

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2004-309965

(P2004-309965A)

(43) 公開日 平成16年11月4日(2004.11.4)

(51) Int.Cl.⁷

G10L 15/28

G10L 15/00

F I

G10L 3/00

511

G10L 3/00

551B

テーマコード (参考)

5D015

審査請求 有 請求項の数 6 O L (全 11 頁)

(21) 出願番号

特願2003-106567 (P2003-106567)

(22) 出願日

平成15年4月10日 (2003. 4. 10)

(71) 出願人 398018021

株式会社アドバンスト・メディア

東京都豊島区東池袋3丁目1番1号サンシ

ャイン60 48F

(74) 代理人 100082669

弁理士 福田 賢三

(74) 代理人 100095337

弁理士 福田 伸一

(74) 代理人 100061642

弁理士 福田 武通

(74) 代理人 100095061

弁理士 加藤 恭介

最終頁に続く

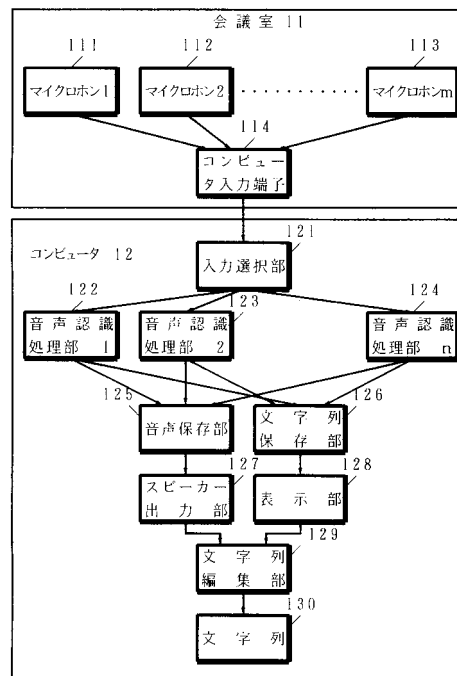
(54) 【発明の名称】 会議録音・書き起こしシステム

(57) 【要約】

【課題】本発明は、会議等に参加した者の発言を議事録という形で時系列的に書き起こすことができる会議録音・書き起こしシステムに関するものである。

【解決手段】マイクロホンは、当該会議の出席者の数だけあり、前記出席者のそれぞれに一つずつが割り当てられている。入力選択部は、前記マイクロホンの内、音声出力のある予め決められた数だけを選択するようになっている。音声認識処理部の数は、前記入力選択部において選択できるマイクロホンの数と同じである。前記それぞれの音声認識処理部によって認識された音声は、そのまま音声保存部に保存され、認識されたデータは、文字列になり、文字列保存部に保存される。文字列保存部に保存された文字列は、表示手段に表示され、編集者が前記表示手段に表示された認識結果と音声を聞きながら会議の議事録を書き起こすことができる。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

複数の話者の発言を文字列として編集することができる会議録音・書き起こしシステムにおいて、
複数のマイクロホンと、
前記マイクロホンの内、予め決められた数以下のマイクロホンの音声出力を選択する入力選択部と、
前記予め決められた数の音声認識処理部と、
前記それぞれの音声認識処理部によって認識された音声およびその文字列を保存する音声保存部および文字列保存部と、
前記音声保存部の音声と、文字列保存部の文字列に基づいて時系列的な文字列に編集する文字列編集部と、
を少なくとも備えたことを特徴とする会議録音・書き起こしシステム。

10

【請求項 2】

前記入力選択部は、前記マイクロホンから入力した零個の信号、一個の信号、複数個の信号を選択し、前記選択された信号を異なる音声認識処理部にそれぞれ出力することを特徴とする請求項 1 に記載された会議録音・書き起こしシステム。

【請求項 3】

前記入力選択部は、前記マイクロホンから入力した信号をデジタルまたはアナログ信号として異なる音声認識処理部にそれぞれ出力することを特徴とする請求項 1 または請求項 2 に記載された会議録音・書き起こしシステム。

