

A THEORETICAL MODEL FOR TIME CODE MODULATION

Dimitris Kalogiros and Vassilis Stylianakis

Wire Communications Laboratory, Dept. of Electrical and Computer Engineering
University of Patras, GREECE

E-mail: kalogiro@wcl.ee.upatras.gr, stylian@wcl.ee.upatras.gr

ABSTRACT

The traditional waveform coding techniques for digital communication systems convert the original analog input signal into a digital bit stream using uniform sampling in the time domain, i.e., PCM, DM, ADPCM. In this paper we propose the Time Code Modulation (TCM) technique as an alternative coding scheme, where information is extracted from the signal, only at the time instants when necessary. This results in a variable sampling rate, where its mean value is significantly less than the Nyquist rate. In addition we suggest a general theoretical model for TCM and we present simulation results for various implementations of TCM coders and decoders. A theoretical estimation of SNR vs. sampling rate performance is also presented.

1. INTRODUCTION

The class of band-limited signals consists of finite energy signals $x(t)$, with their Fourier transforms $X(f)$ vanishing outside the interval $(-w_{\max}, w_{\max})$. However, high frequency components of a band-limited signal do not appear in every time instant of the duration of the signal.

Generally, we can consider that the spectral characteristics of a band-limited signal are time varying [1] and the quantity w_{\max} is also a function of time. A band-limited signal of this category is defined by the following relation:

$$x(t) = \int_{-w(t)}^{w(t)} X(f, t) e^{j2\pi ft} df \quad (1)$$

where $X(f, t)$ denotes the time varying spectra of the signal and $\max(w(t)) = w_{\max}$.

Most conventional (uniform sampling) waveform coders [2]-[4] are designed to remove the redundancy in a waveform for purposes of bit rate reduction. However, sampling at a constant rate, which is imposed by w_{\max} , sets a limit on the system performance. This especially happens when the mean value of $w(t)$ differs strongly from w_{\max} . Obviously, it is more efficient to sample the low frequency regions of the signal at a lower rate than the high frequency regions.

An example of a non-uniform sampling technique applied to speech waveforms is given in [5]. The speech waveform is sampled at the time instants of its maxima and minima. The values of the local extrema (minima or maxima) as well as their

time instants result in an intelligible approximation of the original speech waveform at the decoder output.

Other non-uniform sampling techniques are presented in [1]. However, their performances are not adequate and the achieved mean sampling rates are close to the Nyquist rate [1], [5].

The research presented in this paper introduces a sampling technique which can be regarded as a generalization of the Time Code Modulation techniques, which are presented in [6]-[8]. A waveform coder based on this technique, would be able to decrease inherently its sampling rate whenever the input waveform changes in a predictable way due to the absence of its high frequency components. On the other hand, the sampling rate increases whenever the input waveform changes in a rapid and unpredictable way due to the presence of its high frequency components.

2. SYSTEM DESCRIPTION

The block diagram of the coder is shown in Fig. 1. The input signal $x_a(t)$ is sampled at a rate F_0 ($T_0 = 1/F_0$), much higher than the Nyquist rate. The sampling process produces the sequence of samples x_n .

The coder uses the predictive model M_k to produce an approximation \hat{x}_n for every x_n . The difference $x_n - \hat{x}_n$ is compared to the threshold value Δ .

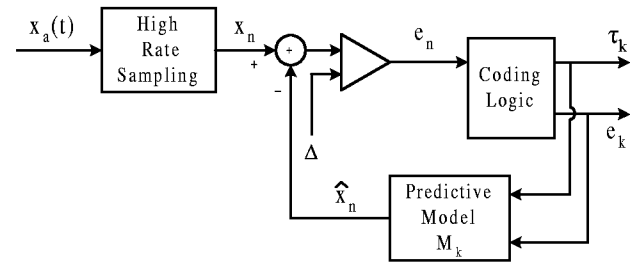


Figure 1 The block diagram of the coder.

