MINISTÉRIO DA DEFESA
EXÉRCITO BRASILEIRO
DEPARTAMENTO DE CIÊNCIA E TECNOLOGIA
INSTITUTO MILITAR DE ENGENHARIA
Seção de Engenharia Elétrica / SE 3

Relatório Final

Submetido ao Programa de Bolsas de Iniciação em Desenvolvimento Tecnológico e Inovação (PIBITI-CNPq/IME) Edital 2018/2019

Título:
Processamento de sinais em tempo real utilizando LabVIEW.

Aluno:
Matheus Shiraga Rasera, 3º ano.

Orientador:
José Antonio Apolinário Junior, D. Sc.

Rio de Janeiro
2019

Conteúdo

1 Introdução 3

2 Fundamentação Teórica 4

2.1. Programação em LabVIEW do *myRIO* 4

2.2. Microfone de Eletreto 5

2.3. Sinal do Tiro 6

2.4. Identificação do Sinal Acústico como Disparo de Arma de Fogo 7

2.5. Escopo do projeto e dos testes 9

3 Resultados 10

3.1. Configuração do FPGA – *Acquisition Loop* (*FPGA Main.vi*) 10

3.2. Inicialização da aquisição e processamento em paralelo – Código principal (*RT Main.vi*) 11

3.3. Rotina de enviar alerta (email) 12

3.4. Rotina de processamento e identificação – Algoritmo GCC-PHAT 14

4 Resultados 14

5 Conclusão 16

6 Referências Bibliográficas 17

Resumo

O presente projeto destina-se a captar um sinal impulsivo, correspondente a um tiro de fuzil, interpretá-lo e enviar um alerta a um servidor da segurança pública, via email. O trabalho foi fundamentado no uso do hardware *NI myRIO* da National Instruments, de modo a propiciar o processamento dos dados obtidos em tempo real.

Abstract

The present project is intended to capture an impulsive signal corresponding to a rifle shot, interpret it and send an alert to a public safety server via email. The work was based on the use of *NI myRIO* hardware from National Instruments, in order to allow the processing of the data obtained in real time.

1. Introdução

A detecção automática de disparos por armas de fogo tem utilidade na segurança pública, de forma a alertar – de forma rápida e eficiente –, quando um confronto armado está ocorrendo. Nesse trabalho, procura-se identificar o som de um disparo de um armamento, por exemplo, um fuzil 7,62 mm, enviando um sinal de alerta aos servidores de segurança pública.

Este projeto busca colaborar com uma solução eficiente ao monitoramento de ruas e praças de cidades, e que consiga efetuar a detecção de disparos de armas de fogo, de modo a ajudar no combate à criminalidade e assegurar a segurança pública.

Diferentemente de projetos anteriores, que buscaram realizar a localização de alvos utilizando arranjos de sensores acústicos para, por exemplo, identificar a posição de um *sniper* [1][2][3], este buscará somente reconhecer o disparo de uma arma de fogo, e utilizará somente um microfone em cada um de seus aparelhos. Embora somente um aparelho não seja capaz de localizar o alvo, com vários espalhados em uma determinada área, é possível estimar a localização do alvo utilizando o método da triangularização, baseado no alcance dos microfones utilizados.

Um aspecto importante do projeto é a capacidade de processar continuamente informações em tempo real, não perdendo informações do ambiente em que for implementado.

O sistema, constituído de um *myRIO* da National Instruments e um microfone eletreto a ele conectado via conector P2, exigirá pouca manutenção e pode ser, potencialmente, um substituto às tecnologias similares comercialmente disponíveis no exterior.

2. Fundamentação Teórica

2.1. Programação em LabVIEW do *myRIO*

Para adquirir e processar em tempo real os sinais acústicos adquiridos pelo microfone, foi necessário o estudo de funções e arquiteturas de programação específicas para o problema.

Inicialmente, foi necessária a configuração do hardware, assim como o *download* de seus respectivos drivers e módulos necessários. Foram utilizados os módulos de *Real-Time (FIFO)* e *myRIO*, além do módulo *Audio Data Plugin*. Na Figura 1, pode-se observar o hardware e na Figura 2, o microfone de eletreto utilizados.

Figura 1: Hardware *NI myRIO* utilizado no projeto.

Figura 2: Microfone de eletreto com cabo conector P2 para conexão no *myRIO*.

O que permite o processamento e a aquisição de sinais ocorrendo constantemente e simultaneamente é a programação em loops paralelos, no qual um dos loops é responsável pelo recebimento do sinal (loop determinístico), e outro responsável pela análise dos sinais recebidos e comparação com o padrão de áudio gravado (loop não-determinístico) [11]. Este último loop é chamado de não-determinístico pois, mediante uma estabelecida condição (sinal recebido é identificado como disparo de arma de fogo), ele envia o sinal de que o evento ocorreu sem ter um limite temporal inflexível.

2.2. Microfone de eletreto

O microfone utilizado foi particularmente projetado para o projeto, tendo um pré-amplificador acoplado ao mesmo. O circuito pré-amplificador pode ser observado na Figura 3. Ele fornece um ganho de, aproximadamente, 100 vezes, com uma saída de $\pm 1\,V$ ($2\,V$ caso utilizasse o circuito opcional). O canal da esquerda (*Audio IN Left*) contém o sinal amplificado a ser utilizado, enquanto o da direita (*Audio IN Right*) mantém somente ruídos do sinal.

