Naval Research Laboratory

Washington, DC 20375-5000

NRL Report 9357

2

ANDVT Rate Conversion Algorithm (from 2400 b/s to 1200 b/s)

G. S. KANG AND L. J. FRANSEN

Information Technology Division

December 27, 1991

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
December 27, 1991

3. REPORT TYPE AND DATES COVERED
Final

4. TITLE AND SUBTITLE
ANDVT Rate Conversion Algorithm
(from 2400 b/s to 1200 b/s)

5. FUNDING NUMBERS
PE — 602232N
PR — R C32A13
X7290-CC

6. AUTHOR(S)
George S. Kang and Larry J. Fransen

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Research Laboratory
Washington, DC 20375-5000

8. PERFORMING ORGANIZATION REPORT NUMBER
NRL Report 9357

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Space and Naval Warfare Systems Command
PMW-151-21
Arlington, VA 20360

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The 2400-b/s linear predictive coder (LPC) is currently being widely deployed to support tactical voice communication over narrowband channels. However, lower-data-rate voice encoders are needed for special applications: improved performance in high-bit-error conditions and implementation of integrated voice/data systems.

In this report, we generated a rate-conversion algorithm to compress voice data rate from 2400 b/s to 1200 b/s. Rate reduction is effected outside the Advanced Narrowband Digital Voice Terminal (ANDVT), requiring no modification to the ANDVT software or hardware. This is a cost-effective way of obtaining 1200 b/s voice data for ANDVT users for special communication requirements.

Speech intelligibility at 1200 b/s as measured by the Diagnostic Rhyme Test is 90.4, which is only 1.7 points below that of the 2400 b/s LPC. Thus, 1200 b/s speech will be acceptable as 2400 b/s speech.

14. SUBJECT TERMS
Speech compression Speech synthesis Speech rate conversion
Speech analysis Low-data-rate speech

15. NUMBER OF PAGES
29

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19 SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

CONTENTS

ANDVT RATE CONVERSION ALGORITHM
(from 2400 b/s to 1200 b/s)

1. INTRODUCTION

The linear predictive coder (LPC) operating at 2400 bits per second (b/s) is being widely deployed to support tactical voice communication over narrowband channels (approximately 3 kHz). The Advanced Narrowband Digital Voice Terminal (ANDVT or AN/USC-43(V)) is an example of a 2400-b/s LPC being deployed for tri-service tactical applications.

This report is written for ANDVT users who also have a need to compress voice data to 1200 b/s for specialized applications, such as improved bit-error performance, voice/data integration, voice/voice integration, to name a few. Although a 1200-b/s voice processor can be designed as a self-contained unit, we intentionally chose an approach in which the 1200-b/s bit stream is derived from the ANDVT bit stream (Fig. 1). This approach, presented in this report, has the following advantages:

- Complex numerical computations imposed by speech analysis/synthesis are not necessary.
- A separate audio front end (handset, bandpass filters, analog-to-digital converter, and digital-to-analog converter) is not needed.
- Both the 1200-b/s and the standard 2400-b/s bit streams are available.

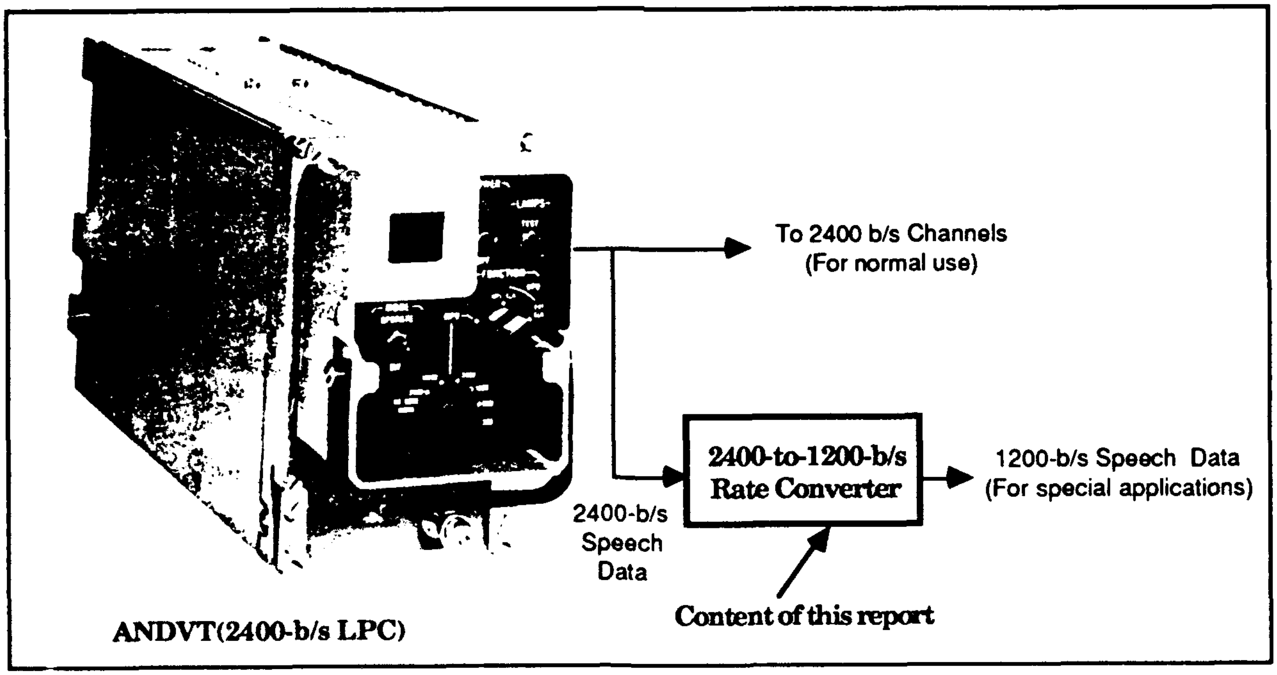


Fig. 1 — A 2400-to-1200-b/s rate converter as an add-on unit to the ANDVT (2400-b/s LPC) . This report describes the rate conversion algorithm. At the receiver, the 1200-b/s bit stream is converted to an operationally-compatible 2400-b/s bit stream for input to the ANDVT receiver. The rate converter removes both interframe and intraframe speech redundancies. Interframe redundancies are those speech parameters that can be interpolated from the speech parameters of the two adjoining frames; intraframe redundancies are speech parameters that are not related to human speech or are not discernible by the human ears from others.

This report is the result of our continuing R&D effort to enhance the capabilities of the widely deployed ANDVT. Previously, we developed a method for modifying the ANDVT bit stream to transmit digital data (up to 80 b/s) *simultaneously* with voice data without degrading speech intelligibility or disabling the interoperability with other ANDVTs [1]. The digital data were introduced outside the ANDVT. Therefore, no software or hardware modifications to ANDVT were required.

In this report, we introduce a method for compressing the ANDVT data rate from 2400 b/s to 1200 b/s. Again, data rate reduction is effected outside the ANDVT, requiring no modification to the ANDVT software or hardware. Speech intelligibility at 1200-b/s (as measured by the Diagnostic Rhyme Test (DRT)) is 90.4, which is only 1.7 points below that of the 2400-b/s LPC. We expect the 1200-b/s speech generated by the rate converter will be as acceptable as the original 2400-b/s LPC speech.

2. 2400-b/s LPC

Because the 1200-b/s voice processor operates in conjunction with the Government-standard 2400-b/s LPC, only relevant information related to rate conversion is summarized in this section.

Specifications and Standards

The Government-standard 2400-b/s LPC [2] (hereafter simply referred to as the 2400-b/s LPC) is a speech analysis/synthesis system based on linear predictive coding. Interoperability requirements are specified in four different documents (listed in Table 1). Both the ANDVT and the Subscriber Terminal Unit - Third Generation (STU-III) adhere to this standard.

Table 1 — Documents that Specify the Government-Standard 2400-b/s LPC

Document Number	Title	Remarks
Federal Standard 1015	Analog to Digital Conversion of Voice by 2400 b/s Linear Predictive Coding	This document specifies the Government-standard 2400-b/s LPC.
MIL-STD-188-113	Common Long Haul and Tactical Standards for Analog-to-Digital Conversion Techniques	This document specifies pulse code modulator (PCM), continuously variable slope delta (CVSD), adaptive predictive coder (APC), and LPC.
STANAG 4198	Parameters and Coding Characteristics That Must be Common to Assume Interoperability of 2400 b/s Linear Predictive Encoded Digital Speech	This NATO document is virtually identical to Federal Standard 1015.
TT-B1-4210-0087B	Performance Specification for the Advanced Narrowband Digital Voice Terminal (ANDVT) Tactical Terminal (TACTERM)	This TRI-TAC document specifies the Government-standard 2400-b/s LPC for ANDVT.

LPC Formulation

The 2400-b/s LPC is based on the principle that a speech sample can be represented as a linear combination of past samples. Thus,

$$\varepsilon_i = x_i - \sum_{k=1}^{K} \alpha_k x_{i-k}, \tag{1}$$

where x_i is the ith speech sample, a_k is the kth prediction coefficient, ε_i is the ith error sample, and K is the order of the predictor (K = 10 for ANDVT). In matrix notation, Eq. (1) can be written as

$$X = H\alpha + \varepsilon. \tag{2}$$

On the basis of unbiased estimation [3], the prediction coefficients are obtained from

$$(H^T H)\ \alpha = H^T X, \tag{3}$$

where H^T implies the transposed H. Through Cholesky decomposition [4], matrix $H^T H$ can be written in the form of:

$$H^T H = LDL^T \tag{4}$$

where L is a lower triangular matrix, D is a diagonal matrix, and (DL^T) becomes an upper triangular matrix. From Eqs. (3) and (4), prediction coefficients are obtained from

$$(LDL^T)\ \alpha = H^T X. \tag{5}$$

From Eq. (5), prediction coefficients are solved in two steps:

$$k = L^{-1}\ (H^T X) \tag{6}$$

and

$$\alpha = (DL^T)^{-1}\ k. \tag{7}$$

The quantity k in Eq. (6) is a set of reflection coefficients, and they are linearly and uniquely related to a set of prediction coefficients by Eq. (7). All current 2400-b/s LPCs transmit reflection coefficients in lieu of prediction coefficients. A merit for using reflection coefficients is that if the magnitude of each reflection coefficient is less than unity, the speech synthesizer is guaranteed to be stable.

