DISTRIBUTIONS OF 3D DCT COEFFICIENTS FOR VIDEO

Malavika Bhaskaranand and Jerry D. Gibson

Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA - 93106
Email: {malavika, gibson}@ece.ucsb.edu

ABSTRACT

The three-dimensional discrete cosine transform (3D DCT) has been proposed as an alternative to motion-compensated transform coding for video content. However, so far no definitive study has been done on the distribution of 3D DCT coefficients of video sequences. This study performs two goodness-of-fit tests, the Kolmogorov-Smirnov (KS) test and the χ^2-test, to determine the distribution that best fits the 3D DCT coefficients of the luminance components of video sequences with low motion or structured motion. The results indicate that the DC coefficient can be well approximated by a Gaussian distribution and a majority of the high-energy AC coefficients can be approximated by a Gamma distribution. Knowledge of the coefficient distributions can be used to design quantizers optimized for 3D DCT coefficients and hence achieve better coding efficiency.

***Index Terms*—** Video compression, Three-dimensional discrete cosine transform, DCT coefficient distributions

1. INTRODUCTION

Most video compression schemes today perform motion compensation followed by a two-dimensional transform (hybrid coding). Video can also be viewed as 3D data with two spatial dimensions and a time dimension. In this context, the 3D DCT was first proposed by Roese *et al.* [1]. However, it was side-lined then because of its high computational cost and memory requirements compared to hybrid coding [2]. Now, with increasingly complex hybrid video coders like H.264 and cheap computational power and memory, interest in the 3D DCT has been renewed [3]- [6].

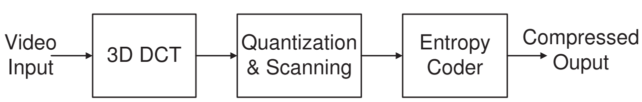

Fig. 1. A 3D DCT based video coder

A block diagram of a 3D DCT coder for video is shown in Figure 1. $N \times N \times N$ data cubes that are N pixels wide, N pixels high, and N frames deep are extracted from the video sequence. The $N \times N \times N$ 3D DCT is applied on each data cube. The data cubes could overlap in the spatial dimensions to reduce blocking artifacts or in the temporal dimension to reduce the dependency of adjacent data cubes along time. The coefficients are quantized and then scanned into a 1D array in order of increasing probability of coefficients being zero, in order to statistically increase the runs of zeros and achieve better compression ratio by run-length encoding.

This research has been supported by the California Micro Program, Applied Signal Technology, Cisco, Sony-Ericsson and Qualcomm, Inc., and by NSF Grant Nos. CCF-0429884, CNS-0435527, and CCF-0728646.

Finally, the quantized and re-ordered coefficients are entropy coded to produce a compressed bitstream. Such a video coding scheme can be expected to work well for video sequences with low motion content. For high motion sequences, a motion-compensated 3D DCT scheme might be required to achieve good compression efficiency.

Several techniques have been proposed to compress video using the 3D DCT. Bauer and Sayood [3] propose the use of dynamic code selection to code 3D DCT coefficients. Servais and de Jager [4] reduce the entropy of the 3D DCT coefficients by exploiting the relationship between them and the 2D DCT coefficients of the first frame in the group. Yeo and Liu [7] propose the 3D zig-zag scan as an extension to the widely used 2D zig-zag scan. Chan and Lee [5] conclude from empirical observations that the AC coefficients can be modeled by a Laplacian distribution and propose a scanning and quantization method based on a shifted complement hyperboloid function. Bozinovic and Konrad [6] show that global, constant velocity, translational motion in a video sequence results in the 3D DCT coefficients occupying a plane and based on this result develop a motion-adaptive scanning and quantization method for 3D DCT coefficients. Although the knowledge of coefficient statistics is essential for designing optimal quantizers, none of the earlier works has studied the distributions of 3D DCT coefficients.

