamma_{16}$	75.00

TAB. 5.40: Resultado do algoritmo K-Means com $K = 10$, após as combinações para o conjunto UMPORCALCADO.

MESMOCALCADO

As características a serem combinadas e submetidas ao teste com o algoritmo K-NN para o grupo MESMOCALCADO, estão reunidas na TAB. 5.41. Devido ao número de combinações, são mostrados na TAB. 5.42 somente os resultados maiores que 65 %.

Característica	Símbolo Representativo	P	% Acertos
MFCC 1	γ_1	-	65.8
MFCC 2	γ_2	-	65
Sonoridade Específica	λ_e	-	54.17
Coef. Cepstrais	φ	-	54.17
MFCC 3	γ_3	-	53.3
Env. Welch	ϖ_w	-	50

TAB. 5.41: Características combinadas para o conjunto MESMOCALCADO, avaliando o K-NN.

Combinação	% Acertos
$\gamma_1 + \gamma_2 + \lambda_e$	73.33
$\gamma_1 + \gamma_2 + \lambda_e + \varphi$	73.33
$\gamma_1 + \lambda_e$	70.83
$\gamma_1 + \gamma_2$	66.67
$\gamma_1 + \gamma_2 + \varphi$	66.67

TAB. 5.42: Resultado do algoritmo K-NN após as combinações para o conjunto MESMOCALCADO.

A TAB. 5.43 mostra as características eleitas para avaliação do K-Means, com $K = 5$.

Os resultados das combinações com os melhores resultados se encontram na TAB. 5.44.

Característica	Símbolo Representativo	P	% Acertos
MFCC 19	γ_{19}	0.2	52.5
MFCC 20	γ_{20}	0.2	49.2
MFCC 2	γ_2	0.1	45.8
MFCC 8	γ_8	0.2	45.8
MFCC 18	γ_{18}	0.2	45.8
Coef. Cepstrais	φ	0.1	34.2
Sonoridade Específica	λ_e	-	30.8

TAB. 5.43: Características combinadas para o conjunto MESMOCALCADO, avaliando o K-Means com $K = 5$.

Combinação	% Acertos
$\gamma_2 + \gamma_8 + \gamma_{18} + \gamma_{19} + \gamma_{20} + \lambda_e$	75.83
$\gamma_2 + \gamma_8 + \gamma_{18} + \gamma_{19} + \gamma_{20} + \lambda_e + \varphi$	75.83
$\gamma_2 + \gamma_8 + \gamma_{18} + \gamma_{19} + \gamma_{20} + \varphi$	75.00
$\gamma_2 + \gamma_8 + \gamma_{18} + \gamma_{19} + \gamma_{20}$	75.00
$\gamma_2 + \gamma_8 + \gamma_{18} + \gamma_{19} + \lambda_e$	73.33
$\gamma_2 + \gamma_8 + \gamma_{18} + \gamma_{19}$	71.67
$\gamma_2 + \gamma_8 + \gamma_{18} + \gamma_{19} + \varphi$	71.67
$\gamma_2 + \gamma_8 + \gamma_{18} + \gamma_{20}$	68.33
$\gamma_2 + \gamma_8 + \gamma_{18} + \lambda_e$	68.33
$\gamma_2 + \gamma_8 + \gamma_{18} + \varphi$	68.33
$\gamma_2 + \gamma_8 + \gamma_{19}$	67.50
$\gamma_2 + \gamma_8 + \gamma_{20}$	65.83
$\gamma_2 + \gamma_8 + \lambda_e$	64.17
$\gamma_2 + \gamma_8 + \varphi$	64.17
$\gamma_2 + \gamma_8$	64.17
$\gamma_2 + \gamma_8 + \lambda_e + \varphi$	64.17

TAB. 5.44: Resultado do algoritmo K-Means com $K = 5$, após as combinações para o conjunto MESMOCALCADO.

A TAB. 5.45 mostra as características eleitas para avaliação do K-Means, com $K = 10$. Os resultados das combinações com os melhores resultados se encontram na TAB. 5.46.

Característica	Símbolo Representativo	P	% Acertos
MFCC 2	γ_2	0.2	51.7
MFCC 3	γ_3	-	50
MFCC 8	γ_8	0.2	51.7
MFCC 18	γ_{18}	0.2	52.5
MFCC 19	γ_{19}	0.2	52.5
Sonoridade Específica	λ_e	0.2	35
Coef. Cepstrais	φ	0.1	36.7

TAB. 5.45: Características combinadas para o conjunto MESMOCALCADO, avaliando o K-Means com $K = 10$.

Combinação	% Acertos
$\gamma_2 + \gamma_8$	69.17
$\gamma_2 + \gamma_{19}$	67.50
$\gamma_2 + \gamma_3 + \gamma_8 + \gamma_{18} + \gamma_{19}$	66.67
$\gamma_2 + \gamma_3 + \gamma_8 + \gamma_{18}$	65.83
$\gamma_2 + \gamma_3 + \gamma_8 + \gamma_{18} + \lambda_e + \varphi$	65.83
$\gamma_2 + \gamma_3 + \gamma_8 + \gamma_{19}$	65.00
$\gamma_2 + \gamma_3 + \gamma_8$	64.17
$\gamma_2 + \gamma_{18}$	63.33

TAB. 5.46: Resultado do algoritmo K-Means com $K = 10$, após as combinações para o conjunto MESMOCALCADO.

IGNORECALCADO

As características a serem combinadas e submetidas ao teste com o algoritmo K-NN para o grupo IGNORECALCADO, estão reunidas na TAB. 5.47. Devido ao número de combinações, são mostrados na TAB. 5.48 somente os resultados maiores que 65 %.

Característica	Símbolo Representativo	P	% Acertos
MFCC 1	γ_1	-	60
Coef. Cepstrais	φ	0.2	58.57
MFCC 3	γ_3	-	54.8
Sonoridade Específica	λ_e	-	54.76
MFCC 2	γ_2	-	52.9

TAB. 5.47: Características combinadas para o conjunto IGNORECALCADO, avaliando o K-NN.

Combinação	% Acertos
$\gamma_1 + \gamma_2 + \lambda_e$	72.86
$\gamma_1 + \lambda_e$	71.90
$\gamma_1 + \gamma_2$	63.33
$\gamma_1 + \gamma_2 + \varphi$	62.86
$\gamma_1 + \varphi$	59.52
$\gamma_1 + \gamma_2 + \gamma_3 + \lambda_e$	59.05
$\gamma_1 + \gamma_2 + \gamma_3 + \lambda_e + \varphi$	59.05
$\gamma_1 + \gamma_2 + \gamma_3$	58.57
$\gamma_1 + \gamma_3$	55.24
$\gamma_1 + \gamma_2 + \gamma_3 + \varphi$	58.57

TAB. 5.48: Resultado do algoritmo K-NN após as combinações para o conjunto IGNORE-CALCADO.

A TAB. 5.49 mostra as características eleitas para avaliação do K-Means, com $K = 5$. Os resultados das combinações com os melhores resultados se encontram na TAB. 5.50.

Característica	Símbolo Representativo	P	% Acertos
MFCC 8	γ_8	-	33.8
MFCC 3	γ_3	-	28.1
MFCC 6	γ_6	-	28.6
MFCC 16	γ_{16}	0.2	25.2
Coef. Cepstrais	φ	0.05	18.1
Sonoridade Específica	λ_e	0.05	18.1

TAB. 5.49: Características combinadas para o conjunto IGNORECALCADO, avaliando o K-Means com $K = 5$.

Combinação	% Acertos
$\gamma_3 + \gamma_8 + \gamma_{16}$	42.38
$\gamma_3 + \gamma_6 + \gamma_8 + \gamma_{16}$	41.90
$\gamma_3 + \gamma_{16}$	40.48
$\gamma_3 + \gamma_6 + \gamma_8$	40.00
$\gamma_6 + \gamma_8$	37.14
$\gamma_8 + \gamma_{16}$	37.14
$\gamma_8 + \lambda_e$	37.14
$\gamma_3 + \gamma_8$	36.19
$\gamma_3 + \gamma_6$	35.24
$\gamma_8 + \varphi$	33.81
$\gamma_6 + \gamma_{16}$	33.33
$\gamma_6 + \lambda_e$	30.48
$\gamma_{16} + \lambda_e$	30.48

TAB. 5.50: Resultado do algoritmo K-Means com $K = 5$, após as combinações para o conjunto IGNORECALCADO.

