

UNIVERSIDADE FEDERAL DE VIÇOSA
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
CURSO DE ENGENHARIA ELÉTRICA

FELIPE PROCÓPIO RIBEIRO

**AUTOMATIZAÇÃO DA PLANTA DIDÁTICA ATRAVÉS DO ACIONAMENTO
DE CARGA POR COMANDO DE VOZ**

VIÇOSA - MG
2012

UNIVERSIDADE FEDERAL DE VIÇOSA
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
CURSO DE ENGENHARIA ELÉTRICA

FELIPE PROCÓPIO RIBEIRO

**AUTOMATIZAÇÃO DA PLANTA DIDÁTICA ATRAVÉS DO ACIONAMENTO
DE CARGA POR COMANDO DE VOZ**

Trabalho de conclusão de curso submetido à
Universidade Federal de Viçosa para obtenção
dos créditos referentes à disciplina Monografia
e Seminário do curso de Engenharia Elétrica.

Orientador: Prof^a. Dr^a. KÉTIA SOARES MOREIRA

VIÇOSA - MG
2012

FELIPE PROCÓPIO RIBEIRO

**AUTOMATIZAÇÃO DA PLANTA DIDÁTICA ATRAVÉS DO ACIONAMENTO
DE CARGA POR COMANDO DE VOZ**

Monografia apresentada ao Departamento de Engenharia Elétrica do Centro de Ciências Exatas e Tecnológicas da Universidade Federal de Viçosa, para a obtenção dos créditos da disciplina ELT 490 – Monografia e Seminário e cumprimento do requisito parcial para obtenção do grau de Bacharel em Engenharia Elétrica.

Aprovada em 18 de Junho de 2012

COMISSÃO EXAMINADORA

Prof^a. Dr^a Kétia Soares Moreira – Orientadora
Universidade Federal de Viçosa

“Há mais pessoas que desistem do que pessoas que fracassam.”
(Henry Ford)

DEDICATÓRIA

Dedico este trabalho aos meus pais Gilson e Andréia.

AGRADECIMENTOS

Agradeço aos meus pais, Gilson e Andréia, pelo apoio incondicional. À Gisele pela parceria que foi crucial para o sucesso dessa caminhada. Aos amigos por alegrarem o dia a dia. Agradeço também os professores Kétia e André pela paciência e conhecimento compartilhado. E mais importante, a Deus pela vida e pelas oportunidades concedidas.

RESUMO

AUTOMATIZAÇÃO DA PLANTA DIDÁTICA ATRAVÉS DO ACIONAMENTO DE CARGA POR COMANDO DE VOZ

O uso da fala tem o papel de agilizar e facilitar a realização de tarefas cotidianas, tanto na manipulação de sistemas pessoais como para controle de equipamentos e utensílios eletrônicos. Nesse contexto, esse trabalho tem como objetivo analisar qualitativamente o acionamento de uma carga por meio de técnicas de reconhecimento da voz. Os sinais adquiridos são processados, com o objetivo de eliminar dados desnecessários ou distorcidos. Após o processamento, são extraídos os Coeficientes Mel-Cepstrais do sinal de áudio que são utilizados como parâmetros de entrada para as Redes Neurais Artificiais (RNA) que tem o papel de reconhecer e classificar cada comando. De acordo com o reconhecimento, são controlados alguns dispositivos da Planta Didática da SMAR no padrão Fieldbus utilizando o protocolo OLE for Process Control (OPC).

Palavras-chave: OPC, MATLAB, Comando de voz, Redes Neurais Artificiais.

ABSTRACT

AUTOMATIZATION OF THE DIDACTIC PLANT THROUGH THE LOAD ACTUATION BY VOICE COMMAND

The use of speech has the role of expedite and facilitate the realization of daily tasks, both in handling personal systems as to control equipment and electronic utensils. In this context, this work aims to analyze qualitatively the load actuation through voice recognition techniques. The acquired signals are processed, with objective of eliminating unnecessary or distorted data. After processing, are extracted coefficients Mel-Cepstrais of the audio signal that are used as input parameters to the artificial neural networks (ANN) whose role is to recognize and classify each command. According to the recognition, are controlled some devices of didactic plant SMAR in standard Fieldbus using the OLE for Process Control (OPC) protocol.

Keywords: OPC, MATLAB, Voice Command, artificial neural networks (ANN).

