California Coding: Early LPC Speech in Santa Barbara, Marina del Rey, and Silicon Valley 1967–1982

Robert M. Gray
Information Systems Laboratory, Department of Electrical Engineering
Stanford, CA 94305

The author's group's work in speech coding was partially supported by the National Science Foundation. Thanks to J. D. Markel, A.H. "Steen" Gray, Jr., John Burg, Charlie Davis, Mike McCammon, Danny Cohen, Steve Casner, Richard Wiggins, Vishu Viswanathan, Jim Murphy, Cliff Weinstein, Joseph P. Campbell, Randy Cole, Rich Dean

A pdf of these slides may be found at http://ee.stanford.edu/˜gray/swim.pdf

Introduction

- The basic problem(s): linear prediction et al.
  - Signal processing: signal estimation, regression, smoothing, spectral shaping
  - modeling/density estimation
  - spectrum/correlation estimation

- The solution(s)

- The California part of the history of linear prediction in speech processing:

- Events
  * The first realtime LPC on the ARPAnet, How IP was influenced by voice
  * The first hardware LPC boxes
  * The publication of *Linear Prediction of Speech*
  * TI's Speak & Spell
- Places, Institutions, and People
  * Santa Barbara: UCSB, SCRL, CHI; Glen Culler, John Markel, Steen Gray, Mike McCammon, Hisashi Wakita
  * Marina del Rey: ISI; Danny Cohen, Steve Casner, Randy Cole
  * Silicon Valley: Time & Space Processing, SRI, Stanford; John Burg, Charlie Davis, Andrès Buzo, Yoseph Linde, Larry Stewart, Bob Gray, Earl Craighill, Tom Magill, Cheong Un

The Problem(s)

Fix m. Observe a data sequence $X^m = \{X_0, X_1, \dots, X_{m-1}\}$.

Optimal 1-step prediction ¿ What is the optimal predictor of the form $\tilde{X}_m = p(X_0, \dots, X_{m-1})$?

Optimal 1-step linear prediction ¿ What is the optimal linear predictor of the form $\tilde{X}_m = -\sum_{l=1}^{m} a_l X_{m-l}$?

Modeling/density estimation ¿ What is the probability density function (pdf) that "best" models X^m?

Spectrum Estimation ¿ What is the "best" estimate of the power spectral density or covariance of the underlying random process?

Wide literature on all of these topics in a speech context and they are intimately related.

See, e.g., Makhoul's classic survey [35] and Markel and Gray's classic book [41].

Clearly problems ill-posed unless define terms like "optimal" and assume some structure.

Deterministic vs. stochastic

Parametric (Gaussian) vs. nonparametric

Some Technical Preliminaries

$X^m = (X_0, X_1, \ldots, X_{m-1})^t$ a random vector

¿ What is best $\tilde{X}_m = p(X^m)$ given X^m in sense of minimizing mean squared prediction error $E[\epsilon_m^2]$, $\epsilon_m = X_m - \tilde{X}_m$, ?

Answer: Conditional expectation $\tilde{X}_m = E[X_m|X^m]$ is MMSE estmator and MMSE is conditional variance $\alpha_m = \sigma^2_{X_m|X^m}$.

If X^{m+1} is Gaussian with autocorrelation $r_{i,j} = E[X_i X_j]$, define $R_n = \{r_{i,j};\ i,j = 0,1,\ldots,n-1\}$, then MMSE is $\alpha_m = \sigma^2_{X_m|X^m} = |\ R_{m+1}\ |\ /\ |\ R_m\ |$ and optimal estimate is

$$\tilde{X}_m = E[X_m|X^m] = (r_{m,0}, r_{m,1}, \ldots, r_{m,m-1}) R_m^{-1} X^m$$

$\Rightarrow$ Form and performance are determined entirely by R_{m+1}!

$\Rightarrow$ optimal estimate is *linear*!!

Nongaussian case: Optimal *linear* estimate and performance are identical to that for Gaussian case with same autocorrelation!

$\Rightarrow$ Optimal linear estimate and performance for nonGaussian same as optimal for Gaussian.

Moral: Gaussian assumption provides short cut proofs in nonGaussian problems — no calculus and global optimality!

Cholesky decomposition for efficient inversion — *covariance method*. If R_{m+1} Toeplitz (e.g., underlying process stationary), Levinson-Durbin algorithm — *autocorrelation method*

For linear estimate: $a \stackrel{\Delta}{=} (a_0 = 1, a_1, \ldots, a_m)$

$$E[\epsilon_m^2] = a^t R_{m+1} a \stackrel{\Delta}{=} D_m(a) \left(= \int_{-1/2}^{1/2} |A(f)|^2 S_X(f)\, df \right)$$

where $A(f) = \sum_{n=0}^{m} a_n e^{-i2\pi nf}$ $\quad \left(S_X(f) = \sum_n r(n) e^{-i2\pi nf}\right)$

(if Toeplitz)

Orthogonality principle $\Rightarrow$ familiar normal (Wiener-Hopf, Yule-Walker) m linear equations in m unknowns.

¿ What if don't know R_m, observe long sequence of actual data $X_0, X_1, \ldots, X_{n-1}$? Under suitable conditions can estimate:

$$\hat{r}_k = \frac{1}{n-m}\sum_{l=m}^{n-1} X_l X_{l-|k|};\ \hat{R}_{m+1} = \{\hat{r}_{i-j};\ i,j = 0,1,\ldots,m\}$$

$$\overline{r}_{i,j} = \frac{1}{n-m}\sum_{l=m}^{n-1} X_{l-i} X_{l-j};\ \overline{R}_{m+1} = \{\overline{r}_{i,j};\ i,j = 0,1,\ldots,m\}$$

and "plug in." $\hat{R}_m$ Toeplitz, $\overline{R}_m$ not.

As $n \to \infty$, $\overline{R}_{m+1} \approx \hat{R}_{m+1}$ LP $\Leftrightarrow \min_{a:a_0=1} D_m(a) = \alpha_m$

Processes and Filters

For $n = m, m+1, \ldots$ find linear least squares estimate $\tilde{X}_n = -\sum_{l=1}^{m} a_l X_{n-l}$: Previous formulation $\Rightarrow$ optimal a, MMSE α_m.

$\epsilon_n = \sum_{k=0}^{m} a_k X_{n-k}$ = output of LTI filter with input X_n, response a_k: *prediction error filter* or *inverse filter* $\Leftrightarrow A(f)$
$X_n = \epsilon_n - \sum_{l=1}^{m} a_l X_{n-l}$ = output of an *autoregressive filter* driven by ϵ_n $\Leftrightarrow 1/A(f)$

Limit: $m \to \infty$, the orthogonality principal $\Rightarrow E(\epsilon_n X_{n-k}) = 0$, $k = 1, 2, \ldots \Rightarrow \epsilon_n$ is uncorrelated with all previous X_n and hence also with all previous $\epsilon_n \Rightarrow$ prediction error is *white*!

For finite m can interpret LP as trying to make prediction error *as white as possible*.

Sample Sequence View

Instead of taking expectations, minimize sample energy of error sequence: $\epsilon_n = \sum_{l=0}^{m} a_l X_{n-l}$ for $n \geq m$:

$$\mathcal{E}_n = \frac{1}{n-m} \sum_{k=m}^{n-1} \epsilon_k^2 = a^t \overline{R}_{m+1} a$$

the same quadratic form minimized using plug-in LP

If n large, $\overline{R}_m \approx \hat{R}_m$

Maximum Likelihood View

Suppose process X_n is order m Gaussian autoregressive source: W_n is iid Gaussian, variance σ^2 and

$$X_n = \begin{cases} -\sum_{k=1}^{m} a_k X_{n-k} + W_n & n = 0, 1, \ldots \\ 0 & n < 0 \end{cases}$$

Observe $X_0, \ldots, X_{n-1}$. ¿ What is maximum likelihood estimate of a, i.e., what a maximizes $f_{X^n|a}(x)$? $A_n X^n = W^n$ where

$$A_n = \begin{bmatrix} 1 & & & & & \\ a_1 & 1 & & & & 0 \\ \vdots & a_1 & 1 & & & \\ a_m & & \ddots & \ddots & & \\ & \ddots & & & & \\ 0 & & a_m & \ldots & a_1 & 1 \end{bmatrix}$$

$$\Rightarrow R_W^{(n)} = A_n R_X^{(n)} A_n^* = \sigma^2 I \;\Rightarrow\; R_X^{(n)} = \sigma^{-2} A_n^{-1} A_n^{-1*}$$

so the pdf for $X = X^n$ is

$$f_{X^n|a}(x) = \left(\frac{1}{2\sigma^2\pi}\right)^{\frac{n}{2}} e^{-\frac{1}{2\sigma^2} x^t A_n^t A_n x}$$

For large n $x^t A_n^t A_n x \approx \mathcal{E}_n = a^t \overline{R}_{m+1} a$

ML estimate does (approximately for large n) same minimization as LP solution a and $\sigma^2 = \alpha_m$.

