

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
19 February 2004 (19.02.2004)

PCT

(10) International Publication Number
WO 2004/015552 A2

(51) International Patent Classification⁷: **G06F 1/00**

3US (GB). **NEWBY, Ian** [GB/GB]; 20 Trevelyan, Bracknell, Berkshire RG12 8YD (GB). **ANDERSON, Alexander** [GB/GB]; Copperfield, West Street, Odiham, Hampshire RG29 1NR (GB).

(21) International Application Number:
PCT/GB2003/003509

(22) International Filing Date: 11 August 2003 (11.08.2003)

(74) Agents: **PIOTROWICZ, Pawel** et al.; Venner, Shipley & Co., 20 Little Britain, London EC1A 7DH (GB).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0218706.0 12 August 2002 (12.08.2002) GB
0222736.1 1 October 2002 (01.10.2002) GB

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(71) Applicant (*for all designated States except US*): **DOMAIN DYNAMICS LIMITED** [GB/GB]; Thames Tower, Station Road, Reading, Oxfordshire RG1 1LX (GB).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **GREEN, Samuel** [GB/GB]; 28 Inkpen Road, Kintbury, Hungerford, Berkshire RG17 9TX (GB). **GRIFFITH, Derek** [GB/GB]; 73a Gresham Road, Staines, Middlesex TW18 2BD (GB). **RODRIGUES, Eduardo** [GB/GB]; North View Farmhouse, 42 Old Village Road, Little Weighton, Yorkshire HU20

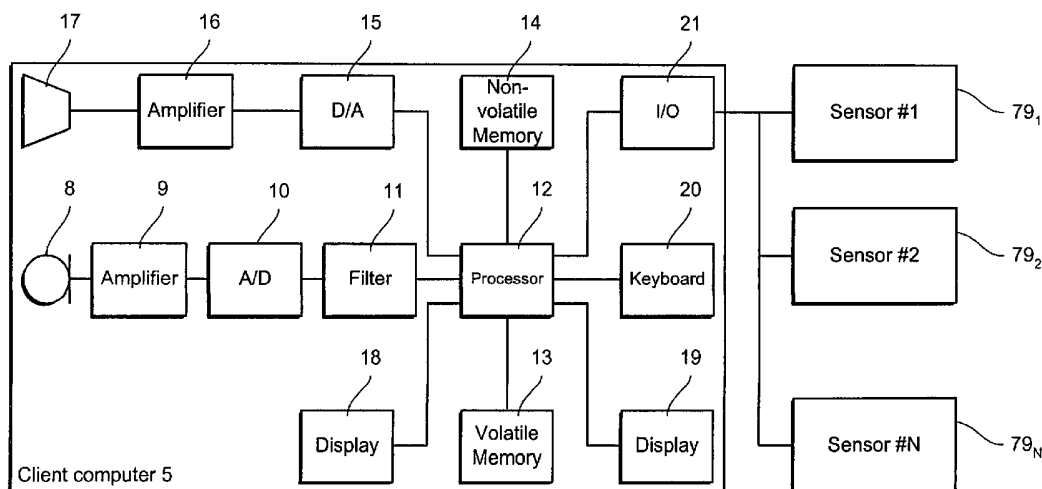
(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: METHOD OF AUTHENTICATION



(57) Abstract: A method of authentication for use over the Internet (7) includes downloading a Java applet (33) from a web server (31) running on a server computer (6) to a web-browser (30) running on a client computer (5). The applet (33) enables the client computer (5) to collect biometric data such as spoken responses or fingerprints, to characterise the biometric data and to send characteristic data (35) to the authentication server (32) for further processing.



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Method of authentication

Field of the invention

The present invention relates to a method of authentication.

5

Background art

In many Internet applications, such as e-commerce, it is desirable to validate the identity of a user. For example, if a user accesses an on-line bank account, the user is usually asked to supply confidential information, such as identification numbers, passwords or personal details. The information is checked against information
10 previously supplied by or given to the user when the account is opened and which is held by the bank. If there is a match between the information currently provided by the user and stored by the bank, then the user's identity is validated. However, this approach relies on the information being known only to a limited group, preferably
15 just the user and the bank.

Authentication using spoken response, fingerprint, retinal configuration, iris pattern and other biometric data, may provide an alternative or addition to conventional methods of validating a user. Biometric authentication has the advantage that it is
20 not simply based on confidential information. However, biometric authentication has several drawbacks. Firstly, even using known compression techniques, the amount of data transmitted through the Internet can be unacceptably high. Secondly, even using current error detection and correction algorithms, the integrity of the information may be lost particularly when transmitting over large distances.
25 Thirdly, the computational power required to validate many users simultaneously is prohibitively high.

The present invention seeks to help overcome these disadvantages and to provide a method of authentication for use over a network, such as the Internet.

30

Summary of the invention

According to a first aspect of the present invention there is provided a method of authentication for use over a network, the method comprising transmitting a

computer program from a server computer to a client computer and executing the computer program at the client computer, the client computer thereafter requesting a user to provide biometric data, obtaining a recorded signal based on the biometric data, deriving characteristic data from the recorded signal for characterising the biometric data and transmitting the characteristic data from the client computer to a server computer.

This has several advantages. It helps to share processing between the client and server computers. It can also reduce the amount of information transmitted to the server computer.

The biometric data may include spoken response, fingerprint, handprint, face pattern, scent, DNA, iris pattern, retinal configuration, handwriting, voice or acoustic signature.

Requesting the user to provide biometric data may comprise requesting the user to provide a response. Obtaining the recorded signal based on the biometric data may comprise obtaining a recorded signal including a recorded signal portion corresponding to the response. Deriving characteristic data from the recorded signal for characterising the biometric data may comprise deriving a set of feature data for characterising the recorded signal portion. Transmitting the characteristic data from the client computer to a server computer may comprise transmitting the set of feature data from the client computer to a server computer. The method may further comprise the client computer calibrating an input device and setting a signal level of the recorded signal. The method may further comprise the client computer determining an endpoint of the recorded signal.

Obtaining the recorded signal may comprise capturing generatable or transient biometric data. Generatable biometric data means biometric data which is capable of being generated by the user, such as a spoken response or handwritten response, and which does not already exist. The transient biometric data may be a spoken response.

Requesting the user to provide biometric data may comprise requesting the user to provide a spoken response to a prompt. Obtaining the recorded signal based on the biometric data may comprise obtaining a recorded signal including a recorded signal portion corresponding to the spoken response. Deriving characteristic data from the recorded signal for characterising the biometric data may comprise deriving a set of feature vectors for characterising the recorded signal portion. Transmitting the characteristic data from the client computer to a server computer transmitting the set of feature vectors from the client computer to a server computer. The method may further comprise the client computer calibrating a microphone and an amplifier for setting a signal level of the recorded signal. The method may further comprise the client computer determining an endpoint of the recorded signal.

Obtaining the recorded signal may comprise capturing permanent biometric data. "Permanent" means not substantially changing over a period of time during which a user may need authenticating. Usually, biometric data is considered to be permanent if does not change substantially over a period of years or tens of years. Nevertheless, biometric data which changes substantially over a relatively long period, such as months or years, may still be considered permanent over a relatively short period, for example days or weeks, if the relatively short period is at least as long as the period from enrolment until final expected potential authentication. Permanent biometric data may be handwriting.

Requesting the user to provide biometric data may comprise requesting the user to provide a written response to a prompt. Obtaining the recorded signal based on the biometric data may comprise obtaining a recorded signal including a recorded signal portion corresponding to the written response.

Obtaining the recorded signal may comprise reading permanent biometric data.

Requesting the user to provide biometric data may comprise requesting the user to submit at least a body portion for sensing by a biometric sensor. Reading the biometric data may comprise capturing an image. Reading the biometric data may comprise recording a pattern or a configuration. Obtaining the recorded signal

based on the biometric data comprises recording a representation of the biometric data. Obtaining the recorded signal based on the biometric data may comprise taking a fingerprint or taking a chemical sample for example by sampling scent.

5 The method may comprise the client computer requesting the computer program from the server computer. The method may comprise the client computer dynamically downloading the computer program from the server computer. The method may comprise the client computer accessing a web page provided by the server computer and requesting the computer program from the server computer
10 without prompt by the user. Executing the computer program may occur substantially immediately after the computer program is transmitted from a server computer to a client computer. "Substantially immediately" means within a few seconds.

15 The method may comprise the client computer requesting the user to provide further biometric data, obtaining recorded signals for respective biometric data, deriving respective characteristic data and transmitting the characteristic data from the client computer to the server computer. The method may further comprise the server computer combining characteristic data so as to provide archetype
20 characteristic data. The method may further comprise the server computer comparing characteristic data with archetype characteristic data so as to determine a score dependent upon a degree of matching.

According to a second aspect of the present invention there is provided a method
25 of operating a server computer comprising receiving a request from a client computer, transmitting a computer program to the client computer, the computer program when executed by a computer causing the computer to request a user to provide biometric data, to obtain a recorded signal including a recorded signal
30 portion based on the biometric data, to derive characteristic data from the recorded signal corresponding to the biometric data and to transmit the characteristic data from the client computer to a server computer.

The method may further comprise receiving the characteristic data from the client computer. The method may further comprise combining characteristic data so as to provide archetype characteristic data. The method may further comprise comparing characteristic data with archetype characteristic data so as to determine a score
5 dependent upon a degree of matching.

According to a third aspect of the present invention there is provided a signal representing control codes for causing computer apparatus to perform the method.

10 According to a fourth aspect of the present invention there is provided apparatus configured to perform the method.

According to a fifth aspect of the present invention there is provided apparatus for authentication comprising a server computer and a client computer; said server
15 computer being configured to transmit a computer program to said client computer and said client computer being configured to execute said computer program and thereafter to request a user to provide biometric data, to obtain a recorded signal based on the biometric data, to derive characteristic data from said recorded signal for characterising said biometric data and to transmit said characteristic data to a
20 server computer.

According to a sixth aspect of the present invention there is provided apparatus comprising a server computer which is configured to receive a request from a client computer and transmit a computer program to the client computer, said computer
25 program when executed by a computer causes said computer to request a user to provide biometric data, to obtain a recorded signal including a recorded signal portion based on said biometric data, to derive characteristic data from said recorded signal corresponding to said biometric data and to transmit said characteristic data from said client computer to a server computer.

30

The server computers may be the same. In other words, the client computer may download the computer and upload the characteristic data to the same computer.

The computer program may be executable on a virtual machine. The computer program may be in Java and may be a Java applet.

5 According to a seventh aspect of the present invention there is provided a signal representing control codes for causing computer apparatus to perform a method comprising requesting a user to provide biometric data, obtaining a recorded signal based on the biometric data, deriving characteristic data from the recorded signal for characterising the biometric data and transmitting the characteristic data from the client computer to another computer apparatus. The signal may represent
10 bytecode of a Java applet.

According to an eighth aspect of the present invention there is provided a data carrier storing the signal.

15 **Brief description of the drawings**

Embodiments of the present invention will now be described by way of example, with reference to the accompanying drawings in which:

Figure 1 is a schematic diagram of an authentication system for performing a method of authentication;

20 Figure 2 is shows a distributed authentication system including a client computer and a server computer;

Figure 3 is a schematic diagram of the client computer shown in Figure 2;

Figure 4 is a schematic diagram of the server computer shown in Figure 2;

Figure 5 is a process flow diagram of a method of authentication;

25 Figure 6 shows a web server transmitting a Java applet to a web browser;

Figure 7 shows a Java applet and an authentication server exchanging information;

Figure 8 shows a Java applet and an authentication server exchanging data during an enrolment stage;

Figures 9a to 9e show messages displayed during an enrolment stage;

30 Figure 10 shows a Java applet and an authentication server exchanging data during an authentication stage;

Figures 11a to 11d show messages displayed during an authentication stage

Figure 12 is an analog representation of a recorded signal;

- Figure 13 is a generic representation of a recorded signal;
Figure 14 is a digital representation of a recorded signal;
Figure 15 illustrates dividing a recorded signal into timeslices;
Figure 16 is a process flow diagram of a method of generating a featuregram;
5 Figure 17 illustrates generation of a feature vector;
Figure 18 illustrates generation of a featuregram from a plurality of feature vectors;
Figure 19 illustrates explicit endpointing;
Figure 20 illustrates creation of a speech featuregram;
Figure 21 illustrates generation of a speech featuregram archetype;
10 Figure 22 shows a probability distribution function;
Figure 23 shows a continuous distribution function;
Figure 24 shows an authentication biometric;
Figure 25 illustrates comparison of a featuregram archetype with an authentication featuregram; and
15 Figure 26 shows a client computer with biometric sensors.

Best modes for carrying out the invention

Authentication using spoken response

- Referring to Figure 1, an authentication system 1 for performing a method of
20 authentication is shown. The authentication system 1 limits access by a user 2 to a secure system 3. The secure system 3 may be an on-line bank account. The authentication system 1 is managed by a system administrator 4.

- An authentication system is described in GB 0211842.0 filed on 22 May 2002 and in
25 PCT/GB2003/002246 filed on 22 May 2003.

- Referring to Figure 2, the authentication system 1 is distributed and includes a client computer 5, a server computer 6 and a network 7. In this case, the network 7 is the Internet. The network may be wired or wireless or include one or more wired and
30 one or more wireless sections. The network may include a personal area network (not shown), a local area network (not shown) and/or a wide area network (not shown). Thus, a Bluetooth™ or WiFi wireless link may connect the client computer

5 to an access node (not shown) which in turn is connected to a local area network (not shown) which is in turn connected to the Internet.