Índice

1	Introdução.....	11
2	Revisão Literária.....	12
	2.1 – Coeficientes Mel-Cepstrais.....	12
	2.2-Redes Neurais Artificiais.....	12
	2.3-OPC.....	13
3	Objetivos.....	15
4	Metodologia.....	16
	4.1 Aquisição.....	16
	4.2 Reconhecimento de Padrões	16
	4.2.1-Filtros.....	16
	4.2.2-Seleção.....	17
	4.2.3- Divisão em Janelas.....	18
	4.3 Extração das Características.....	18
	4.3.1-Espectro de Potência.....	19
	4.3.2-Banco de Filtros.....	19
	4.3.3-Análise Espectral.....	19
	4.4 Reconhecimento.....	19
	4.4.1 Treinamento da RNA.....	19
	4.5 OPC.....	20
	4.5.1- Cliente.....	20
	4.5.2 – Servidor.....	21
	4.5.2.1 Exportação das Tags.....	21
	4.6 Acionamento dos Dispositivos.....	21
5	Resultados e Discussões.....	23
6	Conclusão.....	25
7	Referências Bibliográficas.....	26

1 Introdução

O uso da fala tem o papel de agilizar e facilitar a realização de tarefas cotidianas, tanto na manipulação de sistemas pessoais como para controle de equipamentos e utensílios eletrônicos. Com o advento da tecnologia que muito evoluiu, as máquinas predominam em quase todos os cenários, desde a conquista espacial até às residências. Sendo assim, nada melhor do que dotar tais equipamentos com a capacidade de percepção e compreensão da voz humana, que é a forma mais simples, natural e eficaz do ser humano expressar seus pensamentos, e desta forma humanizar mais o relacionamento homem-máquina[1].

A acústica da fala é uma característica biométrica que carrega informações fonético-linguísticas que podem ser detectadas e classificadas pelos sistemas digitais de reconhecimento de voz [1]. Entretanto, o reconhecimento da voz humana ainda apresenta falhas devido às variações de sotaques (como por exemplo: inglês americano e inglês britânico). Além dessas dificuldades, existem outros fatores externos que podem interferir na interpretação do áudio, como: variabilidade de locutor, ruído do ambiente além da questão da palavra isolada e da palavra concatenada[2 ,3].

Diante dessa situação, ainda muitas pesquisas estão sendo realizadas na tentativa de desenvolver novos métodos e tecnologias que possam contornar imperfeições de sistemas de reconhecimento da fala.

O reconhecimento da fala pode ser utilizado em ambientes industriais onde deve haver maior interação homem-máquina para facilitar a comunicação entre ambos. Na indústria, trocas de informações entre estações de trabalho são frequentes e muito importantes. Entretanto, nestes ambientes, o trabalhador fica exposto a diversas intensidades de ruídos alterando a informação acústica produzida pelo locutor.

Neste contexto, este trabalho tem o objetivo do acionamento de carga por comando de voz. Para isto, utilizará na interface o padrão OLE for Process Control (OPC). Este permite qualquer Interface Homem Máquina (HMI) acessar dados e acionar cargas no padrão Fieldbus.

2 Revisão Literária

Algumas informações são necessárias para uma maior compreensão deste relatório. Assim, nas próximas seções, algumas definições serão elucidadas, como o objetivo de proporcionar maior compreensão da metodologia, execução e resultados do projeto.

2.1 Coeficientes Mel-Cepstrais

Os Coeficientes Mel-Cepstrais são os parâmetros mais utilizados em aplicações de reconhecimento de fala. Esses coeficientes visam aproximar a forma de percepção de sons do sistema auditivo humano[1].

Os coeficientes mel-cepstrais são aqueles obtidos pela representação em frequência deformada na escala mel. Um *mel* é uma unidade de medida da frequência percebida de um tom. Não corresponde às frequências físicas do tom, à medida que o sistema auditivo humano aparentemente não percebe a diferença de frequência de forma linear. Assim, a escala mel foi construída de tal maneira que incrementos iguais na escala mel correspondam a incrementos subjetivos iguais em frequência [11,12]. A Figura 1 mostra o diagrama de blocos para a obtenção dos coeficientes mel-cepstrais.



Figura 1 – Diagrama de blocos para a obtenção dos coeficientes mel – cepstrais.

2.2 Redes Neurais Artificiais

Redes Neurais Artificiais são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência [7].

Uma rede neural artificial é composta por várias unidades de processamento que geralmente são conectadas por canais de comunicação que estão associados a determinado peso. As unidades fazem operações apenas sobre seus dados locais, que são entradas recebidas pelas suas conexões [7,10].

A propriedade mais importante das redes neurais é a habilidade de aprender conforme seu ambiente e com isso melhorar seu desempenho. Isso é feito através de um processo iterativo de ajustes aplicado a seus pesos, o treinamento. O aprendizado ocorre quando a rede neural atinge uma solução generalizada para uma classe de problemas[7].

Uma importante aplicação dessas redes é no reconhecimento de padrões (fala, caracteres, imagens, etc). Nesse caso, vetores são apresentados a uma RNA já treinada e esta é responsável por classificá-los em classes previamente definidas de acordo com o treinamento recebido [9].

A Figura 2 representa uma rede neural com três níveis.

