Description

This invention relates to a speech recognition system capable of dealing with a plurality of users.

Single-speaker dependent speech recognisers are known, for example for connection to a telephone, in which the user "trains" the recogniser with his or her voice patterns. These patterns are stored as "templates" in Random Access Memory (RAM) within the speech recogniser. When a user wishes to make a phone call, he activates the recogniser, for example by lifting the receiver, and the recogniser uses his templates to interpret voiced instructions, for example of the telephone number the user wishes to dial. When other users wish to use the speech recogniser, they have to go through the training procedure, and when this happens the original user's templates are deleted. Therefore, if the original user needs to use the speech recogniser again, he must retrain the recogniser first.

In multi-user environments, for example a Private Automatic Branch Exchange (PABX), one recogniser could be provided to serve all the extensions connected to the PABX. However, in this case, the user at every extension would probably have to go through the tedium of training the recogniser every time he wished to make a call.

An alternative arrangement would be to provide a speech recogniser at each extension, so that each will be trained only by that extension's user. However, such an arrangement would be prohibitively expensive.

A better alternative is to provide a few recognisers, connected to the PABX and shared between the many users. In such a system, a common database is provided containing the templates of all the users. When any particular user activates his extension, one of the quiescent recognisers is arbitrarily selected and assigned to his use, and his templates are then loaded into that recogniser. A micro-processor is provided to control the loading of the templates. The rate at which templates are loaded into the recognisers depends upon the speed of the micro-processor, and the more recognisers wanting access to the database there are, the more complex the micro-processor is required to be. Every recogniser is assigned a priority, and the micro-processor deals with requests for service in strict priority order; so if a lower priority recogniser has access and is loading templates, and a higher priority recogniser requests access, service to the lower device is suspended while the higher device is served. Only when the higher priority device has finished is the lower priority device allowed to continue. This type of interruption may continually occur, so that a low priority device may be delayed while several higher priority devices are served, and the overall system suffers from the disadvantage of not being totally user transparent.

According to the invention there is provided a speech-recognition system for the provision of a speech recognition facility to a multiplicity of user terminals, the system comprising: speech recognition means connected via a switching node to a plurality of user terminals, there being a smaller number of speech recognition means than there are user terminals; and

a database containing respective templates for each of the user terminals;

characterised in that

there is a plurality of speech recognition means, and in that the arrangement is such that, when a user terminal is activated, one of the recognition means is assigned to deal with voiced instructions input via that user terminal; the database being arranged, in use, to provide a copy of relevant speaker-dependent templates to the assigned recognition means and in that there is provided control means connected to the speech recognition means and arranged, in use, to receive requests for templates from the recognition means, and to instruct the database to provide complete copies of the requested templates in the order in which they are requested.

Advantageously, the control means comprises a cyclic channel selector and a clock, the cyclic channel selector and the clock being interconnected by a switch which is controlled by the requests for copies of templates from the recognisers.

A preferred embodiment of the invention will now be described, by way of example, with reference to the accompanying drawings, in which:

Figure 1 shows a schematic illustration of a preferred embodiment of the invention;

Figure 2a shows a schematic illustration of a controller according to a preferred embodiment of the invention; and

Figure 2b shows a realisation of the controller of Figure 2a.

Referring to the drawings, Figure 1 shows a PABX 1 having n extensions 2 connected to a bank of m speech recognisers 3. The bank of speech recognisers 3 has access to a database 4a containing n sets of templates 4, one set corresponding to each extension 2, all stored in a memory device 5, for example a Winchester disk drive. A controller 6 is also connected to the bank of recognisers 3.

When one of the extensions 2 is activated, by a user picking up the receiver, one of the quiescent speech recognisers 3 in the bank is assigned to that extension. That recogniser 3 then ceases to be available for assignment to another extension 2. Once a speech recogniser 3 has been assigned, it requests the templates 4 associated with the rel-

evant activated extension 2, and these templates are fetched from the memory device 5, and loaded into the recogniser. The user may then give voiced instructions to the recogniser 3 for handling the call.

Once no more instructions are expected from an extension 2, for example once a call has been set up, the recogniser 3 assigned to deal with that extension is returned to its quiescent state.

