Packet Speech Program
Review Meeting

June 3, 1982

Sponsored By
Department of Defense
Defense Advanced Research Projects Agency

Hosted By
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY
The Defense Advanced Research Projects Agency sponsored a meeting on June 3, 1982, at Lincoln Laboratory to review the results of the DARPA Packet Speech Program and to demonstrate packet speech technology to a group of invited visitors from the Department of Defense. The purpose of this document is to record the proceedings of the meeting by collecting the visual aids used in the technical presentations and in the demonstration.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>TAB #</th>
<th>Title</th>
<th>Authors/Institutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Packetized Speech Overview</td>
<td>Lt. Col. Duane A. Adams, Defense Advanced Research Projects Agency</td>
</tr>
<tr>
<td>2</td>
<td>Network Voice Protocols</td>
<td>Danny Cohen, Univ. of Southern California Information Sciences Institute</td>
</tr>
<tr>
<td>3</td>
<td>Speech Conferencing</td>
<td>James W. Forgie, MIT Lincoln Laboratory</td>
</tr>
<tr>
<td>4</td>
<td>Wideband Integrated Voice/Data Network</td>
<td>Clifford J. Weinstein, MIT Lincoln Laboratory</td>
</tr>
<tr>
<td>5</td>
<td>Packet Radio Speech</td>
<td>Earl J. Craighill, SRI International</td>
</tr>
<tr>
<td>6</td>
<td>Algorithmic Achievements</td>
<td>John Makhoul and Vishu Viswanathan, Bolt Beranek and Newman, Inc.</td>
</tr>
<tr>
<td>7</td>
<td>Very Low Rate Vocoder</td>
<td>Richard Schwartz, Bolt Beranek and Newman, Inc.</td>
</tr>
<tr>
<td>8</td>
<td>Packet Voice Terminals and the LEXNET Local Network</td>
<td>Gerald C. O'Leary, MIT Lincoln Laboratory</td>
</tr>
<tr>
<td>9</td>
<td>Compact LPC Vocoder</td>
<td>Joel A. Feldman, MIT Lincoln Laboratory</td>
</tr>
<tr>
<td>10</td>
<td>Flexible Array Processors</td>
<td>Glen Culler, CHI Systems, Inc.</td>
</tr>
<tr>
<td>11</td>
<td>VLSI Array Processor</td>
<td>Edward Greenwood, Motorola Government Electronics Group</td>
</tr>
<tr>
<td>12</td>
<td>Single Chip LPC</td>
<td>Robert Brodersen, Univ. of California/Berkeley</td>
</tr>
</tbody>
</table>
# AGENDA FOR DARPA PACKET SPEECH MEETING

Lincoln Laboratory

3 June 1982

<table>
<thead>
<tr>
<th>TIME</th>
<th>TOPIC</th>
<th>SPEAKER</th>
</tr>
</thead>
<tbody>
<tr>
<td>0930-1100</td>
<td>Background, Goals, Key Results, and Demo</td>
<td>Adams (DARPA)</td>
</tr>
<tr>
<td>1100-1115</td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>1115-1135</td>
<td>Network Voice Protocols</td>
<td>Cohen (ISI)</td>
</tr>
<tr>
<td>1135-1150</td>
<td>Speech Conferencing</td>
<td>Forgie (Lincoln)</td>
</tr>
<tr>
<td>1150-1205</td>
<td>Wideband Network</td>
<td>Weinstein (Lincoln)</td>
</tr>
<tr>
<td>1205-1225</td>
<td>Packet Radio Speech</td>
<td>Craighill (SRI)</td>
</tr>
<tr>
<td>1230-1330</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>1330-1350</td>
<td>Summary of Key Algorithm Developments</td>
<td>Makhoul (BBN)</td>
</tr>
<tr>
<td>1350-1405</td>
<td>Very Low-Rate Speech</td>
<td>Schwartz (BBN)</td>
</tr>
<tr>
<td>1405-1425</td>
<td>Packet Voice Terminal</td>
<td>O’Leary (Lincoln)</td>
</tr>
<tr>
<td>1425-1445</td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>1445-1500</td>
<td>Compact LPC Vocoder</td>
<td>Feldman (Lincoln)</td>
</tr>
<tr>
<td>1500-1515</td>
<td>Flexible Array Processor</td>
<td>Culler (CHI Systems)</td>
</tr>
<tr>
<td>1515-1530</td>
<td>VLSI Array Processor</td>
<td>Greenwood (Motorola)</td>
</tr>
<tr>
<td>1530-1545</td>
<td>Single Chip LPC</td>
<td>Brodersen (U.C. Berkeley)</td>
</tr>
<tr>
<td>1545-1630</td>
<td>Future Plans and Discussion</td>
<td>Kahn (DARPA)</td>
</tr>
<tr>
<td>Name</td>
<td>Affiliation</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------</td>
<td></td>
</tr>
<tr>
<td>Abene, Capt. Peter</td>
<td>DCA</td>
<td></td>
</tr>
<tr>
<td>Adams, Lt. Col. Duane</td>
<td>DARPA/IPTO</td>
<td></td>
</tr>
<tr>
<td>Barna, Joseph</td>
<td>DCEC</td>
<td></td>
</tr>
<tr>
<td>Beauchemin, Edward</td>
<td>MIT-LL</td>
<td></td>
</tr>
<tr>
<td>Berger, Robert</td>
<td>MIT-LL</td>
<td></td>
</tr>
<tr>
<td>Bially, Ted</td>
<td>MIT-LL</td>
<td></td>
</tr>
<tr>
<td>Blankenship, Peter</td>
<td>MIT-LL</td>
<td></td>
</tr>
<tr>
<td>Blumenthal, Steven</td>
<td>BBN</td>
<td></td>
</tr>
<tr>
<td>Brodersen, Robert</td>
<td>U. California/Berkeley</td>
<td></td>
</tr>
<tr>
<td>Burchfield, Jerry</td>
<td>BBN</td>
<td></td>
</tr>
<tr>
<td>Casner, Stephen</td>
<td>ISI</td>
<td></td>
</tr>
<tr>
<td>Cohen, Dan</td>
<td>ISI</td>
<td></td>
</tr>
<tr>
<td>Cole, Randy</td>
<td>ISI</td>
<td></td>
</tr>
<tr>
<td>Cowles, Jack</td>
<td>Western Union</td>
<td></td>
</tr>
<tr>
<td>Craighill, Earl</td>
<td>SRI International</td>
<td></td>
</tr>
<tr>
<td>Crawford, Richard</td>
<td>DCA</td>
<td></td>
</tr>
<tr>
<td>Culler, Glen</td>
<td>CHI Systems, Inc</td>
<td></td>
</tr>
<tr>
<td>Deckelman, Francis</td>
<td>NAVELEX</td>
<td></td>
</tr>
<tr>
<td>Falk, Gilbert</td>
<td>BBN</td>
<td></td>
</tr>
<tr>
<td>Falk, Col. Harold</td>
<td>ESD</td>
<td></td>
</tr>
<tr>
<td>Feldman, Joel</td>
<td>MIT-LL</td>
<td></td>
</tr>
<tr>
<td>Forgie, James</td>
<td>MIT-LL</td>
<td></td>
</tr>
<tr>
<td>Forsdick, Harry</td>
<td>BBN</td>
<td></td>
</tr>
<tr>
<td>Friedhoffer, Carol</td>
<td>NSA</td>
<td></td>
</tr>
<tr>
<td>Fussell, Jesse</td>
<td>NSA</td>
<td></td>
</tr>
<tr>
<td>Gilbert, Warren</td>
<td>ESD</td>
<td></td>
</tr>
<tr>
<td>Glazim, Bernard</td>
<td>Western Union</td>
<td></td>
</tr>
<tr>
<td>Gold, Bernard</td>
<td>MIT-LL</td>
<td></td>
</tr>
<tr>
<td>Greenwood, Edward</td>
<td>Motorola</td>
<td></td>
</tr>
<tr>
<td>Heggstedt, Harold</td>
<td>MIT-LL</td>
<td></td>
</tr>
<tr>
<td>Hoversten, Estil</td>
<td>Linkabit</td>
<td></td>
</tr>
<tr>
<td>Johnson, Robin</td>
<td>OSD</td>
<td></td>
</tr>
<tr>
<td>Kahn, Robert</td>
<td>DARPA</td>
<td></td>
</tr>
<tr>
<td>Kang, George</td>
<td>NRL</td>
<td></td>
</tr>
<tr>
<td>Kantrowitz, William</td>
<td>MIT-LL</td>
<td></td>
</tr>
<tr>
<td>Kline, Lee</td>
<td>NRL</td>
<td></td>
</tr>
<tr>
<td>Krasner, Michael</td>
<td>BBN</td>
<td></td>
</tr>
</tbody>
</table>
Lippmann, Richard  
McCann, Albinas 
Lynn, Verne L.

