ENHANCE BIT SYNCHRONIZER BIT ERROR PERFORMANCE WITH A SINGLE ROM

Calvin L. James  
Data Systems Group  
Bendix Field Engineering Corporation  
One Bendix Road  
Columbia, Maryland 21045-1897

ABSTRACT

Although prefiltering prevents the aliasing phenomenon with discrete signal processing, degradation in bit error performance results even when the prefilter implementation is ideal. Degradation occurs when decisions are based on statistics derived from correlated samples, processed by a sample mean estimator. i.e., a discrete linear filter. However, an orthonormal transformation can be employed to eliminate prefiltered sample statistical dependencies, thus permitting the sample mean estimator to provide near optimum performance. This paper will present mathematical justification for elements which adversely affect the bit synchronizer’s decision process and suggest an orthonormal transform alternative. The suggested transform can be implemented in most digital bit synchronizer designs with the addition of a Read Only Memory (ROM).

INTRODUCTION

The degradation in optimum performance of digital bit synchronizer designs results from two central processes: tracking variance, and bit decision or estimation. Here we only consider enhancing the estimation process when the synchronizer processes non-white gaussian samples, while assuming perfect synchronization to the input bit or symbol sequence. Although the original synchronizer input is characterized as an Additive White Gaussian Noise (AWGN) channel, the added prefilter requirement dictated by the sampling theorem introduces non-white samples. The samples are thus correlated by virtue of the bandlimiting process. (See Figure 1) In what follows, we will examine the elements which degrade the performance of the discrete linear filter, and develop a maximum likelihood decision rule that provides optimum performance in the presence of non-white gaussian noise. Finally, we suggest a simple implementation of the proposed decision rule which utilizes a single ROM.
PROBABILITY OF ERROR

The preceding paragraph introduces a connection between the sampling theorem and adjacent sample correlation. In the sequel, we will attempt to exploit this relationship so that we might better understand how the sampling rate influences the bit decision process.

In order to satisfy the requirements of the sampling theorem, a sampling rate greater than or equal to twice the highest frequency component encountered by the sampling process must be chosen. Viewing the autocorrelation function of our prefilter, we note that the minimum correlation among samples is attained at the minimum sample rate. (See Figure 2) When the covariance among samples is zero, i.e., the autocorrelation value is zero, these samples are considered independent random variables, e.g., a white noise process, where the autocorrelation is represented by the Dirac delta function. Although independence implies the covariance among samples is zero, the converse is not necessarily true. However, since the original distribution is gaussian, it can be shown that the converse does indeed hold.
Let us assume for a given decision process that the observation interval $T$ is much larger than $1/2B$, where $B$ is the bandwidth of an ideal low pass filter. The bandwidth is made very large to minimize the effects of the filter’s response time. Additionally, each sample is quantized to a sufficient number of bits to insure the analog-to-digital, and sample mean estimation processes introduce negligible error. Under the stated assumptions, the probability of error for this estimation process is

$$P \left( \frac{\bar{X} - V}{\sigma / \sqrt{m}} > \frac{V}{\sigma / \sqrt{m}} \right) = P \left( \frac{\bar{X} - V}{\sigma / \sqrt{m}} > \frac{V^2 TB}{\eta B} = \sqrt{\frac{2E_s}{\eta}} \right)$$

where $\bar{X} = \frac{x_1 + \cdots + x_m}{m}$, $\text{Var}(\bar{X}) = \frac{\sigma^2}{m} = \frac{\eta B}{m}$

Here $m = 2TB$ is chosen as the maximum number of independent equally distributed samples taken during the observation interval $T$. Since the minimum correlation among samples occurs at the minimum sample rate, the sampling frequency is chosen as $f = 2B$. We also note that the probability of error from equation (1) is identical to that obtained when the decision process is implemented with an analog integrate and dump filter. However, when a single pole low pass prefilter is utilized, the probability of error becomes
Here we note a degradation in signal-to-noise performance, $E_s/N$, of $[1.9 + 10 \log G(m)]$ decibel (db). In the derivation of equation (2), the sample mean, $VG(m)$, is a function of $m$. Which, in the limit for large $m$, converges to one. This implies the prefilter bandwidth $B$, increases with $m$ for a given observation interval $T$, provided the samples are to remain nearly independent. Although the autocorrelation value is not identically zero, the term “nearly independent” is used here to imply that the results obtained by assuming an autocorrelation value of zero introduces negligible error. However, for small $m$, the parameter $G(m)$ degrades the single-to-noise performance by reducing the average attainable signal power. In the above discussion, $m$ was chosen to be the maximum number of independent samples. In the following paragraphs, we will investigate the case where $m$ is chosen to improve the sample mean estimate, while restricting the bandwidth parameter $B$.

