Description

The present invention relates to apparatus for speaker identification ie for recognising or verifying the identity of a person whose speech forms the input to the apparatus.

In their paper given at ICASSP 82, (IEE International conference of acoustics, speech and signal processing, Paris, May 3 - 5, 1982, vol 3, pp 1645 - 1648), H. Ney et al describe a technique for increasing the ability of a textdependent speaker recognition system to discriminate between speaker classes; the technique to be performed in conjunction with non linear time alignment between a reference pattern and the test pattern. Unlike what they refer to as the standard approach, where the training of the recognition system merely consists of storing and averaging or selecting the time normalised training patterns separately for each class, in the Ney system the training phase is extended in that a weight is determined for each individual feature component of the complete reference pattern (for each class) according to the ability of the feature to distinguish between speaker classes. The weights depend on the time axis as well as on the frequency axis. The overall distance computed after non-linear time alignment between a reference pattern for a class and a test pattern thus becomes a function of the given set of weights of the reference class considered.

In IEEE ASSP, October 1986, pp 4 - 17, O'Shaughnessy provides a tutorial article on speaker recognition (ASR), dealing inter alia with the differences between speaker recognition, speaker verification and speech recognition. Again much stress is placed on the need to perform time normalization procedures, such as Dynamic Timewarping, if one is to determine distances between test and reference templates. The paper also deals with methods of selecting a useful subset of acoustic features which can aid recognition. In particular it is suggested that FO, timing cues and spectral cues (such as the first three formants) are the most promising candidates for use in speaker recognition. Attention is directed to the first complete paragraph in the right hand column on page 10. O'Shaughnessy does not disclose that digital words address the store or suggest dispensing with the pattern comparison operation, nor even that such an approach would be feasible.

In US 4,100,370 there is described a voice verification system. A selection process is used to find parameters which may facilitate improved discrimination between speakers. Verification is based on a conventional pattern matching process.

According to the present invention there is provided an apparatus for speaker identification comprising:

means for deriving from each of successive sample periods of speech input thereto a digital word representing the frequency spectrum thereof;

a store for storing codewords indicating for each digital word the relative probability of the frequency spectrum corresponding to that work occuring in the speech of one or more known speakers which the apparatus has been trained to identify, the digital word output of the deriving means being connected to address inputs of the store; and

control means responsive, during input of speech by a speaker to be identified, to codewords read out from the store to output, in accordance with a predetermined criterion related to the codewords, data indicating which of the speakers is deemed to have been identified.

In a preferred arrangement, the apparatus is arranged in operation during a training sequence, for each sample period:

a) If the number of codewords already stored at the addressed location in the store does not exceed a predetermined number to store at that location a codeword identifying the speaker whose speech has given rise to the digital word representing that address,

b) If the number of codewords already stored at the addressed location in the store exceeds the predetermined number, to store at that location a codeword reserved for indication of multiple users.

The predetermined number can be zero.

The present invention also provides methods as claimed in claims 12 and 13, respectively.

Thus, the presence of a known speaker's code at any given location indicates that the probability of the corresponding frequency spectrum occurring in the speech of that speaker is high relative to its probability of occurring in the speech of other speakers. The presence of the multi-user code, on the other hand, indicates that the probability of occurrence is not particularly high for any one speaker rather than another.

In practice, prior to initial training, the store will require to be cleared by writing to it a code not used for any other purpose. The presence after training is complete of such a 'not used' code in any location indicates that the probability of the corresponding spectrum occuring in the speech of any of the known users is low.

In a preferred arrangement the control means comprises a respective counter means for each known speaker and is arranged in operation to respond to each codeword read out from the store by incrementing the counter corresponding to the known speaker (if any) indicated by that codeword as having a relatively high probability of being associated with the frequency spectrum corresponding to the address from which the codeword

was read, and the predetermined criterion is that the speaker who shall be deemed to be recognised shall be the one whose counter has the highest count.