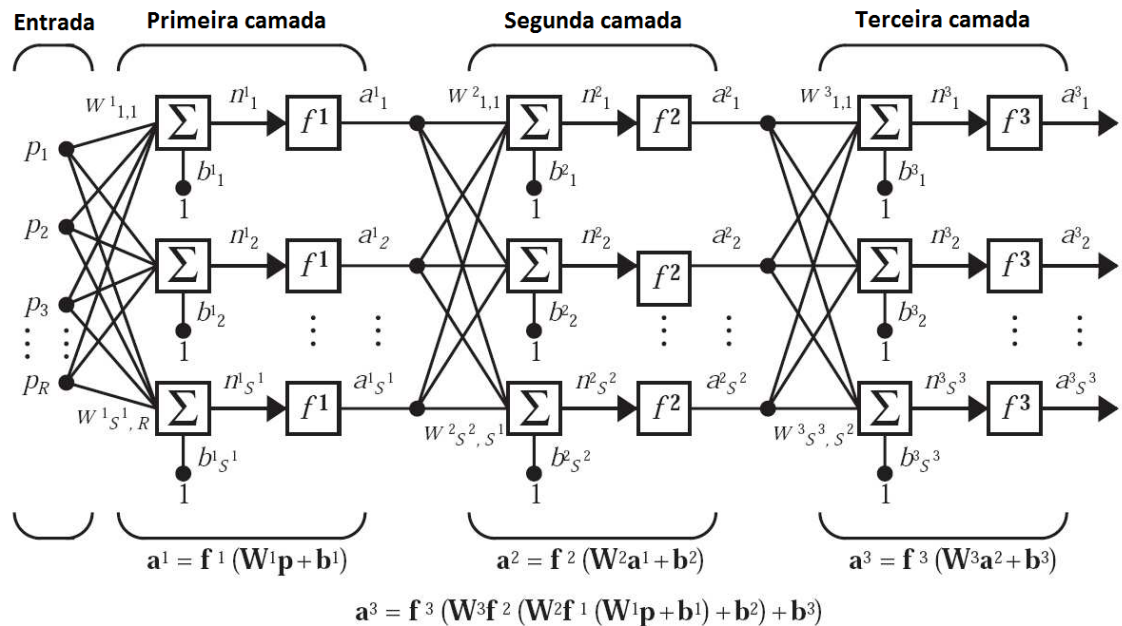


Figura 2 – Rede neural de três níveis

2.3 OPC

OPC (OLE for Process Control) é um padrão industrial publicado para interconectividade de sistema. A interoperabilidade é assegurada através da criação e manutenção de especificações de padrões abertos [8]. Inúmeras são as vantagens de se utilizar o OPC como protocolo de comunicação, entre elas estão:

- Padronização das interfaces de comunicação entre os servidores e clientes de dados de tempo real, facilitando a integração e manutenção dos sistemas;
- Eliminação da necessidade de drivers de comunicação específicos (proprietários);
- Melhoria do desempenho e otimização da comunicação entre dispositivos de automação;
- Interoperabilidade entre sistemas de gestão empresarial, de execução de manufatura e aplicações Windows.
- Redução dos custos e tempo para desenvolvimento de interfaces e drivers de comunicação, com conseqüente redução do custo de integração de sistemas;
- Facilidade de desenvolvimento e manutenção de sistemas e produtos para comunicação em tempo real;
- Facilidade de treinamento.

As plataformas industriais de programação e supervisão de processos utilizam servidores de dados para gerenciamento das informações. Como exemplo destes servidores, existe o OPC (OLE for Process Control), onde através deste, vários programas de configuração, controle e supervisão conseguem comunicar variáveis entre si e com os equipamentos físicos do processo[14].

A SMAR utiliza para configuração de seus equipamentos a plataforma SYSTEM302 e o ProcessView para supervisão, sendo estes clientes OPC. Os recursos de controle destes clientes são voltados para aplicações industriais possuindo confiabilidade e simplicidade, mas é limitado para utilizações científicas de pesquisa[8].

Uma solução para contornar este problema. É utilizar o MATLAB como cliente OPC através do Toolbox específico para este fim. Com isto torna-se possível utilizar todos os recursos de cálculos complexos do MATLAB é aplicá-los no processo físico[13].

3 Objetivos

O objetivo principal deste trabalho é analisar qualitativamente o acionamento de uma carga por meio de técnicas de reconhecimento da voz.

Para atingir a meta principal, outros objetivos específicos devem ser salientados:

- Estudar as principais técnicas utilizadas no reconhecimento de padrões da fala para fazer a identificação da fala;

- Extrair automaticamente características da voz durante a fala por meio de sensores;

- Analisar as variações entre as falas (locutores diferentes);

- Estudar o padrão OLE for Process Control (OPC) e padrão Fieldbus;

- Estudar o acesso de dados e acionamento de cargas para uma Interface Homem Máquina (HMI);

- Fazer a análise dos resultados obtidos.

4 Metodologia

Nesse capítulo, serão abordados os métodos utilizados no sistema de reconhecimento de fala e controle de dispositivos. O sistema implementado pode ser visto na Figura 3.

