Comparison of the Cosine and Fourier Transforms of Markov-1 Signals

MASSIH HAMIDI AND JUDEA PEARL

***Abstract*—This correspondence compares the effectiveness of the discrete cosine and Fourier transforms in decorrelating sampled signals with Markov-1 statistics. It is shown that the discrete cosine transform (DCT) offers a higher (or equal) effectiveness than the discrete Fourier transform (DFT) for all values of the correlation coefficient. The mean residual correlation is shown to vanish as the inverse square root of the sample size.**

I. Introduction

In a recent paper Ahmed *et al.* [1] proposed a new transform called discrete cosine transform (DCT) and, based on empirical evidence, conjecture that its performance is closer to the optimal Karhunen-Lóeve transform (KLT) than the other commonly used transforms (i.e., discrete Fourier, Walsh-Hadamard, Haar). Means *et al.* [2] actually used the DCT for encoding TV pictures in real time.

Pearl showed [3] that for a signal statistic characterized by a covariance matrix T, $|T - T_U|^2$ (to be defined later) constitutes a measure of performance for a transform U, in the sense that the error bounds (in coding and filtering) are increasing functions of $|T - T_U|^2$.

The purpose of this investigation is to determine the relation between $|T - T_C|^2$ (the norm obtained using the DCT) and $|T - T_F|^2$ (the norm obtained using the discrete Fourier transform (DFT)), thus testing the conjecture of Ahmed and his collaborators.

II. Definitions and Nomenclature

Let T be a Töeplitz matrix and U an orthogonal transform. Let $T' = UTU^{-1}$ be the representation of T in the new basis, and $T'_U = \text{diag}\,(T'_{11}, T'_{22}, \cdots, T'_{ii}, \cdots, T'_{MM})$. We define T_U to be the representation of T'_U in the first basis, i.e.,

$$T_U = U^{-1} T'_U U$$

and $|T - T_U|^2$ the Hilbert-Schmidt norm of $T - T_U$, i.e.,

$$|T - T_U|^2 = \frac{1}{M}\left(\sum_{m,n=0}^{M-1} |(T - T_U)_{mn}|^2\right).$$

The cosine transform representation of a Töeplitz matrix T is given by CTC^{-1} where C is an $M \times M$ matrix defined by

$$\begin{cases} C_{oj} = \dfrac{2}{M}\dfrac{1}{\sqrt{2}} & j = 0, 1, \cdots, M-1 \\[2ex] C_{kj} = \dfrac{2}{M}\cos\dfrac{k\pi}{2M}(2j+1) & k = 1, 2, \cdots, M-1 \\ & j = 0, 1, \cdots, M-1. \end{cases}$$

A simple algebraic manipulation shows

$$CC^* = C^*C = \frac{2}{M} I$$

(where X^* indicates the complex conjugate transpose of X). Hence

$$C^{-1} = \frac{M}{2} C^*.$$

Manuscript received December 17, 1975; revised March 5, 1976. This work was supported in part by the Naval Undersea Center, San Diego, CA, under Contract N66001-75-0-226-MJE.

The authors are with the School of Engineering and Applied Science, University of California, Los Angeles, CA 90024.

In contrast, the DFT is defined by a unitary matrix F where

$$F_{kj} = M^{-1/2} \exp\left(i \frac{2\pi}{M} kj\right) \qquad k, j = 0, 1, \cdots, M-1.$$

III. Comparison of $|T - T_C|^2$ and $|T - T_F|^2$

For any orthogonal matrix U (e.g., $U = C$ or $U = F$) we have

$$|T - T_U|^2 = |T' - T'_U|^2 = |T'|^2 - \frac{1}{M}\sum_{m=0}^{M-1} |(UTU^{-1})_{mm}|^2$$

$$= |T|^2 - \frac{1}{M}\sum_{m=0}^{M-1} |(UTU^{-1})_{mm}|^2,$$

i.e., the higher the norm of the diagonal vector of the transformed matrix, the lower $|T - T_U|$ and the better the transform. Hence, to compare $|T - T_C|^2$ and $|T - T_F|^2$, it suffices to compare

$$\sum_{m=0}^{M-1} |(CTC^{-1})_{mm}|^2 \text{ and } \sum_{m=0}^{M-1} |(FTF^{-1})_{mm}|^2.$$

We consider matrices of the form

$$T = \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{M-1} \\ \rho & 1 & & \cdots & \rho^{M-2} \\ \vdots & & \ddots & & \vdots \\ \rho^{M-1} & & & \cdots & 1 \end{pmatrix}$$

which represent covariance matrices of Markov-1 signals, with $0 \leqslant \rho \leqslant 1$ being the covariance coefficient between adjacent samples.