ML method produces a *model* or *density estimate*: a Gaussian autoregressive process fit to data.

Maximum Entropy View

Suppose have estimate $\hat{R}_m$ of correlations to lag m of stationary random process X_n.

¿ What m-step Markov random process maximizes the Shannon differential entropy rate:

$$h(X) = \lim_{n\to\infty} \frac{1}{n} h(X^n)$$

where

$$h(X^n) = -\int f_{X^n}(x^n) \log f_{X^n}(x^n)\, dx^n\, ?$$

Since assume Markov, $h(X) = h(X_m | X^m)$.

Note: No Gaussian assumption, stated as a *variational problem*.

Answer: If n and R_n are fixed, then largest differential entropy is (surprise!) obtained by Gaussian density as

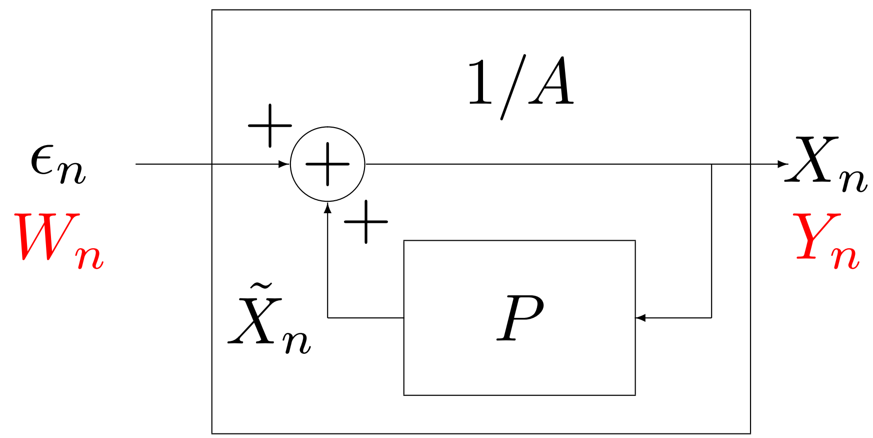
$$h(X^n) = \frac{1}{2}\log(2\pi e)^n \mid R_n \mid .$$

$\Rightarrow$ MAXDET problem, which has a long history and large literature. [69].

As $n \to \infty$, wish to maximize $h(X_m \mid X^m)$, accomplished by a Gaussian conditional density with variance [60]

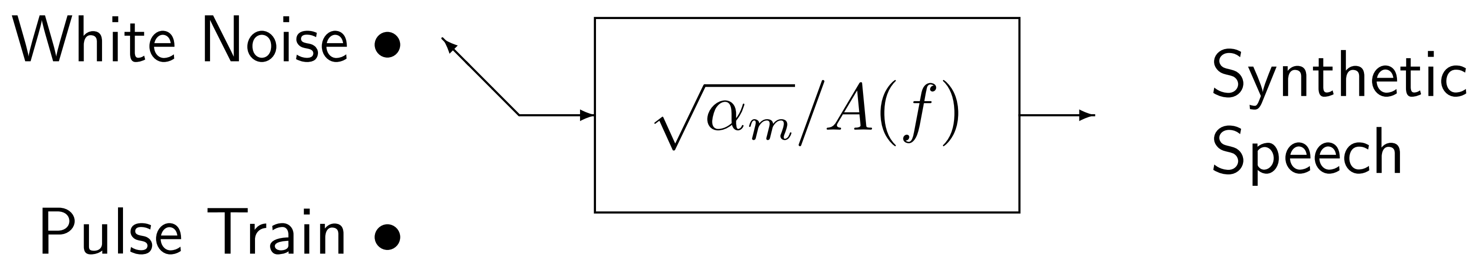
$$\sigma^2_{X_m|X^m} = \frac{\mid R_{m+1} \mid}{\mid R_m \mid} = \alpha_m,$$

achieved by mth order autoregressive process with psd $S(f) = \alpha_m / \mid A(f) \mid^2$ LP problem again



$$\sigma_W^2 = \alpha_m\ ,\ r_X(n) = r_Y(n), n = 0, 1, \ldots, m$$

$\Rightarrow$ LPC *model*



Simplistic: no voicing or pitch estimation details

Minimum Distortion Model Selection

Given process $\{X_n\}$ with autocorrelation function R and psd S (e.g., sample correlation or sample spectrum)

Assume a distortion measure $d(S, S_Y)$ (or $d(R, R_Y)$)

Best model in the class $\mathcal{A}_m$ of all mth order autoregressive processes for an original process chosen by minimum distortion or nearest neighbor rule.

Ditto if have only a finite subset (codebook) of $\mathcal{A}_m$!

Example: Modified *Itakura-Saito distortion measure* by

$$d(S, S_Y) = \int_{-1/2}^{1/2} \left(\frac{S(f)}{S_Y(f)} - \ln \frac{S(f)}{S_Y(f)} - 1 \right) = \frac{a^t R_{m+1} a}{\sigma_Y^2} - \ln \frac{\alpha_m}{\sigma_Y^2} - 1$$

The particular form has many information theoretic and other interpretations. Can be found in Pinsker's classic information theory text[4]. Example of Kullback's *minimum discrimination information* (variational) for density/parameter estimation. [3, 54, 61]

Minimized with value 0 by choosing the a to minimize $D_m(a)$ and by choosing $\sigma_Y^2 = \alpha_m$, that is, by solving the linear prediction problem and choosing the model to be $\alpha_m/|A(f)|^2$.

LPC

Find model (α_m, A) as before. Coding occurs when the final model is selected from a discrete set.

Traditional method: scalar quantization of some equivalent parameter set:

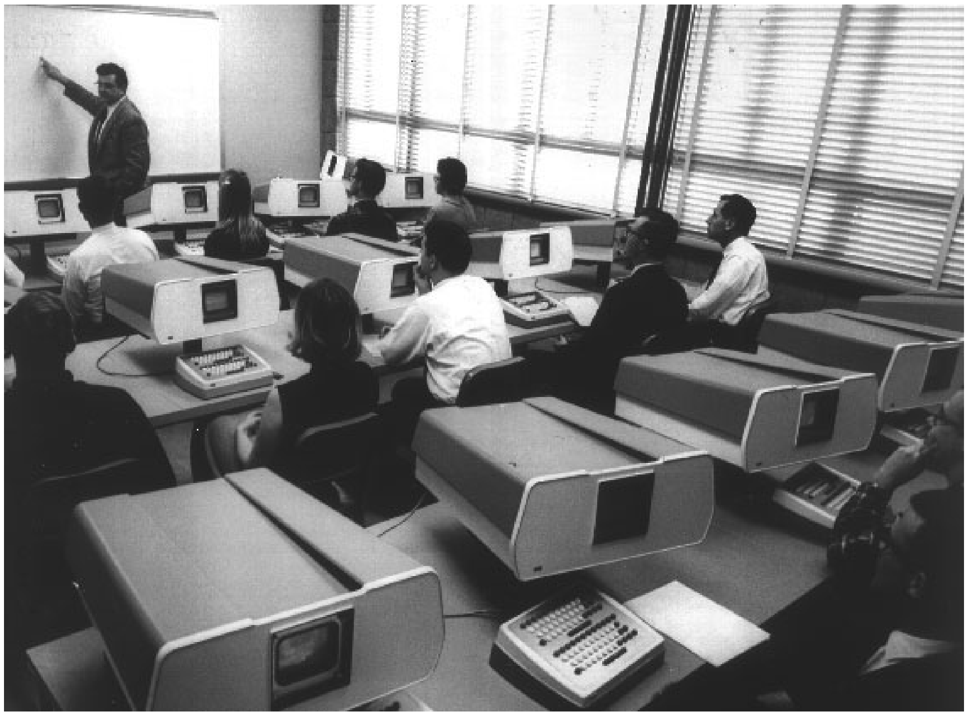
reflection coefficients $\Leftrightarrow$	partial correlation coefficients
inverse filter coefficients	cepstral coefficients
line spectral pairs (LSP)	autocorrelation coefficients

Issues of stability and computational complexity.

Or can quantize in the vector space of all mth order autoregressive models using same distortion measure as previously used to measure quality: Itakura-Saito or equivalent. (LPC-VQ)

California Timeline and Related Events

1966 UCSB Glen Culler introduces On-Line System



(OLS or Culler-Fried) — allows real time signal processing at individual student terminals. Glen Culler founds Culler-Harrison, Inc (CHI) which will eventually develop the FPS array processor.