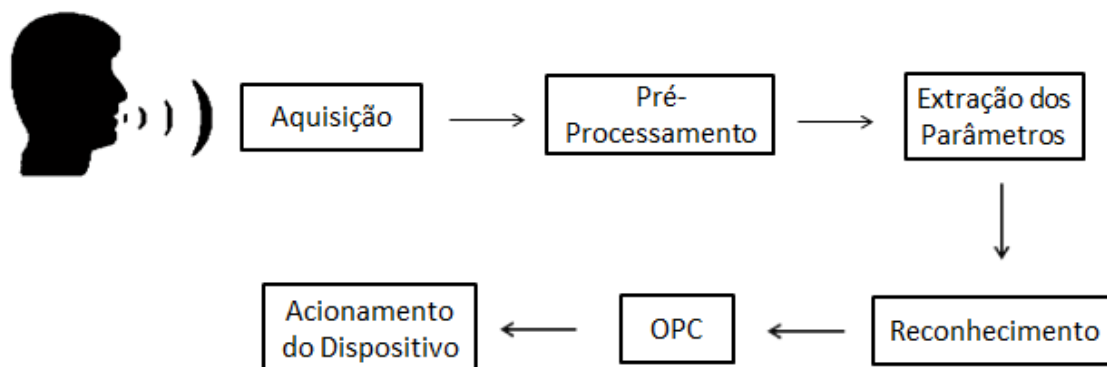


Figura 3 – Sistema de Reconhecimento de Fala e acionamento de dispositivos.

Na aquisição dos dados, o sinal de fala é amostrado. Em seguida, passa por um pré-processamento que elimina informações inúteis e atenua ruídos que possam interferir negativamente na análise dos comandos. Logo depois, são extraídos coeficientes que representam as informações mais significativas do sinal. Estes coeficientes são as entradas do mecanismo reconhecedor e a saída deste é utilizada para o controle de dispositivos.

4.1 Aquisição

A aquisição do áudio foi realizada a uma taxa de 11025 Hz. As gravações foram feitas em um local silencioso utilizando sempre o mesmo microfone.

4.2 Pré-Processamento

Os sinais gravados passaram por um pré-processamento com o objetivo de atenuar ruídos e selecionar as informações úteis. O pré-processamento é composto das etapas: Filtros, Seleção e Divisão em Janelas como mostrado na Figura 4.

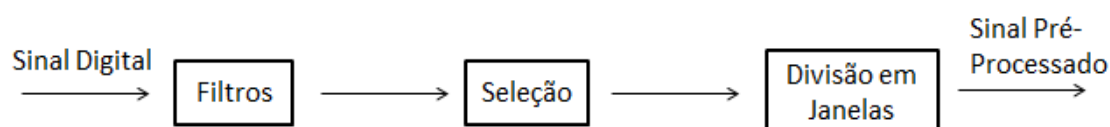


Figura 4 – Etapas do pré-processamento dos sinais.

4.2.1 Filtros

O sinal gravado passa por dois filtros. O primeiro tem a função de eliminar as informações de frequência muito alta, pois estas não pertencem à

voz humana. Foi utilizado um filtro passa-baixa de ordem 21 e frequência de corte igual a 7000Hz. O segundo filtro foi um passa-alta FIR ($1 - 0,97z^{-1}$), com objetivo de compensar a inclinação do sinal espectral. A Figura 5a mostra um sinal de áudio antes de ser filtrado e a Figura 5b um sinal depois de ser filtrado.

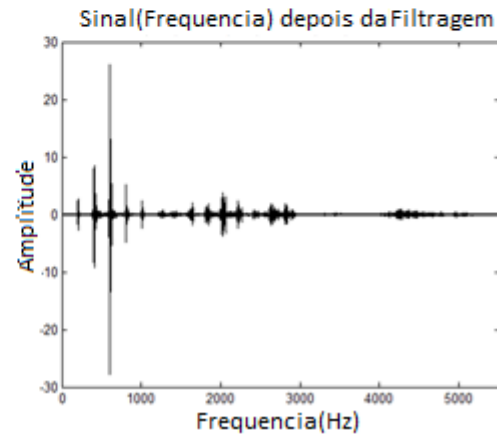
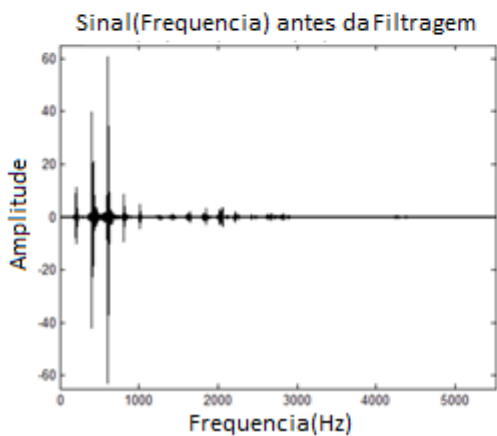
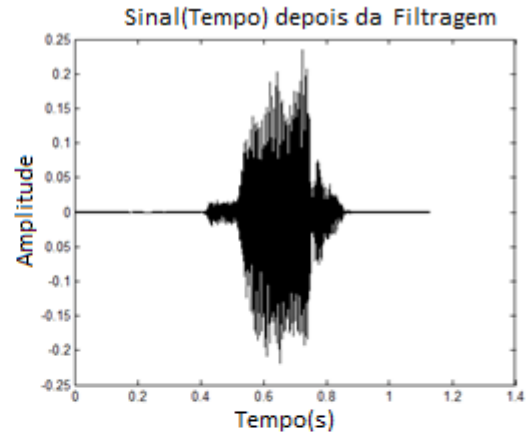
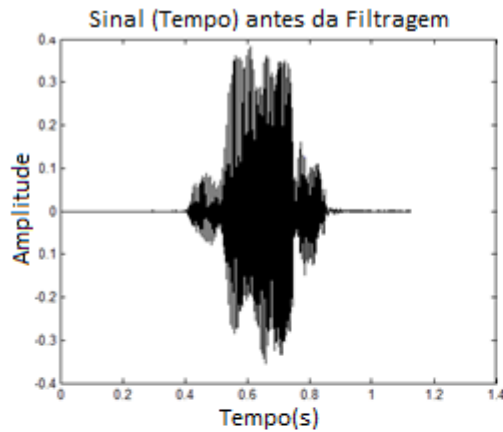