A TAB. 5.51 mostra as características eleitas para avaliação do K-Means, com $K = 10$. Os resultados das combinações com os melhores resultados se encontram na TAB. 5.52.

Característica	Símbolo Representativo	P	% Acertos
MFCC 3	γ_3	-	44.8
MFCC 6	γ_6	-	33.3
MFCC 9	γ_9	0.2	31.9
MFCC 10	γ_{10}	-	30.5
MFCC 16	γ_{16}	0.2	29
Coef. Cepstrais	φ	0.05	30.4
Sonoridade Específica	λ_e	-	26.7

TAB. 5.51: Características combinadas para o conjunto IGNORECALCADO, avaliando o K-Means com $K = 10$.

Combinação	% Acertos
$\gamma_3 + \gamma_6 + \gamma_9 + \gamma_{10}$	63.81
$\gamma_3 + \gamma_6 + \gamma_9 + \gamma_{16}$	61.90
$\gamma_3 + \gamma_6 + \gamma_{10}$	60.95
$\gamma_3 + \gamma_6 + \gamma_9$	59.05
$\gamma_3 + \gamma_6 + \gamma_{16}$	59.52
$\gamma_3 + \gamma_6 + \gamma_9 + \lambda_e$	58.57
$\gamma_3 + \gamma_6 + \gamma_9 + \varphi$	55.71
$\gamma_3 + \gamma_6 + \gamma_9 + \gamma_{10} + \gamma_{16}$	55.71
$\gamma_3 + \gamma_6 + \lambda_e$	54.76
$\gamma_3 + \gamma_6 + \varphi$	54.76
$\gamma_3 + \gamma_6 + \gamma_9 + \gamma_{10} + \gamma_{16} + \lambda_e + \varphi$	54.76
$\gamma_3 + \gamma_{10}$	55.24
$\gamma_3 + \gamma_6$	54.76
$\gamma_3 + \gamma_9$	49.05
$\gamma_3 + \gamma_{16}$	48.57
$\gamma_6 + \gamma_9$	42.38
$\gamma_9 + \gamma_{16}$	40.95

TAB. 5.52: Resultado do algoritmo K-Means com $K = 10$, após as combinações para o conjunto IGNORECALCADO.

Ao analisar os resultados percebe-se que os coeficientes mel cepstrais, a sonoridade específica e os coeficientes cepstrais estiveram presentes dentre as melhores acurácias, segundo a metodologia aplicada. Portanto, no escopo estudado, pode-se afirmar que tais características são importantes na discriminação de indivíduos através do som dos passos. Entretanto, esse fato não descarta as outras características exploradas podendo ser aplicadas de forma diferente ou mesmo extraídas utilizando outros métodos.

6 CONSIDERAÇÕES FINAIS

Neste capítulo são apresentados os comentários finais do trabalho. Na Seção 6.1 são descritos os comentários gerais e contribuições do presente trabalho. Posteriormente, na Seção 6.2 estão reunidas as conclusões alcançadas com a presente pesquisa. Dando sequência a Seção 6.3 reúne alguns desdobramentos possíveis. Por fim, a Seção 6.4 encerra o referido trabalho.

6.1 DISCUSSÕES GERAIS

O presente estudo faz parte do projeto de Casa Inteligente do IME. Primeiramente, em (NASCIMENTO, 2002) foi marcado o início do projeto intitulado inicialmente como SDAR, onde foram levantados os principais requisitos para automação inteligente da casa, destacando-se o sistema de identificação por meio do padrão de caminhar de seus habitantes. Posteriormente, em (LIMA, 2005) a idéia do SDAR foi amadurecida, sendo proposta a modelagem dos serviços da casa, fabricado um sensor passos juntamente com a implementação de um algoritmo de identificação; e implementado e testado o serviço de iluminação adaptativa. Complementando a modelagem dos serviços da casa, em (BOTELHO, 2005) um sistema inteligente residencial foi modelado através de um sistema multiagentes reativo, com agentes serviços e cômodos, agindo cooperativamente para manter o comportamento adaptativo da residência. Além de poder ser instalado em diversos ambientes (escolas, igrejas, escritórios), o modelo proposto por (BOTELHO, 2005) ainda possui a flexibilidade da agregação de novos serviços, sem alteração significativa do modelo.

Deste modo, o trabalho corrente galga mais um passo no projeto da Casa Inteligente, marcando o início da pesquisa em um sistema de identificação baseado no som dos passos. Uma parte dessa investigação envolveu a busca por modelos do processo de geração sonora, ocorrida no impacto do pé com o solo. Onde foi mostrado que apesar dos esforços de alguns trabalhos (EKIMOV, 2006, 2007), ainda não foi possível estabelecer um modelo de produção do som dos passos. Dentre as principais dificuldades observadas para essa tarefa, estão a força de atrito; o tipo de material do solo e calçado; peso do indivíduo caminhante; e modo de executar a trajetória.

A inexistência de um modelo explicativo para a geração do som dos passos inspirou a realização de uma análise da aplicação de características, utilizadas anteriormente em outros sistemas de identificação baseados em som, no sinal acústico dos passos. Para isso, foi realizada uma busca por tais parâmetros e seus respectivos algoritmos e fórmulas de extração. Essa busca gerou a aglomeração de atributos em dois grupos: objetivos e subjetivos. O grupo de predicados objetivos investigados foi composto pelos seguintes itens: taxa de cruzamento por zero - TCZ, energia, envelope espectral baseado e em dois métodos (Welch e Periodograma), coeficientes cepstrais baseados em LPC e frequência dos passos. O grupo dos atributos subjetivos teve como participantes vinte coeficientes mel-cepstrais; e as sonoridades específica e total baseadas no modelo de Zwicker.

A quantidade de dados resultante da extração das características aplicadas implicou na consideração de um método de seleção de atributos. O método de triagem de atributos implementado neste trabalho foi o critério de Fisher, onde se maximiza a razão da variância intra-classes pela variância entre-classes. Tal critério foi aplicado a cada coluna dos dados oriundos da característica em avaliação. Em seguida, os valores são ordenados de forma decrescente. Foi estabelecida uma variável, P , a qual indica a proporção de “melhores indivíduos”, segundo o critério de Fisher, utilizados para teste. Os seguintes valores foram usados: “-” indica a não execução da seleção; 0.2, 0.1 e 0.05 correspondem a 20 %, 10 % e 5 % respectivamente.

Dos trabalhos que tratam o problema de forma direta (identificação de indivíduos a partir do som dos passos) encontrados na literatura (ITAI, 2006; SHOJI, 2004, 2005), foi utilizada uma base de dados muito restrita (cinco indivíduos com um tipo de calçado). Para avaliação dos grupos de predicados estudados, foi realizado no presente trabalho um experimento com 15 indivíduos que usaram ao todo 6 tipos de calçado. Para estudar a qualidade de determinado parâmetro, a base de dados foi separada em seis grupos, a saber: a) COTURNO, TENIS e SAPATO, constituído por indivíduos que caminharam utilizando, respectivamente, o tipo de calçado coturno, tênis e sapato; b) UMPORCALCADO, composto por amostras de um indivíduo de cada tipo de calçado; c) MESMOCALCADO, formado por seis participantes que caminharam utilizando sapato e tênis; e d) IGNORE-CALCADO, grupo composto por todas as amostras de todos os indivíduos e todos os tipos de calçado. Para cada par (indivíduo, calçado), 15 sequências foram capturadas.

Para a tarefa de classificação, dois algoritmos foram avaliados: k vizinhos mais próximos - K-NN e o algoritmo de agrupamento K-Means. A metodologia de partição de dados

utilizada no K-NN foi a *leave-one-out*, onde deixa-se um indivíduo fora da base de treinamento, usando-o como teste do modelo. No K-Means, a metodologia de validação cruzada com k conjuntos foi empregada. O K-NN executou com k variando de 1 até 20. No K-Means, os valores de k escolhidos foram 5 e 10. Nos parâmetros multi-colunas, ambos métodos de classificação executaram variando os parâmetros do critério de Fisher, P , para avaliação da qualidade do método de seleção de atributos.

