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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 March 2010 (11.03.2010)

PCT

(10) International Publication Number
WO 2010/025523 A1

(51) International Patent Classification:
G10L 17/00 (2006.01) *G06F 21/00* (2006.01)

(21) International Application Number:
PCT/AU2009/001165

(22) International Filing Date:
7 September 2009 (07.09.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2008904631 5 September 2008 (05.09.2008) AU

(71) Applicant *(for all designated States except US)*: **AURAYA PTY LTD** [AU/AU]; PO Box 1045, Dickson, Australian Capital Territory 2602 (AU).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SUMMERFIELD, Clive** [AU/AU]; PO Box 1045, Dickson, Australian Capital Territory 2602 (AU). **TALHAMI, Habib, Emile** [AU/AU]; 35B Epping Road, Epping, NSW 2121 (AU).

(74) Agent: **GRIFFITH HACK**; Level 29, Northpoint, 100 Miller Street, North Sydney, NSW 2060 (AU).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

WO 2010/025523 A1

(54) Title: VOICE AUTHENTICATION SYSTEM AND METHODS

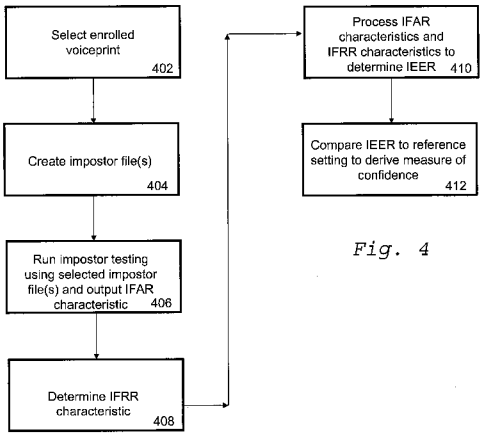

Fig. 4

(57) Abstract: A method for configuring a voice authentication system comprises ascertaining a measure of confidence associated with a voice sample enrolled with the authentication system. The measure of confidence is derived through simulated impostor testing carried out on the enrolled sample.

VOICE AUTHENTICATION SYSTEM AND METHODS

Field of the Invention

The present invention relates generally to a voice authentication system and methods.

Background of the Invention

Voice authentication systems are becoming increasingly popular for providing access control. For example, voice authentication systems are currently being utilised in telephone banking systems, automated proof of identity applications in call centres systems, automatic teller machines, building and office entry access systems, automated password reset, call back verification for highly secure internet transactions, etc.

Voice authentication is typically conducted over a telecommunications network, as a two stage process. The first stage, referred to as the enrolment stage, involves processing a sample of a person's voice presented to a voice authentication engine to generate an acoustic model or "voiceprint" that represents their unique voice characteristics. The second stage, or authentication stage, involves receiving a voice sample of a person to be authenticated (or identified) over the network. Again, the voice authentication engine generates an acoustic model of the sample and compares this with the stored voiceprint to derive an authentication score indicating how closely matched the two samples are (and therefore the likelihood that the person is, in fact, the same as that being claimed). This score is typically expressed as a numerical value and involves various mathematical calculations that can vary from engine to engine.

In the case of the correct, or "legitimate", person accessing the authentication system, the expectation is that their voiceprint (i.e. generated from their voice file) will closely

match the voiceprint previously created for that person, resulting in a high score. If a fraudster (often referred to in the art as an "impostor") is attempting to access the system using the legitimate person's information (e.g. speaking their account number, password, etc), the expectation is that the impostor's voiceprint will not closely match the legitimate person's voiceprint, thus resulting in a low score even though the impostor is quoting the correct information.

Whether a person is subsequently deemed to be legitimate is typically dependent on a threshold set by the authentication system. To be granted access to the system, the score generated by the authentication system needs to exceed the threshold. If the threshold score is set too high then there is a risk of rejecting large numbers of legitimate persons. This is known as the false rejection rate (FRR). On the other hand, if the threshold is set too low there is a greater risk of allowing access to impostors. This is known as the false acceptance rate (FAR).

As one would appreciate, therefore, selecting an appropriate threshold for an authentication system can be difficult to achieve. On one hand the threshold setting needs to be high enough that business security requirements of the secure services utilising the authentication system are met. However, such settings can cause undue service issues with too many legitimate persons being rejected. Similarly, if the threshold is set too low, while achieving good services levels, security may be put at risk. The problem of selecting appropriate threshold settings is compounded by the fact that different authentication engines utilise different attributes or characteristics for voiceprint comparison and as a result may produce a wide range of different scores based on the same type of content provided in the voice samples (e.g. number, phrases, etc.). What is more, a single engine will also produce quite different scores for voice samples of different content types,

for example an account number compared to a date of birth, or a phrase.

Definitions

"Voice Sample" is used herein to denote a sample of a person's voice.

"Voice file" is the storage of a voice sample as a data file.

"Voiceprint" is an acoustic model of a person's voice characteristics (i.e. an acoustic model). Voiceprints are generated from voice samples/files and may be processed by a voice authentication engine to generate probability scores as to how closely the characteristics of an associated speaker match those of another speaker.

"Content Type" - refers to the type of content being provided in the voice sample. For example, the content may be a spoken account number or password. Other content types can include but are not limited to an answer to a question; an unconstrained passage of speech as spoken by a caller to a call centre agent; or a standard phrase (e.g. "At ABC bank, my voice is my password"). In an embodiment, content type can also refer to the type of input device being used to provide the sample (e.g. mobile phone, landline, etc.).

"Impostor" is used herein to refer to a person that is known to the system but is not the "legitimate speaker" under test. The term "impostor" is also used as a technical term used in this document to describe the behaviour of a fraudster or an identity thief that is using a legitimate speaker's content information (e.g. spoken account number and password) in an attempt to gain access to that person's secure services.

"Threshold" refers to a base setting against which an authentication score is compared for determining whether to accept or reject that speakers claimed identity. If the score exceeds the threshold, the person will typically be accepted. If the score is below the threshold, the authentication system typically rejects the person. Multiple thresholds may be utilised associated with different levels of assurance.

"Business Rule" is used herein to refer to one or more risk parameters determined by a secure service associated with allowing customers access to different secure services. For example, a business rule may stipulate that a user only has to receive a moderate authentication score in order to gain access to their account for performing standard transactions (e.g. to pay their electricity bill), but would need to be authenticated to a much higher level of certainty (i.e. produce a high authentication score) to perform high cost high risk transactions such as accessing overseas accounts, etc.

"False-Accept Rate" (FAR) is a measure of the rate at which impostors are incorrectly accepted as a legitimate speaker by an authentication system. In one embodiment, the FAR may be defined as: the number of impostors accepted divided by the total number of speakers tested.

"False-Reject Rate" (FRR) is a measure of the rate at which legitimate speakers are incorrectly rejected by the system. In one embodiment, the FRR may be defined as: the number of speakers rejected divided by the total number of speakers tested.

"Equal-Error Rate" (EER) is a characteristic resulting from a threshold setting of the system where the false-accept rate (FAR) equals the false-reject rate (FRR). The EER is used as a raw measure of how well an authentication system can separate legitimate speakers from impostor speakers. The lower the score,

the better the technology is at performing this function. A 0% EER indicates that the authentication system has been able to separate all legitimate speakers from all the impostors.

"Failure To Acquire" (FTA) means that a particular voice file cannot be used to obtain a suitable verification result, e.g. the noise level is too high.

"Failure To Enroll" (FTE) means that a particular set of voice files cannot be used to compute a voiceprint, e.g. the amount of speech data is insufficient for enrollment.

"IVR" refers to an Interactive Voice Response (system).

"World Models" (also referred to as "Universal" or "Background Models") are speech models generated from a complete population of speakers and may be used in the authentication system to normalise the score generated by each individual speaker models.

Summary of the Invention

In accordance with a first aspect, the present invention provides a method for configuring a voice authentication system, the method comprising the steps of: ascertaining a measure of confidence associated with a voice sample enrolled with the authentication system, the measure of confidence being derived through simulated impostor testing carried out on the enrolled sample.

Through extensive testing, the present inventors have discovered that not all voiceprints enrolled with an authentication system have the same security performance. That is, some voiceprints are more easily broken into by fraudsters than others (i.e. they are more vulnerable). In light of this discovery, embodiments of the present invention are operable to measure the performance of each individual voiceprint (i.e. by deriving individual

confidence measures for each voice sample). In embodiment, the authentication system can then implement various optimisation actions to improve the performance of the voiceprint and thus the overall performance of the authentication system. For example, where an individual is known to have a weak or vulnerable voiceprint (i.e. one with low measure of confidence), then special procedures and rules can be applied to either strengthen the voiceprint or implement special procedures, such as asking additional questions or passing the caller to a call centre for special processing to strengthen the identity authentication process for those known to have weak voiceprints.

In an embodiment the method comprises the further step of implementing an optimisation action for the enrolled voice sample based, at least in part, on the ascertained measure of confidence.

In an embodiment the simulated impostor testing comprises utilising at least one authentication engine to compare at least one impostor voice sample against a voiceprint derived from the enrolled sample, to determine an individual false acceptance rate.

In an embodiment the individual false acceptance rate (IFAR) is utilised to derive the measure of confidence.

In an embodiment the method comprises the further step of determining an individual false rejection rate (IFRR) for the enrolled sample, such that the IFRR is additionally utilised to derive the measure of confidence.

In an embodiment the step of determining the IFRR comprises utilising an authentication engine to compare a legitimate voice sample (i.e. a sample provided by the same speaker to which the enrolled sample belongs) against a voiceprint derived from the enrolled voice sample, to output a score which can be processed

to determine the IFRR.

In an embodiment the IFAR and IFRR are utilised to determined an individual equal error rate (IEER) associated with the enrolled voice sample.

In an embodiment the method comprises the further step of comparing the IEER with a reference setting to derive the measure of confidence.

In an embodiment the reference setting is a mean individual equal error rate for a plurality of other samples enrolled with the system.

In an embodiment a weak measure of confidence is assigned to the enrolled voice sample responsive to determining that the IEER is greater than the mean IEER.

In an embodiment, responsive to establishing that the enrolled voice sample is weak, the method comprises carrying out the optimisation action of re-building a voiceprint associated with the enrolled voice sample to adjust a speaker and/or environmental characteristic associated with the voiceprint.

In an embodiment, responsive to establishing that the enrolled voice sample is weak, the method comprises carrying out the optimisation action of re-building a world model from which the associated voiceprint was derived.

In an embodiment, responsive to establishing that the enrolled voice sample is weak, the method comprises carrying out the optimisation action of re-building the voiceprint.

In an embodiment the optimisation action comprises setting a threshold associated with the enrolled sample, based on the derived measure of confidence.

In an embodiment, upon determining that the measure of confidence does not meet a set threshold, the optimisation action comprises requesting that the voice sample be re-enrolled.

In an embodiment the optimisation step is repeated each time a new voice sample is enrolled with the system.

In an embodiment the optimisation action is carried out for enrolled voice samples until a threshold performance measure for the system has been met.

In an embodiment the threshold performance measure is associated with an overall equal error rate for the system.

In an embodiment the impostor samples have the same content type and/or speaker characteristic as the enrolled sample.

In an embodiment the impostor samples are samples provided by other legitimate persons during either enrolment with the system or during a subsequent authentication session.

In accordance with a second aspect, the present invention provides a method for configuring a voice authentication system, the method comprising the steps of:

ascertaining a measure of confidence associated with a voiceprint of a voice sample enrolled with the authentication system, the measure of confidence being derived through simulated impostor testing carried out on the enrolled sample.

In accordance with a third aspect, the present invention provides a voice authentication system comprising:

an ascertaining module operable to ascertain a measure of confidence associated with a voice sample enrolled with the authentication system, the measure of confidence being derived

through simulated impostor testing carried out on the enrolled sample by an impostor testing module.

In an embodiment the system further comprises an optimisation module operable to implement an optimisation action for the enrolled voice sample based, at least in part, on the ascertained measure of confidence.

In an embodiment the impostor testing module compares at least one impostor voice sample against the enrolled sample, to determine an individual false acceptance rate.

In embodiment the impostor testing module comprises an authentication engine operable to compare the at least one impostor voice sample against a voiceprint derived from the enrolled sample, the resultant scores processed by the testing module to provide the individual false acceptance rate.

In an embodiment the individual false acceptance rate is utilised to derive the measure of confidence.

In an embodiment the impostor testing module is further arranged to determine an individual false rejection rate for the enrolled sample, the individual false rejection rate being additionally utilised to derive the measure of confidence.

In an embodiment the individual false rejection rate is determined utilising an authentication engine which is operable to compare a legitimate voice sample against a voiceprint derived from the enrolled voice sample to output a score which can be processed to determine the IFRR.