Some embodiments of the insertion will now be described, by way of example, with reference to the accompanying drawings, in which:

Figure 1 is a block diagram of one form of apparatus according to the invention;

Figure 2 is a more detailed block diagram of the apparatus of Figure 1; and

Figure 3 illustrates alternatives for the coders shown in Figure 2.

In Figure 1, speech signals from a user are input, for example via a microphone M, to a filter bank F to analyse the frequency spectrum of the input speech. Filter outputs representing respective frequency bands are rectified and smoothed and then sampled (S) and coded (C) to together provide the address input to a read-write memory S. For each sample the contents of the address thus accessed may in general contain one of three types of label:

a) a 'not used' label - eg all zeros

b) a 'multi-user' label - eg all ones

c) a 'valid user' label - eg anything other than all zeros or all ones.

Before being used for speaker identification, the apparatus must undergo a training sequence in which each authorised user speaks into the microphone. The object is to record in the store which spectral combinations are peculiar to certain users, and which are not. Thus, in the training mode, the following action is taken for each sample.

i) if a 'not used' label is found in the store then the code of the particular user is written into that store location

ii) if a 'valid user' label is found and is the same as the current user's code, no action is taken

iii) if a 'valid user' label is found and is different from the code of the current user, then a multi-user label is written into that store location

iv) if a 'multi-user' label is found, no action is taken.

It will be seen that each store location corresponds to a particular spectral combination. After the training sequence has been completed for a number of users (the 'valid users') each store location contains a label indicating whether the corresponding combination is particular to one authorised user (in which case the label is the user code) or is common to two or more such users, or has been used by none of the authorised users.

Once trained, the apparatus can be used in an identification mode A user to be identified speaks into the microphone M. If his utterance contains a significant number of spectral combinations known to be unique to one valid user, there is a high probability that the unknown user is that person. Thus the labels found in the store are acted upon as follows:

v) if a 'not used' label is found in the store it is ignored

vi) if a 'multi-user' label is found it is also ignored

vii) if a 'valid user' label is found, a counter, set aside for that user, is incremented.

At the end of the utterance a count will be held for each of the valid users. The highest count should then indicate which of the valid users has been speaking. Since there will always be at least one highest count, this does not serve to distinguish between valid users on the one hand and unauthorised users on the other. If the object of the recogniser is simply to verify the claimed identity of a user input via a speech recogniser or other means, this may suffice, but preferably the criterion for recognition should be that the highest count should be significantly greater than any other count. Additionally, or instead, it may be required that the highest count exceed a threshold value.

In Figure 1, it is assumed that the training and identification mode functions are performed by a control unit C. A name codec NC is also shown for translating valid user labels into user names.

Figure 2 shows a practical embodiment of the invention. The microphone M is again shown, followed by a speech amplifier A. The filter bank F comprises sixteen bandpass filters F1 to F16 covering the frequency band from 250Hz to 5000Hz. It is preferred that the higher frequency bands are wider than the lower bands - for example in accordance with critical band theory. Rectification and smoothing (with a time constant to suit the following sampler) is illustrated schematically by rectifiers R1-R16.

Sixteen sample and hold circuits S1-S16 are shown, driven by sampling pulses $\phi 1$. It is estimated that a sampling rate of the order of 50 to 100 samples per second will provide satisfactory results, though a higher resolution may be obtained by employing a higher rate - eg up to 1000 samples/sec.

Coders C1 to C16 receive the sampled outputs and convert these into a 16-bit word for application to the address inputs of the store S of 64 Kbyte capacity. The coders may simply be comparators CP as illustrated in figure 3a which compare the samples with a fixed threshold voltage Vt and produce a '1' or a '0' according to whether the sample is above or below the threshold. However it may be preferred to make the coded outputs independent of the input signal level (which will be dependent on microphone position and amplifier gain) by employing differential coding.