In contrast, several works have studied the distributions of 2D DCT coefficients of images. Reininger and Gibson [8] establish that the DC coefficient statistics approximate a Gaussian distribution and the AC coefficients fit a Laplacian distribution, based on results of the Kolmogorov-Smirnov test and simulations. Lam and Goodman [9] mathematically prove that the distribution of DCT coefficients is close to Laplacian using a double stochastic model of images. Smoot and Rowe [10] show that the DCT coefficients of luminance and chrominance components and of the differential signal obtained after motion estimation are Laplacian. Eggerton and Srinath [11] conclude that no single distribution fits all the DCT coefficients for all images, but Laplacian fits the majority and the Cauchy distribution fits the ensemble of all DCT coefficients. Bellifemine *et al.* [12] study the statistics of the DCT coefficients of the differential signal produced after motion estimation and demonstrate that the Laplacian distribution is a good approximation. Muller [13] shows that the Generalized Gaussian function better models the statistics of the DCT coefficients compared to the Laplacian distribution. Joshi and Fischer [14] compare the performance of the Generalized Gaussian and Laplacian models in image coding and conclude that the more complex Generalized Gaussian model does not give a significant advantage over the Laplacian model. Chang *et al.* [15] propose that the DCT coefficients are best modeled by the Generalized Gamma function.

This paper studies the distributions that best approximate the statistics of 3D DCT coefficients with most of the energy content. Kolmogorov-Smirnov (KS) and χ^2 goodness-of-fit tests have been performed considering the Gaussian, Laplacian, Gamma and Rayleigh distributions as probable models. The model distribution that gives the minimum KS or χ^2 statistic is chosen as the best fit. Results for the luminance components of videos with low motion or structured motion show that the 3D DCT DC coefficient can be approximated by a Gaussian distribution and a majority of the significant AC coefficients can be approximated by a Gamma distribution. These results can be used to design optimal quantizers for 3D DCT coefficients. The paper is organized as follows: Section 2 provides definitions of the 3D DCT, goodness-of-fit tests and model distributions. Section 3 gives a summary of the experiments conducted and presents extensive results.

2. DEFINITIONS

2.1. Three dimensional DCT

The discrete cosine transform (DCT) proposed by Ahmed, Natarajan and Rao [16], has energy packing efficiency close to that of the optimal Karhunen-Loeve transform. In addition, it is signal-independent and can be computed efficiently by fast algorithms. For these reasons, the DCT is widely used in image and video compression. Since the common three-dimensional DCT kernel is separable, the 3D DCT is usually obtained by applying the one-dimensional DCT along each of the three dimensions. Thus, the $N \times N \times N$ 3D DCT can be defined as [1]

$$X(u,v,w) = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1} x(i,j,k)C(i,u)C(j,v)C(k,w) \quad (1)$$

$$x(i,j,k) = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\sum_{w=0}^{N-1} X(u,v,w)C(i,u)C(j,v)C(k,w) \quad (2)$$

$$\text{where } C(p,q) = \begin{cases} \frac{1}{\sqrt{N}} & ,q=0 \\ \sqrt{\frac{2}{N}}cos\left(\frac{(2p+1)q\pi}{2N}\right) & ,q \neq 0 \end{cases}$$

2.2. Goodness-of-fit tests

Goodness-of-fit tests are used to examine the hypothesis that a given data set comes from a model distribution with given parameters. The Kolmogorov-Smirnov(KS) test and the χ^2-test are two popular goodness-of-fit tests. For characterizing the statistics of 2D DCT coefficients, the KS test was used in [8], [10] and [11] and the χ^2-test was used in [12], [13] and [15]. Therefore these two goodness-of-fit tests were chosen for this study.

2.2.1. *Kolmogorov-Smirnov test*

This test compares the empirical cumulative distribution function (CDF) with the given model CDF. Given the sample data set $X = \{x_1, x_2, \ldots, x_M\}$ with order statistics $x_{(n)}$, $n = 1, 2, \ldots, M$, the empirical CDF is defined as

$$\widehat{F}_X(x) = \begin{cases} 0, & x < x_{(1)} \\ \frac{n}{M}, & x_{(n)} \leq x < x_{(n+1)}, n = 1, 2, \ldots, M-1 \\ 1, & x \geq x_{(M)} \end{cases} \quad (3)$$

The KS statistic D_n is then defined as [17]

$$D_n = \max_{i=1,2,\ldots,M} |F_X(x_{(i)}) - \widehat{F}_X(x_{(i)})| \quad (4)$$

It can be seen that the KS statistic is a measure of the distance between the empirical CDF and the model CDF and therefore a measure of the goodness-of-fit. If the empirical CDF is tested against several model CDFs, the model that gives the minimum KS statistic can be taken to be the best fit for the data.

