

Universidade Federal de Itajubá
Engenharia Eletrônica

Relatório de desenvolvimento técnico Trabalho Final de Graduação

*"Desenvolvimento de uma babá eletrônica com
reconhecimento de choro infantil utilizando redes
neurais artificiais"*

Ciclo de desenvolvimento:

2016

Aluno:

Gustavo Soares Leal, 18983

Orientador:

Prof. Dr. Giscard Francimeire Cintra Veloso

Itajubá, MG

16 de setembro de 2016

Sumário

1	Introdução	2
2	Proposta de trabalho	3
3	Análise do tema	4
3.1	Projeto de sistemas embarcados	4
3.1.1	Especificação de componentes	4
3.2	Aquisição e tratamento de sinais analógicos	4
3.3	Processamento digital de sinais sonoros	4
3.3.1	Audição humana	4
3.3.2	Amostragem	5
3.3.3	Extração de características	5
3.4	Comunicação via radiofrequência	6
3.4.1	Meios de transmissão	6
3.4.2	Modulação digital	6
3.4.3	Transmissão e recepção de dados	6
3.5	Classificação por redes neurais artificiais	6
4	Metodologia de trabalho	7
5	Cronograma	8

1 Introdução

Atualmente, há dispositivos eletrônicos comercializados para monitoramento de crianças, como a chamada babá eletrônica. Este tipo de dispositivo, em sua maioria, confia na reprodução do som ambiente do quarto onde está uma criança. Os responsáveis são alertados quando escutam a reprodução dos sons emitidos pela criança.

Este trabalho apresenta o desenvolvimento de um sistema embarcado capaz de detectar choro infantil e enviar, remotamente, um alerta vibratório para um dispositivo remoto, funcionando assim como uma babá eletrônica com foco principal em usuários surdos ou com alguma dificuldade de audição.

O projeto inclui o desenvolvimento de hardware e software, desde a escolha de componentes e o projeto da placa de circuito impresso, até o desenvolvimento do software, responsável por todo o processamento, que será embarcado na placa.

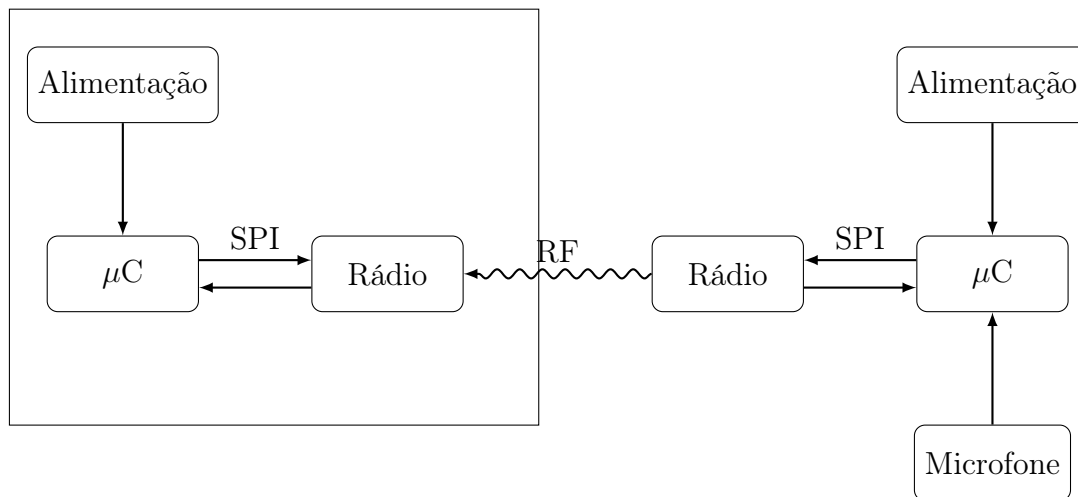
O sistema, na visão do hardware, é constituído de duas partes: uma Central, que é responsável por toda a captura e processamento dos dados e um Receptor, ao qual a Central envia, via radiofrequência, o aviso sobre a detecção do choro.

O software é responsável por fazer a captura de sinais sonoros, processá-los e extrair deles características adequadas para a detecção do choro. O software, munido destes dados, passa a utilizá-los no processamento para realizar a detecção do choro utilizando técnicas de inteligência artificial.

Posteriormente, se o choro for detectado, o software deve utilizar o hardware de rádiofrequência para enviar, da Central para o Receptor, um sinal que indica tal detecção. O receptor, utilizado como uma pulseira, então, encarrega-se de emitir sinais vibratórios, visuais e sonoros para alertar seu usuário.