In an embodiment the individual false acceptance rate and individual false rejection rate are utilised to establish an individual equal error rate (IEER) for the enrolled voice sample.

In an embodiment the impostor testing module is operable to compare the IEER with a reference setting to derive the measure of confidence.

In an embodiment the reference setting is a mean individual equal error rate for a plurality of other samples enrolled with the system.

In an embodiment a weak measure of confidence is assigned to the enrolled voice sample responsive to determining that the IEER is greater than the mean IEER.

In an embodiment responsive to establishing that the enrolled voice sample is weak, the optimisation module re-builds a voiceprint associated with the enrolled voice sample to adjust a speaker and/or environmental characteristic associated with the voiceprint.

In an embodiment the optimisation module re-builds a world model from which the associated voiceprint was derived, responsive to establishing that the enrolled voice sample is weak.

In an embodiment the optimisation module sets an acceptance threshold associated with the enrolled sample, based on the derived measure of confidence.

In an embodiment the optimisation module requests that the voice sample be re-enrolled, upon determining that the measure of confidence does not meet a set threshold.

In an embodiment the optimisation action is carried out each time a new voice sample is enrolled with the system.

In an embodiment the optimisation module continues to carry out optimisation actions until a threshold performance measure for

the system has been met.

In an embodiment the performance measure is associated with an overall equal error rate for the system.

In accordance with a fourth aspect the present invention provides a method for providing a secure service, comprising the steps of:

receiving data indicative of a measure of confidence associated with a user of the secure service, the measure of confidence being derived through simulated impostor testing carried out on an voice sample of the user; and

adjusting a level of authentication required by the user to access the secure service based, at least in part, on the measure of confidence.

In an embodiment the level of authentication is adjusted by setting an acceptance threshold level.

In an embodiment the simulated impostor testing is carried out using the methodology according to the first aspect.

In accordance with a fifth aspect the present invention provides a secure service provider system comprising:

a receiving module operable to receive data indicative of a measure of confidence associated with a user of the secure service, the measure of confidence being derived through simulated impostor testing carried out on a voice sample of the user; and

an adjustment module operable to adjust a level of authentication required by the user to access the secure service based, at least in part, on the measure of confidence.

In accordance with a sixth aspect the present invention provides a computer program comprising at least one instruction for

controlling a computing system to implement a method in accordance with the first aspect.

In accordance with a seventh aspect the present invention provides a computer readable medium providing a computer program in accordance with the fourth aspect.

Brief Description of the Drawings

Features and advantages of the present invention will become apparent from the following description of embodiments thereof, by way of example only, with reference to the accompanying drawings, in which:

Figure 1a is a block diagram of a system in accordance with an embodiment of the present invention;

Figure 1b is a schematic of the individual modules implemented by the third party server of Figure 1a;

Figure 2 is a basic process flow for carrying out an embodiment of the present invention.

Figure 3 is a flow diagram showing the method steps for enrolling, in accordance with an embodiment of the invention;

Figure 4 is a flow diagram for deriving individual confidence measures;

Figure 5 is a schematic illustrating the system components utilised in re-building world and speaker models;

Figure 6 is a screen shot generated by a graphics rendering application, in accordance with an embodiment;

Figure 7 is a screen shot generated by a graphics rendering application, in accordance with an embodiment, showing different thresholds automatically set by the system per speaker and for speech samples with different content types;

Figure 8 is a screen shot generated by a graphics rendering application, in accordance with an embodiment, showing the speaker information generated by the system;

Figures 9 and 10 are screen shots in accordance with further embodiments of the present invention;

Figure 11 is a screen shot generated by a graphics rendering application, in accordance with an embodiment showing the optimisation process in action and reporting optimisation results; and

Figure 12 is a screen shot generated by a graphics rendering application, in accordance with an embodiment showing the Equal Error rate (EER) for the overall system after the optimisation process for speech samples with Content Type 1 (spoken account numbers) and Content Type 8 (a spoken phrase).

Detailed Description of Preferred Embodiments

For the purposes of illustration, and with reference to the figures, embodiments of the invention will hereafter be described in the context of a voice authentication system for a secure service, such as a secure interactive voice response ("IVR") telephone banking service. In the illustrated embodiment, the authentication system is implemented as a third party system independent of the secure service. In the illustrated embodiment, the authentication system is implemented as a third party system independent of the secure service. It will be understood by persons skilled in the art, however, that both the secure service and authentication system may be

integrated as a single service. Persons (hereafter "customers") communicate with the authentication system using an input device in the form of a fixed telephone (although it will be understood that a mobile telephone, VOIP pc-based telephone, or the like may equally be utilised for communicating with the authentication system).

Fig. 1a illustrates an example system configuration 100 for implementing an embodiment of the present invention. The system 100 includes a user input device 102 in the form of a standard telephone; third party authentication system 104 (hereafter "third party system"); secure service provider system 106 in the form of an Internet banking server hosting a secure customer banking web site; and communications system 108, in the form of a public-switched telephone network.

With reference to Fig. 2 there is shown a flowchart illustrating method steps for implementing an embodiment of the present invention. Embodiments are operable to ascertain a measure of confidence associated with a voiceprint of a voice sample which has been enrolled with the third party system 104 (step 202). In an embodiment, once the measure of confidence has been derived, either the third party system 104 and/or secure service provider system 106 are operable to implement various optimisation actions based on the determined confidence measure (204). In an embodiment, the measure of confidence is determined by carrying out simulated impostor attacks on each enrolled voiceprint. Further, by comparing the individual measures against a baseline or reference confidence measure (e.g. such as an average confidence measure for the system, etc), voiceprints that have an increased susceptibility to a real impostor attack (i.e. "weak" voiceprints) can readily be determined. Optimisation actions can then be taken in order to increase the strength of the voiceprints and thus improve the overall system robustness.

The following description will first describe an example process for "enrolling" (i.e. initially storing voice samples with the system) and then go on to describe embodiments for determining the individual confidence measures and optimisation actions that can be taken to improve the performance and robustness of the third party system 104. In this description the word "customer" refers to a person speaking to the system over a communications network.

INITIAL ENROLMENT

With additional reference to Figure 3, at step 302 a customer dials a telephone banking number associated with the secure service 106. The third party system 104 answers the call and enrolment begins. This may involve requesting that the customer utter speech of a particular type of information (i.e. content type) such as, for example, their customer number, password, a common generic phrase, etc. The system 104 may ask the customer to repeat the utterance a number of times until the system 104 has sufficient samples to create a voiceprint.

According to the illustrated embodiment, a customer's voice sample is subsequently recorded as a voice file and processed to create the voiceprint (also referred to as speaker model). The voice file is stored in database 107, whilst the voiceprint is stored in voiceprint database 109 (step 304). The voiceprint is stored in association with a customer identifier; in this case their customer number recorded by the identity management database 111. In an embodiment the voiceprint is derived from one or more generic world or background models, using techniques known to persons skilled in the art. It will be understood that more than one voice sample (e.g. associated with different content types) may be recorded by each customer (step 304a). For example, the customer may provide separate samples for their account number, telephone number, name, pin number, phrase etc. In an embodiment, the customer may also be asked to answer a

shared secret question or utter a standard phrase (such "At ABC bank, my voice is my password"). It will be understood that these phrases may be used not only to effectively build the authentication system, but also to strengthen security by providing additional authentication samples on which to base an authentication decision.

After the customer voiceprint(s) have been successfully "enrolled", the third party system 104 may test both the failure to enrol (FTE) and failure to acquire (FTA) characteristics, using techniques known to persons skilled in the art (step 306). These statistics are logged by the third party system 104. The process ends with the caller hanging up at step 308.

IMPOSTOR FILES

With additional reference to Figure 1b, the third party system 104 is operable to retrieve files of other customers from the voice file database for use in the impostor testing process. The retrieved files may be tested against the selected voice file on the fly, or alternatively stored in an impostor database (not shown) for batch testing at some late time (e.g. during low usage times). In an embodiment, the voice files selected for impostor testing share the same content type as the file under test. For example, where the voice file under test is associated with a male speaker speaking account numbers; only male voice files saying account numbers will be utilised for impostor testing. In a further embodiment, the impostor files are selected from files that have been provided by other customers for an authentication session. In an embodiment, only files which have scored highly in those previous authentication sessions may be utilised for the impostor testing.

Where the authentication session utilises text dependent authentication engines, the impostor files may be processed (e.g. by segmenting and re-ordering) to generate the requisite

content information for the customer file being tested. In other words, in an embodiment, in order to create impostor voice files, the voice files stored by the database 109 are processed to generate the requisite content information for the customer file being tested (i.e. the "legitimate" voice file). Alternatively, for text independent processing (or where a standard/generic phrase is used for authentication), the retrieved voice files can be used directly as impostor voice files.

The number of voice files selected for the simulated impostor testing will depend on the particular implementation. In other words, the third party system 104 may apply as many voice files as required to produce adequate coverage across the authentication system (i.e. to ensure that an accurate measure of the strength of individual voiceprints can be made and hence the measure of confidence associated therewith). Furthermore, the process of storing voice files in the database 109 may be on-going; that is, new voice samples successfully captured during enrolment or extracted from successful authentication sessions, may be stored in the database 109 for subsequent use in the impostor attacks and re-building of world models.

DETERMINING CONFIDENCE MEASURES

As previously mentioned, the third party system 104 is operable to ascertain measures of confidence for each voiceprint so as to identify voiceprints that are weak and susceptible to impostor attack. Action may then be taken to improve security performance of those weak voiceprints.

With reference to Figure 4, the first step in deriving the measure of confidence involves establishing how well the voiceprint performs in response to a simulated impostor attack. The simulated impostor attack process involves selecting a customer voiceprint that has been produced during enrolment

(step 402). The selected customer voiceprint will hereafter be referred to as the "legitimate" speaker voiceprint. At step 404, one or more voice files of other known customers are retrieved from the voice file database 107 hereafter referred to as impostor voice files, using techniques previously described. The impostor voice files are then applied to the voice authentication engine and the resultant authentication scores produced by the engine when referencing the selected voiceprint are stored in association with the voiceprint under test (step 406).

As mentioned above, one technique for creating impostor voice file is to segment and re-order parts of other customer voice files to create a file having the same content information as was present in the sample from which the target voiceprint was derived. This process may involve, for example, passing the other customer files through a speech recognition engine configured to recognise the constituent parts of the files and segment into voice files accordingly. The process then continues by re-ordering the constituent parts to form the same spoken content as was present in the legitimate person's voice sample. In an embodiment, the basic process for generating an authentication score comprises performing an acoustic analysis of the voice file to produce a sequence of acoustic vectors representing the relevant voice characteristics for statistical analysis and comparison. Statistical pattern matching algorithms operating in the authentication engine compare the sequence of acoustic vectors with the voiceprint of the legitimate customer to generate a probability score representing how well the voice signal matches the legitimate voiceprint (i.e. an indication of the likelihood that the customer providing both samples is one and the same). Such pattern matching algorithms may include dynamic time warping (DTW), the hidden Markov model (HMM), among others. Further, the algorithms operating in the authentication engine also compare the acoustic vector sequence with the World Model to provide a

reference score against which to calibrate the probability scores generated by the user voiceprint. The resultant calibrated probability scores thus provides a measure of how well the impostor voice files matched against the legitimate customer's voiceprint. These measures can thus be used to generate a False Accept characteristics for that customer's vocieprint and can be used to compute the false accept rate for that speaker which is hereafter referred to as the individual false acceptance rate (IFAR).

The next step, step 408, in deriving the measure of confidence involves establishing the false rejection rate associated with the voiceprint (hereafter the individual false rejection rate, or "IFRR"). According to the embodiment described herein, the IFRR is determined by testing the voiceprint with other voice samples of the same content type which have been provided by the legitimate speaker (e.g. either other enrolled samples, or samples which have subsequently been provided during authentication session). An interpolation algorithm is used to smooth the IFRR characteristic where only a few voice samples or voice files are available for determining the IFRR. Alternatively, the FRR for the authentication system as a whole can also be used for the IFRR where there are too few samples to produce an accurate IFFR. Also, at step 408, the overall system EER is established and recorded, for reasons which will become apparent in subsequent paragraphs.

At step 410, the IFAR and IFRR score are processed to determine the individual EER (hereafter "IEER") for the voiceprint. The IEER is determined where the IFAR and IFRR characteristics intersect, (i.e. where the IFRR = IFAR). The IEER, in turn, can be utilised to derive a measure of confidence in the performance of the selected voiceprint, as will be described in subsequent paragraphs.