The requests from the recognisers 3 are handled by the controller 6 on a first-come first-served basis. The controller 6 assigns all the recognisers 3 equal priority, and all subsequent requests are stacked, in order of arrival, until previously-requested templates 4 have been loaded.

The controller 6, shown in more detail in Figure 2a, comprises a cyclic channel selector 7 having a parallel output which is connected to all the recognisers 3 in the bank. The controller 6 also comprises a clock 8 having an output connected to the channel selector 7. Every clock pulse received by the selector 7 causes the parallel output to be incremented. Every recogniser 3 has an output, and these outputs are connected together to perform OR-logic, and thus to provide a control signal which stops the channel selector 7 from continuing to cycle. This is shown schematically as a control signal to a switch 9 which opens the circuit between the clock 8 and the selector 7.

Each recogniser 3 has a user-selectable channel identification number. When access to the memory device 5 is required, to load a set of templates 4, a given recogniser 3 scans the parallel output of the channel selector 7, and compares this with the identification number until there is a match. At this point, the recogniser 3 sends a high control signal, and freezes the selector's parallel output. The recogniser 3 then has complete uninterrupted access to the memory device 5. Once the templates 4 have been loaded into the recogniser 3, it ceases to send the high control signal, and the selector 7 continues to cycle until it is again halted.

One of the advantages of this system is that the controller 6 does not require any intelligence, and may be realised with passive devices, as shown in Figure 2b. In this realisation, open-collector drivers 10 are installed on the output lines of the recognisers 3, the lines all being connected together to perform OR logic; and the resulting single line, shown connected to a pull-up resistor 11, is fed to the switch 9, embodied by a D-type flip-flop.

Of course, the system may be arranged to operate when more than one user is expected to use any particular extension 2. In this case, each extension 2 may indicate to a recogniser 3 a series of users in a preferred order. The recogniser 3 will then request the templates 4 of the most preferred user, and check if these match the actual user. If not, the templates 4 of the second most preferred user may be requested, and so on.

It should be appreciated that the user terminals (extensions 2) and the recognisers 3 do not have to be wire-linked as shown in Figure 1. The user terminals 2 could be portable transceivers, for example in a cellular radio system, connected to the recognisers 3 by radio channels. In this case, a PABX would be inappropriate, but any suitable type of switching node could be provided within box 1. In any event, the user terminals 2 would not necessarily be activated by lifting the receiver, but other means, for example a switch, could be used.

Claims

1. A speech-recognition system for the provision of a speech recognition facility to a multiplicity of user terminals (2), the system comprising: speech recognition means (3) connected via a switching node (1) to a plurality of user terminals (2), there being a smaller number of speech recognition means (3) than there are user terminals (2); and

   a database (4a) containing respective templates (4) for each of the user terminals (2);

   characterised in that

   there is a plurality of speech recognition means (3), and in that the arrangement is such that, when a user terminal (2) is activated, one of the recognition means (3) is assigned to deal with voiced instructions input via that user terminal (2); the database (4a) being arranged, in use, to provide a copy of relevant speaker-dependent templates (4) to the assigned recognition means (3) and in that there is provided control means (6) connected to the speech recognition means (3) and arranged, in use, to receive requests for templates (4) from the recognition means (3), and to instruct the database (4a) to provide complete copies of the requested templates (4) in the order in which they are requested.

2. A speech recognition system as claimed in claim 1, wherein voiced instructions input by any one of a known group of users via a particular user terminal (2) can be dealt with.

3. A speech recognition system as claimed in claim 2, wherein when a user terminal (2) is activated by a particular user the speech recognition means (3) assigned to the terminal identifies the templates (4) relevant to the user.

evant activated extension 2, and these templates are fetched from the memory device 5, and loaded into the recogniser. The user may then give voiced instructions to the recogniser 3 for handling the call.

Once no more instructions are expected from an extension 2, for example once a call has been set up, the recogniser 3 assigned to deal with that extension is returned to its quiescent state.

The requests from the recognisers 3 are handled by the controller 6 on a first-come first-served basis. The controller 6 assigns all the recognisers 3 equal priority, and all subsequent requests are stacked, in order of arrival, until previously-requested templates 4 have been loaded.