The method of authentication is performed by client computer 5 and the server
5 computer 6. Thus, the functions of authentication system 1 may be thought as being divided between the client computer 5 and the server computer 6 in contrast to a single computer, for example as described in GB 0211842.0 *supra* and PCT/GB2003/002246 *supra*.

10 A method of authentication using spoken response will now be described.

Referring to Figure 3, the client computer 5 is shown in more detail. The client
computer 5 is a personal computer (PC) and may be desk-top PC, lap-top PC, hand-
held personal digital assistant (PDA) or cellular telephone. The client computer 5
15 includes a microphone 8 into which a user may provide a spoken response and which converts a sound signal into an electrical signal, an amplifier 9 for amplifying the electrical signal, an analog-to-digital (A/D) converter 10 for sampling the amplified signal and generating a digital signal, a filter 11, a processor 12 for performing signal processing on the digital signal, volatile memory 13 and non-
20 volatile memory 14. In this example, the A/D converter 10 samples the amplified signal at 11025 Hz and provides a 16-bit pulse code modulation (PCM) representation of the signal. The digital signal is filtered using a 4th order 100Hz high-pass filter to remove any DC offset. The amplifier 9, the A/D converter 10 and/or filter 11 may be implemented in a sound card or similar device.

25 The client computer 5 may additionally or alternatively be provided with a headset (not shown) for the user which includes a microphone into which the user may provide the spoken response.

30 The client computer 5 further includes a digital-to-analog (D/A) converter 15, another amplifier 16 and a speaker 17 for providing audio prompts and a monitor 18 for providing text prompts to the user 2. The client computer 5 also includes storage 19, such as a hard disk, a keyboard and mouse 20 and input/output (I/O)

circuit 21, for allowing data to be transmitted and received to and from the network 7 (Figure 2). The I/O circuit 21 may be a modem for connection by a telephone line to an Internet Service Provider (ISP) (not shown) and/or a network interface card for connection to a local area network (not shown) which in turn is connected to an ISP.

As will be explained in more detail later, the client computer 5 loads and runs a web-browser 30 (Figure 7) such as Microsoft® Internet Explorer or Netscape® Navigator which is Java enabled.

10

Referring to Figure 4, the server computer 6 is shown in more detail. The server computer 6 is in the form of a personal computer (PC). The server computer 6 includes a processor 23, volatile memory 24, non-volatile memory 25, storage 26, a display 27, an interface 28, such as a keyboard and mouse, and an I/O circuit 29 for allowing data to be transmitted and received to and from the network 7 (Figure 2). The interface 28 allows access by the system administrator 4 (Figure 2). The I/O circuit 21 may be a modem for connection by a telephone line to an Internet Service Provider (ISP) (not shown) and/or a network interface card for connection to a local area network (not shown) which in turn is connected to an ISP. The secure system 4 may be located on the server computer 6, connected to the server computer 6 via a local area network (not shown) or connected to the server computer 6 via the Internet 7.

As will be explained in more detail later, the server computer 6 runs a web server 31 (Figure 6), such as Apache, and an authentication server 32, which in this case is a voice authentication server (Figure 6).

Referring to Figure 5, the authentication process comprises two stages, namely enrolment (Step S1) and authentication (Step S2).

30

The aim of enrolment is to obtain specimens of biometric data, such as, a plurality of specimens of speech, from a user and to process them so as to derive a compact

data structure for example comprising acoustic information-bearing attributes that characterise the way the user speaks.

In this embodiment, enrolment includes asking a user to provide one or more
5 responses to a prompt and to make recordings. Each recording is divided into frames which are converted into feature vectors. The feature vectors may be concatenated to form a so-called "featuregram".

The featuregram is processed so as to isolate a portion which corresponds to the
10 spoken response provided by the user. This is called a "speech featuregram".

By combining two or more speech featuregrams based on speech utterances made by the same user in response to the same prompt, a reliable and distinctive template may be formed for each prompt for each user. The template is referred to as a
15 "speech featuregram archetype" (FGA). One or more speech featuregram archetypes corresponding to different prompts may be stored in an authentication biometric, which is subsequently used in authentication.

During authentication further specimens of biometric data, in this case speech
20 specimens, are obtained. The specimens are used to generate speech featuregrams which are compared with corresponding speech featuregram archetypes so to determine whether the authenticating user is the same as the enrolling user.

Featuregrams and featuregram archetypes are described in more detail in GB
25 0211842.0 *supra* and PCT/GB2003/002246 *supra*, and are also described in more detail later.

Some processes are common to both enrolment and authentication, such as generation of featuregrams. Thus, it is advantageous to generate featuregrams at the
30 client computer 5 and transmit them to the server computer 6 for further processing, such as (during enrolment) deriving speech featuregram archetypes and (during authentication) comparing speech featuregrams with speech featuregram archetypes.

Referring to Figure 6, the client computer 5 is shown running a web-browser 30. A user accesses a web page (not shown) provided a web server 31, such as Apache, running on the server computer 6. The server computer 6 also runs an authentication server 32, such as a Java application.

The web-page includes buttons which, when pressed, begins enrolment or authentication. If a button is pressed, Java applet code 33 is downloaded to the client computer 5.

10

Referring to Figure 7, the Java applet code 33 is run by the web browser 30. The applet 33 establishes a connection with the authentication server 32 which starts a new thread of execution and parameters 34 are exchanged

15 The Java applet 33 may do a number of things. It can cause the client computer 5 to perform a calibration process using specimens of spoken utterances, to capture recordings, to generate featuregrams, to perform endpointing and to perform sanity checks as described in GB 0211842.0 *supra* and PCT/GB2003/002246 *supra*. One or more featuregrams, preferably speech featuregrams 35, are transmitted to the server computer 6. The server computer 6 may (during enrolment) create speech featuregram archetypes, set pass levels and create an authentication biometric for each user and (during authentication) compare speech featuregrams with corresponding speech featuregram archetypes and check for replay attack as described in GB 0211842.0 *supra* and PCT/GB2003/002246 *supra*.

25

In Figure 7, the applet 33 and the authentication server 32 are shown to be communicating without using the web server 31. However, the authentication server 32 may be Java application or other program which exchanges data with the applet 33 through the web server 31.

30

Enrolment

Referring to Figures 8 and 9a to 9e, an enrolment process, performed when an enrolment version of the applet 33 is downloaded and run, will now be described.

Figures 9a to 9e show screen shots 37₁, 37₂, 37₃, 37₄, 37₅ at different stages of the enrolment process.

- 5 Referring to Figure 9a, the applet 33 presents the user with an entry form and asks the user to provide personal details, such as name and postcode (step P1).
Postcodes may also be known as ZIP codes.

Once the user enters their name and postcode in corresponding fields 38₁, 38₂ and
10 presses a "Submit" button 39, the applet 33 establishes a connection with the authentication server 32 for exchanging parameters 34 (Figure 7) (step P2). The authentication server 32 creates a new thread for the transaction. Once a connection has been established, the applet 33 transmits a message "ENR" to inform the authentication server 32 of the type of transaction being performed (step
15 P3). The applet 33 also transmits the user's details, which in this case comprise the user's name and postcode (step P4 & P5). The authentication server 32 returns a message "CAL" to indicate that it has received the user's details and to instruct the applet 33 to move on to the next stage of enrolment (step P6).

- 20 Referring to Figure 9b, the applet 33 performs a calibration process during which the user provides specimen of speech utterances (step P7). The user may guided through the calibration process, for example using a so-called "calibration wizard"
40.

25 A purpose of calibration is to set the gain of the microphone amplifier 9 (Figure 3) for example to avoid saturation. Commonly this is known as setting a recording volume. A calibration process is described in more detail in GB 0211842.0 *supra* and PCT/GB2003/002246 *supra*.

- 30 When the calibration process ends, the applet 33 transmits calibration data to the authentication server 32, which stores the data as part of the authentication biometric (step P8 & P9). The applet 33 may return information including the type and configuration of the client computer 5, the type of microphone 8 (Figure 3) and

speaker 17 (Figure 3) and the type and gain settings of the amplifiers 9, 16 (Figure 3).

The authentication server 32 returns a message "ENRREC" to indicate that it
 5 received the calibration data and to instruct the applet 33 to proceed with the next stage of enrolment, namely recording (step P10).

Referring to Figure 9c, the applet 33 presents the user with a warning 41 that they
 10 are going to be prompted a plurality of times and, when ready, to press "Continue" 42 (step P11).

The authentication server 32 transmits a plurality of parameters 34 (Figure 7) to the applet 33 regarding what to record and how to create featuregrams, such as
 15 sampling frequency, whether to use data compression, a number of prompts to be used, a number of repetitions of prompts to be used and a plurality of prompt (steps P12 to P16). Data compression may comprise reducing or omitting overlapping of timeslices.

Table 1 below provides examples of some typical parameters:
 20

Table 1

Message type	Example
Sample Frequency (Hz)	11025
Compress Flag (Y/N)	N
Number of prompts	4
Number of repeats (R)	4
Prompt 1	52
Prompt 2	29
Prompt 3	43
Prompt 4	37

The authentication server 32 generates a personal identifier (PID), for example using the user's name and postcode (step P17) and transmits it to the applet 33 (step P18).

- 5 The applet 33 transmits a message "START" to the authentication server 32 to inform it that recording has started and to warn it that featuregrams are about to be sent (step P19).

Referring to Figure 9d, the applet 33 displays a text prompt 43, such as "Please say:-
10 52" and displays a timer 44 to indicate a time left for responding (step P20). Typically, the user is allowed three seconds to respond. While the text prompt 43 is displayed, the applet 33 records a spoken response 36 (Figure 7) (step P21).

The applet 33 then generates a speech featuregram (step P22). This comprises
15 dividing the recorded signal into timeslices (overlapping timeslices if the compression flag is not set), converting each timeslice into a feature vector, concatenating feature vectors to form a featuregram and performing endpointing to identify a portion of the recorded signal which contains a spoken utterance and isolate the recorded signal portion to generate the speech featuregram.

20 Once the applet 33 has generated a speech featuregram, it transmits the featuregram 35 to the authentication server 32, together with data identifying the prompt and indicating its duration (steps P23 & P24).

- 25 Steps P20 to P24 are repeated for each prompt and each prompt is repeated a predetermined number of times. Thus, in this case there are 4 prompts and 4 repeats. The order may be determined by the applet 33.

When the authentication server 32 has received all the featuregrams 35 it was
30 expecting, it sends a message to the applet 33 containing the user's PID (step P25).

Referring to Figure 9e, the applet 33 displays a message 45 that enrolment is complete and informs the user of their PID (step P26).

If, at any time, the applet 33 dies, for example due to the web browser closing, the applet 33 sends a message "ABORT" to the authentication server. If the authentication server 32 receives this message, it stops the transaction thread from
5 continuing. This has the advantage of being robust since it helps to prevent server threads from persisting after applet disconnection and thus 'unclogs' the server.

The authentication server 32 combines speech featuregrams corresponding to the same prompt to form a plurality of speech featuregram archetypes (step P27).
10 These are stored together with information relating to the prompts and calibration data in an authentication biometric (step P28).

A method of creating speech featuregram archetypes and a method of generating authentication biometrics is described in GB 0211842.0 *supra* and
15 PCT/GB2003/002246 *supra*.

Authentication

Once enrolment has been successfully completed, the user is registered as a valid user. Access to the secure system 3 or authorisation of a transaction is conditional
20 on successful authentication.

Referring to Figures 10 and 11a to 11d, an authentication process, performed when an authentication version of applet 33 is downloaded and run, will now be described.
25

Figures 11a to 11d show screen shots 46₁, 46₂, 46₃, 46₄, 46₅ at different stages of the authentication process.

Referring to Figure 11a, the applet 33 presents the user with an entry form and asks
30 the user to provide their PID (step Q1).

Once the user enters their PID in the field provided 47 and presses a "Submit" button 48, the applet 33 establishes a connection with the authentication server 32

for exchanging parameters 34 (Figure 7) (step Q2). The authentication server 32 creates a new thread for this transaction. Once a connection has been established, the applet 33 transmits a message "VAL" to inform the authentication server 32 of the type of transaction being performed (step Q3). The applet 33 also transmits the
5 user's PID (step Q4). The authentication server 32 checks the validity of the PID (step Q5) and if valid returns a message "PIDOK" (step Q6). If not a valid PID, the authentication server 32 sends a message "PIDNOTOK" and the applet 33 will finish with the corresponding message.

10 The authentication server 32 sends data relating to the recording device and recording volume (steps Q7 & Q8). It then sends a message "VALREC" to instruct the applet 33 to begin recording (step Q9). Additionally or alternatively, the applet may perform a calibration process.

15 Referring to Figure 11b, the applet 33 presents the user with a warning 49 that they are going to be prompted a plurality of times and, when ready, to press "Continue" 50 (step Q10).

The authentication server 32 transmits a plurality of parameters to the applet 33
20 regarding what to record and how to create featuregrams, such as sampling frequency, whether to use data compression, a number of prompts to be used, a number of repetitions of prompts to be used and a plurality of prompt (steps Q11 to Q15). Table 2 below provides examples of some typical messages:

Table 2

Message type	Example
Sample Frequency (Hz)	11025
Compress Flag (Y/N)	N
Number of prompts	4
Number of repeats (R)	1
Prompt 1	52
Prompt 2	29
Prompt 3	43
Prompt 4	37

The applet 33 transmits a message "START" to the authentication server 32 to
 5 inform it that recording has started and to warn it that featuregrams are about to be
 sent (step Q16).

Referring to Figure 11c, the applet 33 displays a text prompt 51, such as "Please
 say:- 29" and displays a timer 52 to indicate a time left for responding (step Q17).
 10 Typically, the user is allowed three seconds to respond. While the text prompt 51 is
 displayed, the applet 33 records a spoken response (step Q18).