No intuito de avaliar a sensibilidade dos elementos dos grupos objetivo e subjetivo, os testes de classificação foram realizados de duas maneiras: 1) em cada subgrupo, os algoritmos foram executados utilizando cada característica individualmente; 2) após, uma eleição das melhores configurações de execução e características, segundo as melhores taxas de classificação, são combinadas e avaliadas novamente pelos algoritmos.

Um sistema composto de quatro etapas básicas: 1) modelo da base de dados; 2) pré-processamento; 3) extração e seleção de características; e 4) identificação; foi implementado em Matlab, onde foi possível realizar as avaliações propostas. Quando a combinação das melhores características e configurações foi realizada, o sistema proposto obteve as seguintes taxas de reconhecimento, por subgrupo:

- COTURNO: 97.5 % (K-NN), 97.5 % (K-Means com $k = 5$) e 95 % (K-Means com $k=10$);
- TENIS: 86.25 % (K-NN), 81.25 % (K-Means com $k = 5$) e 75 % (K-Means com $k = 10$);
- SAPATO: 70 % (K-NN), 88.33 % (K-Means com $k = 5$) e 90 % (K-Means com $k = 10$);
- UMPORCALCADO: 91.67 (K-NN), 88.33 % (K-Means com $k = 5$) e 85 % (K-Means com $k = 10$);
- MESMOCALCADO: 73.33 (K-NN), 75.83 % (K-Means com $k = 5$) e 69.17 % (K-Means com $k = 10$);
- IGNORECALCADO: 72.86 (K-NN), 42.38 % (K-Means com $k = 5$) e 63.81 % (K-Means com $k = 10$);

As contribuições deste trabalho podem ser resumidas em:

- Análise de características, utilizadas anteriormente para soluções em reconhecimento de locutor, na classificação através do som dos passos;
- União de características objetivas e subjetivas no intuito de aumentar a taxa de acertos do sistema;
- Avaliação de dois algoritmos de classificação, com diferentes parametrizações, aplicados ao problema em questão;
- Investigação da qualidade dos resultados mediante o número variado de calçados e indivíduos.

6.2 CONCLUSÕES

O presente trabalho abordou o problema de identificação de indivíduos a partir do som dos passos, através do estudo de características aplicadas à sinais sonoros, objetivando classificação, porém aplicadas ao sinal sonoro criado pelo impacto do pé com o chão.

Para a realização desse estudo, uma busca por características utilizadas em outros sistemas de identificação baseadas em som, tais como identificação de locutor, como também parâmetros psicoacústicos foram aplicados aos sons dos passos. Os parâmetros considerados geraram um volume considerável de dados, fazendo-se necessária a aplicação de métodos para eleição de atributos que melhor representem o padrão do caminhar.

Nas avaliações das características, a frequência dos passos, outrora usada com sucesso em trabalhos predecessores (LIMA, 2005), não se portou como uma característica discriminativa. Em contrapartida, os 20 coeficientes mel-cepestrais examinados mostraram ser bons parâmetros representativos do padrão de caminhada. Entretanto, não descarta-se a possibilidade de encontrar padrões no período do caminhar.

A fase de identificação foi composta com a avaliação de dois algoritmos classificadores: *k-means* e *k* vizinhos mais próximos - KNN, utilizando validação cruzada com *k* conjuntos e *leave-one-out*, como metodologias de partição de dados, respectivamente. O critério de Fisher foi utilizado como método de seleção de atributos. Em ambos, foram testadas várias configurações utilizando o critério de seleção, com variadas proporções de dados.

Para avaliar o sistema, foi realizado um experimento real através de gravações do caminhar de um conjunto de indivíduos, os quais utilizaram seis tipos de calçados diferentes no total. Uma análise foi feita através da separação do conjunto total em seis subgrupos ca-

racterísticos: COTURNO, TENIS, SAPATO, UMPORCALCADO, MESMOCALCADO e IGNORECALCADO.

A metodologia de testes consistiu em aplicar cada característica, singularmente, nos algoritmos de classificação. Após, as características são selecionadas e combinadas de acordo com a quantidade de acertos obtidos na etapa anterior. Essa metodologia resultou em taxas de identificação que vão de 33.8 % a 97.5 %.

6.3 TRABALHOS FUTUROS

Como foi visto, para se chegar a um sistema de identificação ideal, baseado no som dos passos, ainda requer muitos esforços. Por este trabalho ser o início de um novo sistema de identificação, existem inúmeras possibilidades a se explorar. Dentre elas, a seguir são destacadas algumas.

- **Pré-processamento** - é necessário elaborar uma pesquisa mais profunda no sinal do som dos passos, de maneira a projetar filtros ou banco de filtros para filtragem do ruído ambiente. É preciso também elaborar um algoritmo que consiga identificar os blocos com e sem sinal dos passos, de forma que se possa trabalhar somente com o sinal “vozeado”; e trabalhar com a aplicação das características em cada estímulo, no intuito de encontrar padrões em cada bloco de onda provocado por cada contato do pé com o solo;
- **Extração de características** - existem outras características que podem ser adicionadas a esse sistema. Exemplos são: os vários algoritmos para o cálculo da frequência fundamental e sensação de frequência fundamental *pitch*; outros parâmetros psicoacústicos como *fluctuation strength*, *sharpness*, *roughness* (ITAI, 2006); usar técnicas mais robustas, tais como o *Random Sample Consensus* - RANSAC (FISCHLER, 1987; ZULIANI, 2006), para determinação de modelos que identifiquem padrões no sinal sonoro do caminhar;
- **Seleção de atributos** - estudar e aplicar outros métodos de seleção de atributos, tais como a análise de componentes principais;
- **Identificação** - implementar outros métodos de classificação, tais como diversas variações das redes neurais artificiais;

- **Reconhecimento de vários indivíduos simultaneamente** - existe um atributo subjetivo que através dele é possível notar as diferenças entre dois ou mais instrumentos musicais tocando uma mesma nota em uma mesma frequência. Este atributo se chama timbre. Esforços nesse sentido poderiam ser realizados para tentar diferenciar o som dos passos de dois ou mais indivíduos, caminhando simultaneamente. Sendo esse problema um dos maiores desafios ainda restantes nesta pesquisa.

Além do exposto, o som pode ainda ser usado para aumentar as capacidades de segurança e conforto dos indivíduos. Por exemplo, pode ser implementado um sistema que reconheça os padrões de estresse da voz, ou barulhos de gritos, tiros ou pancadas; para a aplicação de medidas de segurança. Enfim, as possibilidades de pesquisa que se abrem ao estudar a aplicação do sinal sonoro à casa inteligente são diversas, pois através do som é possível desde a identificação de aspectos psicológicos do indivíduo, até a identificação de determinada música apenas por sussurro produzido.

6.4 AGRADECIMENTOS

O autor agradece à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, por financiar parte desta pesquisa.