The controller 6, shown in more detail in Figure 2a, comprises a cyclic channel selector 7 having a parallel output which is connected to all the recognisers 3 in the bank. The controller 6 also comprises a clock 8 having an output connected to the channel selector 7. Every clock pulse received by the selector 7 causes the parallel output to be incremented. Every recogniser 3 has an output, and these outputs are connected together to perform OR-logic, and thus to provide a control signal which stops the channel selector 7 from continuing to cycle. This is shown schematically as a control signal to a switch 9 which opens the circuit between the clock 8 and the selector 7.

Each recogniser 3 has a user-selectable channel identification number. When access to the memory device 5 is required, to load a set of templates 4, a given recogniser 3 scans the parallel output of the channel selector 7, and compares this with the identification number until there is a match. At this point, the recogniser 3 sends a high control signal, and freezes the selector's parallel output. The recogniser 3 then has complete uninterrupted access to the memory device 5. Once the templates 4 have been loaded into the recogniser 3, it ceases to send the high control signal, and the selector 7 continues to cycle until it is again halted.

One of the advantages of this system is that the controller 6 does not require any intelligence, and may be realised with passive devices, as shown in Figure 2b. In this realisation, open-collector drivers 10 are installed on the output lines of the recognisers 3, the lines all being connected together to perform OR logic; and the resulting single line, shown connected to a pull-up resistor 11, is fed to the switch 9, embodied by a D-type flip-flop.

Of course, the system may be arranged to operate when more than one user is expected to use any particular extension 2. In this case, each extension 2 may indicate to a recogniser 3 a series of users in a preferred order. The recogniser 3 will then request the templates 4 of the most preferred user, and check if these match the actual user. If not, the templates 4 of the second most preferred user may be requested, and so on.

It should be appreciated that the user terminals (extensions 2) and the recognisers 3 do not have to be wire-linked as shown in Figure 1. The user terminals 2 could be portable transceivers, for example in a cellular radio system, connected to the recognisers 3 by radio channels. In this case, a PABX would be inappropriate, but any suitable type of switching node could be provided within box 1. In any event, the user terminals 2 would not necessarily be activated by lifting the receiver, but other means, for example a switch, could be used.

Claims

1. A speech-recognition system for the provision of a speech recognition facility to a multiplicity of user terminals (2), the system comprising: speech recognition means (3) connected via a switching node (1) to a plurality of user terminals (2), there being a smaller number of speech recognition means (3) than there are user terminals (2); and

   a database (4a) containing respective templates (4) for each of the user terminals (2);

   characterised in that

   there is a plurality of speech recognition means (3), and in that the arrangement is such that, when a user terminal (2) is activated, one of the recognition means (3) is assigned to deal with voiced instructions input via that user terminal (2); the database (4a) being arranged, in use, to provide a copy of relevant speaker-dependent templates (4) to the assigned recognition means (3) and in that there is provided control means (6) connected to the speech recognition means (3) and arranged, in use, to receive requests for templates (4) from the recognition means (3), and to instruct the database (4a) to provide complete copies of the requested templates (4) in the order in which they are requested.

2. A speech recognition system as claimed in claim 1, wherein voiced instructions input by any one of a known group of users via a particular user terminal (2) can be dealt with.

3. A speech recognition system as claimed in claim 2, wherein when a user terminal (2) is activated by a particular user the speech recognition means (3) assigned to the terminal identifies the templates (4) relevant to the user.

* C L M S *

4. A speech recognition system as claimed in claim 3, wherein the user terminal (2), when activated by any of the known group of users, identifies the members of the known group of users to the assigned speech recognition means (3), indicating the preferred order of these users, the arrangement being such that the recognition means (3) then requests the templates (4) of the most preferred user and, when a copy of these is received from the database (4a), tests if they correspond to the current user and, if they do not, requests the templates (4) of the next most preferred user, tests for correspondence with these and continues this process until the current user's templates (4) are identified.

5. A speech recognition system as claimed in any previous claim, wherein the control means (6) comprises a cyclic channel selector (7) and a clock (8), the cyclic channel selector (7) and the clock (8) being interconnected by a switch (9) which is controlled by the requests for copies of the templates (4) from the recognition means (3).