Makhoul, John  
McCammon, Michael  
McElwain, Constance  
McLaughlin, Alan  
Moran, George  
Morris, Paul  
Morrow, Walter

Northrup, Richard  
O'Leary, Gerald

Parker, James  
Paul, Douglas  
Peatfield, C. Ross  
Peeler, Charles  
Proctor, William

Rettberg, Randy  
Roucos, Salim

Sandy, G. Ferrell  
Schecter, Harold  
Schwartz, Perry  
Schwartz, Richard  
Secrest, Bruce  
Segota, Anton  
Sharpe, Lt. Grady  
Shippee, Bernard  
Sonini, Lt. Col. Joseph  
Sues, Larry

Taylor, Lee  
Thomas, Robert  
Tierney, Joseph  
Tomlinson, Raymond  
Tremain, Thomas

Viswanathan, Vishu  
Weinstein, Clifford

Zimmermann, Lt. Col. Frank

MIT-LL  
Ft. Monmouth  
DARPA

BBN  
CHI Systems, Inc.  
MIT-LL  
MIT-LL  
DCEC  
Ft. Monmouth  
MIT-LL

RADC  
MIT-LL

MITRE  
MIT-LL  
MIT-LL  
Ft. Bragg  
CHI Systems, Inc.

BBN  
BBN

MITRE  
ESD  
MITRE  
BBN  
Texas Instruments, Inc.  
ESD  
ESD  
Western Union  
ESD  
RADC

MIT-LL  
BBN  
MIT-LL  
BBN  
NSA

BBN  
MIT-LL  
DCEC
Organization Addresses

BBN
Bolt Beranek and Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

CHI Systems, Inc.
CHI Systems, Inc.
150-A Aero Camino
Goleta, CA 93017

DARPA
DARPA/IPTO
Architect Building
1400 Wilson Boulevard
Arlington, VA 22209

DCA
Defense Communications Agency
4135 Courthouse Road
Arlington, VA 22209

DCEC
Defense Communications Engineering Center
1860 Wiehle Avenue
Reston, VA 22090

ESD
Electronic Systems Division
L. G. Hanscom Field
Bedford, MA 01731

FT. BRAGG
Dir., ADDS Testbed
Fort Bragg, NC 28307

FT. MONMOUTH
Hqs., U. S. Army
Communications Electronics Command
Ft. Monmouth, NJ 07703

ISI
Information Sciences Institute
University of Southern California
4676 Admiralty Way
Marina Del Rey, CA 90291
PACKETIZED SPEECH OVERVIEW
3 JUNE 1982

LT. COL. DUANE A. ADAMS
DARPA
OBJECTIVES

- Achieve point-to-point and conferenced narrowband digital speech through packet-switched networks
- Develop algorithms for degraded speech environments
- Demonstrate multi-user integrated data/voice transmissions
- Maintain compatibility with end-to-end security and internet protocols
- Develop low-cost packet speech hardware
MAJOR THRUSTS

- ALGORITHMS
- NETWORKS
- HARDWARE
SPEECH FREQUENCY RANGE: 0-4000 Hz
SAMPLING FREQUENCY: 8000 SAMPLES/SEC
PCM DATA RATE: 64,000 BITS/SEC
PACKETIZED SPEECH

VLR

100

LPC

1,000

APC

10,000

CVSD

100,000

PCM

100,000

PHONEMIC

FORMANT

TRACKING

SPEECH

MODELS

WAVEFORM

TRACKING

DIRECT

DIGITIZATION

DIGITAL SPEECH REPRESENTATIONS
WHY NARROWBAND?

- FACTOR OF 25 REDUCTION IN BANDWIDTH OVER PCM
- INTELLIGIBILITY VERY GOOD
- COMPUTATIONALLY FEASIBLE
- MULTIRATE PERMITS SERVICE UNDER OVERLOAD AND DEGRADED NETWORK CONDITIONS
PACKETIZED SPEECH

ANALYSIS OF SPEECH

INPUT SPEECH

SPECTRAL COEFFICIENTS

PITCH

$T = 10$ TO $20$ msec, $150$ SAMPLES OF SPEECH ASSUMED TO BE STATIONARY OVER $T$
PACKETIZED SPEECH

VOCODER ALGORITHMS

- LINEAR PREDICTIVE CODING (LPC)
- CHANNEL VOCODER
- HOMOMORPHIC VOCODER
- SPECTRAL ENVELOPE ESTIMATION
# Packetized Speech

## Control Packets

<table>
<thead>
<tr>
<th>Header</th>
<th>Call Setup</th>
</tr>
</thead>
</table>

## Data Packets

<table>
<thead>
<tr>
<th>Header</th>
<th>Digitized Speech</th>
</tr>
</thead>
</table>
PACKETIZED SPEECH

PACKET VOICE/DATA NETWORK
PACKETIZED SPEECH

NETWORK ISSUES

- NETWORK DELAYS
- VARIANCE IN PACKET DELIVERY
- LOST PACKETS AND PACKET REASSEMBLY
- NETWORK CONTROL STRATEGIES
- PACKET AGGREGATION
PACKETIZED SPEECH

ADVANTAGES OF PACKET SPEECH

- INTEGRATED VOICE AND DATA
- INTERNET
- SILENCE DETECTION
- VARIABLE RATE SPEECH
- BROADCAST/CONFERENCING
- SURVIVABLE SPEECH, IF DESIRED
PACKETIZED SPEECH

TRAFFIC: 2,700 ERLANGS VOICE TRAFFIC
36.15 MBPS DATA TRAFFIC

TOTAL MONTHLY COST ($M/MO)
(BACKBONE SWITCHING & TRANSMISSION)

2.4  8  9.6  16  64

VOICE DIGITIZATION RATE (KBPS)

42  58  62  80

TOTAL THROUGHPUT (MBPS)

15% 37% 42% 54% 83%

PERCENTAGE OF VOICE TRAFFIC
MAJOR MILESTONES

1974 CVSD DEMONSTRATED ON ARPA NET

1975 LPC DEMONSTRATED ON ARPA NET

1976 NVP PROTOCOLS DEVELOPED

1976 LPC CONFERENCE ON ARPA NET

1977 VARIABLE FRAME RATE LPC
MAJOR MILESTONES (CONT.)

1978  CVSD DEMONSTRATED ON PACKET RADIO NET
      SATNET CONFERENCING WITH LPC
      SPECTRAL ENVELOPE ESTIMATION ALGORITHM

1979  REAL-TIME VARIABLE RATE CHANNEL VOCODER
      INTERNET SPEECH DEMONSTRATED

1980  VOICE TERMINAL DEMONSTRATED ON LOCAL NET
MAJOR MILESTONES (CONT.)

1981 NVP-II/ST PROTOCOLS
CHI-5 ARRAY PROCESSOR DEVELOPED
LPC SPEECH ON PACKET RADIO NET
PCM ON WIDEBAND CHANNEL
1982

- COMPACT LPC VOCODER (NEC)
- LPC ON WIDEBAND NETWORK
- MULTI-RATE SPEECH ON WIDEBAND (PCM AND CVSD)
- INTERNET CONFERENCING
- MOBILE LPC ON PACKET RADIO NET
PACKETIZED SPEECH

