

# Sistema de Conversão Português/Libras

Leandro Sabino da Silva <sup>1</sup> e Raissa Bezerra Rocha <sup>1,2</sup>

<sup>1</sup>Departamento de Engenharia Elétrica, Universidade Federal de Sergipe - leandrosabino08@gmail.com

<sup>1,2</sup>Instituto de Estudos Avançados em Comunicações - raissa@iecom.com.br

**Resumo**—Este artigo apresenta o desenvolvimento de um sistema de conversão entre os idiomas português-libras. Desenvolvido para ser utilizado como ferramenta de auxílio a ouvintes e surdos no processo de comunicação entre as modalidades oral e gesto-visual. Por meio da técnica *HMM* utilizada para treinar e reconhecer os sinais de fala é possível implementar o reconhecedor necessário a ser utilizado no sistema de conversão. Dessa maneira, os sinais de fala são convertidos para seus respectivos sinais em libras. Para avaliar o sistema é necessário verificar a correspondência de cada sinal de fala ao seu sinal em libras ao qual se limita a apenas letras do alfabeto.

**Palavras-Chave**—sistema de conversão, reconhecedor de fala, *HMM*

## I. INTRODUÇÃO

Segundo o último Censo realizado pelo IBGE (2010), estima-se que há 9,7 milhões de pessoas com algum grau de deficiência auditiva. Desses, 2.147.366 milhões apresentam deficiência auditiva severa, situação em que há uma perda entre 70 e 90 decibéis (dB) [1].

A LIBRAS (Língua Brasileira de Sinais) é o segundo idioma oficial do Brasil de acordo com a Lei 10.436 de 24 de abril de 2002. Contudo, a difusão da língua entre ouvintes é precária e não atinge todos os espaços.

O sistema de conversão entre idiomas realiza o mapeamento dos sinais de fala para os sinais em libras que o representam por meio de um algoritmo que estabelece esta conexão. Dessa maneira, o algoritmo é capaz de identificar qual fonema o locutor alimenta na entrada do sistema e retorna uma imagem que representa o sinal em libras correspondente.

Para a produção eficiente do conversor, é preciso produzir um reconhecedor de fala capaz de capturar os sinais necessários para realizar a caracterização dos sinais de fala.

O Modelo Oculto de Markov (*HMM*) é a principal técnica utilizada para a etapa de reconhecimento de fala. Utilizando-se da etapa de treinamento do sistema, é possível obter boas taxas de reconhecimento. Ao qual, gera uma boa expectativa para a eficácia do conversor.

Por essas razões é proposto um sistema de acessibilidade a pessoas surdas visto a condição da população com algum grau de surdez no Brasil para promover uma maior interação entre surdos e leigos com a LIBRAS e estreitar a comunicação entre as pessoas.

Esse trabalho está organizado da seguinte maneira: a Seção II descreve o sistema proposto para a conversão entre idiomas e o sistema de reconhecimento de fala, enquanto a Seção III traz as considerações finais.

## II. SISTEMA PROPOSTO

### A. Diagrama de Blocos

A seguir, é apresentado na Figura 1 o diagrama de blocos construído para guiar o desenvolvimento do sistema de conversão.

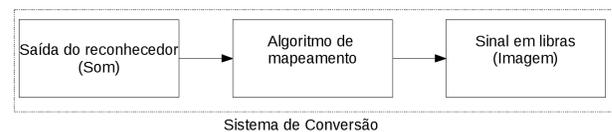


Figura 1: Diagrama de blocos do sistema de conversão.

Após a aquisição do sinal gerado pelo reconhecedor de fala, mapeia-se o sinal para gerar as imagens que correspondem àquele fonema. Esse processo final é relativamente simples tendo em vista que o algoritmo de busca precisa apenas vincular fonema/imagem.

Dessa maneira, a maior parte do trabalho se encontra na produção de um bom reconhecedor de fala que possa gerar informações confiáveis para o conversor gerar o real sinal em libras.

Atualmente, o sistema encontra-se na fase de produção do reconhecedor de fala. Pretende-se fazer testes com surdos e ouvintes para avaliação do sistema e de sua contribuição para a inclusão social.