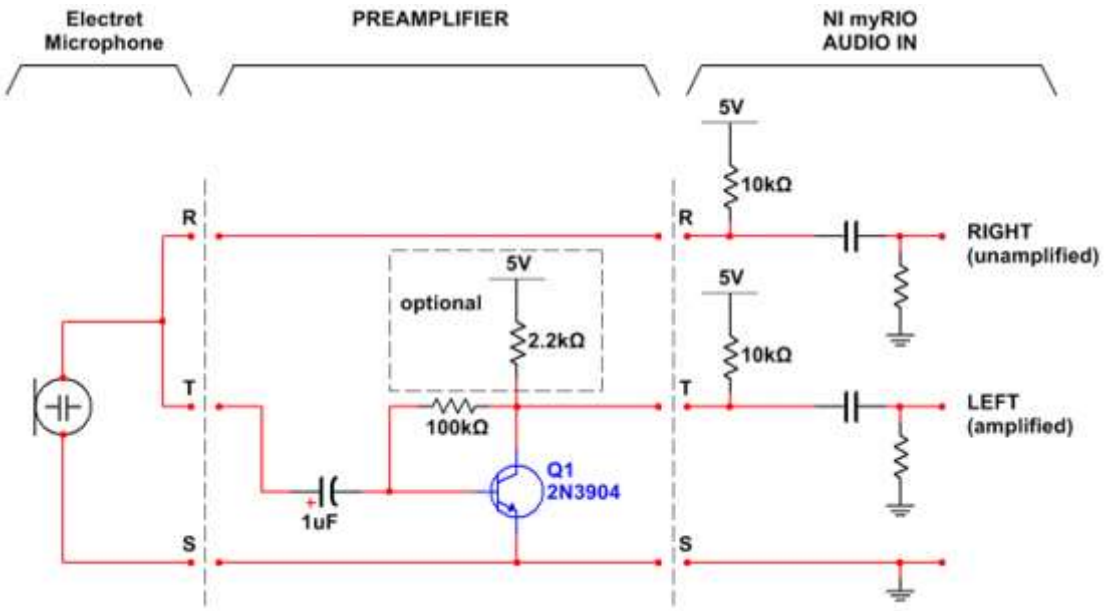

Figura 3: Circuito pré-amplificador conectado entre o microfone de eletreto e a entrada de áudio do *myRIO*. A resistência opcional de 2,2 $k\Omega$ conectada ao fornecimento de 5 V do *myRIO* dobra o sinal para $\pm$ 2 V [13].

2.3. Sinal do Tiro

Um disparo de arma de fogo produz um som gerado pela deflagração do cartucho na arma e outro gerado quando o projétil em movimento dispersa as moléculas do meio que se encontram em sua trajetória [5]. Esse som gerado é identificado por duas principais ondas e características: a *muzzleblast*, proveniente da explosão da carga propelente do projétil; e a *shockwave*, gerada somente enquanto o projétil estiver em velocidade maior que a velocidade do som.

A primeira – *muzzleblast* – é um sinal emitido em todas as direções mas com o mais alto nível sonoro para a frente da arma [4]. Ao se propagar pelo meio, a onda sofre distorção e atenuação em função da distância, reflexões, absorção atmosférica, gradientes de temperatura e ventos [6], sendo assim fatores que dificultam a identificação desse componente da onda a partir de uma certa distância. Essa dificuldade se deve ao aumento do ruído em relação ao sinal que chega ao microfone (*Signal-to-Noise Ratio*), mas que pode ser contornada ao se implantar um número razoável de sistemas de identificação de disparos, assim sendo a distância à fonte seria reduzida.

Já a segunda – *shockwave* – é também conhecida como onda-N, pois apresenta rápida ascensão e queda, tornando-a parecida com a letra "N". Entre suas principais

características, possui uma duração de aproximadamente 0,5 ms, sendo mais impulsiva que a onda *muzzleblast*, que tem duração de aproximadamente 3 ms. Entretanto, diferentemente da *muzzleblast*, a *shockwave* somente propagará seu sinal em um cone de Mach na direção em que o disparo foi executado. O ângulo entre a onda de choque e a trajetória do projétil, também chamado de ângulo de *Mach*, é mais agudo quanto maior for a velocidade do projétil [7]. Comparada com a *muzzleblast*, a *shockwave* apresenta frequências bem maiores e maior energia na medida em que o microfone estiver próximo da trajetória do projétil. Na Figura 4, é possível observar as diferentes componentes do sinal acústico gerado pela realização do tiro.

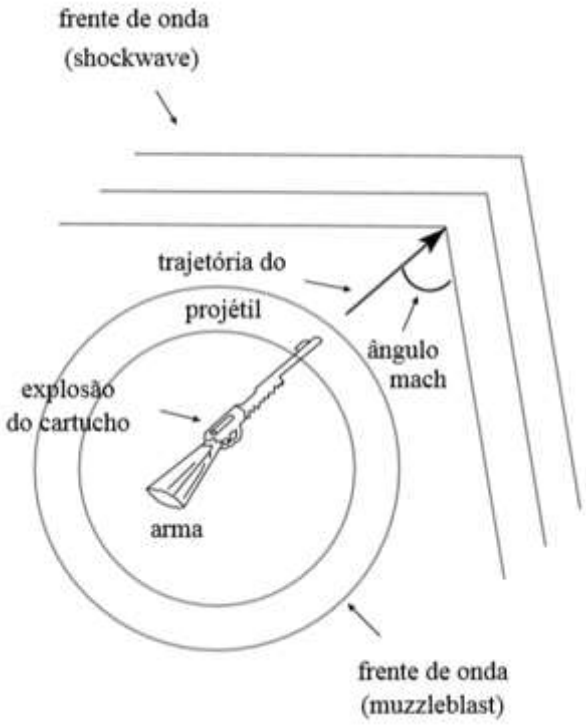

Figura 4: Sinais acústicos originados a partir do disparo de uma arma de fogo [5]. O sinal *shockwave* só é gerado enquanto o projétil tiver uma velocidade maior que a do som.