Figura 5a – Sinal antes de ser filtrado.

Figura 5b – Sinal depois de ser filtrado.

4.2.2 Seleção

O processo de seleção detecta o início e fim do comando, reduzindo assim a quantidade de dados do sinal. Essa etapa é baseada na análise de energia do sinal que é dada por:

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

A partir da curva de energia é possível localizar os limites de uma palavra isolada que apresenta picos de energia em sua duração. A Figura 6a apresenta um sinal de voz no tempo após a aquisição e a Figura 6b, a energia deste sinal.

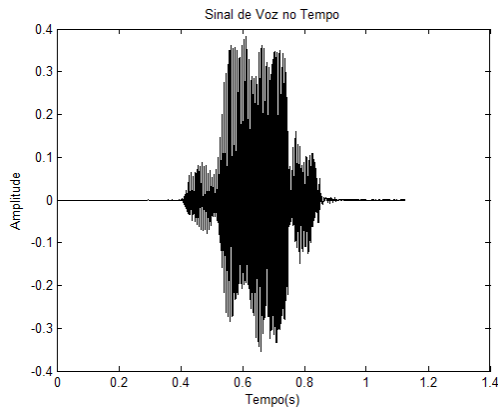


Figura 6a – Sinal de Voz no Tempo.

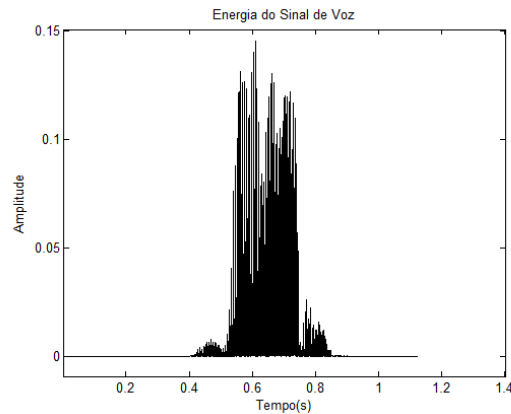


Figura 6b – Energia do Sinal de Voz.

Para encontrar os limites de início e fim de um comando, foi necessário normalizá-lo. Assim, possibilitou-se estipular um limiar fixo de energia mínima do sinal. Em casos onde o ruído é muito baixo, os limites da palavra são facilmente encontrados, pois basta determinar um patamar de energia acima da energia do ruído. No entanto, pode haver ruídos pontuais de amplitude elevada, assim fez-se necessária a implementação do método dos saltos [17]. Esse artifício proporciona maior robustez no sistema de seleção, pois uma vez que o algoritmo de busca encontra um valor que ultrapasse o limiar estabelecido, um outro ponto posterior a esse também passa pelo mesmo teste. Quando os dois pontos testados tiverem energia maior que o patamar estipulado, a amostra é selecionada. Nesse trabalho usou-se um limiar igual a 0,1% da energia máxima do sinal e o salto foi igual a 9ms.

Após encontrar os limites da palavra, criou-se um novo vetor de menor tamanho contendo somente as informações de interesse.

4.2.3 Divisão em Janelas

O sinal de áudio adquirido é fragmentado em janelas de 25ms sobrepostas a cada 10ms. Para proporcionar suavização do sinal em cada quadro é multiplicado por uma função janela de Hamming de comprimento igual a 512 amostras.

4.3 Extração das Características

Essa etapa visa reduzir a quantidade de dados sem que ocorra a perda significativa de informação. Assim, são utilizados os Coeficientes Mel-Cepstrais que se baseiam no princípio da audição humana. Esses coeficientes são obtidos conforme a Figura 7 e explicitados nas subseções seguintes.

