INTERIM REPORT

September 15, 1969 - February 15, 1970

DIGITAL COMMUNICATIONS TECHNIQUES

for

National Aeronautics and Space Administration

under

NASA GRANT NGR-33-006-020

Prepared by

Raymond L. Pickholtz, Associate Professor

PIBEE 70-001

1970

DEPARTMENT OF ELECTRICAL ENGINEERING
POLYTECHNIC INSTITUTE OF BROOKLYN

Introduction

The work reported in this interim report represents a brief summary of three of the activities being carried out at the Polytechnic Institute of Brooklyn under NASA Grant NGR-33-006-020. The three areas are:

I. Study of Convolutional Codes

II. Coding for a Dispersive Channel

III. Recursive Signal Processing

We have obtained some new results for convolutional codes by extending some error bounding techniques due to Forney. In particular, we have developed a unified treatment of erasure and variable list decoding, which, we hope, will be useful in analyzing the performance of proposed practical algorithms for decoding using a list. The main objective of the continued research in this area is to develop a simple and practical algorithm for decoding convolutional codes which does not require as much memory as the viterbi algorithm for maximum likelihood decoding nor as much computation as the various sequential decoding schemes.

The second area attempts to solve an old problem via a different approach then usual - i.e. reducing the effects of intersymbol interference in digital communications by using error correcting codes rather than by linear operators using equalizers. The main idea is to use the dispersion in the channel to perform part of the coding operator. Standard decoding methods are used. Under certain conditions, several orders of magnitude improvement is possible.

The third area, on recursive signal processing, is a continuation of work initiated over a year ago. The main emphasis is on discrete, quantized processing of data. Several new algorithms were developed and preliminary simulation results indicate significant improvement over approaches derived from classical techniques when quantization effects are important.

The common thread of these three areas (and others not reported here) is the emphasis of discrete digital processing for digital communications. Details will be produced in the final report and in expected published papers and theses.

The following faculty and their students are being supported in this program.

Raymond L. Pickholtz, Associate Professor

Jack K. Wolf, Professor

Richard Haddad, Associate Professor

I. Study of Convolutional Codes

We have extended Forney's method for analyzing random tree codes, to allow investigation of a broader class of decoding algorithms. Specifically, considering maximum-likelihood decoding of a terminated tree code, Forney established the equivalence of a decoding error and the existence of a single unmerged span over which some incorrect word has greater likelihood than the correct one. We, then, actually loosen this bound by calculating the probability that some incorrect word has greater likelihood over one or more unmerged spans; the resulting bound is found to be asymptotically equal to Forney's. The advantage we gain is that we are now able to analyze algorithms where the event of an error cannot be equated quite so neatly with the occurrence of a single unmerged span with greater likelihood (viz., list decoding).

As an initial application of the method, we obtain, for tree codes, a unified treatment of Erasure and Variable List Size decoding, analogous to that developed, by Forney, for block codes. Referring to Forney's papers (and using his notation), where he obtains complementary exponents $E_i(R,T)$, $i = 1,2$ for block codes, we find analogous exponents $e_i(r,T)$, $i = 1,2$ for tree codes, given by:

$$e_i(r,T) = \inf_{\mu\epsilon(0,1)} \frac{E_i(\mu n,(1-\mu)T}{1-\mu} \quad ; \; i = 1,2$$

Further, analogous to the Feedback and List exponents, $E_f(R)$, $E_\ell(R)$, respectively, that Forney obtains as limiting cases of $E_i(R,T)$, $i = 1,2$, we find exponents $e_f(r)$, $e_\ell(r)$ given by:

$$e_{f(\ell)}(r) = \inf_{\mu\epsilon(0,1)} \frac{E_{f(\ell)}(\mu r)}{1-\mu}$$

as Feedback and List exponents for tree codes.

In the case of the Very-Noisy channel, we know $e_f(r)$ in closed form; $e_\ell(r)$ is known in closed form for $r < c/2$, but must be calculated by computer for higher rates.

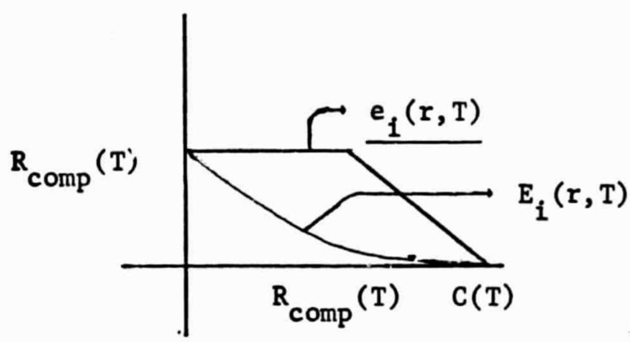
Along the way, we have had to establish certain properties of the exponents $E_i(R,T)$, $i = 1,2$ (viz., convexity in (R,T)). We have not yet characterized the properties of $e_i(r,T)$, $i = 1,2$, but we can lower bound them by relatively meaningful quantities as:

$$e_1(r,T) \geq \underline{e_1(r,T)} \triangleq \inf_{\mu \in (0,1)} \frac{E_1(\mu r, T)}{1-\mu}$$

$$e_2(r,T) \geq \underline{e_2(r,T)} \triangleq \underline{e_1(r,T)} + T$$

For a fixed value of T. we can show that the behavior of $\underline{e_i(r,T)}$ will have the form:

(Ex-very noisy channel)



where $R_{comp}(T)$, $C(T)$ may be calculated for the given value of T (the above result can also be seen quite clearly from Forney's sketch of the $e_i(r,T)$ surface).

II. Coding For A Dispersive Channel

The problem considered is that of reducing intersymbol interference caused by time dispersive channels. This is not a new problem, and much recent work has centered on the use of the Tapped Delay Line (TDL) equalizer. The conventional approach is to choose the no dispersion channel as the desired channel, and then to minimize some measure of the intersymbol interference.

The approach taken in this report recognizes the encoding properties of time dispersive channels. These channels process the transmitted data in much the same way as the generator of a cyclic algebraic code. Two methods of attack are taken. In the first method the code generator coefficients are used as the desired response for an otherwise conventional TDL equalizer. This method is termed the Coded Equalizer method. Transmission of k q-ary symbols through the channel in cascade with the coded equalizer results in a code word. Thus, error correction can be obtained with an ordinary algebraic decoder, and without transmission of parity symbols. In the second method the channel encoding is accepted without further processing by a TDL equalizer, and is subsequently decoded by a channel decoding matrix. This matrix is designed to minimize the additive noise variance subject to the constraint that intersymbol interference be eliminated. A modification of this method allows trading computation time for a limited amount of intersymbol interference.

Upper bounds on the probability of error are derived for both methods. These bounds are in the form of easily calculated error functions. The parameters of code block length, number of information symbols per block, and alphabet size appear explicitly.

Computer simulations confirm the derived bounds and show that under certain conditions orders of magnitude improvement in error rate can be obtained.

III. Recursive Signal Processing

Two different areas in optimal filtering of discrete-time data are under investigation. The first study deals with optimal state estimation based on noisy, quantized data, while the second considers bias effects in mismatched minimum variance polynomial filters.

(I) Nonlinear Recursive Estimation with Quantized Data.

The problem considered here is the optimal estimation of the state of a dynamic system based on quantized, noisy measurements. The results obtained are particularly applicable when the granularity is coarse and the usual quantization model (additive, independent, uniformly distributed noise) is inadequate.

The state of the system is generated by the vector difference equation

$$x(k) = f[x(k-1), k] + G(k-1)\, u(k-1), \quad k \geq 1 \tag{1}$$

and the noisy measurements before quantization are described by

$$y(j) = h[x(j), j] + v(j) \tag{2}$$

Upon quantization, this measurement is degraded, and the information at the quantizer output (upon which the state estimate is to be conditioned) is that

$$a(j), < y(j) \leq b(j); j = 1, 2, \ldots, k \tag{3}$$

In Eqs. (1), (2), (3) u(k) and v(j) are sequences of mutually independent gaussian, vector random variables with zero mean and covariances

$$E\{u(i)\, u'(j)\} = Q(i)\, \delta_{ij} \; ; \quad E\{v(j)\, v'(i)\} = R(i)\, \delta_{ij}$$

We also assume that p[x(o)], the initial state probability density is known; that each component of f(.,.) has continuous derivatives, and that p(x(o)) and h(.,.) have continuous second partials.

A direct approach to the determination of approximate algorithms for the recursive calculation of the mean and mode of x(k) conditioned on the quantized data is taken. An exact equation for the conditional mode estimate is derived which requires the solution of a nonlinear two-point boundary value problem. The latter is linearized to give an approximate recursive state estimation algorithm of the predictor-corrector form

$$\tilde{x}(i) = \bar{x}(i) - P(i)\, hx'[\bar{x}(i)]\, d[\bar{x}(i), R(i)] \tag{4}$$

where $\bar{x}(i) = f(\bar{x}(i-1))$ represents the predicted MLE

and

$$P(i) = \{1-P(i/i-1)\, L[\bar{x}(i)]\}^{-1}\, P(i/i-1) \tag{5}$$

$$P(i/i-1) = fx[\bar{x}(i)]\, P(i-1)\, fx'[\bar{x}(i)] + \Lambda(i-1)$$

In Eq. (5), hx, fx represent matrices of partials of h, and f respectively, $\Lambda(.) = G(.)\, Q(.)\, G'(.)$, P(o) is related to the initial state density; d is calculated from the known sequence of gaussian density functions, and $L[\bar{x}(i)]$ is the matrix obtained from the first partial of the product hx(.) d(.,.) evaluated at $\bar{x}(i)$.

