



## A NEW PARADIGM FOR PLOTTING SPECTROGRAM

ROHINI R. MERGU<sup>1</sup> AND SHANTANU K. DIXIT<sup>2</sup>

Walchand Institute of Technology

\*Corresponding Author: Email- <sup>1</sup>rohinimergu@gmail.com and <sup>2</sup>dixitsk1@yahoo.com

Received: January 12, 2012; Accepted: February 15, 2012

**Abstract-** An important aid in analysis & display of sound is the spectrogram. It represents time-frequency-intensity display of short time spectrum. Spectrogram finds applications in different fields such as biomedical, speech processing, speech enhancement, detecting animal cries, detecting vehicle sounds etc. But to conclude from visual inspection the clarity of spectrogram is also important. Before plotting the spectrogram, the time domain sound signal is converted to frequency domain. The transform domain used plays vital role in clarity of spectrogram. Generally Fast Fourier Transform is used to convert the time domain signal into frequency domain signal before plotting spectrogram. This paper discusses the effect of using Discrete Cosine Transform for converting the speech signal into frequency domain before plotting spectrogram. It is observed that resolution of spectrogram is transform dependent.

**Keywords-** Spectrogram, ECG, DCT, DFT

**Citation:** Rohini R. Mergu and Shantanu K. Dixit (2012) A New Paradigm for Plotting Spectrogram. Journal of Information Systems and Communication, ISSN: 0976-8742 & E-ISSN: 0976-8750, Volume 3, Issue 1, pp-158-161.

**Copyright:** Copyright©2012 Rohini R. Mergu and Shantanu K. Dixit. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### Introduction

The application of spectrogram in the field of music, sonar/radar, speech processing, seismography etc. Another application of spectrogram is in detecting animal cries. Spectrogram finds applications in different fields such as Image processing for feature detection, in the biomedical field for finding abnormalities in the Electro Cardio Gram (ECG) signal, in speech processing to find phonemes and phoneme boundaries, in the field of speech recognition for identification of speaker, in the field of speech enhancement to see the content of noise in the signal and in field of communication etc.

Speech spectrogram reading involves interpreting the acoustic patterns in the image to determine the spoken utterance. One must selectively attend to many different acoustic cues, interpret their significance in light of other evidence, and make inferences based on information from multiple sources. The evidence, obtained from spectrogram reading experiments [1] indicates that the process can be modeled with rules. Formalizing spectrogram reading entails refining the language used to describe acoustic events in the spectrogram, selecting a set of relevant acoustic

events that distinguish among phonemes, and developing rules which map these acoustic attributes into phonemes. Phoneme segmentation by an expert system utilizing spectrogram reading strategy and knowledge [2] can detect phonemes in a spectrogram and determines their boundaries as well as their coarse categories. Reliable feature detection is a prerequisite to higher level decisions regarding image content. Using Spectrogram line detection, multi-scale, linear feature detection of an image with low signal to Noise ratio is possible [3]. Also can detect tracks of varying gradients. Spectrogram based ECG signal & the PSD together with offline evaluation has been observed[4]. The spectrogram found to be more precise over conventional FFT in finding small abnormalities in ECG signal. An iterative algorithm to estimate the instantaneous frequency (IF) and matched spectrogram of nonstationary signals is modeled [5]. The matched spectrogram obtained by this method is concentrated along the IF for mono-component signals. The proposed algorithm is then combined with a simple window adaptation scheme. Spectrograms are commonly used to analyze animal vocalizations. But [7] analyzed how far it is possible to deduce the mechanical origin of sound generation

and modulation from the spectrogram. Also investigated the relationship between simple mathematical events such as transients, harmonics, amplitude and frequency modulation and the resulting structures in spectrograms. Application of Spectrogram in automatic recognition of multi-speaker continuous speech is observed in [8]. Using spectrogram it is possible to see the noise content in the enhanced speech by comparing with clean speech spectrogram [6]. It is observed that a large portion of the spectrogram is practically blank (i.e., unshaded) and the speech energy is concentrated in a few isolated regions. The voiced portion of speech is characterized by dark parallel "stripes" whereas unvoiced portion is characterized by gray patches. Some parallel stripes are horizontal while some are slanting up or down, indicating a change in the pitch of the speech signal. When white Gaussian noise amounting to the clean speech, the blank region of the spectrogram become shaded, and some of the stripes corresponding to voiced speech disappear. The speech is enhanced using spectral subtraction and with the observation of spectrogram it is found that there is a significant reduction of the unwanted short stripes. Thus just from observation of spectrogram [6] one can conclude about speech quality.