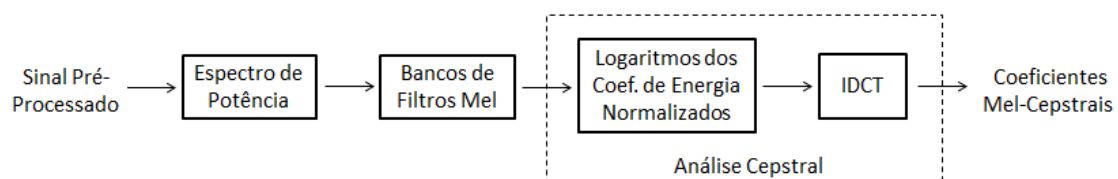


Figura 7 – Obtenção dos Coeficientes Mel-Cepstrais

4.3.1 Espectro de Potência

O Espectro de Potência é o quadrado do módulo da Transformada de Fourier do sinal.

4.3.2 Banco de Filtros

Ao espectro de Potência é aplicado um banco de filtros espaçados linearmente na escala mel. A relação que mapeia a escala MEL é dada pela Equação 2.

$$B(f) = 1125 \cdot \ln\left(1 + \frac{f}{700}\right) \quad (\text{eq. 2})$$

4.3.3 Análise Cepstral

A energia do sinal é multiplicada por cada um dos filtros do banco de filtros. Então calculou-se a log-energia da saída de cada filtro [16].

Através da Transformada Inversa do Cosseno (IDCT), os Coeficientes Mel-Cepstrais são então compactados. Nessa etapa, foram calculados 13 coeficientes.

4.4 Reconhecimento

O reconhecimento é feito pelas Redes Neurais Artificiais (RNA). Nessa etapa, os Coeficientes Mel-Cepstrais são utilizados como dados de entrada da RNA.

A entrada da RNA deve estar no formato de uma coluna única, então deve-se transformar a matriz $13 \times \text{NC}$ em uma matriz $13 \cdot \text{NC} \times 1$, sendo NC o número de colunas. Como NC depende da palavra, foram geradas matrizes de dimensões diferentes de acordo com cada comando.

Uma característica da RNA é possuir todos os dados de entrada com dimensões iguais na etapa do treinamento. Então foram acrescentados “zeros” ao final das colunas menores de modo que todas ficassem com dimensões iguais.

4.4.1 Treinamento da RNA

Nesse trabalho foram usadas duas RNAs diferentes. A primeira é responsável por classificar 3 comandos referentes aos dispositivos enquanto a segunda tem a função classificar 5 comandos que determinam o valor que o dispositivo vai assumir. As palavras a serem reconhecidas como dispositivos são: “bomba”, “vazão” e “temperatura” enquanto as palavras que representam o valor são: “Zero”, “Um”, “Dois”, “Três” e “Quatro”. Essas palavras foram escolhidas por intuição de cada e pelas diferenças fonéticas das mesmas.

Na formação do banco de dados, todos os comandos mencionados foram gravados 60 vezes por três pessoas diferentes, resultando assim em 480 amostras. Também foi gravado o ruído proveniente da bomba hidráulica em duas intensidades diferentes (SNR=4 e SNR=43). A relação sinal ruído SNR foi

calculada pela diferença de energias médias entre os sinais e essa energia pode ser calculada pela equação 3:

$$E = \frac{\sum_{k=1}^N (x^2)}{N} \quad (\text{eq. 3})$$

É importante lembrar que os dados foram normalizados antes de serem apresentados à RNA . A normalização foi calculada pela Equação 4:

$$y(n) = \frac{x(n)}{\max|x|} \quad (\text{eq. 4})$$

A RNA utilizada nesse trabalho possui uma camada oculta e o número de neurônios dessa camada foi pré-estabelecido em 5, 10, 15 e 20 para as duas RNAs. A camada de saída apresenta três neurônios para a RNA referente aos dispositivos e cinco para a RNA de valores.

O algoritmo de treinamento foi o de Levenberg-Marquardt, o de aprendizagem foi o back-propagation e a função de ativação foi a Sigmoidal. As amostras foram divididas aleatoriamente em 70% para treinamento, 30% para validação.

4.5 OPC

O OPC é responsável pela comunicação entre o MATLAB e a planta didática usando o modo Cliente/Servidor. Nesse trabalho o MATLAB exerce a função de cliente e a planta de servidor.

4.5.1 Cliente

Como citado anteriormente, o MATLAB exerce a função de cliente. Assim, tem acesso direto aos dados da planta com permissão para ler e escrever as variáveis que desejar. Para configurar o sistema, deve-se obedecer ao fluxograma apresentado na Figura 8 abaixo.

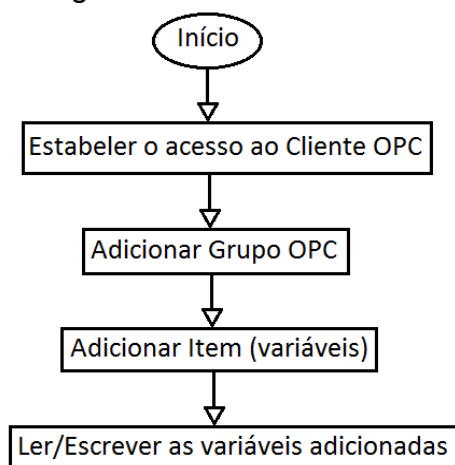


Figura 8 - Fluxograma da configuração do cliente OPC.