In an embodiment, the measure of confidence is based, at least in part, on the relationship between the IEER and some reference, such as the average system EER (i.e. the statistical mean of all IEER scores evaluated and recorded by the authentication system). See step 412. In another embodiment, the IEER may be compared against the median EER, the mode of the EER, or some other statistical EER average value which provides a meaningful reference point for establishing the confidence measure. In a specific embodiment, either a weak or strong measure will be attributed to the voiceprint, based on the relationship between the IEER for that voiceprint the mean EER for the system as a whole. In an embodiment, a strong voiceprint is associated with a voiceprint which has a lower IEER than the mean; whereas a weak measure is attribute to a voiceprint having a higher IEER than the mean. The actual deviation between the IEER and mean EER may further be used to evaluate and record the relative strength or weakness of the voiceprint.

OPTIMISATION ACTIONS

Once the IEER score and confidence measure have been derived for the selected voiceprints, a number of different optimisation actions can be carried out by the authentication system 100 to improve the performance of the enrolled voiceprints and thus the performance of the authentication system as a whole.

One such optimisation action involves assigning appropriate individual speaker thresholds for each customer, based on the derived measure of confidence. By assigning appropriate individual thresholds, the percentage of false acceptances and false rejections can be controlled at a customer (per speaker) level, resulting in improved individual customer security and usability. For example, where an individual voiceprint is deemed to be strong, the threshold setting for that voiceprint can be set high, thus increasing the security level for the

associated customer without affecting the performance of the system. Conversely, where an individual voiceprint is deemed by the system to be weak, then the threshold for acceptance can be set lower, thereby reducing the probability of the customer being falsely rejected by the authentication system 100.

Figure 7 shows a screen shot of the graphical user interface which shows different threshold settings derived by the third party system (104) for a number of different speakers (identified by their "ID") for speech items having different Content Types. In this embodiment, the system has computed two threshold settings (upper and lower) which are used by the application to enhance the user interaction with the system. In this example, Threshold 1 for speaker identity 460005 has an upper threshold setting of 49.49 for speaker item 1, compared to 61.56 for speaker identity 460001.

In addition, since the system 100 has recognised that a voiceprint is weak (which in this case equates to EER which is above a particular percentage, e.g. 5%), additional security measures can be put in place to improve the level of security surrounding that voiceprint. For example, a business rule may be assigned to that customer requiring that a further piece of identification information be provided in the authentication session in order to verify the customer's identity being granted access to the secure service. Alternatively, the system may automatically pass the call to an operator to carry out further authentication checks on the customer.

Furthermore, where the third party system 104 establishes that a selected voiceprint is too weak to provide a suitable authentication result (e.g. by comparing the amount of deviation from the mean to a set threshold), the customer may be asked to re-enrol their voice sample.

In another embodiment, upon detecting a sufficiently weak voiceprint, the speaker model for that voiceprint may be re-built to improve the measure of confidence. With reference to the schematic of Figure 1b and flow chart of Figure 5, the process of re-building the speaker model will now be described. It will be understood that the various functions performed by the process are carried out by the performance and optimisation module 114 implemented by the third party system 104.

The process first involves re-building the world models (step 502). In an embodiment, this process involves selecting all voice files from the voice file database 109 and performing feature extraction on those voice files. Feature extraction involves applying an acoustic signal processing algorithm to extract the acoustic features of the voice file.

The extracted features are subsequently clustered into one or more groups (step 504), where each group shares one or more common features. For example, groups may be formed from according to speaker gender, input class (e.g. landline originating, mobile phone originating), etc. The grouping can be carried out either manually by selecting voice files that are known to share a common feature (e.g. by inspecting data provided by the customer during enrolment), or automatically using a clustering algorithm that groups all voice files which share common acoustic features.

At step 506, a model is built for each group, using techniques known to persons skilled in the art.

At the same time steps 502 to 506 are being carried out, a speaker selection process extracts the voice files associated with the weak voiceprint from the voice file database 109, performs a feature extraction on the extracted voice files and presents the features to a speaker model training process (steps 508 to 512). The speaker model training process uses the

parameters created during the world model creation process (step 506) as seed parameters which are then re-estimated using the relevant extracted feature(s), to re-build the voiceprint. Where more than one world model has been created at step 506, the world model having parameters which are closest to the relevant extracted parameters is used. The re-built model is then stored in the voiceprint database 109 in place of the weak voiceprint.

Figures 8, 9, 10, 11 and 12 show example screen shots of the system performing these steps. Figure 8 is a screen shot showing information derived by the system about speakers' voiceprints enrolled in the system and the IEER score for each speaker voiceprint derived using the testing method described herein.

Figure 9 shows a screen show of the system identifying weak voiceprints as determined by the process which, in this embodiment, are highlighted in brown. In this case voiceprints with IEER scores greater than 5% are considered weak and are therefore selected for optimisation as per any of the optimisation process described herein.

Figure 10 shows a screen shot of the system configuration consol which is operable to set the parameters for selection of voiceprints for optimisation. Figure 11 shows the screen shot of the optimisation process as reported by the system and the completion of the optimisation procedure. Figure 12 shows the EER performance of the overall system once the optimisation procedure has completed. In this case the EER performance for the overall system is 0.65% for speech item 1 (i.e. spoken account numbers) and 0.55% for speech item 8 (which are spoken phrases). This compares to an EER of 2-3% prior to the optimisation process.

It will be understood that any one or more of the above optimisation actions can be carried out each time a new voice

sample is enrolled with the system, or alternatively can be carried out on an ongoing basis until a performance threshold for the authentication system 104 has been met. In an embodiment, the performance of the third party system 104 may be measured by inspecting the overall EER score of the system.

SYSTEM CONFIGURATION

A more detailed explanation of the various modules implemented by the third party system 104 will now be described with reference to Figure 1b.

As mentioned in preceding paragraphs, the third party system 104 comprises a server 105 which functions not only to authenticate customers of the secure service, but in addition to determine measures of confidence for each enrolled voice sample (and, in embodiments, the overall system) and carry out appropriate optimisation actions. To perform this functionality, the server 105 comprises computer hardware including a processor, motherboard, random access memory, hard disk and a power supply. The server 105 also includes an operating system which co-operates with the hardware to provide an environment in which software applications can be executed. In this regard, the hard disk of the server 105 is loaded with voice authentication software, such as the Auraya voice authentication module which is available from Auraya Systems Pty Ltd, Australia. The hard disk is also loaded with an impostor testing module 116 which operates in conjunction with the voice authentication software to carry out the simulated impostor attacks, as herein before described. A performance evaluation and optimisation module 114 is also provided for calculating the confidence measures and implementing the various optimisation actions previously described. A graphics rendering application is also provided for displaying the results of the impostor testing and various confidence measures for each tested voice sample. An example screen shot generated by the graphics rendering application

showing the IEERs for each enrolled voice file is illustrated in Figure 6, 7, 8, 9, 10, 11 and 12.

The server 105 is also coupled to a voice file database 107, voiceprint database 109, identity management database 111 and confidence measure database 113. The hard disk of the server 105 also includes appropriate software and hardware for communicating with the secure service provider system 106. The communication may be made over any suitable communications link, such as an Internet connection, a wireless data connection or public network connection. In an embodiment, the voice samples provided for enrolment and authentication are initially logged with the secure service provider 106 and subsequently passed over the communications link to the third party system 104. Alternatively, the samples may be provided directly to the server 105 (in which case the server 105 would also implement a suitable call answering service).

The customer input device is a standard telephone including a transceiver and suitable for use with a public-switched telephone network.

As discussed, the communication system 108 is in the form of a public switched telephone network. However, in alternative embodiments the communications network may be a packet-switched network, such as the Internet. In such an embodiment customers may use a networked computing device to exchange data (more particularly, XML code and packetised voice messages) with the server 105 using a packet-switched network protocol, such as the TCP/IP protocol. Further details of such an embodiment are outlined in the international patent application PCT/AU2008/000070, the contents of which are incorporated herein by reference. In another alternative embodiment, the communication system may additionally comprise a third generation ("3G") or GPRS enabled mobile telephone network connected to the packet-switched network which can be utilised

to access the server 105. In such an embodiment, the customer input device 102 would include wireless capabilities for transmitting the voice message. The wireless computing devices may include, for example, mobile phones, personal computers having wireless cards and any other mobile communication device which facilitates voice recordal functionality. In another embodiment, the present invention may employ an 802.11 based wireless network or some other personal virtual network.

The other element in the system 100 is the secure service provider system 106 which, according to the embodiment described herein, is in the form of an Internet banking server. The secure service provider system 106 comprises a transceiver in the form of a network card for communicating with both the customers and third party system 104. The server also includes appropriate hardware and/or software for providing an answering service. In the illustrated embodiment, the secure service provider 106 communicates with the customers 102 over a public-switched telephone network 108 utilising the transceiver module.

The secure service provider system 106 may also include an ascertaining module for ascertaining measures of confidence from the third party system and an adjustment module which either adjusts the threshold settings within the secure service provider system 106 or instructs the third party system 104 of the appropriate settings for each type of transaction, dependent on the determined measure of confidence..

Although in embodiments described in preceding paragraphs the authentication system 104 was in the form of a "third party", or centralised system, it will be understood that the system need not be a third party system but instead may be incorporated into the secure service provider system.

Furthermore, it will be understood that any suitable measure of confidence may be associated with the individual voice

samples/voiceprints and need not be limited to the embodiment described herein. It will also be understood that the measure of confidence may be derived through mechanisms other than simulated impostor testing. For example, the individual scores utilised in deriving the measure may be ascertained through normal (i.e. not simulated) operation of the authentication system.

While the invention has been described with reference to the present embodiment, it will be understood by those skilled in the art that alterations, changes and improvements may be made and equivalents may be substituted for the elements thereof and steps thereof without departing from the scope of the invention. In addition, many modifications may be made to adapt the invention to a particular situation or material to the teachings of the invention without departing from the central scope thereof. Such alterations, changes, modifications and improvements, though not expressly described above, are nevertheless intended and implied to be within the scope and spirit of the invention. Therefore, it is intended that the invention not be limited to the particular embodiment described herein and will include all embodiments falling within the scope of the independent claims.

In the claims which follow and in the preceding description of the invention, except where the context requires otherwise due to express language or necessary implication, the word "comprise" or variations such as "comprises" or "comprising" is used in an inclusive sense, i.e. to specify the presence of the stated features but not to preclude the presence or addition of further features in various embodiments of the invention.

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A method for configuring a voice authentication system, the method comprising the steps of:

ascertaining a measure of confidence associated with a voice sample enrolled with the authentication system, the measure of confidence being derived through simulated impostor testing carried out on the enrolled sample.

2. A method in accordance with claim 1, comprising the further step of implementing an optimisation action for the enrolled voice sample based, at least in part, on the ascertained measure of confidence.

3. A method in accordance with claim 1 or claim 2, wherein the simulated impostor testing comprises utilising an authentication engine to compare at least one impostor voice sample against a voiceprint derived from the enrolled sample, to determine an individual false acceptance rate.

4. A method in accordance with claim 3, wherein the individual false acceptance rate is utilised to derive the means of confidence.

5. A method in accordance with any one of the preceding claims, comprising the further step of determining an individual false rejection rate for the enrolled sample, the individual false rejection rate being additionally utilised to derive the measure of confidence.

6. A method in accordance with claim 5, wherein the step of determining the individual false rejection rate comprises utilising an authentication engine to compare at least one legitimate sample from the same speaker who provided the enrolled sample, against a voiceprint derived from the enrolled sample.

7. A method in accordance with claim 5 or claim 6, wherein the individual false acceptance rate and false rejection rate are utilised to determine an individual equal error rate (IEER) associated with the enrolled voice sample.

8. A method in accordance with claim 7, comprising the further step of comparing the IEER with a reference setting to derive the measure of confidence.

9. A method in accordance with claim 8, where in the reference setting is at least one of a mean individual equal error rate for a plurality of other samples enrolled with the system or an average equal error rate for the authentication system.

10. A method in accordance with claim 9, wherein a weak measure of confidence is assigned to the enrolled voice sample responsive to determining that the IEER exceeds the mean IEER or average system EER, by a specified amount.

11. A method in accordance with claim 10 when dependent on claim 2, wherein, responsive to establishing that the enrolled voice sample is weak, carrying out the optimisation action of re-building a voiceprint associated with the enrolled voice sample to adjust a speaker and/or environmental characteristic associated with the voiceprint.

12. A method in accordance with claim 10 or claim 11 when dependent on claim 2, wherein, responsive to establishing that the enrolled voice sample is weak, carrying out the optimisation action of re-building a world model from which the associated voiceprint was derived.

13. A method in accordance with any one of claims 10 to 12 when dependent on claim 2 wherein, responsive to establishing

that the enrolled voice sample is weak, carrying out the optimisation action of re-building the voice-print.

14. A method in accordance with any one of the preceding claims when dependent on claim 2, wherein the optimisation action comprises setting a threshold associated with the enrolled sample, based on the derived measure of confidence.