Reknowned for building fast and effective computer systems, and many of his students and associates will found companies. CHI ideas will be successfully adopted and commercialized by FPS, will replace SPS-41.

December Saito and Itakura [5] describe an approach to

automatic phoneme discrimination and, as part of the development, develop the ML approach to speech coding: LP parameters transmitted to decoder with voicing information. Decoder synthesizes from noise or pulse train driving autoregressive filter.

See also 1968 & 1969 papers [11, 12].

From [5]:

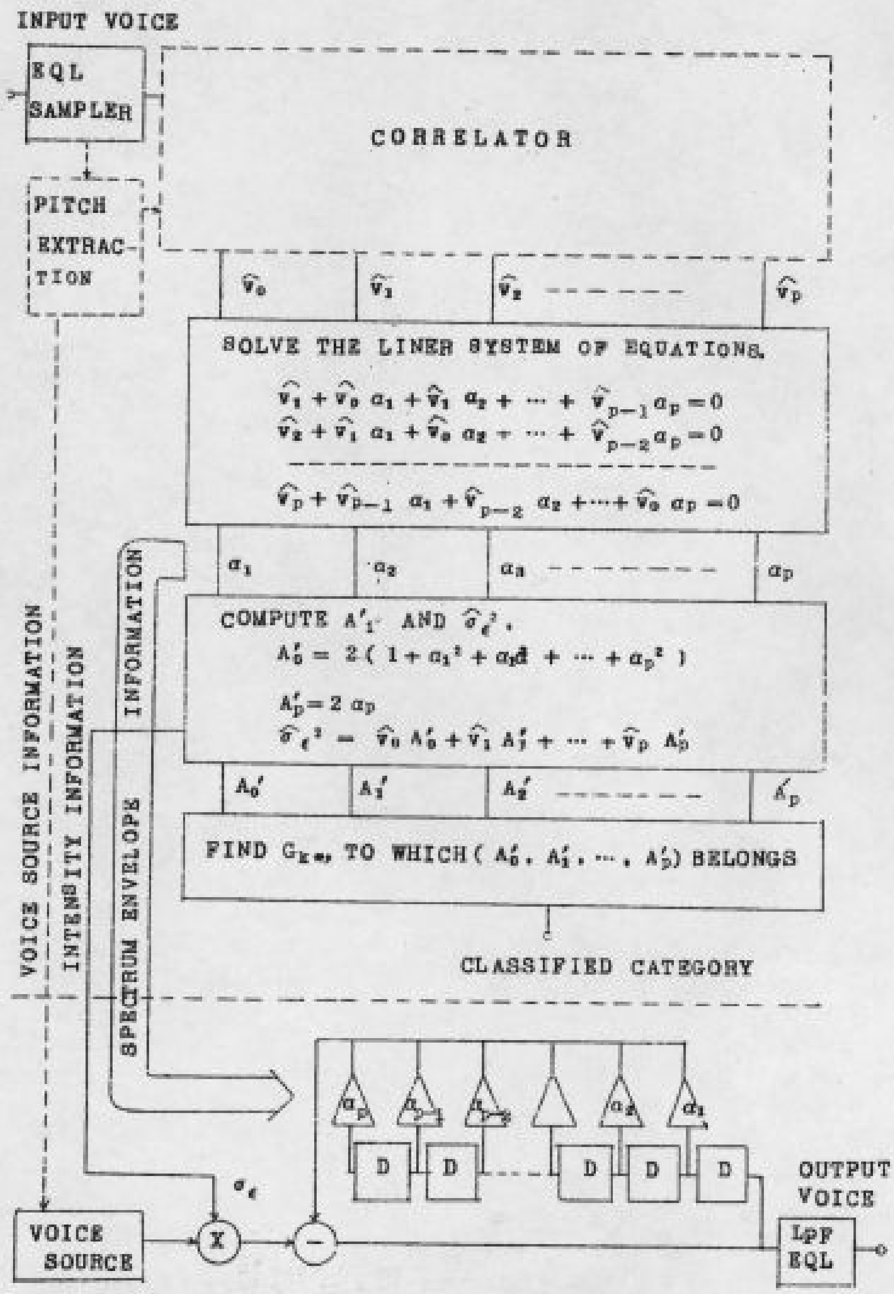


図 5. 新しいパラメータ伝送方式

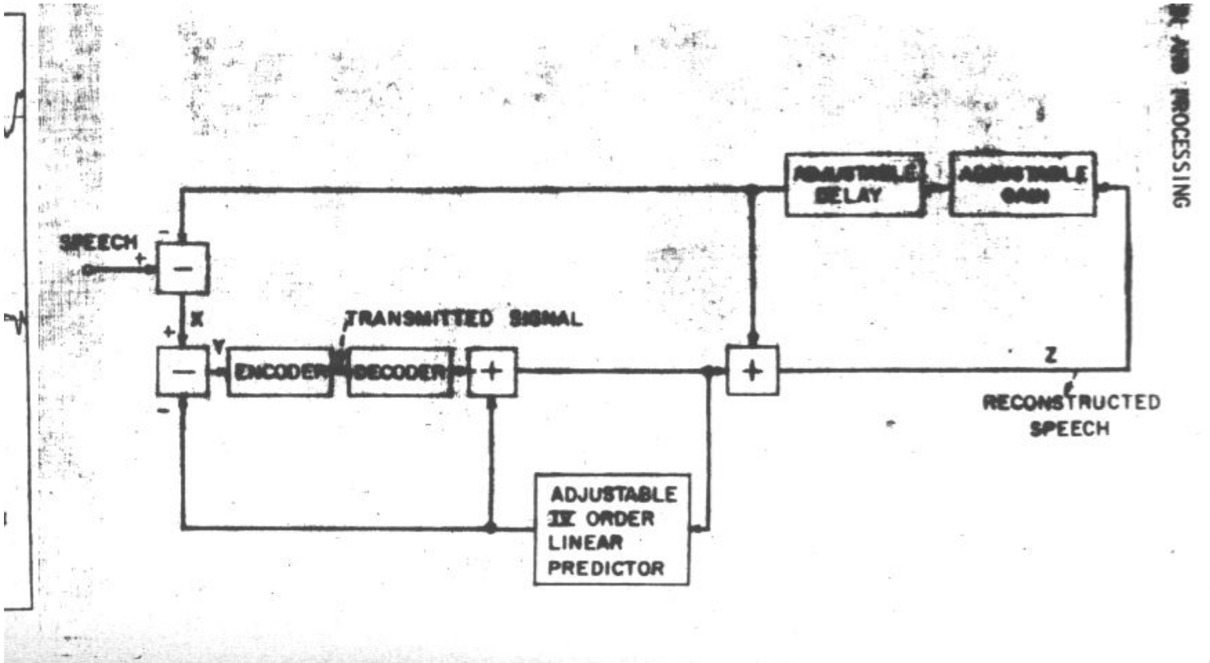
−39−

1967 

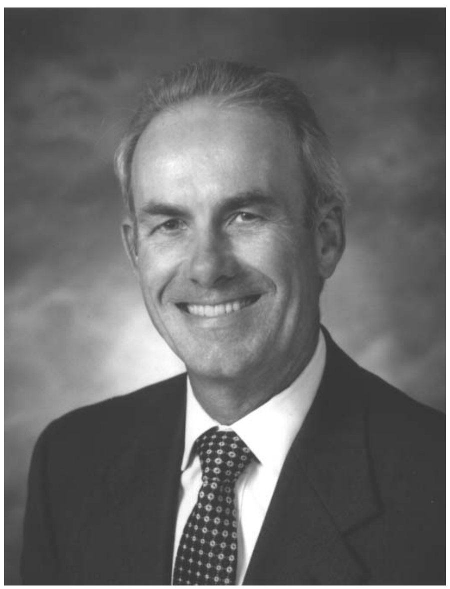
October John Burg presents maximum entropy approach [9] and wins best presentation award at the meeting of the Society of Exploration Geophysists. Dave Sakrison complains about buzz words "maximum entropy." Focus is on prediction error properties. Variational, not parametric.

November B.S. Atal and M.R. Schroeder [6]: LP coefficients used to form prediction residual, which is also coded. Adaptive predictive coding (APC), *residual excited* LPC. No explicit modeling. Elaborated in 1968 [7, 8] using covariance method.

