VITERBI ALGORITHM RECEIVERS FOR NONLINEAR SATELLITE CHANNELS*

T.E. Stern
Dept. of Electrical Engineering &
Computer Science
Columbia University
New York, N.Y. 10027

The problem of maximum likelihood estimation of data sequences transmitted over nonlinear channels is considered. The nonlinearity treated is of the type occurring typically in a satellite repeater. It is shown that the optimal receiver consists of a bank of linear filters whose sampled outputs are processed using the Viterbi Algorithm. Several alternate realizations are discussed as well as some approximations that lead to a reduction in computational load.

1. Introduction

A serious transmission impairment in satellite data communication systems, and one that is particularly difficult to deal with, is the nonlinear transmission characteristic of the traveling-wave tube (TWT) amplifier in the satellite repeater. Considerable effort has been devoted to analyzing the effect of this nonlinearity on the signal characteristics and predicting the degradation produced by it when a conventional receiver is used to detect the data. However, little attention appears to have been directed toward the design of an optimal receiver for this channel. The purpose of this paper is to propose a number of receiver structures for performing maximum likelihood estimation of data sequences carried by signals subject to the type of nonlinear distortion typical of satellite channels. To simplify the exposition, our analysis will treat the case of baseband signals. However, by replacing the real signals used herein by complex signals, the work can be extrapolated directly to the case of modulated (e.g. FSK) systems.

Maximum likelihood sequence estimation (MLSE) in linear channels with additive noise and intersymbol interference has been described in the literature. A special case of a nonlinear channel with binary input has also been treated. In each case, the unique aspect of the proposed receiver is the processing of linear components between the TWT and the receiver. To simplify the exposition, our analysis will treat the case of baseband signals. However, by replacing the real signals used herein by complex signals, the work can be extrapolated directly to the case of modulated (e.g. FSK) systems.

Maximum likelihood sequence estimation (MLSE) in linear channels with additive noise and intersymbol interference has been described in the literature. A special case of a nonlinear channel with binary input has also been treated. In each case, the unique aspect of the proposed receiver is the processing of linear components between the TWT and the receiver. To simplify the exposition, our analysis will treat the case of baseband signals. However, by replacing the real signals used herein by complex signals, the work can be extrapolated directly to the case of modulated (e.g. FSK) systems.

In Section 2 we present the channel model, as well as several equivalent representations of the channel output signal. Based on these representations we describe in Section 3, some receiver structures for optimal estimation of the transmitted data, based on the maximum likelihood criterion. Where possible we adopt the terminology and notation of reference 4 in order to exhibit the relation between the present work and that of Forney. The remainder of the paper is devoted to a discussion of computational considerations and the implications of this work for practical receivers.

2. Channel Model; Signal Representation

The channel model being considered is shown in Fig.1. A data source emits data symbols xk at regular intervals of T sec., drawn from an alphabet of m symbols. The signal carrying the data

\[ v(t) = \sum_{k} x_{k} h(t - kT) \]  

appears at the input to a zero-memory nonlinearity f(·). In a satellite channel, f would represent the nonlinearity of the TWT, and h would represent the impulse response of the combination of all linear elements between the source and the input to the TWT. (We note in passing, that a complex f operating on a complex signal v would be appropriate to represent the effect of the phase and amplitude nonlinearities of the TWT on a modulated signal.) The output u of the nonlinearity is convolved with the function g yielding a signal s to which white Gaussian zero mean noise n is added, so that the signal appearing at the input to the receiver is

\[ r = s + n = (u \circ g) + n \]  

The function g represents the impulse response of all linear components between the TWT and the receiver. Both h and g are assumed to be of finite duration. The most serious limitation of this model is that noise is only assumed to be present in the downlink. We justify this approximation on two grounds: (1) In practice the signal-to-noise ratio on the uplink of a satellite channel is usually much greater than that on the downlink, and (2) noise at the input to the nonlinearity would vastly complicate the problem.

We now develop three equivalent representations of the signal s, each of which will suggest a particular receiver structure.

Representation A.

In this and subsequent developments it will be convenient to resolve various functions into “chips”, that is, functions that are non-zero only over the interval [0, T).

\[ h_{i}(t) = \sum_{i=0}^{\infty} h_{i}(t-iT) \]  

where hi is the i-th chip of h, and h is time-limited to V+1 chips. (This subscript notation will be used throughout to indicate a chip decomposition of any time function.) Decomposing v in a similar fashion we have

\[ v(t) = \sum_{i} v_{i}(t-iT) \]  

† When limits are omitted from summations the index runs from -∞ to +∞.