6. A system as claimed in any previous claim, wherein the user terminals (2) comprise portable transceivers.

7. A system as claimed in any one of the preceding claims wherein the switching node (1) is configured as a telephone exchange.

8. A system as claimed in claim 7, wherein the switching node (1) is a private automatic branch exchange and the user terminals (2) are telephone extensions of that exchange.

9. A telephone exchange for connection to a multiplicity of telephone instruments (2), the exchange including speech recognition means (3) for the provision of a speech recognition facility to telephone instruments (2) connected to the exchange, the speech recognition means (3) being provided with a store (4a) for speaker-dependent templates for use in recognition, characterised in that there is a plurality of speech recognition (3) means but many fewer than the number of telephone instruments (2) which the exchange can serve; the exchange including control means (6) which, in use, provide that when a telephone instrument (2) connected to the exchange is activated one of the recognition means (3) is assigned to deal with voiced instructions input via that instrument (2), and the store (4a), in which, in use, speaker-dependent templates (4) are stored, is arranged to provide copies of relevant speaker-dependent templates (4) to the assigned recognition means (3), the control means (6) being arranged, in use, to receive requests for templates (4) and to instruct the database (4a) to provide complete copies of the requested templates (4) in the order in which they are requested.

Patentansprüche

1. Spracherkennungssystem zur Bereitstellung einer Spracherkennungsanlage für eine Vielzahl von Benutzerendstellen (2), wobei das System aufweist: eine Spracherkennungseinrichtung (3), verbunden über einen Schaltknoten (1) mit einer Vielzahl von Benutzerendstellen (2), wobei es eine kleinere Anzahl Spracherkennungseinrichtungen (3) als Benutzerendstellen sind, und
eine Datenbank (4a), die jeweils Schablonen (4) für jede der Benutzerendstellen (2) enthält;
dadurch gekennzeichnet, daß
es eine Vielzahl von Spracherkennungseinrichtungen (3) gibt, und dadurch, daß die Anordnung so ist, daß, wenn eine Benutzerendstelle (2) aktiviert ist, eine der Erkennungseinrichtungen (3) zugeordnet ist ist sich mit sprachlichen Anweisungen zu befassen, die über diese Benutzerendstelle (2) eingegeben sind; die Datenbank (34a) darauf ausgerichtet ist, bei Benutzung eine Kopie wesentlicher sprecherabhängiger Schablonen (4) den zugeordneten Erkennungseinrichtungen (3) bereitzustellen, und dadurch, daß eine Steuereinrichtung (6) bereitgestellt ist, die mit den Spracherkennungseinrichtungen (3) verbunden ist, und die bei Gebrauch eingerichtet ist, Anforderungen nach Schablonen (4) von den Erkennungseinrichtungen (3) zu empfangen und die Datenbank (4a) anzuweisen, vollständige Kopien der angeforderten Schablonen (4) in der Reihenfolge, in der sie angefordert wurden, bereitzustellen.

2. Spracherkennungssystem nach Anspruch 1, das sich mit Sprachanweisungen, die durch irgendeine einer bekannten Benutzergruppe über eine einzelne Benutzerendstelle (2) eingegeben sind, befassen kann.

3. Spracherkennungssystem nach Anspruch 2, bei dem, wenn eine Benutzerendstelle (2) aktiviert ist durch einen einzelnen Benutzer, die Spracherkennungseinrichtung (3), die der Endstelle zugewiesen ist, die Schablonen (4), die dem Benutzer entsprechen, identifiziert.

* C L M S *

4. Spracherkennungssystem nach Anspruch 3, bei dem die Benutzerendstelle (2), wenn sie von irgendeinem der bekannten Benutzergruppe aktiviert wird, die Mitglieder der bekannten Benutzergruppe der zugewiesenen Spracherkennungseinrichtung (3) identifiziert, wobei die bevorzugte Reihenfolge dieser Benutzer angeben wird, wobei die Anordnung dergestalt ist, daß die Erkennungseinrichtung (3) dann die Schablonen (4) des bevorzugtesten Benutzers anfordert und, wenn eine Kopie derselben von der Datenbank (4a) empfangen wird, prüft, ob sie dem aktuellen Benutzer entsprechen und, wenn sie es nicht tun, die Schablonen (4) des nächsten bevorzugtesten Benutzers anfordert, die Übereinstimmung mit diesem prüft und diesen Prozeß fortsetzt, bis die Schablonen (4) des aktuellen Benutzers identifiziert sind.

