PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : G10L 5/00, G06F 11/00	A1	(11) International Publication Number: WO 93/05502 (43) International Publication Date: 18 March 1993 (18.03.93)

(21) International Application Number: PCT/US92/06352

(22) International Filing Date: 3 August 1992 (03.08.92)

(30) Priority data:
755,241 5 September 1991 (05.09.91) US

(71) Applicant: MOTOROLA, INC. [US/US]; 1303 East Algonquin Road, Schaumburg, IL 60196 (US).

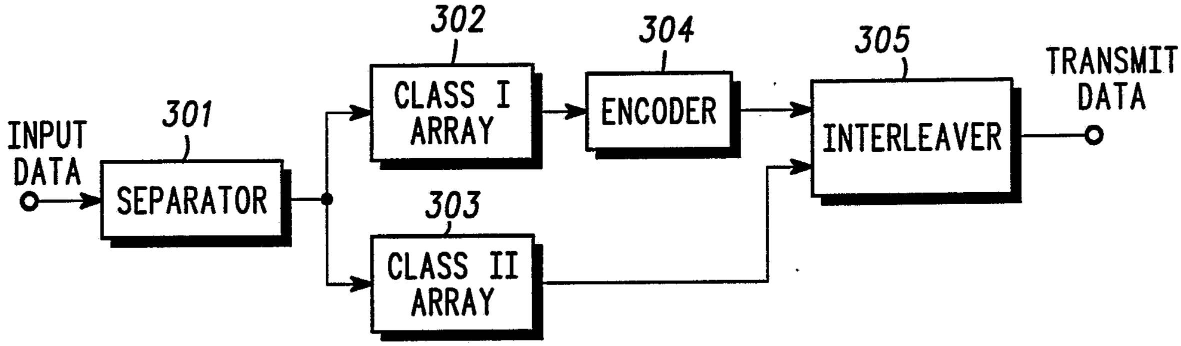
(72) Inventors: GERSON, Ira, A. ; 1120 Nottingham Lane, Hoffman Estates, IL 60195 (US). WINTER, Eric, H. ; 728 South Elmwood Avenue, Oak Park, IL 60304 (US). JASIUK, Mark, A. ; 6221 North Melvina, Chicago, IL 60646 (US).

(74) Agents: PARMELEE, Steven, G. et al.; Motorola, Inc., Intellectual Property Dept./JWH, 1303 E. Algonquin Road, Schaumburg, IL 60196 (US).

(81) Designated States: JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, SE).

Published
With international search report.

(54) Title: ERROR PROTECTION FOR MULTIMODE SPEECH CODERS



(57) Abstract

A method for protecting information bits wherein input data bits, at least some of which are to be protected, are sorted based upon information determined from a subset of the input data bits. An error control coding technique is applied to at least some of the sorted bits. In the preferred embodiment, an input data stream of voice coder bits is separated into arrays of bits. A first array (302) comprises voice coder bits needing error protection, with the bits arranged in order of importance determined by voicing mode. The second array (303) comprises bits that will not be error protected. The bits from the first array are provided to the input of an encoder (304), then the encoded bits are combined (305) with the bits from the second array (303) to form a bit stream.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
MC	Monaco
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SK	Slovak Republic
SN	Senegal
SU	Soviet Union
TD	Chad
TG	Togo
UA	Ukraine
US	United States of America

ERROR PROTECTION FOR MULTIMODE SPEECH CODERS

Technical Field

This invention relates generally to digital data transmission and in particular to speech coders, and is more particularly directed toward a method for providing error protection for appropriate bits.

Background of the Invention

Two-way radio communication of voice signals is an essential capability for modern society. Organizations involved in police work, public safety, and transportation are major users of voice communication, to say nothing of military users. With increasing demands being placed upon available radio frequency spectrum, much research has been done on maximizing spectrum efficiency.

One way to increase spectrum efficiency is to compress speech signals prior to transmission. This compression operation, as is well-known in the art, reduces bandwidth requirements for the transmission of voice signals and permits the assignment of more RF (radio frequency) communication channels within a given range of frequencies. Similarly, speech compression algorithms may also be applied to digital speech storage.

With the increasing emphasis on digital communication, compression schemes peculiar to digital systems have received much attention. Compression schemes dedicated to digitized speech are commonly referred to as voice coders. U.S. Patent No. 4,933,957 to Bottau et al. describes a low bit rate voice coding method and system exemplary of those commonly known in the art.