The output of the comparator is given by the following equation:

$$e_n = \begin{cases} -1 & x_n - \hat{x}_n < -\Delta \\ 0 & |x_n - \hat{x}_n| \leq \Delta \\ 1 & x_n - \hat{x}_n > \Delta \end{cases} \quad (2)$$

Let us assume that index ns_k stands for the critical value of index n where the inequality $|x_n - \hat{x}_n| > \Delta$ becomes valid for the k -th time. Product $ns_k T_0$ stands for the corresponding critical time instant. In this critical time instant the coder creates two source symbols at its output: 1) the value of the current critical e_n , i.e., e_{ns_k} , which is represented by e_k 2) the time interval τ_k between the current critical time instant $ns_k T_0$ and the previous one $ns_{k-1} T_0$, i.e., $\tau_k = ns_k T_0 - ns_{k-1} T_0$. In correspondence with uniform sampling techniques, critical time instants are called “sampling points”.

Due to the fact that F_0 is much higher than the Nyquist rate, we can approximate the value of the original signal at the sampling points, by the means of the following equation:

$$x_a(ns_k T_0) = x_{ns_k} \approx \hat{x}_{ns_k} + e_k \Delta \quad (3)$$

The parameters of the predictive model M_k are optimized before the calculation of $\hat{x}_{ns_{k+1}}$, and a new model, denoted by M_{k+1} , is derived. The coder achieves this optimization by using (3) to calculate the values of the signal in the $ns_1 T_0, ns_2 T_0, \dots, ns_k T_0$ time points.

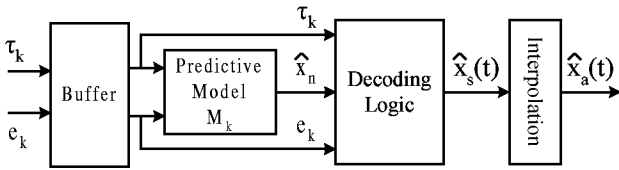


Figure 2 The block diagram of the decoder.

The decoder functions in the same way as the feedback branch of the coder. It uses τ_k and e_k in order to produce the sequence \hat{x}_n and to optimize its predictive model M_k in the same way that the coder does at the feedback branch. The predictive model M_k in both coder and decoder should have the same initialization values. The decoder employs a buffer due to the fact that some delay is required in the decoding process. In theory the size of the buffer should be able to accommodate the maximum possible τ_k , i.e., $T_{\text{delay}} \geq \max(\tau_k) + T_0$.

With the knowledge of τ_k, e_k , and the appropriate initialization values of model M_k , the decoder calculates the values of the signal in the sampling points by the means of (3). We represent this information by the continuous-time signal $\hat{x}_s(t)$ or by the sequence $\hat{x}[k]$:

$$\begin{aligned} \hat{x}_s(t) &= \sum_k (\hat{x}_{ns_k} + e_k \Delta) \delta(t - ns_k T_0) \\ &\stackrel{(3)}{\approx} \sum_k x_a(ns_k T_0) \delta(t - ns_k T_0) \end{aligned} \quad (4a)$$

$$\hat{x}[k] = \hat{x}_{ns_k} + e_k \Delta \approx x_a(ns_k T_0) \quad (3) \quad (4b)$$

From the previous description a conclusion can be drawn that, although the analog input signal is sampled at a constant rate F_0 , the coder generates source symbols only in the sampling points $ns_k T_0$. The sampling period T_s is equal to the mean value of time intervals τ_k :

$$T_s = E[\tau_k] \quad F_s = \frac{1}{E[\tau_k]} \quad (5)$$

The possible values of e_k are by definition discrete (-1 or 1) and only one bit is required to represent them. On the other hand, the sampling at the input of the coder results in the discrete nature of τ_k . This sampling process is equivalent to the quantization process of PCM.

