

*Contribuições para o desenvolvimento  
de um sistema para auxílio a deficientes  
auditivos*

**Bruno do Amaral**

Setembro / 2017

Dissertação de Mestrado em Ciência da  
Computação

**CONTRIBUIÇÕES PARA O DESENVOLVIMENTO DE UM SISTEMA PARA  
AUXÍLIO A DEFICIENTES AUDITIVOS**

Dissertação apresentado à Banca Examinadora para defesa  
no curso de Mestrado em Ciência da Computação da  
Faculdade Campo Limpo Paulista.

Campo Limpo Paulista, 18 de Setembro de 2017

Bruno do Amaral

Prof. Dr. José Hiroki Saito (Orientador)

## FICHA CATALOGRÁFICA

Dados Internacionais de Catalogação na Publicação (CIP)

Câmara Brasileira do Livro, São Paulo, Brasil.

Amaral, Bruno do

Contribuições para o desenvolvimento de um sistema para auxílio a deficientes auditivos / Bruno do Amaral. Campo Limpo Paulista, SP: FACCAMP, 2017.

Orientador: Prof<sup>o</sup>. Dr. José Hiroki Saito.

Dissertação (Programa de Mestrado em Ciência da Computação) – Faculdade Campo Limpo Paulista – FACCAMP.

1. Tecnologias assistivas. 2. Deficientes auditivos. 3. Protótipo. 4. Qualidade de vida. I. Saito, José Hiroki. II. Campo Limpo Paulista. III. Título.

CDD-005.43

## **Agradecimentos**

Agradeço a Deus por todas as coisas boas que me concedeu.

Ao prof. Dr. José Hiroki Saito pela orientação, paciência e todo o apoio dado para o desenvolvimento deste trabalho.

Aos meus pais que sempre me apoiaram e estiveram presentes na minha formação.

A amiga Nayari, por todo o apoio e incentivo durante todos esses anos.

Agradeço a todos que de alguma forma me apoiaram e contribuíram para a execução deste trabalho.

## Resumo

*Buscando contribuir com o desenvolvimento de tecnologias assistivas para o auxílio na execução das atividades diárias dos deficientes auditivos, a presente proposta consiste no delineamento de um protótipo composto das partes que realizam a captação e processamento dos sinais por meio de diversos microfones e sensores, sistema de rádio para a transmissão de dados para o computador realizar o processamento desses sinais com base na Análise de Componentes Independentes, aplicado à Separação Cega de Fontes, e por conseguinte o reconhecimento das expressões de fala; e da parte de sensibilização que se aplica à sinalização do deficiente por indicação luminosa, textos e vibração por meio de um aplicativo para smartphone. Por sua vez, o deficiente auditivo poderá utilizar a função do aplicativo de reconhecimento de texto com a reprodução da fala, possibilitando uma comunicação com ouvintes. Para a validação do sistema na parte de captação e processamento de sinais, foram desenvolvidos experimentos envolvendo dois microfones; e experimentos de misturas simuladas de uma quantidade maior de sinais, seguida da separação de fontes. Para a parte de sensibilização foi desenvolvido um protótipo de hardware e de software em smartphone, em que o deficiente tem a possibilidade de receber informações e realizar interações. Assim, a intenção desta pesquisa é contribuir para o desenvolvimento de uma ferramenta de auxílio para os deficientes auditivos que possa proporcionar melhoria na sua qualidade de vida, independência e inclusão social.*

**Palavras-chave:** tecnologias assistivas, deficientes auditivos, protótipo, qualidade de vida.

## Abstract

*Aiming to contribute to the development of assistive technologies to aid in the daily activities of the hearing impaired, the present proposal consists of the design of a prototype composed of a section that carry out the capture and processing of the signals through several microphones and sensors, radio system for the transmission of data to the computer to perform the processing of these signals based on the Independent Component Analysis applied to the Blind Source Separation, then recognition of the speech expressions. The other section is the sensitization part that informs the handicapped person by luminous indication, texts and vibration through a smartphone applicative. In turn, the hearing impaired can use the function of the text recognition application with the corresponding computerized speech reproduction, enabling communication with listeners. For the validation of the first section of the system, it was developed experiments with two microphones; and simulated signal mixture experiments, with a more quantity of signals, followed by source separations. For the sensitization section, it was developed a hardware prototype, and a smartphone software, intended to be used by the impaired person to receive information and interaction. Thus, the intention of this research is to contribute to the development of a technological tool to aid the hearing-impaired persons, in improvement of their quality of life, independence and social inclusion.*

**Keywords:** assistive technologies, hearing impaired, prototype, quality of life.

## SUMÁRIO

CAPÍTULO 1 INTRODUÇÃO .....	8
1.1. Introdução .....	8
1.2. Contexto e Motivação .....	9
1.3. Objetivos.....	10
1.4. Contribuições .....	10
1.5. Estrutura do Trabalho.....	11
1.6. Considerações Finais.....	12
CAPÍTULO 2 INFORMAÇÕES GERAIS.....	13
2.1. Sistema sensorial da audição.....	13
2.1.1. Sistema Auditivo Humano .....	13
2.2. Deficiência auditiva e as tecnologias de assistência.....	16
2.2.1. Sistemas de Assistência a Pessoas com Deficiência Auditiva.....	16
2.3. Considerações Finais.....	23
CAPÍTULO 3 REVISÃO BIBLIOGRÁFICA .....	24
3.1. Considerações Iniciais.....	24
3.2. Trabalhos relativos à tecnologia assistiva.....	25
3.3. Trabalhos relativos às técnicas de processamento de sinais.....	30
3.4. Considerações Finais.....	35
CAPÍTULO 4 MATERIAIS E MÉTODOS .....	36
4.1. Análise de Componentes Independentes.....	36
4.1.1. Restrições .....	38
4.1.2. Ambiguidades .....	39
4.2. Propriedades fundamentais usadas em ICA.....	39
4.2.1. Independência estatística .....	40

4.2.2. Misturas não Gaussianas e o Teorema do Limite Central .....	40
4.2.3. Medidas para a não gaussianidade .....	41
4.2.3.1. Curtose.....	41
4.2.3.2. Negentropia.....	41
4.3. Métodos para Separação das Fontes .....	42
4.3.1. Pré-processamento .....	42
4.3.2. Algoritmo FastICA .....	44
4.4. Separação cega para misturas convolutivas .....	45
4.4.1. Algoritmo para misturas convolutivas.....	46
4.5. Reconhecimento de Fala.....	47
4.5.1. Pré-processamento .....	48
4.5.2. Análise Espectral .....	49
4.5.3. Extração de Parâmetros.....	49
4.5.4. Reconhecimento.....	51
4.6. Descrições das tecnologias integradas.....	52
4.6.1. Sistema de comunicação por radiofrequência.....	52
4.6.2. Módulo Bluetooth .....	54
4.6.3. Módulo sensor de som .....	56
4.6.4. Plataforma de desenvolvimento Arduino .....	57
4.7. Sistemas Operacionais e Simuladores.....	63
4.7.1. App Inventor .....	63
4.7.2. Ambiente de desenvolvimento integrado ao arduino.....	64
4.7.3. Sistema Operacional Android.....	64
4.8. Considerações finais.....	65
CAPÍTULO 5 DELINEAMENTO DE UM SISTEMA DE AUXÍLIO A DEFICIENTES AUDITIVOS.....	67

5.1. Desenvolvimento do SADA.....	67
5.2. Caracterização do Sistema Físico SADA Ouvitronic-1.....	70
5.3. Ilustração de Uso da Plataforma Arduino para Transmissão e Recepção de RF.....	72
5.4. Caracterização do Software de Comunicação no SADA Ouvitronic-1.....	73
5.4.1. Aplicativo Ouvitronic-1 em Android para <i>smartphone</i> .....	73
5.4.2. Sistema de comunicação Arduino e Aplicativo.....	78
5.5. Considerações Finais.....	81
CAPÍTULO 6 RESULTADOS E VALIDAÇÃO.....	82
6.1. Descrição do Pré-Protótipo de Hardware.....	82
6.2. Descrição do Software Desenvolvido.....	84
6.3. Experimentos sobre o processamento de sinais para misturas instantâneas e reconhecimento de fala.....	86
6.3.1. Experimento I.....	87
6.3.2. Experimento II.....	89
6.3.3. Experimento III.....	90
6.3.4. Experimento IV.....	92
6.4. Experimentos com captação de misturas convolutivas e validação do BSS.....	94
6.4.1. Experimento I.....	96
6.4.2. Experimento II.....	96
6.4.3. Experimento III.....	97
6.5. Considerações Finais.....	98
CAPÍTULO 7 CONCLUSÕES E PROPOSIÇÕES DE TRABALHOS FUTUROS.....	99
7.1. Conclusões.....	99
7.2. Propostas para trabalhos futuros.....	100
REFERÊNCIAS BIBLIOGRÁFICAS.....	102

## GLOSSÁRIO

AASI- Aparelhos de Amplificação Sonora Individual

A/D- Analógico/Digital

ADC- Analog-Digital Converter (Conversor Analógico-Digital)

ANATEL- Agência Nacional de Telecomunicações

AO- *Analogic Output* (Saída Analógica)

API- *Application Programming Interface* (Interface de Programação de Aplicativos)

APP – Application -*software* aplicativo

ASK- *Amplitude Shift Keying* (Modulação por Chaveamento de Amplitude)

ASR- *Automatic Speech Recognition* (Reconhecimento Automático de Fala)

BSS- *Blind Source Separation* (Separação Cega das Fontes)

BTE- *Behind The Ear* (Atrás do ouvido)

CIC- *Completely In Canal* (Completamente no Canal)

CPU- *Central Processing Unit* (Unidade Central de Processamento)

DA- Deficiente Auditivo

DO- *Digital Output* (Saída Digital)

ECG- Eletrocardiograma

EEG- Eletroencefalograma

EEPROM- *Electrically-Erasable Programmable Read-Only Memory* (Memória Somente de Leitura Programável e Apagável Eletronicamente)

FDP- Função de distribuição de probabilidade

FHSS- *Frequency Hopping Spread Spectrum* (Frequência de Espelhamento Espectral)

FIR- *Finite Impulse Response* (Resposta de Impulso Finita)

GND- *Graduated Neutral Density Filter* (Filtro Graduado de Densidade Neutra)

HMM- *Hidden Markov Models* (Modelo Oculto de Markov)

IBGE- Instituto Brasileiro de Geografia e Estatística

I/O- *Input/Output* (Entrada/Saída)

ICA- *Independent Component Analysis* (Análise de Componentes Independentes)

ICSP- *In Circuit Serial Programming* (Programação Serial em Circuito)

IDE- *Integrated Development Environment* (Ambiente de Desenvolvimento Integrado)

ISM- *Industrial, Scientific and Medical* ( Industrial, Científica e Médica)

ITC- *In The Canal* ( Dentro do Canal)

ITE- *In The Ear* ( No ouvido)

LED- *Light Emitting Diode* (Diodo Emissor de Luz)

LIBRAS- Língua Brasileira de Sinais

Mel- *Melody* (melodia)

MFCC- *Mel-Frequency Cepstral Coefficients* (Coeficientes Cepstrais de Frequência-Mel)

MIT- *Massachusetts Institute of Technology* (Instituto de Tecnologia de *Massachusetts*)

MSE- *Mean Squared Error* (Erro Médio Quadrático)

PCA- *Principal Component Analysis* (Análise de Componentes Principais)

PDF- *Probability Density Function* (Função de Densidade de Probabilidade)

PDS- Processamento Digital de Sinais

PP- *Projection Pursuit* (Busca de Projeções)

RF- Radiofrequência

RIC- *Receiver In Canal* ( Receptor no Canal)

RNA- Redes Neurais

RTC- *Real Time Counter* (Contador de Tempo Real)

PWM- *Pulse Width Modulation* (Modulação por largura de pulso)

SADA- Sistema de Auxílio a Deficientes Auditivos

SDK- *Software Development Kit* (Kit de Desenvolvimento de Software)

SEDH- Secretaria Especial de Direitos Humanos

SIG- *Special Interest Group* (Grupo de Interesse Especial)

SMS- *Short Message Service* (Serviço de Mensagens Curtas)

SNC- Sistema Nervoso Central

SNP- Sistema Nervoso Periférico

SO- Sistema Operacional

SPI- *Serial Peripheral Interface*

SPP- *Serial Port Protocol* (Protocolo de Porta Serial)

SRAM- *Static Random Access Memory* (Memória de Acesso aleatório Estática)

SRST- *Speech Recognition and Synthesis Tool* (Reconhecimento de voz e ferramenta de síntese)

TA- Tecnologia Assistiva

TDD- *Telecommunication Device for the Deaf* ( Dispositivo de Telecomunicação para Surdo)

TDF- Transformada Discreta de Fourier

TI-Tecnologia da Informação

TTL- *Transistor-Transistor Logic* (Lógica de Transistor-Transistor)

TTY- *Teletypewriter* (Teclado-Teletipo)

TWI- *Two Wire Interface* (Interface por dois fios)

UART- *Universal Asynchronous Receiver/Transmitter* (Transmissor/Receptor Assíncrono Universal)

ULA- Unidade Lógica Aritmética

USB- *Universal Serial Bus* (Barramento Serial Universal)

USART- *Universal Synchronous/ Asynchronous Receiver/ Transmitter* (Transmissor/Receptor Síncrono/Assíncrono Universal)

VCC- Voltagem em Corrente Contínua

W3C- *World Wide Web Consortium* (Consórcio Internacional de Padrões para a Internet)

WPAN- *Wireless Personal Area Nertwork* (Rede Sem Fio de Área Pessoal)

ZCR- *Zero Crossing Rate* (Taxa de Cruzamento por Zero)

## LISTA DE FIGURAS

Figura 1. Símbolo Internacional de Surdez [DETRAN-PE 2016].	9
Figura 2. Estrutura do ouvido humano (E – externo, M – médio e I – Interno) [Fonseca 2002].	14
Figura 3. Exemplo de um sistema SADA com dispositivo de alerta. Fonte [Oliveira 2015].	18
Figura 4. Modelos de AASI. Fonte [Nicolau 2010].	19
Figura 5. Campanhas auxiliares para telefone. Fonte [Lojamatel 2017].	20
Figura 6. Relógio com dispositivo vibratório. Fonte [PTassistech 2017].	20
Figura 7. TTY.	21
Figura 8. Avatar do <i>App</i> ProDeaf [ProDeaf 2017].	22
Figura 9. Avatar do <i>App</i> Hand Talk [Hand Talk 2017].	22
Figura 10. Tela do <i>App</i> Librazuka [Librazuka 2017].	22
Figura 11. Modelo de funcionamento para estimação das fontes.	38
Figura 12. (A) Distribuição conjunta dos componentes independentes $S_1$ e $S_2$ com distribuição uniforme;	43
Figura 13. (A) Distribuição conjunta das misturas branqueadas uniformemente distribuídas; (B) Distribuição da estimativa dos sinais fonte. Adaptada de Ferreira [Ferreira 2002].	43
Figura 14. Diagrama de blocos de um sistema de análise do sinal de fala, com a obtenção dos parâmetros para o reconhecimento. Adaptada de Nunes [Nunes 1996].	48
Figura 15. Efeito de <i>Aliasing</i> [Barreto 2015].	48
Figura 16. Exemplo de um sinal discretizado.	49
Figura 17. Divisão em quadros do sinal de voz.	49
Figura 18. Diagrama representativo do módulo de um reconhecimento de fala.	52
Figura 19. Diagrama de blocos TX.	53
Figura 20. Módulo transmissor TX.	53
Figura 21. Diagrama de blocos RX.	54

Figura 22. Módulo receptor RX. ....	54
Figura 23. Conexão <i>Bluetooth</i> ponto a ponto e ponto a multiponto.....	55
Figura 24. Módulo <i>Bluetooth</i> HC-05.....	56
Figura 25. Módulo sensor de som. ....	57
Figura 26. Diagrama de blocos do processador AVR. ....	59
Figura 27. Diagrama de blocos do microcontrolador ATmega328. ....	60
Figura 28. Diagrama de pinos do microcontrolador <i>ATmega 328</i> , em empacotamento <i>dual-in-line</i> de 28 pinos (detalhes sobre os pinos em [Atmel 2017]). ....	60
Figura 29. Plataforma de desenvolvimento Arduino. ....	61
Figura 30. Ilustração do <i>Arduino Uno</i> [Blum 2013]. ....	61
Figura 31. Pinos <i>ATmega328</i> correspondentes aos pinos do <i>Arduino Uno</i> . Adaptação [Uno 2016]. ....	62
Figura 32. Programação gráfica no App Inventor 2. ....	64
Figura 33. Diagrama de um Sistema de Auxílio a Deficientes Auditivos (SADA) proposto. .	68
Figura 34. Diagrama ilustrativo do protótipo de SADA Ouvitronic-1, onde na metade esquerda tem-se a parte de captação e processamento e na metade à direita, a parte de sensibilização e interação do DA.....	69
Figura 35. Sistema físico de captação e transmissão de sinais.....	71
Figura 36. Sistema físico para recepção dos dados RF pelo computador central. ....	71
Figura 37. Sistema físico para transmissão de dados do computador central para o celular, por meio do protocolo <i>Bluetooth</i> . ....	71
Figura 38. Diagrama de blocos do sistema SADA Ouvitronic-1 com a disposição dos módulos de transmissão e recepção de RF e computador central para processamento e transmissão de dados para o DA. ....	72
Figura 39. Circuito transmissor RF (módulo TX) do protótipo. ....	73
Figura 40. Circuito receptor RF (módulo RX) e módulo <i>Bluetooth (HC-05)</i> do protótipo. ....	73
Figura 41. Fluxograma do aplicativo.....	75
Figura 42. Fluxograma para o transmissor RF. ....	78

Figura 43. Fluxograma para o receptor RF.....	79
Figura 44. Diagrama ilustrativo da parte do <i>hardware</i> considerada no pré-protótipo. ....	82
Figura 45. PARTE 2 do pré-protótipo do SADA Ouvitronic.....	83
Figura 46. Tela de início do aplicativo OuviTronic. ....	84
Figura 47. Telas do aplicativo: Cadastro de módulos, à direita (a) e Cadastro de palavras, à esquerda (b). ....	85
Figura 48. Tela de conversão de texto em fala. ....	86
Figura 49. Os três sinais, numerados de 1 a 3. Linha superior: sinais originais e linha inferior: sinais das misturas. ....	88
Figura 50. Sinais estimados pelo algoritmo <i>FastICA</i> , obtido pelo Matlab. ....	88
Figura 51 Gráfico da porcentagem de acertos do Experimento II.....	89
Figura 52. Gráfico percentual de acertos variando o número de misturas, Experimento III....	91
Figura 53. Gráfico percentual médio de acertos para cada texto, Experimento III.....	92
Figura 54. Gráfico percentual de acertos variando o número de misturas, experimento IV. ....	93
Figura 55. Gráfico percentual médio de acertos para cada texto, experimento IV. ....	94
Figura 56. Diagrama de cenários entre sensores e as fontes de origem de som, para misturas convolutivas.....	95
Figura 57. Misturas dos áudios obtidos no experimento I.....	96
Figura 58. (a) áudio de voz e (b) áudio musical. ....	96
Figura 59. Gráficos para misturas convolutivas entre dois áudios de voz. ....	97
Figura 60. (a) áudio de voz masculina “contando de 1 a 10” e (b) áudio de voz feminina “ditando de A até L do alfabeto”.....	97
Figura 61. Misturas dos áudios utilizados no experimento III. ....	97
Figura 62. (a) lado esquerdo, áudio de voz, e (b) lado direito, áudio de ruído de liquidificador. ....	98

## LISTA DE TABELAS

Tabela 1. Comparativo entre os diversos aplicativos citados.....	23
Tabela 2. Trabalhos relativos às tecnologias assistivas.....	29
Tabela 3. Trabalhos relativos às técnicas de processamento de sinais.....	34
Tabela 4. Frases com vozes distintas usadas no Experimento I.....	87
Tabela 5. Experimento II.....	89
Tabela 6. Quantidade de palavras para cada texto referente ao número de amostras, experimento III.....	90
Tabela 7. Taxas de acerto dos sinais extraídos e reconhecidos em relação ao aumento na quantidade de misturas, experimento III.....	91
Tabela 8. Quantidade de palavras para cada texto referente ao número de amostras, experimento IV.....	92
Tabela 9. Taxas de acerto dos sinais extraídos e reconhecidos em relação ao aumento na quantidade de misturas, Experimento IV.....	93
Tabela 10. Distância entre sensores e as fontes de origem de som, para misturas convolutivas.....	95

O presente capítulo refere-se à introdução ao texto do projeto de um protótipo de sistema para auxílio a deficientes auditivos, Dissertação de Mestrado. Na Seção 1.1 é apresentada uma introdução geral e é seguida da Seção 1.2, onde são descritas as motivações para o trabalho. Na Seção 1.3 são apresentados os principais objetivos para a proposta de projeto e na Seção 1.4, as principais contribuições. Na Seção 1.5 apresenta-se a descrição dos capítulos. Finalmente, a Seção 1.6 aborda as considerações finais.

Para facilitar a descrição do texto, denotamos doravante **Deficiente Auditivo**, simplesmente pela sigla **DA**, na maioria das ocorrências em que incluem deficientes auditivos e surdos.

### 1.1. Introdução

O número de pesquisas e projetos na área de Tecnologia Assistiva (TA) vem se tornando cada vez maior, buscando melhorias na qualidade de vida, possibilitando independência e inclusão social, por meio da ampliação da comunicação, habilidades, aprendizado, mobilidade e trabalho às pessoas portadoras de deficiência.

Neste contexto, a Tecnologia Assistiva é apresentada, pela Secretaria Especial de Direitos Humanos do Governo Federal (SEDH), como sendo um conjunto de características interdisciplinares que formam um conjunto de produtos, serviços, recursos, tecnologias entre outros, de apoio às pessoas com deficiências ou aqueles que carecem de apoio, visando proporcionar autonomia, qualidade de vida e inclusão social. O suporte de recursos tecnológicos dessa natureza faz-se necessário para milhões de pessoas, sendo a maior parte desse contingente carente economicamente, o que cria uma barreira ao acesso dos bens necessários à melhoria na qualidade de vida dessas pessoas [SEDH 2009].

Com esse intuito, diversas pesquisas sobre a surdez têm ocorrido até o presente, nos quais são estudadas as características fisiológicas, os hábitos, a cultura e a linguagem das pessoas surdas [Danesi 2007]. O Censo Demográfico do Instituto Brasileiro de Geografia e Estatística (IBGE) de 2010 constatou que havia aproximadamente 9,7 milhões de pessoas, com algum tipo de perda auditiva, das quais 2 milhões, completamente surdas.

Por outro lado, diversos campos de pesquisa tiveram avanços significativos até o presente, como em processamento de sinais, onde os sinais elétricos resultantes de falas podem ser processados com diversas finalidades, dentre as quais, a filtragem de fala de um indivíduo e separação de falas individuais, numa mistura de sinais, conhecido como Separação Cega de Fontes (em inglês, *Blind Source Separation-BSS*) [Comon 2010]. Outro campo de pesquisa com avanço significativo é o de Reconhecimento Automático de Fala (em inglês, *Automatic Speech Recognition-ASR*) [Benesty 2008]. Em ASR, os sinais de falas individuais são processados e as palavras são reconhecidas, resultando num texto digitalizado da fala.

## 1.2. Contexto e motivação

A Lei 8.160 torna obrigatória a utilização em todos os locais públicos, procedimentos ou serviços para pessoas portadoras de deficiência auditiva, o símbolo internacional de surdez, conforme a Fig. 1 [Brasil 1991].



Figura 1. Símbolo Internacional de Surdez [DETRAN-PE 2016].

Além disso, segundo o Decreto 5.296/04 (que regulamenta as Leis 10.048 e 10.098, ambos do ano 2000) [Brasil 2004], é previsto o atendimento prioritário aos portadores de deficiência auditiva por meio de intérprete de Libras (Língua Brasileira de Sinais), guias-intérpretes ou pessoas capacitadas para esse tipo de atendimento.

Apesar das providências citadas, os DAs ainda enfrentam muitas dificuldades na execução de tarefas do dia a dia, desde ir a um estabelecimento comercial até se socializarem em um ambiente de trabalho. Dentre suas principais queixas é a falta de intérprete de Libras em locais como bancos, consultórios, hospitais, supermercados entre outros. Os DAs, devido à predominância da comunicação oral, sofrem de problemas sociais e emocionais.

Um dos grandes problemas relatados pelos DAs é a grande dificuldade de encontrar um emprego, tendo sua oportunidade de mostrar sua capacidade e talento negado pelo preconceito ainda disseminado no país. E quando conseguem um emprego, acabam tendo grandes dificuldades em se socializar e compreender os processos, normas e dinâmicas do trabalho. A situação tem um agravante quando o sujeito é o único DA da corporação, na qual acaba sendo vítima de sua deficiência, colocado a um patamar inferior aos demais trabalhadores [Marin 2006].

Um outro exemplo de problema é a dependência constante de DAs, de acompanhantes em consultas médicas. Essa dependência muitas vezes limita o diagnóstico e o tratamento, em função do constrangimento a que são submetidos, além da dificuldade no diálogo com o médico, na falta de domínio da língua oral ou na ausência de um interlocutor.

### **1.3. Objetivos**

Com a intenção de contribuir para a melhoria do dia a dia de um DA, a presente pesquisa consiste no esboço de um protótipo de sistema envolvendo a captação de sinais de ambientes que sejam de interesse para os deficientes auditivos; o processamento dos sinais e reconhecimento de expressões de fala, e o uso de dispositivos de sinalização que convertem sinais sonoros em outros tipos de sinais, como textos e vibrações para possibilitar o auxílio ao DA. Como objetivos específicos do trabalho podem ser enumerados os seguintes itens:

- 1) Aplicação de técnicas de *hardware* para sensoriamento de sinais, transmissão em radiofrequência, e processamento;
- 2) Aplicação de técnicas de programação em dispositivos celulares para o desenvolvimento de aplicativos em sistema operacional *Android*;
- 3) Estudo e aplicação de técnicas de processamento de sinais para a separação cega de fontes;
- 4) Aplicação de técnicas de processamento de sinais para o reconhecimento de falas; e
- 5) Esboço de protótipo de sistema para auxílio a deficientes auditivos.

### **1.4. Contribuições**

Dentre as principais contribuições do presente trabalho citamos:

- 1) esboço de um protótipo de sistema envolvendo a captação de sinais de ambientes ruidosos usando um ou mais microfones, e a separação do sinal de fala da mistura desses sinais;

2) esboço de um protótipo de *software* em smartphone para auxílio ao DA, em função do protótipo do item (1), envolvendo a interação do DA com as informações que são recebidas e transmitidas; e

3) realização de experimentos que validam a captação e separação do sinal de fala de ambiente de múltiplas fontes e ruidoso.

## 1.5. Estrutura do trabalho

O presente trabalho está dividido em capítulos, sendo os seguintes, além do presente capítulo de introdução:

**Capítulo 2** Informações Gerais: onde serão apresentadas as pesquisas referentes às informações gerais sobre o sistema auditivo humano compreendendo sua anatomia e seu funcionamento.

**Capítulo 3** Revisão Bibliográfica: onde serão descritos os trabalhos sobre as técnicas de ICA e BSS, bem como uma ampla abordagem de trabalhos com o enfoque no estudo e desenvolvimento de sistemas de assistência a deficientes auditivos.

**Capítulo 4** Materiais e Métodos: neste capítulo é apresentada a metodologia que será utilizada para o desenvolvimento da dissertação, além da descrição das bases de dados utilizadas. Sendo assim, são descritas a definição do modelo estatístico, ambiguidades e restrições para execução das estimações do modelo da Análise de Componentes Independentes em Separação Cega das Fontes; noções sobre o reconhecimento de fala; e tecnologias para o desenvolvimento de sistemas de *hardware* e *software* pertinentes.

**Capítulo 5** Delineamento de um Sistema de Auxílio a Deficientes Auditivos: neste capítulo será apresentado o delineamento de um Sistema de Auxílio a Deficientes Auditivos, SADA, bem como a caracterização do sistema, apontando os tópicos investigados.

**Capítulo 6** Resultados e Validação: neste capítulo são descritos os experimentos em processamento de sinais para misturas instantâneas e reconhecimento de fala e a descrição e os resultados para os experimentos com misturas convolutivas. Por fim, são detalhados os experimentos práticos de captação de sinais com dois microfones, para efeitos de validação do sistema.

**Capítulo 7** Conclusões e Proposições de Trabalhos Futuros: Neste capítulo são apresentadas as conclusões finais deste trabalho e proposições para trabalhos futuros.

### **1.6. Considerações finais**

No presente capítulo foram descritos o contexto, as motivações para o trabalho proposto, objetivos gerais e específicos, bem como a estruturação dos capítulos. No capítulo seguinte serão apresentadas algumas informações gerais, como o sistema auditivo humano, as deficiências auditivas e os principais dispositivos e tecnologias de assistência aos DAs, atualmente disponíveis.

## **CAPÍTULO 2**

### **INFORMAÇÕES GERAIS**

---

Neste capítulo é apresentada uma descrição de informações gerais relativas à audição, bem como aos deficientes auditivos. Assim, na Seção 2.1 é apresentado o sistema sensorial da audição; a Seção 2.2 refere-se aos deficientes auditivos e tecnologias de assistência atualmente disponíveis; e finalmente, na Seção 2.3 apresenta-se as considerações finais.

#### **2.1. Sistema sensorial da audição**

O sistema nervoso é subdividido em dois sistemas, em que um deles é o Sistema Nervoso Periférico (SNP) que envia as informações para o Sistema Nervoso Central (SNC) pelos neurônios sensoriais (aférentes), e leva as informações do SNC para células específicas por meio dos neurônios eferentes. O outro sistema é o Sistema Nervoso Central, propriamente dito, que é responsável pelo processamento de informações [Silverthorn 2009].

Muitas das atividades do sistema nervoso se iniciam pela excitação dos receptores sensoriais, sejam eles receptores táteis, visuais, auditivos, entre outros. Os receptores funcionam como transdutores convertendo os estímulos em sinais intracelulares, e possuem a função de captar os estímulos do ambiente para serem recebidos pelo SNP, que os encaminha para o SNC, que possui a função de processar, interpretar, classificar, armazenar as informações e elaborar respostas [Guyton 2006].

##### **2.1.1. Sistema Auditivo Humano.**

O som nada mais é que oscilações (ondas) que se propagam por um meio material, como o ar, até atingirem o ouvido externo, produzindo a sensação sonora. O cérebro interpreta a frequência e a amplitude das ondas sonoras, pelo número de picos das ondas que passam em determinado ponto, a cada segundo, e pela entonação do som.

A entonação é diferenciada por ondas de baixa frequência para sons graves, e ondas de alta frequência para sons agudos. O sistema auditivo compreende os sons nas faixas de 20 a 20.000,00 oscilações por segundo (Hz), sendo os valores inferiores correspondentes a sons graves e os valores superiores a sons agudos. A sensação de altura do som depende da

sensibilidade auditiva de cada pessoa, sendo mensurada em decibéis (dB). Portanto, sons acima de 80 dB ou mais podem causar danos nos receptores sensoriais do ouvido resultando ao longo do tempo na perda auditiva.

O ouvido é formado por três partes: ouvido externo (E), ouvido médio (M) e ouvido interno (I). A Fig.2 consiste na ilustração da estrutura do ouvido humano, no qual se observa as três partes E, M e I, descritas a seguir:

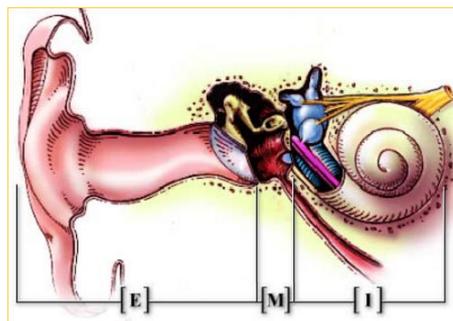


Figura 2. Estrutura do ouvido humano (E – externo, M – médio e I – Interno) [Fonseca 2002].

- **Ouvido externo** - consiste na aurícula (orelha) e o canal auditivo com seu fundo fechado por uma membrana fina denominada de tímpano. A membrana faz a ligação entre o ouvido externo e o médio. Tem por funções recolher e encaminhar a onda sonora até o tímpano, que irá produzir uma vibração (primeira transdução);
- **Ouvido médio** - é formado pela membrana timpânica, uma cavidade com ar conectada à faringe por meio da tuba auditiva. A tuba auditiva regula a pressão do ouvido médio, durante atividades como mastigação, bocejo e deglutição. Também, o ouvido médio é formada por três ossos denominados martelo, bigorna e estribo. Esse arranjo cria uma “alavanca” que amplifica a força de vibração, para diminuir perdas no caminho devido ao atrito, e o som é conduzido até o ouvido interno;
- **Ouvido interno** ou labirinto, também conhecido como cóclea, possui o formato em caracol sendo o responsável por receber as vibrações, e por meio de um líquido criam ondas que abrem e fecham os canais gerando sinais elétricos (segunda transdução). Conforme as ondas se movem no duto coclear, realizam um movimento, abrindo e

fechando células pilosas (“pequenos pelos”), gerando sinais elétricos (terceira transdução). A quarta e a quinta transdução são resultantes da ligação do neurotransmissor com os neurônios sensoriais auditivos que encaminha as informações codificadas do som ao SNC. Além dessas funções, o labirinto é o órgão do sentido de equilíbrio, pois informa ao cérebro a posição do corpo no espaço.

