DFT, DCT, MDCT, DST AND SIGNAL FOURIER SPECTRUM ANALYSIS

L. Yaroslavsky
Dept. of Interdisciplinary Studies, Faculty of Engineering,
Tel Aviv University, Tel Aviv 69978, Israel
yaro@eng.tau.ac.il

Ye Wang
Speech and Audio Systems Lab,
Nokia Research Center, Tampere, Finland
ye.wang@nokia.com

ABSTRACT

DFT, DCT, DST and MDCT are compared in terms of their resolution power for signal Fourier spectrum analysis and energy compaction properties. For test sinusoidal signals with random frequency it was shown by computer simulation that the resolution power of the transforms is not uniform within the frequency band and that on average over the frequency range, DFT, DCT and DST have almost the same resolution power. while MDCT slightly remises them in this respect.

1. INTRODUCTION

Signal Fourier spectrum analysis is one of the major tools of signal processing. For real life continuous signals such as audio signals and images, it is associated with signal integral Fourier transformation. In digital signal processing, integral Fourier transformation is approximated by Discrete Fourier Transforms implemented via Fast Fourier Transform algorithms. From the other side, it has been found that in image and audio coding, restoration and similar applications other transforms such as Discrete Cosine Transform (DCT), Discrete Sine Transform (DST), Modified DCT (Modulated Lapped Transform, MDCT) may perform better then DFT ([1,4]). However, for the appropriate use of these transforms one needs very frequently to establish a correspondence between DFT signal spectra and those of DCT, MDCT, DST and like and evaluate their applicability for signal Fourier analysis. A typical example is MDCT based perceptual audio coding ([2,3]). The paper addresses this issue.

2. INTERRELATION BETWEEN INTEGRAL FOURIER TRANSFORM, DFT, DCT, MDCT, DST

Discrete representation of signal integral transforms parallels that of signals. For signal $a(x)$ and its Fourier spectrum $\alpha(f)$ represented in a discrete form by means of sequences of their samples $\{a_k\}$ and $\{\alpha_r\}$taken at sets of equidistant points $\{(k+u)\Delta x\}$ and $\{(r+v)\Delta f\}$, $k = \ldots,-2,-1,0,1,2,\ldots;$ $r = \ldots,-2,-1,0,1,2,\ldots$ such that

$$a(x) = \sum_k a_k \varphi_x \left(x - (k+u)\Delta x\right), \qquad (1)$$

$$\alpha(f) = \sum_r \alpha_r \varphi_f \left(f - (r+v)\Delta f\right), \qquad (2)$$

where Δx and Δf are discretization intervals and u and v are shifts (in fraction of the corresponding discretization interval) of sample positions from the origin of the corresponding coordinates, discrete representation of the Fourier integral

$$\alpha(f) = \int_{-\infty}^{\infty} a(x) \exp(i2\pi f x) dx \qquad (3)$$

takes form of "Shifted Discrete Fourier Transforms" (SDFT) ([4]):

$$\alpha_r = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} a_k \exp\left(i2\pi \frac{kv}{N}\right) \exp\left(i2\pi \frac{(k+u)r}{N}\right) \qquad (4),$$

the most wide known special case of which (for zero shifts u and v) is DFT:

$$\alpha_r = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} a_k \exp\left(i2\pi \frac{kr}{N}\right) \qquad (5)$$

Popular in digital signal processing DCT, MDCT and DST are yet other special cases of SDFT.

One can show that signal spectra obtained by DCT, MDCT, DST are identical to Shifted Discrete Fourier Transform spectra of signals that are certain permutation modifications of the original signal:

DCT: $$\alpha_r = \sum_{k=0}^{2N-1} a_k \cos\left[\pi \frac{\left(k + \frac{1}{2}\right) r}{N}\right] =$$

$$\frac{1}{2}\sum_{k=0}^{2N-1}\tilde{a}_k \exp\left[i2\pi\frac{\left(k+\frac{1}{2}\right)r}{2N}\right],$$

where

$$\tilde{a}_k=\begin{cases} a_k, & k=0,\dots,N-1 \\ a_{3N-1-k}, & k=N,\dots,2N-1 \end{cases}$$

$$\text{MDCT: } \alpha_r=\sum_{k=0}^{N-1}a_k\cos\left[2\pi\frac{\left(k+\frac{N+2}{4}\right)\left(r+\frac{1}{2}\right)}{N}\right]=$$

$$\frac{1}{2}\sum_{k=0}^{N-1}\tilde{a}_k\exp\left[i2\pi\frac{\left(k+\frac{N+2}{4}\right)\left(r+\frac{1}{2}\right)}{N}\right],$$