Thus (figure 3b) a sample may be compared by a comparator CP with the previous sample delayed in a latch L clocked by pulses $\phi 2$ delayed with respect to the sampling pulses $\phi 1$. Alternatively a sample may be compared with the signal level in an adjacent band (or indeed a non-adjacent band) a '1' being output if (for example) the sample value exceeds that in the frequency band below and a '0' otherwise. The lowest frequency band (F16) can be compared with the highest (F1). If desired both time differential coding and band differential coding might be used - either in parallel (though with a smaller number of bands than that shown since 32 address lines would - at the present time - imply an impractically large memory) or in series, eg by interposing the circuitry of figure 3b in the outputs of figure 3c (or vice versa).

It should be noted that when using a fixed threshold coding or compared band coding, the sixteen sample and hold circuits are not required. This is because the coded digital word may be continuously presented to the address bus of the RAM, when clocking the RAM will provide the sampling means.

In the training mode, the code for the speaker is entered into an input code latch CL. As the store is addressed, by successive coded samples, the labels output at the (8-bit) data output of the store are compared by a comparator CF1 with the speaker's code. If the two are the same, no action is taken (step (ii) above). If they are different, an AND gate A (enabled by an external signal TR only during training) is enabled to pass to the memory S write pulses $\phi 2$ (delayed sufficiently from $\phi 1$ to allow for the memory read access time, and delays in the coder C and comparator CF1). If the label output from the store is zero, the code from the latch CL passes unhindered via OR gates 01-08 and is written into the memory S (step (i) above). If the label is other than all zeros, this is recognised by an OR gate 09 which, via OR gates 01-08, forces the store input to all ones (FF hex) prior to the write operation, thereby writing (or re-writing) a multi-user label to the store as required by step (iii) (or step (iv)).

For the identification mode, counters Z1 to Zn are provided, one for each valid user. An 8-bit store could permit up to 254 users, but naturally counters need only be provided for the maximum number of valid users for which the system is expected to be trained. For each sample, decoder D receives the resulting store output data and decodes it to n enable lines, one for each counter. Dummy outputs 'O' and 'FF' illustrate steps (v) and (vi) above - ie no action is taken. When (step (vii)) a 'valid user' code appears, the corresponding counter is enabled and incremented by pulse $\phi 2$ mentioned previously.

The counting process continues as long as an identification enabling signal IE is present. This might be generated in response to a signal detector at the amplifier A and have a duration sufficient to permit processing of sufficient samples to achieve reasonable recognition accuracy. From $^1/_2$ to 3 seconds duration might be typical, with perhaps an 8-bit counter length (ie a maximum count of 255).

Once counting is complete, the n 8-bit counter outputs are compared in a comparator CF2. This serves to identify the counter with the highest count and to output the corresponding user code to an output latch OL, though, as discussed above it may serve also to check other criteria before outputting a valid code.

It may be noted that it is not necessary to the operation of the apparatus that the words uttered by speakers to be recognised be the same as those uttered by other speakers; or indeed that they be words previously uttered by that particular speaker during the training mode, provided that the words uttered during the training mode encompass a sufficiently comprehensive range of sounds.

The above described embodiment is only one possible implementation, and various modifications are possible; for example:

- rather than invoking the multi-user code when two or more speakers produce the same "sample" during training, this limit may be raised at the expense of a larger memory by allowing the recordal in the store of two (or N) 'valid user' codes against any given store address, the multi-user code being substituted only when three (or N+1) speakers have produced that sample. A retrieval during the identification mode of two valid user codes from one address would then cause both of the corresponding counters to be incremented.

If this option is used, weighting might also be introduced by, for example, incrementing the appropriate counter twice when only a single code appears.

- It has been assumed that no action is taken if a "not used" code is retrieved during the identification mode. Since the appearance of this code implies a reduced probability that the user speaking is a valid user, the occurrences of this code might be counted and the result used to weight the recognition process, for example by increasing the threshold which the highest count must meet before recognition is deemed to be complete.
- Although functions such as coding, decoding and comparison, have been illustrated in figure 2 as performed by dedicated hardware devices, these functions can, if desired, be

performed by a suitably programmed microprocessor or other data processing apparatus. Equally, digital filtering could be employed.