15. A method in accordance with any one of the preceding claims when dependent on claim 2, wherein, upon determining that the measure of confidence does not meet a set threshold, the optimisation action comprises requesting that the voice sample be re-enrolled.

16. A method in accordance with claim 2, wherein the optimisation step is repeated each time a new voice sample is enrolled with the system.

17. A method in accordance with claim 2, wherein the optimisation action is carried out on selected enrolled voice samples until a threshold performance measure for the system has been met.

18. A method in accordance with claim 17, wherein the performance measure is associated with an overall equal error rate for the system.

19. A method in accordance with any one of the preceding claims, wherein the impostor samples have the same content type and/or speaker characteristic as the enrolled sample.

20. A method in accordance with claim 19, wherein the impostor samples are samples provided by other legitimate persons during enrolment with the system and/or during a subsequent authentication session.

21. A voice authentication system comprising:

an ascertaining module operable to ascertain a measure of confidence associated with a voice sample enrolled with the authentication system, the measure of confidence being derived through simulated impostor testing carried out on the enrolled sample by an impostor testing module.

22. A system in accordance with claim 21, further comprising an optimisation module operable to implement an optimisation action for the enrolled voice sample based, at least in part, on the ascertained measure of confidence.

23. A system in accordance with claim 20 or claim 21, wherein the impostor testing module comprises an authentication engine operable to compare at least one impostor voice sample against a voiceprint derived from the enrolled sample, to determine an individual false acceptance rate which is utilised to derive the measure of confidence.

24. A system in accordance with claim 23, wherein the authentication engine is further arranged to compare at least one legitimate sample against a voiceprint derived from the enrolled sample to determine an individual false rejection rate which is additionally utilised to derive the measure of confidence.

25. A system in accordance with claim 24, wherein the individual false acceptance rate and individual false rejection rate are utilised to establish an individual equal error rate (IEER) for the enrolled voice sample.

26. A system in accordance with claim 25, wherein the impostor testing module is operable to compare the IEER against a reference setting to determine whether a strength of the enrolled sample.

27. A system in accordance with claim 26, wherein the reference setting is a mean individual equal error rate for a plurality of other samples enrolled with the system or an overall equal error rate for the authentication system.

28. A system in accordance with claim 27, wherein a weak measure of confidence is assigned to the enrolled voice sample responsive to determining that the IEER exceeds the mean IEER or average system EER by some specified amount.

29. A system in accordance with claim 28, wherein, responsive to establishing that the enrolled voice sample is weak, the optimisation module re-builds a voiceprint associated with the enrolled voice sample to adjust a speaker and/or environmental characteristic of the voiceprint.

30. A system in accordance with claim 28 or claim 29, wherein the optimisation module re-builds a world model from which the associated voiceprint was derived, responsive to establishing that the enrolled voice sample is weak.

31. A system in accordance with any one of the preceding claims when dependent on claim 22, wherein the optimisation module sets an acceptance threshold associated with the enrolled sample, based on the derived measure of confidence.

32. A system in accordance with any one of the preceding claims 21 to 31, wherein the optimisation module requests that the voice sample be re-enrolled, upon determining that the measure of confidence does not meet a set threshold.

33. A system in accordance with claim 22, wherein the optimisation action is carried out each time a new voice sample is enrolled with the system.

34. A system in accordance with claim 22, wherein the

optimisation module continues to carry out optimisation actions until a threshold performance measure for the system has been met.

35. A system in accordance with claim 34, wherein the performance measure is associated with an overall equal error rate for the system.

36. A method for providing a secure service, comprising the steps of:

receiving data indicative of a measure of confidence associated with a user of the secure service, the measure of confidence being derived through simulated impostor testing carried out on an voice sample of the user; and

adjusting a level of authentication required by the user to access the secure service based, at least in part, on the measure of confidence.

37. A method in accordance with claim 36, wherein the level of authentication is adjusted by setting an acceptance threshold level.

38. A method in accordance with claim 36 or claim 37, wherein the simulated impostor testing is carried out using the methodology according to any one of claims 1 to 20.

39. A secure service provider system comprising:

a receiving module operable to receive data indicative of a measure of confidence associated with a user of the secure service, the measure of confidence being derived through simulated impostor testing carried out on a voice sample of the user; and

an adjustment module operable to adjust a level of authentication required by the user to access the secure service based, at least in part, on the measure of confidence.

40. A computer program comprising instructions for controlling a computer system to implement a method in accordance with any one of claims 1 to 20 or 36 to 38.

41. A computer readable medium comprising a computer program in accordance with claim 40.

42. A data signal providing a computer program in accordance with claim 40.

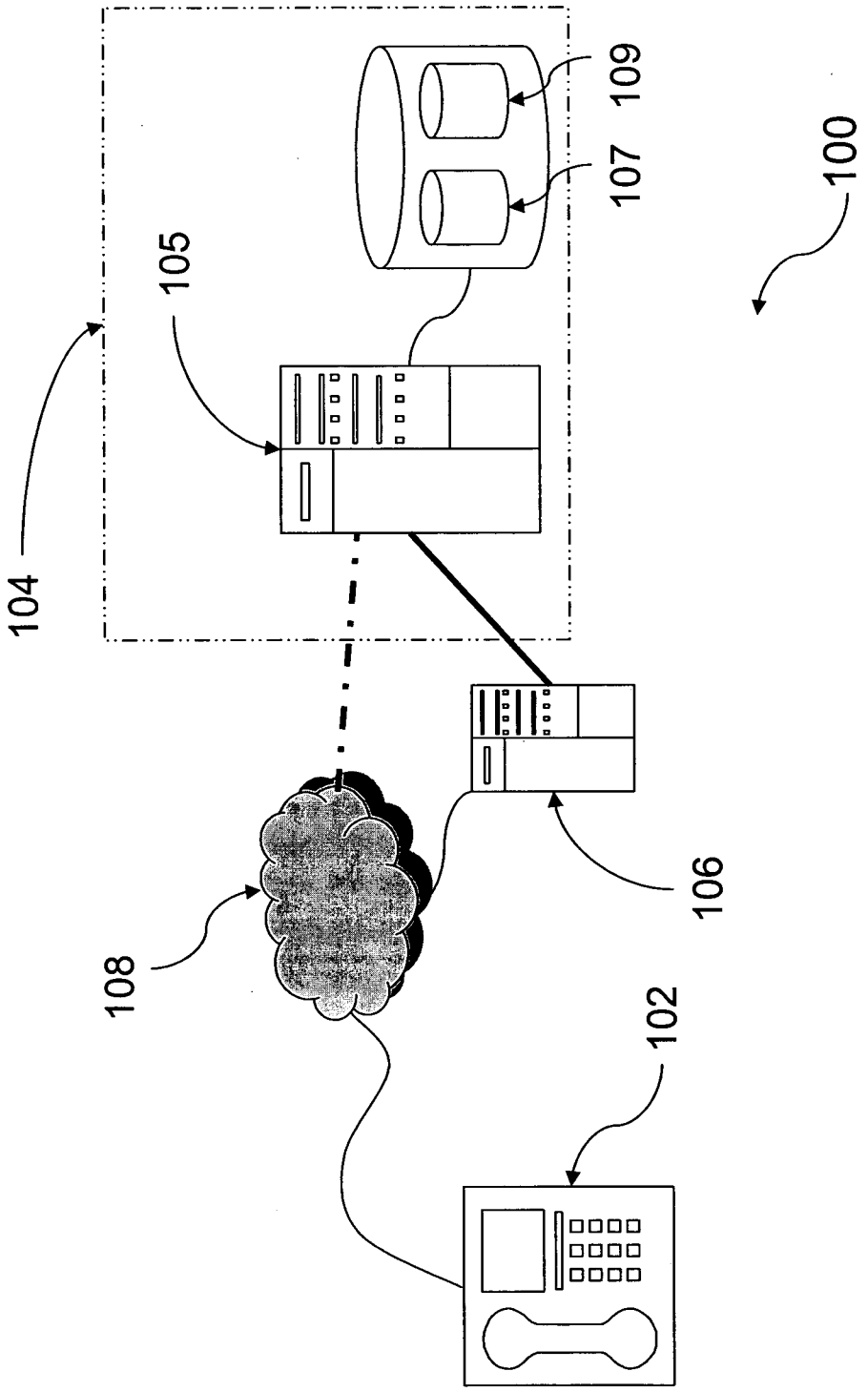

Fig. 1a

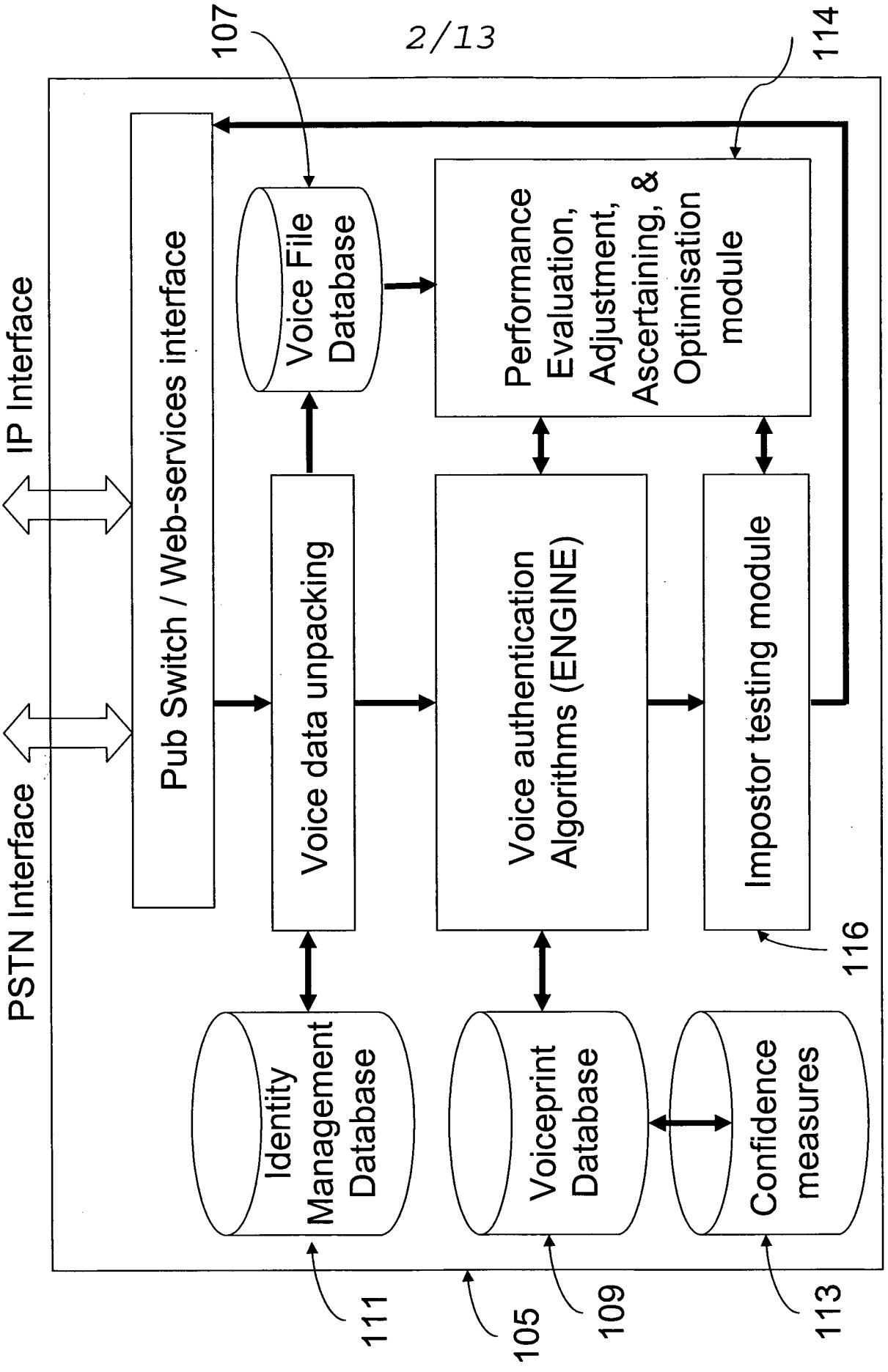
107
114
IP Interface
PSTN Interface
Pub Switch / Web-services interface
Voice File Database
Performance Evaluation, Adjustment, Ascertaining, & Optimisation module
Voice data unpacking
Voice authentication Algorithms (ENGINE)
Impostor testing module
116
Identity Management Database
Voiceprint Database
Confidence measures
111
105
109
113