**The spectrogram**

A spectrogram is a time-varying spectral representation that shows how the spectral density of a signal varies with time. In the field of time–frequency signal processing, it is one of the most popular quadratic Time-Frequency Distribution that represents a signal in a joint time-frequency domain. Also known as spectral waterfalls, sonograms, voiceprints, or voicegrams, spectrograms are used to identify phonetic sounds, to analyze the cries of animals; they were also used in many other fields including music, sonar/radar, speech processing, seismology, etc. The instrument that generates a spectrogram is called a spectrograph. The most common format is a graph with two geometric dimensions: the horizontal axis represents time, the vertical axis is frequency; a third dimension indicating the amplitude of a particular frequency at a particular time is represented by the intensity or colour of each point in the image.

Spectrograms are usually created in one of two ways: approximated as a filter bank that results from a series of band pass filters (this was the only way before the advent of modern digital signal processing), or calculated from the time signal using the short-time Fourier transform (STFT) . The spectrogram of a signal  $s(t)$  can be estimated by computing the squared magnitude of the STFT of the signal  $s(t)$ , as shown below:

$$\text{Spectrogram}(t, \omega) = |\text{STFT}(t, \omega)|^2$$

These two methods actually form two different quadratic Time-Frequency Distributions, but are equivalent under some conditions. Creating a spectrogram using the STFT is usually a digital process. Digitally sampled data, in the time domain, is broken up into chunks, which usually overlap, and Fourier transformed to calculate the magnitude of the frequency spectrum for each chunk. Each chunk then corresponds to a vertical line in the image; a measurement of magnitude versus frequency for a specific moment in time. The spectrums or time plots are then "laid side by side" to form the image or a three-dimensional surface.[5]

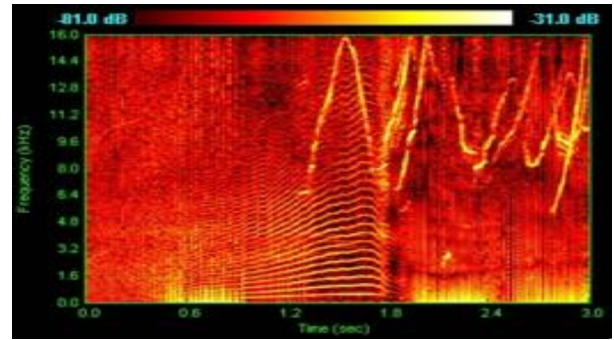


Fig. 1- Spectrogram of Dolphin Vocalization

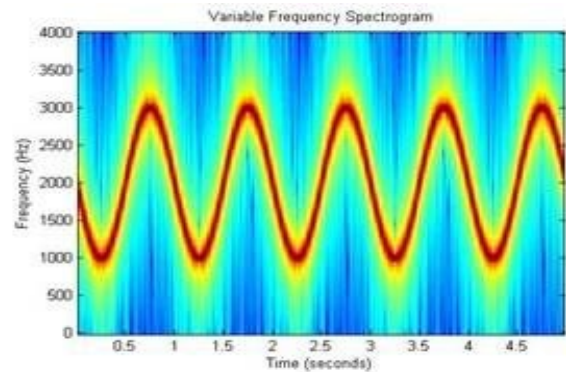


Fig. 2- Spectrogram of an FM signal

The use of color to highlights the important features of a spectrogram. In the spectrogram shown in Figure 1 the shades of red indicates increasing energy along the frequency axis, blue to mean decreasing energy, and yellow and green to mean an energy maximum. Areas which are white do not have enough energy to be of interest.