Segundo Silverthorn [Silverthorn 2009] a localização do som é um processo que envolve a resposta sensorial dos dois ouvidos (esquerdo e direito), que analisa a diferença de tempo da chegada do som em cada um dos ouvidos e executa um processamento complexo no SNC.

A deficiência auditiva no envelhecimento é devida a uma crescente degeneração sensorial do ouvido, incluindo células da cóclea e SNC. A deficiência auditiva também é causada por alguma lesão ou doença. Já a surdez é de origem congênita e não permite ouvir nenhum som.

Com o exame fonoaudiólogo chamado de audiometria tonal é possível identificar o tipo ou o grau de DA de um indivíduo, sendo a unidade de medida em decibéis dentre as seguintes categorias:

- Leve - entre 20 a 40 dB, não impede a aquisição normal da linguagem, mas com dificuldade para entender a fala e a incapacidade em ouvir alguns sons, como, por exemplo o tique-taque do relógio;
- Média - entre 40 a 70 dB, ouve somente palavras e sons com entonação forte. Nesse nível a comunicação já se torna bastante limitada;
- Severa - entre 70 a 90 dB, só consegue ouvir sons muito próximos e reconhecer palavras somente quando o som estiver próximo e alto;
- Profunda- com mais de 90 dB, nenhum som ou palavra é ouvido.

A perda de audição pode ser dividida em três tipos: a de condução, a sensório-neural e a central:

- **Perda auditiva de condução**, em que o som não consegue ser transmitido pelo ouvido externo ou médio, devido a obstrução do canal auditivo por cerume, infecção no líquido do ouvido médio e até traumas que impedem a vibração dos ossos no ouvido médio e pode ser corrigida por desobstrução do canal ou técnicas microcirúrgicas;

- **Perda de audição sensório-neural**, causada por danos em estruturas do ouvido interno, incluindo morte nas células ciliadas da cóclea, devida à idade, exposição a ruídos e enfermidades, podendo ser tratada com aparelhos auditivos;
- **Perda auditiva central**, ocasionada por danos nas vias neurais entre o ouvido e o SNC, tendo como uma das causas, acidente vascular cerebral, sendo uma perda auditiva incomum.

## **2.2. Deficiência Auditiva e as Tecnologias de Assistência**

As pessoas com DAs desenvolvem uma maior habilidade visual aproveitando todos os lados da visão, o que deve ser explorado no processo de aprendizagem. Muitos dos contingentes da comunidade surda têm a capacidade de interpretar a língua de sinais e a oralizada. O surdo oralizado é aquele que faz uso da leitura labial, comunicação oral propriamente dita em voz com sotaque característico e possuem a compreensão da linguagem textual da língua falada. Essas habilidades permitem que o surdo se aproxime da comunidade ouvinte ampliando seu universo de interação [Meirelles 2004].

Muitos esforços vêm sendo desenvolvidos para a difusão da Língua Brasileira de Sinais (Libras), que consiste na comunicação por meio da expressão de sinais com as mãos, entre os DAs e as demais pessoas da sociedade, ou mesmo entre os DAs, oficializada pela Lei 10.436 de 2002 [Brasil 2002]. Todavia, somente uma pequena parcela da população sabe se comunicar por Libras [Ossada & Rodrigues 2016], além dos DAs enfrentarem uma grande barreira no domínio das duas linguagens, a Libras e a linguagem falada ou escrita, como descreve Ferreira [Ferreira 2014].

A fim de proporcionar maior acessibilidade na execução das tarefas diárias, e principalmente ao acesso aos meios de comunicação *Web*, surgiram diversos sistemas de apoio aos DAs, descritos a seguir.

### **2.2.1. Sistemas de Assistência a Pessoas com Deficiência Auditiva**

Nas últimas décadas muitos esforços vêm sendo investidos no desenvolvimento de Tecnologias de Comunicação e Informação. Todo tipo de tecnologia desenvolvida com o objetivo de proporcionar uma melhor qualidade de vida, inclusão social e autonomia aos deficientes, no auxílio à execução de tarefas diárias como alimentação, leitura, locomoção, comunicação, é denominada de Tecnologia Assistiva [Bersch 2013].

O advento da Tecnologia da Informação (TI) vem trazendo grandes benefícios a pessoas deficientes, que se apoiam em inúmeros aplicativos de computadores e celulares *smartphones* que oferecem a essas pessoas a acessibilidade para algumas tarefas do dia a dia. Pode-se citar aplicativos que realizam a conversão de texto para voz, tradutores para Libras, leitura automática do celular com toques na tela *touch screen*, entre outros. Segundo Radabaugh [Radabaugh 1993], “para as pessoas sem deficiência a tecnologia torna as coisas mais fáceis. Para as pessoas com deficiência, a tecnologia torna as coisas possíveis”.

O acesso à informação pela TI por meio de conteúdos digitais, muitas vezes, não são ferramentas acessíveis para os DAs, segundo Torres, Mazzoni e Alves [Torres 2002], que destacam:

- Materiais audiovisuais devem ser legendados, tanto com legendas em texto quanto em Libras;
- Opções para controle de volume, no *hardware* disponibilizado pela biblioteca para utilização desses usuários;
- Acesso visual à informação sonora (transcrição para equivalentes textuais ou pictóricos) e uma sinalização visual para os eventos do sistema em utilização (envio e recepção de mensagens na *Internet*);
- Serviço de transcrição de documentos digitais orais para textuais.

Os Sistemas de Auxílio a Deficientes Auditivos (SADA) desenvolvidos dentro do campo de TA, são sistemas compostos de sensores para captação dos sinais, sistemas computacionais remotos ou locais que realizam o processamento dos sinais recebidos pelos sensores e atualizam as informações para o usuário. Essas informações são fornecidas por meio de comandos que atuam sobre dispositivos, que podem ser indicadores luminosos, por vibração, sistemas de reconhecimento de texto que transformam um texto escrito em voz, sistemas de legenda, sistemas de conversão de texto para Libras, entre outros.

A Fig.3 ilustra um sistema SADA, onde podemos observar um dispositivo que representa um “ouvido eletrônico”, que irá captar os alertas sonoros provindos de uma fonte de som como um bebê em prantos ou, então o toque de uma campainha e, por meio de um

microfone, transmite o sinal elétrico até uma pulseira vestível. A pulseira ao receber o sinal elétrico aciona indicadores luminosos e vibratórios alertando o deficiente.

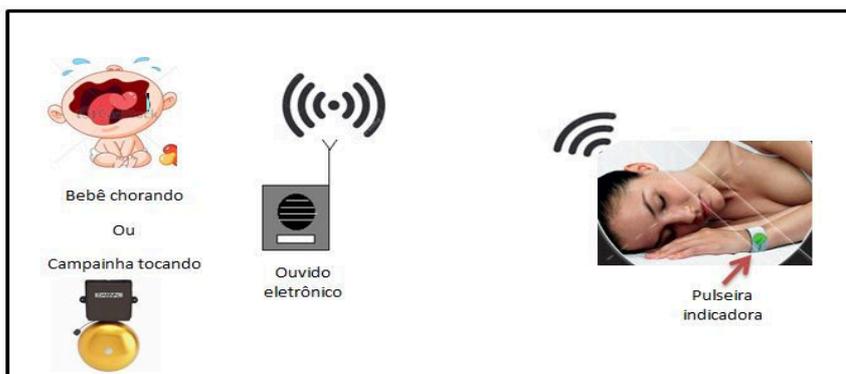


Figura 3. Exemplo de um sistema SADA com dispositivo de alerta. Fonte [Oliveira 2015].

Sendo um dos primeiros aparelhos criados para os DAs, os Aparelhos de Amplificação Sonora Individual (AASI) são aparelhos que realizam o tratamento do sinal acústico e o direcionamento do sinal amplificado para o ouvido. Podem ser utilizados de diversas formas que variam de acordo com a necessidade, formato e tecnologia empregada, podendo ser analógicos, programáveis ou digitais [Pereira & Feres 2005]. Existem basicamente quatro modelos de aparelhos auditivos, sendo eles (Fig.4):

1. **Aparelhos de Canal (*Canal Aids*):** Colocados dentro do canal auditivo. Existem dois tipos sendo o **ITC** (*In The Canal*, dentro do canal) e o **CIC** (*Completely In Canal*, completamente no canal). O ITC é adaptado para a forma do canal auditivo e, o CIC fica oculto dentro do canal auditivo. Ambos os modelos são indicados para perdas leves a severas.
2. **Intra-Auricular (ITE-*In The Ear*, no ouvido):** Indicado para perda auditiva leve a severa, aparelho maior que preenche totalmente “concha” do ouvido e uma parte do canal acústico, ficando visível. Existe o modelo reduzido, denominado Meia-Concha.
3. **Retroauricular (BTE- *Behind the Ear*, atrás do ouvido):** Indicado para perda auditiva leve a profunda, possui uma caixa plástica que é utilizada atrás do ouvido e um molde auricular para dentro do canal auditivo. Pode ser de adaptação aberta ou fechada: (a) aberta é utilizada junto a um tubo fino no lugar do molde para conectar o aparelho ao ouvido; e (b) fechada significa que é usado junto a um molde que praticamente fecha completamente o ouvido externa e o canal auditivo.
4. **Receptor no canal (RIC- *Receiver In Canal*):** Indicador para perda auditiva leve e moderadamente severa, sendo uma evolução do modelo tradicional BTE é pouco ou quase nada visível.



Figura 4. Modelos de AASI. Fonte [Nicolau 2010].

Basicamente, os AASI são dispositivos bastante simples por meio dos quais um pequeno microfone capta o som do ambiente, convertendo-o em um sinal elétrico passando por um amplificador, estágio de pré-processamento do som com a redução de ruídos. O amplificador envia o sinal para o receptor que envia o sinal sonoro para o ouvido interno seguindo o percurso até o cérebro.

Os aparelhos auditivos somente são eficazes para deficientes auditivos e não para surdos. Células ciliadas no ouvido interno devem capturar as vibrações que o aparelho auditivo enviar e converter essas vibrações em sinais nervosos; portanto, é preciso ter pelo menos algumas células ciliadas no ouvido interno para que ele funcione. E, mesmo se algumas células ciliadas permanecerem, um aparelho auditivo não restaurará completamente a audição normal.

Em caso de surdez severa é recomendado o implante coclear, pois as informações sonoras são enviadas diretamente ao nervo auditivo como sinais elétricos. Podendo ser usado sozinho ou acompanhado de aparelhos auditivos tradicionais em pessoas que têm perda auditiva moderada para algumas frequências, com uma perda auditiva mais severa em outras frequências.

Dentre os dispositivos acessíveis, se encontram os dispositivos que podem ser acoplados a certos modelos de telefone para aumentar o seu volume, conhecidos como campainhas auxiliares para telefone, que podem ser utilizados por pessoas com deficiência auditiva leve, ou por qualquer pessoa em um ambiente com ruído excessivo [França 2011]. Este é conectado direto à linha telefônica ou ao ramal analógico; quando o telefone tocar, além de ter seu volume de toque amplificado, uma lâmpada é acesa, como pode ser observado no modelo da Fig.5.



Figura 5. Campainhas auxiliares para telefone. Fonte [Lojamatel 2017].

Outro produto, que possibilita o acesso a DAs, é o relógio com um adaptador para viabilizar vibrações (como uma forma de despertador tátil). Existem diversos modelos que podem possuir luzes de LED de alta intensidade, alarme sonoro interno que aumenta o volume de som, receptor embutido que pode ser utilizado para alertar por indicativo luminoso, por exemplo, se o bebê estiver em prantos ou se o telefone estiver tocando (Fig.6).



Figura 6. Relógio com dispositivo vibratório. Fonte [PTAssistech 2017].

Os Telefones com Teclado-Teletipo (TTY, do inglês *Teletypewriter*) ou Dispositivo de Telecomunicação para Surdo (TDD, do inglês *Telecommunication Device for the Deaf*), surgiram por volta da década de 1960, conceito criado pelo cientista surdo Robert Weitbrecht para que pessoas com sua condição pudessem se comunicar via telefone [Maiorana-Basas 2014].

O telefone TTY (Fig.7) é um dispositivo do tamanho de um pequeno computador portátil com um teclado QWERTY<sup>1</sup> e uma pequena tela que exibe o texto digitado eletronicamente. É necessário conectar um telefone fixo junto ao aparelho, no qual o DA irá digitar a mensagem que deseja transmitir pela linha telefônica, assim como receber mensagens. Grande parte dos telefones móveis do país possui a função TTY que, quando habilitada, é possível utilizar o equipamento TTY, a partir de um cabo específico conectado entre os dispositivos.



Figura 7. TTY [Notisurdo 2017].

Segundo o relatório publicado pela SEDH [SEDH 2009], as pessoas com DA, demandam de aparelhos de amplificação sonora individual (AASI) adaptados à distância (Telemedicina); de tecnologias para reabilitação auditiva com uso de tecnologia da informação; pesquisas em reabilitação auditiva; investimento no desenvolvimento de novos modelos de AASIs e de uma base tecnológica para adaptação à distância.

Contudo, a indústria brasileira não supre a demanda, o que se deve a poucos recursos para o desenvolvimento de novos produtos internos devido à dominação do mercado por empresas internacionais ou importadores independentes, no qual muitas vezes cria-se uma barreira ou limitações que atrasam o desenvolvimento de TA.

Portanto, os recursos de acessibilidades para surdos, no país, ainda são insuficientes, devido a diversos fatores. Nesse contexto, é preciso fazer uso de mecanismos que possibilitam aos DAs o acesso a conteúdos digitais, com autonomia, além de proporcionar um aprendizado contínuo na escrita e leitura da Língua Portuguesa. Dentre esses mecanismos, se destacam os *softwares* que realizam a tradução de textos e áudios da Língua Portuguesa para Libras, com a

---

<sup>1</sup> Layout mundialmente utilizado para o arranjo das letras de teclado.

utilização de animação por avatar, disponíveis para administradores de *site* e para dispositivos móveis. Dentre eles podemos citar:

1. **ProDeaf:** É uma plataforma de tradução do português para Libras, via *Web* e por aplicativo para *smartphone*. Com acesso gratuito, o aplicativo possui o recurso de tradução automática de pequenas frases do português para Libras, a partir de texto ou por reconhecimento de voz, e também oferece dicionário de Libras (Fig.8).
2. **Hand Talk:** Assim como o ProDeaf, é *App* gratuito e realiza a tradução automática de texto e voz para Libras, além de possuir um recurso de vídeos que ensina as expressões e sinais em Libras. Também, possui recursos para empresas e administradores de *site*, com recurso de tradução (Fig.9).
3. **Librazuka:** É um aplicativo desenvolvido dentro de uma universidade, criado com o intuito de popularizar a Libras por meio de módulos teóricos sobre o alfabeto, números, gramáticas e jogos de entretenimento, para um aprendizado dinâmico (Fig.10).



Figura 8. Avatar do *App* ProDeaf [ProDeaf 2017].

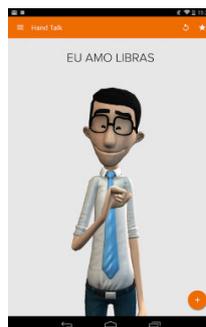


Figura 9. Avatar do *App* Hand Talk [Hand Talk 2017].



Figura 10. Tela do *App* Librazuka [Librazuka 2017].

4. **Acessibilidade Brasil:** a Acessibilidade Brasil é uma sociedade constituída por especialistas da área de educação especial, professores, engenheiros, administradores de empresas, arquitetos, desenhistas industriais, analistas de sistemas e jornalistas, que tem como interesse comum o apoio, ações e projetos que privilegiam a inclusão social e econômica de pessoas com deficiência, idosos e pessoas com baixa escolaridade. Essa sociedade disponibiliza, também, um *App* gratuito, tradutor para Libras a partir da conversão de textos e reconhecimento de voz, fazendo uso de avatar para a interpretação dos sinais [Acessibilidade 2017].
5. **Rybená:** É mais uma opção de acessibilidade, sendo um *App* gratuito de tradução de textos para Libras e a reprodução de textos para voz falada [Rybená 2017].

Nota-se que todos os aplicativos citados acima são gratuitos, têm a função básica de tradução de texto em português para Libras com avatar, tendo também a possibilidade de reconhecimento de fala, traduzindo-a para texto e posteriormente para Libras. Alguns aplicativos fornecem funções adicionais como o dicionário de sinais (ProDeaf, Librazuka e Acessibilidade Brasil), e também funções de conversão de texto em fala (Rybená), conforme Tabela 1.

Tabela 1. Comparativo entre os diversos aplicativos citados.

Funções	ProDeaf	HandTalk	Librazuka	Acessibilidade Brasil	Rybená
Texto --> Libras	sim	sim	sim	sim	Sim
Voz --> Libras	sim	sim	-	sim	sim
Dicionário Libras	sim	-	-	sim	-
avatar	sim	sim	sim	sim	sim

### 2.3. Considerações Finais

Neste capítulo foram apresentadas as informações gerais relativas à audição, deficientes auditivos e tecnologias de assistência. No próximo capítulo, será dada continuidade aos tópicos abordados com uma pesquisa bibliográfica dos principais trabalhos que envolvem o desenvolvimento de tecnologias de assistência aos DAs.

## CAPÍTULO 3

# REVISÃO BIBLIOGRÁFICA

---

O presente capítulo refere-se à revisão bibliográfica sobre os principais trabalhos desenvolvidos na área de tecnologia de assistência a deficientes auditivos. Assim, na Seção 3.1 são apresentadas as considerações iniciais, na Seção 3.2 são descritos trabalhos relevantes sobre tecnologias assistivas aos DAs, e na Seção 3.3 são descritos os trabalhos referentes ao processamento de sinais, relativo às tecnologias assistivas. Finalmente na Seção 3.4 são apresentadas as considerações finais sobre o capítulo.

### 3.1. Considerações Iniciais

A tecnologia da informação tem sido promissora na criação de produtos e serviços em assistência às pessoas deficientes, principalmente pela dedicação de muitos pesquisadores, tornando expectativas em realidade.

Muitos aplicativos e recursos de acessibilidade têm sido desenvolvidos, dentre eles recursos de interface *Web* com conteúdo textual simples e claro, vídeos legendados, conteúdos visuais e auditivos como equivalentes textuais. Também, o desenvolvimento de aplicativos para celulares *smartphone* com avatares que reproduzem o texto escrito, ou o áudio, para linguagem de sinais, aplicativos para a promoção do alfabeto da linguagem brasileira de sinais dentre outros recursos. Também existem disponíveis os dispositivos e recursos, como por exemplo o telefone TTY, campainha luminosa, despertador e babá eletrônica com alerta visual e vibratório.

Apesar dos recursos citados, ainda persistem barreiras na comunicação e inserção social de pessoas com deficiência, em especial a auditiva. Isso se deve, em parte, à quantidade insuficiente e à inacessibilidade desses recursos para boa parte dos DAs, carentes economicamente, como evidencia este capítulo, com base nos trabalhos desenvolvidos por pesquisadores.

Outro tópico a ser abordado neste capítulo são os trabalhos sobre a técnica de Análise de Componentes Independentes sendo amplamente utilizada em uma gama de aplicações em

análises de sinais onde não se conhece os sinais das fontes, conforme o problema de separação cega das fontes.

Dentro das diversas aplicações é possível citar as separações de áudio com base no problema de *Cocktail-party*, no qual temos  $n$  fontes de sons gravados por  $n$  números de microfones, no qual temos interesse em apenas uma das fontes. As vezes apenas um sinal é de interesse como de uma fala de uma pessoa e os demais são considerados ruídos, que para removê-los é possível aplicar técnicas como filtros lineares, *wavelet* e retração de códigos esparsos, algumas aplicadas em trabalhos correlatos, descritos neste capítulo.

### 3.2. Trabalhos Relativos à Tecnologia Assistiva

Plachevski [Plachevski 2014] propôs o desenvolvimento de um dispositivo de reconhecimento de palavras que emite um alerta vibratório, quando palavras pré-determinadas são reconhecidas, com o intuito de melhorar a comunicação dos DAs com pessoas consideradas ouvintes. Para isso, fez uso da tecnologia de reconhecimento de fala, por meio de um aplicativo para *smartphone* com sistema operacional *Android*, que emite um sinal vibratório e luminoso chamando a atenção do DA, sempre quando fossem detectadas palavras previamente registradas no dicionário de dados do dispositivo.

O aplicativo desenvolvido fornece, ao usuário final, diversas possibilidades de configuração, como alerta por vibração ou por tela luminosa, escolha entre um nome único ou nome e sobrenome, entre outros parâmetros. Para comprovar a eficácia do sistema desenvolvido, realizou testes para identificar os tipos de palavras, quanto à gramática, composição e idioma, à distância e ângulo relativo entre o locutor e o microfone, e também, o nível de ruído em cada demonstração.

Os testes foram realizados com a utilização do sistema *Pocketsphinx* para *Android* adaptado para realizar o reconhecimento de palavras e *Software Noise Meter* versão 2.7, responsável por medir o nível de ruído (dB) do local. Com os testes, observou que o sucesso no comportamento do dispositivo está diretamente relacionado aos níveis de ruído no ambiente e a distância do locutor ao dispositivo de reconhecimento, portanto, quanto maior a distância, maior o ruído e menor a taxa de acerto nas palavras reconhecidas.

Ferreira, em [Ferreira 2014] propôs um guia de recomendações avançadas para o *design* e avaliação de interfaces *Web* promovendo o acesso a pessoas surdas. A partir de uma ampla pesquisa bibliográfica envolvendo pessoas surdas em um Centro de Atendimento ao Surdo,

descreveu as principais dificuldades e necessidades que eles enfrentam e possíveis soluções apontadas para melhorias no acesso à *Web*.

Diante disso, constatou-se que pessoas com vários tipos de deficiências, das quais muitas possuem grandes dificuldades no acesso à *Web*, são sujeitas à exclusão e isolamento da maior fonte de conhecimento e comunicação do mundo. A falta de acesso se dá mediante a inexistência ou pouca acessibilidade oferecida, quesito importante para pessoas com deficiência auditiva severa, uma vez que se predomina a linguagem textual o que é de longe uma grande barreira, pois se sabe que os surdos possuem grande dificuldade de compreensão e associação das letras e o que elas significam, uma vez que não ouvem os fonemas das palavras e desde cedo aprendem a linguagem de sinal que difere da linguagem textual [Ferreira 2014].

Destaca-se o padrão W3C (Consórcio Internacional de Padrões para a *Internet* ou em inglês, *World Wide Web Consortium*) para tornar as interfaces acessíveis, o qual aponta, dentre as recomendações, que as páginas *Web* devem possuir conteúdo textual simples e claro, os vídeos devem possuir legendas com textos descritivos dos fatos, usar conteúdos visuais e auditivos através de equivalentes textuais, possibilitando o acesso às informações por usuários cegos e surdos alfabetizados.

A partir das entrevistas concedidas por pessoas com deficiência auditiva, Ferreira [Ferreira 2014] descreve que a principal dificuldade apontada por todos os participantes é o problema de leitura em textos longos, principalmente os que possuem palavras menos comuns. Como solução, indica a inclusão de um dicionário visual com opção para o surdo aprender o significado de cada palavra e contextualizar o seu emprego, que pode ser mediante o uso de avatar que transcreve para Libras. Também, coloca como solução para o problema de escrita do surdo, a utilização de um corretor ortográfico que aponta a palavra errada e por meio de um *link* no aplicativo, mostra a forma correta, que pode ser explicado em Libras por meio do avatar.

Embora já existam alguns sistemas, dispositivos e serviços que contribuam para a melhoria em atividades cotidianas de pessoas surdas, vale identificar se realmente estes estejam contribuindo de forma efetiva para a autonomia das pessoas surdas. A partir de uma pesquisa criteriosa com pessoas surdas, França em [França 2011] investigou como se desenvolve a interação social de pessoas surdas em seu cotidiano a partir de sistemas de produtos e serviços de comunicação, apontando as principais facilidades, dificuldades, necessidades de intervenções entre outras possíveis particularidades de uso.

Dentre vários pontos, França [França 2011] coloca as barreiras de comunicação no dia a dia das pessoas entrevistadas que são impostas pelas inacessibilidades em uma série de produtos e serviços, evidenciando que muitas dessas barreiras seriam inexistentes se houvesse a prestação adequada de serviços específicos às necessidades dos surdos.

Dentre as sugestões de adequações e melhorias estão a disponibilidade de tradutores e intérpretes de sinais para um atendimento adequado em restaurantes, bancos e principalmente em hospitais, pela falta de atendimento especializado em consultas médicas, sendo necessário o acompanhamento de familiares.

Outro apontamento de França [França 2011] é a carência de recursos de sinalização luminosa e de vibração para a comunicação e acessórios básicos como campainhas luminosas, relógios com despertador por vibração, babá eletrônica com sinalizador visual e por vibração, dispositivos sinalizadores de senhas de atendimento entre outros produtos do dia a dia. Devido a essa carência, muitos DAs se veem forçados a procurar alternativas, como recorrerem a outras pessoas como intermediadoras, ou são inibidas ao acesso a determinados produtos e serviços.

A tecnologia tem o potencial de reduzir o isolamento, aumentar a independência e proporcionar oportunidades educacionais, financeiras e sociais aos usuários. Contudo, ainda persistem algumas barreiras, em *hardware* e *software*, bem como na *Internet*, que podem limitar o acesso a informação e oportunidades, para indivíduos com necessidades especiais. Com esse intuito, foi realizada uma pesquisa por Maiorana-Basas e Pagliaro [Maiorana-Basas & Pagliaro 2014] com diferentes trabalhos que apontam as tecnologias que os surdos e deficientes auditivos utilizam em seu dia a dia, nos Estados Unidos. A comunicação por texto é amplamente difundida em serviços de mensagens por SMS, e-mail, aplicativos de mensagens instantâneas, e tecnologias similares que contribuem para uma maior aproximação entre indivíduos ouvintes e deficientes auditivos, além de proporcionar maior independência social e acadêmica, e consequentemente, reduzindo o isolamento social.

Em seu trabalho Maiorana-Basas e Pagliaro concluíram que os participantes da pesquisa, preferem utilizar e-mails a telefones TTY/TDD para se comunicarem, pois, esse tipo de tecnologia de mensagens via texto ainda possibilita a utilização de *emoticons*<sup>2</sup> que servem como expressão de emoção e sentimentos dentro de uma mensagem. Também, concluíram que a utilização de vídeo na comunicação moldou as práticas de indivíduos que são deficientes

---

<sup>2</sup> *Emoticons* são representações gráficas de sinais, comumente de expressões faciais, que podem ser incorporados na comunicação textual em dispositivos eletrônicos.

auditivos, por *smartphones*, *tablets*, computadores com câmera, a partir de sistemas como, por exemplo, *Skype* e *FaceTime*. Os resultados da pesquisa apontaram que 71,6% dos entrevistados fazem uso frequente de *smartphones*, principalmente os mais jovens; enquanto isso, os mais velhos preferem utilização de computadores. Os locais em que utilizam essas tecnologias foram com maior frequência dentro de casa, seguido do local de trabalho e, depois, escola.

Maiorana-Basas e Pagliaro concluíram que a tecnologia tem potencial para melhorar e apoiar a educação entre indivíduos que são deficientes auditivos, especialmente no que diz respeito à aprendizagem de línguas. Embora nenhum dos estudos da literatura abordam especificamente os efeitos que proveem as melhorias nas competências em matéria de alfabetização, os próprios entrevistados relataram que a tecnologia proporciona maior acesso ao idioma inglês, e que conteúdos de vídeos propiciam maior acessibilidade para os educandos em relação aos materiais educativos padrões.

Sharma e Wasson [Sharma & Wasson 2012], com o intuito de promover a comunicação entre duas pessoas com deficiência, sendo uma cega e a outra surda, desenvolveu uma ferramenta de reconhecimento e síntese de fala (*Speech Recognition and Synthesis Tool*, SRST).

O sistema foi desenvolvido para Windows com a ferramenta Microsoft.NET 3.5 utilizando o Microsoft Visual C# e consiste em um ambiente de conversação com a função de conversão de fala para texto e de texto para fala. Foram realizados experimentos com estudantes cegos e estudantes surdos treinados no reconhecimento de fala em ferramentas e janelas. Observou-se que eles se comunicaram efetivamente somente em ambiente livre de ruído, e os estudantes cegos tiveram maior dificuldade de familiarização com a ferramenta.

Mutemwa e Tucker [Mutemwa & Tucker 2010] desenvolveram um aplicativo para smartphones com o intuito de possibilitar a comunicação entre pessoas surdas e ouvintes para diagnóstico de doenças médicas. O aplicativo usa a comunicação interpretada entre linguagem gestual e textual na língua inglesa, a partir de um banco de vídeos de linguagem de sinais pré-gravados para a pessoa surda.

O usuário surdo cria texto em inglês respondendo a perguntas enquanto a pessoa ouvinte lê e responde usando um dicionário de pesquisa. O sistema foi projetado para suportar a independência de conteúdo como a visita de um médico, a compra de medicamentos em uma farmácia, entre outros. A criação de conteúdo é feita com facilidade ao carregar vídeos em linguagem de sinal e escrevendo textos a partir da navegação *web*.

Amarasinghe e Wimalaratne [Amarasinghe & Wimalaratne 2017] desenvolveram um *framework* para comunicação com pessoas surdas. Fornece uma interface de comunicação por celular para sistemas Android, que converte a voz da pessoa sem deficiência em texto, em seguida fornece um *feedback* tátil para a pessoa com DA.

O *feedback* tátil é feito em Código Morse, por ser uma ferramenta de auxílio a pessoas com deficiência. Então, o trabalho teve como objetivo usar um conjunto de padrões de vibração de taquigrafia de Código Morse para traduzir a mensagem de texto em uma resposta tátil para auxílio ao DA.

O sistema funciona da seguinte maneira: (1) uma pessoa comum faz uma chamada de voz, (2) então o sistema converte a voz para texto e exibe no display para o DA, (3) o texto é convertido em abreviações em Código Morse e é enviado ao DA por vibrações.

A Tabela 2 mostra um resumo dos trabalhos descritos sobre tecnologias assistivas.

Tabela 2. Trabalhos relativos às tecnologias assistivas.

Autores	Resumo	Caracterização do sistema proposto
[Plachevski 2014]	Envolve um sistema de tecnologia assistiva para DAs.	Sistema que envolve detecção de palavras pré-determinadas e emissão de alerta por sinais vibratórios ou luminosos
[Ferreira 2014]	Envolve o <i>design</i> inclusivo e participativo na <i>Web</i> incluindo os DAs.	Descreve as principais dificuldades e necessidades que os DAs enfrentam e possíveis soluções para melhorias no acesso à <i>Web</i> .
[França 2011]	Relativo à interação social de pessoas surdas no cotidiano, mediante sistemas e serviços de comunicação.	Descreve as barreiras impostas pelas inacessibilidades e carência de recursos de sinalização luminosa e de vibração.
[Maiorana-Basas & Pagliaro 2014]	Trata-se de uma revisão sobre diferentes tecnologias utilizadas no dia a dia pelos DAs, nos Estados Unidos.	A maioria dos entrevistados preferem usar <i>e-mails</i> a telefones TTY para se comunicarem. Os resultados de pesquisa apontaram que 71,6% dos entrevistados fazem uso frequente de <i>smartphones</i> .
[Sharma & Wasson 2012]	Concentra-se em promover a comunicação entre duas pessoas com deficiência, sendo uma cega e a outra surda.	Desenvolveu uma ferramenta de reconhecimento e síntese de fala (SRST). Observou-se que os usuários se comunicaram efetivamente somente em ambiente livre de ruído, e os estudantes cegos tiveram maior dificuldade de familiarização com a ferramenta.
[Mutemwa & Tucker 2010]	Possibilitar a comunicação entre pessoas surdas e ouvintes para diagnóstico de doenças médicas.	Desenvolveu um aplicativo que usa a comunicação interpretada entre linguagem gestual e textual na língua inglesa, a partir de um banco de vídeos de linguagem de sinais pré-gravados para a pessoa surda.
[Amarasinghe & Wimalaratne 2017]	Desenvolver um <i>framework</i> para comunicação com pessoas surdas.	Fornece uma interface de comunicação por celular para sistemas Android, que converte a voz da pessoa sem deficiência em texto, em seguida fornece um <i>feedback</i> tátil por Código Morse para a pessoa com deficiência auditiva (DA).

### 3.3. Trabalhos Relativos às Técnicas de Processamento de Sinais

Abreu [Abreu 2003] propôs uma técnica de redução de ruído e recuperação de sinal de voz em um ambiente ruidoso, baseado no método de decomposição do sinal usando filtros e ICA, juntamente com um algoritmo de erro médio quadrático (*Mean Squared Error*, MSE), que funciona como um filtro recursivo, tendo como objetivo realizar a extração de sinais de voz, num ambiente de simulação do *Cocktail Party*, com apenas um canal de entrada.