3. VARIOUS IMPLEMENTATIONS

A number of different versions of the TCM is presented next, depending on the various implementations of the predictive model and the interpolation process of the general TCM model.

3.1 Sample and Hold prediction

In the first implementation of TCM [6],[7] a Sample and Hold model is used as the predictive model. The operation of such a model can be represented by the following equation:

$$\hat{x}_n = \hat{x}[k-1] \quad ns_{k-1} < n \leq ns_k \quad (6)$$

At the decoder, the sequence $\hat{x}_a(nT_0)$ is created by the means of a linear interpolation [9] between the values of the sequence $\hat{x}[k]$:

$$\begin{aligned} \hat{x}_a(nT_0) &= \hat{x}[k-1] + \frac{\hat{x}[k] - \hat{x}[k-1]}{\tau_k} (n - ns_{k-1}) T_0 \\ ns_{k-1} &\leq n < ns_k \end{aligned} \quad (7)$$

Just as uniform time sampling techniques select amplitude values at equally spaced time points, the previous implementation of TCM seems to perform a uniform amplitude sampling, where time values are selected at equally spaced amplitude points [6],[7].

3.2 First Order Hold prediction

A second version [8] of the system is derived when the prediction model is implemented using a First Order Hold (FOH) [9] model. In this case, \hat{x}_n is given by the following equation set:

$$\begin{aligned} \hat{x}_n &= \hat{x}[k-1] + \alpha_k (n - ns_{k-1}) T_0 \quad ns_{k-1} < n \leq ns_k \\ \alpha_k &= \frac{\hat{x}[k-1] - \hat{x}[k-2]}{\tau_{k-1}} \end{aligned} \quad (8)$$

As we can see from (8), the signal derivative's mean value is calculated in the interval $[n_{k-2}, n_{k-1}]$. The derived value α_k is used for the calculation of \hat{x}_n in the interval $(ns_{k-1}, ns_k]$.

At the decoder, interpolation can be accomplished according to the following three techniques.

3.2.1 First Order Hold interpolation

In this case [8], we simply use the output of the FOH predictive model, provided that its values at the sampling points have been corrected by the means of (3). Consequently, $\hat{x}_a(nT_o)$ will be:

$$\begin{aligned} \hat{x}_a(nT_o) &= \hat{x}[k-1] + \alpha_k(n - ns_{k-1})T_o \quad ns_{k-1} \leq n < ns_k \\ \alpha_k &= \frac{\hat{x}[k-1] - \hat{x}[k-2]}{\tau_{k-1}} \end{aligned} \quad (9)$$

3.2.2 Linear interpolation

In this case [8], the sequence $\hat{x}_a(nT_o)$ is produced by the means of a linear interpolation [9] between the values of $\hat{x}[k]$. For the interpolation we use the same interpolation method that was employed in 3.1. The values of $\hat{x}_a(nT_o)$ are given by (7).

3.2.3 Sinc Pulses interpolation

This third interpolation technique involves sinc pulses. For each interval $[ns_{k-1}, ns_k]$, the values of the sequence $\hat{x}_a(nT_o)$ are produced as a superposition of three sinc pulses, according to the following equation:

$$\begin{aligned} x_a(nT_o) &= \hat{x}[k-1] \text{sinc}(nT_o - ns_{k-1}T_o) \\ &+ \frac{\hat{x}[k-1] + \hat{x}_{ns_k}}{2} \text{sinc}\left(nT_o - (ns_{k-1}T_o + \frac{\tau_k}{2})\right) \\ &+ \hat{x}[k] \text{sinc}(nT_o - ns_kT_o) \end{aligned} \quad ns_{k-1} \leq n < ns_k \quad (10)$$

4. SYSTEM PERFORMANCE AND SIMULATION RESULTS

The versions of the TCM system, which are described in 3.2, have been simulated by the means of the Matlab program language. Their function has been tested using speech signals, music signals and band-limited white Gaussian noise as the system inputs.