**Computation of Spectrogram**

In different fields as mentioned in the introduction the graphical representation spectrogram plays vital role. The conclusion of ECG abnormality can be done using spectrogram[4]. The conclusion of speech quality [6] and phoneme boundary detection [1] can be done from observation of spectrogram. Speaker recognition is also can be made [8] by visual inspection of spectrogram. But to conclude from visual inspection the clarity or the resolution of spectrogram is also important. Before plotting the spectrogram the time domain speech signal is converted to frequency domain. The domain used for transformation plays vital role in clarity of spectrogram. Generally Fast Fourier Transform is used to convert the time domain signal into frequency domain signal. This paper discusses the effect of using discrete cosine transform for converting the time signal into frequency domain before plotting spectrogram.

**A. Using Discrete Fourier Transform**

Discrete Fourier Transform (DFT) can be computed efficiently using a Fast Fourier Transform (FFT) algorithm. The discrete Fourier transform (DFT) is a specific kind of Fourier Transform, used in Fourier Analysis. It transforms the time domain function into frequency domain representation. FFT algorithms are so commonly employed to compute DFTs that the term FFT is often used to mean DFT in colloquial settings.

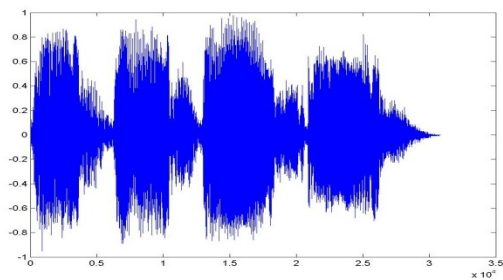


**B. Using Discrete Cosine Transform**

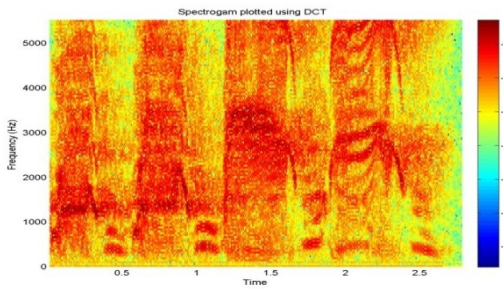
Discrete Cosine Transform (DCT) expresses a sequence of finitely many data points in terms of a sum of cosine functions oscillating at different frequencies. It turns out that cosine functions are much more efficient as fewer terms are needed to approximate a typical signal. In particular, a DCT is a Fourier related transform similar to the Discrete Fourier Transform (DFT), but using only real numbers. DCTs are equivalent to DFTs of roughly twice the length, operating on real data with even symmetry.

**Results**

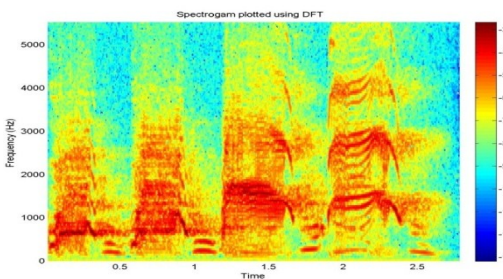
Here spectrograms are plotted for different sounds such as animal, transportation so, instrumentation and people sound etc. The spectrograms plotted using 256 point DFT & 256 point DCT are shown in figures below.



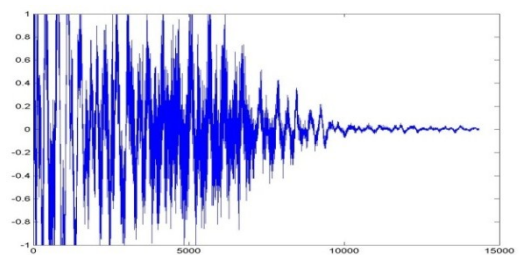
**Fig. 3-** Chimpanzee voice



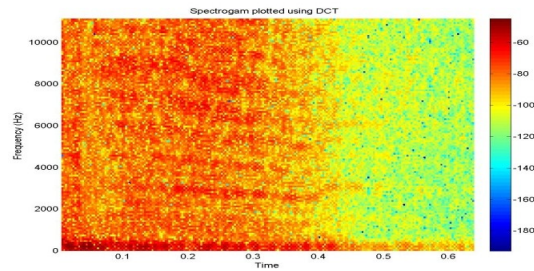
**Fig. 4-** Spectrogram of Chimpanzee voice plotted using DCT