Linear predictive coding, or LPC, is an example of a popular voice coding technique. In LPC, an attempt is made to approximate human speech by deriving appropriate models for both the human vocal tract and for excitations applied to the vocal tract. Since speech is a very repetitive type of signal, the amount of information required to allow a decoder to accurately reproduce a speech waveform can be much reduced. Depending upon the nature of the speech being transmitted, some bits may be more perceptually significant to the reconstructed speech than others.

As with any type of digital signal, decisions must be made at the decoder regarding whether a logic "1" level or a logic "0" level was originally transmitted. Error control coding, a concept well-understood in the art, is often employed to increase the likelihood that the decoder will make correct decisions. Of course, it is self-defeating to compress digital speech for transmission only to add a large number of error control bits. A compromise must be reached in order to maximize the effectiveness of a given speech compression algorithm while trying to ensure speech quality by guaranteeing error-free reception of critical bits. Of course, critical bits can be identified for a variety of data transmission scenarios that do not involve speech coding applications.

Accordingly, a need arises for a method for error protecting critical bits for transmission, where the specific bits requiring protection may be dependent upon a subset of input bits such as, for example, the set of bits identifying the type of speech waveform being coded. The specific bits requiring protection could also be determined by expected communication channel conditions.

Summary of the Invention

This need and others are satisfied by the method of the present invention for error protecting information bits The method comprises providing input data bits, at least some of which are to be protected, sorting the input data bits based upon information determined from a subset of the input data bits, and applying an error control coding technique to at least some of the sorted bits.

Brief Description of the Drawings

FIG. 1 is an illustration of a typical speech waveform;

FIG. 2A is a block diagram of a VSELP speech coder;

FIG. 2B illustrates, in more detail, VSELP codebooks used in the preferred embodiment;

FIG. 3 is a block diagram of an encoder system that may be used to implement the method of the present invention;

FIG. 4 is a graph of decoded bit error rate versus bit position for a variable rate convolutional encoder; and

FIG. 5 shows the use of a pre-decoder to select the appropriate post-decoder.

Description of a Preferred Embodiment

FIG. 1 illustrates a typical speech waveform. Human speech may be generally classified as either "voiced" or "unvoiced." In voiced speech, a perceptible, periodic, excitation is applied to the vocal tract. Voiced speech is commonly associated with vocalization of long vowel sounds, such as the long "a," "i," and "o" sounds in the word "radio." Consonant sounds not involving substantial periodic excitation are considered unvoiced.

FIG. 1 is also illustrative of the fact that voiced speech exhibits observable periodicities. There are long-term periodicities, such as those seen in segment 102, and short term correlations like those of segment 101. From a standpoint of probabilistic analysis, these tendencies for speech waveform characteristics to exhibit themselves result in relatively high short-term and long-term correlations that make linear prediction a viable voice coding concept. The relationship between these correlations and LPC techniques will be discussed in more detail below.

FIG. 2A depicts, in block diagram form, a vector-sum excited linear prediction (VSELP) voice coder. Speech signals, even when they can be characterized as voiced, exhibit stochastic (random) properties as well as periodic properties. In a VSELP system, vocal tract excitation is modeled by a combination of a first waveform vector selected from a fixed set of excitation vectors called a codebook (201), and a second vector selected by extracting a portion of a waveform based upon the past history of the speech being coded. This past history is stored in a long term predictor memory (202).

Long-term predictor (202) information is especially useful in coding voiced speech, where long-term correlations are predominant. Codebook-derived vectors (201) help in coding speech waveforms that are either voiced or unvoiced. In order to accommodate speech waveforms with varying degrees of voiced characteristic, both codebook vectors (201) and long term predictor vectors (202) are scaled by applying multiplicative gain factors (204 and 205, respectively).

These scaled vectors are summed (207) and the resulting excitation is applied to an LPC filter (208). In the preferred embodiment, the LPC filter is an IIR (infinite impulse response) filter implemented in a DSP (digital signal processor). The LPC filter (208) is primarily intended to model the vocal tract to which the excitation is applied. Reprogramming of LPC filter (208) coefficients may be effected periodically to optimize the speech coder output. In fact, the speech coder output is compared (209) with digitized speech (210) and the resulting error is minimized (211) by altering vector selection from both the codebook (201) and the long-term predictor (202).