Para as simulações, realizadas pelo *software* Matlab, foram utilizados sinais de voz proferindo locuções de números de zero a nove, a uma frequência de amostragem de 8kHz, preservando apenas a fala, removendo partes de expressões vocais.

Pela introdução de diferentes tipos de ruídos, foi constatada a robustez do método frente a sinais indesejáveis, notando a diminuição de erros e redução de ruídos a cada interação do processo, com o aumento do número de funções bases. Foi testada a sensibilidade do método quanto à reverberação, concluindo que o erro aumenta com o comprimento do filtro utilizado.

Pham et al. [Pham et al. 2003] desenvolveu um método para misturas convolutivas, que são misturas de sinais que ocorrem com atrasos diferentes, devido à característica do ambiente, diferentes das misturas instantâneas, que consistem na mistura linear das fontes sem atrasos. O método de Pham faz uso do domínio da frequência, onde as operações de convolução tornam-se produtos, fazendo com que o problema de BSS seja tratado como um caso de mistura instantânea em cada raia de frequência amostrada e, considera a não-estacionariedade do ambiente.

Há de se apontar que, do ponto de vista computacional é um tanto simples esta solução, mas existem problemas de ambiguidade de escala e de permutação, pois ao aplicar o ICA em cada raia de frequência separadamente, a ordem das escalas é aleatória para cada um dos sinais obtidos. Então, ao transformar para o domínio do tempo, os componentes de uma mesma raia de frequência podem não ser da mesma fonte ou não possuir uma escala ideal. Também, o número de raias a empregar acaba sendo bem maior do que o comprimento dos filtros aplicados [Araki et al. 2003].

Cavalcante [Cavalcante 2004] propôs o estudo de técnicas de BSS aplicada ao contexto de processamento multiusuário em comunicações digitais, a partir de estratégias de estimação da função de densidade de probabilidade. A partir deste método realizou uma análise sobre os impactos da estatística de ordem superior em algoritmos adaptativos para BSS.

Em seu trabalho, Rahbar e Reilly [Rahbar & Reilly 2005], descrevem que basicamente existem dois métodos para o problema de BSS em misturas convolutivas que são o método no domínio da frequência e o método no domínio do tempo. Em geral, o algoritmo no método do domínio da frequência é mais simples e possui melhor convergência que no domínio do tempo. Também, aponta que, as principais dificuldades da aplicação do método no domínio da frequência, é a permutação arbitrária e ambiguidades escalares para a estimativa da resposta em cada raia de frequência no processo de separação.

Rahbar e Reilly propuseram uma abordagem para o problema de BSS no domínio da frequência para misturas de sinais em um ambiente fechado reverberante. Demonstraram um procedimento de diagonalização conjunta de matrizes de densidade espectral cruzada, de sinais na saída do sistema de mistura para a identificação do sistema de mistura a cada frequência. A diagonalização conjunta no domínio da frequência foi feita utilizando um novo e rápido algoritmo de convergência com o uso de um método alternativo de otimização por mínimos-quadrados (*Alternating Least-Squares*, ALS). O inverso do sistema de mistura é então usado para separar as fontes.

Para validar o método, Rahbar e Reilly realizaram dois experimentos em ambientes reais reverberantes diferenciando entre eles a dimensão do ambiente e a distância entre os microfones e os alto falantes. Constataram que o algoritmo melhora significativamente a qualidade da estimativa dos sinais de áudios das misturas com um ganho de 20dB no sinal proporcional para uma sala de escritório moderadamente reverberante.

Baseado no estudo do funcionamento do sistema auditivo humano, bem como os problemas associados aos DAs e métodos de amenizar seus problemas, Pedersen apresenta em seu trabalho [Pedersen 2006] o estudo da separação de sinais acústicos aplicados aos aparelhos auditivos, como objetivo principal.

Um novo algoritmo é proposto para separação de sinais acústicos, em que o número de fontes de sinais excede o número de sensores. A fim de separar as fontes a partir das misturas, este método combina iterativamente duas técnicas: a separação cega das fontes pela análise de componentes independentes (ICA) e mascaramento de tempo e frequência. O algoritmo proposto foi aplicado para a separação de sinais de voz, bem como sinais de música estéreo.

O algoritmo proposto por Pedersen utiliza gravações a partir de dois microfones estreitamente espaçados, semelhantes aos microfones usados em próteses auditivas. Pedersen,

também, apresenta um resultado teórico relativo à convergência em algoritmos de análise de componentes independentes usando gradiente descendente.

Para melhorar a inteligibilidade da fala em ambientes difíceis, aumentando o sinal de áudio desejado e suprimindo o ruído de fundo, as técnicas de captação tanto podem ser baseadas em uma única gravação de microfone ou gravações multi-microfone.

O trabalho de Pedersen [Pedersen 2006] teve dois objetivos principais:

1. Revisão de técnicas de separação da fonte: Com vistas a levantar conhecimentos sobre os métodos existentes dentro das técnicas para a separação de voz multi-microfone. Estas técnicas incluem: a separação cega das fontes, *beamforming*<sup>3</sup> e análise computacional de cenas auditivas.
2. Investigação de BSS para aparelhos auditivos: propor algoritmos para separação de sinais, especialmente sinais registrados por um único aparelho auditivo. As técnicas utilizadas são baseadas na separação cega das fontes por meio ICA e máscara de tempo-frequência.

Suyama [Suyama 2007] propôs novos métodos para solução dos problemas recorrentes em BSS, para o caso de misturas convolutivas e não-lineares. Para as misturas convolutivas o método proposto é baseado em previsão não-linear com o objetivo de eliminar as características convolutivas da mistura. Já para o caso de misturas não-lineares, é proposto um método para separação de sinais denominado de modelo de não-linearidade posterior.

No desenvolvimento do trabalho, empregou um algoritmo com características evolutivas como meio de minimizar os problemas de convergência para mínimos locais no processo de adaptação do sistema de separação. Fez uso de um código para estimar a entropia baseado em estatísticas de ordem superior para avaliar a função de custo.

Concluiu que o método proposto baseado em filtros de erros para o problema de misturas convolutivas é aplicável a sinais com distribuição discreta, no qual pelo emprego dos filtros foi possível reduzir as redundâncias nos sinais observados, transformando o problema em uma mistura instantânea e equivalente. Para o problema de separação de misturas não-lineares obteve uma ferramenta flexível para a separação dos sinais, com estrutura não linear e possibilidade de inverter o processo de mistura de modo adequado.

---

<sup>3</sup> Beamforming é uma técnica de processamento de sinal utilizado em conjunto de sensores para a direção do sinal de transmissão ou recepção [Van Veen & Buckley 1988].

O trabalho de Moreto [Moreto 2008] concentra-se no estudo de algumas das principais técnicas de ICA para misturas instantâneas e na sua aplicação para a separação cega de sinais de áudio, abordando os aspectos da análise estatística dos métodos existentes para combinações linear das fontes.

Para a separação de misturas instantâneas de sinais para aplicações de áudio, propôs o algoritmo de busca de projeções (*Projection Pursuit*, PP) para separação de múltiplas fontes. Além disso, realizou uma análise e comparação dos algoritmos FastICA [Hyvärinen et al. 2001], o algoritmo PP e o PearsonICA [Moreto 2008]. Avaliou o método ICA baseando na maximização da curtose de ordem superior e na negentropia. O PearsonICA é um método baseado na estimativa da distribuição da fonte com o auxílio de curvas de Pearson.

Os experimentos práticos foram sintetizados e avaliados com o auxílio do *software* Matlab, com misturas instantâneas. Também, foi realizada uma simulação real do problema de *Coctail Party*, obtendo os sinais misturados em sensores (microfones) permitindo a avaliação do comportamento dos algoritmos em situações reais, a partir de sensores reais.

A principal vantagem do método proposto por Moreto é permitir a rápida adaptação em cenários não estacionários, mas sendo necessário conhecer aspectos das fontes para o tipo de aplicação.

Batalheiro et al. também propuseram métodos, como a análise no domínio do tempo e no domínio da frequência baseados em ICA [Batalheiro et al. 2008].

Para auxílio a idosos, Harvey [Harvey 2010] desenvolveu um dispositivo com o objetivo de monitorar o estado de saúde, fornecendo lembretes diários e alarmes, além de serviços de comunicação, integrado ao sistema de telefonia móvel. O projeto aborda a acessibilidade do telefone móvel, fornecendo uma extensão de interface de usuário com os fundamentos e processos de reconhecimento de fala, contribuindo para os portadores de deficiência.

Nesta et.al. [Nesta et al. 2011] propuseram um algoritmo baseado no método do domínio da frequência com uma implementação recursiva regularizada do ICA (RR-ICA), em que exploram a continuidade da matriz de separação ao longo da frequência e a continuidade da atividade no tempo das fontes. O algoritmo permite estimar um grande número de matrizes de separação mesmo com uma pequena quantidade de dados.

Com o algoritmo proposto, Nesta et al. realizaram dois experimentos com dois sensores e dois alto-falantes com uma variação na distância entre os sensores e os alto-falantes e na

dimensão do ambiente, sendo que para um ambiente menor foi constatada uma reverberação moderada, já para o ambiente maior a reverberação foi intensa, dificultando a separação.

Os resultados dos experimentos também demonstraram sucesso na separação rápida de sinais curtos (por exemplo, de 0,5 a 1s), estimando filtros de separação longos para lidar com ambientes altamente reverberantes.

Wang et al. [Wang et al. 2011] descreve que a solução no domínio da frequência do problema de BSS em misturas convolutivas, aplicada separadamente em cada raia de frequência, deve ter a ambiguidade de permutação em cada raia de frequência resolvida para que os componentes de frequência da mesma fonte sejam agrupados. Assim, propõe um método de alinhamento para resolver o problema de permutação, em que se divide a banda de frequência em pequenas regiões com base no resultado de alinhamento de permutação *bin-wise*.

A Tabela 3 refere-se a um quadro resumo sobre os principais trabalhos relativos às técnicas de processamento de sinais para separação de fontes, descritos na presente seção.

Tabela 3. Trabalhos relativos às técnicas de processamento de sinais.

Autores	Resumo	Caracterização do sistema
[Abreu 2003]	Método de redução de ruído e recuperação de sinal de voz em ambiente ruidoso.	Usa apenas um canal de entrada de sinal.
[Pham 2003]	Método que faz uso do domínio da frequência, para misturas convolutivas, em problemas de BSS e considera a não-estacionariedade do ambiente.	Resolve os problemas de ambiguidade de escala e de permutação.
[Cavalcante 2004]	Foi proposto o estudo de técnicas de BSS no contexto de processamento de vários usuários a partir da estimação de função de densidade de probabilidade.	Os métodos propostos usam uma análise sobre os impactos das estatísticas de ordem superior em algoritmos adaptativos para BSS.
[Rahbar & Reilly 2005]	Propôs um método para o problema de BSS no domínio da frequência para misturas de sinais em ambientes reverberantes.	Concentra-se na resolução das ambiguidades de permutação dependentes de frequência ao explorar a não-estacionariedade inerentes a misturas convolutivas
[Pedersen 2006]	Estudos sobre a separação de sinais usando vários microfones muito próximos, para auxílio aos DAs.	Descreve técnicas de separação de fontes
[Suyama 2007]	Proposição de métodos para solução de problemas recorrentes em BSS, para o caso de misturas convolutivas e não-lineares.	Para a separação de misturas não-lineares foi obtida uma ferramenta flexível com possibilidade de inverter o processo de mistura.
[Moreto 2008]	Concentra-se no estudo de principais técnicas de ICA para misturas instantâneas, aplicadas à separação cega de sinais de áudio.	Apresenta análises experimentais das principais técnicas de ICA e a comparação entre elas, apontando as limitações, características e o seu potencial de aplicação em situações práticas.

[Batalheiro et al. 2008]	Propõe método de BSS para misturas convolutivas usando banco de filtros não-uniformes.	Análise no domínio do tempo e no domínio da frequência baseados em ICA
[Harvey 2010]	Um <i>framework</i> de monitoração e auxílio a idosos.	Aborda o uso de telefone móvel e procedimentos com reconhecimento de fala.
[Nesta et al. 2010]	Propõe um algoritmo baseado no método do domínio da frequência com uma implementação recursiva regularizada do ICA (RR-ICA).	Explora a continuidade da matriz de separação em frequência e a continuidade do tempo ativo das fontes.
[Wang et al. 2011]	Descreve uma solução para o problema de ambiguidade de permutação no domínio da frequência do problema de BSS em misturas convolutivas.	Concentra-se na resolução do problema em que se divide a banda de frequência em pequenas regiões com base no resultado de alinhamento de permutação.

### 3.4. Considerações Finais

Neste capítulo foram apresentados diversos trabalhos relativos às tecnologias assistivas a DAs; dentre eles se destacam os trabalhos de França [França 2011] e Ferreira [Ferreira 2014], que desenvolveram um amplo estudo a partir de entrevistas com indivíduos surdos. Por meio do estudo bibliográfico e das respostas dos voluntários participantes, apresentaram um grande repertório que foi de grande importância no desenvolvimento do trabalho descrito nesse Exame.

Por fim, sobre as técnicas de processamento de sinais para eliminação de ruídos e separação de fontes os estudos sobre misturas convolutivas serão importantes no desenvolvimento deste trabalho.

No próximo capítulo serão apresentados detalhamentos sobre a metodologia adotada para o desenvolvimento do trabalho proposto.

## CAPÍTULO 4

### MATERIAIS E MÉTODOS

---

Neste capítulo será abordada a descrição dos principais materiais e métodos utilizados para o trabalho de Dissertação de Mestrado. Uma das ferramentas básicas para o BSS é a Análise de Componentes Independentes, ICA, que será definida na Seção 4.1. Na Seção 4.2 serão apresentados alguns conceitos fundamentais usados em ICA. Na Seção 4.3 serão apresentados os métodos de pré-processamento como a centralização e o branqueamento; após os quais será detalhado o funcionamento do algoritmo FastICA, que realiza as operações do ICA, em tempo reduzido. Nota-se que quando a mistura dos sinais ocorre instantaneamente à geração, o ICA permite realizar a sua separação. No entanto, quando a mistura ocorre com atrasos distintos em cada um dos sinais, o que é denotado mistura convolutiva, a separação dessa mistura exige outros métodos. Na Seção 4.4 é abordado o processamento para misturas convolutivas. Já na Seção 4.5 serão descritos os principais conceitos relacionados ao reconhecimento automático de fala e, na Seção 4.6, serão apresentados detalhes sobre a tecnologia de *hardware* e *software* que são usados no trabalho. Na Seção 4.7 são apresentados os Sistemas Operacionais e os simuladores utilizados. Na Seção 4.8 são apresentadas as considerações finais.

#### 4.1. Análise de Componentes Independentes

O termo Separação Cega das Fontes (*Blind Source Separation*, BSS) é relativo ao procedimento de obtenção dos sinais fonte, quando os mesmos não são conhecidos, uma vez que os sensores detectores de sinais são localizados espacialmente, de forma que os sinais detectados são misturas instantâneas ou convolutivas de vários sinais fonte. O objetivo é estimar os sinais fonte sem o conhecimento do processo de mistura, utilizando os vetores de observações, que são os sinais misturados captados pelos sensores. As misturas podem ser sons, sinais elétricos, sinais de eletroencefalograma (EEG), eletrocardiograma (ECG) e imagens em geral, no caso de sinais bidimensionais [Pedersen 2006].

Encontram-se duas situações distintas no problema de BSS. A primeira é caracterizada pela mistura instantânea, a partir de uma combinação linear das fontes, e a segunda, pela mistura convolutiva, na qual cada fonte aparece convoluída com a resposta impulsiva, do trajeto entre a origem e o ponto de mistura.

Um dos métodos mais difundidos para BSS é a Análise de Componentes Independentes (em inglês, *Independent Component Analysis*- ICA). Com esse método, podem-se encontrar fontes que deram origem a uma mistura observada pelos detectores, com a suposição de independência estatística entre elas. O ICA tem aplicações em muitas áreas diferentes como na bioengenharia para identificação de fontes em sinais do cérebro, na análise de sinais em sistema de telecomunicação, entre outras áreas de engenharia.

Um exemplo amplamente estudado em BSS é a separação das diversas fontes de sinais durante um coquetel (*Cocktail Party* em inglês) [Hyvärinen & Oja 1997]. O problema consiste em aplicar o método para identificar a fala de cada uma das pessoas.

O modelo da mistura consiste em  $M$  sinais de misturas, capturados pelos sensores (microfones), em função do tempo  $X_1(t), X_2(t) \dots X_M(t)$ , que resultam de uma combinação linear dos  $N$  sinais de fontes em função do tempo  $S_1(t), S_2(t) \dots S_N(t)$ , estatisticamente independentes entre si.

Para ilustrar o modelo considera-se o caso em que  $M = N = 2$ , duas variáveis aleatórias das misturas observadas  $X_1$  e  $X_2$ , em função de duas fontes  $S_1$  e  $S_2$ , representadas pelo sistema de equação:

$$X_1(t) = a_{11}S_1(t) + a_{12}S_2(t) \quad (4.1)$$

$$X_2(t) = a_{21}S_1(t) + a_{22}S_2(t) \quad (4.2)$$

onde  $a_{ij}$ , com  $i, j = 1; 2$  são os coeficientes da matriz  $\mathbf{A}$  da mistura, dependentes da distância dos microfones ao sensor. Na forma matricial, o sistema de equação está descrito pelas equações (4.3) e (4.4).

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} \quad (4.3)$$

$$\mathbf{X} = \mathbf{A}\mathbf{S} \quad (4.4)$$

sendo que  $\mathbf{X}$  é o vetor de mistura com dimensão  $M$  e  $\mathbf{S}$  é o vetor de fontes independentes. A matriz de mistura é dependente da posição em que os sensores estão em relação às fontes.

O procedimento de BSS consiste na estimação da matriz inversa de  $\mathbf{A}$  denominada de  $\mathbf{W}$ . Se  $\mathbf{W} = \mathbf{A}^{-1}$ , então teremos a solução do problema, encontrando as fontes originais, reescrevendo a equação (4.4):

$$\mathbf{S} = \mathbf{A}^{-1}\mathbf{X} \quad (4.5)$$

$$\mathbf{S} = \mathbf{W}\mathbf{X} \quad (4.6)$$

A Fig.11 ilustra a captação de três sinais de origem  $\mathbf{S}_1$ ,  $\mathbf{S}_2$  e  $\mathbf{S}_3$ , por três sensores que obtêm os sinais observados  $\mathbf{X}_1$ ,  $\mathbf{X}_2$  e  $\mathbf{X}_3$ . Dos sensores, as misturas  $\mathbf{X}$  entram para o bloco de separação ICA, tendo por fim os sinais separados.

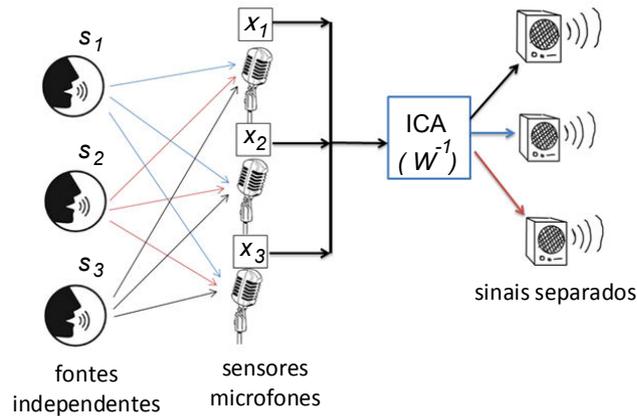


Figura 11. Modelo de funcionamento para estimação das fontes.

Se o número de sensores for menor que o número de fontes, teremos um sistema de equações incompleto, havendo perdas e a possibilidade de não ser possível estimar as fontes. Já nos casos em que o número de sensores é maior que o número de fontes teremos sistemas com mais equações do que incógnitas, sendo assim a solução seria facilitada.

Para simplicidade, foram omitidos os termos para ruído no modelo básico, já que este modelo é considerado suficiente para muitas aplicações. Para um problema mais realista deve-se assumir que nas imediações dos sensores há presença de ruídos.

#### 4.1.1. Restrições

Para que seja possível aplicar a técnica ICA é preciso fazer certas suposições e restrições segundo Hyvärinen et al. [Hyvärinen et al. 2001]:

1. As fontes devem ser estatisticamente independentes: Este é o princípio básico do ICA, que assegura que o modelo possa estimar as fontes. Na Seção 4.2.1 será abordada a definição de independência estatística;
2. O componente independente deve ter uma distribuição não-Gaussiana, sabendo-se que os cumulantes de ordem superior são nulos para distribuições Gaussianas, tratado na Seção 4.2.2. Essas informações de ordem superior são essenciais para estimar as fontes

usando o modelo ICA. Na Seção 4.3.4 será introduzido o algoritmo de ponto-fixo baseado nos conceitos de independência estatística e não-Gaussianidade; e

3. O número de componentes independentes é igual ao número de misturas observadas.

A suposição do número de componentes independentes ser igual ao número de misturas observadas simplifica muito a estimativa, pois consideramos que a matriz de mistura é quadrada podendo assumir a inversa.

#### 4.1.2. Ambiguidades

O ICA possui algumas ambiguidades como segue:

1. Não é possível determinar as variâncias dos componentes independentes. Devido  $\mathbf{S}$  e  $\mathbf{A}$  serem desconhecidos, matematicamente qualquer multiplicação escalar em uma das fontes  $S_i$  poderá ser cancelado ao dividir a coluna correspondente de  $a_i$  em  $\mathbf{A}$ . Ao assumir que os componentes independentes possuam variâncias unitárias:  $E\{S_i^2\} = 1$ , a matriz  $\mathbf{A}$  será adaptada nos métodos de ICA, levando em conta essa restrição. É possível multiplicar por -1 o componente independente que não afetaria o modelo. Essa ambiguidade é insignificante na maioria das aplicações.
2. Não é possível determinar a ordem dos componentes independentes. Novamente, como  $\mathbf{S}$  e  $\mathbf{A}$  são desconhecidos, obviamente a ordem dos termos da soma pode ser trocada, obtendo o mesmo resultado na combinação linear.

#### 4.2. Propriedades Fundamentais Usadas em ICA

Como não há qualquer informação sobre os sinais e suas distribuições nas fontes ou na matriz de mistura são utilizadas estatísticas baseadas em momentos, para caracterizar as distribuições dos sinais estimados.

A estimativa de uma variável aleatória  $X$ ,  $E\{X\}$ , é calculada pela função de densidade de probabilidade (*fdp*) que descreve a probabilidade de uma variável aleatória contínua, dada pela equação (4.7), onde  $p(X)$  é a probabilidade de  $X$ .

$$E\{X\} = \int_{-\infty}^{\infty} Xp(X)dX \quad (4.7)$$

O  $k$ -ésimo momento de uma variável aleatória  $X$  é dado pela equação (4.8):

$$E\{X^k\} = \int_{-\infty}^{\infty} X^k p(X) dX \quad (4.8)$$

Na maioria das distribuições, os momentos possuem valor finito.

#### 4.2.1. Independência Estatística

Duas variáveis aleatórias  $X$  e  $Y$  são independentes entre si, quando a função *fdp* conjunta é dada pelo produto das respectivas funções *fdp* marginais. Essa independência é definida matematicamente em termos da *fdp* marginal,  $p_X(X)$  e  $p_Y(Y)$ , e conjunta  $p_{XY}(X, Y)$ , pela equação (4.9).

$$p_{X,Y}(X, Y) = p_X(X)p_Y(Y) \quad (4.9)$$

Entretanto, se as duas variáveis não forem independentes, essa equação não é válida. Para os momentos de ordem  $k$ , a equação (4.10) é válida quando as variáveis  $X$  e  $Y$  são estatisticamente independentes:

$$E[X^k Y^k] = E[X^k]E[Y^k] \quad (4.10)$$

#### 4.2.2. Misturas não-Gaussianas e o Teorema do Limite Central

Uma das restrições para aplicação do ICA é que as fontes não podem ter uma distribuição normal ou Gaussiana. Os sinais misturados tendem a ter uma *fdp* Gaussiana, porém, as fontes dos sinais devem ter uma *fdp* não-Gaussiana. Para explicar esse fato recorreremos ao Teorema do Limite Central

O Teorema do Limite Central diz que uma soma de variáveis aleatórias independentes tendendo ao infinito, a função de densidade de probabilidade dessa soma tenderá a uma distribuição Gaussiana. Seja  $X_k$  a soma de uma sequência de  $k$  variáveis aleatórias independentes  $Z_i$  e identicamente distribuídas, conforme equação (4.11).

$$X_k = \sum_{i=1}^k Z_i \quad (4.11)$$

Como a média e a variância de  $X_k$  pode crescer infinitamente, para  $k \rightarrow \infty$ , consideremos ao invés de  $X_k$  as variáveis  $Y_k$ , conforme equação (4.12)

$$Y_k = \frac{X_k - m_{Xk}}{\sigma_{Xk}} \quad (4.12)$$

onde  $m_{X_k}$  e  $\sigma_{X_k}$  são a média e a variância de  $X_k$ . É possível mostrar que  $Y_k$  converge para uma distribuição Gaussiana com média zero e variância unitária quando  $k \rightarrow \infty$ . Esse resultado é conhecido como o Teorema do Limite Central.

### 4.2.3. Medidas para a não-Gaussianidade

Como vimos, a estimativa das fontes numa mistura usando ICA é possível a partir da não-Gaussianidade de uma variável aleatória. Para isso, uma das funções de pré-processamento do algoritmo de ICA é obter uma medida quantitativa de não-Gaussianidade de uma variável. Esta seção abordará as duas funções clássicas: (1) curtose de quarta ordem não-comulativa e (2) negentropia baseada na entropia diferencial.

#### 4.2.3.1. Curtose

A curtose refere-se ao grau de achatamento de uma distribuição em relação ao de uma distribuição normal ou Gaussiana. O coeficiente da curtose  $C_x$  é obtido a partir do quociente entre o momento de quarta ordem  $E\{X^4\}$  e o quadrado do momento de segunda ordem  $E\{X^2\}$  conforme equação (4.13)

$$C_x = \frac{E\{X^4\}}{[E\{X^2\}]^2} - 3 \quad (4.13).$$

É possível provar que para uma distribuição Gaussiana a curtose é igual a zero. Já para distribuições não-Gaussianas, a curtose é diferente de zero [Hyvärinen et al. 2001].

#### 4.2.3.2. Negentropia

Um segundo método muito importante para a medida da não-Gaussianidade é dado pela negentropia, que consiste em uma medida quantitativa baseada na entropia diferencial [Hyvärinen et al. 2001]. A entropia é baseada na quantidade de informações a partir de uma observação, que uma variável aleatória pode fornecer. Quanto mais aleatória, imprevisível e desestruturada for a variável, maior a entropia. A entropia é referente a variáveis discretas, sendo que no caso de variáveis contínuas denota-se entropia diferencial. A negentropia é obtida pela diferença entre a entropia diferencial de uma distribuição Gaussiana e a entropia diferencial da distribuição em questão, para uma mesma variância, e serve como medida de não-Gaussianidade. Uma variável aleatória Gaussiana possui a maior entropia diferencial entre

todas as variáveis aleatórias de igual variância. Desse modo, a negentropia pode ser utilizada como uma medida de não-Gaussianidade.

A medida de não-Gaussianidade por negentropia possui como vantagem ser bem definida pela teoria estatística. Entretanto, a negentropia apresenta grau elevado de dificuldade computacional, por isso utilizam-se aproximações mais simples da negentropia.

### 4.3. Método para Separação das Fontes

Esta seção trata-se da descrição do método de separação de fontes usando ICA. Descreve-se inicialmente o pré-processamento e posteriormente o algoritmo de separação de fontes propriamente dito.

#### 4.3.1. Pré-processamento

Antes da aplicação do ICA é muito útil a realização de algoritmos de pré-processamento, como o mais básico que é a centralização e o mais robusto que é o branqueamento.

O pré-processamento mais básico é a centralização de  $\mathbf{X}$ . Observa-se, que a centralização de  $\mathbf{X}$  corresponde à operação que consiste em subtrair o valor médio, para cada elemento  $X_i$  do vetor  $\mathbf{X}$ , ou seja, se  $M = E\{X_i\}$ , fazer  $X_i = X_i - M$ , para torná-lo uma variável de média zero. Isto implica que as fontes estimadas devem possuir média zero. Portanto, assumimos que tanto as variáveis da mistura e os componentes independentes possuem média zero.

O branqueamento é outro pré-processamento que possui um grau de complexidade maior que a centralização, mas ainda constitui um procedimento simples de ser implementado, e que diminui a complexidade do problema.

O processo de branqueamento significa aplicar uma dada transformação linear ao vetor  $\mathbf{X}$  observado, pela sua multiplicação linear por uma matriz  $\mathbf{V}$ , obtendo um novo vetor  $\mathbf{Z}$ , conforme equação (4.14).

$$\mathbf{Z} = \mathbf{V}\mathbf{X} \quad (4.14)$$

tal que a  $C_z = \text{covar}(\mathbf{Z}) = (\mathbf{Z}\mathbf{Z}^T) = \mathbf{I}$ .

O branqueamento é sempre possível, sendo um método popular, o da decomposição de autovalores (*Eigenvalue Decomposition*), aplicando-se o algoritmo de PCA (*Principal Component Analysis*) [Hyvärinen & Oja 1997].

Observa-se que o branqueamento reduz o número de parâmetros da estimativa, e resolve boa parte dos problemas de ICA, devido a sua simplicidade em relação a outros algoritmos de ICA.

Na Fig.12 (A) é mostrada a distribuição conjunta dos componentes independentes  $S_1$  e  $S_2$  antes do branqueamento, e na Fig.12 (B), a distribuição dos componentes das misturas  $X_1$  e  $X_2$ . Na Fig.13 (A) é ilustrada a distribuição dos componentes das misturas após o branqueamento com os componentes  $Z_1$  e  $Z_2$ ; e na Fig.13 (B), a distribuição da estimativa dos sinais fonte  $S_1'$  e  $S_2'$ , que resulta da aplicação do algoritmo *FastICA*, descrito em seguida.

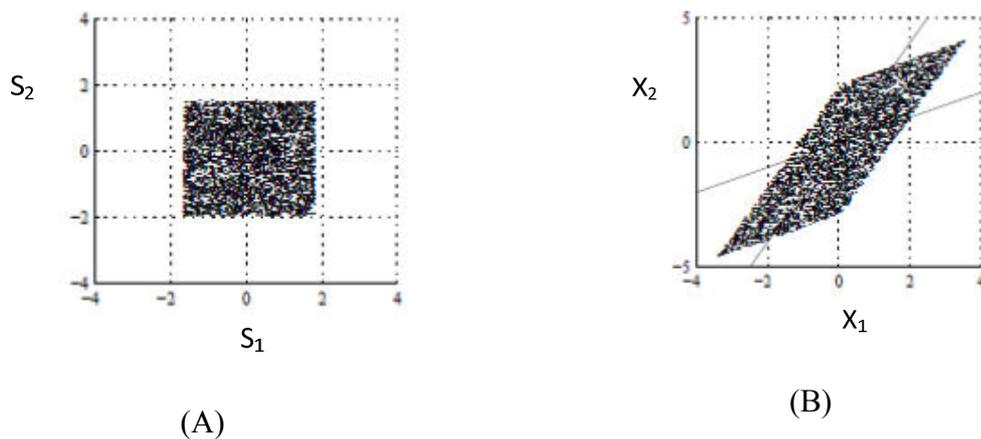


Figura 12. (A) Distribuição conjunta dos componentes independentes  $S_1$  e  $S_2$  com distribuição uniforme; (B) Distribuição dos componentes de mistura  $X_1$  e  $X_2$ . Adaptada de Ferreira [Ferreira 2002].

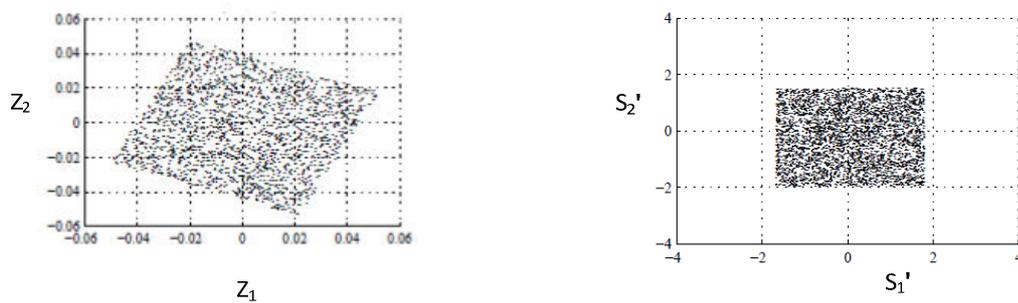


Figura 13. (A) Distribuição conjunta das misturas branqueadas uniformemente distribuídas; (B) Distribuição da estimativa dos sinais fonte. Adaptada de Ferreira [Ferreira 2002].