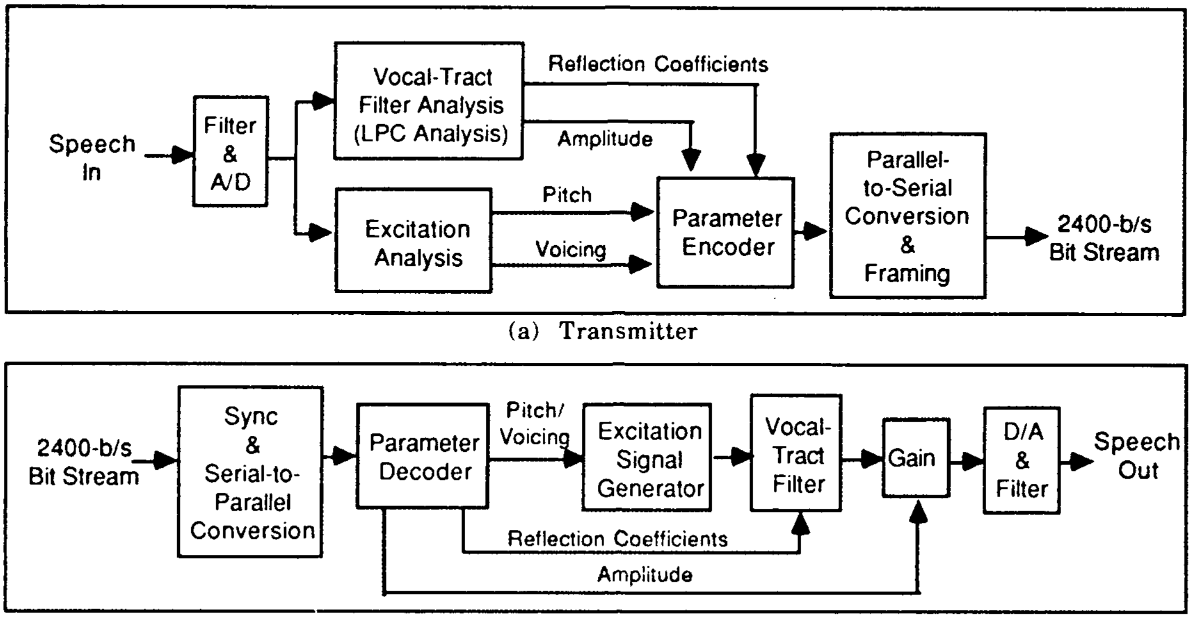
Based on Eq. (1), speech is synthesized by

$$x_i = \sum_{k=1}^{K} \hat{\alpha}_k x_{i-k} + \varepsilon_i, \tag{8}$$

where ^ signifies a quantized parameter. For the 2400-b/s LPC, ε_i is drastically simplified as either a pulse train for generating vowels or as random noise for generating consonants. The pitch period controls the pulse repetition rate.

Block Diagram

A block diagram of the 2400-b/s LPC is shown in Fig. 2. Both the vocal-tract filter coefficients (reflection coefficients) and the excitation signal parameteres are extracted from the speech waveform.



(a) Transmitter

(b) Receiver

Fig. 2 — Block diagram of the 2400-b/s LPC. For further information, see the documents listed in Table 1.

Speech Parameters

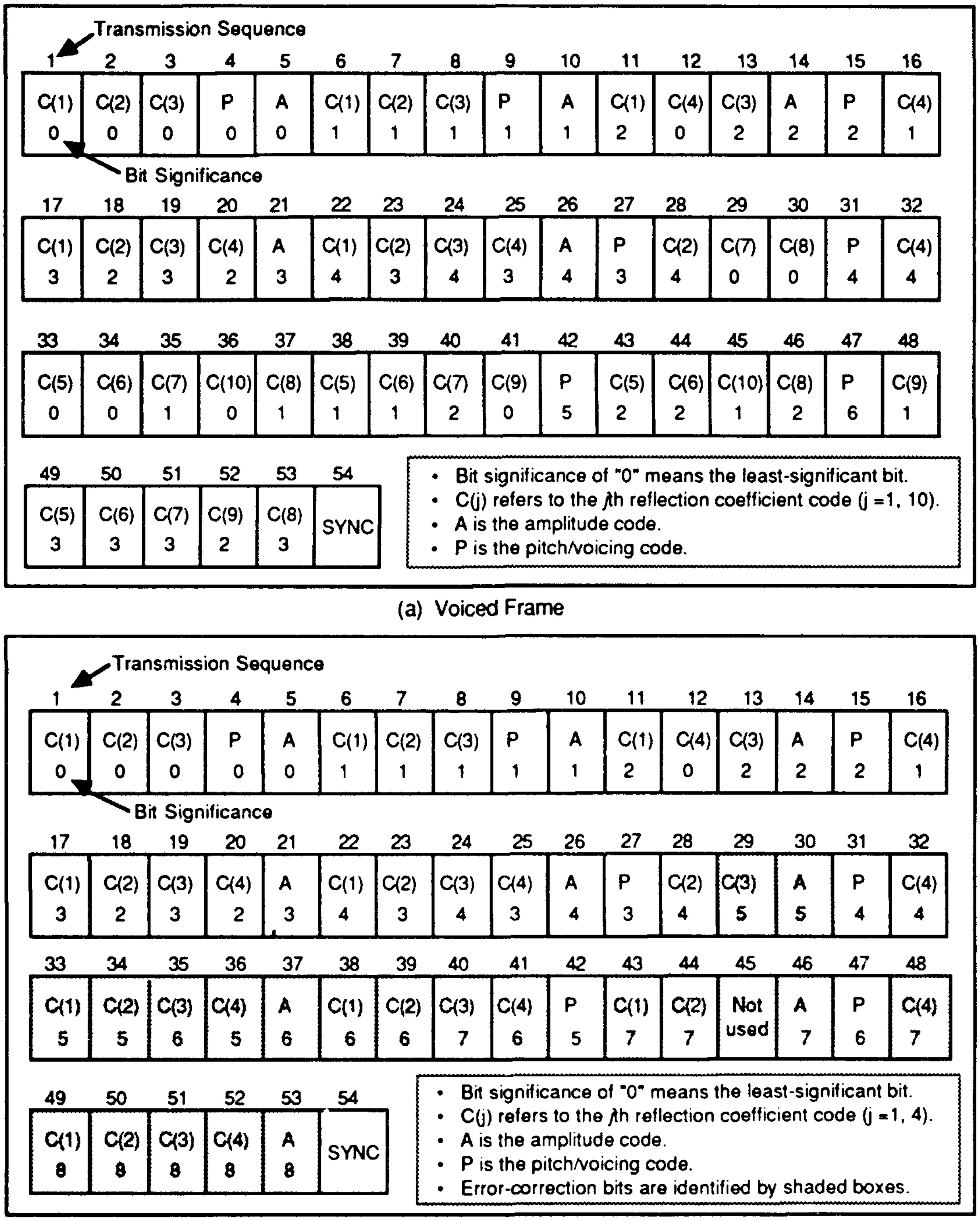
The following four sets of parameters are extracted from the speech waveform once every 22.5 ms; they are encoded as described in Table 2.

Table 2 — 2400-b/s LPC Parameters

Speech Parameter	No. of Bits Per Frame	Remarks
Amplitude	5	Root-mean-square value of preemphasized speech wavefom quantized semi-logarithmically over a 60-dB dynamic range
Pitch	6	Quantized logarithmically from a pitch interval of 20 to 160 samples at 20 steps per octave
Voicing	1	Two-state voicing decision
Reflection Coefficients	41 (Voiced frames)	The first through tenth reflection coefficients are quantized to 5, 5, 5, 5, 4, 4, 4, 4, 3, and 2 bits, respectively.
	20 (Unvoiced frames)	The first through fourth reflection coefficients are quantized to 5, 5, 5, and 5 bits, respectively. Twenty bits are used for protecting the four most significant bits of the reflection coefficients and amplitude parameters by Hamming (8,4) codes. One bit is unused.
Sync	1	Alternating "1" and "0"

Bit Stream

Bit-stream format of the 2400-b/s LPC is defined by the documents listed in Table 1 (see Fig. 3).



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
C(1)	C(2)	C(3)	P	A	C(1)	C(2)	C(3)	P	A	C(1)	C(4)	C(3)	A	P	C(4)
0	0	0	0	0	1	1	1	1	1	2	0	2	2	2	1

17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
C(1)	C(2)	C(3)	C(4)	A	C(1)	C(2)	C(3)	C(4)	A	P	C(2)	C(7)	C(8)	P	C(4)
3	2	3	2	3	4	3	4	3	4	3	4	0	0	4	4

33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
C(5)	C(6)	C(7)	C(10)	C(8)	C(5)	C(6)	C(7)	C(9)	P	C(5)	C(6)	C(10)	C(8)	P	C(9)
0	0	1	0	1	1	1	2	0	5	2	2	1	2	6	1

49	50	51	52	53	54
C(5)	C(6)	C(7)	C(9)	C(8)	SYNC
3	3	3	2	3	

- Bit significance of "0" means the least-significant bit.
- C(j) refers to the jth reflection coefficient code (j =1, 10).
- A is the amplitude code.
- P is the pitch/voicing code.

(a) Voiced Frame

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
C(1)	C(2)	C(3)	P	A	C(1)	C(2)	C(3)	P	A	C(1)	C(4)	C(3)	A	P	C(4)
0	0	0	0	0	1	1	1	1	1	2	0	2	2	2	1

17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
C(1)	C(2)	C(3)	C(4)	A	C(1)	C(2)	C(3)	C(4)	A	P	C(2)	C(3)	A	P	C(4)
3	2	3	2	3	4	3	4	3	4	3	4	5	5	4	4

33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
C(1)	C(2)	C(3)	C(4)	A	C(1)	C(2)	C(3)	C(4)	P	C(1)	C(2)	Not used	A	P	C(4)
5	5	6	5	6	6	6	7	6	5	7	7		7	6	7

49	50	51	52	53	54
C(1)	C(2)	C(3)	C(4)	A	SYNC
8	8	8	8	8	

- Bit significance of "0" means the least-significant bit.
- C(j) refers to the jth reflection coefficient code (j =1, 4).
- A is the amplitude code.
- P is the pitch/voicing code.
- Error-correction bits are identified by shaded boxes.

(b) Unvoiced Frame

Fig. 3 — Bit stream of the 2400-b/s LPC. If speech is voiced, 10 reflection coefficients are transmitted, and none of the speech parameters are error-protected. If speech is unvoiced, only the first four reflection coefficients are transmitted; 20 bits are used for protecting the four most-significant bits of the amplitude code A and reflection coefficient codes C(1), C(2), C(3), and C(4). In both the voiced and unvoiced frames, pitch and voicing are combined into a seven-bit quantity. See the next section for the description of error protection.

Speech Intelligibility

According to ANDVT users during Desert Storm operations, the 2400-b/s LPC provided much superior speech quality than previous narrowband secure phones. As noted from Fig. 4, DRT scores of the 2400-b/s LPC is in the "good" to "very good" regions.