20

【請求項 4】

前記入力選択部は、前記信号の強さが予め決められたレベルを一定時間連続して超えているものを先着順に選択することを特徴とする請求項 1 ないし請求項 3 のいずれか 1 項に記載された会議録音・書き起こしシステム。

【請求項 5】

前記入力選択部は、コンピュータソフトウェアであることを特徴とする請求項 1 ないし請求項 4 のいずれか 1 項に記載された会議録音・書き起こしシステム。

【請求項 6】

前記各音声認識処理部における音声認識結果を時系列に並べて表示する機能と、音声認識結果に対応する音声を再生する機能とを有することを特徴とする請求項 1 ないし請求項 5 のいずれか 1 項に記載された会議録音・書き起こしシステム。

30

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、会議等に出席した者の発言を議事録という形で時系列的に書き起こすことができる会議録音・書き起こしシステムに関するものである。本明細書においていう「会議」は、広い意味のものであり、通常の会議以外に、話し合い、集団面談、団体交渉等多数の発言者がいることを前提としている。

【0002】

40

【従来の技術】

図 3 は従来例の会議における議事録を時系列で書き起こすための概念図である。図 3 において、マイクロホン a 3 1 1、マイクロホン b 3 1 2、・・・、マイクロホン n 3 1 3 でひろい上げた音声は、たとえば、テーブルコードのような音声記憶手段 a 3 1 4、音声記憶手段 b 3 1 5、・・・、音声記憶手段 n 3 1 6 によって一時記憶される。

【0003】

前記各音声記憶手段 3 1 4 ないし 3 1 6 の内容は、たとえば、速記者によるワードプロセッサ、あるいは音声認識手段（図示されていない）等の文字列編集手段 a 3 1 7、・・・、n 3 1 9 によって文字列とした後、時系列編集手段 3 2 0 によって、文字列を時系列的に並べることができる。

50

【 0 0 0 4 】

さらに、一般的に、複数の話者が一つのマイクロホンを通じて音声記憶手段に記憶され、前記音声記憶手段に記憶されている話者の声を聞きながら、たとえば、速記者がワードプロセッサ等を用い、議事録として書き起こしていた。

【 0 0 0 5 】

また、特開平 5 - 1 2 2 4 6 号公報に記載されている音声文書作成装置は、複数の入力を選択して一つとし、無意味語を削除して漢字変換を容易にするとともに、口語調文章を所定の書式の文語調文章に作成する会議議事録編集の機能を備えている。

【 0 0 0 6 】

【 特許文献 1 】

特開平 5 - 1 2 2 4 6 号公報

【 0 0 0 7 】

【 発明が解決しようとする課題 】

一般に、会議議事録は、テープレコーダによって録音された録音テープを聞くことにより、ワードプロセッサやパーソナルコンピュータのキーボード入力を行って、文字列として、書き起こされていた。また、前記人手による書き起こしの代わりに、前記録音された録音テープの音声は、音声認識により、文字列に変換することが行われていた。

【 0 0 0 8 】

しかし、前記録音テープに録音されている会議中の音声は、一つのマイクロホンで複数人の発言を同時に記録しているため、人手または音声認識技術によっても、多くの問題を有していた。特に、発言者とマイクロホンとの距離が離れている場合、録音テープに録音されている発言者の録音内容は、周囲のノイズとともに、他の発言者の内容が重なり、内容を明瞭に文字列化することが困難であった。

【 0 0 0 9 】

図 3 に示すように、発言者一人一人にマイクロホンを用意して、議事録を録音テープにとることが考えられた。

【 0 0 1 0 】

しかし、複数のマイクロホンを用いて録音された録音テープは、発言内容を書き下ろす場合、発言のない無音区間を含む複数の音声を全て聞き取る必要があるため、議事録の作成が困難であるだけでなく、効率が非常に悪くなるという問題があった。また、異なるマイクロホンに入力された発言は、発言順に時系列を揃えて書き下ろす必要があり、作業を煩雑にするという問題があった。

【 0 0 1 1 】

前記問題は、たとえば、複数のマイクロホンからの音声をミキシングして、一つの録音テープに録音することで解消されるが、複数の発言者の発言が重なる場合、書き起こしが困難になる。