Observa-se ainda que o componente *shockwave* somente é observado quando o microfone estiver a frente do armamento, no seu campo de visão (FOV, ou *Field Of View*).

2.4. Identificação do Sinal Acústico como Disparo de Arma de Fogo

Existem diversas pesquisas que buscam detectar e classificar os tiros, assim como comparar a eficiência dessas diversas teorias [7][8][9][10]. Como grande parte delas é custosa em termos energéticos e computacionais, optou-se por uma abordagem mais simples, de forma a privilegiar o processamento em tempo real dos dados adquiridos.

A partir da gravação de um "tiro limpo", no qual as condições do ambiente são conhecidas e sabe-se que o sinal é proveniente de uma arma de fogo, obtém-se um padrão para comparar os próximos sinais a serem adquiridos. Com um banco de dados de sinais-padrão mais abrangente, a eficiência da detecção poderia ser aprimorada, diminuindo o número de falsos-positivos e negativos.

Obtendo-se uma amostra de mesmo tamanho do sinal-padrão, realiza-se uma correlação cruzada entre ambos os sinais. O valor máximo dessa correlação é comparado com o valor limite (*threshold*) para identificação de ocorrência de picos. Este limite será inicialmente obtido a partir do ponto comum de dois histogramas formados por uma amostra de correlações máximas de cinquenta sinais acústicos previamente definidos como tiros ou não. O valor obtido será um padrão para avaliar o sinal acústico e definirá taxas de falso-positivo e falso-negativo. A Figura 5 mostra, como exemplo, um padrão em que a taxa determina valores iguais de falso-positivos e falso-negativos, taxa esta chamada de *Equal Error Rate* (EER) [12]. A linha em vermelho poderia indicar uma taxa em que os índices de falso-positivos são menores, com respectiva adição aos índices de falsos-negativo.

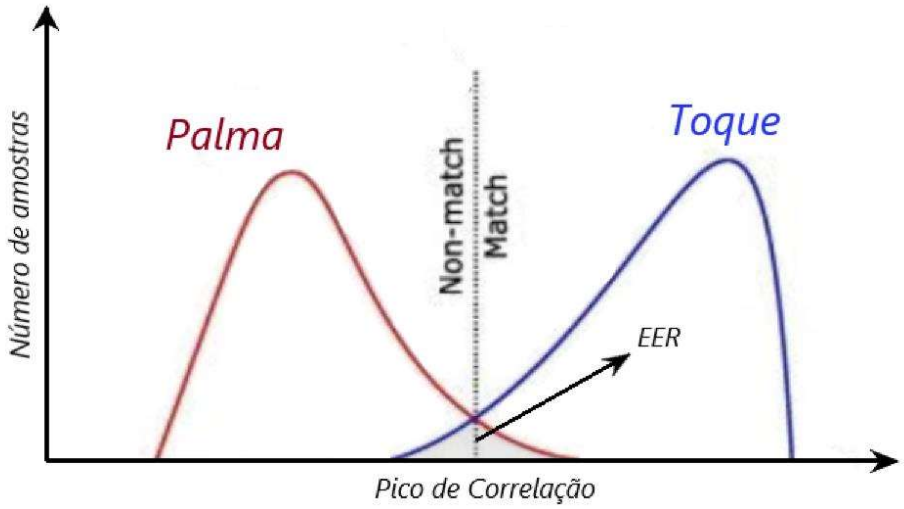

Figura 5: Indicação do EER a partir de trechos de dois histogramas. O EER define uma taxa igual de falsos-positivo e falso-negativos [12].

É importante destacar que, como os dados para a formação dos histogramas não são necessariamente uniformes, a taxa obtida inicialmente pode ser diferente da EER. Assim, pode-se verificar posteriormente a variação do valor limite para atingir um resultado que almeje uma menor taxa de falsos-positivo ou falsos-negativo, consequentemente variando a taxa de erro. Deste modo, é possível adequar os resultados acusados pelo sistema de acordo com as necessidades dos usuários.

Por fim, sendo ambas as componentes *muzzleblast* e *shockwave* do sinal impulsivas, seria necessário diferenciá-las de modo a realizar o processo acima. Para isso, seria necessária a implementação de um algoritmo de tomada de decisões, classificando as ondas em suas componentes baseadas na frequência que atingia o microfone. Deste modo, frequências acima da faixa de 1.6 kHz como *shockwave*; e as abaixo desse limite como *muzzleblast*. Contudo, levando em conta a dificuldade em gravar sinais de tiros reais, tal característica do sistema detector automático de tiros não foi implementado no desenvolvimento desse projeto.

2.5. Escopo do projeto e dos testes

Dada a já mencionada inviabilidade de ir com frequência a um campo de avaliações especializado para realizar testes de detecção de tiros, o projeto contempla a detecção de um pequeno sinal impulsivo, caracterizado por um pequeno "toque" no microfone, conforme a Figura 6 abaixo.

