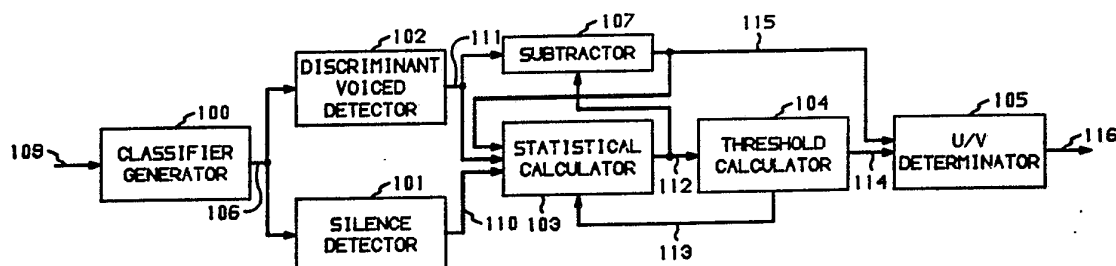




## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification<sup>4</sup> : <b>G10L 3/00</b></p>	<p><b>A1</b></p>	<p>(11) International Publication Number: <b>WO 88/ 07739</b> (43) International Publication Date: 6 October 1988 (06.10.88)</p>
<p>(21) International Application Number: PCT/US88/00031 (22) International Filing Date: 12 January 1988 (12.01.88) (31) Priority Application Number: 034,298 (32) Priority Date: 3 April 1987 (03.04.87) (33) Priority Country: US (71) Applicant: AMERICAN TELEPHONE &amp; TELEGRAPH COMPANY [US/US]; 550 Madison Avenue, New York, NY 10022 (US). (72) Inventor: THOMSON, David, Lynn ; 29W543 Country Ridge Drive, Apartment D, Warrenville, IL 60555 (US). (74) Agents: HIRSCH, A., E., Jr. et al.; Post Office Box 679, Holmdel, NJ 07733 (US).</p>		<p>(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent). <b>Published</b> <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p>

## (54) Title: AN ADAPTIVE THRESHOLD VOICED DETECTOR



## (57) Abstract

Apparatus for detecting a fundamental frequency in speech by statistically analyzing a discriminant variable generated by a discriminant voiced detector (102) so as to determine the presence of the fundamental frequency in a changing speech environment. A statistical calculator (103) is responsive to the discriminant variable to first calculate the average of all of the values of the discriminant variable over the present and past speech frames and then to determine the overall probability that any frame will be unvoiced. In addition, the calculator informs two values, one value represents the statistical average of discriminant values that an unvoiced frame's discriminant variable would have and the other value represents the statistical average of the discriminant values for voice frames. These latter calculations are performed utilizing not only the average discriminant value but also a weight value and a threshold value which are adaptively determined by a threshold calculator (104) from frame to frame. An unvoiced/voiced determinator (105) makes the unvoiced/voiced decision by utilizing the weight and threshold values.

**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.

<b>AT</b> Austria	<b>FR</b> France	<b>ML</b> Mali
<b>AU</b> Australia	<b>GA</b> Gabon	<b>MR</b> Mauritania
<b>BB</b> Barbados	<b>GB</b> United Kingdom	<b>MW</b> Malawi
<b>BE</b> Belgium	<b>HU</b> Hungary	<b>NL</b> Netherlands
<b>BG</b> Bulgaria	<b>IT</b> Italy	<b>NO</b> Norway
<b>BJ</b> Benin	<b>JP</b> Japan	<b>RO</b> Romania
<b>BR</b> Brazil	<b>KP</b> Democratic People's Republic of Korea	<b>SD</b> Sudan
<b>CF</b> Central African Republic	<b>KR</b> Republic of Korea	<b>SE</b> Sweden
<b>CG</b> Congo	<b>LI</b> Liechtenstein	<b>SN</b> Senegal
<b>CH</b> Switzerland	<b>LK</b> Sri Lanka	<b>SU</b> Soviet Union
<b>CM</b> Cameroon	<b>LU</b> Luxembourg	<b>TD</b> Chad
<b>DE</b> Germany, Federal Republic of	<b>MC</b> Monaco	<b>TG</b> Togo
<b>DK</b> Denmark	<b>MG</b> Madagascar	<b>US</b> United States of America
<b>FI</b> Finland		

- 1 -

## AN ADAPTIVE THRESHOLD VOICED DETECTOR

Technical Field

This invention relates to determining whether or not speech contains a fundamental frequency which is commonly referred to as the unvoiced/voiced decision. More particularly, the unvoiced/voiced decision is made by a two stage  
5    voiced detector with the final threshold values being adaptively calculated for the speech environment utilizing statistical techniques.