【 0 0 1 2 】

また、録音テープに録音された音声の聞き取りや書き起こしは、書き起こしたい発言者の頭出しや、前記発言者の不明瞭な部分を繰り返して聞く等の操作が必要になる。そして、前記操作は、再生、早送り、巻き戻し等、頻繁に行う必要があり、効率を悪くするという問題があった。

【 0 0 1 3 】

前記問題を解決する方法としては、多数のマイクロホンに対して、それぞれ音声認識処理を行うことである。しかし、一台のコンピュータで、発言者の数と同じ数の音声認識処理手段を備えておくことは、2 人ないし 3 人の会議を除いて、前記コンピュータにとって重い負荷となるため、現状では無理である。

【 0 0 1 4 】

また、音声認識処理は、普通、音量が最大のものを一つ選択して入力しているので、複数の発言がある場合、声の小さい発言者が無視される恐れがある。さらに、会議の議事録は、無音時間が問題になる場合もあるが、従来の書き起こし方法を使用すると、会議の時系

10

20

30

40

50

列が正確に表現できない。

【0015】

以上のような課題を解決するために、本発明は、複数の発言者から同時に発言があっても、重複録音がなく別々に、正確に録音され、かつ、一台のパーソナルコンピュータで音声認識処理ができるとともに、書き起こしおよび編集が容易である会議録音・書き起こしシステムを提供することを目的とする。

【0016】

本出願人は、発言者が同時に発言する場合、多くとも2人ないし3人程度であり、これ以上多くの発言者が発言を行った場合、会議が紛糾して、議事録をとる必要がないということに着目した。すなわち、本発明は、多数の発言者にそれぞれマイクロホンが備えられていても、マイクロホンの数だけ音声認識処理手段を設ける必要がなく、多くとも2ないし3程度にした点にある。 10

【0017】

【課題を解決するための手段】

(第1発明)

第1発明の会議録音・書き起こしシステムは、複数の話者の発言を文字列として編集することができるものであり、複数のマイクロホンと、前記マイクロホンの内、予め決められた数以下のマイクロホンの音声出力を選択する入力選択部と、前記予め決められた数の音声認識処理部と、前記それぞれの音声認識処理部によって認識された音声およびその文字列を保存する音声保存部および文字列保存部と、前記音声保存部の音声と、文字列保存部の文字列に基づいて時系列的な文字列に編集する文字列編集部と、を少なくとも備えたことを特徴とする。 20

【0018】

(第2発明)

第2発明の会議録音・書き起こしシステムにおいて、前記入力選択部は、前記マイクロホンから入力した零個の信号、一個の信号、複数個の信号を選択し、前記選択された信号を異なる音声認識処理部にそれぞれ出力することを特徴とする。

【0019】

(第3発明)

第3発明の会議録音・書き起こしシステムにおいて、前記入力選択部は、前記マイクロホンから入力した信号をデジタルまたはアナログ信号として異なる音声認識処理部にそれぞれ出力することを特徴とする。 30

【0020】

(第4発明)

第4発明の会議録音・書き起こしシステムにおいて、前記入力選択部は、前記信号の強さが予め決められたレベルを一定時間連続して超えているものを先着順に選択することを特徴とする。

【0021】

(第5発明)

第5発明の会議録音・書き起こしシステムにおいて、前記入力選択部は、コンピュータソフトウェアであることを特徴とする。 40

【0022】

(第6発明)

第6発明の会議録音・書き起こしシステムにおいて、前記各音声認識処理部における音声認識結果は、時系列に並べて表示されると同時に、音声認識結果に対する音声を再生する機能を有することを特徴とする。

【0023】

【発明の実施の形態】

(第1発明)

第1発明は、複数の話者の発言が同じ時間に重複していても、発言内容が正確な文字列と 50

して、時系列通りに編集することができる会議録音・書き起こしシステムである。マイクロホンは、当該会議の出席者の数だけあり、前記出席者のそれぞれに一つずつが割り当てられている。入力選択部は、前記マイクロホンの内、音声出力のある予め決められた数以下のマイクロホンの音声出力を選択するようになっている。また、前記マイクロホンの数は、必ずしも、出席者と同じにする必要がない。その際に、出席者は、名前を言ってから発言するとか、あるいは発言者の声紋によって識別される。