Fig. 1b

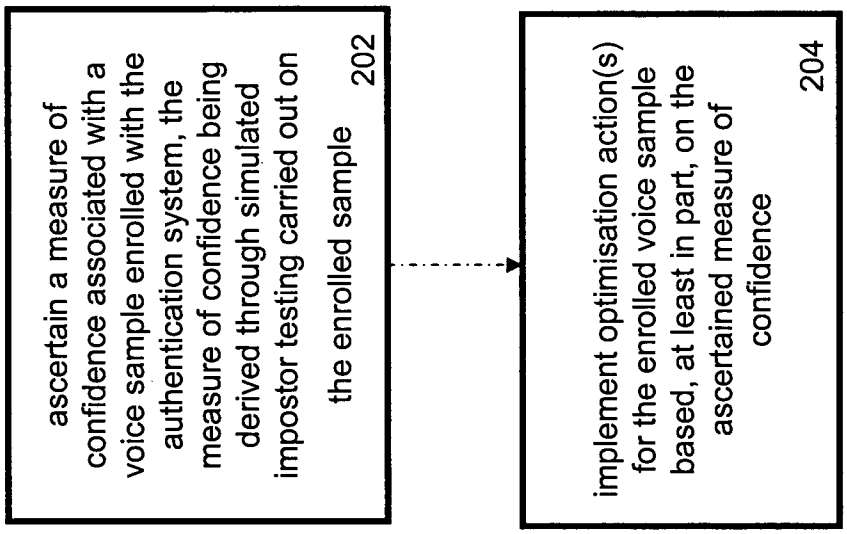
ascertain a measure of confidence associated with a voice sample enrolled with the authentication system, the measure of confidence being derived through simulated impostor testing carried out on the enrolled sample
202
implement optimisation action(s) for the enrolled voice sample based, at least in part, on the ascertained measure of confidence
204

Fig. 2

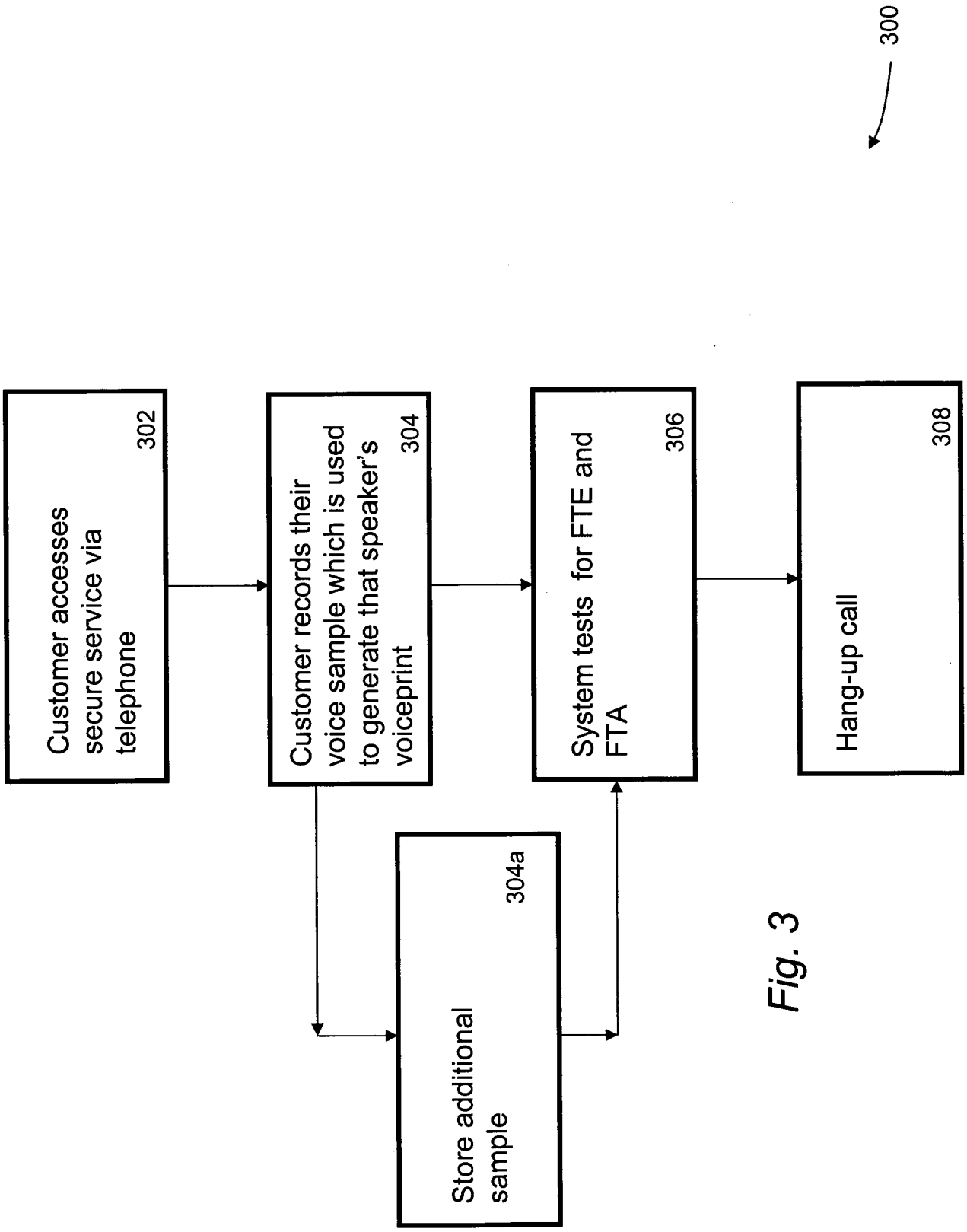
300
Customer accesses secure service via telephone
302
Customer records their voice sample which is used to generate that speaker's voiceprint
304
System tests for FTE and FTA
306
Hang-up call
308
Store additional sample
304a

Fig. 3

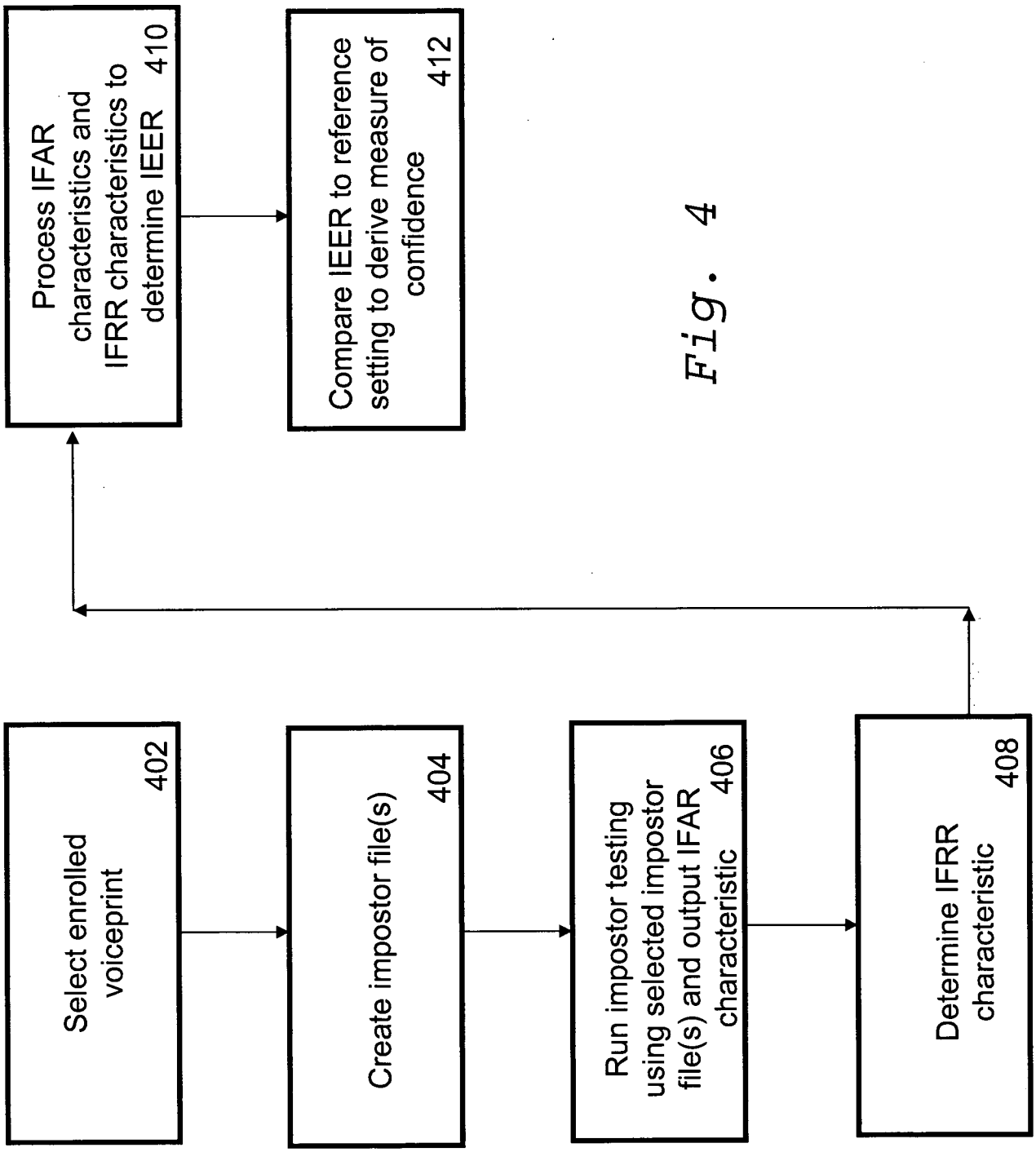
Select enrolled voiceprint 402
Create impostor file(s) 404
Run impostor testing using selected impostor file(s) and output IFAR characteristic 406
Determine IFRR characteristic 408
Process IFAR characteristics and IFRR characteristics to determine IEER 410
Compare IEER to reference setting to derive measure of confidence 412

Fig. 4

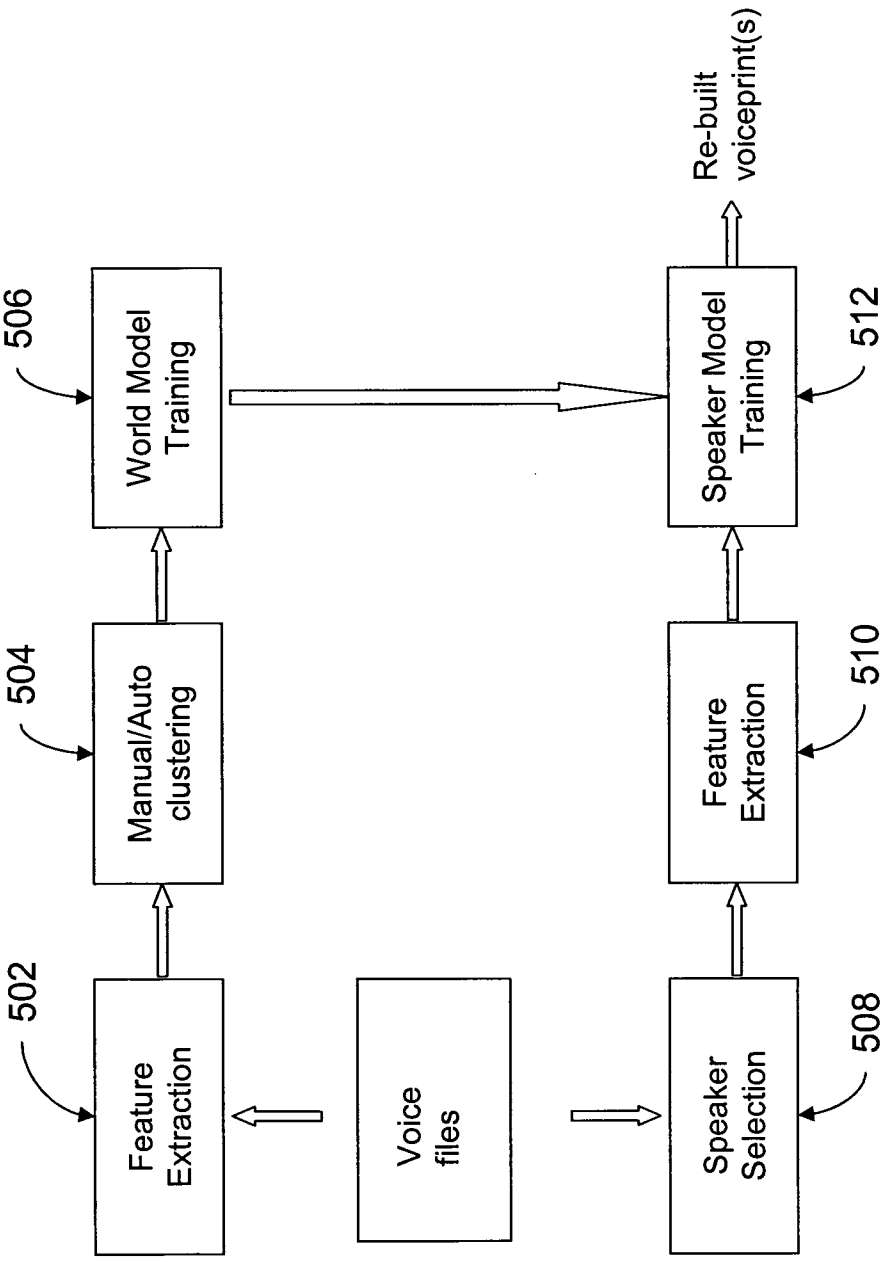
Re-built voiceprint(s)
506
World Model Training
Speaker Model Training
512
504
Manual/Auto clustering
Feature Extraction
510
502
Feature Extraction
Voice files
Speaker Selection
508