Background and Problem

In low bit rate voice coders, degradation of voice quality is often due to inaccurate voicing decisions. The difficulty in correctly making these voicing  
10    decisions lies in the fact that no single speech parameter or classifier can reliably distinguish voiced speech from unvoiced speech. In order to make the voice decision, it is known in the art to combine multiple speech classifiers in the form of a weighted sum. This method is commonly called discriminant analysis. Such a method is illustrated in D. P. Prezas, et al., "Fast and Accurate Pitch Detection  
15    Using Pattern Recognition and Adaptive Time-Domain Analysis," Proc. IEEE Int. Conf. Acoust., Speech and Signal Proc., Vol. 1, pp. 109-112, April 1986. As described in that article, a frame of speech is declared voice if a weighted sum of classifiers is greater than a specified threshold, and unvoiced otherwise. The weights and threshold are chosen to maximize performance on a training set of  
20    speech where the voicing of each frame is known.

A problem associated with the fixed weighted sum method is that it does not perform well when the speech environment changes. The reason is that the threshold is determined from the training set which is different from speech subject to background noise, non-linear distortion, and filtering.

25        One method for adapting the threshold value to changing speech environment is disclosed in the paper of H. Hassanein, et al., "Implementation of the Gold-Rabiner Pitch Detector in a Real Time Environment Using an Improved Voicing Detector," IEEE Transactions on Acoustic, Speech and Signal Processing, 1986, Tokyo, Vol. ASSP-33, No. 1, pp. 319-320. This paper discloses  
30    an empirical method which compares three different parameters against independent thresholds associated with these parameters and on the basis of each comparison either increments or decrements by one an adaptive threshold value. The three parameters utilized are energy of the signal, first reflection coefficient,

- 2 -

and zero-crossing count. For example, if the energy of the speech signal is less than a predefined energy level, the adaptive threshold is incremented. On the other hand, if the energy of the speech signal is greater than another predefined energy level, the adaptive threshold is decremented by one. After the adaptive threshold has been calculated, it is subtracted from an output of a elementary pitch detector. If the results of the subtraction yield a positive number, the speech frame is declared voice; otherwise, the speech frame is declared on voice. The problem with the disclosed method is that the parameters themselves are not used in the elementary pitch detector. Hence, the adjustment of the adaptive threshold is ad hoc and is not directly linked to the physical phenomena from which it is calculated. In addition, the threshold cannot adapt to rapidly changing speech environments.

#### Solution

The above described problem is solved and a technical advance is achieved by a voicing decision apparatus that adapts to a changing environment by utilizing adaptive statistical values to make the voicing decision. The statistical values are adapted to the changing environment by utilizing statistics based on an output of a voiced detector. The statistical parameters are calculated by the voiced detector generating a general value indicating the presence of a fundamental frequency in a speech frame in response to speech attributes of the frame. Second, the mean for unvoiced ones and voiced ones of speech frames is calculated in response to the generated value. The two means are then used to determine decision regions, and the determination of the presence of the fundamental frequency is done in response to the decision regions and the present speech frame.

Advantageously, in response to speech attributes of the present and past speech frames, the mean for unvoiced frames is calculated by calculating the probability that the present speech frame is unvoiced, calculating the overall probability that any frame will be unvoiced, and calculating the probability that the present speech frame is voiced. The mean of the unvoiced speech frames is then calculated in response to the probability that the present speech frame is unvoiced and the overall probability. In addition, the mean of the voiced speech frame is calculated in response to the probability that the present speech frame is voiced and the overall probability. Advantageously, the calculations of probabilities are performed utilizing a maximum likelihood statistical operation.

- 3 -

Advantageously, the generation of the general value is performed utilizing a discriminant analysis procedure, and the speech attributes are speech classifiers.

Advantageously, the decision regions are defined by the mean of the unvoiced and voiced speech frames and a weight and threshold value generated in response to the general values of past and present frames and the means of the voiced and unvoiced frames.