【0024】

通常の会議において、たとえば、3名以上の者が同時に発言することは少ないと仮定している。2名の発言者は、ある発言の終わりと始まり、合の手等において重複発言が発生するので、書き起こす必要がある。この場合、会議の参加者が多くても、音声認識処理部の数は、2個で済むことになる。

10

【0025】

前記入力選択部におけるマイクロホンの選択は、多数出席している会議であっても、同時に発言を行うのは2人ないし3人程度であるということを前提にしている。もし、これ以上多くの者が同時に発言した場合、会議は、紛糾状態であり、それらの発言を書き起こしても意味がない。

【0026】

音声認識処理部の数は、前記入力選択部において選択できるマイクロホンの数と同じである。前記それぞれの音声認識処理部によって認識された音声は、そのまま音声保存部に保存される。また、前記それぞれの音声認識処理部によって認識されたデータは、文字列になり、文字列保存部に保存される。

20

【0027】

前記音声保存部に保存された音声、および文字列保存部に保存された文字列は、コンピュータの表示手段に音声の波形と認識結果としての文字列が表示される。音声は、必要に応じて、コンピュータのスピーカーから録音された通りに出力される。会議の議事録を書き起こす者は、コンピュータの表示手段に表示されている音声の認識結果が正しいと判断した場合、そのまま採用し、前記音声の認識結果が間違いであると判断した場合、実際の音声を聞きながら修正した文字列とする。

【0028】

たとえば、前記会議の議事録を書き起こす者は、前記音声の認識結果が不明であると判断した場合、表示手段に表示されている音声の波形部分またはその近傍のボタンをクリックすることにより、会議の発言内容をスピーカーによって再現する。前記会議の議事録を書き起こす者は、前記音声の認識結果と実際の録音内容とに基づいて、正しい文字列を時系列的に作成する。

30

【0029】

会議の議事録を書き起こす者は、音声認識処理部における認識結果が文字列保存部に保存されており、前記認識結果を表示手段により表示させると同時に、音声保存部に保存された音声を聞きながら、文字列編集部によって、正確な文字列にするだけでなく、時系列的にも正しく編集することができる。

【0030】

40

(第2発明)

第2発明は、会議中に誰も発言しない場合であっても、無音の入力があるとみなしている。すなわち、入力選択部は、前記マイクロホンから入力した零個の信号(無音)、一個の信号、複数個の信号を選択し、前記選択された信号を異なる音声認識処理部にそれぞれ出力する。1ないし2個の重複した発言がある場合、異なる音声認識処理部は、それぞれを認識結果として、異なる発言者毎の文字列とする。

【0031】

(第3発明)

第3発明は、マイクロホンから入力した信号をデジタル信号またはアナログ信号とすることで、入力選択部にソフトウェアを使用したり、あるいは、オーディオ機器に使用するハ

50

ードウェアを使用することができる。前記入力選択部によって選択されたそれぞれの信号は、異なる音声認識処理部によって音声認識される。前記マイクロホンから入力されたアナログ信号は、音声認識処理部の前でA/D変換されるようにすることもできる。

【0032】

(第4発明)

第4発明は、入力選択部における選択順位が付けられている。すなわち、前記選択順位は、入力された信号の強さが予め決められたレベルを一定時間以上連続して超えているものを先着順としている。第4発明は、複数の発言者が重複して発言していても、ある一定レベルの強さと連続時間で優先順位をつけているため、隣り合った席の者が小さな声で内輪話をしても、選択されないようにすることができる。

10

【0033】

(第5発明)

第5発明は、前記入力選択部にコンピュータソフトウェアを使用することで、複数の発言者が同時であった場合、それぞれのマイクロホンで録音した内容を別々に録音した後、音声認識を正確にすることができる。

【0034】

(第6発明)

第6発明は、各音声認識処理部における音声認識結果を時系列に並べて表示すると同時に、音声認識結果に対する音声を再生する機能を有する。会議の議事録を書き起こす者は、必要に応じて、再生された音声を聞きながら、前記音声認識結果を時系列的に修正して、議事録を容易に作成することができる。