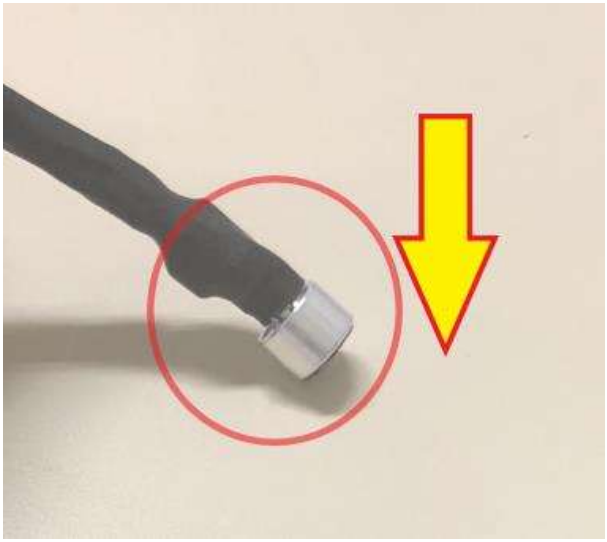

Figura 6: Caracterização do "toque" do microfone na mesa, utilizado para os testes dos áudios.

Essa necessidade gera algumas diferenças com o projeto proposto, destacando-se a maior dificuldade de se detectar o sinal, uma vez que o tiro apresenta um comportamento bem característico e picos de amplitudes maiores (melhorando a relação sinal-ruído); e tirando-se a necessidade de colocar uma estrutura de tomada de decisão para distinguir os sinais *muzzleblast* e *shockwave*.

3. Resultados

3.1. Configuração do FPGA – *Acquisition Loop* (*FPGA Main.vi*)

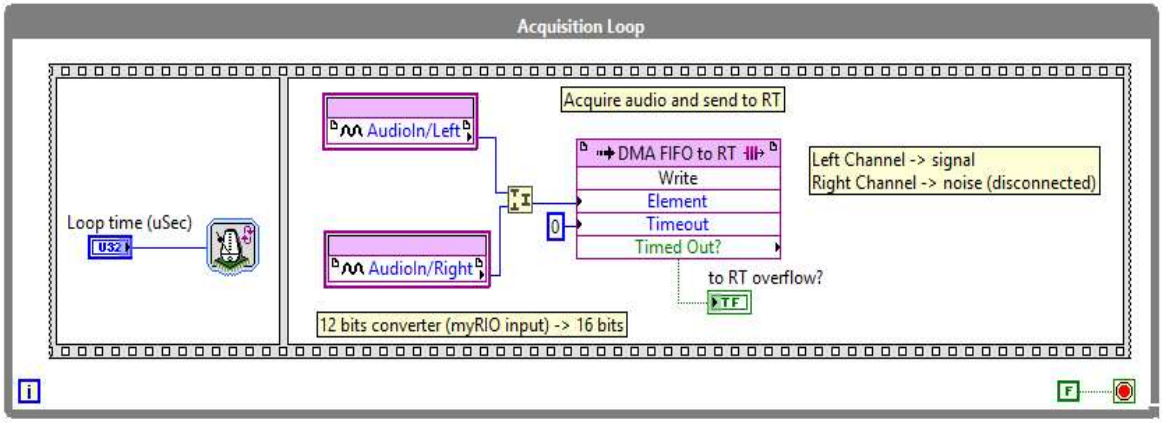

Figura 7: Código de configuração do FPGA: *Acquisition Loop* da *FPGA Main.*

Configurou-se o FPGA do *myRIO* para adquirir, pelo método DMA (*Direct Memory Access*), os sinais de áudio capturados pelo microfone de eletreto. Esse método busca fornecer um acesso rápido e igualmente eficiente aos dados adquiridos, possibilitando o processamento em tempo real da aplicação. O código pode ser visto na Figura 7.

O período de amostragem, definido pelo controle *"Loop Time (uSec)"*, foi inicializado no código principal, sendo de 23 μs, equivalente a uma frequência de amostragem de, aproximadamente, 43.478 kHz.

Conforme visto no circuito interno do microfone de eletreto, o canal esquerdo possui a informação desejada (*amplified*).

3.2. Inicialização da aquisição e processamento em paralelo – Código principal (*RT Main.vi*)

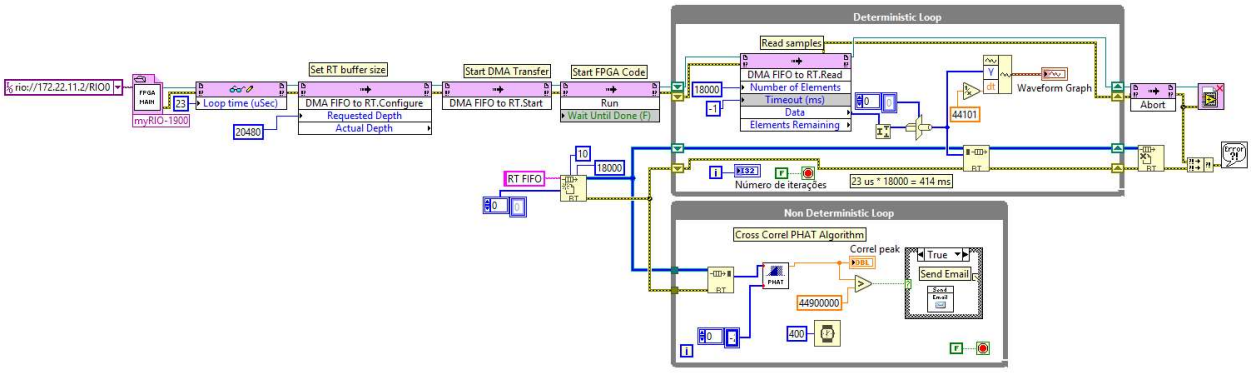

Figura 8: Visão geral do código implementado.