The applet 33 then generates a speech featuregram (step Q19). This comprises
 dividing the recorded signal into timeslices (if the compression flag is not set then
 15 the timeslices overlap), converting each timeslice into a feature vector,
 concatenating feature vectors to form a featuregram and performing endpointing to
 identify a portion of the recorded signal which contains a spoken utterance and
 isolate the recorded signal portion to generate the speech featuregram.

20 Once the applet 33 has generated a featuregram, it transmits the featuregram to the
 authentication server 32, together with data identifying the prompt and indicating its
 duration (steps Q20 & Q21).

Steps Q17 to Q21 are repeated for each prompt. Thus, in this case there are 4 prompts.

When the authentication server 32 has received all the featuregrams it was
5 expecting, it compares each featuregram with a corresponding featuregram
archetype for the same prompt (step Q22). It collects the scores for each
comparison and determines whether there is a match or not (step Q23). The
authentication server 32 sends a message to the applet 33 informing it whether the
user passed or not (step Q24).

10

Referring to Figure 11d, the applet 33 displays an appropriate result (step Q25).

Generating featuregrams at the client computer has several advantages. It reduces
the amount of information transmitted to the authentication server. For example, a
15 spoken response typically comprises 64 kB of data, whereas a featuregram may
comprises only 2 kB of data. It is more robust and it helps to share processing
between the client and server computers.

The processes used (during enrolment) to generate featuregrams 35 and
20 authentication biometrics and (during authentication) to generate featuregrams 35
and to compare featuregrams 35 with featuregram archetypes will now be described
in more detail:

Recording

25 Referring again to Figure 3, at the client computer 5, a spoken response is recorded
by the microphone 8, amplified by amplifier 9 and sampled using A/D converter 10
at 11025 Hz to provide a 16-bit PCM digital signal. Preferably, the recording lasts
about 3 seconds. The signal is then filtered to remove any d.c. component. The
signal may be stored in volatile memory 13.

30

Referring to Figures 12, 13, 14, an example of a recorded signal 54 is shown in
analog, partitioned and digital representations. The partitioned representation helps
show that that the recorded signal 54 may comprise different sections 55, 56, 57.

Referring particularly to Figure 12, the recorded signal 54 may comprise one or more speech utterances 55, one or more background noises 56 and/or one or more silence intervals 57. A speech utterance 55 is defined as a period in a recorded
5 signal 54 which is derived solely from the spoken response of the user. A background noise 56 is defined as a period in a recorded signal arising from audible sounds, but not originating from the speech utterance. A silence interval 57 is defined as a period in a recorded signal which is free from background noise and speech utterance.

10

As explained earlier, the purpose of the enrolment is to obtain a plurality of specimens of speech so as to generate an authentication biometric. To help achieve this, recorded responses are characterised by generating "featuregrams" which comprise sets of feature vectors. The recordings are also examined so as to isolate
15 speech from background noise and silences.

The recordings are inspected to identify spoken utterances. This is known as "endpointing". By identifying speech utterances, a speech featuregram may be generated which corresponds to portions of the recorded signal comprising speech
20 utterances.

Referring to Figure 15, a portion 54' of the recorded signal 54 is shown. The recorded signal 54 is divided into frames, referred to herein as timeslices 58. The recorded signal 54 is divided into partially-overlapping timeslices 58 having a
25 predetermined period. In this example, timeslices 58 have a period of 50 ms, i.e. $t_1 = 50\text{ms}$, and overlap by 50%, i.e. $t_2 = 25\text{ ms}$. However, if compression is used then there is no overlap.

Featuregram generation

30 Referring to Figures 16, 17 and 18, a process by which a featuregram is generated at the client computer 5 will be described in more detail:

The recorded signal 54 is divided into timeslices 58 (step T1). Each timeslice 58 is converted into a feature vector 59 using a feature transform 60 (step T2).

5 The content of the feature vector 59 depends on the transform 60 used. In general, a feature vector 59 is a one-dimensional data structure comprising data related to acoustic information-bearing attributes of the timeslice 58. Typically, a feature vector 59 comprises a string of numbers, for example 10 to 50 numbers, which represent the acoustic features of signal comprised in the timeslice 58.

10 In this example, a so-called cepstral transform 60 is used. In this example, for a sampling rate of 11025Hz, each feature vector 59 comprises twelve signed 8-bit integers, typically representing the second to thirteenth calculated cepstral coefficients. Data relating to energy (in dB) may be included as a 13th feature. This has the advantage of helping to improve the performance of a word spotting routine
15 that would otherwise operate on the feature vector coefficients alone.

Further details regarding cepstral transforms may be found on page 115 in "Fundamentals of Speech Recognition" by Rabiner & Juang (Prentice Hall, 1993).

20 Other transforms may be used. For example, a linear predictive coefficient (LPC) transform may be used in conjunction with a regression algorithm so as to produce LPC cepstral coefficients. Alternatively, a TESPAP transform may be used.

Linear predictive coefficient (LPC) transform is described by B.S. Atal,
25 "Effectiveness of linear prediction characteristics of the speech wave for automatic speaker identification and verification", Journal of Acoustical Society of America, Vol. 55, pp-1304-1312, June 1974. Further details regarding the TESPAP transform may be found in GB-B-2162025.

30 Referring to Figure 18, a featuregram 61 comprises a set or concatenation of feature vectors 59. The featuregram 61 includes speech utterances, background noise and silence intervals.

The featuregram 61 may be sent from the client computer 5 to the server computer 6 for endpointing for determining a speech featuregram. However, it is preferable to perform endpointing at the client computer 5 and to transmit a speech featuregram 35.

5

Endpointing

Endpointing seeks to identify portions of a recorded signal which contains spoken utterances. This allows generation of speech featuregrams which characterise the spoken utterances. In this case, explicit endpointing is used.

10

Explicit endpointing seeks to locate approximate endpoints of a speech utterance in a particular domain without using any *a priori* knowledge of the words that might have been spoken. Explicit endpointing tracks changes in signal energy profile over time and frequency and makes boundary decisions based on general assumptions regarding the nature of profiles that are indicative of speech and those that are
15 representative of noise or silence. Explicit endpointing cannot easily distinguish between speech spoken by the enrolling user and speech prominently forming part of background noise. Therefore, it is desirable that no one else speaks in close proximity to the valid user when enrolment takes place.

20

Referring to Figure 19, an explicit endpointing process 62 generates a plurality of pairs 63 of possible start and stop points for a stream of timeslices 58.

Endpointing is described in more detail on pages 143 to 149 of "Fundamentals of
25 Speech Recognition" *supra*.

Creating a speech featuregram

Once the endpoints of the recorded signal 54 (Figure 12) have been identified and, optionally, the recorded signal (Figure 12) passes a plurality of sanity checks, such
30 checking signal-to-noise ratio, then a speech featuregram may be created.

Referring to Figure 20, a speech featuregram 35 is created using a process 64 by concatenating feature vectors 59 extracted from the section of the featuregram 61

that originates from the speech utterance. The speech section of the featuregram 61 is located using the speech endpoints 63.

The speech featuregram 35 is then transmitted from the client computer 5 to the server computer 6.

Creating a speech featuregram archetype

The aim of the enrolment is to provide a characteristic voiceprint for one or more words or phrases. However, specimens of the same word or phrase provided by the same user usually differ from one another. Therefore, it is desirable to obtain a plurality of specimens and derive a model or archetypal specimen. This may involve discarding one or more specimens that differ significantly from other specimens.

Referring to Figure 21, a speech featuregram archetype 65 is calculated at the server computer 6 (Figure 6) using an averaging process 66 using w-featuregrams 35₁, 35₂, ..., 35_w. In this case, four featuregrams 35 are used, the average of the three most similar featuregrams being used to create the featuregram archetype 65.

Setting an appropriate pass level

A featuregram archetype 65 is obtained for each prompt. Thus, during subsequent authentication, a user is asked to provide a response to a prompt. A speech featuregram 35 is obtained and compared with the speech featuregram archetype 65 at the server computer 6 (Figure 6) using a dynamic time warping process which is described in more detail later. The comparison produces a score and the score is compared with a preset pass level. A score which falls below the pass level indicates a good match and so the user is accepted as being a valid user.

A valid user is likely to provide a response that results in a low score, falling below the pass level, and which is accepted. However, there may be occasions when even a valid user provides a response that results in a high score and which is rejected. Conversely, an impostor may be expected to provide poor responses which are usually rejected. Nevertheless, they may occasionally provide a sufficiently close-matching response which is accepted. Thus, the pass level affects the proportion of

valid users being incorrectly rejected, i.e. the “false reject rate” (FRR) and the proportion of impostors which are accepted, i.e. “false accept rate” (FAR).

In this example, a neutral strategy is adopted which shows no bias towards
5 preventing unauthorised access or allowing authorised access.

A pass level for a fixed-word or fixed-phrase prompt is determined using previously acquired captured recordings taken from a wide range of representative speakers.

10 A featuregram archetype is obtained for each of a first set of users for the same prompt in a manner hereinbefore described. Thereafter, each user provides a spoken response to the prompt from which a featuregram is obtained and compared with the user’s featuregram archetype using a dynamic time warping process so as produce a score. This produces a first set of scores corresponding to valid users.

15

The process is repeated for a second set of users, again using the same prompt. Once more, each user provides a spoken response to the prompt from which a featuregram is obtained. However, the featuregram is compared with a different user’s featuregram archetype. Another set of scores is produced, this time
20 corresponding to impostors.

Referring to Figure 22, frequency of scores for valid users and impostors are fitted to first and second probability density functions 67_1 , 67_2 respectively using:

$$25 \quad p(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right\}$$

where, p is probability, x is score, μ is mean score and σ is standard deviation. Other probability density functions may be used.

The mean score μ_1 for valid users is expected to be lower than the mean score μ_2 for the impostors. Furthermore, the standard deviation σ_1 for the valid users is usually smaller than the standard deviation σ_2 of the second density function

- 5 Referring to Figure 23, the first and second probability density functions 67_1 , 67_2 are numerically integrated to produced first and second continuous density functions 68_1 , 68_2 . The point of intersection 69 of the first and second continuous density functions 174_1 , 174_2 is the equal error rate (ERR), wherein $FRR = FAR$. The score at the point of intersection 69 is used as a pass score for the prompt.

10

Creating an authentication biometric

- Referring to Figure 24, an authentication biometric 70 is shown. The authentication biometric 70 comprises sets of data 71_1 , 71_2 , ... 71_q corresponding to featuregram archetypes 65 and associated prompts 72. The authentication biometric 70 may
15 further comprise ancillary information including the number of prompts to be issued during authentication 73, scoring strategy 74 and gain settings 75. The biometric 70 may include further information, for example related to high-level logic for analysing scores.

- 20 The authentication biometric 70 is stored in storage 26 (Figure 4).

Matching authentication featuregrams with the authentication biometric

- Referring to Figure 25, a dynamic time warping process 77 is used to compare a speech featuregram 35 obtained during authentication with a speech featuregram archetype 65 obtained during enrolment. This is achieved by compressing and/or
25 expanding different sections of the speech featuregram 35 until a region inside the speech featuregram 35 matches the speech featuregram archetype 65. The best fit is known as the winning path and a "cost of alignment" 78 is output which specifies how close the fit is. The cost 78 is used to determine whether the speech
30 featuregram 35 is sufficiently "close" to the speech featuregram archetype 65 and thus whether to validate the user.

Dynamic time warping is described in more detail on pages 221 to 226 of "Fundamentals of Speech Recognition" *supra*.

Other types of biometric data may be used instead of voice. These may include
5 permanent biometric data such as physical characteristics such as fingerprint, handprint, face pattern, scent, DNA, iris pattern and retinal configuration. Permanent biometric data do not change or substantially do not change. These may include generatable biometric data such as handwriting or acoustic signature.

10 Referring to Figure 26, the client computer 5 may include one or more sensors 39₁, 39₂, 39_N for capturing biometric data. The sensors 39₁, 39₂, 39_N may be peripheral devices connected via the input/output circuits 21 or incorporated into the server computer 5. If authentication using voice is not used then the microphone 8, amplifier 9, A/D converter 10 and filter 11 may be omitted.

15

The server computer 6 runs an authentication server appropriate to the, or each, type of biometric data.

Authentication using fingerprint

20 A method of authentication using fingerprint will now be described.

The method is similar to the method of authentication using spoken responses. However, there are some differences including the type of biometric data used, namely fingerprint, the manner in which it is captured, the number of specimens
25 taken, the manner in which it is characterised, the form of data which is returned to the authentication server 32 and the manner in which the data is processed by the server 32. The method may also include appropriate calibration and sanity checks.

Referring to Figure 26, the first sensor 39₁ is suitable for recording fingerprints and
30 may include optical or capacitive arrangements for recording a fingerprint pattern. Examples of fingerprint sensors are given in GB-A-1377797, WO-A-9712340 and EP-A-1239404.

The client computer 5 downloads a Java applet 33 as described earlier. The Java applet code 33 causes the client computer 5 to request the user to provide a fingerprint, for example by asking them to place a finger against the sensor 39₁. It also causes the client computer 5 generate characteristic data for characterising the fingerprint, for example as described in chapter 5 of "Introduction to Fingerprint Comparison" by Gary Jones (2000) [ISBN 0-9661970-3-8] and in GB-A-1577797. For example, this may include determining the locations of an end of a ridge or valley, or bifurcation. Thus, the characteristic data may comprise a plurality of sets of co-ordinates. The Java applet code 33 causes the client computer 5 to transmit the characteristic data to the server computer 6.