The method for detecting the presence of a fundamental frequency in speech frames comprises the steps of: generating a general value in response to a set of classifiers defining speech attributes of a present speech frame to indicate the presence of the fundamental frequency, calculating a set of statistical parameters in response to the general value, and determining the presence of the fundamental frequency in response to the general value and the calculated set of statistical parameters. The step of generating the general value is performed utilizing a discriminant analysis procedure. Further, the step of determining the fundamental frequency comprises the step of calculating a weight and a threshold value in response to the set of parameters.

#### Brief Description of the Drawing

FIG. 1 illustrates, in block diagram form, the present invention; and FIGS. 2 and 3 illustrate, in greater detail, certain functions performed by the voiced detection apparatus of FIG. 1.

#### Detailed Description

FIG. 1 illustrates an apparatus for performing the unvoiced/voiced decision operation by first utilizing a discriminant voiced detector to process voice classifiers in order to generate a discriminant variable or general variable. The latter variable is statistically analyzed to make the voicing decision. The statistical analysis adapts the threshold utilized in making the unvoiced/voiced decision so as to give reliable performance in a variety of voice environments.

Consider now the overall operation of the apparatus illustrated in FIG. 1. Classifier generator 100 is responsive to each frame of voice to generate classifiers which advantageously may be the log of the speech energy, the log of the LPC gain, the log area ratio of the first reflection coefficient, and the squared correlation coefficient of two speech segments one frame long which are offset by one pitch period. The calculation of these classifiers involves digitally sampling analog speech, forming frames of the digital samples, and processing those frames

- 4 -

and is well known in the art. Generator 100 transmits the classifiers to silence detector 101 and discriminant voiced detector 102 via path 106. Discriminant voiced detector 102 is responsive to the classifiers received via path 106 to calculate the discriminant value,  $x$ . Detector 102 performs that calculation by solving the equation:  $x = c'y+d$ . Advantageously, "c" is a vector comprising the weights, "y" is a vector comprising the classifiers, and "d" is a scalar representing a threshold value. Advantageously, the components of vector c are initialized as follows: component corresponding to log of the speech energy equals 0.3918606, component corresponding to log of the LPC gain equals -0.0520902, component corresponding to log area ratio of the first reflection coefficient equals 0.5637082, and component corresponding to squared correlation coefficient equals 1.361249; and d initially equals -8.36454. After calculating the value of the discriminant variable  $x$ , the detector 102 transmits this value via path 111 to statistical calculator 103 and subtracter 107.

Silence detector 101 is responsive to the classifiers transmitted via path 106 to determine whether speech is actually present on the data being received on path 109 by classifier generator 100. The indication of the presence of speech is transmitted via path 110 to statistical calculator 103 by silence detector 101.

For each frame of speech, detector 102 generates and transmits the discriminant value  $x$  via path 111. Statistical calculator 103 maintains an average of the discriminant values received via path 111 by averaging in the discriminant value for the present, non-silence frame with the discriminant values for previous non-silence frames. Statistical calculator 103 is also responsive to the signal received via path 110 to calculate the overall probability that any frame is unvoiced and the probability that any frame is voiced. In addition, statistical calculator 103 calculates the statistical value that the discriminant value for the present frame would have if the frame was unvoiced and the statistical value that the discriminant value for the present frame would have if the frame was voiced. Advantageously, that statistical value may be the mean. The calculations performed by calculator 103 are not only based on the present frame but on previous frames as well. Statistical calculator 103 performs these calculations not only on the basis of the discriminant value received for the present frame via path 106 and the average of the classifiers but also on the basis of a weight and a threshold value defining whether a frame is unvoiced or voiced received via

- 5 -

path 113 from threshold calculator 104.

Calculator 104 is responsive to the probabilities and statistical values of the classifiers for the present frame as generated by calculator 103 and received via path 112 to recalculate the values used as weight value a, and threshold value  
 5 b for the present frame. Then, these new values of a and b are transmitted back to statistical calculator 103 via path 113.

Calculator 104 transmits the weight, threshold, and statistical values via path 114 to U/V determinator 105. The latter detector is responsive to the information transmitted via paths 114 and 115 to determine whether or not the  
 10 frame is unvoiced or voiced and to transmit this decision via path 116.