### 4.3.2. Algoritmo *FastICA*

Nesta sub-seção será abordado o algoritmo de *FastICA*, com diferentes estruturas para processar os componentes de ICA. O algoritmo *FastICA* é uma técnica computacional muito utilizada, proposto por Hyvärinen e Oja [Hyvärinen e Oja 1997], baseado na iteração do ponto fixo para maximizar a não-Gaussianidade de  $\mathbf{W}^T \mathbf{Z}$ .

Quando comparado com os métodos existentes para ICA, o algoritmo baseado na iteração do ponto fixo possui as seguintes características:

1. A convergência é cúbica, ou quadrática e, portanto, mais rápida que outros algoritmos, como o método de gradiente descendente com convergência linear;
2. Não há definição de parâmetros, tornando-o fácil de usar;
3. O algoritmo é capaz de extrair, de praticamente qualquer distribuição não-Gaussiana, os componentes independentes;
4. A escolha adequada de uma medida de não-linearidade pode otimizar o método;
5. Os componentes podem ser estimados um a um, facilitando a análise exploratória de dados, diminuindo o custo computacional; exigindo espaço de memória reduzido.

O algoritmo de *FastICA* pode ser adequado para se maximizar a não-Gaussianidade tanto pelo método da curtose, quanto pelo método da negentropia.

Para o método da negentropia, o algoritmo *FastICA* realiza uma combinação de propriedades superiores que resulta em uma iteração do ponto fixo com propriedades estatísticas devido à negentropia. O **Algoritmo-1** é baseado no método da negentropia, com desenvolvimento em seis passos [Hyvärinen e Oja 1997].

Algoritmo-1: *FastICA* baseado no método da negentropia.

#### **Algoritmo-1: FastICA iteração em ponto-fixo**

1. Fazer com que os dados de entrada  $\mathbf{X}$  possuam média zero.
2. Fazer o branqueamento dos dados obtendo  $\mathbf{Z}$ .
3. Escolher um vetor de peso  $\mathbf{W}$  inicial (aleatório).
4.  $\mathbf{W}^+ = E\{\mathbf{Z}g(\mathbf{W}^T \mathbf{Z})\} - E\{g'(\mathbf{W}^T \mathbf{Z})\}\mathbf{W}$
5. Normalizar, dividindo  $\mathbf{W}^+$  por sua norma.

$$\mathbf{W}_{normalizado} = \frac{\mathbf{W}^+}{\|\mathbf{W}^+\|}$$

6. Se não convergir, voltar ao passo 4.

Inicialmente, no passo-1 o Algoritmo-1 faz uma avaliação dos sinais de mistura  $\mathbf{X}$ , se os dados possuem média zero. Caso isso não ocorra, deve ser realizado o processo de transformação para a média zero, centralização. Em seguida, no passo-2 é realizado o processo de branqueamento dos dados que é um pré-processamento para a aplicação do método de ICA. Com o branqueamento, o processo de ICA passa a ser um simples processo de rotação. Durante uma iteração do algoritmo, um novo valor da matriz,  $\mathbf{W}^+$  é obtido, até que haja a convergência. Caso não haja convergência repete-se a iteração.

A otimização baseia-se no método de Newton [Hyvärinen et al. 2001], onde  $g$  é uma das seguintes funções:

$$g_1(y) = \tanh(a_1 y), g_2(y) = y \exp\left(-\frac{y^2}{2}\right) \text{ e } g_3(y) = y^3 \quad (4.15)$$

onde  $a_1$  é uma constante e  $g'$  é a derivada de  $g$ .

#### 4.4. Separação Cega para Misturas Convolutivas

Conforme abordado em 4.1, a Análise de Componentes Independentes (ICA) é o método mais utilizado para a solução do problema em BSS, ao assumir que as fontes são independentes entre si, e as misturas são instantâneas. Contudo, ao se deparar com misturas convolutivas a separação cega entre sinais de áudio em um ambiente real é mais difícil por muitos aspectos, como Hyvärinen et al. (2001) aponta:

1. As misturas não são instantâneas, pois cada sinal fonte têm propagações lentas, chegando aos microfones em diferentes tempos; e sofrem reverberações no ambiente;
2. O ruído pode ser considerável;
3. Pode existir não-estacionariedade das misturas devido às possíveis alterações nas posições entre os microfones que implica diretamente na matriz de mistura.

Portanto, nas misturas convolutivas o número de parâmetros a serem analisados pode resultar numa tarefa complexa. Para o tratamento dos dados, convencionalmente, é necessária a utilização de filtros FIR (*Finite Impulse Response*) de separação, que devido a um número grande de coeficientes exige-se um grande processamento.

Neste trabalho de dissertação, para o desenvolvimento dos experimentos em misturas convolutivas, foi utilizado um algoritmo de código aberto que faz uso do domínio da frequência, em que as convoluções tornam-se produtos, fazendo com que o problema de BSS seja tratado

como um caso de mistura instantânea em cada raia de frequência amostrada e, considera a não-estacionariedade do ambiente.

#### 4.4.1. Algoritmo para Misturas Convolutivas

Nesta sub-seção será abordado o algoritmo de código aberto, para a separação de misturas convolutivas, desenvolvido por Pham et al. [Pham et al. 2003], usado nesta dissertação. O algoritmo resolve o problema de ambiguidade de permutação com base na continuidade da resposta de frequência do filtro de separação, dando suporte no domínio do tempo.

O algoritmo, **Algoritmo-2**, executa a separação  $W(n)$  no domínio da frequência com transformada de Fourier, atuando sobre uma matriz de mistura  $N \times M$ , para  $N$  fontes e  $M$  sensores.

Antes, porém, de se aplicar o Algoritmo-2, existe o pré-processamento com o filtro FIR, de tamanho  $p$ , modelado conforme a equação (4.16):

$$X(n) = H(n) * S(n) = \sum_0^{P-1} H(i)S(n - i) \quad (4.16)$$

onde  $H(n)$  corresponde à sequência da matriz  $N \times M$  contendo a resposta ao impulso dos canais da mistura,  $n$  é o índice do tempo, o operador “\*” indica a convolução da matriz. O filtro FIR é utilizado para estimar os sinais de origem, antes de sofrer o efeito de mistura convolutiva.

A matriz de separação  $W(n)$  é estimada no domínio da frequência, aplicando a transformada de Fourier, em que os sinais no domínio do tempo observados são convertidos no tempo-domínio da frequência (passo 1). Então o BSS instantâneo é aplicado a cada raia de frequência (passo 2).

Em seguida, resolve-se a permutação inerente (passo 3) e ambiguidades de escala (passo 4), e os sinais separados, em todas as raias de frequência são combinados e transformados inversamente para o domínio do tempo (passo 5). Após a obtenção da matriz de separação  $W(n)$  aplica-se o cálculo para a obtenção da estimativa das fontes independentes.

Maiores detalhes do algoritmo no domínio da frequência em misturas convolutivas, podem ser obtidos em *Blind Source Separation Advances in Theory, Algorithms and Applications* de Naik e Wang [Naik & Wang 2014].

Algoritmo-2: Separação no domínio da frequência para misturas convolutivas.

Entrada: M sinais de sensores de M sensores

1. Decompõe a mistura convolutiva no domínio do tempo em misturas instantâneas no domínio da frequência
2. Executa a separação instantânea em cada raia de frequência com o ICA de valor complexo
3. Aplica o alinhamento de permutação em cada região e sub-região da banda de frequência
4. Aplica o princípio de distorção para resolver a indeterminação da escala de frequência
5. Obtém a matriz de separação  $W(n)$  a partir da transformada inversa de Fourier.
6. Após a obtenção dos filtros, estima-se as fontes independentes.

Saída: N sinais estimados de N fontes

#### 4.5. Reconhecimento de Fala

A produção da voz é possível pelo aparelho fonador, constituído pelos pulmões, brônquios, traqueia, laringe, glote, cordas vocais, faringe, úvula, boca e órgãos anexos, e fossas nasais [Zemlin 2000], cujos sons são classificados entre vocálicos e não vocálicos. Os sons vocálicos se dão pela passagem contínua do ar pelo trato vocal, associado à reprodução das vogais da linguística. Já os sons não vocálicos ocorrem quando o trato vocal impõe resistência à passagem do ar, associada a determinadas consoantes da linguística [Luft 1991].

Com base na análise dos sinais acústicos da voz humana, os sistemas de reconhecimento automático de fala (ASR- *Automatic Speech Recognition*), têm por objetivo reconhecer a mensagem contida na fala e executar uma ação programada, como a transcrição de um texto, acionamento de dispositivos ou tradução de línguas, por exemplo.

Os mecanismos de reconhecimento se tornaram eficientes com a utilização de técnicas de processamento digital de sinais e a partir de redes neurais artificiais, com restrições de operação e implementação.

A aquisição do sinal de fala se dá por um transdutor eletroacústico como o microfone, dada pela variação de pressão das ondas sonoras. Para que se possam extrair os parâmetros desse sinal para o reconhecimento de fala, é preciso que o sinal analógico da fala passe pelas três etapas apresentadas na Fig. 14, quais sejam, pré-processamento, análise espectral e extração de parâmetros.

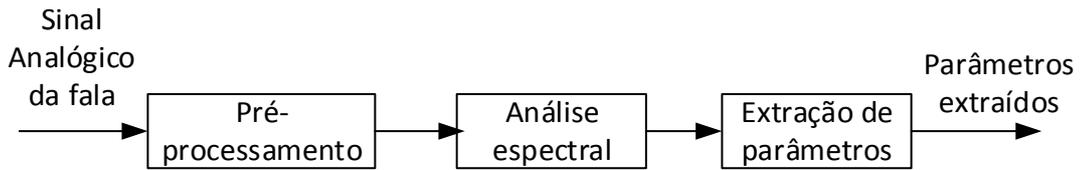


Figura 14. Diagrama de blocos de um sistema de análise do sinal de fala, com a obtenção dos parâmetros para o reconhecimento. Adaptada de Nunes [Nunes 1996].

#### 4.5.1. Pré-processamento

Inicialmente o sinal de fala analógico é filtrado para eliminar frequências acima da metade da frequência de amostragem utilizada, reduzindo o efeito denominado de *aliasing*.

Pelo teorema de amostragem de Nyquist, a frequência de amostragem deve ser maior ou igual a duas vezes a maior frequência do sinal para que não ocorra distorção do sinal. Na Fig. 15, ilustra-se o caso em que é aplicada uma taxa de amostragem de 24Hz inferior ao mínimo de 40 Hz, para um sinal senoidal de 20Hz (20 períodos em 1s). O sinal resultante do efeito de *aliasing* tem frequência de 4Hz (4 períodos em 1s).

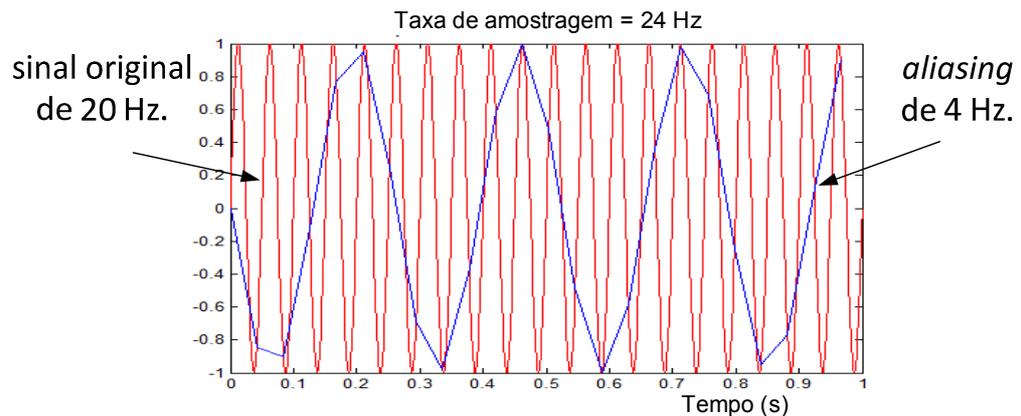


Figura 15. Efeito de *Aliasing* [Barreto 2015].

Aplicando-se a frequência de amostragem observando-se o Teorema de Nyquist [Bresolin 2003], garante-se a estimação de pontos que ao todo formam a curva do sinal, sem perda das características principais, quando se tem pelo menos duas amostras por período de frequência máxima do sinal.

No caso do pré-processamento do sinal de fala, o *hardware* que recebe esse sinal analógico possui um conversor analógico digital (A/D) em que o sinal é convertido em digital e amostrado a uma taxa entre 8 a 16kHz, e o transforma em dados discretos possibilitando sua manipulação pelo computador.

Na Fig.16 encontra-se uma ilustração de um sinal discretizado, nos tempos discretos  $x_1$  a  $x_{10}$ , com amplitudes  $y_1$  a  $y_{10}$ , respectivamente.

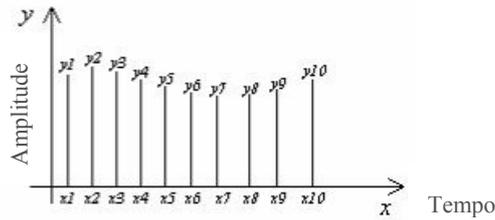


Figura 16. Exemplo de um sinal discretizado.

#### 4.5.2. Análise Espectral

A análise espectral do sinal sobre pequenos intervalos de tempo é dada pelas etapas a seguir:

1. **Divisão em quadros:** O sinal é dividido em quadros, sendo cada um desses quadros composto por  $N$  amostras analisadas individualmente. A uma sobreposição de  $M$  amostras, é criado um novo quadro, o que garante a detecção de variações do sinal, conforme Fig.17, onde na coordenada horizontal tem-se a variável tempo e na coordenada vertical a amplitude do sinal de voz.

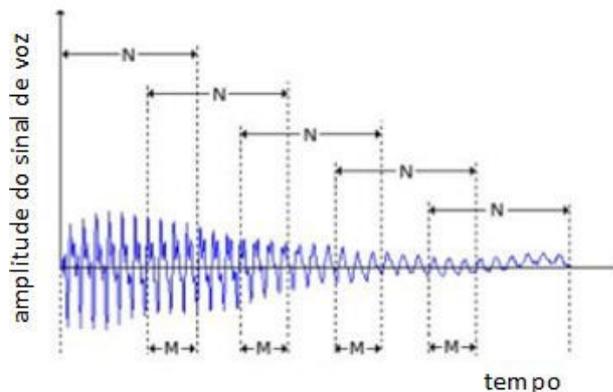


Figura 17. Divisão em quadros do sinal de voz.

2. **Janela:** Os quadros ou *frames* da sequência de fala discreta  $x(n)$  são multiplicados por uma função discreta  $w(n)$ , deslocada de  $i$ , sendo  $n$  a variável do tempo discreto, na equação (4.17):

$$f(n) = x(n)w(i - n) \quad (4.17)$$

Após a aplicação dessa técnica, as amostras de voz selecionadas são utilizadas para a obtenção das características representativas do sinal de voz e seus respectivos parâmetros.

#### 4.5.3. Extração de Parâmetros

Em seguida serão apresentadas brevemente algumas das técnicas que são utilizadas para a extração de parâmetros [Petry 2002].

1. **Banco de filtros:** Consiste na aplicação de filtros passa-banda com deslocamento em frequência, no qual se utiliza os parâmetros de energia na saída de cada filtro;
2. **Energia:** Analisa a variação de amplitude do sinal de voz em um quadro ao longo do tempo, fornecendo informações a respeito da geração da amostra de voz. Também, pode ser utilizado para a determinação de limites das palavras, por meio de um valor limiar que quando abaixo é classificado como ruído de fundo e acima como voz. Portanto, pode-se medir a energia do sinal,  $E(n)$ , a partir do valor da janela  $w(m)$  com  $N$  pontos aplicados ao sinal pré-enfatizado  $y(n)$ , sendo  $n$  o índice de amostragem no tempo discreto do centro da janela, equação (4.18).

$$E(n) = \frac{1}{N} \sum_{m=0}^{N-1} \left[ w(m)y\left(n - \frac{N}{2} + m\right) \right]^2 \quad (4.18)$$

3. **Taxa de cruzamento de zero:** Consiste na análise no domínio do tempo, pela contagem do número de vezes em que o sinal cruzou o nível médio zero. Essa medida de variação de amplitude do sinal de voz entre positivo e negativo, irá indicar se o segmento analisado corresponde a um segmento com fala do sinal ou não [Timoszczuk 2004]. A partir da equação (4.18) obtém-se a taxa de cruzamento por zero ( $ZCR$ ), sendo  $y(n)$  a  $n$ -ésima amostra do sinal de voz pré-enfatizado.

$$ZCR = \frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{2} |sign(y(n)) - sign(y(n-1))| \quad (4.19)$$

4. **Coefficientes cepstrais:** os coeficientes cepstrais são os que garantem que as informações relevantes do sinal de fala sejam extraídas representando as principais especificações do sinal de origem. Para isso, se aplica a Transformada Discreta de Fourier (TDF) que é um método para conversão de um sinal discreto no domínio do tempo para o domínio da frequência. A representação no domínio da frequência pela TDF, da equação que descreve o *frame*  $f(n)$  do sinal vocal pré-enfatizado, pela convolução do sinal de excitação  $u(n)$  com a resposta impulsiva do trato da fala  $h(n)$  [Petry 2000] é dada pela equação (4.20)

$$f(n) = u(n) * h(n) \quad (4.20)$$

E aplicando a função logarítmica, sendo  $F\{\}$  a representação da TDF, resulta na equação (4.21).

$$\log(F\{f(n)\}) = \log(F\{u(n)\}) + \log(F\{h(n)\}) \quad (4.21)$$

Portanto, ao aplicar a transformada inversa de Fourier do logaritmo do sinal, conforme equação (4.21), obtém-se os coeficientes cepstrais do sinal de voz.

5. **Coefficientes cepstrais de frequência-mel (MFCC- Mel- frequency cepstral coeficientes):** A obtenção dos coeficientes mel-cepstrais dar-se-á por meio da extração da representação cepstral do sinal utilizando a escala *mel*, a partir da Transformada Discreta de Fourier no espectro de frequência do sinal. A escala *mel* relaciona a forma com que os seres humanos percebem as frequências pelo sistema auditivo humano com as frequências físicas, valores obtidos a partir de experimentos subjetivos. O nome *mel* refere-se à palavra melodia (*melody*, em inglês) [Petry 2002]. A equação 4.22 define a conversão da escala de frequências (*f*) em *mel*

$$\text{mel}(f) = 2595 \log_{10} \left( 1 + \frac{f}{700} \right) \quad (4.22).$$

Para isso, mapeiam-se as frequências acústicas em hertz para a escala de frequências percebidas em *mel*, e após isso, aplica-se um banco de filtros digitais espaçados dados pela escala *mel* [Petry et al. 2000]. Matematicamente, isso pode ser descrito pela equação (4.23):

$$c(n) = \sum_{k=1}^K \log |S(k)| \cdot \cos \left[ n \left( k - \frac{1}{2} \right) \frac{\pi}{K} \right], \quad 0 \leq n < P \quad (4.23)$$

sendo  $c(n)$  o  $n$ -ésimo coeficiente *mel*-cepstral,  $P$  o número de coeficientes extraídos,  $k$  o número de filtros digitais e  $S(k)$  o sinal de saída do banco de filtros digitais.

#### 4.5.4. Reconhecimento

Com os parâmetros extraídos, é feito o mapeamento acústico correspondente às locuções de entrada, em que o objetivo é gerar uma dada palavra, ou conjunto de palavras, por meio de um processo de busca. Como Nunes [Nunes 1996] descreveu, a busca é dada pela comparação entre a sequência de vetores de características acústicas e os modelos das palavras que estão no vocabulário utilizado pelo sistema. A Fig.18 é um diagrama simplificado do relacionamento de um módulo de reconhecimento de fala com a extração de parâmetros e vocabulário de reconhecimento.

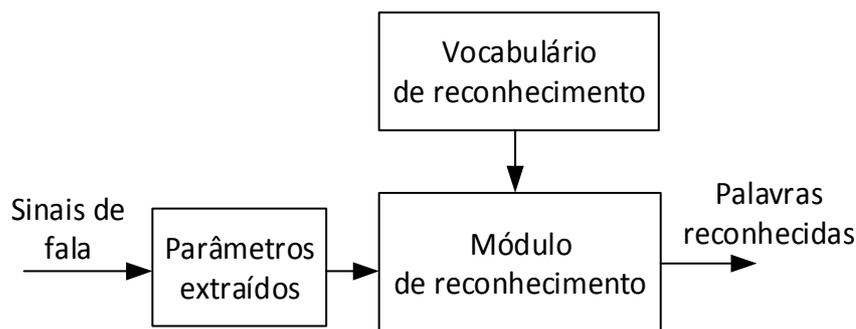


Figura 18. Diagrama representativo do módulo de um reconhecimento de fala.

O vocabulário de reconhecimento nada mais é, que um banco de dados que contém o universo de palavras, no qual toda e qualquer locução será comparada com um conjunto de palavras deste universo. Cada palavra é obtida a partir da concatenação de subunidades fonéticas para cada palavra e a respectiva descrição de cada fonema (duração, classe, entre outros).

Existe uma grande variabilidade de fala em uma população, em que se têm diferenças linguísticas, condições emocionais, físicas e do próprio trato vocal dos locutores. Para tanto, o reconhecimento de palavras deve seguir algoritmos estatísticos, tais como o Modelo Oculto de Markov (*Hidden Markov Models, HMM*) [Rabiner 1989] e as técnicas de redes neurais por *Deep Learning* [Yu & Li 2015].

#### 4.6. Descrições das Tecnologias Integradas

Nesta seção serão tratadas as características de cada um dos componentes que possam integrar dispositivos de auxílio aos DAs, sendo eles os módulos de radiofrequência, o módulo Bluetooth, módulo KY-038 sensor de som e a plataforma de desenvolvimento Arduino.

##### 4.6.1. Sistema de Comunicação por Radiofrequência.

Os sistemas de telecomunicação utilizando ondas eletromagnéticas de radiofrequência significaram um grande passo no desenvolvimento de tecnologias sem fio. A ANATEL (Agência Nacional de Telecomunicações) é o órgão responsável pela homologação e gerenciamento das faixas de frequências que operam no Brasil. Portanto, através da Resolução Nº 365 de 10 de maio de 2004, que regulamenta os equipamentos de radiocomunicação de radiação restrita, disponibiliza a faixa de 433MHz para a operação de equipamentos para telecommando de baixa emissão de potência [Ribeiro 2006].

O sistema de Transmissão por Radiofrequência (RF) é formado por um receptor (RX) e um transmissor (TX), que realiza a transmissão de dados serial ponto a ponto pela modulação por chaveamento de amplitude *ASK* (*Amplitude Shift Keying*), operando em frequência de 433MHz e distância de alcance máxima sem obstruções de 100 metros.

A transmissão pode ser descrita conforme diagrama de blocos da Fig.19 [Ribeiro 2006]. O bloco de lógica TX corresponde ao controle da transmissão, que fica a cargo de um microcontrolador. Na presente proposta de projeto, um processador *Arduino* será o responsável pelo bloco lógico TX realizando a multiplexação no tempo do sinal a ser transmitido, introduzindo a codificação necessária para possibilitar a recepção do sinal.

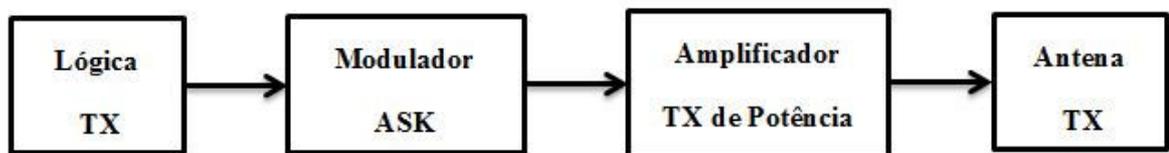


Figura 19. Diagrama de blocos TX.

O *Modulador ASK* consiste em gerar sinais a serem transmitidos, com níveis de amplitude da portadora em função de um sinal de entrada com níveis de amplitude discreta, sendo suas principais características a facilidade de modulação e demodulação, pequena largura de faixa e baixa imunidade a ruídos. Além disso, os blocos de amplificação e antena compõe o diagrama de blocos TX [Baptista 2007].

O diagrama da Fig.20, ilustra um exemplo de um módulo de transmissão TX onde é destacado o terminal da antena, o cristal de 433MHz, a alimentação (Vcc e GND) e a entrada de dados (*Data*), que corresponde aos três blocos da Fig.20, com exceção do primeiro bloco à esquerda, Lógica TX.

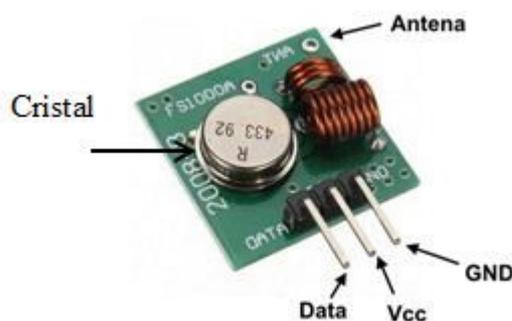


Figura 20. Módulo transmissor TX.

A Fig.21 corresponde ao diagrama de blocos do módulo e recepção RX, onde o primeiro bloco é relativo à Antena RX de recepção do sinal RF. Como o sinal recebido pela antena tem

uma amplitude muito baixa e com ruídos, é necessário passar pelo estágio de amplificação RX, que é o próximo bloco. O estágio de amplificação irá amplificar o sinal na faixa de frequência projetada e realizar a filtragem dos ruídos. No bloco de demodulação, *Demodulador ASK*, o sinal passa por um retificador, em seguida por um filtro passa baixa e por fim por um circuito de recomposição para concluir a demodulação do sinal, sendo o circuito integrado (LM358M) de dois amplificadores, integrante principal desta etapa [Ribeiro 2006].

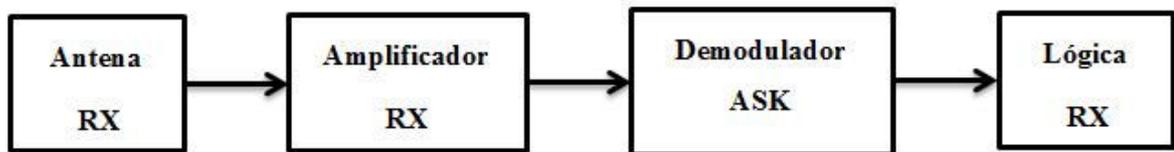


Figura 21. Diagrama de blocos RX.

A partir da demodulação, um microcontrolador, que na presente proposta de projeto consiste num módulo de processamento *Arduino*, irá realizar a função do bloco de Lógica RX, que inclui primeiramente a demultiplexação do sinal.

A Fig.22 apresenta uma ilustração de um módulo de recepção RX, formado por uma antena para a recepção do sinal, um circuito amplificador e filtragem do sinal recebido a partir de um circuito integrado com dois amplificadores (LM358M), e da demodulação, correspondente aos três primeiros blocos do diagrama da Fig.21. Os sinais de RF são capturados pela antena na posição situada à esquerda inferior, e os dados demodulados são entregues ao microcontrolador de Lógica RX, pelos terminais *Data* mostrados à direita inferior, onde também são mostrados os terminais de alimentação (Vcc e GND).

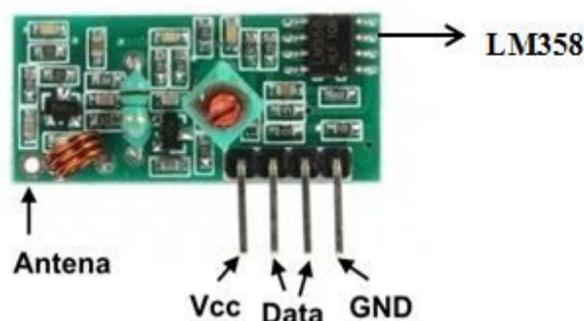


Figura 22. Módulo receptor RX.

#### 4.6.2. Módulo *Bluetooth*

O *Bluetooth* é um sistema de comunicação sem fio de curto alcance, desenvolvido também com base na tecnologia RF, buscando promover uma solução de baixo custo e baixo

consumo de energia na transmissão de dados, possibilitando a conexão de diferentes tipos de dispositivos e de diferentes fabricantes. Essa tecnologia se enquadra no cenário de rede sem fio de área pessoal, denominada WPAN (*Wireless Personal Area Network*), ao substituir cabos de dispositivos portáteis num raio de alcance de até 10 metros.

Com o objetivo de padronizar o desenvolvimento da tecnologia *Bluetooth*, um grupo de empresas criou o SIG (*Special Interest Group*, ou Grupo de Interesse Especial) para essa tecnologia. Esse grupo criou uma especificação industrial aberta, pelo qual definiu uma pilha de protocolos que impõe as regras de como os *softwares* e aplicativos devem proceder para a localização e a utilização dos serviços disponíveis nos dispositivos *Bluetooth*.

O *Bluetooth* provê uma conexão ponto-a-ponto ou ponto-a-multiponto, usando o *piconet*, que é a denominação do protocolo de rede especificada para a essa tecnologia, formando entre 2 a 8 dispositivos conectados na rede, no qual um dos dispositivos funciona como mestre e os demais como escravos, Fig.23. O dispositivo mestre, dentre outras funções, controla o acesso dos dispositivos escravos, determinando o *clock*. Os dispositivos *Bluetooth* operam na banda livre de frequência ISM (*Industrial, Scientific and Medical*) em 2,4 GHz, com uma taxa de transmissão de dados de até 3Mbps. Emprega a técnica de saltos de frequência com 1600 saltos por segundo com frequência de espelhamento espectral FHSS (*Frequency Hopping Spread Spectrum*). Seu espectro de frequências é dividido em 79 canais de radiofrequência (23 para alguns países), cada um com 1MHz de largura de banda [Priess 2003].

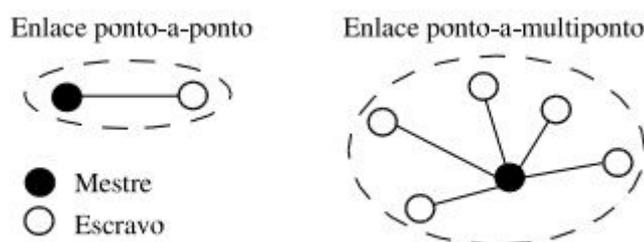


Figura 23. Conexão *Bluetooth* ponto a ponto e ponto a multiponto.

Um módulo no sistema *Bluetooth* realiza a comunicação por rádio de curto alcance na classe 2 com 4dBm de potência (2,5mW) e transmissão via porta serial (*Serial Port Protocol-SPP*). Possui como vantagem seu baixo custo, baixo consumo de energia e tamanho reduzido. Tem como desvantagem a transferência lenta de arquivos e suporta poucos dispositivos conectados de uma vez. Sua segurança é feita por autenticação do tipo desafio-resposta para verificar a validade da unidade de recepção e encriptação de dados, com chave de criptografia com tamanho de 128 bits de estabelecimento de transmissão entre os dispositivos [Kobayashi

2004]. Opera com uma tensão de alimentação de 3,3 V e corrente de pareamento em torno de 35 mA.

Na Fig.24 consta uma fotografia de um exemplo de módulo de comunicação *Bluetooth*, HC-05, que contém 6 pinos de conexão sendo:

1. **STATE:** Determina quando há conexão com o módulo;
2. **RX:** Recebe os dados seriais do *Arduino*. Deve ser conectado em um divisor de tensão, de modo a ser alimentado com 3,3V;
3. **TX:** Transmite os dados serial para o *Arduino*;
4. **GND:** Pino terra.
5. **VCC:** Alimentação 3.6 a 6 V;
6. **KEY ou EN:** Pino colocado em alto para utilização do modo de comando AT.

O módulo *Bluetooth* HC-05 conta com um circuito integrado rádio-transceptor *Bluetooth* versão 2.0 (BC417) e antena integrada.

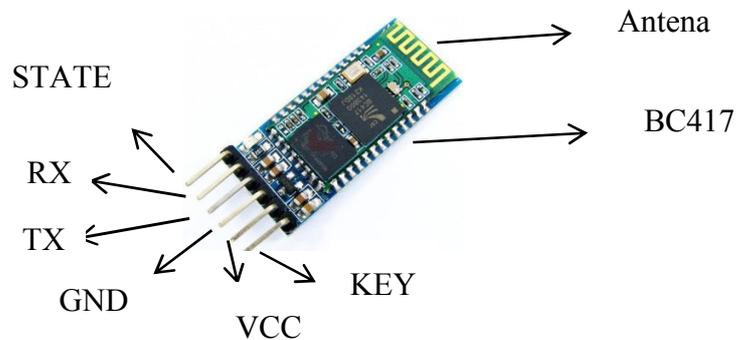
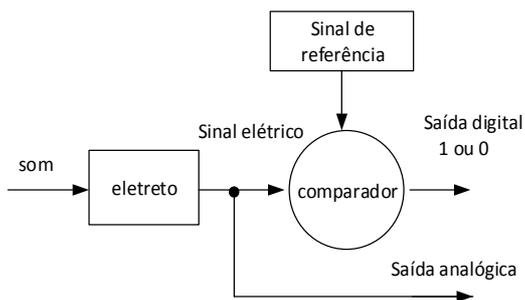


Figura 24. Módulo *Bluetooth* HC-05.