In order to estimate a measure for the performance of TCM versions presented in 3.2, let us consider the version which is described in 3.2.1

In every interval of the form $[ns_{k-1}, ns_k]$, the difference $\tilde{x}_n = x_a(nT_o) - \hat{x}_a(nT_o)$ represents the error which occurs during the reconstruction of the signal samples. Obviously, due to (4b), (9), as the value of n approaches ns_{k-1} , the error

approaches 0. On the other hand, when n approaches ns_k , the absolute value of error approaches the value of threshold Δ , but never exceeds it. Based on these observations we can consider that the random variable \tilde{X} , which denotes the error in each reconstructed sample, has a uniform pdf:

$$f(\tilde{x}) = \frac{1}{2\Delta} \quad -\Delta \leq \tilde{x} \leq \Delta \quad (11)$$

If random variable X stands for the value of $x_a(nT_o)$ for any value of n , the SNR is given by the following equation:

$$\text{SNR} = \frac{E[X^2]}{E[\tilde{X}^2]} = \frac{E[X^2]}{\frac{1}{3}\Delta^2} \quad (12)$$

Let us define N as the ratio of the peak-to-peak signal value to the threshold value Δ :

$$N \equiv \frac{x_{\max} - x_{\min}}{\Delta} \quad (13)$$

Then, equations (12) and (13) result in

$$\begin{aligned} \text{SNR}|_{\text{dB}} &= 10 \log_{10} \left(3N^2 E \left[\frac{X^2}{(x_{\max} - x_{\min})^2} \right] \right) \\ &= 10 \log_{10} (3N^2 E[\tilde{X}^2]) \end{aligned} \quad (14)$$

where the normalized X is denoted by \tilde{X} .

Figures 3,4,5 summarize the simulation results. Figure 3 illustrates the relation between N and the relative reduction in the sampling rate. F_N denotes the Nyquist rate of the input signal and F_s the sampling rate achieved by TCM. Figure 4 illustrates the dependence of the SNR on the relative reduction of the sampling rate.

By following the interpolator of Fig. 2 with a low-pass filter that has a sharp cutoff beyond the frequency $F_N/2$, the SNR is improved, especially in cases 3.2.1, 3.2.3. Figure 4 illustrates this improvement. Furthermore, we observe that our estimated SNR does not approach the simulation results. This is due to the fact that (11) is not valid in this case.

The simulation results indicate that (3) is valid if $F_o/F_N \geq 20$. The curves of figures 3,4,5 correspond to $F_o/F_N = 40$.

To convert the output of the coder into a bit stream we employ a simple coding scheme, where each pair (e_k, τ_k) is represented by a fixed length code-word. The simulations prove that if $20 \leq F_o/F_N \leq 40$ then $6 \leq R \leq 8$, where R denotes the mean code-word length per pair (e_k, τ_k) . However, a significant reduction in the value of R can be achieved by the use of entropy coding techniques.

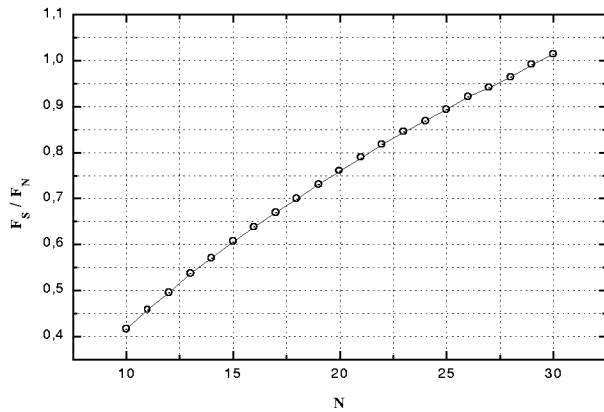


Figure 3 Normalized mean sampling rate vs. N.