2 Proposta de trabalho

O sistema desenvolvido será apresentado em suas duas partes: *hardware* e *software*. O *hardware* é constituído por duas placas, sendo um transmissor, responsável pela captação e processamento do som, e receptor, que receberá um sinal gerado pelo transmissor, ativando, assim, os mecanismos físicos de alerta. O transmissor é constituído por quatro blocos principais: aquisição de som, rádio, processamento e alimentação. O bloco de aquisição de som é constituído por um microfone do tipo MEMS e componentes passivos, sendo os últimos necessários para o ajuste de ganho e frequência de corte do microfone. O sinal de saída do microfone é ligado a uma porta do microcontrolador com função de conversor analógico/digital. O bloco de rádio possui um transceptor de radiofrequência de baixa potência trabalhando na frequência de 915MHz. Há também, neste bloco, uma antena monopolo desenhada na própria placa de circuito impresso, bem como os componentes passivos necessários para o casamento de impedância entre a antena e o transceptor. Este bloco comunica-se com o microcontrolador via protocolo SPI. O bloco de processamento contém um microcontrolador responsável por receber o sinal do microfone, realizar seu processamento e classificação e comunicar-se com o rádio. O bloco de alimentação foi projetado de forma que a placa possa ser alimentada de duas maneiras: utilizando uma fonte DC ou pilhas do tipo AAA. Adicionou-se, portanto, diodos de proteção e uma chave liga-desliga. O bloco contém um regulador de tensão para garantir que 3,3V sejam supridos aos componentes.



3 Análise do tema

3.1 Projeto de sistemas embarcados

Um sistema embarcado é um sistema computacional geralmente dedicado a uma tarefa específica, sendo parte de um sistema com outros componentes, como partes mecânicas, sensores, etc. Neste trabalho, os requerimentos para o sistema prezam pela capacidade de processamento, economia de energia, baixo custo e segurança

3.1.1 Especificação de componentes

Uma vez definidos os requerimentos para o sistema a ser desenvolvido, a especificação dos componentes a serem utilizados devem levar em conta fatores diversos: a tensão de operação do equipamento, a tolerância dos componentes a variações elétricas e de temperatura, a capacidade de processamento, o consumo de energia e o custo.

Ao começar pela escolha do microcontrolador, optou-se por utilizar um Freescale (NXP) MKL25Z128, que conta com um núcleo ARM Cortex-M0+. Esta decisão levou em conta o baixo custo e a vasta gama de recursos (suporte e software) disponíveis para a arquitetura ARM.

Na parte analógica, o microfone utilizado foi escolhido por ter um *footprint* reduzido, permitindo maior área útil na placa de circuito impresso, e boa relação sinal-ruído.

3.2 Aquisição e tratamento de sinais analógicos

3.3 Processamento digital de sinais sonoros

É ampla a utilização de microprocessadores ou microcontroladores para processamento digital de sinais. Entre as vantagens desta aplicação, destacam-se a facilidade de realizar modificações - bastando apenas alterar o código - e a maior precisão, uma vez que não são utilizados componentes cujos valores podem variar em função de pequenas alterações de temperatura ou tempo. [1]

3.3.1 Audição humana

O ouvido humano é constituído por estruturas responsáveis por captar vibrações do ar e transmiti-las ao cérebro. No ouvido externo está localizado o canal auditivo, um pequeno tubo que se estende em direção à cabeça. Sua função é direcionar o som ambiente para os ouvidos médio e interno, que se localizam dentro do crânio. Na extremidade final do canal auditivo está o tímpano, uma fina camada que vibra quando atingida pelas ondas sonoras. O ouvido médio constitui-se de pequenos ossos que transferem a vibração do tímpano para a cóclea, localizada no ouvido interno.[2] A cóclea

é uma pequena estrutura espiral que possui um líquido em seu interior. O movimento do líquido coclear causado pelas vibrações deforma a membrana basilar. As características destas deformações dependem da região atingida da membrana, conferindo-a a capacidade de distinguir diferentes frequências.[3]

3.3.2 Amostragem

Dado um sinal contínuo e real $x(t)$, o processo de amostragem consiste em mensurar o valor desta função a cada T segundos, obtendo assim uma função $x(nT)$, onde n é um valor inteiro maior ou igual a zero correspondente ao número da amostra. O valor T é chamado *tempo de amostragem* e o valor $\frac{1}{T} = f_s$ é a *frequência de amostragem*. De acordo com o teorema de Nyquist, a frequência de amostragem f_s deve ser, no mínimo, duas vezes maior que a maior frequência presente no sinal. Desta forma, é possível reconstruir o sinal a partir de suas amostras.

3.3.3 Extração de características

A extração de características de um sinal é um processo onde identifica-se alguma propriedade do sinal e suas componentes - frequência, por exemplo, para gerar uma representação paramétrica deste sinal.