20

【0035】

【実施例】

図1は本発明の一実施例で、会議室における会議の議事録を書き起こすシステムを説明するための概略ブロック構成図である。図1において、会議室11には、発言者(会議参加者)の数だけマイクロホン1、マイクロホン2、・・・、マイクロホンmが設置されている。そして、前記各マイクロホンには、コンピュータ12の音声入力端子に接続できるコンピュータ入力端子114が接続されている。

【0036】

コンピュータ12には、前記コンピュータ入力端子114に接続されている入力選択部121と、前記入力選択部121によって選択された音声を認識できる音声認識処理部(1) 122、音声認識処理部(2) 123、・・・、音声認識処理部(n) 124と、音声をそのまま保存する音声保存部125と、前記各音声認識処理部(1) ないし(n) で認識した文字列を保存する文字列保存部126と、前記音声保存部125によって保存された音声を出力するスピーカ出力部127と、前記文字列保存部126によって保存された文字列を表示する表示部128と、スピーカ出力部127の出力、および/または表示部128に表示された文字列を編集する文字列編集部129と、編集された文字列130とから構成される。

30

【0037】

会議の出席者数=マイクロホン数m>>音声認識処理部nである。そして、前記音声認識処理部(1) ないし(n) は、発言者の音声信号が一定レベルを超えたものを採用し、一定レベル以下のものを無視する。また、前記音声認識処理部(1) ないし(n) は、前記優先順位にしたがい、発言者の1ないしmの中のn個を同時に選択することができる。

40

【0038】

図2は本発明の一実施例におけるコンピュータの表示部を使用した編集の例を説明するための図である。図2において、表示部128(図1参照)は、編集領域21と、音声パターン表示部22と、会議情報表示部23とから構成されている。

【0039】

編集領域21には、会議情報表示部23の一部を選択することによって、音声認識処理部

50

によって認識された認識結果 2 3 3 が表示される。会議情報表示部 2 3 は、発言者の氏名が表示される発言者表示部 2 3 1 と、発言時間が表示される時間表示部 2 3 2 と、発言が音声認識処理部 1 ないし n によって認識された結果を表示する認識結果表示部 2 3 3 と、前記編集領域 2 1 で編集された結果を表示する編集結果表示部 2 3 4 とから構成されている。

【0040】

図 1 および図 2 を参照して会議における発言者の音声認識に基づく文字列の編集について説明する。会議は、m 人が参加しており、それぞれにマイクロホン 1、2、・・・m が備えられている。前記各マイクロホンには、コンピュータ入力端子 1 1 4 が設けられており、コンピュータ 1 2 の端子（図示されていない）に接続されている。

10

【0041】

今、二人の発言者は、同時に、マイクロホン 1 とマイクロホン 2 を通して発言したとする。入力選択部 1 2 1 は、マイクロホン 1 およびマイクロホン 2 の出力を同時に選択して、マイクロホン 1 の出力を音声認識処理部（1） 1 2 2 に、マイクロホン 2 の出力を音声認識処理部（2） 1 2 3 に割り当てる。

【0042】

前記音声認識処理部（1） 1 2 2 および音声認識処理部（2） 1 2 3 によって認識された音声は、音声保存部 1 2 5 にそれぞれ保存される。前記音声認識処理部（1） 1 2 2 および音声認識処理部（2） 1 2 3 によって認識された文字列は、文字列保存部 1 2 6 にそれぞれ保存される。

20

【0043】

前記音声保存部 1 2 5 および文字列保存部 1 2 6 に保存された音声および文字列は、編集する際にコンピュータ 1 2 の表示部 1 2 8 における編集領域 2 1 に表示される（図 2）。たとえば、編集者は、1 4 時 5 4 分 2 4 秒の発言を選択（2 3 4 1）すると、発言者ヤスモトマイが、「大議論をもの中で忘れててました。」という認識結果が編集結果表示部 2 3 4 に表示される。また、同じ時間に、アベユウコは、「はは」と笑ったと認識されている。

