

REALabs-BOT (Remotely Accessible Laboratory)

Navegação de Robôs Móveis por Meio de Comandos de Voz no REAL WebLab



Autor: Pedro Pereira Ribeiro – pedro.pr88@gmail.com

Orientadores: Prof. Dr. Eleri Cardozo (FEEC/Unicamp) e Dra. Eliane Gomes Guimarães (DRVC/CTI)

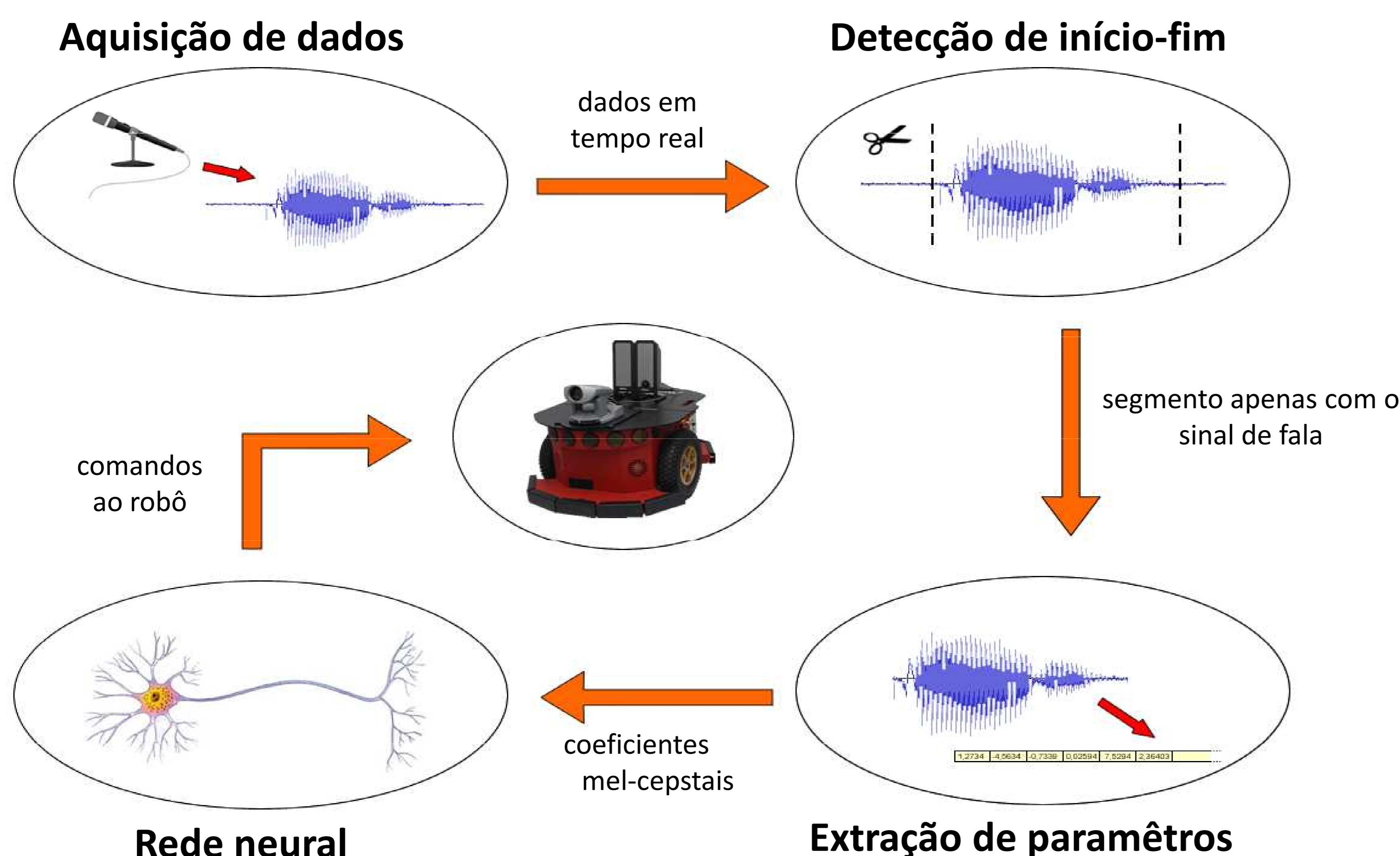
Introdução

O projeto REAL (Remotely Accessible Laboratory) tem como objetivo desenvolver um ambiente (WebLab) no qual pesquisadores situados em qualquer lugar do país tenham acesso à infraestrutura e aos equipamentos robóticos disponíveis no Departamento de Computação e Automação Industrial (DCA/FEEC) da Unicamp e na Divisão de Robótica e Visão Computacional (DRVC) do CenPRA, valendo-se de redes de alta velocidade, como as redes dos projetos RNP/Giga e Fapesp/Kyatera, ou da Internet.



A aplicação, desenvolvida neste contexto, é um sistema de reconhecimento de fala baseado na técnica de comparação de padrões, projetado de forma a ser dependente do locutor e capaz de identificar palavras isoladas. Para isso, fez-se uso de redes neurais artificiais do tipo “Perceptron Multi-Camadas”, além de certas técnicas de análise e extração de parâmetros da voz, chamados Mel-Cepstrais. Esta aplicação visa incrementar os recursos do WebLab, com a inclusão da possibilidade de controle de robôs móveis por meio de comandos falados. O vocabulário considerado aqui consiste, portanto, em cinco palavras destinadas ao controle da movimentação destes robôs. São elas: “esquerda”, “direita”, “anda”, “recua”, “para”.

Reconhecimento de Fala



No primeiro estágio, os dados vindos do microfone são captados pelo programa, e colocados no formato desejado. Esses dados são analisados em tempo real pelo detector início-fim. Este é responsável por identificar quando uma palavra está sendo pronunciada, copiar apenas o trecho de fala e acionar os módulos subsequentes.

Tendo o sinal de fala com a palavra completa, o próximo passo é realizar uma série de operações de modo a extrair dados relevantes, que possibilitem a diferenciação entre as diferentes palavras do vocabulário.

Cada palavra do vocabulário tende a produzir um vetor de parâmetros característico. Embora os valores contidos no vetor de duas elocuições de uma mesma palavra sejam diferentes, eles seguem um determinado padrão. É função da rede neural, detectar esses padrões e decidir, assim, qual foi a palavra pronunciada.

Resultados

Os resultados alcançados foram bons. Apesar de terem sido usadas técnicas relativamente simples em comparação às utilizadas em sistemas maiores, obteve-se uma taxa de acertos de 90%. Esse resultado pode ser alcançado sem os benefícios de técnicas mais avançadas, pois a aplicação exigia características simples do reconhecedor, como o vocabulário reduzido e palavras pronunciadas isoladamente. Futuramente, é possível ainda evoluir o sistema para que seja capaz de reconhecer palavras independentemente do locutor, incorporando técnicas mais avançadas ao projeto. Desta maneira, ampliam-se ainda mais as possibilidades de navegação dos robôs.