INITIAL NETWORK LPC DEMONSTRATION
(December 1974)
SUMMARY OF
PACKET SPEECH DEMONSTRATION SEQUENCE
JUNE 3, 1982

LOCAL CALLS ON LINCOLN LEXNET

POINT-TO-POINT
1. PCM (64 KBPS)
2. ECVSD (16-64 KBPS)
3. LPC (2.4 KBPS) USING SINGLE-CARD LPC IN PVT

CONFERENCE
4. PCM 4-PARTY CONFERENCE

CALLS OVER WIDE BAND SATNET

POINT-TO-POINT
5. PCM LL - ISI
6. PCM LL - SWITCHED TELEPHONE NETWORK INTERFACE AT ISI
   (CALL TO LOCAL L.A. WEATHER)

CONFERENCE
7. 3-SITE, 4-PARTY LPC CONFERENCE
   LL (2 PARTIES ON LEXNET)
   ISI (ON LEXNET)
   SRI (ON PRNET UNIT IN SPEECH LAB, USING CHI-V LPC AND SIU)

POINT-TO-POINT (MOBILE)
8. LPC SRI (MOBILE PRNET UNIT) - LL (LEXNET PVT)
   MOBILE VAN RUN DEMONSTRATING PRNET SPEECH AND ALTERNATE ROUTING
TERMINAL OPTIONS

PACKET VOICE TERMINAL (PVT) (on LEXNET)
SPEECH INTERFACE UNIT (SIU) (on PRNET at SRI)
PACKET/CIRCUIT INTERFACE (PCI) (for Multiplexed PCM Interoffice Telephone Trunks)
SPEECH-45 (PDP-11/45 at ISI)

VOICE CODER OPTIONS

64 kbps $\mu$-LAW PCM
NEC SINGLE-CARD LPC
16 kbps CVSD CARD
VARIABLE-RATE EMBEDDED CVSD CARD (16-64 kbps)
LPC-10 IN CHI-5 ON SIU
LPC-10 IN FPS AP120B ON SPEECH-45
SWITCHED TELEPHONE NETWORK INTERFACE (STNI Card) (for Interfacing a Telephone Extension to a PVT)

INTERNETWORK OPTIONS

LEXNET (Lincoln Experimental Network)
WB SATNET (Wideband Satellite Network)
PRNET (Packet Radio Network)
Switched Telephone Net
LOCAL POINT-TO-POINT 64 kbps PCM CALL

PCM SPEECH PACKETS

144-bit HEADER

1440 SPEECH BITS (22.5 ms)

LEXNET

PVT

PCM CODEC

PVT

PCM CODEC
VARIABLE-RATE EMBEDDED CVSD CALL

PACKET PRIORITIES

<table>
<thead>
<tr>
<th>I</th>
<th>16 kbps CVSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>16-32 kbps</td>
</tr>
<tr>
<td>III</td>
<td>32-48 kbps</td>
</tr>
<tr>
<td>IV</td>
<td>48-64 kbps</td>
</tr>
</tbody>
</table>

CVSD PACKET AND EMBEDDED DIFFERENCE PACKETS

90 ms

PVT

ECVSD CARD

PVT

ECVSD CARD
LOCAL POINT-TO-POINT 2.4 kbps LPC CALL

LEXNET

LPC SPEECH PACKETS

144-bit HEADER
(Ten 22.5-ms Parcels)

480 SPEECH BITS

PVT

LPC CARD

PVT

LPC CARD
POINT-TO-POINT PCM CALL USING WIDEBAND CHANNEL

PVT GATEWAY

PSAT

WB SATNET

PSAT

PVT

DELAY 22.5 ms

45 75 157.5 270 45 45

+5 ms

+25 +20 +10

ACCUMULATED VARIATION

TOTAL MEAN DELAY 640 ms
LPC CALL, LINCOLN TO SRI MOBILE TERMINAL
NETWORK VOICE PROTOCOLS

Danny Cohen
The History of Network Voice Protocols

In the beginning ARPA created the ARPANet. Now the Network was formless and empty, darkness was over the surface of the Deep...

And LGR said "Let there be protocols!" and there was NCP. ARPA saw that the NCP was good and said "Use it for Packet Voice!"
Well, ..... NCP was not exactly the right Host/Host protocol for online speech.

The NCP provided more favors than online-voice needed, at a non-trivial expense.

Therefore the need for NVP was recognized.
The NCP was not the only friendly agent doing unwanted favors —
—So were the IMPs, with their reliability trick: Packets of Type 0
We had a problem...
The Arpanet protocols were not designed for real-time applications.
So we had to think about it until the next meeting...
Objectives for NVP

* Real time data
* Extensible (vocoder indep)
* Network independent
* Separation of Data/Control
NVP-0 was the the first voice protocol, a precursor to NVP-1. It was used for some CVSD experiments.
NVP-I was used for most of the experiments: CVSD, LPC-I, LPC-II
NVP-I was extended to support conferencing NVCP and voice-mail
When the PRNet was born we were expected to run online realtime voice to it by using TCP.
The thought of using TCP for this purpose was so overwhelming that the original TCP became IP+TCP
NVP-I was modified to run on top of IP, and was also improved in some areas. It became NVP-II
However, even the new team of NVP-II/IP has some problems with satellite links. ST was born....
LESSONS

Packet communication is practical for online speech.
Real-time data needs a different kind of protocol than the standard computer communication (reliability, flow control....)
Motivated people and small committees can accomplish a lot.
Milestones in Packet Speech Communication

8/74: CVSD/ARPANet: ISI + LL
12/74: LPC/ARPANet: LL + CH1

1/76: LPC-Conference/ARPANet: ISI + LL + CH1 + SRL

4/77: Flanagan (BTL) applies for a patent on packet transmission of speech

7/78: USA patent 4,100,377 granted
SUMMARY

The interest of the various phone companies in packet voice is a significant testimonial to our success.
SPEECH CONFERENCING

JAMES W. FORGIE
MIT LINCOLN LABORATORY
OUTLINE

1. VOICE CONFERENCING SYSTEM DESIGN ISSUES

2. HISTORY OF PACKET CONFERENCING

3. CURRENT WB NET CONFERENCING SYSTEM
VOICE CONFERENCING SYSTEM DESIGN CHOICES

1. SUMMATION OR SIGNAL SELECTION

2. VOICE OR PUSH-BUTTON CONTROL

3. CENTRALIZED OR DISTRIBUTED CONTROL
<table>
<thead>
<tr>
<th>PACKET CONFERENCE DEMONSTRATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SITES</strong></td>
</tr>
<tr>
<td><strong>PACKET ADDRESSING</strong></td>
</tr>
<tr>
<td><strong>CONTROL TECHNIQUES</strong></td>
</tr>
<tr>
<td><strong>NETWORKS</strong></td>
</tr>
<tr>
<td><strong>I</strong></td>
</tr>
<tr>
<td><strong>II</strong></td>
</tr>
<tr>
<td><strong>III</strong></td>
</tr>
<tr>
<td><strong>IV</strong></td>
</tr>
<tr>
<td><strong>V</strong></td>
</tr>
</tbody>
</table>
 PACKET CONFERENCING EXPERIMENTS
SATNET/INTERNET VOICE CONFERENCING

ATLANTIC SATNET

LL (Lexington, MA)

ARPANET

LCC

BBN (Cambridge, MA)

CHAIRMAN

LCC

CCP

ISI (Los Angeles, CA)

LCC = LOCAL CONFERENCE CONTROLLER

CCP = CONFERENCE CONTROL PROGRAM
ADVANTAGES OF PACKET TECHNIQUES FOR CONFERENCING

1. EFFICIENT USE OF COMMUNICATION RESOURCES

2. CONCURRENT EXCHANGE OF VOICE AND CONTROL MESSAGES

3. EASY AUGMENTATION FOR MULTIMEDIA CONFERENCING

4. SIMPLE IMPLEMENTATION OF DISTRIBUTED CONTROL FOR SURVIVABILITY
WIDEBAND INTEGRATED VOICE/DATA NETWORK

CLIFFORD J. WEINSTEIN
MIT LINCOLN LABORATORY
WIDEBAND INTEGRATED VOICE/DATA NETWORK

• MOTIVATION FOR WIDEBAND EXPERIMENT

• EXPERIMENTAL SYSTEM DESCRIPTION

• CHANNEL ASSIGNMENT AND PACKET MULTIPLEXING

• EXPERIMENT STATUS
WIDEBAND INTEGRATED VOICE/DATA NETWORK

EARTH STATION ANTENNA

PACKET SATELLITE DEMAND ASSIGNMENT PROCESSOR AND BURST MODEM

TO OTHER NETWORKS

VOICE/DATA TRAFFIC CONCENTRATOR

LOCAL NETWORK (LEXNET)