5. Spracherkennungssystem nach einem der vorhergehenden Ansprüche, bei dem die Steuereinrichtung (6) einen zyklischen Kanalauswähler (7) und einen Taktgeber (8) aufweist, wobei der zyklische Kanalauswähler (7) und der Taktgeber (8) durch einen Schalter (9) miteinander verbunden sind, der durch die Anforderungen nach Kopien der Schablonen (4) von der Erkennungseinrichtung (3) gesteuert ist.

6. System nach einem der vorhergehenden Ansprüche, bei dem die Benutzerendgeräte (2) tragbare Sender-Empfänger aufweisen.

7. System nach einem der vorhergehenden Ansprüche, bei dem der Schaltknoten (1) als eine Fernsprechvermittlung konfiguriert ist.

8. System nach Anspruch 7, bei dem der Schaltknoten (1) eine Wähl-Nebenstellenanlage ist und die Benutzerendstellen (2) Fernsprechernebenanschlüsse dieser Vermittlung sind.

9. Fernsprechvermittlung zur Verbindung mit einer Vielzahl von Fernsprechvorrichtungen (2), wobei die Vermittlung Spracherkennungseinrichtungen (3) zur Bereitstellung einer Spracherkennungsanlage für die Fernsprecheinrichtungen (2), die mit der Vermittlung verbunden sind, einschließt, wobei die Spracherkennungseinrichtungen (3) mit einem Speicher (4a) für sprecherabhängige Schablonen zum Gebrauch bei der Erkennung versehen sind, dadurch gekennzeichnet, daß es eine Vielzahl von Spracherkennungseinrichtungen (3) gibt, aber viel weniger als die Anzahl von Fernsprecheinrichtungen (2), die die Vermittlung bedienen kann; die Vermittlung eine Steuereinrichtung (6) einschließt, die bei Gebrauch gewährleistet, daß wenn eine Fernsprecheinrichtung (2), die mit der Vermittlung verbunden ist, aktiviert ist, einer der Erkennungseinrichtungen (3) zugewiesen ist, um sich mit den Sprachanweisungen, die über diese Einrichtung (2) eingegeben sind, zu befassen, und daß der Speicher (4a), in dem bei Gebrauch sprecherabhängige Schablonen (4) gespeichert sind, eingerichtet ist, um Kopien wesentlicher sprecherabhängiger Schablonen (4) der zugewiesenen Erkennungseinrichtung (3) bereitzustellen, wobei die Steuereinrichtung (6) eingerichtet ist, bei Gebrauch Anforderungen für Schablonen (4) zu empfangen und die Datenbank (4a) anzuweisen vollständige Kopien der geforderten Schablonen (4), in der Reihenfolge, in der sie angefordert werden, bereitzustellen.

Revendications

1. Un système de reconnaissance de la parole destiné à conférer une capacité de reconnaissance de la parole à une multiplicité de terminaux d'utilisateurs (2), le système comprenant : un moyen de reconnaissance de la parole (3) relié par l'intermédiaire d'un noeud de commutation (1) à une série de terminaux d'utilisateurs (2), le nombre de moyens de reconnaissance de la parole (3) étant inférieur à celui des terminaux d'utilisateurs (2); et

   une base de données (4a) contenant des gabarits respectifs (4) pour chacun des terminaux d'utilisateurs (2);

   caractérisé en ce que

   il existe une série de moyens de reconnaissance de la parole (3), et en ce que l'agencement est tel que l'un des moyens de reconnaissance est assigné, lorsqu'un terminal d'utilisateur (2) est activé, au traitement des instructions phoniques entrées par l'intermédiaire de ce terminal d'utilisateur; la base de données étant agencée pour fournir en utilisation, au moyen de reconnaissance assigné (3), une copie des gabarits appropriés (4) en fonction de la personne qui parle et en ce qu'il est prévu un moyen de commande (6), relié au moyen de reconnaissance de la parole (3) et agencé pour recevoir, en utilisation, des demandes de gabarit (4) provenant du moyen de reconnaissance (3) et pour des instructions à la base de données (4a) afin de compléter des copies complètes des gabarits demandés (4) dans l'ordre dans lequel elles sont demandées.