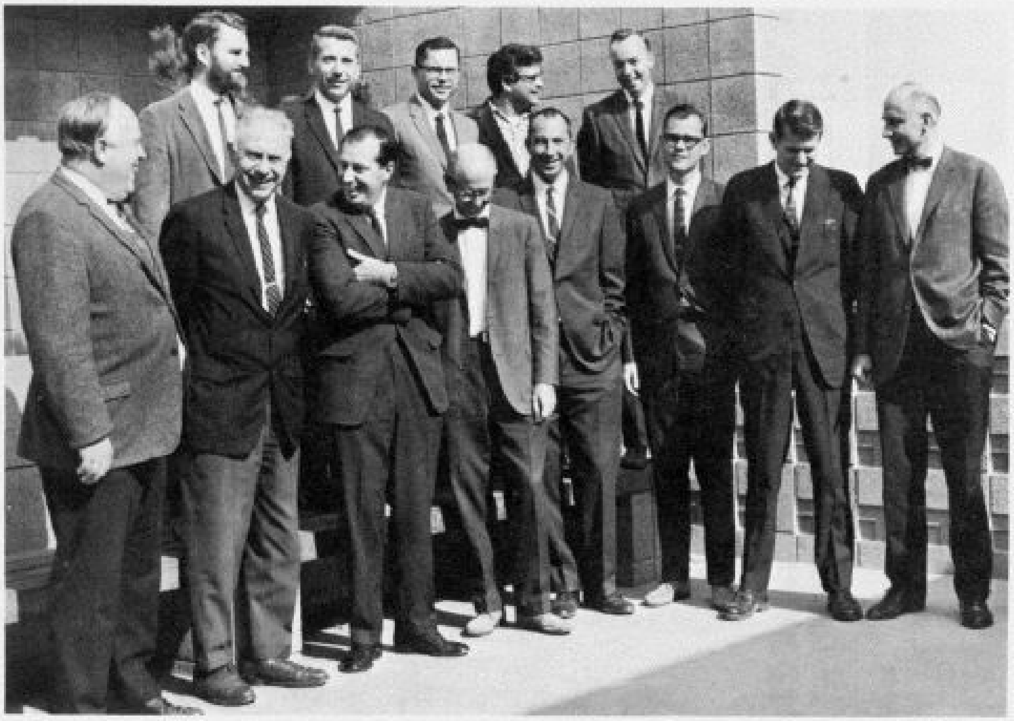
From [6]



1968



John Markel drops required language course in French for PhD program at Arizona State. Moves to UCSB (Fortran is accepted there). Joins Speech Communications Research Lab (SCRL). Reads Flanagan's book and sets goal to someday write and publish a book in the same series with the same publisher.



Front row: Philip Ordung, Albert Conrad, Jorge Fontana, George Matthaei, Roger Wood, Kenneth Kotzebue, John Skalnik, Glen Wade. **Second row:** Augustine Gray Jr., Glen Heidbreder, John Baldwin, Glen Culler, James Howard.

UCSB Faculty in 1968, including A.H. Gray, Jr., and Glen Culler

John Burg [10] presents “A new analysis technique for time series data” at Nato Advanced Study Institute — the Burg algorithm. Finds reflection coefficients from original data using a forward-backward algorithm. Later dubbed “covariance lattice” approach in speech[46, 55].

1969 Itakura and Saito[12] introduce partial correlation (PARCOR) [1969] variation on autocorrelation method, finds partial correlation [1] coeffients. Similar to Burg algorithm, but based on classical statistical ideas and lower complexity.

May Glen Culler proposes for online speech system aimed at real-time speech encoding based on a signal decomposition that would now be called a Gabor wavelet analysis. [13]

November B.S. Atal presents LPC speech coder at Annual Meeting of the Acoustical Society of America. [14]. Abstract published in 1970, full paper with Hanauer in 1971[16], uses covariance method.

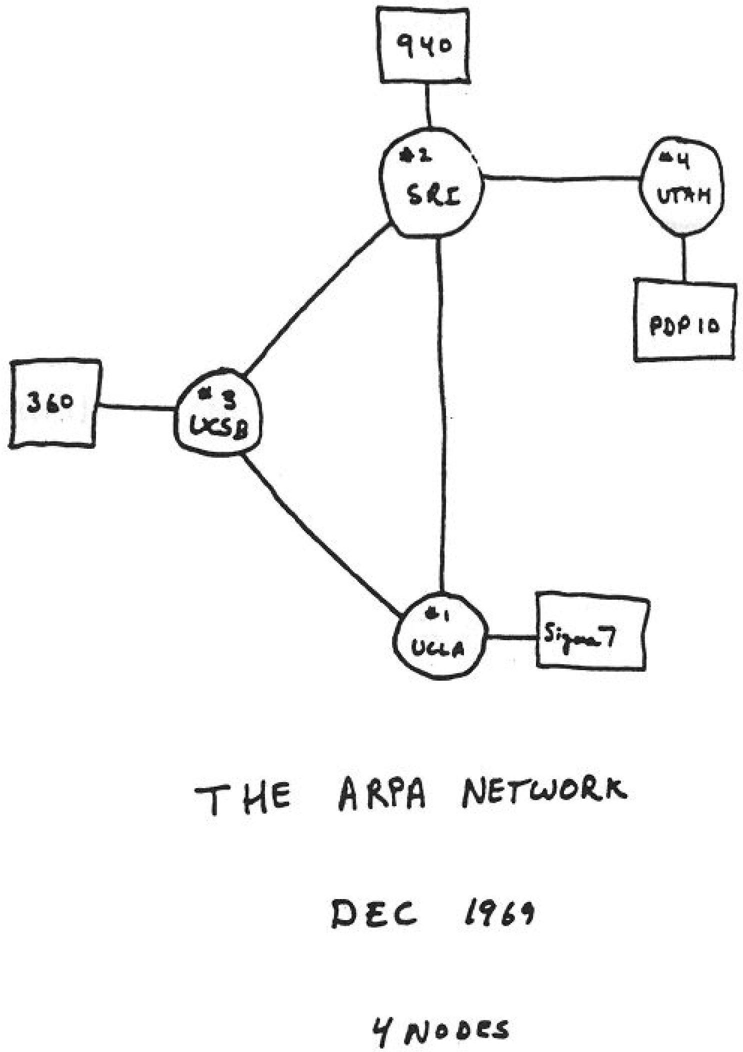


FIGURE 6.2 Drawing of 4 Node Network
(Courtesy of Alex McKenzie)

Thanks to Culler, UCSB becomes the third node on the ARPAnet (joining #1 UCLA, #2 SRI, #4 University of Utah)
Drawing by Jon Postel of ISI

1971 Real time LPC using Cholesky/covariance at Philco-Ford in PA. LONGBRAKE II Final Report in 1974[33]. 16 bit fixed point LPC. Four were sold (Navy and NSA), they weighed 250 lbs each. Used PFSP signal processing computer.

1972 Danny Cohen working at Harvard on realtime visual flight

simulation. Richard Wiggins at Harvard studying applied math, meets Danny Cohen. Bob Kahn (ARPA) suggests to Danny that similar ideas would work for real time speech communication over developing ARPAnet; advises Danny to move to Information Sciences Institute (ISI) in Marina del Rey.

1973 Danny moves to ISI, works with Steve Casner, Randy Cole, and others and with SCRL on real time operating systems. Kahn realizes bandwidth will not allow ordinary PCM or even ADM on ARPAnet, serious compression needed. Forms Network Secure Communications (NSC) group (later called Network Speech Compression and Network Skiing Club). Bob Kahn becomes éminence grise of ARPA speech project. Danny learns of LPC.

DSP chips did not yet exist. Every node on ARPAnet had different equipment and software. Focus on interface.

ARPA Network Information Center
Stanford Research Institute
Menlo Park, California 95025

RECEIVED
DEC 6 1973

Marcia Keeney
SRI-ARC
November 14, 1973

Network Speech Compression Note #3
NIC 19946

NETWORK SPEECH COMPRESSION GROUP MEMBERSHIP LIST

Steve F. Boll University of Utah Computer Science Dept. 3160 Merrill Engineering Building Salt Lake City, Utah 84112	SFB (801) 581-8576 UTAH-10
Dan Cohen University of Southern California Information Sciences Institute 4676 Admiralty Way Marina Del Rey, California 90291	DC (213) 822-1511 USC-ISI
Glenn J. Culler Culler-Harrison, Inc. 150-A Aero Camino Goleta, California 93017	GJC (805) 968-1813 CHI2
Robert E. Kahn Advanced Research Projects Agency 1400 Wilson Boulevard Arlington, Virginia 22209	REK2 (202) 694-5921 or 694-5922 ARPA-TIP
D. T. (Tom) Magill Stanford Research Institute 333 Ravenswood Avenue Menlo Park, California 94025	DTM (415) 326-6200 ext 2664 SRI
John Makhoul Bolt Beranek and Newman Inc. 50 Moulton Street Cambridge, Mass. 02138	JM (617) 491-1850 ext 234 BBN-TENEX
John D. Markel Speech Communications Research Lab, Inc. 800 Miramonte Drive Santa Barbara, California 93109	JDM (805) 965-3011 SCRL
Joseph Tierney MIT Lincoln Laboratory Lexington, Mass. 02173	JT3 (617) 862-5500 ext 277 IND

Original Members of NSC: ISI, University of Utah, BBN, MIT-LL, SRI. Soon joined by SCRL, CHI.