Consider now in greater detail the operations of blocks 103, 104, 105, and 107 illustrated in FIG. 1. Statistical calculator 103 implements an improved EM algorithm similar to that suggested in the article by N. E. Day entitled  
 "Estimating the Components of a Mixture of Normal Distributions", *Biometrika*,  
 15 Vol. 56, No. 3, pp. 463-474, 1969. Utilizing the concept of a decaying average, calculator 103 calculates the average for the discriminant values for the present and previous frames by calculating following equations 1, 2, and 3:

$$n = n+1 \text{ if } n < 2000 \quad (1)$$

$$z = 1/n \quad (2)$$

20 
$$X_n = (1-z) X_{n-1} + zx_n \quad (3)$$

$x_n$  is the discriminant value for the present frame and is received from detector 102 via path 111, and n is the number of frames that have been processed up to 2000. z represents the decaying average coefficient, and  $X_n$  represents the

average of the discriminant values for the present and past frames. Statistical calculator 103 is responsive to receipt of the  $z$ ,  $x_n$  and  $X_n$  values to calculate the variance value,  $T$ , by first calculating the second moment of  $x_n$ ,  $Q_n$ , as follows:

$$Q_n = (1-z)Q_{n-1} + zx_n^2. \quad (4)$$

5 After  $Q_n$  has been calculated,  $T$  is calculated as follows:

$$T = Q_n - x_n^2. \quad (5)$$

The mean is subtracted from the discriminant value of the present frame as follows:

$$x_n = x_n - X_n \quad (6)$$

10 Next, calculator 103 determines the probability that the frame represented by the present value  $x_n$  is unvoiced by solving equation 7 shown below:

$$P(u|x_n) = \frac{1}{1 + \exp(ax_n + b)}. \quad (7)$$

After solving equation 7, calculator 103 determines the probability that the discriminant value represents a voiced frame by solving the following:

$$15 \quad P(v|x_n) = 1 - P(u|x_n). \quad (8)$$

- 7 -

Next, calculator 103 determines the overall probability that any frame will be unvoiced by solving equation 9 for  $p_n$ :

$$p_n = (1-z) p_{n-1} + z P(u|x_n) . \quad (9)$$

After determining the probability that a frame will be unvoiced, calculator 103 determines two values,  $u$  and  $v$ , which give the mean values of discriminant value for both unvoiced and voiced type frames. Value  $u$ , statistical average unvoiced value, contains the mean discriminant value if a frame is unvoiced, and value  $v$ , statistical average voiced value, gives the mean discriminant value if a frame is voiced. Value  $u$  for the present frame is solved by calculating equation 10, and value  $v$  is determined for the present frame by calculating equation 11 as follows:

$$u_n = (1-z) u_{n-1} + z x_n P(u|x_n)/p_n - zx_n \quad (10)$$

$$v_n = (1-z) v_{n-1} + z x_n P(v|x_n)/(1-p_n) - zx_n \quad (11)$$

Calculator 103 now communicates the  $u$ ,  $v$ , and  $T$  values, and probability  $p_n$  to threshold calculator 104 via path 112.

Calculator 104 is responsive to this information to calculate new values for  $a$  and  $b$ . These new values are then transmitted back to statistical calculator 103 via path 113. This allows rapid adaptations to changing environments. If  $n$  is greater than advantageously 99, values  $a$  and  $b$  are calculated as follows. Value  $a$  is determined by solving the following equation:

$$a = T^{-1} \frac{(v_n - u_n)}{1 - p_n(1 - p_n) T^{-1} (u_n - v_n)^2} . \quad (12)$$

- 8 -

Value  $b$  is determined by solving the following equation:

$$b = -\frac{1}{2} a(u_n + v_n) + \log[(1-p_n)/p_n] . \quad (13)$$

After calculating equations 12 and 13, calculator 104 transmits values  $a$ ,  $u$ , and  $v$  to block 105 via path 114.

- 5 Determinator 105 is responsive to this transmitted information to decide whether the present frame is voiced or unvoiced. If the value  $a$  is positive, then, a frame is declared voiced if the following equation is true:

$$ax_n - a(u_n + v_n)/2 > 0 ; \quad (14)$$

- 10 or if the value  $a$  is negative, then, a frame is declared voiced if the following equation is true:

$$ax_n - a(u_n + v_n)/2 < 0 . \quad (15)$$

Equation 14 can also be expressed as:

$$ax_n + b - \log[(1-p_n)/p_n] > 0 .$$

Equation 15 can also be expressed as:

15  $ax_n + b - \log[(1-p_n)/p_n] < 0 .$

If the previous conditions are not met, determinator 105 declares the frame unvoiced.