‡ The symbol \( \circ \) denotes convolution.

* This work was partially supported by the National Science Foundation under Grant NSF-ENG 74-17152.

† This work was partially supported by the National Science Foundation under Grant NSF-ENG 74-17152.
where from (1), (3), (4) we deduce
\[ \varphi(t) = \sum_{i=0}^{n-1} h_i(t) x_{i-1} \]  
(5)
and define
\[ \psi(t) = \sum_{i=0}^{n-1} h_i(t) x_{i-1} \]
(6)
and which
\[ \sum_{i=0}^{n-1} x_{i-1} = \sum_{i=0}^{n-1} x_{i-1} \]
(7)
Again, \( u(t) \) can be expressed in terms of its chips
\[ u(t) = \sum_{i=0}^{n-1} h_i(t) \]
(8)
where, because of the zero-memory property of \( f \),
\[ f(t) = f(t-kT) \]
(9)
Finally, we can substitute (14) into (9) and obtain
\[ s(t) = \sum_{k=0}^{n-1} g(t) U(t-kT, x_k) dT \]
(10)
and
\[ \sum_{k=0}^{n-1} g(t) U(t-kT, x_k) dT \]
(11)
where \( U(t, x_k) \) is an \( n \times 1 \) vector defined by
\[ U(t, x_k) = U(t-kT) \]
(12)
and
\[ \sum_{k=0}^{n-1} g(t) U(t-kT, x_k) dT \]
(13)
which is the representation we are seeking: an expansion of \( s \) in terms of \( \psi \) where the coefficient vectors \( \psi_k \) are functions of the data vectors \( x_k \).

Representation B.
As will be shown in Section 3, the fact that the \( \psi_k \) generally have a duration longer than one chip causes certain complications in receiver implementation. These difficulties can be circumvented by representing \( s \) in terms of a larger set of functions, each of one chip duration. To determine this representation we resolve \( g \) into its chips,
\[ g(t) = \sum_{i=0}^{n-1} g_i(t-kT) \]
(14)
Substituting (14) into (9) we find
\[ s(t) = \sum_{k=0}^{n-1} g_k(t-kT) U(t-kT, x_k) dT \]
(15)
which can be written in terms of chips as
\[ s(t) = \sum_{k=0}^{n-1} g_k(t-kT) \]
(16)
A new orthonormal basis
\[ \tilde{\psi} = (\tilde{\psi}_1, \tilde{\psi}_2, \ldots, \tilde{\psi}_n) \]
(17)
can now be chosen (each \( \tilde{\psi}_1 \) of duration one chip), which spans the space containing the \( S \)'s. Since each \( S(t, x_k) \) is a sum of terms of the form \( g_k \tilde{\psi}_k \), the dimension of the required signal space is bounded above by
\[ N_s \leq 2(n+1) \]
(18)
Now \( s \) can be written in the form
\[ s(t) = \sum_{k=0}^{n-1} g_k(t-kT) \]
(19)
where \( \tilde{\psi}_k \) is an \( n \times 1 \) vector defined by
\[ \tilde{\psi}_k = \int_0^T g_k(t) \theta(t) dt \]
(20)
Substituting (20) into (9) we obtain
\[ s(t) = \sum_{k=0}^{n-1} g_k(t-kT) \]
(21)
which is the representation we are seeking: an expansion of \( s \) in terms of \( \psi \) where the coefficient vectors \( \psi_k \) are functions of the data vectors \( x_k \).

Representation C.
Since the nonlinear effects will usually be small in practical cases, it is useful to consider another representation in which the nonlinear distortion is separated from the signal that would have been transmitted had the nonlinearity been absent. Thus, in this case, we resolve the signal into the form
\[ s(t) = \varphi(t) + \psi(t) \]
(22)
An indicated in Fig. 2, \( \varphi \) represents the signal that would have been transmitted if \( f \) were replaced by a unity gain element, and \( \psi \) represents the residual distortion due to the nonlinearity. As before we have the option of representing \( \varphi \) and \( \psi \) in terms of either "long" or "short" orthogonal functions. To avoid needless
repetition, only the latter approach will be chosen here. Let
\[ u(t) = \sum_{i=0}^{\infty} u_i(t-iT) \]  
(19)
be the overall impulse response of the linear path in Fig.2, where \( w \) has been resolved into \( v+1 \) chips with \( v \) given in (17). Following our previous techniques, \( \nu \) and \( \mu \) can be written in terms of their respective chips:
\[ \nu(t) = \sum_{i=0}^{\infty} \nu_i(t-iT) \]
(20)
and
\[ \mu(t) = \sum_{i=0}^{\infty} \mu_i(t-iT) \]
(21)
with
\[ \nu_i(t) = \mathbb{E}[\nu(t,T_k)] - V(t_{k-1}) \]
(22)
\( V \) is defined in (6).)