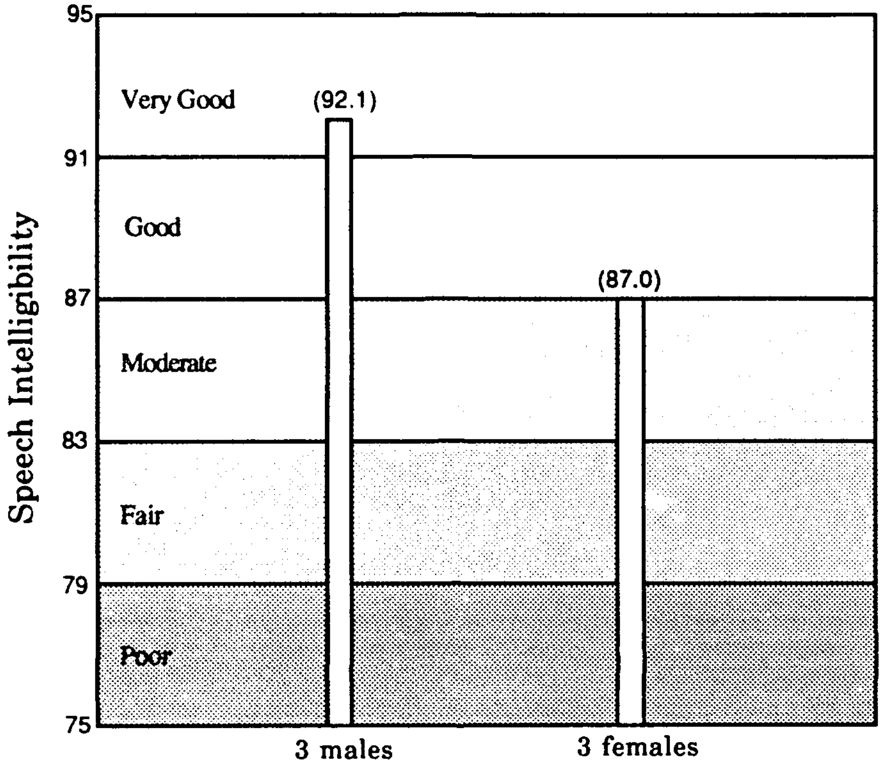


Fig. 4 — DRT scores of the current 2400-b/s LPC. The descriptors "very good," "good," "fair," etc. have been adopted by the DoD Digital Voice Processor Consortium.

3. 1200-b/s VOICE ALGORITHM

The main embodiment of the 1200-b/s LPC is the rate converter that converts the 2400-b/s bit stream to a 1200-b/s bit stream. Likewise, it converts the 1200-b/s bit stream to a 2400-b/s bit stream for input to the ANDVT receiver to regenerate speech.

Block Diagram

Figure 5 is a block diagram of the rate converter. At the transmitter, the 2400-b/s bit stream from the 2400-b/s LPC is synchronized, and the individual speech parameters are demultiplexed. Then speech parameters are requantized to generate the 1200-b/s bit-stream based on the algorithm described in this section.

At the receiver, the 1200-b/s bit stream is synchronized, and the individual speech parameters are demultiplexed. Demultiplexed parameters are decoded based on the algorithm presented in this section. Decoded speech parameters are converted to a 2400-b/s bit stream so that the 2400-b/s LPC can generate speech.

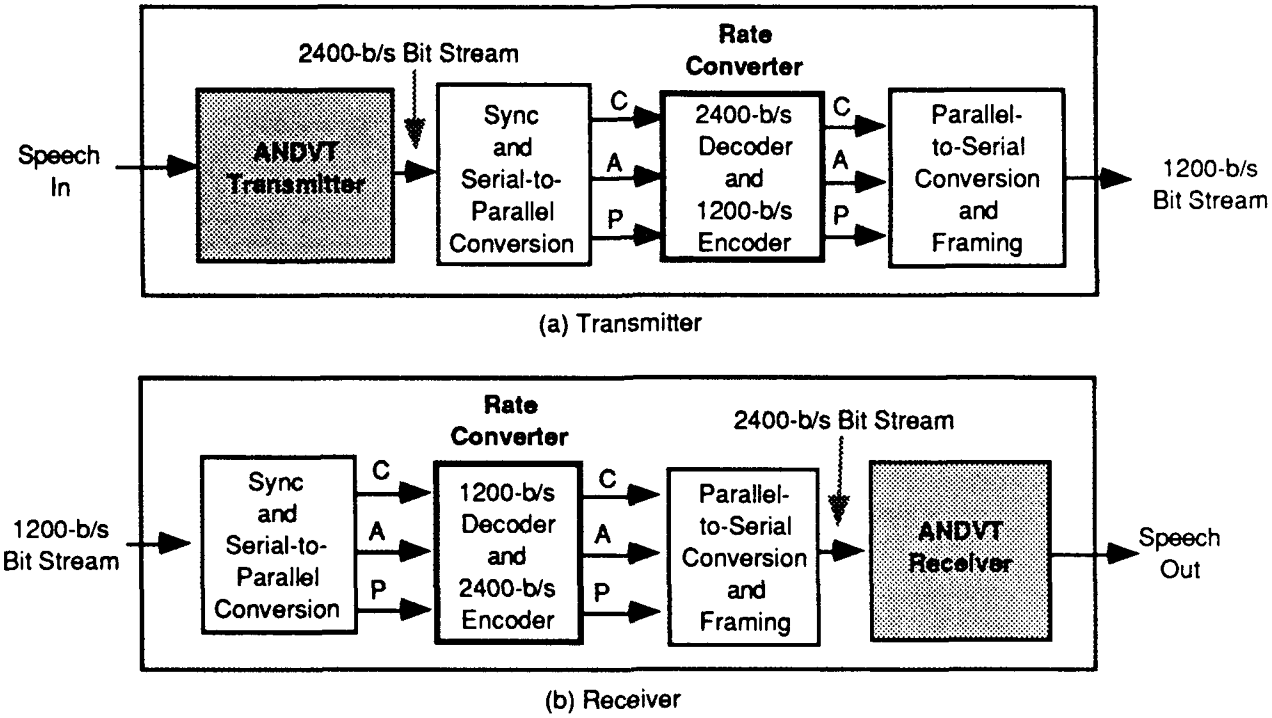

2400-b/s Bit Stream
Rate Converter
Speech In
ANDVT Transmitter
Sync and Serial-to-Parallel Conversion
C
A
P
2400-b/s Decoder and 1200-b/s Encoder
Parallel-to-Serial Conversion and Framing
1200-b/s Bit Stream
(a) Transmitter
Rate Converter
2400-b/s Bit Stream
1200-b/s Bit Stream
Sync and Serial-to-Parallel Conversion
1200-b/s Decoder and 2400-b/s Encoder
Parallel-to-Serial Conversion and Framing
ANDVT Receiver
Speech Out
(b) Receiver

Fig. 5 — Block diagram of the 1200-b/s LPC. The rate converter, indicated by the heavy-lined blocks, is discussed in this section. As defined in Fig. 3, C is a set of reflection coefficient codes, A is amplitude code, and P is pitch/voicing code.

Figure 6 summarizes the processes of rate conversion from 2400 to 1200 b/s. We made rate conversion computationally simpler by retaining the original ANDVT codes in the 1200-b/s bit stream where possible. The only exception is for reflection coefficients.

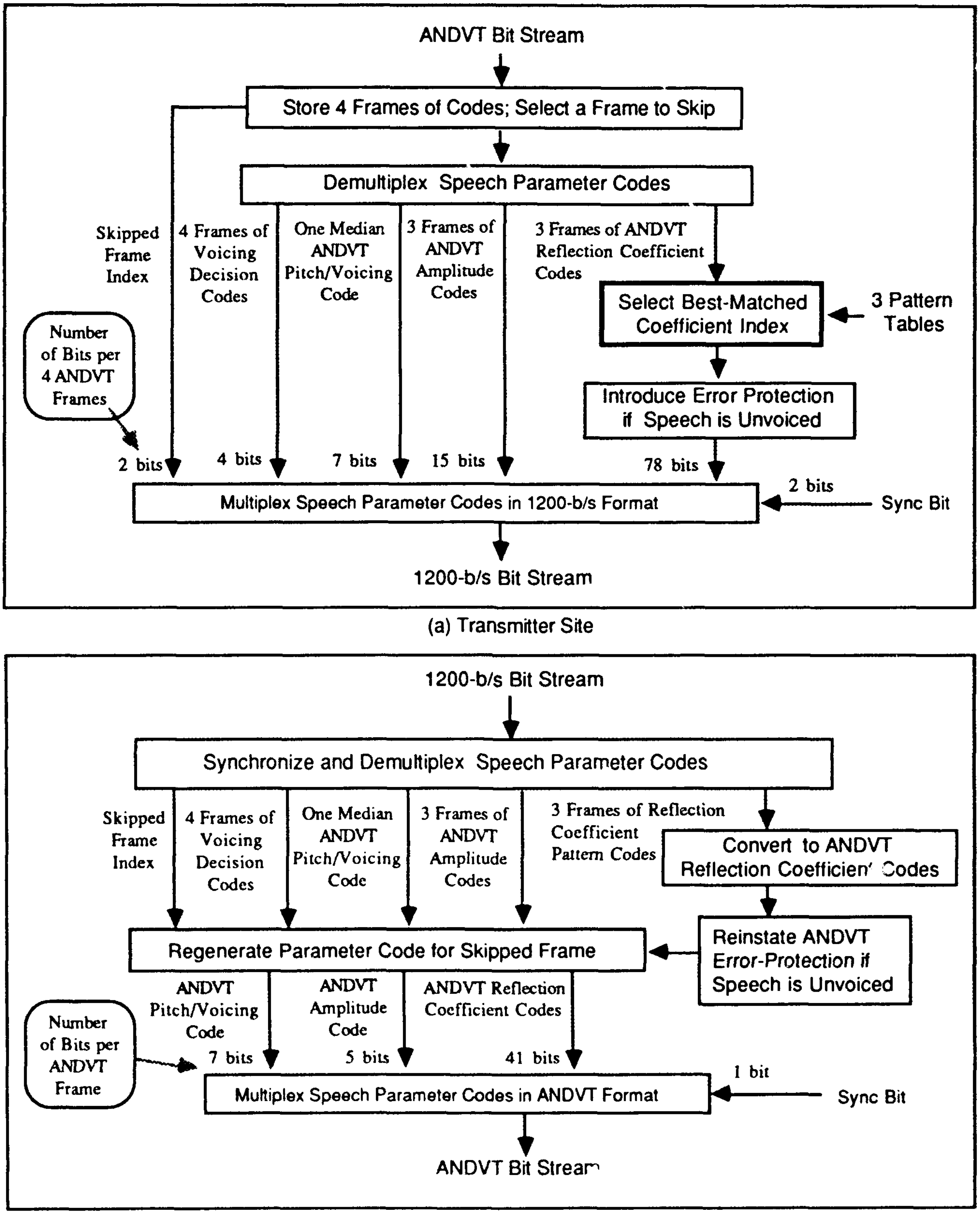
Reflection coefficients are represented by two vectors: one representing the first through fourth coefficients, and the other representing the remaining coefficients only when speech is voiced. If speech is unvoiced, the second vector is replaced by error-protection codes. We made the pattern-matching process simpler by limiting the number of patterns to be searched to only 1000 for each vector by sorting templates in ascending order of the spectrally most sensitive first reflection coefficient.

Interframe Redundancy Reduction by Frame Skipping

The voice data rate is compressed from 2400 b/s to 1200 b/s by reducing two different speech redundancies: interframe redundancies and intraframe redundancies. First, we discuss how to reduce interframe speech redundancies.

A significant characteristic of speech is that information content varies with time. Information content is large at speech onset where speech parameters change abruptly (Fig. 7). On the other hand, information content is low during sustained vowels where speech parameters are not changing significantly (Fig. 7).

An ideal way to transmit speech, therefore, is by a variable data rate. As early as 1977, researchers updated parameters at variable rates prior to converting to a fixed rate [5]. According to their observations, an average data rate of 1500 b/s had the same speech quality as a fixed rate of 2350 b/s.