The server computer 6 may (during enrolment) store the characteristic data in the authentication biometric 70. The server computer 6 (during authentication) may compare characteristic data stored in the authentication biometric with characteristic data generated during authentication and determine whether there is a match.

Generating characteristic data at the client computer 5 has several advantages. It reduces the amount of information transmitted to the server computer 6, it is more robust and it helps to share processing between the client and server computers 5, 6.

Although the embodiments hereinbefore described using the single type of biometric, two or more biometrics may be used. Thus, the authentication biometric may store data relating to two or more different types of biometric. For example, enrolment may involve recording spoken responses and fingerprints. Subsequent authentication may require the user to provide spoken responses and/or fingerprints. Using two or more biometrics has the advantage of providing additional security.

30

It will be appreciated that many modifications may be made to the embodiments described above. For example, a Java application or an executable file may be downloaded to the client computer and run. Other types of code which are

dynamically downloadable and executable may be used. The code may be an interpreted or compiled code. A single or many Java applets may be downloaded, for example one applet for recording, one applet for endpointing etc. A single applet may be used for both enrolment and authentication. Separate server
5 computers may be used for the web server and the authentication server. The client computer may unload the computer program from memory after execution and this may be done automatically. It will also be appreciated that the terms “authenticating” and “identifying” may be used interchangeably.

10 Although claims have been formulated in this application to particular combinations of features, it should be understood that the scope of the disclosure of the present invention also includes any novel features or any novel combination of features disclosed herein either explicitly or implicitly or any generalisation thereof, whether or not it relates to the same invention as presently claimed in any claim and whether
15 or not it mitigates any or all of the same technical problems as does the present invention. The applicants hereby give notice that new claims may be formulated to such features and/or combinations of such features during the prosecution of the present application or of any further application derived therefrom.

Claims

1. A method of authentication for use over a network, the method comprising:
transmitting a computer program from a server computer to a client computer
5 and
executing said computer program at said client computer, said client computer
thereafter requesting a user to provide biometric data, obtaining a recorded signal
based on said biometric data, deriving characteristic data from said recorded signal
for characterising said biometric data and transmitting said characteristic data from
10 said client computer to a server computer.
2. A method according to claim 1, wherein requesting said user to provide
biometric data comprises requesting said user to provide a response.
- 15 3. A method according to claim 1 or 2, wherein obtaining said recorded signal
based on said biometric data comprises obtaining a recorded signal including a
recorded signal portion corresponding to said response.
4. A method according to any preceding claim, wherein deriving characteristic
20 data from said recorded signal for characterising said biometric data comprises
deriving a set of feature data for characterising said recorded signal portion.
5. A method according to claim 4, wherein transmitting said characteristic data
from said client computer to a server computer comprises transmitting said set of
25 feature data from said client computer to a server computer.
6. A method according to claim 5, further comprising said client computer
calibrating an input device and setting a signal level of said recorded signal.
- 30 7. A method according to claim 5 or 6, further comprising said client computer
determining an endpoint of said recorded signal.

8. A method according to claim 1, wherein obtaining the recorded signal comprises capturing generatable biometric data.
9. A method according to claim 1 or 2, wherein obtaining the recorded signal
5 comprises capturing transient biometric data.
10. A method according to any preceding claim, wherein requesting said user to provide biometric data comprises requesting said user to provide a spoken response to a prompt.
- 10 11. A method according to claim 10, wherein obtaining said recorded signal based on said biometric data comprises obtaining a recorded signal including a recorded signal portion corresponding to said spoken response.
- 15 12. A method according to claim 11, wherein deriving characteristic data from said recorded signal for characterising said biometric data comprises deriving a set of feature vectors for characterising said recorded signal portion.
- 20 13. A method according to claim 12, wherein transmitting said characteristic data from said client computer to a server computer comprises transmitting said set of feature vectors from said client computer to a server computer.
- 25 14. A method according to claim 13, further comprising said client computer calibrating a microphone and an amplifier for setting a signal level of said recorded signal.
15. A method according to claim 13 or 14, further comprising said client computer determining an endpoint of said recorded signal.
- 30 16. A method according to claim 1 or 8, wherein obtaining the recorded signal comprises capturing permanent biometric data.

17. A method according to claim 1, 8 or 16, requesting said user to provide biometric data comprises requesting said user to provide a written response to a prompt.

5 18. A method according to claim 17, wherein obtaining said recorded signal based on said biometric data comprises obtaining a recorded signal including a recorded signal portion corresponding to said written response.

19. A method according to claim 1, wherein obtaining the recorded signal
10 comprises reading permanent biometric data.

20. A method according to claim 1 or 19, wherein requesting the user to provide biometric data comprises requesting said user to submit at least a body portion for sensing by a biometric sensor.

15 21. A method according to claim 1, 19 or 20, wherein reading said biometric data comprises capturing an image.

22. A method according to claim 1 or any one of claims 19 to 21, wherein reading
20 said biometric data comprises recording a pattern.

23. A method according to claim 1 or any one of claims 19 to 22, wherein reading said biometric data comprises recording a configuration.

25 24. A method according to claim 1 or any one of claims 19 to 23, wherein obtaining said recorded signal based on said biometric data comprises recording a representation of said biometric data.

25. A method according to claim 1 or any one of claims 19 to 24, wherein
30 obtaining said recorded signal based on said biometric data comprises taking a fingerprint.

26. A method according to claim 1, 19 or 20, wherein reading said biometric data comprises taking a chemical sample.
27. A method according to claim 26, wherein taking the chemical sample
5 comprises sampling scent.
28. A method according to any preceding claim, comprising said client computer requesting said computer program from said server computer.
- 10 29. A method according to any preceding claim, comprising said client computer dynamically downloading said computer program from said server computer.
30. A method according to any preceding claim, comprising said client computer accessing a web page provided by the server computer and requesting said computer
15 program from said server computer without prompt by the user.
31. A method according to any preceding claim, wherein executing said computer program occurs substantially immediately after said computer program is transmitted from a server computer to a client computer.
- 20 32. A method according to any preceding claim, further comprising said client computer requesting the user to provide further biometric data, obtaining recorded signals for respective biometric data, deriving respective characteristic data and transmitting said characteristic data from said client computer to said server
25 computer.
33. A method according to any preceding claim, further comprising said server computer combining characteristic data so as to provide archetype characteristic data.
- 30 34. A method according to any preceding claim, further comprising said server computer comparing characteristic data with archetype characteristic data so as to determine a score dependent upon a degree of matching.

35. A method of operating a server computer comprising:
receiving a request from a client computer;
transmitting a computer program to the client computer, said computer
5 program when executed by a computer causing said computer to request a user to
provide biometric data, to obtain a recorded signal including a recorded signal
portion based on said biometric data, to derive characteristic data from said
recorded signal corresponding to said biometric data and to transmit said
characteristic data from said client computer to a server computer.
- 10 36. A method according to claim 35, further comprising:
receiving said characteristic data from said client computer.
37. A method according to claim 35 or 36, further comprising:
15 combining characteristic data so as to provide archetype characteristic data.
38. A method according to claim 35 or 36, further comprising:
comparing characteristic data with archetype characteristic data so as to
determine a score dependent upon a degree of matching.
- 20 39. Apparatus configured to perform the method according to any preceding
claim.
40. Apparatus for authentication comprising:
25 a server computer and a client computer; said server computer being
configured to transmit a computer program to said client computer and said client
computer being configured to execute said computer program and thereafter to
request a user to provide biometric data, to obtain a recorded signal based on the
biometric data, to derive characteristic data from said recorded signal for
30 characterising said biometric data and to transmit said characteristic data to a server
computer.

41. Apparatus comprising a server computer which is configured to receive a request from a client computer and transmit a computer program to the client computer, said computer program when executed by a computer causes said computer to request a user to provide biometric data, to obtain a recorded signal
5 including a recorded signal portion based on said biometric data, to derive characteristic data from said recorded signal corresponding to said biometric data and to transmit said characteristic data from said client computer to a server computer.
- 10 42. A method or apparatus according to any preceding claim, wherein the server computers are the same.
43. A method or apparatus according to any preceding claim, wherein said computer program is executable on a virtual machine.
- 15 44. A method or apparatus according to any preceding claim, wherein said computer program is in Java.
45. A method or apparatus according to any preceding claim, wherein said
20 computer program is a Java applet.
46. A signal representing control codes for causing computer apparatus to perform a method comprising requesting a user to provide biometric data, obtaining a recorded signal based on said biometric data, deriving characteristic data from said
25 recorded signal for characterising said biometric data and transmitting said characteristic data from said client computer to another computer apparatus.
47. A signal according to claim 46 representing bytecode of a Java applet.
- 30 48. A signal representing control codes for causing computer apparatus to perform a method according to any one of claims 35 to 38.
49. A data carrier storing a signal according to any one of claims 35 to 38.