PVT

TRAFFIC EMULATOR

PACKET VOICE TERMINAL
PODA FRAME

INFORMATION SUBFRAME

CONTROL SUBFRAME

CENTRALIZED ASSIGNMENT

DISTRIBUTED ASSIGNMENT

REQUESTS AND SYNC

GUARD INTERVAL

INFORMATION BURSTS FROM VARIOUS PSATs

DATAGRAMS OR AGGREGATED STREAM PACKETS FROM VARIOUS CONCENTRATORS

BURST PREAMBLE AND CONTROL

MESSAGE No. 1

MESSAGE No. 2

MESSAGE No.N

DATA OR VOICE PACKETS FROM VARIOUS LOCAL NETS AND TERMINALS

HEADER  DATAGRAM  OR  AGGREGATED HEADER  VOICE PACKET No.1  ...  VOICE PACKET No.M
CONCENTRATOR/GATEWAY

FUNCTIONS

• GATEWAY BETWEEN WB SATNET AND LEXNETs, PRNET, ARPANET, EDN

• STREAM (ST) PROTOCOL FOR EFFICIENT PACKET SPEECH

• DoD STANDARD INTERNET PROTOCOL (IP) FOR DATA AND CONTROL

IMPLEMENTATIONS

• MINICONCENTRATOR (LINCOLN)
  - PDP-11/44 WITH Z80-BASED PACKET PROCESSOR
  - DEVELOPMENT TESTBED FOR INTERNET STREAM PROTOCOLS

• VOICE FUNNEL (BBN)
  - MULTIPLE MC68000 PROCESSES CONNECTED VIA BUTTERFLY SWITCH
  - VERY HIGH THROUGHPUT CAPACITY (2-20 MBPS)
PACKET SPEECH EXPERIMENT STATUS – JUNE 1982

EARTH STATION INTERFACE

PACKET SATELLITE
DEMAND ASSIGNMENT
PROCESSOR

G

VOICE FUNNEL

F

WB SATNET

ESI
PSAT

ESI
PSAT

ESI
PSAT

ESI
PSAT

G

PDP-11/44
GATEWAY

PCI

PACKET/CIRCUIT
INTERFACE

TELEPHONE
OFFICE
EMULATOR

ARPANET CONNECTIONS
NOT SHOWN
WIDEBAND NETWORK SPEECH EXPERIMENT MILESTONES

18 NOV 81  LL TO ISI (PCM, 2 CONVERSATIONS, ONE USING STNI AT ISI)
22 JAN 82  PCI – PCI VIA SATELLITE  
            PCI – LEXNET  
            2 HOSTS ON LL PSAT  
            2 CONVERSATIONS
3 FEB 82  FUNNEL AT LL TO GATEWAY AT ISI (PCM)
22 FEB 82  PCI AT LL TO PCI AT DCEC (2 CONVERSATIONS)
4 MAR 82  3-SITE SPEECH (LL, ISI, DCEC) IN 2-PARTY COMBINATIONS
9 APR 82  FUNNEL AT LL TO FUNNEL AT ISI
22 APR 82  SRI (LEXNET) TO ISI (LEXNET)
7 MAY 82  SIMULTANEOUS LPC CALLS
            PVT (LL) – PRNET (CHI-V) AT SRI
            PVT (LL) – LEXNET (AP120B) AT ISI
13 MAY 82  2-SITE (LL/ISI), 3-PARTY PCM CONF
21 MAY 82  3-SITE LPC CONF: PVTs AT LL, ISI, SRI
1 JUNE 82  3-SITE LPC CONF: LEXNET PVTs AT LL, ISI; PRNET AT SRI
FUTURE DIRECTIONS

• MULTI-USER, MULTI-SITE EXPERIMENTS
  - MULTIPLEXING TO TEST SYSTEM LIMITS
  - ADAPTIVE TRAFFIC CONTROL OF VOICE AND DATA
  - INTERNET VOICE CONFERENCING
  - VOICE-CONTROLLED SYSTEMS

• PACKET VIDEO DEVELOPMENT/EXPERIMENTS

• SECURE PACKET SPEECH SYSTEM DEVELOPMENT/EXPERIMENTS
PACKET RADIO SPEECH

EARL J. CRAIGHILL
SRI INTERNATIONAL
EXPERIMENTAL WIDEBAND VOICE/DATA SATELLITE NETWORK

Subsystems at Each Satellite Node
- Earth Station (RF)
- Flexible Burst Modem
- Demand Assignment Processor

FEATURES
- Multi-user Voice/Data
- Satellite/Terrestrial
- Distributed Control
- Internet/Security Protocols
- Mobile Access for Voice
- Flexibility for Experiments
THE PACKET RADIO UNIT (PRU)
THE COMMON NETWORK ELEMENT

Terminal Device
Repeater/Packet Radio Unit
Station

THE PACKET RADIO UNIT

TERMINAL DEVICE
HIGH SPEED PORT
DIGITAL UNIT
POWER SUPPLY
PACKET RADIO CONCEPTS

- BROADCAST RADIO
- PACKET-SWITCHED
- MULTIPLE ACCESS
- MULTIPLE-HOP
- MOBILE OPERATION
- AUTOMATED NETWORK MANAGEMENT AND CONTROL
REAL TIME TRANSPORT

Microphone → Vocoder → Packet (H, PL) → Network → Receiver Queue → Vocoder → Speaker

- Processing Delay
- Packet Loading Time
- Network Transient Time: 340 ms
- Receiver Buffering Time: 150 ms
- Processing Delay: 5 ms

ETE Delay

Through put = (PL-H)/T

Reliability

- Transmit Queue Trimming
- Network Nodes Discard Packets
- Receiver Discards Packet if "Too Late"
TRANSMITTER FLOW CONTROL

Variable Packet Length
Controls the offering rate of packets to the network

Header

135 ms Speech
New Packet Length = 720 b

Transmitter Queue

Empty

338 ms Speech
New Packet Length = 1296 b

Disgard Oldest

585 ms Speech
New Packet Length = 2000 b

PR net
Voice Transport Performance in PRnets

<table>
<thead>
<tr>
<th>Protocol Version</th>
<th>Packet Length</th>
<th>Vocoder</th>
<th>Percent Packet Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>Fixed</td>
<td>CVSD</td>
<td>44.4% — 47.9%</td>
</tr>
<tr>
<td>Data</td>
<td>Variable</td>
<td>CVSD</td>
<td>35.4% — 42.2%</td>
</tr>
<tr>
<td>Voice</td>
<td>Fixed</td>
<td>CVSD</td>
<td>8.6% — 17.7%</td>
</tr>
<tr>
<td>Voice</td>
<td>Variable</td>
<td>CVSD</td>
<td>4.7% — 9.3%</td>
</tr>
<tr>
<td>Voice</td>
<td>Variable</td>
<td>LPC</td>
<td>1.8% — 3.1%</td>
</tr>
</tbody>
</table>
Need Type of Service to Support Data/Voice on the Same Network

- Different Requirements for Data and Voice
- Network Transport Parameters Changed on a Packet-by-Packet Basis
- Adjustments for Dynamic Network Conditions
ALGORITHMIC ACHIEVEMENTS

JOHN MAHKOUL,
VISHU VISWANATHAN
BOLT BERANEK AND NEWMAN INC.
MAJOR CONTRIBUTIONS

- Fundamental Understanding of the Basic Theoretical Issues of Speech Coding at Data Rates Ranging From 64,000 bits/s Down to 100 bits/s
- Development of Practical Coding Algorithms That Maximize Speech Quality and Intelligibility
- Dissemination of Information to DoD Agencies and Scientific Community at Large
 COMPONENTS OF A SPEECH COMPRESSION SYSTEM

