

# 12º SEMINÁRIO DE INICIAÇÃO CIENTÍFICA E EXTENSÃO DA UEMG

## Processamento das Características da Voz no domínio da Frequência Baseado em Método com Resolução Reduzida

Alita de Oliveira Duarte

Sylvio Barbon Junior

Email para contato: [alita.o.duarte@gmail.com](mailto:alita.o.duarte@gmail.com)

Palavras chave: Processamento Digital de Sinais, Transformada Discreta de Fourier, vogal, Pitch

A voz tem sido objeto de estudo em diversas áreas, o processamento deste tipo de sinal unidirecional tem evoluído junto com a computação. Algumas técnicas são propostas com o intuito de incrementar o desempenho das análises deste sinal.

Este trabalho descreve a proposta de uma abordagem para melhoria do processamento de sinais, particularmente voz. Diversas aplicações voltadas à manipulação do som têm como requisito o processamento de um sinal discretizado, ou seja, de análise realizável por computadores. Este processo mesmo considerando os níveis perceptuais humanos, que contribuem para uma diminuição da somatória de elementos, é abordado pelas técnicas de quantização e amostragem.

Visando melhorias, não a redução da complexidade do algoritmo, mas sim na diminuição da quantidade de dados a serem processados, porém mantendo a mesma informação, foram feitos testes com sinais no domínio da frequência utilizando a frequência fundamental (Pitch) como parâmetro de consistência da informação.

O algoritmo para o experimento foi desenvolvido em linguagem Java onde o cálculo do Pitch, uma importante característica na área de pesquisa da fala, neste caso foi utilizado para verificar se o sinal manteve-se semelhante. Para o cálculo do Pitch, utilizou-se a Transformada de Fourier Discreta ferramenta matemática usada para converter sinal no domínio do tempo para o domínio da frequência.

Os testes foram realizados com resoluções 512, 1024, 4096, 8192, 16384 e 32768 elementos, comparadas com o resultado do sinal completo (máxima resolução), usando as vogais /a/, /e/ e /i/, amostradas durante 3 segundos. Estas vogais foram utilizadas, pois representam diferentes características da voz, onde a vogal /a/ tem uma maior uniformidade no sinal seguidas por /i/ e /e/.

Para a avaliação da variabilidade do Pitch foram realizados diversos testes, sendo que a diferença entre a nova resolução, inferior ao sinal completo, e a máxima foi obtida através da média das diferenças, em modulo. Observando os testes todas as resoluções analisadas apresentaram aproximações do valor de Pitch em relação ao sinal original. Considerando a média das diferenças foi observado que as resoluções em potência de 2 maiores que 4096 elementos apresentaram os melhores resultados.

Ao final dos testes foi possível observar a pequena variação no Pitch, que possibilitam a interpretação de que existe alto um grau de compatibilidade e possibilidade de abstração das características pelo sinal de menor quantidade de pontos. O que reduz drasticamente o tempo envolvido no processamento, pois são necessárias menos amostras.