【0044】

前記文字列保存部 1 2 6 に保存されていた前記ヤスモトマイの発言の認識結果は、内容が判らないため、編集者は、音声パターン表示部 2 2 またはその近傍のボタンをクリックすることにより、音声保存部 1 2 5 に保存されていた音声出力がコンピュータ 1 2 のスピーカー出力部 1 2 7 から音声信号が出力される。また、発言者表示部 2 3 1、時間表示部 2 3 2、認識結果表示部 2 3 3、および編集結果表示部 2 3 4 のいずれかを選択した際に、自動的に音声出力することもできる。

30

【0045】

前記編集者は、前記音声信号を聞きながら、前記「大議論をもの中で忘れててました。」が「大事なものを書くのを忘れてました。」であると判断して、編集領域 2 1 において、正しく編集する。前記発言と同時に、全く異なる「はは」というアベユウコの発言は、別のマイクロホン 2 を介して音声認識処理部（2） 1 2 3 で処理されたため、正確に認識されている。

40

【0046】

図 2 に示す画面には、1 4 時 5 4 分 4 3 秒に、ヤスモトマイが「こうしようとおもうしを大島まで全部作ってください」という認識結果が選択され、編集領域 2 1 において、「仕様書まで全部作ってください。発想は、・・・」と編集した状態が示されている。

【0047】

発言者とマイクロホンとの対応関係は、最初に自己紹介として氏名をいうことにより、次回から、声紋により識別することができるため、発言毎に氏名をいう必要がない。あるいは、マイクロホン 1 は〇〇さん、マイクロホン 2 は□□さんと予め対応させておくこともできる。

【0048】

50

以上、本実施例を詳述したが、本発明は、前記本実施例に限定されるものではない。そして、本発明は、特許請求の範囲に記載された本発明を逸脱することがなければ、種々の設計変更を行うことが可能である。本発明の音声認識処理部等ブロック構成図の具体的技術は、周知または公知の技術を使用することができるため、詳細が省略されている。

【0049】

【発明の効果】

本発明によれば、会議において、同時に発言する者の数は、2人ないし3人程度であることに着目し、出席者の数に対応するマイクロホンからの出力を選択して、2ないし3個の音声認識処理部を有するだけで、重複した発言を明瞭に書き起こすことができる。

【0050】

本発明によれば、音声保存部と文字列保存部とをそれぞれ設けることにより、音声を聞きながら文字列を時系列的に書き起こすことができる。

【0051】

本発明によれば、発言者の氏名は、対応しているマイクロホンの番号または声紋によって識別ができるだけでなく、発言時間とともに、編集結果が書き起こされる。前記発言時間は、会場内の無音時間が判り、時系列としても正確な編集結果が書き起こされる。

【0052】

本発明によれば、発言者の数とマイクロホンの数が一致しているため、重複発言があっても、正確かつ容易に編集結果が書き起こされる。本発明によれば、前記マイクロホンの数が多くても、入力選択部によって、発言のあったマイクロホンの出力のみを選択できるため、音声認識処理部の数をパーソナルコンピュータ一台で処理できる程度にすることができ

【0053】

本発明によれば、多数の発言者の発言が時系列的に編集されているため、特定の発言者の発言内容に着目して編集することができる。また、本発明によれば、テーマに対する賛成派、反対派の意見を別々に対応して編集する等、いろいろな目的に適った編集が可能である。