7 REFERÊNCIAS BIBLIOGRÁFICAS

- ABU-SHIKHAH, N. M. e ALOQUILI, O. A new spectrum envelope estimation technique based on sample projections. *Signal Process.*, 87(10):2446–2454, 2007. ISSN 0165-1684.
- ADDLESEE, M. D., JONES, A., LIVESEY, F. e SAMARIA, F. The orl active floor [sensor system]. *Personal Communications, IEEE [see also IEEE Wireless Communications]*, 4(5):35–41, 1997. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=626980.
- ALLARD, P., STOKES, I. A. e BLANCHI, J.-P. *Three - Dimensional Analysis of Human Movement*. Human Kinetics Publishers, 1 edition, 1995. ISBN 0873226232.
- ANALOG DEVICES, I. Ad1986a ac97 2.3 and hd audio soundmax audio codec w/jack sensing, 2007. URL <http://www.analog.com/en/prod/0,2877,AD1986A,00.html> (visitado em 09 de janeiro de 2008).
- APPELL, J.-E. *Loudness models rehabilitative audiology*. Tese de Doutorado, Universität Oldenburg, 2002.
- BAKER, R. The history of gait analysis before the advent of modern computers. *Gait & Posture Volume*, 26:331–342, Setembro 2007.
- BARBEDO, J. G. A. *Avaliação objetiva de qualidade de sinais de áudio e voz*. Tese de Doutorado, Universidade Estadual de Campinas, 2004.
- BOTELHO, W. T. Um sistema de identificação e adaptação pervasivo para a casa inteligente utilizando sistemas multiagentes. Dissertação de Mestrado, Instituto Militar de Engenharia, 2005.
- CAMPOS, T. E. Técnicas de seleção de características com aplicações em reconhecimento de faces. Dissertação de Mestrado, Instituto de Matemática e Estatística/USP, 2001.
- CARVALHO, R. L., GOLDSCHMIDT, R. R., ROSA, P. F. F. e DE LIMA, S. S. Knn aplicado à classificação de indivíduos na casa inteligente. Em *Simpósio de Pesquisa Operacional da Marinha*, 2007.
- CATTIN, P. C. *Biometric authentication system using human gait*. Tese de Doutorado, Swiss Federal of Technology, Zurich, Switzerland, 2002.
- CHAN, M. J., CHAN, M. . V., HUI, M. L. e KUO, M. K. Normal gait analysis. volume 4, págs. 13–15. The Journal of the Hong Kong Physiotherapy Association, 1981. URL <http://sunzi1.lib.hku.hk/hkjo/view/25/2500012.pdf>.

- CRESS, D. H. Terrain considerations and data base development for the design and testing of devices to detect intruder-induced ground motion. Technical Report M-78-1, U.S. Waterways Experimental Station, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS, 1978.
- DAVIS, S. e MERMELSTEIN, P. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing]*, *IEEE Transactions on*, 28(4):357–366, 1980.
- DE LIMA, C. B. Sistemas de verificação de locutor independente do texto baseados em gmm e ar-vetorial utilizando pca. Dissertação de Mestrado, Instituto Militar de Engenharia, 2001.
- DE RUYTER, B., AARTS, E., MARKOPOULOS, P. e IJSSELSTEIJN, W. *Ambient Intelligence*, chapter Ambient Intelligence Research in HomeLab Engineering the User Experience, págs. 49–62. Springer Berlin Heidelberg, 2005. ISBN 978-3-540-23867-6 (Print) 978-3-540-27139-0 (Online).
- EKIMOV, A. e SABATIER, J. M. Vibration and sound signatures of human footsteps in buildings. *The Journal of the Acoustical Society of America*, 120(2):762–768, 2006. URL <http://link.aip.org/link/?JAS/120/762/1>.
- EKIMOV, A. e SABATIER, J. M. Ultrasonic wave generation due to human footsteps on the ground. *The Journal of the Acoustical Society of America*, 121(3):EL114–EL119, 2007.
- ELLIS, D. P. W. PLP and RASTA (and MFCC, and inversion) in Matlab, 2005. URL <http://www.ee.columbia.edu/~dpwe/resources/matlab/rastamat/>. online web resource.
- EVANS, L. B., BASS, H. E. e SUTHERLAND, L. C. Atmospheric absorption of sound: Theoretical predictions. *The Journal of the Acoustical Society of America*, 51(5B): 1565–1575, 1972.
- FARRELL, K., MAMMONE, R. e ASSALEH, K. Speaker recognition using neural networks and conventional classifiers. *Speech and Audio Processing, IEEE Transactions on*, 2(1):194–205, 1994. ISSN 1063-6676.
- FASTL, H. e ZWICKER, E. *Psychoacoustics: Facts and Models*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006. ISBN 3540231595.
- FECHINE, J. M. *Reconhecimento Automático de Identidade Vocal Utilizando Modelagem Híbrida: Paramétrica e Estatística*. Tese de Doutorado, Universidade Federal da Paraíba, 2000.
- FISCHLER, M. A. e BOLLES, R. C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. págs. 726–740, 1987.

- FLETCHER, H. e MUNSON, W. A. Loudness, its definition, measurement and calculation. *The Journal of the Acoustical Society of America*, 5(2):82–108, 1933. URL <http://link.aip.org/link/?JAS/5/82/1>.
- FONTIJN, W. e MENDELS, P. Storytoy the interactive storytelling toy. Em *In Proc. W3 PerGames*, 2005.
- GALBRAITH, F. W. e BARTON, M. V. Ground loading from footsteps. *The Journal of the Acoustical Society of America*, 48(5B):1288–1292, 1970. URL <http://link.aip.org/link/?JAS/48/1288/1>.
- GERHARD, D. Audio signal classification: History and current techniques. Technical report, Department of Computer Science University of Regina, Saskatchewan, CANADA, November 2003.
- GOLDSCHMIDT, R. R. e PASSOS, E. *Data Mining um guia prático*, volume 1. Editora Elsevier, 2005. ISBN 8535218777.
- HAN, J. e KAMBER, M. *Data Mining. Concepts and Techniques*. Morgan Kaufmann, San Francisco, LA, 2006. ISBN 1-55860-901-6.
- INTILLE, S. S., LARSON, K., BEAUDIN, J. S., TAPIA, E. M., KAUSHIK, P., NAWYN, J. e MCLEISH, T. The placelab: a live-in laboratory for pervasive computing research (video). Em *Proceedings of Pervasive 2005 Video Program*, 2005.
- ITAI, A. e YASUKAWA, H. Footstep recognition with psycho-acoustics parameter. Em *APCCAS IEEE Asia Pacific Conference on Circuits and Systems 2006*, págs. 992–995, Singapore, December 2006. ISBN 1-4244-0387-1.
- JR., J. R. D., PROAKIS, J. G. e HANSEN, J. H. *Discrete Time Processing of Speech Signals*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1993. ISBN 0780353862.
- JÚNIOR, J. E. F. N. *Síntese evolutiva de segmentos sonoros*. Tese de Doutorado, Universidade Estadual de Campinas, 2003.
- KUDO, T. N. Computação ciente de contexto aplicada ao monitoramento de condições críticas em ambientes físicos. Dissertação de Mestrado, Universidade Federal de São Carlos, 2004.
- LAUNER, S. *Loudness Perception in Listeners with sensorineural hearing impairment*. Tese de Doutorado, Universität Oldenburg, 1995.
- LEITE, I. e PEREIRA, A. icat, the affective chess player. Em *Paolo Petta and Carlos Martinho (eds) proceedings of the systems demonstrations. Second international conference on Affective Computing and Intelligent Interaction (ACII 07)*, págs. 29–33, 2007. ISBN 978-989-20-0799-1.
- LI, X., LOGAN, R. J. e PASTORE, R. E. Perception of acoustic source characteristics: Walking sounds. *The Journal of the Acoustical Society of America*, 90:3036–3049, 1991.