In flow chart form, FIGS. 2 and 3 illustrate, in greater detail, the operations performed by the apparatus of FIG. 1. Block 200 implements  
5 block 101 of FIG. 1. Blocks 202 through 218 implement statistical calculator 103. Block 222 implements threshold calculator 104, and blocks 226 through 239 implement block 105 of FIG. 1. Subtractor 107 is implemented by both block 208 and block 224. Block 202 calculates the value which represents the average of the discriminant value for the present frame and all previous  
10 frames. Block 200 determines whether speech is present in the present frame; and if speech is not present in the present frame, the mean for the discriminant value is subtracted from the present discriminant value by block 224 before control is transferred to decision block 226.

However, if speech is present in the present frame, then the statistical  
15 and weight calculations are performed by blocks 202 through 222. First, the average value is found in block 202. Second, the second moment value is calculated in block 206. The latter value along with the mean value  $X$  for the present and past frames is then utilized to calculate the variance,  $T$ , also in block 206. The mean  $X$  is then subtracted from the discriminant value  $x_n$  in  
20 block 208.

Block 210 calculates the probability that the present frame is unvoiced by utilizing the present weight value  $a$ , the present threshold value  $b$ , and the discriminant value for the present frame,  $x_n$ . After calculating the probability that the present frame is unvoiced, the probability that the present frame is voiced is  
25 calculated by block 212. Then, the overall probability,  $p_n$ , that any frame will be unvoiced is calculated by block 214.

Blocks 216 and 218 calculate two values:  $u$  and  $v$ . The value  $u$  represents the statistical average value that the discriminant value would have if the frame were unvoiced. Whereas, value  $v$  represents the statistical average value  
30 that the discriminant value would have if the frame were voiced. The actual discriminant values for the present and previous frames are clustered around either value  $u$  or value  $v$ . The discriminant values for the previous and present frames are clustered around value  $u$  if these frames had been found to be unvoiced; otherwise, the previous values are clustered around value  $v$ . Block 222 then  
35 calculates a new weight value  $a$  and a new threshold value  $b$ . The values  $a$  and  $b$

are used in the next sequential frame by the preceding blocks in FIG. 2.

Blocks 226 through 239 implement U/V determinator 105 of FIG. 1. Block 226 determines whether the value  $a$  for the present frame is greater than zero. If this condition is true, then decision block 228 is executed. The latter  
5 decision block determines whether the test for voiced or unvoiced is met. If the frame is found to be voiced in decision block 228, then the frame is so marked as voiced by block 230 otherwise the frame is marked as unvoiced by block 232. If  
the value  $a$  is less than zero for the present frame, blocks 234 through 238 are executed and function in a similar manner to blocks 228 through 232.

10 It is to be understood that the afore-described embodiment is merely illustrative of the principles of the invention and that other arrangements may be devised by those skilled in the art without departing from the spirit and the scope of the invention.

Claims

1. An apparatus for detecting the presence of a fundamental frequency in frames of speech, comprising:  
means responsive to a set of classifiers defining speech attributes of one said frames of speech for generating a general value indicating said presence of said fundamental frequency;  
means responsive to said general value for calculating a set of statistical parameters; and  
means responsive to said general value and the calculated set of statistical parameters for determining said presence of said fundamental frequency.
2. The apparatus of claim 1 wherein said generating means comprises means for performing a discriminant analysis to generate said general value.
3. The apparatus of claim 1 wherein said determining means comprises means for calculating a threshold value in response to said set of said parameters;  
means for calculating a weight value in response to said set of said parameters; and  
means for communicating said weight and threshold values to said means for calculating said set of parameters to be used for calculating another set of parameters for another one of said frames of speech.
4. The apparatus of claim 3 wherein said mean for calculating said set of parameters further responsive to the communicated weight and threshold values and another general value of said other one of said frames for calculating another set of statistical parameters.
5. The apparatus of claim 4 wherein said means for calculating said set of parameters further comprises means for calculating the average of said general values over said present and previous ones of said speech frames; and  
means responsive to said average of said general values for said present and previous ones of said speech frames and said communicated weight and threshold values and said other general value for determining said other set of statistical parameters.
6. An apparatus for detecting the presence of a fundamental frequency in frames of speech, comprising:  
means responsive to a set of classifiers defining speech attributes of one of said frames of speech for generating a general value indicating said presence of said fundamental frequency;

means for estimating a mean for unvoiced ones of said frames in response to said general value;

means for estimating a mean for voiced ones of said frames in response to said general value;

5 means responsive to said mean for unvoiced ones of said frames and said mean of voiced ones of said frames for determining decision regions; and

means for making the determination of said presence of said fundamental frequency in response to said decision regions and a present one of said frames.