- **Transmit**
  - Analyzer: $x(t)$
  - Encoder: $y(t)$
  - Transmission Channel: $y'(t)$
  - Decoder: $x'(t)$
  - Synthesizer: $s'(t)$
- **Receive**
  - Sampled Speech: $s(t)$
  - Receiver: $s'(t)$

The diagram illustrates the process of digitizing speech and transmitting it through a channel before reconstructing it at the receiver.
<table>
<thead>
<tr>
<th>TYPE</th>
<th>BIT - RATE</th>
<th>REMARKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIGH BIT RATE CODER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 - BIT LOG PCM</td>
<td>64 KBITS/S</td>
<td>HIGH QUALITY SPEECH</td>
</tr>
<tr>
<td>ADAPTIVE PREDICTIVE CODERS (APC)</td>
<td>16 KBITS/S</td>
<td>HIGH QUALITY SPEECH</td>
</tr>
<tr>
<td>BASEBAND CODERS (BBC) OR VOICE - EXCITED CODERS</td>
<td>9.6 KBITS/S</td>
<td>COMMUNICATIONS QUALITY, WIRELINE COMMUNICATIONS</td>
</tr>
<tr>
<td>LOW BIT RATE CODERS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPC, CHANNEL VOCODER, ETC.</td>
<td>1.5 - 2.4 KBITS/S</td>
<td>HF COMMUNICATIONS</td>
</tr>
<tr>
<td>VECTOR QUANTIZATION</td>
<td>400 - 800 BITS/S</td>
<td>NAVY AND MINING SYSTEM APPLICATIONS</td>
</tr>
<tr>
<td>SEGMENT VOCODER</td>
<td>150 - 300 BITS/S</td>
<td>LOW TRANSMISSION POWER.</td>
</tr>
</tbody>
</table>
A GENERAL SYNTHESIS MODEL

\[ x'(t) \]

SOURCE OR EXCITATION FUNCTION

EXCITATION SIGNAL \[ u(t) \]

SPECTRAL SHAPING FILTER

SYNTHETIC SPEECH \[ s'(t) \]
SOURCE MODELS FOR LOW BIT RATE VocoderS

1. Pulse/Noise Model

- Pulse Source for Voiced Sounds
- Random Noise for Unvoiced Sounds
- Developed Robust Pitch and Voicing Extraction Algorithms
- Model too Idealized for Certain Sounds — Buzzy Speech Quality

2. Mixed-Source Model

- Uses Pulse and Noise Sources Simultaneously
- Eliminates Buzzy Quality and Produces More Natural-Sounding Speech
- Application in Phonetic Synthesis and Text-to-Speech Systems
SPECTRAL MODELS

- Linear Prediction (All-Pole) Model: LPC Vocoder
- Filter-Bank Model: Channel Vocoder
- Cepstral Model: Homomorphic Vocoder
- Formant Model: Formant Vocoder
- Spectral Amplitude Representation: Spectral Envelope Estimation Vocoder
- Pole-Zero Model

WHAT WE LEARNED:

An Accurate Representation of the Short-Term Speech Spectral Envelope, Especially at Spectral Energy Peaks, Is Necessary for Good-Quality Speech Synthesis
Each Speech Sample is Approximated by a Weighted Linear Summation of Previous Speech Samples:

\[ s(n) \approx \sum_{k=1}^{p} a(k) s(n-k). \]

**ANALYSIS**

\[ A(z) = 1 - \sum_{k=1}^{p} a(k) z^{-k} \]

\[ e(n) = s(n) - \sum_{k=1}^{p} a(k) s(n-k) \]

**SYNTHESIS**

\[ H(z) = \frac{1}{A(z)} = \frac{1}{1 - \sum_{k=1}^{p} a(k) z^{-k}} \]

\[ s'(n) \]
SPECTRAL MODELING BY LINEAR PREDICTION

\[ \hat{P}(\omega) \]

[\varepsilon]

\[ p = 14 \]

RELATIVE ENERGY (dB)

FREQUENCY (kHz)

25 ms

00250
CONTRIBUTIONS
IN LINEAR PREDICTION

1. Fundamental Understanding of Linear Prediction
   – Relation Between Time Domain and Frequency Domain
   – Spectral Matching Properties

2. Lattice Analysis and Synthesis Structures
   – Filter Stability
   – Quantization and Roundoff Properties
   – Hardware Implementation
   – Adaptive Algorithms

3. Fixed-Point Implementation

4. Results Used In Other Fields
QUANTIZATION OF SPECTRAL PARAMETERS

- Quantization Properties of Different Parametric Representations
  - Reflection Coefficients \( \{ K_i \} \)
- The Notion of Spectral Sensitivity for Optimal Quantization
  - Log Area Ratios \( \{ L_i \} \)
  - Arc Sine Parameters

\[ L_i = \log \frac{A_{i+1}}{A_i} = \log \frac{1+K_i}{1-K_i} \]
VARIABLE FRAME RATE TRANSMISSION

IDEA:

- Coder Parameters Are Transmitted Only if They Have Changed Sufficiently Since the Last Transmission

METHOD:

- Transmit Selected Frames of Parameters, with the Selection Made Independently for Pitch, Gain and Spectral Parameters
- Reconstruct the Untransmitted Parameters by Linear Interpolation
VARIABLE FRAME RATE TRANSMISSION - RESULTS

<table>
<thead>
<tr>
<th></th>
<th>Bits/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Fixed Frame Rate</td>
<td></td>
</tr>
<tr>
<td>100 Frames/s</td>
<td>5600</td>
</tr>
<tr>
<td>2. Variable Frame Rate</td>
<td></td>
</tr>
<tr>
<td>Average: 31 Frames/s</td>
<td>2075</td>
</tr>
</tbody>
</table>

RESULT:
Systems 1 and 2 Produce the Same Speech Quality
ROBUST PERFORMANCE
UNDER ACOUSTIC BACKGROUND NOISE

THREE APPROACHES:

1. Noise-Cancelling Microphones (Sponsored by Other DoD Agencies)

2. Two-Microphone Input for Adaptive Noise Cancelling

3. Single Microphone Input
   - Preprocessing by Spectral Subtraction Method
   - Preprocessing Enhances the Quality and Intelligibility of Vocoded Speech in Acoustic Background Noise
ROBUST PERFORMANCE UNDER TELEPHONE SPEECH INPUT

- Frequency Distortion Caused by Telephone Access to the Packet Voice Network
- A Blind Deconvolution Approach to Remove the Effect of the Telephone Channel
- Adaptive Filtering Techniques to Reduce the Effect of Echo on Vocoder Speech
MULTIRATE SPEECH CODING

- Transmission at a Wide Range of Bit Rates to Accommodate Varying Traffic Loads on the Packet Network
- Embedded-Code Systems
- Specific Designs Considered:
  - CVSD/PCM (16 to 64 kbits/s)
  - LPC Vocoder/Adaptive Transform Coder (2.4 to 16 kbits/s)
  - Channel Vocoder/Sub-Band Coder (2.4 to 9.6 kbits/s)
LOW BIT RATE VOCODERS

1. Scalar Quantization of Spectral Parameters (1.5-2.4 kbits/s)
   - Independent Quantization of Each Parameter

2. Vector Quantization of Spectral Parameters (400-800 bits/s)
   - Parameters Quantized as a Vector in Multi-Dimensional Space
   - Use Clustering to Determine Quantization Templates

3. Segment Quantization (150-300 bits/s)
   - Quantize a Sequence of Spectra as One Unit

4. Phonetic Vocoder (100 bits/s)
   - Use Speech Recognition Techniques
STATE-OF-THE-ART
IN SPEECH COMPRESSION

• Virtually Transparent Speech Quality at 16 kbits/s

• Quality and Intelligibility Degrade With Decreasing Bit-Rate, Especially
  – Below 4.8 kbits/s
  – For Female Speech
  – Under Acoustic Noise and Channel Distortion Conditions

• Substantially Improved Intelligibility and Quality of 2.4 kbits/s Vcoders
AREAS OF FURTHER RESEARCH

PROBLEM: How to Maximize Speech Quality and Intelligibility for Every bit That We Use in Transmission?