Claims

1. An apparatus for speaker identification comprising:
means for deriving from each of successive sample periods of speech input thereto a digital word representing the frequency spectrum thereof;
a store for storing codewords indicating for each digital word the relative probability of the frequency spectrum corresponding to that word occurring in the speech of one or more known speakers which the apparatus has been trained to identify, the digital word output of the deriving means being connected to the address inputs of the store; and
control means responsive, during input of speech by a speaker to be identified, to codewords read out from the store, to output, in accordance with a predetermined criterion related to the codewords, data indicating which of the known speakers is deemed to have been identified.

2. An apparatus according to claim 1, arranged in operation during a training sequence, for each sample period,
a) If the number of codewords already stored at the addressed location in the store does not exceed a predetermined number to store at that location a codeword identifying the speaker whose speech has given rise to the digital word representing that address,
b) If the number of codewords already stored at the addressed location in the store exceeds the predetermined number, to store at that location a codeword reserved for indication of multiple users.

3. An apparatus according to claim 2 in which the predetermined number is zero.

4. An apparatus according to claim 1, 2 or 3 in which the control means comprises a respective counter means for each known speaker and is arranged in operation to respond to each codeword read out from the store by incrementing the counter corresponding to the known speaker (if any) indicated by that codeword as having a relatively high probability of being associated with the frequency spectrum corresponding to the address from which the codeword was read, and in which the predetermined criterion is that the speaker who shall be deemed to be recognised shall be the one whose counter has the highest count.

5. An apparatus according to claim 4 in which the predetermined criterion includes the requirement that identification shall not be deemed to have occurred unless the said highest count exceeds the counts of the other counters by a set margin.

6. An apparatus according to claim 4 or 5 in which predetermined criterion includes the requirement that identification shall not be deemed to have occurred unless the said highest count exceeds a threshold value.

7. An apparatus according to claim 6 in which the control means contains a further counter, is responsive to the occurence of a codeword indicating a low probability of any known speaker to increment the further counter, and is responsive to the content of the further counter to adjust the said threshold value.

8. An apparatus according to any one of the preceding claims in which the deriving means comprises filter means for dividing the frequency spectrum of the speech input into a plurality of frequency band outputs, means for measuring the average amplitude of, and sampling those outputs, and coding means for obtaining the digital word from the resulting samples.

9. An apparatus according to claim 8 in which the coding means is arranged for each band to compare each sample with a threshold value and produce a '1' or '0' signal in dependence thereon.

10. An apparatus according to claim 8 in which the coding means is arranged for each band to compare each sample with the previous sample obtained for that band and produce a '1' or '0' signal in dependence thereon.

11. An apparatus according to claim 8 in which the coding means is arranged for each band to compare each sample with the sample obtained for another band and produce a '1' or '0' signal in dependence thereon.

12. A method for identifying the speaker of an unknown input speech signal as one of a group of previously identified speakers, said method

performed by a suitably programmed microprocessor or other data processing apparatus. Equally, digital filtering could be employed.

Claims

1. An apparatus for speaker identification comprising:
means for deriving from each of successive sample periods of speech input thereto a digital word representing the frequency spectrum thereof;
a store for storing codewords indicating for each digital word the relative probability of the frequency spectrum corresponding to that word occurring in the speech of one or more known speakers which the apparatus has been trained to identify, the digital word output of the deriving means being connected to the address inputs of the store; and
control means responsive, during input of speech by a speaker to be identified, to codewords read out from the store, to output, in accordance with a predetermined criterion related to the codewords, data indicating which of the known speakers is deemed to have been identified.

2. An apparatus according to claim 1, arranged in operation during a training sequence, for each sample period,
a) If the number of codewords already stored at the addressed location in the store does not exceed a predetermined number to store at that location a codeword identifying the speaker whose speech has given rise to the digital word representing that address,
b) If the number of codewords already stored at the addressed location in the store exceeds the predetermined number, to store at that location a codeword reserved for indication of multiple users.