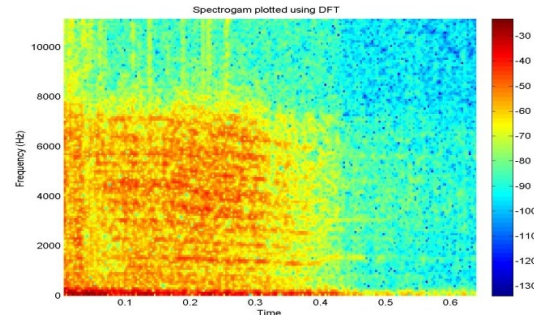
**Fig. 5-** Spectrogram of Chimpanzee voice plotted using DFT



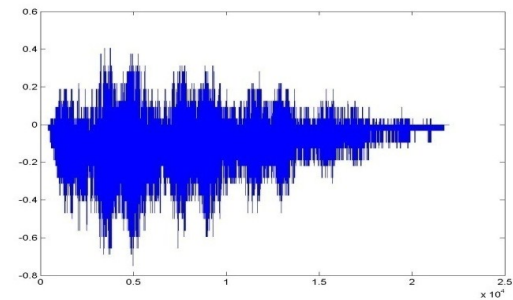
**Fig. 6-** Plot of musical Instrument sound (Synthesizer sound)



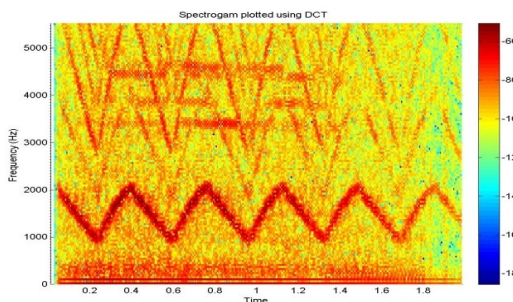
**Fig. 7-** Spectrogram for synthesizer sound plotted using DCT



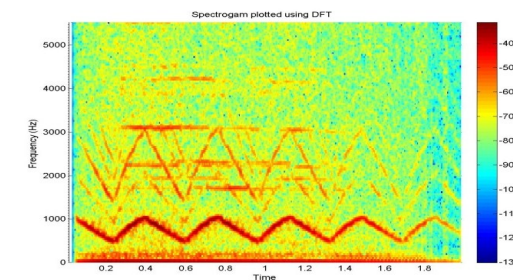
**Fig. 8-** Spectrogram for synthesizer sound plotted using DFT



**Fig. 9-** Plot of Transportation Vehicle Ambulance Siren sound



**Fig.10-** Spectrogram for Ambulance Sire sound plotted using DCT



**Fig.11-** Spectrogram for Ambulance Sire sound plotted using DFT

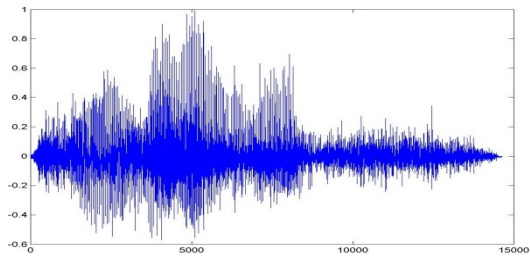


Fig. 12- Plot of Human sound (Blowing Nose)