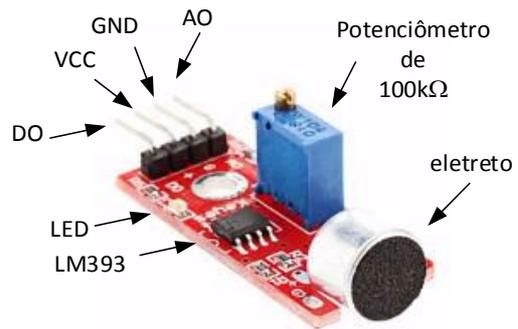
#### 4.6.3. Módulo Sensor de Som

Um módulo sensor de som deve receber na sua entrada um sinal de som e caso o mesmo seja detectado, fornece na sua saída digital a indicação de detecção (nível lógico 1). Caso o som não seja detectado, a saída deve indicar o nível lógico 0 Fig.25 (A). Além disso, a saída analógica fornece o sinal elétrico do som.

Um exemplo de módulo sensor de som é o módulo KY-038, Fig. 25 (B), constituído por um microfone eletreto condensador de alta sensibilidade, um comparador diferencial circuito integrado LM393, potenciômetro de 100k $\Omega$  para ajuste de sensibilidade do eletreto, LED indicativo de alimentação e LED de saída do comparador.



(A) Diagrama de blocos do sensor de som



(B) Sensor de som KY-038

Figura 25. Módulo sensor de som.

O módulo KY-038 possui os seguintes pinos:

- (1) AO: sensor de saída analógica;
- (2) GND: *Ground*;
- (3) VCC: Fonte de alimentação faixa de entrada: 3V-5V;
- (4) DO: Saída Digital (saída do comparador).

Quando o microfone eletreto detecta uma onda sonora, tem-se na saída analógica (AO) o sinal de tensão proporcional à intensidade do som aplicado, conforme a sensibilidade do eletreto. Já a saída digital (DO) é resultado da tensão do comparador de tensão, que conforme o ajuste do potenciômetro o sinal de tensão de saída do comparador é alterado, pois o circuito comparador neste caso, nada mais é que um amplificador em que uma das entradas tem-se um sinal de tensão de referência fixa e na outra entrada o sinal de tensão do microfone. A tensão na entrada será comparada com a tensão de referência, e se for um valor inferior à tensão de referência, a saída do comparador será nível baixo perto de 0V; se o valor for superior ao valor de referência, a tensão de saída do comparador estará em nível alto até 5V.

#### 4.6.4. Plataforma de Desenvolvimento *Arduino*

*Arduino* é uma plataforma de desenvolvimento de dispositivos de processamento em placa única, projetada com um microcontrolador *Atmega AVR*, com suporte de entrada/saída e linguagem de programação padrão, normalmente C/C++. Pode ser usado para o projeto de dispositivos independentes, ou ainda para ser conectado a um computador hospedeiro. A plataforma contém algumas linhas de E/S digital e analógica, além de uma interface serial ou USB, para interligação com o hospedeiro, usado para programação e interação em tempo real.

Uma placa *Arduino* não possui quaisquer recursos de rede, porém é possível combinar uma ou mais placas usando extensões apropriadamente projetadas chamadas de *shields*.

A plataforma *Arduino Uno*, adotada para o presente trabalho, conta com um microcontrolador *ATmega328* que contém internamente o processador denominado *AVR* da família de 8 bits, com arquitetura RISC (*Reduced Instruction Set Computer*), mostrado no diagrama de blocos da Fig.26.

O processador *AVR* usa uma arquitetura *Harvard*, com memória de programa e memória de dados separados, em barramentos próprios. As instruções são executadas em um único nível de *pipelining*. Quando uma instrução está sendo executada, a próxima instrução é pré-carregada da memória de programa, permitindo que as instruções sejam executadas a cada ciclo. A memória de programa consiste em tecnologia *flash*, reprogramável. Um banco de registradores de acesso rápido, contém 32 registradores de 8 bits de propósito geral, o que viabiliza uma operação na Unidade Lógica e Aritmética (ULA), num único ciclo. Numa operação aritmética típica, dois operandos são lidos do banco de registradores, a operação é executada na ULA, e o resultado é guardado no banco de registradores, num único ciclo.

Seis dos 32 registradores podem ser usados como três registradores de 16 bits usados como ponteiros de endereçamento indireto para o espaço de dados, permitindo cálculo de endereços eficiente.

A arquitetura *AVR* com as características descritas permite um desempenho 10 vezes maior em relação a uma arquitetura de microcontrolador *CISC (Complex Instruction Set Computer)* convencional.

O microcontrolador *ATmega 328* é constituído pelo processador *AVR* descrito, com os dispositivos periféricos, conforme ilustrado em diagrama de blocos na Fig.26, cujos detalhes podem ser obtidos na literatura [Atmel 2017].

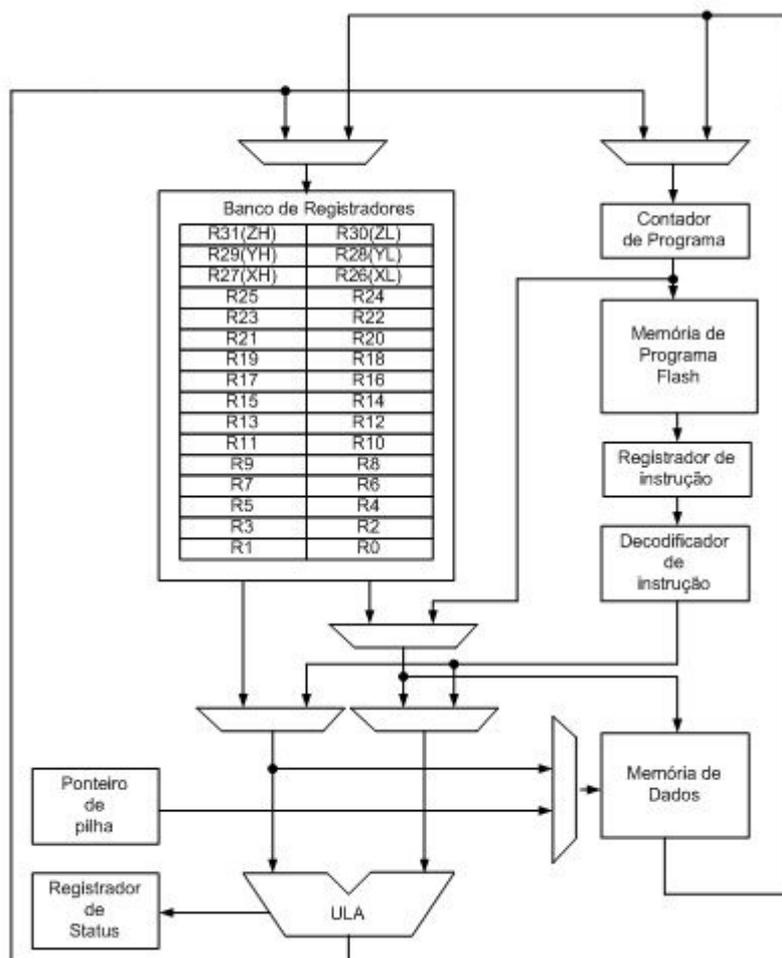


Figura 26. Diagrama de blocos do processador AVR.

O microcontrolador *ATmega328* contém os seguintes fatores, além do processador AVR: 32Kbytes de memória *flash* programável, 1 Kbytes de EEPROM, 2Kbytes de SRAM, e 23 linhas de entrada/saída (I/O) de propósito geral, contador em tempo real (*Real Time Counter*, RTC), três contadores/temporizadores (*Timer/Counters*) flexíveis com modos de comparação com PWM (*Pulse Width Modulation*), um USART (*Universal Synchronous/Asynchronous Receiver/Transmitter*) serial programável, uma interface serial de 2-fios (I2C) orientado a *byte*, um ADC (*Analog to Digital Converter*) de 10 bits de 6 canais, um *Watchdog Timer* programável com oscilador interno, uma porta serial SPI, e seis modos selecionáveis por *software* de uso de energia.

O modo ocioso (*idle*) para a CPU enquanto permite o funcionamento da SRAM, *Timer/Counters*, porta SPI e sistema de interrupção. O modo *Power-down* salva o conteúdo dos registradores, mas congela o oscilador, desabilitando todas as outras funções do *chip* até uma nova interrupção ou *reset*. Em modo *Power-save*, o temporizador assíncrono (*asynchronous timer*) continua funcionando, permitindo o usuário manter uma base de tempo, enquanto o resto

do dispositivo repousa. O modo *ADC Noise Reduction* para a CPU e todos os módulos de entrada/saída exceto o temporizador assíncrono e o ADC para minimizar o ruído de chaveamento durante as conversões ADC. No modo *Standby*, o oscilador cristal/ressonador fica funcionando enquanto o resto do dispositivo repousa. Isso permite um rápido reinício combinado com o baixo consumo. Em modo *Extended Standby*, ambos o oscilador principal e o temporizador assíncrono continuam funcionando.

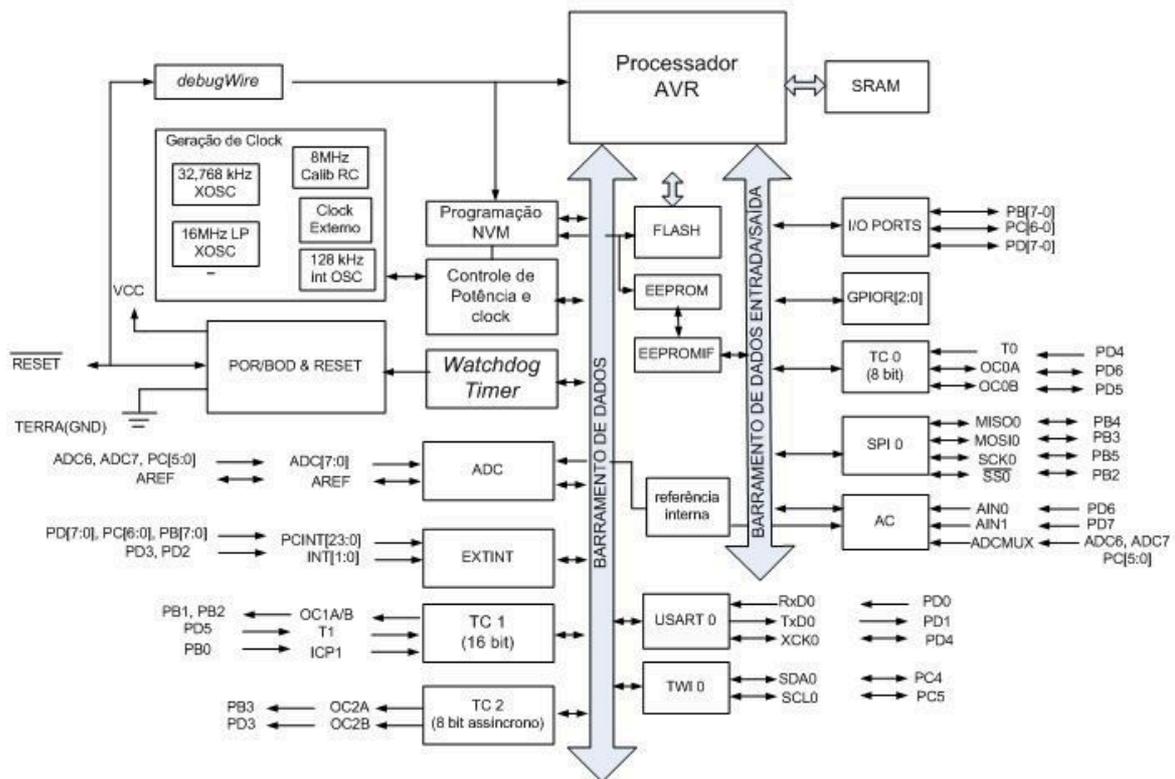


Figura 27. Diagrama de blocos do microcontrolador ATmega328.

(PCINT14/RESET) PC6	1	28	PC5 (ADC5/SCL/PCINT13)
(PCINT16/RXD) PD0	2	27	PC4 (ADC4/SDA/PCINT12)
(PCINT17/TXD) PD1	3	26	PC3 (ADC3/PCINT11)
(PCINT18/INT0) PD2	4	25	PC2 (ADC2/PCINT10)
(PCINT19/OC2B/INT1) PD3	5	24	PC1 (ADC1/PCINT9)
(PCINT20/XCK/T0) PD4	6	23	PC0 (ADC0/PCINT8)
VCC	7	22	GND
GND	8	21	AREF
(PCINT6/XTAL1/TOSC1) PB6	9	20	AVCC
(PCINT7/XTAL2/TOSC2) PB7	10	19	PB5 (SCK/PCINT5)
(PCINT21/OC0B/T1) PD5	11	18	PB4 (MISO/PCINT4)
(PCINT22/OC0A/AIN0) PD6	12	17	PB3 (MOSI/OC2A/PCINT3)
(PCINT23/AIN1) PD7	13	16	PB2 (SS/OC1B/PCINT2)
(PCINT0/CLK0/ICP1) PB0	14	15	PB1 (OC1A/PCINT1)

Figura 28. Diagrama de pinos do microcontrolador *ATmega 328*, em empacotamento *dual-in-line* de 28 pinos (detalhes sobre os pinos em [Atmel 2017]).

A plataforma de desenvolvimento *Arduino*, Fig.29, é um dispositivo que usa o microcontrolador *ATmega328*, e contém os principais recursos periféricos de um sistema de desenvolvimento, como entrada e saída (E/S) geral, entradas para conversão analógico-para-

digital (ADC), entradas de alimentação, interface e conector USB, e botão de *reset*, além de conector para programação.

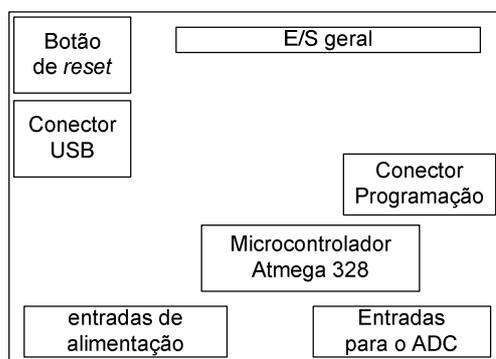


Figura 29. Plataforma de desenvolvimento Arduino.

Um exemplo de implementação da plataforma *Arduino* é mostrado na Fig.30, constituído de um microcontrolador com velocidade de *clock* de 16 MHz, memória *flash* de 32 KB, memória SRAM 2 KB e EEPROM 1 KB; um circuito para a comunicação serial USB formado por um cristal de 16MHz, dois LEDs para indicação de transmissão e recepção de dados serial e um *chip* de comunicação serial; um LED de comando (pino 13) possui um circuito integrado regulador de tensão, conector USB, conector *jack* para alimentação externa de 7 a 12V DC, um botão de *reset*, 6 pinos para programação serial em circuito (ICSP, *in-Circuit Serial Programming*) para o microcontrolador *ATmega 328*; além de 14 pinos digitais de entrada e saída, sendo seis para saída com modulação PWM e seis para entrada analógica.

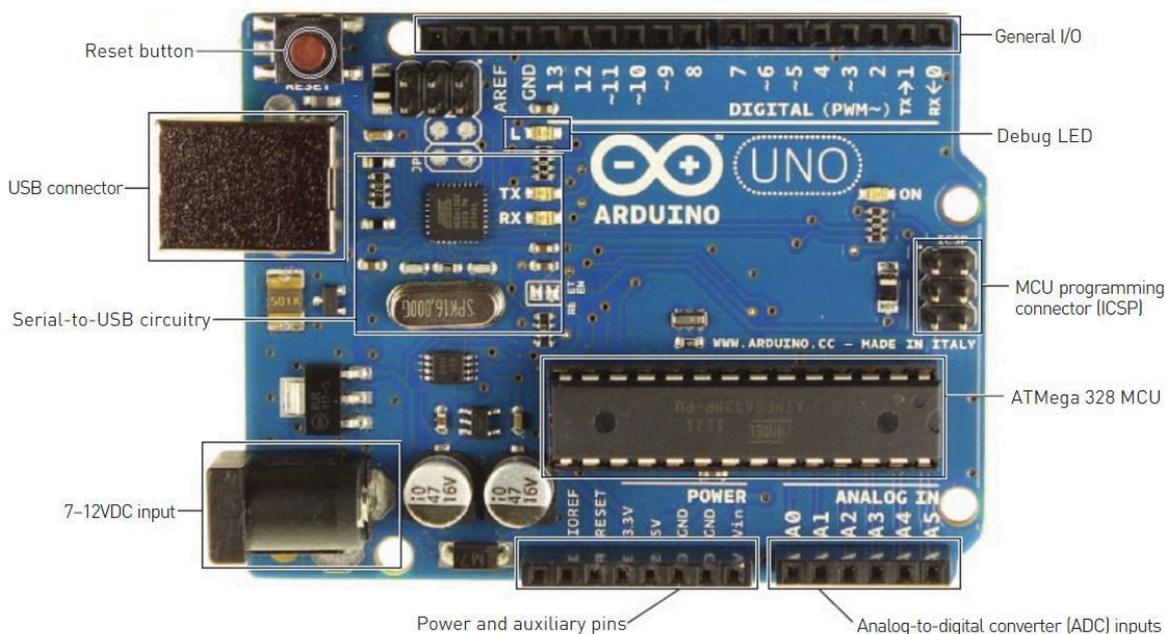


Figura 30. Ilustração do *Arduino Uno* [Blum 2013].

A descrição dos principais pinos são:

- (1) **Serial:** 0 (RX) recebe e 1 (TX) transmite dados seriais. são conectados aos pinos do *ATmega328* porta USB para TTL *serial*;
- (2) **Pino 2 e 3:** são interruptores externos, podendo ser configurados para disparar uma interrupção por borda de subida ou de descida ou por uma mudança de valor;
- (3) **Pinos 3, 5, 6, 9, 10 e 11- PWM:** Proveem 8 bits para saída PWM pela função *analogWrite()*;
- (4) **SPI: Pinos 10 (SS), 11 (MOSI), 12 (MISO), 13 (SCK).** Estes pinos suportam a comunicação SPI (Interface Serial Periférica);
- (5) **Pino 13 LED:** Este pino digital acende um LED quando colocado em alto; e
- (6) **Entrada analógica A0 a A5:** Cada uma com 10 bits de resolução (1024 valores diferentes).

Possui também outros pinos com funções específicas sendo eles:

- (1) **TWI: Pino A4 ou SDA e pino A5 ou SCL.** Comunicação Interface serial por duas linhas (em inglês, *Two Wire Interface-TWI*);
- (2) **AREF.** Pino de referência analógica para o conversor A/D. Usado com *analogReference()*;
- (3) **Reset.** Tipicamente usado para adicionar um botão de reposição para *shields* que bloqueiam o que há na placa.

A correspondência dos pinos descritos do *Arduino Uno* em relação aos pinos do microcontrolador *ATmega328* é ilustrada na Fig.31, onde na parte central podemos notar a descrição dos pinos do *ATmeg328* e nas extremidades os pinos do *Arduino Uno* em cor vermelha.

Função Arduino						Função Arduino
reset	(PCINT14/RESET) PC6	1	28	PC5 (ADC5/SCL/PCINT13)		analog input 5
digital pin 0 (RX)	(PCINT16/RXD) PD0	2	27	PC4 (ADC4/SDA/PCINT12)		analog input 4
digital pin 1 (TX)	(PCINT17/TXD) PD1	3	26	PC3 (ADC3/PCINT11)		analog input 3
digital pin 2	(PCINT18/INT0) PD2	4	25	PC2 (ADC2/PCINT10)		analog input 2
digital pin 3 (PWM)	(PCINT19/OC2B/INT1) PD3	5	24	PC1 (ADC1/PCINT9)		analog input 1
digital pin 4	(PCINT20/XCK/T0) PD4	6	23	PC0 (ADC0/PCINT8)		analog input 0
VCC	VCC	7	22	GND		GND
GND	GND	8	21	AREF		Referência analógica
(cristal)	(PCINT6/XTAL1/TOSC1) PB6	9	20	AVCC		VCC
digital pin 5 (PWM)	(PCINT7/XTAL2/TOSC2) PB7	10	19	PB5 (SCK/PCINT5)		digital pin 13
digital pin 6 (PWM)	(PCINT21/OC0B/T1) PD5	11	18	PB4 (MISO/PCINT4)		digital pin 12
digital pin 7	(PCINT22/OC0A/AIN0) PD6	12	17	PB3 (MOSI/OC2A/PCINT3)		digital pin 11 (PWM)
digital pin 8	(PCINT23/AIN1) PD7	13	16	PB2 ( $\overline{SS}$ /OC1B/PCINT2)		digital pin 10 (PWM)
	(PCINT0/CLK0/ICP1) PB0	14	15	PB1 (OC1A/PCINT1)		digital pin 9 (PWM)

Figura 31. Pinos *ATmega328* correspondentes aos pinos do *Arduino Uno*. Adaptação [Uno 2016].

A plataforma necessita de uma alimentação externa podendo ser via conector USB tipo B ou conector *jack* com tensões limites de entrada variando entre 6 a 20 V, com correntes de entrada e saída na faixa de 40mA. É recomendado utilizar uma tensão entre 7 a 12 V, pois se menor que 7 V o pino com regulador de saída de 5V pode operar com uma tensão menor. Já com uma alimentação maior que 12V o regulador de tensão pode superaquecer e danificar a placa. Os pinos de alimentação são:

- (1) **VIN:** é a tensão de entrada para a placa quando ele está usando uma fonte de alimentação externa;
- (2) **5V:** a saída é regulada em 5V pelo regulador de tensão na placa a partir de uma alimentação entre 7 a 5V;
- (3) **3,3V:** pino de saída regulada em 3,3V com corrente máxima de 50mA;
- (4) **GND:** Pinos de aterramento.

#### 4.7. Sistemas Operacionais e Simuladores

Nesta seção serão apresentados os *softwares* utilizados para o desenvolvimento do trabalho, que compreende o *App Inventor 2*, ambiente de desenvolvimento integrado do *Arduino* e o Sistema Operacional *Android*.

##### 4.7.1. App Inventor

A Plataforma de *software App Inventor* é uma aplicação de código aberto disponibilizada gratuitamente pela empresa Google em parceria com o Instituto de Tecnologia de Massachusetts (MIT), Estados Unidos, que permite, aos iniciantes de programação em computadores, criarem os aplicativos de *software* para o sistema operacional *Android*. Sua criação teve início em 2009, com a missão de democratizar o ensino de programação, de modo aos usuários não serem meros consumidores, mas sim criadores de tecnologia.

O *App Inventor* baseia-se na criação de aplicativos visualmente, por arrastar e soltar blocos, evitando uma linguagem complexa baseada em texto, permitindo a criação de aplicativos básicos e totalmente funcionais [Inventor 2016].

A parte gráfica é construída na tela de *designer* em cima de uma "tela para celular", no qual se escolhe a opção pela paleta no lado esquerdo, como pode ser observado na Fig.32, e arrasta até a tela do celular. Automaticamente, cria-se um bloquinho na outra tela do *software* para fazer as especificações do componente.



Figura 32. Programação gráfica no App Inventor 2.

#### 4.7.2. Ambiente de Desenvolvimento Integrado do *Arduino*

O ambiente de desenvolvimento integrado do *Arduino* (IDE, *Integrated Development Environment*) é um *software* de código aberto no qual qualquer pessoa pode contribuir e seu código fonte escrito em Java encontra-se hospedado no site GitHub. Esse *software* é disponível para os sistemas operacionais *Windows*, *Mac OS X* e *Linux*, e pode ser utilizado em qualquer placa *Arduino*. Outras informações sobre o *software* podem ser obtidas através do site *Arduino* [Arduino 2017].

#### 4.7.3. Sistema Operacional *Android*

Com a ideia de obter um sistema operacional de plataforma livre que pudesse ser adotada por vários fabricantes, sendo a solução para o aumento da utilização de seus recursos nos *smartphones*, a *Google* adquiriu a *Android Inc.* em 2005. O Sistema Operacional (SO) *Android* surgiu em meados de 2003 pelos desenvolvedores que fundaram a *Android Inc.*

O SO *Android*, baseado no *Kernel Linux*, é uma plataforma código aberto que possibilita o desenvolvimento de inúmeros aplicativos de forma simples pelos usuários, a partir do *kit* de desenvolvimento de *software* (*Software Development Kit*, SDK). As empresas se beneficiaram com o custo reduzido do desenvolvimento de *software* para o mercado de *smartphones* e em

consequência, para os consumidores adquirirem aparelhos com mais recursos e a custos baixos [Morimoto 2009].

A arquitetura do SO *Android* é estruturada em camadas, do mais alto nível que são as aplicações (*applications*) até a de mais baixo nível (*Linux Kernel*), sendo:

- (1) Aplicações (*Applications*): São desenvolvidas em Java, tendo como aplicações nativas calendário, calculadora, agenda, bússola, entre outros;
- (2) *Framework* de Aplicações (*Application Framework*): Nesta camada se encontram as Interfaces de Programação de Aplicativos (*Application Programming Interfaces, APIs*) que são os elementos necessários para a manipulação de aplicativos. Todas as aplicações presentes no sistema podem ter seus recursos utilizados por outras aplicações, integrando as *APIs Android*.
- (3) Bibliotecas (*Libraries*): É um conjunto de bibliotecas em linguagem C/C++ utilizadas pelo sistema, bibliotecas para as áreas de: multimídia, visualização de camadas 2D e 3D, funções para navegadores *Web*, funções para gráficos, funções de aceleração de *hardware*, entre outros;
- (4) Ambiente de Execução (*Android Runtime*): Dividido em Máquina Virtual *Dalvik* (*Dalvik Virtual Machine*), em que para cada aplicação executada é criada uma pequena camada da máquina virtual, ou seja, a aplicação executa dentro de seu próprio processo encapsulada no contexto da sua máquina virtual; e as bibliotecas (*Core Libraries*) que são responsáveis pela atribuição das funcionalidades disponíveis nas principais bibliotecas em linguagem Java como estrutura de dados, acesso a arquivos, rede e gráficos;
- (5) *Linux Kernel*: É o núcleo do SO, em torno do qual os programas são escritos. Executa o gerenciamento dos serviços centrais do sistema como segurança, gestão de memória, gestão de processos, pilha de protocolos de rede e modelo de *drivers*. O *kernel* exerce ainda a atividade de separação entre o *hardware* e o restante da arquitetura.

O SO *Android* surgiu visando permitir aos desenvolvedores criarem aplicativos capazes de explorar todos os recursos disponíveis, uma vez que não diferencia os aplicativos nativos e o de terceiros.

#### **4.8. Considerações Finais**

Neste capítulo foram apresentados os principais conceitos sobre uma das ferramentas básicas para a BSS, a Análise de Componentes Independentes, ICA. Foi apresentado um algoritmo comumente utilizado para ICA, *FastICA*, que realiza as operações de ICA, em tempo reduzido e também, foi apresentado um algoritmo que explora o domínio da frequência para a separação de fontes de sinais de misturas convolutivas. Foram descritos alguns principais conceitos relacionados ao reconhecimento automático de fala; e conceitos sobre a tecnologia

de *hardware* e *software* utilizados no desenvolvimento do protótipo. A seguir, no capítulo seguinte, será apresentado o desenvolvimento do SADA.

## CAPÍTULO 5

# DELINEAMENTO DE UM SISTEMA DE AUXÍLIO A DEFICIENTES AUDITIVOS

---

Este capítulo apresenta o delineamento de um Sistema de Auxílio a Deficientes Auditivos, SADA. Será apresentada a caracterização do sistema, apontando os tópicos investigados. Na Seção 5.1 está descrito o sistema de auxílio a deficientes auditivos com as características gerais do protótipo de sistema proposto. Na Seção 5.2 estão descritos os aspectos sobre o *hardware* de comunicação no SADA; na Seção 5.3 são descritos alguns detalhes sobre o software de comunicação e sensibilização do DA; na Seção 5.4 é apresentada as considerações finais.

### 5.1 Desenvolvimento do SADA

Pode-se definir que um Sistema de Auxílio a Deficientes Auditivos (SADA) genérico deve fazer uso de sensores; de sistemas computacionais localizados, ou remotos, como em computação em nuvem, para o processamento de sinais captados pelos sensores; de seleção dos sinais captados com o uso de um classificador; e de dispositivos atuadores que sinalizam as pessoas, com o resultado do processamento de sinais, além da interação dos DAs com as pessoas envolvidas. Ao mesmo tempo o sistema, bem como os DAs, podem informar os departamentos pertinentes como corpo de bombeiros, hospitais, ambulâncias e polícia, sobre algum acontecimento, se for o caso. A Fig.33 ilustra um sistema de SADA proposto pelo autor, constituído das seguintes atividades principais: (A) captação de sinais de vários sensores embarcados; (B) processamento de sinais para os sinais captados no item A; (C) classificação dos sinais captados e processados no item B; e (D) sinalização aos DAs pelos dispositivos atuadores e interação dos DAs com as pessoas usando sistemas como de tradução de texto em falas, com o uso de *smartphones*.

O protótipo de SADA proposto para o presente trabalho implementa essas quatro atividades, com as seguintes características:

**Atividade A** – (a) captação dos sinais de vários microfones com a intenção de obter em um ambiente ruidoso a fala de um único indivíduo que esteja transmitindo uma informação dirigida ao DA, como sinalização de atenção a algum evento relevante em ocorrência, ou prestes a

ocorrer; (b) e captação de sinais de diversos dispositivos sensores específicos como sirenes ou campainhas, que não são normalmente ouvidos pelos DAs; (c) os sinais dos microfones e sensores diversos são transferidos a um computador por meio de um sistema de rádio (RF), composto de módulos de transmissão TX e de recepção RX, como descritos na Seção 4.5.

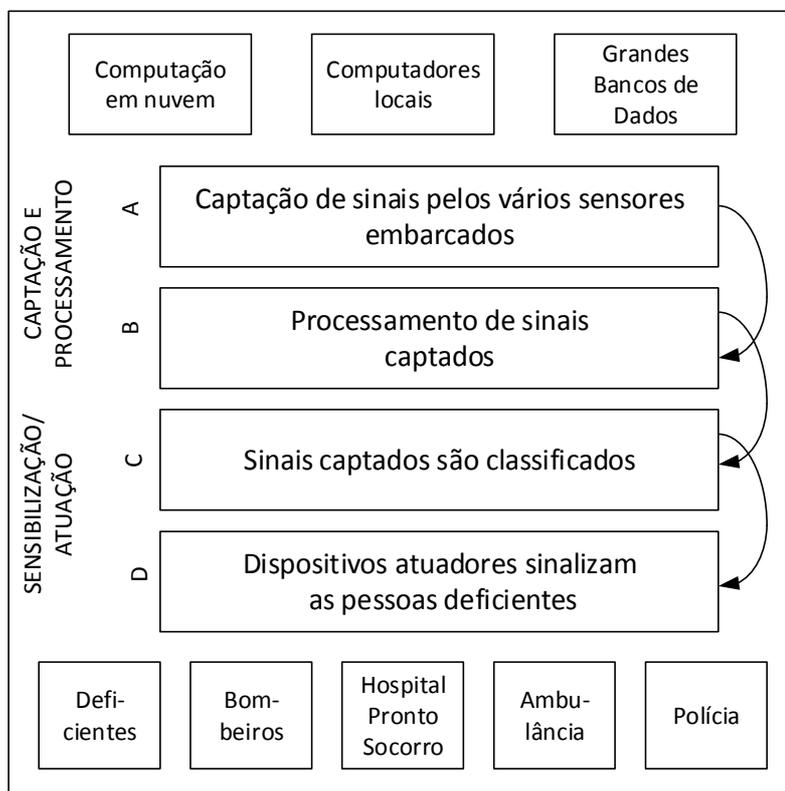


Figura 33. Diagrama de um Sistema de Auxílio a Deficientes Auditivos (SADA) proposto.

**Atividade B** - consiste em realizar os processamentos de sinais de áudio captados pelos microfones, usando pré-processamentos, ICA, e reconhecimento de fala, para posterior classificação de expressões relevantes para o SADA;

**Atividade C** - consiste no processamento para a detecção de expressões de interesse para o SADA, usando amostras pré-cadastradas e armazenadas em banco de dados; e

**Atividade D** - envolve o aparelho celular (*smartphone*) de posse dos DAs, que deve receber os sinais e expressões de alerta classificados pelo SADA na atividade C, além dos sinais captados pelos sensores específicos como sirenes ou campainhas, usando um meio de transmissão sem fio como *Bluetooth*, descrito na Seção 4.5.2. O celular deve usar o recurso de vibração (*vibracall*) para possibilitar a sinalização. Além disso, os DAs podem interagir com as pessoas usando o *smartphone* e sistemas como o de conversão de texto em fala.