No código principal, configurou-se a rotina do FPGA, arbitrando uma taxa de amostragem de 43.478 kHz, ou seja, com um período de amostragem de 23 μs, e capacidade para processar até 20480 amostras de cada vez, *buffer* este configurado para ser um pouco maior que o número de elementos no frame para evitar *overflow*. O código pode ser visto em seu todo na Figura 8 acima.

No loop determinístico, foram enviados frames de 18000 elementos a uma fila do RT FIFO (fila *Real Time First In First Out*), esta que permite o acesso em tempo real dos elementos nela armazenados e que garante que não haja perda de dados. Com isso, espera-se que o loop tenha uma duração máxima total de 414 ms. Já no loop não determinístico, como o tempo de *benchmarketing* de 5 ms foi calculado para realizar o processamento do algoritmo de identificação dos padrões dos sinais acústicos (*PHAT Algorithm*), o tempo de duração está condicionado ao envio do email ao servidor de segurança. De forma a poder se observar o pico de correlação cruzada calculado a cada frame, colocou-se um tempo de espera de 400 ms, tempo próximo ao da aquisição do loop determinístico.

Nas Figuras 9 e 10, pode-se observar mais detalhadamente os loops determinístico e não determinístico, respectivamente.

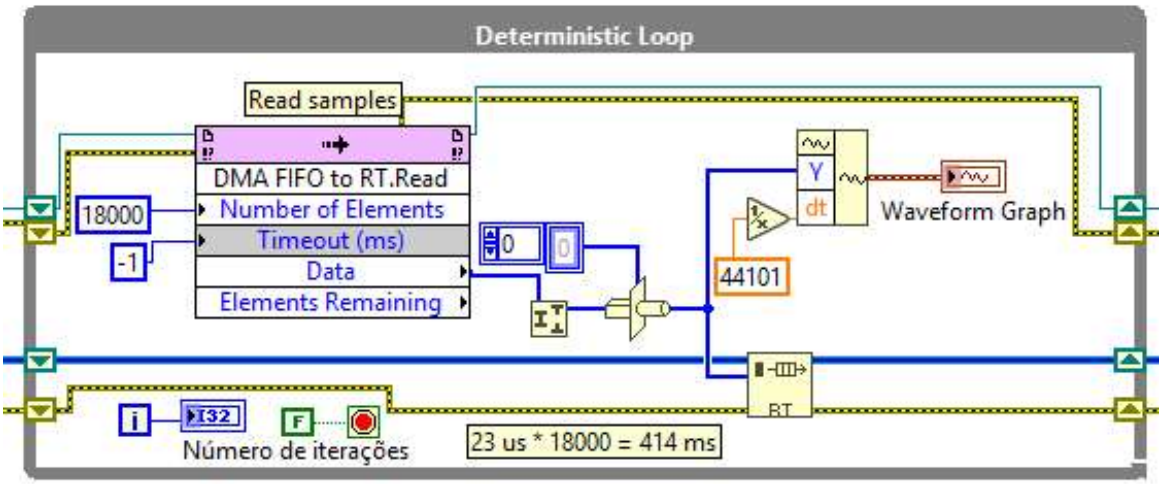

Figura 9: Loop determinístico, responsável pela aquisição dos áudios e enfileiramento dos mesmos.

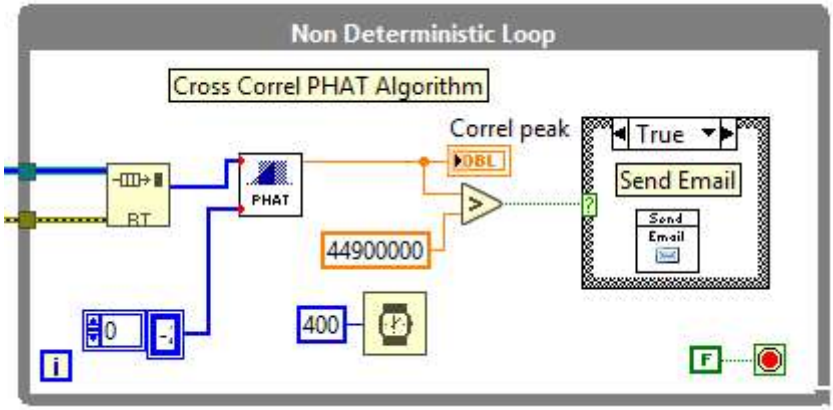

Figura 10: Loop não determinístico, responsável pelo processamento dos áudios e envio do alerta.

3.3. Rotina de enviar alerta (email)

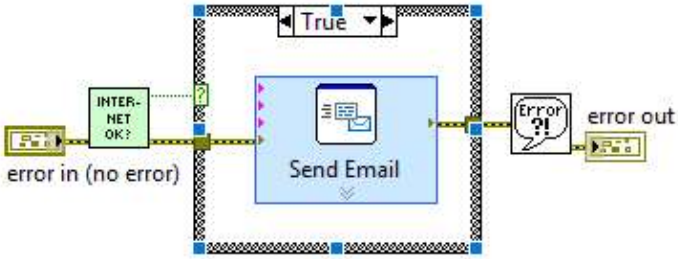

Figura 11: Rotina de envio de email utilizada para comunicação com servidor de segurança pública.

A rotina de envio de email – que pode ser vista na Figura 11 –, consistiu em um teste de acesso à internet, no qual o sistema tentava acessar o domínio "www.msftncsi.com"

via conexão TCP na porta 80, e então enviava a mensagem conforme as configurações mostradas na Figura 12.

O email chegava assim que o código era executado, garantindo que o servidor de segurança pública receba a mensagem em tempo real.