- Source Modeling
  - Pitch and Voiced/Unvoiced Decision

- Spectral Modeling
  - Female Speech

- Pattern Recognition and Information Theoretic Coding Techniques

- Understanding and Modeling of the Effects of Coding Techniques on Human Speech Perception
VERY LOW RATE VOCODER

RICHARD SCHWARTZ
BOLT BERANEK AND NEWMAN INC.
VERY LOW RATE VOCODER

• Requirements
  – 100-200 bits/s
  – Intelligibility in a Conversation
  – Naturalness

• Applications
  – Low Power Communication
  – Degraded Channel Conditions
  – "Burn Through" Jamming Networks
  – Speech Storage
VERY LOW RATE VOCODERS

1. Phonetic Vocoder ~100 bits/s

2. Segment Vocoder ~150-300 bits/s
PHONETIC Vocoder

"BAT"

DURATION: B 70 ms  Á 130 ms  T 80 ms

Number of Phonemes in English: 42
Typical Speaking Rate: 12 Phonemes/sec.

SPEECH → ANALYSIS (Phonemes, Durations, Pitch) → TRANSMISSION CHANNEL ≈ 100 BITS/S → PHONETIC SYNTHESIS (Phonemes, Durations, Pitch) → SYNTHESIZED SPEECH
FEASIBILITY STUDY

• Approach
  – Feature-Based Phonetic Recognition
  – Phonetic Synthesis-by-Rule

• Conclusions
  – Over 80% Correct Phonetic Recognition Is Necessary for Intelligibility in Context
  – Phonetic Synthesis Must be Improved to Obtain Desired Natural Speech Quality
DIPHONE SYNTHESIS

"BAT"

PHONEMES

ENERGY

DIPHONES

TIME

B - Ā

ā - T

SPEECH IS SYNTHESIZED BY CONCATENATING DIPHONE TEMPLATES
DIPHONE RECOGNITION NETWORK

\[ \text{TAP} = \hat{T \hat{A} \hat{P}} \]

\begin{align*}
\text{TAP} & \quad \text{PAT} \\
\text{PAT} & \quad \text{TIP} \\
\text{TIP} & \quad \text{PIT} \\
\text{PIT} & \quad \text{TREE} \\
\text{TREE} & \quad \text{TREAT} \\
\text{TREAT} & \quad \text{TRIP} \\
\text{TRIP} & \quad \text{RIP} \\
\text{RIP} & \quad \text{IE} \\
\text{IE} & \quad \text{T} \\
\text{T} & \quad \text{A} \\
\text{A} & \quad \text{I} \\
\text{I} & \quad \text{R} \\
\text{R} & \quad \text{T/R} \\
\text{T/R} & \quad \text{TAP} 
\end{align*}
PHONETIC VOCODER

- Requires Phonetic Recognition
  - Still a Difficult Problem
- Requires Manual Labeling of Diphones
  - Labor Intensive
SEGMENT VOCODER

- A Segment Is a "Diphone-Like" Sequence of Spectra Between Two Steady State Sounds
- Segment Templates Are Selected Automatically (Unsupervised)
  - There Is No Phonetic Labeling
- Result: Segment Vocoder Speech Is Intelligible at About 150 bits/s
AUTOMATIC TEMPLATE SELECTION

- Training Data: 15 Minutes of Unconstrained Speech
- Segmentation Algorithm:
  - Divides Training Speech Into "Diphone-Like" Segments (15 Minutes Contains About 8,000 or $2^{13}$ Segments)
- Random Quantizer
  - All Segments are Used as Templates
- Template Network:
  - Determines Constraints on Template Sequences by Requiring Spectral Continuity Between Successive Templates
  - Only 1024 or $2^{10}$ Templates Can Follow Any Given Template
# BIT ALLOCATION FOR SEGMENT VOCODER

<table>
<thead>
<tr>
<th></th>
<th>Bits/Segment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segment Template Code</td>
<td>10</td>
</tr>
<tr>
<td>Segment Duration</td>
<td>$1^{1/2}$</td>
</tr>
<tr>
<td>Pitch</td>
<td>1</td>
</tr>
<tr>
<td>Energy</td>
<td>$1^{1/2}$</td>
</tr>
</tbody>
</table>

| Total Bits/Segment       | 14           |
| Average Segment Rate     | 11 per sec   |

Average Bit Rate = $11 \times 14 = 154$ bits/s
PACKET VOICE TERMINALS
AND
THE LEXNET LOCAL NETWORK

GERALD C. O’LEARY
MIT LINCOLN LABORATORY
PACKET VOICE TERMINAL AND LEXNET

• PVT REQUIREMENTS

• ARCHITECTURE

• SPEECH PROCESSOR OPTIONS

• LEXNET

• CURRENT STATUS

• FUTURE DIRECTIONS
PACKET VOICE TERMINAL REQUIREMENTS

• ACCOMMODATE PCM AND NARROWBAND SPEECH PROCESSORS

• IMPLEMENT PACKET VOICE PROTOCOLS

• SUPPORT CONNECTION TO DIFFERENT PACKET NETS

• TELEPHONE-LIKE USER INTERFACE

• SOURCE OF PACKET SPEECH TRAFFIC FOR WIDEBAND PACKET VOICE EXPERIMENTS
TERMINAL DESIGN

• MICROPROCESSOR CONTROLLED MODULES (8085 SERIES)

• FUNCTIONAL PARTITIONING FOR FLEXIBILITY
  (A) LOCAL NET INTERFACE (2K ROM, 2K RAM)
  (B) PROTOCOL PROCESSOR (2K ROM, 10K RAM, 32K RAM/ROM)
  (C) PCM CODEC AND VOCODER INTERFACE
  (D) TELEPHONE INSTRUMENT
  (E) NARROWBAND VOCODER (INTERNAL/EXTERNAL)

• PROTOCOL PROCESSOR SUPPORTS NVP-11 AND ST PROTOCOLS AND CONFERENCING (~ 28K BYTES)

• INTERFACE TO LINCOLN EXPERIMENTAL NETWORK (LEXNET)
  (A) EXPLOITS LOCAL NETWORK TECHNOLOGY (SIMILAR TO ETHERNET)
  (B) CSMA/CD TRANSMISSION PROTOCOL

• 200 INTEGRATED CIRCUITS, 40 WATTS, 8-1/2" X 10-1/2" X 16-1/2" PACKAGE
SPEECH PROCESSOR OPTIONS

• 64 KBIT/SEC PCM

• 16 - 64 KBIT/SEC EMBEDDED CVSD

• 2400 BIT/SEC LPC

• SWITCHED TELEPHONE NETWORK INTERFACE CARD

• NETWORK TIMING MEASUREMENTS CARD
EMBEDDED VOCODER

- VARIABLE RATE VOCODER
- RATE-QUALITY TRADE-OFF
- GENERATE PACKETS OF DIFFERENT PRIORITIES
- ASSUMES A NETWORK WHICH DISCARDS LOW PRIORITY PACKETS WHEN CONGESTED
- RECEIVER WILL MAKE BEST RECONSTRUCTION WITH WHAT IS RECEIVED
- EXPERIMENTAL REALIZATION

  16 KBPS CVSD MINIMUM RATE
  ERROR SIGNAL QUANTIZED TO 2, 4, OR 6 BITS
  RATES 16, 32, 48, OR 64 KBPS
  FORMS PACKETS OF DIFFERENT PRIORITIES
LINCOLN EXPERIMENTAL PACKET VOICE NETWORK (LEXNET)

- CONNECTS PVTs TO FORM DISTRIBUTED TELEPHONE SYSTEM
- CONCENTRATOR INTERFACE BOX PROVIDES HIGH RATE CONNECTION TO SATELLITE CHANNEL
- CONFERENCE ACCESS CONTROLLER
- LINCOLN'S NET SPANS ABOUT 1000 FEET
- 1 MBIT/SEC ON RG-59/U COAXIAL CABLE
SUMMARY