where

$$\tilde{a}_k=\begin{cases} a_k-a_{N-1-k}, & k=0,\dots,N-1 \\ a_k+a_{3N-1-k}, & k=N,\dots,2N-1 \end{cases};$$

$$\text{DST: } \alpha_r=\sum_{k=0}^{N-1}a_k\sin\left[\pi\frac{(k+1)(r+1)}{N+1}\right]=$$

$$\frac{1}{2}\sum_{k=0}^{2N}\hat{a}_k\exp\left[i2\pi\frac{(k+1)(r+1)}{N+1}\right],$$

where

$$\tilde{a}_k=\begin{cases} a_k,\ k=0,\dots,N-1 \\ 0,\ k=N,2N+1 \\ -a_{2N-k},\ k=N+1,\dots,2N \end{cases}$$

These relationship mnemonically illustrated in Fig. 1 lucidly explain relationship between the above trigonometric bases and their similarity and dissimilarity.

2. COMPARISON OF TRANSFORM SPECTRAL RESOLUTION POWER AND ENERGY COMPACTION CAPABILITY

In this section, we compare above trigonometric bases in terms of their energy compaction capability and of their resolution power in Fourier spectrum analysis. Transform energy compaction capability means the capability of the transform to redistribute signal energy into small number of transform coefficients. It can be characterized by the fraction of total number of signal transform coefficients that carry certain (substantial) percentage of the signal energy. The lower is this fraction for a given energy percentage, the better is the transform energy compaction capability. This property is the decisive one in most of applications.

The transform resolution power in signal spectral estimation characterizes sharpness of spectral peaks of

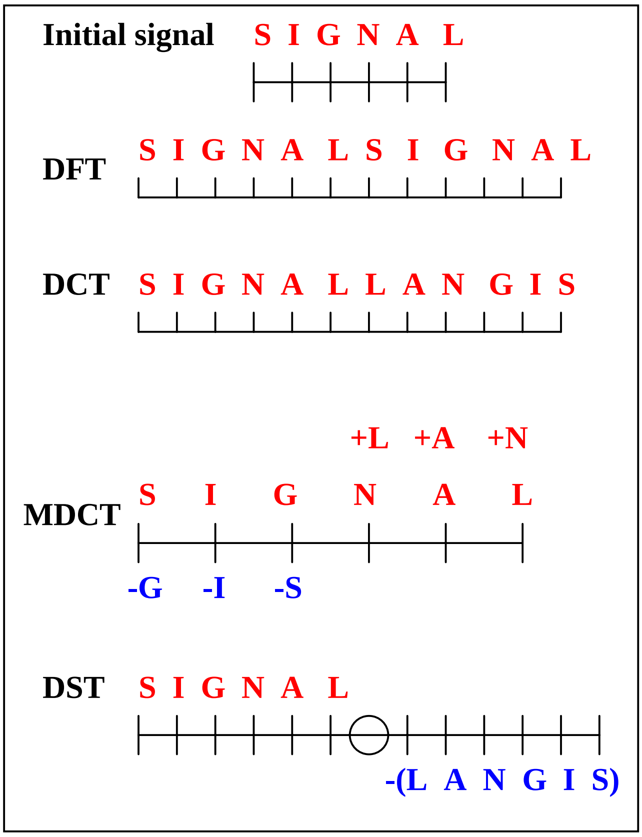

Fig. 1 Signal and its corresponding representations for DFT, DCT, MDCT and DST

sinusoidal signals measured in these bases. It can be evaluated numerically as the width, in fractions of the discretization interval, of the spectral peak within which a given (substantial) percentage of the energy of a sinusoidal signal is contained. From the sampling theory it follows that the width of the spectral peaks in signal discrete spectrum is, in general, proportional to the discretization interval in frequency domain. However, the proportionality coefficient is different for different discrete representations of the Fourier integral which the above trigonometric transforms are.

Evaluation of transform spectral resolution power requires testing spectral peak width for sinusoidal signals of different and arbitrary frequencies within the frequency range defined by the signal discretization rate. Although the evaluation can, in principle, be carried out analytically, obtaining numerical data will anyway require numerical analysis. The same results can be obtained by numerical simulation of the transforms. To this goal, sinusoidal test signals with random frequency uniformly distributed within the corresponding frequency discretization interval should be selected and the results of spectrum estimation should, for each central frequency, be averaged over the realizations of such signals. We demonstrate the results of the simulation in which 100 realizations were used for each frequency sampling interval and the resolution power of spectrum analysis was evaluated in terms of the width