In the preferred embodiment, the speech coder operates at 5.9 kilobits per second (kbps) with a frame length of 20 milliseconds (ms). The length of the frames or packets approximates the period of speech over which at least some parameters normally remain relatively constant. Examples of these relatively constant parameters include LPC filter coefficients and voicing mode. In the preferred embodiment, the voicing modes are unvoiced, slightly voiced, moderately voiced, and strongly voiced. The 20 ms frame is divided into four 5 ms subframes to accommodate parameters that change more frequently in speech waveforms. These more volatile parameters include excitation vector information and multiplicative gain values.

In actual practice, a frame contains 118 information bits. But, as discussed previously, not all bits are of equal importance. The long-term predictor vector adds little to the coding of unvoiced speech, but it is very important to proper speech waveform reconstruction in the strongly voiced voicing mode. Because of this variation in significance, the present invention permits speech coder bits to be selectively error protected based upon voicing mode, thus achieving the needed compromise between desirable error control coding of information bits and unwanted bandwidth expansion that would result from the addition of too many overhead bits.

FIG. 2B shows the various codebooks of the preferred embodiment in more detail. As discussed previously, the long term predictor vector adds little to the coding of unvoiced speech. Thus, for unvoiced mode, vectors from two VSELP codebooks (214 and 215) are selected instead. If the speech waveform being voice coded is slightly, moderately, or strongly voiced, the important long term predictor (212) vector is transmitted along with a single vector from a VSELP codebook (213). Of course, just as in the simpler example used above, each vector has an associated multiplicative gain value (220 through 223) applied to an appropriate multiplier (216 through 219) to adjust the amplitudes for optimum speech coding. Just as before, after the selected vectors are summed (224), they are applied to an LPC filter.

In the preferred embodiment, convolutional coding is used to afford error protection. Convolutional encoders and decoders are well-known in the art, and are also very easy to implement, particularly with the power of a DSP at the designer's disposal. Of course, this does not obviate the use of well-known block encoders, or even a combination of the two, to provide the necessary error protection.

FIG. 3 shows, in block diagram form, the encoder system of the present invention. The 118 bits of frame data are applied to the input of a data separator (301). Based on the voicing mode, the data separator (301) places the bits that are considered most important for that particular voicing mode in a memory array called the Class I array (302). The important bits assigned to the Class I array (302) are arranged in order from most important to least important. The data bits that are not considered important for the particular voicing mode are placed in a memory array designated the Class II array (303). These bits will not be subject to error control coding.

The bits from the Class I array (302) are input to a convolutional encoder (304). The encoder (304) of the preferred embodiment is switchable from a rate 1/3 encoder to a rate 1/2 encoder. This is often termed a multirate encoder. As is well-known in the art, data that is encoded using a rate 1/3 convolutional code will have a lower decoded bit error rate than data encoded with a rate 1/2 code. Of course, the lower bit error rate is accomplished at the expense of more overhead bits. FIG. 4 illustrates a well-known peculiarity of convolutional codes. The best recovered bit error rate occurs for bits that are either near the first bit to be encoded or the last bit to be encoded. Since the early bits are the best protected, the voicing mode bits are placed in this position.

Since the voicing mode bits are preferably always in the same position, and encoded in the same way, a single pre-decoder (501), as illustrated in FIG. 5, can be used to decode these bits. The resulting decoded mode information can be used to select (502) the appropriate post-decoder (503 through 506) in order to ensure that the appropriate mode-determined ordering and encoding algorithm is used to decode the protected voice coder data bits.

In digital transmission systems, some transmitted bits have a higher probability of decoded error than other bits, principally because of their position in the channel. For example, in TDM (time-division multiplex) systems, probability of decoded error of a particular bit may be related to the bit's proximity to a transmitted synchronization sequence.

Returning to FIG. 3, the output of the encoder (304) is applied to the input of an interleaver (305), as is the output of the Class II array (303). The interleaver (305) simply combines the encoded and unencoded bits so that the resultant data stream can be transmitted.

What is claimed is:

Claims

1. A method for providing error protection for voice coder information bits to be transmitted, the method comprising:

(a) providing input data bits comprising voice coder information bits;

(b) separating the input data bits, based upon voicing mode, into first and second arrays of bits;

(c) applying an error control coding technique to the bits of the first array to produce encoded bits;

(d) combining the encoded bits with the bits of the second array to form a bit stream; and

(e) transmitting the bit stream.

2. The method in accordance with claim 1, wherein the bits of the first array are arranged, dependent upon voicing mode, such that bits having greater perceptual significance in reconstructing speech are encoded such that these more perceptually significant bits have lower probability of decoded error.