An approximate conditional mean estimate is derived upon making the simplifying assumption that at each step the a priori density of x(i) given the priori measurements is gaussian. This approximate estimator also has a predictor-corrector form

$$\tilde{x}(i) = \tilde{x}(i/i-1) + W(i)\, d[\tilde{x}(i/i-1), Py(i)]$$

where

$$d[x(i/i-1), Py(i)] = Py^{-1}(i)\, \frac{\int_{a(i)}^{b(i)} [y(i)-h[x(i/i-1)]g\{y(i)-h(x(i/i-1), Py(i)]\}dy(i)}{\int_{a(i)}^{b(i)} g\{y(i)-h[x(i/i-1)], Py(i)\}dy(i)}$$

and

$$Py(i) = hx[x(i/i-1)]\, W(i)\, hx'[\quad + R(i)$$

$$W(i) = P(i/i-1)$$

and h_{jxx} is the matrix formed from the second partials of $h_j(.)$ and where $g(.,.)$ represents a gaussian density.

Monte Carlo simulation of these algorithms demonstrated the superiority of these over the Kalman filter in which the quantization error is approximated by additive measurement noise, at least for the examples considered (first-order and a second-order system) wherein the granularity is coarse. The improvement in the RMS error is by a factor of 10 for one example.

II Bias Effects in Mismatched Minimum Variance Polynomial Filters

A steady-state dynamic error, or bias, results whenever the model of the signal process (and the filter designed on the basis of that model) is an inadequate representation of the actual filter input. Suppose a finite-memory polynomial filter is correctly designed to smooth polynomial inputs of degree M. Bias errors result when this filter is excited by non-polynomial inputs, or by polynomial inputs of degree L > M. Our study is devoted to the latter aspect, which we view as resulting from the fact that the state vector of the input is (L+1) dimensional whereas the filter state is only (M+1) dimensional.

Under matched conditions, the filter algorithm for a p-unit prediction interval is

$$\tilde{\underline{\xi}}(n + p(n)) = H'(p)\, \underline{f}$$

where $\underline{f}' = [f(n)\ f(n-1)\ \ldots\ f(n-\ )]$ is the noisy data sequence, and $\xi(n)$ is the state vector of the input, and $\xi(.)$ its estimate, and

$$H(p) = H(o)\, \underline{\Psi}(p)$$

where $\underline{\Psi}(p)$ is the state transition matrix, and

H(o) is the product of known matrices of the form

$$H(o) = P^{1}\, C^{-2}$$

By partitioning the state vector, and all the associated matrices according to

$$\underline{\xi}(t) : \begin{bmatrix} \xi_o \\ \cdot \\ \cdot \\ \cdot \\ \xi_m \\ \hline \xi_{m+1} \\ \cdot \\ \cdot \\ \cdot \\ \xi_c(t) \end{bmatrix} = \begin{bmatrix} \underline{\xi}_1(t) \\ \\ \xi_2(t) \end{bmatrix}$$

we can determine the bias according to

$$\underline{b}_1(n + p/n) = (p)\ \underline{N}_2(n)$$

where $\underline{N}_2(t)$ represents the bottom part of the partitioned state vector

$$\underline{N}_2(t) = \begin{bmatrix} \nabla^{m+1}\ x(t)/(M+1)! \\ \cdot \\ \cdot \\ \cdot \\ \nabla\ \ x(t)/L! \end{bmatrix}$$, where $x(t)$ is the noise free input.

and $B(p) = [\bar{\Psi}_{11}(p) + \bar{\Psi}_{12}(p)\ \Gamma_{22}]$, where $\bar{\Psi}(p) = \begin{bmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{bmatrix}$ is the transition matrix

Γ_{22} is a submatrix of Γ, a known matrix

and $= (-1)^{m+1} J_1 \left\{ S_{21} + \ _{22}\ \Sigma_{21}\ S_{11} \right\}'\ J_1$, all known matrices

and $\underline{b}_1(n+p/n) = E\{\xi_1(n+p) - \xi_1^n(n+p/n)\}$ is the bias error in prediction.

These results are generalizations of specific work found in texts by Morrison and Blackman. Specific cases can be evaluated to minimize or keep within tolerable bounds, the bias by suitable choices of p, the prediction interval, and N the filter memory. Usually $p = -N/2$ is a good choice; but this represents a smoothing estimate as distinguished from filtering (present-time estimating) or prediction.