

# Análise comparativa entre as Transformada Discreta de Fourier e a Transformada Discreta Cosseno na compressão e recuperação espectral de sinais de voz.

Priscila Lima Rocha

Washington Luís Santos Silva

Ginalber Luiz de Oliveira Serra

**Resumo**—Neste artigo é proposta uma análise comparativa prática entre a Transformada Discreta Cosseno (TDC) e a Transformada Discreta de Fourier (TDF) em aplicações de processamento digital de sinais de voz. A análise pretende mostrar qual das duas Transformadas é mais eficiente, no que diz respeito à utilização do número mínimo de amostras espectrais para recompor um dado sinal. Os resultados experimentais são apresentados em forma de gráficos. Utilizou-se para a simulação e obtenção dos resultados, exemplos de sinais de voz de palavras e números com adição de ruído branco.

**Palavras-Chave**—Processamento digital de sinais, Composição de sinais, Transformada discreta.

**Abstract**— In this paper is proposed a comparative practice analysis between the Discrete Fourier Transform (DFT) and the Discrete Cosine Transform (DCT) in application of digital processing of speech signals. The analysis is based in to check which of two Transform is more efficient, as regard use of fewer samples of the spectrum for reconstruct a given signal. The experiment results are presented in form of graphs. It was used for the simulation and obtain of the results, example of speech signal of spoken words and digits with added of white noise.

**Keywords**—Digital processing of signals, Composition of signals, Discrete transform.

## I. INTRODUÇÃO

Em processamento digital de sinal em tempo real um esforço computacional significativo pode ser evitado se os sinais de entrada forem representados por suas coordenadas (não-correlacionadas) obtidas com a Transformada de Karhunen-Loève [1]. Por esta razão, a transformada de Karhunen-Loève poderia ser usada em aplicações como processamento de imagens, codificação de voz e reconhecimento de padrões [2]. Porém a carga computacional da transformada de Karhunen-Loève inviabiliza sua utilização. Para solucionar tal problema, utilizam-se outras transformadas que se equivalem assintoticamente a Transformada de Karhunen-Loève para uma classe muito grande de processos estacionários. Entre as transformadas mais utilizadas estão a Transformada Discreta de Fourier e a Transformada Discreta Cosseno [3]. Neste artigo é proposta uma análise comparativa entre as Transformada Discreta Cosseno (TDC) e a Transformada Discreta de Fourier (TDF) em aplicações de processamento digital de voz para a telefonia digital móvel, para TV digital e voz sobre IP. A análise baseia-se em verificar

qual das duas Transformadas é mais eficiente na utilização de um número mínimo de amostras espectrais para recompor um dado sinal de voz. Esta eficiência é verificada através do erro quadrático médio dos sinais recuperados e a relação entre a potência do sinal original e a potência do sinal comprimido utilizado para a reconstituição.

## II. METODOLOGIA

Para a análise de desempenho das transformadas, desenvolveu-se um algoritmo para medir o erro quadrático médio dos sinais recuperados e a relação entre a potência do sinal original e a potência do sinal comprimido utilizado para a reconstituição. Utilizaram-se exemplos de palavras e números como sinais de voz de entrada do algoritmo. A cada sinal de voz foi adicionado um ruído branco e a relação sinal/ruído foi variada em 10dB, 30dB, 40dB e 60dB. Os valores para a relação sinal/ruído foram escolhidos para observar a robustez ao ruído, do trabalho proposto, na recuperação dos sinais de voz. Para cada variação da relação sinal/ruído, o número total de amostras foi diminuído de 0 a 90% do sinal original.

## III. RESULTADOS OBTIDOS

Com o algoritmo para a análise de desempenho da Transformada Discreta de Fourier e da Transformada Discreta Cosseno testou-se um conjunto de vinte e três sinais de voz com as seguintes pronúncias: “Casa”, “Aberto”, “Fechado”, “Frente”, “Costa”, “Liga”, “Desliga”, “Sim”, “Não”, “Direita”, “Esquerda”, “Volta”, “Zero”, “Um”, “Dois”, “Três”, “Quatro”, “Cinco”, “Seis”, “Sete”, “Oito”, “Nove”, “Dez”. Será apresentado o resultado para a palavra “Aberto”. As conclusões apresentadas neste artigo referem-se aos testes com todos os exemplos. Pôde-se observar durante os testes que a variação do ruído de um nível mais crítico (10dB) para um nível razoável (60dB) não altera significativamente a recomposição do sinal, a potência e erro quadrático médio. Isto demonstra a robustez do sistema ao ruído, já que as altas frequências inseridas pelo ruído nos sinais a serem recompostos não interferem no processamento das transformadas.

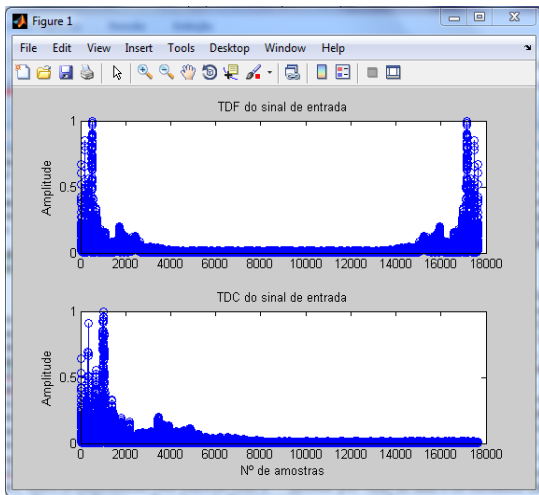


Fig. 1. Espectro de Frequência da TDC e TDF para a palavra “Aberto”