2.2.2. *χ^2-test*

Unlike the KS test, the χ^2-test compares probability density functions (pdf). The range of the data is divided into k disjoint and exhaustive bins A_i, $i = 1, 2, \ldots, k$. Let $E_i = np_i$ be the expected frequency in bin A_i with $p_i = P(x \in A_i)$ and n the total number of data samples. Let O_i be the observed frequency in bin A_i. Then the χ^2-statistic is defined as [17]

$$V_k = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \quad (5)$$

Since this is a measure of deviation of the empirical frequencies from the expected frequencies, the model pdf that gives the minimum χ^2-statistic can be considered as the best fit.

2.3. Probability Distributions

The Gaussian, Laplacian, Gamma and Rayleigh distributions are commonly used for modelling distributions of DCT coefficients. Therefore, they were chosen for this study and are specified below.

The Gaussian pdf with mean μ and variance σ^2 is defined as

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (6)$$

The Laplacian pdf with mean μ and variance $\frac{2}{\lambda^2}$ is defined as

$$f_X(x) = \frac{\lambda}{2}\exp\left(-\lambda|x-\mu|\right) \quad (7)$$

The Gamma pdf with mean μ and variance σ^2 is defined as [18]

$$f_X(x) = \frac{\sqrt[4]{3}}{\sqrt{8\pi\sigma|x-\mu|}}\exp\left(-\frac{\sqrt{3}|x-\mu|}{2\sigma}\right) \quad (8)$$

The Rayleigh pdf with mean $\sigma\sqrt{\frac{\pi}{2}}$ and variance $\frac{4-\pi}{2}\sigma^2$ is defined as

$$f_X(x) = \frac{x}{\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2}\right), x \geq 0 \quad (9)$$

3. EXPERIMENTS AND RESULTS

$8 \times 8 \times 8$ data cubes were extracted from the luminance components of video sequences. Adjacent cubes were chosen without any overlap in the spatial or temporal domains because overlapping along spatial dimensions is not popular for 2D DCT and overlap along time dimension is not widely used in earlier work on the 3D DCT. The $8 \times 8 \times 8$ 3D DCT was applied to each of the data cubes. The KS and χ^2-tests were performed on all coefficients with more than 0.01% of the total energy. Coefficients with less than 0.01% of the total energy were not taken into consideration because they had pdfs narrowly concentrated around zero and hence were insignificant. The Gaussian, Laplacian and Gamma distributions with mean and variance equal to the sample mean and sample variance were considered. In addition, for coefficients with positive means, the Rayleigh distribution with variance equal to the sample variance was used.

Table 1 gives the details of the video sequences used in this study. Video sequences with complex motion and scene changes were excluded because they would have had sudden changes in coefficient statistics which could have skewed the conclusions drawn from the statistics computed over the entire sequence. Moreover, for such sequences, we do not expect the 3D DCT coding scheme to give good compression ratios and hence intend to use a motion-compensated 3D DCT scheme. The sequences chosen provide

Table 1. Table of test sequences

Test number	Sequence	Resolution	Number of frames
1	akiyo	176 x 144	296
2	foreman	176 x 144	160
3	container	176 x 144	296
4	news	176 x 144	296
5	silent	176 x 144	296
6	mother-daughter	176 x 144	296
7	bus	352 x 288	144
8	flower	352 x 288	248
9	foreman	352 x 288	160
10	paris	352 x 288	1064

7920–210672 samples per coefficient; therefore the empirical distributions derived from these sequences can be assumed to be close to the actual distributions.

The KS and χ^2 statistics were computed for the coefficients under consideration against the model distributions. The distribution that gave the minimum KS statistic was chosen as the one with the best fit under the KS criterion. Similarly for the χ^2 statistic. The results for the first few coefficients are given in Tables 2 and 3. The minimum statistics and the best fit for each coefficient has been indicated in bold. The differences in the choice of best fit model for few coefficients, based on the goodness-of-fit statistic are only statistical. From the tables it can be seen that the Gamma distribution well approximates 13 of 18 AC coefficients based on the KS test and 15 of 18 AC coefficients based on the χ^2-test. Thus, it can be concluded that for the AC coefficients under consideration, the Gamma distribution fits best for a majority of the coefficients. Results for the DC coefficient are discussed later in the paper.