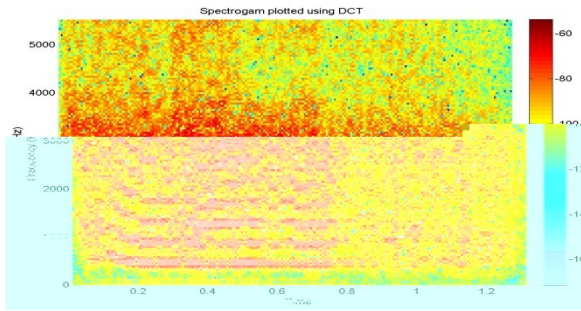


Fig. 13- Spectrogram for Blowing Nose Signal plotted using DCT

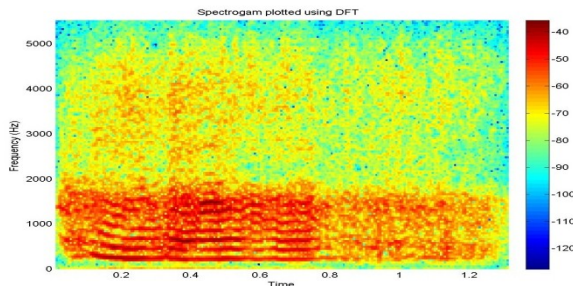


Fig. 14- Spectrogram for Blowing Nose Signal plotted using DFT

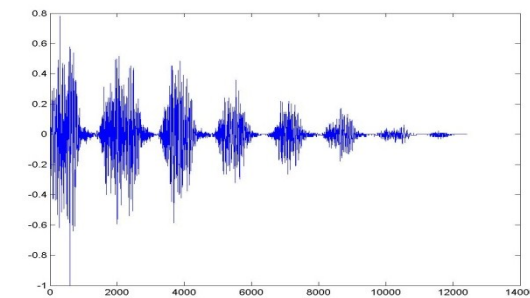


Fig. 15- Plot of Miscellaneous Sound (Axe Throw)

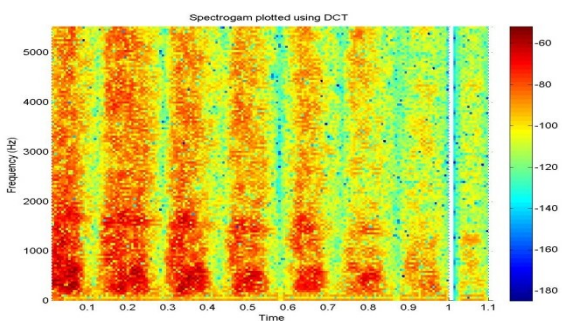


Fig. 16- Spectrogram for Axe Throw Signal plotted using DCT

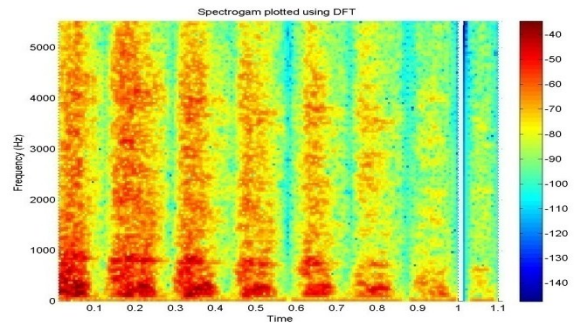


Fig. 17- Spectrogram for Axe Throw Signal plotted using DFT

**Conclusion**

From visual inspection of spectrogram we can conclude in different fields like, about abnormality in biomedical field, about noise content in speech processing field, about speaker in speech recognition field, about the vehicle horn sound in transportation vehicles etc. But for this the necessary condition is that the plotted spectrograms should have higher resolution. This is possible if we plot the spectrograms using DCT. From the results shown above we can conclude that the spectrograms plotted using DCT are with higher resolution than that plotted using DFT. The spectrograms plotted using DCT having all the information of magnitude of sound than that plotted using same point DFT. So, it will be wise to use DCT for frequency domain conversion before plotting spectrogram than using DFT in the cases where the magnitude of sound is important and not the phasor part. It can be concluded that the resolution of spectrogram is transform dependent.

**References**

- [1] Lori F. Lamel (1993) *Science Direct, Computer Speech & Language*, 7(2), 169-191.
- [2] Hatazaki k., Komori Y., Kawabata T., Shinkano K. (1989) *IEEE, Acoustics, speech & Signal processing*, 1, 393-396.
- [3] Thomas A. Lampert, Nick E. Pears, Simon E.M. O'Keefe (2009) *Advanced Video and Signal Based Surveillance, IEEE Conference*, 330-335.
- [4] Fazlul Haque A.K.M., Hanif Ali Md., Adnan Kiber M. (2010) *IJACSA*, 1(3).
- [5] Mustafa K. Emresoy and Amro El-Jaroudi (1998) *Science Direct, Signal Processing*, 64(2), 157-165.
- [6] Zenton Goh, Kah-Chye Tan and Tan B.T.G. (1998) *IEEE trans. on Speech and Audio Processing*, 6(3), 287-292.
- [7] Elemans C.P.H., Heeck K., Muller M. (2008) *International Journal of Animal Sound and its Recording*, 18(2) 183-212.
- [8] Connolly J.H., Edmonds E.A., Guzya J.J., Johnson S.R. and Woodcock A. (1986) *International Journal of Man-Machine Studies*, 24(6), 611-621.