Now let us choose an orthonormal basis, \( \phi^V \) for the \( V \)'s and augment this set by additional functions \( \phi^P \) such that the set
\[ \tilde{\phi} = (\phi^V, \phi^P_1, \phi^P_2, \ldots, \phi^P_m) \]
(23)
is orthonormal and spans the space containing the \( V \)'s and \( P \)'s. It is easily seen that the dimension of the required space is bounded above by
\[ N_v+N_p = (v+1)+2(v+1)m \]
(24)
and the functions \( \phi^P \) are zero outside the interval \([0,T]\). The signal \( s \) can now be represented in terms of this set as
\[ s(t) = \sum_{k=0}^{N_v} \tilde{s}_k(t-kT) \]
(25)
where
\[ \tilde{s}_k(t) = \int_{0}^{T} \phi^P(t,T_k)dt = \mathbb{E}[\phi^P_k] \]
and only the first \( N_v \) components of \( \tilde{s}_k \) are non-zero.

The representations of (13), (18) and (24) are in forms that are directly applicable to techniques of MLSE using the Viterbi Algorithm. We consider some receiver structures based on these representations in the next section.

3. Receiver Structures

In order to perform MLSE, the first requirement of the receiver is that it extract from the received signal, \( r(t) \), a set of observations which constitute a sufficient statistic for estimating the data sequence. If the Viterbi Algorithm is to be used for processing these observations, the second requirement is that the sequence of observations presented to the VA processor contain uncorrelated noise components. For each of the representations developed above, we shall now present a receiver structure which has both of these attributes.

**Representation A.**

The structure of the receiver is shown in Fig.3. The block labelled \( \phi(-t) \) represents a bank of \( N_u \) filters in parallel, each excited by \( r(t) \) and matched to a component of \( \phi \). (Of course, a delay is required for physical realizability.) The filter outputs are sampled at times \( kT \), giving the vector sequence
\[ \mathbf{x}_k = \int_{-\infty}^{\infty} \phi(-kT) \mathbb{E}[r(t)]dt = \sum_{l=0}^{N_u} \mathbf{R}_{k-l} \mathbf{\phi}^P \]
(26)
where \( \mathbf{x}_k \) is an \( N_u \)-vector, and \( \mathbf{R}_{k-l} \) is the square matrix
\[ \mathbf{R}_{k-l} = \int_{-\infty}^{\infty} \phi(-kT) \mathbb{E}[r(t)] \mathbb{E}[r(t)]^T \]
(27)
with
\[ \mathbf{R}_{k-l} = 0 \quad \text{for} \quad |k-l| > v + 2 \]
(28)
and
\[ \mathbf{n}_k = \int_{-\infty}^{\infty} \mathbb{E}[n(t)] \mathbb{E}[n(t)]^T \]
(29)
Assuming that \( n \) is white with two-sided spectral density \( N_s \), the correlation matrix for \( \mathbf{n}_k \), the noise samples at the outputs of the filters, is given by
\[ \mathbf{R}_{k-l} = \mathbb{E}[(\mathbf{n}_k^T \mathbf{n}_l)] = N_s \mathbf{I} \]
(30)
Since the elements of \( \phi \) span the signal space, the sequence \( \{\mathbf{x}_k\} \) is a sufficient statistic for estimating \( \{\mathbf{n}_k\} \). However, \( \phi \) cannot be processed directly by the VA because, as indicated in (27) the filters \( \phi \) create a statistical dependence among the noise samples. To "undo" this effect we follow the procedure of Ref.5, inserting a (matrix) whitening filter (MWF) between the sampler and the VA processor. (See Fig.3.) The form of the MWF is determined most easily using the D-transform, defined as
\[ \mathbb{E}[\mathbf{n}_k] = N_s \mathbf{I} \]
(31)
for the sequence \( \{\mathbf{n}_k\} \), and analogously for the other vector and matrix sequences appearing below. Eq.(25) transforms to
\[ \mathbb{E}[\mathbf{x}_k] = N_s \mathbf{I} \]
(32)
where the spectral matrix of the noise' (the D-transform of \( \phi \)) is \( N_s \mathbf{I} \).