Table 2. KS statistics for few coefficients of 2 test sequences

	Video #4				Video #10			
Coeff.	Gau.	Lap.	Gam.	Ray.	Gau.	Lap.	Gam.	Ray.
0,0,0	0.149	0.209	0.275	**0.125**	**0.089**	0.124	0.197	0.208
0,1,0	0.236	**0.213**	0.233	-	0.161	**0.133**	0.166	-
1,0,0	0.190	0.143	**0.090**	0.754	0.185	0.152	**0.127**	0.787
0,0,1	0.350	0.326	**0.276**	0.916	0.327	0.301	**0.209**	-
0,2,0	0.248	**0.240**	0.307	0.794	0.195	0.156	**0.077**	-
1,1,0	0.197	0.161	**0.106**	-	0.177	0.125	**0.083**	0.788
2,0,0	0.195	**0.139**	0.198	0.756	0.188	0.138	**0.070**	0.802
0,1,1	0.343	0.319	**0.285**	0.910	0.333	0.313	**0.228**	0.894
1,0,1	0.333	0.314	**0.292**	0.902	0.335	0.314	**0.228**	0.897
0,0,2	0.407	**0.404**	0.466	-	0.353	0.330	**0.242**	0.911

Table 3. χ^2 statistics for few coefficients of 2 test sequences

	Video #4				Video #10			
Coeff.	Gau.	Lap.	Gam.	Ray.	Gau.	Lap.	Gam.	Ray.
0,0,0	1976	4407	-	**1016**	**32k**	49k	-	-
0,1,0	9476	3931	**674**	-	97k	27k	**5773**	-
1,0,0	6144	1947	**396**	-	-	15k	**3573**	-
0,0,1	**573**	1614	632	15k	**1329**	65k	3941	-
0,2,0	5059	2409	**635**	38k	193k	32k	**7050**	-
1,1,0	4921	1856	**136**	-	67k	16k	**1134**	-
2,0,0	4566	1700	**230**	265k	284k	36k	**4042**	980k
0,1,1	5461	3389	**1976**	18k	415k	60k	**29k**	-
1,0,1	2950	2153	**1453**	107k	28k	54k	**8820**	68k
0,0,2	7063	5693	**3226**	-	102k	102k	61k	**796**

A 3D-zigzag scan has been proposed by Yeo and Liu [7] as an extension to the 2D-zigzag scan widely used in image and video compression. The basic idea is that coefficients are ordered based on the sum of their indices $(u + v + w)$; the smaller the sum, the lower the frequency. Therefore, based on the 3D-zigzag scan, the $8 \times 8 \times 8$ 3D DCT coefficients can be grouped into 22 frequency groups. Tables 4 and 5 give the number of coefficients in each frequency group which best fit the given distribution for two video sequences based on the two goodness-of-fit tests. The maximum number of coefficients across the different distributions is highlighted in bold. Although in some frequency groups the Gamma distribution is not the best fit for a majority of the AC coefficients, across the whole sequence it can be seen that the Gamma distribution fits best for a majority of the AC coefficients. As in the previous tables, the differences in results based on the goodness-of-fit test are statistical. Another interesting observation is that all coefficients beyond frequency group 10 have very little energy. Therefore, the 3D-zigzag scan might be a simple yet efficient method for scanning 3D DCT coefficients.