ANDVT Bit Stream
Store 4 Frames of Codes; Select a Frame to Skip
Demultiplex Speech Parameter Codes
Skipped Frame Index
4 Frames of Voicing Decision Codes
One Median ANDVT Pitch/Voicing Code
3 Frames of ANDVT Amplitude Codes
3 Frames of ANDVT Reflection Coefficient Codes
Select Best-Matched Coefficient Index
3 Pattern Tables
Introduce Error Protection if Speech is Unvoiced
Number of Bits per 4 ANDVT Frames
2 bits
4 bits
7 bits
15 bits
78 bits
2 bits
Sync Bit
Multiplex Speech Parameter Codes in 1200-b/s Format
1200-b/s Bit Stream
1200-b/s Bit Stream
Synchronize and Demultiplex Speech Parameter Codes
Skipped Frame Index
4 Frames of Voicing Decision Codes
One Median ANDVT Pitch/Voicing Code
3 Frames of ANDVT Amplitude Codes
3 Frames of Reflection Coefficient Pattern Codes
Convert to ANDVT Reflection Coefficient Codes
Regenerate Parameter Code for Skipped Frame
Reinstate ANDVT Error-Protection if Speech is Unvoiced
ANDVT Pitch/Voicing Code
ANDVT Amplitude Code
ANDVT Reflection Coefficient Codes
Number of Bits per ANDVT Frame
7 bits
5 bits
41 bits
1 bit
Sync Bit
Multiplex Speech Parameter Codes in ANDVT Format
ANDVT Bit Stream

(a) Transmitter Site

(b) Receiver Site

Fig. 6 — Summary of processes involved in rate conversion from 2400 to 1200 b/s.

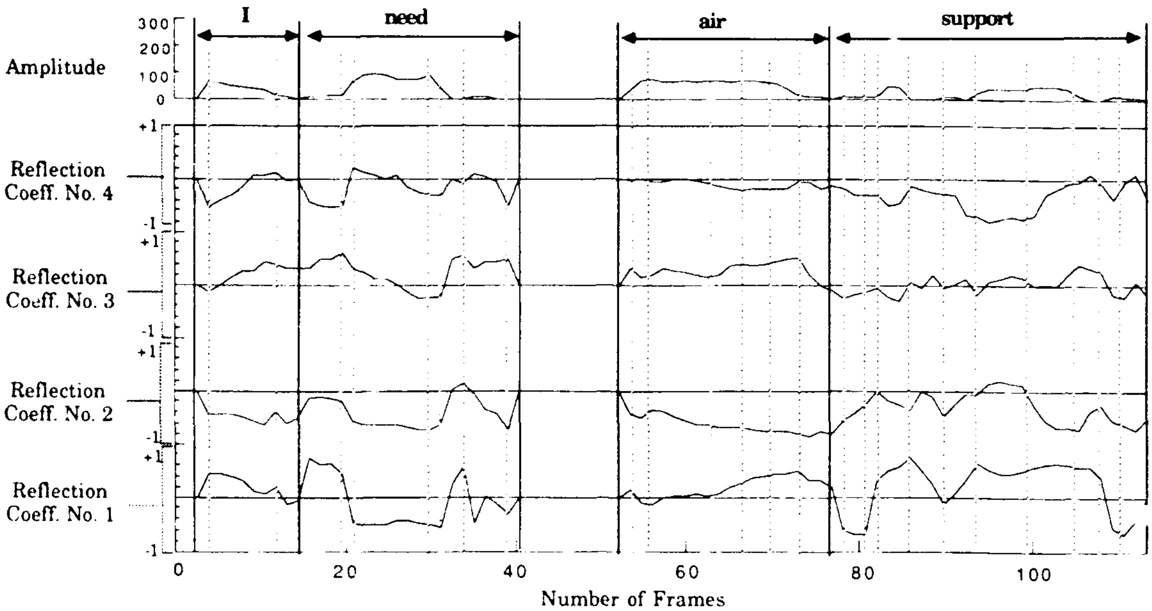


Fig. 7 — Trajectories of 2400-b/s LPC parameters extracted from the speech waveform of "I need air support."

Our approach combines several frames which we call a *superframe*. The superframe data rate is fixed, but within the superframe, the speech data rate is variable. A greater superframe size leads to a more efficient interframe redundancy elimination, but the superframe size cannot be arbitrarily large because of the direct relationship between the number of frames in the superframe and the corresponding speech delay (L). The amount of speech delay in a two-way link is $2L$ (i.e., a delay of L at both the transmitter and receiver). This aspect is discussed in the next section.

Superframe Size and Speech Delay

In a half-duplex system, speech is transmitted in one direction at a time; it is not an interactive system. Furthermore, in a secure half-duplex system (such as ANDVT), pressing the push-to-talk button initiates an exchange of preamble signals between the two parties; during this time period, the speaker must wait. Thus, the normal half-duplex mode of operation entails interruptions and delays. A small amount of signal delay (a fraction of a second) introduced by the 1200-b/s LPC does not cause half-duplex communication difficulties.

For a full-duplex system, however, an excessive signal delay does cause aggravation to both parties, as we occasionally experience with long-distance telephone calls. Human factors related to voice communication in the presence of speech delay has been studied [6]. Figure 8 shows the percentage of users reporting difficulty communicating when confronted with round-trip delay.

We chose a superframe size of 4 frames (L = 90 ms), which results in an overall one-way signal delay of only 360 ms (see Table 3). A delay of 360 ms is acceptable. By using a superframe size of 4 frames, we can omit transmission of one frame of speech data to reduce the overall data.

Frame Skip

Often the speech data in one frame is nearly identical to the interpolated values of the speech data of the two adjoining frames. If so, that particular speech data can be regarded as redundant information that need not be transmitted. We skip one frame of data within each superframe, and the index of the omitted frame is transmitted. At the receiver the omitted data are regenerated by interpolating the data from the two neighboring frames. To minimize speech delay and make interpolation easier, we do not skip the fourth frame (see Fig. 9).

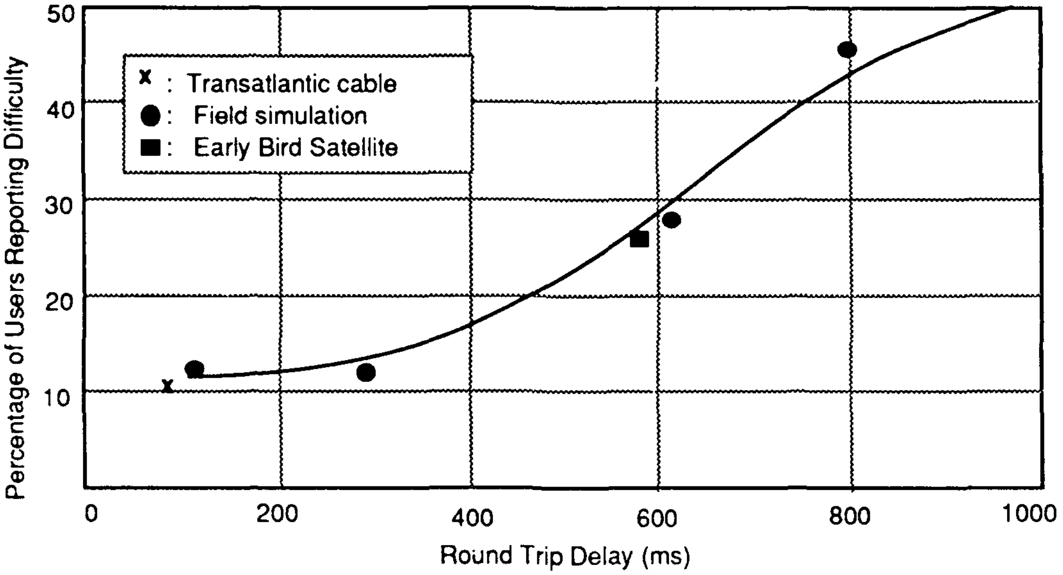


Fig. 8 — Communication difficulties in a full-duplex link. In the presence of excessive speech delays, the talker tends to get annoyed because the listener responds too slowly. Likewise, the listener tends to get upset because the talker appears to ignore the listener's interruption.

Table 3 — One-Way Speech Delay through Rate Conversion

Delay Source		No. of Frames	Delay (ms)
2400-b/s LPC		8, typical	180, typical
Rate Converter	Superframe	4	90
	Parallel-to-Serial and Serial-to-Parallel Conversions	4	90
Total		16	360 ms

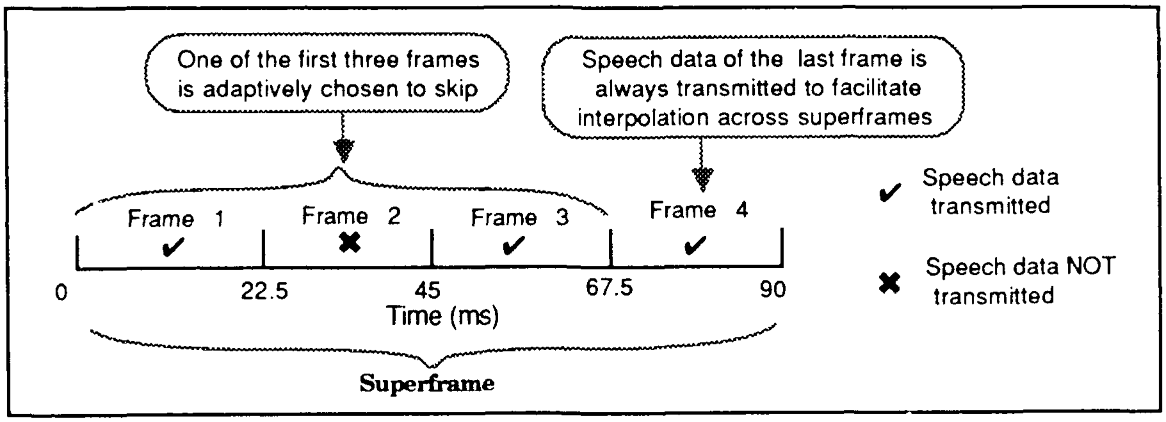

One of the first three frames is adaptively chosen to skip
Speech data of the last frame is always transmitted to facilitate interpolation across superframes
Frame 1 ✓
Frame 2 ✖
Frame 3 ✓
Frame 4 ✓
0
22.5
45
67.5
90
Time (ms)
Superframe
✓ Speech data transmitted
✖ Speech data NOT transmitted

Fig. 9 — Superframe structure. The superframe is made of four successive frames. For each superframe, a frame containing redundant speech data is adaptively determined based on whether it can be interpolated from the speech data of the two neighboring frames, and those speech parameters are not transmitted. Because one out of the first three frames is skipped, we must transmit two bits of overhead data to indicate the location of the skipped frame.

To determine which frame is to be skipped, we compute the error between the actual speech parameters of one frame with the interpolated values based on the speech parameters of the two adjoining frames. The speech parameters include the first four reflection coefficients and amplitude. Both pitch and voicing decision are not included in determining the most redundant frame. We chose the first four reflection coefficients because they are the most significant contributors to the speech spectral envelope. The error at the *i*th frame is defined as

$$\Delta E_i = Gc \sum_{j=1}^{4} \left| C_j(i) - \overline{C}_j(i) \right| + Ga \left| A(i) - \overline{A}(i) \right| \quad i = 1,\ 2,\ 3 \tag{9}$$

where $A(i)$ and $C_j(i)$ are the amplitude code and the *j*th reflection coefficient code of the *i*th frame, respectively; *Ga* and *Gc* are weighting factors, each somewhere between 0 and 1. We used encoded parameters in Eq. (9) to avoid decoding (i.e., to reduce computations). The quantity with the upper bar denotes the interpolated value from the two adjacent frames. We let $Ga > Gc$ at speech onsets because speech amplitude changes significantly. Elsewhere, we let $Gc > Ga$ because we would like to detect where speech resonant frequencies change significantly. The speech parameters for the frame having the smallest error will not be transmitted (Fig. 9).