It can be shown that, except for certain degenerate cases, \( R(D^{-1}) \) can be factored into the form
\[ R(D^{-1}) = F(D)F^{-1}(D^{-1}) \]
(33)
where \( F(D^{-1}) \) and \( F^{-1}(D^{-1}) \) are square matrices whose elements represent stable and non-anticipative discrete time transfer functions.

From (26a) and (29) it can be deduced that
\[ F(D) = \frac{1}{N_u} \sum_{l=0}^{N_u} \mathbf{R}_{k-l} \]
(34)
The factorization of \( R(D^{-1}) \) given in (29) indicates that the matrix transfer function for the MWF should be \( F^{-1}(D^{-1}) \); i.e.,
\[ \mathbf{n}_k = F^{-1}(D^{-1}) \mathbf{x}_k \]
(35)
Using (28), (29), (30) and inverse transforming (31) we find
\[ x_k = y_k + \delta_k \]
where
\[ y_k = \sum_{j=0}^{v+1} f_j z_{k-j} \]
(32)
and \( z_{k-j} \) is a sequence of noise vectors in which each component of each \( z_{k-j} \) is an independent zero mean Gaussian random variable of variance \( \sigma_0 \). Since the whitening filter is invertible, it is information-lossless, and therefore the output \( x_k \) remains a sufficient statistic for the estimation problem.

Considering the source as a finite state machine with state defined by
\[ s_k = (x_{k-1}, x_{k-2}, \ldots, x_{k-v}) \]
(33)
and noting from (32), (12) and (7) that \( x_k \) is a function of \( x_{k-1}, \ldots, x_{k-v} \) we may write
\[ x_k = \mathcal{E}(s_k, \sigma_{k+1}) \]
(34)
Each \( s_k \) may assume one of \( m^v \) possible values, representing the number of distinct combinations of the \( x_i \) in (33).

Clearly, estimating the sequence \( [s_k] \) is equivalent to finding \( \hat{S}_k \), the estimate of \( x_k \). The Viterbi Algorithm is an efficient recursive method for obtaining maximum likelihood estimates of \( [s_k] \) given the observations \( [z_k] \). Since the VA has been treated thoroughly elsewhere, we shall discuss it no further here. It is sufficient for our purposes to note that the vector version of the VA requires the computation of quantities of the form
\[ \gamma_k(s_{k-1}, s_j) = \|z_k - v(s_{k-1}, s_j)\|^2 \]
(34)
for each \( k \), where \( s_{k-1} \) ranges over the \( m^v \) possible states, and \( s_j \) ranges over the \( m \) states that are the possible successors to each \( s_{k-1} \). The \( v \)'s are the essential ingredients in the construction of the log likelihood functions associated with the state sequences that could have produced the observations \( z_k \). Once the \( v \)'s have been computed, the remaining computation for MLSE proceeds in a straightforward fashion.

Representation C.

Considering the representation of (18) and proceeding as above, we obtain a sufficient statistic using a bank of \( N_s \) filters, each matched to one of the components of \( v \). (Fig.4). Since \( v \) has duration \( T \), the sampled filter output in this case takes the particularly simple form
\[ \tilde{z}_k = \delta_k + \delta_k \]
(35)
where the \( \delta_k \) are now uncorrelated. Therefore, no whitening filter is necessary. In this case, \( \delta_k \) is a function of \( \tilde{x}_k \), defined in (17), so that, using the definition of \( \sigma \) in (33) we may write
\[ \delta_k = \mathcal{G}(s_k, \sigma_{k+1}) \]
(36)
\[ \|\delta_k\|^2 = \sum_{i=1}^{w} \sum_{j=1}^{w} \delta_{k,i} \delta_{k,j} \]

\( w \) denotes the sum of the squares of the components of \( \tilde{x}_k \).

The expression (34) required in the VA therefore becomes
\[ \gamma_k(s_{k-1}, s_j) = \|\tilde{z}_k - v(s_{k-1}, s_j)\|^2 \]

Representation C.