Table 4. Results for frequency groups based on KS test

	Video #4				Video #7			
(u+v+w)	Gau.	Lap.	Gam.	Ray.	Gau.	Lap.	Gam.	Ray.
0	0	0	0	**1**	**1**	0	0	0
1	0	1	**2**	0	0	0	**3**	0
2	0	**3**	**3**	0	0	1	**5**	0
3	0	2	**8**	0	0	1	**9**	0
4	0	3	**8**	0	0	5	**10**	0
5	1	3	**5**	0	0	5	**16**	0
6	1	2	**4**	0	0	10	**14**	0
7	0	3	**4**	0	0	**13**	12	0
8	0	0	**5**	0	0	**13**	9	0
9	0	0	**1**	0	0	4	**5**	0
10	0	0	0	0	0	0	**2**	0
11	0	0	0	0	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
21	0	0	0	0	0	0	0	0

Table 5. Results for frequency groups based on χ^2-test

	Video #4				Video #7			
(u+v+w)	Gau.	Lap.	Gam.	Ray.	Gau.	Lap.	Gam.	Ray.
0	0	0	0	**1**	0	0	0	**1**
1	1	0	**2**	0	0	0	**3**	0
2	0	0	**6**	0	0	0	**6**	0
3	2	0	**7**	1	0	1	**9**	0
4	1	0	**10**	0	0	1	**14**	0
5	0	0	**9**	0	0	1	**20**	0
6	0	0	**7**	0	0	3	**21**	0
7	0	0	**7**	0	0	4	**21**	0
8	0	0	**5**	0	0	2	**20**	0
9	0	0	**1**	0	0	1	**8**	0
10	0	0	0	0	0	0	**2**	0
11	0	0	0	0	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
21	0	0	0	0	0	0	0	0

The results of the KS and χ^2-tests for all the video sequences used in this study are consolidated in Tables 6 and 7. Each number in the tables represents the number of coefficients for which the corresponding distribution is the best fit. The distribution which best fits a majority of the coefficients for each sequence is highlighted in

Table 6. Consolidated results from KS test

Test #	DC Gau.	DC Lap.	DC Gam.	DC Ray.	AC Gau.	AC Lap.	AC Gam.	AC Ray.
1	0	0	0	**1**	0	5	**19**	0
2	**1**	0	0	0	0	8	**23**	0
3	**1**	0	0	0	2	7	**21**	0
4	0	0	0	**1**	2	17	**40**	0
5	**1**	0	0	0	2	**13**	12	0
6	**1**	0	0	0	7	**8**	6	0
7	**1**	0	0	0	0	52	**85**	0
8	**1**	0	0	0	0	25	**50**	0
9	**1**	0	0	0	0	6	**20**	0
10	**1**	0	0	0	0	3	**55**	0
Total	**8**	0	0	2	13	144	**331**	0

Table 7. Consolidated results from χ^2-test

Test #	DC Gau.	DC Lap.	DC Gam.	DC Ray.	AC Gau.	AC Lap.	AC Gam.	AC Ray.
1	0	0	0	**1**	0	1	**23**	0
2	**1**	0	0	0	0	3	**28**	0
3	0	**1**	0	0	0	1	**29**	0
4	0	0	0	**1**	4	0	**54**	1
5	0	**1**	0	0	0	8	**18**	1
6	0	**1**	0	0	1	2	**18**	0
7	0	0	0	**1**	0	13	**124**	0
8	**1**	0	0	0	0	0	**75**	0
9	**1**	0	0	0	0	0	**26**	0
10	**1**	0	0	0	3	0	**53**	2
Total	**4**	**3**	0	**3**	8	28	**448**	4

bold. Based on the KS test, 8 of 10 DC coefficients can be approximated by the Gaussian distribution and 331 of 488 AC coefficients can be approximated by the Gamma distribution. Based on the χ^2-test, 4 of 10 DC coefficients can be approximated by the Gaussian distribution and 448 of 488 AC coefficients can be approximated by the Gamma distribution. Hence it can be concluded that the Gaussian distribution is a good fit for a majority of the DC coefficients and the Gamma distribution is a good approximation for a majority of the AC coefficients.

4. CONCLUSIONS

This study performs two goodness-of-fit tests to determine the distribution of the 3D DCT coefficients for the luminance components of video sequences with low motion or structured motion. The results show that no single distribution can be used to model the distributions of all coefficients for all video sequences. However, the distributions of a majority of the significant AC coefficients can be modeled by the Gamma distribution and the distribution of the DC coefficient can be approximated by a Gaussian distribution in most cases. This knowledge can enable the design of optimal quantizers for 3D DCT coefficients that produce minimum distortion and thus achieve close to optimal compression efficiency.