In Eq. (9), the parameters are expressed in terms of codes (as used by ANDVT) rather than parameter values so that the dynamic range of each parameter is identical. We include both spectral parameters (reflection coefficients) as well as amplitude (speech root-mean-square value) parameters in Eq. (9) because they are both significant to speech intelligibility yet their trajectories are often uncorrelated.

Eliminating interframe redundancy reduces the data rate from 2400 b/s to approximately 1800 b/s. Further data rate reduction is achieved by reducing intraframe redundancies; this is discussed in the next section.

Intraframe Redundancy Reduction by Vector Quantization

The term "intraframe redundancy" refers to the redundancy within a frame of encoded speech parameters. Two significant reasons why encoded reflection coefficients of the 2400-b/s LPC have intraframe redundancies are:

- *Just-Noticeable Differences:* As listed in Table 2, the reflection coefficients for voiced speech are quantized to 5, 5, 5, 5, 4, 4, 4, 4, 3, and 2 bits, respectively. Thus, the 2400-b/s LPC has the capability to reproduce 2^{41} (or 2.2 trillion) speech sounds. Numerous combinations of reflection coefficients are redundant in the sense that we cannot hear any difference between them.

- *Nonspeech Sounds:* Not only can we not hear the difference, many combinations of reflection coefficients do not represent human voice. Note that voiced speech always has the first resonant frequency somewhere between 200 and 800 Hz. The rooster's crow, for example, has no resonant frequency within that frequency region. The 2400-b/s LPC is capable of producing a crow sound. We will implement the 1200-b/s LPC in such a way that encoded reflection coefficients are selected from the reflection coefficients derived from human speech.

Intraframe redundancy can be reduced by jointly quantizing all reflection coefficients by pattern matching (i.e., vector quantization) because the quantized reflection coefficients are derived only from human voices Furthermore, no two sets of quantized reflection coefficients generate sounds which are perceptually indistinguishable. Three steps are involved in implementing a vector quantization process: template collection, template storage, and template matching. These are discussed next.

Template Collection

We chose three tables of templates to encode the reflection coefficients (Table 4): two tables are derived from templates collected from voiced speech, and the remaining table is constructed from unvoiced speech templates.

Table 4 — Number of Spectral Patterns for 2400-b/s LPC and 1200-b/s LPC

Reflection Coefficients		2400-b/s LPC	1200-b/s LPC	
		Total Number of Spectral Patterns	Total Number of Spectral Patterns (Template Size)	Template Library
Unvoiced Speech	1st through 4th Coefficients	2^{20} = 1,048,576	2^{13} = 8,192	Table No. 1
	5th through 10th Coefficients	Not transmitted*	Not transmitted*	
Voiced Speech	1st through 4th Coefficients	2^{20} = 1,048,576	2^{13} = 8,192	Table No. 2
	5th through 10th Coefficients	2^{21} = 2,097,152	2^{13} = 8,192	Table No. 3

* Freed bits are used to protect acoustically more sensitive parameters. Section 2 describes the 2400-b/s LPC; Section 3 describes the 1200-b/s rate converter.

We generate a representative number of reflection coefficient templates by analyzing many representative voice samples. As depicted in Fig. 10, templates are generated by the following steps:

Step 1: The first incoming reflection coefficient set is the first template, and it is stored in memory.

Step 2: The second incoming coefficient set is compared with the stored template. If the spectral difference is within 2 dB, the incoming coefficient set is regarded as being the same family; thus it is discarded. Otherwise, it is stored as a new template. We used a 2 dB spectral error criterion because we begin to hear the difference between two sets of reflection coefficients having a spectral difference greater than 2 dB.

Step 3: Step 2 is repeated until the maximum allowable template size (i.e., 2^{13} = 8,192) is reached. Actually we collected more than the maximum number, pending elimination of the least-frequently-used templates later on to meet the required maximum table size.

The speech source for table generation was 420 speakers uttering 8 sentences each, excerpted from the "Texas Instrument - Massachusetts Institute of Technology (TIMIT) Acoustic-Phonetic Speech Data Base." The TIMIT database was designed to provide acoustic phonetic speech data for the development and evaluation of speech recognizers. TIMIT was digitally recorded by TI, transcribed by MIT, and formatted and mastered for CD-ROM by National Institute of Standard and Technology [7].

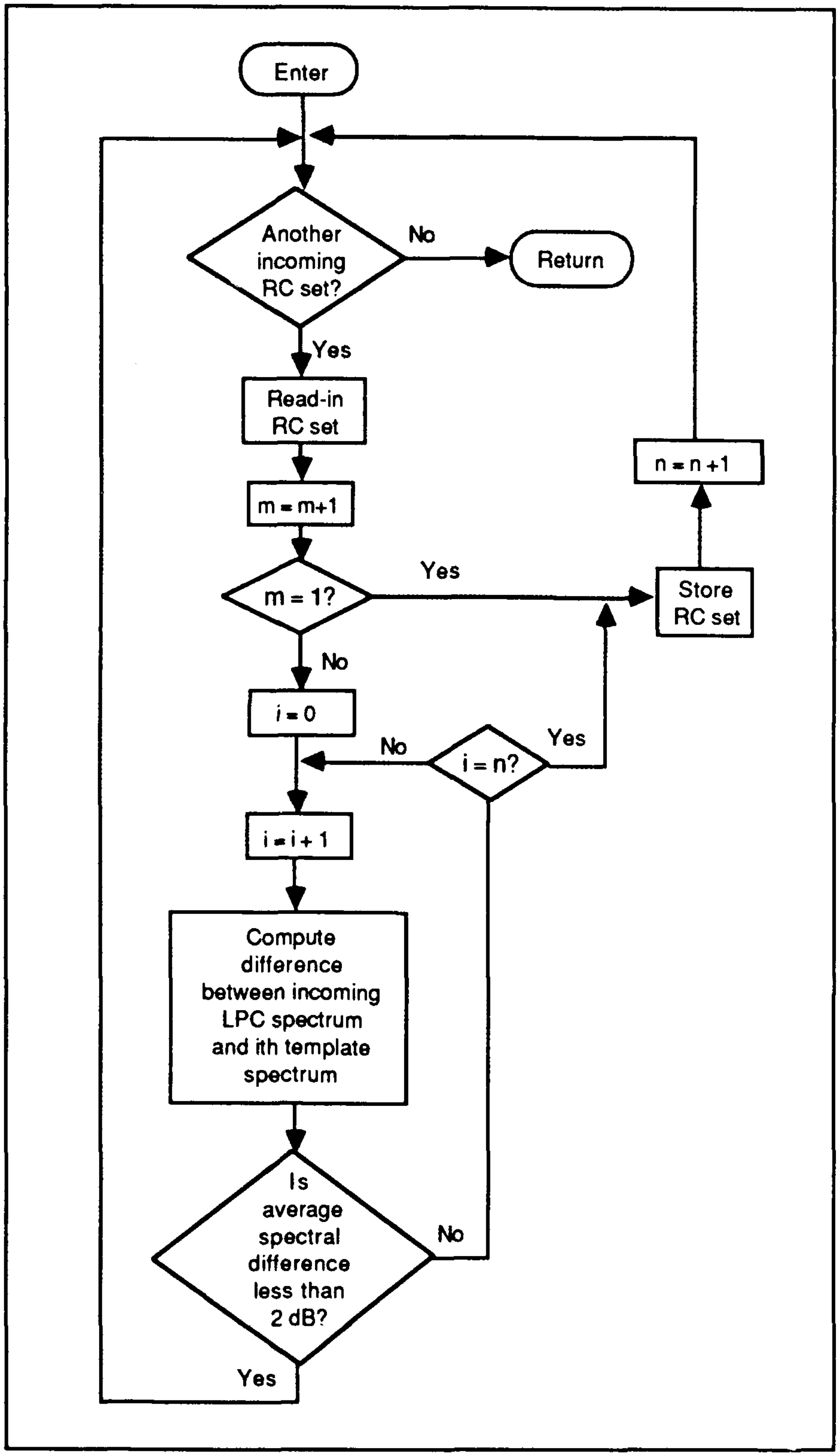


Fig. 10 — Reflection coefficient (RC) template collection process. This process is executed off-line only once prior to rate converter implementation. Index *m* denotes a frame of training speech samples, index *n* denotes the template size (maximum number of *n* is 8,192), and index *i* denotes a count of the pattern matching loop ($1 \leq i \leq n$).

Certain features are desirable in a speech database. It must contain:

- all the phonetic elements (i.e., vowels, fricatives, stops, liquids, glides, etc.),
- a fair amount of dialects that introduce variations in formant distributions for a given phonetic element, and
- a representative number of voice characteristics (male, female, clear, resonant, anti-resonant, nasal, raspy, etc.).

As far as we can determine, the TIMIT database has more of these features than any other collection. Ten typical sentences found in the TIMIT database are:

- Pizzerias are convenient for a quick lunch.
- Put the butcherblock table in the garage.
- Drop five forms in the box before you go out.
- Her wardrobe consists of only skirts and blouses.
- Elderly people are often excluded.
- Objects made of pewter are beautiful.
- The morning dew on the spider web glistened in the sun.
- Cheap stockings run the first time they're worn.
- Don't do Charles' dirty dishes.
- Calcium makes bones and teeth strong.

Template Storage

Once templates are collected, they are stored so that templates with similar spectral patterns are clustered. Then, computation for selecting the best-matched templates is reduced by searching only a portion of the 8192 templates. To accomplish this, we store all templates in ascending order of coefficients. Table 5 lists the population counts within each subgroup that have the same leading reflection coefficient index.

The stored patterns are continuously numbered from 0 to 8191 or represented by a 13-bit code (b1, b2, b3, ..., b13). We store templates in a binary tree (Fig. 11) to make the bit b1 most significant and the bit b13 least significant. An unequal sensitization of bits is beneficial for error protection in the voice processor as well as in the modem.

Template Matching

The reflection coefficient indices from each frame are compared with the templates. We used a weighted Euclidian distance, denoted by *d(i)*, to select the best-matching reflection coefficient set:

$$d(i) = \sum_j w(j) \left| k_j - \kappa_j(i) \right| \qquad (10)$$

where k_j is the *j*th incoming reflection coefficient, and $\kappa_j(i)$ is the *j*th reflection coefficient of the *i*th template. Both k_j and $\kappa_j(i)$ are actual coefficient values (i.e., decoded values). The quantity *w(j)* is the weighting factor associated with the *j*th incoming reflection coefficient and is proportional to the spectral-error sensitivity of the *j*th reflection coefficient to the LPC spectrum.

We do not exhaustively search all 8,192 templates to find the best match; rather, we search only 1000 templates around the leading coefficient index (i.e., index associated with the first reflection coefficient for Tables 1 and 2 or fifth reflection coefficient for Table 3). Figure 12 is the flow diagram of the template-matching process.