Fig. 5

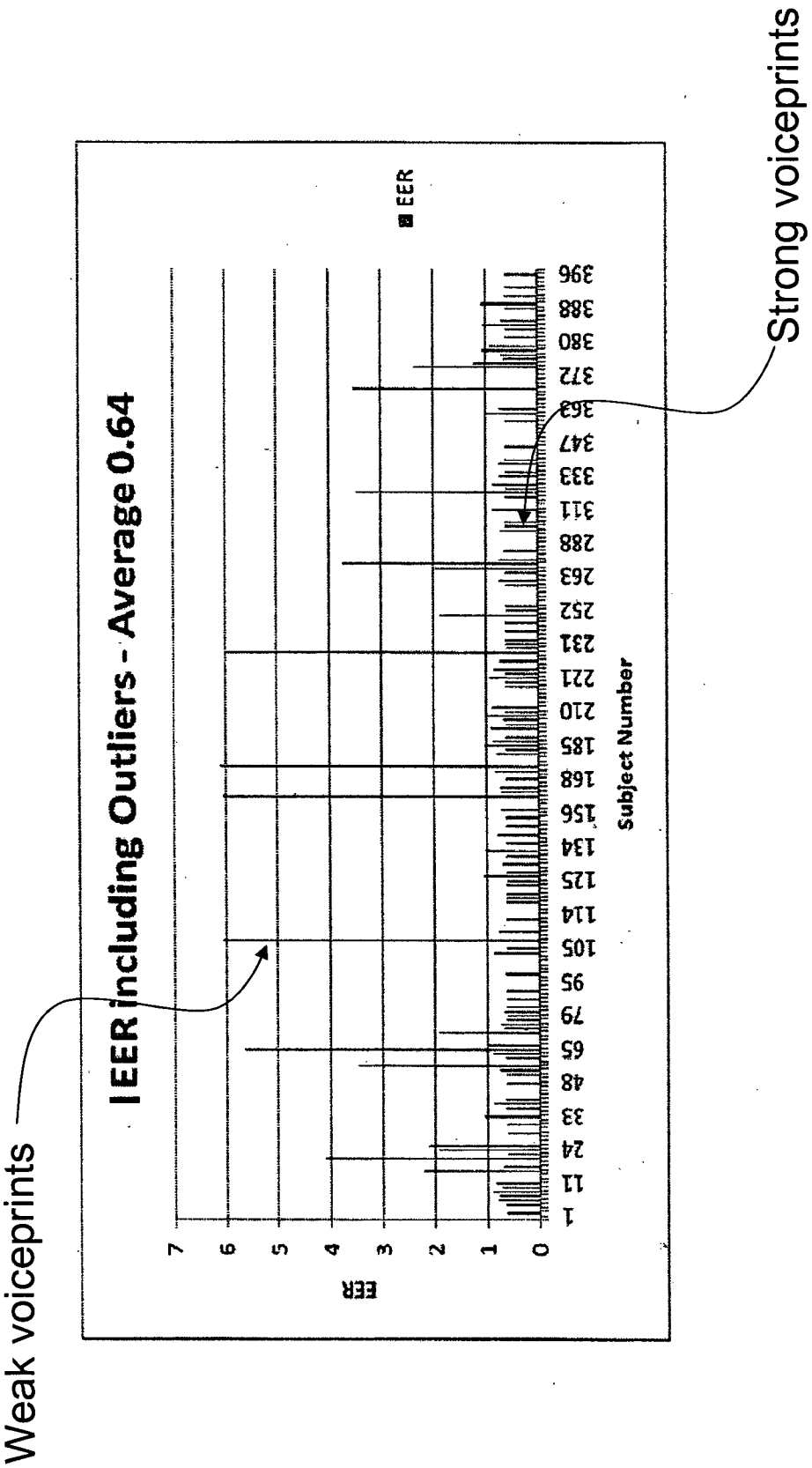
Strong voiceprints
Weak voiceprints
EER including Outliers - Average 0.64
EER
Subject Number
EER
0
1
2
3
4
5
6
7
1
11
24
33
48
65
79
95
105
114
125
134
156
168
185
210
221
231
252
263
288
311
333
347
363
372
380
388
396

Fig. 6

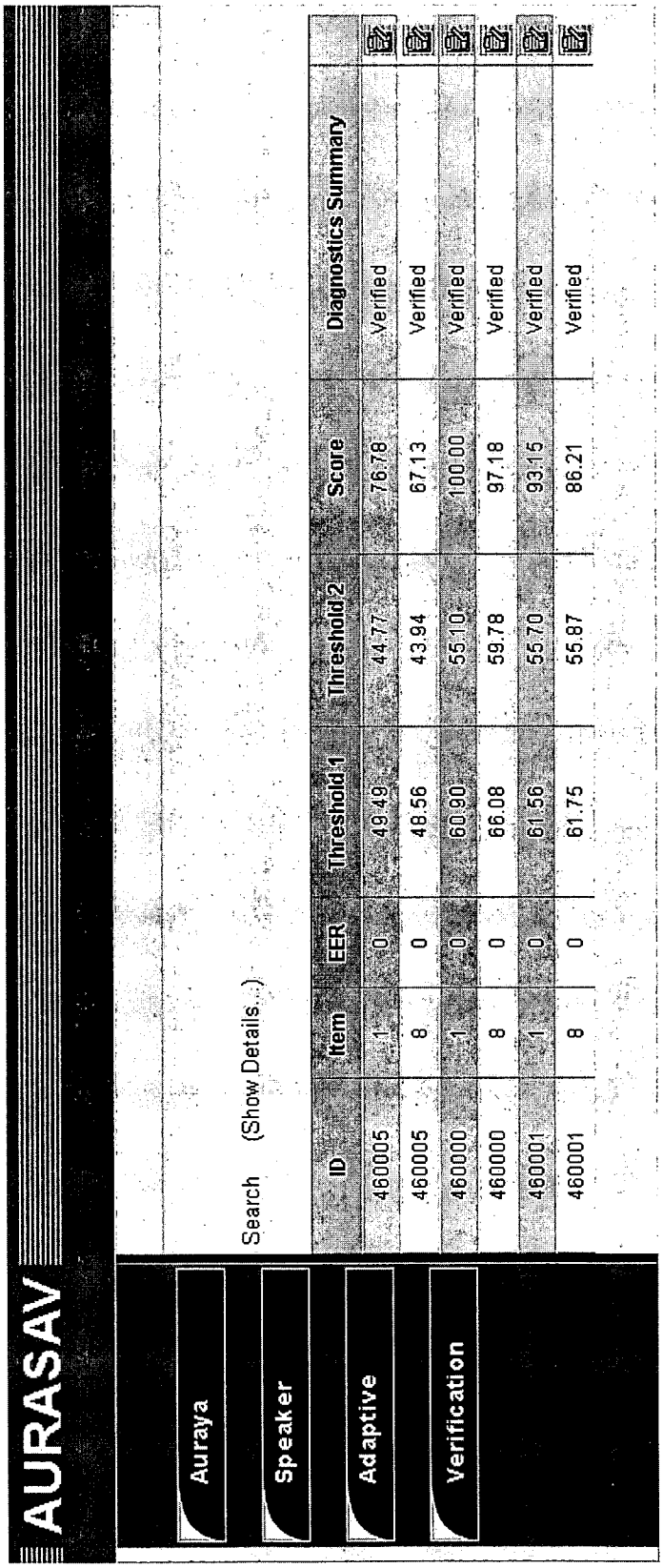
AURASAV
Auraya
Speaker
Adaptive
Verification
Search (Show Details...)
ID
Item
EER
Threshold 1
Threshold 2
Score
Diagnostics Summary
460005 1 0 49.49 44.77 76.78 Verified
460005 8 0 48.56 43.94 67.13 Verified
460000 1 0 60.90 55.10 100.00 Verified
460000 8 0 66.08 59.78 97.18 Verified
460001 1 0 61.56 55.70 93.15 Verified
460001 8 0 61.75 55.87 86.21 Verified

Fig. 7

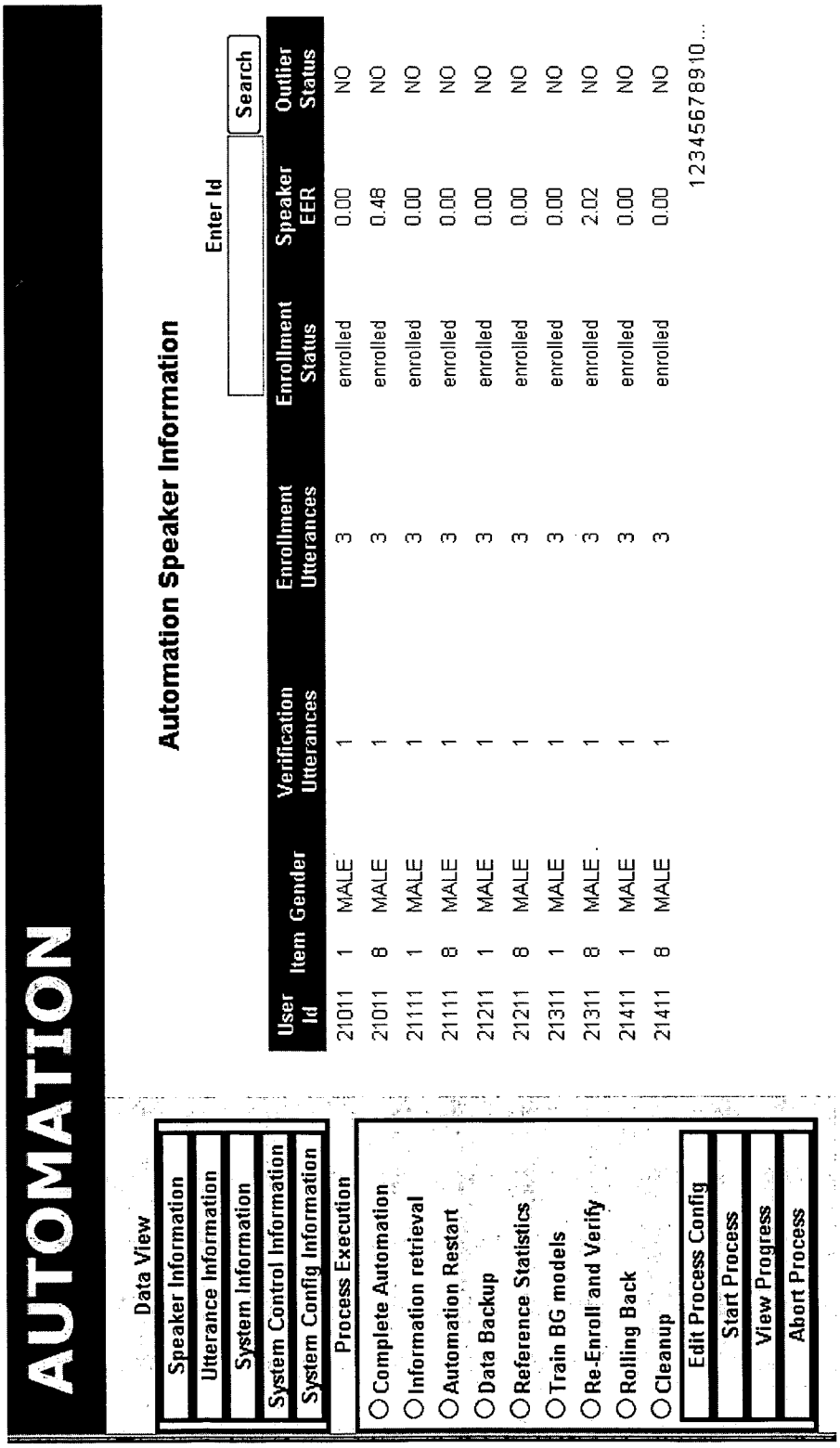
AUTOMATION
Data View
Speaker Information
Utterance Information
System Information
System Control Information
System Config Information
Process Execution
Complete Automation
Information retrieval
Automation Restart
Data Backup
Reference Statistics
Train BG models
Re-Enroll and Verify
Rolling Back
Cleanup
Edit Process Config
Start Process
View Progress
Abort Process
Automation Speaker Information
Enter Id
Search
User Id
Item
Gender
Verification Utterances
Enrollment Utterances
Enrollment Status
Speaker EER
Outlier Status
21011 1 MALE 1 3 enrolled 0.00 NO
21011 8 MALE 1 3 enrolled 0.48 NO
21111 1 MALE 1 3 enrolled 0.00 NO
21111 8 MALE 1 3 enrolled 0.00 NO
21211 1 MALE 1 3 enrolled 0.00 NO
21211 8 MALE 1 3 enrolled 0.00 NO
21311 1 MALE 1 3 enrolled 0.00 NO
21311 8 MALE 1 3 enrolled 2.02 NO
21411 1 MALE 1 3 enrolled 0.00 NO
21411 8 MALE 1 3 enrolled 0.00 NO
1 2 3 4 5 6 7 8 9 10 ...