Clearly, for $\rho = 0$ and $\rho = 1$ the cosine and Fourier transforms are equivalent, since T is diagonal in both representations. For an intermediate value of ρ we obtained

$$(CTC^{-1})_{00} = \frac{1+\rho}{1-\rho} - \frac{2}{M}\frac{\rho(1-\rho^M)}{(1-\rho)^2}$$

and

$$(CTC^{-1})_{mm} = \frac{1}{2}\left(\frac{e^{i\alpha}+\rho}{e^{i\alpha}-\rho} + \frac{e^{-i\alpha}+\rho}{e^{-i\alpha}+\rho}\right) - \frac{\rho(1-(-1)^m\rho^M)}{M} \cdot \left(\frac{e^{-i(\alpha/2)}}{e^{-i\alpha}-\rho} + \frac{e^{i(\alpha/2)}}{e^{i\alpha}-\rho}\right)^2$$

for $m \neq 0$, where $\alpha = m\pi/M$.

An elementary (but tedious) computation leads to

$$\sum_{m=0}^{M-1} |(CTC^{-1})_{mm}|^2 = \frac{M(1+\rho^2)}{1-\rho^2} - \frac{4\rho^2}{(1-\rho^2)^2} + \frac{2(1-\rho^{2M})\rho^2}{M(1-\rho^2)^3}[3(1+\rho^2)+4\rho] - \frac{4\rho^2(1-\rho^M)^2}{M^2(1-\rho)^4}$$

for $M = 2k$, $k > 1$ (i.e., M even $\geqslant 4$).

Combined with

$$M|T|^2 = M\frac{1+\rho^2}{1-\rho^2} - \frac{2\rho^2(1-\rho^{2M})}{(1-\rho^2)^2},$$

we finally obtain the desired norm

$$M|T - T_c|^2 = \frac{2\rho^2(1+\rho^{2M})}{(1-\rho^2)^2} - \frac{2\rho^2(1-\rho^{2M})}{M(1-\rho^2)^3}(3+4\rho+3\rho^2) + \frac{4\rho^2(1-\rho^M)^2}{M^2(1-\rho)^4}.$$

Note that

$$\lim_{M \to \infty} |T - T_c| = 0$$

implying that the DCT is asymptotically equivalent [4] to the Karhunen-Lóeve transform (KLT) of Markov-1 processes. Moreover, since for large M and $\rho \neq 1$, we have

$$|T - T_c| = \sqrt{2}\,\frac{\rho}{1-\rho^2}\,0(M^{-1/2})$$

we conclude that the degradation in performance in filtering and coding [3] vanishes like $M^{-1/2}$.

An asymptotic equivalence between the KLT and the DCT was also argued by Shanmugam [5] using a circulant extension of T. His argument, however, remains incomplete as the relation between the DCT and the circulant matrices used in [5] is rather unclear.

In order to calculate $|T - T_F|^2$, recall [3] that for $T_{ij} = t(|i-j|)$ we have

$$(T - T_F)_{ij} = \frac{|i-j|}{M}\,[t(|i-j|) - t(M - |i-j|)]$$

and substituting $t(|i-j|) = \rho^{|i-j|}$ we obtain

$$M|T - T_F|^2 = \frac{2\rho^2(1+\rho^{2M})}{(1-\rho^2)^2} - \frac{2(1+\rho^2)\rho^2(1-\rho^{2M})}{M(1-\rho^2)^3} - \frac{\rho^M(M^2-1)}{3}.$$

It shows that the asymptotic behavior of $|T - T_F|$ for large M is identical to that of $|T - T_c|$. Thus, the performance difference between the DCT and the DFT must vanish like M^{-1}. Indeed, for large M one obtains the positive difference

$$|T - T_F|^2 - |T - T_c|^2 \cong \frac{4\rho^2}{M^2(1-\rho^2)(1+\rho)^2} \qquad \rho < 1,$$

indicating that the cosine transform is closer to optimal than the Fourier transform over the entire range of $0 < \rho < 1$.