### B. Reconhecedor de Fala

O principal meio de comunicação do homem é a fala [2]. Seu processamento para um sistema de reconhecimento sofre interferência de algumas variabilidades, tais como: fonética, acústica, intra-locutor e entre-locutor. Além da disponibilidade do sistema funcionar em condições de ruído de fundo, o que exige robustez do sistema [5] [4].

O reconhecimento de fala permite que computadores equipados com microfones interpretem a fala humana para transcrição ou comandos [3]. Para isto, mapeia-se um sinal acústico em um conjunto de palavras; ou seja, converte um sinal acústico em sua representação ortográfica [4] [5] [6]. No reconhecimento, a sequência das observações da elocução em teste é aceita como verdadeira se possuir alguma medida de similaridade (verossimilhança) acima de um limiar estipulado [7].

Portanto, o sistema de reconhecimento de fala tem por objetivo identificar qual a palavra ou frase foi pronunciada pelo locutor. Dessa forma, o reconhecimento de padrões precisa de

uma fase de treinamento e outra de reconhecimento a fim de, inicialmente, gerar uma base de dados de treinamento (fonemas, trifones, palavras, frases) como modelos de referência para o que pretende ser reconhecido. Na etapa de reconhecimento, os modelos obtidos da fase de treinamento são usados para comparação cuja regra de decisão estipula o que mais se assemelha àquele padrão [4] [8] [9].

Nesse contexto, os modelos ocultos de Markov ou HMM (*Hidden Markov Models*) são utilizados por apresentarem uma maior robustez ao sistema, tendo em vista que é capaz de modelar as variabilidades acústicas e temporais dos sinais de voz [10] [7] [8] [4].

As etapas para o desenvolvimento do reconhecedor incluem: processamento do sinal de voz, que compreende as etapas de pré-ênfase e segmentação do sinal de voz, extração de características, desenvolvimento do modelo acústico e decodificação [11]. A Figura 2 ilustra essas etapas.

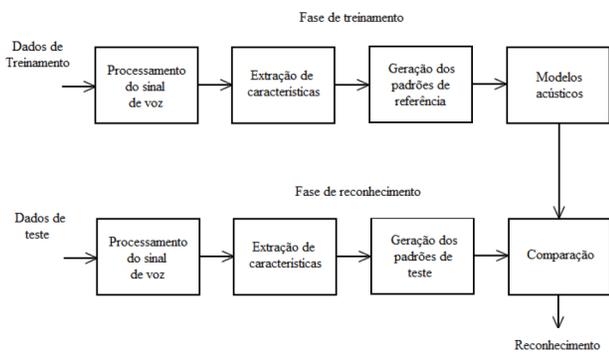


Figura 2: Diagrama de blocos de um sistema reconhecimento de fala [11].

Uma das características que torna o sinal de voz vulnerável a ruído é as baixas amplitudes nas altas frequências, tendo em vista que sua energia concentra-se nas baixas frequências. Devido a isso, é necessário pré-enfatizar o sinal passando-o por um filtro de primeira ordem para acentuar as frequências mais altas e deixar seu espectro mais plano [2] [11].

Na etapa seguinte, é preciso segmentar o sinal de voz em janelas de duração definida, desde que se respeite a característica quase estacionária do sinal de voz [11].

Após isso, extrai-se as informações mais relevantes do sinal de voz. Os coeficientes MFCC's (*Mel Frequency Cepstral Coefficients*) atendem a esse requisito por ser capaz de representar o sinal de voz baseado no ouvido humano [12] [11] [2].

Dessa forma, é possível criar um modelo matemático que represente cada segmento da fala, pois o algoritmo utiliza das características extraídas da etapa anterior. Devido às fontes de variabilidades do sinal de voz, é necessário submeter o sistema ao treinamento de um modelo que melhor generalize a descrição de um fonema a partir de um extenso conjunto de sentenças faladas pela maior diversidade de pessoas possíveis [13] [11].