Em resumo, o protótipo de sistema SADA proposto é composto de: (1) Captação e Processamento - uma parte relativa à captação de sinais por diversos sensores e microfones, um

sistema de rádio para transmissão de dados para o computador o processamento de sinais; e (2) Sensibilização - uma parte relativa à sinalização dos dispositivos celulares de posse dos DAs, por alerta visual por texto e vibração. O protótipo do sistema SADA será denotado neste trabalho de **SADA Ouvitronic-1**, pois o sistema é proposto para substituir em parte a função do ouvido do DA.

Se a mensagem recebida pelo celular do DA corresponder a uma das frases de alerta, salvas no banco de dados do aplicativo, o celular irá vibrar com maior intensidade, permitindo assim, captar a atenção da pessoa com DA. O sistema pode também conter frases prontas no banco de dados, para reprodução sonora como meio facilitador de comunicação com pessoas ouvintes, caso o DA tenha também necessidade desse tipo de auxílio.

O sistema é ilustrado no diagrama da Fig.34, onde se observa a interação entre o sistema e o usuário por meio do celular (ou *smartwatch*), na recepção do alerta visual e sensorial e na função do aplicativo de conversão de texto para áudio. Nota-se que o sistema suporta um número de K sensores de som/ transmissores RF para z alarmes, bem como o suporte para N microfones.

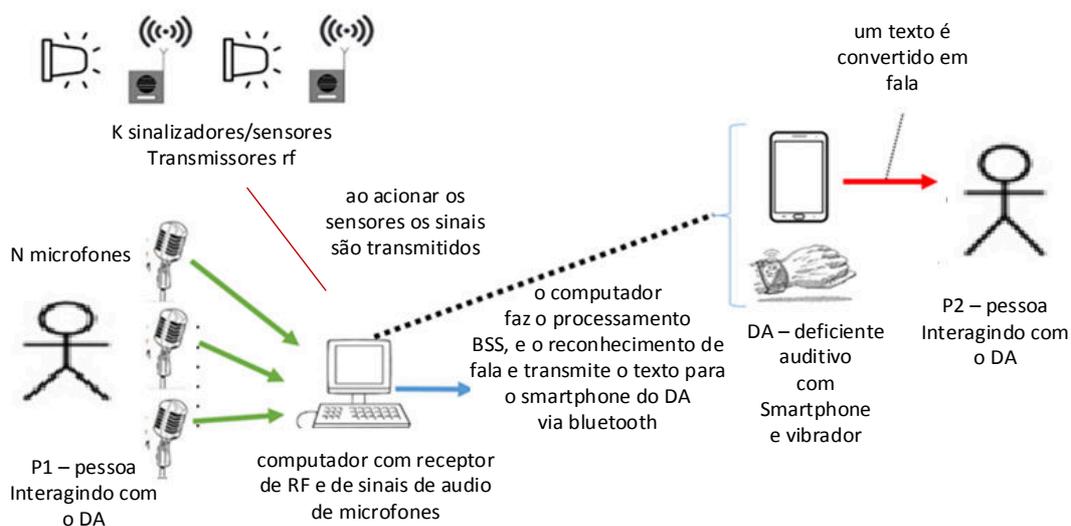


Figura 34. Diagrama ilustrativo do protótipo de SADA Ouvitronic-1, onde na metade esquerda tem-se a parte de captação e processamento e na metade à direita, a parte de sensibilização e interação do DA.

O processamento dos sinais recebidos dos microfones dar-se-á pela aplicação do método no domínio da frequência para misturas convolutivas, Algoritmo-2, para se extrair os componentes independentes dos sinais da mistura. O objetivo é que o sistema tenha a capacidade de obter os componentes resultantes dos sinais de fala do ambiente do DA, filtrando os ruídos e separando a fala de algum indivíduo (P1) que esteja querendo transmitir uma informação importante para o DA. Com a separação da fala do indivíduo e aplicação do

reconhecimento de fala, é extraído o conteúdo textual de interesse. De posse do texto, este é transmitido, via serial, para a plataforma *Arduino* que se encarregará de enviar, por comunicação *Bluetooth*, para o aparelho celular do DA.

Para que ocorra o processamento dos sinais recebidos dos microfones, é necessário que o celular permaneça com o *Bluetooth* pareado com o módulo, para que haja o recebimento de mensagens. O sistema se encarrega de notificar o usuário por meio de vibração e de exibir e salvar no banco de dados os textos recebidos. O usuário DA também poderá apagar as mensagens salvas.

Durante a interação com uma pessoa, por exemplo (P2), o DA pode também escrever um texto e convertê-lo em fala com o auxílio de um aplicativo no smartphone.

O sistema proposto poderá ser usado em qualquer ambiente fechado, seco e sem exposição direta ao sol, seja em uma sala de trabalho ou em uma residência, para atender a um DA, desde que respeitem as limitações físicas dos componentes.

Os componentes supracitados que se comunicam por radiofrequência, como o transmissor e receptor de som para alarmes, devem estar posicionados numa distância máxima de aproximadamente 100 metros. Já o *hardware* que conta com o módulo *Bluetooth*, não deve estar a uma distância maior que 10 metros com o celular, sendo este o alcance máximo para que o módulo se comunique com o celular.

A seguir será caracterizado o sistema físico SADA **Ouvitronic-1**, com o detalhamento de funcionamento, circuito elétrico de ligação entre os componentes e o projeto em seu todo. Logo depois, será descrito o sistema de *software* composto do detalhamento para o desenvolvimento do aplicativo para *Android* e o algoritmo em Linguagem C para a comunicação entre os componentes.

## **5.2. Caracterização do Sistema Físico do SADA Ouvitronic-1**

O sistema físico de transmissão de dados sem fio entre os sensores de som posicionados diante a uma sirene de um sistema de segurança, por exemplo, contém basicamente um sensor que capta os sinais do ambiente, uma plataforma de processamento *Arduino* e um transmissor de sinais em radiofrequência, conforme diagrama da Fig.35.



Figura 35. Sistema físico de captação e transmissão de sinais.

O módulo de RF com banda de 433MHz foi escolhido para realizar a transmissão de dados sem fio entre os sensores, pois possibilita a transmissão de dados a uma boa distância com um ótimo custo benefício.

Para a recepção dos dados RF será utilizado um sistema conforme mostra o diagrama de blocos da Fig.36, onde um módulo RX recebe os sinais de RF da antena e, por intermédio da plataforma *Arduino*, os sinais são enviados para o computador central.

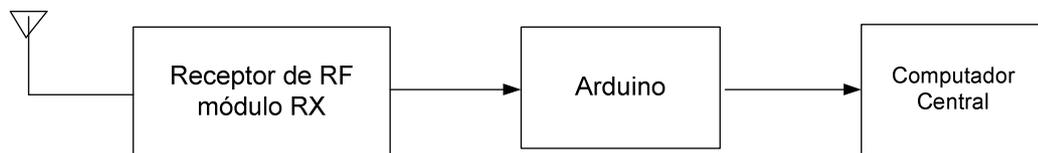


Figura 36. Sistema físico para recepção dos dados RF pelo computador central.

O sistema de transmissão de dados para o celular utiliza o protocolo *Bluetooth* conforme ilustrado na Fig.37, onde o computador central através da plataforma *Arduino* envia os sinais para o módulo de transmissão *Bluetooth*, que por sua vez, transmite o sinal por meio de radiofrequência apropriada.

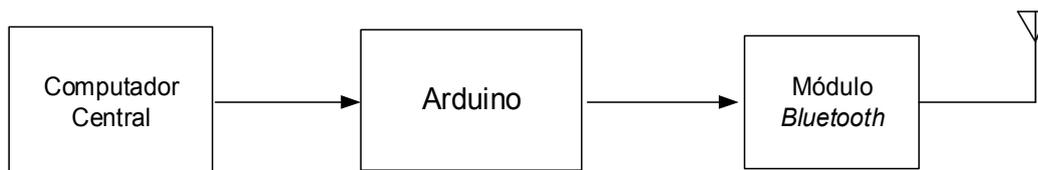


Figura 37. Sistema físico para transmissão de dados do computador central para o celular, por meio do protocolo *Bluetooth*.

Conforme descrito no Capítulo 4, o *Bluetooth* é um meio de comunicação de radiofrequência com um espectro de frequência de 2,4GHz, sendo um tipo de transmissão de dados eficiente com ampla adaptação no mercado, para o qual os equipamentos de uso pessoal, tais como *smartphones* e *smartwatch*, possuem o suporte para a transmissão de dados ponto a ponto em uma rede WPAN.

A Fig.38 é um detalhamento da Fig.34, em que apresenta em detalhes o sistema SADA Ouvitronic-1 proposto, com a disposição dos módulos de transmissão e recepção de RF; com o computador central para a recepção de sinais de fala dos microfones, para a recepção dos sinais

recebidos dos módulos RF, e para o processamento e transmissão de dados para o DA, em protocolo *Bluetooth*.

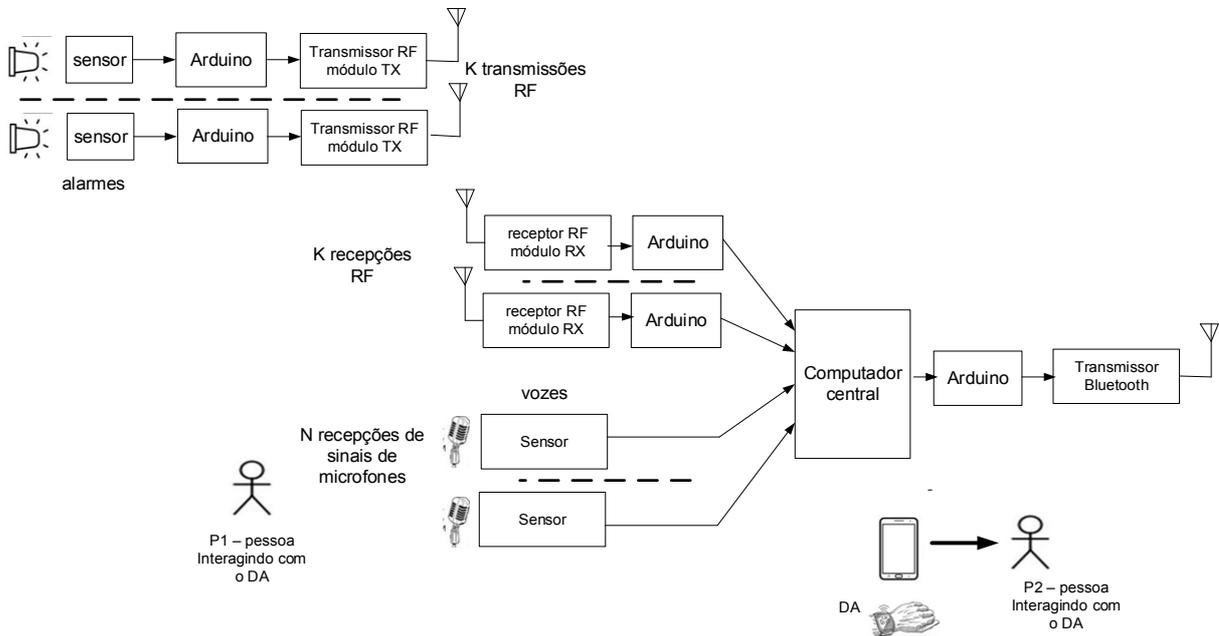


Figura 38. Diagrama de blocos do sistema SADA Ouvitronic-1 com a disposição dos módulos de transmissão e recepção de RF e computador central para processamento e transmissão de dados para o DA.

A plataforma de desenvolvimento *Arduino* foi escolhida porque além de seu tamanho ser reduzido, possui o suporte de fácil acesso para a integração de tecnologias como RF, módulo *Bluetooth*, sensor de som, dentre outras.

### 5.3. Ilustração de Uso da Plataforma *Arduino* para Transmissão e Recepção de RF

A Fig.39 apresenta a ligação do circuito transmissor RF (módulo TX) e a plataforma *Arduino*. Observa-se a utilização de uma chave digital denominada *logicstate* que ativa os estados 0 (nível baixo 0V) e 1 (nível alto 5V) para a plataforma *Arduino*, simulando a saída digital do módulo sensor de som.

A Fig.40 ilustra o desenho do circuito receptor RF (módulo RX) e de comunicação *Bluetooth*. Para conectar o módulo de comunicação *Bluetooth* HC-05 na porta RX é necessário utilizar um divisor de tensão, de modo a sua entrada seja o valor de 3,3V, conforme mostrado. Foram utilizados dois resistores R1 e R2, de 2kΩ e 1kΩ, respectivamente, com base na equação (4.24):

$$V_{out} = \frac{V_{in} * R1}{R1 + R2} \quad (4.24)$$

Vin=5V (Pino 3 *Arduino* denotado como TX)

$$V_{out} = \frac{5 * 2000}{2000 + 1000} = 3,3V \quad (4.25)$$

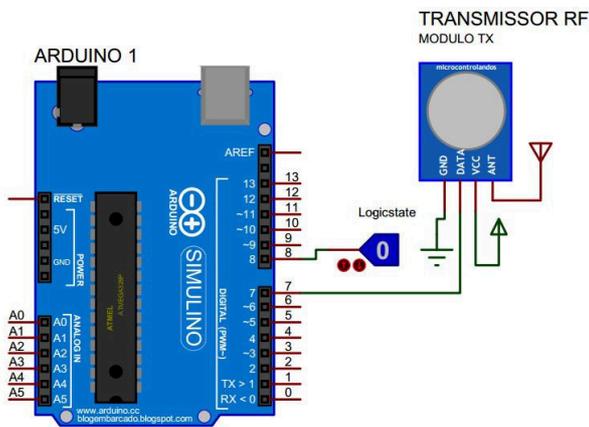


Figura 39. Circuito transmissor RF (módulo TX) do protótipo.

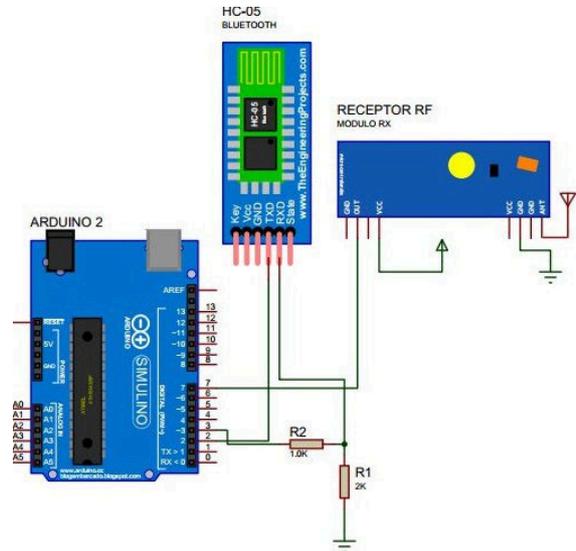


Figura 40. Circuito receptor RF (módulo RX) e módulo *Bluetooth* (HC-05) do protótipo.

#### 5.4. Caracterização do *Software* de Comunicação no SADA Ouvitronic-1

O sistema de *software* do SADA Ouvitronic-1 será composto de duas partes, sendo: (1) o aplicativo Ouvitronic-1, em *Android*, para *smartphone*, e (2) o programa em linguagem C que faz a comunicação entre a plataforma *Arduino* e o celular. A seguir serão apresentados os fluxogramas e a descrição da lógica de funcionamento.

##### 5.4.1. Aplicativo SADA Ouvitronic-1 em *Android* para *Smartphone*

Para o desenvolvimento do aplicativo, em *Android*, foi utilizada a plataforma de desenvolvimento *Web App Inventor 2*, de acordo com o fluxograma na Fig.41. Observa-se que a primeira ação assim que o sistema se inicia é o carregamento das bibliotecas e variáveis internas.

Logo, surge a primeira interação do usuário. Quando o *Bluetooth* do celular estiver desativado o sistema gera uma mensagem: “Um aplicativo quer ativar o *Bluetooth* neste dispositivo”, momento em que o usuário deve permitir ou negar.

Se a interação for permitida o celular acionará o *Bluetooth*, caso contrário o aplicativo volta para a tela de início e permanecerá em espera até que o usuário faça uma ação. Para que o aplicativo funcione na íntegra, será necessária a seleção do módulo *Bluetooth* a ser pareado, associando-o a uma frase para identificação em “*Cadastro de módulos*”, no caso o módulo que integra o *hardware* do protótipo. Para tanto o celular deverá estar com o *Bluetooth* acionado.

Após a seleção do módulo, aparecerá a seguinte mensagem “*Conexão Bluetooth estabelecida*”. O usuário também deverá configurar a mensagem pré-cadastrada com o dispositivo receptor. Por exemplo, dentre as mensagens pré-cadastradas para recepção do som tem-se “*A campanha está tocando*”, “*Alerta de alarme acionado*” e “*Bebê em prantos*”. Para cada módulo receptor o usuário deverá associar a mensagem correspondente sendo por exemplo:

Módulo 1 “*A campanha está tocando*”;

Módulo 2 “*Bebê em prantos*”;

Módulo 3 “*Alerta de alarme acionado*”.

Assim o aplicativo ficará em modo de espera até o usuário realizar uma ação ou ao ser interceptado pela mensagem recebida do módulo ou do *hardware* com o texto convertido a partir dos sinais de áudio captados dos microfones.

Também, o usuário poderá pré-cadastrar algumas palavras para o banco de dados do aplicativo, de modo que ao receber uma mensagem do *hardware* o aplicativo compare com as palavras cadastradas do banco de dados e se for igual o celular irá vibrar, intensamente, até o usuário interagir desativando a notificação. Caso a palavra não esteja cadastrada no banco, apenas irá notificar por vibração por um determinado tempo e a mensagem ficará salva.

Dentre outras ações, o usuário poderá utilizar o modo de conversão de texto para voz, a partir da tela denominada “*Falar*”. Nesta tela será possível salvar frases no banco de dados do aplicativo ao clicar no botão “*Adicionar lista de frases*”, no qual caso o DA desejar reproduzi-la em voz computadorizada, basta clicar em cima do texto. Também, será possível reproduzir as frases digitadas no campo de texto ao clicar no botão “*Falar*”.



**Algoritmo-3:**

**% parte de inicialização e controle**

1. Início
2. Inicializa as variáveis e bibliotecas
3. Se *Bluetooth* estiver ligado
4. Exibe a tela de início
5. Sistema fica em espera
6. Senão exibe a mensagem de solicitação “*Um app quer ativar o Bluetooth no dispositivo*”
7. Se for permitido à solicitação
8. Ativa o *Bluetooth*
9. Retorna a verificação do *Bluetooth*
10. Senão exibe a tela de início
11. Se for recebido uma mensagem
12. Procura texto correspondente no banco de dados
13. Se texto for correspondente ao módulo cadastrado
14. Aciona o *vibracall* por 10 segundos
15. Exibe a mensagem no visualizador de texto
16. Salva mensagem no banco de dados
17. Se texto recebido for igual a texto cadastrado no banco de dados
18. Aciona o *vibracall* por 5 segundos
19. Exibe a mensagem no visualizador de texto
20. Salva mensagem no banco de dados
21. Senão Aciona o *vibracall* por 1 segundo
22. Exibe a mensagem no visualizador de texto
23. Salva mensagem no banco de dados
24. Se botão “Falar” for acionado
25. Exibe a tela função “Falar”
26. Se o botão “Cadastro de Módulo” for acionado
27. Abre a tela “Cadastro de Módulo”
28. Se o botão “Cadastro de palavras” for acionado
29. Abre a tela “cadastro de palavras”

**% fim da parte de inicialização e controle**

### **% Tela Função “Falar”**

1. Se botão “Falar” for acionado
2. Aciona o reconhecimento de texto para fala
3. Se botão “adicionar lista de frases” for acionado
4. Adiciona o texto no banco de dados
5. Se for selecionado uma frase na lista de texto salva
6. Aciona o reconhecimento de texto para fala
7. Se botão “remoção de frases for acionado”
8. Limpa a mensagem selecionada no banco de dados
9. Se botão “voltar” for acionado
10. Retorna a tela de início

### **% Fim da função “Falar”**

### **% Tela Função Cadastro de Módulos**

1. Se o botão “Cadastro de Módulo” for acionado
2. Abre a tela “Cadastro de Módulo”
3. Se botão “Adicionar” for acionado
4. Incrementa o contador de módulo
5. Exibe a mensagem “Módulo cadastrado com sucesso”
6. Adiciona o texto ao banco de dados
7. Exibe na lista de visualização
8. Se for habilitado o campo remoção de palavras
9. Apaga as frases selecionadas
10. Se botão “voltar” for acionado
11. Retorna a tela de início

### **% Fim da Função Cadastro de Módulos**

### **% Tela Função Cadastro de Palavras**

1. Se botão “Adicionar” for acionado
2. Adiciona o texto ao banco de dados
3. Exibe na lista de visualização
4. Se for habilitado o campo remoção de palavras
5. Apaga as frases selecionadas
6. Se botão “voltar” for acionado
7. Retorna a tela de início

### **% Fim da Função Cadastro de Palavras**

### 5.4.2. Sistema de Comunicação *Arduino* e Aplicativo

O algoritmo que faz a transmissão por RF do sinal do microfone para o receptor, tem como base o fluxograma do transmissor RF, conforme Fig.42.

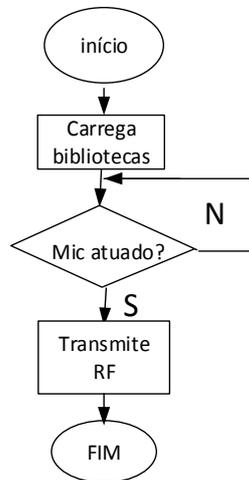


Figura 42. Fluxograma para o transmissor RF.

Como pode ser observado no fluxograma do transmissor RF, o sistema inicialmente carregará as bibliotecas para a transmissão. Em seguida verifica se o pino digital do microfone está recebendo 0 (baixo) ou 1 (alto). Caso esteja em alto o *Arduino* atua o pino do transmissor RF, caso contrário o sistema fica em espera.

O algoritmo em linguagem natural para a transmissão em RF, com base no fluxograma da Fig.42 é dado por **Algoritmo-4**.

Algoritmo-4: Possibilita a transmissão em RF.

#### **Algoritmo-4: Transmissão de RF**

1. Início
2. Carrega as bibliotecas
3. Se microfone for atuado
4. Transmite o sinal em RF
5. Senão verifica novamente
6. Fim

A Fig.43 é o fluxograma do receptor RF, onde é possível observar que a primeira ação do programa será de carregar as bibliotecas, em seguida verificar se o sinal RF foi recebido, caso contrário, o sistema ficará em espera, e se sim, verificará se o módulo *Bluetooth* estará pareado. Se o módulo não estiver pareado o sistema ficará em espera, se pareado o módulo transmitirá a seguinte mensagem para o celular “A campanha está tocando”.

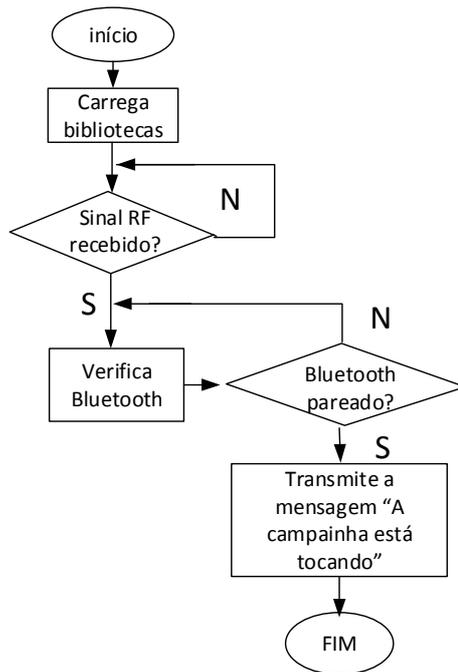


Figura 43. Fluxograma para o receptor RF.

O algoritmo em linguagem natural para a recepção em RF e transmissão *Bluetooth*, com base no fluxograma da Fig.43 é dado pelo **Algoritmo-5**.

Algoritmo-5: Recepção do sinal em RF e transmissão de dados por *Bluetooth*.

**Algoritmo-5: Receptor de sinal por RF e transmissão *Bluetooth***

1. Início
2. Carrega as bibliotecas
3. Se sinal de transmissão por RF for recebido
4. Verifica a conexão entre o módulo *Bluetooth* e o celular
5. Se dispositivos estiverem pareados
6. Transmite a mensagem “A campanha está tocando” por *Bluetooth*
7. Senão verifica a conexão entre o módulo *Bluetooth* e o celular
8. Senão verifica se o sinal de transmissão por RF foi recebido
9. Fim

Para que o *Arduino* se comunique com o aplicativo para o celular e com os dispositivos que integram o protótipo, foi necessário utilizar uma biblioteca para a comunicação RF denominada de *VirtualWire*, e outra para a comunicação *Bluetooth* denominada “*softwareSerial*”.

A biblioteca *VirtualWire* como o nome já induz, permite a comunicação sem fio. Faz uso da interrupção *Timer1*, o que é necessário utilizar uma porta digital do *Arduino* [McCauley 2013]. Dentre os principais comandos da biblioteca, devem ser utilizadas as funções a seguir:

1. `vw_set_tx_pin` (`pino_transmissor`): configura o pino de transmissão ( pino 7);
2. `vw_set_rx_pin` (`pino_receptor`): configura o pino de Recepção ( pino 7);
3. `vw_setup` (2000): inicializa a biblioteca, mas todos os pinos devem estar configurados antes de chamar essa função;
4. `vw_send` (`mensagem`, `comprimento`): transmite a mensagem, no qual a “*mensagem*” é uma matriz de *bytes* a enviar e “*comprimento*” é o número de *bytes* armazenados na matriz. Essa função retorna imediatamente, enquanto a mensagem é enviada lentamente pelo processo de interrupção de fundo;
5. `vw_wait_tx` (): aguarda a mensagem ser totalmente transmitida;
6. `vw_rx_start` (): ativa o processo de recepção. Essa função deve ser chamada antes de qualquer outra função de recepção. A recepção dos dados é monitorada por um processo de interrupção de fundo;
7. `vw_get_message` (`message`, `&msgLength`): pega a última mensagem recebida. A função retorna verdadeiro se a mensagem foi verificada e falso se a mensagem recebida estiver corrompida. Dentro dos parênteses, “*message*” é uma matriz onde a mensagem é copiada e “*&msgLength*” é o tamanho máximo da matriz de entrada.

O *hardware* do *Arduino*, possui o suporte para a comunicação serial nos pinos 0 e 1 (RX e TX, respectivamente) e via conexão USB com o computador. O suporte se dá por um componente de *hardware* Transmissor/Receptor Assíncrono Universal, UART (*Universal Asynchronous Receiver/Transmitter*), que permite que o microcontrolador *Atmega* receba a comunicação serial, mesmo quando trabalha em outras tarefas.

Contudo, pode ser criada a biblioteca *SoftwareSerial* a fim de permitir a comunicação serial por meio dos demais pinos digitais do *Arduino*, que realiza a função de duplicar a funcionalidade dos pinos seriais por meio de *software*. A velocidade alcançada é de até 115.200 *bytes* por segundo. A utilização de várias portas com *software* serial impossibilita o recebimento de todos os dados de uma vez, portanto apenas uma por vez recebe os dados [Arduino 2016].

A conexão física deve possuir as seguintes conexões:

1. O pino TX do módulo *Bluetooth* conectado ao pino do *Arduino* que estiver sendo usado como RX pela biblioteca *SoftwareSerial*;
2. O pino RX do módulo *Bluetooth* conectado ao pino do *Arduino* que estiver sendo usado como TX pela biblioteca *SoftwareSerial*;

Para facilitar o desenvolvimento da programação, foi utilizada a biblioteca para a comunicação serial para os pinos 2 e 3 (RX e TX), sendo a porta TX do módulo conectado ao pino 2 do *Arduino* e a porta RX do módulo conectado ao pino 3 do *Arduino* através de um divisor de tensão, conforme anteriormente detalhado.

## **5.5. Considerações Finais**

Foi apresentado o delineamento de um Sistema de Auxílio a Deficientes Auditivos, SADA, como trabalho de Dissertação de Mestrado, descrevendo os aspectos sobre o *hardware* de comunicação; e software de comunicação e sensibilização do DA, com sinais pertinentes de assistência. No capítulo a seguir serão apresentados os experimentos realizados para validação da proposta, como os resultados de processamento de sinais para BSS e reconhecimento de fala, seguido dos resultados do protótipo de *hardware* do SADA Ouvitronic-1.

## CAPÍTULO 6 RESULTADOS E VALIDAÇÃO

Neste capítulo serão apresentados os resultados de atividades previstas no projeto proposto. Na Seção 6.1 são descritas as partes do pré-protótipo de *hardware* do SADA, consideradas. Na Seção 6.2 são descritos os experimentos em processamento de sinais para misturas instantâneas e reconhecimento de fala; e em seguida, a descrição e os resultados para os experimentos com misturas convolutivas, na Seção 6.3, realizando os experimentos práticos de captação de sinais com dois microfones, para efeitos de validação do sistema. Na sequência, são apresentadas as considerações finais, Seção 6.4.

### 6.1. Descrição do Pré-Protótipo de *Hardware*

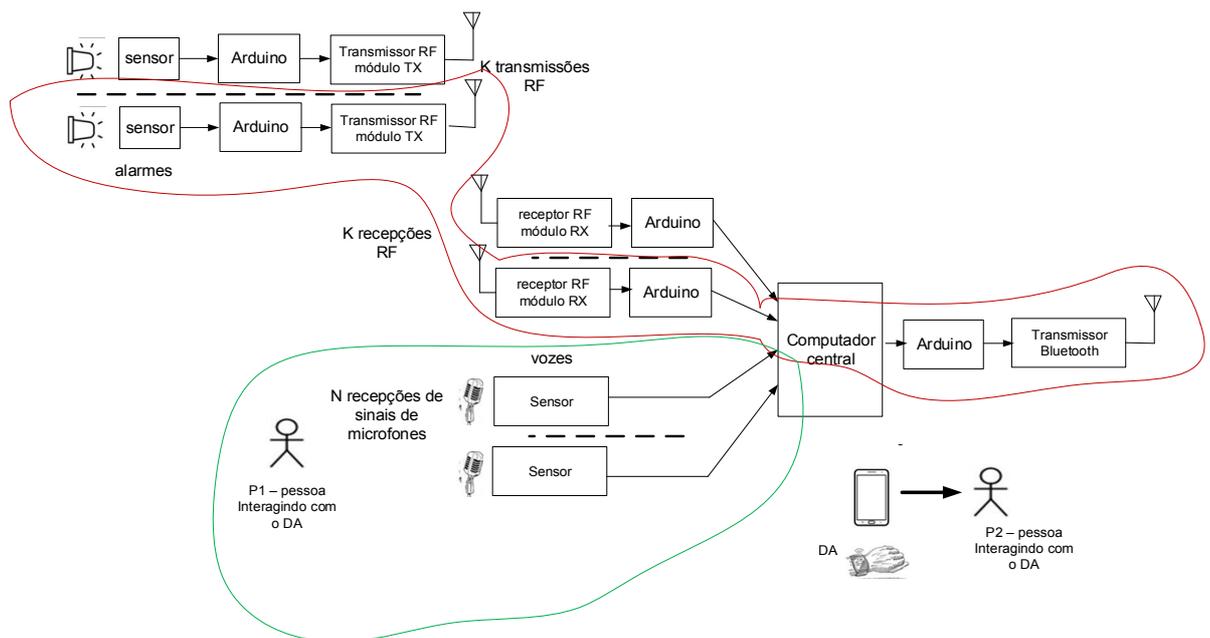


Figura 44. Diagrama ilustrativo da parte do *hardware* considerada no pré-protótipo.

Na Fig.44 é feita uma ilustração das partes do SADA Ouvitronic desenvolvidas como pré-protótipo para testes: a) PARTE 1, corresponde aos blocos contornados em linha vermelha, que consiste na captação de alarme por sensor, seguida de transmissão e recepção RF para ser enviado ao DA por meio de comunicação *Bluetooth*; e b) PARTE 2, corresponde ao bloco contornado em verde que consiste de dois microfones interligados ao computador, por meio de dois canais de áudio, para os testes de BSS, com sinais convolutivos.

Na Fig.45 (a) é feita uma ilustração com uma foto que caracteriza os componentes da PARTE 1 do pré-protótipo, de sensibilização do SADA Ouvitronic-1, onde constam um sensor de som, um transmissor RF, conectado a uma plataforma de desenvolvimento (Arduino 1) e um receptor de RF conectado a uma outra plataforma (Arduino 2). A mesma plataforma de desenvolvimento, Arduino 2, está sendo também usada para a transmissão do alarme, via Bluetooth, para o DA. Na Fig.45 (b) mostra-se uma foto da PARTE 2 do pré-protótipo, mostrando dois sensores microfones tipo de mesa, marca Knup modelo Kp-903, sensibilidade 54db  $\pm$ 2db, Impedância 2,2k e Resposta de Frequência 20Hz a 20kHz, interligados por dois canais independentes de áudio, a um computador com SO *Windows 7*, processador Intel modelo Core i7 5<sup>o</sup> geração de 2,40 GHz e memória RAM 8GB.