10 7. The apparatus of claim 6 wherein said means for determining comprises means for calculating the variance of said general values over said present and previous ones of said speech frames;

means for generating said decision regions in response to said mean for unvoiced ones of said frames and said mean of voiced ones of said frames and

15 said variance;

said means for estimating said mean of said unvoiced frames comprises means responsive to present and past ones of said frames for calculating the probability that said present one of said frames is unvoiced;

20 means responsive to said present and past ones of said frames and said probability that said present one of said frames is unvoiced for calculating the overall probability that any frame will be unvoiced;

said means for estimating said mean of said voiced frames further comprises means for calculating the probability that said present one of said frames is voiced;

25 means responsive to said probability that said present one of said frames is unvoiced and said overall probability and said variance for calculating the said mean of said unvoiced ones of said frames; and

means responsive to said probability that said present one of said frames is voiced and said overall probability and said variance for calculating the said mean of said voiced ones of said frames.

30 8. The apparatus of claim 7 wherein said means for calculating said probability that said present one of said frames is unvoiced performs a maximum likelihood statistical operation.

35 9. The apparatus of claim 8 wherein said means for calculating said probability that said present one of said frames is unvoiced further responsive to a

- 13 -

weight and threshold value to perform said maximum likelihood statistical operation.

10. A method for detecting the presence of a fundamental frequency in frames of speech comprising the steps of:

5 generating a general value in response to a set of classifiers defining speech attributes of one of said frames of speech to indicate said presence of said fundamental frequency;

calculating a set of statistical parameters in response to said general value; and

10 determining said presence of said fundamental frequency in response to said general value and the calculated said set of statistical parameters.

11. The method of claim 10 wherein said step of generating comprises the step of performing a discriminant analysis to generate said general value.

15 12. The method of claim 11 wherein said step of determining comprises the steps of calculating a threshold value in response to said set of said parameters;

calculating a weight value in response to said set of said parameters; and

20 calculating said set of parameters to be used for calculating another set of parameters for another one of said frames of speech.

25 13. The method of claim 12 wherein said step of calculating said set of parameters further responsive to the communicated weight and threshold values and another general value of said other one of said frames for calculating another set of statistical parameters.

14. The method of claim 13 wherein said step of calculating said set of parameters further comprises the steps of calculating the average of said general values over said present and previous ones of said speech frames; and

30 determining said other set of statistical parameters in response to said average of said general values for said present and previous ones of said speech frames and said communicated weight and threshold values and said other general values.