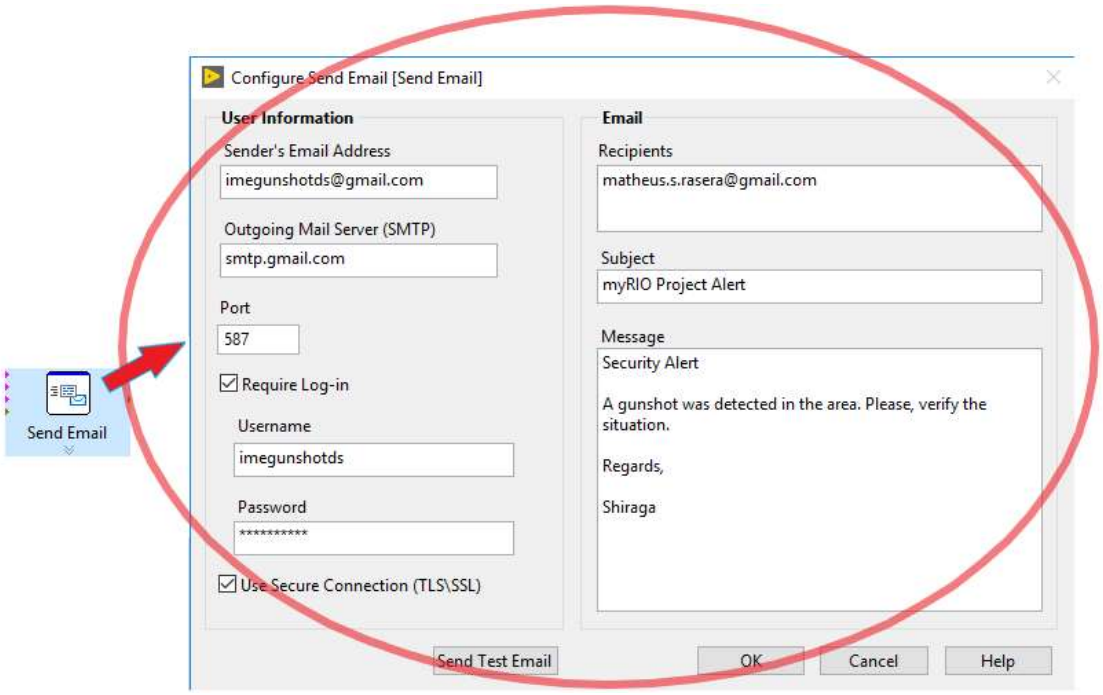

Figura 12: Configurações do email de alerta de segurança a ser enviado no caso de detecção de sinal compatível.

3.4. Rotina de processamento e identificação – Algoritmo GCC-PHAT

Para se detectar o grau de proximidade dos frames analisados com o sinal de áudio de referência, utilizou-se o algoritmo GCC-PHAT (*Generalized Cross Correlation Phase Transform*). Para tal, foram utilizadas a entrada do sinal adquirido e do sinal de referência, aplicado FFTs (*Fast Fourier Transforms*) com o dobro de elementos de cada sinal – de modo a garantir a eficiência do método de reconhecimento [2] –, e, em seguida, multiplicando elemento a elemento a saída de uma com o conjugado da saída de outra. Finalmente, aplicando-se a transformada inversa de Fourier (IFFT) e obtendo-se o valor máximo do módulo destas saídas – o qual se denominou de pico de correlação cruzada (*"Cross Correl Peak"*), foi possível verificar o grau de proximidade do sinal analisado com o nosso sinal de referência.

O código pode ser observado na Figura 13 abaixo.

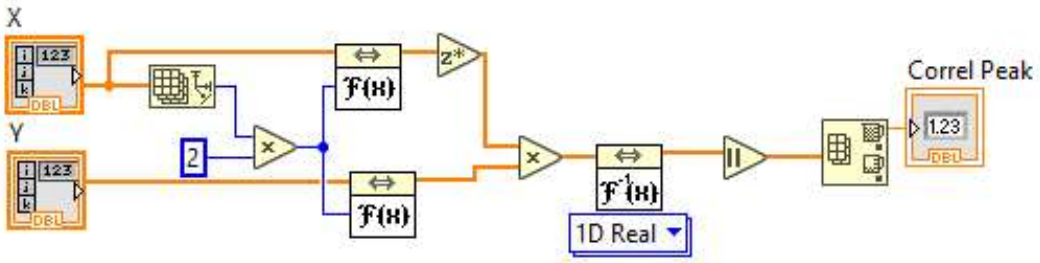

Figura 13: Algoritmo GCC-PHAT utilizado para análise do sinal. Na entrada *"X"*, acabamos utilizando o sinal de entrada; e na entrada *"Y"*, o sinal de referência.

4. Resultados

De forma a se determinar um *threshold* aceitável para a aplicação, foram mapeados os picos de correlação cruzada de 100 sinais de áudio referentes ao "toque" no microfone, que apresentam grande semelhança com o sinal de referência; de 50 sinais de áudio de conversas e músicas; e de outros 100 sinais de áudio referentes a uma palma.

Com os picos de correlação cruzada mapeados, plotou-se os mesmos em um histograma, onde as barras azuis, com maiores picos de correlação cruzada, são dos sinais de "toque"; e as barras vermelhas são de picos um pouco menores, que são dos sinais de conversas e músicas, e das palmas, conforme as Figuras 14 e 15, respectivamente.