Fig. 8

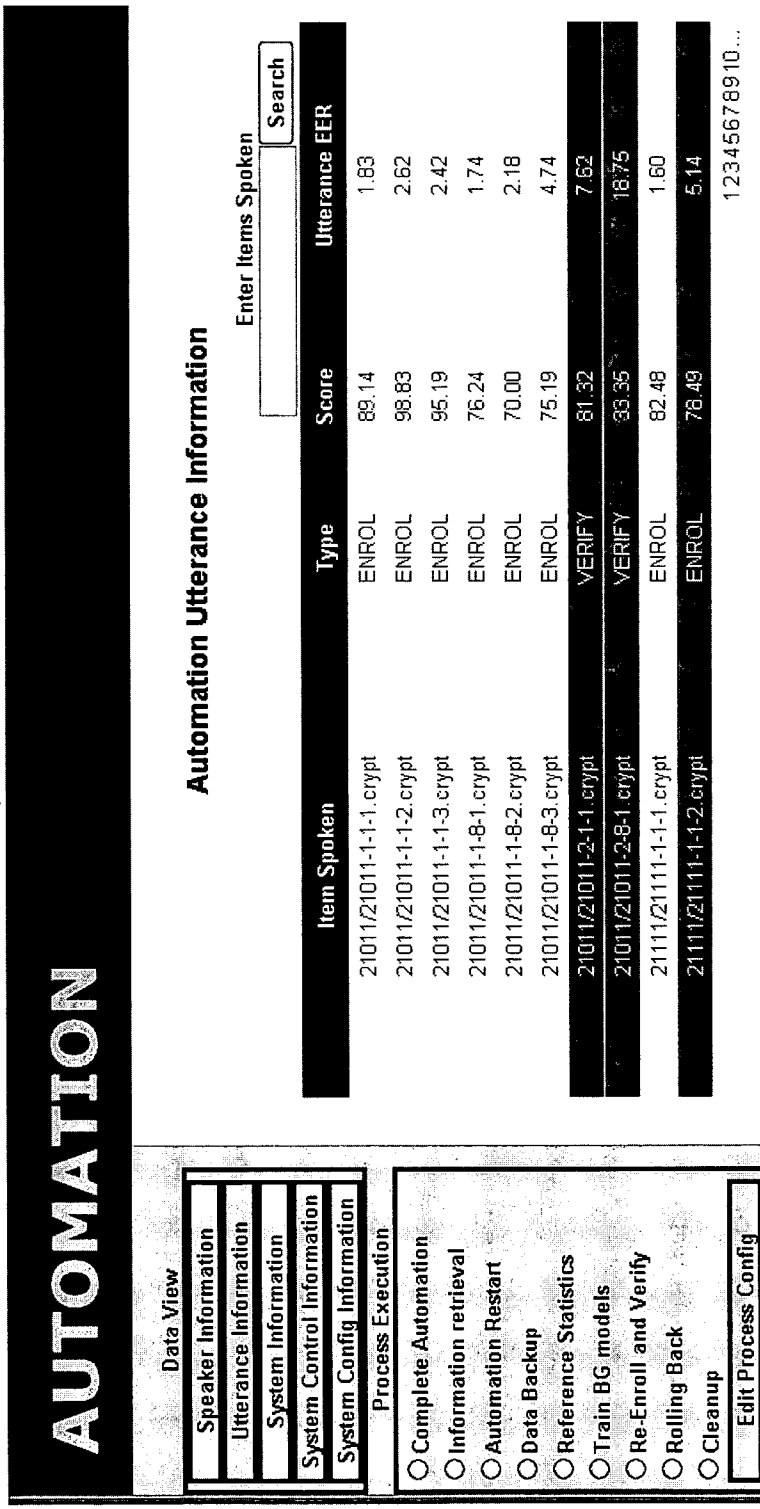
AUTOMATION
Data View
Speaker Information
Utterance Information
System Information
System Control Information
System Config Information
Process Execution
Complete Automation
Information retrieval
Automation Restart
Data Backup
Reference Statistics
Train BG models
Re-Enroll and Verify
Rolling Back
Cleanup
Edit Process Config
Start Process
View Progress
Abort Process
Automation Utterance Information
Enter Items Spoken
Search
Item Spoken
Type
Score
Utterance EER
21011/21011-1-1-1.crypt ENROL 89.14 1.83
21011/21011-1-1-2.crypt ENROL 98.83 2.62
21011/21011-1-1-3.crypt ENROL 95.19 2.42
21011/21011-1-8-1.crypt ENROL 76.24 1.74
21011/21011-1-8-2.crypt ENROL 70.00 2.18
21011/21011-1-8-3.crypt ENROL 75.19 4.74
21011/21011-2-1-1.crypt VERIFY 81.32 7.62
21011/21011-2-8-1.crypt VERIFY 93.35 18.75
21111/21111-1-1-1.crypt ENROL 82.48 1.60
21111/21111-1-1-2.crypt ENROL 78.49 5.14
1 2 3 4 5 6 7 8 9 10 ...

Fig. 9

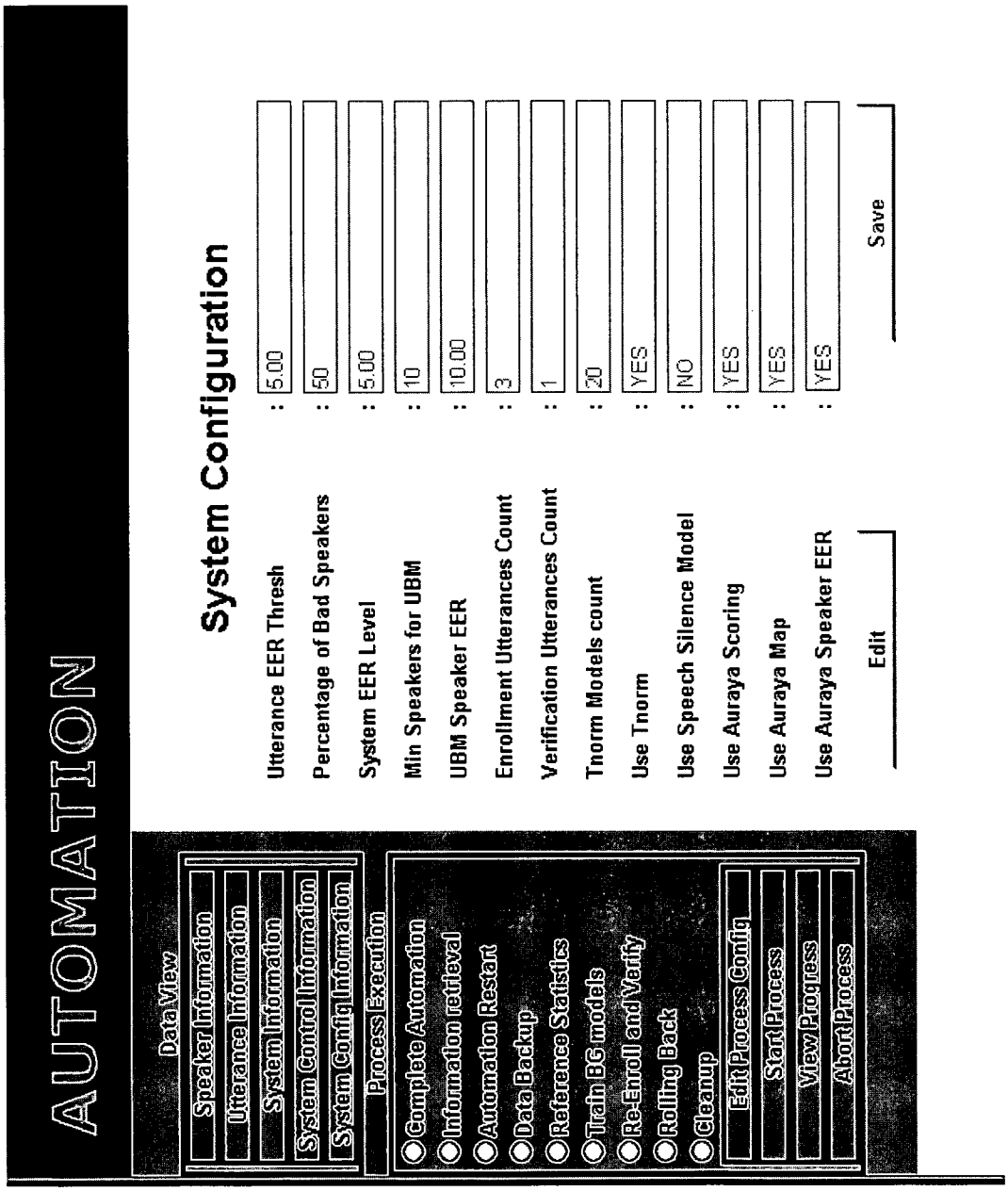
AUTOMATION
Data View
Speaker Information
Utterance Information
System Information
System Control Information
System Config Information
Process Execution
Complete Automation
Information retrieval
Automation Restart
Data Backup
Reference Statistics
Train BG models
Re-Enroll and Verify
Rolling Back
Cleanup
Edit Process Config
Start Process
View Progress
Abort Process
System Configuration
Utterance EER Thresh : 5.00
Percentage of Bad Speakers : 50
System EER Level : 5.00
Min Speakers for UBM : 10
UBM Speaker EER : 10.00
Enrollment Utterances Count : 3
Verification Utterances Count : 1
Tnorm Models count : 20
Use Tnorm : YES
Use Speech Silence Model : NO
Use Auraya Scoring : YES
Use Auraya Map : YES
Use Auraya Speaker EER : YES
Edit
Save