4.5.2 Servidor

A planta didática tem a função de “obedecer” os comandos vindos do cliente. O cliente deve acessar os blocos do servidor, no entanto deve-se configurar o servidor de modo que a comunicação com o cliente possa ser realizada [15,18]. Para configurar esse sistema, basta executar os passos abaixo:

1. Criar um novo projeto em “*HSE Project*”.
2. Inserir a “Bridge” em “*HSE Network 1*” na *CPU DF51*.
3. Criar o Canal Fieldbus H1 em *DF51*.
4. Inserir os “Devices” em *H1*
5. Adicionar os “Blocos” em *FB VFI*.
6. Caracterizar os blocos no modo Offline.
7. Parametrizar os blocos.
8. Fazer o download para a planta.
9. Exportar as Tags.

4.5.2.1 Exportação das Tags

Uma vez e o procedimento de download foi concluído, um último procedimento é exigido para habilitar a visualização da comunicação On Line com os parâmetros dos blocos funcionais. Este comando é chamado de *Export Tag*.

O Export Tag gera um arquivo contendo todos os tags (devices de blocos funcionais) presentes na configuração. Este arquivo é usado pela interface OPC (OLE for Process Control) com o propósito de supervisão.

4.6 Acionamento do Dispositivo

Tendo em vista que a comunicação com o OPC foi estabelecida, a RNA reconhece os comandos do locutor, acessa os dispositivos desejados e altera o valor dos mesmos. A Tabela 1 mostra a relação entre o comando e ação realizada pela RNA.

Tabela 1 – Comandos e Ações da RNA

RNA para os Dispositivos		RNA para os Valores	
Comando	Ação	Comando	Ação
Bomba	Liga/Desliga a bomba hidráulica	Zero	0%
		Um	25%
Temperatura	Fonte de Corrente da Planta	Dois	50%
		Três	75%
Vazão	Válvula de Vazão	Quatro	100%

Ao dizer “temperatura”, por exemplo, a RNA dos dispositivos reconhece esse comando e acessa a fonte de corrente da planta. Então deve-se dizer um número (de zero a quatro) para que a RNA dos valores possa reconhecer o comando e escrever no dispositivo previamente acessado o valor relacionado ao número proferido.

Ao proferir o comando bomba, o programa altera o valor da variável que representa a bomba hidráulica, ou seja, caso a bomba esteja desligada, ela liga e vice-versa. Isso acontece porque a bomba só pode estar ligada ou desligada, pois não apresenta modos de funcionamento intermediários.

5 Resultados e Discussões

Pela análise das figuras 3a e 3b nota-se claramente a eficácia do processamento do sinal da voz devido à acentuada atenuação dos sinais de baixa frequência.

As Tabelas 2 e 3 mostram a porcentagem de acertos dos testes realizados para a RNA dos dispositivos e a RNA dos valores respectivamente. NNCO significa número de neurônios na camada oculta.

Tabela 2 – Testes da RNA dos dispositivos.

NNCO	COMANDO	SEM RUÍDO		COM RUÍDO (SNR=4)		(COM RUÍDO) SNR=43	
		ACERTO POR COMANDO (%)	ACERTO MÉDIO (%)	ACERTO POR COMANDO (%)	ACERTO MÉDIO (%)	ACERTO POR COMANDO (%)	ACERTO MÉDIO (%)
5	BOMBA	100	100	100	100	100	100
	TEMPERATURA	100		100		100	
	VAZÃO	100		100		100	
10	BOMBA	100	100	100	100	100	100
	TEMPERATURA	100		100		100	
	VAZÃO	100		100		100	
15	BOMBA	100	100	100	100	100	100
	TEMPERATURA	100		100		100	
	VAZÃO	100		100		100	
20	BOMBA	100	100	100	100	100	100
	TEMPERATURA	100		100		100	
	VAZÃO	100		100		100	

Tabela 3 – Testes da RNA dos valores.

NNCO	COMANDO	SEM RUÍDO		SNR=4		SNR=43	
		ACERTO POR COMANDO (%)	ACERTO MÉDIO (%)	ACERTO POR COMANDO (%)	ACERTO MÉDIO (%)	ACERTO POR COMANDO (%)	ACERTO MÉDIO (%)
5	ZERO	100	100	100	100	95,2	99
	UM	100		100		100	
	DOIS	100		100		100	
	TRÊS	100		100		100	
	QUATRO	100		100		100	
10	ZERO	100	100	100	100	100	100
	UM	100		100		100	
	DOIS	100		100		100	
	TRÊS	100		100		100	
	QUATRO	100		100		100	
15	ZERO	100	100	100	100	100	100
	UM	100		100		100	
	DOIS	100		100		100	
	TRÊS	100		100		100	
	QUATRO	100		100		100	
20	ZERO	100	100	100	99,7	100	100
	UM	100		100		100	
	DOIS	100		100		100	
	TRÊS	100		98,4		100	
	QUATRO	100		100		100	

Pela análise da Tabela 3, pode-se perceber que um aumento no número de neurônios na camada oculta não implica numa melhora do desempenho da rede.