1/16

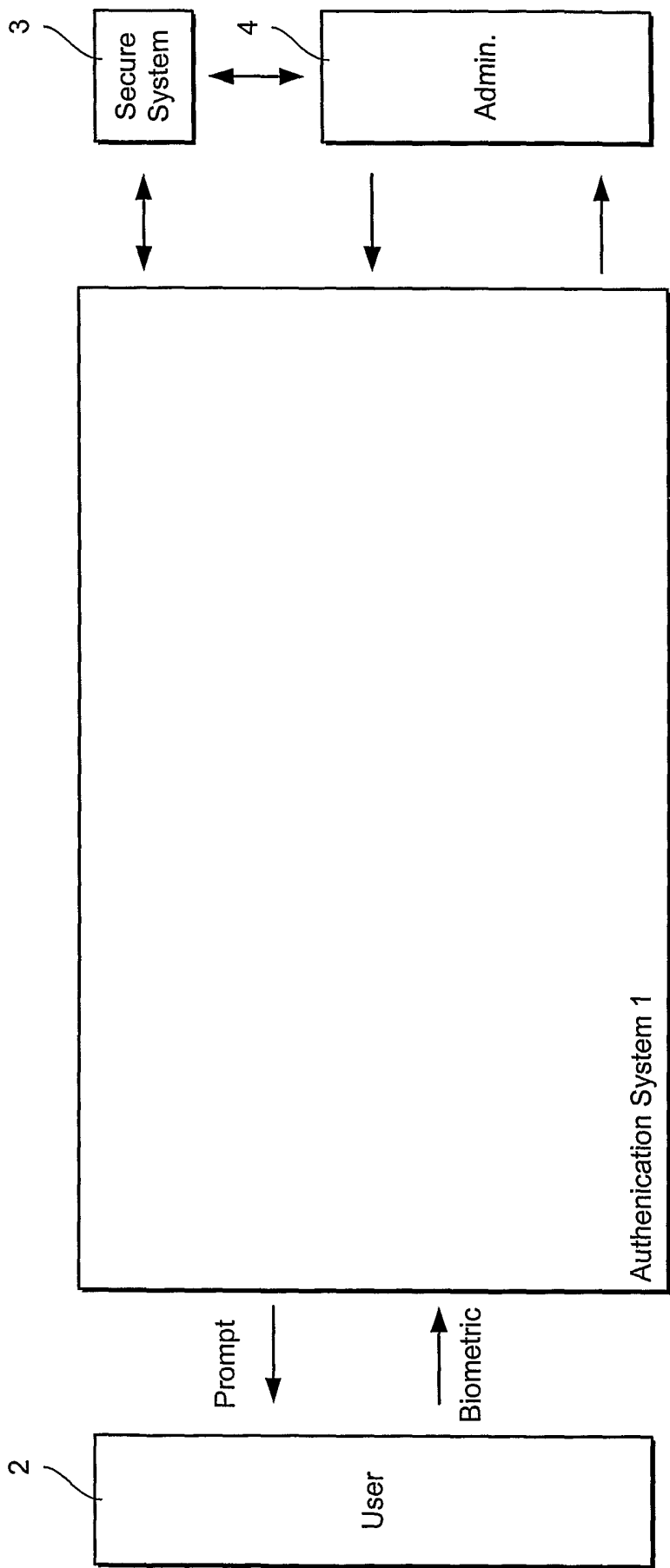


Fig. 1

2/16

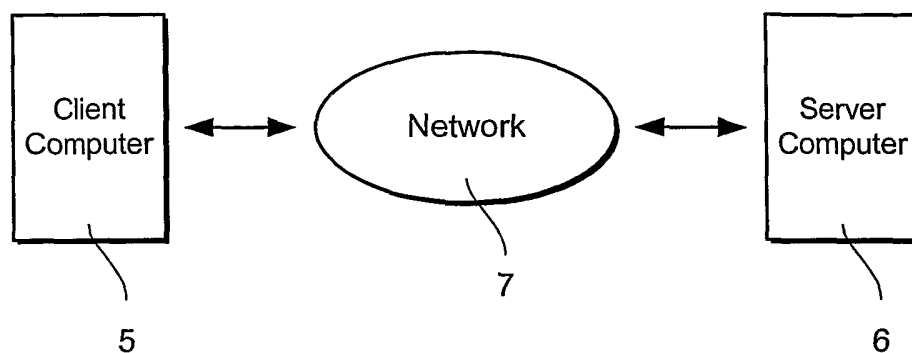


Fig. 2

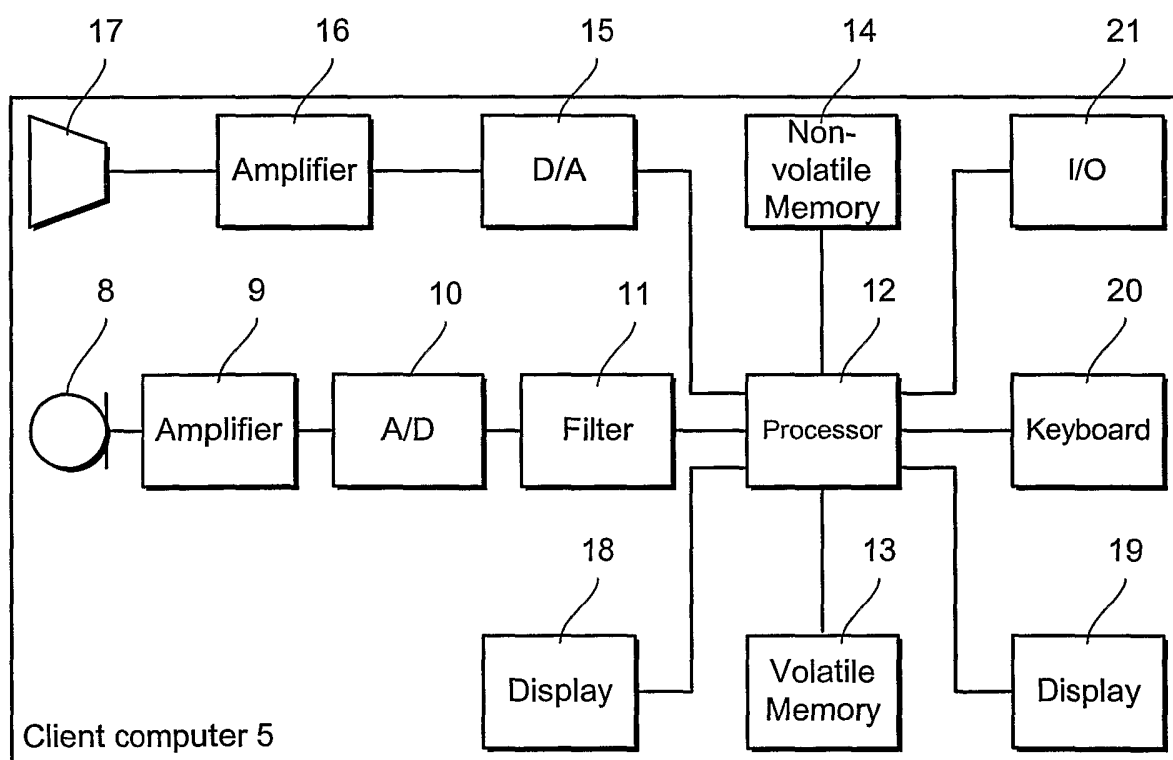


Fig. 3

3/16

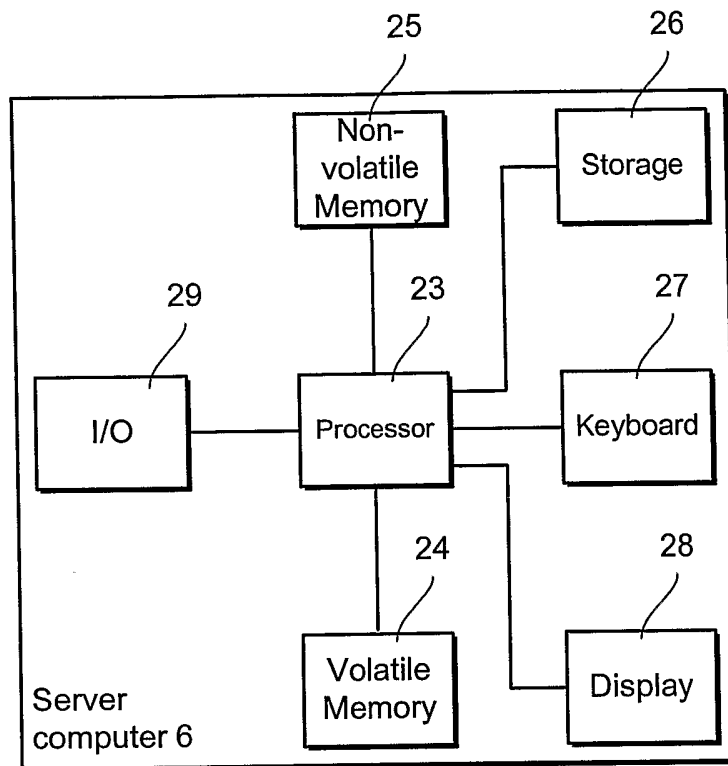


Fig. 4

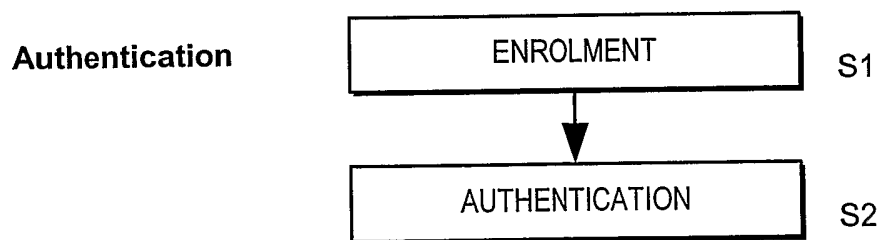


Fig. 5

4/16

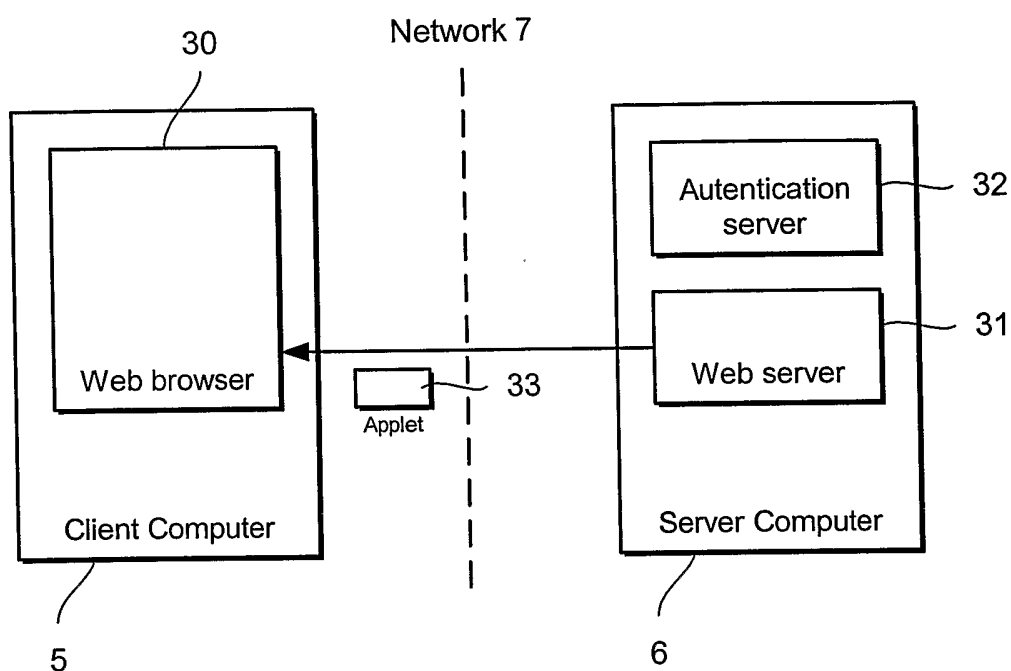


Fig. 6

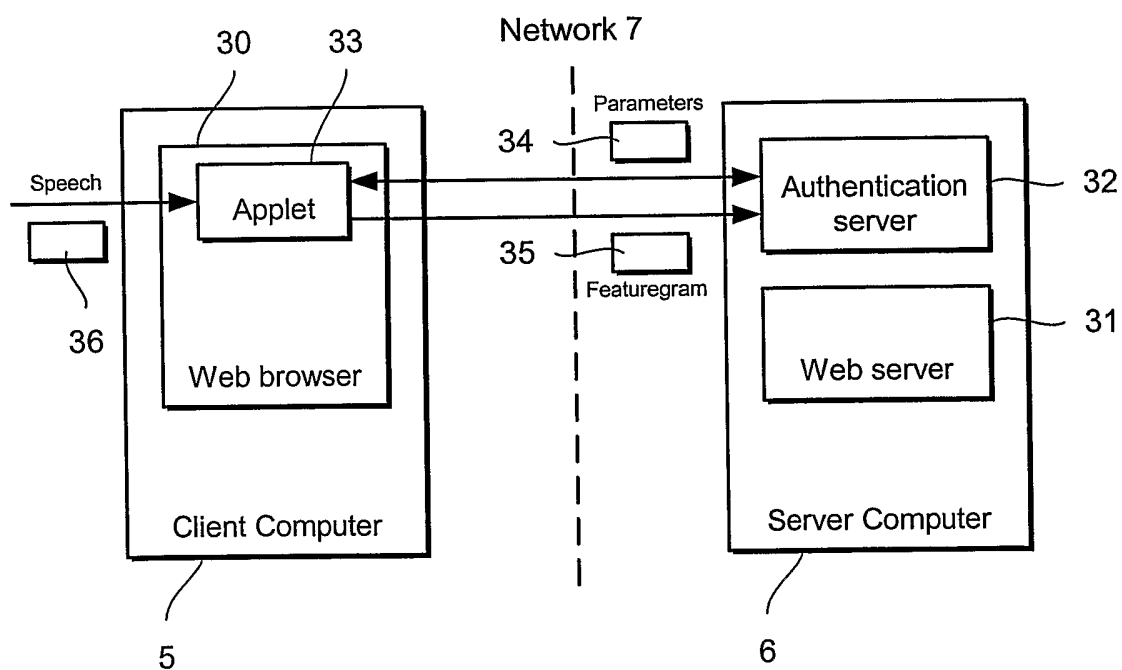


Fig. 7

5/16

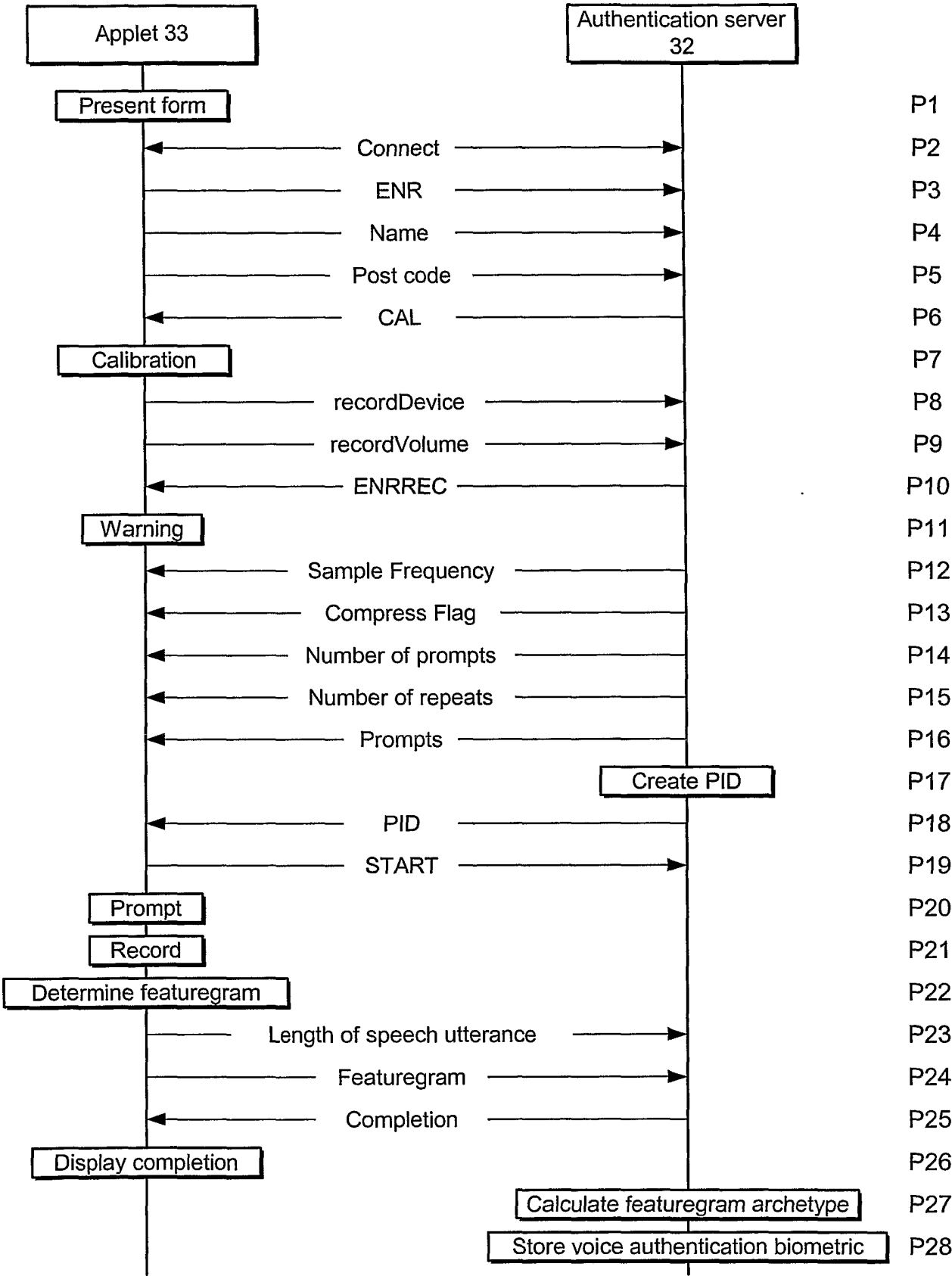


Fig. 8

6/16

Please enter your details below and then press the 'Submit' button.