3. An apparatus according to claim 2 in which the predetermined number is zero.

4. An apparatus according to claim 1, 2 or 3 in which the control means comprises a respective counter means for each known speaker and is arranged in operation to respond to each codeword read out from the store by incrementing the counter corresponding to the known speaker (if any) indicated by that codeword as having a relatively high probability of being associated with the frequency spectrum corresponding to the address from which the codeword was read, and in which the predetermined criterion is that the speaker who shall be deemed to be recognised shall be the one whose counter has the highest count.

5. An apparatus according to claim 4 in which the predetermined criterion includes the requirement that identification shall not be deemed to have occurred unless the said highest count exceeds the counts of the other counters by a set margin.

6. An apparatus according to claim 4 or 5 in which predetermined criterion includes the requirement that identification shall not be deemed to have occurred unless the said highest count exceeds a threshold value.

7. An apparatus according to claim 6 in which the control means contains a further counter, is responsive to the occurence of a codeword indicating a low probability of any known speaker to increment the further counter, and is responsive to the content of the further counter to adjust the said threshold value.

8. An apparatus according to any one of the preceding claims in which the deriving means comprises filter means for dividing the frequency spectrum of the speech input into a plurality of frequency band outputs, means for measuring the average amplitude of, and sampling those outputs, and coding means for obtaining the digital word from the resulting samples.

9. An apparatus according to claim 8 in which the coding means is arranged for each band to compare each sample with a threshold value and produce a '1' or '0' signal in dependence thereon.

10. An apparatus according to claim 8 in which the coding means is arranged for each band to compare each sample with the previous sample obtained for that band and produce a '1' or '0' signal in dependence thereon.

11. An apparatus according to claim 8 in which the coding means is arranged for each band to compare each sample with the sample obtained for another band and produce a '1' or '0' signal in dependence thereon.

12. A method for identifying the speaker of an unknown input speech signal as one of a group of previously identified speakers, said method

* C L M S *

comprising:
storing in a random access memory stored codewords uniquely corresponding to said previously identified speakers at address locations corresponding to the input speech's frequency spectrum, previously supplied by such speakers, a given one of said codewords being supplied at a memory output in response to input address signals;
applying to said memory a succession of addresses corresponding to the frequency spectrum of the unknown input speech signal; and
identifying the speaker of the unknown input speech signal from the resulting succession of codewords output from the memory.

13. A method for receiving spoken input and for recognising the identity of the speaker comprising the steps of:
deriving from each of a succession of input speech samples a digital word representing the frequency spectrum thereof;
storing a plurality of codewords in a memory device, each codeword indicating for each said digital word the relative probability of the frequency spectrum of said digital word occurring in the speech of one or more known speakers;
connecting the derived digital words to the addressing circuits of said memory device;
obtaining a succession of codewords accessed by said addressing circuits; and
using said succession of codewords, in accordance with a predetermined criterion related to the codewords, to indicate which speaker is identified.

Patentansprüche

1. Vorrichtung zur Identifizierung eines Sprechers, mit
   - Mitteln zum Erhalten eines digitalen Wortes, welches das Frequenzspektrum desselben darstellt, aus aufeinanderfolgenden Abtastperioden von einem eingegebenen Spracheingang,
   - einem Speicher zum Speichern von Codewörtern, die für jedes Digitalwort die relative Wahrscheinlichkeit des Frequenzspektrums angeben, das dem in der Sprache auftretenden Wort von einem oder mehreren bekannten Sprechern entspricht, auf die die Vorrichtung zum Identifizieren trainiert wurde, wobei der digitale Wortausgang der Mittel zum Erhalten desselben mit den Adreßeingängen des Speichers verbunden sind, und
   - Steuermitteln, die während der Spracheingabe durch einen zu identifizierenden Sprecher auf Codewörter reagiert, die aus dem Speicher ausgelesen werden, zum Ausgeben, gemäß einem vorbestimmten Kriterium entsprechend den Codewörtern, von Daten, welche angeben, welcher von den bekannten Sprechern als identifiziert angenommen wird.