Table 5 — Population Counts Based on the Leading Reflection Coefficient Index

(1st through 4th coefficients)

First Ref. Coeff.		Population Counts	
Index	Code	Table 1 (Unvoiced)	Table 2 (Voiced)
-15	10001	0	0
-14	10010	52	0
-13	10011	189	6
-12	10100	399	30
-11	10101	386	37
-10	10110	414	59
-9	10111	520	113
-8	11000	567	150
-7	11001	603	234
-6	11010	536	232
-5	11011	504	254
-4	11100	565	370
-3	11101	490	395
-2	11110	433	398
-1	11111	419	491
0	00000	490	603
1	00001	330	443
2	00010	302	541
3	00011	292	585
4	00100	273	681
5	00101	169	569
6	00110	137	605
7	00111	84	586
8	01000	29	392
9	01001	6	266
10	01010	2	106
11	01011	1	39
12	01100	0	7
13	01101	0	0
14	01110	0	0
15	01111	0	0
Total		8192	8192

(5th through 10th coefficients)

Fifth Ref. Coeff.		Population Counts
Index	Code	Table 3 (Voiced)
-8	1000	369
-7	1001	335
-6	1010	405
-5	1011	452
-4	1100	451
-3	1101	555
-2	1110	524
-1	1111	602
0	0000	613
1	0001	604
2	0010	597
3	0011	577
4	0100	519
5	0101	477
6	0110	381
7	0111	731
Total		8192

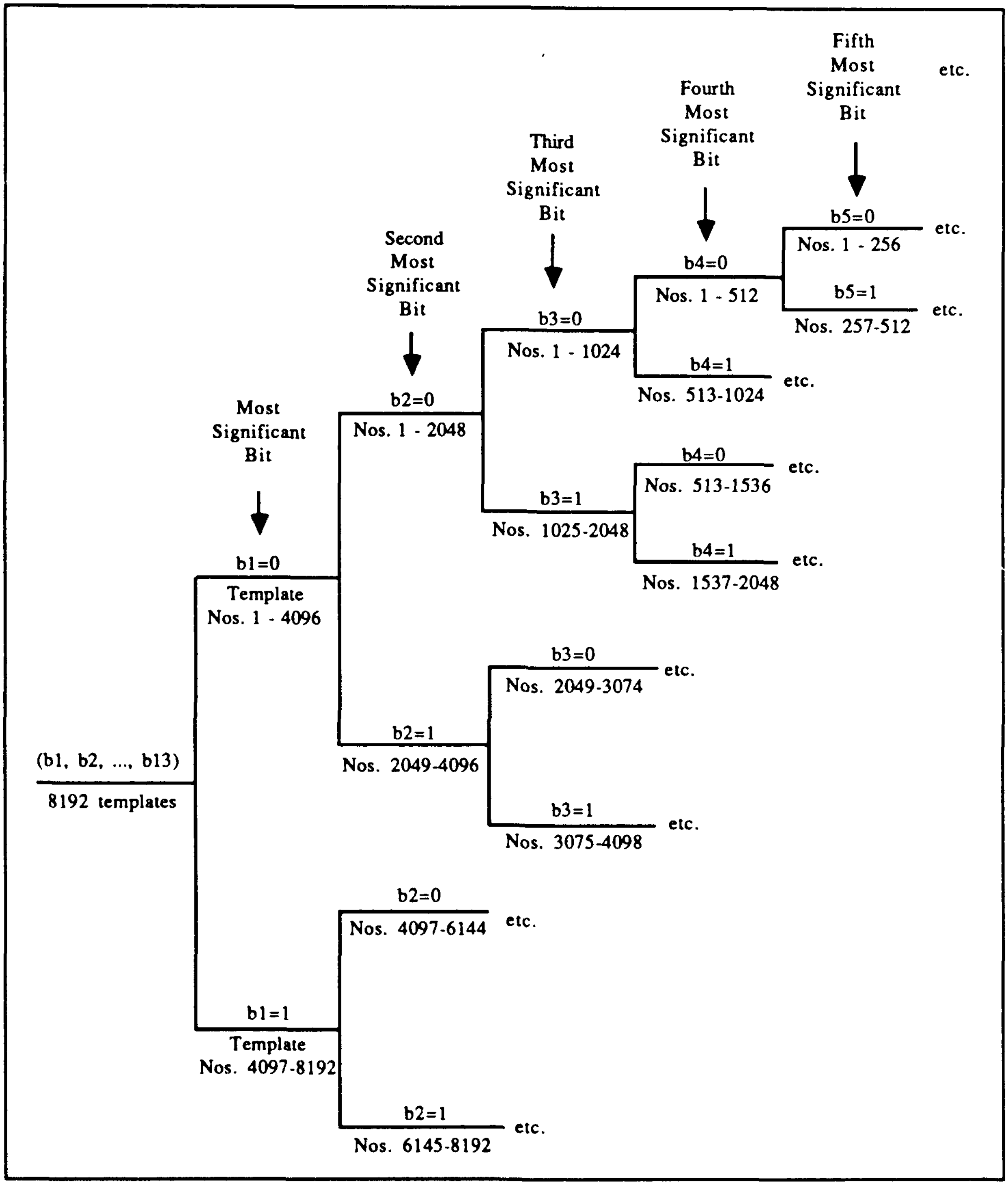


Fig. 11 — Tree arrangement to group reflection coefficient templates. Since the first reflection coefficient is spectrally most sensitive, templates are arranged based on the ascending order of the first reflection coefficient value.

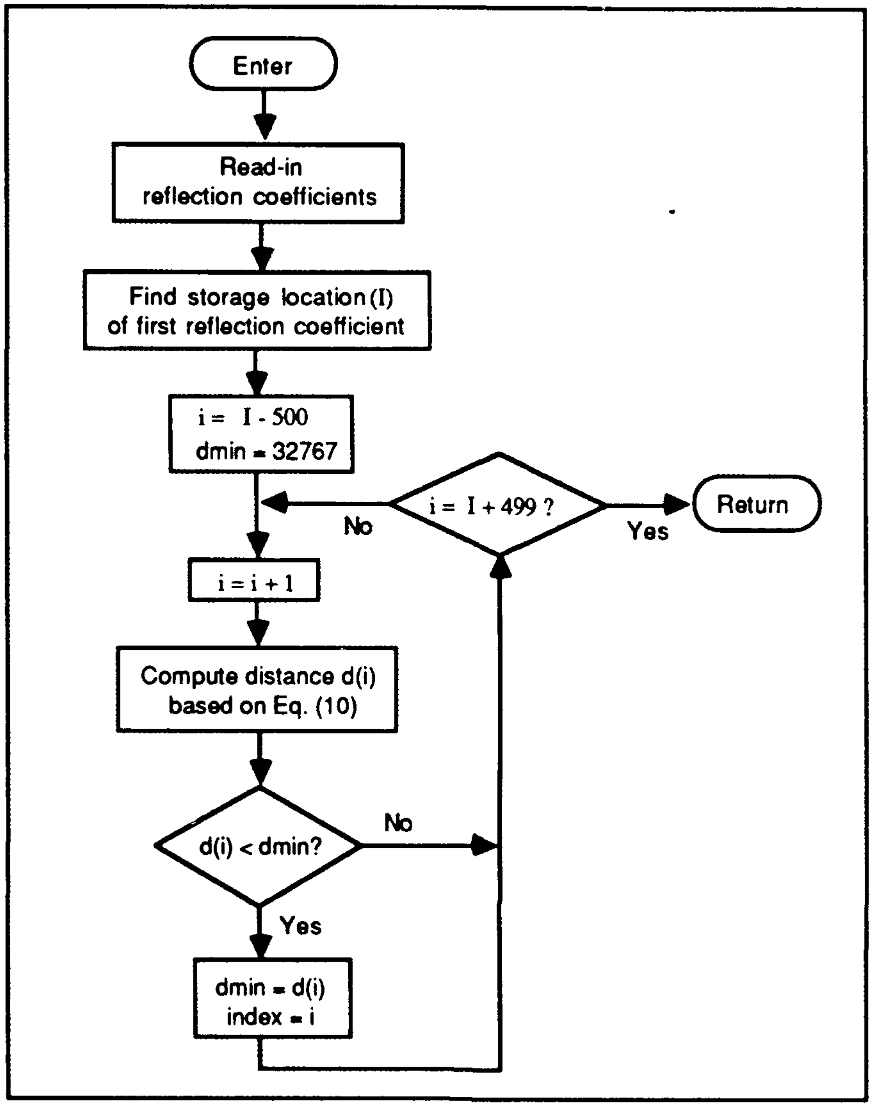

Enter
Read-in reflection coefficients
Find storage location (I) of first reflection coefficient
i = I - 500
dmin = 32767
i = I + 499 ?
No
Yes
Return
i = i + 1
Compute distance d(i) based on Eq. (10)
d(i) < dmin?
No
Yes
dmin = d(i)
index = i

Fig. 12 —Template matching process.

Weighting Factors for Template Matching

No two reflection coefficients contribute equally to the LPC spectrum. Thus, the selection of the best-matched reflection coefficient set from a collection of templates must be based on the individually weighted coefficient difference, as indicated in Eq. (10). The spectral-error sensitivity of one reflection coefficient is dependent on all the reflection coefficients. In other words, for a given reflection coefficient set, the spectral-error sensitivity of each coefficient is predetermined. Unfortunately, the derivation of such an analytical expression is untractable because reflection coefficients are related to the LPC spectrum via a many-fold transformation, as will be shown.

Thus, we resort to a numerical method for determining the spectral-error sensitivity for a given coefficient set. This approach is feasible because we are interested in a *small error* in Eq. (10) for choosing the best-matched reflection coefficient set. Thus, the spectral-error sensitivity of each coefficient is obtained by the LPC spectral change (average value over a 4 kHz passband) caused by a small reflection coefficient perturbation. We repeat these computations off-line for each reflection coefficient set in the template, and the spectral-error sensitivities are stored with the reflection coefficients.

We do not generate weighting factors for the fifth through tenth coefficients because the spectral-error sensitivity of these coefficients does not vary as much as for the first four reflection coefficients. Thus, we let weighting factors for these coefficients be unity. For the first through fourth reflection coefficients, the weighting factor is proportional to the spectral-error sensitivity. We follow these steps to determine their spectral-error sensitivities:

1. *Conversion from reflection coefficients to prediction coefficients*: The first through fourth reflection coefficients are converted to four prediction coefficients by the recursive expressions derived from Eq. (7):

$$\left[\alpha_{1/1}\right] = [1]\left[k_1\right] , \tag{11}$$

$$\begin{bmatrix} \alpha_{1/2} \\ \alpha_{2/2} \end{bmatrix} = \begin{bmatrix} 1 & -\alpha_{1/1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} , \tag{12}$$

$$\begin{bmatrix} \alpha_{1/3} \\ \alpha_{2/3} \\ \alpha_{3/3} \end{bmatrix} = \begin{bmatrix} 0 & -\alpha_{1/1} & -\alpha_{2/2} \\ 0 & 1 & -\alpha_{1/2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} , \tag{13}$$

and

$$\begin{bmatrix} \alpha_{1/4} \\ \alpha_{2/4} \\ \alpha_{3/4} \\ \alpha_{4/4} \end{bmatrix} = \begin{bmatrix} 1 & -\alpha_{1/1} & -\alpha_{2/2} & -\alpha_{3/3} \\ 0 & 1 & -\alpha_{1/2} & -\alpha_{2/3} \\ 0 & 0 & 1 & -\alpha_{1/3} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{bmatrix} , \tag{14}$$

where $a_{j/4}$ is the jth reflection coefficient of the fourth order system, which was previously denoted simply by α_j.