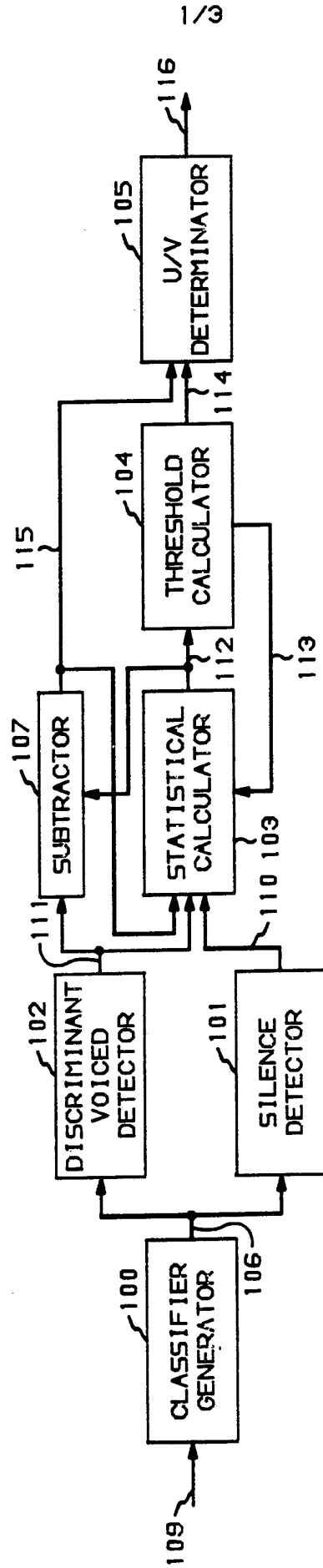


FIG. 1

2/3

FIG. 2

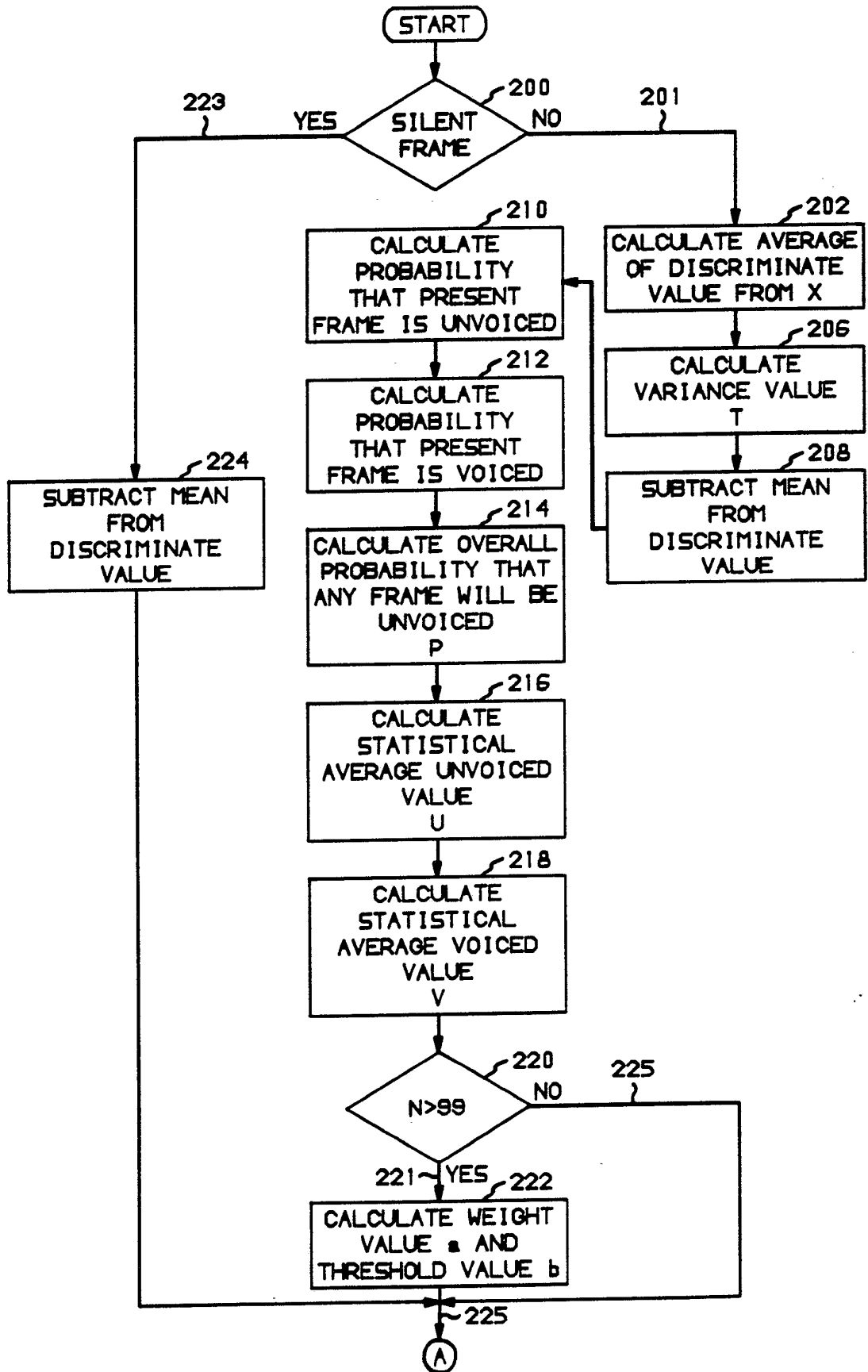
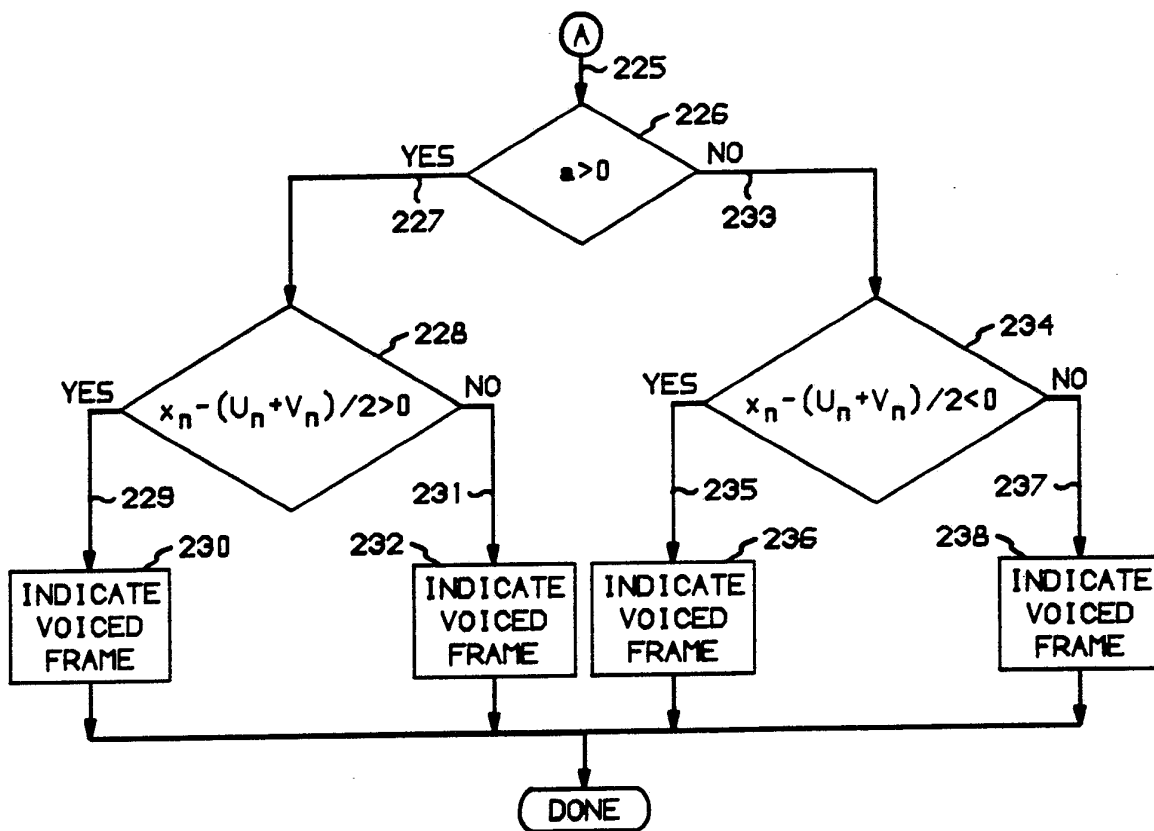