- LIMA, E. L. *Elementos de Topologia Geral*. EDUSP, 1970.
- LIMA, S. S. Análise e desenvolvimento de um ambiente para aplicações domóticas. Dissertação de Mestrado, Instituto Militar de Engenharia, 2005.
- MAMMONE, R. J., ZHANG, X. e RAMACHANDRAN, R. P. Robust speaker recognition: a feature-based approach. *Signal Processing Magazine, IEEE*, 13(5):58–71, 1996. ISSN 1053-5888.
- MARZINZIK, M., APPELL, J. E., HOHMANN, V. e KOLLMEIER, B. *Psychoacoustics, Speech and Hearing Aids*, chapter Evaluation of dynamic compression algorithms using a loudness model for hearing impaired listeners, págs. 203–204. World Scientific, 1996.
- MOORE, B. C. J., GLASBERG, B. R. e VICKERS, D. A. *Psychoacoustics, Speech and Hearing Aids*, chapter Factors influencing loudness perception in people with cochlear hearing loss, págs. 7–18. World Scientific, 1996.
- MOZER, M. C. The neural network house: An environment that adapts to its inhabitants. Em COEN, M., editor, *Proceedings of the American Association for Artificial Intelligence Spring Symposium on Intelligent Environments*, págs. 110–114, Menlo Park, CA, 1998. AAAI Press.
- NASCIMENTO, A. F. D. Sistema dinâmico de automação residencial. Dissertação de Mestrado, INSTITUTO MILITAR DE ENGENHARIA, 2002.
- OETZMANN, A. e MAZZONI, D. Audacity manual, 2008. URL <http://audacity.sourceforge.net/manual-1.2/> (Visitada em 09 de janeiro de 2008).
- OLSEN, H. F. *Music, Physics and Engineering*. Dover Publishing, Canada, 2nd edition, 1967. ISBN 486-21769-8.
- ORR, R. J. e ABOWD, G. D. The smart floor: a mechanism for natural user identification and tracking. Em *CHI '00: CHI '00 extended abstracts on Human factors in computing systems*, págs. 275–276, New York, NY, USA, 2000. ACM. ISBN 1-58113-248-4.
- PAPPA, G. L., FREITAS, A. A. e KAESTNER, C. A. A. A multiobjective genetic algorithm for attribute selection. Em LOFTI, A., GARIBALDI, J. e JOHN, R., editores, *Proc. 4th Int. Conf. on Recent Advances in Soft Computing (RASC-2002)*, págs. 116–121. Nottingham Trent University, December 2002. ISBN 1842330764. URL <http://www.cs.kent.ac.uk/pubs/2002/1789>.
- PICONE, J. Signal modeling techniques in speech recognition. *Proceedings of the IEEE*, 81(9):1215–1247, 1993. ISSN 0018-9219.
- RABINER, L. e JUANG, B.-H. *Fundamentals of speech recognition*. Prentice Hall PTR, 1993. ISBN 0130151572.

- REYNOLDS, D. A. e ROSE, R. C. Robust text-independent speaker identification using gaussian mixture speaker models. *Speech and Audio Processing, IEEE Transactions on*, 3(1):72–83, 1995.
- ROSA, P. F. F., VIDAL, F. S. e LIMA, S. S. Identificação de indivíduos a partir das características do seu caminhar. Em *Congresso Latino-americano de Controle Automático*, págs. 360–365, 2006.
- RUYTER, B. D. 365 days of ambient intelligence research in homelab. Neroc Publishers, 2003.
- SAVIO, D. e LUDWIG, T. Smart carpet: A footstep tracking interface. Em *AINAW '07: Proceedings of the 21st International Conference on Advanced Information Networking and Applications Workshops*, págs. 754–760, Washington, DC, USA, 2007. IEEE Computer Society. ISBN 0-7695-2847-3.
- SCHWARZ, D. Spectral envelopes in sound analysis and synthesis. Diplomarbeit nr. 1622, Universität Stuttgart, Fakultät Informatik, 1998.
- SHANNON, C. E. Communication in the presence of noise. *Proc. Institute of Radio Engineers, Reprint as classic paper in: Proc. IEEE*, 37(2):10–21, Jan. 1949.
- SHOJI, Y., TAKASUKA, T. e YASUKAWA, H. Personal identification using footstep detection. Em *Proceedings of 2004 International Symposium on Intelligent Signal Processing and Communication Systems - ISPACS 2004*, págs. 43–47, November 2004. ISBN 0-7803-8639-6.
- SHOJI, Y., TAKASUKA, T. e YASUKAWA, H. A study on footstep detection for personal recognition. *Proceedings of the IEICE General Conference*, págs. 53–54, 2005. URL <http://ci.nii.ac.jp/naid/110004738138/en/>.
- SILVA, D. G. Estudo de compensação de canal e análise fractal aplicada ao reconhecimento de locutor. Dissertação de Mestrado, Instituto Militar de Engenharia, 2002.
- SMITH, J. O. *Mathematics of the Discrete Fourier Transform (DFT), with Music and Audio Applications*, volume III. W3K Publishing, 2003. ISBN 0-9745607-0-7.
- STEVENS, S. S. e VOLKMANN, J. The relation of pitch to frequency. *The American Journal of Psychology*, 53(3):329–353, Jul. 1940.
- STOFFER, D. S., TYLER, D. E. e WENDT, D. A. The spectral envelope and its applications. *Statistical Science*, 15(3):224–263, 2000.
- SUUTALA, J. e RÖNING, J. Combining classifiers with different footstep feature sets and multiple samples for person identification. Em *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2005.
- TAPIA, E. M., MARMASSE, N., INTILLE, S. S. e LARSON, K. Mites: Wireless portable sensors for studying behavior. Em *Proceedings of Extended Abstracts Ubicomp 2004: Ubiquitous Computing*, 2004. URL <http://web.media.mit.edu/intille/papers-files/MunguiaTapiaETAL04.pdf>.

- TIMONEY, J., MACMANUS, L., LYSAGHT, T. e SCHOENWIESNER, M. Implementing loudness models in matlab. Em *Proc. of the 7th Int. Conference on Digital Audio Effects (DAFX-04)*, number 1, págs. 177–188, Naples, Italy, 2004.
- VAN BREEMEN, A. icat: Experimenting with animabotics. Em *AISB 2005 Creative Robotics Symposium*, Hatfield, England, April 2005a.
- VAN BREEMEN, A., YAN, X. e MEERBEEK, B. icat: an animated user-interface robot with personality. Em *AAMAS '05: Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, págs. 143–144, New York, NY, USA, 2005b. ACM. ISBN 1-59593-093-0.
- WATTERS, B. G. Impact-noise characteristics of female hard-heeled foot traffic. *The Journal of the Acoustical Society of America*, 37(4):619–630, 1965. URL <http://link.aip.org/link/?JAS/37/619/1>.
- WELCH, P. The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *Audio and Electroacoustics, IEEE Transactions on*, 15(2):70–73, Jun 1967. ISSN 0018-9278.
- WINTER, D. A. Kinematic and kinetic patterns in human gait: Variability and compensating effects. *Human Movement Science*, 3:51–76, March 1984.
- ZERNICKE, R. F. The emergence of human biomechanics. *Perspectives on the Academic Discipline of Physical Education*, págs. 124–136, 1981.
- ZULIANI, M. *Computational Methods for Automatic Image Registration*. Tese de Doutorado, University of California, 2006.
- ZWICKER, E., FASTL, H. e DALLMAYR, C. Basic program for calculating the loudness of sounds from their 1/3 oct band spectra according to iso532b. *Acustica*, 55:63–67, 1984.
- ZWICKER, E. e FASTL, H. *Psychoacoustics: Facts and Models*. Springer-Verlag New York, Inc., Berlin, 1990. ISBN 3540231595.