3. The method in accordance with claim 1, wherein the step (c) of applying an error control coding technique comprises applying the bits of the first array as input to an encoder to provide encoded bits.

4. The method in accordance with claim 3, wherein the encoder comprises a convolutional encoder.

5. The method in accordance with claim 4, wherein the convolutional encoder encodes at least some bits at a first rate and at least some bits at a second rate.

6. The method in accordance with claim 5, wherein the first rate corresponds to a rate 1/3 convolutional code, and the second rate corresponds to a rate 1/2 convolutional code.

7. The method in accordance with claim 1, wherein the bits of the first array are arranged in order of importance.

8. The method in accordance with claim 7, wherein the order of importance in which the bits of the first array are arranged is determined by voicing mode.

9. The method in accordance with claim 1, wherein the step (e) of transmitting the bit stream includes storage and retrieval of the bit stream from a memory device.

10. Apparatus for providing error protection for voice coder information bits comprising:

means for providing input data bits comprising voice coder information bits;

means for separating the input data bits, based upon voicing mode, into first and second arrays of bits;

means for applying an error control coding technique to the bits of the first array to produce encoded bits;

means for combining the encoded bits with the bits of the second array to form a bit stream.

FIG.1

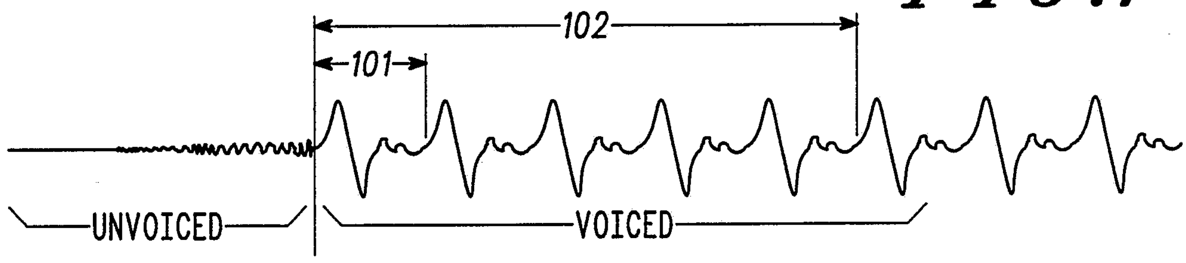


FIG.2A

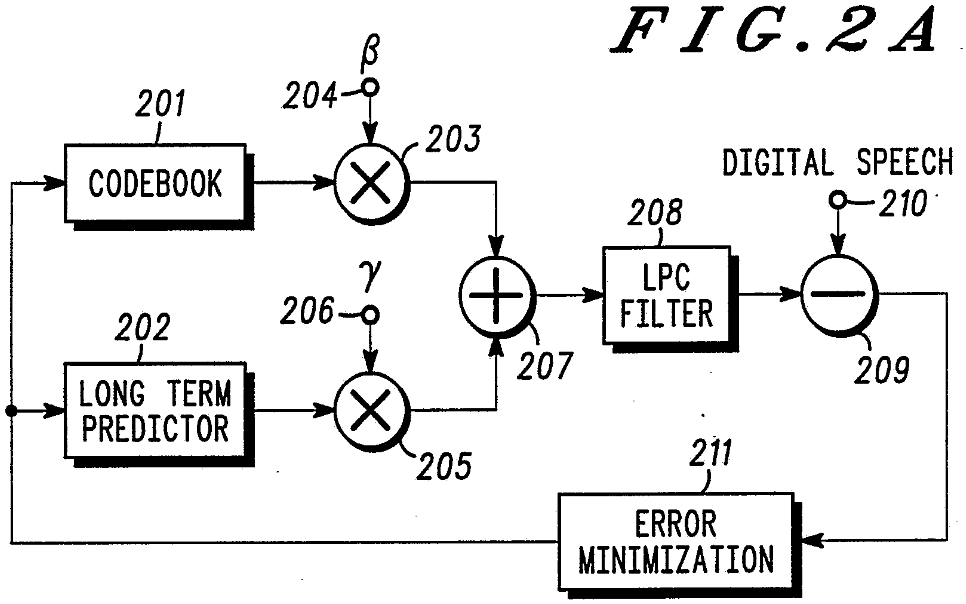


FIG.2B

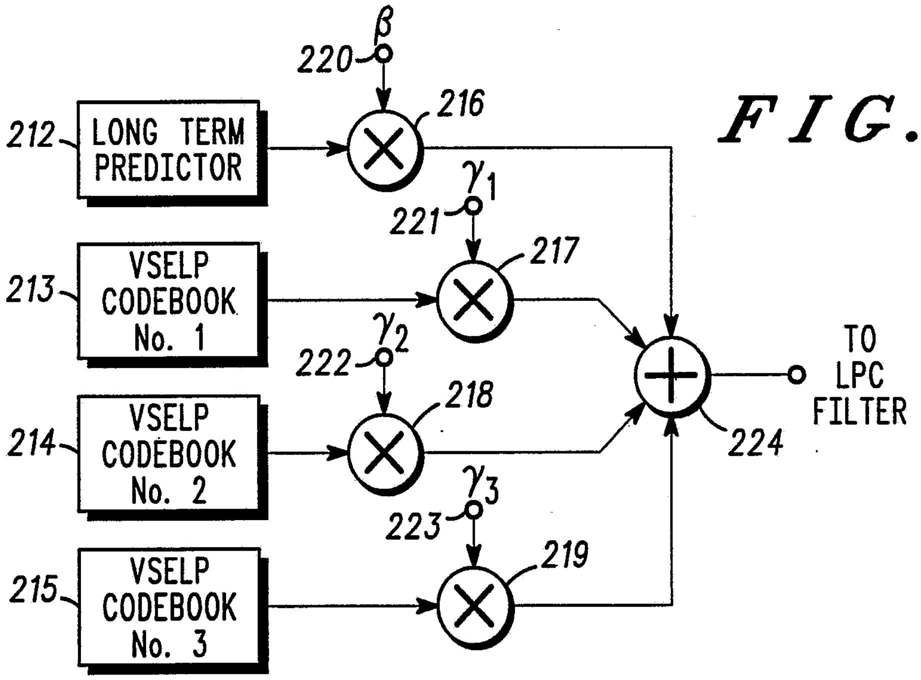


FIG.3

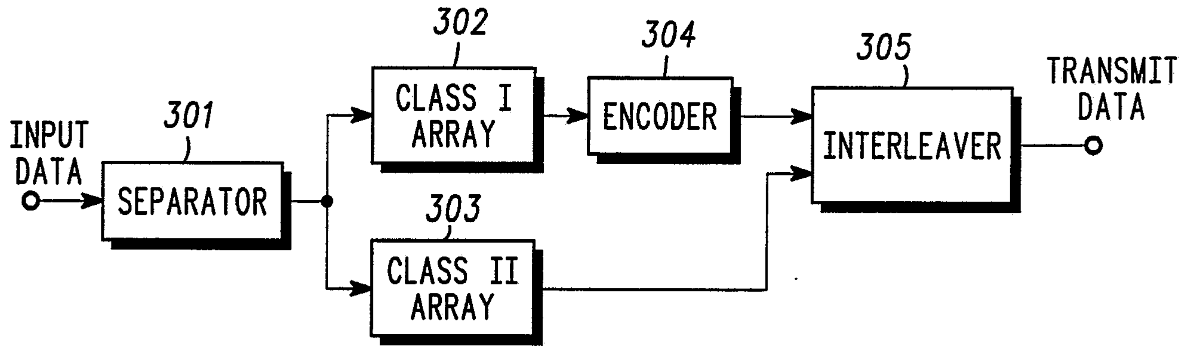


FIG.4

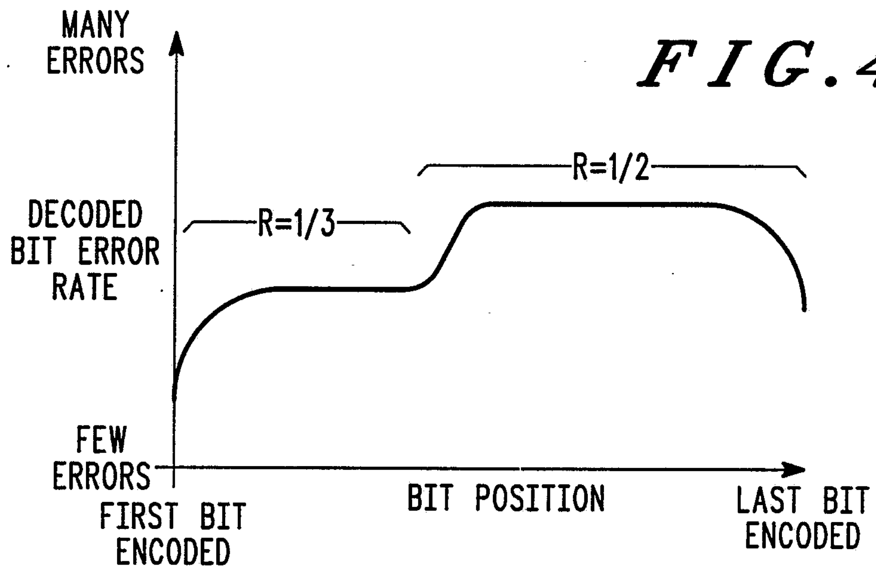
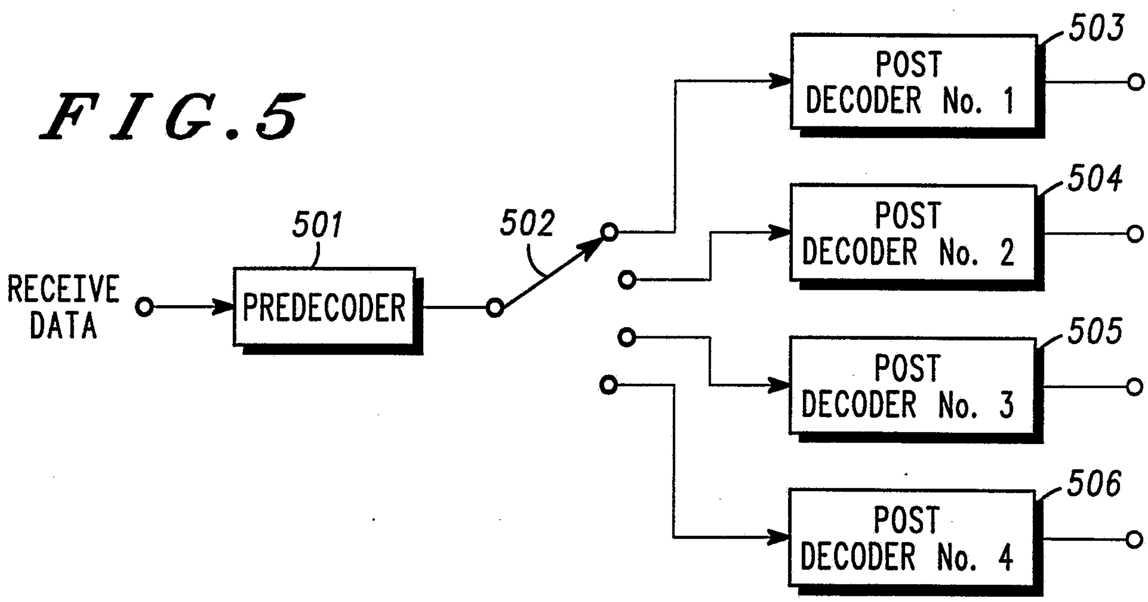


FIG.5



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US92/06352

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :G10L 5/00; G06F 11/00
US CL :U.S. 371/43
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 371/37.1; 381/31,46

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P Y	US,A, 5,097,507 (ZINSER ET AL.) 17 MARCH 1992, SEE ABSTRACT, COL.2, LINES 52-58.	1,2,3,7,8,10 4-6 & 9
Y	US,A, 4,802,171 (RASKY) 21 JANUARY 1989, SEE ABSTRACT, COLUMN 2, LINES 38-47, COLUMN 3, LINES 67+.	4-6
A	US,A, 4,831,624 (MCLAUGHLIN ET AL) 16 MAY 1989, SEE ENTIRE DOCUMENT.	1-10
A,P	US,A, 5,070,503 (SHIKAKURA) 03 DECEMBER 1991, SEE ENTIRE DOCUMENT.	1-10

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date		
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search 03 SEPTEMBER 1992	Date of mailing of the international search report 07 DEC 1992
Name and mailing address of the ISA/ Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. NOT APPLICABLE	Authorized officer CHARLES E. ATKINSON Telephone No. (703) 308-0856

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

APS

L1	6423	S 371/?/CCLS
L2	19	S L1 AND (VOICE OR SPEECH) (5A) (ERROR)
L3	15	S L2 AND (SIGNIFICANT OR IMPORTANT)
L4	2	S SPEECH COPER (5A) ERROR