**THE MAXIMUM LIKELIHOOD DECISION RULE**

Although the estimated average value of the impinging signal is improved by increasing the number of samples per observation interval, statistically dependent samples result when the prefilter bandwidth is restricted. Statistical dependence, influences our choice of an optimum decision rule from the decision space. The derivation of the maximum likelihood decision rule best explains this influence.

If decisions are made using a rule which decides $V$ in favor of $V*$ based on the likelihood ratio $L(x)$, i.e., the ratio of the logarithm of probability density functions, this rule can be considered an optimum choice if it minimizes the average risk. Risk here implies bit or symbol errors. “Bayes”, with knowledge of “a priori” (or prior) probability distribution i.e., the probability of transmitting $V$ is $P(V)$, is such a rule. If we assign $V$ and $V*$ equal probability, i.e., $P(V) = P(V*) = 1/2$, then the likelihood ratio can be expressed as

$$L(x) = \sum_{i=1}^{m} \sigma^{-1}_{\alpha_i} x_i y_{i'} + 2 \sum_{i<j} \sigma^{-1}_{\alpha_i} x_i y_{i'} x_{j'} > 0$$

In the above expression, antipodal signalling is assumed. i.e., $V^* = -V$. The resulting maximum likelihood, or Bayes decision rule, $H(x)$ is defined as
The interpretation of \( H(x) \) is as follows: decide \( V \) was transmitted if \( L(x) \) is greater than zero, otherwise choose \( V^* \). When the noise is white, the second summation in equation (3) is identically zero, since the covariance among independent samples is zero. Hence, the maximum likelihood ratio reduces to

\[
L(x) = \sigma^{-2} \sum_{i=1}^{m} x_i v_i > 0 \quad \text{where} \quad \text{Var}(x_i) = \sigma_{kk} = \sigma^2
\]

From the preceding discussion, we observe when processing samples in the presence of non-white noise, decisions based on the rule described by equation (5) are inferior to those utilizing (3). This fact results from the loss of power or information provided by the second summation in (3). In the next, section, we will describe a process which permits the correlated samples from the prefilter to be processed by a linear discrete filter process similar to that described by (5).

ORTHONORMAL TRANSFORMATION

The maximum likelihood decision rule, where \( L(x) \) defined in (3) is used by (4), becomes quite difficult to implement due to the number of product terms which must be accumulated for each sample. The first summation of the likelihood ratio of (3) represents the accumulated products formed by the prefiltered samples, and the elements along the main diagonal of the inverted covariance matrix. While the accumulated products formed with the off diagonal elements are represented by the second summation. Reducing the number of computations required by the maximum likelihood decision process, provides the motivation for an orthonormal transformation. However, our first task is to generate the covariance matrix of the desired prefilter. This is done by noting the relationship between covariance and autocorrelation.

Given a signal at the output of our prefilter of the form \( x(t) = v(t) + n(t) \), the covariance is defined by the following relationship

\[
\sigma_{yy} = E((x_i - E(x_i))(x_j - E(x_j))) = E(x_i x_j - x_i E(x_j) - x_j E(x_i) + E(x_i)E(x_j))
\]

\[
= E(x_i x_j) - E(x_i)E(x_j) \quad \text{But} \quad E(x_i) = V
\]
Therefore

\( (6) \quad \sigma_y = E(x_i x_j) - V^2 \)

On the other hand, the autocorrelation is defined as

\( (7) \quad E(x_i x_j) = E((E(x_i + n_j)(E(x_j) + n_j)) = V^2 + E(n_i n_j) \)

Here we use the fact that the noise has a mean value of zero i.e., \( E(n_i) = 0 \). Combining the above results yield

\( (8) \quad \sigma_y = E(x_i x_j) - V^2 = V^2 + E(n_i n_j) - V^2 = E(n_i n_j) \)

But \( E(n_i n_j) \) is just the autocorrelation of the non-white zero means noise samples from the output of our prefilter. The sampled autocorrelation function \( R(k) \), can be derived from the continuous case by computing the inverse Fourier transform of the power spectral density of our desired prefilter, while assuming a white noise input. Then for the continuous variable \( t \) in \( R(t) \), make the substitution, \( t = kT/m \), where \( E(n_i n_j) = R(k) \), and \( k = (i-j) \) is contained in the set, \( \{0,1,\ldots,m-1\} \). Note, due to the symmetric nature of the autocorrelation function, \( R(-t) = R(t) \), the domain of \( k \) is represented as a subset of the positive integers. Now the covariance matrix can be formed with each row-column entry corresponding to \( E(n_i n_j) \). After generating the covariance matrix, we observe that its structure has the form of a real symmetric matrix. From matrix algebra we note the following:

If \( S \) is a real symmetric matrix, there exists an orthogonal matrix \( A \) such that \( A^T S A = D \) is a diagonal matrix. Define \( y \) such that \( y = A^T x \). Then,

\( (9) \quad x = (A^T)^{-1} \quad y = (A^T)^T y = A y \)

Here we used the fact that \( A^T = A^{-1} \), since \( A \) is orthogonal. Then it follows that

\( (10) \quad x^T S x = (A y)^T S (A y) = y^T (A^T S A) y = y^T D y \)

It is interesting to note, that the square magnitude of the sample vector \( x \) is invariant under a orthogonal transformation. This can be shown as follows:

\( (11) \quad |x|^2 = x^T x = (Ay)^T (Ay) = y^T A^T A y = y^T y = |y|^2 \)

Again we use the fact that \( A^{-1} = A^T \) as noted in (9). Using (10), it can be shown that the original samples \( x_i \), can be transformed by the matrix \( A^T \), such that the resulting maximum
likelihood decision rule can be implemented with a linear discrete filter process similar to that described by (5).

**IMPLEMENTATION**

In the preceding paragraphs, material presented permitted the diagonalization of the real symmetric covariance matrix $S$, utilizing the orthogonal matrix $A$. Since our problem involves the inverted covariance matrix, we state the following relationships

(a) Given $S$ is a real symmetric matrix, i.e., $S = S'$. This implies, $S^{-1}S' = I$. Thus, $S^{-1} = (S')^{-1} = (S^{-1})'$, i.e., the inverted covariance matrix is symmetric also.

(b) Given $D$ is a diagonal matrix, then $D^{-1}$ exists, and $D^{-1}=(A'SA)^{-1} = A^{-1}S^{-1}(A')^{-1} = A'S^{-1}(A'^{-1}) = A'^{-1}S^{-1}A$

Here we find, if the entries of the diagonal matrix $D$ are $d_{ij}$, when $i=j$, otherwise zero, then the entries of $D^{-1}$ are $1/d_{ij}$, when $i=j$. Additionally, the orthogonal matrix $A$ that diagonalized $S$, also diagonalizes $S^{-1}$. Therefore equations (9) and (10) hold when substituting $S^{-1}$ and $D^{-1}$.

Using equation (3), and the fact that under our orthogonal transform the summation representing the off diagonal products equals zero, simplified the stated likelihood ratio of (3) to

\[
L(y) = \sum_{i=1}^{n} \lambda_i y_i e_i > 0 \quad \text{But} \quad y_i e_i = \left( \sum_{j=1}^{n} \alpha_y x_j \right) \left( \sum_{k=1}^{n} \alpha_k v_k \right) \\
\text{Since} \quad z_i = \sum_{j=1}^{m} \alpha_y q_j 
\]

Here we let $E(y_i)=e_i$, and $z=A'q$. Therefore,

\[
L(x) = \sum_{j=1}^{n} \sum_{i=1}^{m} \sum_{k=1}^{n} \lambda_i \alpha_y \alpha_k x_j v_k = \sum_{j=1}^{n} x_j w_j > 0
\]
where

\[ w_j = \sum_{l=1}^{m} \sum_{k=1}^{m} \lambda_{lk} \alpha_{lj} \alpha_{lk} v_k \]

*note, \( \alpha \) are elements of \( A^t \), while \( \lambda \) are elements of \( D^{-1} \)*

Equation (12) represents a discrete linear filter which processes the products formed with the contaminated samples, \( x_j \) and the “weighting function”, \( w_j \). The weighting function, is a linear function of known constants and its value is independent of the sample \( x_j \). The only remaining procedures which must be developed in order to implement the above outlined maximum likelihood decision process, are the generation of the orthogonal matrix \( A \), and the diagonal matrix \( D^{-1} \). The elements of matrices \( A \) and \( D^{-1} \) can be obtained by first solving the characteristic equation of \( S^{-1} \) for the associated eigenvalues. Each of the \( m \) distinct eigenvalues, form the entries along the main diagonal of \( D^{-1} \). While the corresponding eigenvectors form the column vectors of the orthogonal matrix \( A \).

The orthonormal transform outlined in the preceding paragraph, can be performed utilizing a look-up table approach implemented in a single ROM. Here, the address space of the ROM consist of the quantized sample value \( x_j \), obtained from the analog-to digital process, and the sample index \( j \), derived from the synchronizer’s tracking loop. The contents specified by the address space, i.e., the ROM output, form the product of the quantized prefilter sample value \( x_j \), and the weighting function \( w_j \). This product is then accumulated by the discrete linear filter implementation of equation (12).
REFERENCES