2. Un système de reconnaissance de la parole selon la revendication 1, dans lequel peuvent être traitées des instructions phoniques entrées, par un utilisateur quelconque inclus dans

* C L M S *

un groupe connu, par l'intermédiaire d'un terminal particulier d'utilisateur (2).

3. Un système de reconnaissance de la parole selon la revendication 2, dans lequel le moyen de reconnaissance de la parole (3) assigné à un terminal d'utilisateur (2) identifie, lorsque le terminal est activé par un utilisateur particulier, les gabarits appropriés (4) à l'utilisateur.

4. Un système de reconnaissance de la parole selon la revendication 3 dans lequel le terminal d'utilisateur (2) identifie au moyen de reconnaissance de la parole assigné (3), lorsqu'il est activé par un utilisateur quelconque inclus dans un groupe connu, les membres du groupe connu d'utilisateurs en indiquant l'ordre préféré de ces utilisateurs, l'agencement étant tel que le moyen de reconnaissance (3) demande ensuite les gabarits (4) de l'utilisateur particulièrement préféré et essaie, lorsqu'une copie de ceux-ci est reçue depuis la base de données (4), s'ils correspondent à l'utilisateur actuel, et demande, si tel n'est pas le cas, les gabarits (4) de l'utilisateur suivant dans l'ordre de préférence décroissant, essaie quant à la correspondance avec ceux-ci et continue ce processus jusqu'à ce que les gabarits de l'utilisateur actuel (4) soient identifiés.

5. Un système de reconnaissance de la parole selon l'une quelconque des revendications précédentes, dans lequel le moyen de commande comprend un sélecteur (7) de canal cyclique et une horloge (8), le sélecteur (7) de canal cyclique et l'horloge (8) étant reliés entre eux par un commutateur (9) qui est commandé par les demandes de copies des gabarits (4) provenant du moyen de reconnaissance (3).

6. Un système selon une revendication précédente quelconque dans lequel les terminaux d'utilisateurs (2) comprennent des émetteurs-récepteurs portables.

7. Un système selon l'une quelconque des revendications précédentes dans lequel le noeud de commutation (1) est configuré sous forme de standard téléphonique.

8. Un système selon la revendication 7 dans lequel le noeud de commutation (1) est un central automatique privé et les terminaux d'utilisateurs sont des postes téléphoniques de ce central.

9. Un central téléphonique à relier à une série d'instruments téléphoniques (2), le central incluant un moyen de reconnaissance de la parole (3) pour conférer à des instruments téléphoniques (2) reliés au central une possibilité de reconnaissance de la parole, le moyen de reconnaissance de la parole (3) étant pourvu d'une mémoire (4a) de gabarits, dépendant de la personne qui parle, à utiliser pour la reconnaissance, caractérisé en ce qu'il existe une série de moyens de reconnaissance (3) de la parole, mais en nombre beaucoup moindre que le nombre des instruments téléphoniques (2) que le standard peut desservir; le standard incluant un moyen de commande (6) qui assure, en utilisation, que l'un des moyens de reconnaissance (3) est assigné, lorsqu'un instrument téléphonique (2) relié au central est activé, au traitement d'instructions phoniques entrées par l'intermédiaire de cet instrument (2), et la mémoire (4a) dans laquelle des gabarits (4) dépendant des personnes qui parlent sont mémorisées en utilisation, est agencée pour fournir, au moyen de reconnaissance assigné (3), des copies des gabarits appropriés (4) dépendant de la personne qui parle, le moyen de commande (6) étant agencé en utilisation pour recevoir des demandes de gabarits (6) et pour donner, à la base de données (4), des instructions de fournir des copies complètes des gabarits demandés (4) dans l'ordre dans lequel elles sont demandées.

* C L M S *