

**Supressão de ruído acústico no sinal de áudio da fala durante
exames de ressonância magnética**

Vinícius Pessoa Duarte
Nº USP: 8941043

Departamento de Ciência da Computação
Instituto de Matemática e Estatística
Universidade de São Paulo

Proposta apresentada à disciplina
MAC0499 - Projeto de conclusão de curso supervisionado

Curso: Ciência da Computação
Supervisor: Prof. Dr. Marcel Jackowski

São Paulo
Abril, 2018

Sumário

1	Introdução	2
2	Objetivos	3
2.1	Estudo e implementação de métodos para processamento de áudio	3
2.1.1	Subtração espectral	3
2.1.2	Método dos mínimos quadrados normalizado (NLMS)	3
2.1.3	<i>Dictionary learning</i>	3
2.2	Testes comparativos	4
2.2.1	Teste de complexidade	4
2.2.2	Teste de qualidade	4
2.3	Acompanhamento de sessões de RM	4
2.4	Elaboração de software livre	5
2.5	Cronograma	6

1 Introdução

A captura do áudio da fala juntamente à imagens de ressonância magnética (RM) que retratam o formato do trato vocal pode facilitar a caracterização diferencial de desordens de fala.

O sinal de áudio dentro da sala de RM é normalmente poluído por diversos ruídos de interferência, como aqueles gerados pela movimentação do paciente, a interferência dos campos magnéticos gerados máquina, ou o ruído acústico gerado pelas bobinas que produzem os campos magnéticos supracitados (Ravicz, Melcher e Kiang 2000). Esses ruídos podem dificultar a compreensão e análise da fala original, impondo a necessidade de supressão de ruído. Dado que tais ruídos possuem características periódicas, em frequências específicas, esta supressão pode ser concretizada por algoritmos que visam identificar essas características e, a partir dos padrões encontrados, extrair a fala "limpa" do paciente, i.e., a fala sem ruído proveniente máquina de RM.

Ainda, dada a impossibilidade do uso de objetos que contenham partes metálicas (NessAiver et al. 2006), se torna necessário o uso de microfones especiais de fibra ótica para capturar a fala do paciente. A baixa qualidade de captura desses microfones, quando comparados aos microfones padrão, também demandam o melhoramento dos dados capturados.

2 Objetivos

A seguir, descritas as tarefas que esperamos realizar durante este projeto.

2.1 Estudo e implementação de métodos para processamento de áudio

Para realizar a supressão de ruído acústico e melhoramento do áudio captado, três métodos serão estudados e implementados, sendo eles: subtração espectral; métodos dos mínimos quadrados normalizado e supressão de ruído baseada em *dictionary learning*.

2.1.1 Subtração espectral

Para realização satisfatória de subtração espectral, é necessária a suposição de que o ruído na fala gravada é aditivo, i.e., no domínio das frequências, a fala e o ruído podem ser separados em componentes distintas que, quando somadas, resultam no espectrograma do áudio com ruído original (Ravicz, Melcher e Kiang 2000).

Partindo de tal suposição, o algoritmo de subtração espectral consiste em, a partir de uma estimativa do ruído presente no áudio, realizar, no domínio das frequências, a subtração entre o espectrograma do áudio original e do ruído estimado. Assim, a versão melhorada do áudio é obtida após realizar a transformação do espectrograma resultante para o domínio de sinais (Pribil, Pribilova e Frollo 2016).

Esse método, ao supor aditividade do ruído, pode gerar artefatos no áudio resultante, quando da utilização de dados de fala obtidos em situações de sessões de RM, pois o ruído gerado pelas bobinas que criam os campos magnéticos possuem frequências que conflitam com o espectro de fala.

2.1.2 Método dos mínimos quadrados normalizado (NLMS)

Este método consiste em, a partir do áudio captado durante a sessão de RM e uma captação do ruído gerado durante esse tipo de procedimento, realizar a supressão do ruído acústico com a utilização um filtro adaptativo (Chinaboina et al. 2011).

Tal filtro é treinado com base em pesos de ativação, os quais são atualizados iterativamente de acordo com o erro estimado entre o ruído de referência recebido como entrada do algoritmo e a estimativa do áudio suprimido calculado na etapa anterior, uma função de passo, a qual visa regular o tempo para convergência do filtro, uma amostra do áudio corrompido original o próprio valor dos pesos durante a iteração anterior (Bresch et al. 2006).

2.1.3 *Dictionary learning*

Esse método de supressão de ruído, baseado no algoritmo apresentado em (Vaz, Ramanarayanan e Narayanan 2013), consiste em duas fases: a aplicação do algoritmo *probabilistic latent component analysis* (PLCA) e realização de análise *wavelet*.

Na primeira etapa, o algoritmo é inicialmente treinado com o ruído a ser suprimido da fala gravada, gerando um dicionário de ruídos, juntamente de seus pesos de ativação. Esse ruído pode ser obtido tanto pela captação de áudio externa à máquina de RM, quanto utilizando os primeiros segundos de áudio do arquivo a ser processado.

Após esse passo, o algoritmo recebe o espectrograma do áudio a ser processado, calculando para cada quadro dessa entrada, uma medida de divergência em relação ao ruído estimado no passo anterior, i.e, um cálculo de semelhança entre as distribuições das frequências de ambas as entradas, atualizando então, um dicionário de fala, seus pesos de ativação, bem como os pesos de ativação do dicionário de ruídos, dependendo do valor calculado.