FIG. 3



# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 88/00031

<b>I. CLASSIFICATION OF SUBJECT MATTER</b> (if several classification symbols apply, indicate all) <sup>6</sup>		
According to International Patent Classification (IPC) or to both National Classification and IPC		
IPC <sup>4</sup> : G 10 L 3/00		
<b>II. FIELDS SEARCHED</b>		
Minimum Documentation Searched <sup>7</sup>		
Classification System	Classification Symbols	
IPC <sup>4</sup>	G 10 L 3/00	
Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched <sup>8</sup>		
<b>III. DOCUMENTS CONSIDERED TO BE RELEVANT <sup>9</sup></b>		
Category <sup>10</sup>	Citation of Document, <sup>11</sup> with indication, where appropriate, of the relevant passages <sup>12</sup>	Relevant to Claim No. <sup>13</sup>
P, X	Proceedings: ICASSP 87, 1987 International Conference on Acoustics, Speech, and Signal Processing, 6-9 April 1987, Dallas, Texas, volume 1 of 4, IEEE, (New York, US), D.L. Thomson: "A multivariate voicing decision rule adapts to noise, distortion, and spectral shaping", pages 197-200 see abstract	1, 2, 6-11
A	IEEE Transactions on Acoustics, Speech, and Signal Processing, volume ASSP-31, no. 3, June 1983, IEEE, (New York, US), P. De Souza: "A statistical approach to the design of an adaptive self-normalizing silence detector", pages 678-684 see paragraph III: "Training and adaption"	1-5, 10-14
-----		
<p><sup>10</sup> Special categories of cited documents: <sup>10</sup></p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"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)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"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</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step</p> <p>"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.</p> <p>"Z" document member of the same patent family</p>		
<b>IV. CERTIFICATION</b>		
Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report	
4th August 1988	30 AUG 1988	
International Searching Authority	Signature of Authorized Officer	
EUROPEAN PATENT OFFICE	 P.C.G. VAN DER PUTTEN	