Por fim, a etapa de decodificação transcreve as amostras de voz desconhecidas em sua forma textual. [4]. Para isso, o decodificador precisa de um dicionário para indicar ao

decodificador as possíveis saídas e uma rede de fonemas que indica as possíveis transcrições [11] [5].

Para a etapa de construção do reconhecedor de fala, foi utilizado neste trabalho o software livre HTK (*Hidden Markov Models Toolkit*) que consiste em um conjunto de ferramentas que modelam o HMM [13].

### III. CONSIDERAÇÕES FINAIS

O presente trabalho possui uma taxa de reconhecimento de fala em torno de 73,5% para as frases de treinamento; utilizando modelos acústicos de monofones. Ao qual é boa considerando um resultado preliminar.

Como continuação, é previsto a obtenção dos modelos acústicos dos trifones, avaliação do reconhecedor, implementação do conversor e de seu algoritmo de busca e por fim a avaliação dos resultados finais por meio do uso do conversor por surdos e ouvintes.

Dessa maneira, o presente trabalho propõe uma diminuição nas barreiras de comunicação entre surdos e ouvintes e contribuir para uma maior integração dos surdos e ouvintes nos canais de comunicação.

### REFERÊNCIAS

- [1] Apesar de avanços, surdos ainda enfrentam barreiras de acessibilidade. **Cidadania e Justiça**. set.2016. Disponível em: <<http://www.brasil.gov.br/cidadania-e-justica/2016/09/apesar-de-avancos-surdos-ainda-enfrentam-barreiras-de-acessibilidade>>. Acesso em: 20 fevereiro 2019.
- [2] Silva, A. C. da; Camilo, F. M. E.; Ferraz, T. V. D. **Reconhecimento de Fala Dependente de Locutor Utilizando Redes Neurais Artificiais**. Trabalho de Conclusão de Curso. Ouro Branco - MG. Universidade Federal de São João Del Rei - Campus Alto Paraopeba - Departamento de Telecomunicações. 2013.
- [3] Fechine, J. M. **Reconhecimento Automático de Identidade Vocal Utilizando Modelagem Híbrida: Paramétrica e Estatística**. Doutorado em engenharia elétrica, Universidade Federal da Paraíba, Campina Grande, 2000.
- [4] Ynoguti, C. A. **Reconhecimento de Fala Contínua Usando Modelos Ocultos de Markov**. Doutorado em engenharia elétrica, Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas, Campinas, 1999.
- [5] Tevah, R. T. **Implementação de um sistema de reconhecimento de fala contínua com amplo vocabulário para o português brasileiro**. Mestrado em ciências em engenharia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.
- [6] Selmini, A. M. **Sistema Baseado em Regras para o Refinamento da Segmentação Automática de Fala**. Doutorado em engenharia elétrica, Universidade Estadual de Campinas Faculdade de Engenharia Elétrica e de Computação, Campinas, 2008.
- [7] Paranaçuá, E. D. S. **Reconhecimento de locutores utilizando modelos de markov escondidos contínuos**. Mestrado em ciências em engenharia, Instituto Militar de Engenharia, Rio de Janeiro, 1997.
- [8] Rabiner, L. R. **A tutorial on hidden markov models and selected applications in speech recognition** jan. 1988.
- [9] Raphael, L. J.; Borden G. J.; Harris K. S. **Speech Science Primer**. Lippincott Williams Wilkins, 5<sup>a</sup>ed, 2011.
- [10] Silva, Carlos Patrick A. da. **Um software de reconhecimento de voz para português brasileiro**. Mestrado em engenharia elétrica, Universidade Federal do Pará, Belém, 2010.
- [11] Rocha, Raissa Bezerra. **Desenvolvimento de um codificador de voz pessoal de baixa taxa baseado em modelos de markov escondidos**. Mestrado em engenharia elétrica, Universidade Federal de Campina Grande - UFPB, Campina Grande - PB, 2012.
- [12] Picone. J. **Signal Modeling Techniques in Speech Recognition**. Proceedings of the IEEE, 1993.
- [13] S. Young et al. **The HTK Book**. Cambridge University Engineering Department, 2009.