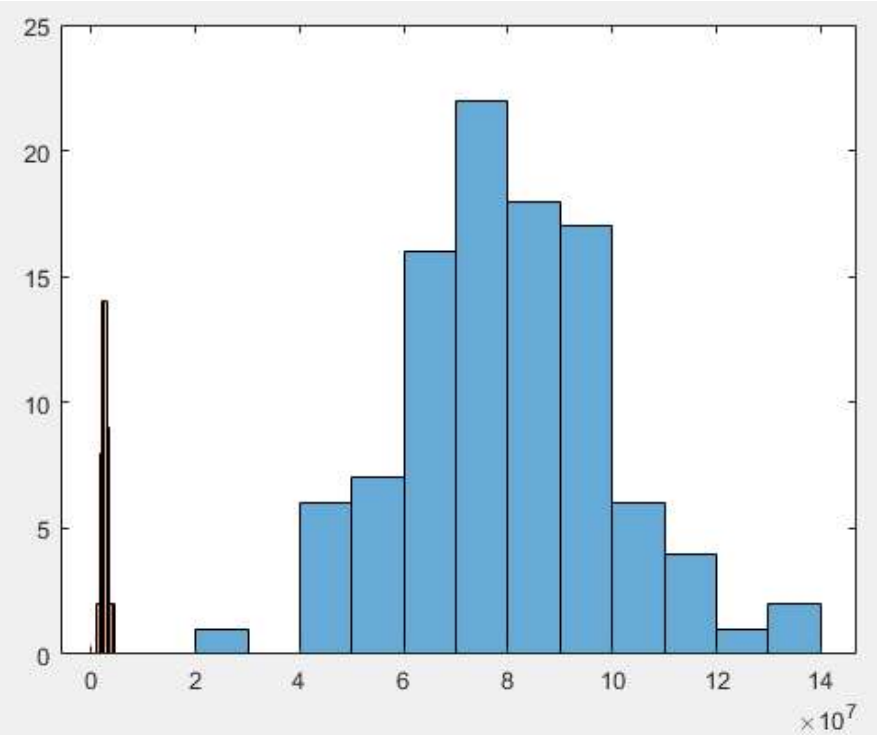

Figura 14: Histograma dos picos de correlação cruzada dos sinais das conversas e músicas (vermelho) com o dos toques (azul).

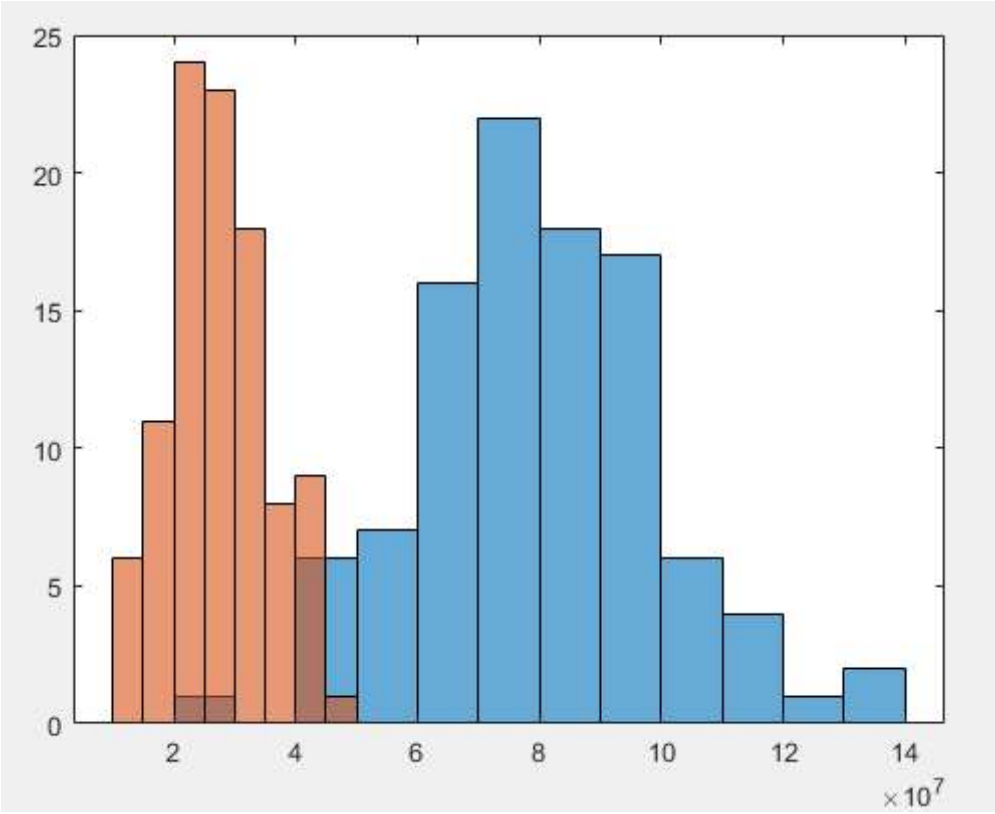

Figura 15: Histograma dos picos de correlação cruzada dos sinais das palmas (vermelho) com o dos toques (azul).

Desta forma, escolheu-se, a partir do histograma da Figura 15, um valor de *threshold* de pico de correlação cruzada de $4{,}5\ x\ 10^7$.

A partir desse valor, foram realizados alguns testes de reconhecimento do sinal. Na Tabela 1 abaixo, pode se verificar os resultados obtidos:

Tabela 1: Taxa de acerto na detecção dos "toques" no microfone.

	"Toque"	Conversa	Palma
Acertos	44	50	43
Total	50	50	50
Taxa de identificação	88%	100%	86%

Nos testes, 44 dos 50 "toques" foram identificados, representando uma taxa de identificação de 88%, com 6 falso-negativos; para as conversas, nenhuma foi identificada como "toque", gerando uma taxa de acerto de 100%; e, finalmente, as palmas

apresentaram taxa de identificação de 86%, sendo reconhecidas como “toques” (7 falso-positivos).