Others attend as well, including TI, NRL, Harris, NSA, Bell Labs

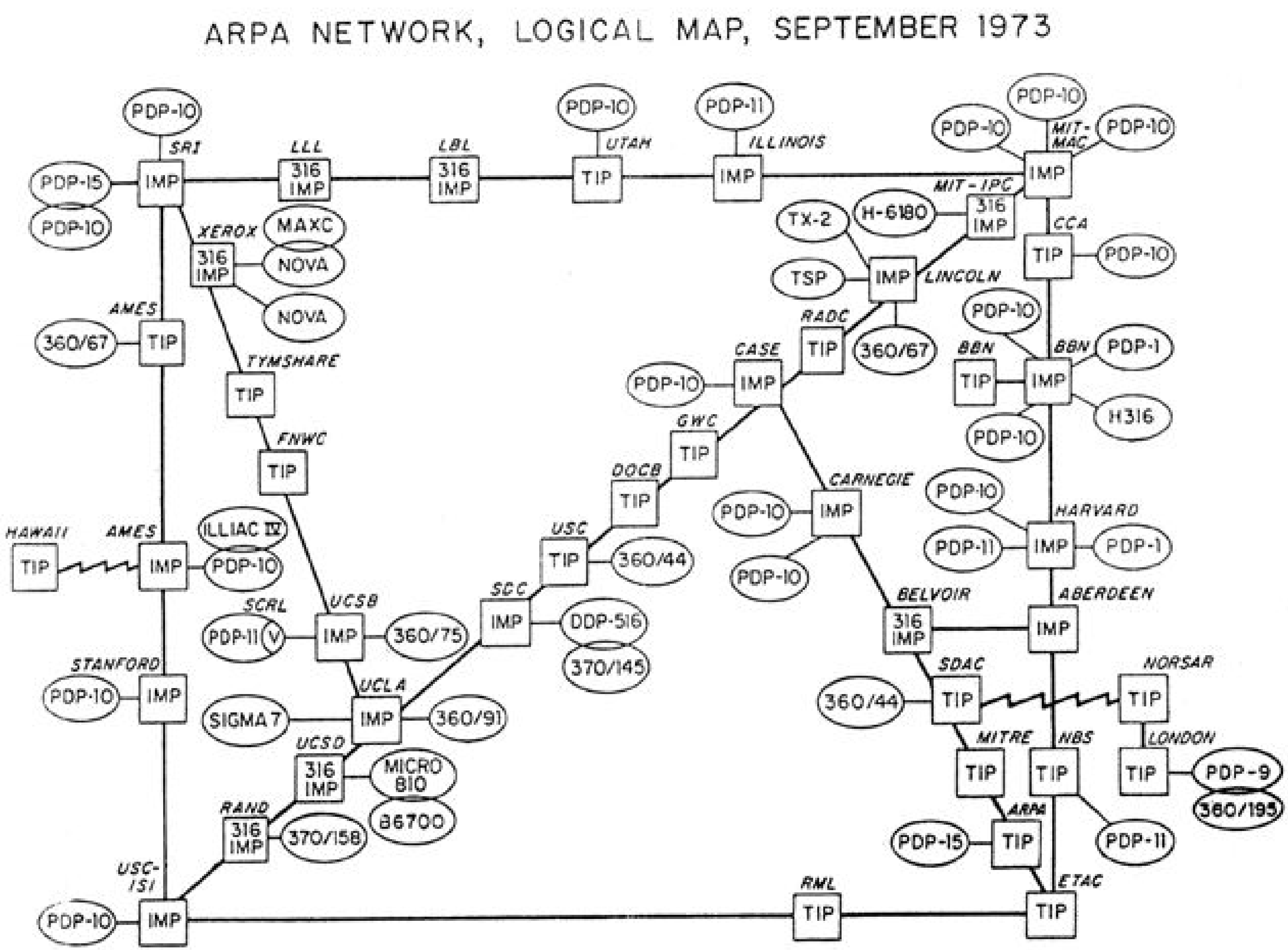
LL tradition is that first realtime 2400 bps LPC on the FDP done by Ed Hofstetter using Markel LPC formulation.

John Markel becomes Vice President of SCRL

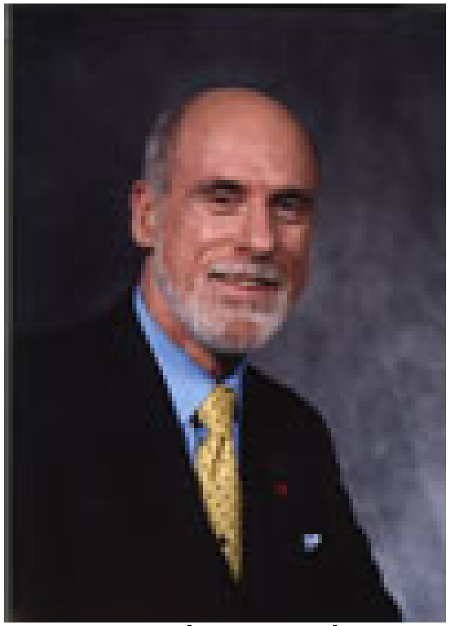
Markel, Gray, and Wakita [21, 22, 23] publish SCRL reports and papers describing their implementations of Itakura's algorithms plus several of their own (e.g., for pitch estimation). Make Fortran code freely available for LPC and associated signal processing algorithms.

ISI adopts basic Markel/Gray software [22] as vocoder technique for network speech project. SPS-41 chosen for implementation (LL and CHI excepted) . Divided software development among ISI, BBN, LL, SRI

John Burg meets Bishnu Atal, learns of LPC and similarity to his own algorithms.


ARPA NETWORK, LOGICAL MAP, SEPTEMBER 1973
PDP-10
SRI
IMP
PDP-15
PDP-10
LLL
316 IMP
LBL
316 IMP
UTAH
TIP
ILLINOIS
PDP-11
IMP
MIT-MAC
MIT-IPC
XEROX
MAXC
NOVA
NOVA
AMES
TIP
360/67
TYMSHARE
FNWC
H-6180
TX-2
TSP
LINCOLN
CCA
RADC
CASE
360/67
BBN
PDP-1
H316
GWC
DOCB
CARNEGIE
HARVARD
HAWAII
AMES
ILLIAC IV
USC
360/44
SCC
DDP-516
370/145
UCSB
SCRL
PDP-11
V
360/75
BELVOIR
ABERDEEN
STANFORD
UCLA
SIGMA 7
360/91
SDAC
NORSAR
MITRE
NBS
LONDON
UCSD
MICRO 810
B6700
RAND
370/158
PDP-9
360/195
ARPA
USC-ISI
RML
ETAC

1974



TCP invented by Bob Kahn and Vint Cerf (ISI) [25].

Network Voice Protocol (NVP) developed and written by Danny Cohen in coordination with others in NSC [30, 39, 43, 63], independent of TCP, uses only ARPANET message header. Successfully tested using CVSD in August between ISI and LL.

December First realtime two-way LPC packet speech communication. 3.5 kbs over ARPAnet between CHI and MIT-LL. [29, 34, 37, 63, 64] Uses basic M&G LPC algorithms

[37, 23, 26] coupled with NVP. CHI: MP-32A signal processor + AP-90 array/arithmetic coprocessor, LL: TX2 and FDP.

NSC realized that early ARPAnet protocols and TCP were incompatible with real time communication because of packet and reliability constraints. Argued for changes (for years!)

1975 John Burg's company Time and Space Processing (TSP) begins work on real time speech box using the Burg algorithm. Charlie Davis manages hardware and algorithm implementations.

John Markel notes many algorithms published for "optimal quantization" of LPC parameters. Observes any system is optimal according to some criterion. Other Markelisms:
Never believe demonstrations based on the training data. Never trust quality if the original is played first.

Un and Magill (SRI): residual excited linear prediction (RELP) [38]. Downsampled adaptive delta modulator codes open loop LP residuals + nonlinear spectral flattener in decoder. Similar scheme developed at LL by Weinstein (VELP)

October Chaffee and Omura (UCLA) algorithm for designing a codebook of LPC models based on rate distortion theory ideas for speech coding at under 1000 bps. [31, 36]

1976 *Linear Prediction of Speech* by J.D. Markel and A.H. Gray, Jr. published, fulfilling Markel's goal.