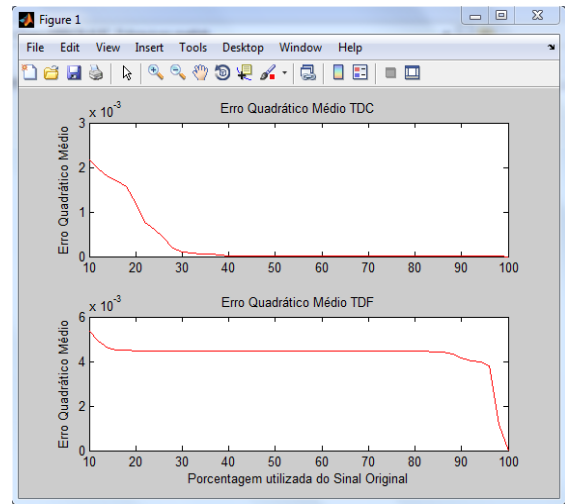


Fig. 4. Erro quadrático médio para a TDC e TDF

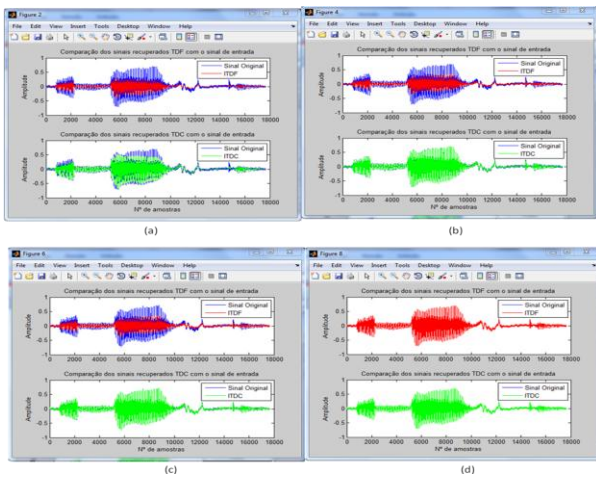


Fig. 2. Comparação do sinal original com o sinal reconstruído: (a) 10% das amostras utilizadas; (b) 40% das amostras utilizadas; (c) 70% das amostras utilizadas; (d) 100% das amostras utilizadas.

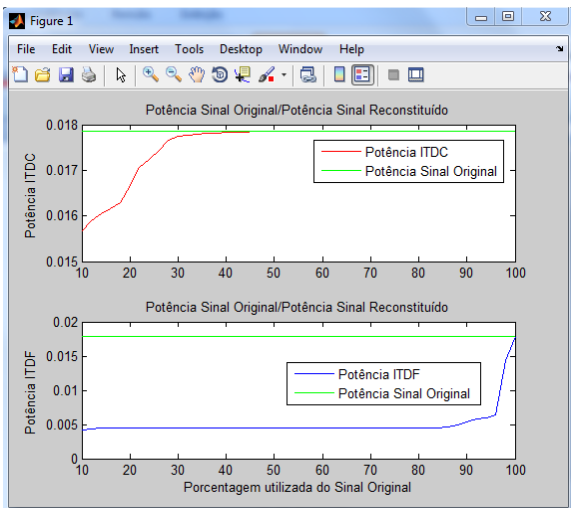


Fig. 3. Relação entre potência do sinal original e a potência utilizada para a reconstrução pela ITDC e ITDF.

#### IV. CONCLUSÕES

Após os ensaios computacionais, comprovou-se, um desempenho melhor da TDC em relação à TDF para toda faixa de coeficientes do espectro. Observou-se na análise, que o espectro da TDC apresenta as frequências de maiores amplitudes do sinal bem concentradas, facilitando a retirada de amostras não significativas para a recomposição do sinal. Isto já não acontece para a TDF. Esta característica do espectro influencia nos resultados da potência e do erro quadrático médio, onde se notou que com poucas amostras utilizadas do sinal pela TDC, estas continham grande parte da potência do sinal original. Para a TDF, entretanto, foi necessário à utilização de praticamente todas as amostras para que o sinal reconstruído atingisse a potência do sinal original. Para todos os testes, observou-se que o erro quadrático médio da TDC é bem menor que o da TDF e que este diminui significativamente para um pequeno número de amostras. Ao contrário, a TDF apresenta um erro elevado que se mantém praticamente constante para quase toda a análise, diminuindo somente quando se utiliza quase todas as amostras do sinal. Assim, observa-se que para processamento de sinais em tempo real, onde o tempo de processamento é fundamental e inversamente proporcional ao número de amostras, a Transformada Discreta Cosseno é mais aconselhável, pois tem resultados altamente positivos com um número de amostras reduzido, no sentido que aproxima melhor a TKL (transformada de Kahunem-Loève) que a Transformada Discreta de Fourier.

#### AGRADECIMENTOS

Ao Instituto Federal do Maranhão por fornecer infraestrutura e materiais necessários à pesquisa. Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq por fomentar este projeto. Aos professores colaboradores que orientam o desenvolvimento deste trabalho.

#### REFERÊNCIAS

- [1] G. Rajesh, A. Kumar; K. Ranjeet, “Speech compression using different transform techniques”. ICCCT. pp. 146-151, Sept. 2011.
- [2] W. A. Jassim, R. Paramesran, “Speech signal representation by Discrete Transform” TECHPOS. pp. 1-5, Dec. 2009.
- [3] J. Pearl, “On coding and filtering stationary signals by Discrete Fourier Transform”, IEEE Trans. Informl. Theory., vol IT-19, pp. 229-232, 1973