Cepstrum O cepstrum resulta da transformada inversa de Fourier do logaritmo natural da densidade espectral de potência de um sinal, que, por sua vez, corresponde à magnitude da transformada de Fourier de um sinal, elevada ao quadrado. Portanto, para um sinal $x(t)$ real, temos:

$$cepstrum = \mathfrak{F}^{-1} \left\{ \ln(|\mathfrak{F}\{x(t)\}|^2) \right\}$$

Em aplicações de reconhecimento de voz, utiliza-se o cepstrum de potência, que é dado pelo quadrado da magnitude do cepstrum, isto é:

$$cepstrum \text{ de potência} = |cepstrum|^2 = \left| \mathfrak{F}^{-1} \left\{ \ln(|\mathfrak{F}\{x(t)\}|^2) \right\} \right|^2$$

MFCC Os MFCC (*Mel Frequency Cepstral Coefficients*) são características largamente utilizadas em aplicações de reconhecimento de voz.[4]

A implementação para obtenção destes coeficientes começa pela divisão de um sinal discreto no tempo $S(n)$ em i *frames* curtos, cada um com N amostras, gerando assim os *frames* $S_f(k)$, onde f representa o número do *frame* e k , suas amostras. Para cada um destes *frames*, os seguintes processos são realizados:

- Transformada discreta de Fourier (DFT)

$$S_f(k) = \sum_{n=1}^N S(n)$$

3.4 Comunicação via radiofrequência

3.4.1 Meios de transmissão

3.4.2 Modulação digital

3.4.3 Transmissão e recepção de dados

3.5 Classificação por redes neurais artificiais

4 Metodologia de trabalho

O desenvolvimento começa pelo projeto de hardware, incluindo seleção de componentes, esquemático e layout da placa de circuito impresso. Posteriormente, com a placa de circuito impresso fabricada, inicia-se o desenvolvimento de software.

Para realizar o reconhecimento de choro infantil, deve-se extrair características de sons previamente gravados. Neste caso, a técnica utilizada consiste em extrair os MFCC (Mel-Frequency Cepstrum Coefficients). Estes coeficientes, utilizados em aplicações de reconhecimento de voz e fala, constituem uma representação paramétrica compacta dos sinais de áudio e são resultado da transformada cosseno do logaritmo da parte real do espectro de energia de curto-prazo, expressos na escala de frequência mel.

mfcc_implementation

O algoritmo para a extração dos coeficientes é implementado em um microcontrolador da família ARM Cortex-M0+. Utilizando uma porta *serial* como saída, o microcontrolador é conectado a um computador que recebe os coeficientes calculados.

Para extrair as características do choro infantil, utiliza-se amostras sonoras pré-gravadas. A placa é colocada próxima a uma caixa de som, numa distância análoga a que estaria se posicionada num berço. Os sons são reproduzidos e os coeficientes gravados. Repete-se o procedimento para os sons negativos, isto é, que não representam choro, como ruídos de carros, chuva, animais domésticos e construções.

Os coeficientes, então, são identificados e importados para a ferramenta *MATLAB*®, onde utiliza-se a *Neural Network Toolbox* para treinar uma rede neural para reconhecimento de padrões.

Após ajustes e configuração do algoritmo, obtidos resultados satisfatórios, isto é, um percentual o mais próximo possível a 100% de acertos, obtém-se os pesos e vieses da rede. Um algoritmo para identificação do choro utilizando a rede neural previamente treinada é implementado no microcontrolador.

Quando o choro é detectado, o microcontrolador deve enviar um sinal para um módulo de radiofrequência, que transmite a informação para um receptor em forma de pulseira. Este receptor, também dotado de um módulo de radiofrequência e de um microcontrolador da família AVR, então, emite sinais vibratórios, visuais e sonoros para alertar o usuário.

5 Cronograma

	Março	Abril	Maio	Junho	Julho	Agosto	Setembro	Outubro
Projeto de placas de circuito impresso	X	X						
Fabricação de placas de circuito impresso	X	X						
Programação e testes de rotinas para captação de som		X	X	X				
Desenvolvimento e treinamento de redes neurais		X	X	X				
Testes em situações reais				X	X			
Avaliação de resultados					X	X		
Propostas de otimização						X	X	X

Referências

- [1] J. L. Broge. (). The benefits and advantages of digital signal processing, endereço: <http://articles.sae.org/14409/>.
- [2] S. W. Smith, *The scientist and engineer's guide to digital signal processing*. California Technical Publishing, 1999, ISBN: 0-9660176-6-8.
- [3] D. Purves, D. Fitzpatrick, L. Katz, A. Lamantia, J. McNamara, S. Williams e G. Augustine, *Neuroscience*, 3^a ed. Sinauer Associates, 2004, pp. 290–291, ISBN: 0878937269. endereço: <https://books.google.com.br/books?id=F4pTPwAACAAJ>.
- [4] J. Benesty, M. Sondhi e Y. Huang, *Springer handbook of speech processing*, sér. Springer Handbook of Speech Processing. Springer Berlin Heidelberg, 2007, ISBN: 9783540491255. endereço: <https://books.google.com.br/books?id=Slg10ekZBkAC>.