Fig. 10

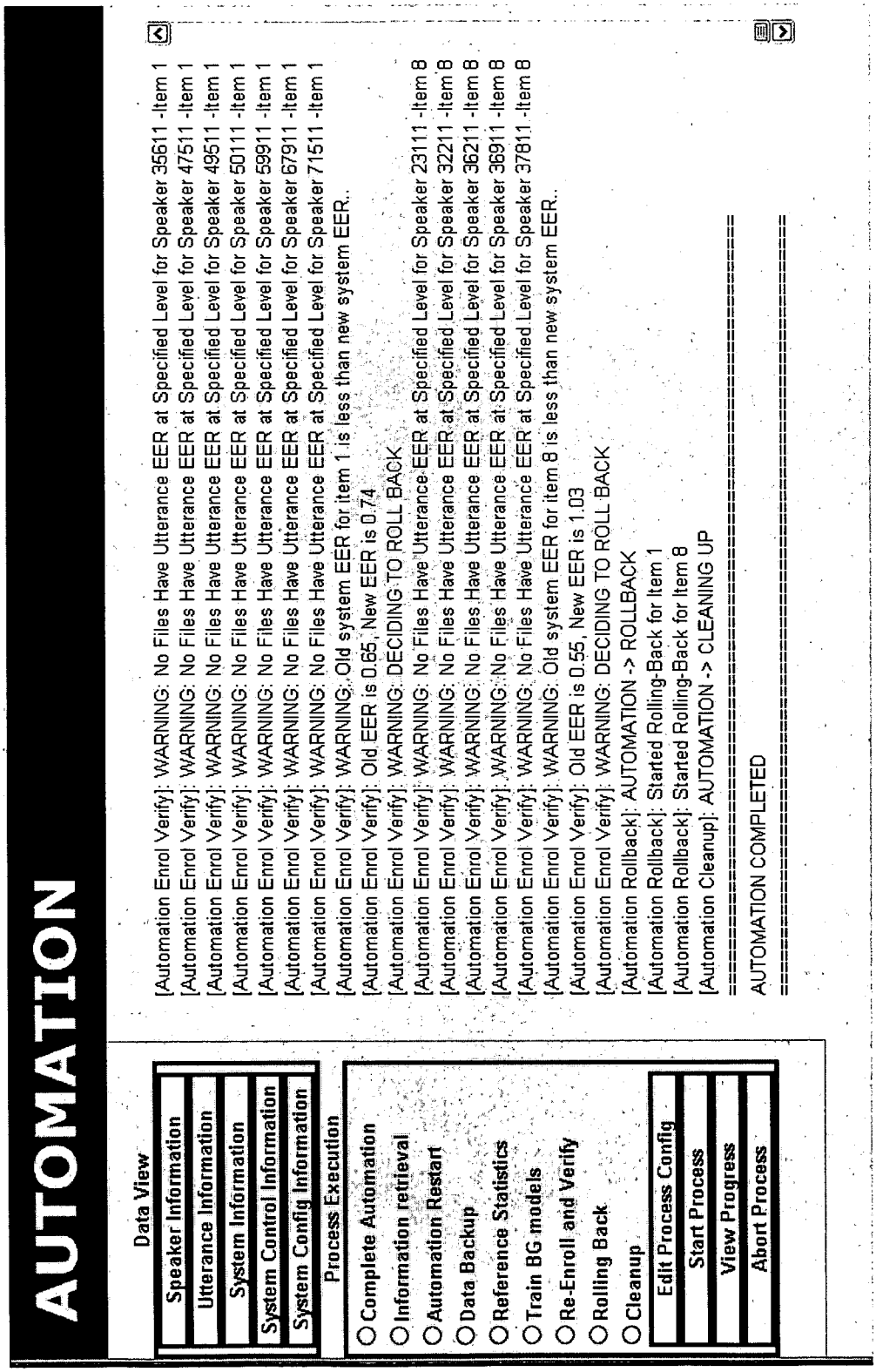
AUTOMATION
Data View
Speaker Information
Utterance Information
System Information
System Control Information
System Config Information
Process Execution
Complete Automation
Information retrieval
Automation Restart
Data Backup
Reference Statistics
Train BG models
Re-Enroll and Verify
Rolling Back
Cleanup
Edit Process Config
Start Process
View Progress
Abort Process
[Automation Enrol Verify]: WARNING: No Files Have Utterance EER at Specified Level for Speaker 35611 -Item 1
[Automation Enrol Verify]: WARNING: No Files Have Utterance EER at Specified Level for Speaker 47511 -Item 1
[Automation Enrol Verify]: WARNING: No Files Have Utterance EER at Specified Level for Speaker 49511 -Item 1
[Automation Enrol Verify]: WARNING: No Files Have Utterance EER at Specified Level for Speaker 50111 -Item 1
[Automation Enrol Verify]: WARNING: No Files Have Utterance EER at Specified Level for Speaker 59911 -Item 1
[Automation Enrol Verify]: WARNING: No Files Have Utterance EER at Specified Level for Speaker 67911 -Item 1
[Automation Enrol Verify]: WARNING: No Files Have Utterance EER at Specified Level for Speaker 71511 -Item 1
[Automation Enrol Verify]: WARNING: Old system EER for item 1 is less than new system EER..
[Automation Enrol Verify]: Old EER is 0.65, New EER is 0.74
[Automation Enrol Verify]: WARNING: DECIDING TO ROLL BACK
[Automation Enrol Verify]: WARNING: No Files Have Utterance EER at Specified Level for Speaker 23111 -Item 8
[Automation Enrol Verify]: WARNING: No Files Have Utterance EER at Specified Level for Speaker 32211 -Item 8
[Automation Enrol Verify]: WARNING: No Files Have Utterance EER at Specified Level for Speaker 36211 -Item 8
[Automation Enrol Verify]: WARNING: No Files Have Utterance EER at Specified Level for Speaker 36911 -Item 8
[Automation Enrol Verify]: WARNING: No Files Have Utterance EER at Specified Level for Speaker 37811 -Item 8
[Automation Enrol Verify]: WARNING: Old system EER for item 8 is less than new system EER..
[Automation Enrol Verify]: Old EER is 0.55, New EER is 1.03
[Automation Enrol Verify]: WARNING: DECIDING TO ROLL BACK
[Automation Rollback]: AUTOMATION -> ROLLBACK
[Automation Rollback]: Started Rolling-Back for Item 1
[Automation Rollback]: Started Rolling-Back for Item 8
[Automation Cleanup]: AUTOMATION -> CLEANING UP
AUTOMATION COMPLETED

Fig. 11

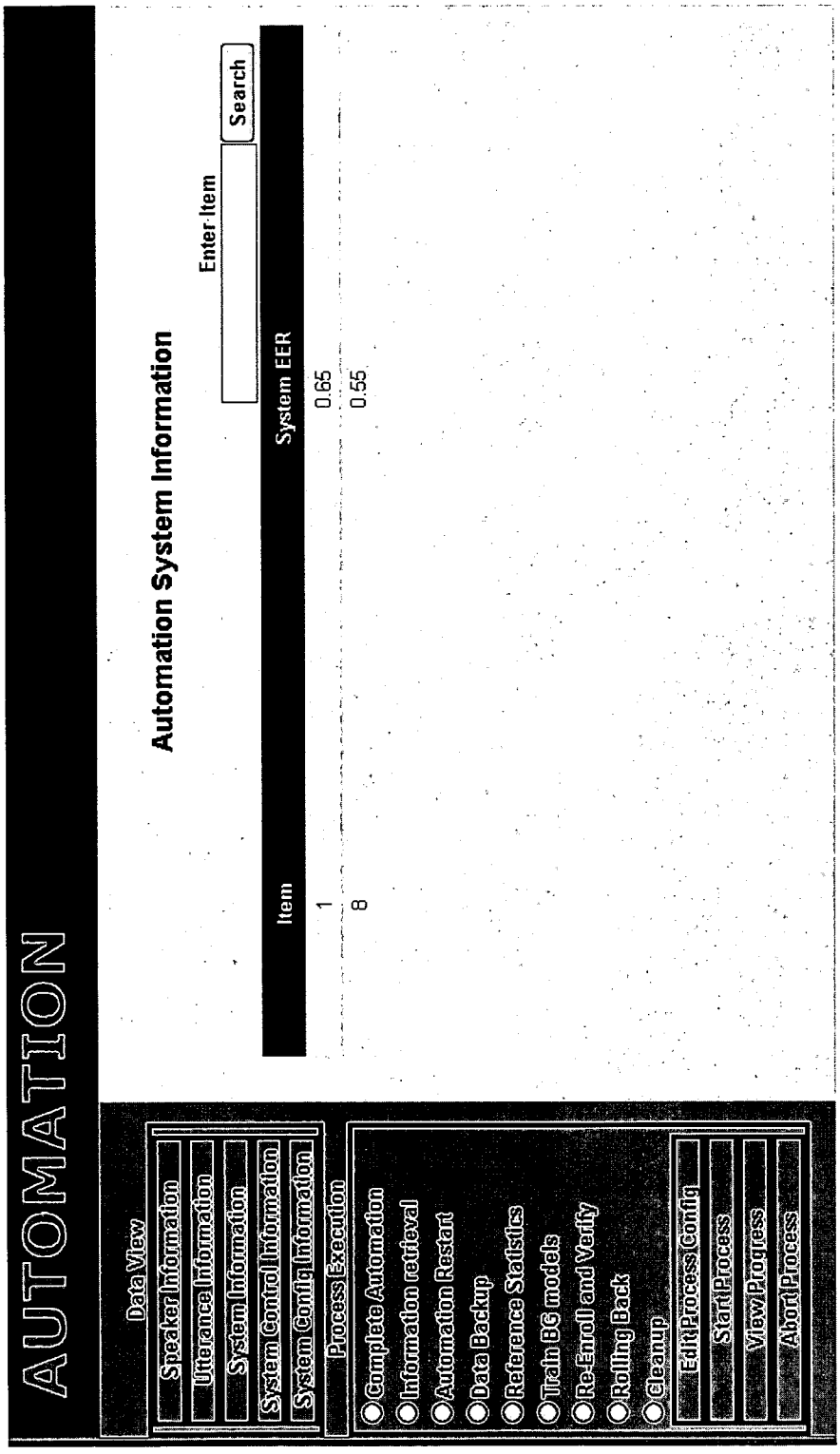
AUTOMATION
Data View
Speaker Information
Utterance Information
System Information
System Control Information
System Config Information
Process Execution
Complete Automation
Information retrieval
Automation Restart
Data Backup
Reference Statistics
Train BG models
Re-Enroll and Verify
Rolling Back
Cleanup
Edit Process Config
Start Process
View Progress
Abort Process
Automation System Information
Enter Item
Search
Item
System EER
1
0.65
8
0.55

Fig. 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2009/001165

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

G10L 17/00 (2006.01) *G06F 21/00* (2006.01)

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)

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)
Epodoc, WPI, Google Patents, Es@cenet, Patent Lens: keywords: (voice, authentication, voiceprint, threshold) and similar terms

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 4761807 (MATTHEWS et al.) 2 August 1988 column 24 lines 53-66; column 25 line 42-column 26 line 51	1-2, 14-17, 21-22, 31-34, 36, 38-42
Y	column 24 lines 53-66; column 25 line 42-column 26 line 51	18, 35
X	US 2007/0038868 A1 (YU et al.) 15 February 2007 para. [0013-0015]	1, 21, 36, 38-42
X	WO 2007/098039 A1 (MICROSOFT CORPORATION) 30 August 2007 Page 20 line 7-page 21 line 20	1, 21, 36, 38-42

[X] Further documents are listed in the continuation of Box C [X] See patent family annex

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but 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)

"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

"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

"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

"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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
12 November 2009	25 NOV 2009
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustralia.gov.au Facsimile No. +61 2 6283 7999	Authorized officer **SURYA PRAKASH** AUSTRALIAN PATENT OFFICE (ISO 9001 Quality Certified Service) Telephone No : (02) 6283 2101

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2009/001165

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2005/0041783 A1 (TIMMINS et al.) 24 February 2005 para. [0094-0095]	1, 21, 36, 38-42
X	Campbell Jr, J.P. 'Speaker Recognition: A Tutorial.' In : Proceeding of the IEEE, September 1997, vol. 85, no. 9, pages 1437--1462. page 1455 column 1 lines 1-55; page 1456 column 2 lines 11-28	1, 3-10, 19-21, 23-28, 36, 37-42
Y	page 1455 column 1 lines 1-55; page 1456 column 2 lines 11-28	11-13, 29-30
X	Reynolds, D.A. 'An Overview of Automatic Speaker Recognition Technology.' In : IEEE International Conference on Acoustics, Speech and Signal Processing, 2002, vol. 4, pages IV-4072--IV-4075. page iv-4073 column 2 line 52-page iv-4074 column 1 line 10; page iv-4074 column 1 lines 45-49	1-2, 21-22, 36, 38-42
Y	page iv-4073 column 2 line 52-page iv-4074 column 1 line 10; page IV-4074 column 2 lines 44-57	11-13, 18, 29-30, 35
X	WO 2003/098373 A2 (DOMAIN DYNAMICS LIMITED) 27 November 2003. page 39 line 20-page 42 line 2; page 45 line 7-page 46 line 9	1-7,14-17, 19-25, 31-34, 36, 37-42
NOTE:	• Claims 11-13, 29-30 do not involve an inventive step in light of disclosures of **Campbell** when combined with **Reynolds**. • Claims 18, 35 do not involve an inventive step in light of disclosures of **US 4761807** when combined with **Reynolds**.	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/AU2009/001165

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member					
US	4761807	AU	19598/83	CA	1157551	CA	1213026
		EP	0029938	EP	0087849	EP	0106575
		EP	0336524	EP	0341800	ES	8607655
		JP	59134958	JP	59134959	US	4371752
		US	4580012	US	4581486	US	4585906
		US	4602129	US	4640991	US	4652700
		US	4757525				
US	2007038868	NONE					
WO	2007098039	AU	2007217884	CA	2643481	CN	101385074
		EP	1989701	KR	20080102373	MX	2008010478
		NO	20083580	US	7539616	US	2007198257
US	2005041783	AU	15983/01	AU	2002227182	CA	2390063
		CA	2431154	CA	2450116	CA	2453499
		CA	2453501	CA	2456667	CA	2504295
		CA	2504738	CA	2520879	EP	1362315
		US	6870921	US	6944279	US	6985569
		US	7388950	US	7466805	US	7499537
		US	2002055351	US	2003026405	US	2003119492
		US	2004029567	US	2004058710	US	2004096043
		US	2004127193	US	2004170260	US	2004170261
		US	2004190688	US	2004223593	US	2004247092
		US	2004258231	US	2004259535	US	2005002501
		US	2005002507	US	2005002508	US	2005002509
		US	2005002510	US	2005041784	US	2005058262
		US	2006018441	US	2006141982	US	2007121882
		US	2009110178	US	2009110179	US	2009156178
		US	2009252304	WO	200135621	WO	2002052368
		WO	2004095811				
WO	2003098373	AU	2003230039	GB	2388947		

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

END OF ANNEX