Comment on M&G from Joseph P. Campbell of LL (former NSA employee): "it was considered basic reading, and code segments (translated from Fortran) from this book (e.g., STEP-UP and STEP-DOWN) are still running in coders that are in operational use today (e.g., FED-STD-1016 CELP)."

January • First LPC conference over ARPANET based on LPC and NVP successfully tested.: CHI, ISI, SRI, LL 3.5 kbps

Texas Instruments begins development of Speak & Spell toy. Larry Brantingham, Paul Breedlove, Richard Wiggins, and Gene Frantz. While previously at MITRE, Wiggins worked on speech algorithms, worked closely with LL, and visitied Itakura and Atal at Bell, NSC, Makhoul and Viswanathan at BBN, George Kang at NRL. While at TI visits Markel at SCRL and ISI in summer of 1977. Meets Burg on an airplane.

March NVP published by Danny Cohen. Excerpt:
"The Network Voice Protocol (NVP), implemented first in December 1973, and has been in use since then for local and transnet real-time voice communication over the ARPANET at the following sites:

- Information Sciences Institute, for LPC and CVSD, with a PDP-11/45 and an SPS-41.
- Lincoln Laboratory, for LPC and CVSD, with a TX2 and the Lincoln FDP, and with a PDP-11/45 and the LDVT.
- Culler-Harrison, Inc., for LPC, with the Culler-Harrison MP32A and AP-90.
- Stanford Research Institute, for LPC, with a PDP-11/40 and an SPS-41.

The NVP's success in bridging the differences between the above systems is due mainly to the cooperation of many people in the ARPA-NSC community, including Jim Forgie (Lincoln Laboratory), Mike McCammon (Culler-Harrison), Steve Casner (ISI) and Paul Raveling (ISI), who participated heavily in the definition of the control protocol; and John Markel (Speech Communications Research Laboratory), John Makhoul (Bolt Beranek & Newman, Inc.) and Randy Cole (ISI), who participated in the definition of the data protocol. Many other people have contributed to the NVP-based effort, in both software and hardware support."

1977 John Markel founds Signal Technology Inc. (STI). First product is Interactive Laboratory System (real time signal processing packages based on SCRL software).

WHY YOU SHOULDN'T WRITE YOUR OWN SIGNAL PROCESSING SOFTWARE.

BY A.H. "STEEN" GRAY, Jr., Ph.D.
Vice President, Signal Technology, Inc.

I've been in this business long enough to know that some things never change. The "make or buy" quandary as it applies to software is a good example. "We've got some expensive programmers; let them earn their keep." How many times have you heard that when you've suggested buying a program package that seems perfectly tailored to your application?

HOW TO ANSWER

If your line of work is signal processing, the answer should be reasonably simple. Just say, "It would take us ten years and a bundle of money to come up with a package as good as the one already available from some experts out in California." That

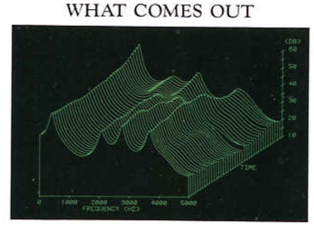
might be just a slight exaggeration, but your point would be well taken. You see, we do have the last word in interactive signal processing software. It's called ILS, and with over 200 installations worldwide, it is often referred to (and not just by us) as the "world standard."

WHAT ILS IS

ILS is a highly modular set of FORTRAN programs that make up a sophisticated interactive software system with standard file structures, documentation and ongoing support. To date, it is performing with excellent results in a wide variety of industries and technologies, including: **speech—noise and vibration—acoustics—biology—medicine—simulation—digital filtering—sonar—radar—seismic—** and some we aren't being told about.

Many of our users also find DACS, our Data Acquisition and Conversion Software, and APAS, our Array Processor Application Software, of great value in their applications.

WHAT COMES OUT



With any compatible computer system and the appropriate terminals, ILS will give you: **pattern analysis—digital filtering—signal editing—3-D displays—modeling—correlation—convolution—spectral density—signal displays—coherence—** and maybe even a picture of your Aunt Sally, if that's what you want.

THE POINT IS

The important thing is, ILS is available *now*. It has been proven in a multitude of applications around the world. And it can cost you a lot less in time and money to buy it from us rather than developing a comparable package yourself.

FREE DEMO

If you have a compatible graphics terminal and modem, we can arrange to give you an on-line demonstration of ILS. Simply call (toll free) and ask for our ILS marketing representative at (800) 235-5787. We also have a video tape that illustrates many of the features of ILS and its capabilities. We'll be happy to send it to you.

IF YOU'RE STILL WITH US

If you've read this far, you probably have some kind of interest in signal processing. I'd be delighted to send you a reprint of the series of three articles on digital filtering that I co-authored with John Markel. Write to me at the address below.

SIGNAL TECHNOLOGY, INC.

15 West De La Guerra Street, Santa Barbara, CA 93101
California (805) 963-1552 • **Outside California Call Toll Free (800) 235-5787**
TWX 910-334-3471—SIGNAL TEC SNC

Circle No. 29

May: TSP demonstrates Model 100 LPC box at Armed Forces Communication and Electronics Association annual show. Quality superior to competition. TSP developed Model 200 (sealed version) sold in 1979 in conjunction with new Codex full duplex, synchronous, 2400bps modem. Bright orange. Sold (mostly to spooks) for $12K/box. Continued as viable product through mid 80s. Competition from Secure Terminal Unit (STU) I (development completed in 1974, produced 1977–1979. APC using Levinson, $35K), STU II (LPC/APC) (developed 1977-1980, produced 1982–1986, $13K), and the eventually dominant STU III (development begun in 1984, production begun 1987, $2K).

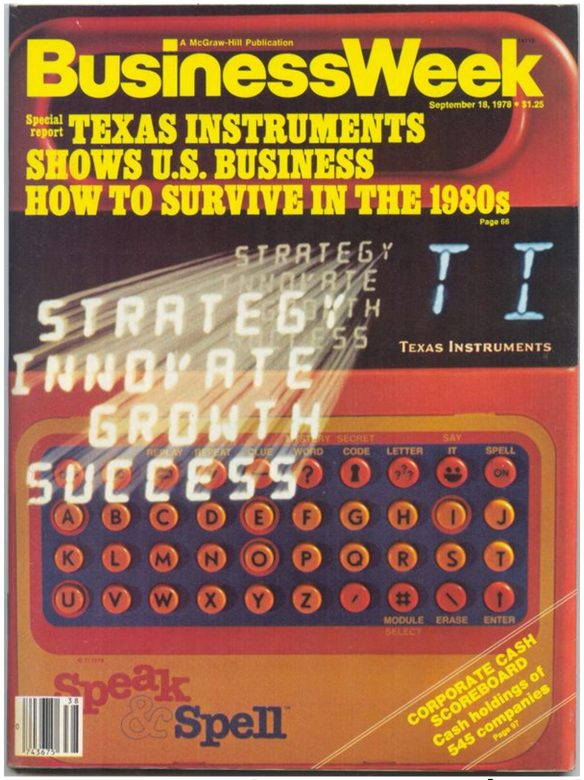
April Bell Labs applies for a patent for "'packet transmission of speech" four years after ARPA/NSC LL/CHI demonstration. Granted USA Patent 4,100,377 in 1978.[63, 70]

IP extracted from TCP to allow for real time speech and other applications. Eventually ATM developed for similar reasons.

Irony in current popular view of VOIP as novel, *IP was in fact specifically designed to handle packet speech and other realtime data!!*

1978 May–April LPC conferencing over ARPANET using variable frame-rate (2–5 kbps) among CHI, ISI, and LL (Vishwanath et al. of BBN developed algorithm)

June Texas Instrument *Speak & Spell* toy hits the market.



1st consumer product incorporating LPC and 1st single chip speech synthesizer and early DSP chip. Speech synthesis from stored LPC words and phrases using TMC 0280 one-chip LPC speech synthesizer. Speech data in up to sixteen 128K ROM chips (TMC 0350). Seminal to the development of DSP chips. Later refinements led to TMS 5100, 5200, and 5220 voice synthesis processors. Before announcement, Wiggins calls Markel, Makhoul, Atal, and Gold to acknowledge their contributions to speech and to announce the Speak & Spell. Markel asked where his royalties were — Wiggins sent him a Speak & Spell.