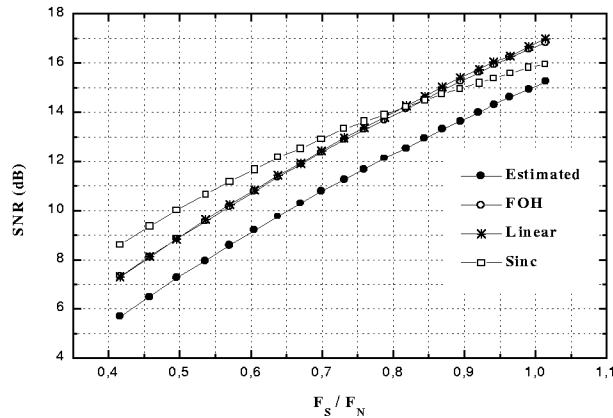


Figure 4 SNR Performance without using low-pass filter.

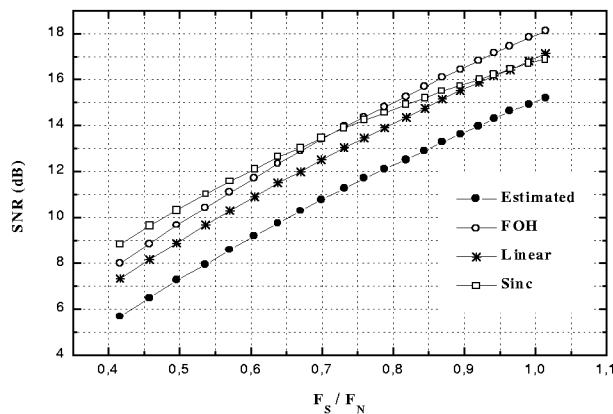


Figure 5 SNR Performance using low-pass filter.

5. CONCLUSIONS

In this paper, we introduced a theoretical model for Time Code Modulation, a non-uniform sampling technique, in which information is extracted from the input waveform only when necessary.

Previous versions of TCM arise as different implementations of this general model. Furthermore, a new version of TCM is introduced, by the means of a new interpolator at the decoder, based on the superposition of sinc pulses. In addition, a theoretical measure for the expected systems performance was estimated.

6. REFERENCES

- [1] Abdul J. Jerri, "The Shannon Sampling Theorem—Its Various Extensions and Applications: A Tutorial Review", *Proc. of the IEEE*, Vol. 65, No. 11, pp. 1565-1598, November 1977.
- [2] N.S. Jayant, "Digital Coding of Speech Waveforms: PCM, DPCM, and DM Quantizers", *Proc. of the IEEE*, Vol. 62, No 5, pp. 611-632, May 1974.
- [3] N.S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, Englewood Cliffs: Prentice-Hall, NJ, 1984.
- [4] A.S. Spanias, "Speech Coding: A Tutorial Review". *Proc. of the IEEE*, Vol. 82, No 10, pp. 1539-1582, October 1994.
- [5] Jae Yeol Rheem, BeomHun Kim, SoulGuil Ann, "A nonuniform sampling method of speech signal and its application to speech coding", *Signal Processing* vol. 41, No 1, pp. 43-48, January 1995.
- [6] J.Foster, T-K. Wang, "Waveform Coding using Time Code Multiplexing", *Conf. Proc.- IEEE Southeastcon*, vol. 3, pp. 1054-1055, April 1990.
- [7] J.Foster, T-K. Wang, "Speech Coding using Time Code Modulation", *Conf. Proc.- IEEE Southeastcon*, vol. 2, pp. 861-863, April 1991.
- [8] V. Stylianakis and D. Kalogiros "An Adaptive Method for Waveform Time Code Modulation", *Proc. Conv.104th Aud. Eng. Soc.*, preprint #4674, May 1998.
- [9] John P. Proakis, Dimitris Manolakis. *Digital Signal Processing – Principles, Algorithms, and Applications*, Third Edition. Prentice Hall 1996.