The $8 \times 8 \times 8$ 3D DCT was used as a starting point because it has been widely used in earlier studies. Further examination of the possibility of using different sizes of the 3D DCT cube needs to be done. Also, future studies could be done on the motion-compensated 3D DCT coefficients for video sequences with complex motion.

5. ACKNOWLEDGEMENT

The authors would like to acknowledge the discussions with Dr. S. Rao Jammalamadaka on goodness-of-fit tests.

6. REFERENCES

[1] J. Roese, W. Pratt, and G. Robinson, "Interframe cosine transform image coding," *IEEE Trans. Commun.*, vol. 25, no. 11, pp. 1329–1339, Nov 1977.

[2] R. Westwater and B. Furht, "Three-dimensional DCT video compression technique based on adaptive quantizers," *Proc. of IEEE International Conf. on Engineering of Complex Computer Systems, 1996*, pp. 189–198, Oct 1996.

[3] M. Bauer and K. Sayood, "Video coding using 3 dimensional DCT and dynamic code selection," *Proceedings of Data Compression Conference, 1995*, p. 451, Mar 1995.

[4] M. Servais and G. de Jager, "Video compression using the three dimensional discrete cosine transform (3D-DCT)," *Proceedings of the South African Symposium on Commun. and Signal Process. 1997, COMSIG '97*, pp. 27–32, Sep 1997.

[5] R.K.W. Chan and M.C. Lee, "3D-DCT quantization as a compression technique for video sequences," *Proc. of International Conf. on Virtual Sys. and MultiMedia*, pp. 188–196, Sep 1997.

[6] N. Bozinovic and J. Konrad, "Motion analysis in 3D DCT domain and its application to video coding," *Signal Processing: Image Communication*, vol. 20, pp. 510–528, July 2005.

[7] B.L. Yeo and B. Liu, "Volume rendering of DCT-based compressed 3D scalar data," *IEEE Trans. Visualization and Computer Graphics*, vol. 1, no. 1, pp. 29–43, Mar 1995.

[8] R. Reininger and J. Gibson, "Distributions of the two-dimensional DCT coefficients for images," *IEEE Trans. Commun.*, vol. 31, no. 6, pp. 835–839, Jun 1983.

[9] E.Y. Lam and J.W. Goodman, "A mathematical analysis of the DCT coefficient distributions for images," *IEEE Trans. on Image Process.*, vol. 9, no. 10, pp. 1661–1666, Oct 2000.

[10] S.R. Smoot and L.A. Rowe, "Study of DCT coefficient distributions," *Proc. SPIE*, vol. 2657, pp. 403–411, Jan. 1996.

[11] J.D. Eggerton and M.D. Srinath, "Statistical distributions of image DCT coefficients," *Comput. Elect. Eng.*, vol. 12, pp. 137–145, 1986.

[12] F. Bellifemine, A. Capellino, A. Chimienti, R. Picco, and R. Ponti, "Statistical analysis of the 2D-DCT coefficients of the differential signal for images," *Signal Process. Image Commun.*, vol. 4, pp. 477–488, Nov. 1992.

[13] F. Muller, "Distribution shape of two-dimensional DCT coefficients of natural images," *Electron. Letters*, vol. 29, no. 22, pp. 1935–1936, Oct. 1993.

[14] R.L. Joshi and T.R. Fischer, "Comparison of generalized gaussian and laplacian modeling in DCT image coding," *IEEE Signal Process. Letters*, vol. 2, no. 5, pp. 81–82, May 1995.

[15] J.H. Chang, J.W. Shin, Kim N.S., and S.K. Mitra, "Image probability distribution based on generalized gamma function," *IEEE Signal Process. Letters*, vol. 12, no. 4, pp. 325–328, April 2005.

[16] N. Ahmed, T. Natarajan, and K.R. Rao, "Discrete cosine transfom," *IEEE Trans. Computers*, vol. C-23, no. 1, pp. 90–93, Jan. 1974.

[17] V.K Rohatgi and A.K.E Saleh, *An introduction to probability & statistics*, pp. 504–505,608–609, John Wiley & Sons, 2001.

[18] N.S. Jayant and P. Noll, *Digital coding of waveforms*, p. 34, Prentice Hall, 1984.