Name:

Postcode:

38₁ 37₁ 38₂ 39

Fig. 9a

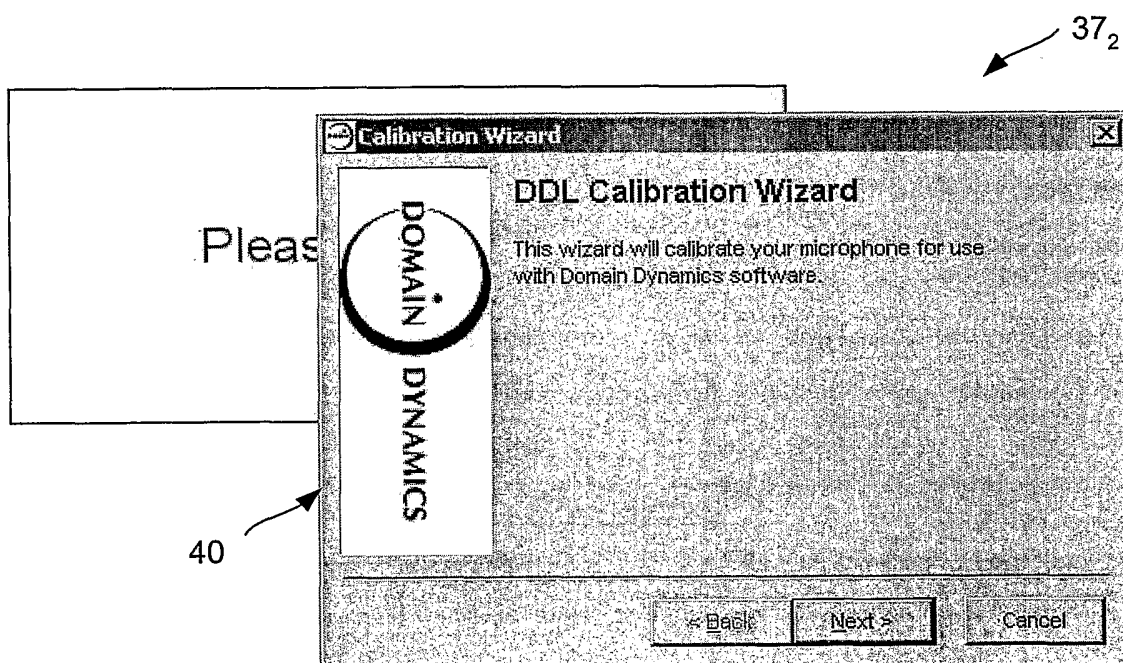


Fig. 9b

The next screen is be a series of prompts necessary to enrol your voice on the system. Please speak clearly, but naturally, and avoid background noise. Speak each prompt as it appears. Please press the 'Continue' button to proceed.