2. Vorrichtung nach Anspruch 1, die im Betrieb während einem Trainingsdurchlauf für jede Abtastperiode,
   a) wenn die Zahl der bereits an einer adressierten Stelle im Speicher gespeicherten Codewörter eine vorbestimmte Anzahl nicht überschreitet, an dieser Stelle ein Codewort speichert, welches den Sprecher identifiziert, dessen Sprache das Digitalwort ergeben hat, das diese Adresse repräsentiert,
   b) wenn die Zahl der schon an der adressierten Stelle im Speicher gespeicherten Codewörter die vorbestimmte Anzahl überschreitet, an dieser Stelle ein Codewort speichert, das für die Angabe von mehreren Benutzern reserviert ist.

3. Vorrichtung nach Anspruch 2, in der die vorbestimmte Anzahl Null ist.

4. Vorrichtung nach Anspruch 1, 2 oder 3, in der die Steuermittel eine entsprechende Zähleinrichtung für jeden bekannten Sprecher enthalten, die im Betrieb derart ausgebildet ist, daß für jedes aus dem Speicher ausgelesene Codewort der Zähler entsprechend dem bekannten Sprecher (wenn ein solcher vorliegt) inkrementiert wird, wodurch angegeben wird, daß das Codewort eine relativ hohe Wahrscheinlichkeit des in Verbindungbringens mit dem Frequenzspektrum entsprechend der Adresse aus der das Codewort gelesen wurde, aufweist und in der das vorbestimmte Kriterium ist, daß der Sprecher der als erkannt angenommen werden soll, der sein soll, dessen Zähler den höchsten Zählerstand aufweist.

5. Vorrichtung nach Anspruch 4, in der das vorbestimmte Kriterium das Erfordernis enthält, daß die Identifikation nicht als aufgetreten gelten soll, außer wenn der höchste Zählerstand den Zählerwert der anderen Zähler um einen gesetzten Bereich überschreitet.

6. Vorrichtung nach Anspruch 4 oder 5, in der das vorbestimmte Kriterium das Erfordernis enthält, daß die Identifikation nicht als aufgetreten gewertet werden soll, außer wenn der höchste Zählerstand einen Schwellwert überschreitet.

* C L M S *

7. Vorrichtung nach Anspruch 6, in der die Steuermittel einen weiteren Zähler enthalten, der auf das Auftreten eines Codeworts reagiert, das eine niedrige Wahrscheinlichkeit eines jeden bekannten Sprechers angibt, zum Inkrementieren des weiteren Zählers und auf den Inhalt des weiteren Zählers zum Einstellen des Schwellwerts reagieren kann.

8. Vorrichtung nach einem der vorangehenden Ansprüche, in der die Mittel zum Erhalten des Digitalworts Filtermittel zum Trennen des Frequenzspektrums der Spracheingabe in mehrere Frequenzbandausgänge enthält, weiterhin Mittel zum Messen der Durchschnittsamplitude der Ausgänge und ein Abtasten derselben und Codiermittel zum Erhalten des Digitalworts aus den sich ergebenden Abtastungen.

9. Vorrichtung nach Anspruch 8, in der die Codiermittel für jedes Band derart vorgesehen sind, daß jede Abtastung mit einem Schwellwert verglichen wird und ein "1"- oder "0"-Signal in Abhängigkeit davon erzeugt wird.

10. Vorrichtung nach Anspruch 8, in der die Codiermittel für jedes Band derart angeordnet sind, daß jede Abtastung mit einer vorangehenden für dieses Band erhaltenen Abtastung verglichen wird und in Abhängigkeit davon ein "1"- oder "0"-Signal erzeugt wird.

11. Vorrichtung nach Anspruch 8, in der die Codiermittel für jedes Band derart angeordnet sind, daß jede Abtastung mit der Abtastung verglichen wird, die für ein anderes Band erhalten wurde, und in der in Abhängigkeit davon ein "1"- oder "0"-Signal erzeugt wird.