8 APÊNDICES

8.1 APÊNDICE 1: RESULTADOS COMPLEMENTARES

8.1.1 K-MEANS APLICADO A TODOS OS COEFICIENTES MEL-CEPESTRAIS

K-Means - MFCC (Coef. 01)						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)
-	5/30.00	5/25.00	5/23.33	5/16.67	5/26.67	5/13.81
0.05	5/25.00	5/12.50	5/16.67	5/16.67	5/18.33	5/9.52
0.10	5/25.00	5/12.50	5/20.00	5/16.67	5/16.67	5/9.52
0.20	5/25.00	5/12.50	5/38.33	5/16.67	5/16.67	5/10.00
-	10/47.50	10/33.75	10/35.00	10/16.67	10/41.67	10/21.43
0.05	10/25.00	10/15.00	10/30.00	10/26.67	10/22.50	10/9.52
0.10	10/35.00	10/13.75	10/36.67	10/23.33	10/20.83	10/10.48
0.20	10/37.50	10/15.00	10/40.00	10/21.67	10/22.50	10/10.48
K-Means - MFCC (Coef. 02)						
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)
-	5/32.50	5/31.25	5/23.33	5/16.67	5/37.50	5/12.38
0.05	5/42.50	5/25.00	5/23.33	5/20.00	5/35.00	5/14.76
0.10	5/45.00	5/30.00	5/41.67	5/20.00	5/45.83	5/15.24
0.20	5/42.50	5/33.75	5/41.67	5/18.33	5/42.50	5/16.67
-	10/62.50	10/46.25	10/38.33	10/36.67	10/48.33	10/25.71
0.05	10/62.50	10/40.00	10/40.00	10/43.33	10/46.67	10/19.52
0.10	10/72.50	10/37.50	10/48.33	10/40.00	10/49.17	10/20.95
0.20	10/67.50	10/52.50	10/53.33	10/41.67	10/51.67	10/23.33
K-Means - MFCC (Coef. 03)						
-	5/60.00	5/35.00	5/38.33	5/43.33	5/30.00	5/26.67
0.05	5/55.00	5/27.50	5/40.00	5/43.33	5/25.83	5/20.00
0.10	5/62.50	5/32.50	5/35.00	5/51.67	5/26.67	5/22.86
0.20	5/67.50	5/33.75	5/45.00	5/51.67	5/24.17	5/28.10
-	10/60.00	10/61.25	10/50.00	10/68.33	10/50.00	10/44.76
0.05	10/60.00	10/45.00	10/48.33	10/51.67	10/30.00	10/29.52
0.10	10/62.50	10/47.50	10/50.00	10/53.33	10/35.83	10/33.33
0.20	10/70.00	10/48.75	10/58.33	10/58.33	10/43.33	10/40.95
K-Means - MFCC (Coef. 04)						
-	5/57.50	5/28.75	5/33.33	5/26.67	5/33.33	5/18.57
0.05	5/67.50	5/25.00	5/38.33	5/35.00	5/27.50	5/16.19
0.10	5/80.00	5/27.50	5/35.00	5/36.67	5/25.83	5/11.43
0.20	5/70.00	5/31.25	5/31.67	5/31.67	5/25.83	5/12.38
-	10/77.50	10/52.50	10/46.67	10/55.00	10/37.50	10/28.10
0.05	10/82.50	10/40.00	10/46.67	10/50.00	10/40.83	10/24.76
0.10	10/90.00	10/50.00	10/41.67	10/56.67	10/41.67	10/23.33
0.20	10/85.00	10/55.00	10/45.00	10/53.33	10/39.17	10/25.24

TAB. 8.1: K-Means aplicado aos coeficientes 01 ao 04.

K-Means - MFCC (Coef. 05)						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)
-	5/60.00	5/30.00	5/38.33	5/35.00	5/31.67	5/23.33
0.05	5/62.50	5/40.00	5/53.33	5/43.33	5/36.67	5/14.76
0.10	5/75.00	5/48.75	5/48.33	5/43.33	5/36.67	5/18.57
0.20	5/80.00	5/45.00	5/46.67	5/48.33	5/38.33	5/21.43
-	10/65.00	10/47.50	10/53.33	10/50.00	10/30.83	10/29.52
0.05	10/72.50	10/42.50	10/51.67	10/50.00	10/37.50	10/18.57
0.10	10/77.50	10/58.75	10/50.00	10/53.33	10/38.33	10/24.76
0.20	10/77.50	10/52.50	10/55.00	10/48.33	10/40.00	10/25.71
K-Means - MFCC (Coef. 06)						
-	5/40.00	5/51.25	5/25.00	5/51.67	5/34.17	5/28.57
0.05	5/52.50	5/33.75	5/41.67	5/46.67	5/40.00	5/20.95
0.10	5/62.50	5/41.25	5/35.00	5/58.33	5/35.83	5/24.29
0.20	5/57.50	5/51.25	5/30.00	5/60.00	5/37.50	5/28.10
-	10/52.50	10/58.75	10/41.67	10/63.33	10/33.33	10/33.33
0.05	10/55.00	10/35.00	10/50.00	10/51.67	10/40.83	10/24.76
0.10	10/60.00	10/45.00	10/53.33	10/65.00	10/41.67	10/31.43
0.20	10/72.50	10/60.00	10/48.33	10/60.00	10/44.17	10/32.86
K-Means - MFCC (Coef. 07)						
-	5/30.00	5/26.25	5/30.00	5/35.00	5/25.00	5/16.19
0.05	5/32.50	5/26.25	5/38.33	5/36.67	5/30.83	5/10.95
0.10	5/30.00	5/37.50	5/36.67	5/33.33	5/30.00	5/9.52
0.20	5/27.50	5/37.50	5/35.00	5/35.00	5/32.50	5/10.48
-	10/47.50	10/42.50	10/43.33	10/55.00	10/29.17	10/28.10
0.05	10/47.50	10/35.00	10/48.33	10/51.67	10/39.17	10/23.33
0.10	10/47.50	10/52.50	10/53.33	10/58.33	10/40.00	10/23.33
0.20	10/52.50	10/51.25	10/55.00	10/66.67	10/36.67	10/26.67
K-Means - MFCC (Coef. 08)						
-	5/52.50	5/55.00	5/40.00	5/48.33	5/38.33	5/33.81
0.05	5/62.50	5/38.75	5/43.33	5/40.00	5/42.50	5/22.38
0.10	5/70.00	5/48.75	5/48.33	5/43.33	5/45.00	5/28.57
0.20	5/72.50	5/56.25	5/50.00	5/58.33	5/45.83	5/30.48
-	10/67.50	10/60.00	10/45.00	10/51.67	10/44.17	10/37.62
0.05	10/72.50	10/38.75	10/51.67	10/46.67	10/44.17	10/22.38
0.10	10/77.50	10/45.00	10/60.00	10/45.00	10/43.33	10/26.19
0.20	10/82.50	10/55.00	10/63.33	10/63.33	10/51.67	10/35.24

TAB. 8.2: K-Means aplicado aos coeficientes 05 ao 08.

K-Means - MFCC (Coef. 09)						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)
-	5/47.50	5/50.00	5/38.33	5/63.33	5/26.67	5/26.67
0.05	5/65.00	5/55.00	5/55.00	5/55.00	5/35.83	5/16.67
0.10	5/70.00	5/60.00	5/56.67	5/68.33	5/35.00	5/28.10
0.20	5/62.50	5/56.25	5/53.33	5/70.00	5/34.17	5/26.67
-	10/45.00	10/57.50	10/35.00	10/68.33	10/30.00	10/27.62
0.05	10/60.00	10/53.75	10/55.00	10/58.33	10/36.67	10/16.67
0.10	10/72.50	10/61.25	10/56.67	10/70.00	10/40.83	10/28.10
0.20	10/62.50	10/57.50	10/60.00	10/70.00	10/45.83	10/31.90
K-Means - MFCC (Coef. 10)						
-	5/40.00	5/40.00	5/36.67	5/33.33	5/36.67	5/15.24
0.05	5/57.50	5/45.00	5/35.00	5/50.00	5/35.00	5/20.00
0.10	5/62.50	5/41.25	5/41.67	5/45.00	5/40.00	5/17.14
0.20	5/60.00	5/42.50	5/33.33	5/38.33	5/39.17	5/16.67
-	10/55.00	10/47.50	10/33.33	10/50.00	10/45.83	10/30.48
0.05	10/67.50	10/47.50	10/38.33	10/60.00	10/39.17	10/20.48
0.10	10/75.00	10/53.75	10/51.67	10/60.00	10/45.83	10/23.81
0.20	10/80.00	10/48.75	10/48.33	10/60.00	10/48.33	10/24.76
K-Means - MFCC (Coef. 11)						
-	5/45.00	5/38.75	5/38.33	5/30.00	5/34.17	5/21.43
0.05	5/52.50	5/33.75	5/38.33	5/38.33	5/33.33	5/15.71
0.10	5/70.00	5/41.25	5/41.67	5/41.67	5/31.67	5/14.29
0.20	5/75.00	5/41.25	5/53.33	5/40.00	5/39.17	5/17.62
-	10/60.00	10/50.00	10/46.67	10/51.67	10/43.33	10/26.67
0.05	10/65.00	10/40.00	10/48.33	10/45.00	10/35.00	10/19.52
0.10	10/72.50	10/51.25	10/51.67	10/53.33	10/40.00	10/21.43
0.20	10/80.00	10/50.00	10/61.67	10/60.00	10/44.17	10/25.71
K-Means - MFCC (Coef. 12)						
-	5/65.00	5/37.50	5/33.33	5/25.00	5/34.17	5/17.62
0.05	5/62.50	5/31.25	5/50.00	5/55.00	5/42.50	5/19.05
0.10	5/62.50	5/43.75	5/50.00	5/48.33	5/40.83	5/15.24
0.20	5/75.00	5/53.75	5/45.00	5/46.67	5/40.00	5/18.57
-	10/65.00	10/38.75	10/40.00	10/45.00	10/37.50	10/26.19
0.05	10/62.50	10/37.50	10/55.00	10/60.00	10/43.33	10/26.19
0.10	10/75.00	10/47.50	10/48.33	10/58.33	10/46.67	10/23.81
0.20	10/80.00	10/58.75	10/55.00	10/56.67	10/43.33	10/22.86

TAB. 8.3: K-Means aplicado aos coeficientes 09 ao 12.