Segundo os autores do artigo no qual o método é baseado, essa primeira etapa, além de ser eficiente em remover ruído em momentos sem fala, consegue diminuir a potência do ruído no arquivo de áudio, viabilizando o uso de análise *wavelet* na segunda etapa.

Tal etapa consiste em, a partir da análise *wavelet*, limiarizar os coeficientes das componentes geradas, a partir de alguma métrica de divergência. O áudio com ruído suprimido é obtido, então, com a reconstrução do sinal a partir das componentes limiarizadas.

2.2 Testes comparativos

Com os métodos implementados, serão realizados teste de complexidade e qualidade dos resultados.

2.2.1 Teste de complexidade

Com esses testes, é esperado poder determinar qual método obtém melhor performance diante de arquivos de áudio obtidos nas sessões de RM. Para obter subsídios para tal análise, serão realizados 20 testes, para cada método, sendo extraídas métricas da velocidade de execução dos algoritmos, como: média; desvio-padrão; variância; quartis.

2.2.2 Teste de qualidade

Para supressão de ruído, o teste de qualidade vai se dar pela comparação do erro entre um arquivo de áudio, constituído pela adição de ruído em um arquivo "limpo", processado pelas técnicas e o arquivo original, sem adição de ruído.

Com essa comparação, é possível medir a qualidade de supressão de ruído para cada método implementado.

2.3 Acompanhamento de sessões de RM

Para entender melhor a forma da coleta do áudio a ser processado pelos algoritmos implementados, serão realizados acompanhamentos em sessões de RM. Dado que tais acompanhamentos dependem do

calendário de consultadas atreladas a pesquisa do Prof. Dr. Marcel P. Jackowski, estes não estão listados no cronograma de atividades.

2.4 Elaboração de software livre

A partir dos estudos, implementações e testes comparativos, deseja-se elaborar um software, em formato livre, capaz de realizar a supressão de ruído acústico e melhoria de áudio em dados na forma daqueles coletados durante os acompanhamentos de sessões de RM.

O software será produzido em python, por ser uma linguagem facilmente adaptável para diferentes sistemas operacionais, além de possuir uma ampla gama de bibliotecas para computação científica, como numPy, scikit-learn e sciPy.

A ferramenta elaborada receberá um arquivo de áudio a ser processado e, opcionalmente, um segundo arquivo de áudio, contendo o ruído obtido a partir de um segundo microfone. O produto desse processamento será um arquivo de áudio referente ao processamento do arquivo inserido.

O usuário, ainda, poderá escolher dentre os métodos estudados durante esse projeto para realizar o processamento dos dados inseridos. Para auxiliar a escolha do método a ser utilizado, os dados obtidos nos testes comparativos serão utilizados para informar ao usuário qual opção é mais rápida e qual gera melhor resultado.

Ainda, será permitido que o software seja usado tanto em linha de comando, quanto em interface gráfica. Assim, o primeiro modo de utilização facilita a incorporação da ferramenta em *scripts*, enquanto o segundo modo permite um uso mais intuitivo da ferramenta.

Com a elaboração dessa ferramenta, é esperado obter um barateamento da integração procedimento com a análise de áudio, o que permite uma maior adoção dessas análises, por parte de médicos e pesquisadores.

2.5 Cronograma

Com base nas tarefas descritas no capítulo anterior, é possível fragmentar as tarefas nas seguintes etapas:

1. Revisão bibliográfica
2. Elaboração de resumo
3. Elaboração da proposta
4. Familiarização com análise de Fourier e filtros
5. Implementação do algoritmo de subtração espectral
6. Familiarização com o método LMS
7. Implementação do algoritmo LMS
8. Familiarização com o algoritmo PLCA e *dictionary learning*
9. Familiarização com análise *wavelet* e filtros
10. Implementação do algoritmo baseado em *dictionary learning*
11. Implementação do algoritmo de melhoria de áudio
12. Realização dos testes
13. Montagem do software
14. Elaboração de monografia
15. Elaboração de pôster

Tabela 1: Cronograma

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
mar	x	x		x											
abr	x		x	x											
mai					x	x	x								
jun							x	x	x						
jul										x					
ago										x	x				
set												x	x		
out													x	x	
nov														x	x

Referências

- [1] Erik Bresch et al. “Synchronized and noise-robust audio recordings during realtime magnetic resonance imaging scans”. Em: *The Journal of the Acoustical Society of America* 120.4 (2006), pp. 1791–1794.
- [2] Radhika Chinaboina et al. “Adaptive algorithms for acoustic echo cancellation in speech processing”. Em: *International Journal of Research and Reviews in Applied Sciences* 7.1 (2011), pp. 38–42.
- [3] Moriel S NessAiver et al. “Recording high quality speech during tagged cine-MRI studies using a fiber optic microphone”. Em: *Journal of Magnetic Resonance Imaging* 23.1 (2006), pp. 92–97.
- [4] Jiri Pribil, Anna Pribilova e Ivan Frollo. “Analysis of acoustic noise and its suppression in speech recorded during scanning in the open-air MRI”. Em: *Advances in Noise Analysis, Mitigation and Control*. InTech, 2016.
- [5] Michael E Ravicz, Jennifer R Melcher e Nelson Y-S Kiang. “Acoustic noise during functional magnetic resonance imaging”. Em: *The Journal of the Acoustical Society of America* 108.4 (2000), pp. 1683–1696.
- [6] Colin Vaz, Vikram Ramanarayanan e Shrikanth Narayanan. “A two-step technique for MRI audio enhancement using dictionary learning and wavelet packet analysis.” Em: *INTERSPEECH*. 2013, pp. 1312–1315.