A receiver based on the representation (24) has essentially the same form as that of Fig.4, except that in this case we use \( (N_v \times N_s) \) receiving filters matched to the components of \( \tilde{x}_k \). Again, no whitening filter is required. The sampled filter output is
\[ \tilde{x}_k = x_k + \delta_k + \delta_k \]
From (24), (17) and the definition of \( v \) we may write
\[ \tilde{x}_k = \mathcal{G}(s_k, \sigma_{k+1}) \]
so that (34) now takes the form
\[ \gamma_k(s_{k-1}, s_j) = \|\tilde{z}_k - v(s_{k-1}, s_j)\|^2 \]

4. Computational Considerations, Approximations

The three receiver structures proposed above have similar characteristics: a bank of linear filters followed by a VA processor. Their most important common feature, however, is the fact that each system produces an observation vector which is a function of \( v \) past inputs, where \( v \) represents the memory of the channel, and is independent of the nonlinearity. Thus \( m^v \) the number of states involved in VA computation for these systems is identical to the number that would be required for a linear system with the same memory. Since the VA computational load is proportional to the number of states, this is a comforting result.

Less comforting is the fact that the number of required linear filters is apparently prodigious. (The worst case appears in Eq.(23).) It must be recalled, however, that the figures presented on dimensionality of the signal space are only upper bounds. Generally certain symmetries will exist in the signal set and in the nonlinearity which will tend to reduce the dimensionality of the space. Furthermore, it is likely that a very good approximation to the optimal receiver can be achieved by using a relatively small set of matched filters, which represent the signal space in some approximate fashion. For example, using Representation C, it may be that the \( N_v \) filters required to represent the \( v \)'s would also provide an adequate representation of the \( p_i \)'s. This notion of approximation may be made more precise by recognizing that the downlink (represented by \( g \) must be at least approximately band-limited to some bandwith \( W \). In such a case the results of Landau and Pollak\(^9\) indicate that the dimensionality of the space of \( w \) will not be much greater than \( 2WT(v_g+2) \). Typically, \( W \) will be of the order of \( \frac{1}{T} \) so that if, for example, \( g \) has a chip duration, \( v_g+4 \), a set of ten filters should be ample, even though the upper bounds presented above suggest that much larger numbers might be necessary. To keep the number of required filters to a minimum, it appears that the option of "long" onthonormal signal sets (i.e., Representation A) is to be preferred over sets of one chip duration. A question then arises as to the practical realization of the filter-whitener combination. Clearly, one option is the realization indicated in Fig.3, in which \( N_s \) recursive digital filters are required to implement the MWF. Another possibility, which avoids having to realize \( N_s \) filters, would be to realize the combination of the matched filter bank and the MWF as a...
set of \( N_k \) continuous time filters followed by samplers. Note that just the numerators of the transfer functions \( F^{-1} (D^{-1}) \) can be combined with the matched filters to yield \( N_k \) continuous time finite impulse response filters. The denominators of the elements of \( F^{-1} (D^{-1}) \) are all identical and can be realized as \( N_k \) identical recursive digital filters following the samplers. Finally, it should be observed that a modification of the VA developed by Ungerboeck\(^{11}\) dispenses with the whitening filter altogether, in effect replacing it by operations performed within the modified VA itself. Since these additional operations can be performed once, with the results stored for use in real time computation, the modified algorithm is of the same order of complexity as the original.

5. Conclusions

It has been shown that maximum likelihood sequence estimation for a nonlinear data channel can be performed using a receiver consisting of a bank of linear filters followed by a VA processor. In a practical implementation the computational load in the linear part of the receiver might be one order of magnitude larger than that required for an equivalent linear channel, but the VA computation would require the same number of states as in the linear case.

It is worth noting that these results can be extended to the case of nonlinearities with memory, in which case the number of states would have to be augmented appropriately.

Investigations are currently in progress to assess the performance improvement achievable with the proposed receivers as compared with receivers designed for linear channels.

References


\[ r \rightarrow g \]

\[ s + 2 \]

\[ + v \]

\[ h \]

\[ x_k \]

\[ \text{(m-ary)} \]

\[ \text{SOURCE} \]

\[ \text{CHANCEL} \]

\[ \text{FIG. 1} \]
CHANNEL WITH NONLINEARITY EXTRACTED
FIG. 2

OPTIMAL RECEIVER (WITH MWF)
FIG. 3

OPTIMAL RECEIVER (WITHOUT MWF)
FIG. 4