• CURRENT STATUS

NETWORKS AT LINCOLN (2),
ISI
SRI
DCEC
10 TERMINALS

• FUTURE DIRECTIONS

INTEGRATE PACKET VOICE CAPABILITY TO OTHER
DoD DATA NETWORKS
SECURE SPEECH TERMINALS
TECHNOLOGY TRANSFER
VOICE CONTROLLED CONFERENCE SET UP
NETWORK TIMING MEASUREMENTS
COMPACT LPC VOCODER

JOEL A. FELDMAN
MIT LINCOLN LABORATORY
PVT SINGLE BOARD 2.4 KBPS LPC VOCODER

- HARDWARE

- ARCHITECTURE

- PROCESSORS' TASKS

- PROCESSORS' REAL-TIME AND MEMORY USAGE

- INITIALIZATION OPTIONS
COMPACT VOCODER 'STREAMED' PROGRAM STRUCTURE

ONE SAMPLING INTERVAL

TIME

SAMPLE N
FOREGROUND PROGRAM ACTIVE
NOMINAL LOAD
SAMPLE N + 1
NOMINAL LOAD
SAMPLE N + 2
WORST CASE LOAD
SAMPLE N + 3
NOMINAL LOAD
SAMPLE N + 4
NOMINAL LOAD

BACKGROUND PROGRAM ACTIVE
LPC ANALYZER TASKS (1 N.E.C. μPD7720)

- EXECUTED EACH SAMPLE (Interrupt Driven Foreground)
  
  (a) INPUT SAMPLE, MU-LAW TO LINEAR CONVERSION, HAMMING WINDOW, DOWNSCALE, UPDATE CORRELATION COEFFICIENTS
  
  (b) CHECK FOR FRAME MARK FROM CONTROL PROCESSOR
  
  (c) IF FRAME TIME, ACTIVATE BACKGROUND ROUTINE AND SWAP OUTPUT PARAMETER BUFFERS

- EXECUTED EACH FRAME (Background)
  
  (a) BLOCK FLOATING POINT CORRELATION COEFFICIENTS AND COMPUTE REFLECTION COEFFICIENTS
  
  (b) TRANSFER REFLECTION COEFFICIENTS AND ENERGY ESTIMATE TO CONTROL PROCESSOR
GOLD PITCH DETECTOR TASKS (1 N.E.C. μPD7720)

- EXECUTED EACH SAMPLE (Interrupt Driven Foreground)
  
  (a) INPUT SAMPLE, MU-LAW TO LINEAR CONVERSION, LOW-PASS PRE-FILTER, PEAK-DETECTION, UPDATE SIX PARALLEL PITCH PERIOD ESTIMATES
  
  (b) CHECK FOR FRAME MARK FROM CONTROL PROCESSOR
  
  (c) IF FRAME TIME, ACTIVATE BACKGROUND ROUTINE

- EXECUTED EACH FRAME (Background)
  
  (a) PITCH PERIOD SCORING AND FINAL VOICING DECISION AND PITCH ESTIMATION
  
  (b) OUTPUT VOICING DECISION AND PITCH ESTIMATE TO CONTROL PROCESSOR
LPC SYNTHESIZER TASKS (1 N.E.C. \(\mu\)PD7720)

- EXECUTED EACH SAMPLE (Interrupt Driven Foreground)
  
  (a) UPDATE EXCITATION GENERATOR, UPDATE LATTICE FILTER, OUTPUT SAMPLE

  (b) CHECK FOR FRAME MARK FROM CONTROL PROCESSOR

  (c) IF FRAME TIME, ACTIVATE BACKGROUND ROUTINE AND SWAP INPUT PARAMETER BUFFERS

- EXECUTED EACH FRAME (Background)

  (a) CONVERT ENERGY ESTIMATE TO PITCH PULSE AMPLITUDES (Voiced Frames) OR PSEUDO-RANDOM NOISE AMPLITUDES (Unvoiced Frames)

  (b) INPUT PITCH/VOICING PARAMETER, ENERGY ESTIMATE AND REFLECTION COEFFICIENTS FROM CONTROL PROCESSOR
CONTROL MICROCOMPUTER TASKS (1 INTEL 8085)

• EXECUTED EACH FRAME

(a) GIVE FRAME MARK TO ANALYZER, PITCH DETECTOR AND SYNTHESIZER

(b) INPUT ANALYSIS PARAMETERS FROM ANALYZER AND PITCH DETECTOR, CODE TO 2400 bps.

(c) INPUT SYNTHESIS PARAMETERS FROM HOST TERMINAL, DECODE TO 16-bit VALUES, OUTPUT TO SYNTHESIZER
<table>
<thead>
<tr>
<th>Percentage of:</th>
<th>Program ROM (512 x 23)</th>
<th>Data RAM (128 x 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gold Pitch Detector</td>
<td>Linear Predictive Analysis</td>
</tr>
<tr>
<td></td>
<td>95%</td>
<td>70%</td>
</tr>
<tr>
<td></td>
<td>3%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>10%</td>
<td>0%</td>
</tr>
</tbody>
</table>
LPC VOCODER REAL-TIME USAGE

(a) 10th ORDER LINEAR PREDICTIVE MODEL
(b) 8-kHz SAMPLING FREQUENCY
(c) 22.5-ms SPEECH FRAMES

PERCENTAGE OF REAL-TIME USED

GOLD PITCH DETECTOR  ≈35%
LINEAR PREDICTIVE ANALYSIS  63%
SYNTHESIS  46%
CONTROL MICROCOMPUTER

REAL-TIME AND MEMORY REQUIREMENTS

- 67% REAL-TIME (22.5 ms Frames, 8 MHz X-TAL)
- 2013 BYTES ROM
- 139 BYTES RAM
COMPACT LPC VOCODER INITIALIZATION OPTIONS

- LINEAR PREDICTIVE MODEL ORDER (up to 15)
- ANALYSIS AND SYNTHESIS FRAME SIZE
- SPEECH SAMPLING FREQUENCY
- SPEECH INPUT/OUTPUT CODING FORMAT
  (a) 16-bit LINEAR
  (b) 8-bit MU-255 LAW
- PERFORMANCE OPTIMIZATION FOR INPUT SPEECH BACKGROUND NOISE LEVEL
SUMMARY

• VERY FLEXIBLE, COMPACT LPC Vocoder based on a commercial signal processing microcomputer

• IMPLEMENTATION ACHIEVES ORDER OF MAGNITUDE DECREASE IN
  (a) POWER DISSIPATION
  (b) INTEGRATED CIRCUIT AREA
  (c) PRODUCTION COST

• Vocoder successfully implemented and demonstrated using EPROM versions of the N.E.C. \( \mu \)PD7720
FLEXIBLE ARRAY PROCESSORS

GLEN CULLER
CHI SYSTEMS, INC.
DEVELOPMENT OF FLEXIBLE ARRAY PROCESSORS

1. ARRAY PROCESSOR OVERVIEW
2. SIGNAL PROCESSING FRO ARC
3. THE FIRST LPC SPEECH ON ARPA NET
4. THE FPS-120B COMMERCIALIZATION
5. THE PLASMA SIMULATION SPINOFF
6. THE LPCAP
7. THE CHI-5, A GENERAL PURPOSE ARRAY PROCESSOR
8. CHI-5 ARCHITECTURE ONTO VLSI CHIPS
CHI-5 CHARACTERISTICS