Para o teste da palma, foi identificado que um certo padrão de palma se aproximava do sinal de referência do “toque”, resultando em picos de correlação elevados o suficiente para acusar falso-positivos. Deste modo, para garantir que o teste não fosse comprometido, foram escolhidos 5 voluntários que não conheciam o funcionamento do sistema para realizar as palmas, ou seja, fornecendo palmas com padrões aleatórios.

Já a rotina de email permite ao sistema conseguir continuamente capturar e processar os áudios adquiridos pelo microfone em tempo real, e enviar o alerta caso seja identificado o padrão procurado.

5. Conclusão

Tendo em vista os histogramas mostrados anteriormente nas Figuras 14 e 15, foi possível verificar que seria muito improvável o projeto confundir uma conversa ou música com o toque no microfone; enquanto seria mais plausível que o mesmo confundisse a entrada com um outro sinal impulsivo como uma palma.

Note que há dois *outliers*, com o número de amostras dado, para os sinais dos “toques”. Isso pode ser devido tanto à irregularidade dada no momento de realizar os toques, quanto do fato de ser uma análise estatística.

O sistema está funcional, sendo capaz de cumprir com seu objetivo inicial de enviar um alerta em tempo real a um servidor de segurança pública, com taxa de acerto consideravelmente boa.

Os erros de identificação dos toques no microfone podem ser justificados no escopo desse projeto, uma vez que um sinal de tiro de arma de fogo possui um padrão muito mais característico, com frequências e componentes já conhecidas e mapeadas. O mesmo algoritmo de detecção poderia, se adaptado à identificação real de disparos de armas de fogo, alcançar taxas de acerto ainda superiores aos dos testes aqui realizados. Além disso, o local em que o sistema está posicionado também afeta os valores da correlação, sendo necessário realizar um novo estudo dos valores de *threshold* para cada posição a se colocar o mesmo, a fim de se obter a máxima eficiência.

Como possíveis melhorias a serem implementadas, destacam-se uma ampla gama de testes de algoritmos de detecção de sinais impulsivos adicionais que podem ser implementados (por exemplo, o teste por intensidade sonora que chega ao microfone),

aprimorando sua eficiência; uma maior coleta de dados para arbitrar um valor de *threshold* que atenda melhor ao projeto, mapeando os picos de correlação no local; um programa de identificação de alvo por triangularização, baseado no horário em que cada *myRIO* capturou o áudio, supondo um cenário de uma cidade inteligente com vários desses sistemas instalados nas ruas; e o desenvolvimento de uma base de dados contendo diversos disparos de armas de fogo, com possível utilização de inteligência artificial para ajudar no reconhecimento.

6. Referências Bibliográficas

1 Izabela L. Freire e José A. Apolinário Jr., "GCC-based DoA Estimation of Overlapping Muzzlebast and Shockwave Components of Gunshot Signals," IEEE Second Latin American Symposium on Circuits and Systems (LASCAS), 2011.

2 R. T. Ahmed, M. Uppal e A. Muhammad, "Improving Efficiency and Reliability of Gunshot Detection Systems," IEEE International Conference on Audio Speech, and Signal Processing (ICASSP), 2013.

3 Juan R. Aguilar, "Gunshot Detection Systems in Civilian Law Enforcement," Journal of the Audio Engineering Society, 2015.

4 S. D. Beck, H. Nakasone e K.W. Marr, "Variations in recorded acoustic gunshot waveforms generated by small firearms," *The Journal of the Acoustical Society of America*, vol. 129, no. 4, pp. 1748–1759, 2011.

5 Paulo C. Prandel e José A. Apolinário Jr., "Estimação da Direção de Disparo de Arma de Fogo: Uma Implementação em Tempo Real", Revista de Tecnologia da Informação e Comunicação, vol. 3, no. 2, 2014.

6 T. Embleton, "Tutorial on sound propagation outdoors," *The Journal of the Acoustical Society of America*, vol. 100, no. 1, pp. 31-48, 1996. [Online]. Available: http://link.aip.org/link/?JAS/100/31/1.

7 T. Mäkinen e P. Pertilä, "Shooter localization and bullet trajectory, caliber, and speed estimation based on detected firing sounds," *Applied Acoustic*, vol. 10, no. 71, pp. 902-913, 2010.

8 A. Chacon e P. Julian, "Evaluation of gunshot detection algorithms," *IEEE Transactions on Circuits and Systems-I: regular papers*, vol. 58, no. 2, 2011.

9 A. Dufaux, "Detection and recognition of impulsive sound signals," PhD thesis, Institute of Microtechnology. University of Neuchâtel, 2001.

10 Izabela L. Freire e José A. Apolinário Jr., "Localização de atirador por arranjo de microfones," *X Simpósio Brasileiro de Automação Inteligente*, 2011.

11 C. Hill, "Deterministic and non-deterministic loops," *VCSU-MEP on Euler*. Disponível em: http://euler.vcsu.edu:7000/13509/. Acesso em: 09 de novembro de 2018.

12 B. Lockhart, "In Biometrics, Which Error Rate Matters?," *Tractica*. Disponível em: https://www.tractica.com/biometrics/in-biometrics-which-error-rate-matters/. Acesso em: 13 de novembro de 2018.

13 Ed Doering, "NI myRIO Project Essentials Guide," *National Technology and Science Press*, 2013. Disponível em https://download.ni.com/evaluation/academic/myRIO_project_essentials_guide__Feb_09_2016___optimized.pdf. Acesso em: 15 de agosto de 2019.