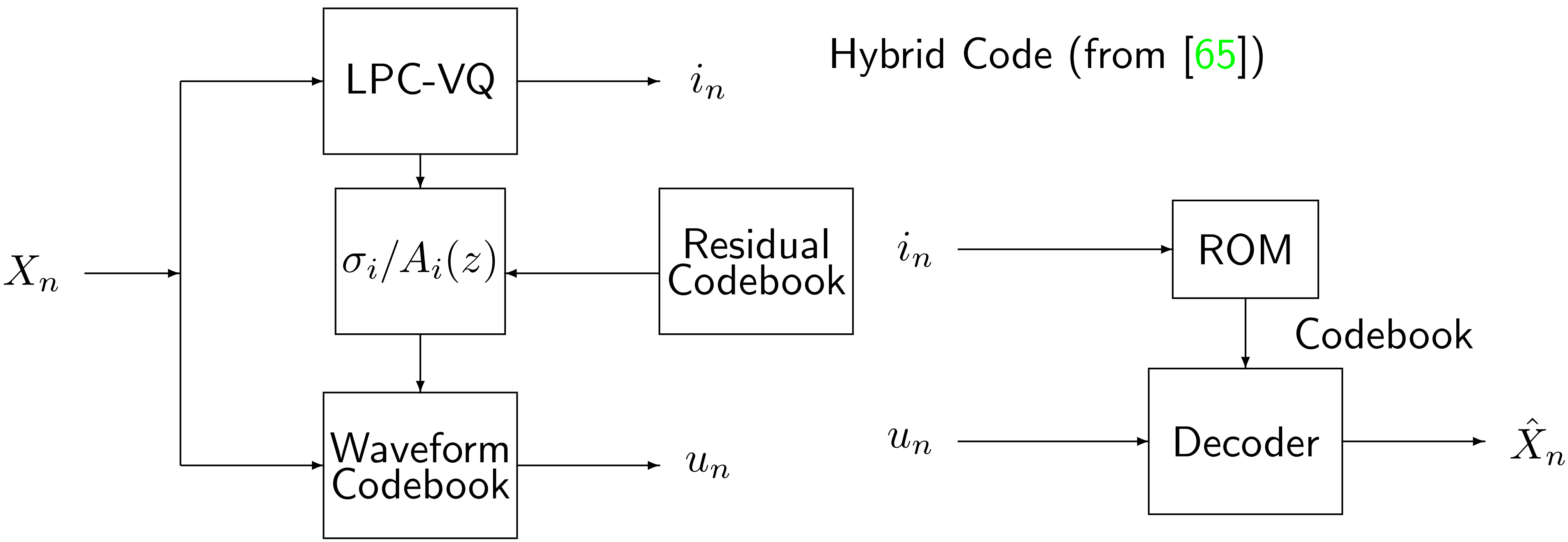
Steen Gray gives talk on "Speech compression and speech distortion measures" and plays tapes of 800bps vector quantized LPC speech at 1978 Information Theory Workshop in Lake Tahoe. They sound awful. John Markel phones to explain the original speech was awful.

November–December Gray, Buzo, Matsuyama, Gray, and Markel present clustered LPC VQ using Itakura-Saito distortion at conferences [49, 50]. In 1979–1980 expand in [51, 52, 56].

1979 ARPANET/Atlantic SATNET conferencing using LPC10 at 2.4kbps: ISI, LL, NDRE (Norway), UCL (London)

1981 Larry Stewart [59, 62, 65] designs trellis codes (low complexity VQ) for closed loop prediction residuals for coded LPC model. "Hybrid" code using LPC and excitation codebooks. Simulations using legendary Altos systems at Xerox PARC.

Similar ideas reported by Matsuyama et al. in 1978 [48] and 1981 [58].



Hybrid Code (from [65])

1982 LL's single-card device using 3 NEC7720 chips, 18 sq.in. and 5.5 w shown to NSA Director. In combination with NSA's in-house development of a DSP-chip-based 2.4 kbps modem$\Rightarrow$ that secure desktop telephones were feasible, and led directly to the decision to go ahead with the STU-III development–eventually displaced TSP and other LPC boxes.

Epilog

• Residual codebook excitation combined with perceptual weighting filters and random codebook VQ to produce CELP (1985) [66] + postfiltering from Chen and Gersho (UCSB, 1987) [67] ⇒ Fed-Std-1016 CELP coder [Tremain, Campbell, Welch [68], which included some old Markel/Gray software. Last NSA speech coding standard thus blended West and East. STU III still in use, now being replaced by MELP at 2400 bps and G729 CS ACELP at 8,000 bps. Voice coding at NSA died in 2003.

•Randy Cole: "it's hard to understate the influence that the NSC work had on networking. . . . the NSC effort was the first real exploration into packet-switched media, and we all know the effect that's having on our lives 30 years later."

This talk is dedicated to Tom Stockham
12/22/33–1/6/2004

who played a minor role in this particular story as member of the Utah branch of the NSC, but a major role in the development of audio and image processing and coding at MIT, Lincoln Lab, and the University of Utah — winner of an Emmy, Grammy, and Oscar. He will be missed.

References

[1] U. Grenander and M. Rosenblatt, *Statistical Analysis of Stationary Time Series*, John Wiley & Sons, NY, 1957.

[2] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications,* University of Calif. Press, Berkeley and Los Angeles, 1958.

[3] S. Kullback, *Information Theory and Statistics*, Dover New York, 1968. Reprint of 1959 edition published by Wiley

[4] M. S. Pinsker, "Information and information stability of random variables and processes," Holden Day, San Francisco, 1964.

[5] S. Saito and F. Itakura, "The theoretical consideration of statistically optimum methods for speech spectral density," Report No. 3107, Electrical Communication Laboratory, NTT, Tokyo, December 1966.

[6] B.S. Atal and M. R. Schroeder, "Predictive coding of speech signals," *Proc. 1967 AFCRL/IEEE Conference on Speech Communication and Processing*, pp. 360–361, Cambridge, Mass, 1967.

[7] B.S. Atal and M. R. Schroeder, "Predictive coding of speech signals," *Rep. 6th Int. Congr. Acoust.*, Y. Konasi, Ed., Tokyo Japan, Rep.C-5-5, August 1968,

[8] B.S. Atal and M. R. Schroeder, "Predictive coding of speech signals," *WESCON Tech. Papers*, Paper 8/2, 1968.

[9] John Parker Burg, "Maximum entropy spectral analysis," presented at the 37th Meeting of the Society of Exploration Geophysisists, Oklahoma City, Oklahoma, October 1967.

[10] John Parker Burg, "A new analysis technique for time series data," presented at the NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics," Enschede, The Netherlands, Aug. 1968, reprinted in *Modern Spectrum Analysis*, D. G. Childers, ed., IEEE Press, New York, 1978.

[11] F. Itakura and S. Saito, "Analysis synthesis telephony based upon the maximum likelihood method," Reports of 6th Int. Cong. Acoust., ed. by Y. Kohasi, Tokyo, C-5-5, C17–20, 1968.

[12] F. Itakura and S. Saito, "Analysis synthesis telephony based on the partial autocorrelation coefficient," Acoust. Soc. of Japan Meeting, 1969.

[13] Glen J. Culler, "An Attack on the Problems of Speech Analysis and Synthesis with the Power of an On-Line System," *Proceedings of the 1st International Joint Conference on Artificial Intelligence*, Donald E. Walker, Lewis M. Norton (Eds.): Washington, DC, May 1969. pp. 41–48.

[14] B.S. Atal, "Speech analysis and synthesis by linear prediction of the speech wave," presented at the 78th Meeting of the Acoustical Society of America, San Diego, November 1969. Abstract in *J. Acoust. Soc. Am.*, Vol 47, p. 65, 1970.

[15] "Adaptive predictive coding of speech signals," *Bell Sys. Tech. J.*, Vol. 49, No. 8, pp. 1973–1986, October 1970.

[16] B. S. Atal and S.J. Hanauer "Speech analysis and synthesis by linear prediction of the speech wave," *J. Acoustic Society of America*, Vol. 50, pp. 637–655, August 1971.

[17] F. Itakura and S. Saito, "Digital filtering techniques for speech analysis and synthesis," *Conf. Rec. 7th Int Congr Acoustics*, Paper 25C1, 1971.

[18] F. Itakura and S. Saito, "On the optimum quantization of feature parameters in the parcor speech synthesizer," *Conference Record, 1972 International Conference on Speech Communication and Processing*, Boston, MA, pp. 434–437, April 1972.

[19] J.D. Markel, "Digital inverse filtering–a new tool for formant trajectory estimation," *IEEE Trans. on Audio and Electro Acoustics*, pp. 129–13?, June 1972. (autocorrelation method)

[20] H. Wakita, "Estimation of the vocal tract shape by optimal inverse filtering and acoustic/articulatory conversion methods," *SCRL Monograph No. 9*, SCRL, Santa Barbara, CA 1972.

[21] J.D. Markel, A.H. Gray, and H. Wakita, *Linear Prediction of Speech — Theory and Practice*, SCRL Monograph No. 10, Speech Communications Research Laboratory, Santa Barbara, California, 1973.

[22] "Documentation for SCRL Linear Prediction Analysis/Synthesis Programs," Speech Communications Research Labs, Inc." Nov. 1973.