K-Means - MFCC (Coef. 13)						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)
-	5/57.50	5/52.50	5/33.33	5/51.67	5/35.00	5/19.05
0.05	5/80.00	5/43.75	5/31.67	5/51.67	5/39.17	5/21.43
0.10	5/82.50	5/55.00	5/36.67	5/63.33	5/33.33	5/20.48
0.20	5/85.00	5/58.75	5/43.33	5/63.33	5/38.33	5/22.38
-	10/62.50	10/48.75	10/36.67	10/50.00	10/36.67	10/21.43
0.05	10/82.50	10/42.50	10/36.67	10/53.33	10/36.67	10/21.90
0.10	10/80.00	10/51.25	10/46.67	10/60.00	10/35.00	10/19.05
0.20	10/77.50	10/57.50	10/46.67	10/58.33	10/40.00	10/21.90
K-Means - MFCC (Coef. 14)						
-	5/45.00	5/31.25	5/35.00	5/48.33	5/30.83	5/21.43
0.05	5/67.50	5/41.25	5/46.67	5/55.00	5/34.17	5/17.14
0.10	5/67.50	5/45.00	5/50.00	5/58.33	5/35.83	5/21.43
0.20	5/65.00	5/47.50	5/55.00	5/60.00	5/35.00	5/24.29
-	10/42.50	10/30.00	10/31.67	10/55.00	10/32.50	10/21.90
0.05	10/67.50	10/45.00	10/48.33	10/56.67	10/32.50	10/17.14
0.10	10/65.00	10/46.25	10/53.33	10/65.00	10/39.17	10/23.81
0.20	10/67.50	10/48.75	10/55.00	10/61.67	10/40.00	10/26.67
K-Means - MFCC (Coef. 15)						
-	5/50.00	5/30.00	5/35.00	5/31.67	5/29.17	5/20.95
0.05	5/57.50	5/36.25	5/50.00	5/33.33	5/30.00	5/14.76
0.10	5/62.50	5/38.75	5/53.33	5/46.67	5/37.50	5/20.00
0.20	5/55.00	5/40.00	5/61.67	5/48.33	5/40.83	5/18.10
-	10/52.50	10/28.75	10/40.00	10/41.67	10/30.00	10/24.76
0.05	10/62.50	10/38.75	10/51.67	10/58.33	10/30.00	10/20.48
0.10	10/70.00	10/42.50	10/55.00	10/46.67	10/37.50	10/21.43
0.20	10/72.50	10/46.25	10/60.00	10/60.00	10/38.33	10/23.81
K-Means - MFCC (Coef. 16)						
-	5/55.00	5/31.25	5/30.00	5/38.33	5/30.83	5/24.29
0.05	5/65.00	5/43.75	5/41.67	5/58.33	5/35.83	5/20.48
0.10	5/75.00	5/47.50	5/50.00	5/68.33	5/43.33	5/23.81
0.20	5/82.50	5/53.75	5/51.67	5/71.67	5/41.67	5/25.24
-	10/55.00	10/31.25	10/26.67	10/33.33	10/30.83	10/23.81
0.05	10/65.00	10/46.25	10/40.00	10/55.00	10/39.17	10/22.86
0.10	10/72.50	10/50.00	10/45.00	10/66.67	10/44.17	10/25.71
0.20	10/75.00	10/51.25	10/53.33	10/68.33	10/40.00	10/29.05

TAB. 8.4: K-Means aplicado aos coeficientes 13 ao 16.

K-Means - MFCC (Coef. 17)						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)
-	5/50.00	5/20.00	5/30.00	5/33.33	5/22.50	5/20.48
0.05	5/65.00	5/28.75	5/53.33	5/48.33	5/35.00	5/20.95
0.10	5/75.00	5/36.25	5/58.33	5/45.00	5/33.33	5/22.38
0.20	5/77.50	5/47.50	5/61.67	5/56.67	5/36.67	5/20.48
-	10/52.50	10/17.50	10/28.33	10/36.67	10/28.33	10/22.86
0.05	10/70.00	10/33.75	10/48.33	10/46.67	10/38.33	10/22.38
0.10	10/72.50	10/35.00	10/58.33	10/46.67	10/39.17	10/25.24
0.20	10/82.50	10/43.75	10/63.33	10/56.67	10/39.17	10/23.33
K-Means - MFCC (Coef. 18)						
-	5/37.50	5/28.75	5/21.67	5/26.67	5/32.50	5/17.14
0.05	5/70.00	5/50.00	5/50.00	5/35.00	5/38.33	5/17.62
0.10	5/67.50	5/57.50	5/50.00	5/50.00	5/41.67	5/16.67
0.20	5/70.00	5/48.75	5/53.33	5/46.67	5/45.83	5/21.90
-	10/42.50	10/28.75	10/33.33	10/33.33	10/35.00	10/20.48
0.05	10/67.50	10/46.25	10/53.33	10/38.33	10/35.83	10/22.38
0.10	10/70.00	10/55.00	10/53.33	10/53.33	10/43.33	10/21.43
0.20	10/75.00	10/60.00	10/58.33	10/55.00	10/52.50	10/29.05
K-Means - MFCC (Coef. 19)						
-	5/55.00	5/23.75	5/38.33	5/31.67	5/35.83	5/20.00
0.05	5/65.00	5/45.00	5/65.00	5/45.00	5/33.33	5/16.19
0.10	5/75.00	5/36.25	5/63.33	5/50.00	5/44.17	5/20.48
0.20	5/77.50	5/33.75	5/60.00	5/51.67	5/52.50	5/21.43
-	10/60.00	10/21.25	10/33.33	10/35.00	10/40.83	10/24.76
0.05	10/72.50	10/50.00	10/61.67	10/43.33	10/35.83	10/15.71
0.10	10/82.50	10/45.00	10/66.67	10/45.00	10/46.67	10/22.86
0.20	10/82.50	10/43.75	10/66.67	10/55.00	10/52.50	10/22.38
K-Means - MFCC (Coef. 20)						
-	5/40.00	5/36.25	5/41.67	5/46.67	5/39.17	5/22.86
0.05	5/72.50	5/36.25	5/48.33	5/55.00	5/41.67	5/20.48
0.10	5/72.50	5/42.50	5/63.33	5/46.67	5/46.67	5/22.38
0.20	5/77.50	5/46.25	5/70.00	5/53.33	5/49.17	5/24.76
-	10/52.50	10/46.25	10/46.67	10/46.67	10/37.50	10/23.81
0.05	10/77.50	10/35.00	10/50.00	10/58.33	10/40.00	10/22.86
0.10	10/77.50	10/46.25	10/68.33	10/50.00	10/47.50	10/25.24
0.20	10/77.50	10/51.25	10/70.00	10/56.67	10/47.50	10/26.19

TAB. 8.5: K-Means aplicado aos coeficientes 17 ao 20.