12. Verfahren zum Identifizieren des Sprechers eines unbekannten Eingangssprachsignals als einer aus einer Gruppe von vorangehend identifizierten Sprechern, mit den folgenden Schritten:
- Speichern von in einem Speicher mit wahlfreiem Zugriff gespeicherten Codewörtern gleichmäßig entspechend den vorangehend identifizierten Sprechern an Adreßstellen, entsprechend dem Frequenzspektrum der eingegebenen Sprache, das vorangehend von diesen Sprechern zugeführt wurde, wobei ein gegebenes dieser Codewörter einem Speicherausgang zugeführt wird, entsprechend den Eingangsadreßsignalen,
- Anwenden einer Abfolge von Adressen auf den Speicher, entsprechend dem Frequenzspektrum des unbekannten Eingabesprachsignals und
- Identifizieren des Sprechers des unbekannten Eingangsprachsignals aus den sich ergebenden aufeinanderfolgenden Codewortausgängen von dem Speicher.

13. Verfahren zum Erhalten des gesprochenen Eingangs und zum Erkennen der Identität des Sprechers mit den folgenden Schritten:
- Erhalten aus jeder Abfolge von Eingangssprachabtastungen ein Digitalwort, welches das Frequenzspektrum darstellt,
- Speichern mehrerer Codewörter in einer Speichervorrichtung, wobei jedes Codewort für jedes der Digitalworte die relative Wahrscheinlichkeit des Frequenzspektrums des Digitalworts angibt, welches in der Sprache von einem oder mehreren bekannten Sprechern auftritt,
- Verbinden des erhaltenen Digitalworts mit der Adreßschaltung der Speichervorrichtung,
- Erhalten einer Abfolge von Codewörtern, die durch die Adreßschaltungen zugänglich geworden sind und
- Verwenden der Abfolge von Codewörtern zum Angeben welcher Sprecher identifiziert wurde, gemäß einem vorbestimmten Kriterium entsprechend den Codewörtern.

Revendications

1. Dispositif pour l'identification d'un locuteur comprenant :
des moyens pour obtenir à partir de chacune des périodes d'échantillonnage successives de l'entrée de la parole dans ceux-ci, un mot numérique représentant son spectre de fréquence,
une mémoire pour mémoriser des mots de code indiquant pour chaque mot numérique la probabilité relative du spectre de fréquence correspondant à ce mot à se produire dans la prononciation d'un ou plusieurs locuteurs connus que le dispositif a appris à identifier, la sortie du mot numérique des moyens d'obtention étant reliée aux entrées d'adresse de la mémoire, et
des moyens de commande répondant pendant l'entrée de la prononciation par un locuteur qui doit être identifié au mot de code extrait de la mémoire pour sortir en conformité avec un critère prédéterminé par rapport au mot de code, des données indiquant lequel des locuteurs connus est estimé avoir été identifié.

2. Dispositif selon la revendication 1, aménagé, durant le fonctionnement pendant une séquen-

* C L M S *

ce d'apprentissage pour chaque période d'échantillonnage,

a) Si le nombre des mots de code déjà mémorisés à l'emplacement adressé dans la mémoire ne dépasse pas un nombre prédéterminé, de manière à mémoriser à cet emplacement un mot de code identifiant le locuteur dont la parole a donné naissance au mot numérique représentant cette adresse,

b) Si le nombre des mots de code déjà mémorisés à l'emplacement adressé dans la mémoire dépasse le nombre prédéterminé, de manière à mémoriser à cet emplacement un mot de code réservé pour indication d'utilisateurs multiples.