1. GENERAL PURPOSE COMPUTER WITH ARRAY PROCESSOR INTERIOR

2. TWO ARRAY MEMORIES 1024 WORDS EACH

3. MATRIX ADDRESSING FOR ARRAY MEMORIES

4. FOUR STAGE PROGRAMMABLE PIPELINE

5. 16 PORT MEMORY
   8 PROCESS PORTS
   8 IO PORTS

6. Q-BUS AND UNIBUS INTERFACES

7. TWO CHANNEL RS-232 INTERFACE
<table>
<thead>
<tr>
<th>Operation</th>
<th>Time (μsec)</th>
<th>Operation</th>
<th>Time (μsec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOT PRODUCT</td>
<td>N/4</td>
<td>LATTICE REDUCTION</td>
<td>N/2</td>
</tr>
<tr>
<td>2 MULTIPLIER LATTICE FILTER</td>
<td>9N</td>
<td>RUNNING AVERAGE</td>
<td>0.75N</td>
</tr>
<tr>
<td>(10 POLE)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLOATING POINT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADD</td>
<td>3</td>
<td>MULTIPLY</td>
<td>3</td>
</tr>
<tr>
<td>DIVIDE</td>
<td>8</td>
<td>LOGARITHM</td>
<td>10</td>
</tr>
<tr>
<td>FFT:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>128 COMPLEX</td>
<td>1.24 - 1.5</td>
<td>1024 COMPLEX</td>
<td>11.50 - 15.0</td>
</tr>
<tr>
<td>1024 COMPLEX</td>
<td>6.20 - 8.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPC (164 POINT FRAME):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COEFFICIENT ANALYSIS</td>
<td>2.3</td>
<td>PITCH ANALYSIS</td>
<td>1.6</td>
</tr>
<tr>
<td>(BERG REDUCTION)</td>
<td></td>
<td>SYNTHESIS</td>
<td>2.5</td>
</tr>
<tr>
<td>TRANSFER RATE:</td>
<td></td>
<td></td>
<td>1 MILLION WORDS/SECOND</td>
</tr>
</tbody>
</table>

PROGRAM OBJECTIVES

- DELIVER VLSI ARITHMETIC PROCESSOR CHIPS
- DELIVER TWO DEMONSTRATION SYSTEMS
- DEVELOP SINGLE MODULE ARRAY PROCESSOR CONCEPT
- DELIVER SOFTWARE DEVELOPMENT TOOLS
  - TRANSPORTABLE - FORTRAN 77
  - LAYERED - ARITHMETIC PROCESSOR
    - DEMONSTRATION SYSTEM
  - SIMULATORS AND ASSEMBLERS
<table>
<thead>
<tr>
<th>ALGORITHM</th>
<th>TIME (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1024 POINT COMPLEX FFT</td>
<td>5632 μSECS</td>
</tr>
<tr>
<td>512 POINT REAL FFT</td>
<td>1152 μSECS</td>
</tr>
<tr>
<td>1024 POINT REAL VECTOR MULTIPLICATION</td>
<td>260 μSECS</td>
</tr>
<tr>
<td>1024 POINT COMPLEX MULTIPLY</td>
<td>520 μSECS</td>
</tr>
<tr>
<td>DARPA 2400 BIT FULL DUPLEX VOICE PROCESSING</td>
<td>≤5263 μSECS</td>
</tr>
<tr>
<td>COMPLEX CONJUGATE POLE-ZERO PAIR FILTER WITH GAIN</td>
<td>1.25 μSECS/POINT</td>
</tr>
<tr>
<td>32 BIT MANTISSA FLOATING POINT ADDITION</td>
<td>1.5 TO 2.5 μSECS</td>
</tr>
<tr>
<td>32 BIT MANTISSA FLOATING POINT MULTIPLICATION</td>
<td>.75, 1.0 OR 1.75 μSECS</td>
</tr>
<tr>
<td>1 MULTIPLICATION AND 3 DATA ADDITIONS</td>
<td>250 nSECS</td>
</tr>
</tbody>
</table>

(1) BASED ON 4 MHz SYSTEM CLOCK FREQUENCY AND BLOCK FLOATING POINT ARITHMETIC FOR PROCESSES
DEVICE CHARACTERISTICS

- OPERATION CYCLE TIME $\leq 250$ NSECS; 200 NSECS GOAL
- POWER DISSIPATION $< 250$ MWATTS
- SUPPLY VOLTAGE 5 VOLTS
- DEVICE SIZE 298 X 305 MILS
- DEVICE PINOUTS 100
- PACKAGE PIN GRID ARRAY
FUTURE ARRAY PROCESSOR MODULE

SIZE - LESS THAN 50 SQUARE INCHES
POWER - LESS THAN 10 WATTS
NUMBER OF TOTAL PARTS* - LESS THAN 30

* WITH THE FOLLOWING MEMORY(1)

I/O MEMORY - 8K X 16 RAM
MACRO MEMORY - 1K X 16 PROM
VECTOR SCRATCH MEMORY - 4K X 16 RAM
MICROCODE MEMORY - 2K X 128 PROM

(1) MEMORY SIZES CAN BE EXPANDED
PROGRAM STATUS

- CHIP IN MASK SHOP
- DEMONSTRATION SYSTEM IN TEST
- SOFTWARE TASKS COMPLETED
SINGLE CHIP LPC

ROBERT BRODERSEN
UNIV. OF CALIFORNIA/BERKELEY
Single Chip LPC

- Combine analog & digital functions for a completely integrated, full duplex LPC vocoder

- Use standard foundry processing available from MOSIS

- Use most advanced CAD tools under development at Berkeley

- Large macro design methodology to give flexibility to a special purpose design
PROPOSED LPC Vocoder CHIP

MICROPHONE → PRE-AMP → ANTI-ALIASING FILTER → AGC

LPC ANALYSIS FILTER → PARCOR CALC. → 8 BIT LOG A/D

GOLD PITCH TRACKER → RESIDUAL

FORMATTING

RECEIVED DIGITAL DATA → FORMATTING & INTERPOLATION → LPC SYNTHESIS FILTER

TRANSMITTED DIGITAL DATA

SPEAKER

RECONSTM. FILTER & PWR. AMP.
Signal Processor

Bus Switch (Sig)

Adder

Multiplier Input

Saturating Adder

16

Shift Left

16

Comp

Shift Right Load

16

3:1 Mux

Accumulator

Read Memory

Area = 3600 mil^2 (x=2 μm)
PARALLEL PROCESSORS

Processor 1: Analysis Filter, Synthesis Filter, Pre-Emphasis, De-Emphasis, Pitch Tracker Filter

Processor 2: Reflection Coefficient Calculation and Residual Energy Estimate

Processor 3: Gold-Rabiner Algorithm and Voicing Decision

Misc. Blocks

Excitation Generator: Forms Excitation Waveform

Parameter I/O: Communicates with MPU

Controller: Program Counter, Microcode RAM, Addressing Arithmetic, Generates All Control Signals
<table>
<thead>
<tr>
<th>Component</th>
<th>Area (sq. mils)</th>
<th>Power (milli watts)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-amp</td>
<td>400</td>
<td>5</td>
</tr>
<tr>
<td>Anti-aliasing filter</td>
<td>2400</td>
<td>8</td>
</tr>
<tr>
<td>AGC</td>
<td>2000</td>
<td>10</td>
</tr>
<tr>
<td>LPC filters (5000)</td>
<td>2 x 5000 = 10,000</td>
<td>2 x 10 = 20</td>
</tr>
<tr>
<td>LPC digital correlator (7500)</td>
<td>10,000</td>
<td>40</td>
</tr>
<tr>
<td>A/D converter</td>
<td>10,000</td>
<td>150</td>
</tr>
<tr>
<td>Pitch tracker filters (10,000)</td>
<td>4200</td>
<td>15</td>
</tr>
<tr>
<td>Pitch tracker estimators</td>
<td>4800</td>
<td>15</td>
</tr>
<tr>
<td>Pitch tracker scoring</td>
<td>7800</td>
<td>20</td>
</tr>
<tr>
<td>Formatting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpolation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clocks (2000)</td>
<td>8000</td>
<td>32</td>
</tr>
<tr>
<td>Power amp &amp; output filt.</td>
<td>2000</td>
<td>15</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>61,600</strong></td>
<td><strong>330</strong></td>
</tr>
</tbody>
</table>
**Signal Processor Execution Cycle**

**Preamble:** Pre-emphasis, De-emphasis and low-pass filter for pitch processor; nominally 32 instructions.

**Analysis/Synthesis:** 10 cycles of 32 instructions = 320 j/16 instructions for each filter.

**Clock Rate:** \( 352 \times 8 \text{MHz} = 2.816 \text{MHz} \)
**SUMMARY**

- 20,000 TRANSISTORS
- .85 WATT POWER DISSIPATION (NMOS)
- 40,000 mil² - 4μm CHANNEL LENGTH
- 6K CONTROL ROM
- 2K DYNAMIC RAM
- 3 PARALLEL PROCESSORS