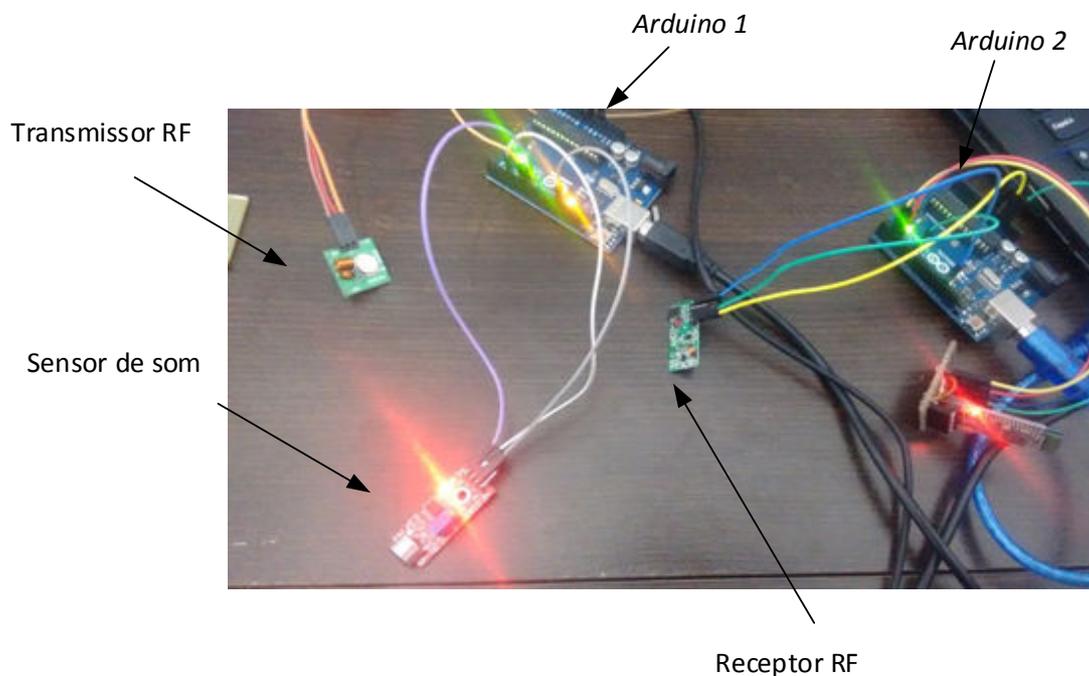


Figura 45. (a) PARTE 2 do pré-protótipo do SADA Ouvitronic.

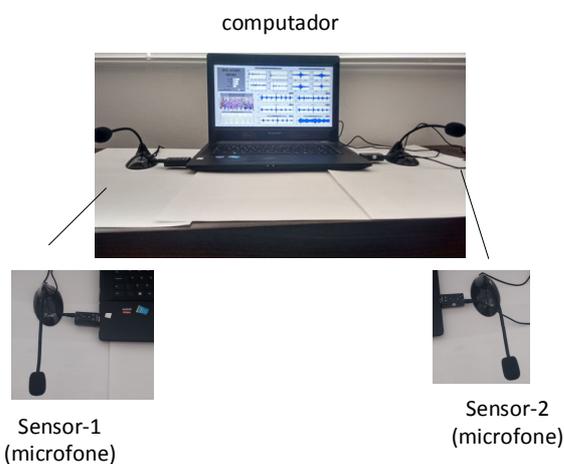


Figura 45. (b) PARTE 2 do pré-protótipo do SADA Ouvitronic.

## 6.2. Descrição do *Software* Desenvolvido

Com base no fluxograma da Fig.41, capítulo 5, foi desenvolvido o *App* para dispositivos *Android*, tendo a tela inicial do aplicativo *OuviTronic* apresentado pela Fig.45.

Assim que o aplicativo é inicializado, o mesmo exibe a pergunta na tela do *smartphone*: “Um *app* quer ativar o *Bluetooth* neste dispositivo” e aguarda a permissão ou recusa do usuário. Para que o DA localize os módulos *Bluetooth* disponíveis é necessário clicar no ícone de uma engrenagem e selecionar o dispositivo. Com o pareamento entre o celular e o módulo *Bluetooth*, o aplicativo exibe na tela uma mensagem na cor verde “Conexão estabelecida”, caso contrário “*Bluetooth* não conectada” conforme a Fig. 46.

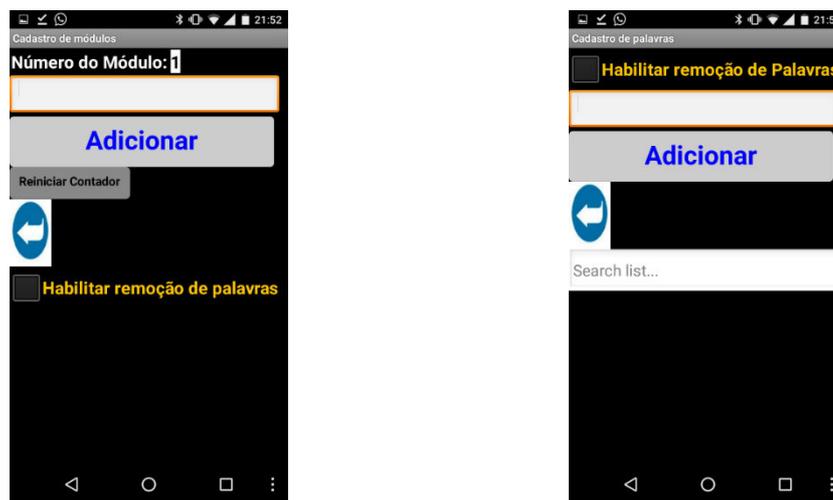


Figura 46. Tela de início do aplicativo OuviTronic.

Na fig.46, parte superior nota-se um campo de texto de mensagem. Logo abaixo, nota-se os botões denotados: “<”, “1”, “>”, “Limpar” e “Parar”. O segundo botão com o conteúdo “1” corresponde a um contador que incrementa a cada mensagem recebida do *Bluetooth*. No momento do recebimento da mensagem, o celular vibra intensamente, chamando a atenção do DA por meio de vibração. A cada mensagem recebida, o contador é incrementado, e a mensagem é salva no banco de dados, e conforme o DA clicar nos botões antes e depois, denotados por “<” e “>”, respectivamente, o DA poderá visualizar a mensagem correspondente. Já o botão “Parar” desativa a notificação de recebimento de mensagens, porém a mensagem é salva no banco de dados, e posteriormente poderá ser visualizada ao clicar nos botões de avanço ou de retorno. O usuário também poderá limpar o campo de texto da tela inicial ao clicar no botão “Limpar”.

Ao clicar no botão “Cadastro de módulos”, aparece no smartphone uma nova tela, Fig.47 (a), em que o DA pode adicionar textos ao clicar na caixa de texto e depois em adicionar. Cada texto adicionado incrementa o contador que se refere ao número do módulo correspondente. Quando, por exemplo, o módulo 1 for acionado o dispositivo envia o número 1 para o aplicativo e o mesmo compara com a primeira frase cadastrada e a exibe na tela inicial no campo de texto.

Para que o aplicativo se torne eficaz no auxílio aos DAs, é importante que o usuário cadastre as palavras que ele quer ser notificado com maior intensidade, seja para reconhecer quando receber a mensagem com seu nome ou para um pedido de socorro, por exemplo, conforme Fig.47 (b). Quando habilitada a função “Coincidir a frase inteira”, mostrada na tela da Fig.46, ao receber a mensagem o aplicativo associa o texto recebido com as palavras cadastradas, e se no caso de coincidência, vibra mais intensamente o celular. Quando essa função é desabilitada, o aplicativo apenas verifica as correspondências nas palavras salvas.



(a)

(b)

Figura 47. Telas do aplicativo: Cadastro de módulos, à direita (a) e Cadastro de palavras, à esquerda (b).

Dentre as opções de interação com o *App*, o DA poderá utilizar a ferramenta de conversão de texto em fala, ao clicar no botão “Falar”. Na tela da Fig.48, o DA poderá digitar um texto ao clicar na caixa de texto, ou também, poderá fazer uso de frases salvas, com o intuito de facilitar sua comunicação no dia-a-dia. Ao clicar numa frase salva, o aplicativo reproduz o texto em fala, imediatamente. Para somente reproduzir o texto digitado em fala, basta clicar no botão de “rosto emitindo ondas sonoras”.

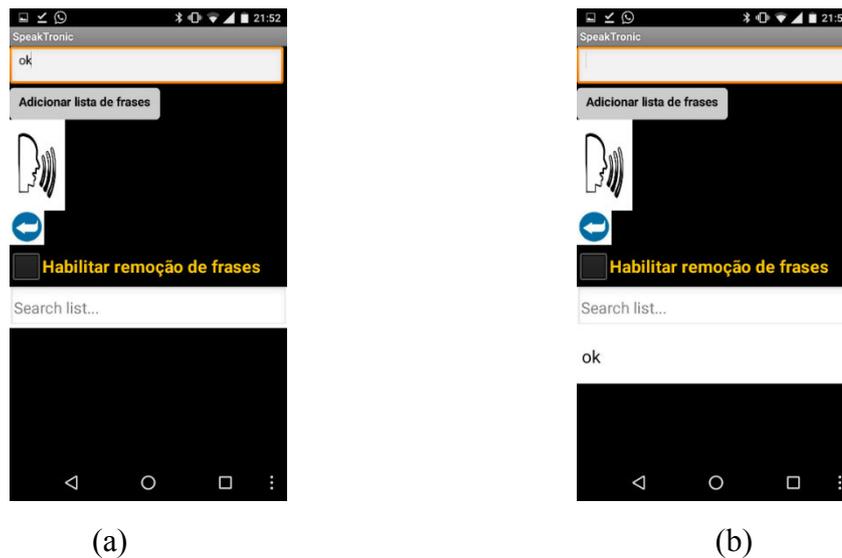


Figura 48. Tela de conversão de texto em fala.

Observa-se, na Fig.48 (a), que a palavra “ok” está selecionada, e foi reproduzida em fala, na Fig.48 (b), pela conversão de texto em fala. É importante ressaltar que, mesmo que o aplicativo esteja na tela de conversão de texto em fala, se o celular receber uma notificação do módulo *Bluetooth* o aplicativo volta à tela inicial e realiza a notificação via vibração e LED intermitente.

Por fim, observa-se na Fig.48 que há um campo “Habilita remoção de frases”, que quando estiver selecionado, ao clicar em alguma das frases salvas na lista, automaticamente a exclui tanto da tela, quanto do banco de dados do dispositivo. Dessa forma, o DA poderá armazenar e apagar as frases que desejar.

### 6.3. Experimentos sobre o Processamento de Sinais para Misturas Instantâneas e Reconhecimento de Fala

Neste tópico são apresentados quatro experimentos comparativos de reconhecimento de áudio para texto, nos quais por meio do algoritmo *FastICA*, se extraíram as componentes independentes da mistura instantânea inserida.

Para validar os experimentos foram utilizados três textos de referências, os quais foram convertidos em gravações computacionais com vozes de diferentes timbres. Os áudios então são misturados e inseridos no algoritmo. Após a extração dos componentes, os áudios são convertidos novamente para texto por meio de um aplicativo de reconhecimento de voz. O texto convertido é comparado com o texto original, obtendo a quantidade de erros e acertos referente

ao texto original. Essa comparação foi feita manualmente, e por conseguinte não foi considerada a existência de falsos acertos ou erros na apresentação dos resultados.

Foram realizados quatro experimentos. O primeiro experimento (Experimento I), consistiu na verificação da funcionalidade do algoritmo *FastICA*, aplicando os métodos de BSS e ASR. O segundo experimento (Experimento II) foi desenvolvido com o objetivo de verificar o desempenho do método, com um número significativo de palavras nos dados utilizados. O terceiro experimento (Experimento III) consistiu em duplicar o Experimento II, mas com um aumento gradativo de 25 misturas para cada teste, iniciando com 3 e finalizando com 103 misturas, considerado como um número grande de misturas. Devido à quantidade de misturas, a taxa utilizada foi de 300.000 amostras/segundo. Por fim, o quarto experimento (Experimento IV) igual ao Experimento II e III, com o número de amostras por segundo de 6.500 e o número de misturas foi elevado até a perda dos áudios extraídos por não reconhecimento.

O objetivo dos experimentos a seguir foi de testar a capacidade de reconhecimento de voz para texto, resultantes dos componentes independentes, a fim de averiguar a capacidade de filtragem e reconhecimento do algoritmo *FastICA* para misturas instantâneas.

### 6.3.1. Experimento I

Foi realizado o experimento no *Matlab* com a utilização do algoritmo *FastICA* [FastICA 2016], com última atualização datada em 2005. Inicialmente foram gravadas três frases com vozes distintas, denotadas F1, F2 e F3, com o conteúdo conforme Tabela 4.

Tabela 4. Frases com vozes distintas usadas no Experimento I.

F1	<i>“O verdadeiro sentido da existência humana não é simplesmente nascer, viver e morrer, mas sim, deixar um pouco de si em cada momento em que se vive.”</i>
F2	<i>“As pessoas costumam dizer que a motivação não dura sempre. Bem, nem o efeito do banho, por isso recomenda-se diariamente.”</i>
F3	<i>“Escolha uma ideia. Faça dessa ideia a sua vida. Pense nela, sonhe com ela, viva pensando nela. Deixe cérebro, músculos, nervos, todas as partes do seu corpo serem preenchidas com essa ideia. Esse é o caminho para o sucesso.”</i>

A partir dos sons obtidos das gravações foram extraídos os sinais digitais, a uma taxa de 300.000 amostras/segundo. Na Fig.49 são ilustrados, na linha superior, três sinais fonte

denotados "sinal de origem", numerados de 1 a 3, da esquerda para a direita; e na linha inferior são mostrados os sinais resultantes de misturas, numerados de 1 a 3, da esquerda para a direita.

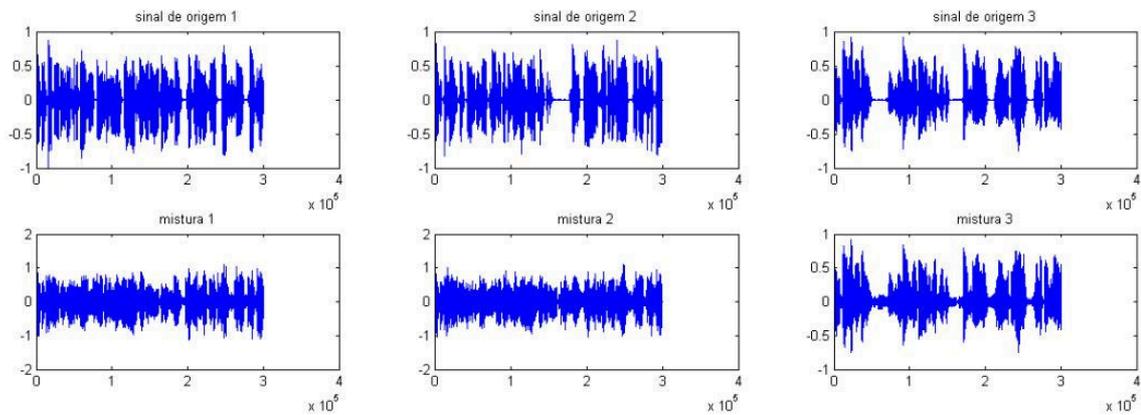


Figura 49. Os três sinais, numerados de 1 a 3. Linha superior: sinais originais e linha inferior: sinais das misturas.

As três misturas obtidas a partir da transformação linear entre a matriz da mistura de valores aleatórios, foram gravados em sinais de áudio (*wav*) utilizando a função do *Matlab* “*audiowrite*”. Os sons das misturas foram colocados, um áudio por vez, no aplicativo de plataforma livre *Dictanote* [Dictanote 2016] para o reconhecimento de fala, e como era previsto, não foi possível o reconhecimento.

Aplicando o Algoritmo-1, o primeiro componente independente foi calculado após 15 iterações, enquanto o segundo foi calculado após 4 iterações e o ultimo componente após 2 iterações. Na Fig.50 são apresentados os componentes extraídos.

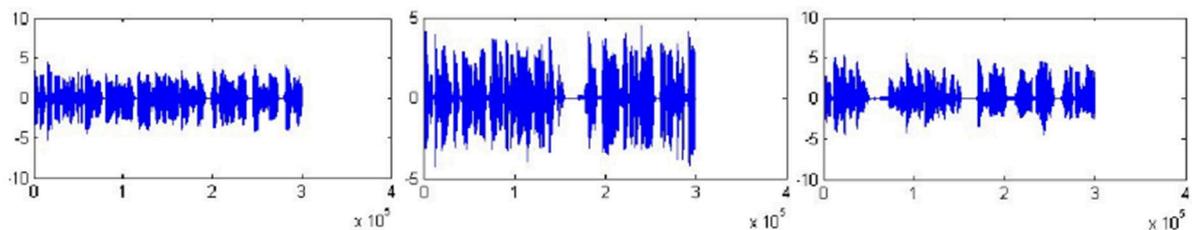


Figura 50. Sinais estimados pelo algoritmo *FastICA*, obtido pelo *Matlab*.

O áudio gerado é praticamente idêntico ao áudio original, com um perceptível aumento do volume do som. De posse dos sinais fonte estimados, foi utilizado o sistema ASR. O reconhecimento perfeito das falas, comprovaram a eficácia do algoritmo *FastICA* em estimar os sinais das fontes originais.

### 6.3.2. Experimento II

Neste experimento foram considerados três textos literários com 212, 215 e 206 palavras, respectivamente, o primeiro obtido de A.Huxley [Huxley 1979], e os seguintes de M.Assis [Assis 1899] [Assis 1891], conforme Tabela 5.

Na aplicação do Algoritmo-1, os sinais extraídos apresentaram 14, 17 e 21 erros, respectivamente, quanto ao texto original, na contagem de palavras erroneamente extraídas, representando 6,60%, 7,91% e 10,19%, do texto original, respectivamente; isso significa 93,40%, 92,09% e 89,81% de acerto.

A Tabela 5 e a Fig.51, de porcentagem de acerto, ilustram o desempenho obtido no Experimento II. Nota-se que a contagem de palavras erroneamente extraídas, foi feita manualmente pelo autor.

Tabela 5. Experimento II.

	Qtde total de palavras	Qtde de erros no sinal extraído pelo Algoritmo-1	Porcentagem de erros (%)	Porcentagem de acertos (%)
Texto1 (A.Huxley (1979))	212	14	6,60	93,40
Texto2 (Machado de Assis (1899))	215	17	7,91	92,09
Texto3 (Machado de Assis (1891))	206	21	10,19	89,81

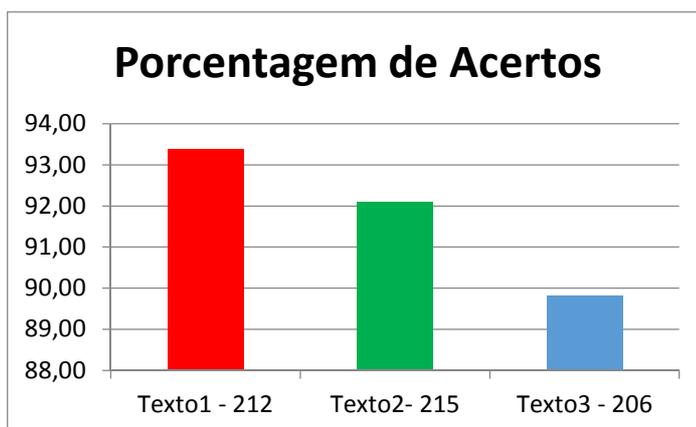


Figura 51 Gráfico da porcentagem de acertos do Experimento II

### 6.3.3. Experimento III

Para o experimento três, foi utilizada uma taxa de amostragem de 300.000 amostras/segundo. Portanto, o tamanho dos áudios dos três textos de referência foi encurtado para 52 segundos diminuindo a quantidade de palavras para cada texto conforme Tabela 6.

Tabela 6. Quantidade de palavras para cada texto referente ao número de amostras, experimento III.

Textos	Qtde palavras referente ao número de amostras
A.Huxley (Texto1)	129
Machado de Assis-Dom Casmurro (Texto 2)	136
Machado de Assis-Quincas Borba (Texto 3)	125

Inicialmente foram utilizados os três áudios de referências obtendo três novas misturas para esse novo número de amostras por segundo. Após os processos de reconhecimento e conversão de áudio em texto, foram analisadas as taxas de acerto em comparação com o texto de origem. Foi obtido um percentual para os textos 1, 2 e 3 respectivamente 89,92%; 96,32% e 92%.

Para a segunda etapa do experimento, foram utilizadas 28 misturas compostas dos três textos de referência, mais áudios de músicas distintas, todos com o mesmo tempo de duração, correspondendo a uma mistura quadrada. Com isso obtemos as respectivas porcentagens 90,70%; 96,32% e 92,8%.

Na terceira etapa o número de misturas passou para 53, composto dos áudios de referência mais 50 outros áudios musicais. Com essa composição obtemos as seguintes porcentagens de acertos: 89,92%; 94,12% e 89,60%. Já é possível verificar que não houve mudança significativa em comparação com a etapa anterior.

Para a quarta etapa o número de misturas foi aumentado de 25 passando para 78 misturas com a mesma composição que as etapas anteriores. Foi obtida uma taxa de acerto de 86,82%; 97,06% e 90,40%.

A última etapa do experimento foi composta de 103 misturas, no qual foram gerados os seguintes percentuais de acerto dos áudios de referência reconhecidos com o texto original sendo: 89,15%; 92,65% e 93,60%.

Na Tabela 7 é apresentada uma comparação entre os percentuais das taxas de palavras reconhecidas pelo *software* de reconhecimento de voz, em relação ao texto original. Com isso

foi possível extrair a porcentagem média para cada texto, 89,92%, 96,32% e 92%, respectivamente. A partir da Tabela 7, gerou-se o gráfico de linhas. Com base nos valores, obtidos, pode-se observar que não houve mudança significativa conforme o aumento do número de misturas, em relação ao reconhecimento de voz.

Tabela 7. Taxas de acerto dos sinais extraídos e reconhecidos em relação ao aumento na quantidade de misturas, experimento III.

Qtidade. de Misturas	Taxa de acerto dos sinais extraídos e reconhecidos (%)		
	TEXTO1	TEXTO2	TEXTO3
3	89,92	96,32	92,00
28	90,70	96,32	92,80
53	89,92	94,12	89,60
78	86,82	97,06	90,40
103	89,15	92,65	93,60
<b>Médias</b>	<b>89,30</b>	<b>95,29</b>	<b>91,68</b>

Na Fig.52 é ilustrado o percentual de acertos variando o número de misturas, do experimento III e na Fig.52, o percentual médio de acertos para cada texto.

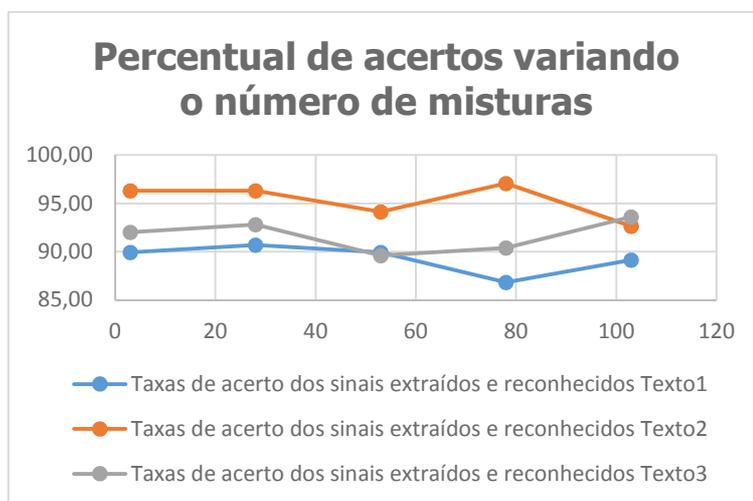


Figura 52. Gráfico percentual de acertos variando o número de misturas, Experimento III.

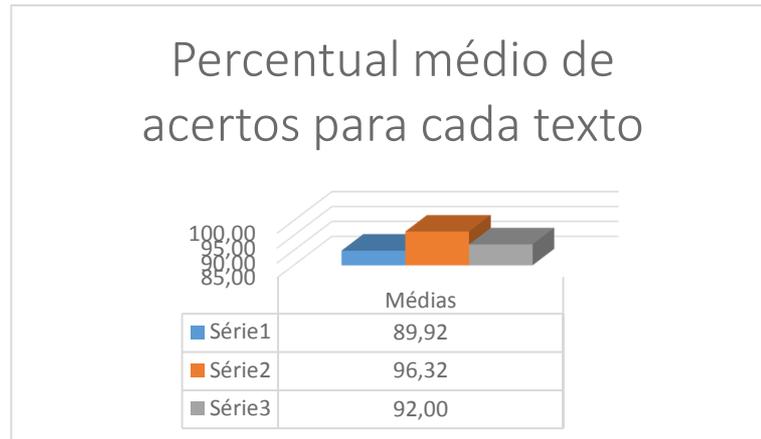


Figura 53. Gráfico percentual médio de acertos para cada texto, Experimento III.

### 6.3.4. Experimento IV

Com o intuito de chegar ao limite, da possibilidade de extração dos três áudios literários, o número de misturas foi sendo gradativamente aumentado até chegar a 303 misturas, no qual não foi possível encontrar os componentes de áudio referentes aos textos considerados.

Para essa simulação foram utilizadas 6.500 amostras/segundo, devido à grande exigência de processamento e memória do computador na realização do experimento. Portanto, todos os áudios extraídos possuem 14 segundos resultando na quantidade de palavras para cada texto entre 32 a 40 palavras conforme a Tabela 8.

Tabela 8. Quantidade de palavras para cada texto referente ao número de amostras, experimento IV.

TEXTOS	Qtde palavras referentes ao número de amostras
A.Huxley (Texto1)	35
Machado de Assis-Dom Casmurro (Texto 2)	40
Machado de Assis-Quincas Borba (Texto 3)	32

Os resultados obtidos com o reconhecimento dos textos da Tabela 8, podem ser observados na Tabela 9, no qual se destaca a pouquíssima variação no reconhecimento entre os textos em relação ao número de amostras nas misturas. A queda na taxa de acertos mais expressiva antes do declínio total ocorreu com a quantidade de 228 misturas, principalmente nos textos 1 com 62,86% e texto 3 com 71,88%. Ao aumentar o número de misturas observa-se o não reconhecimento de nenhum dos áudios dos textos, nem foi possível encontrar entre os áudios extraídos o som correspondente, já que os sinais extraídos são fornecidos aleatoriamente.

Tabela 9. Taxas de acerto dos sinais extraídos e reconhecidos em relação ao aumento na quantidade de misturas, Experimento IV.

Qtidade. de Misturas	Taxas de acerto dos sinais extraídos e reconhecidos (%)		
	Texto1	Texto2	Texto3
3	82,86	97,50	87,50
28	82,86	97,50	87,50
53	77,14	97,50	81,25
78	77,14	92,50	87,50
103	82,86	82,50	87,50
128	80,00	97,50	81,25
153	74,29	95,00	78,13
178	80,00	95,00	81,25
203	71,43	95,00	75,00
228	62,86	87,50	71,88
253	0,00	0,00	0,00
278	0,00	0,00	0,00
303	0,00	0,00	0,00

Com base na Tabela 9 foram extraídos dois gráficos, sendo o primeiro, o da Fig.54, que ilustra o percentual individual com o aumento gradativo do número de misturas, onde observa-se quase uma constante, até a quantidade de 228 misturas; após isso observa-se uma queda abrupta no gráfico.

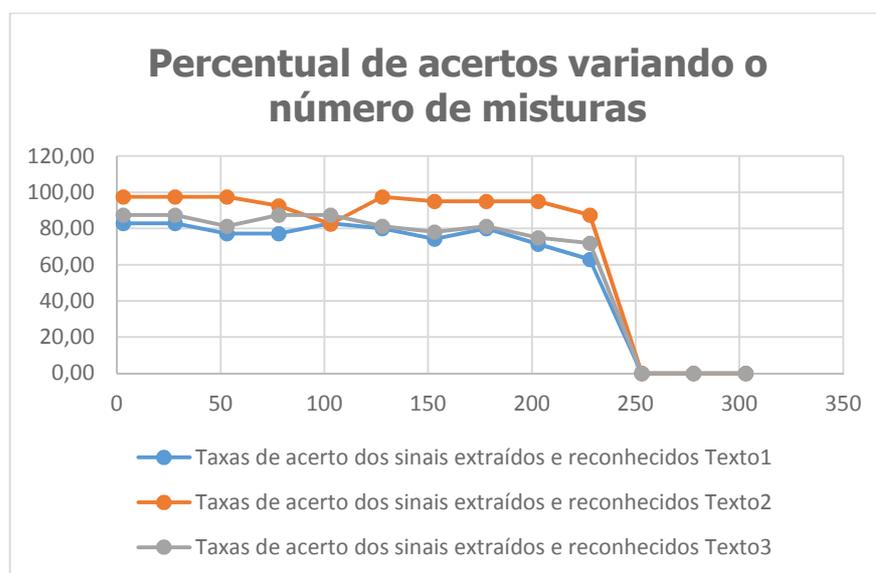


Figura 54. Gráfico percentual de acertos variando o número de misturas, experimento IV.

No segundo gráfico, o da Fig.55, ilustra-se o percentual médio entre todas as taxas de acerto para cada texto, em que pode ser observada uma pequena variação no percentual, o que

novamente acentua os resultados favoráveis quanto à qualidade no reconhecimento de voz com a metodologia do ICA por meio do algoritmo *FastICA*.

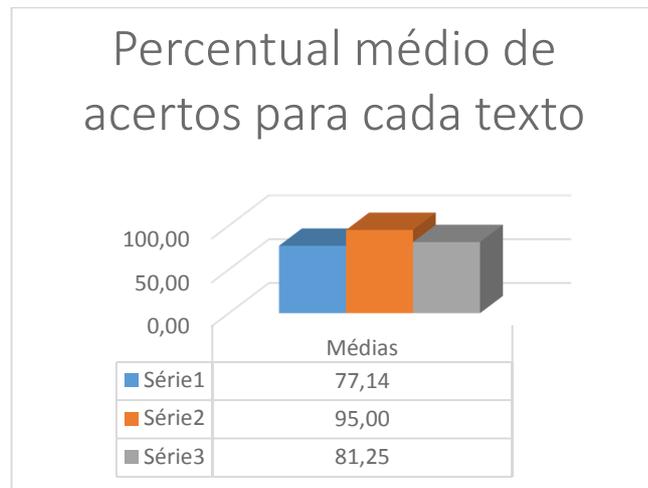


Figura 55. Gráfico percentual médio de acertos para cada texto, experimento IV.

Comparando-se os resultados dos experimentos III e IV, nota-se que até a quantidade de 103 misturas, as taxas de acertos tiveram alterações para valores menores no experimento IV, porém com redução menor que 10%. A diferença deve-se ao fato de que no experimento III, a taxa de amostragem foi de 300.000 amostras/segundo e tamanho do texto com duração de 52 segundos; e no experimento IV, a taxa de amostragem foi de 6.500 amostras/segundo, com duração de texto de 14 segundos.

#### 6.4. Experimentos com Captação de Misturas Convolutivas e Validação do BSS

Ao realizar experimentos práticos, em ambiente real, com captação de sinais usando dois microfones, os sinais captados foram diferentes de simples misturas instantâneas lineares de fontes originais. Esses sinais são conhecidos como misturas convolutivas. Utilizando misturas convolutivas, com o algoritmo *FastICA*, os resultados demonstraram incompatibilidade com as fontes originais, mesmo após os algoritmos convergirem para uma solução, confirmando-se que o algoritmo *FastICA* é adequado somente para misturas instantâneas e, portanto, para misturas convolutivas é necessário aplicar outras técnicas para a separação.

Com o objetivo de obter uma validação sobre as misturas convolutivas contextualizando o sistema SADA delineado nesta dissertação, suscetível a ruídos e reverberações do ambiente de aplicação, nesta seção são descritos experimentos realizados com duas gravações

simultâneas usando dois microfones, em um ambiente real, porém com o algoritmo proposto por PHAN para a separação de misturas convolutivas (Algoritmo-2, seção 4.4.1).

São apresentados três experimentos de misturas convolutivas, em que os áudios das misturas obtidas por meio de gravação de dois microfones em um ambiente real, são inseridos no algoritmo para se extraírem os componentes independentes de origem. Todos os áudios utilizados nos experimentos, detalhados a seguir, possuem 9 segundos de gravação, a uma frequência de amostragem de 8kHz. A distância entre os dois microfones foi de 45 cm, e as distâncias entre a fonte de origem do som (emissor) e o sensor receptor (microfone) foram conforme descritos na Tabela 10, onde a os dois emissores se posicionam de tal forma que as linhas imaginárias que ligam o Emissor-1 ao Sensor-1 e o Emissor-2 ao Sensor-2, são paralelas, formando ângulos retos com a linha imaginária que liga os dois sensores, conforme ilustração da Fig.56.