3. Dispositif selon la revendication 2, dans lequel le nombre prédéterminé est égal à zéro.

4. Dispositif selon la revendication 1, 2 ou 3 dans lequel le moyen de commande comprend un moyen de compteur respectif pour chaque locuteur connu et est placé en fonctionnement pour répondre à chaque mot de code extrait de la mémoire en incrémentant le compteur correspondant au locuteur connu (s'il en existe un) indiqué par ce mot de code comme ayant une probabilité relativement élevée d'être associé au spectre de fréquence correspondant à l'adresse à partir de laquelle le mot de code a été extrait et dans lequel le critère prédéterminé est que le locuteur que l'on estime être reconnu sera celui dont le compteur présente le compte le plus élevé.

5. Dispositif selon la revendication 4, dans lequel le critère prédéterminé comporte l'exigence que l'identification ne soit pas estimée s'être produite à moins que le compte le plus élevé dépasse les comptes des autres compteurs par une marge établie.

6. Dispositif selon la revendication 4 ou 5, dans lequel le critère prédéterminé comporte l'exigence que l'identification ne soit pas estimée s'être produite à moins que le compte le plus élevé dépasse une valeur de seuil.

7. Dispositif selon la revendication 6, dans lequel le moyen de commande contient un autre compteur, lequel est sensible à la présence d'un mot de code indiquant une faible probabilité de tout locuteur connu quelconque pour incrémenter l'autre compteur et est sensible au contenu de l'autre compteur pour ajuster ladite valeur de seuil.

8. Dispositif selon l'une quelconque des revendications précédentes, dans lequel le moyen d'obtention comprend des moyens de filtre pour diviser les spectres de fréquence de l'entrée vocale en une pluralité de sorties de bande de fréquence, des moyens pour mesurer l'amplitude moyenne et échantillonner ces sorties et des moyens de codage pour obtenir le mot numérique à partir des échantillons obtenus.

9. Dispositif selon la revendication 8, dans lequel le moyen de codage est prévu pour chaque bande afin de comparer chaque échantillon à une valeur de seuil et produit un signal à "1" ou à "0" en fonction de celui-ci.

10. Dispositif selon la revendication 8, dans lequel le moyen de codage est prévu pour chaque bande afin de comparer chaque échantillon à l'échantillon précédent obtenu pour cette bande et produire un signal à "1" ou à "0" en fonction de celui-ci.

11. Dispositif selon la revendication 8, dans lequel le moyen de codage est prévu pour chaque bande afin de comparer chaque échantillon à l'échantillon obtenu pour une autre bande et produire un signal à "1" ou à "0" en fonction de celui-ci.

12. Procédé pour identifier le locuteur d'un signal vocal d'entrée non connu comme un d'un groupe de locuteur précédemment identifiés, ledit procédé comprenant les étapes consistant à :

mémoriser dans une mémoire vive des mots de code mémorisés correspondant particulièrement auxdits locuteurs précédemment identifiés à des emplacements d'adresse correspondant au spectre de fréquence de la parole d'entrée, précédemment délivrée par ces locuteurs, un mot de code donné parmi les mots de code étant délivré à une sortie de la mémoire en réponse aux signaux d'adresse d'entrée,

appliquer à ladite mémoire une suite d'adresses correspondant au spectre de la fréquence du signal vocal d'entrée non connu, et

identifier le locuteur du signal vocal d'entrée non connu à partir de la suite résultante des mots de code sortis de la mémoire.

13. Procédé pour recevoir une entrée parlée et pour reconnaître l'identité de locuteur comprenant les étapes consistant à :

obtenir à partir de chacune d'une suite d'échantillons vocaux d'entrée un mot numéri-

* C L M S *

que représentant son spectre de fréquence,
mémoriser une multitude de mots de code dans un dispositif mémoire, chaque mot de code indiquant pour chaque mot numérique la probabilité relative du spectre de fréquence dudit mot numérique se produisant dans la parole d'un ou plusieurs locuteurs connus,
connecter les mots numériques obtenus au circuit d'adressage dudit dispositif mémoire,
obtenir une succession de mots de code auxquels on a accédé par lesdits circuits d'adressage, et

utiliser ladite suite de mots de code, en conformité avec un critère prédéterminé se rapportant aux mots de code afin d'indiquer quel locuteur est identifié.

* C L M S *