2. *Frequency response of the LPC analysis filter*: The LPC analysis filter transforms speech samples to prediction residual. From Eq. (1), the transfer function of the LPC analysis filter, denoted by $A(z)$, is

$$A(z) = 1 - \alpha_1 z^{-1} - \alpha_2 z^{-2} - \alpha_3 z^{-3} - \alpha_4 z^{-4} - \ldots \tag{15}$$

in which

$$z = \varepsilon^{j\omega\tau} \tag{16}$$

where ω is the frequency in radians/second, τ is the speech sampling time interval (125 μs), and $j = \sqrt{-1}$. Thus, the amplitude response of the LPC analysis filter is

$$A(\omega) = \left| 1 - \alpha_1 \varepsilon^{-j\omega\tau} - \alpha_2 \varepsilon^{-2j\omega\tau} - \alpha_3 \varepsilon^{-3j\omega\tau} - \alpha_4 \varepsilon^{-4j\omega\tau} - \ldots \right| . \tag{17}$$

3. *Computation of LPC spectrum*: The LPC spectrum, denoted by H(ω), is an inverse of $A(\omega)$ expressed by Eq. (16). Thus,

$$H(\omega) = -20 \log \left\{A(\omega)\right\} . \tag{18}$$

We evaluate the spectrum with a frequency step of 10 Hz from 100 to 3800 Hz. Figure 13 is an example of the resultant spectrum.

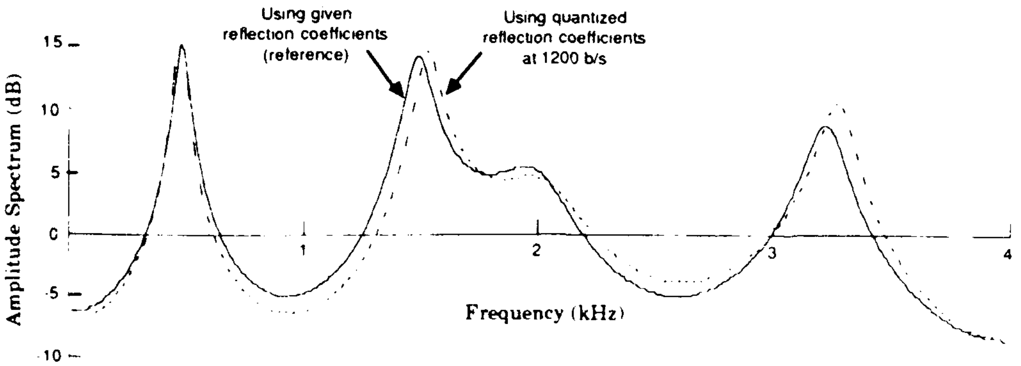


Fig. 13 — LPC spectra using quantized and unquantized reflection coefficients at 1200 b/s.

4. *Spectral-Error Sensitivity*: The spectral-error sensitivity is the gradient of the LPC spectrum $H(\omega)$ that results from a small change in a reflection coefficient. It is analytically intractable to derive such an expression. Thus, we derive tne spectral-error sensitivities numerically for a given set of reflection coefficients. First, we compute the LPC spectrum $H(\omega)$ based on the aforementioned procedures. Then, the jth reflection coefficient is perturbed by a small amount, and the resultant LPC spectrum $H'(\omega)$ is computed. The root-mean-square difference between the two LPC spectra (ΔH) is computed by

$$\Delta H = \frac{1}{N} \sum_{n=1}^{N} \left[H(w) - H'(w)\right]^2 . \tag{19}$$

The ratio of ΔH to Δk_j is the spectral-error sensivity $S(j)$. Thus,

$$S(j) = \frac{\Delta H}{\Delta k_j} = \frac{\frac{1}{N} \sum_{n=1}^{N} \left[H(w) - H'(w)\right]^2}{\Delta k_j} , \tag{20}$$

where N is the number of spectral points. We compute the LPC spectrum from 100 to 3800 Hz at a 10 Hz separation. Thus, $N = 381$. Figure 14 is a plot of the spectral-error sensitivity of each reflection coefficient of the 10th-order LPC.

Bit Allocation

For a data rate of 1200 b/s, the number of bits in each superframe of 90 ms (four LPC frames) is 108 bits (Table 6). We allocate bits to each parameter based on the following considerations:

Reflection Coefficients: A total of 26 bits are available per frame or 78 bits per superframe. If speech is voiced, the first four reflection coefficients are encoded in 13 bits per frame, and the remaining coefficients are encoded by another 13 bits per frame. If speech is unvoiced, the first through fourth reflection coefficients are encoded by 13 bits per frame (similar to voiced speech), but the remaining 13 bits per frame are used for protecting the first through fourth coefficients.

Amplitude: The ANDVT amplitude parameter is 5 bits per frame. It is passed directly to the rate converter (i.e., no code conversion is necessary). Thus, 15 amplitude bits are allocated per super frame.

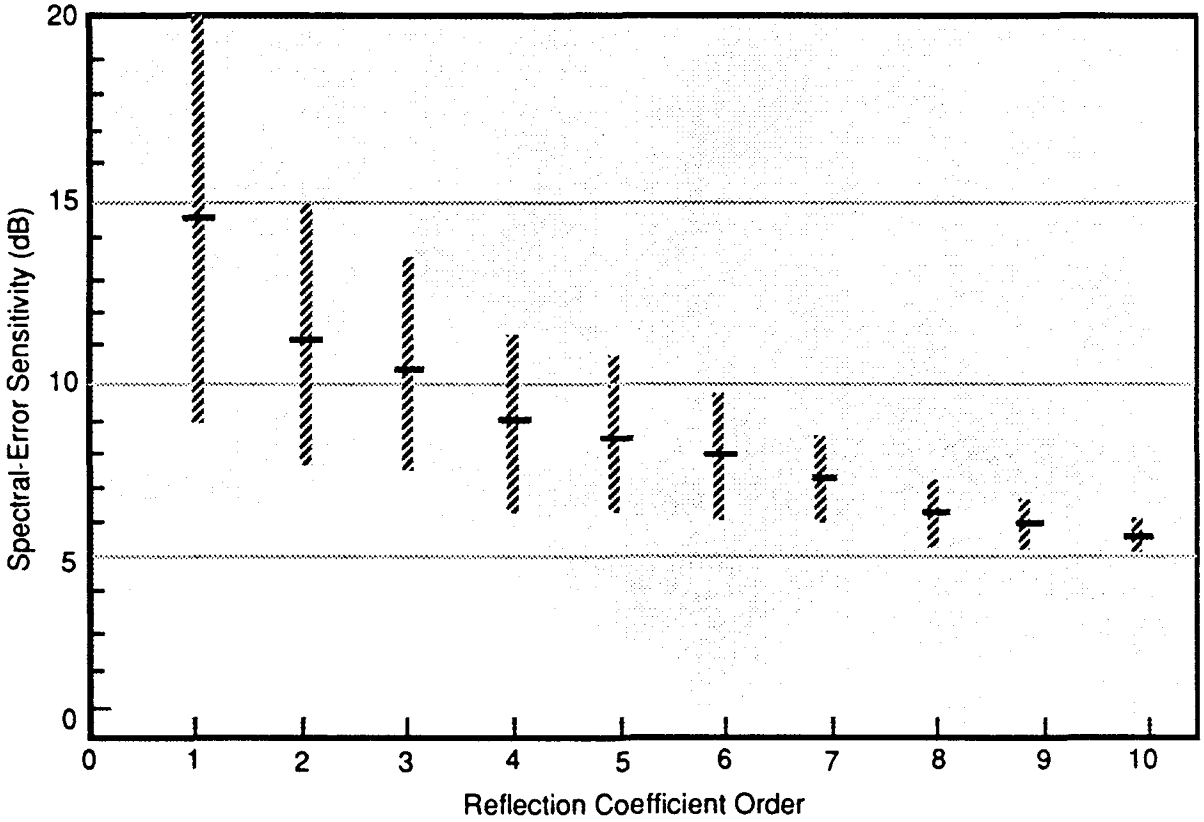


Fig. 14 — Spectral-error sensitivities of reflection coefficients of the 10th-order LPC. This figure plots the distribution (mean and standard deviation) of spectral-error sensitivity of each reflection coefficient.

Table 6 — Bit Allocation to Generate 1200-b/s Speech Data

	Number of Updates per Superframe	No. of Bits per Superframe	Remarks
Reflection Coefficients	3	3 x 26 = 78	If speech is voiced, 26 bits per frame represent 10 coefficients. If speech is unvoiced, 13 bits per frame represent the first four coefficients; the remaining 13 bits per frame are for error protection.
Amplitude	3	3 x 5 = 15	The 2400-b/s LPC amplitude code is transmitted directly.
Voicing Decision	4	4 x 1 = 4	One bit per frame (including skipped frame) is transmitted.
Pitch Period	1	1 x 7 = 7	The 2400-b/s LPC pitch/voicing code is transmitted directly.
Skipped Frame Index	1	1 x 2 = 2	There are three possibilities for choosing one out of 3 frames.
Sync	1	1 x 2 = 2	Two "1s" and two "0s" are alternately transmitted per superframe.
TOTAL		108	

Voicing Decision: The voicing decision is 1 bit per frame. We transmit a separate voicing decision for each frame, including the voicing decision for the skipped frame. Thus, we allocate four bits per superframe.

Pitch Period: The pitch period in normal conversation does not vary as rapidly as the other parameters. Particularly, tactical communication is delivered without normal pitch inflections. Thus, we transmit the pitch period only once per superframe. The transmitted pitch period is a median value of the individual pitch periods within the superframe. To simplify the rate conversion process, the ANDVT 7-bit pitch/voicing code (the 6-bit pitch code is combined with 1-bit voicing code to provide 1-bit error protection cability) is transmitted directly. The decoded voicing bit from this pitch/voicing information is discarded at the receiver (one unused bit per superframe).

Frame Selection: Because one of the first three frames is skipped in each superframe, two bits are assigned to indicate which frame is skipped.

Synchronization: Two bits are allocated in each superframe (i.e., one sync bit for every 53 bits of data). Thus, the sync-bit density is identical to that of the current 2400-b/s LPC.

Error Protection

If speech is voiced, no parameters are error-protected. If speech is unvoiced, 13 bits are used to protect the amplitude parameter and reflection coefficients. The amplitude code is protected in the same manner as the 2400-b/s LPC. Thus, the ANDVT amplitude code (including the error protection code) is directly transfered to a 1200-b/s bit stream. The remaining nine bits protect four reflection coefficients.

Reflection coefficients are protected by the Hamming (8,4) code (i.e., Hamming (7,4) code with an overall parity bit) for the amplitude parameters because decoding can be carried out efficiently by a table look-up (i.e., no computation is necessary). The most significant eight bits of the reflection coefficient vector are protected by two sets of Hamming (8,4) codes, and the five least significant five bits are left unprotected. We use the Hamming (8,4) listed in Table 7.