[23] J.D. Markel and A.H. Gray, Jr., "On autocorrelation equations as applied to speech analysis," *IEEE Trans. on Audio and Electroacoustics*, Vol. 21, pp. 69–79, April 1973.

[24] D.T. Magill, "Adaptive speech compression for packet communication systems," *Conf. Rec. 1973 IEEE Telecom. Conf. Proc*, 29D-129D-5. DELCO algorithm, LPC with removal of silence and sending only changes.

[25] Vinton Cerf, Robert Kahn, "A protocol for packet network internetworking," *IEEE Trans. Commun.*, Vol. 22, May 1974.

[26] J.D. Markel and A.H. Gray, Jr., "A linear prediction vocoder simulation based upon the autocorrelation method," *IEEE Trans on Acoustics, Speech, and Signal Processing*, Vol. 22, pp. 124–134, April 1974.

[27] Draft Summary of First CHI–LL LPC Attempt Mike McCammon 11/26/74

[28] Draft Experiment #2 CHI–LL LPC Voice Contact Mike McCammon 11/26/74

[29] Report on Third CHI–LL LPC Voice Experiment 26 November 10:00–1:00 AM PST Mike McCammon

[30] ISI/SR-74-2 Annual Technical report May 1973-June 1974

[31] D.L. Chaffee and J.K. Omura, "A very low rate voice compression system," *Abstracts of Papers, International Symposium on Information Theory*, p. 69, October, 1974.

[32] NSC Note 47 Two NSC Group Proposals and a Status Report J.D. Markel SCRL 11 November 1974

[33] J. Welch, "LONGBRAKE II Final Report," Contract No. DAAB03-74-C-0098, Philco-Ford Corporation, Willow Grove, Pennsylvania, 1974.

[34] ISI/SR-75-3 Annual Technical report May 1974-June 1975

[35] John Makhoul, "Linear prediction: a tutorial review," *Proceedings of the IEEE*, Vol. 63, No. 4, April 1975.

[36] D.L. Chaffee, *Applications of rate distortion theory to the bandwidth compression of speech signals*, PhD Dissertation, UCLA.

[37] Glen J. Culler, Michael McCammon, and J.F. McGill, "Real time implementation of an LPC algorithm," *Culler/Harrison Inc. Quarterly Technical Report on Speech Signal Processing Research at CHI, during Nov. 1974–April 1975*, May 1975.

[38] Cheong Un and D. Thomas Magill, "The residual-excited linear prediction vocoder with transmission rate below 9.6kbs, *IEEE Trans on Communications*, Vol 23, pp. 1466-1474.

[39] D. Cohen "Specifications for the Network Voice Protocol," USC/Information Sciences Institute, ISI/RR-75-39, March 1976. Also NSC Note 68, January 1976.

[40] ISI/SR-76-6 Annual Technical report July 1975 – June 1976

[41] J.D. Markel and A.H. Gray, Jr, *Linear Prediction of Speech*, Springer-Verlag, 1976.

[42] A.H. Gray and J.D. Markel, "Distance Measures for Speech Processing," *IEEE Trans. on ASSP*, Vol. 24, no. 5, Oct. 1976.

[43] D. Cohen, "RFC0741: Specifications for the Network Voice Protocol," 22 Nov 1977.

[44] A. H. Gray, Jr., R. M. Gray and J. D. Markel, "Comparison of optimal quantizations of speech reflection coefficients," *IEEE Trans. on Acous., Speech & Signal Process.*, Vol. ASSP–25, pp. 9–23, Feb. 1977.

[45] Gold, Bernard (invited paper), "Digital Speech Networks", Proc. IEEE Vol. 65, No. 12, Dec. 1977. (Discusses ARPA voice project.)

[46] J. Makhoul, "Stable and lattice methods for linear prediction," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 25, pp. 423–428, October 1977.

[47] Wiggins, R. and L. Brantingham, "Three-chip system synthesizes human speech," Electronics, Aug 31, 1978, 109-116. Uses TMC 0280, LPC-10, 600-2400 bps.

[48] Y. Matsuyama, "Speech compression systems using a set of inverse filters," PhD Dissertation, Stanford University, August 1978.

[49] R.M. Gray, A.Buzo, Y. Matsuyama, A.H. Gray, Jr. and J.D. Markel, "Source coding and speech compression," *Proceedings of the 1978 Int'l. Telemetering Conf., 24*, Los Angeles, CA, pp. 871–878, Nov. 1978.

[50] A. Buzo, R.M. Gray, A.H. Gray, Jr., and J.D. Markel, "Optimal Quantizations of Coefficient Vectors in LPC Speech," *1978 Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan*, Honolulu, HI, Dec. 1978.

[51] A. Buzo, A.H. Gray, Jr., R.M. Gray, and J.D. Markel, "A Two–Step Speech Compression System with Vector Quantizing," *Proceedings of the 1979 Int'l. Conf. on Acoustics, Speech and Signal Processing*, pp. 52–55, Wash. DC, 1979.

[52] Y. Linde, A. Buzo and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. on Comm.*, Vol. COM–28, pp. 84–95, Jan. 1980.

[53] J.-P. Adoul, J.L. Debray, and D. Dalle, "Spectral distance measure applied to the design of DPCM coders with L predictors," *Conf. Record 1980 IEEE ICASSP*, Denver, CO, pp. 512–515, April 1980.

[54] R. M. Gray, A. Buzo, A. H. Gray, Jr., and Y. Matsuyama, "Distortion measures for speech processing," *IEEE Trans. Acous., Speech and Sign. Proc.*, Vol. ASSP–28, pp. 367–376, Aug. 1980.

[55] A.H. Gray, Jr., and D.Y. Wong, "The Burg algorithm for LPC Speech analysis/synthesis," *IEEE Trans on Acoustics, Speech, and Signal Processing,"*Vol. 28, No. 6, December 1980, pp. 609–615.

[56] A. Buzo, A.H. Gray, Jr., R.M. Gray, and J.D. Markel, "Speech coding based upon vector quantization," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP–28, pp. 562–574, October 1980.

[57] Cohen, D., "A voice message system," in R. P. Uhlig (ed.), Computer Message Systems, pp. 17-27, North-Holland, 1981. (Discusses ARPA voice project.)

[58] Y. Matsuyama and R.M. Gray, "Universal tree coding for speech," *IEEE Trans. on Information theory*, Vol. 27, No. 1, pp. 31–40, January 1981.

[59] L. C. Stewart, *Trellis Data Compression*, Ph.D. Dissertation, Department of Electrical Engineering, Stanford University, June 1981.

[60] J. van Campenhout and T. Cover, "Maximum entropy and conditional probability," *IEEE Transactions on Information Theory*, Volume: 27 , No 4 , July pp 483 – 489, 1981.

[61] R.M. Gray, A.H. Gray, Jr., G. Rebolledo, and J.E. Shore, "Rate distortion speech coding with a minimum discrimination information distortion measure," *IEEE Transactions on Information Theory*, vol. IT–27, no. 6, pp. 708–721, Nov. 1981.

[62] L.C. Stewart, R.M. Gray, and Y. Linde, "The design of trellis waveform coders," *IEEE Transactions on Communications*, Vol. COM–30, pp.702–710, April 1982.

[63] Exerpts from "Packet Speech Program Review Meeting" 3 June 1982 Sponsored by DARPA Hosted by MIT LL "Network voice protocols," Danny Cohen ISIpp. 40–59 'Packetized Speech Overview" pp. 17–23

[64] C.J. Weinstein and J.W. Forgie, "Experience with speech communication in packet networks," *IEEE Journal on Selected Areas in Communications*, Vol. 1, No. 6, December 1983.

[65] R.M. Gray, "Vector Quantization,"*IEEE ASSP Magazine*, Vol. 1, pp. 4–29, April 1984.

[66] M.R. Schroeder and B. Atal, Code-Excited Linear Prediction (CELP): High Quality Speech at Very Low Bit Rates, Proc. ICASSP-85, p. 937, Tampa, Apr. 1985.

[67] Chen and A. Gersho, "Real-Time Vector APC Speech Coding at 4800 bps with Adaptive Postfiltering," *Proc. ICASSP '87*, pp. 2185–2188.

[68] J. P. Campbell, T. E. Tremain, and V. C. Welch, "The Federal Standard 1016 4800 bps CELP Voice Coder," *Digital Signal Processing*, Vol. 1, no. 3 (1991): 145 – 155.

[69] L. Vandenberghe, S. Boyd, and S.-P. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM Journal on Matrix Analysis and Applications*, Vol. 19, 499-533, 1998.

[70] Danny Cohen, "Realtime Networking and Packet Voice," SIGCOM'99