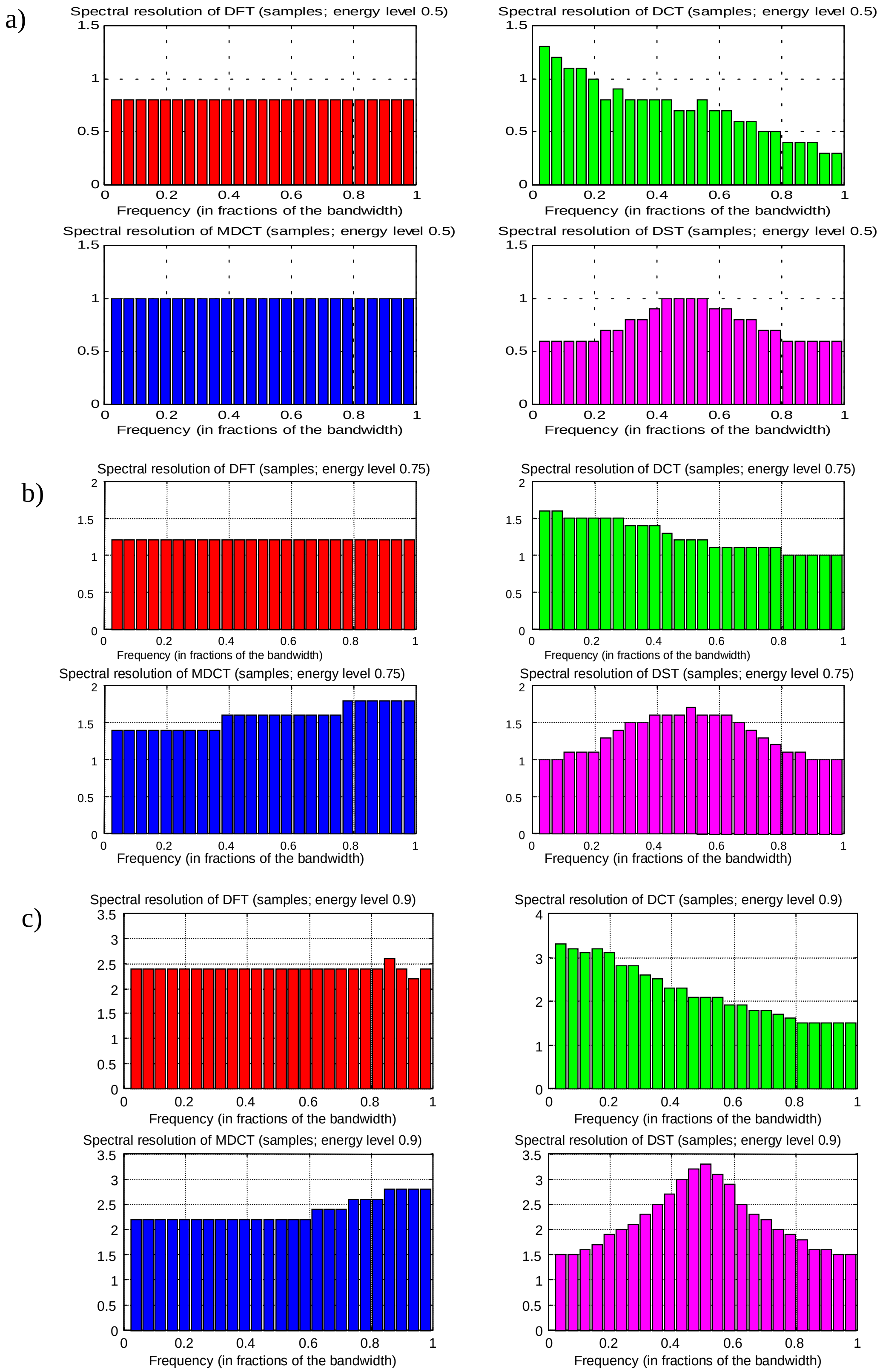

Fig. 2 4 Comparison of spectral resolution (width of signal spectrum) of DFT, DCT, MDCT and DST for sinusoidal signals of 512 samples as a function of signal frequency for different energy levels: a) 0.5; b) 0.75; c) 0.9.

of the signal spectral peaks on three levels of the precentage of signal energy: 50%. 75% and 90%. In order to measure the peak width with a subpixel accuracy, power spectra of test signals were sinc-interpolated with a zoom factor of 5. Results obtained are shown in Figs. 2 and 3.