Em geral, a RNA criada não foi gravemente afetada pelos ruídos uma vez que na Tabela 3 houve acerto de 99% dos testes para SNR=43. Contudo, isso não implica na constância do desempenho em locais mais ruidosos.

O desempenho da RNA é afetado quando o tamanho do vocabulário (classes a serem classificadas pela RNA) aumenta, pois a RNA dos dispositivos que possui somente três classes teve taxa de acerto maior que a RNA dos valores que possui cinco categorias de classificação.

Os vetores utilizados como entrada da RNA eram compostos por um grande número de elementos. No entanto, não foram encontradas dificuldades na fase de treinamento.

O algoritmo realiza todas as funções num intervalo de tempo de 0,5 segundos, que é considerado razoavelmente curto.

De um modo geral, o sistema proposto obteve uma ótima eficiência na classificação dos comandos.

6 Conclusão

O método proposto foi considerado eficiente em suas etapas de extração das características da fala e reconhecimento dos comandos, o que possibilitou realizar o acionamento e controle dos dispositivos.

Assim, considerando-se o tempo de processamento do algoritmo, dentro de parâmetros compatíveis com a operação em tempo real, espera-se sua utilização na interface Humano-Computador por meio de comandos de voz.

Com os conhecimentos adquiridos nesse projeto, seria interessante aumentar o vocabulário e pesquisar técnicas de reconhecimento diferentes como Modelos Ocultos de Markov por exemplo.

7 Referências Bibliográficas

- [1] Huang, Xuedong; Acero, Alex ; Hon, Hsiao-Wuen. Spoken Language Processing. Prentice-Hall, 2001.
- [2] S. Furui. Digital speech processing, synthesis, and recognition. New York, Marcel Dekker, INC, 1989.
- [3] A.L.Trindade. Estudo de técnicas de processamento de sinais para a geração de gráficos espectrais de sinais digitais. Porto Alegre: CPGCC da UFRGS, 1994.
- [4] Nelson Alexandre Estevão. Aplicações de Redes Neurais Artificiais no Reconhecimento de Fonemas da Língua Portuguesa. Master's thesis, Universidade Federal de Minas Gerais, Belo Horizonte, Novembro 2000.
- [5] Braga, A.P.L.F. Carvalho, and T.B. Ludemir, "Redes Neurais Artificiais, Teoria e Aplicações. LTC - Livros Técnicos e Científicos Editora S. A., 2000.
- [6] Haykin, S.; "Redes Neurais, Princípios e Prática". Bookman, 2ª Edição, 2001.
- [7] Castro, A. A. M.; Prado, P. P. L.; "Algoritmos para reconhecimento de padrões".
Revista Ciências Exatas de Taubaté, vol.5-8, p. 129-145, 1999-2002.
- [8] Manuais Técnicos retirado <http://www.smar.com.br>
- [9] Ishi, Carlos T. Análise de um Sistema de Reconhecimento de Voz baseado em Fonemas. Instituto Tecnológico de Aeronáutica, Departamento de Engenharia Eletrônica e Computação, 1998.
- [10] Petry, Adriano; Reconhecimento Automático de Locutor Utilizando Medidas de Invariantes Dinâmicas Não-Lineares. Universidade Federal do Rio Grande do Sul, Agosto 2005.
- [11] Combrick, H. P.; Botha, E. C. On The Mel-scaled Cepstrum. Department of Electrical and Electronic Engineering, University of Pretoria, South Africa.
- [12] Rodrigues, G. F. Reconhecimento das Vogais Orais do Português Brasileiro utilizando Normalização de Locutores. Universidade Federal de Minas Gerais, Escola de Engenharia, Belo Horizonte, Dezembro de 2002.
- [13] Lieping, Zhang; Aiqin, Zeng; Yunsheng Zhang. On Remote Real-time Communication between MATLAB and PLC Based on OPC Technology. School of Information Engineering and Automation, Kunming University of Science and Technology. China, July 2007.

[14] Huang, Qingbao; Song, Shaojian. Real- time Control System Based on MATLAB PCAudio via OPC Techology. College of Electrical Engineering, Guangxi University. Nanning, China.

[15] OPC Foundation. OPC Data Access Custom Interface Standard Version 2.05[EB/OL](2001-12). <http://www.opcfoundation.org>.

[16] Zanotelli, Tiago. Reconhecimento de Fala de Locutor Restrito para Acionamento de Dispositivos usando Modelos Ocultos de Markov. Departamento de Engenharia Elétrica da Universidade Federal de Viçosa. Viçosa-MG, Dezembro 2008.

[17] Brandão, Alexandre S., Redes Neurais Artificiais aplicadas ao Reconhecimento de Comandos de Voz. Departamento de Engenharia Elétrica da Universidade Federal de Viçosa. Viçosa-MG, Junho 2005.

[18] Ferrari, Naraiel P., Controle de Processos Através do Smartphone. Departamento de Engenharia Elétrica da Universidade Federal de Viçosa. Viçosa-MG, Dezembro 2011.