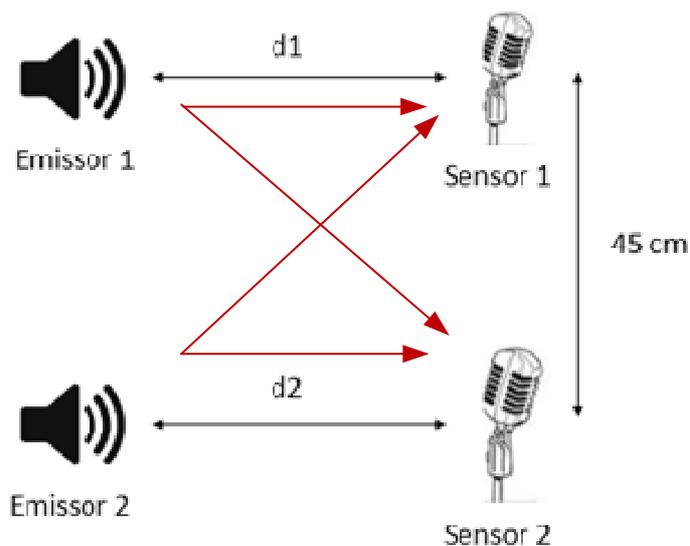


Figura 56. Diagrama de cenários entre sensores e as fontes de origem de som, para misturas convolutivas.

Tabela 10. Distância entre sensores e as fontes de origem de som, para misturas convolutivas.

Experi- mentos	Distância-1 (d1) entre o Emissor-1 e Sensor-1 (microfone) (cm)	Distância-2 (d2) entre o Emissor-2 e Sensor-2 (microfone) (cm)	Distância entre os dois sensores (microfones) (cm)
I	30	30	45
II	30	30	45
III	30	150	45

### 6.4.1 Experimento I

Para o Experimento I foram utilizados dois microfones (sensores) para gravar dois áudios ao mesmo tempo, sendo um áudio de música de fundo e um áudio de voz de uma pessoa do sexo masculino “contando de 1 a 10”, com distância entre as fontes para os sensores de 30 cm. Os áudios das misturas obtidas pelos dois microfones são mostrados nos gráficos da Fig. 57.

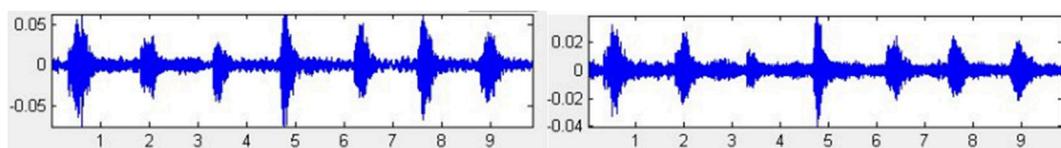


Figura 57. Misturas dos áudios obtidos no experimento I.

Com as misturas, é executado o algoritmo em que são extraídos os componentes independentes, como se pode observar nos gráficos da Fig.58.

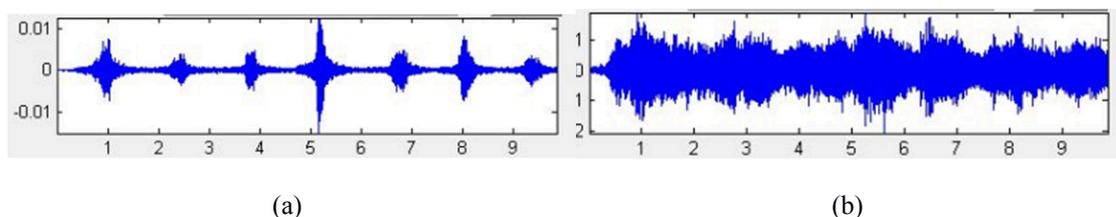
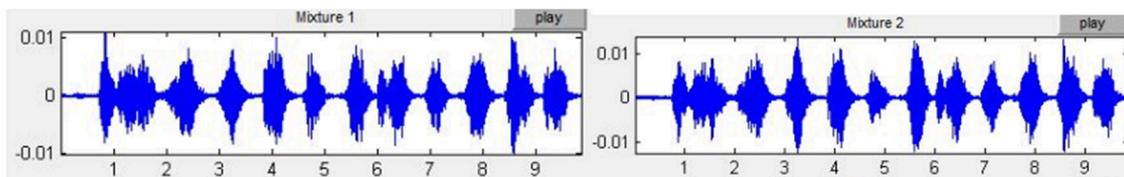


Figura 58. (a) áudio de voz e (b) áudio musical.

Observa-se que o gráfico da Fig.58 (a), referente ao áudio de voz, alterna entre amplitudes altas e baixas, em que as baixas representam as pausas entre uma palavra e outra, e as altas representam as palavras. Já o gráfico da Fig.58 (b) não possui uma diminuição de amplitude brusca, uma vez que a música reúne um conjunto de instrumentos musicais.

### 6.4.2. Experimento II

Neste Experimento II foram utilizados um áudio de voz feminina “ditando letras do alfabeto de A até L” e uma voz masculina “contando de 1 a 10”, com distância entre as fontes para os sensores, de 30 cm. A Fig.59 corresponde aos gráficos das misturas, nos quais nota-se que ambos são muito semelhantes.

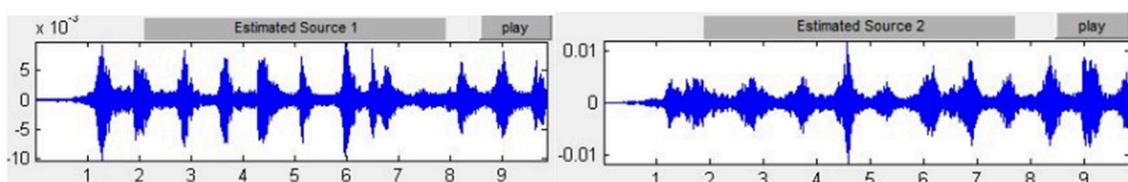


(A)

(B)

Figura 59. Gráficos para misturas convolutivas entre dois áudios de voz.

As fontes estimadas são apresentadas na Fig.60, sendo o gráfico da Fig.60 (a), correspondente ao áudio de voz masculina “contando de 1 a 10”, em que é possível visualizar a contagem pelo pico na amplitude para cada número. Já no gráfico da Fig.60 (b), correspondente à voz feminina “ditando as letras do alfabeto da língua portuguesa de A até L”. Como a voz feminina é mais aguda em relação à masculina, a frequência é maior e a amplitude é menor que o gráfico da Fig.59 (a).



(a)

(b)

Figura 60. (a) áudio de voz masculina “contando de 1 a 10” e (b) áudio de voz feminina “ditando de A até L do alfabeto”.

### 6.4.3. Experimento III

Para o Experimento III, foi utilizado um áudio de voz masculina “contando de 1 a 10” e um ruído de fundo de um liquidificador, sendo que o liquidificador ficou a uma distância de 150 cm do sensor e a pessoa de sexo masculino a uma distância de 30 cm do sensor. As misturas podem ser observadas na Fig.61.

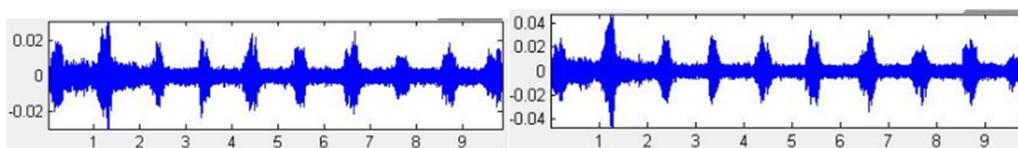


Figura 61. Misturas dos áudios utilizados no experimento III.

Os sinais estimados das misturas, podem ser observados na Fig.62 que evidencia no gráfico da Fig.62 (a), a voz masculina, devido às pausas naturais entre a fala de um número a outro. Já o gráfico da Fig.62 (b) é basicamente um sinal de ruído contínuo em que não há diferenças bruscas de amplitude e frequência.

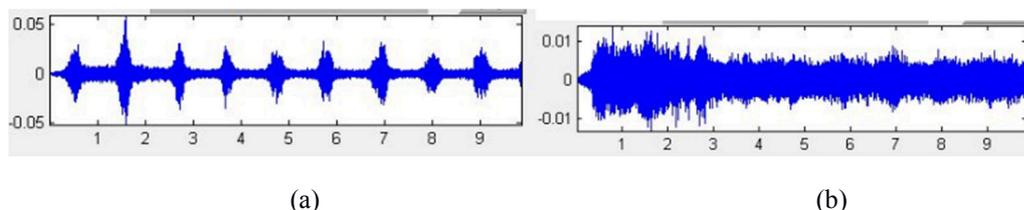


Figura 622. (a) lado esquerdo, áudio de voz, e (b) lado direito, áudio de ruído de liquidificador.

## 6.5. Considerações Finais

Neste capítulo foram descritas: as montagens de protótipos para os experimentos de validação do SADA Ouvitronic-1, denominados de pré-protótipos; e os experimentos realizados para a validação do pré-protótipo. Os testes realizados com algoritmo de separação cega das fontes, *FastICA*, de pequenas frases e consequente reconhecimento de fala com ASR, tiveram resultados satisfatórios. Os sinais de áudio extraídos das misturas foram submetidos ao reconhecimento automático de fala, obtendo o texto escrito confirmando a eficácia do algoritmo *FastICA*. Quanto ao desempenho obtido na separação das fontes e reconhecimento da fala, pode-se concluir que para o número de três falas distintas, foi satisfatório.

Já ao tentar extrair os componentes independentes de misturas convolutivas com o algoritmo *FastICA*, como previsto, não foi possível obter a estimativa dos sinais de origem mesmo convergindo para uma solução. Portanto, foi comprovado que o algoritmo *FastICA* é adequado somente para misturas instantâneas.

Como solução para misturas convolutivas, foi utilizado um algoritmo baseado em ICA que utiliza técnicas que consideram a não-estacionaridade das misturas, com filtros FIR e domínio da frequência para se extrair os componentes independentes. Com esse algoritmo, foram obtidos resultados satisfatórios nos experimentos realizados.

## CAPÍTULO 7

# CONCLUSÕES E PROPOSIÇÕES DE TRABALHOS FUTUROS

---

Neste capítulo são apresentadas as conclusões finais deste trabalho dividida nas seguintes seções: a seção 7.1 apresenta as conclusões; a seção 7.2 apresenta as proposições para trabalhos futuros.

### 7.1 Conclusões

Todos nós, humanos, merecemos ter o acesso aos meios que contribuem para a melhoria da qualidade de vida como cidadão em meio a sociedade civilizada. Deste modo, é pertinente insistir em pesquisas que promovam a tecnologia necessária para a assistência aos deficientes, em todos os aspectos, decifrando as suas principais dificuldades. Neste sentido, esta dissertação foi desenvolvida com a ideia de desenvolver ferramentas que possam servir como contribuição para a melhoria da comunicação dos DAs, que representa uma fração dos problemas enfrentados no dia a dia por essas pessoas.

Com esse intuito, no Capítulo 2 foram descritas as informações que nortearam o desenvolvimento deste trabalho, como alguns sistemas de auxílios aos DAs desenvolvidos dentro do campo da Tecnologias Assistiva, dentre os quais os diversos Aparelhos de Amplificação Sonora Individual (AASI), cada vez mais sofisticados na supressão de ruídos.

Capítulo 3 foi feita uma revisão bibliográfica dos trabalhos relacionados ao desenvolvimento de sistemas de auxílio a deficientes auditivos; e de pesquisas sobre processamento de sinais, de misturas convolutivas e análise de componentes independentes.

No Capítulo 4 foi apresentado o problema de Separação Cega de Fontes e uma das ferramentas mais utilizadas que é a Análise de Componentes Independentes, com o algoritmo *FastICA* utilizado nos experimentos em misturas instantâneas. Foram descritas as misturas convolutivas, que nada mais são do que as misturas em ambientes reais. Em seguida, foram detalhados os materiais e os *softwares* utilizados no desenvolvimento desta Dissertação.

O Capítulo 5 foi dedicado à descrição de todas as partes que envolveram o delineamento do protótipo SADA, aspectos sobre o *hardware* de comunicação; e *software* de comunicação e sensibilização do DA, com sinais pertinentes de assistência. Destaca-se a caracterização de um sistema, composto de captação, processamento, sensibilização e interação do DA.

Por conseguinte, no Capítulo 6 são apresentados os resultados obtidos com os experimentos em ICA e sobre o desenvolvimento do pré-protótipo SADA. A respeito dos experimentos em ICA, foram divididos em duas fases sendo: Fase 1- os experimentos comparativos de reconhecimento de áudio para texto, a partir dos áudios extraídos das misturas instantâneas simuladas; Fase 2- os experimentos com misturas convolutivas em simulação de momentos em um ambiente real.

A respeito dos resultados obtidos na Fase 1, pode-se destacar a eficácia do algoritmo *FastICA* na extração dos componentes independentes em que não foi possível mais o reconhecimento do texto a partir de 228 misturas a 6500 amostras por segundo, em que ocorreu o declínio total. Essa primeira fase, teve um caráter analítico sobre a capacidade máxima de extração dos componentes das fontes de origem.

Nota-se no entanto, que apesar desse número grande de sinais de áudio para misturas, foram analisadas as separações de apenas três sinais de origem, nos diversos experimentos.

A Fase 2, os resultados dos experimentos com o algoritmo que faz uso da análise no domínio da frequência considerando a não-estacionariedade de misturas convolutivas, foram satisfatórios em que se pode concluir a possível aplicabilidade do protótipo descrito nessa dissertação.

Por fim, os resultados obtidos com o pré-protótipo estão de acordo com o que foi proposto, obtendo resultados satisfatórios de funcionamento tanto da captação dos sinais pelos sensores quanto no processamento e sensibilização do DA.

## **7.2 Propostas para Trabalhos Futuros**

Para uma continuação deste trabalho, podemos sugerir para pesquisas futuras os seguintes itens:

1. Realizar maiores experimentos com o protótipo com instalações reais de mais sensores, a fim de melhorar os testes para o algoritmo de misturas convolutivas;
2. Possibilitar a portabilidade do protótipo para qualquer ambiente;
3. Melhorar a interface do aplicativo para *Android*;
4. Possibilitar a utilização do aplicativo para o sistema operacional *IOS*, da *Apple Inc.*

5. Delinear um novo protótipo do SADA Ouvitronic com o computador hospedeiro de processamento de BSS portátil, junto ao DA.

## REFERÊNCIAS BIBLIOGRÁFICAS

[Abreu 2003] Abreu, N. C. L. (2003) **Extração de sinais de voz em ambientes ruidosos por decomposição em funções bases estatisticamente independentes**. Dissertação de Mestrado, Universidade Federal Maranhão UFMA, São Luís-MA.

[Acessibilidade 2017] Acessibilidade Brasil. Disponível em: <[www.acessibilidadebrasil.org.br/versao\\_anterior/](http://www.acessibilidadebrasil.org.br/versao_anterior/)> Acesso em janeiro de 2017.

[Amarasinghe & Wimalaratne 2017] Amarasinghe, A. e Wimalaratne, P. (2017). **An Assistive Technology Framework for Communication with Hearing Impaired Persons**. GSTF Journal on Computing (JoC), vol.5, n.2, janeiro de 2017.

[Android 2017] Android, D. (2017) **Android Architecture**. Disponível em: <<https://developer.android.com/images/system-architecture.jpg>>. Acesso em março de 2017.

[Araki et al. 2003] ARAKI, S. et al. (2003) **The fundamental limitation of frequency domain blind source separation for convolutive mixtures of speech**. IEEE Transactions on Speech and Audio Processing.2003, vol. 11, n.2.

[Arduino 2016] Arduino (2016) **SoftwareSerial Library**. Disponível em <<https://www.arduino.cc/en/Reference/softwareSerial>>. Acesso em novembro de 2016.

[Arduino 2017] Arduino (2017). Disponível em: < <https://www.arduino.cc/en/main/software>>. Acesso em maio de 2017.

[Assis 1891] Assis, M. (1891) **Quincas Borba**. Rio de Janeiro-RJ, B.L.Garnier, Livreiro-Editor.

[Assis 1899] Assis, M. (1899) **Dom Casmurro**. Rio de Janeiro-RJ, B.L.Garnier, Livreiro-Editor.

[Atmel 2017] Atmel (2017) **Atmel 8-Bit Microcontroller With 4/8/16/32kbytes In-System Programmable Flash. Datasheet**. Disponível em:< [http://www.atmel.com/images/Atmel-8271-8-bit-AVR-Microcontroller-ATmega48A-48PA-88A-88PA-168A-168PA-328-328P\\_datasheet\\_Complete.pdf](http://www.atmel.com/images/Atmel-8271-8-bit-AVR-Microcontroller-ATmega48A-48PA-88A-88PA-168A-168PA-328-328P_datasheet_Complete.pdf)>. Acesso em março de 2017.

[Baptista 2007] Baptista, M.C.P. (2007) **Identificação e caracterização da modulação dos sinais digitais em RF**. Dissertação de Mestrado, Universidade de Aveiro, Portugal.

[Batalheiro et al. 2008] Batalheiro, P. B. et al. (2008). **Separação Cega de Fontes para misturas convolutivas usando um banco de filtros não-uniforme**. XXVI Simpósio Brasileiro de Telecomunicações, Rio de Janeiro, RJ.

[Benesty 2008] Benesty, J.; Sondhi, M. M.; Huang, Y. (2008) **Springer Handbook of Speech Processing**. Heidelberg-Alemanha, Springer-Verlag.

[Barreto 2015] Barreto, M. A. (2015). **Aplicação de Ressonância Estocástica em Medidas Eletrofisiológicas**. Dissertação de Mestrado em Ciência da Computação, Faculdade Campo Limpo Paulista.

[Bersch 2013] Bersch,R. (2013) **Introdução à tecnologia assistiva**. Porto Alegre, CEDI. Disponível em: <<http://www.haasfretes.com.br/arquivos/introducao-tecnologia-assistiva.pdf>>. Acesso em setembro de 2016

[Brasil 2004] Brasil (2004) **Decreto nº 5.296, de 02 de dezembro de 2004**. Estabelece normas gerais e critérios básicos para promoção da acessibilidade das pessoas com deficiência ou com mobilidade reduzida. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/\\_ato2004-2006/2004/decreto/d5296.htm](http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2004/decreto/d5296.htm)>. Acesso em outubro de 2016.

[Brasil 2002] Brasil (2002) **Lei nº 10.436, de 24 de abril de 2002**. Dispõe sobre língua Brasileira de Sinais Libras e dá outras providências. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/leis/2002/l10436.html](http://www.planalto.gov.br/ccivil_03/leis/2002/l10436.html)>. Acesso em outubro de 2016.

[Brasil 1991] Brasil (1991) **Lei nº 8.160, de 08 de janeiro de 1991**. Dispõe sobre a caracterização de símbolo que permita a identificação de pessoas portadoras de deficiência auditiva. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/leis/L8160.htm](http://www.planalto.gov.br/ccivil_03/leis/L8160.htm)> Acesso em novembro de 2016.

[Bresolin 2003] Bresolin, A.A.(2003) **Estudo do reconhecimento de voz para o acionamento de equipamentos elétricos via comandos em português**. Dissertação (Mestrado em Automação Industrial), Universidade Do Estado De Santa Catarina, Centro De Ciências Tecnológicas – CCT, Joinville-SC.

[Blum 2013] Blum, J. (2013) **Exploring Arduino: tools and techniques for engineering wizardry**. Nova Jersey-USA, John Wiley & Sons.

[Cavalcante 2004] Cavalcante, C.C. (2004) **Sobre Separação Cega de Fontes: Proposições e Análise de Estratégias para Processamento Multi-Usuário**. Tese, Universidade Estadual de Campinas, Campinas-SP.

[Comon 2010] Comon, P. & Jutten, C. (2010) **Handbook of Blind Source Separation: Independent Component Analysis and Applications**. Burlington, MA, USA, Academic Press.

[Danesi 2007] Danesi, M. C. (2007) **O admirável mundo dos surdos: novos olhares do fonoaudiólogo sobre a surdez, 2.ed.** . Porto Alegre-RS, EDIPUCRS.

[Detran-PE 2016] Detran-PE (2016) **Símbolo Internacional de Surdez** Disponível em: <[http://www.detran.pe.gov.br/index.php?option=com\\_content&id=390&Itemid=296](http://www.detran.pe.gov.br/index.php?option=com_content&id=390&Itemid=296)>. Acesso em outubro de 2016.

[Dictanote 2016] Dictanote (2016) Disponível em <<https://dictanote.com/>>. Acesso em julho de 2016.

[FastICA 2016] FastICA (2016) Disponível em <<http://research.ics.aalto.fi/ica/fastica/code/dlcode.shtml>> . Acesso em junho de 2016.

[França 2011] França, A. C. C. V. (2011) **Interação de pessoas surdas mediada por sistemas de produtos e serviços de comunicação**. Dissertação de Mestrado em Tecnologia, Universidade Tecnológica Federal do Paraná, Curitiba-PR.

[Ferreira 2002] Ferreira, A. J. (2002) **Aplicação da Análise em Componentes Independentes na Compressão de Imagem**. Dissertação de Mestrado, Centro de Cálculo do Instituto Superior de Engenharia de Lisboa, Lisboa, Portugal.

[Ferreira 2014] Ferreira, M. A. M. (2014) **Design Inclusivo e Participativo na Web: Incluindo Pessoas Surdas**. Dissertação de Mestrado em Ciência da Computação, Faculdade de Campo Limpo Paulista, Campo Limpo Paulista-SP.

[Fonseca 2002] Fonseca, H.; Santos, V.; Ferreira, A. (2002) **Audição Constituição do sistema auditivo humano**. Disponível em: <[http://telecom.inescn.pt/research/audio/cienciaviva/constituicao\\_audicao.html](http://telecom.inescn.pt/research/audio/cienciaviva/constituicao_audicao.html)> Acesso em agosto de 2016.

[Guyton 2006] Guyton, A. C. & Hall, J. E. (2006) **Tratado de fisiologia médica**. Rio de Janeiro, Elsevier.

[Hand Talk 2017] Hand Talk (2017) **Application Hand Talk**. Disponível em <<https://www.handtalk.me/app>>. Acesso em abril de 2017.

[Huxley 1979] Huxley, A. (1979) **Admirável mundo novo**. Trad. V. Oliveira e L. Vallandro, Porto Alegre-RS, Editora Globo.

[Harvey 2010] Harvey, A. P. (2010) **Automatic speech recognition for assistive technology devices**. Proceedings of the 8th International Conference on Disability Virtual Reality And Associated Technologies, 31 de agosto a 2 de setembro de 2010, Valparaíso, Chile.

[Hyvärinen et al. 2001] Hyvärinen, A. et al. (2001) **Independent Component Analysis**. Nova Jersey, USA, John Wiley & Sons, Inc,

[Hyvärinen & Oja 1997] Hyvärinen, A. e Oja, E. (1997) **A fast fixed-point algorithm for independent component analysis**, Neural Computation, v.9, p.1483-1482.

[IBGE 2010] IBGE (2010) **Censo Demográfico 2010 Características gerais da população, religião e pessoas com deficiência**. ISSN 0104-3145 Censo demográfico, Rio de Janeiro, p.1-215, 2010. Disponível em [http://biblioteca.ibge.gov.br/visualizacao/periodicos/94/cd\\_2010\\_religiao\\_deficiencia.pdf](http://biblioteca.ibge.gov.br/visualizacao/periodicos/94/cd_2010_religiao_deficiencia.pdf). Acesso em julho de 2016.

[Inventor 2016] App Inventor (2016) Inventor MIT *App*. Disponível em: <<http://appinventor.mit.edu/explore/about-us.html>>. Acesso em outubro de 2016.

[Kobayashi 2004] Kobayashi, C. Y. (2004) **A Tecnologia Bluetooth e aplicações**. USP. São Paulo-SP. Disponível em: <[http://grenoble.ime.usp.br/movel/monografia\\_bluetooth.pdf](http://grenoble.ime.usp.br/movel/monografia_bluetooth.pdf)>. Acesso em dezembro de 2016.

[Librazuka 2017] Application Librazuka, (2017). Disponível em <[www.economista.com.br/d/librazuka-o-aplicativo-de-libras/](http://www.economista.com.br/d/librazuka-o-aplicativo-de-libras/)>. Acesso em abril de 2017.

[Luft 1991] Luft, J. A. (1991) **Reconhecimento Automático de Voz para Palavras Isoladas e Independente de Locutor**. Tese de Doutorado. Universidade Federal do Rio Grande do Sul-Escola de Engenharia, Porto Alegre-RS.

[Lojamatel 2017] Lojamatel (2017). Disponível em <<http://www.lojamatel.com.br/produtos/campainhas/153>>. Acesso em abril de 2017.

[McCauley 2013] McCauley, M. (2013). **VirtualWire**. Disponível em:<<http://www.airspayce.com/mikem/arduino/VirtualWire.pdf>> Acesso em julho de 2017.

[Maiorana-Basas & Pagliaro 2014] Maiorana-Basas e Pagliaro, C. M. (2014) **Technology use among adults who are deaf and hard of hearing: A national survey**, Journal of Deaf Studies and Deaf Education, v. 19, n. 3, p. 400-410.

[Marin 2006] Marin, C. R. & Góes, M. C. R. (2006) **A experiência de pessoas surdas em esferas de atividade do cotidiano**, Cad Cedes, v. 26, n. 69, p. 231-49.

[Moreto 2008] Moreto, F.A.L. (2008) **Análise de Componentes Independentes Aplicada à Separação de Sinais de Audio**. Dissertação de Mestrado, Escola Politécnica da Universidade de São Paulo, São Paulo-SP.

[Morimoto 2009] Morimoto, C. E. (2009) **Smartphones: guia prático**. Porto Alegre-RS, Sul Editores.

[Mutemwa & Tucker 2010] Mutemwa, M. e Tucker, W. D. (2010). **A mobile Deaf-to-hearing communication aid for medical diagnosis**. Department of Computer Science, University of the Western Cape, Bellville, Africa do Sul. Disponível em:<<https://repository.uwc.ac.za/handle/10566/738>>. Acesso em julho de 2017.

[Naik & Wang 2014] Naik, G. R. e Wang, W. (2014). **Blind source separation**. Berlin, Heidelberg: Springer.

[Nesta et al. 2011] Nesta, F. et al. (2011). **Convolutive BSS of short mixtures by ICA recursively regularized across frequencies**. IEEE transactions on audio, speech, and language processing. Disponível em:<<https://pdfs.semanticscholar.org/52e5/3b96436efd889f0b0fdf3a6350b284b59088.pdf>> Acesso em julho de 2017.

[Nicolau 2010] Nicolau, V. B. (2010) **Cancelamento de realimentação acústica em aparelhos auditivos através do método de erro de predição com comprimento deficiente**. Dissertação de Mestrado, Universidade Federal de Santa Catarina, Florianópolis-SC.

[Notisurdo 2017] Notisurdo (2017). **História da tecnologia para surdo**. Disponível em:<<http://notisurdo.com.br/tecnohist.html>> Acesso em setembro de 2017.

[Nunes 1996] Nunes, H. F. (1996) **Reconhecimento de fala baseado em HMM**. Dissertação de Mestrado, Universidade Estadual De Campinas, Faculdade de Engenharia Elétrica e de Computação, Campinas-SP.

[Oliveira 2015] Oliveira, A. F. (2015) **Interfaces Visuais e Perceptivas para apoio a Surdos**, Anais do XI Workshop de Computação da FACCAMP, Campo Limpo Paulista-SP.

[Ossada & Rodrigues 2016] Ossada, S. A. R. e Rodrigues, S. C. M. A. (2016) **Tecnologia da informação em colaboração na comunicação dos deficientes auditivos**, FaSCi-Tech, Periódico Eletrônico da Fatec São Caetano do Sul, v. 1, n. 10.

[Pham et al. 2003] Pham, D.T. et al. (2003). **Blind Separation of convolutive audio mixtures using nonstationarity**. Nara, Japan. Disponível em: < <http://bsp.teithe.gr/members/downloads/bssaudio/references/sepa-audioR.pdf> >. Acesso em junho de 2017

[Plachevski 2014] Plachevski, M. S. (2014). **Sistema de Tecnologia Assistiva para Captar a Atenção de Deficientes Auditivos e Surdos**. Dissertação (Mestrado em Tecnologia). Universidade Estadual de Campinas, Faculdade de Tecnologia em Tecnologia, Limeira-SP.

[Pedersen 2006] Pedersen, M. S. (2006) **Source Separation for Hearing Aid Applications**. Tese de Doutorado, Technical University of Denmark. Kongens Lyngby-Dinamarca.

[Pereira & Feres 2005] Pereira, M. B. e Feres, M. C. L. C. (2005) **Próteses auditivas**. Medicina (Ribeirão Preto. Online), v. 38, n. 3/4, p. 257-261.

[Petry et al. 2000] Petry, A.; Zanuz, A.; Barone, D. A. C. (2000) **Reconhecimento Automático de Pessoas Pela Voz Através de Técnicas de Processamento Digital de Sinais**. Instituto de Informática-Universidade Federal do Rio Grande do Sul, Porto Alegre-RS.

[Petry 2002] Petry, A. (2002) **Reconhecimento Automático de Locutor Utilizando Medidas de Invariantes Dinâmicas Não-lineares**. Tese de Doutorado, Universidade Federal Do Rio Grande Do Sul, Porto Alegre- RS.

[Priess 2003] Priess, W. (2003) **Mecanismos de escalonamento com qualidade de serviço em redes com tecnologia Bluetooth**. Dissertação de Mestrado, Universidade Federal do Rio de Janeiro, Rio de Janeiro-RJ.

[ProDeaf 2016] ProDeaf (2016). Disponível em: < <http://prodeaf.net/pt-br/Contato> >. Acesso em abril de 2016.

[Rabiner 1989] Rabiner, L.R. (1989). **A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition**. Proceedings of IEEE, vol. 77, n.2, p. 257-286.

[Radabaugh 1993] Radabaugh, M. P. (1993) **NIDRR's Long Range Plan-Technology for access and function research section two: NIDRR Research Agenda Chapter 5: Technology for access and function**. United States, Disponível em: <<http://www.ncd.gov/newsroom/publications/1993/assistive.htm#5>> Acesso em setembro de 2016.

[Rahbar, K. & Reilly 2005] Rahbar, K. e Reilly, J. P. (2005). **A frequency domain method for blind source separation of convolutive audio mixtures**. IEEE Transactions on Speech and Audio Processing, vol. 13, n. 5, setembro de 2005.

[Ribeiro 2006] Ribeiro, D. A. G. (2006). **Rádio Freqüência para Identificação e Apoio a Deficientes**. Disponível em: < <http://www.teleco.com.br/DVD/PDF/tutorialradiofiad.pdf>>. Acesso em 04 de março de 2017.

[Rybená 2017] **Reprodução de texto para voz falada**. Disponível em: <[portal.rybena.com.br/site-rybena/conheca-o-rybena](http://portal.rybena.com.br/site-rybena/conheca-o-rybena)> . Acesso em janeiro 2017.

[SEDH 2009] SEDH (2009) **Tecnologia Assistiva**. Secretaria Especial de Direitos Humanos, Subsecretaria Nacional de Promoção dos Direitos das Pessoas com Deficiência, Comitê de Ajudas Técnicas, Brasília-DF.

[Sharma & Wasson 2012] Sharma, R.F e Wasson, S. G. (2012) **Speech Recognition and Synthesis Tool: Assistive Technology for Physically Disabled Persons**. International Journal of Computer Science and Telecommunications, vol. 3, n.4, abril de 2012.

[Silverthorn 2009] Silverthorn, D. U. (2009) **Fisiologia humana: uma abordagem integrada**. 5. ed. Porto Alegre-RS, Artmed editora.

[Suyama 2007] Suyama, R. (2007) **Proposta de métodos de separação cega de fontes para misturas convolutivas e não-lineares**. Tese de Doutorado, Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas. Campinas-SP.

[Timoszczuk 2004] Timoszczuk, A. P. (2004) **Reconhecimento automático do locutor com redes neurais pulsadas**. Tese de Doutorado, Universidade de São Paulo, São Paulo-SP.

[Torres 2002] Torres, E. F.; Mazzoni, A. A.; Alves, J. B. M. (2002) **A acessibilidade à informação no espaço digital**. Ciência da Informação, Brasília, v. 31, n. 3, p. 83-91.

[Uno 2016] Arduino UNO (2016) Disponível em: <<https://www.arduino.cc/en/Main/ArduinoBoardUno>>. Acesso em setembro de 2016.

[Wang, L. et al. 2011] Wang, L. et al. (2011). **A region-growing permutation alignment approach in frequency-domain blind source separation of speech mixtures**. IEEE transactions on audio, speech, and language processing, vol. 19, n.3, março de 2011.

[Yu & Li 2015] Yu, D. e Li, D (2015) Automatic Speech Recognition: A Deep Learning Approach. Londres, Inglaterra, Springer-Verlag.

[Zemlin 2000] Zemlin, W.R.(2000) **Princípios de anatomia e fisiologia em fonoaudiologia**. Porto Alegre-RS, Editora Artes Médicas Sul.