8.1.2 K-NN APLICADO A TODOS OS COEFICIENTES MEL-CEPESTRAIS

K-NN - MFCC (Coef. 1)						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)
-	1/65.00	1/68.75	1/61.67	1/78.33	1/65.83	1/60.00
0.05	1/70.00	1/41.25	1/20.00	1/48.33	1/28.33	1/25.71
0.10	3/62.50	1/51.25	1/45.00	1/60.00	1/35.83	1/28.57
0.20	1/57.50	1/52.50	1/60.00	1/68.33	1/55.00	1/35.24
K-NN - MFCC (Coef. 2)						
-	3/72.50	1/66.25	1/56.67	1/51.67	1/65.00	1/52.86
0.05	3/75.00	1/48.75	1/15.00	3/48.33	1/45.00	1/20.95
0.10	1/82.50	1/66.25	1/50.00	1/45.00	1/50.00	1/33.33
0.20	1/85.00	1/61.25	1/45.00	1/50.00	1/53.33	1/34.76
K-NN - MFCC (Coef. 3)						
-	1/70.00	1/68.75	1/51.67	1/80.00	1/53.33	1/54.76
0.05	1/52.50	3/28.75	1/41.67	1/45.00	1/29.17	1/26.67
0.10	1/67.50	1/36.25	1/48.33	1/58.33	1/37.50	1/30.95
0.20	1/65.00	1/57.50	1/51.67	1/76.67	1/35.00	1/48.10
K-NN - MFCC (Coef. 4)						
-	1/82.50	1/50.00	1/41.67	1/58.33	1/40.00	1/35.71
0.05	3/95.00	1/28.75	1/46.67	1/55.00	1/35.83	1/23.81
0.10	5/90.00	1/36.25	1/48.33	1/58.33	1/35.83	1/25.71
0.20	1/90.00	1/48.75	1/33.33	3/58.33	1/34.17	1/27.14
K-NN - MFCC (Coef. 5)						
-	1/65.00	1/43.75	1/38.33	1/36.67	1/39.17	1/31.43
0.05	1/55.00	1/26.25	3/38.33	1/38.33	1/26.67	1/16.67
0.10	1/67.50	1/38.75	1/43.33	1/43.33	1/25.00	1/19.05
0.20	3/75.00	1/45.00	1/55.00	1/36.67	1/21.67	1/23.81
K-NN - MFCC (Coef. 6)						
-	1/50.00	1/40.00	1/38.33	1/46.67	1/25.83	1/25.24
0.05	5/55.00	1/33.75	1/36.67	1/50.00	1/28.33	1/13.33
0.10	3/57.50	1/33.75	1/38.33	1/60.00	1/34.17	1/17.62
0.20	1/65.00	1/36.25	1/41.67	1/63.33	1/30.83	1/27.14
K-NN - MFCC (Coef. 7)						
-	1/60.00	1/46.25	1/38.33	1/73.33	1/36.67	1/32.86
0.05	1/52.50	1/43.75	3/30.00	1/55.00	1/37.50	1/17.62
0.10	1/67.50	1/52.50	1/36.67	1/63.33	1/43.33	1/23.81
0.20	1/65.00	3/42.50	1/46.67	1/61.67	1/32.50	1/27.62
K-NN - MFCC (Coef. 8)						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)

TAB. 8.6: KNN aplicado aos coeficientes mel-cepestrais 01 ao 08.

K-NN - MFCC (Coef. 9)						
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)
-	1/32.50	1/26.25	1/28.33	1/36.67	1/30.00	1/16.67
0.05	1/47.50	1/48.75	1/43.33	3/40.00	1/30.83	1/11.90
0.10	1/47.50	1/40.00	1/50.00	3/50.00	1/28.33	1/12.86
0.20	1/50.00	1/36.25	1/50.00	3/55.00	1/33.33	1/19.05
K-NN - MFCC (Coef. 10)						
-	1/65.00	1/38.75	1/48.33	1/45.00	1/43.33	1/27.62
0.05	3/55.00	1/38.75	1/26.67	1/38.33	1/25.83	1/20.00
0.10	3/67.50	1/38.75	1/30.00	1/41.67	1/38.33	1/22.38
0.20	1/77.50	1/33.75	1/41.67	1/45.00	1/45.83	1/19.52
K-NN - MFCC (Coef. 11)						
-	1/57.50	1/32.50	1/38.33	1/53.33	1/34.17	1/25.24
0.05	3/52.50	1/23.75	1/43.33	1/25.00	1/24.17	1/13.33
0.10	1/57.50	1/22.50	1/46.67	1/46.67	1/31.67	1/13.81
0.20	1/67.50	1/43.75	1/55.00	3/50.00	1/33.33	1/25.71
K-NN - MFCC (Coef. 12)						
-	1/60.00	1/28.75	1/30.00	1/38.33	1/29.17	1/20.48
0.05	3/55.00	1/30.00	1/41.67	1/48.33	1/27.50	1/13.33
0.10	1/75.00	1/31.25	3/36.67	1/55.00	1/20.83	1/17.62
0.20	3/67.50	1/40.00	1/40.00	1/50.00	1/24.17	1/16.67
K-NN - MFCC (Coef. 13)						
-	1/50.00	1/43.75	1/41.67	1/46.67	1/40.83	1/25.71
0.05	3/50.00	1/40.00	1/45.00	3/43.33	1/25.00	1/13.81
0.10	5/62.50	1/33.75	1/53.33	1/46.67	1/32.50	1/20.95
0.20	3/60.00	1/48.75	1/56.67	1/55.00	1/42.50	1/23.81
K-NN - MFCC (Coef. 14)						
-	1/47.50	1/32.50	1/30.00	1/48.33	1/30.00	1/21.43
0.05	3/55.00	1/28.75	1/38.33	1/50.00	1/17.50	1/12.38
0.10	1/47.50	1/37.50	1/40.00	1/46.67	1/25.00	1/17.62
0.20	1/65.00	1/31.25	1/45.00	1/46.67	1/34.17	1/20.48
K-NN - MFCC (Coef. 15)						
-	1/40.00	1/17.50	1/36.67	3/36.67	1/25.83	1/18.57
0.05	1/55.00	1/28.75	1/30.00	1/31.67	1/18.33	1/15.24
0.10	1/65.00	1/26.25	1/35.00	3/35.00	1/15.83	1/14.29
0.20	1/65.00	1/32.50	1/33.33	1/41.67	1/20.83	1/13.81
K-NN - MFCC (Coef. 16)						
-	1/32.50	1/17.50	1/25.00	1/23.33	1/22.50	1/13.33
0.05	3/42.50	1/27.50	1/35.00	1/38.33	1/37.50	1/15.71
0.10	3/67.50	1/30.00	1/41.67	1/53.33	1/35.00	1/16.19
0.20	1/62.50	1/35.00	1/38.33	1/45.00	1/26.67	1/10.95

TAB. 8.7: KNN aplicado aos coeficientes mel-cepestrais 09 ao 16.

K-NN - MFCC (Coef. 17)						
-	1/40.00	1/12.50	1/33.33	1/25.00	1/23.33	1/11.43
	COTURNO	TENIS	SAPATO	UMPORCALCADO	MESMOCALCADO	IGNORECALCADO
P. Fisher	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)	K/(%)
0.05	3/55.00	1/23.75	1/41.67	1/35.00	1/26.67	1/13.81
0.10	1/52.50	1/20.00	5/35.00	3/28.33	1/35.00	1/19.52
0.20	5/62.50	1/30.00	1/45.00	1/45.00	1/30.00	1/15.71
K-NN - MFCC (Coef. 18)						
-	3/30.00	1/18.75	1/16.67	1/20.00	1/22.50	1/12.86
0.05	1/55.00	1/31.25	1/41.67	1/35.00	1/27.50	1/14.29
0.10	1/65.00	1/37.50	1/36.67	1/28.33	1/32.50	1/16.19
0.20	5/60.00	1/33.75	1/33.33	1/41.67	1/28.33	1/12.86
K-NN - MFCC (Coef. 19)						
-	1/55.00	1/23.75	1/18.33	1/30.00	1/24.17	1/17.62
0.05	5/65.00	1/28.75	1/40.00	1/35.00	1/29.17	1/14.76
0.10	1/70.00	1/26.25	1/41.67	1/25.00	1/30.83	1/14.76
0.20	5/77.50	1/25.00	1/56.67	1/31.67	1/27.50	1/14.76
K-NN - MFCC (Coef. 20)						
-	1/42.50	1/18.75	1/30.00	1/38.33	1/32.50	1/19.05
0.05	1/57.50	1/26.25	1/38.33	1/41.67	1/32.50	1/13.81
0.10	1/70.00	1/33.75	1/50.00	1/45.00	1/36.67	1/16.67
0.20	1/65.00	1/32.50	1/45.00	1/45.00	1/28.33	1/20.00

TAB. 8.8: KNN aplicado aos coeficientes mel-cepestrais 17 ao 20.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)