For moderate values of M we should examine the expressions for $|T - T_F|^2$ and $|T - T_C|^2$ over the range $0 \leqslant \rho \leqslant 1$. The two are plotted, in a normalized form, in Fig. 1. We chose $|T - I|^2$ as a common normalizing factor, where I is the identity matrix, and so

$$|T - I|^2 = \frac{2\rho^2}{M(1-\rho^2)^2}\,[M - 1 - M\rho^2 + \rho^{2M}].$$

It measures the degree of cross correlation contained in the unprocessed signal, and, therefore, the maximum amount of decorrelation that can be accomplished by any transform (i.e., the KLT). The ratio

$$\frac{|T - T_U|^2}{|T - I|^2}$$

represents the fractional correlation left "undone" by a transformation U.

Fig. 1 shows that for $M = 8, 16, 64$, and for the entire range of $0 < \rho < 1$, $|T - T_F|^2$ is higher than $|T - T_C|^2$. The difference between the two are quite noticeable, occasionally reading a ratio of 2:1.

CONCLUSIONS

We established that the DCT is asymptotically equivalent to the KLT of Markov-1 signals and demonstrated that the rate of convergence is on the order of $M^{-1/2}$. $|T - T_C|^2$ is shown to be smaller than $|T - T_F|^2$ for all values of M and ρ, i.e., the discrete cosine transform offers a better approximation to the KLT of Markov-1 signals than the DFT.

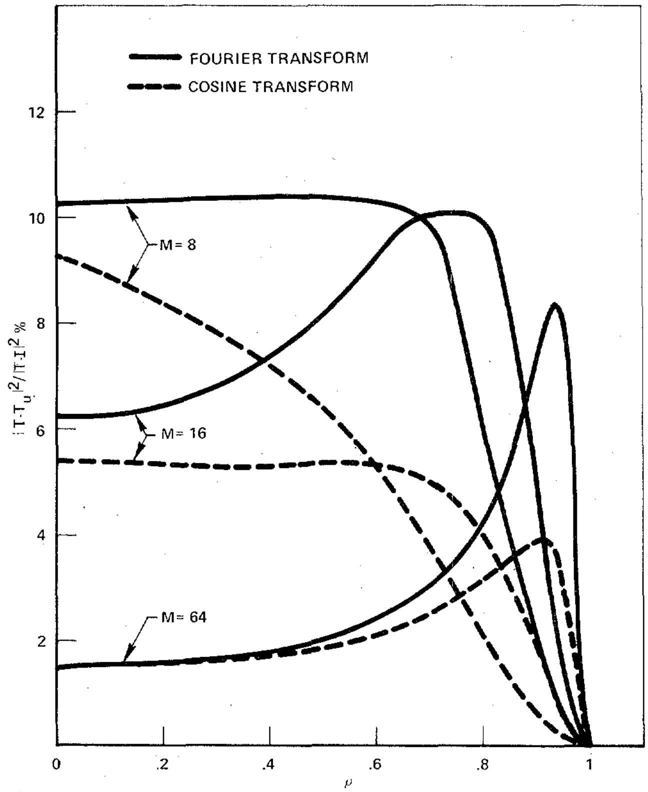

Fig. 1. Normalized correlation measures for Fourier (solid lines) and cosine (broken lines) transforms.

REFERENCES

[1] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE Trans. Comput.* (Corresp.), vol. C-23, pp. 90-93, Jan. 1974.

[2] R. W. Means, H. J. Whitehouse, and J. M. Speiser, "Television encoding using a hybrid discrete cosine transform and a differential pulse code modulator in real time," in *Proc. IEEE Nat. Telecommunications Conf.*, San Diego, CA, Dec. 1974.

[3] J. Pearl, "On coding and filtering stationary signals by discrete Fourier transforms," *IEEE Trans. Inform. Theory* (Corresp.), vol. IT-19, pp. 229-232, Mar. 1973.

[4] —, "Asymptotic equivalence of spectral representations," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-23, pp. 547-551, Dec. 1975.

[5] K. S. Shanmugam, "Comments on discrete cosine transform," *IEEE Trans. Comput.* (Corresp.), vol. C-24, p. 759, July 1975.

Two Preliminary Studies of the Intelligibility of Predictor-Coefficient and Formant-Coded Speech

LYNN O. KEELER, GARY L. CLEMENT, WILLIAM J. STRONG, AND E. PAUL PALMER

Abstract—**Two preliminary studies comparing the intelligibilities of predictor-coefficient versus formant-frequency-coded speech and the intelligibilities of predictor-coefficient-coded speech using different numbers of coefficients are reported.**

INTRODUCTION

Much of the motivation for research in speech analysis-synthesis stems from a desire to better understand the essential

Manuscript received July 16, 1975; revised January 9, 1976.

The authors are with the Department of Physics and Astronomy, Brigham Young University, Provo, UT 84602.