【図面の簡単な説明】

【図1】本発明の一実施例で、会議室における会議の議事録を書き起こすシステムを説明するための概略ブロック構成図である。

【図2】本発明の一実施例におけるコンピュータの表示部を使用した編集の例を説明するための図である。

【図3】従来例の会議における議事録を時系列で書き起こすための概念図である。

【符号の説明】

11・・・会議室

111・・・マイクロホン1

112・・・マイクロホン2

113・・・マイクロホンm

114・・・コンピュータ入力端子

12・・・コンピュータ

121・・・入力選択部

122・・・音声認識処理部1

123・・・音声認識処理部2

124・・・音声認識処理部n

125・・・音声保存部

126・・・文字列保存部

127・・・スピーカー出力部

128・・・表示部

129・・・文字列編集部

130・・・文字列

10

20

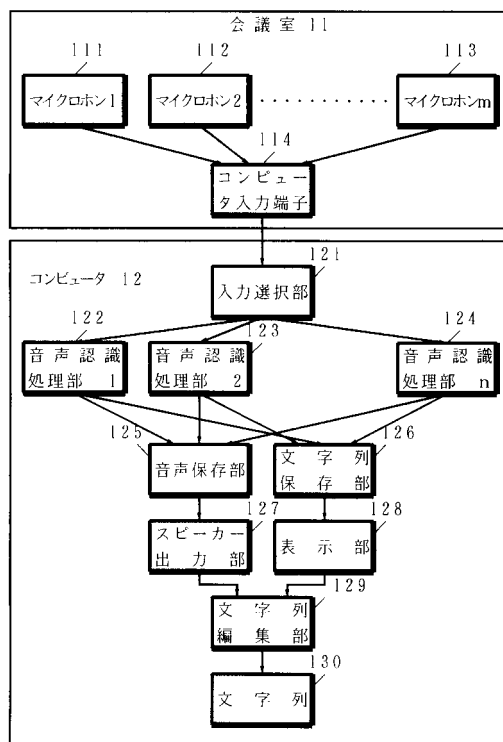
30

40

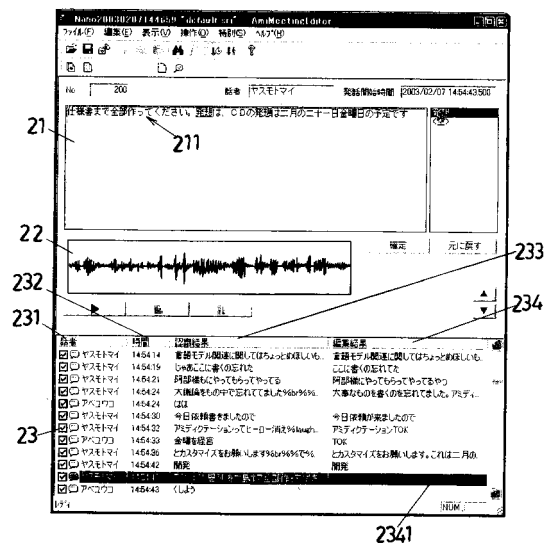
50

- 2 1 . . . 編集領域
- 2 1 1 . . . 編集 中の 文字 列
- 2 2 . . . 音声 パターン 表示 部
- 2 3 . . . 会議 情報 表示 部
- 2 3 1 . . . 発言 者 表示 部
- 2 3 2 . . . 時間 表示 部
- 2 3 3 . . . 認識 結果 表示 部
- 2 3 4 . . . 編集 結果 表示 部
- 2 3 4 1 . . . 認識 結果 選択 部

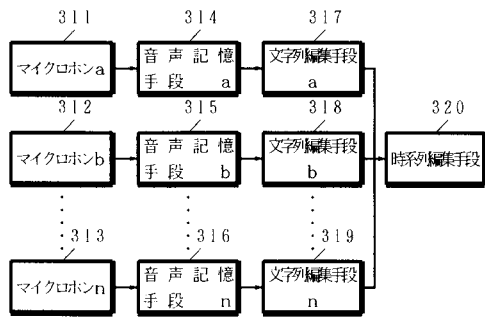
【 図 1 】



【 図 2 】



【 図 3 】



フロントページの続き

- (72)発明者 中村 雅巳
東京都豊島区東池袋三丁目1番1号 サンシャイン60 48階 株式会社アドバンスト・メディア内
- (72)発明者 山田 実一
東京都豊島区東池袋三丁目1番1号 サンシャイン60 48階 株式会社アドバンスト・メディア内
- (72)発明者 新開 正男
東京都豊島区東池袋三丁目1番1号 サンシャイン60 48階 株式会社アドバンスト・メディア内
- (72)発明者 青木 秀仁
東京都豊島区東池袋三丁目1番1号 サンシャイン60 48階 株式会社アドバンスト・メディア内

Fターム(参考) 5D015 DD02 KK03