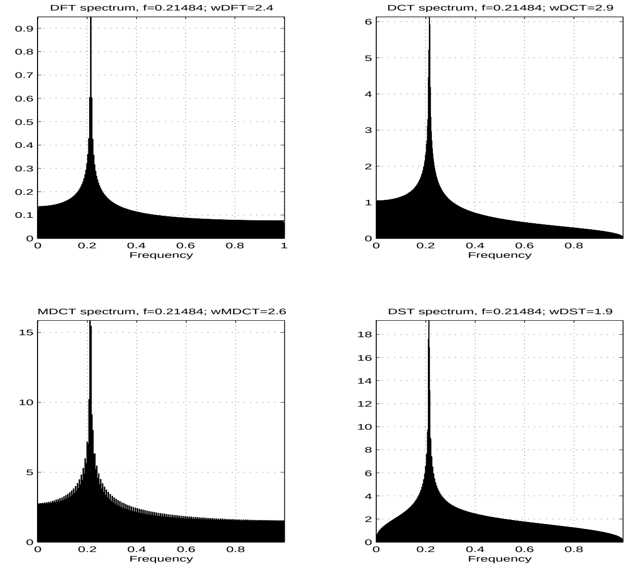

Fig. 3 Examples of spectra of sinusoidal signals with random frequency in the range [54.5-55.5] in the window of 512 samples (no window function). For the display purpose, spectral magnitudes are risen to the power 0.3. wDFT, wDCT, wMDCT and wDST show peak width on the energy level 0.9.

Pure analytical evaluation of the transform energy compaction capability is also problematic since it is feasible only for very limited mathematical models of signals. Another option is to evaluate it experimentally for a number of "typical" test signals. We demonstrate the results of such an evaluation for a set of MPEG audio test signals and some classic and pop music signals. For these signals, energy compaction capability of transforms was compared for different window size (256, 512 and 1024 samples) with and without window functions and on average over large signal sequences of several hundreds of thousands of samples. The results are illustrated in Fig. 4 in frequency coordinates normalized by the signal frequency bandwidth.

3. CONCLUSION

- All above mentioned transforms can be used for signal Fourier analysis.
- The transforms exhibit different Fourier spectrum analysis resolution power; the resolution power is not uniform over the frequency band and it varies from approximately 1.5 to 3.5 frequency samples on the energy level 0.9. Resolution power of DFT is almost uniform in the bandwidth. Resolution power of DCT encreases with frequency and reaches its highest value (the lowest peak width) at high frequencies. Resolution power of DST is lower in the middle of the frequency range. On average over the frequency range, DFT, DCT and DST have almost the same resolution power. MDCT slightly remises them in this respect.
- Within their resolution power, the transforms provide practically identical results of spectrum analysis.
- For audio signals, DCT, MDCT and DST exhibit, on average over large signal sequences, practically similar energy compaction capability. More then 95% energy is concentrated within 10% of the normalized frequency scale for most of the test signals for all transforms concerned. DFT, on average, performs in this respect worse then DCT, DST and MDCT. The energy compaction property of different transforms gets more unified with increasing the window size.

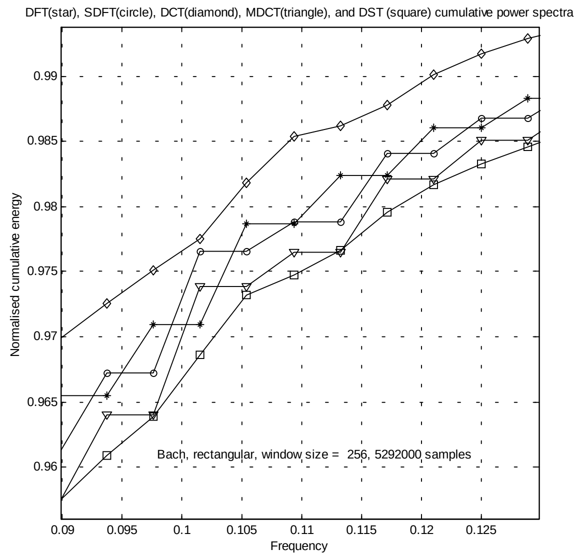

Fig. 4. An example of comparison of convergence of spectra in different bases of a piece of 5292000 samples of Bach music for window size of 256 samples. For each frequency, the data show fraction of total signal energy that is contained in the bandwidth bounded by this frequency

4. REFERENCES

1. Malvar, H., Signal Processing with Lapped Transform, Artech House, Boston, 1991
2. N. Jayant, J. Johnson, R. Safranek, Signal Compression based on Models of Human Perception, Proc. of IEEE, v. 81, No 10, Oct. 1993
3. Ye Wang, Leonid. Yaroslavsky, Miikka Vilermo, Mauri Vaeaenaenen, Restructured audio encoder for improved computational efficiency, AES 108-th Int. Convention, Paris, Febr. 2000
4. L. Yaroslavsky, M. Eden, Fundamentals of Digital Optics, Birkhauser, Boston, 1996.