41 37₃ 42

Fig. 9c

7/16

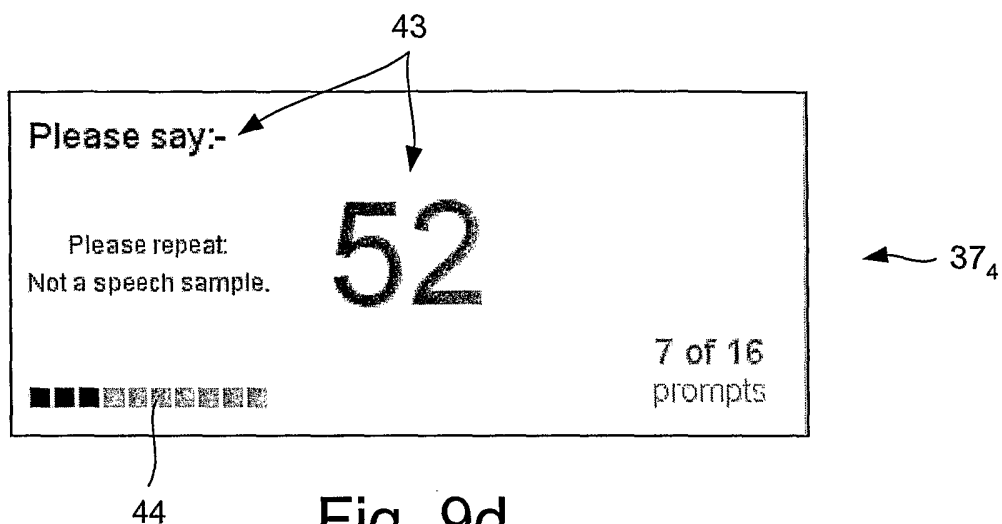


Fig. 9d

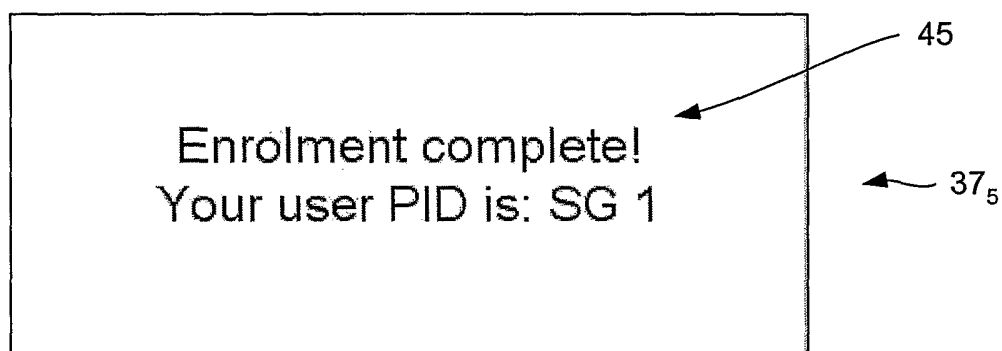


Fig. 9e

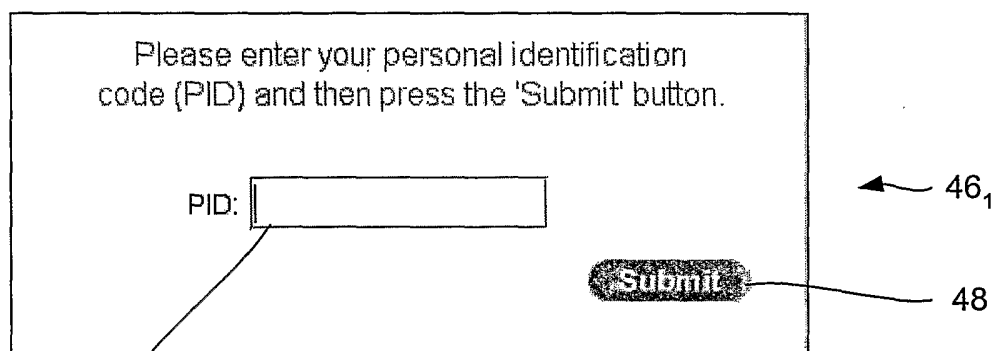


Fig. 11a

8/16

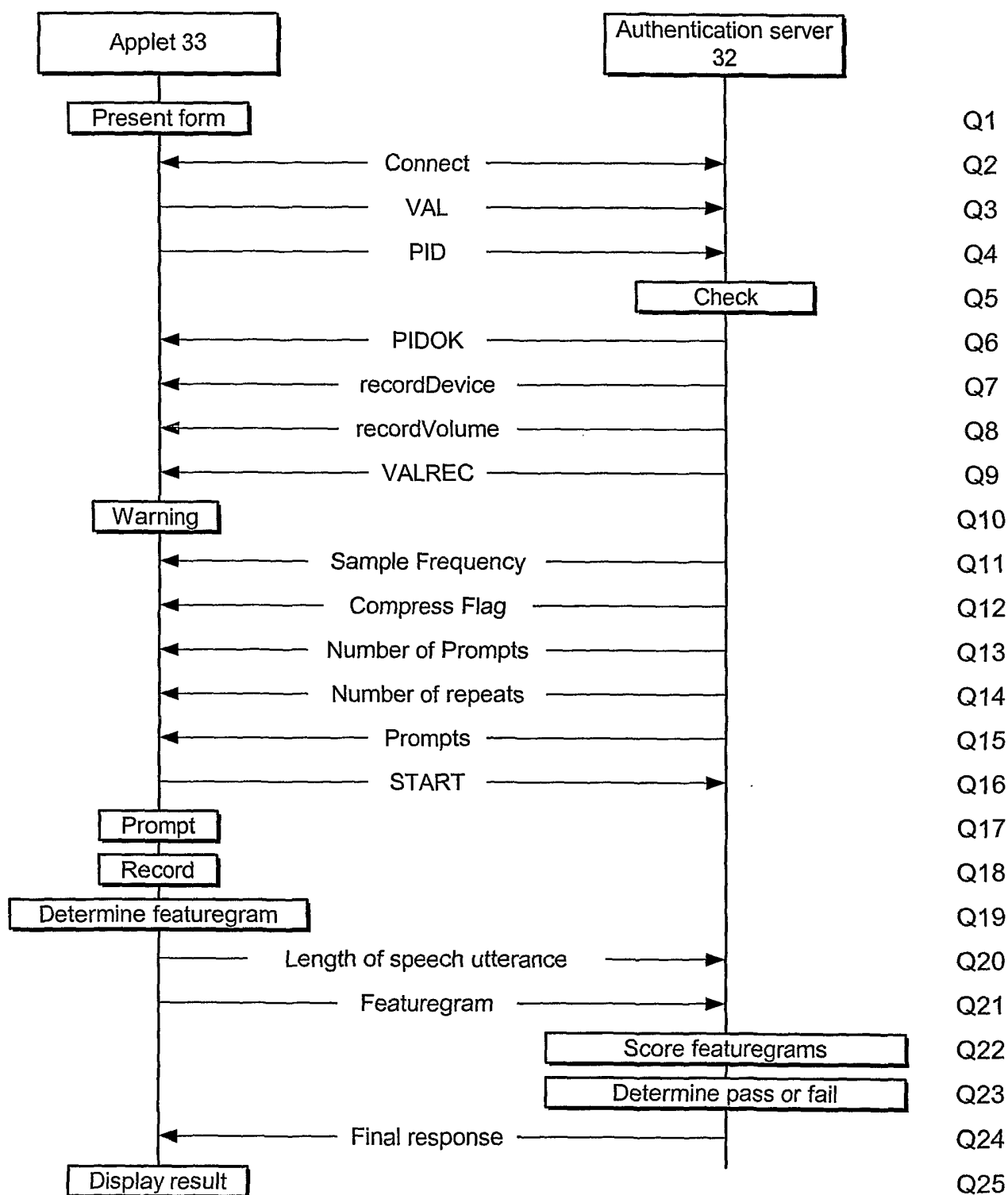


Fig. 10

9/16

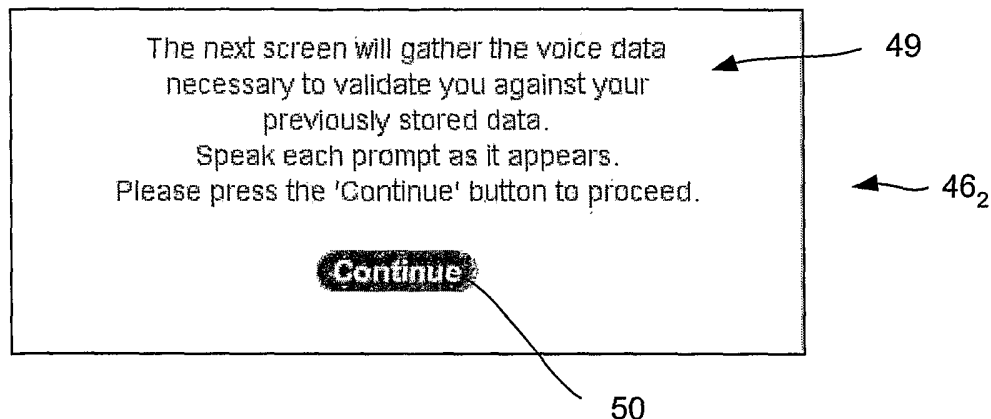


Fig. 11b

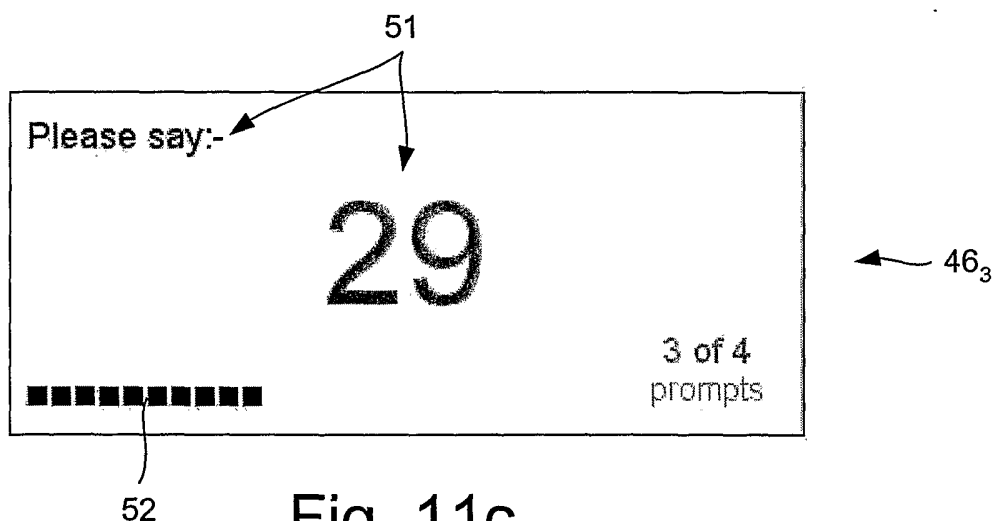


Fig. 11c

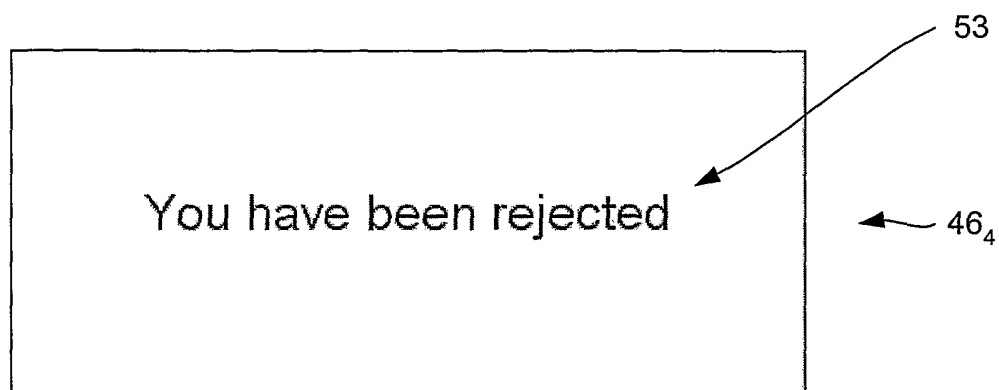
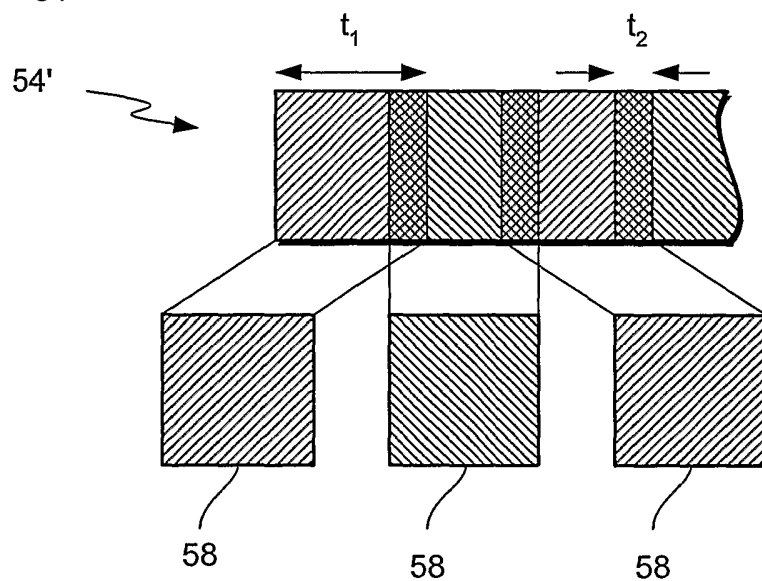
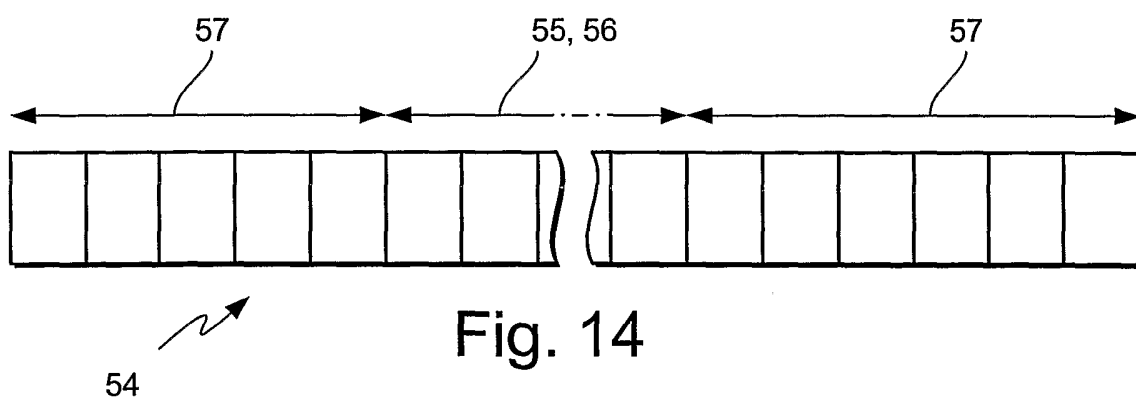
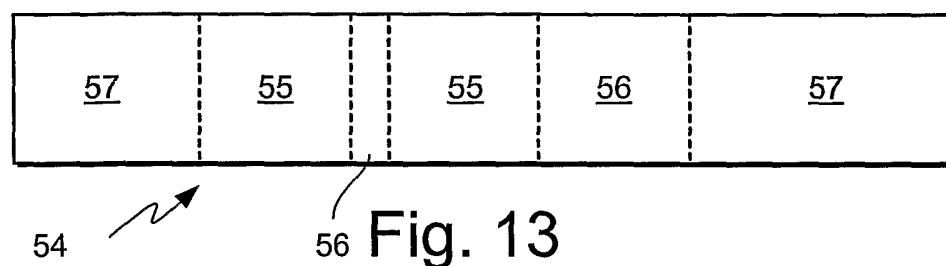
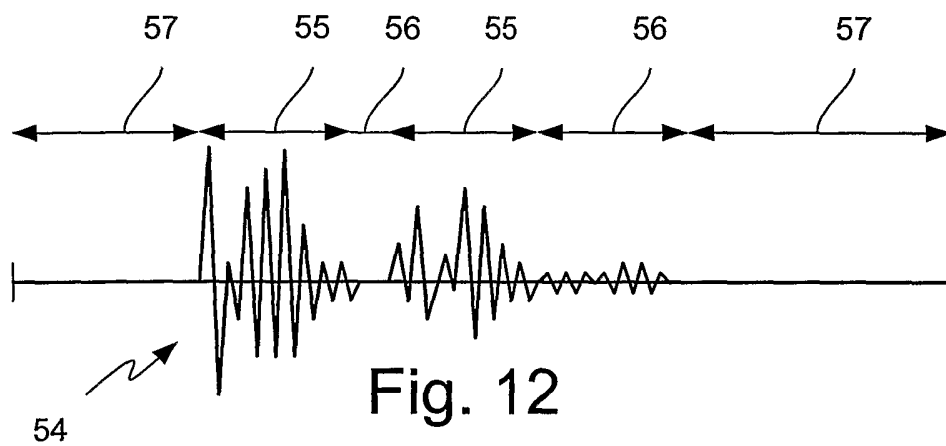