Table 7 — Hamming (8,4) Code to Protect Four Reflection Coefficient Bits
(This table is generated by using the generator matrix given in Eq. (21))

Speech Data	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
Error-Protection Code	0	7	B	C	D	A	6	1	E	9	5	2	3	4	8	F

The information contained in Table 7 (with an exception of data of "0" and code of "0") can be generated by adding appropriate rows of the generator matrix G(8,4) given in Eq. (21). The left four bits represent speech data, and the right four bits represent error-protection code with parity bit. For each partitioned matrix, the far left bit is the most-significant bit, and the far right bit is the least-significant bit.

$$G(8,4) = \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{array}\right) \quad (21)$$

A Hamming code has the capability of correcting a single error and detecting a multiple of two errors. Figure 15 shows the word-error rate in terms of the random bit-error rate.

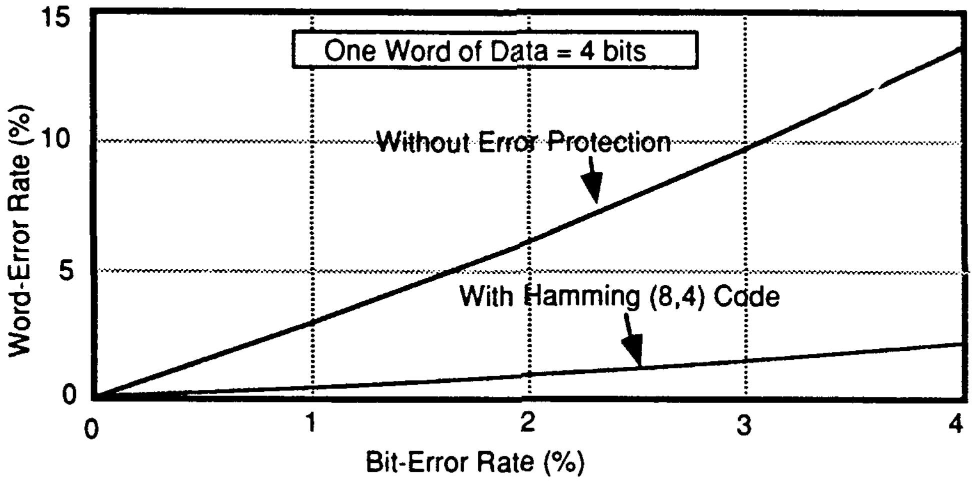


Fig. 15 — Word-error rate in terms of random bit error rate. As noted from figure, Hamming (8,4) reduces the word-error rate (having more than one error in a 4-bit word) by one order of magnitude.

To reduce decoding computations, the decoded output for all possible input words ($2^8 = 256$ words) is precomputed by the well-established decoding procedure [8] to generate a table. Then, input data can be used as an address to read the decoded data directly from the table. Table 8 is the decoding table.

4. SPEECH INTELLIGIBILITY AT 1200 b/s

The Diagnostic Rhyme Test (DRT) evaluates the discriminability of initial consonants of monosyllable rhyming word pairs. For many years, DRT scores have been widely used as a diagnostic tool to refine voice processors. It has also been effectively used to rank several competing voice processors. Over the years, a large amount of DRT data has been collected from different voice processors under varied operating conditions. According to our experience, DRT scores are dependable (i.e., scores are repeatable under retesting), and they often reveal latent defects of synthetic speech that are not easily discernible through casual listening. Figure 16 shows DRT scores for both 1200 and 2400 b/s.

5. CONCLUSIONS

This report has shown that highly intelligible speech at 1200 b/s can be generated by converting the 2400-b/s LPC bit stream. Such a rate conversion approach can provide the ANDVT user with 1200-b/s speech data for specialized applications such as improved bit-error performance at 2400 b/s or voice/data integration at 2400 b/s. Rate conversion is performed external to the ANDVT and requires no modifications to ANDVT software or hardware.

Through rate conversion, speech intelligibility is degraded by only 1.8 points (for male speakers) and 1.5 points (for female speakers) as measured by the DRT. In other words, speech intelligibility at 1200 b/s is nearly as good as speech intelligibility at 2400 b/s.

Table 8 — Decoding Table for Hamming (8,4) Code

Received Code word	Decoded Data	Received Code word	Decoded Data	Received Code word	Decoded Data	Received Code word	Decoded Data	Received Code word	Decoded Data	Received Code word	Decoded Data	Received Code word	Decoded Data	Received Code word	Decoded Data
0	0	32	0	64	0	96	-	128	0	160	-	192	-	224	E
1	0	33	-	65	-	97	7	129	-	161	A	193	C	225	-
2	0	34	-	66	-	98	6	130	-	162	B	194	C	226	-
3	-	35	2	67	C	99	-	131	C	163	-	195	C	227	C
4	0	36	-	68	-	100	6	132	-	164	A	196	D	228	-
5	-	37	8	69	4	101	-	133	A	165	A	197	-	229	A
6	-	38	6	70	6	102	6	134	8	166	-	198	-	230	6
7	1	39	-	71	-	103	6	135	-	167	A	199	C	231	-
8	0	40	-	72	-	104	E	136	-	168	E	200	E	232	E
9	-	41	2	73	4	105	-	137	9	169	-	201	-	233	E
10	-	42	2	74	5	106	-	138	8	170	-	202	-	234	E
11	2	43	2	75	-	107	2	139	-	171	2	203	C	235	-
12	-	44	3	76	4	108	-	140	8	172	-	204	-	236	E
13	4	45	-	77	4	109	4	141	-	173	A	205	4	237	-
14	8	46	-	78	-	110	6	142	8	174	8	206	8	238	-
15	-	47	2	79	4	111	-	143	8	175	-	207	-	239	F
16	0	48	-	80	-	112	7	144	-	176	B	208	D	240	-
17	-	49	7	81	7	113	7	145	9	177	-	209	-	241	7
18	-	50	B	82	5	114	-	146	B	178	B	210	-	242	B
19	1	51	-	83	-	115	7	147	-	179	B	211	C	243	-
20	-	52	3	84	D	116	-	148	D	180	-	212	D	244	D
21	1	53	-	85	-	117	7	149	-	181	A	213	D	245	-
22	1	54	-	86	-	118	6	150	-	182	B	214	D	246	-
23	1	55	1	87	1	119	-	151	1	183	-	215	-	247	F
24	-	56	3	88	5	120	-	152	9	184	-	216	-	248	E
25	9	57	-	89	-	121	7	153	9	185	9	217	9	249	-
26	5	58	-	90	5	122	5	154	-	186	B	218	5	250	-
27	-	59	2	91	5	123	-	155	9	187	-	219	-	251	F
28	3	60	3	92	-	124	3	156	-	188	3	220	D	252	-
29	-	61	3	93	4	125	-	157	9	189	-	221	-	253	F
30	-	62	3	94	5	126	-	158	8	190	-	222	-	254	F
31	1	63	-	95	-	127	F	159	-	191	F	223	F	255	F

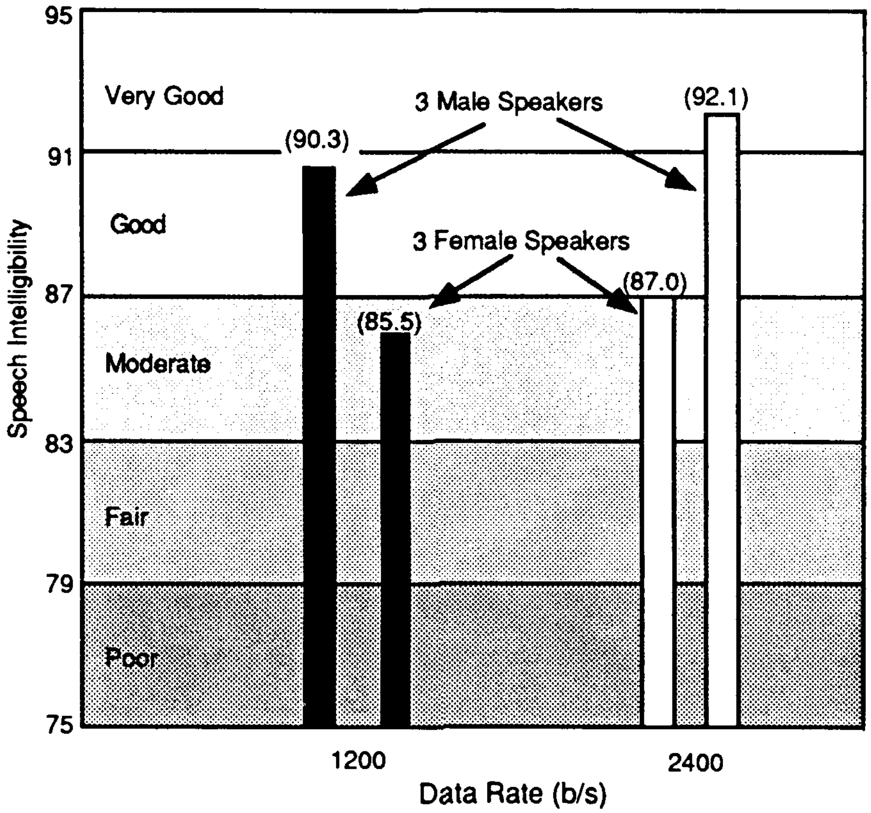


Fig. 16 — DRT scores at 1200 and 2400 b/s. As noted, the intelligibility at 1200 b/s is only one or two points below that of 2400 b/s. Casual listening cannot discern the difference between 1200- and 2400-b/s speech.

6. ACKNOWLEDGMENTS

The effort reported in this report was supported in part by Dennis McGregor of NRL, and also Timothy McChesney and Sharon James of SPAWAR. The rate-conversion algorithm presented in this report will be used for *the narrowband voice/data integration demonstration* that is being undertaken by McGregor at NRL.

7. REFERENCES

1. G.S. Kang, "Narrowband Integrated Voice/Data System Based on the 2400-b/s LPC," NRL Report 8942, December 1985.

2. Federal Standard 1015, "Analog to Digital Conversion of Voice by 2,400 bits/s Linear Predictive Coding," published by General Services Administration (GSA), November 28, 1984. Copies are for sale at the GSA Specification Unit (WFSIS), Room 6039, 7th and D Street SW, Washington, DC 20407.

3. T.O. Lewis and P.L. Odell, *Estimation in Linear Models* (Prentice, Hall, Englewood Cliffs, NJ, 1971).

4. G.W. Stewart, *Introduction to Matrix Computations* (Academic Press, New York, 1973).

5. E. Blackman, R. Viswanathan, and J. Makhoul, "Variable-to-Fixed Conversion of Narrowband LPC Vocoder," 1977 IEEE International Conference on Acoustics, Speech, and Signal Processing, 409-412 Hartford, CT, May 9-11, 1977.

6. E.T. Klemmer, "Human Factor Problems in Satellite Telephoning," *Human Factors*, 475-840, December 1966.

7. J.S. Carofolo, "DARPA TIMIT Acoustic-Phonetic Speech Database," National Institute of Standards and Technology, Gaithersburg, MD 20899.

8. M.Y. Rhee, *Error-Correcting Coding Theory* (McGraw-Hill Publishing Company, New York, New York, 1989).