Fig. 11d

10/16



11/16

**Featuregram
generation**

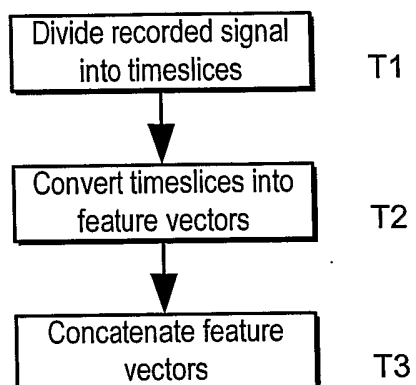


Fig. 16

**Featurevector
generation**

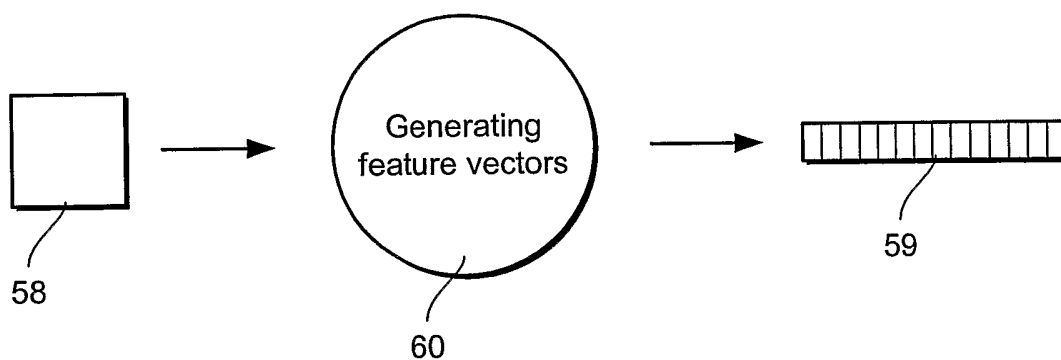


Fig. 17

**Featuregram
generation**

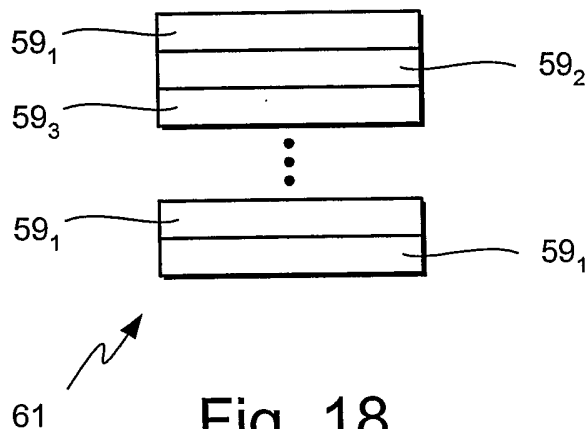


Fig. 18

12/16

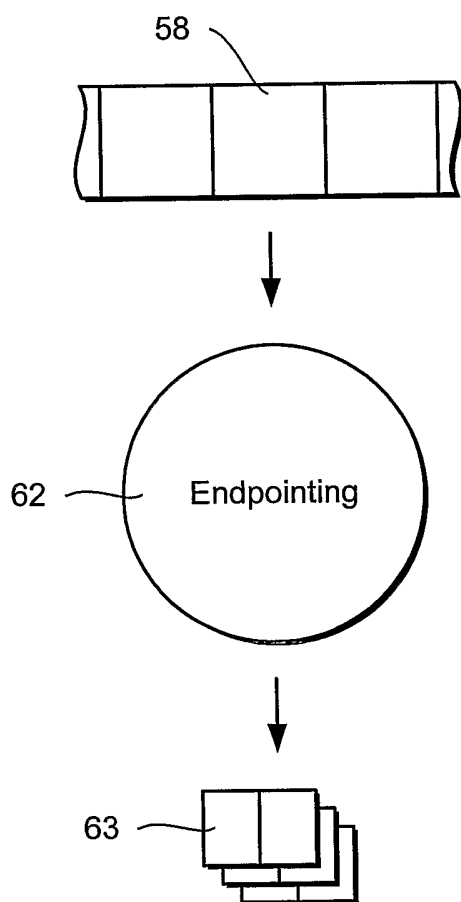
**Explicit
Endpointing**

Fig. 19

13/16

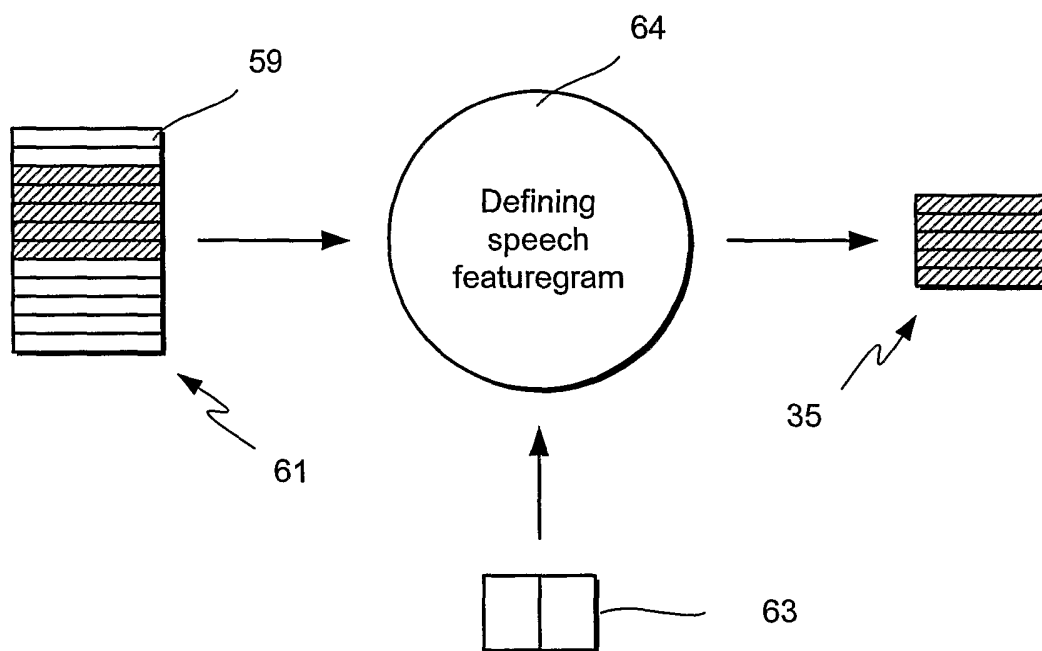


Fig. 20

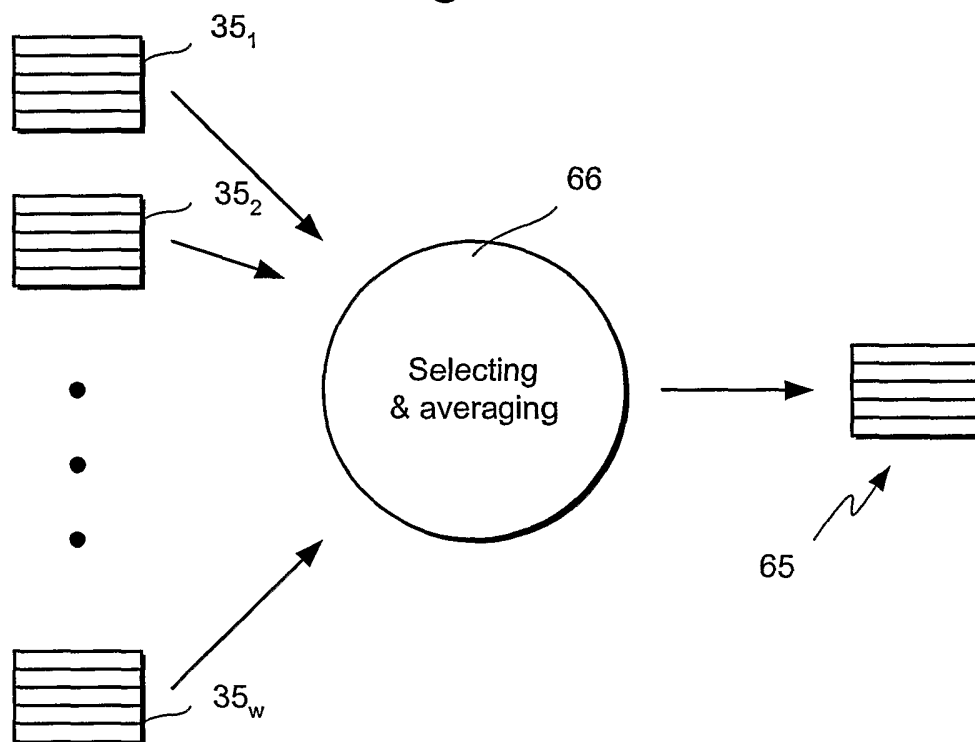


Fig. 21

14/16

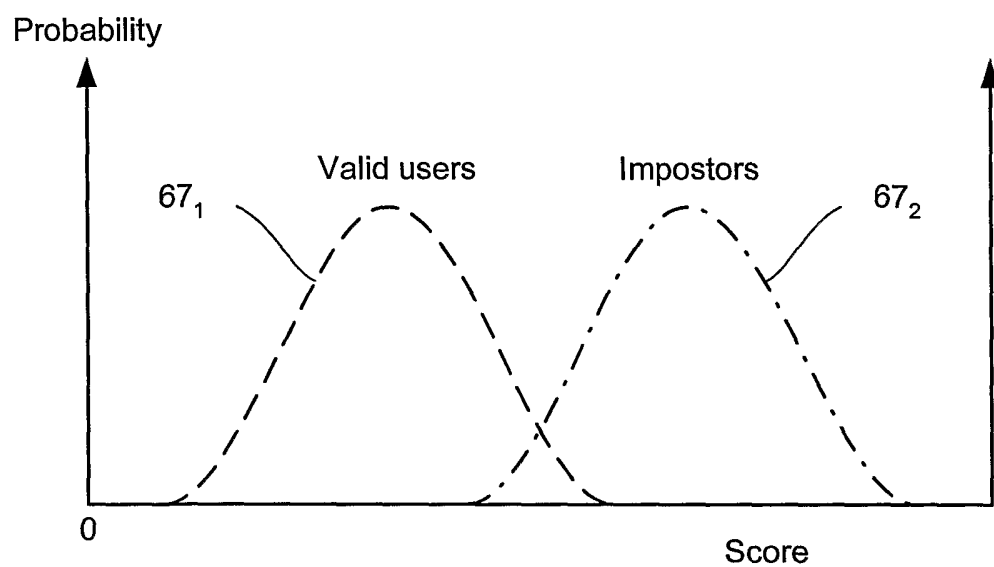


Fig. 22

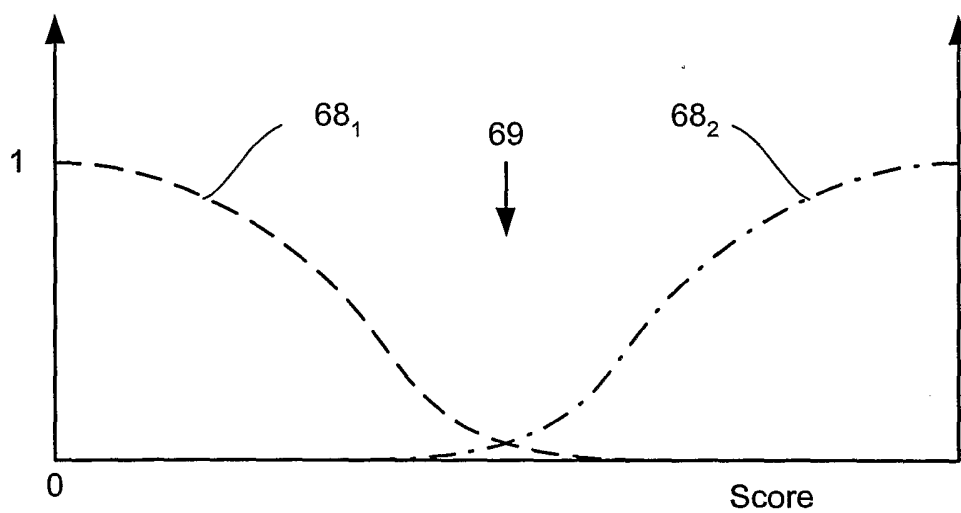


Fig. 23

15/16

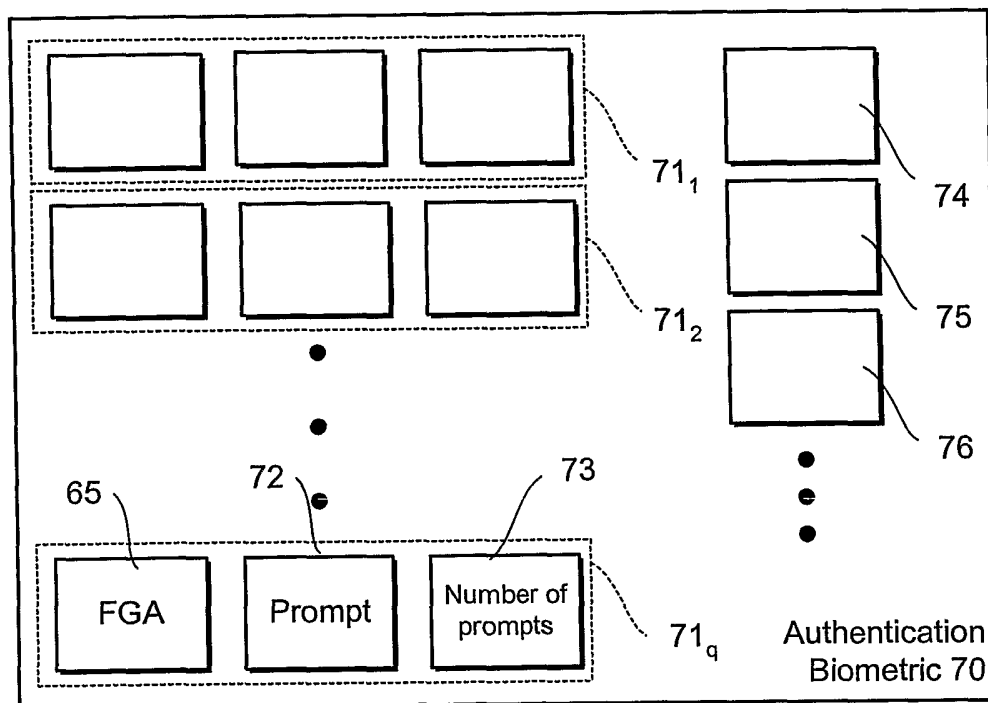


Fig. 24

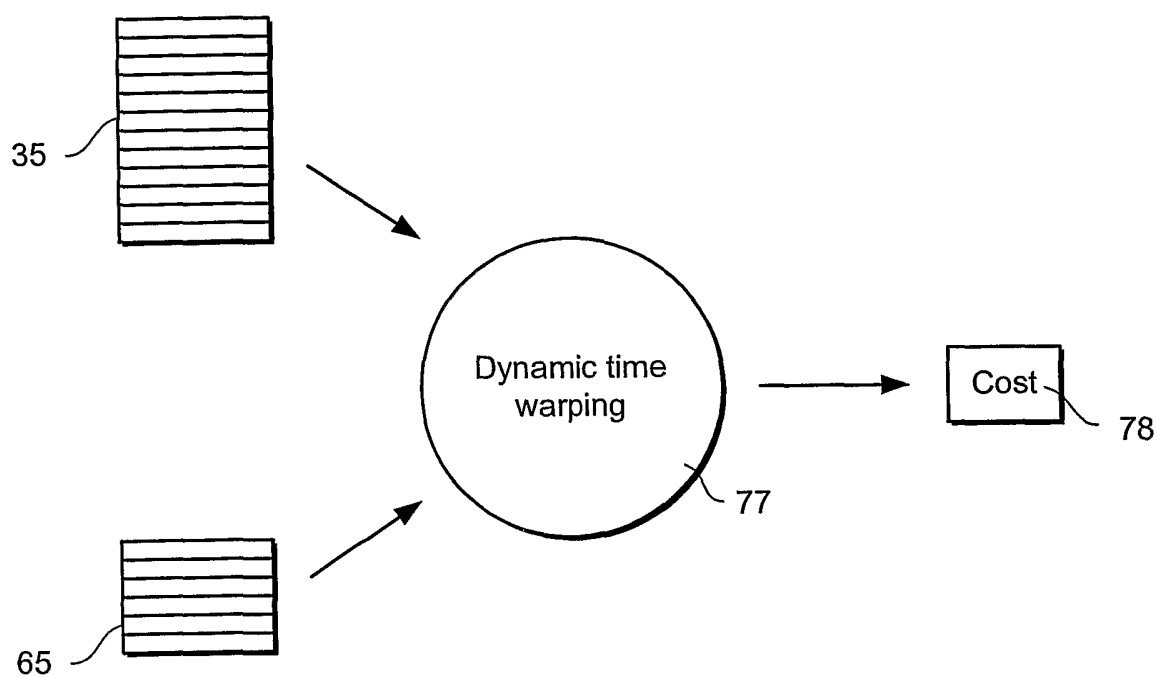


Fig. 25

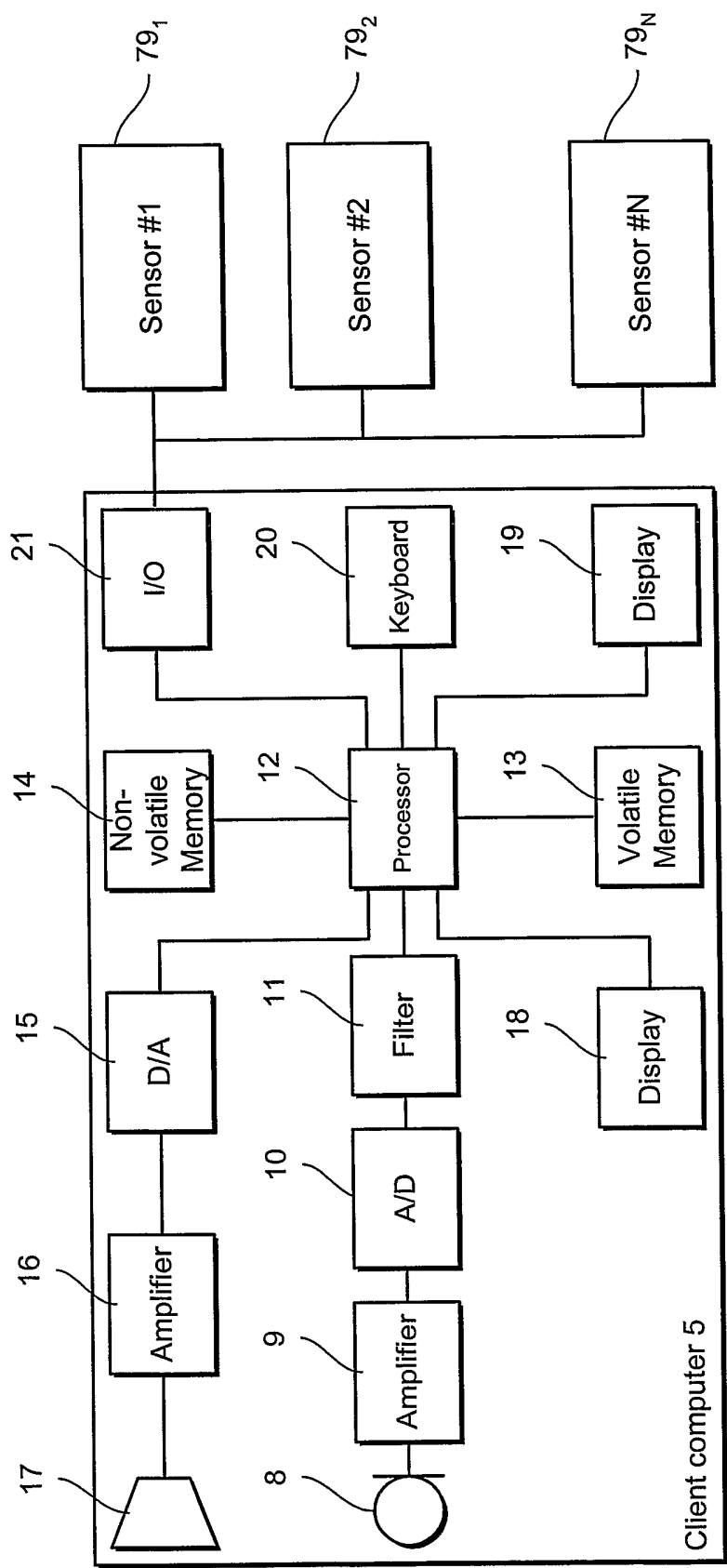


Fig. 26